October 2010

IC-CAP Modeling Handbook

Author Name: Franz Sischka

© Agilent Technologies, Inc. 2000-2010
3501 Stevens Creek Blvd., Santa Clara, CA 95052 USA
No part of this documentation may be reproduced in any form or by any means (including electronic storage and retrieval or translation into a foreign language) without prior agreement and written consent from Agilent Technologies, Inc. as governed by United States and international copyright laws.

Acknowledgments
UNIX ® is a registered trademark of the Open Group.
Windows ®, MS Windows ® and Windows NT ® are U.S. registered trademarks of Microsoft Corporation.

Errata The IC-CAP product may contain references to "HP" or "HPEESOF" such as in file names and directory names. The business entity formerly known as "HP EEsof" is now part of Agilent Technologies and is known as "Agilent EEsof." To avoid broken functionality and to maintain backward compatibility for our customers, we did not change all the names and labels that contain "HP" or "HPEESOF" references.

Warranty The material contained in this documentation is provided "as is", and is subject to being changed, without notice, in future editions. Further, to the maximum extent permitted by applicable law, Agilent disclaims all warranties, either express or implied, with regard to this manual and any information contained herein, including but not limited to the implied warranties of merchantability and fitness for a particular purpose. Agilent shall not be liable for errors or for incidental or consequential damages in connection with the furnishing, use, or performance of this document or of any information contained herein. Should Agilent and the user have a separate written agreement with warranty terms covering the material in this document that conflict with these terms, the warranty terms in the separate agreement shall control.

Technology Licenses The hardware and/or software described in this document are furnished under a license and may be used or copied only in accordance with the terms of such license.

Restricted Rights Legend U.S. Government Restricted Rights. Software and technical data rights granted to the federal government include only those rights customarily provided to end user customers. Agilent provides this customary commercial license in Software and technical data pursuant to FAR 12.211 (Technical Data) and 12.212 (Computer Software) and, for the Department of Defense, DFARS 252.227-7015 (Technical Data - Commercial Items) and DFARS 227.7202-3 (Rights in Commercial Computer Software or Computer Software Documentation).

Contents 13
Manufacturing Integrated Circuits 14
Understanding the Semiconductor Process 14
Process and Device Modeling 15
Device Modeling Measurement Tutorial 30
Basics of DC and AC Characterization of Semiconductors 30
PlotOption Access Name and Value Examples 30
The Guarding (Triax Cable) Principle for Fast and Accurate Measurements 31
Measurement Environment 31
Measurements 55
Contents 55
Measurement Techniques 56
Contents 56
Prerequisites 57
Contents 57
Wafer Prober 58
FMAX Cables 68
Agilent RF and Microwave Test Accessories 69
Overview 69
Adapters, Cables and Connectors - General Connector Information 69
Metrology/Instrument Grade Adapters 70
Metrology Grade Adapters 71
Instrument Grade Adapters 71
General Purpose Grade Adapter Selection Guide 72
Adapter Kit Selection Guide 75
Flexible Microcircuit Packaging 76
Triax Feedthru and Triax-Coax Converters from Trompeter Electronics 79
Introduction 80
Probe Features Affect Layout 80
Layout Rules, Mechanical 82
Layout Rules, Electrical 85
Specific Applications 89
Digital Devices 91
Characterization Devices 92
Summary 95
DC 96
Contents 96
DC Characterization for Device Modeling 103
Oscillations during DC measurements 110
How to avoid oscillations 111
Monitoring the SMUs by an oscilloscope 111
Connecting triax SMU outputs to coax inputs 111
The Effect of Big Capacitive Loads 111
SMUs and Capacitive loads - The Effect of Long Cables 112
Conclusion 113
Miscellaneous Information 114
Process Control Monitoring (PCM) Measurements versus Modeling Measurements 115
CV 117
Contents 117
CV Characterization 118
How do capacitance meters work? 118
On-Wafer Measurements 119
CV on-wafer measurements - What to do with the unused pins? 120
3 Contacting Cases - for CV Calibration and Measurement, Applying the Auto-Balancing Method 120
Simulations of a Multi-pin Device CV Measurement 122
Considerations regarding the signal level and the CV-frequency 124
Obtainable CV measurement resolution 124
Conclusion 125
Miscellaneous Information 125
CV Measurement And Calibration Techniques 129
Network Analyzers 139
Contents 139
Measurement 140
Contents 140
Displaying Phase Shift 141
Linear Vector Network Analyzer Measurements 142
Standalone VNA's 142
DC Biased VNA's 143
Calibration 150
Contents 150
Calibration Procedures 151
Checking the max. Applicable RF Signal Level 152
Important Note on the Agilent N5250A 110GHz PNA Network Analyzer 154
Considering the DC Bias Losses During S-Parameter Measurements 154
Entering Calkit Data for Connectorized Devices 156
Entering Calkit Data for On-Wafer Probes 157
Entering Calkit Data for Packaged Devices 160
Entering CalKit Data into the Network Analyzer 160
Entering SOLT Calkit Data (Agilent PNA-NetworkAnalyzer-specific) 162
Entering SOLT Calkit Data HP NetworkAnalyzer specific 168
Entering TRL Calkit Data Agilent NetworkAnalyzer specific 170
NWA Calibration in General 172

NWA Calibration Procedure For At-Temperature Measurements . . . 174
NWA Calibration Verification . . . 174
NWA Instrument Options in IC-CAP . . . 175
Performing the NWA Calibration from IC-CAP . . . 178
Defining Custom Calkits . . . 179
Introduction to NWA Calibration . . . 188
VNA Cal Kit Manager 2.0 . . . 201
S-Parameters . . . 203
Contents . . . 203
Amplifier Characteristics from S-Parameters . . . 204
Input Reflection Coefficient with arbitrary load impedance Z L . . . 204
Output Reflection Coefficient with arbitrary source impedance Z S . . . 204
Voltage Gain with arbitrary Z S and Z L . . . 204
Stability: . . . 204
Power Gain from S-Parameters . . . 204
Power Gain . . . 205
Available Power Gain . . . 205
Transducer Power Gain . . . 205
Maximum Available (stable) power amplification Gain . . . 205
Maximum Unilateral Transducer Power Gain . . . 206
Mason's power gain . . . 206
Definition of the Transit Frequency . . . 206
Definition of the Maximum Oscillation Frequency fmax: . . . 207
Notes on Estimating fmax vs fT by Extrapolation, and Comparison of fT vs fmax Measurements . 208
Considerations About the Validity of Unilateral Gain Calculations . . . 210
Calculating S-parameters from Voltages . . . 211
Converting N-Port S-Parameters to M-Port . . . 212
Converting from 2-PORT -> 3-PORT . . . 212
Once these 3-port S-parameters are known, they can be used to calculate the common-base or common-collector 2-port S-parameters of a bipolar transistor etc. . . . 213
Converting from 2-PORT -> 1-PORT and its Application for the Q factor of RF-passive Components . . . 213
Definition of S-Parameters . . . 214
Differential S-Parameters . . . 218
Calculating differential and Common-mode S-parameters directly from S-parameters . . . 218
Related to the Q-Factor of Spiral Inductors Operated in Differential Mode . . . 218
Differential 1-Port S-Parameters . . . 219
Publications on Differential/Common Mode Z-parameters . . . 219
Interpreting S-Parameter Plots . . . 220
INTRODUCTION TO UNDERSTANDING S-PARAMETER PLOTS . . . 220
Phase Shift of Inductors and Capacitors . . . 229
Attenuation of Resistors Related to S-Parameters . . . 229
About the Interpretation of S-Parameters of an Asymmetrical Tee Structure . . . 231
Interpreting S-Parameter Plots of a Transistor . . . 233
Starting Points of Transistor S-Parameter Plots . . . 233
Interpreting S-Parameters of Strip Line . . . 235
S-PARAMETERS OF SERIAL LOSSLESS DELAY LINES WITH DIFFERENT IMPEDANCES . . . 237
Q-Factor Calculation . . . 240
Small Signal Versus Large Signal S-Parameters . . . 241
S-Parameter Basics for Modeling Engineers . . . 241
For Experts . . . 248
S-Parameter Impedance Matching (for obtaining max. power gain) . . . 250
Smith Chart Circuit Elements Paths . . . 253
Forbidden Regions of the Smith Chart . . . 254
S-Parameters Measurements for Modeling . . . 255
10 Rules for Successful S-parameter Handling . . . 255
So, How To Deal With S-Parameters? . . . 265
Appendix . . . 266
S-Parameter Utilities . . . 270
Calculation the radiation losses from S-parameters . . . 270
Understanding S-Parameter Plots . . . 271
The Smith Chart for Sxx Parameters . . . 271
THE POLAR PLOTS FOR THE S12 AND S21 PARAMETERS . . . 273
De-embedding Techniques . . . 275
Contents . . . 275
Basics of De-embedding Techniques . . . 276
Why De-embedding ? . . . 276
Chip carrier . . . 276
Converting the S-Parameters to Y, we can strip-off the capacitors . . . 277
At a glance: . . . 277
A note on adding of Y matrices . . . 277
A typical representation of Matrix Yparas . . . 278
A note on adding of Z Matrices . . . 278
For on-wafer OPEN dummy structures, Yparas looks like this: . . . 279
The Lumped Components Values are: . . . 279
Physical Interpretation of the OPEN dummy . . . 280
Reminder on the de-embedding of the complete Y matrix . . . 280
For on-wafer SHORT dummy structures, Zparas looks like this: . . . 281
about the de-embedding of complete Y and Z matrices . . . 281
De-embedding of complete Y and Z matrices . . . 281
After so much theory, a look at the reality . . . 282
Verifying the de-embedding for on-Wafer transistor measurements . . . 285
Model The On-Wafer OPEN Dummy . . . 285
Model The On-Wafer SHORT Dummy . . . 286

- Verify The De-Embedding With Modeling The THRU Dummy . . . 286
- An Example of bad THRU deembedding results . . . 287
- De-Embedding Conclusions . . . 287
- Try yourself in IC-CAP . . . 287

De-Embedding Techniques Tutorial . . . 288
- De-embedding tutorial . . . 288
- About the de-embedding . . . 288
- MATRIX MANIPULATIONS IN DETAILS . . . 289
- DE-EMBEDDING OF SERIES PARASITICS (Z-MATRIX) . . . 289
- DE-EMBEDDING OF PARALLEL PARASITICS (Y-MATRIX) . . . 290
- DE-EMBEDDING BY SUBTRACTING THE Z AND Y MATRICES . . . 290
- HOW TO DETERMINE THE SEQUENCE OF Y AND Z MATRIX DE-EMBEDDING . . . 291
- DE-EMBEDDING OF LOSSLESS 50 Ohm DELAY LINES (S-MATRIX) . . . 292
- DE - EMBEDDING USING ABC - MATRICES . . . 292
- DE - EMBEDDING OF CHAINS OF ABC - MATRICES . . . 293
- DE-EMBEDDING OF LOSSY DELAY LINES WITH GENERAL CHARACTERISTIC IMPEDANCE . . . 293

ft Modeling of a Transistor Affected by Parasitics . . . 296
- An introduction to de-embedding . . . 296

Interpreting OPEN-Dummy Measurements . . . 298
- Effect of R1, C1 and R2, C2 on Sxy . . . 298
- Effect of R12 and C12 on Sxy . . . 298
- Effect of R3 on Sxy . . . 298
- Effect of L3 on Sxy . . . 298
- QUIZ: . . . 299
- Open Dummy High Frequency Modeling . . . 299
- How the Silicon Substrate and the Silicon Oxide parameters affect the 'knee' in the OPEN Dummy Structure S11 . . . 301
- A study on the effect of the oxide loss tangent and the substrate conductivity . . . 301
- The signal pad of the OPEN sees basically a sequence of SiO2 > silicon substrate > SiO2, i.e. a circuit schematic like below . . . 302
- Interpreting the OPEN Sxx related to substrate and oxide params . . . 302

Twoport Matrices of Basic Schematics . . . 303

Twoport Matrix Conversions . . . 305
- H to Y Conversion . . . 305
- Y to Z and Z to Y Conversion . . . 305
- S to H Conversion . . . 305
- Matrix Conversions at a Glance . . . 305
- TEE to PI Conversion . . . 306

Twoport Matrix Definitions . . . 308
- Terms and Definitions on Series and Parallel Circuits . . . 308
- Twoport Matrix Signal Flow Definition . . . 308

Verifying the De-embedding Procedure . . . 310
- Modeling of the Open Dummy . . . 311

Vector Network Analyzer - Basics for Modeling Engineers . . . 312
- Measurement Errors . . . 314
- Vector Error Correction . . . 314
- What to do for Vector-Error Correction? . . . 314
- Practical Considerations before starting with NWA Measurements . . . 315
- SOLT calibration for Ground-Signal-Ground probes (G_S_G) . . . 315
- Other Calibration Methods . . . 316
- Verify the Calibration . . . 316
- Verify the Calibration . . . 316
- Summary . . . 316

Nonlinear Network Analyzers . . . 317
- Contents . . . 317

Nonlinear De-Embedding . . . 318
- 2-Step Nonlinear De-Embedding . . . 318
- 3-Step Nonlinear De-Embedding . . . 319

Nonlinear RF Measurements and Device Modeling . . . 321
- From DC > CV > NWA . . . 321
- Lightwave Analogy to RF Energy . . . 321
- Linear Vector Network Analyzer . . . 322
- Considering a 2-port: stimulated at port1, then at port 2 . . . 322
- S-Parameters = linear theory . . . 323
- Beyond S-Parameters . . . 324
- Extending linear to RF-power-dependent signals . . . 324
- Linear versus nonlinear behavior . . . 325
- Phase Reference . . . 328
- Phase: The critical element . . . 328
- NVNA measurement . . . 328
- What is Harmonic Balance Simulation? . . . 331
- X-Parameters . . . 335
- What are HOT S-Parameters? . . . 336
- Ideal X-Parameter Extraction Steps (Perfect match, no source harmonics) . . . 337
- X-Parameter Measurement Principle . . . 338
- X-Parameters . . . 338
- Intermezzo: NVNA X-Parameter Display . . . 338
- Intermezzo: Number of X-Parameter measurements . . . 339
- Example of an X-Parameter MDIF file, created by the NVNA . . . 339
- Simulating with X-Parameters . . . 340
- ADS:X-Parameters become the PHD Model . . . 340
- PHD Model . . . 341
- Performing the NVNA on-wafer Measurements . . . 343

NVNA Calibration . . . 343
NVNA Calibration Plane . . . 344
Importing the NVNA (DC, i1, v1, i2, v2) mdif data file into IC-CAP . . . 344
Case Study Example . . . 345
Notes about De-Embedding . . . 346
Calculate the Fourier components of the vx and ix . . . 347
Handling NVNA measurements and standard device modelslike PSP or Hicum, Gummel-Poon, Angelov etc. in IC-CAP . . . 348
A possible Modeling Sequence: . . . 348
The IC-CAP Nonlinear RF Tutorial ModelFile (1) . . . 348
The IC-CAP Nonlinear RF Tutorial ModelFile (2) . . . 349
The IC-CAP Nonlinear RF Tutorial ModelFile (3) . . . 349
Conclusions . . . 350
Modeling Tips and Tricks . . . 350
vDS vGS in the time domain . . . 350
Time Domain . . . 352
Contents . . . 352
TDR Measurements Basics . . . 353
Modeling the Complex Load Out of TDR Plots . . . 354
About the Context Between Time and Location . . . 355
TDR Calibration Techniques . . . 356
Calibration Techniques . . . 356
TDR Plots Tutorials . . . 358
Introduction . . . 358
Expected Reflectograms of Typical Parasitic Components . . . 358
TDR Plots for Experts . . . 363
Differential TDR Measurements . . . 365
Summary Differential TDR . . . 365
Summary Differential TDT . . . 365
Noise . . . 366
Contents . . . 366
Noise Measurement Golden Rules . . . 367
Verifying the Measurement Data of the IC-CAP 1/f Noise Toolkit . . . 367
A Note on Data Consistency Checks for MOS Transistors . . . 372
Appendix . . . 372
Spectrum Analyzer . . . 375
Data Verification . . . 376
Data Verification and Data Consistency Checks . . . 376
Prologue . . . 376
DC vs. DC . . . 376
DC Measurements vs. Spar DC Bias . . . 377
Spar DC bias data mirrored into foutput . . . 377
CV vs. Spar converted to CV . . . 378
Spar Starting Points vs. Existing DC Modeling . . . 378
Data Management . . . 382
Measurement Standardization and Data Consistency Checks . . . 382
Required Naming Conventions for the concept of separated measurements and extractions . . . 382
Project Flow Suggestion for Reusable Model Files . . . 383
Examples for Data Consistency Checks . . . 384
Advantages of Standardization . . . 385
Extraction . . . 386
Contents . . . 386
Curve Fitting Techniques . . . 387
Introduction . . . 387
Contents . . . 387
Synthesizing Measurement Data for Verification of Extraction Methods . . . 388
Regression Analysis . . . 389
Mathematical Basics of Curve Fitting . . . 389
Linear Curve Fitting . . . 389
General Formula for Regression Analysis . . . 390
Application: Calculating the quadratic polynomial parameters from measurement data . . . 391
Applying Linear Regression Analysis to Modeling . . . 393
DC curves: . . . 393
General CV modeling . . . 394
CV modeling with MJ=const, but including offset capacitance C offs . . . 395
Total CV modeling including offset capacitance . . . 395
Regression Analysis for S-Parameters . . . 396
Summary . . . 396
Direct, Visual Parameter Extraction . . . 398
General Modeling Techniques . . . 401
Diode DC modeling . . . 401
MOS Gate Oxide Capacitance (simplified model) . . . 402
Data Interpolation . . . 404
Simulations . . . 405
Introduction . . . 405
Contents . . . 405
History of Circuit Simulators . . . 406
The origin of Berkeley SPICE . . . 406
The origin of ADS . . . 406
References . . . 406
Simulator Interface of IC-CAP . . . 407
How IC-CAP interfaces with the simulators. . . . 407
Netlist Syntax for ADS and SPICE . . . 410
Contents . . . 410

Comparing UCB Spice and ADS Netlist Syntax 411
Table 1: summary of the most commonly used components 411
Common Parameter List 412
How to have a common Parameter List for Spice and ADS 412
UCB SPICE Simulators Netlists (detailed) 413
User-defined Model Equations in Berkeley SPICE 3 414
ADS Simulator Netlist Syntax (detailed) 415
Getting a list of all ADS components/models: 415
User-defined Model Equations in ADS 417
Using ADS sdd's (Symbolic defined devices) 417
Using Verilog-A 417
Netlist Syntax for ADS and SPICE 420
CV and S-Parameters 421
How to Obtain CV Curves from SPICE 422
Calculating CV Curves from Complex Currents 422
How to Obtain S-Parameter Curves from SPICE 423
Calculating CV Curves from Complex Voltages 423
Harmonic Balance 427
ADS 427
What is Harmonic Balance (HB) Simulation? 427
What happens when a nonlinear harmonic balance simulation is executed? 427
Nonlinear High Frequency Characteristics 429
Harmonic Balance Simulations for Modeling 431
History 435
History of Circuit Simulators 435
Device Modeling 436
Contents 436
Diode 437
Contents 437
Modeling a Diode 438
Introduction 438
DC Characterization Parameters IS, N AND RS 439
Determination of the DC parameters IS and N 439
Parameter RS 440
CV Characterization 440
HF Modeling: Parameter TT 442
Model Limitations 443
Details of the Berkeley SPICE Diode Modeling for Experts 443
SPICE Diode Model Equations 444
Publications 445
ADS Junction Diode Model 446
INSTANCE 446
MODEL CARD 446
Step Recovery Diode 447
Standard Diode Model 449
Non Quasi-static Diode Model 450
Parameter Extraction 450
Model Implementation 451
Modelling Results 452
Summary 453
References 454
Transistors 455
Bipolar in General 456
Contents 456
Advanced Bipolar Transistor Modeling 457
Why BJT Modeling Is Still Hot? 457
What is a Compact Model? 457
Why Standard Compact Model? 457
History of BJT Modeling 457
Example of Model Evolution 457
The Spice Gummel-Poon Model 458
SGP-Model: DC Formulation 458
SGP-Model: normalized Base Charge q b 458
SGP-Model: Parasitic Resistors 459
SGP-Model: Capacitors 459
SG-P Model: Additional Phase Shift and Temperature 460
The Spice Gummel-Poon Model 460
SGP: Parameter Extraction Strategy 460
SGP Extraction: Space Charge Capacitors 461
SGP Extraction: Parasitic Resistors: RE 461
SGP Extraction: DC biasing 462
G-P Extraction: Early-Voltages 462
G-P Extraction: Gummel-Poon Plot Parameters 462
SGP Extraction: TFF Parameters 462
SGP Extraction: Final Parameter Tuning 463
SGP Extraction: Final Result - Example 463
Bipolar Transistor Modeling 464
History of Bipolar Transistor Modeling 464
The Gummel-Poon Model 464
GP-Model: normalized Base Charge q b 464
GP-Model: normalized Base Charge q b 465
GP-Model: DC Formulation 465
GP-Model: Parasitic Resistors 465

GP-Model: Capacitors . . . 466
G-P Model: Additional Phase Shift and Temperature . . . 466
GP-Model: Model Parameters . . . 466
G-P Extraction: Notes . . . 467
G-P: Parameter Extraction Strategy . . . 467
G-P Extraction: Space Charge Capacitors . . . 467
G-P Extraction: Parasitic Resistors: RE . . . 467
G-P Extraction: Parasitic Resistors: RBM . . . 467
G-P Extraction: DC Biasing . . . 468
G-P Extraction: Early-Voltages . . . 468
G-P Extraction: Gummel-Poon Plot Parameters . . . 469
G-P Extraction: S-Parameter De-embedding . . . 469
G-P Extraction: S-Parameter: Data Management . . . 469
G-P Extraction: Data Management ASCII File . . . 470
G-P Extraction: RB Parameters . . . 470
G-P Extraction: TFF Parameters . . . 470
G-P Extraction: Final Parameter Tuning . . . 471
BSIM . . . 472
Contents . . . 472
BSIM3.3v2 - Parameter Extraction Strategy . . . 472
Fully Scalable BSIM4 Modeling From DC to RF . . . 476
GaAs Transistor Models in General . . . 490
General Transistor Modeling Strategies . . . 491
Measurements . . . 491
Data Consistency Checks DC . . . 491
Data Consistency Checks S-Parameters . . . 491
Wrap-up: The applied checktools and verification model files: . . . 492
After all the data have been checked, we can finally start the modeling . . . 492
Example: CV Modeling Gummel-Poon . . . 493
Transfer Curve(s) Modeling . . . 495
Output Characteristics Modeling . . . 496
S-parameter Fitting . . . 497
Overview of the Hicum Level 2 model . . . 500
Forewords . . . 500
What is compact modeling? . . . 500
Need for accurate models . . . 500
HiCUM can help circuit design . . . 500
HiCUM is also better for Scaling and Statistics . . . 502
Statistics can help designers . . . 502
HiCUM Core Model . . . 502
Transfer current . . . 502
Base Current . . . 503
Depletion charges and capacitances . . . 503
BE Capacitance . . . 504
Transit Time . . . 505
Transit Frequency . . . 505
Transit Frequency and Critical current . . . 506
Critical current ICK . . . 506
Scalable Modeling . . . 507
Extraction Example: Junction Capacitance . . . 510
Example results: f T (30GHz Process) . . . 511
Statistical Modeling: The SMACH Example . . . 511
Definition . . . 512
Statistical Modeling Procedure . . . 512
Check Correlation between PCMs . . . 513
Statistical Modeling Procedure . . . 513
Learning the Angelov Model . . . 514
IPK0 Current for maximum transconductance . . . 514
VPKS Gate voltage for maximum transconductance . . . 514
P1 Polynomial coefficient for channel current . . . 514
P2 Polynomial coefficient for channel current . . . 514
P3 Polynomial coefficient for channel current . . . 515
LAMBDA Channel length modulation parameter . . . 515
ALPHAR Saturation parameter . . . 515
ALPHAS Saturation parameter . . . 515
UCB MOS Level-3 MODEL . . . 517
MOS Transistors for Dummies . . . 522
MOS Modeling for Dummies . . . 523
PSP . . . 525
Contents . . . 525
Background of the PSP Model Extraction Package . . . 526
Introduction to PSP . . . 526
Internal model structure of PSP . . . 526
Parameter extraction principle for scalable models . . . 527
Binning in PSP . . . 529
Example extraction flow . . . 529
Summary . . . 530
Global PSP Model . . . 534
PSP Model Extraction Package . . . 535
Key Features . . . 535
PSP Extraction Flow . . . 535
Extensions to the PSP Model Extraction Package from the Modeling Bench . . . 537
PSP Parameter Extraction Strategy . . . 541

CV Modeling . . . 541
Diode . . . 542
Individual Transistor DC Modeling . . . 543
The Curtice Mesfet Model . . . 546
Common parameters for all Curtice model levels: . . . 546
Diode model parameters: . . . 546
Piecewise linear model parameters: . . . 546
Level 1 parameters: . . . 546
Level 2 parameters: . . . 546
DC Modeling . . . 547
Dynamic performance . . . 548
Publications for the Curtice model . . . 550
The Gummel-Poon model . . . 551
Contents . . . 551
AC Small Signal Modeling . . . 552
Overview . . . 552
Appendices . . . 553
Linear Curve Fitting: Regression Analysis . . . 553
Verifying the quality of extraction routines . . . 554
Direct visual parameter extraction of BF, ISE and NE . . . 554
RBM . . . 556
VAF . . . 556
IS . . . 557
NE . . . 558
ISE . . . 558
IKF . . . 558
Calculating IKF . . . 559
BF . . . 559
Calculation of h21 of the Gummel-Poon Model . . . 560
Simplicification . . . 562
Modeling the output characteristics quasi-saturation effect with Gummel-Poon . . . 562
Gummel-Poon Bipolar Model . . . 564
Structure of this Manual . . . 564
About This Manual . . . 564
Gummel-Poon CV Modeling . . . 566
The Gummel-Poon Capacitor Equations . . . 566
Modeling the Space Charge Capacitors . . . 566
Measuring and modeling the Base-Emitter capacitance. . . . 567
The Equation . . . 567
Determination of the CV Parameters . . . 567
What to do in IC-CAP . . . 568
Introduction . . . 569
Contents . . . 569
Operating Modes of the Bipolar Transistor . . . 569
The Gummel-Poon Equivalent Schematic . . . 569
The Gummel-Poon Model Equations . . . 571
Proposed Global Extraction and Optimization Strategy . . . 575
Limitations of the Gummel-Poon Model . . . 577
Ohmic Effects . . . 577
Forward DC Modeling . . . 577
Reverse DC Modeling . . . 577
AC Modeling . . . 577
Temperature Modeling . . . 577
Applications in Integrated Circuits . . . 577
Conclusion . . . 577
Modeling the Base Resistor rBB' . . . 578
Extraction of RB, IRB and RBM . . . 578
Modeling the Parameter XCJC . . . 582
A final remark on the S-Parameter Extraction and Optimization Strategy . . . 582
Modeling the Resistors . . . 584
Contents . . . 584
Modeling the Emitter Resistor . . . 584
Extraction of RBM from DC measurements . . . 584
Modeling the Collector Resistor . . . 585
An alternate method to calculate the ohmic parasitic resistors from s-parameter measurements . . . 586
Modeling the Temperature Effects . . . 588
Non Linear DC Modeling . . . 589
Contents . . . 589
Modeling the Output Characteristic . . . 590
Modeling the Collector Current . . . 592
Modeling the Base Current . . . 594
Modeling the Current Amplification at High Current . . . 596
Extraction of the remaining Reverse Parameters . . . 599
Preface . . . 600
About Gummel Poon User Information . . . 600
The IC-CAP model file "GP_CLASSIC_NPN.mdl" features . . . 600
Organization of Information . . . 600
Publications . . . 601
Paras.resistor modeling . . . 601
Transit Time Modeling . . . 602
Modeling the Diffusion Capacitor CDBC . . . 602
Pre-considerations concerning the model equation: . . . 604

Performing the measurement 604
Extracting the parameters 605
How to proceed 605
Extraction of VTF 606
What to do in IC-CAP 606
How to optimize the S-parameter Setup in IC-CAP 607
Extraction of PTF 607
Extraction of TR 607
YParameter Modeling 609
Contents 609
Modeling of Transistors in the Operating Point 610
Modeling with Parameter Tuning 613
Developing A Small Signal Model For Mos Transistors 615
Publications: 616
S2Y-Parameter Transistor Modeling GUI 617
How to Proceed 617
How to use the GUI 618
What if the locus curve becomes inductive !! 619
Modeling Result Examples 620
Testing S-Parameters for Small Signal Transistors 621
Converting S-parameters into a small signal schematic of a bipolar transistor: 621
Converting S-parameters into a small signal schematic of a bipolar transistor, ... 621
Converting S-parameters into a small signal schematic of a MOS transistor: 622
VBIC 623
Contents 623
Introduction to VBIC 624
VBIC Background Information 627
VBIC Model History Notes 627
SDD - Definition of the VBIC 1.1.4 Model in MDS 627
Information on the VBIC Code, Release 1.2. 630
Equivalent Circuit Network for VBIC 1.2 631
Default Parameters rev.1.2 631
Recapitulating the Gummel-Poon Model 633
Definitions 633
DC Performance 633
Resistors 634
Dynamic Behavior 634
A Quick Tutorial on the Gummel-Poon Parameters 635
Gummel-Poon Model Parameter List 637
The VBIC Model 639
The Normalized Base Charges q b and q bp 639
Space Charge Capacitors 640
DC Performance of the Main NPN Transistor 640
DC Performance of the Parasitic PNP Transistor 641
Quasi-Saturation 642
Dynamic Behavior 643
Space Charge Capacitors 643
Diffusion Charge Capacitors of the NPN Transistor: 643
Additional Phase Shift 643
Self-Heating Modeling 644
Temperature Modeling 644
VBIC Model Parameter List 644
Comparing the VBIC and Gummel-Poon Parameters 645
VBIC Modeling Strategy 648
CV 648
DC 648
DC Finetuning 648
S-parameters 648
S-Parameter Finetuning 648
Measuring and Extracting the VBIC Parameters 649
Measurements 649
Parameter Extractions 649
Space Charge Capacitances 649
Parasitic Resistors From DC Measurements 650
Early Voltages 652
Diode Parameters 653
IS, NF 654
IKF 656
Output Characteristics 656
WSP (DC Current Distribution Between Main NPN and Parasitic NPN) 657
AVC1 and AVC2 (Avalanche Effect Modeling) 658
RTH (Output Characteristics Thermal Effect Modeling) 658
S-Parameter Modeling 659
Parametrized Modeling vs. Transistor Geometry 662
Regression Analysis 664
Mathematical Basics of Curve Fitting 664
Linear Curve Fitting 664
General Formula for Regression Analysis 665
Application: Calculating the quadratic polynomial parameters from measurement data 666
Acknowledgements, Web info, and Publications 668
Acknowledgements 668
Web Info 668
Publications 668
RF Passive Linear Components 669

Contents . . . 669
Modeling of a Chip Capacitor up to 26 GHz . . . 670
Modeling of a Dummy Chip on a Wafer . . . 671
Physical Interpretation of the Equivalent Schematic: . . . 673
Modeling the Second Order Effects of an Open Dummy Structure . . . 673
Modeling Frequency Dependent Losses . . . 675
Time Domain: . . . 675
Frequency Domain: . . . 675
Modeling . . . 675
Time Domain . . . 676
Test Fixture Modeling in Frequency Domain . . . 678
Modeling of Spiral Inductors . . . 680
De-embedding the measured data before modeling . . . 680
1-port modeling of a spiral inductor . . . 680
2-port modeling of a spiral inductor . . . 683
Calculating the Q-factor of spiral inductors . . . 686
Modeling of a 3.5mm Connector . . . 687
Modeling Vias . . . 688
Package Modeling Workshop . . . 691
Introduction . . . 691
Limitations with Standard Models . . . 691
Two Types of Characterization Measurements Required . . . 691
Modeling of a SMT Resistor . . . 697
Developing Models . . . 698
Introduction . . . 698
Developing the Equivalent Schematic . . . 698
Modeling of a SOT-23 Package . . . 703
The SOT23 package . . . 703
S-Parameters of Strip Lines . . . 707
Strip Lines Tutorial . . . 707
Modeling a lossy strip line from S-parameters . . . 707
Modeling a Strip Line Transition . . . 710
Testing S-parameters For Typical Passive Circuits . . . 711
Test Fixture Modeling in Time Domain . . . 712
Modeling of a TSOP44 Packagae . . . 717
TDR-related issues . . . 717
Low Frequency (1/f) Noise Modeling for Semiconductors . . . 725
Contents . . . 725
Introduction . . . 726
Noise Term Definitions . . . 726
Types of Noise . . . 728
Thermal Noise . . . 728
Shot Noise . . . 728
Generation-Recombination Noise . . . 729
1/f Noise . . . 729
Noise Models in Simulation Programs . . . 730
Resistors . . . 730
Inductors, Capacitors, and Lines . . . 730
Diodes . . . 730
Bipolar Transistors . . . 730
Junction FETs . . . 731
MOSFETs . . . 732
MESFETS . . . 733
Low (1/f) Noise Measurement Setup . . . 734
1/f Noise Measurements and Modeling . . . 736
Introduction . . . 736
Noise Mechanism . . . 738
Noise Models . . . 739
1/f Noise Measurement Toolkit in IC-CAP . . . 742
1/f Parameter Extraction . . . 744
Examples . . . 745
AF, KF and BF/EF Noise Parameter Extraction and Verification . . . 751
Bipolar Transistors: . . . 751
MOS Transistors . . . 753
Noise floor information on the toolkit measurement setup . . . 757
Parts list for 1/f noise toolkit . . . 757
Explanation of Spice circuit deck for 1/f noise simulations . . . 758
Thermal Modeling . . . 761
Modeling Procedure . . . 761
Static and Dynamic Thermal Modeling . . . 761
Temperature Modeling . . . 765
On Target Modeling . . . 767
Contents . . . 767
Using PCM Data for Device Modeling . . . 768
Introduction . . . 768
Application Scenarios . . . 768
On Target Modeling working . . . 770
IC-CAP Structure . . . 770
Supported simulators and models . . . 774
Applying PlotOptimizer . . . 775
Tutorial . . . 777
How to start the Target Modeling Tool . . . 777
Loading a Predefined Example . . . 777

Exploring the Example:Device Table . . . 778
Exploring the Example:Available PCM data . . . 779
Exploring the Example:Simulation Templates for PCM . . . 779
Understanding the Example:Creating a subset of devices . . . 780
Understanding the Example:Viewing subset of devices . . . 780
Understanding the Example:Defining I-V Setups from PCM data . . . 781
Understanding the Example:THE PLOTS . . . 781
Understanding the Example:Using PlotOptimizer to tune the parameters . . . 781
Understanding the Example:Report Generation . . . 781
Direct Measurement Data Reduction for Scalable Parameter Extraction . . . 783
Introduction . . . 783
New Approach for Direct Extraction of Geometrically Normalized Device Parameters . . . 783
Summary . . . 786
Statistics . . . 787
Contents . . . 787
Parametric Statistics . . . 788
Contents . . . 788
Measurement Related Statistics . . . 789
About The Difficulties To Identify The Golden Die Using Measurement-Related Statistics . . . 791
Model Parameter Related Statistics . . . 793
Principal Component Analysis . . . 793
Factor Analysis . . . 794
Practical Example . . . 795
Description of the Steps . . . 797
Literature . . . 800
Conclusions . . . 801
Publications . . . 801
Non-Parametric Statistics . . . 802
Contents . . . 802
Introduction to Non-Parametric Statistics . . . 803
Applying Non-Parametric Statistics Directly To Measured Data . . . 807
Literature about non-parametric statistics . . . 808
Identify the 'golden device' and the 'limits devices' . . . 809
Conventional Statistics Methods . . . 809
What are Corners, Boundaries, and Monte Carlo statistical models, obtained with conventional statistics? . . . 809
Introducing The Non-Parametric Statistics Concept . . . 809
What means typical component and limits components identification ? . . . 809
Non-Parametric Statistics Works on Gaussian Distribution . . . 810
Non-Parametric Statistics Works on Non-Gaussian Distributions . . . 810
Conventional Statistics vs. Non-Parametric Statisticswrap-up . . . 810
How does 'Non-Parametric' Statistics work? . . . 810
The Result . . . 812
An Example: Diode with measurements from DC to S-parameters . . . 813
Modeling Results . . . 815
A simple method to identifythe 'golden device' and the 'limits devices'has been proposed. . . . 815
Non-Parametric Boundary Analysis . . . 816
Motivation . . . 816
Circuit and System Design Methods . . . 816
What are Corners, Boundaries, and Monte Carlo? . . . 816
Non-Parametric Boundary Analysis . . . 816
Secondary Inputs . . . 817
Outputs . . . 817
Normalize the Dataset . . . 817
Estimating Density Using Nearest Neighbors Method . . . 818
Nominal and Boundary Point Selection . . . 818
Spatial Diversity---An Illustration . . . 818
Validating the Algorithm . . . 819
2-D Joint Gaussian Distribution with Nominal and Boundary Points . . . 819
Hypershell in 320-D Spherical Gaussian Distribution . . . 819
Conditional Density Histogram on 4-D Bimodal Gaussian Distribution . . . 819
Applying the Tool . . . 820
2-D Analysis of BSIM3 Data . . . 820
Analysis of Diode Data . . . 820
Conclusion . . . 820
Improving The Manufacturability Of Electronic Designs . . . 822

Contents

- *Manufacturing ICs* (iccapmhb)
- *Measurement for Device Modeling* (iccapmhb)

Manufacturing Integrated Circuits

Understanding the Semiconductor Process

The semiconductor manufacturing process comprises:

1. Creating a CMOS Inverter Circuit
2. Main Steps of the Semiconductor Process
3. Semiconductor Process Flow

Creating a CMOS Inverter Circuit

The process of making a CMOS inverter circuit is a complicated one. Figure 1 illustrates an example of a simple CMOS inverter circuit.

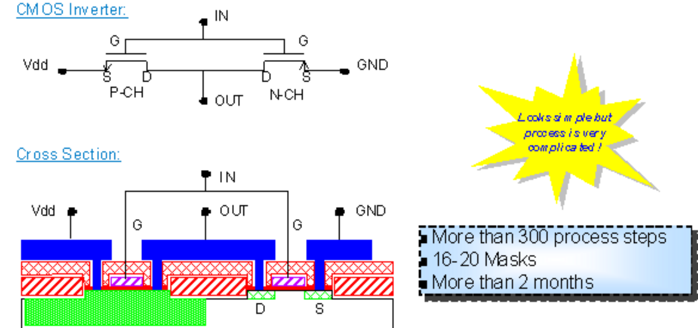
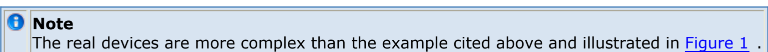


Figure 1 : A Simple CMOS Inverter Circuit

See the Cross section view in Figure 1 . The process of making a circuit consists of many layer depositions, doping, and etchings. Different elements can be created by slightly varying the parameters and device geometries.



Note
The real devices are more complex than the example cited above and illustrated in Figure 1 .

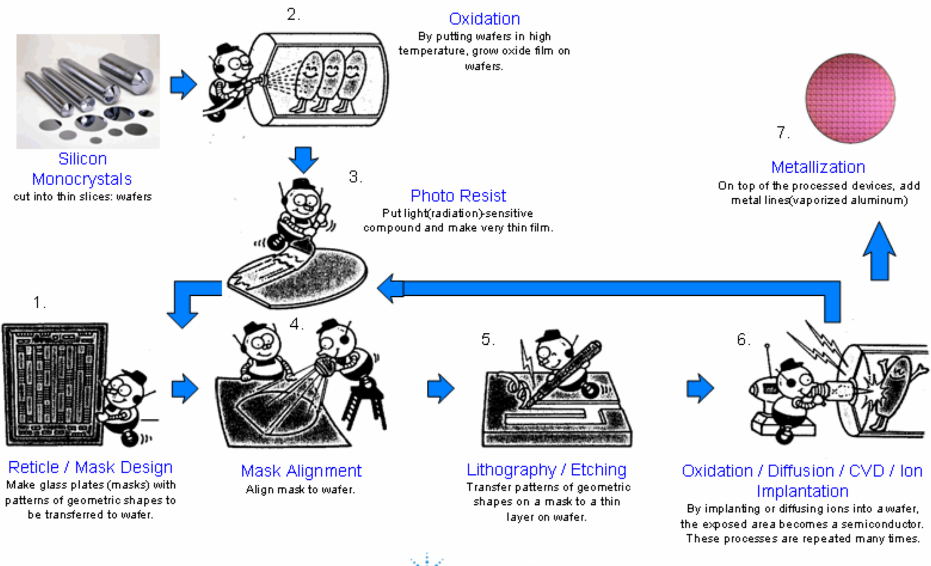
A fairly advanced CMOS process may contain over 300 steps using around 20 mask sets, and may take more than 2 months from wafer start to finish. Expenses for producing one wafer by a high volume process can exceed several $1000.

Main Steps of the Semiconductor Process

The semiconductor process consist of combining and repeating six main steps which are described in detail in the following section. These steps are complex and interact with each other. Refer Table 1: for further details.

Step	Semiconductor process	Description
1.	Photolithography	A very important process, which defines feature (transistor, diode etc.) size and geometry.
2.	Oxidation/Diffusion	This process creates insulating layers, or resistive (doped layers)at high temperature.
3.	Ion Implant	This process is more precise than the diffusion for doping layers with very tight control.
4.	Thin-film deposition	This process consists of low temperature deposition which protects the previous processes.
5.	Plasma Etch	It is a precise way of removing unwanted portions of layers
6.	Metallization	This process provides device interconnection using up to 6 layers. It is also used for making MIM capacitors, spiral inductors etc.

Semiconductor Process Flow

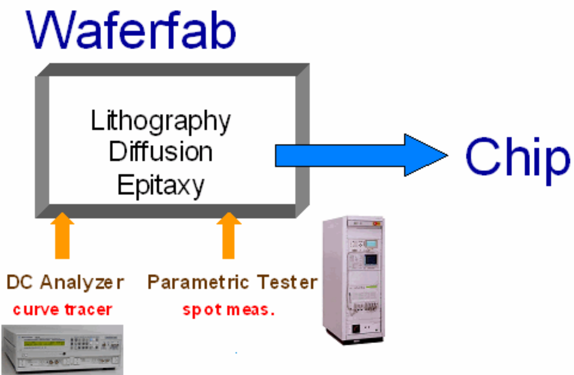


The following sequential steps describe the semiconductor process flow:

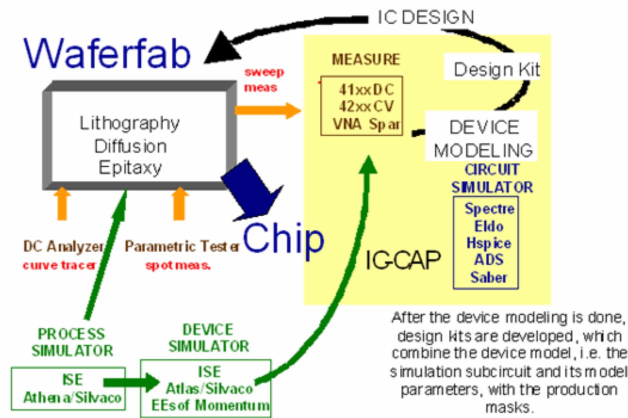
1. **Designing the Recticle Mask Set:** The recticle mask set is designed to transfer the design to the wafer. This concept is similar to traditional photography negative.
2. **Oxidation:** It creates insulating layers or resistive doped layers by growing an oxide film at high temperature.
3. **Photolithography:** It consists of the photo resist layer that creates the mask on the wafer so that the metal and oxide layers can be formed in the required places. The resist prevents the layers to be formed in certain places. It can be removed later.
4. **Mask Alignment:** This process aligns the mask to the wafer.
5. **Lithography:** It places the mask onto the wafer using the resist layer (Refer Photolithography). Etching removes the mask resist layer once the deposition or diffusion step is complete.
6. **Oxidation/ Diffusion/ Implantation:** This process introduces the ions into appropriate layers to create the semiconductor material in the exposed areas. The previous mask, lithography, etching, diffusion, and implant steps are repeated many times to create many layers of the devices.
7. **Metallization:** This process is repeated several times to create the conductive connections between devices and layers and to create some passive components like MIM capacitors, spiral inductors etc. The metal is usually deposited by evaporation.

Process and Device Modeling

This section provides an overview of the modeling process that is carried out during manufacturing (the process and device simulations)and for the individual components such as ADS, Spectre Spice simulations. It also describes how all this is assorted when the chip is finally designed.

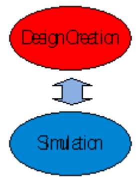


This is a typical scheme for control measurements during a chip production. Curve-tracer-like measurements are made by DC analyzers and fast spot measurements (single point measurements) are made by Parametric Testers.

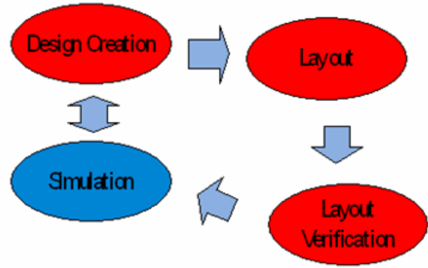


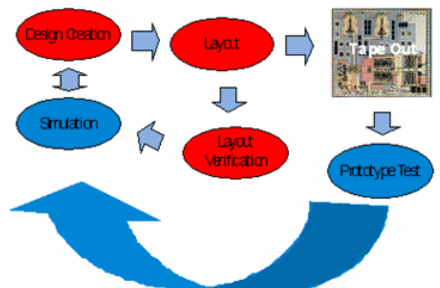
The process itself, i.e. the Diffusion and the Epitaxy, can also be modeled using Process Simulators. Once these simulators are calibrated to the process, the performance of devices can then be simulated using Device Simulators, which also take the photolithography into account. Either such simulation results can be imported into IC-CAP for modeling (using .mdm data files) or IC-CAP performs measurements directly on the components

When designing the chip, simulations of the complete chip circuit are performed based on these design kits



In a next step, linking the circuit schematic to a layout, the parasitic effects of that layout are taken into account and fed back into the simulation.
This way it is assured that also including these parasitic effect, the design will meet its specifications.





Slide 10

Now that we understand the semiconductor manufacturing process, and its modeling aspects, we will look at typical measurements.

IV (current-voltage) Characteristics

Measured parameter examples:

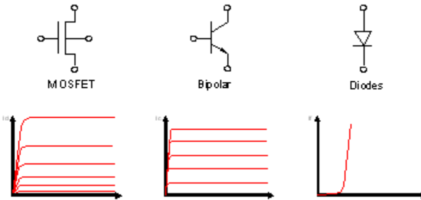
- Transistor gain
- Thresholds
- Delta L, Delta W
- Modeling Parameters

Process parameters monitored:

- Lithography
- Oxidation / Diffusion
- Ion Implantation
- Deposition
- Etching

Current voltage (or IV) testing is probably the best known type of parametric test. Since it is also the most important parametric test, we need to understand what kind of parameters are measured and which processes are evaluated by it. Measured parameters , like gain and threshold voltage are very popular. However more than just the standard curves are measured. To extract useful process information, variations of the standard measurement are often made. Of course the IV measurements relate back to the process parameters .

Active Region IV Curves
(typically used for device modeling):



The MOSFET (Metal Oxide Semiconductor - Field Effect Transistor) dominates today's technologies due to its performance (device geometries have shrunk over the past decades drastically, thus allowing to manufacture MOS transistors with transit frequencies in the GHz range). Another argument in favor of MOS is its lower cost compared to e.g. bipolar. Bipolar technology is also still popular for analog and high frequency digital applications.

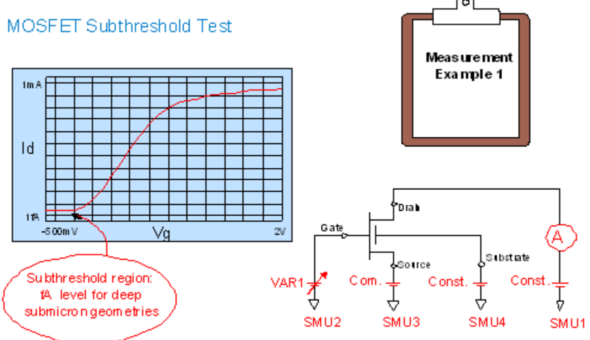
Transistor characteristics are critical measurements for the R&D engineer, and here are several popular curves which are measured. For example, here we have:

- IdVd of a MOSFET
- IcVc of a bipolar transistor
- diode's forward bias

These curves display device characteristics in an active state at relatively high current levels.

Low Level IV Curves

MOSFET Subthreshold Test



Measurement Example 1

Drain
Gate
Source
Substrate
VAR1
Com.
Const.
Const.
A
SMU2
SMU3
SMU4
SMU1

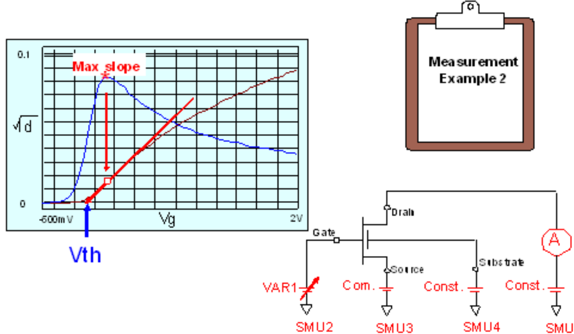
Importantly, low level IV sweeps must also be performed to measure leakage, threshold, gain and off state etc.

Let's look at one example, and see its application. A MOS transistor drain current vs. gate voltage (IdVg) curve is shown above. This characteristic is of great importance to the evaluation of semiconductor processes and device design. It allows the extraction of device gain and threshold voltage.

The lowest current region is called the "sub-threshold region". In this region the MOSFET is in the off state, where the current should be kept at a very low level to reduce power consumption and noise levels.

Typically the current is in the fA region, if processes are good. However, if the oxide is damaged during the process, the current may become much larger. Defects due to improper oxidation processes can be revealed only by an accurate monitoring of the sub-threshold current.

Threshold Voltage (Vth)



Measurement Example 2

Drain
Gate
Source
Substrate
VAR1
Com.
Const.
Const.
A
SMU2
SMU3
SMU4
SMU1

The threshold voltage (or Vth) is not only another important device parameter, it is also a very important parameter for circuit simulation. Vth can be controlled by changing doping profile and oxide thickness. Alternatively, it can be applied to evaluate processes to control doping concentration and oxide thickness.

There are several common ways to determine the threshold voltage. This example shows the saturation region threshold voltage. The IdVg curve is obtained using the measurement setup shown, and the square root IdVg curve is drawn. This curve is derivated vs. vg to find the maximum slope. At this maximum slope, a tangential line is drawn to SQRT(id), and the threshold voltage is then extracted as the X axis intercept of this tangent.

Name	Value
IDOFF	1.08009E-011
IGOFF	0
ISOFF	-9.63797E-012
IBOFF	-1.16291E-012
____________1	
IDLIN	3.03792E-006
IDLIN_BB	2.50114E-006
IDSAT	3.83689E-005
IDSAT_BB	2.78835E-005
____________2	
VTLIN	1.61396
NVTLIN	5.98765
ISVTLIN	5.93618E-012
VTLIN_BB	1.87587
NVTLIN_BB	5.93923
ISVTLIN_BB	9.40393E-013
VTSAT	1.53884
NVTSAT	6.02311
ISVTSAT	1.00489E-011
VTSAT_BB	1.7981
NVTSAT_BB	6.05216
ISVTSAT_BB	1.81961E-012
____________3	
GM	5.96973E-006
GD	2.07511E-011

Process Control Monitoring

Spot Measurements

Drain
Gate
Source
Substrate
A
Const.
Const.
Const.
Const.
SMU2
SMU3
SMU4
SMU1

Other measurements performed in a production environment are so-called spot-measurements.
As an example, instead of measuring the complete idvd output characteristics, only IDLIN (vg=vg_max, vd=50mV, vb=0V) and IDSAT (vg=vg_max, vd=vd_max, vb=0V) are measured. Commonly, applying the max. neg. Bulk voltage, IDLIN_BB (vg=vg_max, vd=50mV, vb=-vBB) and
IDSAT (vg=vg_max, vd=vd_max, vb=-vBB) are measured too.

These high-speed measurements are performed on as many as available devices on the wafer, and displayed vs. their distribution. This allows to easily- check the performance and stability of the process- check the variance of the process

	Resistivity	**Continuity/Bridging**	**Leakage/Breakdown**	**Capacitance/ CV**	**IV Curves**
Typical Measurement	• Poly resistors • Metal contacts • Current capability • Substrate resistors	• Contact resistance • Via continuity • Inter-metal bridging	• Breakdown voltage • Oxide damage • Leakage current	• Capacitance • CV curve	• Threshold voltage • Drain current (On, off) • Max Gain (Gmmax) • Punchthrough voltage
Features Maintained	• Film thickness *Doping concentration • Film purity • Line width • Mask overlay	• Particles / defects in films • Incomplete etches • Defect densities • Dielectric integrity • Contact integrity	• Ion implantation damage • Isolation-edge effects • Contamination • Dielectric integrity	• Film thickness • Doping concentration and profiles • Carrier lifetimes • Thresholds	• Transistor gain • Thresholds
Processes Monitored	• Oxidation • Deposition • Ion implantation • Metallisation	• Lithography • Deposition • Etch • Metallisation	• Oxidation • Diffusion • Ion implantation • Deposition	• Oxidation • Diffusion • Ion implantation • Metallisation	• Lithography • Oxidation • Diffusion • Ion implantation • Etch
Structures					

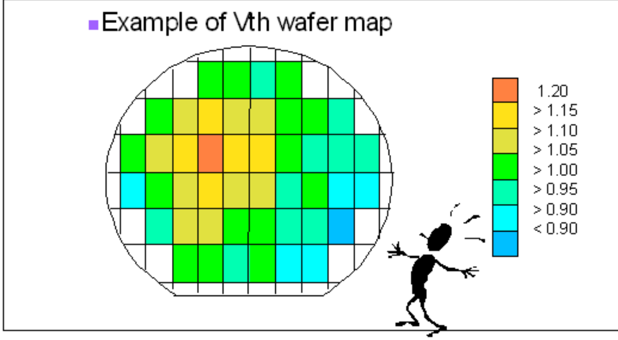
This summary table shows the relationship between typical parametric tests and the processes monitored by them. The basic test types are: resistivity, continuity/bridging, leakage/breakdown, capacitance/CV, and DC/IV curves. By using a combination of these tests, nearly all process aspects can be covered.

There is more information about the tests in the Appendix.

Wafer Map

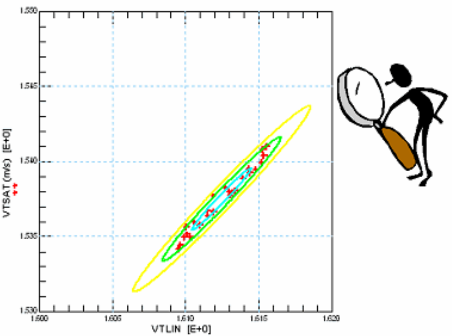
Different types of analysis can be performed on the results obtained, in order to understand what is happening with the devices and process.

Here the uniformity of a measurement across the wafer is analyzed by drawing a wafer map for the Threshold Voltage VTH. Wafer maps for the other PCM measurements (ID_LIN, ID_SAT etc.) are obtained in the same way.



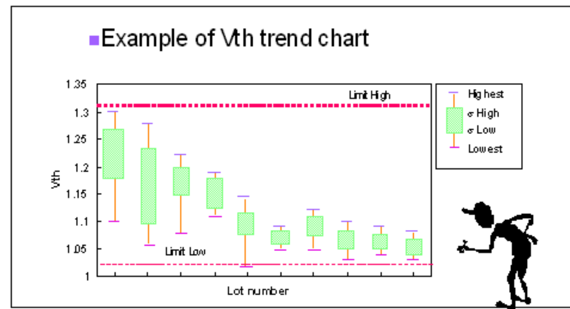
Correlation between process parameters and measurement results will allow you to find the best process condition for producing the best end product.

Correlation Data



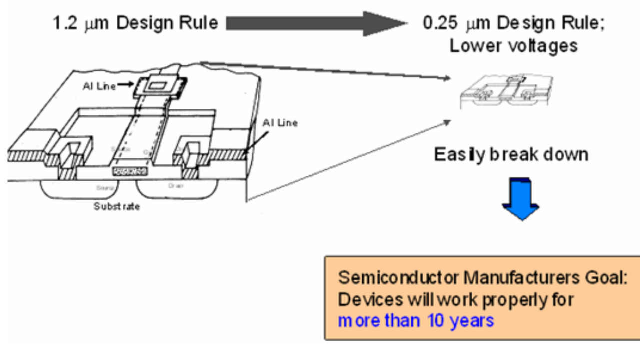
The trend chart is critical for capturing potential process problems. Key test parameters must stay within certain limits to ensure that functional devices are made. Parameter values near to or over the limits can trigger urgent action for process engineers.

Trend Chart



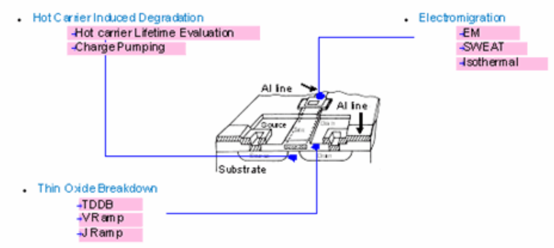
In addition to the standard traditional IV and CV measurements and curves, reliability testing is very important.

When Devices are Scaled Down ....



MOS transistors operate with faster and faster speed, and this means that they must become smaller and smaller, and therefore, operate at lower voltages and currents. However this also means that they can be much more easily degraded, damaged or even broken. Additionally, this degradation may happen over long periods of time, even years. Therefore, lifetime and reliability is a big issue, and the general goal is to ensure a minimum of a 10 year lifetime of devices.
Reliability uses Parametric tests for evaluation, and therefore parametric instruments are important for this application.

Three problems of Deep Sub-micron Devices:

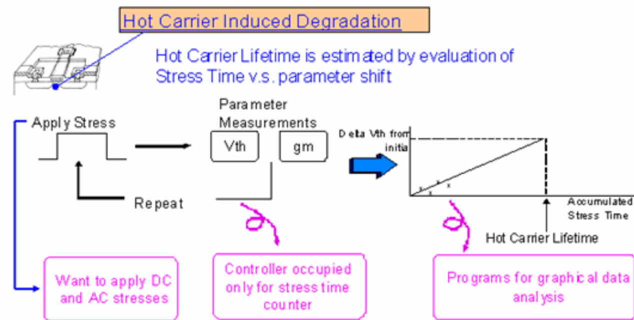


WLR (Wafer Level Reliability) is used to perform reliability evaluation at wafer level.
It needs special test structures and test methods to perform in very short time.

The following are three main areas of concern in Reliability Testing:

- Hot Carrier Injection
- Oxide Breakdown
- Electromigration

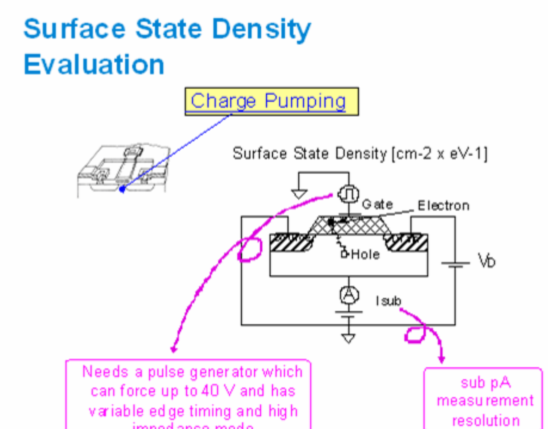
We will look at these in more detail on the next few slides, but some of the common measurement names are also listed here so that you can identify them more easily. These are tests such as Charge Pumping , TDDB , V Ramp , and SWEAT tests.
When performed on wafer the reliability tests are called WLR , ie Wafer Level Reliability.



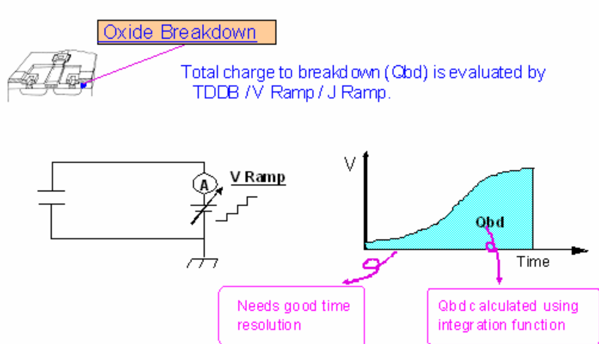
Hot carrier injection happens when small gate sizes lead to high electric fields close to the drain terminal, creating highly energized carriers known as "Hot Carriers". These are injected into the oxide layer and trapped, causing degradation of the device due to the trapped charge.

This effect can cause device gain and speed degradation and finally malfunctions.

Evaluation is done by repeatedly applying a DC or AC stress for a defined time, and then re-measuring the appropriate parameters and plotting the change vs. accumulated stress .

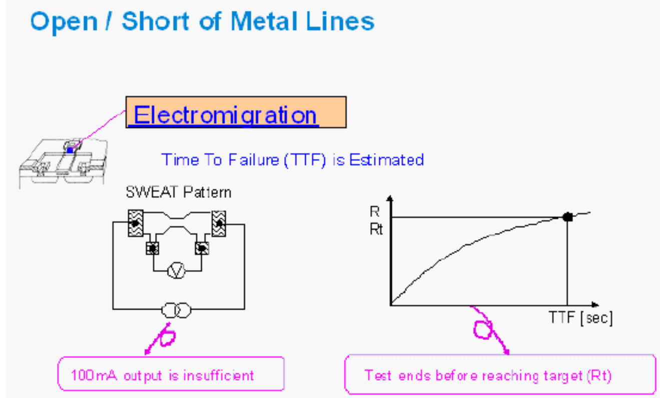


The trapped charges can also be evaluated using Charge Pumping techniques, applying a pulse to the Gate. DC analyzers therefore feature a Pulse Generator option with the necessary functions .

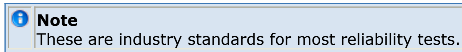


Thin Oxide Breakdown is caused by a combination of the thin oxide layer and imperfections or defects in the oxide itself. The defects cause a higher current to flow, which then causes further damage, and so the oxide breaks down.

Here the V Ramp test applies an increasing voltage to the oxide and accurately measures the time and voltage at breakdown. It can also be evaluated using I Ramp (current ramp), or TDDB which is a constant current or constant voltage Time Dependant Dielectric Breakdown.



Electromigration is again aggravated by small sizes. If the metal interconnects are too small, high current densities induce the migration of metal atoms. Then voids or hillocks are formed, which results in increased resistance, opens circuits, or shorts to adjacent structures.
SWEAT and Isothermal methods are used for evaluating electromigration in shorter times.

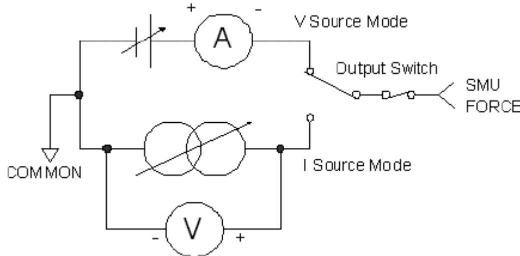


Now let's move to the measurement instrument requirements which result from all the test parameters which we have discussed.

Understanding the SMU

→ **SMU = Source Monitor (Measurement) Unit**

Simplified SMU schematic:

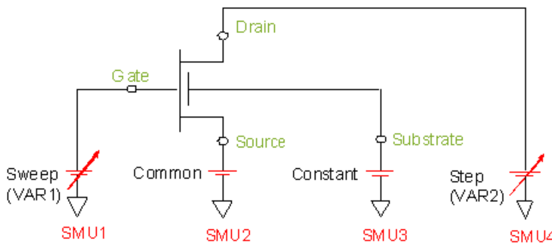


Firstly, let's make sure we understand what the SMU is. SMU stands for Source Monitor Unit, or Source Measurement Unit.
The SMU has 3 modes of operation:
Voltage source (or force) and current monitor mode.
Current source and voltage monitor mode.
Source common mode.
Each SMU can operate in any combination of sink or source, allowing full 4-quadrant (negative and positive) source and measure capability.

SMUs are the fundamental hardware component of parametric test. Try to imagine what you would have to do to duplicate SMU functionality using discrete equipment (eg voltmeters, ammeters, power supplies, etc.). Just appreciate the powerful simplicity of the SMU.

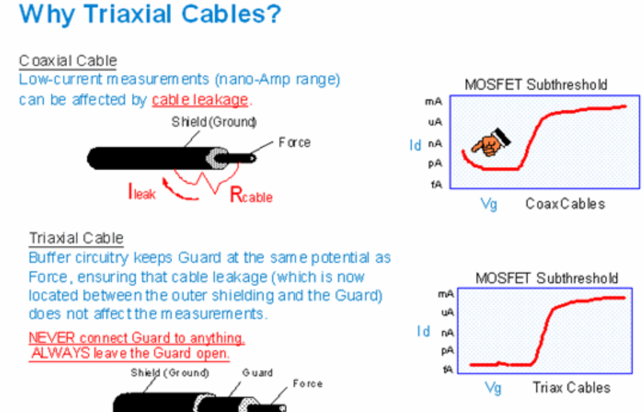
Why 4 SMU's?

- Four terminal MOS device
- Only one contacting is required to perform all required tests of the MOS transistor
- Speeding up testing



Why a DC Analyzer has 4 SMUs?

Well, transistors are actually 4 terminal devices . As well as the Gate, Drain and Source, the substrate under the transistor must be biased to a known voltage in order for the device to function properly. By assigning one SMU to each terminal, any measurement can be made without having to change the device connections. Thus, the minimum number of SMUs required to make effective MOS transistor measurements is 4.

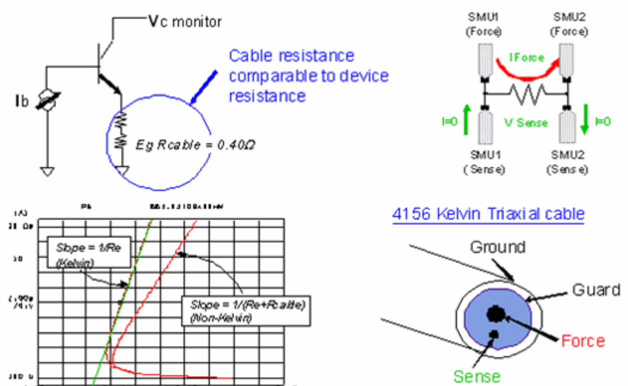


Importance of Triaxial Cables
In a standard coaxial cable (where the signal core is surrounded by a grounded shield) there is too much leakage to make effective low current measurements, since there is actually a very small current flowing through the insulator. Also, parasitic capacitance slows down the measurement speed.

To make truly accurate low current measurements, triaxial cables must be used. In triaxial cables a guard is added around the force core line to isolate it from the grounded shield. A special buffer circuitry ensures that the Guard and Force lines are at the same voltage potential, thus removing the effects of any stray leakage and capacitance. Triaxial cables are required for low current measurements (ie below 1nA).

Remember that the Guard is always at Force voltage, and therefore must NEVER be connected to ground. Always leave it open.

The Agilent DC Analyzers are designed for accuracy, so they only support triaxial cables.

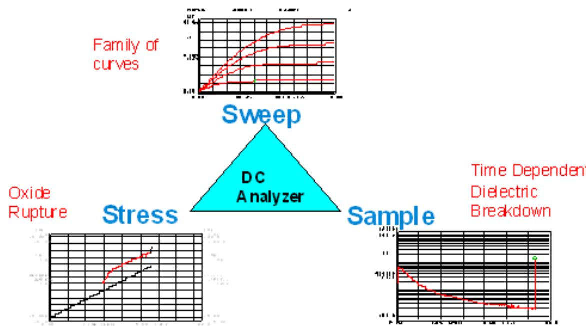


Kelvin Measurements, otherwise known as 4-wire measurements, are required for low resistance measurements, when the resistance being measured is comparable to the resistance of the cables . This is therefore also true when high currents are applied. The graph illustrates the error which can occur when not using Kelvin cables.

The concept is that 2 wires do the normal force and measure , but the additional 2 wires sense the voltage at the device and correct the SMU output to compensate for any voltage drop in the cables. Consequently the set conditions are applied at the device, rather than at the SMU internally, and the measurement is more accurate.

So Kelvin measurements are done for low device resistance or high current, whereas triaxial cables are necessary for low current measurements.

Agilent DC Analyzers have Several Operating Modes



In addition to the traditional sweep functionality, the modern Agilent DC ANalyzers have powerful sampling and stress capabilities, as well important modes such as Standby and Knob Sweep.

A Spot measurement forces a single voltage or current and measuring during the force. However, Sweep measurements involve incrementing the forced voltage or current for a certain number of steps, and measuring at each point. The user can also define a secondary sweep to create families of curves.

Sampling measurements involve repeating a constant (rather than incrementing) forced value over period of time, possibly until some pre-defined stop condition is achieved.

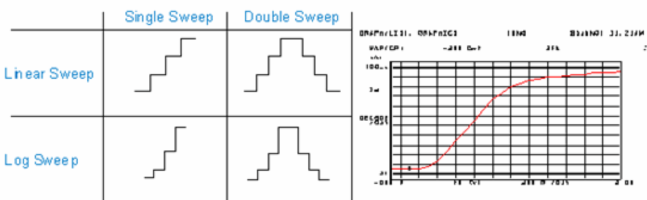
Stress involves forcing a DC or AC signal across a device, and measuring after the stress application has completed. The Agilent DC analyzers offer repeated stress/measure cycles either manually or automatically. Note that the stress settings are completely independent of the measurement settings.

In a sweep measurement, voltages or currents are increasingly swept between each measurement point to create a device curve . The sweep mode has a lot of flexibility. It can be set up for single or double sweeps, in combination with linear or log sweeps. Additionally, an SMU can also be put into Pulse Mode where the output of the pulsed source is made as a typically 1us pulse at each step rather than being fixed for that step. This feature is valuable in cases where a normal sweep would cause too much device self-heating.

With Standby mode SMUs stay at their output value, even between measurements so that devices can remain powered up constantly.

Knob sweep mode allows Curve Tracer like functionality by repeatedly sweeping up to the value set by the front panel knob.

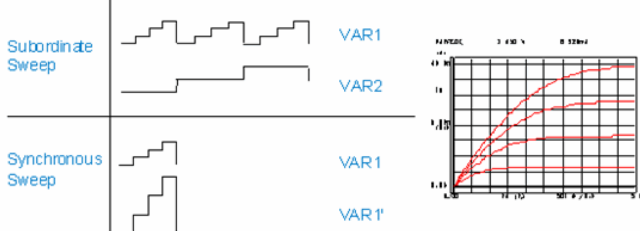
Sweep Measurement Basics



- **Auto-analysis** - calculates results automatically
- **Pulse Mode** - reduces device heating
- **Standby Mode** - devices remain powered on
- **Knob Sweep** - curve tracer like function

As well as the primary sweep (called VAR1), the user can define a secondary step source (VAR2) which will increment in value each time the primary sweep source (VAR1) completes its sweep sequence.
Furthermore, the user can define a secondary synchronous source (VAR1'), which tracks the primary sweep source (VAR1) timing, but with a user-determined ratio and offset. This is useful for characteristics such as Gummel plots.

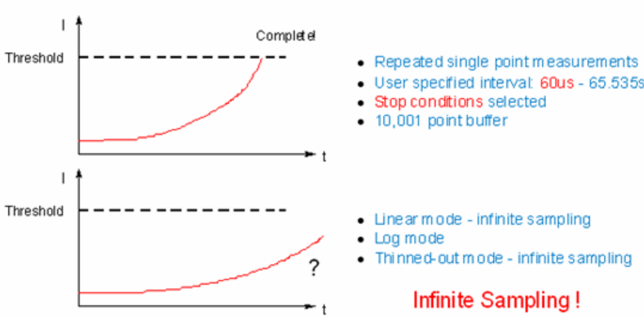
Additional Sweep Functions



Sampling is a measurement process done over time , where the measurement is repeated at specified time intervals. The voltage and current settings are not varied for sampling mode. Many sophisticated features greatly simplify the task of taking repeated time measurements.

As well as setting the time interval, the user can set either a given number of measurement points, or no limit, or continue until some pre-defined stop condition is reached, such as reaching a certain current.

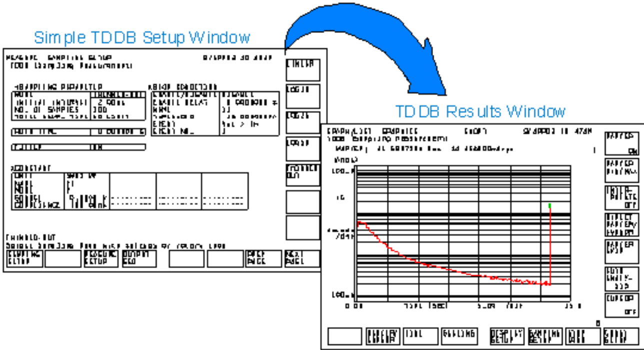
Sampling Measurement Basics



Here is a practical example of the sampling measurement. Earlier we discussed the TDDB measurement, which determines how long it takes to rupture the insulating dielectric layer. This gives a measure of the reliability of the dielectric.

In this case, only 300 sample points are retained, and thinned-out sampling is used, with an interval of 2ms. An important point about thinned-out sampling is that the sample points near the actual event are still spaced at the specified 2ms spacing. Points earlier on are the ones which get "thinned-out", so that there is still fine sampling near the event of interest. This has many obvious benefits, and illustrates a key difference between this instrument and a sampling oscilloscope.

TDDB - Time Dependent Dielectric Breakdown



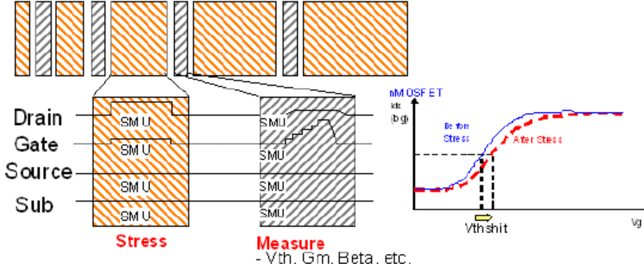
Now let's look at the Stress function. A stress is simply a DC or an AC signal applied to the device for a specified amount of time , however it may be at a higher voltage or current than the device would normally work at. The purpose is to see how the device behaviour changes after a certain amount of stress , possibly beyond it's normal operating limits. It gives an indication about how robust the device is, and of it's lifetime and reliability, without waiting the years it would normally take to see an actual device fail.

The stress menu is a separate menu from the sweep and sampling measurement menus, since typically the user would want to configure a stress as a measurement cycle which is repeated over and over. The beauty of the stress function is that the stress settings for the defined stress sources do not have any relationship to the measurement settings for those same resources; they are totally independent.

Up to 4 resources (ie SMUs, VSUs, or PGUs) can be defined as DC or AC stress sources as appropriate. For example, an SMU could be forcing current in stress mode but be forcing voltage in measure mode. This gives the user a great deal of flexibility in setting up stress/measure routines.

Stress Function Basics

- Alternatively measure and stress a device
- Stressing = outputting from SMU/VSU/PGU for a specified time



The stress feature is somewhat of an advanced topic, but it is used most often for reliability testing. A major benefit is that the stress mode is built into the DC Analyzer itself, and allows use of the built-in PGU channels.
We just mentioned that the stress settings are independent of the measure settings. This is critical for practical flexibility. Importantly, the stress can be applied in precise amounts , and accurately recorded so the reliability data is actually meaningful.

Benefits of the Stress Function integrated in the DC Analyzer

Built-in Feature
- No external power supplies or pulse generators needed

Greatly Improved Measurement Flexibility
- Stress settings separate from measure settings

Better Test Accuracy
- Stress applied in precise amounts; stress time accurately recorded
- Stress time - 5000us to 365days
- PGU pulse count - 1 to 65535 pulses

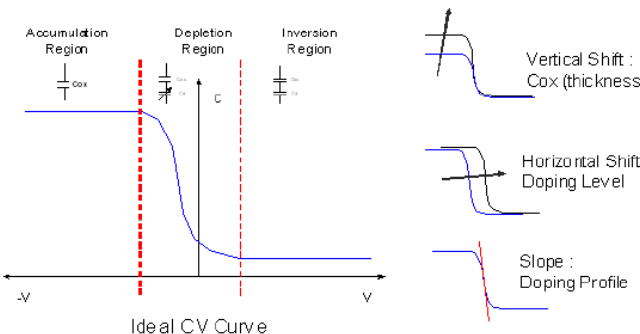
Automated Reliability Testing
- Users can automate the stress/measure procedure
- External gate trigger

The plot above depicts the capacitance of a MOS transistor Gate oxide. It is a bias-dependent capacitor.
A vertical displacement of the CV curve relates directly to the oxide thickness (Cox).
A horizontal shift relates to the doping concentration level.

The slope of the transition part of the curve tells us about the doping profile. The steeper the slope the more abrupt the change in doping level with depth.

CV measurement can reveal :
Doping concentrations and profiles
Oxide and metal impurities
Film thickness
Carrier lifetime
Thresholds

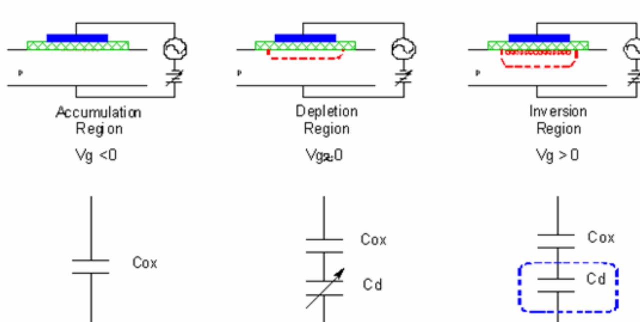
Monitored processes:
Oxidation / Diffusion
Ion Implantation
Metallisation



As the gate bias is made less negative, fewer holes are attracted to it. The substrate works similarly to metal. In consequence, the MOS Gate oxide capacitance value will be equal to the oxide capacitor Cox.
At some point (usually around Vg = 0V), a depletion region starts to build-up below the gate area because the holes are repelled from the surface between the oxide and the silicon. As the bias voltage increases, this depletion region continues to grow. In this state, the MOS capacitance consists of Cox and depletion layer capacitance Cd. Therefore the MOS capacitance will decrease.

As the bias increases further, electrons (minority carriers) will be attracted to the region under the Gate. It is said that this region has become inverted. Any further increase in bias will not change the width of the depletion region, it will only add more electrons to the inverted regions. The capacitance measured in this condition will depend on the method used to conduct the measurement.

If a high frequency (>100kHz) measurement is made, the minority carriers (electrons) cannot react fast enough to the test signal so the capacitor will look like the depletion region capacitance in series with the oxide capacitance. Since the depletion region width is constant in inversion, the capacitance is a constant.



If a low frequency capacitance measurement is used, an entirely different result is obtained. Iin this case, the minority carriers under the gate region can keep up with the test signal so the capacitance is again just the oxide capacitance (Cox).

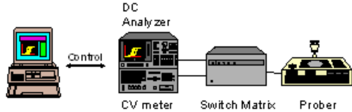
The common method for low frequency CV measurement is the Quasi-Static method, where the MOS gate current responses to a slow linear DC voltage ramp, yielding the CV characteristics. The capacitance value can be calculated by
C = Ig / (dV / dt)
where Ig is the Gate current.

The Quasi-Static method has become one of the more common methods of looking at interface trapped charge. A shift to the right in a n-channel process can be due to sodium

contamination (Na+ ion trapping). This increases a transistors threshold voltage. The slope of the rising edge relates directly to Vth if the oxide characteristics are the same as the final transistor.

Why a Switching Matrix?

Create a mini "Test System"

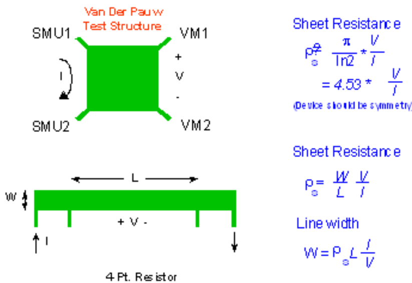


Our last topic regarding measurement instrument requirements is about the switch matrix .

The equipment required to do accurate current-voltage (ie IV) measurements and high frequency capacitance measurements (ie CV) is different. So without some means to switch between the two, users had to spend a great deal of time changing cables every time they wanted to change between IV and CV measurements. This obviously also applies when other instruments are used additionally (eg HF, PGUs or DVMs).

Very often, the test device structures are organized on the wafer into groups, and probe cards with e.g. 24 pins are applied, rather than manual needles. A swithcin matrix allows to switch between the different devices within a group without re-cabling.

Resistivity



Resistivity measurements are used to evaluate processes related to creation of conductive layers and lines or to evaluate mask alignment accuracy.
Resistivity of a film can reveal:

- The Film thickness, cross-section, purity
- Doping concentration and uniformity
- Line widths
- Mask overlay

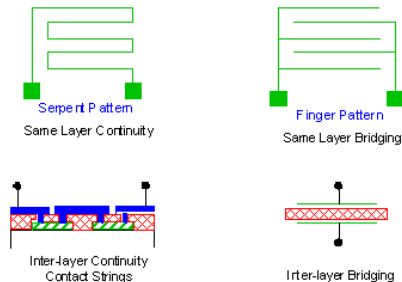
Monitored processes:

- Oxidation / diffusion
- Deposition
- Ion implantation
- Metallisation

A Van der Pauw measurement is done by passing a current through two adjacent corners of a square and measuring the voltage drop across the remaining two electrodes. The sheet resistivity is calculated by a very simple formula. The main advantage of the Van der Pauw structure is that it is symmetrical. The leads can be switched and the measurement can be performed again. If the obtained results are not equal, non-uniformities exist within the films and can be detected.

A simpler structure to measure the resistivity is a basic 4-point resistor. The resistivity can easily be calculated if the dimensions are known. On the other hand, the line width can be calculated if the sheet resistance is known.

Continuity/Bridging



These tests are used to determine whether the conductive lines are correctly open or correctly short. These tests provide a powerful tool for investigating defect levels in processes related to the creation of conductive lines and contacts.

Continuity and bridging tests can reveal:

- Particles / defects in film
- Incomplete etches
- Defect densities
- Dielectric integrity
- Contact integrity

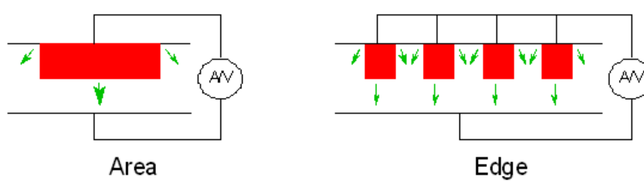
Monitored processes:

- Lithography
- Deposition
- Etching
- Metallisation

Measurements:
Various structures are used to measure continuity and bridging. A set of closely spaced interigitated fingers provides a good tool for detecting shorts. A similar structure is used to measure continuity, here it is a long continuous line.

Bridging and continuity are also looked for between layers. A simple parallel plate structure is used to look for dielectric problems. Contact strings are used to look for continuity of the contacts between conducting layers. Very often these contact strings have as many as 10000 contacts or more.

Leakage / BreakdownIsolation Edge Effects



Dielectrics and diffused junctions can be very sensitive to contamination or defects introduced during processing. Because of this sensitivity, the leakage current and/or breakdown voltage are key process parameters. The reverse-bias characteristics of a junction is often more important than the forward-biased one.

Reverse-bias testing can reveal:

- Ion implant damage
- Isolation-edge effects
- Contamination
- Dielectric integrity

Monitored processes:

- Oxidation / diffusion
- Ion implantation
- Deposition
- Isolation edge effects

When the leakage current exceeds a predetermined level, the production process should be examined as a cause. One method for how to isolate process problems is shown above. By using two structures with the same area but different perimeters, the contribution to the leakage currents from edge effects can be isolated from the one caused by area effects. These two sources of leakage currents often have different causes.

Device Modeling Measurement Tutorial

The trend to higher integration and higher transmission speed challenges modeling engineers to develop accurate device models up to the Gigahertz range. An absolute prerequisite for achieving this goal are reliable measurements, which have to be checked for data consistency and plausibility.
This is especially true for radio-frequency (RF, >100MHz) and microwave (GHz) measurements, and also for checking and verifying the applied de-embedding techniques. If there are (hidden) problems with the measurement data, RF characterizations can become quite time consuming, with a lot of guesswork and adhoc judgments, and, basically, frustrating and not correct.
If, however, the underlying measurements are flawless and consistent, and provided the applied the models are understood well, RF characterization and device modeling becomes very effective and provides accurate design kits which will satisfy the chip designer's main goal, right the fist time.

Basics of DC and AC Characterization of Semiconductors

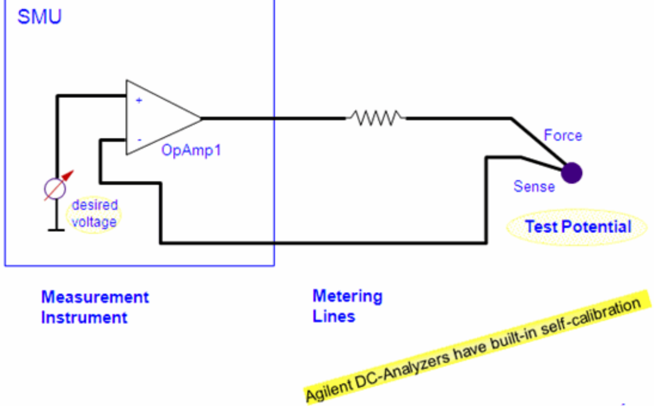
- Basics of device measurement and modeling techniques from DC to RF
- Special aspects of network analyzer calibration
- De-embedding and required dummy structures

While the characterization of electronic components in the DC domain is relatively simple and only requires a Voltmeter and an Ampèremeter, the frequency performance of the device is affected by magnitude dependence and phase shift of the currents and voltages. Furthermore, non-linearities will lead to a spectrum of frequencies, although the device is only stimulated with a single, sinusoidal frequency.
Last not least, inevitable capacitive and inductive parasitics, with values close to those of the very device under test (DUT), will contribute to the measurements and degrade the measured performance of the 'inner' DUT.

In this presentation, we will go step by step through the individual characterization issues and develop measurement strategies which will provide the base of accurate device characterizations.

Modeling of a nonlinear component always begins with the characterization of its DC performance.
Instead of power supplies, precision DC parametric analyzers with source-monitor-unit (SMU) plugins are applied. They offer current ranges from femto-Ampère (1E-15) to several Ampères, and voltage potentials from micro-Volt to hundreds of Volts. This allows to fully characterize the DUT (device under test) in all four I-V quadrants. I.e. forward and reverse currents and voltages, are measured with the same SMU unit.
Usually, in case of a transistor, all 4 terminals (including substrate) are connected to individual SMUs in order to avoid recabling during the forward and reverse measurements.

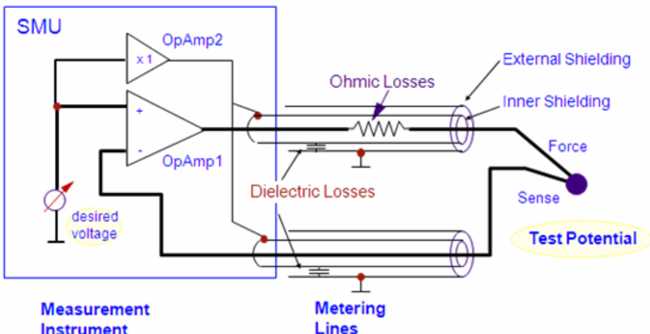
PlotOption Access Name and Value Examples



SMUs apply a Kelvin measurement to avoid parasitic series resistances. This measurement procedure, also known as the four-wire method, consists of a stimulating line (Force) with a second one in parallel (Sense) for every pin of the DUT. The slide above illustrates this. Ohmic losses on the Force line are eliminated by the operational amplifier (OpAmp1) in

voltage follower mode. This means this OpAmp1 output will exhibit a somewhat higher voltage than the desired test voltage at the DUT, because the test current generates some ohmic losses along the Force line. The Sense line, connected to the minus input of the OpAmp1, assures that the DUT is biased with exactly the desired test voltage.

The Guarding (Triax Cable) Principle for Fast and Accurate Measurements

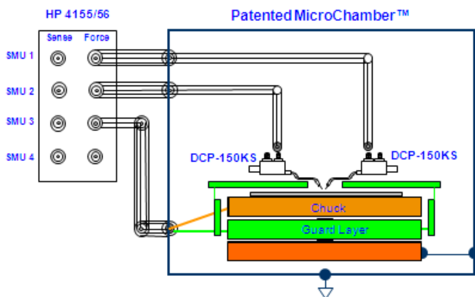


While the Kelvin method compensates the DC errors, it does not cover dynamic DC measurement problems. For example, to avoid external electro-magnetic influences, both the Force and Sense cables are shielded. But such cable shieldings exhibit parasitic capacitance. Due to charging problems, these capacitances will affect the measurement speed and accuracy of our Kelvin measurement.
As a simple example: assume we want to measure the reverse characteristics of a semiconductor diode. This means we need measure very low currents. Before the voltage steps to e.g. -20 V, the quiescent voltage at the diode is zero. That is, the cable capacitors are not charged. When the negative voltage step occurs, these capacitances have to be charged, and the required current is provided by the OpAmp1. This could lead to either a mis-measurement (DUT current plus charging current) or a delay in the triggering of the actual current measurement (by some intelligent firmware in our measurement).
To solve this problem, an extra inner shielding is applied between the hot metering lines and the outer cable shielding, called 'Guard'. This extra shielding is connected to a separate, second OpAmp2 which follows exactly the value of the desired test voltage. Now it is this auxiliary OpAmp2 which supplies the charging current for the test cables, while the main OpAmp1 can start current measurements without being affected by this charging problem. That is, the inner measurement loop does not see the charging problem any more.
Of course the point where Force and Sense are tied together must be as close as possible to the DUT. In case they aren't connected, an internal 10kOhm resistor at the output of the SMU acts as the Kelvin point. Another important fact is that the Guard contact should *never* be connected to Force or Sense. Otherwise, the inner loop OpAmp1 of the SMU would measure the DUT current *plus* the charging current of the auxiliary, second OpAmp2!
Calibration:
In order to maintain the DC measurement accuracy, SMUs perform periodically an auto-calibration. This means that the SMU disconnects its outputs from the DUT, measures possible offset voltages and currents and corrects it. This type of calibration does not require any action from the user.

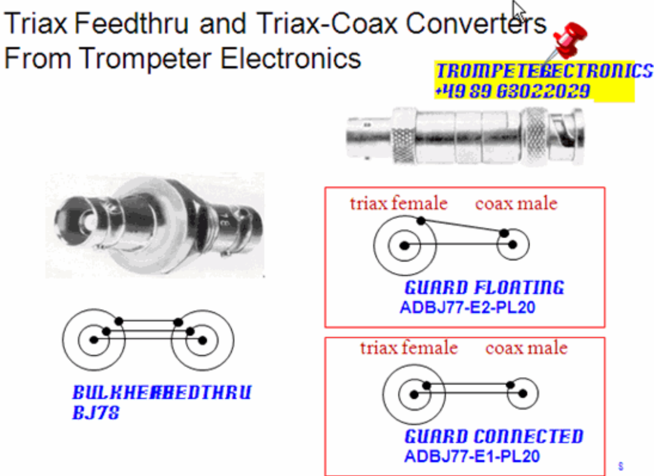
Measurement Environment

Complete triaxial system

Typical on-wafer IV measurement setups require probes and chuck bias. A careful shielding completely surrounds the wafer with guard.

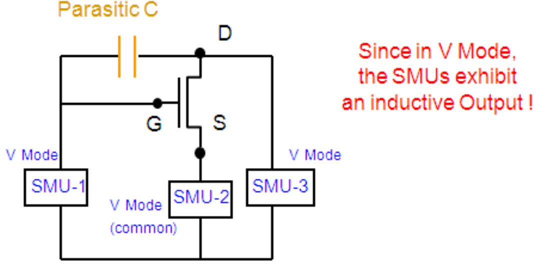


Triax Feedthru and Triax-Coax Converters From Trompeter Electronics



When connecting the Triax connectors, and particularly when applying an adapter, make sure the Guard is left open and floating !

Oscillations during DC measurements



Therefore, the measurement environment
is equivalent to a Hartley oscillator !!

Under certain circumstances, SMUs may oscillate. Super-beta, wide- bandwidth bipolar transistors are especially susceptible to oscillation. But also GaAs transistors.
Oscillation can become a problem when using older DC probe needles. A common way to avoid such oscillations is using HF probes (Ground-Signal-Ground GSG probes), or the newer shielded DC needles.
If you have to live with your DC needles, and when oscillations occur, here some ideas on causes for such oscillations:

1. SMU Induced oscillations. These oscillations occur when an inductive load is connected to the SMU's output. The SMU has often an inductive load because the output impedance of the SMU is inductive during V-Mode operation and usually several SMUs are indirectly connected together through the DUT.
2. Oscillation due to strays. The measurement system including the DUT, stray capacitance and residual inductance of the connection cables, switching matrix, probe card and/or test fixture can be recognized as an oscillation circuit.
   The oscillation detector of SMU may not detect this type of oscillation. This is understandable when thinking of the SMU output as a low-pass filter, while the oscillation frequency may be in the 100MHz ... to several GHz range. Also, if like in most cases the oscillation is located at the DUT, the SMU itself cannot do anything to prevent this oscillations, since the cables are long compared to the oscillation wavelength.
   As a general rule: the oscillation has to be avoided where it happens.

Related to the example of the slide above, we should be aware of the total circuit: The SMUs connected to the MOSFET's gate and drain are operating in the V-Mode. Since SMUs typically appear to be inductive in V-mode, this makes this configuration equivalent to a Hartley oscillator !
The SMU may oscillate if an unusually large inductance is connected to it. This could occur if the DUT is a superbeta transistor (big hFE) and the SMU connected to the emitter is set to one of the low current ranges.

For more details on conditions for oscillations, refer to the application note 356 -1 (publication number Agilent 5950-2954)

Methods to eliminate oscillations

- Ferrite beads placed as close as possible to the DUT to prevent high frequency oscillations (Part Number 9170-0029)
- Keep cables as short as possible - Use High quality cable (low cable inductance)

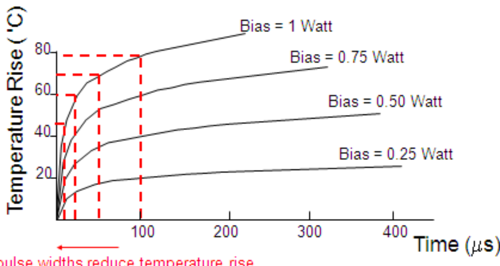
- In some cases, more than 1 Ferrite Bead may be required!

There are some effective methods to eliminate oscillations:

1. For FETs, add resistive ferrite beads as close as possible to the gate.
2. For bipolar transistors, add resistive ferrite beads as close as possible to the base or emitter.
3. Keep cables as short as possible. Long cables cause oscillation because of their large inductance.

In some cases, it is necessary to use more than one ferrite beads (Agilent Part Number 9170-0029).
Or, apply shielded DC probes, even better apply RF probes.

Self-heating during DC measurements



Device Temperature vs. Time at different bias levels

As a general observation, currents above ~25mA will cause self-heating effects with transistors. Rule of thumb: 25mA with typically 2V equals already 50mW for that tiny transistor on the wafer!
Self-heating can become a severe problem with device characterization, because, when self-heating occurs, the measurement results depends on the measurement speed!
This can be verified when e.g. measuring a Gummel plot for a bipolar transistor, once sweeping from low to high voltages different measurement speeds, and then sweeping from high to low voltages. Calculate beta out of your different measurements.If self-heating occured, you will get as many beta curves as you have performed measurements!

The plot in the slide above gives the chip temperature increase of a packaged transistor as a function of the pulse width and the applied bias power. The pulse period is 1s. As can be seen, self-heating can only be avoided when applying very short, pulsed measurements, below 1us pulse width. Such pulsed measurement systems, also included pulsed S-parameter measurements, are commercially available.
However, such systems are quite complex and expensive. Therefore, if you have to live with self-heating, make sure your device suffers always from the same self-heating. This means in general, apply the slowest measurement speed for you're your DC measurements, since your biased network measurements will be slow as well, and self-heating *will* definitively occur there !

What to do when self-heating becomes an issue for your device modeling:

- from the slide above, we identify a min. pulse width of ~1us or below. This means the *DC analyzers in pulse mode cannot be used*(pulses are 100us or longer)
- use the IC-CAP pulsed modeling system
- live with self-heating, but make sure your data are consistent:
  - all DC measurements are performed with the same self-heating (slow measurements!)
  - same self-heating for biased S-parameter measurements as for the DC measurements.

As discussed in the previous chapter, the DC voltages and currents can be measured directly. The calibration is periodically auto-executed by the instrument.

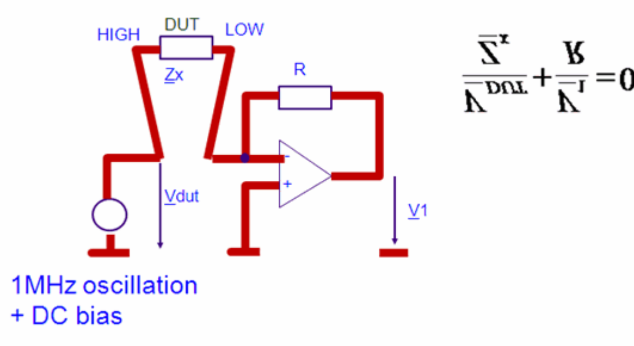
After such a DC characterization, modeling engineers usually perform a so-called CV (c apacitance versus voltage) measurement in order to characterize the device capacitances at a standard frequency of 1MHz. This frequency is high enough to allow a resolution down to a few femto-Ampere (provided shielded probes are applied for e.g. on-wafer measurements), yet still low enough to neglect second order parasitics like resistors in series with the capacitors, or like inductances.
For such CV measurements, the DC-bias is swept, a test frequency (1MHz) is applied to the DUT, and the instrument calculates the capacitance between the 2 pins of the DUT from the magnitude and phase of the device voltage and current. This means, an impedance meter interprete the measurement result *always* with respect to a user-specified schematic: either a capacitor in series with a resistors, or both in parallel. This explains, why capacitances and resistor values may vary with frequency when measured with such a device. In other words, these frequency-variations are due to a too simplistic analysis model behind the measurement. A better way is therefore to measure the capacitances with network analyzers. In this case, it is up to the user to interpret the measurement result (S-parameters).

CV Measurements - Calibration considerations

Test cables and fixtures contribute and affect the device characterization. For CV measurements, the calibration consists of unconnecting the DUT, assuming an *ideal* OPEN condition and measuring the cables and their OPEN parasitics (CV-Meter calibration). After that, the corresponding capacitance is automatically subtracted from the DUT measurement by the CV meter.

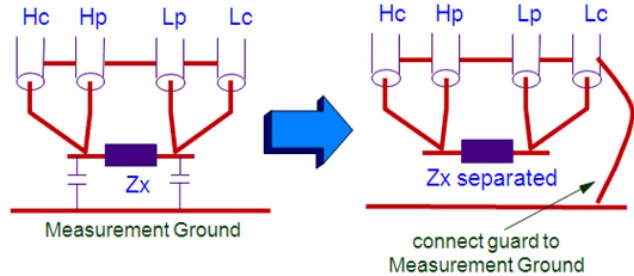
> **Note**
> If we are interested in the inner DUT's CV curves, i.e. without its surrounding test pads capacitances, we need to connect to an OPEN dummy structure during CV meter calibration instead of simply leaving the cables unconnected. Such an OPEN dummy consists of all connection pads, lines to the DUT etc, but *without* the inner DUT itself.

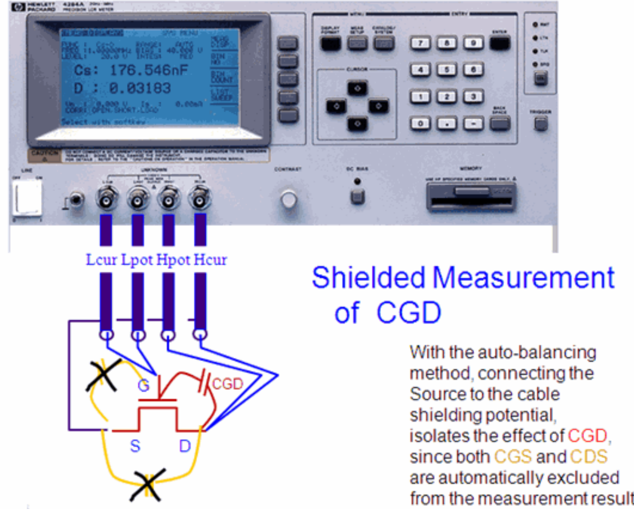
CV Meter Applying the Auto-balancing Method



For CV meters, the measurement principle is typically a so-called auto-balancing method. The slide above depicts the simplified measurement scheme.
The DUT is inserted in the feedback loop of an operational amplifier, and the system is stimulated with a 1MHz sinusoidal signal plus a DC bias. The feedback resistor R is precisely known, and the complex voltages V1 and Vdut are measured. From the formula given above in the slide, the capacitance of the DUT can be calculated, assuming an equivalent schematic of either a resistor in series with the capacitor, or, commonly for modeling, a capacitor in parallel with a resistor (which is the bias-dependent diode resistance for example).

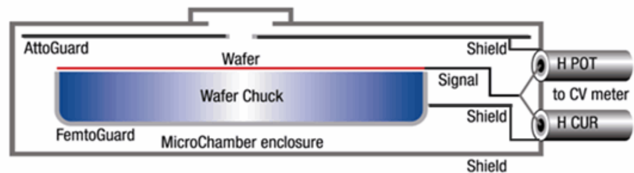
Avoiding Measurement of Parasitic Capacitance





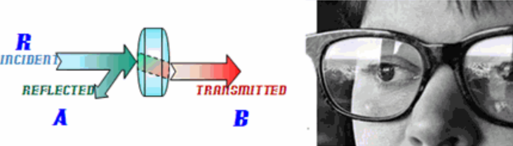
A so-called four-wire method is used for CV measurements. Similarly to the DC Kelvin measurement procedure (Force and Sense), both the low and high pin have to be tied together at the location of the DUT during CV measurements.
In the auto-balancing impedance measurement method, the shieldings of these four wires (which are not instrument chassis ground!) are connected to the virtual ground of the instrument's OpAmp. This eliminates any influences caused by the cables. Including this potential into the measurement setup allows the elimination of further stray capacitances. Like sketched above, it is possible to eliminate stray capacitances against ground (measurement plate).

AttoGuard™ Enhanced CV Measurements



- FemtoGuard surrounds the chuck at shield ground
- Patented AttoGuard above the chuck at shield ground
- Creates a virtual double-shielded Faraday enclosure
  - 10 atto Farad CV measurement resolution
  - Zero CV meter only one time

The two shields insure equal potential everywhere inside the waferprober MicroChamber regardless of chuck locations.
(slide with courtesy from Cascade Microtech)



While the CV measurement is considered as a specific two-pin test condition, the situation changes for frequencies above 100MHz. The device is now operated under its originally intended environment conditions: DC bias is applied to all the pins, and an additional small-signal RF excitation is applied. Now, the sinusoidal currents and voltages at *all* pins of the DUT are to be measured, *with magnitude and phase*.
A natural choice for such characterizations would be Z-, Y- or H-parameters from linear two-port theory. These two-port parameters can be used to completely describe the electrical behavior of our device (or network), including any source and load conditions. For such parameters, we have to measure the voltage or current as a function of

frequency and bias at the ports of the device.
At high frequencies, however, it is very hard to measure voltage and current at the device ports. One cannot simply connect a voltmeter or current probe and get accurate measurements due to the impedance of the probes themselves. Additionally, there is a difficulty to place the probes at the desired positions. Furthermore we have to apply either (AC-wise) OPEN or SHORT circuits as part of the Z-, Y- or H-parameter measurement. Active devices may oscillate or self-destruct with such terminations.
To avoid these problems, twoports are described by S-parameters.
IMPORTANT NOTE: like Y, Z, H parameters, S-Parameters are *linear* and belong to *linear circuit theory*. I.e. they represent the *small-signal behavior* of a device at a certain bias point, and for a certain frequency. Therefore, when measuring them, it must be assured that the *linear device operation* is maintained. As a consequence, linear S-parameters are independent of the applied RF signal power. Since twoport theory is restricted to linear circuit theory, matrix conversions (S to Y, S to Z etc. for de-embedding) are only applicable for linear operation of the device.
When characterizing the capacitance of transistors, the open 3rd transistor terminal should be connected to the shielding potential, eliminating the effect of the unwanted capacitors.

Definition of S-parameters

Referring to the spectacles examples from above, i.e. power-wise, the S-parameters are defined as:

Definition of S-parameters

Referring to the spectacles examples from above,
i.e. power-wise, the S-parameters are defined as:

$$\begin{pmatrix} |\underline{b}_1|^2 \\ |\underline{b}_2|^2 \end{pmatrix} = \begin{pmatrix} |\underline{S}_{11}|^2 & |\underline{S}_{12}|^2 \\ |\underline{S}_{21}|^2 & |\underline{S}_{22}|^2 \end{pmatrix} * \begin{pmatrix} |\underline{a}_1|^2 \\ |\underline{a}_2|^2 \end{pmatrix}$$

with

| $|\underline{a}i|^2$ power towards the two-port gate
| $|\underline{b}i|^2$ power away from the two-port gate

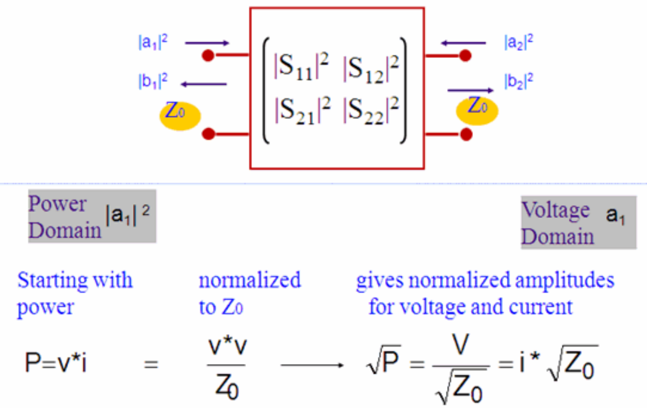
and

$|\underline{S}11|^2$	power reflected from port1
$|\underline{S}12|^2$	power transmitted from port2 to port1
$|\underline{S}21|^2$	power transmitted from port1 to port2
$|\underline{S}22|^2$	power reflected from port2

Agilent Technologies 18

Over the years there have been many methods developed to model the behavior of a circuit or network. H, Z, and Y parameters all have benefits to visualizing and predicting circuit behavior.
However, as the circuit frequency climbs higher, many of these models become complex and unruly. As a result, it is hard to model the network in a straightforward manner. Scattering (S)-parameters give a simple way to quantify the ratio of voltages exiting or scattering from a network relative to the voltages incident upon the network. By cascading* the parameters, models for individual circuit blocks may be put together to predict the operation of the whole system. As we will find out these ratios are fairly easy to measure at high frequencies. And difficult measurements of complex short-circuit currents and open-circuit voltages are not required to measure S-parameters.
Since S-parameters are defined as a voltage ratio, their result maps well into familiar RF and microwave characteristics of gain, loss, an reflection coefficient. They are also easily converted to other parameter forms such as H, Y, or Z, and this transportability makes S-parameters a good choice for importation into most modern simulation and design software programs. For readers acquainted with matrix algebra:cascading involves matrix multiplication of "T" matrices which are derived by a transformation of the S parameter matrices.

S-Parameters and Characteristic Impedance Z0



Power Domain $|a_1|^2$

Voltage Domain a_1

Starting with power | normalized to Z_0 | gives normalized amplitudes for voltage and current

$$P = v^* i = \frac{v^* v}{Z_0} \longrightarrow \sqrt{P} = \frac{V}{\sqrt{Z_0}} = i^* \sqrt{Z_0}$$

This means that S-parameters relate traveling waves (power) to a two-port's (DUT) reflection and transmission behavior. Since the two-port is imbedded in a characteristic impedance of Z0, and since we apply linear circuit theory , these 'waves' can be interpreted in terms of normalized voltage or current amplitudes.

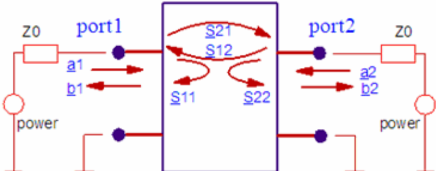
> **Note**
> Think of the spectacles again. Also here, the lens is imbedded on both sides with the same 'characteristic impedance', i.e. air.

Because of the characteristic impedance, we can convert the power towards the two-port into a normalized voltage amplitude of

$$a_i = \frac{v_{towards_twoport}}{\sqrt{Z_0}}$$

and the power away from the two-port can be interpreted in terms of voltages like

$$b_i = \frac{v_{away_from_twoport}}{\sqrt{Z_0}}$$



When determining the linear characteristics of a device, what we need to know is fairly simple. Given a sinewave signal incident upon the DUT, what of that signal is reflected from the device or transmitted through it? From the arrows in the illustration, there are four possibilities for this 2-port device. Knowing these four parameters completely characterizes this device for its linear operation.
For an n port device there is a reflection coefficient at each port and a transmission coefficient between each pair of ports in each direction.
So what we need to measure is

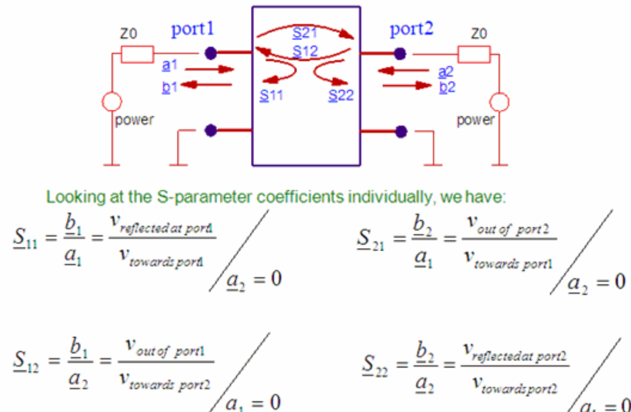
- Transmission Coefficient = Transmitted / Incident
- Reflection Coefficient = Reflected / Incident
  For a complete characterization of the linear behavior of an RF device.

To quantify and express these reflected and transmitted coefficients, the scattering or S-parameter were defined. S-parameters give a simple way to quantify and organize these ratios of voltages exiting or scattering from a network relative to the voltages incident upon the network.
S-parameters are always a ratio of two complex (magnitude and phase) quantities. S-parameter notation identifies these quantities using the numbering scheme shown above. The first number refers to the test-device port where the signal is emerging, or another way to look at it, which network analyzer port is the signal being measured. The second number refers to which test-device port the signal is incident or which network analyzer port the signal is coming from. For example, the S-parameter, S11, identifies the measurement as the complex ratio of the signal emerging from port 1 of the device to the signal applied to port 1 of the device (a reflection measurement).

A two-port device or network has four S-parameters. Two of the terms are related to the reflection from the input and output ports of the DUT. The other two terms are related to the transmission through the DUT in the forward and reverse directions. These concepts can be expanded to multi-port devices and the number of S-parameters is a function of 2 n , where n = the number of ports. For example, a four port device would have 16 S-parameters.

For the two port device, two independent equations may be written, expressing the variable b in terms of the variable a. In the above diagram b1 comprises the sum of a

quantity reflected from port 1 and a quantity that is the result of transmission through the device in the reverse direction. The quantities are scaled to be proportional to the voltage wave amplitude and phase such that $|b_n|^2$ = power emerging from the n'th port and $|a_n|^2$ is the power incident on the n'th port.



S11 and S21 are determined by measuring the magnitude and phase of the incident, reflected and transmitted signals when the output is terminated with a perfect Z0 load. This condition guarantees that a2 is zero. S11 is equivalent to the input complex reflection coefficient or impedance of the DUT, and S21 is the forward complex transmission coefficient.
Likewise, by placing the source at port 2 and terminating port 1 in a perfect load (making a1 zero), S22 and S12 measurements can be made. S22 is equivalent to the output complex reflection coefficient or output impedance of the DUT, and S12 is the reverse complex transmission coefficient.
The accuracy of S-parameter measurements depends greatly on how good a termination we apply to the port not being stimulated. Anything other than a perfect load will result in a1 or a2 not being zero (which violates the definition for S-parameters).

Interpreting S-parameters

S11 and S22

value	interpretation
-1	all voltage amplitudes towards the twoport are inverted and reflected (0 Ω)
0	impedance matching, no reflections at all (50 Ω)
+1	voltage amplitudes are reflected (infinite Ω)

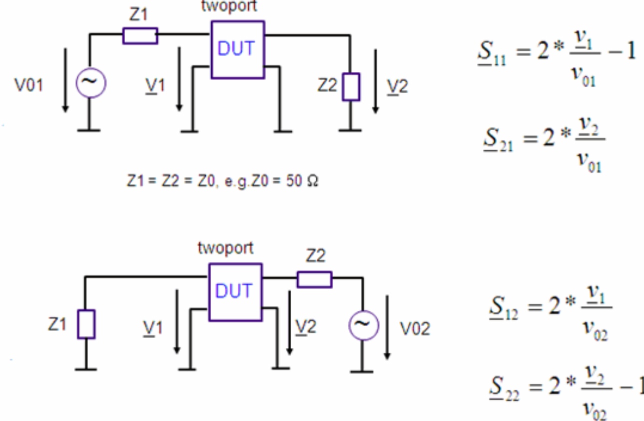
S21 and S12

magnitude	interpretation
0	no signal transmission at all
0 ... +1	input signal is damped in the Z0 environment
+1	unity gain signal transmission in the Z0 environment
> +1	input signal is amplified in the Z0 environment

The magnitude of S11 and S22 is always less than 1. Otherwise, it would represent a negative ohmic impedance value.

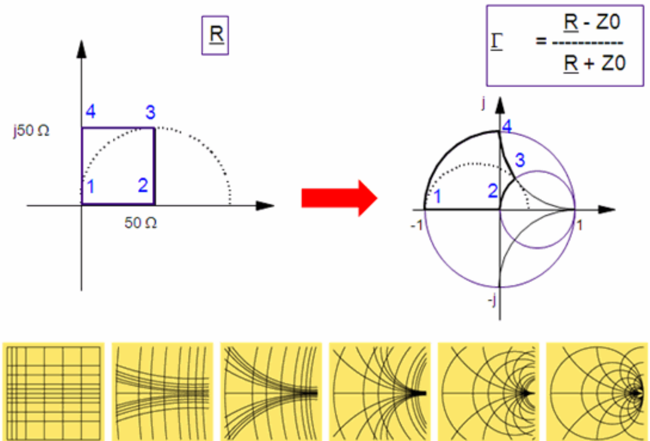
The magnitude of S21 (transfer characteristics) respectively S12 (reverse) can exceed the value of 1 in the case of active amplification. Furhtermore, S21 and S12 can be positive and negative. If they are negative, there is a phase shift. Example: S21 of a transistor starts usually at about S21 = -2 ... -20. This means signal amplification within the Z0 environment *and* 180' phase shift.

Calculating S-Parameters From Voltages



The slide above presents a method to calculate S-parameters from complex voltages.The DUT is imbedded in an characteristic impedance environment.
Provided the required complex forward and reverse voltages $\underline{V}_1$, $\underline{V}_{01}$, $\underline{V}_2$, $\underline{V}_{02}$ can be obtained (from measurements or simulations), we can calculate the S-parameters from the equations given above.

Understanding the Smith Chart



What makes Sxx-parameters especially interesting for modeling, is that S11 and S22 can be interpreted as complex input or output resistances of the two-port. That's why they are usually plotted in a Smith chart.

> **Note**
> Do not forget that *included* in Sxx is the termination at the opposite side of the two-port, usually Z0.

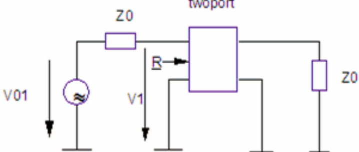
The Smith chart is a transformation of the complex impedance plane R into the complex reflection coefficient P (rho), following the formula given above.
This means that the right half of the complex impedance plane R is transformed into a circle in the P domain. The circle radius is '1' .

In order to get more familiar with interpreting the Smith chart, let's consider a square with the corners (0/0)W, (50/0)W, (50/j50)W and (0/j50)W in the complex impedance plane 'R' and its equivalent in the Smith chart with Z0=50W. Watch the angel-preserving property of this transform (rectangles stay rectangles close to their origins). Also watch how the positive and negative imaginary axis of the R plane is transformed into the Smith chart domain (W), and where (50/j50)W is located in the Smith chart. Also verify that the center of the Smith chart represents Z0, i.e. for Z0 = 50W, the center of the Smith chart is (50/j0)W.

Why a Smith Chart for the Sxx Parameters?

The voltage-related equation for the S11 parameter is $S_{11} = 2 \cdot \frac{v1}{v01} - 1$ (1)

where $\underline{v}1$ is the complex voltage at port 1 and v01 the stimulating AC source voltage (which is typically normalized to '1'). The corresponding circuit schematic is:



Under the assumption that $\underline{R}$ is the complex input resistance at port 1 and Z0 is the system impedance, we get applying the resistive divider formula for equation (1) from above:

$$\underline{S}_{11} = 2 \cdot \frac{\underline{R}}{\underline{R} + Z0} - 1 = \frac{\underline{R} - Z0}{\underline{R} + Z0}$$

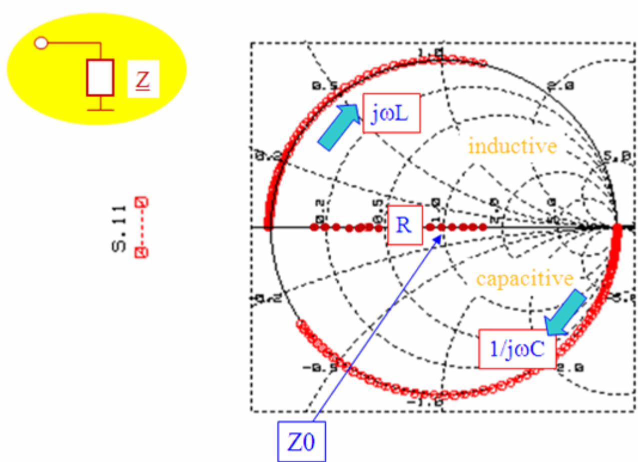
And this is the reflection coefficient $\underline{\Gamma}$ from the Smith Chart definition !!

Or, solved for the complex impedance $\underline{R}$:

$$\underline{R} = Z0 \cdot \frac{1 + \underline{\Gamma}}{1 - \underline{\Gamma}} = Z0 \cdot \frac{1 + \underline{S}_{11}}{1 - \underline{S}_{11}}$$

This explains how we can get the complex input/output resistance of a two-port directly from S11 or S22, if we plot these S-parameters in a Smith chart.

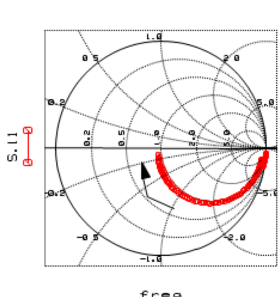
Interpreting S11 and S22 (Sxx)

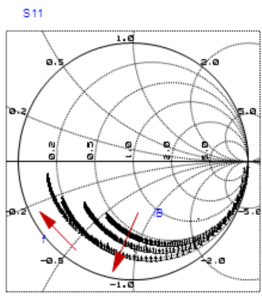


Rules for Smith charts for NWA measurements:

- Sxx on the real axis represent ohmic resistors
- Sxx above the real axis represent inductive impedances
- Sxx below the real axis represent capacitive impedances
- Sxx curves in the Smith chart turn clock-wise with increasing frequency (because in the $\underline{R}$ plane,all curves turn clock-wise too !).

S11 of a Capacitor and of a Bipolar Transistor





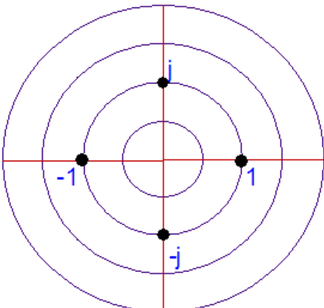
As an example for interpreting Smith charts, the left plot shows the S11 curve of a capacitor located between the two ports of the network analyzer (NWA). The capacitor represents an OPEN for DC, thus S11 = 1 = *Z0. For highest frequencies, it behaves like a SHORT, and we see the 50!measurement_dm_symbol1.gif! of the opposite port2 . The transition between the DC point and infinite frequency follows a circle, and the increasing frequency turns the curve again clockwise.
The right plot in the slide above shows the S11 plot of a bipolar transistor. In this case, the locus curve stars with S11>>1 at low frequencies corresponding to RBB' Rdiode (1+beta)*RE. For increasing frequencies, the curves then turn into the lower half-plane of the Smith chart, the capacitive region. Here, the CBE shorts Rdiode, and beta becomes smaller with frequency. For infinite frequency, when the capacitors represent ideal shorts,

and ß->0, the end point of S11 lies on the middle axis, i.e. the input impedance is completely ohmic, representing RBB' + RE. Since RBB' is bias dependent, and decreasing with increasing iB, the end points of the curves represent this bias-dependency.

Note
Keep in mind: For increasing frequency, the Sxx locus curves turn always clockwise.

The Polar diagram for Sxy

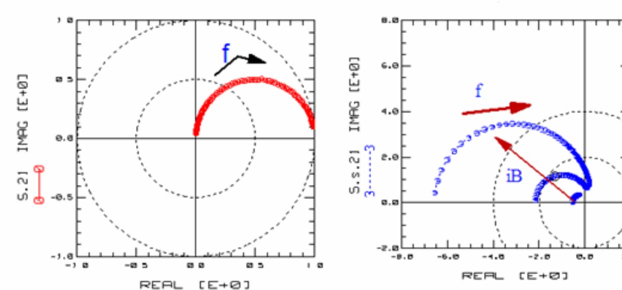
S21 corresponds to the forward transmitted voltage and S12 to the reverse transmitted voltage in a Z0 environment. In the polar diagram, you can easily identify MAGNITUDE and PHASE of these voltages.



The S21 parameter represents the power transmission from port 1 to port 2, if the two-port device is inserted into a matching network with characteristic impedance Z_0 of e.g.

50W. This means, if no signal is transmitted, then S21=0 (located in the center of the polar plot). If the signal is transmitted, then MAG(S21)>0. The magnitude of the S21 curve will be below '1' for damping between the port 1 and port 2, and above '1' for amplification. If the phase is inverted (transistors), we are basically in the left half-plane of the polar plot (REAL[S21]<0).
Like with the Smith chart, all S21 and S12 curves *turn clock-wise with increasing frequency*.

S21 (transmission) of a capacitor and of a bipolar transistor

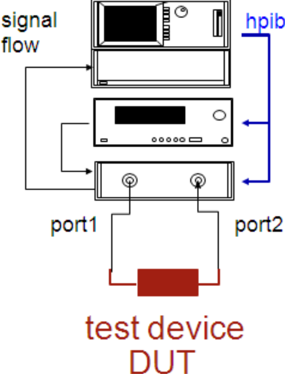


As a tutorial example, the capacitor on the left, exhibits no power transmission for lowest frequencies, but an ideal short (S21=1) for highest frequencies. Thus, the locus curve for S21 represents a circle from infinite to 0 W.

The right plot from above shows the S21 plot of a bipolar transistor between port 1 and port 2. The trace starts with REAL(S21) < -1 at low frequencies (voltage amplification in a 50 W system, plus phase inversion), and then tends towards S21 = 0 for highest frequencies (no voltage transmission, the transistor capacitances short all voltage transmission). Since the current amplification ß is bias depending, the start point of the S21 curve at lowest frequencies reflects this ß(iB) dependency: more ß for higher iB, i.e. more amplification magnitude with S21 for higher iB too.

Network Analyzer Measurements

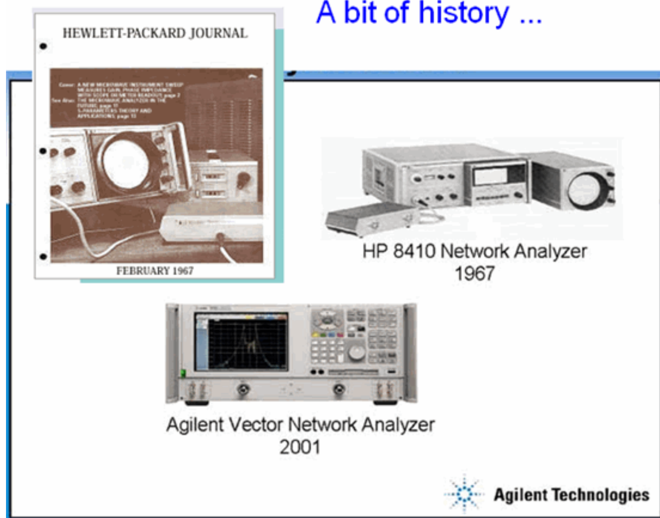
Network Analyzer Mainframe
RF Synthesizer
S-Parameter Testset



After the introduction to the S-parameters, it is time to consider how to measure them. A network analyzer (NWA), also sometimes abbreviated by VNA (vector network analyzer), is applied. This instrument measures S-parameter vectors, i.e. the magnitude and phase, of all four S-parameters of a two-port. This 'full two-port measurement' capability is important, because only in this case are we able to convert the measured S-parameters to Y and Z parameters etc., what is a requirement for de-embedding etc.

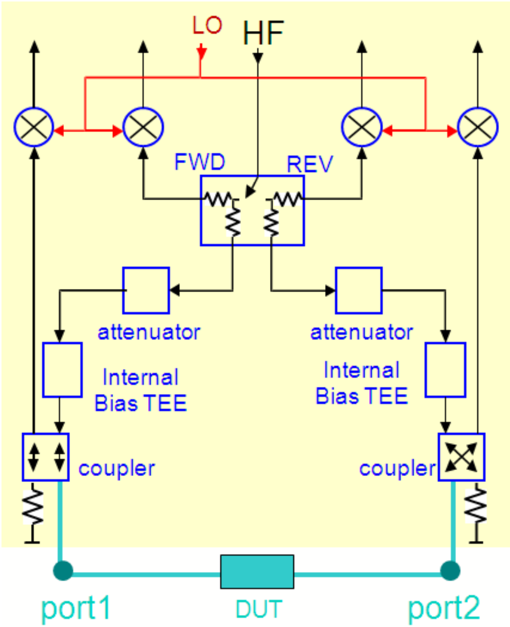
When applying network analyzers for S-parameter measurements, it is important to remember that we measure linear circuit performance and circuit performance for a given frequency (fundamental frequency), ignoring harmonics.

On the other hand, network analyzers can also be applied to specific non-linear measurements, e.g. sweeping the RF power, measuring the transfer characteristics and evaluating for example the 1dB compression point of amplifiers. In this case, however, signal distortion happens and harmonic frequencies show up. In our case, when measuring linear S-parameters with the NWA, always the *base* or *fundamental* frequency and is measured, and harmonics should not occur. Otherwise, they would be ignored !! Therefore, if we are interested in the modeling of device nonlinearities, we should rather apply a spectrum analyzer after the conventional DC-CV-NWA modeling, and use harmonic balance simulation (e.g. Agilent's ADS) to model the RF-power dependent spectrum. Alternatively, one of the currently introduced commercial Nonlinear NWAs (Agilent N4463A Large Signal Network Analyzer) can be applied as well. Such instruments measure both, the magnitude *and* phase of the transmitted and reflected, fundamental *and* harmonics frequencies.



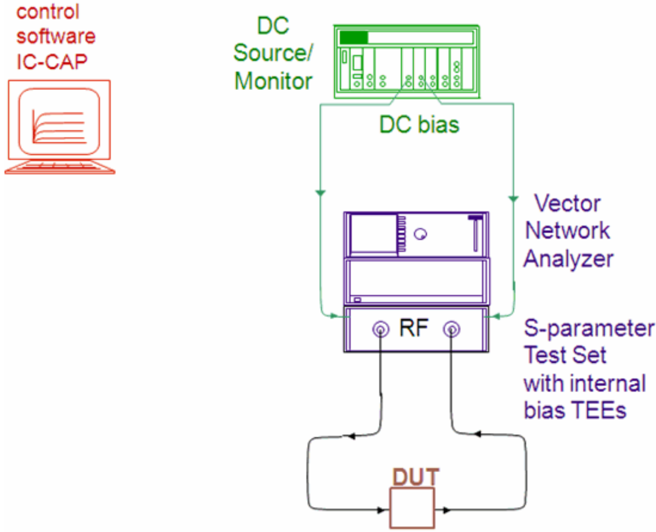
When Hewlett-Packard introduced the HP 8410 in 1967, it revolutionized microwave design. It used the a new 1430A sampler together with a superheterodyne receiver architecture to provide a calibrated microwave receiver. Together with the test sets, it featured measurements of the transmission and reflection coefficients for any twoport device.

Block Diagram of the S-Parameter Testset



The block diagram in the slide above shows the core of this meter combination, the S-parameter testset. The RF Input source at the top, connected to the RF synthesizer, provides the stimulus power. The PIN switch directs the signal to either a forward or a reverse S-parameter measurement. Directional couplers then detect the injected and reflected power of the DUT. The detected signals are downcoverted into four IF signals for further analysis in the VNA mainframe, where each input is digitized and signal processed in order to give the S-parameters.

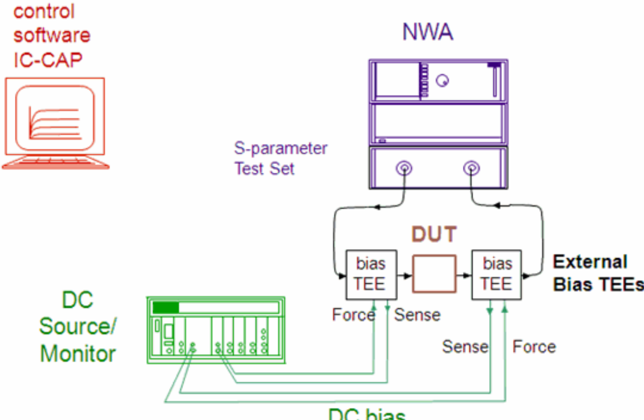
NWA Measurements Using the Internal DC bias TEEs of the S-parameter Testset



For devices like transistors and diodes, an additional DC bias has to be applied to the device.
This can be done by using the DC bias inputs of the NWA's S-parameter testset. Keep in mind that there are typically 1-2W ohmic losses due to the internal bias TEE's inductor. This causes a voltage drop and a reduced bias voltage at the device! Many NWA's also have an internal 1MW resistor to ground, which prevents electrostatic discharge damage to the internal NWA circuitry.

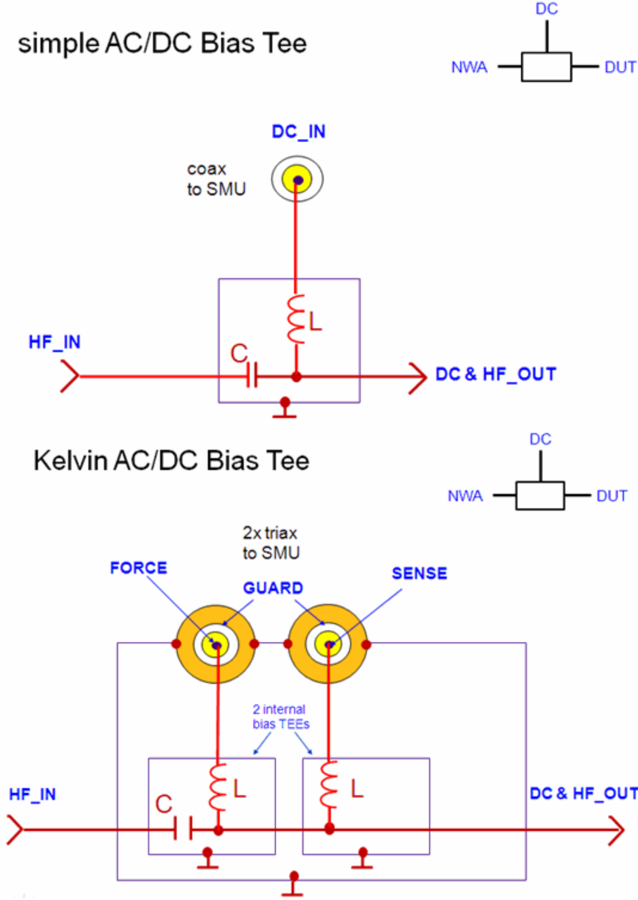
> **Note**
> In the setup presented above, make sure to use the right triax (SMU) to coax (S-parameter testset) adapter, which leaves the middle shield of the SMU triax cable (Guard) unconnected.

NWA measurements with external DC bias TEEs



If we want to avoid the ohmic losses of the NWA's internal bias TEE, we can use external bias TEEs and apply the previously mentioned force-sense DC biasing (Kelvin measurements).However, the ohmic losses for the bias are now *not* zero, but rather ~1W?due to losses in the cables and the connectors. Also, these Kelvin bias TEEs should be placed as close as possible to the DUT.

In the measurement setup above, two external bias TEEs are inserted between the two DUT connections and the network analyzer ports.
Keep in mind to not connect the guard shield of the triax Kelvin cables.



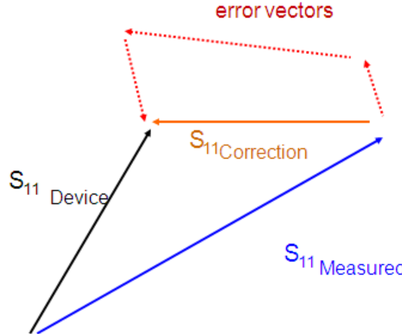
Basics of DC and AC Characterization of Semiconductors

Contents

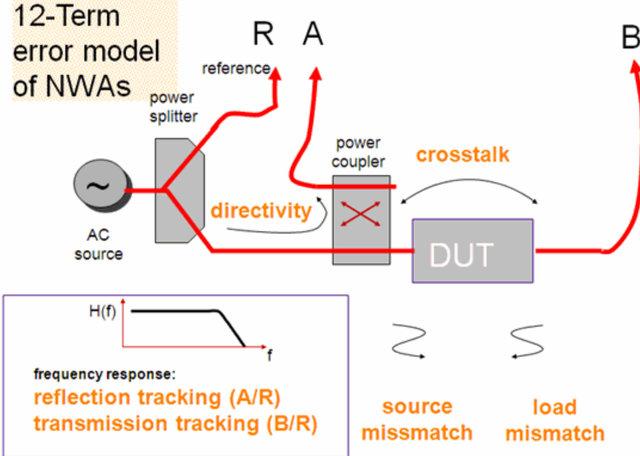
- Basics of device measurement and modeling techniques from DC to RF
- Special aspects of network analyzer calibration
- De-embedding and required dummy structures

Vector Error Correction

Accounts for all major sources of systematic error.

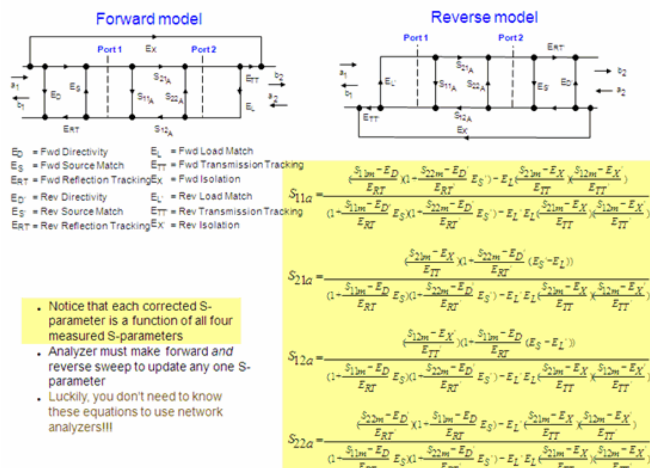


Since S-parameters are vectors, all measurement errors contribute with magnitude and phase, and can be considered as additional error vectors. These errors have to be calibrated out with the correction vector $Sxx_{correction}$, as depicted above.



What is a 12-term error correction? As can be seen above, there are 6 error contribution terms in forward direction, related to the characterizing signals R, A and B of the NWA:1. Directivity:cross-talk of the power splitter in the NWA testset2. Crosstalk:cross-talk inside the S-parameter test set, overlying the DUT3. Source Mismatch:multiple reflections due to impedance mismatch of cables and connectors4. Load Mismatch:the same for the opposite port5. Reflection Tracking A/R:frequency dependence of signal path R->A6. Transmission Tracking A/R:same for signal path R>B

For the reverse calibration, another 6 error terms add up to a total of 12 terms. The procedure to get rid of these 12 terms is called the 12-term error correction.



The formulae above are applied in the NWA in order to correct the measured S-parameter $Sxy_{\mathbf{M}}$ with the correction terms **E**xx, and to finally obtain the requested $Sxy_{\mathbf{A}}$ parameters of the device under test (12 term error correction).

It is of particular interest that e.g. the resulting $S11_{\mathbf{A}}$ parameters are affected by *all* measured $S11_{\mathbf{M}}$ parameters ! This means, if there is a problem with *'only one'* S-parameter index during measurements (or calibration), this will affect *all* S-parameters. This means, an absolute clean calibration and also an absolute proper measurement is

required in order to get good S-parameter results!

Calibrating a NWA

- Full 2-port calibration
  (reflection and transmission measurements)
- 12 systematic error terms measured usually requires 12 measurements on four known standards (SOLT)
- Standards are defined in cal kit definition table; these cal kit definitions are entered to the network analyzer

The internal NWA CAL KIT Definition must watch the actual kit used

There are many different calibration techniques for network analyzers. Such are Short-Open-Load-Thru (SOLT), Thru-Reflection-Load (TRL) or Load-Refection-Match (LRM). For the different calibration procedures, specific known standard terminations have to be measured.

The slide above depicts such standards for connectorized measurements (e.g. 3.5mm connectors).
Although there are many publications on the pros and contras of the different calibration methods, the SOLT is most commonly used for on-wafer measurements of silicon devices. One of the reasons is that due to the electrical losses of silicon, microstrip standards as required for LRM and TRL are difficult to manufacture on the wafer. Another reason for using SOLT is that this calibration is a wide-band calibration and not limited to a frequency band.

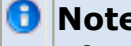
Some Naming Conventions: The Calkit

While in the case of the CV meter, the calibration corrects for a *single, ideal offset capacitor, a NWA calibration relates to cal standards (OPEN, SHORT, LOAD, THRU etc.) from a calkit. These cal standards do not represent ideal standards. They represent the _real, existing standard*, including its nonidealities! It means that a SHORT is not an ideal SHORT, but instead represents rather a small inductance. The same applies to the THRU, which has a non-ideal delay time. The OPEN corresponds rather to a capacitor than to an ideal OPEN. Therefore, these non-idealities of the G-S-G probes have to be entered into the NWA *before* calibration. This is called 'entering' or 'modifying the calkit'.

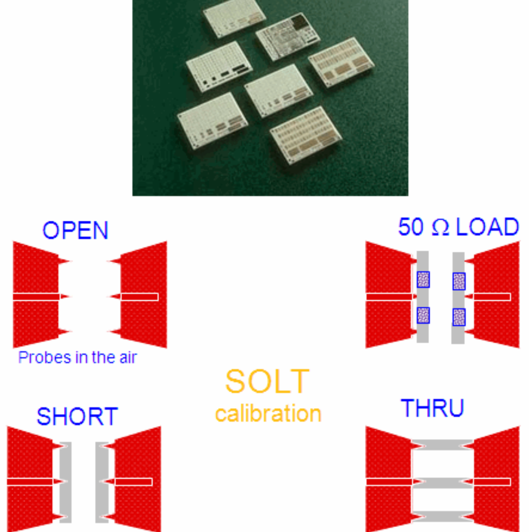
THE CALSET

While entering the calkit refers to the non-idealities of the calibration standards, i.e. the termination of the NWA cables during calibration, the subsequent **calibration** is based on this information, and then related to the selected - frequency range, - the RF power, - the averaging of the NWA etc. After the calibration has been performed, the correction terms are stored in the **calset** of the NWA. In other words, the 12-term error vectors are 'filled up', and can be used with the measurements afterwards.

Finally, when the measurements are performed, the raw measured data arrays will be corrected inside the NWA, using a correction technique related to the selected calibration method, and referring to the specified calset. When using a NWA driver software, this corrected measurement result is transferred into the software and displayed there.

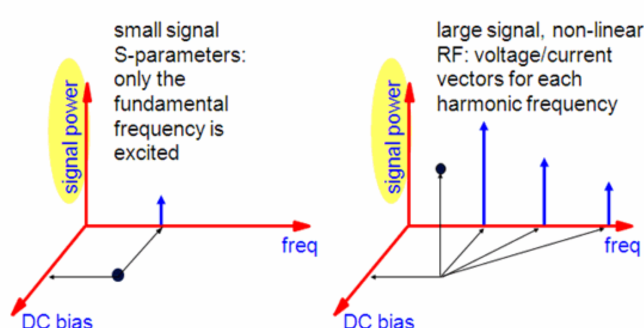


Note
After the calibration, a re-measurement of the OPEN will not represent an ideal open, but instead exactly those parasitic components as described in the documentation of the OPEN. In the same way, a THRU shows up after calibration with its real delay time, and a SHORT represents its inductive behavior! If the calibration was ok, the remeasured standards should give exactly the same parameter values as previously entered into the NWA calkit.



In case of a SOLT calibration, and for on-wafer measurements using Ground-Signal-Ground (GSG) coplanar probes, the slide above depicts the corresponding test structures, which are usually available on a RF-high-performance ceramic substrate, including accurate description of the non-idealities of these standards. These standards (called ISS standards) are provided by the GSG probe manufacturers. Both, the GSG probe and the ISS substrate belong together.

Extending linear to RF power dependent signals

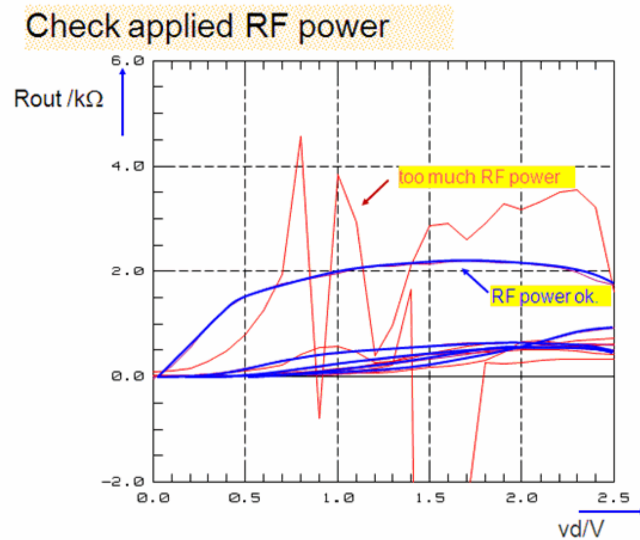


An absolute prerequisite for using S-parameter two-port matrices is the linear, time invariant behavior of the circuit. Only in this case, matrix conversions like for example for de-embedding, are possible. Nonlinear high-frequency measurements cannot be de-embedded by Y- and Z-matrix subtractions.

- What if characterizing and modeling a transistor at higher RF signal levels where it behaves non-linear. If you need to model e.g. a power RF transistor at signal levels where distortion occurs, and where the Kirchhoff law is not fulfilled for the data measured by the NWA, it is absolutely mandatory to replace the linear SPICE S-parameter simulation by a nonlinear_ harmonic balance (HB) simulation. Only with this kind of simulation, we can emulate the conditions of the power RF transistor measurement with the NWA. From the simulation data of harmonic balance, we can derive the S-parameters of the base frequency and compare these S-parameters with those obtained from the NWA measurements. Only in this case the obtained model parameters will be correct for the power RF transistor.

- If distortion occurs, harmonic frequencies show up, which by themselves shift the DC operating point. With the NWA plus SMUs, we measure only the (shifted) DC bias current and with the NWA the fundamental frequency current vector. Their sum is *not* zero. Kirchhoff's law states that the sum of currents, for all frequencies (including DC) into a node is zero. But with linear NWA, we do not measure the harmonics frequencies (with magnitude and phase). Only a very special Nonlinear Network Analyzer, not to confuse with standard NWA's, can measure this complete frequency spectrum with respect to magnitudes and phases.

- Conclusions: If the DUT behaves non-linearly:- signal compression occurs- harmonics show up- the DC operating point may be shifted- loadlines and transfer curves become dynamic, i.e. RF-power dependent- matrix conversions are no longer possible- de-embedding by S-to-Y and S-to-Z matrix conversion (see further below) is no longer valid.



Selecting the right RF power for nonlinear devices before starting calibration.+ When measuring S-parameters of nonlinear devices with a network analyzer, it must be assured that these devices operate in small-signal, linear mode. Otherwise, the high frequency test signals will no longer be sinusoidal, and the occurrence of harmonics

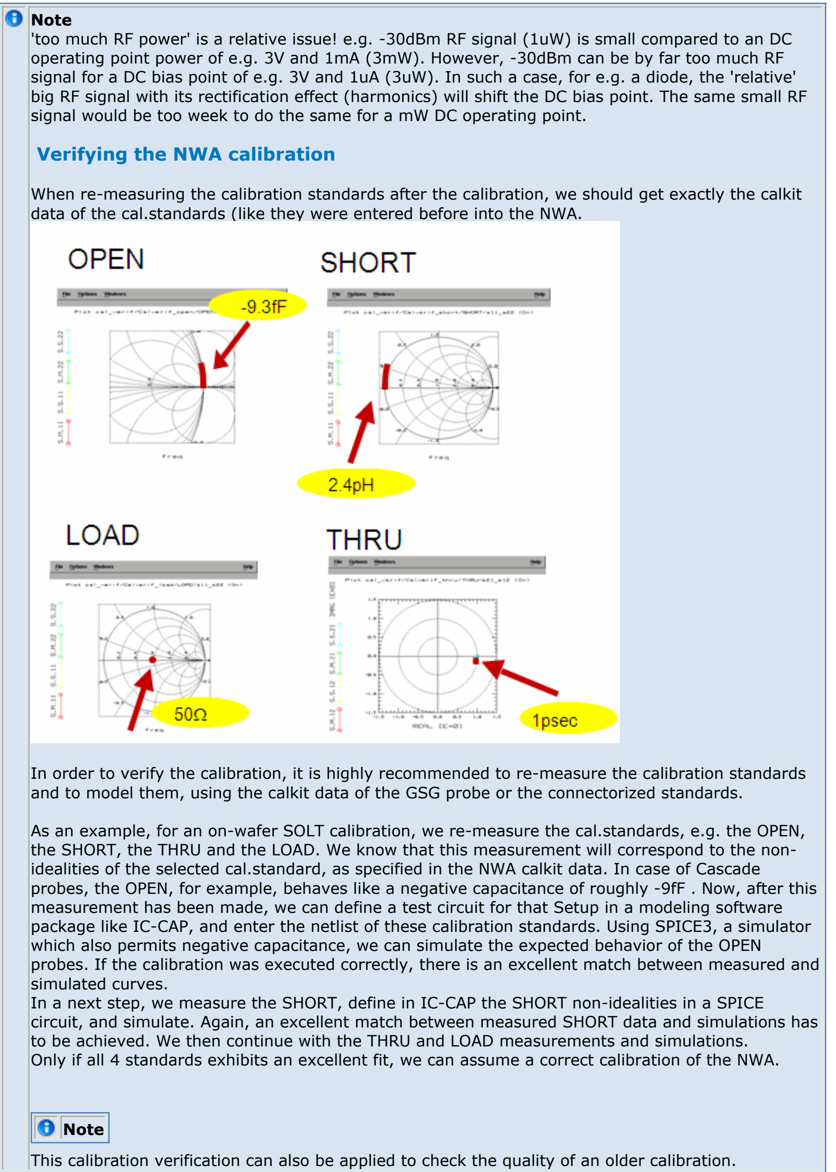
will lead to wrong S-parameter measurements and shifted DC bias conditions.

A smart method to check the correct port power settings is like this: When measuring a DC output characteristics and calculating the output resistor 'Rout' from of it, the resulting curve is very sensitive. Therefore, we can use this plot to identify possible effects of too big an AC power applied to the transistor. This means, we measure the DC output characteristics, and let the NWA perform un-triggered measurements, in continuous mode, i.e. un-synchronized to the DC measurement. Then, we increase the Port power manually (or decrease the port attenuation) until we see an effect on the next 'Rout' measurement. We then know the maximum allowed RF power for the S-parameter measurements of this device!
The plot above reflects such a test. The disturbed curve happens when too much RF power is applied to the transistor.

Note
'too much RF power' is a relative issue! e.g. -30dBm RF signal (1uW) is small compared to an DC operating point power of e.g. 3V and 1mA (3mW). However, -30dBm can be by far too much RF signal for a DC bias point of e.g. 3V and 1uA (3uW). In such a case, for e.g. a diode, the 'relative' big RF signal with its rectification effect (harmonics) will shift the DC bias point. The same small RF signal would be too week to do the same for a mW DC operating point.

Verifying the NWA calibration

When re-measuring the calibration standards after the calibration, we should get exactly the calkit data of the cal.standards (like they were entered before into the NWA.



In order to verify the calibration, it is highly recommended to re-measure the calibration standards and to model them, using the calkit data of the GSG probe or the connectorized standards.

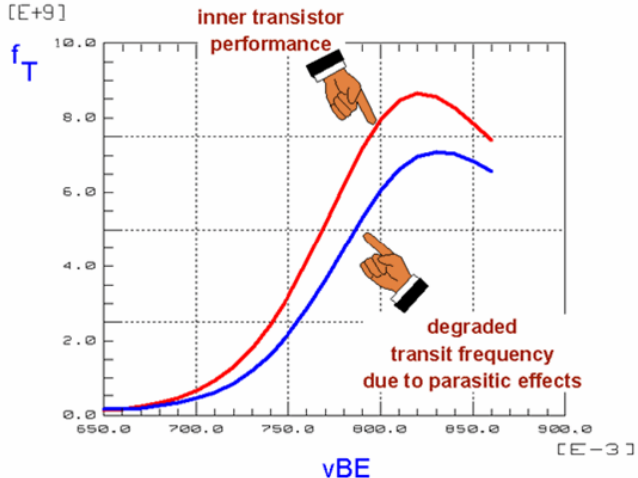
As an example, for an on-wafer SOLT calibration, we re-measure the cal.standards, e.g. the OPEN, the SHORT, the THRU and the LOAD. We know that this measurement will correspond to the non-idealities of the selected cal.standard, as specified in the NWA calkit data. In case of Cascade probes, the OPEN, for example, behaves like a negative capacitance of roughly -9fF . Now, after this measurement has been made, we can define a test circuit for that Setup in a modeling software package like IC-CAP, and enter the netlist of these calibration standards. Using SPICE3, a simulator which also permits negative capacitance, we can simulate the expected behavior of the OPEN probes. If the calibration was executed correctly, there is an excellent match between measured and simulated curves.
In a next step, we measure the SHORT, define in IC-CAP the SHORT non-idealities in a SPICE circuit, and simulate. Again, an excellent match between measured SHORT data and simulations has to be achieved. We then continue with the THRU and LOAD measurements and simulations.
Only if all 4 standards exhibits an excellent fit, we can assume a correct calibration of the NWA.

Note
This calibration verification can also be applied to check the quality of an older calibration.

The results depicted in the slide above give an example for Cascade G-S-G probes, 100um pitch.

Once again, only if the fitting between simulated and measured data is in the few-percent range, for all 4 re-measured calibration standards, the NWA calibration can be considered as good. If only one fit is bad, re-perform a new NWA calibration.

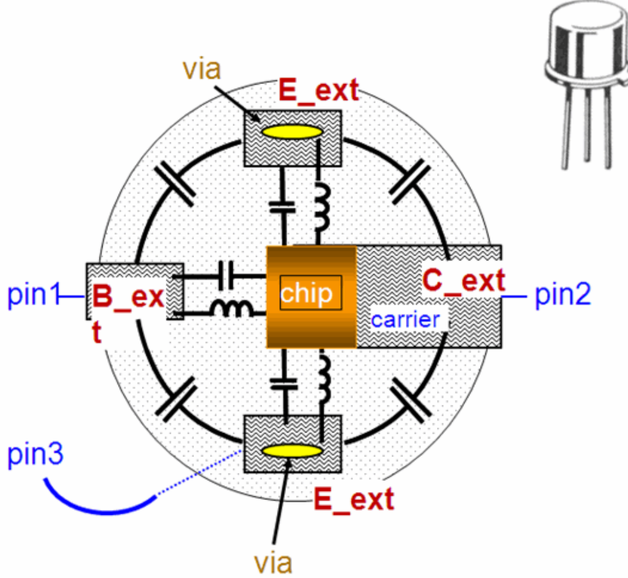
How parasitics degrade the device performance?



After having performed a network analyzer calibration, the calibration plane is located at the ends of the NWA cable connectors (connectorized devices) or at the ends of the probe contacts (on-wafer measurements). The device itself, including its surrounding parasitics, is then connected to this calibration plane.
In the case of a packaged device, S-parameter measurements would now include the test fixture, the package and the very inner DUT. For on-wafer device characterization, using e.g. ground-signal-ground probes (GSG), the test pads (where the probes touch down) degrade the performance of the inner DUT by their layout specific capacitive and pad parasitics.
In order to extend the calibration plane to either the beginning of the package, or to the inner DUT, these outer parasitic effects have to be stripped off. This is called de-embedding.
In the slide above, a brief example on how de-embedding returns the real, inner DUT performance without its degradation due to the measurement environment is given. We can clearly see how the transistor cutoff frequency fT is degraded due to these parasitics.

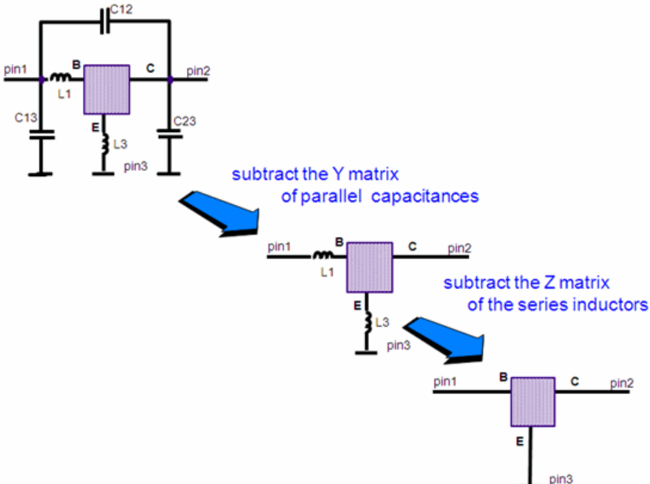
Chip Carrier of a Packaged Transistor and its Parasitic Components

The bipolar transistor is mounted with its Collector to the carrier, Base is bonded, and the Double-Emitter contact (bonded twice) goes to the backside of the carrier.



If we are going to characterize a transistor mounted to a chip carrier, this carrier will distort the performance of the inner transistor. Yet, the chip carrier can be described with series inductors of the bond wires and the parallel capacitors of the pads, as sketched above.

Z and Y Matrix Manipulations to De-embed the Inner Transistor



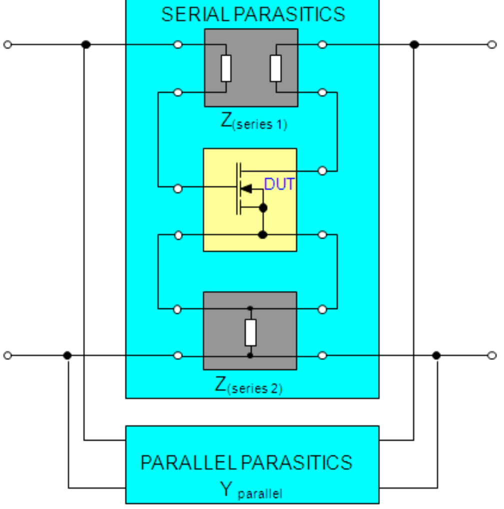
By rearranging the schematic components, we obtain a simplified equivalent schematic as given above.
In order to strip off the parasitic components from the outside towards the inner DUT, the slide shows the corresponding de-embedding procedure.

De-embedding Procedure

We have measured the S-parameters Stotal of the transistor including the carrier parasitics. With the assumption of having *nothing but* parallel capacitors as 'outer' parasitic components, we transform the S-parameters to Y, because a Y matrix represents a PI structure of components. A simple subtraction will de-embed the parasitic capacitor effects.
Now, the new 'outer' parasitic components are the two inductors, which are in series with the chip connections. Series parasitics can be easily eliminated by subtracting a Z matrix. Therefore, we transform the resulting Y-parameters from above into Z-parameters and subtract the inductors.
These Z-parameters are finally transformed back into S-parameters which now describe exclusively the performance of the 'inner' chip.

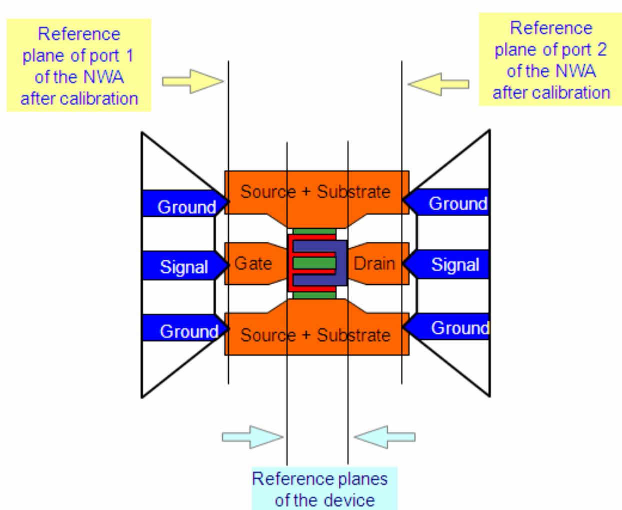
Particularly for on-wafer measurements, de-embedding by subtracting the complete Y matrix (OPEN) and Z matrix (SHORT) is often applied. However, in this case, it must *absolutely* be assured that

- There are no hidden series components present for Y-matrix subtractions (mixed cross-talk represented by chains of C-L-C-L etc.).
- There are no hidden parallel components present for Z-matrix subtractions.
- Related to a two-step de-embedding:
  - For on-wafer measurements (OPEN > SHORT sequence): the SHORT has been de-embedded from the OPEN dummy.
  - For packaged measurements (SHORT > OPEN sequence): the OPEN has been de-embedded from the SHORT dummy.
  - There are no hidden delay line effects present when subtracting the Y or Z-matrices.

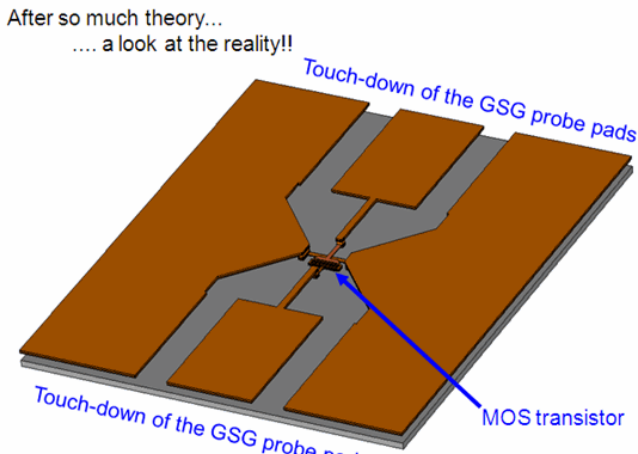


The slide above sketches the de-embedding situation for a transistor with first capacitive parasitics, followed by inductive parasitics.

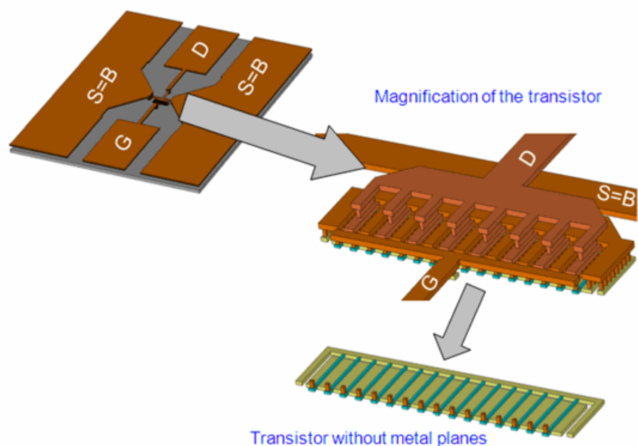
De-embedding shifts the Calibration plane



And this slide above once again visualizes the shift of the calibration plane due to de-embedding.

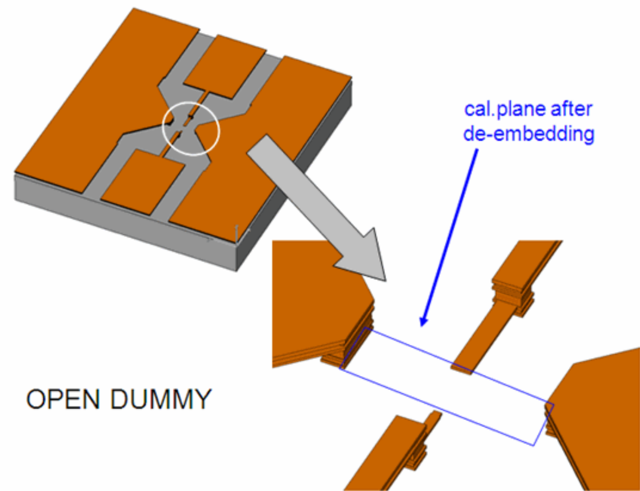


This slide depicts a three-dimensional representation of such a MOS transistor in its on-wafer test environment. It is interesting to note that the size of the transistor itself is very small compared to the size of the pads!
The goal of de-embedding is to shift the NWA calibration plane, which was (after the NWA calibration) at the end of the probe tips, down to the beginning of the transistor. This *'beginning'* is one of the key points for a good de-embedding and a key point for good dummy structures.

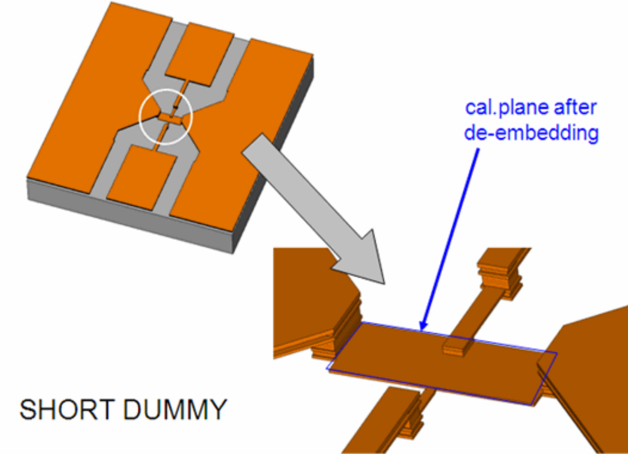


In the above slide, in the section 'Magnification of the transistor' is the inner transistor, which we want to model. All the rest has to be de-embedded, i.e. to be stripped-off.
In other words, the NWA calibration plane has to be shifted down here.

Note
Watch how 'high' the metal planes are relative to the active silicon area. A real mountain of metalization.

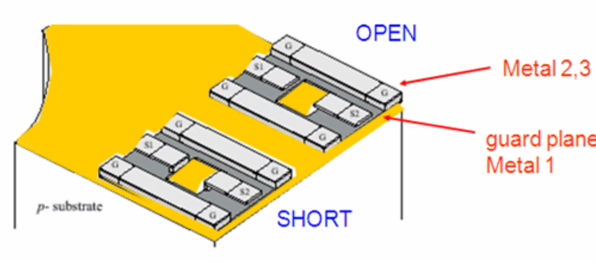


The pre requisite for a correct de-embedding is that certain test structures are available on a wafer together with the device under test (DUT) itself. Depending on the selected de-embedding method, an OPEN and SHORT dummy structure is required and must be measured. For a de-embedding verification, also a THROUGH dummy structure is necessary.
The principle layouts of these structures are given in the actual and the next slides. These layouts are for Ground-Signal-Ground Probes (GSG).
Please check the 'after-de-embedding' calibration plane as marked in the figures. Everything, every part of the DUT included in these cal. plane will be part of the DUT model. In other words, you can think of de-embedding as a shift of the current calibration plane to these new limits on the wafer. The limits of the blue surrounded area in the OPEN dummy will become the shifted calibration plane. All what is inside becomes part of the transistor model.



The SHORT dummy again refers to the desired shifted calibration plane. All what is inside this plane is now filled up with metal, so that we can consider this part to behave ideal, while the striplines from Gate and Drain will behave like inductors and will show up with the SHORT dummy measurement.

Another smart idea for Dummy Test Structures, avoiding the drawbacks of the SHORT metal plane

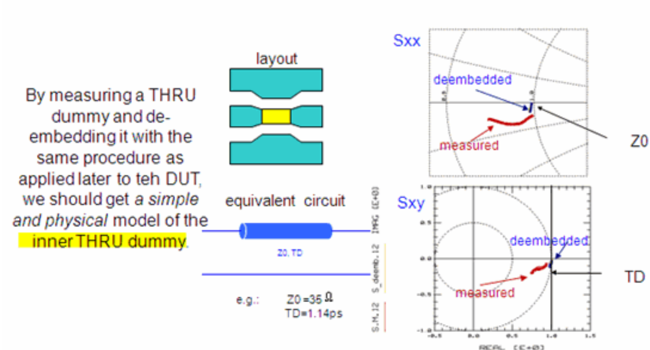


From:
T.E.Kolding, O.K.Jensen, T.Larsen,
Ground-Shielded Measuring Techniques for Accurate On-Wafer Characterization of RF CMOS Devices
IEEE International Conference on Microelectronic Test Structures, March 2000, p.246-251

This is the THRU dummy, for de-embedding verification purpose. Instead of the transistor,

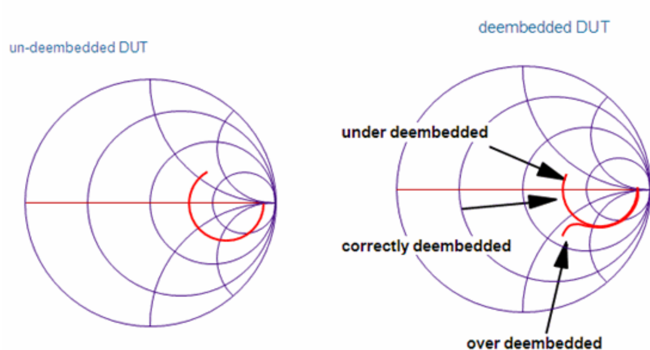
we now have a strip line (metal 1).
See the next slides for a proposed de-embedding verification procedure..

Verifying the de-embedding



Instead of the 'classical' layout shown in the previous slides, today's layouts for silicon wafers look more and more often like the one shown in the slide above.
It avoids the effects of the lossy silicon substrate by using the 1st metal plane to shield the contact pads and the lines to the DUT against the lossy silicon substrate. Especially interesting is the drastically improved SHORT dummy performance, since it applies a series of vias to ground (metal 1) at the end of the SHORT, rather than having - as with the conventional SHORT- a large metal plane (the SHORT) which is considered as ideal, while the two microstrip lines (to the contact pads) of the SHORT dummy are considered as non-ideal. Also, the OPEN is much more ideal with this alternate approach.
Altogether, this layout suggestion of using the 1st metal as a shield features lower parasitics. Therefore, the de-embedding is not so much the 'difference of big numbers' like in the conventional layout suggestions, and therefore more robust.

Under and over de-embedding



Verifying the de-embedding procedure is very important before applying it to the very DUT, i.e. the transistor, the passive RF component etc.
Without this verification step, errors or problems associated with the de-embedding will add to the performance of the inner, de-embedded device, and, thus, lead to a wrong measurement result, and finally to a wrong device model. This is especially critical when de-embedding (subtracting) the complete Y and Z matrices of the OPEN and SHORT dummy, because, in this case, possible problems may not show up compared to de-embedding of individual, lumped components. (With de-embedding lumped components, we would have detected such problems because they would show up with non-physical values).

Therefore, it is suggested to model every dummy structure before simply subtracting their total matrices in order to verify its de-embedding prerequisites:

- for the OPEN dummy: no series parasitics allowed.
- for the SHORT dummy: no parallel parasitics allowed
- verify the de-embedding procedure with a well-known 'golden device', e.g. the proposed THRU strip line, before applying the de-embedding procedure to the very DUT.

Some remarks on the under- and over-de-embedding.
UNDER-DEEMBEDDING
If under-deembedding occurred, i.e. we deembed 'not enough parasitics', what means the lumped components values were estimated too small, the de-embedded traces will not be 'turned back' sufficiently, resulting in a too big a delay time. The de-embedding error will add up to the inner device and distort its model.
The better visible effect is
OVER-DEEMBEDDING

If over-deembedding occurred, however, the deembedded DUT curves will 'turn backwards' for higher frequencies or show some other non-physical effects. They may even turn outside the Smith chart!
This emphasizes once again the need to accurately verify the de-embedding procedure *before* applying it to the very DUT.

Measurements

Contents

- *Measurement Techniques* (iccapmhb)
- *Data Verification* (iccapmhb)
- *Data Management* (iccapmhb)

On the Web

- Articles
- Knowledge Center
- Technical Support
- Training
- Videos

Measurement Techniques

Contents

- *Prerequisites* (iccapmhb)
- *DC* (iccapmhb)
- *CV* (iccapmhb)
- *Network Analyzers* (iccapmhb)
- *Nonlinear Network Analyzers* (iccapmhb)
- *Time Domain* (iccapmhb)
- *Noise* (iccapmhb)

On the Web

- Articles
- Knowledge Center
- Technical Support
- Training
- Videos

Prerequisites

Contents

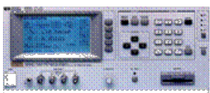
- *Wafer Prober* (iccapmhb)
- *Switching Matrix* (iccapmhb)
- **Test Fixtures**
  - *Intercontinental Microwave Test Fixtures* (iccapmhb)
- **Cables**
  - *Physical Properties of Triax and Quadrax Cables* (iccapmhb)
  - *FMAX Cables* (iccapmhb)
- **Adapter Connectors**
  - *Agilent RF Connectors General Information* (iccapmhb)
  - *Agilent RF Connectors Adapters* (iccapmhb)
  - *Connecting Triax SMU Outputs to Coax Inputs* (iccapmhb)

On the Web

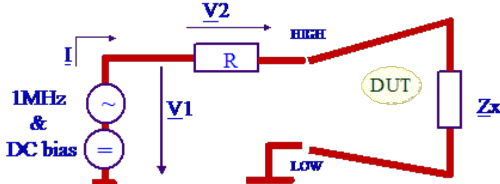
- Articles
- Knowledge Center
- Technical Support
- Training
- Videos

Wafer Prober

Slide 1

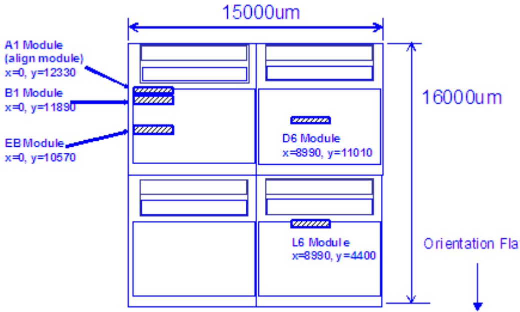


Slide 2



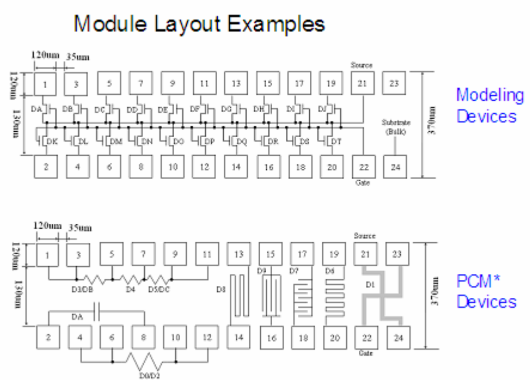
Let's commence with some nomenclature definitions first:
> Wafer> Die (an assembly of repeatable test structures)> Module (e.g a set of MOS transistors of different geometry)
In the slide above, you see a Wafer, with the orientation flat at the bottom. Furthermore, a Die is marked and also zoomed in.
Within the dies are the Modules, see the next slide.
Slide 3

First some nomenclature: Die - Module

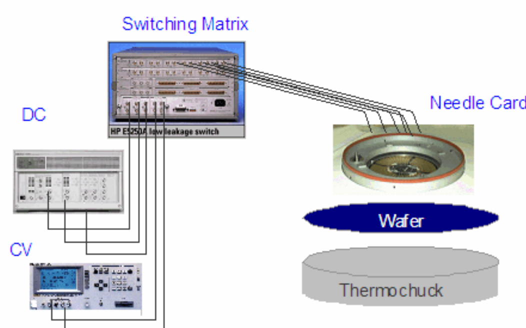


A Die contains several Modules. Each Module contains a set of Test Structures (for production tests, for device modeling etc.)
Their location is typically specified with respect to a corner of the Die.
Several Modules may be grouped within a Block. But Blocks are organizazion-only.

Slide 4

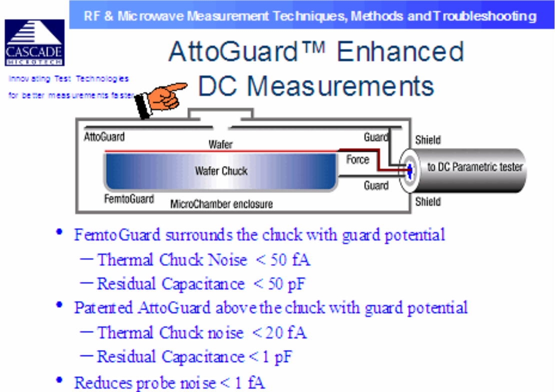


A Module, finally, represents an area with typically several devices for modeling (for MOS, a set of transistors of different Ls and Ws).
It can be considered as a wafer probing position, on which by use of a switching matrix the different measurements are performed:- DC and CV for device modeling (parametric test, multiple bias sweeps)- Spot measurements (single bias points) for PCM devices
Slide 5



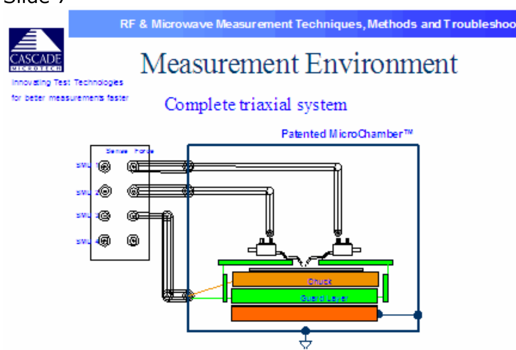
This is a typical DC-CV measurement environment:
> a wafer is placed on the chuck of a wafer prober> the chuck can be heated or cooled (thermochuck)> applying a needle card together with a switching matrix allows to contact many test structures at a time, and to connect the individual test structure to the DC analyzer or the CV meter.
IC-CAP distinguishes between measurement instruments and auxiliary measurement equipment. Therefore, the DC analyzer and the CV meter are controlled from the built-in IC-CAP instruments drivers, while the thermochuck, the waferprober and the switching matrix are controlled by macros.

TASK:
> load the file
3_MEAS_ORGANIZE_n_VERIFY_DATA/5_wafertest/wafertest_DC_CV_PELdep.mdl
and find out how the thermochuck, the wafer prober and the switching matrix are controlled.
Slide 6

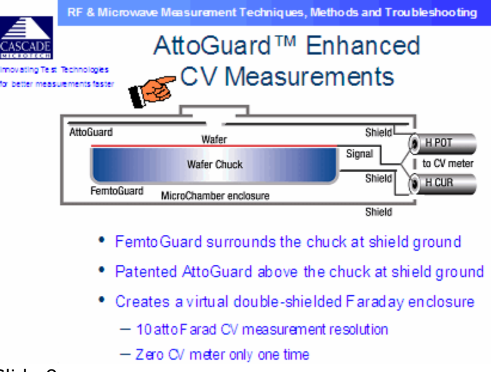


The following slides about wafer prober and the corresponding text are a copy of sections of the file 'DC-CV Training Presentation.pdf' of the Cascade Microtech, Winter 2002 European Probing Seminar CD-ROM, with friendly permission of Cascade Microtech, Europe

Cascade adds two additional chuck guards to minimize stray capacitance, leakage and noise for IV measurements. The Cascade FemtoGuard surrounds the chuck bottom and sides with guard voltage to eliminate coupling the force to the shield below. The new patented AttoGuard extends the guard over the chuck to minimize force-to-shield coupling from above.
Slide 7

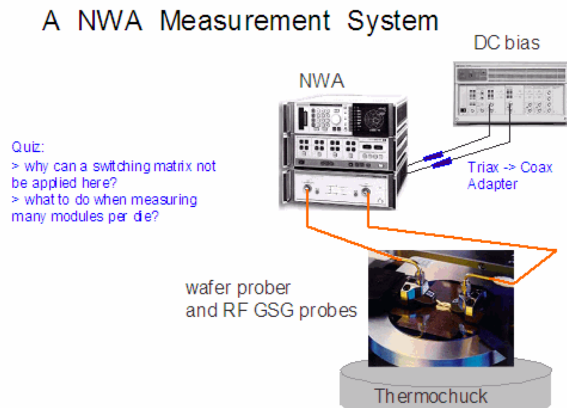


A typical i/v measurement setup requires probes and chuck bias. The patented AttoGuard & FemtoGuard completely surround the wafer with guard allowing only a small orifice for the guarded probes to enter.
Slide 8

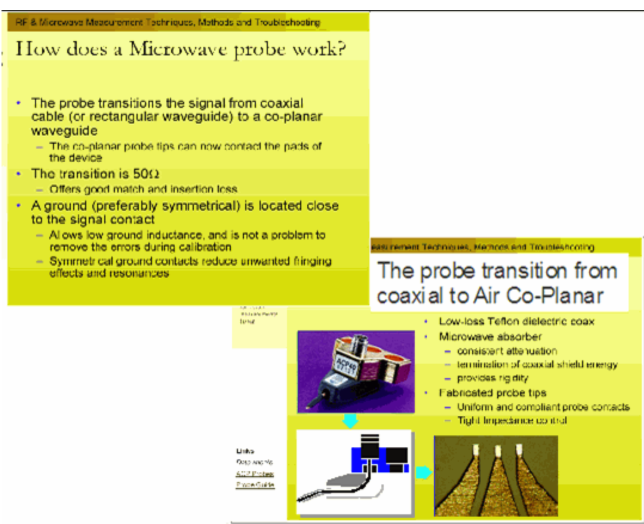


Slide 9

Since the MicroChamber is at shield potential, only guarded coax probes are required inside. Solid PTFE dielectric coax exhibiting < 1 femtoamp noise and <50 femtofarad residual capacitance is used. The probe bodies come with single, quasi-kelvin or kelvin connectors. The quasi-kelvin style probe is most popular because it is common to both IV and CV measurement systems.
Slide 10

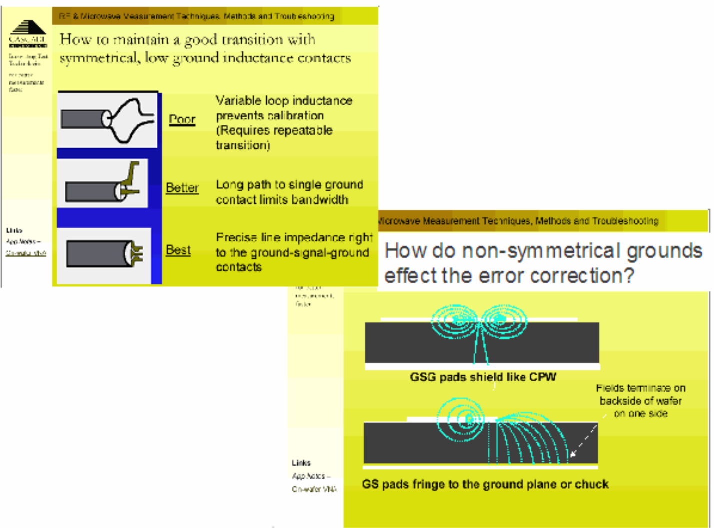


Different to the so far discussed DC-CV measurement scenario, this is a typical setup for S-parameter measurements.
NOTE:As a special remark, watch out for the triax->coax adapter used to feed the SMU bias (triax) into the S-parameter testset bias input (coax) of the NWA.These adapters must not connect the middle shield of the triax to the outside common shield. The middle shield of the triax cables must be *absolutely* unconnected!
Slide 11

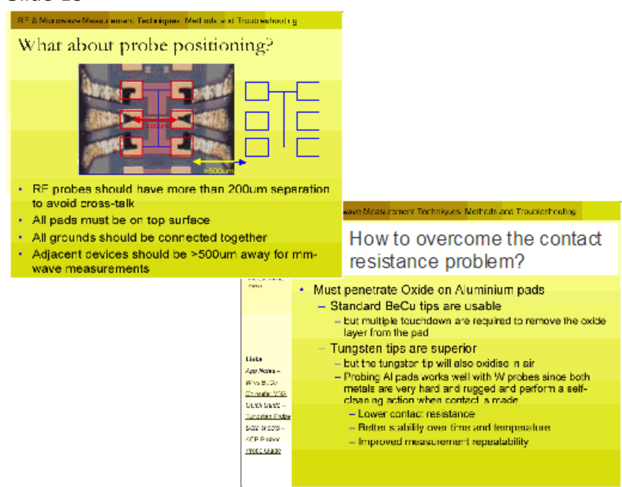


Different to making DC and CV measurements on a prober, several high-frequency aspects have to be considered for successful and correct GHz measurements.
As a general rule, for GHz measurements, go for the most reliable and best equipment available! Do not make improvisations!

Slide 12

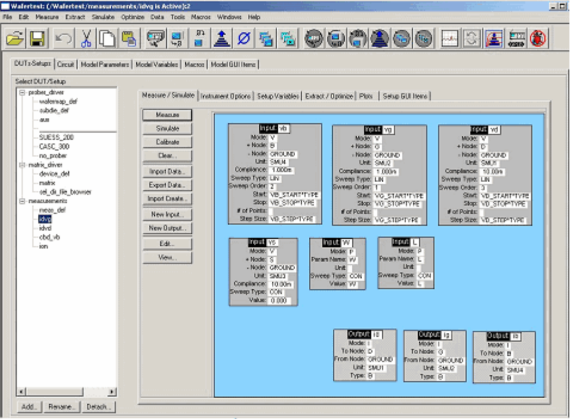


Slide 13



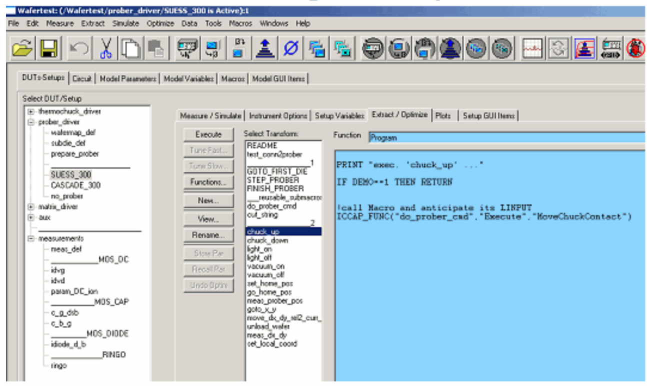
Slide 14

Example of an IC-CAP Wafertest Modelfile



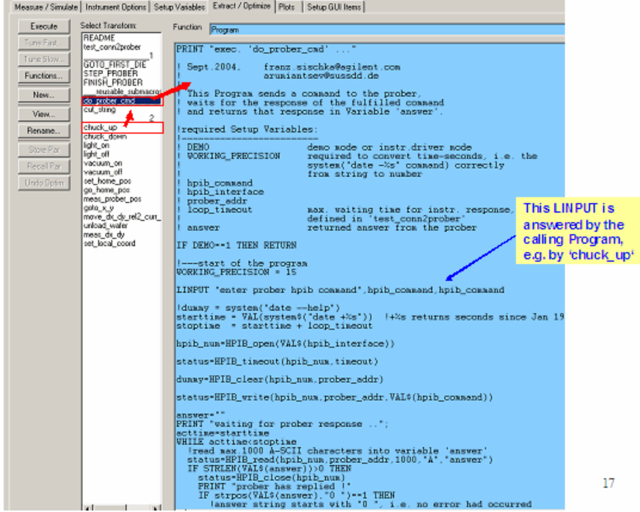
Slide 15

Detail of controling a wafer prober in IC-CAP

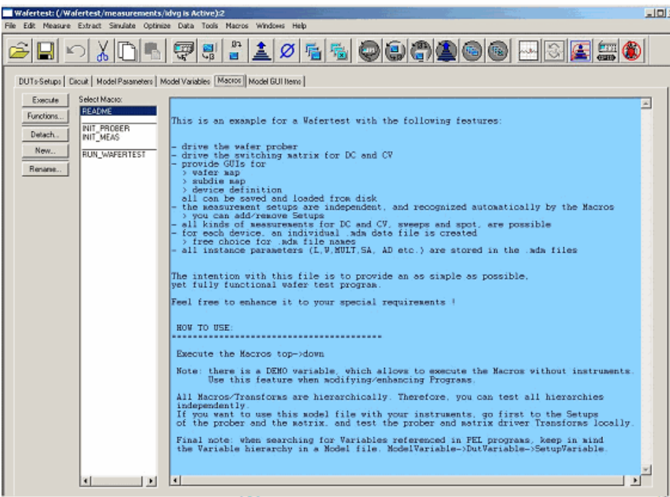


The screenshots above and of the next slides is from the Modelfile
demo_features\3_MEAS_ORGANIZE_n_VERIFY_DATA\5_WAFERTEST\
1_wafertest_DC_CV_indiv_mdm_files_2006\1_wafertest_DC_CV_indiv_mdm_files_2006.mdl
Slide 16

The central wafer prober driver PEL program



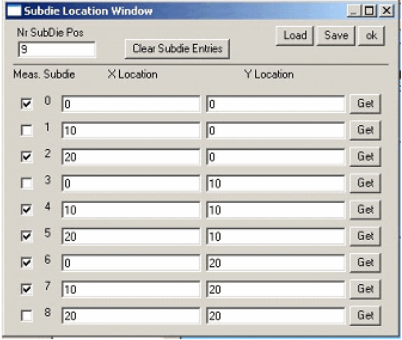
Slide 17



Wafer probers are best controlled by using PEL programs in IC-CAP.
For this purpose, the PEL commands
HPIB_open()
HPIB_write()
...HPIB_close() are available. This PEL programming allows to support any wafer prober.
The screenshot above is of the Modelfile
demo_features\3_MEAS_ORGANIZE_n_VERIFY_DATA\5_WAFERTEST\
1_wafertest_DC_CV_indiv_mdm_files_2006\1_wafertest_DC_CV_indiv_mdm_files_2006.mdl
All different prober-control Macros (Transforms) like chuck_up, chuck_down, goto_x_y
etc. call a central wafer prober control Program 'do_prober_cmd', which finally performs
all the communication with the prober.

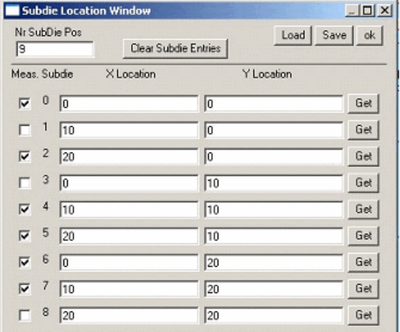
Slide 18

Slide 19

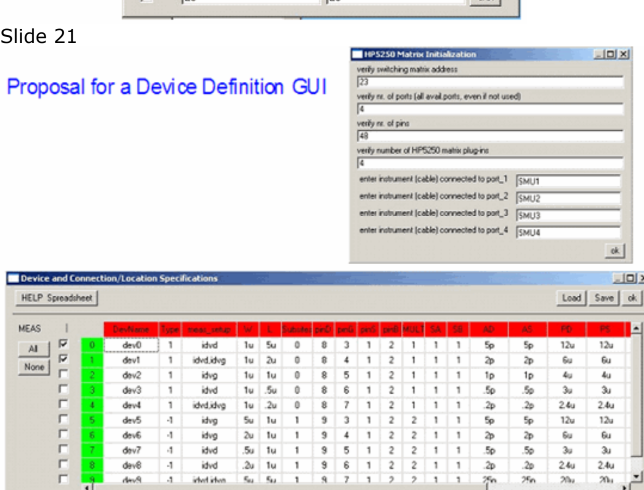


Slide 20

Proposal for a Subdie Definition GUI

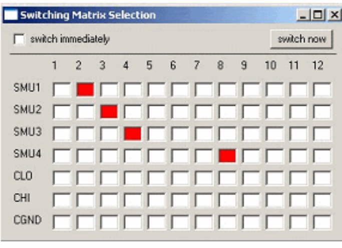


Slide 21



Slide 22

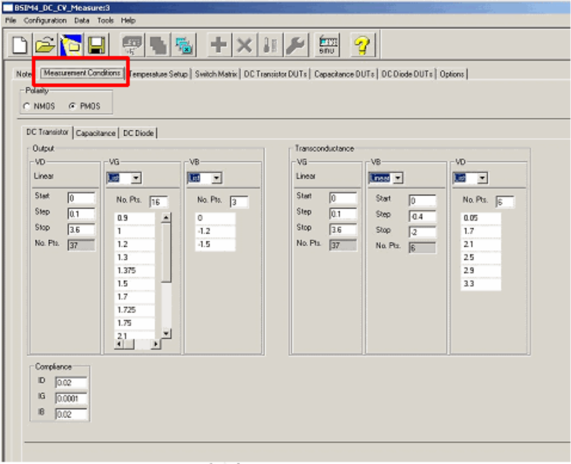
Proposal for a Switching Matrix GUI



Slide 23

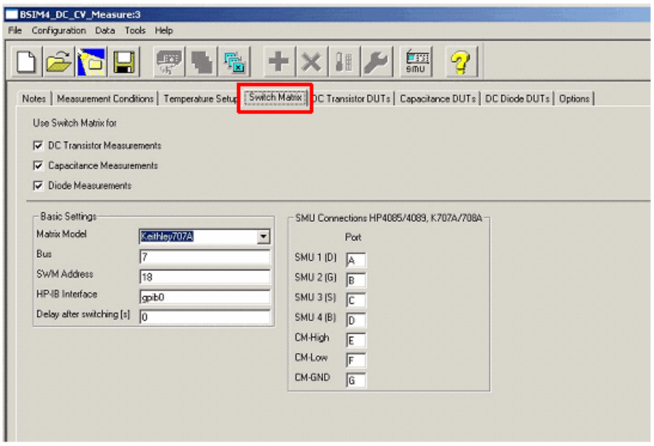
The Ready-To-Go Solution:

IC-CAP MOS Measurement Toolkit 85194K

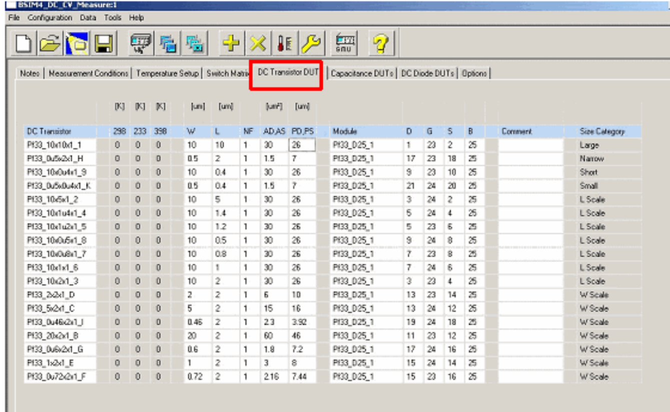


The screenshots of this series of slides is from the IC-CAP MOS Measurement Toolkit 85194K

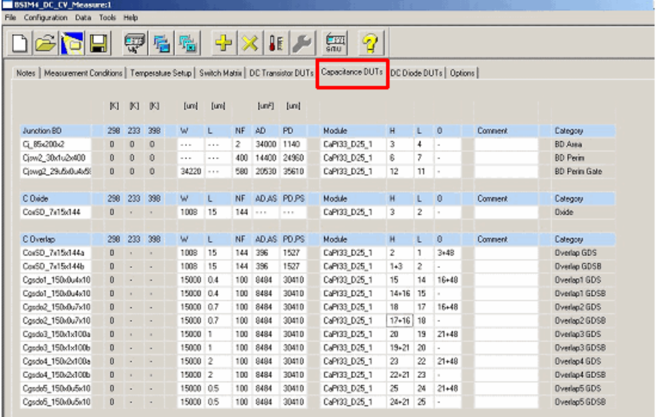
Slide 24



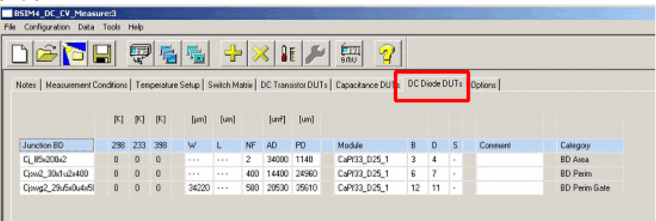
Slide 25



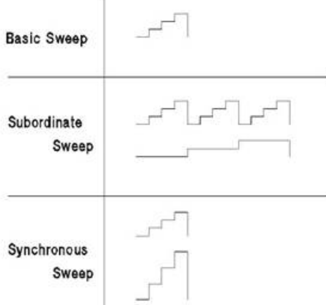
Slide 26



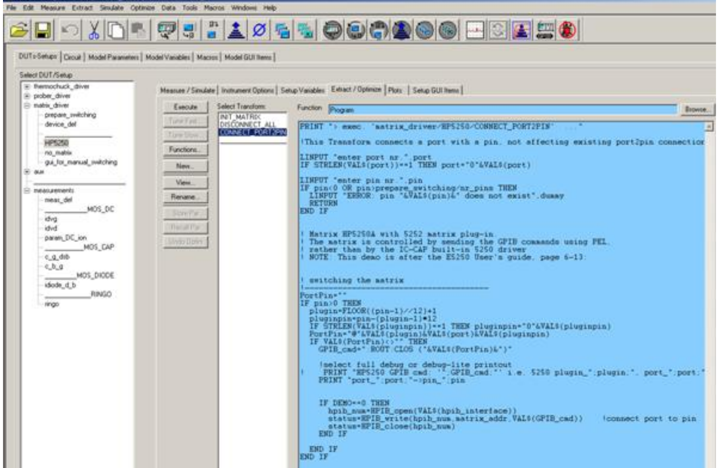
Slide 27



Switching Matrix

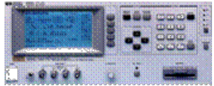


Agilent B2200
There are two possibilities to control switching matrices in IC-CAP.
Either by using the built-in functions, or (giving you more flexibility) by simple PEL programs.
For this purpose, the PEL commands HPIB_open() HPIB_write()...HPIB_close() are available. This PEL programming allows to support *any* switching matrix.
Below is an example screenshot of the Modelfile
demo_features\3_MEAS_ORGANIZE_n_VERIFY_DATA\5_WAFERTEST\
1_wafertest_DC_CV_indiv_mdm_files_2006\1_wafertest_DC_CV_indiv_mdm_files_2006.mdl

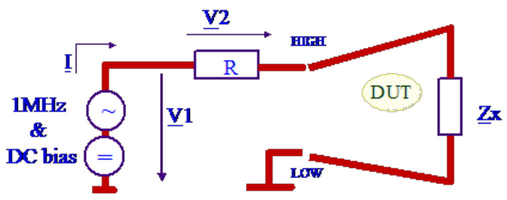


More PEL examples for driving switching matrices are available under
demo_features\2_MEAS_INSTRUMENTS\3_PROBER_MATRIX_THERMOCHUCK\1_Matrix\

Slide 1



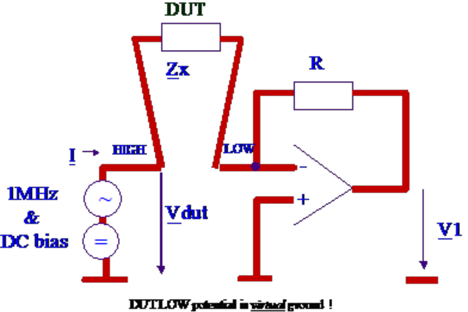
Slide 2



Slide 3

$$Z_x = \frac{V_1}{I} - R = \left(\frac{V_1}{V_2} - 1\right) * R$$

Slide 4

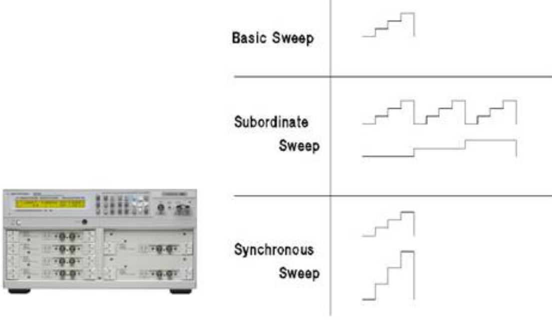


Slide 5

$$\frac{V_{DUT}}{Z_x} + \frac{V_1}{R} = 0$$

Physical Properties of Triax and Quadrax Cables

Triax CableKelvin Triaxial Cable (Quadrax)



ModelLengthCapacitance
16494A Triaxial Cable
Option 0011.5 m125 pF
Option 0023.0 m240 pF
Option 00380 cm 75 pF
16494B Kelvin Triaxial (Quadrax) Cable
Option 0011.5 m140 pF
Option 0023.0 m260 pF
Option 00380 cm 90 pF
16494C Kelvin Triaxial (Quadrax) Cable for 4142
Option 0011.5 m140 pF
Option 0023.0 m260 pF
Notes:
How to Select Options for Cable Length:
The Agilent 4155C/4156C or Agilent 4142B have low current forcing and measurement capability by their SMU (Source Monitor Unit) technology. The SMU uses a kind of feedback circuit for maintaining very high stability of the current/voltage sourcing. Therefore, the guard capacitance (a load for SMUs)of connections to SMU must be within certain limits to assure the accurate functioning.
From experimental measurements, the limit is about 900 pF for the Agilent 4155C/4156C or the Agilent 4142B, for example.
You should select proper cable lengths to keep small guard capacitance for the SMUs.
A system using switching matrices sometimes needs longer cable lengths than for the standalone instruments. If you plan to connect more than ~5 m total cable length (both instrument to switching matrix and switching matrix to the DUTs), make sure the cable guard capacitance does not exceed 900 pF.

FMAX Cables

Slide 1

Agilent RF and Microwave Test Accessories

Overview

Many coaxial connector types are available in the RF and microwave industry, each designed for a specific purpose and application. For measurement applications, it is important to consider the number of connects/disconnects, which impact the connector's useful life.

The frequency range of any connector is limited by the excitation of the first circular waveguide propagation mode in the coaxial structure. Decreasing the diameter of the outer conductor increases the highest usable frequency; filling the air space with dielectric lowers the highest usable frequency and increases system loss.

Performance of all connectors is affected by the quality of the interface for the mated pair. If the diameters of the inner and outer conductors vary from the nominal design, if plating quality is poor, or if contact separation at the junction is excessive, then the reflection coefficient and resistive loss at the interface will be degraded.

A few connectors, such as the APC-7, are designed to be sexless. Most are female connectors that have slotted fingers, which introduce a small inductance at the interface. The fingers accommodate tolerance variations but reduce repeatability and may ultimately break after 1000 connections. Agilent offers slotless versions of connectors in certain measuring products, which decrease inductance and increase repeatability.

The following is a brief review of common connectors used in test and measurement applications:

APC-7 (7 mm) connector

The APC-7 (Amphenol Precision Connector-7 mm) offers the lowest reflection coefficient and most repeatable measurement of all 18 GHz connectors. Development of the connector was a joint effort between HP and Amphenol, which began in the 1960s. This is a sexless design and is the preferred connector for the most demanding applications, notably metrology and calibration.

Type-N connector

The type-N (Navy) 50-ohm connector was designed in the 1940s for military systems operating below 4 GHz. In the 1960s, improvements pushed performance to 12 GHz and later, mode-free, to 18 GHz. Agilent offers some products with slotless type-N center conductors for improved performance to 18 GHz. Agilent type-N connectors are completely compatible
with MIL-C-39012. Certain 75-ohm products use a type-N design with smaller center conductor diameters, and thus are not compatible with 50-ohm connectors.

SMA connector

The SMA (Subminiature A) connector was designed by Bendix Scintilla Corporation and is one of the most commonly used RF/microwave connectors. It is intended for use on semirigid cables and in components that are connected infrequently. Most SMA connectors have higher reflection coefficients than other connectors available for use to 24 GHz because of the difficulty to anchor the dielectric support.

3.5-mm connector

The 3.5-mm connector was primarily developed at Hewlett Packard—now Agilent Technologies, with early manufacturing at Amphenol. Its design strategy focused on highly-rugged physical interfaces that would mate with popular SMA dimensions, allowing thousands of
repeatable connections. It is mode-free to 34 GHz.

1.0-mm launch

The launch adapter has a 1.0-mm female connector on one end and a glass to metal seal interface on the other end. This is for transition of ultra-high frequency (up to 110 GHz) signals from coax into a microstrip package or onto a circuit board.

Adapters, Cables and Connectors - General Connector Information

2.92-mm connector

The 2.92-mm connector mates with SMA and 3.5-mm connectors and offers mode-free performance to 40 GHz.

2.4-mm connector

The 2.4-mm connector was developed by HP, Amphenol, and M/A-COM for use to 50 GHz. This design eliminates the fragility of the SMA and 2.92-mm connectors by increasing the outer wall thickness and strengthening the female fingers. It can mate with SMA, 3.5-mm and 2.92-mm with the use of precision adapters. The 2.4-mm product is offered in three quality grades; general purpose, instrument, and metrology. General purpose grade is intended for economy use on components, cables, and microstrip, where limited connections and low repeatability is acceptable. Instrument grade is best suited for measurement applications where repeatability and long life are primary considerations. Metrology grade is best suited for calibration applications where the highest performance and repeatability are required.

1.85-mm Connector

The 1.85-mm connector was developed in the mid-1980s by Hewlett Packard— now Agilent Technologies—for mode-free performance to 65 GHz. HP offered their design as public domain in 1988 to encourage standardization of connector types; a few devices are available from various manufacturers for research work. The 1.85-mm connector mates with the 2.4-mm connector and has the same ruggedness. Many experts have considered this connector to be the smallest possible coaxial connector for common usage up to 65 GHz.

1.0-mm connector

Designed to support transmission all the way to 110 GHz, this 1.0-mm connector is a significant achievement in precision manufacturing resulting in a reliable and flexible interconnect.

BNC connector

The BNC (Bayonet Navy Connector) was designed for military use and has gained wide acceptance in video and RF applications to 2 GHz. Above 4 GHz, the slots may radiate signals. Both 50-ohm and 75-ohm versions are available. A threaded version (TNC) helps resolve leakage for common applications up to 12 GHz.

SMC connector

The SMC (Subminiature C) is much smaller than an SMA connector, making it suitable for some applications with size constraints. It is often used up to 7 GHz where low leakage and few connections are required.

Connector care and signal performance

While many Agilent RF/microwave connectors have been designed for rugged mechanical interfaces, the user must be aware that cleanliness of the surfaces and care in applying torque to the connector nut are crucial to long life and full signal performance. The following table shows the recommended torque for various connector types.

Table: Recommended torque values for connectors

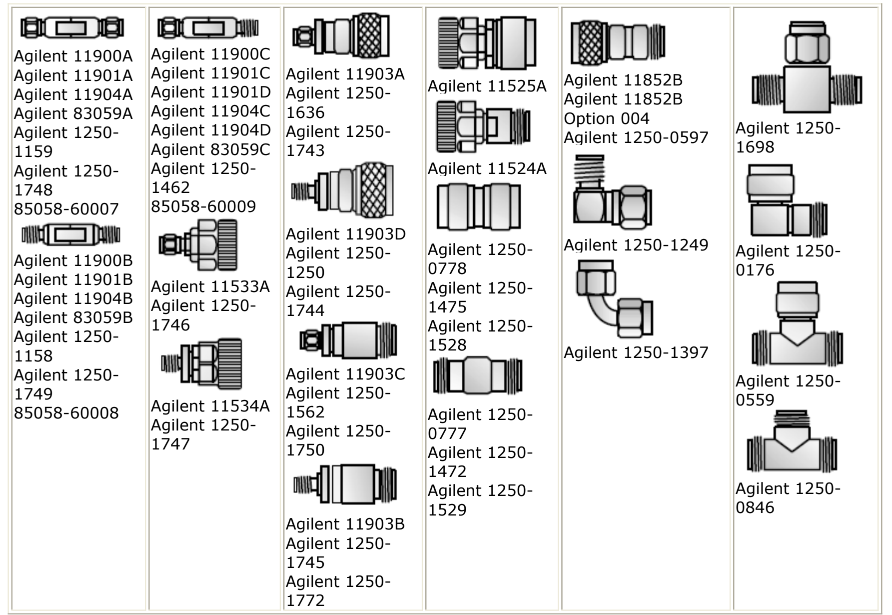
Connector type	**Torque (lb-inch (N-cm))**
Precision 7 mm 12	(136)
Precision 3.5 mm 8	(90)
SMA	5 (56) Use the SMA torque value to connect male SMA connectors to female precision 3.5-mm connectors. Use the 3.5-mm torque value to connect male 3.5-mm connectors to the female SMA (8 lb-inch).
Precision 2.4 mm 8	(90)
Precision 1.85 mm 8	(90)
Type-N	Type-N connectors may be connected finger tight. If a torque wrench is used, 12 lb-inch (136 N-cm) is recommended.

Metrology/Instrument Grade Adapters

Table: Metrology/instrument grade adapter selection guide

Connector type	1.0 mm	1.85 mm	2.4 mm	2.92 mm	3.5 mm	7 mm	Type-N (50 Ω)	Type-N (75 Ω)
1.0 mm	11920A, B, C	11921E, F, G, H	11922A, B, C, D					
1.85 mm		85058-60007 85058-60008 85058-60009						
2.4 mm			11900A, B ,C	11904A, B, C, D 11904S	11901A, B, C, D 1250-2277	11902A, B	11903A, B, C, D	
3.5 mm					83059A, B, C 1250-1748 1250-1749	1250-1746 1250-1747	1250-1743 1250-1744 1250-1745 1250-1750	
Type N (50 Ω)								11852B 11852B Option 004

Typical Configuration

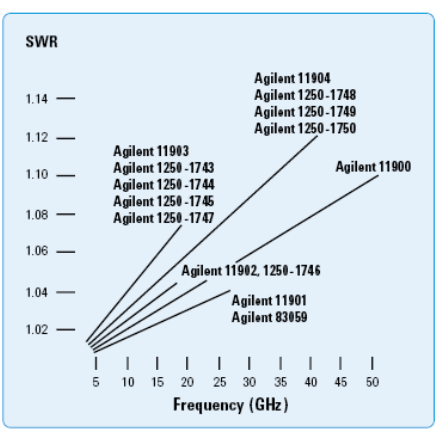


Metrology Grade Adapters

Agilent model	Type [2]	Frequency range	Return loss	Repeatability [3] (min)	Overall length (nom) mm (in)	Ref. plane to ref. plane length (nom) mm (in)	Diameter (nom) mm (in)
11900A	2.4 mm (m), 2.4 mm (m)	DC to 50 GHz	>26 dB	–44 dB	16.2 (0.64)	12.4 (0.49)	9 (0.35)
11900B	2.4 mm (f), 2.4 mm (f)	DC to 50 GHz	>26 dB	–44 dB	18.5 (0.73)	12.4 (0.49)	8 (0.31)
11900C	2.4 mm (m), 2.4 mm (f)	DC to 50 GHz	>26 dB	–44 dB	17.4 (0.69)	12.4 (0.49)	9 (0.35)
11901A	2.4 mm (m), 3.5 mm (m)	DC to 26.5 GHz	>26 dB	–54 dB	20.9 (0.82)	16.1 (0.63)	9 (0.35)
11901B	2.4 mm (f), 3.5 mm (f)	DC to 26.5 GHz	>32 dB	–54 dB	21.1 (0.83)	16.1 (0.63)	8 (0.31)
11901C	2.4 mm (m), 3.5 mm (f)	DC to 26.5 GHz	>32 dB	–54 dB	20.2 (0.80)	16.1 (0.63)	9 (0.35)
11901D	2.4 mm (f), 3.5 mm (m)	DC to 26.5 GHz	>32 dB	–54 dB	21.8 (0.86)	16.1 (0.63)	9 (0.35)
11903A	2.4 mm (m), Type-N (m)	DC to 18 GHz	>28 dB	–48 dB	49.1 (1.93)	46.1 (1.82)	22 (0.86)
11903B	2.4 mm (f), Type-N (f)	DC to 18 GHz	>28 dB	–48 dB	58.3 (2.30)	46.1 (1.82)	15.7 (0.62)
11903C	2.4 mm (m), Type-N (f)	DC to 18 GHz	>28 dB	–48 dB	57.4 (2.26)	46.1 (1.82)	15.7 (0.62)
11903D	2.4 mm (f), Type-N (m)	DC to 18 GHz	>28 dB	–48 dB	50.0 (1.97)	46.1 (1.82)	22 (0.86)
11904A	2.4 mm (m), 2.92 mm (m) [4]	DC to 40 GHz	>24 dB	–40 dB	16.4 (0.64)	11.3 (0.45)	9 (0.35)
11904B	2.4 mm (f), 2.92 mm (f)	DC to 40 GHz	>24 dB	–40 dB	16.3 (0.64)	11.3 (0.45)	8 (0.31)
11904C	2.4 mm (m), 2.92 mm (f)	DC to 40 GHz	>24 dB	–40 dB	13.3 (0.52)	11.3 (0.45)	9 (0.35)
11904D	2.4 mm (f), 2.92 mm (m)	DC to 40 GHz	>24 dB	–40 dB	17.0 (0.67)	11.3 (0.45)	9 (0.35)
11904S	2.4 mm to 2.92 mm matched set						

[1] *Agilent 1190x adapters are phase matched within each family.*
[2] *f = jack, m = plug.*
[3] *Repeatability = –20 Log | Δr |, where | Δr | = | r m₁ – r m₂ |.*
[4] *2.92 mm is compatible with 3.5 mm.*

Typical Precision Adapter Performance



Instrument Grade Adapters

Agilent model	Type [1]	Frequency range	Return loss (typ)	Overall length (nom) mm (in)	Ref. plane to ref. plane length (nom) mm (in)	Diameter (nom) mm (in)
83059A	3.5 mm (m), 3.5 mm (m)	DC to 26.5 GHz	32 dB	28.4 (1.12)	23.1 (0.91)	10 (0.39)
83059B	3.5 mm (f), 3.5 mm (f)	DC to 26.5 GHz	32 dB	26.9 (1.06)	23.1 (0.91)	10 (0.39)
83059C	3.5 mm (m), 3.5 mm (f)	DC to 26.5 GHz	32 dB	25.7 (1.01)	23.1 (0.91)	10 (0.39)
83059K	Set of Agilent 83059A, B, C in wood case					
1250-1743	3.5 mm (m), Type-N (m)	DC to 18 GHz	28 dB	44.2 (1.74)	40.8 (1.61)	20.8 (0.82)
1250-1744	3.5 mm (f), Type-N (m)	DC to 18 GHz	28 dB	43.6 (1.72)	40.8 (1.61)	20.8 (0.82)
1250-1745	3.5 mm (f), Type-N (f)	DC to 18 GHz	28 dB	42.7 (1.68)	31.6 (1.24)	15.8 (0.62)
1250-1746	3.5 mm (m), APC-7	DC to 18 GHz	34 dB	37.9 (1.49) [2]	33.1 (1.30)	22.0 (0.87)
1250-1747	3.5 mm (f), APC-7	DC to 18 GHz	28 dB	37.0 (1.46) [2]	33.1 (1.30)	22.0 (0.87)
1250-1748	3.5 mm (m), 3.5 mm (m)	DC to 26.5 GHz	25 dB	45.1 (1.78)	39.6 (1.56)	9.2 (0.36)
1250-1749	3.5 mm (f), 3.5 mm (f)	DC to 34 GHz	23 dB	43.5 (1.71)	39.6 (1.56)	9.2 (0.36)
1250-1750	3.5 mm (m), Type-N (f)	DC to 18 GHz	24 dB	43.4 (1.71)	31.6 (1.24)	15.8 (0.62)
85058-60007	1.85 mm (m), 1.85 mm (m) [3]	DC to 65 GHz	22 dB	29.5 (1.16)	25.2 (0.99)	9.1 (0.36)
85058-60008	1.85 mm (f), 1.85 mm (f) [3]	DC to 65 GHz	22 dB	31.3 (1.23)	25.2 (0.99)	9.1 (0.36)
85058-60009	1.85 mm (m), 1.85 mm (f) [3]	DC to 65 GHz	22 dB	30.4 (1.20)	25.2 (0.99)	9.1 (0.36)
11852B [4]	50 Ω Type-N (f), 75 Ω Type-N (m)	DC to 3 GHz	30 dB	60.1 (2.37)	50.2 (1.98)	22 (0.87)
11852B Option 004 [4]	50 Ω Type-N (m), 75 Ω Type-N (f)	DC to 3 GHz	30 dB	60.1 (2.37)	50.2 (1.98)	22 (0.87)

[1] *f = jack, m = plug.*
[2] *Overall length with threaded coupling sleeve extended.*
[3] *1.85 mm is compatible with 2.4 mm. To adapt 1.85 mm to other connector types, use Agilent 1190x series adapters.*
[4] *Insertion loss is 5.7 dB typical.*

Selected Instrument Grade Adapters

1 Agilent 1250-1744 Adapter, 3.5 mm (f) to Type-N (m), DC to 18 GHz
2 Agilent 1250-1743 Adapter, 3.5 mm (m) to Type-N (m), DC to 18 GHz
3 Agilent 1250-1747 SMA (f) to APC-7 Adapter
4 Agilent 1250-1746 SMA (m) to APC-7 Adapter
5 Agilent 1250-1750 3.5 mm (m) to Type-N (f)
6 Agilent 1250-1745 3.5 mm (f) to Type-N (f)
7 Agilent 1250-1748 3.5 mm (m) to 3.5 mm (m) Instrument-Grade Adapter
8 Agilent 1250-1749 3.5 mm (f) to 3.5 mm (f)

General Purpose Grade Adapter Selection Guide

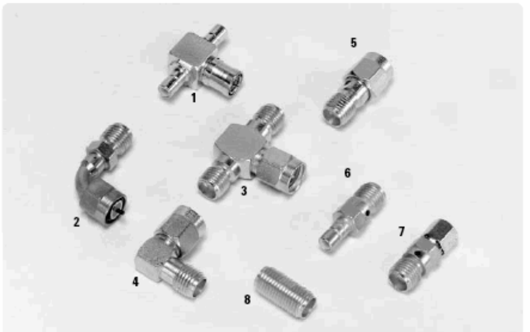
Connector type	1.85 mm	SMA	SMA Tee	SMB	SMC	Type-N (50 Ω)	Type-N (75 Ω)	BNC (75 Ω)	Type-N Tee	BNC (50 Ω)
1.85 mm	N5520A, B, C									
SMA		1250-1158 1250-1159 1250-1462 1250-1694		1250-0674	1250-0675					1250-0562 1250-1200
Right Angle, SMA		1250-1249 1250-1397 1250-1741								
SMA Tee			1250-1698							
SMB		1250-0674		1250-0672 1250-1391		1250-0671				1250-1857
SMC		1250-0675			1250-0827 1250-0837 1250-0838 1250-1113	1250-1152				
7 mm		11533A 11534A 1250-1468				11524A 11525A				
BNC (50 Ω)		1250-1200 1250-0562		1250-1236 1250-1237 1250-1899	1250-0831 1250-0832					
Type-N (50 Ω)		1250-1250 1250-1404 1250-1636 1250-1772			1250-1152	1250-1529 1250-0777 1250-0778 1250-1472 1250-1475	1250-0597			1250-1473 1250-1474 1250-1476 1250-1477
Type-N (75 Ω)								1250-1533 1250-1534 1250-1535 1250-1536		
Right Angle, Type-N (50 Ω)						1250-0176				
Type-N Tee									1250-0559 1250-0846	
BNC (75 Ω)								1250-1286 1250-1287		
BNC Trixial										1250-0595 1250-1830 1250-1930

Adapters APC-7 [1]	
11524A	APC-7 to Type-N (f)
11525A	APC-7 to Type-N (m)
11533A	APC-7 to SMA (m)
11534A	APC-7 to SMA (f)
Adapters Type-N, standard 50 Ω	SWR <1.03 to 1.3 GHz
1250-1472	Type-N (f) to Type-N (f)
1250-1473	Type-N (m) to BNC (m)
1250-1474	Type-N (f) to BNC (f)
1250-1475	Type-N (m) to Type-N (m)
1250-1476	Type-N (m) to BNC (f)
1250-1477	Type-N (f) to BNC (m)
Adapters SMA	
1250-1158	SMA (f) to SMA (f)
1250-1159	SMA (m) to SMA (m)
1250-1249	SMA right angle (m) (f)
1250-1397	SMA right angle (m) (m)
1250-1462	SMA (m) to SMA (f)
1250-1698	SMA tee (m) (f) (f)
1250-1200	BNC (f) to SMA
E9633A	SMA (m) to BNC (m)
1250-1899	BNC (f) to SMB (m)
E9634A	SMA (f) to BNC (m)

Adapters Type-N, standard 50 Ω	
1250-0077	Type-N (f) to BNC (m)
1250-0082	Type-N (m) to BNC (m)
1250-0176	Type-N (m) to Type-N (f) right angle (use below 12 GHz)
1250-0559	Type-N tee, (m) (f) (f)
1250-0777	Type-N (f) to Type-N (f)
1250-0778	Type-N (m) to Type-N (m)
1250-0780	Type-N (m) to BNC (f)
1250-0846	Type-N tee (f) (f) (f)
1250-1250	Type-N (m) to SMA (f)
1250-1562	Type-N (f) to SMA (m)
1250-1636	Type-N (m) to SMA (m)
1250-1772	Type-N (f) to SMA (f)
Adapters Type-N, standard 75 Ω [2]	
1250-0597	Type-N (m) (50 Ω) to Type-N (f) (75 Ω)
1250-1528	Type-N (m) to Type-N (m)
1250-1529	Type-N (f) to Type-N (f)
1250-1533	Type-N (m) to BNC (m)
1250-1534	Type-N (f) to BNC (m)
1250-1535	Type-N (m) to BNC (f)
1250-1536	Type-N (f) to BNC (f)
Adapters type BNC, standard 50 Ω	
1250-0076	Right angle BNC (UG-306/D)
1250-0080	BNC (f) to BNC (f) (UG-914/U)
1250-0216	BNC (m) to BNC (m)
1250-0556	BNC (f) to WECO video (m)
1250-0595	BNC (f) to BNC triaxial (m)
1250-0781	BNC tee (m) (f) (f)
1250-1830	BNC (f) to BNC triaxial (f)
1250-1930	BNC (m) to BNC triaxial (f)
Adapters BNC, standard 75 Ω [3]	
1250-1286	Right angle BNC (m) (f)
E9628A	BNC (f) to BNC (f)
1250-1288	BNC (m) to BNC (m)
Adapters SMB, SMC [4]	
1250-0670	SMC tee (m) (m) (m)
1250-0671	SMB (m) to Type-N (m)
1250-0672	SMB (f) to SMB (f)
1250-0674	SMB (m) to SMA (f)
1250-0675	SMC (m) to SMA (f)
1250-0827	SMC (m) to SMC (m)
1250-0831	SMC (m) to BNC (m)
1250-0832	SMC (f) to BNC (f)
1250-0837	SMC tee (m) (m) (m)
1250-0838	SMC tee (f) (m) (m)
1250-1023	SMC (m) to Type-N (m)
1250-1113	SMC (f) to SMC (f)
1250-1152	SMC (f) to Type-N (m)
1250-1236	SMB (f) to BNC (f)
1250-1237	SMB (m) to BNC (f)
1250-1391	SMB tee (f) (m) (m)
1250-1857	SMB (f) to BNC (m)

[1] APC-7 is a registered trademark of the Bunker Ramo Corporation.
[2] Type-N outer conductor, center pin sized for 75 Ω characteristic.
[3] BNC outer conductor, center pin sized for 75 Ω characteristic.
[4] SMB and SMC are often used inside Agilent instruments for inter-module RF connections. SMB is snap-on configuration. SMC is screw-on configuration.

1 Agilent 1250-1200 Adapter, BNC (f) to SMA (m)
2 Agilent 1250-1899 Adapter, BNC (f) to SMB (m)
3 Agilent 1250-0556 Adapter, BNC (f) to WECO Video (m)
4 Agilent 1250-1477 Standard, N (f) to BNC (m), Precision 50 Ω
5 Agilent 1250-1473 Standard, N (m) to BNC (m), Precision 50 Ω Adapter
6 Agilent 1250-0595 Adapter, BNC (f) to BNC Triaxial (m)
7 Agilent 1250-1930 Adapter, BNC (m) to BNC Triaxial (f)
8 Agilent 1250-1830 Adapter, BNC (f) to BNC Triaxial (f)
9 Agilent 1250-1857 Adapter, SMB (f) to BNC (m)
10 Agilent 1250-0562 Adapter, BNC (f) to SMA (f)
11 Agilent 1250-1236 Adapter, SMB (f) to BNC (f)



1 Agilent 1250-1391 Adapter, SMB Tee (f) (m) (m)
2 Agilent 1250-1741 SMA (f) to SMA (m) Right Angle Adapter
3 Agilent 1250-1698 Adapter, SMA Tee (m) (f) (f)
4 Agilent 1250-1249 Adapter, SMA Right Angle (m) (f)
5 Agilent 1250-1462 Adapter, SMA (m) to SMA (f)
6 Agilent 1250-0674 Adapter, SMB (m) to SMA (f)
7 Agilent 1250-1694 SMA (m) to SMA (f) Adapter
8 Agilent 1250-1158 SMA (f) to SMA (f) Adapter

1 Agilent 1250-0597 Adapter, Type-N (m) 50 Ω to Type-N (f) 75 Ω
2 Agilent 1250-1778 Standard N (m) to Standard N (m) Adapter, 50 Ω
3 Agilent 1250-1529 Standard N (f) to Standard N (f) Adapter, 75 Ω
4 Agilent 1250-1152 Adapter, SMC (f) to Type-N (m)
5 Agilent 1250-1404 Adapter, SMA (f) to Type-N (f)
6 Agilent 1250-1023 Adapter, SMC (m) to Type-N (m)
7 Agilent 1250-1535 Adapter, N (m) to BNC (f) Adapter, 75 Ω
8 Agilent 1250-1533 Standard N (m) to BNC (m) Adapter, 75 Ω
9 Agilent 1250-1250 Adapter, Type-N (m) to SMA (f), 50 Ω
10 Agilent 1250-0846 Tee Adapter, Standard N (f) (f) (f)
11 Agilent 1250-1636 Adapter, Type-N (m) to SMA (m) 50 Ω
12 Agilent 1250-0559 Tee Adapter, Standard N (m) (f) (f)
13 Agilent 1250-0176 Right Angle Standard N (m) to Standard N (f)

Adapter Kit Selection Guide

Connector type	3.5 mm	7 mm	Type-N (50 Ω)	Type-N (75 Ω)	BNC (75 Ω)	Type-F (75 Ω)	BNC (50 Ω)	7-16
3.5 mm	83059K		11878A					
Type-N (50 Ω)			11853A				11854A	
Type-N (75 Ω)				11855A 86213A	11856A	86211A		
7-16	11906D	11906C	11906B					11906A

1.0 mm Adapters

- Increased measurement versatility
- Ease-of-use for on-wafer and coaxial measurements

Increased Measurement Versatility

For Microwave and RF engineers making coaxial measurements at 50, 67 or 110 GHz, the

Agilent 11920/1/2 series 1.0 mm adapters provide an easy way of measuring coaxial devices at high frequencies. The Agilent 11920 A/B/C 1.0 mm to 1.0 mm are designed for the measurement of components with 50 Ω 1.0 mm connectors. The Agilent 11921 A/B/C/D, 1.0 mm to 1.85 mm, and the Agilent 11922 A/B/C/D, 1.0 mm to 2.4 mm, are intended to be used as general purpose adapters that are versatile and interchangeable. These adapters increase the capability needed to use test systems, such as the Agilent N5250A.

Ease-of-use for on-wafer and Coaxial Measurements

Each connector has an air dielectric interface and a center conductor that is supported by a low-loss plastic bead. Available with male and female connectors, these Agilent 1.0 mm adapters provide ease-of-use for microwave engineers who need to connect their test systems. The Agilent 1.0 mm adapters allow engineers to make fewer connections directly to their test port while maintaining the accuracy of their test system.

1.0 mm Adapters Table

Agilent model	11920A 11920B 11920C	11921E 11921F 11921G 11921H	11922A 11922B 11922C 11922D	11923A
Features	← Excellent accuracy and measurement versatility →			
Frequency range	DC to 110 GHz	DC to 67 GHz	DC to 50 GHz	DC to 110 GHz
Frequency response				
Insertion loss	–0.5 dB	–0.5 dB	–0.7 dB	–1.0 dB
Return loss	–24 dB DC to 20 GHz –20 dB 20 to 50 GHz –18 dB 50 to 75 GHz –14 dB 75 to 110 GHz	–20 dB	–20 dB	–16 dB
Input power				
Max CW power	10 W	10 W	10 W	6 W
Repeatability [1]	–35 dB	–35 dB 1.0 mm –40 dB 1.85 mm	–35 dB 1.0 mm –44 dB 2.4 mm	
RF connectors				
A, E:	1 mm (m) to 1 mm (m)	1 mm (m) to 1.85 mm (m)	1 mm (m) to 2.4 mm (m)	1 mm (f) to circuit card launch
B, F:	1 mm (f) to 1 mm (f)	1 mm (f) to 1.85 mm (f)	1 mm (f) to 2.4 mm (f)	
C, G:	1 mm (m) to 1 mm (f)	1 mm (m) to 1.85 mm (f)	1 mm (m) to 2.4 mm (f)	
D, H:		1 mm (f) to 1.85 mm (m)	1 mm (f) to 2.4 mm (m)	

[1] *Measured at 25° C.*

Flexible Microcircuit Packaging

The Agilent 11923A 1.0 mm female connector launch threads into a package or fixture housing to transition a microwave circuit from microstrip to coaxial connector. The Agilent 11923A connector launch is intended for use with the N5250A and other test systems up to 110 GHz. The Agilent 11923A 1.0 mm female connector has an air dielectric interface and center conductor that is supported by a low-loss plastic bead on one end and a glass-tometal seal interface on the other end. This interface consists of a 0.162 mm diameter pin that extends inside the package or fixture for connection onto a microwave circuit. The Agilent 11923A is pre-assembled and supplied with a machining detail for mounting the launch and assembly instructions. The user is responsible for making the connection onto the circuit card, machining the package, and installing the connector. If a quasi-hermetic seal is desired, epoxy may be applied to threads of the launch prior to installation. The procedure describing the necessary dimensions for the package and installation is provided with the launch assembly.

Specifications

Specifications describe the instrument's warranted performance over the temperature range 0 to 55° C (except where noted). Supplemental characteristics are intended to provide information for applying the instrument by giving typical but nonwarranted performance parameters. These are noted as "typical", "nominal", or "approximate".

1.0 mm (f) Connector Launch

Supplemental Characteristics

Model number	Return loss	Max CW power
11923A	−16 dB	better than: 6 W

Environmental Specifications

	Operating	Non-operating
Temperature	0° to 55°C	−40° to 75° C
Altitude	<15.000 meters (<50.000 feet)	<15.000 meters (<50.000 feet)

Note
The operating temperature is a critical factor in the performance during measurements and between calibrations. Storage or operation within an environment other than that specified above may cause damage
to the product and void the warranty.
Non-operating environmental specifications apply to storage and shipment. Products should be stored in a clean, dry environment. Operating environmental specifications apply when the product is in use. Products should
not be operated in a condensing environment.

Key literature

Agilent 11923A Operating and Service Guide 11923-90001

Slotless Connectors

Precision Slotless sockets (female connectors) were developed by Agilent to provide the most accurate traceable calibration possible. Connectors that use precision slotless sockets are metrology grade connectors. The outside diameter of the socket does not change
when mated with pins of varying diameters, within the tolerance requirements of a metrology grade connector. Conventional slotted sockets are flared by the inserted pin. Because physical dimensions determine connector impedance, electrical characteristics of the connector pair are dependent upon the mechanical dimensions of the pin. While connectors are used in pairs, their pin and socket halves are always specified separately as part of a standard, instrument, or device under test. Because the slotted socket's outer diameter changes with different pin diameters, it is very difficult to make precision measurements with the conventional slotted socket connector. The measurement of the device is a function of its connector.

Slotless sockets are used in the following calibration kits:

- Agilent 85052B/C/D
- Agilent 85054B/D
- Agilent 85056A/D

Coaxial Mechanical Calibration kits

Connector	Frequency range	Type	VNA calibration accuracy	Agilent model	Available options
Type-F (75 Ω)	DC to 3 GHz	Economy	5% – 1%	85039B	1A7, A6J, UK6, 00M, 00F, MOF
Type-N (75 Ω)	DC to 3 GHz	Economy	5% – 1%	85036E	1A7, A6J, UK6
Type-N (75 Ω)	DC to 3 GHz	Standard	5% – 1%	85036B	1A7, A6J, UK6
Type-N (50 Ω)	DC to 6 GHz	Economy	5% – 1%	85032E	1A7, A6J, UK6
Type-N (50 Ω)	DC to 9 GHz	Standard	5% – 1%	85032F	1A7, A6J, UK6, 100, 200, 300, 500*
Type-N (50 Ω)	DC to 18 GHz	Economy	5% – 1%	85054D	1A7, A6J, UK6
Type-N (50 Ω)	DC to 18 GHz	Standard	2% – 0.3%	85054B	1A7, A6J, UK6
7-16	DC to 7.5 GHz	Standard	2%	85038A	UK6
7-16 (female)	DC to 7.5 GHz	Standard	2%	85038F	UK6
7-16 (male)	DC to 7.5 GHz	Standard	2%	85038M	UK6
7 mm	DC to 6 GHz	Economy	2% – 0.3%	85031B	1A7, A6J, UK6
7 mm	DC to 18 GHz	Economy	5% – 1%	85050D	1A7, A6J, UK6
7 mm	DC to 18 GHz	Standard	2% – 0.05%	85050B	1A7, A6J, UK6
7 mm	DC to 18 GHz	Precision	0.3% – 0.05%	85050C	1A7, A6J, UK6
3.5 mm	DC to 9 GHz	Standard	5% – 1%	85033E	1A7, A6J, UK6, 100, 200, 300, 400, 500
3.5 mm	DC to 26.5 GHz	Economy	5% – 1%	85052D	1A7, A6J, UK6
3.5 mm	DC to 26.5 GHz	Standard	3% – 0.5%	85052B	1A7, A6J, UK6
3.5 mm	DC to 26.5 GHz	Precision	2% – 0.5%	85052C	1A7, A6J, UK6
2.92 mm	DC to 50 GHz	Economy	11% – 4% (Option 001 65% – 3%)	85056K	1A7, A6J, UK6, 001
2.4 mm	DC to 50 GHz	Economy	5% – 1%	85056D	1A7, A6J, UK6
2.4 mm	DC to 50 GHz	Standard	4% – 0.5%	85056A	1A7, A6J, UK6
1.85 mm	DC to 67 GHz	Economy		85058E	1A7, A6J, UK6
1.85 mm	DC to 67 GHz	Standard		85058B	1A7, A6J, UK6
1 mm	DC to 110 GHz	Precision	5% – 1%	85059A	1A7, A6J, UK6

Option description:

1A7	ISO 17025 compliant calibration
A6J	ANSI Z540 compliant calibration
UK6	Commercial calibration certificate with test data
00M	Includes male standards & male-male adapter
00F	Includes female standards and female-female adapter
MOF	Includes male and female standards and adapters
001	Adds 2.4 mm sliding load and 2.4 mm gauges
100	Includes female-female adapter
200	Includes male-male adapter
300	Includes male-female adapter
400	Adds four 3.5 mm to Type-N adapters
500	Adds four 7 mm to 3.5 mm adapters
500*	Adds four 7 mm to Type-N adapters

Waveguide Mechanical Calibrations Kits

Connector	Frequency range	Type	VNA calibration accuracy	Agilent model	Available options
WR-90	8.2 to 12.4 GHz	Precision	0.3% – 0.05%	X11644A	1A7, A6J, UK6
WR-62	12.4 to 18 GHz	Precision	0.3% – 0.05%	P11644A	1A7, A6J, UK6
WR-42	18 to 26.5 GHz	Precision	0.3% – 0.05%	K11644A	1A7, A6J, UK6
WR-28	26.5 to 40 GHz	Precision	0.3% – 0.05%	R11644A	1A7, A6J, UK6
WR-22	33 to 50 GHz	Precision	0.3% – 0.05%	Q11644A	1A7, A6J, UK6
WR-19	40 to 60 GHz	Precision	0.3% – 0.05%	U11644A	1A7, A6J, UK6
WR-15	50 to 75 GHz	Precision	0.3% – 0.05%	V11644A	1A7, A6J, UK6
WR-10	75 to 110 GHz	Precision	0.3% – 0.05%	W11644A	1A7, A6J, UK6

Coaxial Electronic Calibration Kits (ECal)

Option Description

1A7	ISO 17025 compliant calibration
A6J	ANSI Z540 compliant calibration
UK6	Commercial calibration certificate with test data
00M	Includes male standards & male-male adapter
00F	Includes female standards and female-female adapter
MOF	Includes male and female standards and adapters
00A	Add type-N adapters
00A	[1] Add 7 - 16 adapters
00A	[2] Add 3.5 mm adapters
00A	[3] Add 2.92 mm adapters
00A	[4] Add 2.4 mm adapters
00A	[5] Add 1.85 mm adapters
001	Adds data for Agilent 8702 Lightwave Component Analyzer

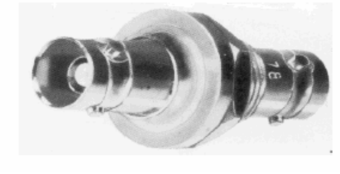
Mechanical Verification Kits

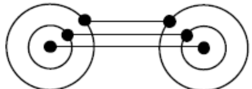
Connector	Frequency range	Type	VNA calibration accuracy	Agilent model	Available options
Type-N (50 Ω)	300 kHz to 18 GHz	Precision	N/A	85055A	1A7, A6J, UK6
7 mm	300 kHz to 6 GHz	Precision	N/A	85029B	1A7, A6J, UK6, 001
7 mm	300 kHz to 18 GHz	Precision	N/A	85051B	1A7, A6J, UK6
3.5 mm	300 kHz to 26.5 GHz	Precision	N/A	85053B	1A7, A6J, UK6
2.4 mm	0.045 to 50 GHz	Precision	N/A	85057B	1A7, A6J, UK6
1.85 mm	0.010 to 67 GHz	Precision	N/A	85058V	1A7, A6J, UK6
WR-28	26.5 to 40 GHz	Precision	N/A	R11645A	1A7, A6J, UK6
WR-22	33 to 50 GHz	Precision	N/A	Q11645A	1A7, A6J, UK6
WR-19	40 to 60 GHz	Precision	N/A	U11645A	1A7, A6J, UK6
WR-15	50 to 75 GHz	Precision	N/A	V11645A	1A7, A6J, UK6
WR-10	75 to 110 GHz	Precision	N/A	W11645A	1A7, A6J, UK6

Option Description

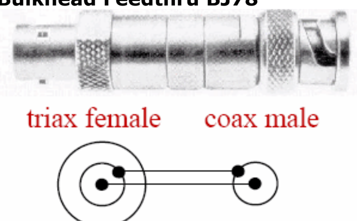
1A7	ISO 17025 compliant calibration
A6J	ANSI Z540 compliant calibration
UK6	Commercial calibration certificate with test data
00M	Includes male standards & male-male adapter
00F	Includes female standards and female-female adapter
MOF	Includes male and female standards and adapters
00A	Add type-N adapters
00A	[1] Add 7 - 16 adapters
00A	[2] Add 3.5 mm adapters
00A	[3] Add 2.92 mm adapters
00A	[4] Add 2.4 mm adapters
00A	[5] Add 1.85 mm adapters
001	Adds data for Agilent 8702 lightwave component analyzer

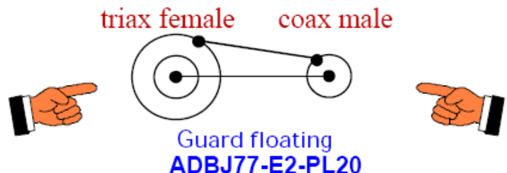
Triax Feedthru and Triax-Coax Converters from Trompeter Electronics





Bulkhead Feedthru BJ78





Introduction

Coplanar GHz probes have been available only since 1983, and have significantly advanced wafer probing capabilities. Wafer probing at GHz frequencies has developed over the past several years from a simple R&D toy to a sophisticated high-volume manufacturing tool used in GaAs, silicon, packaging, and hybrid industries. Successful GHz probing requires that consideration be given to layout and design before design completion and mask fabrication. Failure to observe specific layout requirements can result in the inability to test devices with GHz probes. This Application Note focuses on layout and testing issues as they apply to the use of GHz coplanar probes, Discussed first are specific mechanical and electrical layout rules, rules that will to assure that a fabricated device can be successfully probed. Following the layout rules are design/layout guidelines for current major applications of GHz-wafer-probing technology-microwave monolithic integrated circuits (MMICs), high-speed digital ICs, individual characterization devices for process characterization and modeling, and high-speed hybrids and packages. The rules in this Application Note will help you achieve the best possible tests.

It is useful to stop and consider what you really wish to measure. When GHz probes are used with a network analyzer (after calibration), what is measured will be what is contacted by the probe tips. This includes the pad parasitics, and parasitics associated with the interconnects from the pad to the other devices on your substrate. If you want to measure a device independent of the pad effects, then it is useful to make the pads small, so their effects will be negligible or easily subtracted from the measured data. In any event, it is important to be clear that when using a network analyzer you will be measuring everything past the end of the probe tips, unless you use special calibration/correction techniques. In most applications this is exactly what you want; a measurement the circuit's electrical performance.

This Application Note applies to the following Cascade Microtech probe series:

- WPH-Oxx -18 GHz probes
- WPH-1 xx- 26 GHz probes
- RTP-105-xxx-26 GHz replaceable-tip probes
- WPH-2xx-50 GHz probes
- WPH-3xx -`40 GHZ probes
- WPH-7xx -multicontact probes
- WPH-9xx -multiple-needle probes

Probe Features Affect Layout

A coplanar probe's features (Fig. 1) include the probe body, the coaxial connector, the probe tip, and the contacts at the probe tip end. Coplanar transmission lines carry the signal between the coaxial connector and the probe tip contacts.

Figure 1: Coplanar probe

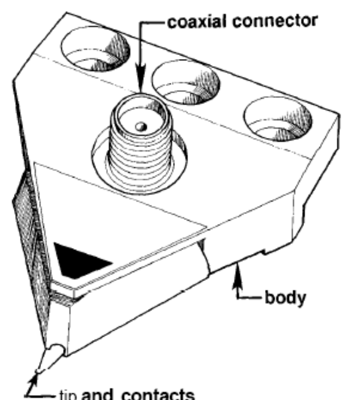
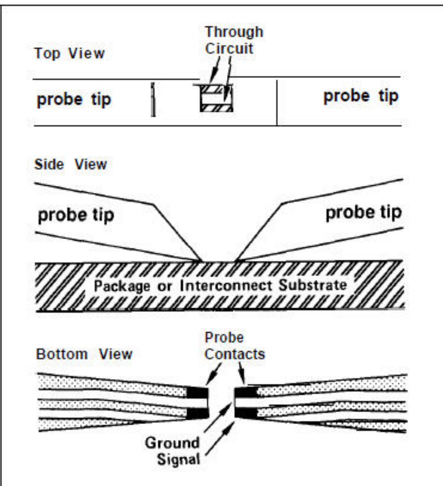


Figure 2: Probe tip: top, bottom, and side view



The transition from coaxial, with a radial electrical field pattern (Fig. 3a), to a coplanar transmission line with a much different electrical field pattern (Fig. 3b) is made within the probe. This is a difficult transition to make at GHz frequencies, and presented difficulties in successfully designing this type of probe. The best DUT tests result if the electric field patterns at the probe tip are similar to the field patterns in the DUT. If transitions are necessary, they should be made cleanly.
The typical probe contact is a signal (S) or ground (G) contact. The signal contacts are electrically connected to a coaxial connector center pin, and the ground contacts are electrically connected to the coaxial-connector body. At high frequencies the ground is not an equipotential reference, as in low frequency circuits, but is a part of a transmission line that contains time-varying electric fields. The 700-series multicontact probes also permit specified contacts to be bypassed ("P"-discussed later in this Application Note), and terminated ("T"). (See the Cascade Microtech Probe Head Selection Guide for additional details).

Figure 3: Electric field patterns in coaxial and coplanar transmission lines.

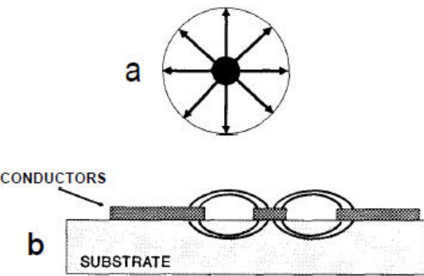
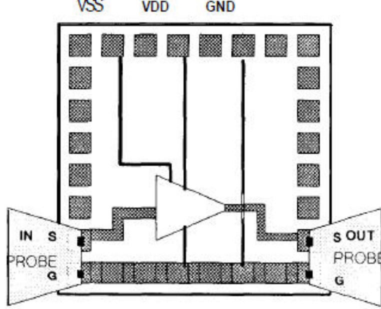


Figure 4 shows an example MMIC pad footprint. Note the
input is on the left, the output on the right, and that the power and control signals at the top are provided by low-frequency needle-type probes. The GHz probes are GS (ground-signal) probes, useful to 26 GHz. The bottom row of pads are on the same bus, providing a low-impedance ground for the two GHz probe ground contacts. When packaged, there will be several bond wires connecting the ground buses to the package ground. This minimizes common-ground inductance. Also, note that the amplifier is positioned close to the ground bus in order to minimize common-ground inductance. In practice, the individual transistors would be positioned very close to the ground bus.

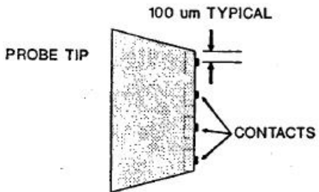
Figure 4: Typical MMIC padout, ground-signal probes in contact.



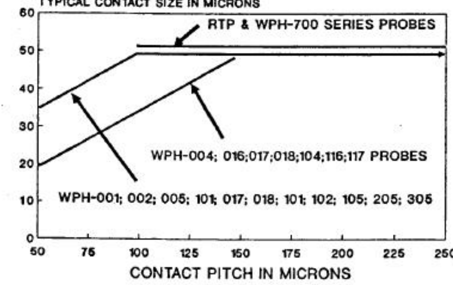
There are several physical features of the coplanar probes that effect the layout. The first of these is the extension of the probe tip substrate material past the center of the contact closest to the edge. This dimension is typically 100µm.

Note
This page (after Figure 4, is a screen shot and this need to be typed.

There are several physical features of coplanar probes that effect the layout. The first phase of these (Fig.5) is the extension of the probe tip substrate material past the center of the contact closest to the edge. This dimension is typically 100µm. Probes must not be allowed to collide, so this 100µm extension must be considered when laying out pads.



Another physical layout consideration is the proble contact size (Fig.6), typically 50 x 50µm. Note that these contacts are much larger than needle probes, and the passivation windows must be large enough to accomodate these contacts plus typical probe system variations.



The final consideration is the probe skating distance. For every 50µm of overtravel (overtravel is the continued downward movement after the probe tip has made initial contact with the wafer) the probe contacts will skate laterally 10µm. If the probes are too close together, they could skate into each other and be damaged. The layout rules will help you prevent probe collisions.

Sequence of Layout Events

The following are guidelines to the typical sequence of events involved in layout for testing with GHz probes. In many cases these guidelines will improve the packaged device performance. Specific requirements may result in these guidelines being altered.

1. **Select basic pad footprint** for signal, power, control, and ground pads to be consistent with input/output type (see step 2) and package requirements. Example pad footprints are shown throughout this Application Note.
2. **Select high frequency input/output type** to meet your specific frequency, design, and signal count requirements; either signal ground (GSG), or signal-signal (SS).
3. **Orient high-gain signal pads** away from each other. Typical microwave circuits place inputs on the left, and outputs on the right.
4. **Use at least one ground pad for each GHz probe**. The grounds pads for each GHz probe must be electrically connected together on the DUT to provide a low-impedence, common-ground node.
5. **Use additional ground pads between signal pins** to obtain reduced crosstalk (as required) within multicontact GHz probes.
6. **Verify that pad pitch matches available probe pitch** (see Cascade Microtech Probe Head Selection Guide). Also, verify that the pad footprint meets packaging requirements.
7. **Verify that pad pitch matches avaialble probe pitch** (see Cascade Microtech Probe Head Selection Guide). Also, verify that the pad footprint meets packaging requirements.
8. **Call the Cascade Microtech Application Group** to verify that the footprint is a good layout for probing. If you want to violate some of the rules, please call the Cascade Microtech Applications Group regarding the trade-offs. Most of the rules can be successfully violated, provided you meet certain conditions.

Layout Rules, Mechanical

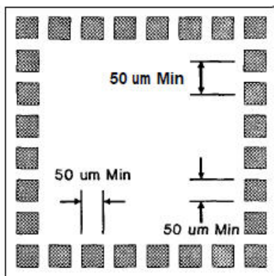
This section describes the mechanically related rules regarding probe pad placement and sizing. In most designs the probe pads are also used as bond pads, so the final layout should satisfy requirements for both probing and bonding. A good layout for GHz probing is usually also a good layout for packaging.
There are two important GHz-probe mechanical parameters which directly affect the mechanical layout rules; the contact dimensions and the substrate extension past the contact. Typical contact size for WPH-series probes (Fig. 6) is 50 x 50µm for contact pitch greater than 100 µm,tapering to 35 x 35 µm as the contact pitch approaches 50 µm.The contact size is important because the passivation window must be large enough to allow the entire contact to touch the pad. The substrate extension is typically 100 µm past the

contact center (Fig. 5).
Coplanar-waveguide GHz probes are constructed with all contacts in a straight line, and with equal spacing between contacts. This means that all pads, arranged for an individual GHz probe measurement, must be in a straight line, have equal pitch, and be on the substrate top. Non-square or very large pads may be laid out more randomly, as long as there is a minimum-size land area, spaced at an equal pitch for the probe contacts to land on.
Top-side pads connected to substrate vias must be used to contact backside grounds. All pads must have as small as possible height variation, because the GHz-probe construction allows little contact height variation. If GHz probes crash into each other, they can be damaged. The pad layout must allow for sufficient damage-avoidance space between GHz probes. Provide sufficient space for probes that are oriented 90 degrees to each other, as well as between probes that face each other. There must be sufficient space between probes to allow for probe skating. The following pad layout rules allow for vertical overtravel of 20 mils (500 µm).Note that 20 mils is the maximum rated overtravel for Cascade Microtech GHz probes, and that the recommended nominal overtravel (to achieve assured contact) for the typical probe is 2-4 mils (50-l00µm).

The mechanical layout rules are given three-digit numbers, starting with 100. This number also indicates the applicable dimension within the referenced figures. Note that the minimum and maximum values given are the recommended values for general testing conditions. If you have a specific need outside these values, please call the Cascade Microtech Applications group for consultation.

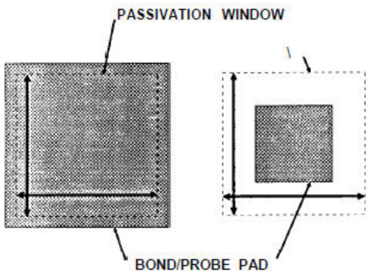
Rule 100 Pad size The minimum pad size is 50 x 50 µm(Fig. 7). The recommended minimum pad size for general use is 100 x 100 µm.Note that because the probe contacts are fairly large, in many situations pads smaller than 50 x 50 µm can be successfully probed. Unless you have a good reason to do otherwise, use 100 x 100 µm pads with a 150 +rnpitch. This provides for easy probing, and parasitic are low enough for many applications.

Figure 7 Rules 100 and 102 Minimum pad size is 50 x 50µm Minimum pad pitch is 50µm



Rule 101 Passivation window The minimum passivation window size is 96 x 96 km(Fig. 8). Smaller passivation windows are acceptable for probe pitch values less than 100 km,because the contact size decreases as contact pitch becomes less than 100 wrn(Fig. 6). Note that for small pad sizes, the passivation window will be significantly larger than the pad. If the pad metalization is above the final passivation layer, this rule does not apply.

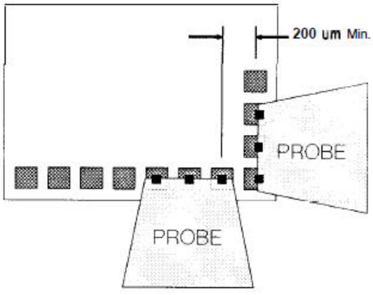
Figure 8 Rule 101 Minimum passivation window size is 96 x 96 µm.



Rule 102 Pad pitch The minimum center-to-center pad pitch is 50µm I*m(Fig. 7). The recommended minimum is 150 µm.Note that many probes have a 100 µm minimum pitch specification. (See the Cascade Microtech Probe Head Selection Guide for additional details.)

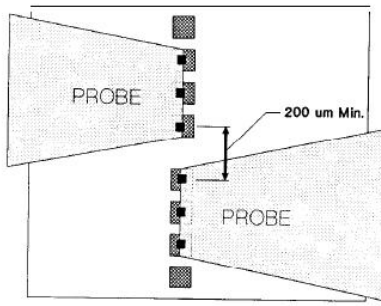
Rule 103 Orthogonal-row pad spacing The minimum center-to-center pad spacing for orthogonally oriented probes (in a corner) is 200 µm(Fig. 9). Sufficient clearance is often achieved by eliminating the corner pad (Fig. 15).

Figure 9 Rule 103 Minimum center-to-center pad spacing in a corner is 200 µm.



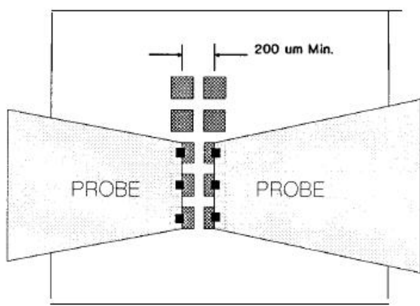
Rule 104 Single-row pad spacing The minimum center-to-center pad spacing for opposing, side-by-side probes that are contacting the same line of pads, (Fig. 10)is 200 µm.

Figure 10 Rule 104 Minimum pad spacing with opposing, side-by-side probes is 200 µm.



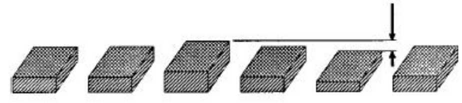
Rule 105 Parallel-row pad spacing The minimum center-to-center pad spacing between facing probes on parallel rows of pads (Fig. 11) is 200 km. Note that this spacing is based on the assumption of 20 mils (500 µm)of probe over-travel. If the probe overtravel can be limited to 10 mils (250 µm), then the minimum spacing is 100 brn.

Figure 11 Rule 105 Minimum center-to-center pad spacing for parallel rows of pads with facing probes is 200 µm.



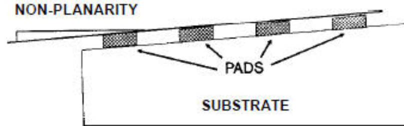
Rule 106 Pad height variation The maximum pad height variation in a row of pads contacted by one probe is 0.5 µm (Fig. 12). Pad height variation usually occurs because pads are constructed with different metal stacks. Be sure to use the same metal layers to construct all pads.

Figure 12 Rule 106 Maximum pad height variation in a row of pads contacted by one probe is 0.5 µm.



Rule 107 Planarity requirements The maximum overall planar deviation of a row of pads contacted by one probe, with respect to the backside of the substrate, is 2 parts per 1,000 for WPH-700 series multicontact probes, and 5 parts per 1,000 for other probes (Fig. 13). Non-planarity can be caused by using different metal stacks for the pads within a row, or incorrectly backlapping a wafer. Also note that during probing, non-planarity between the probe contacts and the pads should not exceed these maximum planar deviations.

Figure 13 Rule 107 Maximum non-planarity from a row of pads to the backside of substrate is 2/1,000 for WPH-700 probes, and 5/1,000 for all other probes.

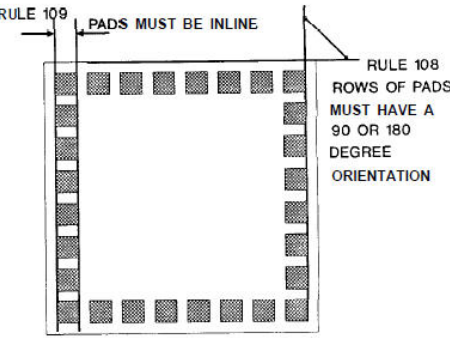


Rule 108 Pads at 90 degrees only Rows of pads meeting at an angle (Fig. 14) must be orthogonal (at 90 degrees).

Rule 109 Pads are collinear All pads contacted by an individual probe must be collinear (Fig. 14).

Figure 14 Rule 108-ROWS of pads meeting at an angle must be at 90 degrees. Rule 109- Pads

contacted by one probe must be collinear.



Rule 110 Pads have equal pitch All pads contacted by an individual probe must have constant pitch. Note that the WPH-700 series multi-contact probes allow for contact removal, and therefore pad removal.

Rule 111 Pads on top All pads must be on the substrate top.

Layout Rules, Electrical

This section describes the electrically related rules regarding pad placement and sizing. In most designs the pads are also used as bond pads, so the final layout should satisfy requirements for both probing and bonding.
The following rules are given three-digit numbers starting with 200.
Compared with mechanically related layout rules, the electrical considerations are more complex. Included in the electrical considerations are:

- Probe grounds;
- Signal input/output;
- Crosstalk between probes;
- Crosstalk between signal lines within a probe;
- Maximum operating frequency for each probe type:
- Power and ground to DUT during probing;
- Maximum current per contact.

Rules 200-205

Rule 200 Pads on top All pads must be on the substrate top surface, and all nodes to be measured or stimulated must connect to these pads. Rule 201 Probe Grounds Each probe must have at least one ground contact.

Rule 202 All grounds connected All probe ground contacts must be electrically connected together on the DUT (Fig 15) This connection should be as short a connection as possible, presenting a low impedance at high frequencies. A useful method for connecting probe grounds is shown in Fig. 16. This approach uses L-shaped metalization to connect the probe grounds. Note that this method requires the edge contacts to be defined as ground for each probe.

Figure 15 Rule 201 -All probes must have at least one ground. Rule 202-All probe grounds must be connected together on the DUT.

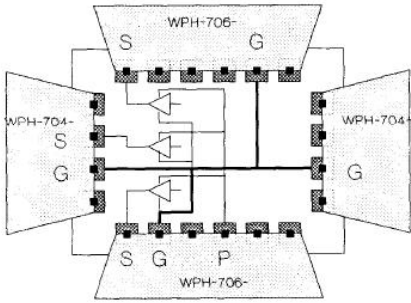
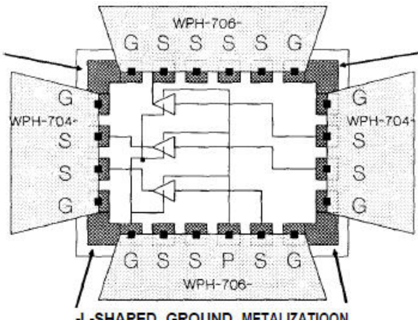


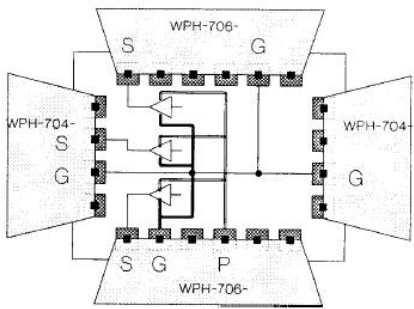
Figure 16 Alternative method of connecting the probe grounds together, using L-shaped metalization connecting ground pads. Requires probes with outside contacts defined as ground.



Rule 203 Signal input/output Signal input/outputs are either signal-ground (SG), ground-signal-ground (GSG), or signal-signal (SS). Any of these input/outputs may be used for measurements up to 18 g h Z The GSG arrangement is advised for measurements beyond 18
ghZ, and required beyond 26 GHz. The SS input/output requires differential signals and a mirrored layout, resulting in a virtual ground. The WPH-700 series probes can be factory programmed to provide a wide variety of contact sequences, including signal, ground, bypassed power, and terminated signal. (See the Cascade Microtech Probe Selection Guide for details.)

Rule 204 High dynamic-current grounds The lowest impedance ground return for high dynamic-current outputs must be connected to the same probe head that is providing the power to those outputs (Fig 17).

Figure 17 Rule 202 Where high dynamic currents exist, provide low-impedance grounds to the same probe which supplies the power.



Rule 205 Maximum rated current The maximum dc current per contact is one-half ampere. Do not apply more than 1 ampere total to a probe that is in continuous contact, i.e., no make and break contact.

Crosstalk

If there is good shielding design, crosstalk is generally not a problem, except when testing very high gain amplifiers, A/D and D/A circuits, or when there are severe reverse isolation requirements. When using coplanar probes, crosstalk may be observed in the following situations:

Crosstalk between probes:

- Capacitive or radiative coupling between probes.
- Common ground inductance in DUT.

Crosstalk between signal lines within a probe:

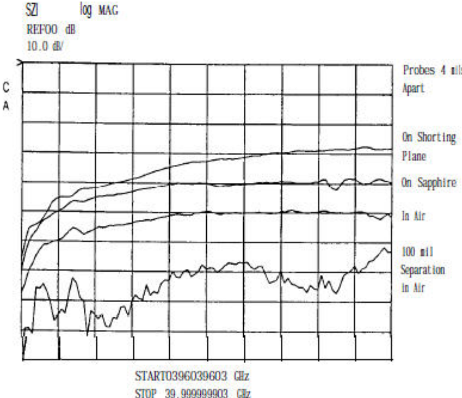
- Common ground inductance in probe.
- Mutual coupling between signal lines within a probe.

There are layout decisions that will affect crosstalk:

- Spacings between input and output pads will affect probe-to-probe crosstalk.
- Common-ground inductance in the device layout will affect probe-to-probe crosstalk.
- There is the question of whether or not to place an intervening ground pad between signal pads when laying out circuits to be probed by the WPH-700 series multicontact probes. The addition of the intervening ground pad will require a corresponding ground contact in the probe, and therefore will reduce crosstalk between signal lines within the probe.

Crosstalk Between Probes- The crosstalk between probes due to capacitive and radiative coupling is shown in Fig. 18. Note that the crosstalk depends on the spacing between probes, and the material on which the probes are landed. When laying out circuits where crosstalk between probes is important, space the pads as far apart as possible.

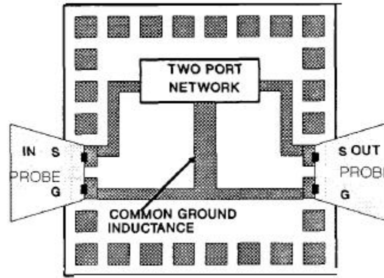
Figure 18 Crosstalk between two GSG probes



Another factor affecting probe-to-probe crosstalk is the common ground inductance, as shown in Fig. 19. Current from both the input and output will flow through the common inductance, resulting in a voltage drop across the common inductance, and part of the output signal will appear at the input, and vice versa. This internally generated crosstalk cannot be corrected by calibration or probing techniques. The guideline for layout is to make the common ground metalization as short and as wide as possible to minimize inductance.

Figure 19 Schematic layout showing common-ground inductance, resulting in crosstalk

between input and output probes.



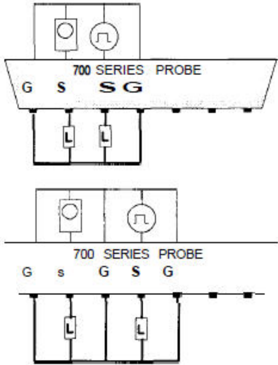
Contact sequence and crosstalk Crosstalk between signals within a probe is generally a consideration only with the WPH-700 series multicontact probes. The main contribution to crosstalk within a probe, over which you have some control, is the contact sequence. Specifically, do you place an intervening ground between adjacent signals? For example, one can specify the following sequences for a WPH-708 multicontact probe (eight contacts): (1) GSSSSSSG, (2) GSSGSSGP, and (3) GSGSGSGS. An intervening ground will obviously lower
the crosstalk between the signal lines.

Select the contact sequence that meets your requirements. Typical voltage crosstalk performance (measured with an HP 54120 digitizing oscilloscope, and using 100 ps rise time signals for all data):

sequence*	crosstalk (%) signal-signal	termination
GSGSG GSSG	1.0 % 4.0 %	sig lines open sig lines open
GSGSG GSSG	0.3 % 3.5 %	sig lines 50 1L-gnd sig lines 50 R-gnd
GSGSG GSSG	2.0 % 14.0 %	sig lines short-gnd sig lines short-gnd

*See Fig. 20 for a test circuit schematic.

Figure 20 Circuit used to measure crosstalk between S1 and S2 in WPH-700 series multi-contact probes.



Load L was either an open, short, or 50 ohms. Probe ground-line common inductance Another
factor affecting signal-to-signal crosstalk within a probe is the common inductance in the probe ground line. This situation is analogous to the common-ground inductance in a DUT, as previously discussed. Figure 21 shows a simplified equivalent circuit of a two-signal probe with common lead inductance Zg12. The amount of crosstalk is shown in Fig. 22 as a function of frequency and common-lead inductance value.

Figure 21 Simplified equivalent circuit of a SGS probe, showing how common ground inductance Zg12 contributes to crosstalk between the two signal lines.

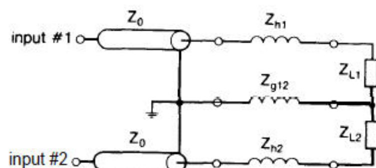
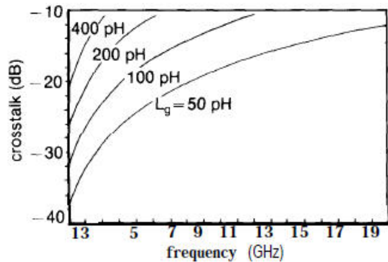


Figure 22 Common-ground inductance affects crosstalk between signal lines within a probe.



The guideline in this situation is to be aware of this source of crosstalk, choose probe families which meet your needs, and to lay out your IC accordingly. The probe families have different common-lead inductance values and you need to review the Cascade Microtech Probe Head Selection Guide as you make your decisions. Note that these considerations also apply to bonding wires used in packaging. The 0.3 nH of the WPH-700 series multicontact probe ground inductance corresponds to a bonding wire 300 km (12 mils) long. Also note that
multiple grounds and power contacts reduce the probe's equivalent common-ground inductance.
DUT Power There are four typical methods to provide power to the DUT during probing: (1) Bias tee, (2) WPH-900 multiple-needle probes, (3) WPH-700 multicontact probes, and (4) single-needle positionable probes. Bias tees, which are high-frequency chokes, are typically used when probing individual transistors or FETs, and allow the signal lines to be dc biased. Bias tees that have been installed in the S-parameter test set are usually used to characterize individual transistors or FETs.

WPH-900 needle-type probes are typically used to provide power to ICs being probed (Fig. 23). They work well, except when their inductance (8 nH typical) affects the circuit

performance. An inductance of 8 nH will adversely affect many circuits (model your circuit to check the result), including high-gain amplifiers and fast digital devices. In these cases, the WPH-700 multi-contact probes are recommended, because one of the available contact configurations has internal power-supply bypassing right at the probe tip (Fig. 24). The bypassing network has been specifically designed to be non-resonant over a very wide frequency when probing ICs. The typical impedance is 0.01 FF in series with 4 ohms and 300 pH. See the WPH-700 multicontact probe brochure for additional details on how to select and use these probes for this application.

Note that in Fig. 24, the inductance associated with the VSS, VDD and GND metalization will be about 1 nH each, which is three times that of probe inductance. If you need * lower inductance, use much wider lines, or make the die smaller, or move the WPH-700 multicontact probe to the side of the chip.

Figure 23 Typical MMIC with WPH-900 probe needles providing power and ground. The needle

inductance can result in circuit malfunction.

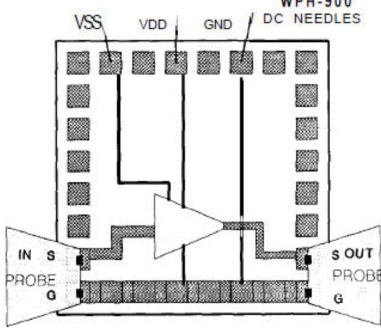
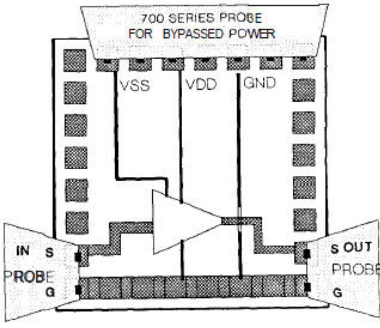


Figure 24 Typical MMIC with WPH-700 series probes, configured to provide bypassed dc power.

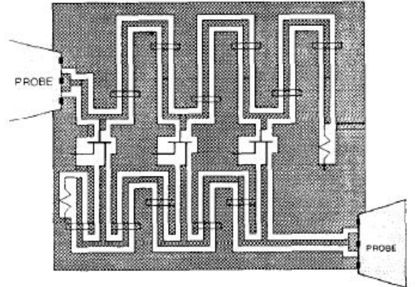


Specific Applications

MMICs

The use of Cascade Microtech probes to measure MMIC performance is a natural application. There are four different MMIC layout topologies which encompass most of the MMICs being manufactured today: (1) coplanar waveguide, (2) lumped element, (3) microstrip, and (4)
differential or balanced. For the examples shown here, the typical pad size is 100 x 100 µm.
A schematic representation of a coplanar waveguide (CPW) topology is shown Fig. 25. This topology is very easy to probe by using Cascade Microtech probes, and the FETs have minimal ground inductance because both the circuit and the probe are fabricated with CPW
topology. The signal input/output should always be GSG (Fig. 25). The circuit is biased either through the signal lines with bias tees, or through separate pads contacted with either the WPH-900 multiple-needle probes or the WPH-700 series multicontact probes. Note the second-level metal crossovers connecting the grounds together, eliminating ground moding.

Figure 25: MMIC layout, with padouts for probes.



A simplified layout of a lumped-element MMIC topology is shown in Fig. 26. Note that in this topology lumped inductors, capacitors, and resistors are used, rather than transmission lines. The interconnects on the die are modeled as inductors. Because lumped element designs are typically used below 18 GHz, the input/output of Fig. 26 is GS. If you wish to use a GSG input/output, the other ground should be connected to the DUT ground bus with a large, minimum length of metal (Fig. 27). Power is provided with WPH-900 multiple-needle probes (Fig. 27)or WPH-700 multicontact probes (Fig. 26).

Figure 26 Schematic of lumped-element MMIC. WPH-700 multicontact probe provides power, and GS probes provide measurements.

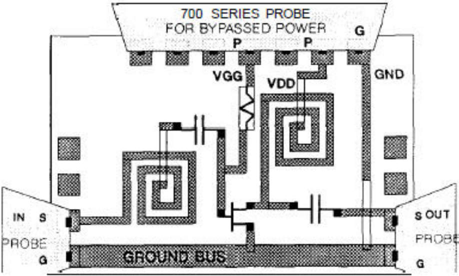
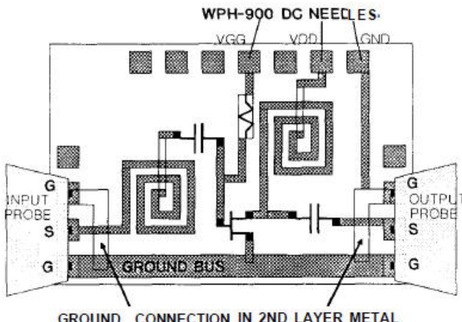


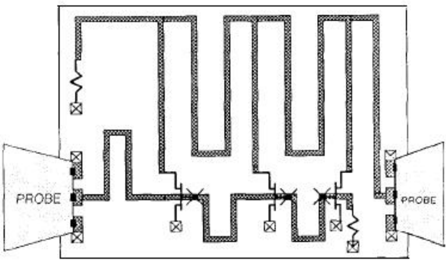
Figure 27 Schematic of lumped-element MMIC. WPH-900 needles provide power, and GSG probes provide measurements. Note short, wide metal used to connect the two ground pads.



When these die are packaged, multiple-bond wires are used to connect the die ground bus to the package ground. Additionally, bond wires will connect the input and output pads to the appropriate package pins. None of these inductive bond wires will be present during probing, and their effect on circuit performance will not be included in the measurements. The guideline is to model your circuit without the package and bond wire parasitics present, and this is what you should measure with the probes.

A simplified example of a microstrip layout is shown in Fig. 28. Microstrip designs always have a backside ground, and usually have substrate vias to connect the top-side devices to ground. Because the Cascade Microtech probes require ground and signal to be in the same plane, substrate vias are necessary to launch signals into a microstrip design.

Figure 28 Typical microstrip-MMIC layout. Substrate vias are required to connect the probe grounds to the IC backside ground.

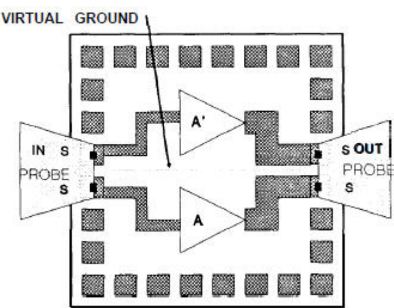


The parasitic inductance in the coplanar-to-microstrip transition should be minimized by making the metal lines to the via as short as possible. Do not land on the substrate vias, as they are typically not flat enough for a probe to reliably contact the vias and a pad at the same time. Note that you will measure, using the probes, the MMIC performance and the coplanar-to-microstrip transition formed by the substrate vias. A correction can be done by measuring and removing the effects of this transition; or you can use a TRL calibration technique, by moving the reference plane just past the transition so it will not be included in the measurement. If substrate vias are not available, a technique using capacitive coupling to the backside is available 1 . Power and control signals are typically sent through the signal lines with bias tees, or WPH-900 multiple-needle probes, or WPH-700 multicontact probes.

An example of a differential-layout MMIC is shown in Fig. 29. The unique feature of this layout is the mirrored symmetry, which results in a virtual ground. Differential input and output signals are typically used without any ground. Single-ended input signals, output signals, and power sources require grounds. The use of balanced topologies results in balanced current requirements, easing power supply bypassing requirements. This topology has been successfully probed up to 18 GHz.

This technique's limitation is that it requires accurately balanced signals. This is generally a limitation with the signal generation equipment, and not of the probe phase match. For WPH-003 18 GHz probes, the phase match is better than 1 ps, and better than 5 ps for WPH-700 multicontact probes.

Figure 29 A differential layout. Circuit A is mirror-imaged about the virtual ground to give circuit A'. The SS input requires differential signals.



The preceding examples show typical ways of providing power and ground to the DUT. In many cases, such as high-gain amplifiers, the use of WPH-900 multiple-needle probes to provide power will result in unwanted oscillations during probing, due to the inductance of the
needle (typically 8 nH0. In such cases, use WPH-700 multicontact probes to provide power.
Additionally, remember that when probing your ICs you will be measuring the IC performance without the package and bond wires, and that the measured performance in the package will be slightly different than measured on the wafer.

Digital Devices

Most digital GHz-probing applications involve GaAs ICs, high-speed ECL, or mixed analog/digital with high-speed digital inputs or outputs. Most digital DUTs are probed with WPH-700 multicontact probes, because of the larger number of required contacts. Additional reasons to use the WPH-700 multicontact probes are the availability of bypassed non-resonant power, and ECL output terminations (50 ohms to VTT).

A typical, and recommended, padout is shown in Fig. 30, with typical pad size being 100 x 100 pm. Note that all probe grounds are connected together on the DUT (Fig. 15), and that the high-current output driver grounds are connected by the lowest impedance path to the same probe which supplies the power to these drivers (Fig. 17). If situations arise where high input/output isolation is required, lay out the IC so that the inputs and outputs are

contacted by separate probes (Fig. 31).

Figure 30 Typical digital IC padout

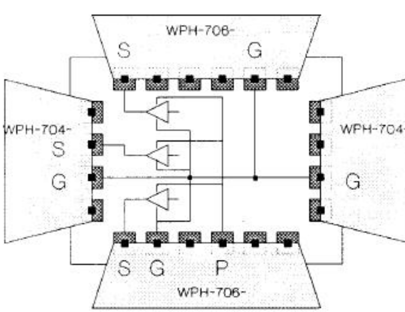
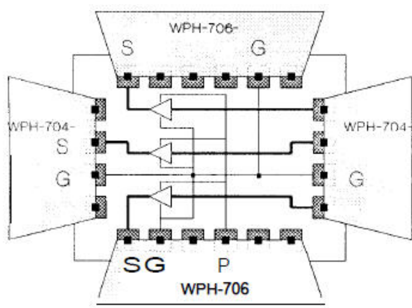


Figure 31 Digital IC example, showing inputs on a probe separate from the outputs. This maximizes isolation between inputs and outputs.



Characterization Devices

Individual FETs, transistors, capacitors, and inductors are often measured in order to develop models, or as part of wafer qualification. In laying out these devices it is important to minimize parasitics associated with the interconnects to these devices, and the pads. In general, if one is measuring low-impedance devices, then the series-inductance parasitics should be minimized by using fat, short bars of metal. If high-impedance devices are to be measured, then the parallel-capacitance parasitics should be minimized by using narrow interconnects and small pads. Also, the substrate will affect the final layout.

For example, the parasitic capacitance to ground is much smaller with semi-insulating GaAs substrates than for conductive silicon substrates. The GaAs industry typically measures fairly large devices (300 µm gate length) as characterization devices. Many device layouts are constructed with 100 x 100 µm pads, and wide metal buses connecting the pads to the DUT, as shown in Fig. 32a and b. However, within the silicon community the characterization devices are often minimum-size devices, presenting high impedances to the
test system. This requires that parasitic capacitances be minimized. To minimize parasitic capacitance to the substrate (Fig. 32c and d) the pads are small, usually 50 x 50 µm, and the interconnects from the pads to the DUT are narrow. Note that with GSG configurations the DUT should connect the two ground buses together (Fig. 32a and c).

Figure 32 Layout examples of active characterization devices.

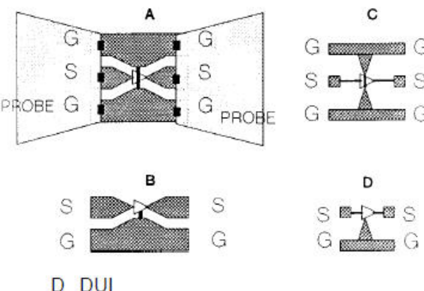


Figure 33 shows typical optimization steps in laying out a characterization pattern for a GaAs FET. Figure 33a shows the initial layout. The initial layout has the problems of narrow, high-inductance, connections to the gate and drain, and the common-ground inductance in the source connection has not been minimized. In Fig. 33b these problems have been resolved by repositioning the DUT, and by widening the interconnect metal to the gate and drain. Finally, in Fig. 33c the pads are moved closer together to further minimize the series-inductance parasitics associated with the metal connecting the gate

and drain to the pads. Larger interdigitated FETs are typically measured as shown in Fig. 34.

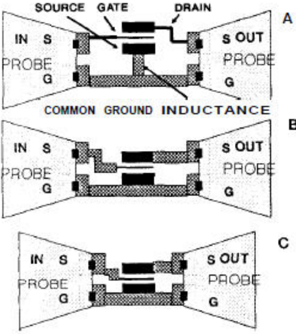


Figure 33 Optimization of large gAaS FET characterization device.

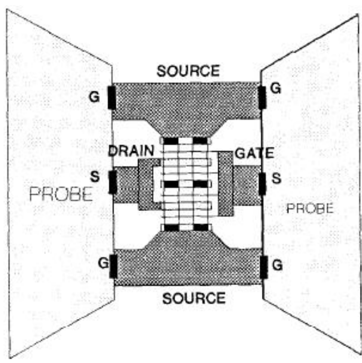


Figure 34 Example layout of four-gate interdigitated GaAs FET. Note the second-layer metal connecting the sources together.

In Fig. 35 the analogous situation is shown for minimum-size transistors on silicon substrates. In Fig. 35a the layout contains large (100 x 100 µm) pads, very closely spaced together. This layout, while excellent for FETs on GaAs is not workable for minimum-size devices on silicon. The problems with this layout are the parasitic capacitances (Fig. 35b). Because silicon is conductive, the parasitic capacitances are much higher than with GaAs or alumina substrates, and can contribute to significant measurement errors, a -50% error in ft, for example. See the paper "GHz On-Silicon-Wafer Probing Calibration Methods" 2 for more details.

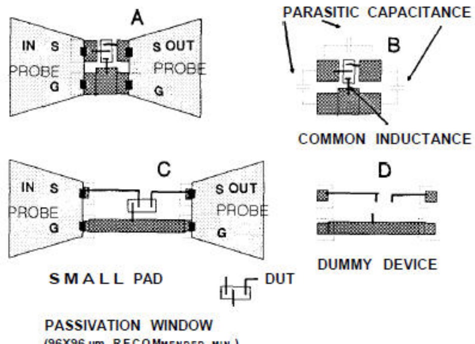


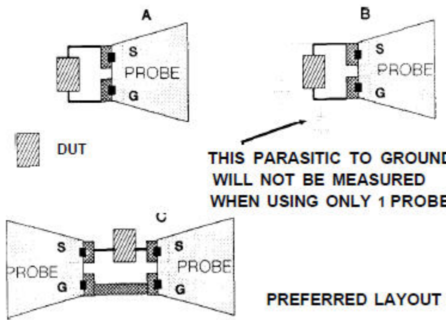
Figure 35 Optimization of small silicon transistor characterization device. Dummy device is used to correct measurement for remaining parallel-capacitive parasitics.

Parasitics are minimized in the Fig. 35c layout by using smaller pads (50 x 50 µm), spacing the pads out further (150 to 250 µm center-to-center spacing), and narrowing the metal lines to minimize capacitance to ground. The DUT is located near the probe ground bus to minimize the common inductance. Also note that when laying out 50 µm pads, the passivation window should be at least 96 x 96 µm, and preferably larger, to accommodate the 50 x 50 µm probe contacts, plus typical placement errors. (References 2 and 3 describe techniques to correct measurement errors due to parasitic capacitances.) The recommended method of minimizing parasitics involves the use of a "dummy" device, which is simply the DUT layout without the DUT (Fig. 35d) 3 . It is recommended that a dummy device be included with each group of DUTs. In this way the wafer variations affecting the parasitic capacitances, such as oxide thickness variations, will also be corrected.

When laying out passive-characterization devices for probing, the above discussion

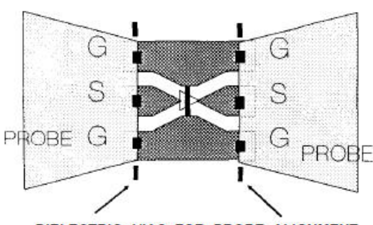
applies. The simplest layout is shown in Fig. 36a, using one probe. With this layout, however, part of the DUT parasitic capacitance to ground will not be measured (Fig. 36b). The solution is to lay out the DUT so that a two-port measurement can be made (Fig. 36c).

Figure 36 Layout examples of passive characterization devices



One final tip. To obtain maximum repeatability, use a dielectric etch to mark where the front of the probe tips should land (Fig. 37).

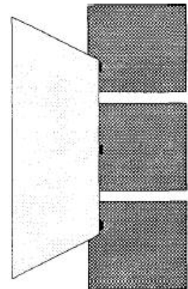
Figure 37 Dielectric vias for probe alignment marks, used to increase measurement repeatability



Hybrids and Packages

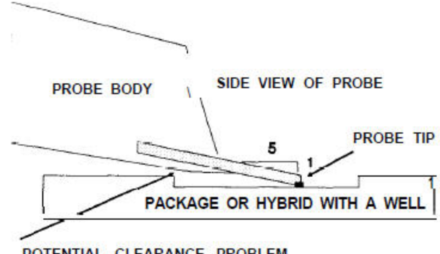
The typical pad and interconnect dimensions in packages and hybrid substrates are much larger than with ICs. The larger pad and interconnect dimensions require short-length DUT connections, especially the ground connections, for good GHz-probing. The probe contacts do not need to land in the center of pads, but can land near the edge (Fig. 38). Be sure that the passivation has been removed from where the probes will land. Also, refer to the Cascade Microtech Probe Head Selection Guide to verify that probes with a wide enough pitch are available.

Figure 38 Off-center probe, landing on large hybrid or package pads. Probing into wells has a potential problem (Fig. 39).



The probe can collide with the top corner of the well. If you have a potential situation like this, please call the Cascade Microtech Applications group and request a side view of the probe you plan to use. Generally the WPH and RTP line of probes approach the DUT at about an 11 degree angle.

Figure 39 Side view of probe and package with a well, illustrating potential clearance problem



Summary

This Application Note is intended to help you lay out your IC hybrid, or package so that Cascade Microtech coplanar probes can be used for testing. These probes are widely used to test individual devices, MMICs digital ICs hybrids, and packages. If you have any questions
pertaining to layout or testing please call the Cascade Microtech Applications Group.

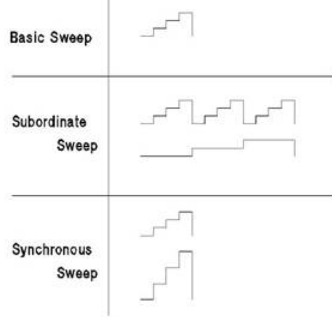
DC

Contents

- *DC Measurements and Calibration Techniques* (iccapmhb)
- *DC Characterization for Device Modeling* (iccapmhb)
- *Power DC Measurements* (iccapmhb)
- *Process Control Monitoring (PCM) Measurements* (iccapmhb)

DC Measurements and Calibration Techniques

Large signal modeling of a nonlinear component always begins with the characterization of its DC performance. Instead of power supplies, DC parametric analyzers with source-monitor-unit plugins (SMU) are used. This allows to fully characterize the DUT (device under test) from femto-Ampere up to its maximum current, and in all four quadrants. I.e. forward and reverse currents and voltages, are measured with the same SMU unit. Usually, in case of a transistor, all 4 terminals (including substrate) are connected to individual SMUs in order to avoid recabling during the forward and reverse measurements. To begin with, there are 3 basic sweep types available for the SMUs of DC analyzers, as depicted below:



SMUs apply Kelvin measurements to avoid parasitic series resistances. This measurement procedure, also known as the four-wire method, consists of a stimulating line (Force) with a second one in parallel (Sense). Figure 1 illustrates this. Ohmic losses on the Force line are eliminated by the OpAmp in voltage follower mode. This means the OpAmp output will exhibit a somewhat higher voltage than the desired test voltage at the DUT, because the test current generates some ohmic losses along the Force line. The Sense line, connected to the minus input of the OpAmp, assures that the DUT is exactly the desired test voltage.

$$\underline{Z}_x = \frac{\underline{V}_1}{\underline{I}} - R = \left(\frac{\underline{V}_1}{\underline{V}_2} - 1\right) * R$$

detailed schematic:

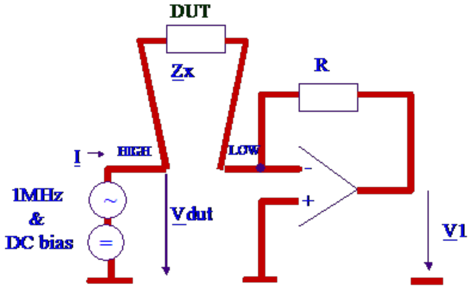


Fig.1: The principle of Kelvin measurements

While this method eliminates DC errors, it does not cover dynamic measurement problems. For example, to avoid external electro-magnetic influences, both the Force and Sense cables consist of shielded cables. Such cables exhibit some parasitic capacitances. Due to charging problems, these capacitances will affect the measurement speed and accuracy of our Kelvin measurement.

A simple example: assume we want to measure the reverse characteristics of a semiconductor diode. This means we need measure very low currents. Before the voltage steps to, for example, -20 V, the quiescent voltage at the diode is zero. That is, the cable capacitors are not charged. When the negative voltage step occurs, these capacitances have to be charged, and the required current is provided by the OpAmp. This could lead to either a mis-measurement (DUT current plus charging current) or a delay in the triggering of the actual current measurement (by some intelligent firmware in our measurement).

To solve this problem, an extra inner shielding is applied between the hot metering lines

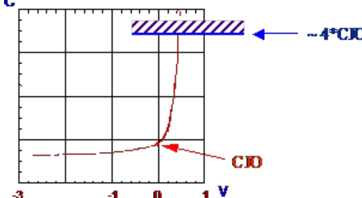
and the outer cable shielding, called 'Guard'. This extra shielding is connected to a separate, second OpAmp which follows exactly the value of the desired test voltage. Now it is this auxiliary OpAmp which supplies the charging current for the test cables, while the main OpAmp can start current measurements independently of this charging problem. That is, the inner measurement loop does not see the charging problem any more.
Of course the point where Force and Sense are tied together must be as close as possible to the DUT. In case they aren't connected, an internal 10kOhm resistor at the output of the SMU acts as the Kelvin point. Another important fact is that the Guard contact should *never* be connected to Force or Sense. Otherwise, the inner loop OpAmp of the SMU would measure the DUT current *plus* the charging current of the auxiliary, second OpAmp!
In order to maintain the DC measurement accuracy, SMUs perform periodically an auto-calibration. This means that the SMU disconnects its outputs from the DUT, measures possible offset voltages and currents and corrects it. This type of calibration does not need any action from the user.

Hints on DC measurements
For transistor measurements: If an *individual* SMU (source monitor unit) is employed for *each* individual pin of the DUT, and the 'Ground Unit' of the 4142 or 415x is *not* used at all, it is possible to perform all the forward and reverse measurements *without reconnecting* the DUT. In this case, it is recommended to configure the 4142 with twice a 1Ampere SMU (for Collector-Emitter , Drain-Source correspondingly) and to employ two 100mA SMUs for the other transistor pins.

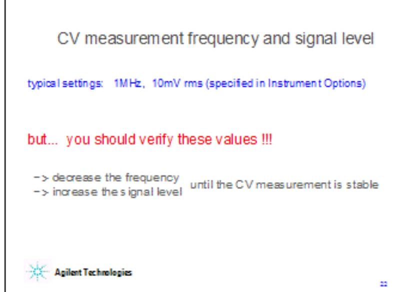
Connecting triax to another connector type
The following table gives an overview about connector adapters from Agilent. For converting a triax SMU output to a coax connection, for example to connect a SMU to the coax input of a NWA S-parameter testset, use the **Agilent 1250-2652 triax-coax converter**. This connector leaves the middle shield open, what is a prerequisite for the shielded Kelvin measurement principle.

$$\frac{\mathbf{V_{DUT}}}{\underline{\mathbf{Z}}_{\mathbf{x}}} + \frac{\mathbf{V_1}}{\mathbf{R}} = 0$$

NOTE: A very good overview about how to connect the Agilent DC analyzers for on-wafer measurements can be found in the 4155B/4156B Product Note-3, Feb.98, Agilent Lit.no. 5966-4185E. See the Semiconductor Test Assistant CD, available from Agilent, dept. Semiconductor Test.
Another vendor of Adapters and Feedthrus is **Trompeter Electronics**, on the web at: http://www.trompeter.com/shop/home.asp
Note: If you want to download the Trompeter Adapter Catalog, select the catalog T20 Military/Aerospace:
The figure below shows the available converters. Again, for connecting the triax cable end to the coax input of a network analyzer testset, we need a triax-female to coax-male converter:the **ADBJ77-E2-PL20.**



If you need to access the force/sense middle shielding contact, e.g. to extend the shielding to the shielding of the device on wafer, use the connector plate of the Agilent4142, opt.012, Agilent part number P/N 04142-60021



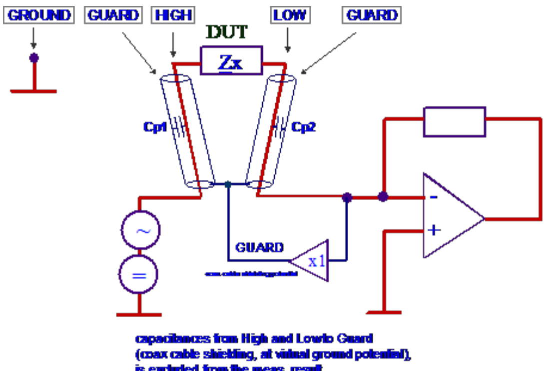
Monitoring the SMU voltages
Applying the measurement setup given in fig.2, an oscilloscope can be connected to monitor the voltage at the transistor pins. This is done by using the sense line as a high impedance scope probe. This allows to measure the voltage versus time accurately.

Fig.2: monitoring the SMU voltage at the transistor pin.

Expanding Current to >1A
If used in combination with an external DC power supply, the Agilent semiconductor parameter analyzers can evaluate the DC characteristics of a power device that requires a current of several amperes. Application note 4156-5 describes how to connect the cost-effective Agilent 665x DC power supply in series to the DC analyzer SMU output. Therefore, no extra HPIB programming is required. These power supplies are fast enough to follow the SMUs output voltages. Furthermore, an internal small shunt resistor is used to monitor the amplified current. This shunt resistor can be measured once by IC-CAP, and the voltage drop across it serves then as current monitor.
Recommended external power supplies:- 50 A / 8 V (Agilent 6651A)- 25 A / 20 V (Agilent 6652A)- 10 A / 50 V (Agilent 6654A- J05)
See the docu in the Modeling Handbook at
ICCAP_ModelingBook_pdf\1MEASUREMENTS\1DC\DC_Power\Publications\Agilent\415x_plus_PowerSupplies\2004_AN4156_5_external_PowerSupply.pdf

---

On-Wafer DC Measurements
The following slide of Cascade Microtech explains recommended grounding and guarding



Thermal Effects During DC Measurements
As a general rule, we have to account for self-heating effects during DC measurements of packaged devices already at currents above ~10mA (for typical Vmax = 5V). For on-chip measurements using a thermochuck, the critical current is about 50mA. This can be seen in fig.1. Besides possible self-damage of the DUT, self-heating is a problem with modeling. This is because the simulation model assumes a *constant* temperature TNOM, at which the simulation is performed. In reality, however, the DUT suffered from self-heating and exhibited therefore different measurement curves than those curves obtained from the simulator. Therefore, a curve fitting with physical parameter values can not be achieved.

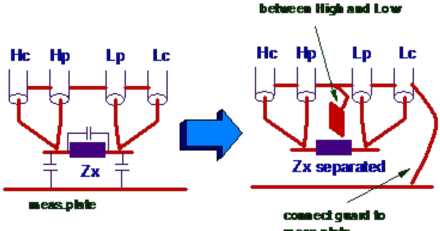


Fig.1: chip temperature increase of a packaged transistor as a function of the pulse width and the applied bias power. The pulse period is 1s.

To prevent from self-heating, pulsed DC measurements could be applied. This gives isothermal measurement conditions, and would allow to extract really those model parameters which are correct for the constant modeling temperature TNOM. However, if these pulses exceed a few milliseconds, self-heating might occur again, see fig.2.

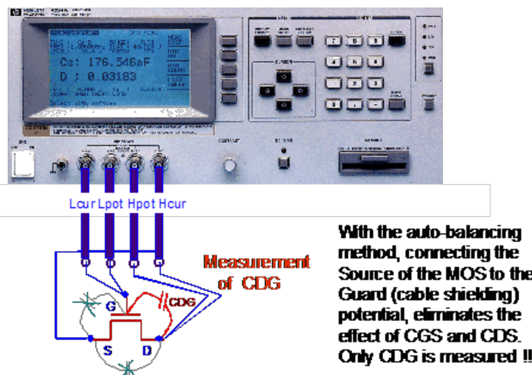
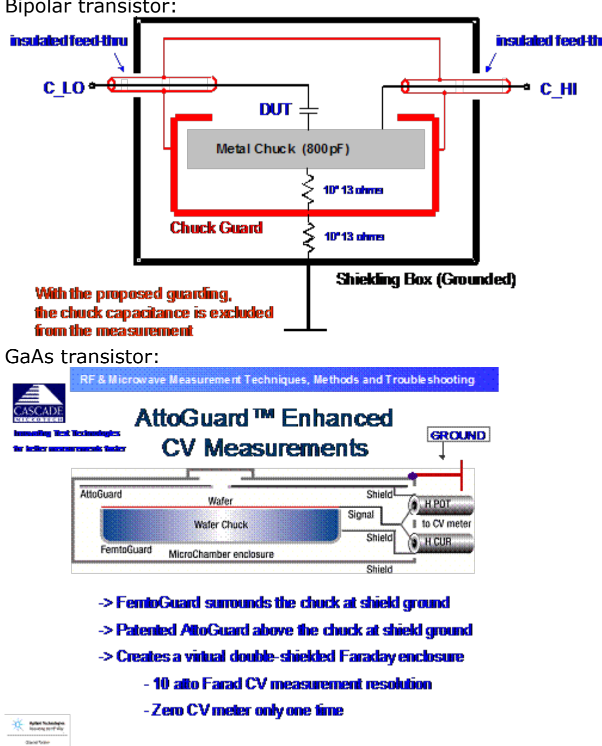


Fig.2: a self heating effect may occur even for pulsed bias conditions

Fig. 3 gives an example for a bipolar transistor, measured with different pulse widths. It can be seen that only pulses in the millisecond range can lead to isothermal measurement results. For bipolar silicon transistors, the Collector current increases with increasing pulse width. For GaAs transistors, it's the opposite.

Bipolar transistor:



GaAs transistor:

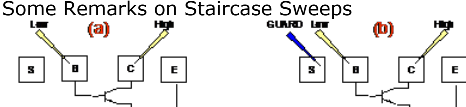
Fig.3: Thermal self-heating for pulsed bias measurements as a function of the pulse width
This is the reason that the minimum pulse with of the Agilent 414x and 415x DC analyzers are not short enough for isothermal modeling measurements. The main application for that type of pulsed measurements is for the other big market of those instruments, the production measurements. For this kind of application, pulsing is used for some specific spot measurements at high currents which otherwise could destroy the component. The isothermal issue is not of interest hereby.

How to check for isothermal measurement conditions:
Let's take a bipolar transistor for example. We measure the so-called Gummel-Poon plot and calculate BETA = ic / ib. The Collector voltage is kept at vCE = MAX(vCE / 2). We then perform these measurements: > sweep vBE from low to high voltage, long integration mode > sweep vBE from high to low voltage, long integration modeIf we realize a different BETA curve, there is thermal self-heating.

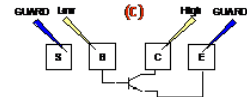
What to do if pulsed measurements are not possible:
If pulsed DC measurements are not available (Agilent pulsed measurement system 85124A, allowing 1ms pulses, see the chapter on network analyzers), we have these possibilities:
DC:Characterize the device in those DC ranges where it will be used later in the design application. In many cases, this is not in the high current range.If the application is in the self-heating range, and the model doesn't include self-heating, it might be smart to measure using long integration range, in order to really heat up the device like in the later application, and to fit the model to the non-isothermal measurements.If your selected

model includes self-heating, then see the chapter on thermal modeling for more details.
AC:If S-parameters are to be fitted too, it is recommendable to measure the DC curves under the same conditions like the biased S-parameters. Since the S-parameter measurements are usually much slower than the DC measurements, we have more self-heating for AC than for DC. Therefore, long integration should be applied to the DC measurements.

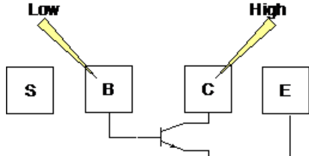
Some Remarks on Staircase Sweeps



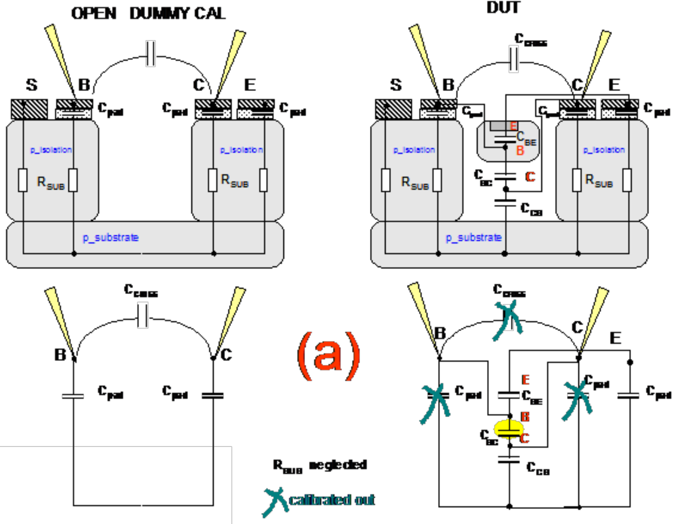
3 Contacting Cases
for CV Calibration and Measurement,
Applying the Auto-Balancing Method



It is important to note that DC analyzers, like other instruments, have two operation modes: - either the controlling software sends the start-stop-step stimulus info and the 'what-to-measure' GPIB commands to the instrument and finally sends a trigger command ('User Sweep' = no), - or every individual measurement point is triggered individually by the software ('User Sweep' = yes).
>> *As a consequence, the measurement vs. time is affected from these two possible settings.*
In the slide above, for 'User Sweep'=no, the CPU of the instrument is the owner, or controller of the measurement, and not the controlling software (like IC-CAP). In case of a DC analyzer, the CPU's firmware itself will control the measurements during the staircase sweep. After having received the 'trigger' command, it will wait for the measurement result to converge (end of charging effects), it will set the ranging of the amplifier (to have the max. resolution of the AD-converter) etc. This means that the time stepping of the staircase sweep is not monotonic (!). However, the measurement result (in e.g. IC-CAP) looks like a straight line, since we plot e.g. current vs. voltage, and ignore the time stepping. This explains some very special effects where e.g. the transfer curve current is different by some % from the output characteristics current, for identical bias voltages, and where self-heating can be excluded to explain the differences.
If the user wants to make sure that the staircase sweep stepping represents a more or less linear stepping in time, he should set the 'User Sweep' = yes, see the next slide. However, due to the huge activity on the GPIB bus, the measurement may take much longer in this case.

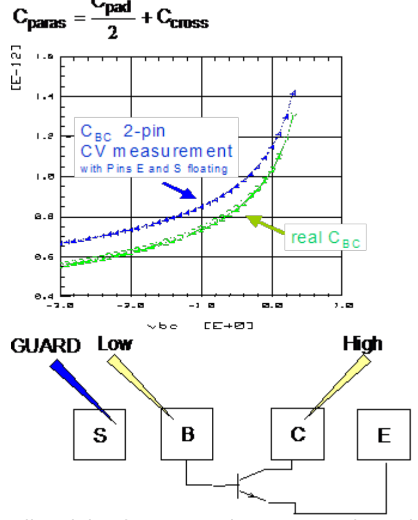


When setting 'User Sweep' = yes, the controlling software (e.g. IC-CAP) is the 'owner of the sweep'. Due to the relatively slow performance of the GPIB bus compared to the instrument's measurement time, the resulting staircase sweep is time-wise pretty monotonic. Apply this instrument mode if you must assure such time dependence of your measurement.
NOTE: in 95% of modeling applications, 'User Sweep' = no is appropriate.
Acknowledgements:
special thanks to Fridolin Illien of ETH University Zuerich, Switzerland, for the oscilloscope screen shots !

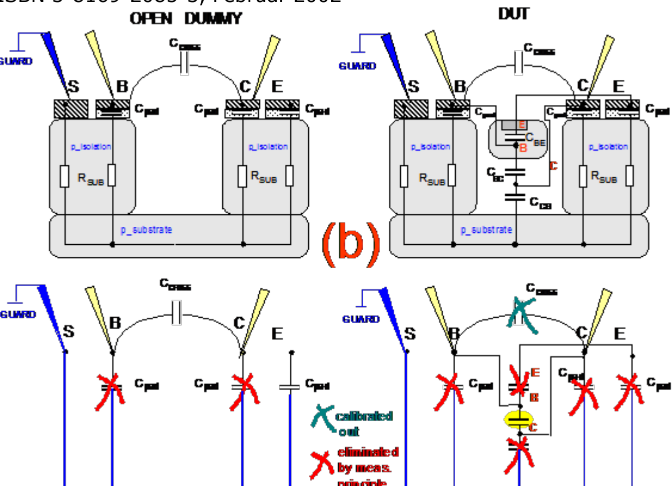


Some Remarks on Oscillations During DC Measurements

Active components like transistors may self-oscillate during DC measurements. Typically, the self-oscillation condition is fulfilled happens only for certain bias conditions. the following two slides show such phenomena of a bipolar transistor. The first slide depicts a linearization effect of iC and iB at vBE >0.85V, caused by self-oscillation. The other two slides show typical self-oscillation at certain bias levels.



All 3 slides from: J.Berkner, Kompaktmodelle für Bipolartransistoren, Expert-Verlag Renningen (Germany),
ISBN 3-8169-2085-3, Februar 2002



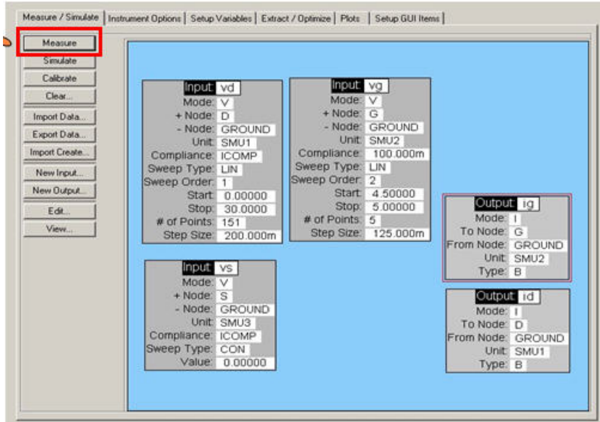
Under certain circumstances, SMUs may oscillate. Superbeta, wide- bandwidth bipolar transistors are especially susceptible to oscillation. But also GaAs transistors.
Oscillation can become a problem when using older DC probe needles. A common way to

avoid such oscillations is using HF probes (Ground-Signal-Ground GSG probes), or the newer shielded DC needles.
If you have to live with your DC needles, and when oscillations occur, here some ideas on causes for such oscillations:
1) SMU Induced oscillationsThese oscillations occur when an inductive load is connected to the SMU's output. The SMU has often an inductive load because the output impedance of the SMU is inductive during V-Mode operation and usually several SMUs are indirectly connected together through the DUT.
2) Oscillation due to straysThe measurement system including the DUT, stray capacitance and residual inductance of the connection cables, switching matrix, probe card and/or test fixture can be recognized as an oscillation circuit.
The oscillation detector of SMU may not detect this type of oscillation. This is understandable when thinking of the SMU output as a low-pass filter, while the oscillation frequency may be in the 100MHz ... to several GHz range. Also, if ~~like in most cases~~ the oscillation is located at the DUT, the SMU itself cannot do anything to prevent this oscillations, since the cables are long compared to the oscillation wavelength.
As a general rule: the oscillation has to be avoided where it happens.
Related to the example of the slide above, we should be aware of the total circuit: The SMUs connected to the MOSFET's gate and drain are operating in the V-Mode. Since SMUs typically appear to be inductive in V-mode, this makes this configuration equivalent to a Hartley oscillator !
The SMU may oscillate if an unusually large inductance is connected to it. This could occur if the DUT is a superbeta transistor (big hFE) and the SMU connected to the emitter is set to one of the low current ranges.
For more details on conditions for oscillations, refer to the application note 356 -1 (publication number Agilent 5950-2954)

Some effective methods to eliminate oscillations.
> For FETs, add resistive ferrite beads (Agilent Part Number 9170-0029) as close as possible to the gate.
> For bipolar transistors, add resistive ferrite beads as close as possible to the base or emitter.
> In some cases, it is necessary to use more than one ferrite beads.
> Keep cables as short as possible. Long cables cause oscillation because of their large inductance.
>>> Or, apply shielded DC probes, even better apply RF probes.

Publications
Agilent 4155B/4156B Product Note No. 3: Prober Connection Guide, Agilent Technologies Literature Number 5966-4185E, 1998
Ultra Low Current dc Characterization of MOSFETs at the Wafer Level,Agilent Technologies Application Note 4156-1, 1/1998, Lit.Nr. 5963-2014E

DC Characterization for Device Modeling

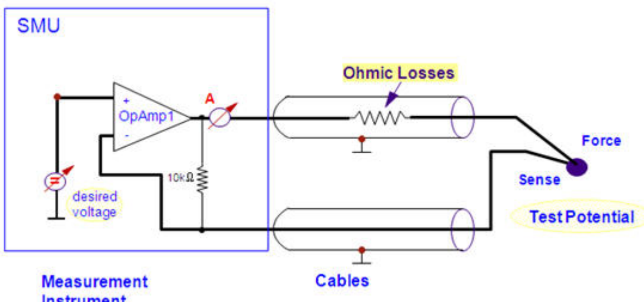


Modeling of a nonlinear component always begins with the characterization of its DC performance.
Instead of power supplies, precision DC parametric analyzers with source-monitor-unit (SMU) plugins are applied. They offer current ranges from femto-Ampère (1E-15) to several Ampères, and voltage potentials from micro-Volt to hundreds of Volts. This allows to fully characterize the DUT (device under test) in all four I-V quadrants. I.e. forward and reverse currents and voltages, are measured with the same SMU unit.
Usually, in case of a transistor, all 4 terminals (including substrate) are connected to individual SMUs in order to avoid re cabling during the forward and reverse measurements.
When performing these measurements, care must be taken for accurate shielding, thermal selfheating, contact resistance etc.

- 4-quadrant DC analyzers are applied
- Several measurement plug-ins (SMU Source Monitor Unit)
- Measurement ranges: 10 fA ... 1A, 2 mV .. 200V
- Kelvin measurement capabilities
- Automatic self-calibration

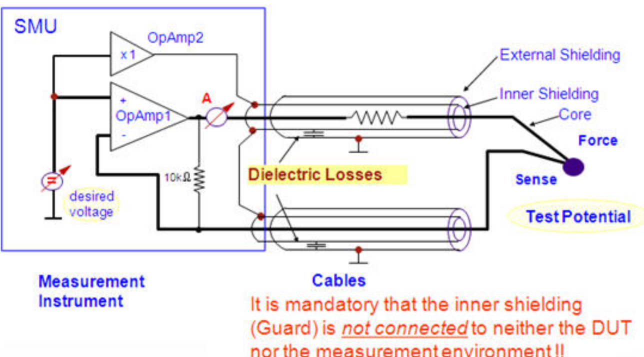
Modeling of a nonlinear component always begins with the characterization of its DC performance. Instead of power supplies, precision DC parametric analyzers with source-monitor-unit (SMU) plugins are applied. They offer current ranges from femto-Ampère (1E-15) to several Ampères, and voltage potentials from micro-Volt to hundreds of Volts. This allows to fully characterize the DUT (device under test) in all four I-V quadrants. I.e. forward and reverse currents and voltages, are measured with the same SMU unit. Usually, in case of a transistor, all 4 terminals (including substrate) are connected to individual SMUs in order to avoid recabling during the forward and reverse measurements.

When performing these measurements, care must be taken for accurate shielding, thermal selfheating, contact resistance etc.



SMUs apply a Kelvin measurement to avoid parasitic series resistances. This measurement procedure, also known as the four-wire method, consists of a stimulating line (Force) with a second one in parallel (Sense) for every pin of the DUT. The slide above illustrates this. Ohmic losses on the Force line are eliminated by the operational amplifier (OpAmp1) in voltage follower mode. This means this OpAmp1 output will exhibit a somewhat higher voltage than the desired test voltage at the DUT, because the test current generates some ohmic losses along the Force line. The Sense line, connected to the minus input of the OpAmp1, assures that the DUT is biased with exactly the desired test voltage.

> **Note**
> when you only connect the Force cable and leave the Sense open, the internal 10kOhm resistor provides the required closing-of-the-loop for the OpAmp.



While the Kelvin method compensates the DC errors, it does not cover dynamic DC measurement problems. For example, to avoid external electro-magnetic influences, both the Force and Sense cables are shielded. But such cable shieldings exhibit parasitic capacitance. Due to charging problems, these capacitances will affect the measurement speed and accuracy of our Kelvin measurement.
As a simple example: assume we want to measure the reverse characteristics of a semiconductor diode. This means we need measure very low currents. Before the voltage steps to e.g. -20 V, the quiescent voltage at the diode is zero. That is, the cable capacitors are not charged. When the negative voltage step occurs, these capacitance have to be charged, and the required current is provided by the OpAmp1. This could lead to either a mis-measurement (DUT current plus charging current) or a delay in the triggering of the actual current measurement (by some intelligent firmware in our measurement).
To solve this problem, an extra inner shielding is applied between the hot metering lines and the outer cable shielding, called 'Guard'. This extra shielding is connected to a separate, second OpAmp2 which follows exactly the value of the desired test voltage. Now it is this auxiliary OpAmp2 which supplies the charging current for the test cables, while the main OpAmp1 can start current measurements without being affected by this charging problem. That is, the inner measurement loop does not see the charging problem any more.
Of course the point where Force and Sense are tied together must be as close as possible to the DUT. In case they aren't connected, an internal 10kOhm resistor at the output of the SMU acts as the Kelvin point. Another important fact is that the Guard contact should never be connected to Force or Sense. Otherwise, the inner loop OpAmp1 of the SMU would measure the DUT current plus the charging current of the auxiliary, second OpAmp2!
Calibration:
In order to maintain the DC measurement accuracy, SMUs perform periodically an auto-calibration. This means that the SMU disconnects its outputs from the DUT, measures possible offset voltages and currents and corrects it. This type of calibration does not require any action from the user.

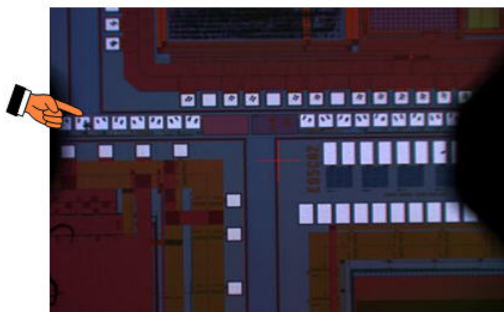
RF & Microwave Measurement Techniques, Methods and Troubleshooting
On-Wafer Measurement Environment
triax cables: ground, gard and core
coax cables: guard and core
Sense Force
SMU 1
SMU 2
SMU 3
SMU 4
Agilent DC Analyzer
Patented MicroChamber™
DCP-150KS
DCP-150KS
Chuck
Guard Layer

Typical on-wafer IV measurement setups require probes and chuck bias. A careful shielding completely surrounds the wafer with guard.

Make sure that the Guard contact of the SMUs is left open. If it was connected either to the inner 'hot' measurement cable layer or the outer ground layer, this would result in wrong current measurements!!

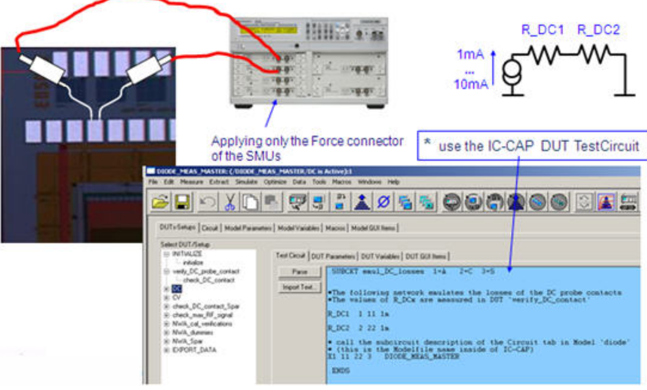
(slide with courtesy from Cascade Microtech)

Contacting a Wafer using ForceSense DC Needles

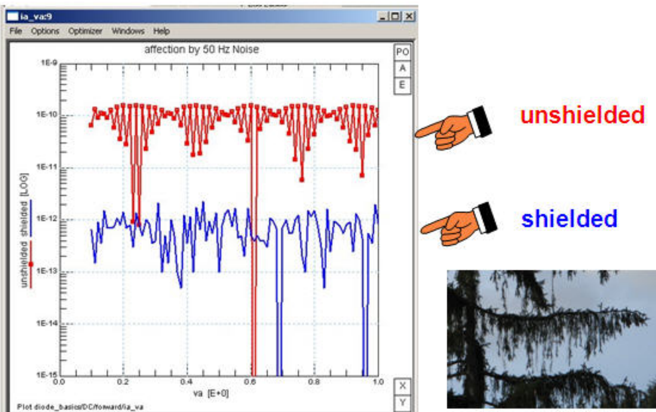


When you can not apply Force-Sense needles, you need to measure the contact resistance.

In this case, enclose your model in a subcircuit* and add the contact resistances there.



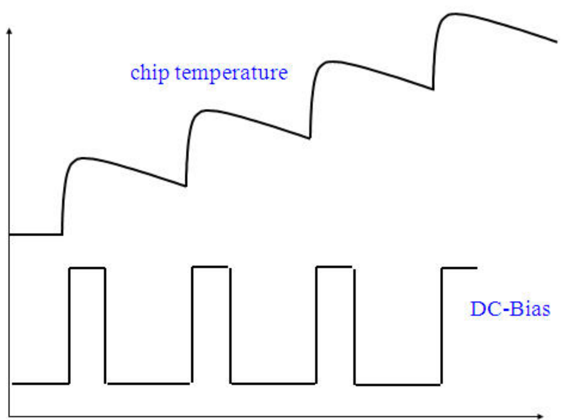
Shielding problem:50Hz noise overlaying the measurement



50Hz noise shows up like a regular pattern, distorting the measurement result.
Uncorrelated noise does not exhibit a regular pattern.

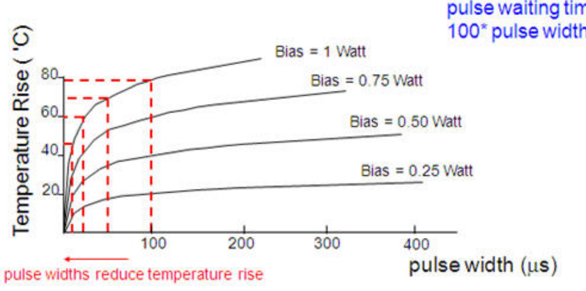
> **Note**
> When your noise looks like the pine (fir) tree branch shown in the photograph, it is uncorrelated noise.

Warming-up during DC biasing



When making DC measurements, self-heating has to be taken into account. Even when applying pulsed DC signals.
And the selfheating of the previous impulse can overlay the selfheating of the following impulse, like sketched above.

Self-heating during DC measurements



Device Temperature Rise vs. Pulse Width at different bias levels

Self-heating can become a severe problem with device characterization, because, when self-heating occurs, the measurement results depends on the measurement speed!
This can be verified when e.g. measuring a Gummel plot for a bipolar transistor, once sweeping from low to high voltages different measurement speeds, and then sweeping from high to low voltages. Calculate beta out of your different measurements.If self-heating occured, you will get as many beta curves as you have performed measurements.

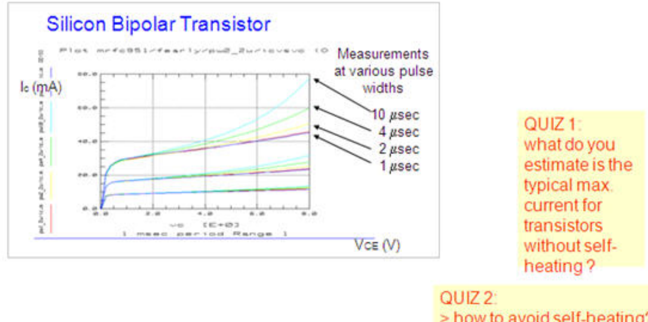
The plot in the slide above gives the chip temperature increase of a packaged transistor as a function of the pulse width and the applied bias power. The pulse period is 1s. As can be seen, self-heating can only be avoided when applying very short, pulsed measurements, below 1us pulse width. Such pulsed measurement systems, also included pulsed S-parameter measurements, are commercially available.

However, such systems are quite complex and expensive. Therefore, if you have to live with self-heating, make sure your device suffers always from the same self-heating. This means in general, apply the slowest measurement speed for you're your DC measurements, since your biased network measurements will be slow as well, and self-heating will definitively occur there.

What to do when self-heating becomes an issue for your device modeling:

- A min. pulse width of ~1us or below (see figure above). This means the DC analyzers in pulse mode cannot be used (pulses are 100us or longer !)
- Use the IC-CAP pulsed modeling system
- Live with self-heating, but make sure your data are consistent:
  - all DC measurements are performed with the same self-heating (slow measurements!)
  - same self-heating for biased S-parameter measurements as for the DCmeasurements.

Self-heating with DC measurements

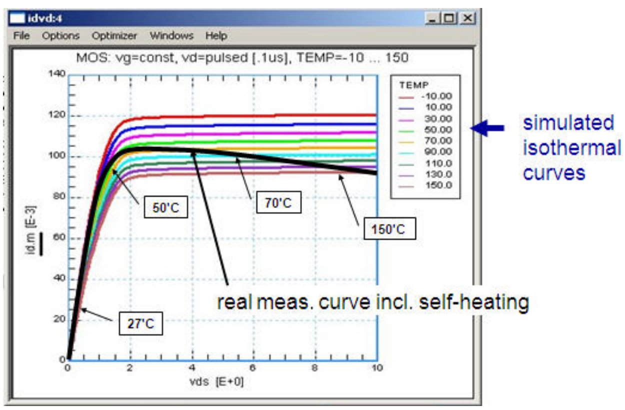


As a general observation, currents above ~25mA will cause self-heating effects with transistors. Rule of thumb: 25mA with typically 2V equals already 50mW for that tiny transistor on the wafer!

Self-heating can only be avoided when applying pulsed DC measurements. However, the max. duration for such a DC pulse is 1us, followed by typically 1ms wait time.

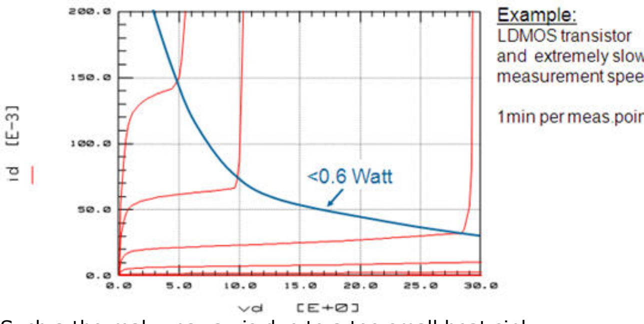
If you have to live with self-heating, make sure to have repeatable measurements for the DC Settings (currents for the transfer and for the output characteristics must be identical for identical bias conditions !!). Furthermore, the self-heating during the (fast) DC measurements must be the same as with the (slow) S-parameter measurements!!

Output Characteristics for a Single vg, at Different Chip Temperatures



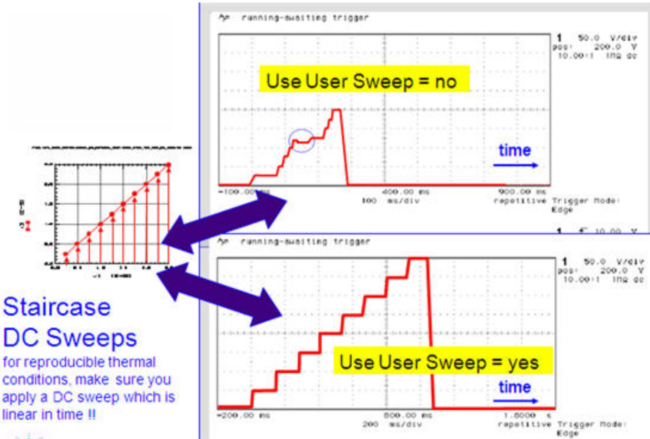
Note
for a clear picture, the isothermal data (-10'C to 150'C) above come from simulations.

Worst Case: Thermal Runaway due to a weak heat sink



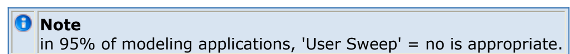
Such a thermal runaway is due to a too small heat sink.
A bigger practical problem, however, is the self-heating before the runaway happens !
Making thermally repeatable measurements in this bias conditions requires extremely slow measurements with long HOLD and DELAY settings.

Dealing with Self-Heating

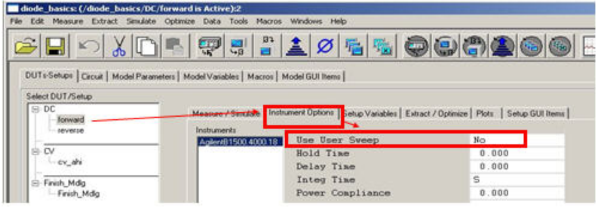


It is important to note that DC analyzers, like other instruments, have two operation modes: - either the controlling software sends the start-stop-step stimulus info and the 'what-to-measure' GPIB commands to the instrument and finally sends a trigger command ('User Sweep' = no), - or every individual measurement point is triggered individually by the software ('User Sweep' = yes).
As a consequence, the measurement vs. time is affected from these two possible settings. In the slide above, for 'User Sweep'=no, the CPU of the instrument is the owner, or controller of the measurement, and not the controlling software (like IC-CAP). In case of a DC analyzer, the CPU's firmware itself will control the measurements during the staircase sweep. After having received the 'trigger' command, it will wait for the measurement result to converge (end of charging effects), it will set the ranging of the amplifier (to have the max. resolution of the AD-converter) etc. This means that the time stepping of the staircase sweep is not monotonic. However, the measurement result (in e.g. IC-CAP) looks like a straight line, since we plot e.g. current vs. voltage, and ignore the time stepping. This explains some very special effects where e.g. the transfer curve current is different by some % from the output characteristics current, for identical bias voltages, and where self-heating can be excluded to explain the differences.
If the user wants to make sure that the staircase sweep stepping represents a more or less linear stepping in time, he should set the 'User Sweep' = yes, see the next slide. However, due to the huge activity on the GPIB bus, the measurement may take much longer in this case.

When setting 'User Sweep' = yes, the controlling software (e.g. IC-CAP) is the 'owner of the sweep'. Due to the relatively slow performance of the GPIB bus compared to the instrument's measurement time, the resulting staircase sweep is time-wise pretty monotonic. Apply this instrument mode if you must assure such time dependence of your measurement.


Note
in 95% of modeling applications, 'User Sweep' = no is appropriate.

'Use User Sweep' can be set in Instrument Options

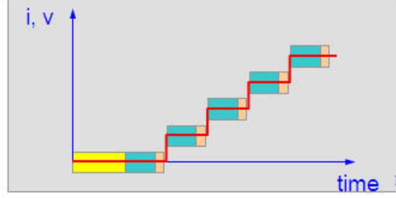


Some Quick Remarks on Trigger Statements

Hold Time
Time to allow for DC settling before starting internal or user sweep.

Delay Time
Time the instrument waits before taking a measurement at each step of an internal or user sweep.

Integ Time
Sets the integration mode for A/D converter (number of samples. It can be set to Long, Medium or Short.

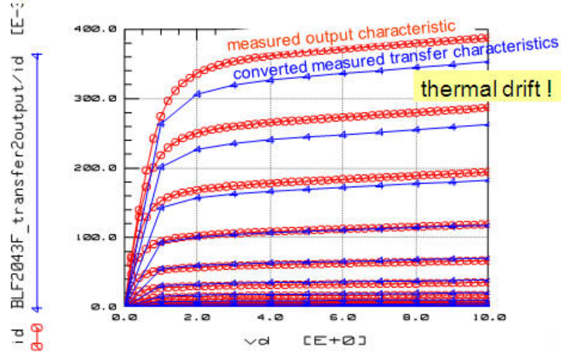


For CW measurements of Power Devices, make sure that at each measurement point, the final temperature rise has settled.
This is achieved by setting the DELAY time to an appropriate value, e.g. 1 minute.

For output characteristics, when the Drain voltage is re-swept at every incrementation of Vg, make sure to set the HOLD time to e.g. 20 minutes.
This will allow the device to cool down from the last vDmax bias condition to the new vD=0V condition at the next vG bias point !

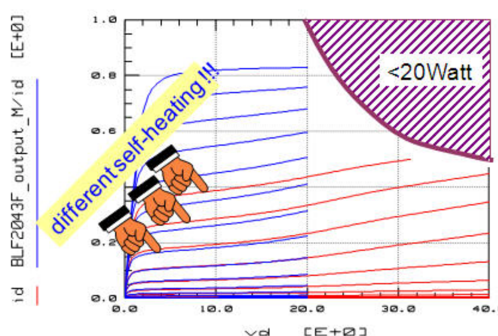
Thermal Effects - DC data consistency

In the plot, you see a conversion of a measured transfer characteristics into an output characteristic and comparing it with the measured output characteristic itself.



If a conversion of a transfer characteristic into a pseudo output measurement (or the other way round) exhibits exactly the same currents than the output characteristics measurement itself, the thermal conditions were identical. In the example above, this is not the case !

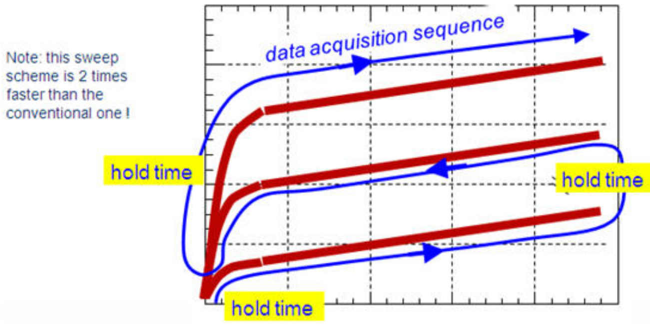
Output Characteristics Data Inconsistency for Power Devices



In the plot, you can see two output characteristics with different vdmax overlaid. The one with the bigger vdmax represents a more self-heated transistor than the other. To avoid this problem, set the Hold Time to a big value (minutes).

When two output characteristics have to be measured for the same device, due to power restrictions, make sure to have the same self-heating in both measurements.

A new measurement approach for power devices"up_down ramping" - faster measurements by making use of the cooling-down phase



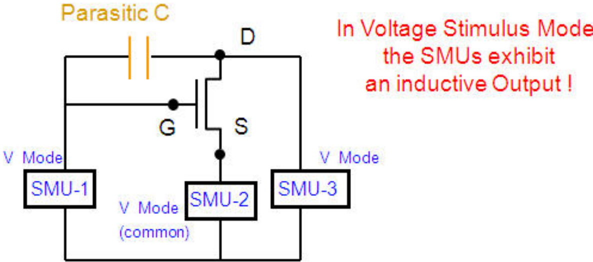
Note
This sweep scheme is 2 times faster than the conventional one!

Especially for packaged devices, the cooling down after reaching the max. of the 1st order sweep can take up to several 10 minutes.

When applying an up-down-ramping bias, the extremely long Hold time of a conventional back-to-zero 1st-order-sweep can be reduced when the 2nd order sweep steps to the next higher value and the 1st order sweeps just steps down.

An example of an instrument driver for such a sweep is in demo_features\2_meas_instruments\2user_def_meas_instr_drv\N67xx_100W_SMU_driver_by_PEL.mdl and then Setup 'Drive_67xx_updownramping_OutputCharact'.

Oscillations during DC measurements



Therefore, the measurement environment is equivalent to a Hartley oscillator !!

Under certain circumstances, SMUs may oscillate. Superbeta, wide- bandwidth bipolar transistors are especially susceptible to oscillation. But also GaAs transistors.
Oscillation can become a problem when using older DC probe needles. A common way to avoid such oscillations is using HF probes (Ground-Signal-Ground GSG probes), or the newer shielded DC needles.
If you have to live with your DC needles, and when oscillations occur, here some ideas on causes for such oscillations:

1. SMU Induced oscillationsThese oscillations occur when an inductive load is connected to the SMU's output. The SMU has often an inductive load because the output impedance of the SMU is inductive during V-Mode operation and usually several SMUs are indirectly connected together through the DUT.
2. Oscillation due to straysThe measurement system including the DUT, stray capacitance and residual inductance of the connection cables, switching matrix, probe card and/or test fixture can be recognized as an oscillation circuit.

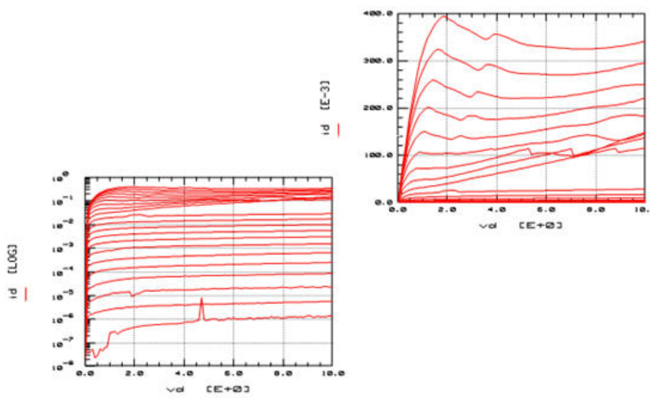
The oscillation detector of SMU may not detect this type of oscillation. This is understandable when thinking of the SMU output as a low-pass filter, while the oscillation frequency may be in the 100MHz ... to several GHz range. Also, if ~~like in most cases~~ the oscillation is located at the DUT, the SMU itself cannot do anything to prevent this oscillations, since the cables are long compared to the oscillation wavelength.
As a general rule: the oscillation has to be avoided where it happens.

Related to the example of the slide above, we should be aware of the total circuit: The SMUs connected to the MOSFET's gate and drain are operating in the V-Mode. Since SMUs typically appear to be inductive in V-mode, this makes this configuration equivalent to a Hartley oscillator !

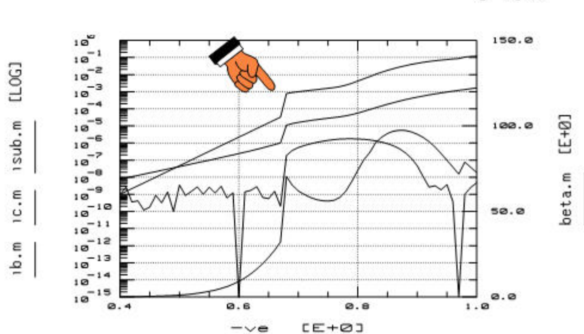
The SMU may oscillate if an unusually large inductance is connected to it. This could occur if the DUT is a superbeta transistor (big hFE) and the SMU connected to the emitter is set to one of the low current ranges.

For more details on conditions for oscillations, refer to the application note 356 -1 (publication number Agilent 5950-2954)

Self-oscillation with an LDMOS output characteristics



Self-Oscillation with a bipolar transistor



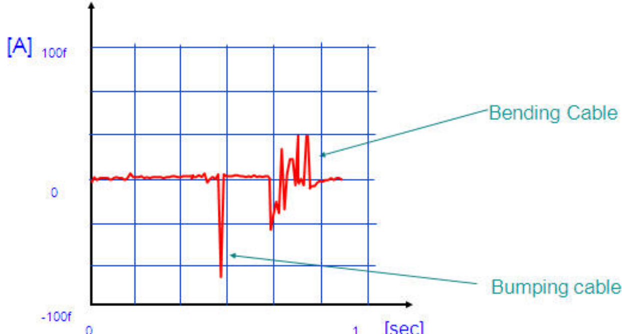
Note
The slide is from the book:
J.Berkner, Kompaktmodelle für Bipolartransistoren, Expert-Verlag Renningen (Germany), ISBN 3-8169-2085-3, February 2002

How to avoid oscillations

- Keep unshielded part of DC contacts as short as possible(Use shielded DC probes)
- Use high quality cable (low cable inductance)
- Use GSG (Ground-Signal-Ground) RF probes.

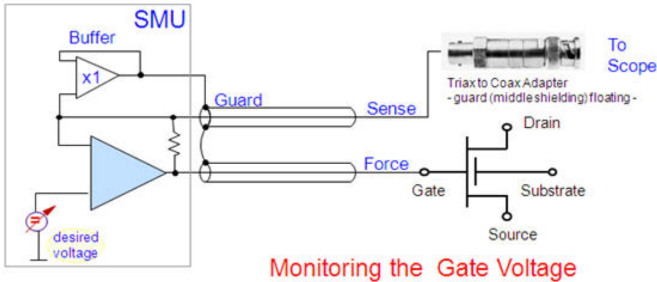
Note
Applying ferrite beads is not recommended for device modeling measurements since that trick does not work for S-parameter measurements.

Effect of cable movement



Monitoring the SMUs by an oscilloscope

Using the Sense Line as a High Impedance Scope Probe



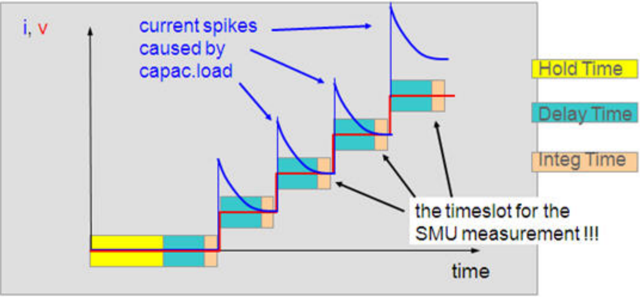
Besides for Kelvin measurements, the sense line can be used also for monitoring the SMU voltage on the device with an oscilloscope. The sense line tracks the force line within 1mV.

All you need is a floating guard triax->coax adapter attached to the sense line at the back of the 4156. Then use any BNC cable to direct connect the SMU sense line to the oscilloscope input. The adapter shown is the Trompeter Electronics AD-BJ20-E2-PL75.

Connecting triax SMU outputs to coax inputs

See *Connecting Triax SMU Outputs to Coax Inputs* (iccapmhb)

The Effect of Big Capacitive Loads



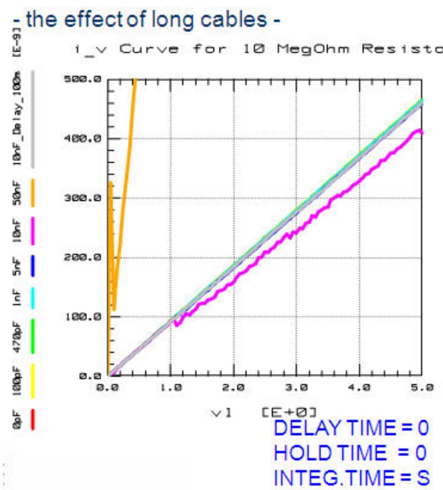
The Delay Time must be big enough to avoid measuring non-static charging effects

The Delay Time must be big enough to avoid measuring non-static charging effects. Big capacitive loads can lead to the case that the SMU begins triggering the measurement, while the charging has not yet finished.

See the DC Analyzer manuals for the max. applicable capacitive load. For example, the E5270A specifications are:

- Maximum capacitive load:
  - For 1 nA to 10 nA ranges:1000 pF
  - For 100 nA to 10 mA ranges:10 nF
  - For 100 mA to 1 A ranges:100 µF
- Maximum guard capacitance:900 pF
- Maximum shield capacitance:5000 pF

SMUs and Capacitive loads - The Effect of Long Cables



a 10MW resistor was measured with different capacitors in parallel. Capacitors exceeding ~1nF affect the measurement result due to a too long charging (settling) time, during which the SMUs already start to measure.

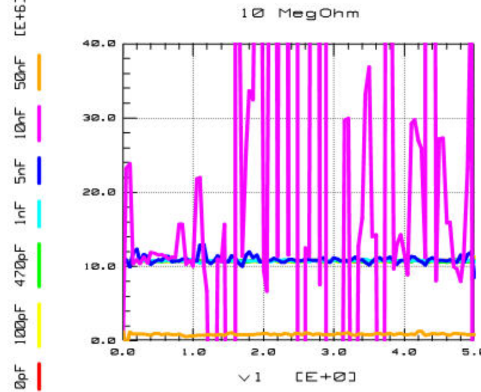
> **Note**
> The grey curve, i.e. the 10nF case "with extended SMU delay time (100ms)" brings the faulty 10nF trace "without delay time" (purple) back to the expected measurement trace. For 50nF, delay time extension did not help any more.

Triax and Quadrax Cables exhibit a capacitance between the signal and Guard. This capacitance should be kept below ~1nF.

> **Note**
> when using twice a single-triax cable for a Kelvin measurement, watch out that the capacitance of the twice a triax cable is doubled. If the DUT capacitance is high, Quadrax cables may therefore be a smarter solution.

Model	Length	Capacitance
16494A Triaxial Cable		
Option 001	1.5 m	125 pF
Option 002	3.0 m	240 pF
Option 003	80 cm	75 pF
16494B Kelvin Triaxial (Quadrax) Cable		
Option 001	1.5 m	140 pF
Option 002	3.0 m	260 pF
Option 003	80 cm	90 pF
16494C Kelvin Triaxial (Quadrax) Cable for 4142		
Option 001	1.5 m	140 pF
Option 002	3.0 m	260 pF

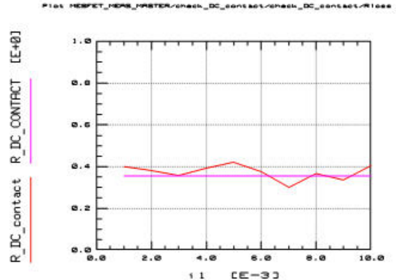
Note
How to Select Options for Cable Length?
The Agilent 4155C/4156C or Agilent 4142B have low current forcing and measurement capability by their SMU (Source Monitor Unit)technology.The SMU uses a kind of feedback circuit for maintaining very high stability of the current/voltage sourcing. Therefore,the guard capacitance (a load for SMUs)of connections to SMU must be within certain limits to prevent SMU instability.
The limit is about 900 pF in the Agilent 4155C/4156C or the Agilent 4142B 's case, for example.
You should select proper cable lengths to keep small guard capacitance for the SMUs.A system using switching matrices sometimes needs longer cable lengths than for the standalone instruments. If you plan to connect more than ~5 m total cable length (both instrument to switching matrix and switching matrix to the DUTs), make sure the cable guard capacitance does not exceed 900 pF.



The measurement result from above (i_vs_v) converted to the resistance value R = dv / di
In the slide above, the 10 MOhm resistor was stimulated with a swept voltage, the current was measured and the resistance was calculated by R = dv/di.
For capacitive loads above ~1nF, the calculated resistor value is distorted, a hint for resolution problems with the SMU.

Do not forget to evaluate your DC contact resistance and account for it when extracting your model DC resistor parameter values !

This is especially true when measuring Silicon devices on the wafer.



The contact resistance for probe contacts on silicon wafers ranges typically between 1-5 Ohm and can be bigger than the transistor resistances.

In IC-CAP, you can add a Test Circuit on the DUT level to account for the DC losses, without affecting the extracted model (which is specified in the Circuit tab).
For an example, load
demo_features\3_MEAS_ORGANIZE_n_VERIFY_DATA\0_MASTER_FILES\DIODE_MEAS_MASTERFILE_demodata_PELdep.mdl
and see the TestCircuit in DUT 'DC'. See also how this contact resistance is measured, in DUT 'verify_DC_probe_contact'.

Conclusion

When making DC measurements for device modeling,take into account:- shielding

technologies

- self heating- contact resistance
- self-oscillation challenge
- capacitive load effect

Miscellaneous Information

The IC-CAP SystemVariable MEASURE_FAST

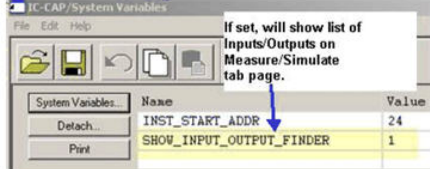
When defined as Yes, IC-CAP will attempt to minimize instrument re-initialization during repeated measurements on the same Setup. See Speeding Up Repetitive Measurements in the IC-CAP Manual, chapter on Measurement. Default is No.

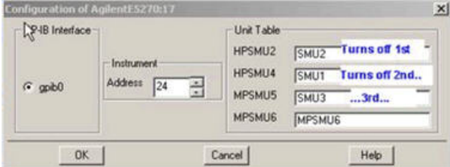
MEASURE_FAST will invoke an instrument initialization for the first measurement of a Setup. I.e. subsequent measurements of **the same Setup** will be faster, because non instrument initialization is done any more.

In order to suppress this initialization between Setups, you also have to set MDS_MEASURE_FAST.

Note
If you set MDS_MEASURE_FAST=1 at the top level, then even your very first measurement ever will not call an instrument initialization.

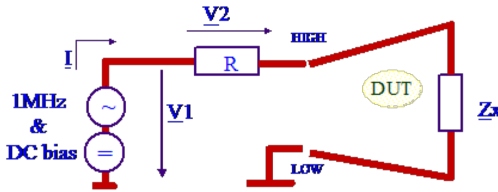
SMU turn-on and turn-off sequence





Slide 1

Slide 2



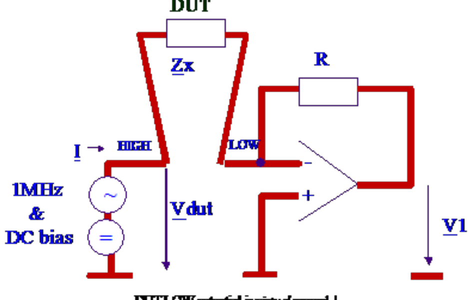
For CW measurements of Power Devices, make sure that at ***each*** measurement point, the ***final temperature rise*** has settled.
This is achieved by setting the DELAY time to an appropriate value, e.g. 1 minute.
For output characteristics, when the Drain voltage is re-swept at every incrementation of Vg, make sure to set the HOLD time to e.g. 20 minutes.
This will allow the device to cool down from the last vDmax bias condition to the new vD=0V condition at the next vG bias point !

Slide 3

$$Z_x = \frac{V_1}{I} - R = \left(\frac{V_1}{V_2} - 1\right) * R$$

Such a thermal runaway is due to a too small heat sink.
A bigger practical problem, however, is the self-heating before the runaway happens !
Making thermally repeatable measurements in this bias conditions requires extremely slow measurements with long HOLD and DELAY settings.

Slide 4



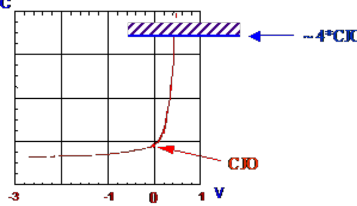
This is the case of a NPN power transistor mounted on a big heatsink:
In this example, we compare the 0 -> 10V output characteristics from a 0 -> 10V vCE sweep and a 0 -> 20V vCE sweep, in CW mode.
As can be seen, in the case of the vCE -> 20V sweep, when the vBE bounces back and is incremented for the next branch of the output characteristics, the transistor is still hot from the vCE=20V case and cools down. The iC-traces drop down at low vCE.

Slide 5

$$\frac{V_{DUT}}{Z_x} + \frac{V_1}{R} = 0$$

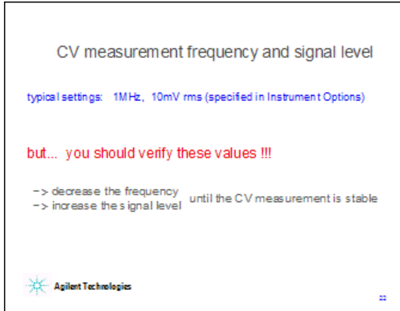
This is once again the case of a NPN power transistor mounted on a big heatsink:
In this example, we compare the 0 -> 10V output characteristics from a 0 -> 10V vCE sweep and a 0 -> 20V vCE sweep,
pulsed (80ms on, 40ms off) and CW.

Slide 6

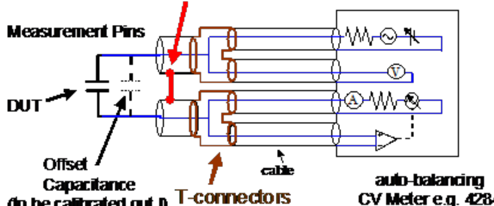


Especially for packaged devices, the cooling down after reaching the max. of the 1st order sweep can take up to several 10 minutes.
When applying an up-down-ramping bias, the extremely long Hold time of a conventional back-to-zero 1st-order-sweep can be reduced when the 2nd order sweep steps to the next higher value and the 1st order sweeps just steps down.

Slide 7



This is the best result obtained for the NPN 2N3055 power transistor in CW mode, but it is obvious that the applied delay time (120s) and the hold time of 15s was still at the lower limit.
Slide 8



But also up-down ramping does not prevent from faulty measurements. Make sure to have set the hold and delay times sufficiently high for consistent measurements !

Process Control Monitoring (PCM) Measurements versus Modeling Measurements

There is occasionally some confusion—between process engineers and modeling

engineers—about the term 'parameters'.

As an example, process engineers may define the threshold voltage of a MOS transistor as the gate voltage that is required to generate a certain Drain current of 1uA, while modeling engineers relate it to the formulation of the selected MOS model. Often, yet not quite model-specific, they define the threshold voltage to be the x-intersect of a line fitted to the steepest slope of the transfer curve Id(Vg).

Therefore, it is obvious that these two parameter values are different. However, both engineering groups refer to it as VTH. With bipolar transistors, it is the Early voltage VAF, that process engineers define as the x-intersect of a tangent to the output characteristics, while modeling engineers consider VAF together with other modeling parameters to achieve a fit of the simulated curves to the measured ones.

It is in the statistical analysis that both worlds come together. Using factor or principal components analysis, we attempt to explore statistical dependencies from a huge amount of data (process and modeling parameters) together. It can be shown that if the model parameters have been carefully determined (without using much optimization, as stated previously), they have a direct relationship to the process parameters. This means that after such a statistical analysis of process and modeling parameters (of both PMOS and NMOS transistors for example), the model parameters can be related to the fewer process parameters. These are measured continuously in production. This implies a considerable reduction of modeling times.

Another means of distinguishing between the process and modeling parameters is the method applied to measure these parameters. In production, spot measurements are usually performed, that is, individual measurement points. The process parameters are then calculated from these spots. These kinds of measurements are also called PCM measurements (process control monitoring) or Parametric Test measurements. Contrary to that, it is always curve sweeps that are required to do component modeling. The measurement curves and the simulation curves should be made identical.

The following tables are intended to further clarify this difference between both parameter worlds.

- Parametric Test (spot measurements):{*}Determination of parameters to describe and monitor a semiconductor production process measurement type: parameters to describe the process current-voltage (IV):gain, threshold voltages, delta_w, delta_L etc.capacitance-voltage (CV):doping concentrations and profiles, oxide and metal impurities,film thickness, carrier lifetimes etc.
- Modeling (sweep measurements):{*}Determination of modeling parameters to achieve a fit of mathematical curves to measured datameasurement type: parameters to fit the model curves to the measured curves MOS bipolar GaAscurrent-voltage (IV):VTH, U0, KP...IS, NF, BF...IS, NF, VTO, BETA...capacitance-voltage (CV):CJ, MJ, CGDO...CJO, MJ, VJ...CGSO, CGDO...S-parameters:ELM...TF, ITF, VTF...TAU, A5...

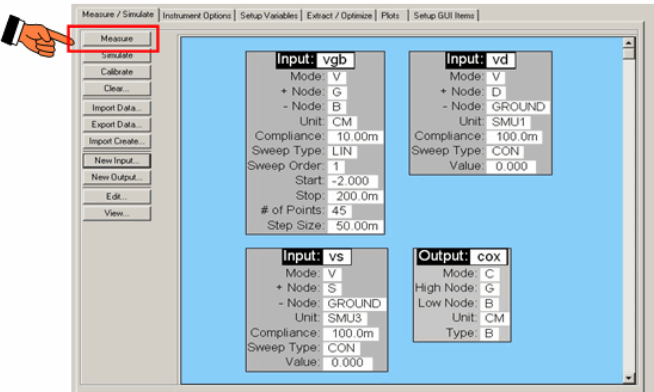
!
Figure: Process Control Measurements during chip manufacturing

CV

Contents

- *CV Characterization* (iccapmhb)
- *CV Measurement And Calibration Techniques* (iccapmhb)

CV Characterization



After a DC characterization, modeling engineers usually perform a so-called CV (capacitance versus voltage) measurement in order to characterize the device capacitances at a standard frequency of 1MHz. This frequency is high enough to allow a resolution down to a few femto-Ampere (provided shielded probes are applied for e.g. on-wafer measurements), yet still low enough to neglect second order parasitic like resistors in series with the capacitors, or like inductances.

For such CV measurements, the DC-bias is swept, a test frequency (1MHz) is applied to the DUT, and the instrument calculates the capacitance between the 2 pins of the DUT from the magnitude and phase of the device voltage and current. This means, an impedance meter interprete the measurement result always with respect to a user-specified schematic: either a capacitor in series with a resistors, or both in parallel. This explains, why capacitances and resistor values may vary with frequency when measured with such a device. In other words, these frequency-variations are due to a too simplistic analysis model behind the measurement. A better way is therefore to measure the capacitance with network analyzers. In this case, it is up to the user to interpret the measurement result (S-parameters).

Compared to DC, CV measurements are 4-terminal measurements too. The CV meter outputs are:* low current and potential

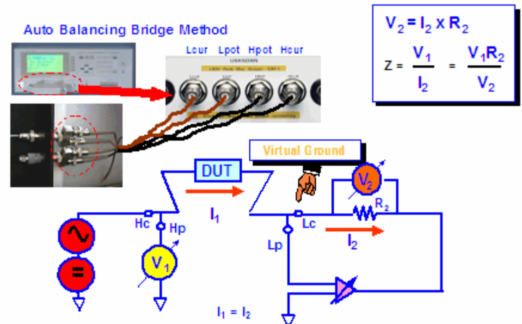
- high current and potential And similarly to the DC Kelvin measurement procedure, both the low and high pin are connected together as closely as possible to the DUT. In the auto-balancing method (next slide), the shielding of these four wires correspond to the virtual ground of the instrument's OpAmp, and not the chassis ground . This eliminates any influences caused by the cables.

CV Measurements - Calibration considerations

Test cables and fixtures contribute and affect the device characterization. For CV measurements, the calibration consists of unconnecting the DUT, assuming an ideal OPEN condition and measuring the cables and their OPEN parasitics (CV-Meter calibration). After that, the corresponding capacitance is automatically subtracted from the DUT measurement by the CV meter.

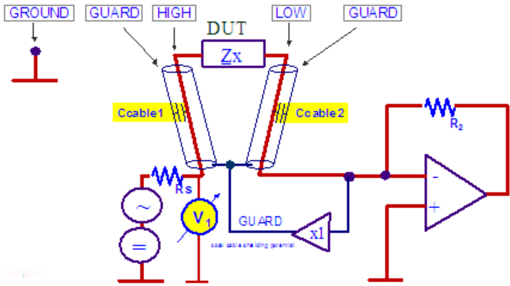
> **Note**
> If we are interested in the inner DUT's CV curves, i.e. without its surrounding test pads capacitances, we need to connect to an OPEN dummy structure during CV meter calibration instead of simply leaving the cables unconnected. Such an OPEN dummy consists of all connection pads, lines to the DUT etc, but without the inner DUT itself.

How do capacitance meters work?



For today's CV meters, the measurement principle is typically a so-called auto-balancing method. The slide above depicts the simplified measurement scheme.
The DUT is inserted in the feedback loop of an operational amplifier, and the system is stimulated with typically a 1MHz sinusoidal signal plus a DC bias. The feedback resistor R2 is precisely known, and the complex voltages V1 and V2 are measured. From the formula given above in the slide, the capacitance of the DUT can be calculated, assuming an equivalent schematic of either a resistor in series with the capacitor, or, commonly for modeling, a capacitor in parallel with a resistor (which is the bias-dependent diode resistance for example)

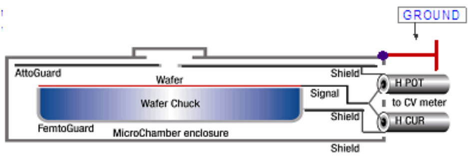
When applying the Auto-Balancing Bridge, capacitance from High and Low Terminal to Guard (e.g. coax cable shielding) are automatically excluded from the measurement result !!!



In the auto-balancing method, the shielding of these four coax cables are connected to the virtual ground of the instrument's OpAmp. A special feature of this auto-balancing method is that it is insensitive to stray capacitance towards (virtual) ground. As depicted in fig.3 below, this becomes clear since the parasitic capacitance Cp1 is tied directly to the supplying bias and stimulus voltages, and the parasitic capacitance Cp2 is located between virtual ground (minus input of the OpAmp) and ground. This means, both parasitic capacitance are not included in the measurement!

On-Wafer Measurements

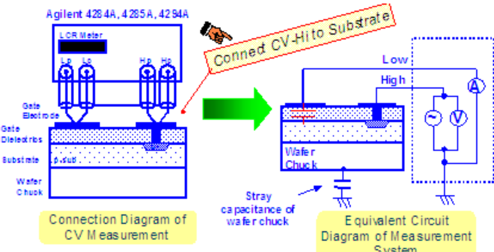
AttoGuard™ Enhanced CV Measurements



- FemtoGuard surrounds the chuck at shield ground
- Patented AttoGuard above the chuck at shield ground
- Creates a virtual double-shielded Faraday enclosure
  - 10 atto Farad CV measurement resolution
  - Zero CV meter only one time

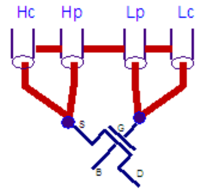
On-Wafer CV Measurement: where to connect the Hi and Lo of the CV meter

Do not connect the instrument's Low terminal to the substrate. The substrate is electrically connected to the prober's noisy ground (wafer chuck stray capacitance), and the Low terminal of the auto balancing bridge instrument is sensitive to noise. If the wafer chuck (stage) of the prober is isolated from the ground and effectively guarded, the shielding conductor of the 4TP cable can be connected to the prober's guard terminal to minimize stray capacitance around the probes.

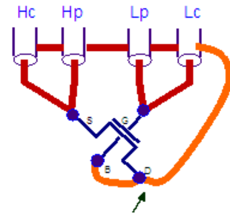


CV on-wafer measurements - What to do with the unused pins?

CV Shielding Techniques - avoiding the measurement of parasitic capacitance



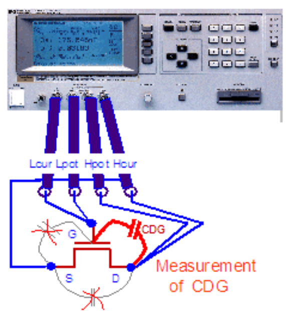
Connecting only 2 pins of a multi-pin device means that the capacitance between these 2 pins plus any other combination of capacitances between the 2 pins (CGD, CDB, CSB) will be measured !



Connecting the unused pins to the coax guard excludes the parasitic capacitances from the measurement. Only CGS is measured.

Due to the auto-balancing CV measurement principle, parasitic capacitances of the DUT pins (CGS, CDB, CGB) can either be included with the measurement of CGS (above left) or they can be excluded if the parasitic capacitance pins are connected to the guard (shielding) of the CV meter coax cables (above right).

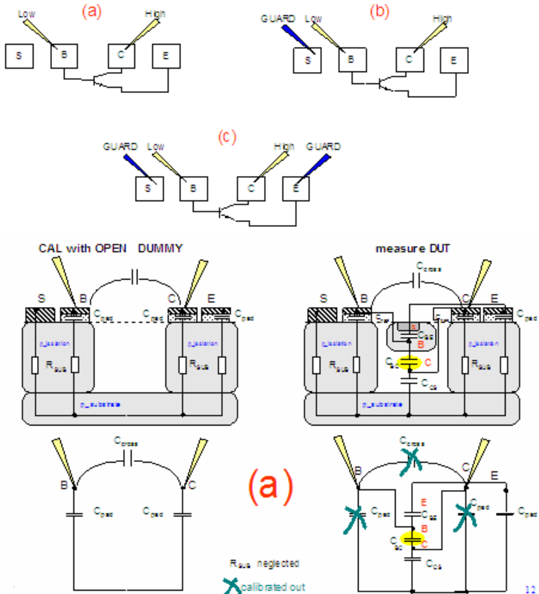
Application Example: Measuring CGD of a packaged MOS



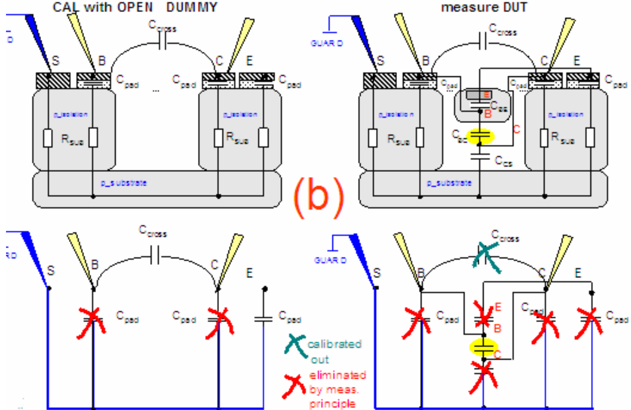
With the auto-balancing method, connecting the Source of the MOS to the Guard, the cable shielding potential, eliminates the effect of CGS and CDS.
Only C_{GD} is measured !!

> **Note**
> The shield of the 4 coax outputs of a LCRZ meter is virtual ground, not chassis ground ! Due to the applied measurement principle (autobalancing bridge), all capacitors between the measurement pins and that virtual ground are excluded from the measurement result, Only the capacitance between the pins is measured. However, if in the example above, the Source contact is left open, the capacitances CGS and CDS will be in series and altogether, will add as a parasitic capacitance to the originally desired capacitance CDG. In this case, the capacitance values measured with a LCRZ meter appear bigger compared to being measured with e.g. a network analyzer.

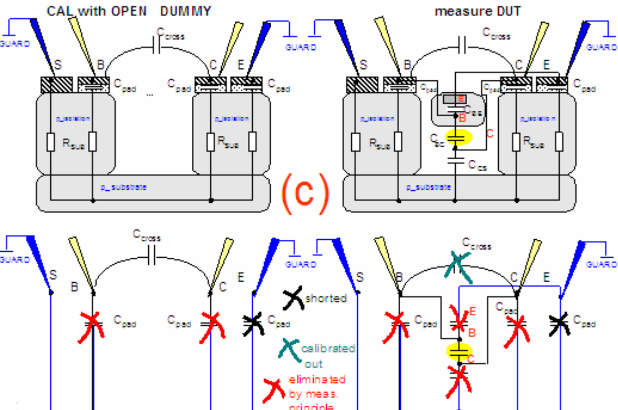
3 Contacting Cases - for CV Calibration and Measurement, Applying the Auto-Balancing Method



In this case, only the pads of the capacitance under test are contacted. For the calibration of the capacitance meter (and thus the compensation of the cables plus the contact pads capacitances), the pins are connected to an empty dummy structure first.
As shown, the calibration including the dummy structure reflects a parasitic capacitance of Cpad/2, while there is a bigger total parasitic capacitance present during the measurement of the DUT, where also C_{CS}, C_{BE} as well as the pad of the Emitter play a role !Since only $C_{pad/2}$ is calibrated out, the DUT measurement represents a too big capacitance.



In this case, the substrate is additionally connected to CV-meter ground. All other unused pins are left unconnected.



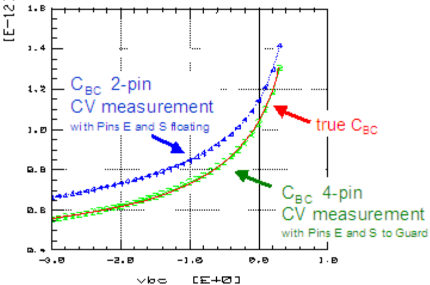
In this 3rd case, all unconnected pads of the 2-pin measurement, including the chuck, are connected with additional pins to the shielding of the CV meter coax cables (the Guard, which, again, represents virtual ground for the CV meter auto-balancing measurement principle).

With this forced guarding setup, all capacitances from the active pins towards guard are

not included in the measurement result. As can be seen in the sketch below, the measurement is the cleanest method to measure nothing but the desired CBC alone.

Proof related to the previous slides:Measuring a CV curve with floating Guard

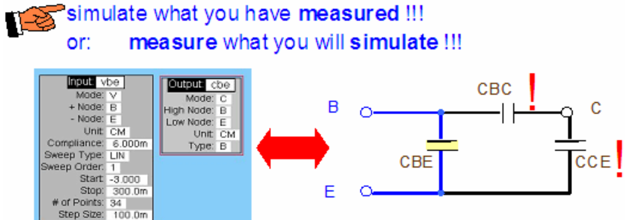
leads to a too big capacitance



Simulations of a Multi-pin Device CV Measurement

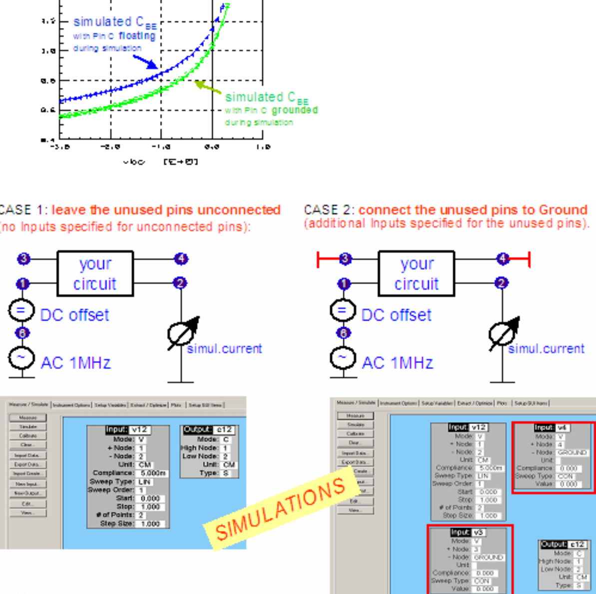
When doing modeling, you need to make sure that you simulate what you have measured.

CV simulations for devices with >2 pins simulate what you have measured or measure what you will simulate



Don't forget about the open pins of your DUT !
A CV simulation will always return the total capacitance between Base and Emitter , including CBC and CCE!

Simulating a CV curve with floating unused pins leads to a too big capacitance



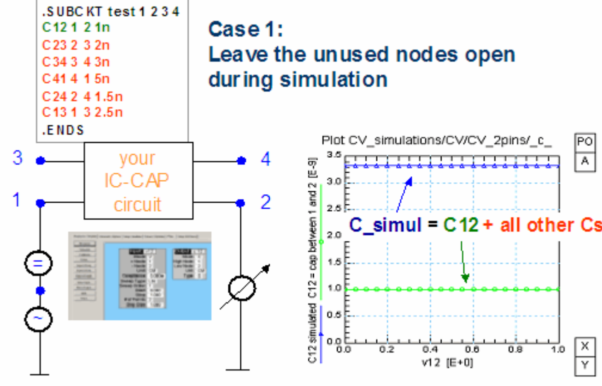
Note
How IC-CAP lets the simulator perform CV simulations:
From an inspection of the Simulation Debugger, IC-CAP surrounds your Circuit Netlist by the extra circuitry given above.
There are 2 cases:

- Case 1: leave the unused pins unconnected (no Inputs specified for unconnected pins).
- Case 2: connect the unused pins to Ground (additional Inputs specified for the unused pins).
  The simulated current value is then interpreted like this:
  From the simulated current of the schematic in the slide above, we obtain the complex admittance Y = iac / v.
  Since v=1, the equation is simply Y = iac
  Therefore, the capacitance value is calculated within IC-CAP from the imaginary part of the current after
  C = IMAG(Y) / (2*PI*1MHz) (1)
  and the conductance is
  G = REAL(Y) (2)

IC-CAP applies formula (1) when 'Mode' is set to 'C', and formula (2) when 'Mode' is set to 'G'. IC-CAP assumes a situation like with the default LCRZ meter setup: a capacitance in parallel with an conductance.

BACK TO THE PROBLEM 'WHAT TO DO WITH THE UNUSED PINS' DURING SIMULATIONS:
It now becomes obvious that in the multipin case, with the unused pins connected to Ground (by corresponding Inputs), these additional capacitors cannot contribute to the simul. result, since all their pins are tied to Ground. In the case of the pins left open during simulations (no Inputs specified for them in the IC-CAP Setup), all capacitances inside the DUT will be stimulated at pin1 and will therefore contribute to the simulated current at pin2. If you want a simulation at another frequency than the default 1MHz, specify the Model Variable CV_FREQ. If you want a measurement at another frequency, specify the meas. frequency in the Instrument Options.

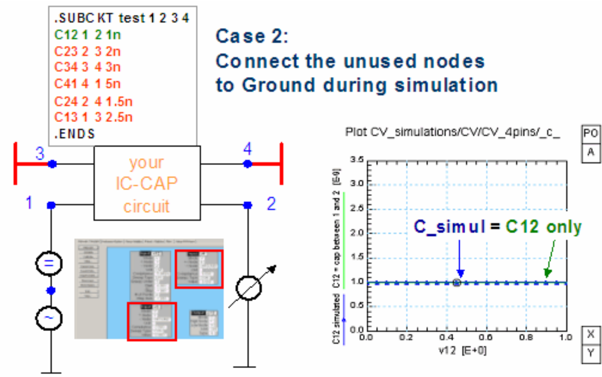
Case 1:Leave the unused nodes open during simulation



* additional surrounding network for SPICE CV simulations:

With the slide above and the next slide, we check the two scenarios. First, we try an IC-CAP Setup in which only one single Input referring to the two CV simulation nodes is specified. Also, a single Output is specified, referring to the same nodes as the Input. For our test, we specify a netlist with 4 nodes, having individual capacitances between all nodes, each with a different capacitance for easier identification of the result. See the slide above.
As can be seen in the Plot, the simulation result now includes the very capacitance between the two pins **plus** the influence of all other capacitances which are in parallel and in series-and-parallel to the very capacitance between the investigated nodes!!!

Case 2:Connect the unused nodes to Ground during simulation



* additional surrounding network for SPICE CV simulations:

In the other scenario, case 2, (see the slide above), the IC-CAP Setup has again an Input referring to the two CV nodes, and an Output referring to the same nodes as this Input. However, additional Inputs are specified this time, which tie the unused (open) nodes to Ground.

In this case, the simulation returns only the capacitance **between** the two CV Inputs, and thus, the simulation result corresponds to the recommended measurement of connecting the unused pins to the virtual ground of the CV meter, i.e. to the shielding of the coax

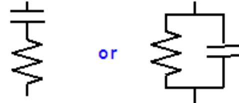
cables.

Considerations regarding the signal level and the CV-frequency

1. CV measurement signal level
   Typical settings: 10mV rms (specified in Instrument Options. You need to verify this value.

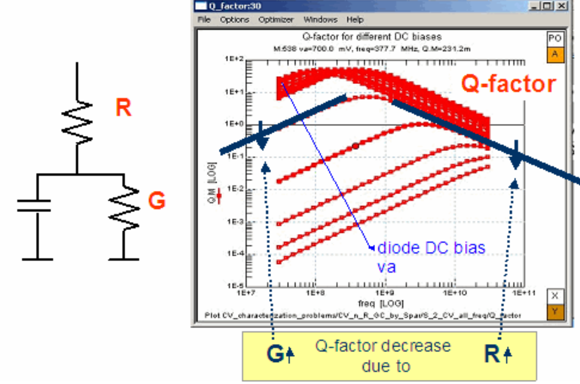
   > **Note**
   > increase the signal level until the CV measurement begins to change ! When this happens, reduce the signal level.

   Like with S-parameters, a CV measurement is a linear measurement, and therefore, harmonics should be avoided. As a consequence the applied CV signal level should not be too big, but also not too small (reduce signal-noise ratio). Since there is usually no DC current flowing at CV measurements, the reason for harmonics is the nonlinear capacitance itself.
   To find out the max. applicable signal level, start with a low signal value and increase it until you see a change in the measured capacitance value. Select a signal value with does not (yet) change the measurement result.
2. CV measurement frequency value
   Typical settings: 1MHz (specified in Instrument Options). The problem with CV meters and especially silicon devices is that the silicon may exhibit considerable losses, which are not represented by the CV meter's simple underlying schematic:

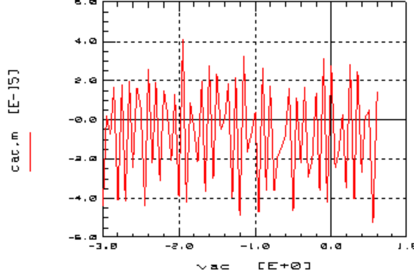
   > **Note**
   > Decrease the frequency until the CV measurement is stable.

   A CV (LCRZ) meter interprets the measurement result with respect to a capacitor in series with a resistor R, or as a capacitor in parallel with a conductor G. If this pre-requisition is not met by the DUT, the capacitance and/or the resistor or conductor become frequency dependent. In many cases, especially for silicon substrates, this can be the case. Therefore, check reducing the frequency from default 1MHz down to e.g. 100kHz and check if the measured CV curve is affected or not. If it is not or no longer affected, the mentioned simple underlying schematic is correct, since second-order effects are switched-off due to the reduced measurement frequency. We now can perform a correct CV measurement.
   Effect of a series resistor with CV measurements

   The parallel diode conductance G is responsible for the decrease of the Q factor to the left of its maximum, whereas the series resistor R is responsible to the right.
3. Maximum applicable DC bias for diode CV measurements
   Avoid DC saturation of the instrument !when measuring a CV curve into the On-state of the diode, don't apply a DC bias bigger than what corresponds to a capacitance ~4 times the 0V capacitance CJ0. When applying a bias voltage to this measurement method, however, special care most be taken to ensure that the measurement OpAmp is not saturated by a low resistance state DUT. This can easily happen when measuring the CV curves of diodes into the ON-state.
   Referring to the equation in the slide before, the real (ohmic) part of the resistance Zx of the DUT, i.e. the semiconductor resistance in parallel to the CV capacitance, has to be bigger than the reference resistor R. Otherwise, the output of the OpAmp can reach saturation (e.g. ±15V supply voltage). This happens easily when measuring semiconductor capacitances in forward bias mode. Here, the ohmic part of Zx tends towards 1W and lower!
   A practical approach: when measuring a CV curve into the On-state of the diode, don't apply a DC bias bigger than what corresponds to a capacitance ~3-5 times the 0V capacitance CJ0.

Obtainable CV measurement resolution



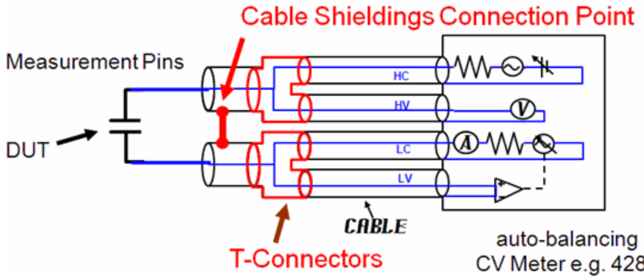
Provided that shielded CV probes are applied, a resolution down to Femto-Farad (fF) can be achieved. The slide above shows the measurement resolution of an OPEN dummy re-measuring, after calibration.

Conclusion

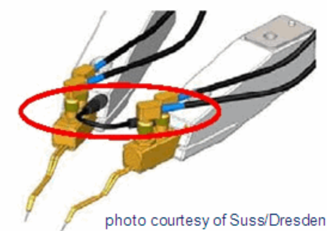
- When making CV measurements for device modeling, take into account what to do with the unused device pins
- check the max. applicable RF signal level
- check the max. applicable CV frequency
- when measuring conductive devices (diode in on-state), make sure to not measure higher than ~4*CJO.
- wafer substrate connected to the HI pin of CV meter

Miscellaneous Information

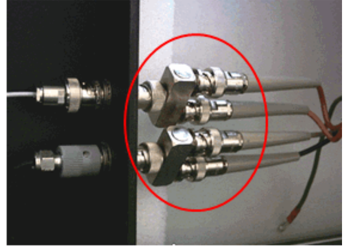
On-Wafer capacitance measurements with the auto-balancing bridge



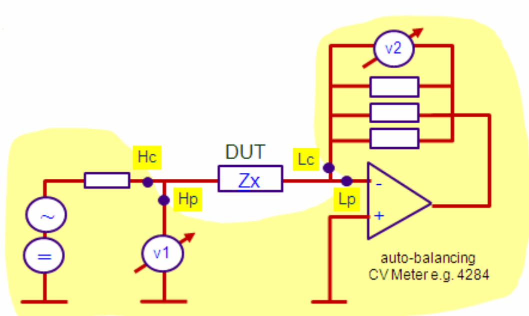
The coax shieldings should be connected together as close to the DUT as possible
Cable Shielding shorted at Probes



Cable Shieldings shorted at Prober-Input

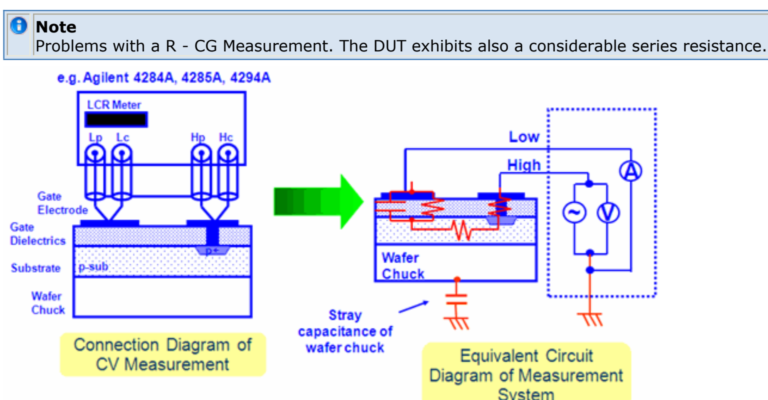


A more circuit-oriented view inside the CV meter

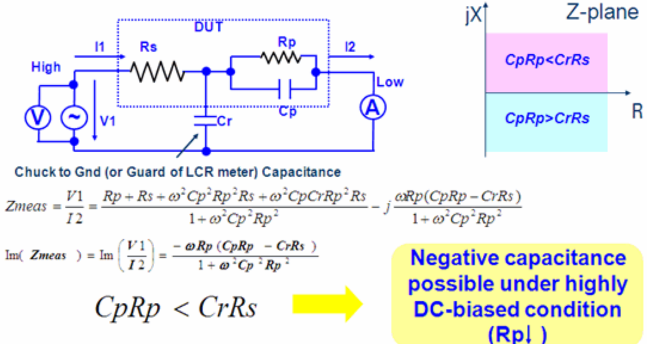


Negative Capacitance Effect Observed for On-Wafer CV Measurement*

Note
Problems with a R - CG Measurement. The DUT exhibits also a considerable series resistance.



This is the circuit diagram of the previous page.The impedance Zmeas that the LCR meter measures is expressed by V1 / I2



$$Zmeas = \frac{V1}{I2} = \frac{Rp + Rs + \omega^2 Cp^2 Rp^2 Rs + \omega^2 Cp Cr Rp^2 Rs}{1 + \omega^2 Cp^2 Rp^2} - j\frac{\omega Rp(CpRp - CrRs)}{1 + \omega^2 Cp^2 Rp^2}$$

$$\mathrm{Im}(Zmeas) = \mathrm{Im}\left(\frac{V1}{I2}\right) = \frac{-\omega Rp(CpRp - CrRs)}{1 + \omega^2 Cp^2 Rp^2}$$

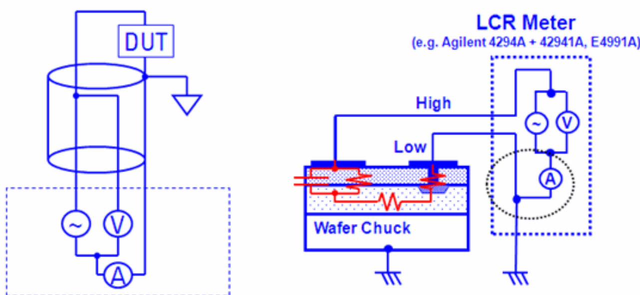
$$CpRp < CrRs$$

This is the circuit diagram of the previous page to show circuit analysis. The impedance Zmeas that the LCR meter measures is expressed by V1/I2.

Impedance Meter using IV Method On-Wafer*

- Advanced IV method : Agilent 4294A + 42941A
- RF IV method : Agilent E4994A

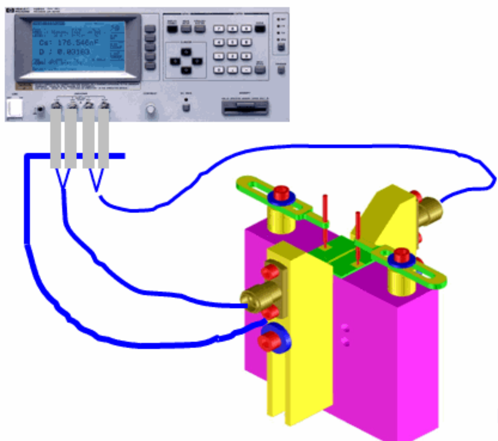
The advance I-V method is modified from the 4 terminal pair auto-balancing bridge to perform grounded measurement. The measurement resources, i.e. signal source, volt meter, and current meter, are positioned at the Gate side. In this case, one terminal of the on-wafer devices is connected to the substrate of wafer, and therefore, it is also connected to the wafer chuck. This grounded measurement config. is sometimes necessary to reduce the effects of wafer chuck



Packaged Devices

Measurement

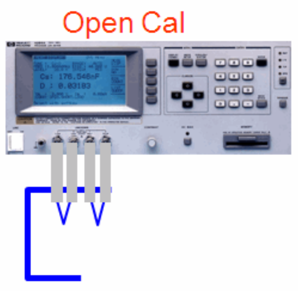
Test fixture ground and unused pinsconnected to CV meter Guard.

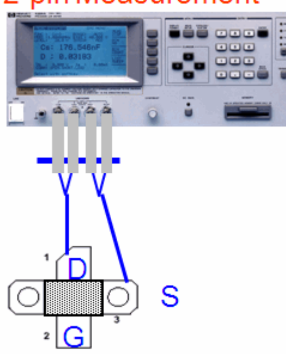


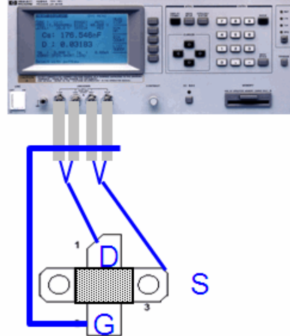
Test fixture ground and unused pinsconnected to CV meter Guard. For packaged devices, we first perform an OPEN CV calibration including the test fixture.

Measuring packaged devices

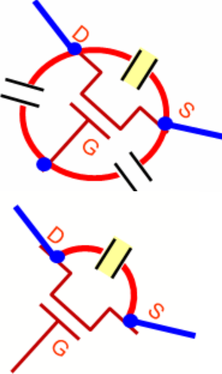
Measure







Simulate
Apply either 2-pin or fully guarded measurements,but later, simulate **exactly** what you have measured.

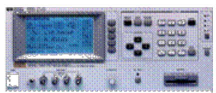


Then, for the measurement, we can either apply the 2-pin measurement (what yields the total capacitance between the measurement pins), or a guarded measurement yielding nothing but the desired capacitance between the pins D and S in the example below. Whatever measurement principle is more adequate in your case, make sure, however, that you simulate later **exactly** what you have measured before !!!

References

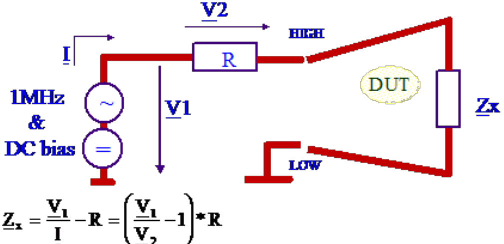
- Agilent Technologies Impedance Measurement Handbook July 2006P/N 5950-3000 : URL http://cp.literature.agilent.com/litweb/pdf/5950-3000.pdf
- Agilent AN 369-5 Multi-frequency C-V Measurement of SemiconductorsP/N 5950-2953 : URL http://cp.literature.agilent.com/litweb/pdf/5950-2953.pdf
- Agilent Evaluation of MOS Capacitor Oxide C-V Characteristics Using the 4294AP/N 5988-5102EN : URL http://cp.literature.agilent.com/litweb/pdf/5988-5102EN.pdf
- On-wafer Balanced Component Measurement using the ENA RF Network Analyzerwith the Cascade Microtech Probing SystemP/N 5988-5886EN : URL http://cp.literature.agilent.com/litweb/pdf/5988-5886EN.pdf
- IV and CV Measurement Using the Agilent B1500A MFCMU and SCUU P/N 5989-3608EN : URL http://cp.literature.agilent.com/litweb/pdf/5989-3608EN.pdf
- Agilent 4070 Series Accurate Capacitance Characterization at the Wafer LevelP/N 5968-2499E : URL http://cp.literature.agilent.com/litweb/pdf/5968-2499E.pdf

CV Measurement And Calibration Techniques



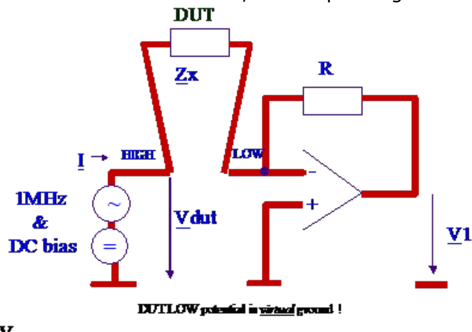
For CV (capacity versus voltage) measurements, two instrument settings are basically possible:- a capacitance in series with a resistor- a capacitance in parallel with a conductor.
This implies that when the device exhibits a more complex CV circuit than what is anticipated with one of these settings, the capacitance (and the resistance/conductance) will become a function of the measurement frequency. Therefore, for device modeling, a frequency of 1MHz is typically applied. This frequency is sufficiently high enough to achieve a good resolution with the capacitance meter (sufficient phase shift for the resolution of the instrument), but also sufficiently low to avoid second order effects (secondary capacitances, inductive influences etc.).

Regarding the measurement principle itself, there are two different methods available: -> the current/voltage method e.g. HP 4280 -> the auto-balancing method e.g. Agilent 4284etc. Both can include an additional DC bias voltage for the required CV measurements.
With the **current/voltage method** (fig.1), the unknown impedance Zx of the component under test is calculated from the stimulating voltage, V1, and the measured current, I. This current I is determined from the voltage drop, V2, along a known precision resistor, R. In practice, R is replaced by a low-loss transformer in order to circumvent the problems associated with a low-valued resistor. However, this limits the lower measurement frequency. On the other hand, this method has advantages over the auto-balancing method, in that it allows measurements of components connected to ground. More details see the literature /Honda/, /Haruta/.



$$\underline{Z}_x = \frac{V_1}{I} - R = \left(\frac{V_1}{V_2} - 1\right) * R$$

(1)Fig. 1: Capacitance measurement using the I-V method, e.g. HP 4280
In the **auto-balancing method**, shown in fig.2, the meter current again flows through both the DUT and the reference resistor, R. Yet, the negative input of the OpAmp acts as a virtual ground. The advantage of this method over the I-V is that the measurement range can be changed by selecting a suitable value for R. In practice, the OpAmp is replaced by a zero detector and a modulator, thus expanding the frequency range into the 100MHz



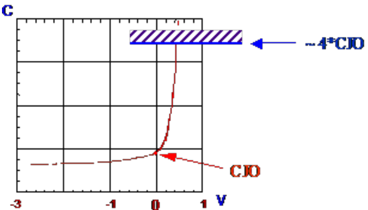
region.

$$\frac{V_{DUT}}{\underline{Z}_x} + \frac{V_1}{R} = 0$$

(2)Fig.2: Capacitance measurement using the auto-balancing method, e.g. Agilent 4284

Important hint about CV measurements of biased semiconductor capacitances using an auto-balancing bridge measurement

When adding applying a bias voltage to the auto-balancing method, however, special care most be taken to ensure that the measurement OpAmp is not saturated by a low resistance DUT. This can easily happen when measuring the CV curves of diodes into the ON-state.
Referring to equation (2) of fig.2, the real (ohmic) part of the resistance Zx of the DUT, i.e. the semiconductor resistance in parallel to the CV capacitance, has to be much bigger than the reference resistor R. Otherwise, the output of the OpAmp can reach saturation (?15V supply voltage). This happens e.g. when measuring semiconductor capacitances in forward bias mode, where the ohmic part of Zx falls rapidly from quasi-infinite to a few ohms !
A practical approach: when measuring a CV curve into the On-state of the diode, don't apply a DC bias bigger than what corresponds to a capacitance ~3-5 times the 0V capacitance CJ0.



CV Meter Settings:

CV measurement frequency and signal level

typical settings: 1MHz, 10mV rms (specified in Instrument Options)

but... you should verify these values !!!

-> decrease the frequency
-> increase the signal level
until the CV measurement is stable

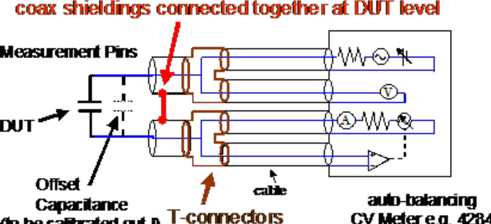
Agilent Technologies

Like with S-parameters, a CV measurement is a linear measurement, and therefore, harmonics should be avoided. Therefore the applied CV signal level should not be too big, but also not too small (reduce signal-noise ratio).
(Since there is usually no DC current flowing at CV measurements, the reason for harmonics is the nonlinear capacitance itself !!).
You can verify the max. applicable signal level by increasing/reducing its value until you get independent CV-curves.

Furthermore, a CV (LCRZ) meter interprets the measurement result with respect to a capacitor in series with a resistor R, or as a capacitor in parallel with a conductor G. If this pre-requisition is not met by the DUT, the capacitance and/or the resistor or conductor become frequency dependent. In many cases, especially for silicon, this is true.

Therefore, check reducing the frequency from default 1MHz down to e.g. 100kHz and check if the measured CV curve is affected or not. If it is not or no longer affected, the mentioned simple underlying schematic is correct, since second-order effects are switched-off due to the reduced measurement frequency. We now can perform a correct CV measurement.
Hint:in IC-CAP, for the CV measurement, specify the measurement frequency in Instrumnet Options, and declare the Setup Variable CV_FREQ for the simulation frequency.

CONNECTING THE DEVICE
A so-called four-wire method is used for both CV measurement methods. Similarly to the DC Kelvin measurement procedure, both the low and high pin of each port have to be tied together at the DUT.
Hint: For proper 4-wire CV measurements, it is mandatory to have a proper guard connection at the end of the 4 coax cables: connect all 4 coax shieldings together at the DUT level, as depicted below.



Capacitance measurement with the auto-balancing bridge
In the auto-balancing method, the shieldings of these four coax cables are connected to the virtual ground of the instrument's OpAmp. A special feature of this auto-balancing method is that it is insensitive to stray capacitances towards Guard, the coax cables shielding potential (virtual ground). As depicted in fig.3 below, this becomes clear since the parasitic capacitance Cp1 is tied directly to the supplying bias and stimulus voltages, and the parasitic capacitance Cp2 is located between virtual ground (minus input of the OpAmp) and ground. This means, both parasitic capacitances are ***not*** included in the measurement !

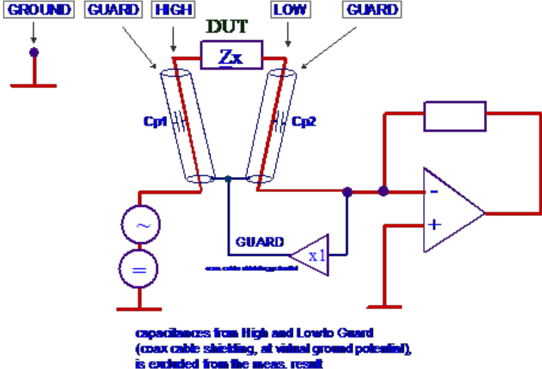


Fig.3: parasitic capacitances towards the coax cable shielding do not affect the Zx measurement.

Extending the Guard shield to the DUT itself:if DUT Zx exhibits not only a capacitance between its 2 connections, but also capacitances towards ground, these capacitances can either be included with the measurement (fig.4a) or they can be excluded if this ground is connected to the guard (shielding) of the CV meter coax cables (fig.4b).

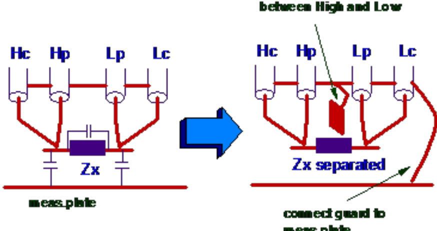


Fig. 4: Shielding techniques for auto-balancing CV measurements:

In the same way, parasitic capacitances from the 'hot' CV pins to the other pins of the DUT can be excluded frothe measurement result by connecting these extra pins to Guard, like depicted in fig. 5 below.

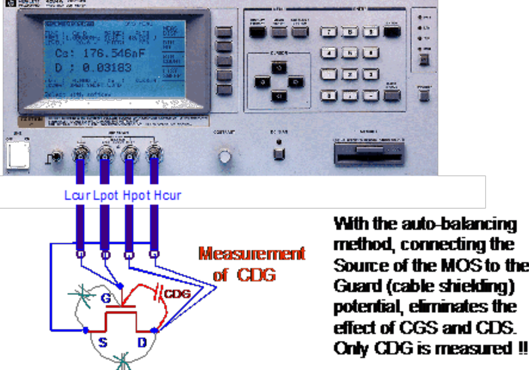


Fig. 5: Measuring semiconductor capacitances

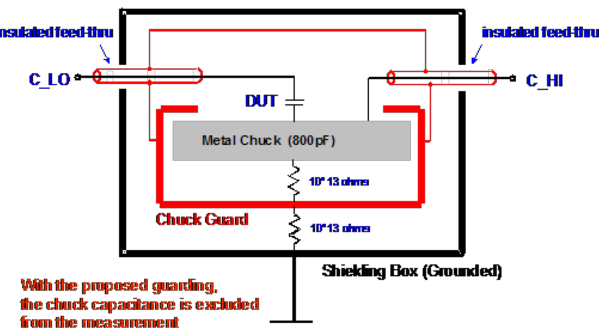
NOTE: **How to connect a MOS transistor for Gate Oxide CV measurements** see further down in section CV On-Wafer Measurement Considerations For MOS Transistors

Calibrating the CV Meter

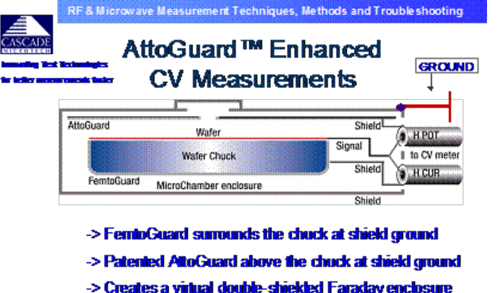
For modeling purposes, it is usually fully sufficient to perform an OPEN calibration (a SHORT calibration would be required if inductors were measured). This means to measure first the test fixture without the DUT (OPEN) and then, after this calibration was done, to automatically (within the instrument) subtract this OPEN capacitance value from the measurement. For measurements on the chip, an open dummy structure is contacted during the OPEN calibration.

Shielding Of On-Wafer CV Measurements

For on-wafer measurements, connect the chuck as sketched below:



The following slide of Cascade Microtech emphasizes this guarding technique once again:

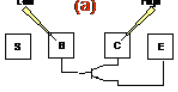


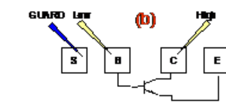
A study about 2-pin, 3-pin an 4-pin CV On-Wafer Measurements

This section is after the book:J.Berkner, Kompaktmodelle für Bipolartransistoren, Expert-Verlag Renningen (Germany),ISBN 3-8169-2085-3, Februar 2002
Notes:

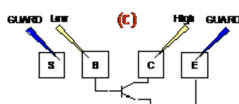
- This section refers to CV measurements applying the auto-balance measurement method.- The results of these considerations can also be applied to any other device like MOS or GaAs transistors.

  When measuring a transistor on wafer, there are three contacting scenarios possible, as depicted below: a 2-pin only method (leaving the chuck floating), a guarded Chuck situation (as recommended by wafer prober manufacturers) as well as a full 4-pin connection.

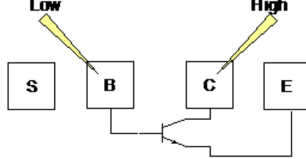


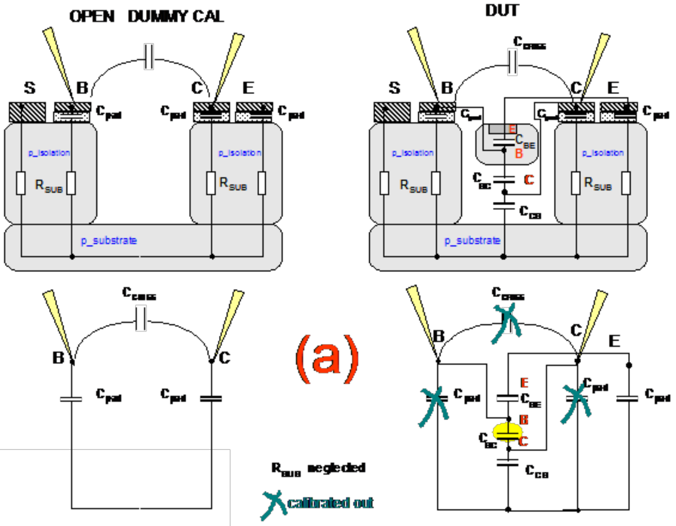


3 Contacting Cases
for CV Calibration and Measurement,
Applying the Auto-Balancing Method



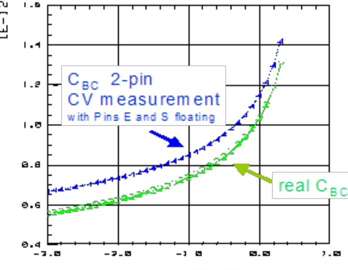
2-pin Measurements, case (a)
In this case, only the pads of the capacitance under test are contacted. For the calibration of the capacitance meter (and thus the compensation of the cables plus the contact pads capacitances), the pins are connected to an empty dummy structure first. However, applying this 2-pin method to the DUT measurement, all parallel capacitances are included in the measurement result as well, as shown below.



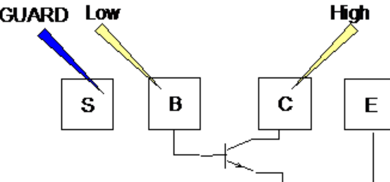


CV measurement of the BC junction capacitance with a 2-pin setup. left: measuring the Dummy-Deviceright: measuring the DUT.

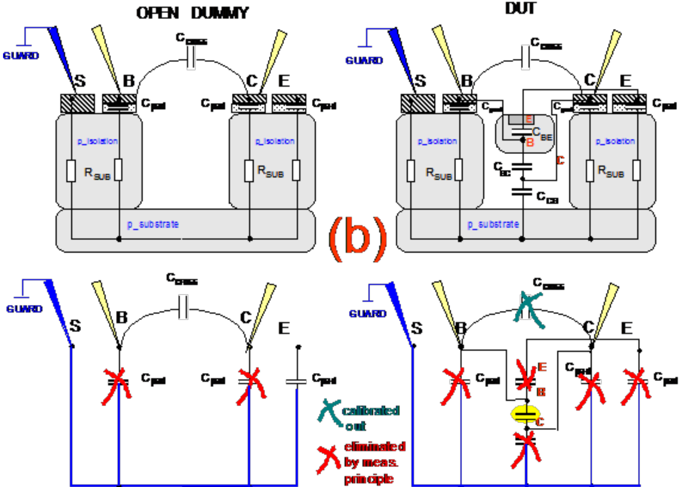
As depicted, the calibration including the dummy structure reflects a parasitic capacitance of $C_{paras} = \frac{C_{pad}}{2} + C_{cross}$
while there is a bigger capacitance present during the 'calibrated' measurement of the DUT, where also the other capacitances C_{CS}, C_{BE} as well as Cpad of the Emitter and the out-calibrated Cpad of the Collector play a role !Since only $C_{pad}/2 + C_{cross}$ is calibrated out, the DUT measurement represents a too big capacitance.



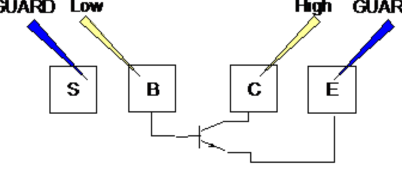
CV Measurements with guarded chuck, case (b)
This is the recommended setup by wafer prober manufacturers: -> the chuck is connected to the CV meter Guard:



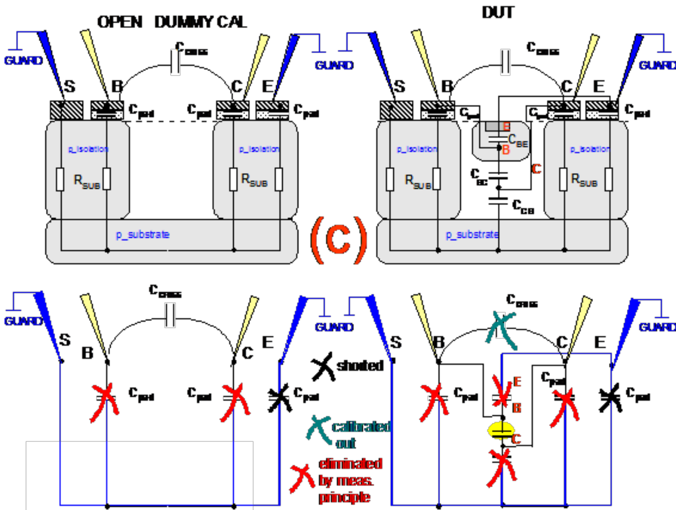
In this case, during the OPEN calibration, the two pads capacitances are towards the Guard potential of the CV meter, and, thus, cancelled out. The calibration basically accounts for C_{cross} and second order capacitance effects. The Emitter pin is open and not included in the calibration.
During the measurement, the Emitter pin plays a role: it gives a path to Guard for C_{BE} , and, thus, cancels that capacitance out. C_{CS} , with its connection to Substrate, is cancelled out due to the measurement principle. The measurement result is the capacitance C_{BC}, as desired.



CV measurement of the BC junction capacitance with guarded chuck (3-pin setup)left: measuring the Dummy-Deviceright: measuring the DUT.
4-pin Measurements, case (c)
In this case, all unconnected pads of the 2-pin measurement, including the chuck, are connected with additional pins to the Guard, i.e. shielding of the CV meter coax cables.



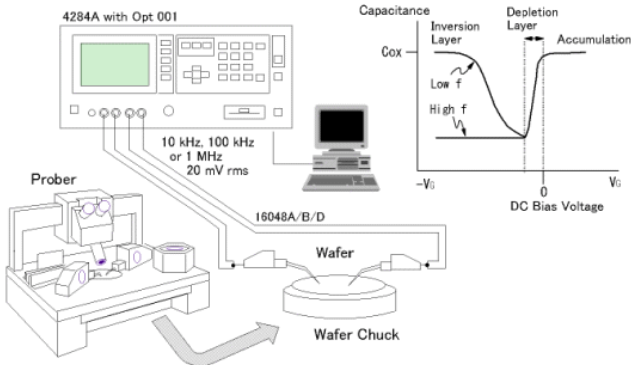
With this forced guarding setup, all capacitances from the active pins towards guard are **ex**cluded from the measurement result without any compromise. As can be seen in the sketch below, the measurement setup is the cleanest method to measure nothing but the desired C_{BC} alone.



CV measurement of the BC junction capacitance with a 4-pin setup.left: measuring the Dummy-Deviceright: measuring the DUT.

Important Hint:

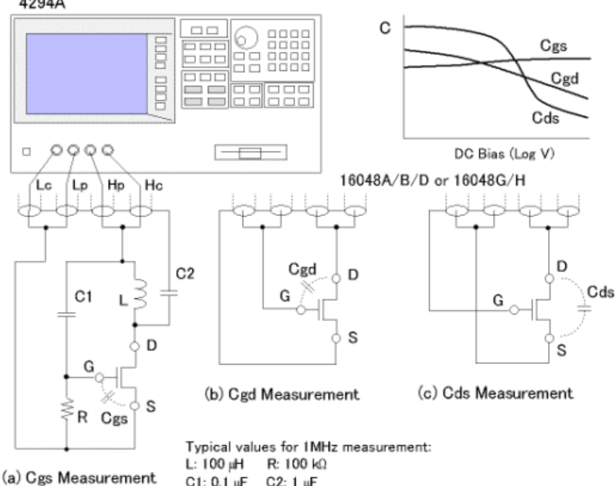
You may now think, 'why should I calibrate when all parasitics are eliminated anyway'.Do not forget that the CV meter calibration eliminates the offset capacitance *between* B and C, which is not eliminated by the CV measurement principle !
CV On-Wafer Measurement Considerations For MOS Transistors
This section is based on Hiroshi Haruta, Agilent Technologies Japan, June 22, 2000
CV characteristic measurement is the standard method of analyzing MOS device parameters including the thickness of gate insulation film, impurity concentration of the substrate, flat band voltage, surface carrier density, threshold voltage, etc. It is also performed to analyze the doping profile of a PN junction device.
How to set up the measurement:



An LCR Meter is connected to a wafer prober with test cables as illustrated.
What parameters are measured?
Voltage-capacitance characteristics are measured with a low-level signal (typically 20 mVrms) at 10 kHz, 100 kHz or 1 MHz where DC bias voltage is swept. Bias voltage is typically varied from -5 V to 5 V in 10 mV minimum steps.
Key measurement points and issues:
Very low capacitance needs to be measured with high accuracy and high resolution (1 fF resolution) at a low test signal level. Compensation must be properly performed at the prober tips. The AC (RF) prober with a guarded wafer stage should be used to avoid noise interference and to reduce the effects of stray capacitance around the probes. It is often difficult to measure MOS devices below the 0.1 mm design level because gate tunneling leakage current causes bias voltage error and test signal distortion.

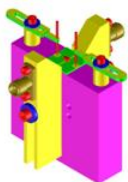
NOTE: **How to connect a MOS transistor for Gate Oxide CV measurements** when measuring a MOS Gate with a **balanced bridge instrument**, it is proposed to connect the **Gate to the low connectors, and the Substrate to the high connectors of the CV meter**.When measuring it using the **I-V method**, the Substrate is at low potential, and the Gate is connected to the signal terminal of the CV meterFor more details, see the publication: H.Suto et.al, 'Methodology for Accurate C-V Measurements of Gate Insulators below 1.5nm EOT', Extended Abstracts of the 2002 Int'l Conf. on Solid State Devices and Materials, Nagoya, pp. 748-749

In the example below, from the Agilent 4294A app.note, a special Setup for measuring the capacitances Cgs, Cgd or Cds of a MOS FET is proposed, with the drain voltage varied.

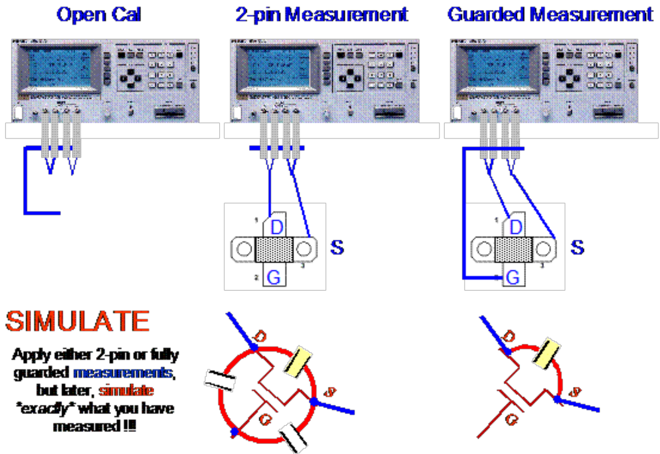


Measurement setups for Cgs, Cgd and Cds measurements are illustrated. The DC bias function of the instrument is used to apply the required drain voltage to the device. To measure the intended capacitance without the influence of other capacitances, the electrode of the device that is not connected to the measurement terminals must be connected to the guard (outer shield conductor of the test cable).

Key measurement points and issues.
The impedance of the guard connection lead should be as low as possible to minimize measurement error. The capacitor C2 used in Cgs measurement must have a sufficiently low impedance to thoroughly de-couple the signal at the drain electrode.
CV Measurement of Packaged Devices
For packaged devices, we first perform an OPEN calibration including the test fixture.



Example of a text fixture (Interconnect Microwave: www.icmicrowave.com)
Then, for the measurement, we can either apply the 2-pin measurement (what yields the total capacitance between the measurement pins), or a guarded measurement yielding nothing but the desired capacitance between the pins D and S in the example below.



IMPORTANT NOTE:Apply either 2-pin or fully guarded measurements, but later, simulate **exactly** what you have measured !!!

For device modeling, decide carefully what you want to measure!! You need to measure what will be simulated later!! Or, the other way around, make sure to simulate what you have measured before!!

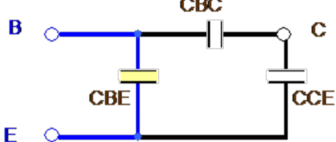
Device Modeling of CV Curves and Corresponding Measurement Aspects

1. Comparing CV measurements directly with CV model equations, without calling a simulator.
In this case, when you enter for example in IC-CAP the CV formula directly as a PEL program (without using a simulator), you must ensure that your measurement represents this situation as well. This is especially important for components with more than 2 nodes, e.g. a transistor. For a bipolar transistor for example, you may have programmed with a PEL function the CV formula

$$C_{BE} = \frac{C_{JE}}{\left(1 - \frac{V_{BE}}{V_{JE}}\right)^{M_{JE}}}$$

Therefore you need to make sure that the measured C_{BE} does not include effects of the C_{BC} nor the C_{CE} capacitance. This means, the open pins of the transistor (Collector and Substrate) have to be connected to the Guard, i.e. the shielding of the CV meter coax cables (virtual ground).
Of course, a proper CV meter calibration has to exclude the contact pad capacitances !
However, there is a caveat: don't forget that the DC behavior itself as well as the additional parasitics (RB, RE, RC etc) can also influence the CV measurement and simulation result !
The reason is that for a CV measurement, and also for the CV model equations as stand-alone, it is assumed that - there is nothing in series with the capacitance- there is nothing but a parallel resistance with the capacitanceBut with real measurements, - the conductivity in the ON state of diodes and transistors acts like a SHORT to the CV curve- and the transit time is overlying the CV measurement in the ON state.

2. Comparing CV measurements with CV simulations
This case is again a bit special for devices with more than 2 pins, e.g. transistors.In this case, remember that in the model itself (e.g. a bipolar transistor), there is also a capacitance between Base and Collector, and another one between Collector and Emitter !
I.e a CV simulation, with Input vBE (voltage) between nodes B and E, will return also the Base-Collector capacitance in series with the Collector-Emitter capacitance. Like depicted below, these 2 are in series, and together they are in parallel to CBE !!



Therefore, a standard CV simulation will always return the total capacitance between Base and Emitter , including C_{BC} and C_{CE}!

NOTE:
Measurement-wise, make sure that this simulation, i.e. the total effective capacitance between the two nodes, is respected with your measurement. Therefore, for the example of C_{BE}, it is ***not*** recommended to connect the open Collector to the virtual Ground of the capacitance meter !

HINT:
capacitances to Ground (e.g. for our bipolar example from above: C to GND, B to GND, E to GND and/or S to GND) do not contribute to CV simulations of the BE, BC, CE, and the CS capacitance !! Therefore, once again, make sure to also have them excluded with the measurements !!

HINT: if you want to check the simulated capacitance vs. the model equation itself, switch-off the Base-Collector and the Collector-Emitter capacitance, i.e. set CJC=0, CJS=0, or divide their actual values by 1E6 (to keep the previously extracted numbers !!).And like in the case before, don't forget that the DC model itself and the additional parasitics (RB, RE, RC etc) can also influence the CV simulation result! Therefore, for a bipolar transistor, divide the actual parameter values of IS, RB, RE and RC by 1E6 each for these CV simulations.
3. Comparing CV meter measurements with CV curves calculated from S-parameter measurements
Here, the CV curves are calculated out of (properly de-embedded) S-parameter curves. I.e. those S_deemb data are converted to Y-parameters, and we get

$$C_{BE} = \frac{1}{2 \cdot PI \cdot freq} \cdot IMAG(Y_{11} + Y_{12})$$

$$C_{BC} = \frac{1}{2 \cdot PI \cdot freq} \cdot IMAG(-Y_{12})$$

$$C_{CE} = \frac{1}{2 \cdot PI \cdot freq} \cdot IMAG(Y_{22} + Y_{12})$$

for a bipolar and the corresponding C_{GS}, C_{GD} and C_{DS} for a MOS.

These matrix conversion formulae assume that there is no component in series with the capacitances, and that there is nothing in parallel but a resistor.
Measurement-wise, these converted measurements should be compared with guarded or 4-pin CV measurements. See /D.MacSweeny/. Simulation-wise, they should be compared either

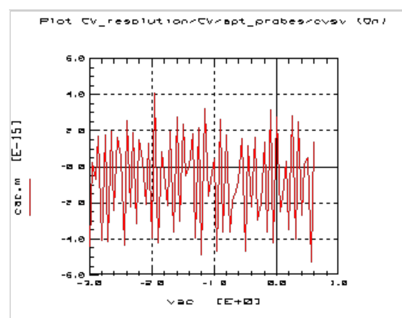
- with the pure CV model equations, - with a CV simulation where the other capacitors and resistors are switched-off, - or with S-parameter simulations converted to CV curves like the measurement data.

  Simulating converted S-parameter simulations is best, since it takes everything into account (the DC model and the outer parasitics), what corresponds exactly to the measurement situation as well!

  -> A Practical Approach for Modeling Transistor Capacitances:
  It has become clear, that CV measurements look easy and simple, but can be quite cumbersome depending on the actual device situation. Whatever method you prefer, keep in mind that the model parameters obtained from CV measurements should be considered as *first order CV model parameters*. The real world is non-linear RF, or at least S-parameter fitting. Therefore, a practical and accurate solution is- to de-embed the measured S-parameters, - convert them to Y-parameters, - show a plot of MAG and PHASE (or REAL/IMAG) of *all* Yxx parameters- simulate S-parameters, and convert them to these kinds of Y-parameters too,- and tune the capacitance parameters (and the other model parameters also).

  Hints on CV Meters:When configuring the CV meter, make sure that it is capable of measuring the complete voltage range, and not merely some individual voltages: e.g.: HP4284, **opt.001**

  Probes using specific shielded CV probes, a measurement resolution for on-wafer CV measurements down to a few Femto-Farad (fF) can be obtained with a CV meter, see the following calibration verification result (after calibration, remeasured OPEN dummy).



<u>Publications:</u>
Hiroshi Haruta, Agilent Technologies Impedance Measurement Handbook, 2nd edition,
Agilent Technologies product number 5950-3000
Note: the older version was:M.Honda, The Impedance Measurement Handbook, Yokogawa-Hewlett-Packard LTD,HP Literature Number 5950-3000
D. MacSweeney, Influence of Probing Conguration and Data Set Size for Bipolar Junction Capacitance Determination, ICMTS Conference March 2002,

Network Analyzers

Contents

- *Measurement* (iccapmhb)
- *Calibration* (iccapmhb)
- *S-Parameters* (iccapmhb)
- *De-embedding Techniques* (iccapmhb)
- *Vector Network Analyzer - Basics for Modeling Engineers* (iccapmhb)

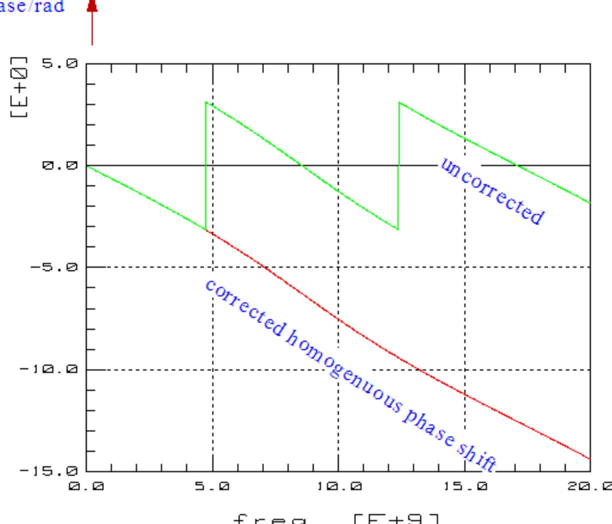
Measurement

Contents

- *Linear Vector Network Analyzer Measurements* (iccapmhb)
- *Displaying Phase Shift* (iccapmhb)

Displaying Phase Shift

This section covers a special PEL routine which corrects the steps in the phase of S-parameters. The following figure gives an example:



The PEL program given below shows how to implement this feature into IC-CAP.

```
!this routine adds 2*PI, if the phase sign changes and the phase step is
!more than 1.5*PI (to prevent from adding 2*PI at an ordinary phase
!cross-over!)
tmp=S.B
COMPLEX res.B.22[SIZE(tmp)]
!=============S11==========S11==========
!--------------treat meas.data.11 first
n = 0 ! counter for how many phase shifts have to be applied
i = 1 ! index
res.M.11[0] = ph(tmp.M.11[0])
WHILE i < SIZE(res)
 phi1 = ph(tmp.M.11[i-1])
 phi = ph(tmp.M.11[i])
 IF ((phi*phi1) < 0  AND (ABS(phi-phi1) >= 1.5*PI)) THEN n = n + 2
 res.M.11[i] = phi  -n*PI
 i =i + 1
END WHILE
!-------------------now treat simul.data.11
n = 0
i = 1
res.S.11[0] = ph(tmp.S.11[0])
WHILE i < SIZE(res)
 phi1 = ph(tmp.S.11[i-1])
 phi = ph(tmp.S.11[i])
 IF ((phi*phi1) < 0  AND (ABS(phi-phi1) >= 1.5*PI)) THEN n = n + 2
 res.S.11[i] = phi  -n*PI
 i =i + 1
END WHILE
!do the same with the partial arrays res.12, res.21, res.22 and finally return the whole array
RETURN res
```

Linear Vector Network Analyzer Measurements

In this section, you will learn about:

- Standalone VNA's
- DC biased VNA's
- Pad layout suggestions for transistor DC and S-parameter measurements
- Frequency Limits for Connectors

Standalone VNA's

In order to characterize components above some hundred Megahertz, vector network analyzers (VNA) are used. They measure the complex power transmission and reflection coefficients at two locations, Port 1 and Port 2. The VNA basically consists of a frequency-swept, high frequency generator (RF synthesizer), an S-Parameter Test Set to acquire the injected and reflected power at the DUT at both ports, and a control and display unit (Mainframe).

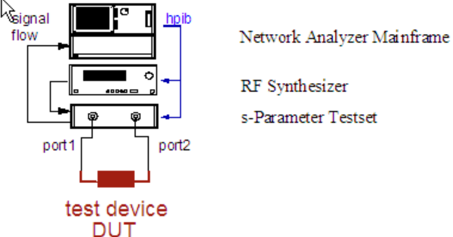


Figure: Vector Network Analyzer Structure

For the measurement, the DUT is connected to the two ports. The RF synthesizer signal is fed into the S-Parameter Testset and applied alternately to port 1 and port 2. Measured is the complex power signal of that specific frequency reflected back from the DUT to port 1, the power signal transmitted from port 1 to port 2, and then, when the RF signal is applied to port 2, the reverse behavior.
The block diagram of fig.2 shows the core of this meter combination, the S-parameter testset. The RF Input source at the top, connected to the RF synthesizer, provides the stimulus power. The PIN switch directs the signal to either a forward or a reverse S-parameter measurement. Directional couplers then detect the injected and reflected power of the DUT. The detected signals are then downcoverted into four IF signals for further analysis in the VNA mainframe, where each input is digitized and signal processed in order to give the S-parameters.

Blockdiagram of the S-Parameter Testset

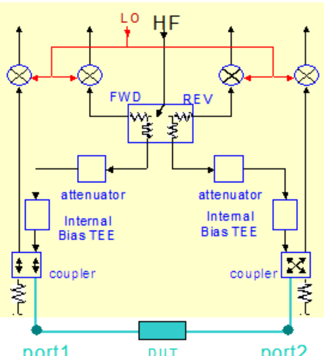
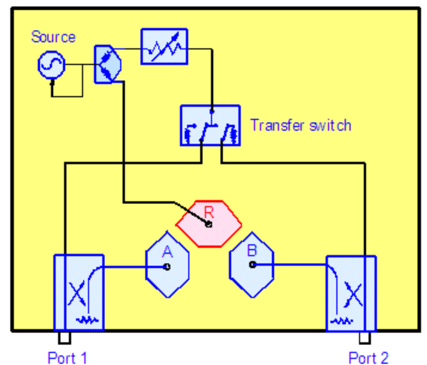


Figure: Block diagram of an S-Parameter Testset (4-sampler)

There are two main operating principles for VNAs: 3-sampler and 4-sampler VNAs. In case of a 3-sampler, the reference signal is coupled-out before the Port1-Port2 switch. Therefore, the switching errors are not covered by the calibration. For a 4-sampler VNA, however, where the reference signals are detected after the switch, a better calibration can be performed, including all error terms. Also, with the 4-sampler VNA, the user has complete freedom with regard to his own calibration techniques and does not need to use simplified procedures such as TLR* and LRM*. (Refer to Agilent Application Notes).



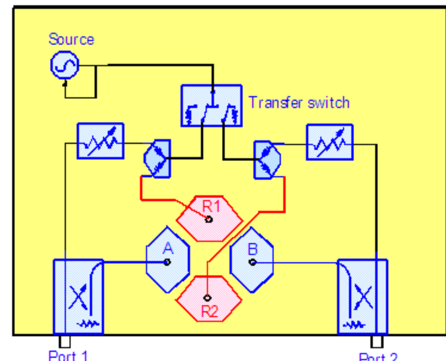


Figure: Principle of a 3-sampler and a 4-sampler VNA

> **Note**
> In both cases, the signal power at the reference port R should not be too small. Otherwise, the downconverter might have phase locking problems. The error message reads in this case: "phase lock lost". However, if the signal is too big, signal distortion may/will happen when measuring nonlinear devices such as Transistors or diodes. This has to be absolutely avoided ! Otherwise, the VNA only measures the fundamental frequency, ignores the harmonic frequencies, and the measured S-parameters are completely wrong !!
> They will not look distorted. They exhibit typical traces, however, the magnitude and the phase is wrong if the measurement was overdriven, and thus, non-linear!!!

DC Biased VNA's

For S-Parameter measurements of active components, the operating point (DC bias) must also be set and provided. Applying the DC bias Force-Sense technique, the measurement setup then appears as shown in fig. 4. In this example, we apply external bias TEEs. Their internal schematic is given in fig. 5. These bias TEEs allow Force-Sense techniques down to some 10 nano-Ampere.
Using external bias TEEs reduces the 2? internal series resistance to typical remaining 1?, and avoids the 1M? ESD protection resistance of the older VNA S-Parameter Testsets (electro-static discharge).

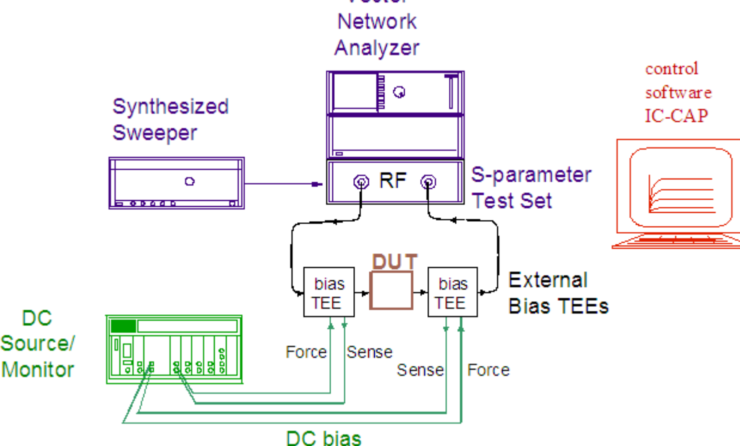


Figure: VNA measurements with DC bias

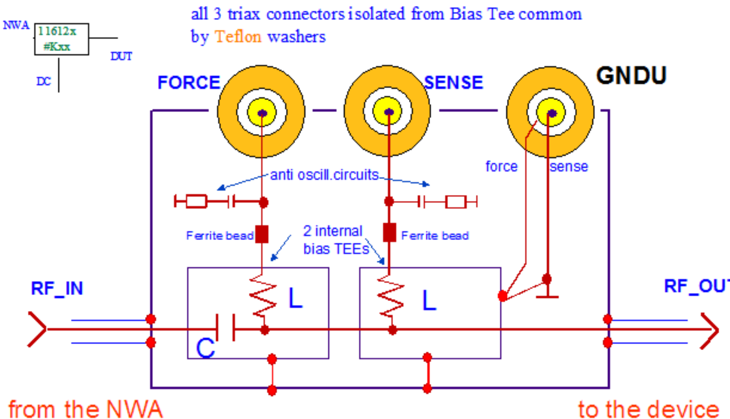


Figure: The inner circuitry of the external bias TEE HP11612

However, also for these special bias TEEs, the total ohmic loss is not 0? Furthermore, as can be seen from the schematic above, the ground unit GNDU of the DC analyzer has to be connected always in order to have the DC ground tied together with the NWA ground !

About The Resolution of the External Bias Tees

The reason for the roughly 1nA DC current resolution limit of the external bias TEEs HP11612 is that with the internal anti-oscillation circuits, there is a capacitor of roughly 5nF to ground for each DC port. This may confuse the internal firmware of the DC analyzer. Because these instruments are optimized for use with semiconductors, they assume a capacitive load of less than ~500pF. The charging/uncharging of such parasitic capacitors is taken into account by the firmware's algorithm which decides, when to begin with the DC current measurements. If there are bigger capacitive loads associated with the DUT, the internal logic considers the current to be settled and starts the current measurement. In fact, it still charges the parasitic capacitor and thus measures a current which belongs rather to charging the capacitor rather than the DC current of the semiconductor! Because the anti-oscillation capacitors are selected for rather worst-case conditions, they may be replaced by smaller ones (~500pF). Of course, this depends on the specific transistor. Also, it could be considered to add ferrite beads to the bias-TEE's DC path.

Conclusion About BIAS Tee

GaAs modeling engineers usually simply apply standard bias TEEs without the force-sense biasing. Silicon modeling engineers sometimes prefer the force-sense bias TEEs. As it was mentioned above, in both case the DC bias loss is not 0?. For this reason, the DC losses have to be taken into account anyway, with the Dut Test Circuit in the IC-CAP ModelFile.

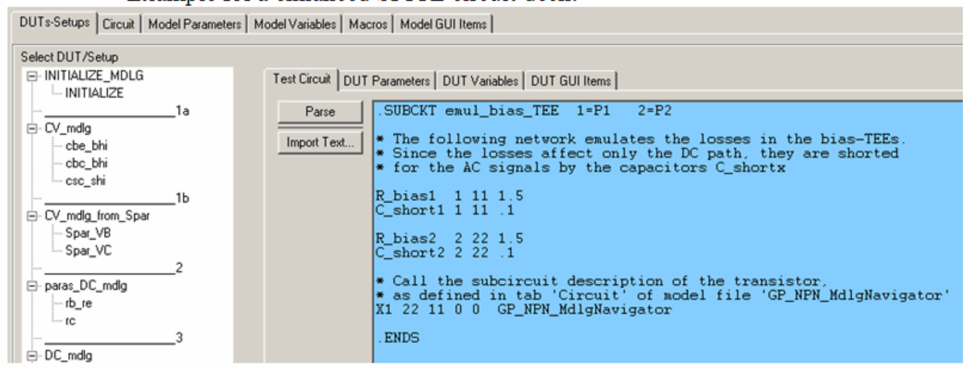
As a consequence of all this, if the additional ~1.5? loss in the DC bias path can be tolerated, you can also connect the triax cables of your DC analyzer via a triax-coax adapter (triax middle shield (guard) left OPEN!), directly to the rear DC connectors of your NWA.
Simply cover the ohmic loss by specifying an appropriate TestCircuit in the DUT level of your IC-CAP ModelFile.

About Using the Internal Bias Tees of the VNA's S-Parameter Testset

Note that there are a few things to consider: For adapting the DC bias to the rear panel of the S-parameter testset ,make sure to apply a pair of TRIAX(f) - BNC(m) ADAPTERs with the triax-middle-shielding unconnected.

Triax to Coax Adapter
- guard (middle shielding) floating -

The internal bias TEEs do not allow the employment of the Force-Sense technique, however, this is no problem when specifying these DC losses in the IC-CAP Test Circuit level. Typically, the S-Parameter Testset's internal bias TEEs have an overall resistance between 2 .. 2.5? modern Agilent PNAs typically show 1? Thiswill lead to a lower test voltage at the DUT than desired, if applying more than a few mA bias current. This effect has to be taken into account by adding such a series resistor to each connection of the transistor in the simulator circuit description, AC-wise shorted by a big capacitor (without these big capacitors, the simulator would not use the correct bias setting for its S-parameter simulations). Remember: only with these additional resistors in place, curve fitting can be correct! Otherwise, we would compare simulated to measured S-parameters whose bias conditions were different!)
Example for a enhanced SPICE circuit deck:



> **Note**
> For older HP VNAs, each port has an internal ESD-protection resistance of 1M? to ground. When measuring both, the S-parameters and the DC bias current, this would reduce the min. current resolution to a few micro-Ampere.

Measurement Hints

It may be smart to use 2 additional SMUs of the DC analyzers for the S-Parameter DC bias.
In this case, no recabling is required when measuring S-parameters. All what is needed is to tell the IC-CAP software to use the SMUs for DC measurements and to use the other SMUs for the S-parameter biasing.

VNA Dynamic Range

Terminate e.g. Port2 with a 50 Ohm LOAD cal.standard and measure S21.
The VNA dynamic range is then the difference between max. applickable RF signal level dBm and the measured noise level of MAG(S21)dBm .
You can improve the dynamic range by reducing the IF bandwith. Typical settings for modern Agilent PNAs is 10Hz. This results in a typical signal-noise ration of ~100dB.
For older HP VNAs, you should increase the averaging instead.

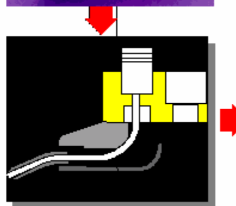
Note
Measurement resolution improvement for PNAs is recommended via IF bandwidth reduction, and not increased averaging.

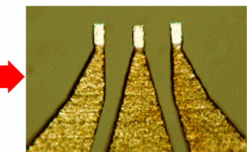
Probing on the Wafer

The slides and text of this section are copied with friendly permission from the file 'RF-Microwave Training Presentation.pdf' of the Cascade Microtech Winter 2002 European Probing Seminar CD-ROM

The probe transitions from coaxial to Air Co-Planar

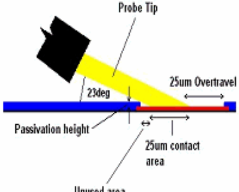
- Low-loss Teflon dielectric coax
- Microwave absorber
  - consistent attenuation
  - termination of coaxial shield energy
  - provides rigidity
- Fabricated probe tips
  - Uniform and compliant probe contacts
  - Tight Impedance control

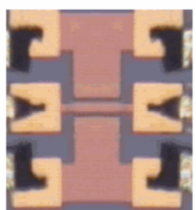




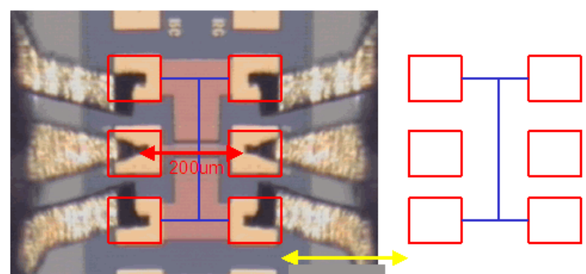
What pad sizes should I use?

- Recommended minimum pad is 80um x 80um for ACP Probes when performing automated measurements
- Smaller pad dimensions can be used for manual probing
- HPC Probe Allows 40um x 70um manual probing
- Passivation height must be considered
- Pad height variation must not exceed 25um





What about probe positioning?



- RF probes should have more than 200um separation to avoid cross-talk
- All pads must be on top surface
- All grounds should be connected together
- Adjacent devices should be >500um away for mm-wave measurements

What are the problems with probing Silicon wafers with Aluminium pads?

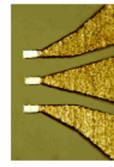
- Aluminium Oxide on Aluminium pads
  - A layer of Al Oxide will grow on the pad surface when left in air
    - This leads to possible contact resistances and variable contact resistance with time
- Conductive substrate increases parasitic reactance
  - Pad and interconnect capacitance and inductances become more significant during device measurement
    - De-embedding of pads and interconnects is required
- Limitations of Pad Parasitic Removal methods
  - The larger the pads and smaller the device, makes de-embedding more difficult to achieve

How do I overcome the contact resistance problem?

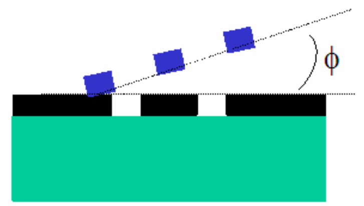
- Must penetrate Oxide on Aluminium pads
  - Standard BeCu tips are usable
    - but multiple touchdown are required to remove the oxide layer from the pad
  - Tungsten tips are superior
    - but the tungsten tip will also oxidise in air
    - Probing Al pads works well with W probes since both metals are very hard and rugged and perform a self-cleaning action when contact is made
      - Lower contact resistance
      - Better stability over time and temperature
      - Improved measurement repeatability

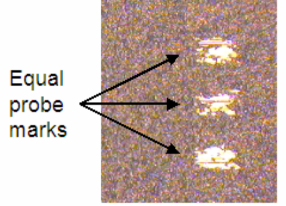
Where do I start?

- Ensure that the probes are in place
- Clean and connect the cables and torque using relevant wrench
  - Use IPA and swab to clean connectors and allow to dry
- Visually inspect the probe tips and clean if contaminated
  - Use IPA and swab, brushing away from the probe body and allow to dry
- Planarize the probes on the Contact Substrate inspecting the probe marks for even GSG contacts
  - Adjust the positioner planarity until all tips make even contact



What is planarization?



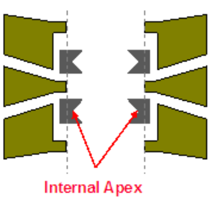


*definition: *planarization* - the ability to ensure all contacts are at the same height

- Use cable strain relief on positioners
- Use Contact Substrate to planarise probes

Use the ISS substrate to check the recommended 25um overtravel of your probing contacts:

Where do I need to place my probes?

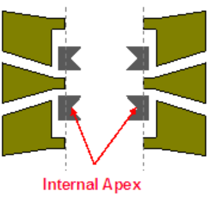


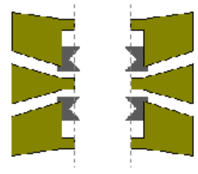
Initial contact **Full skate and overtravel**

Used to set 'skate' and probe separation

The following figure gives and idea, how transistor layouts should look like to cover both, DC measurements and biased S-parameter measurements. For bipolar transistors, it is valid correspondingly.

Where do I need to place my probes?





Used to set 'skate' and probe separation

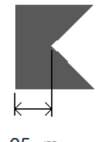


Figure: Layout for DC and S-parameter measurements of transistors

For S-parameter measurements, a good Grounding is essential, and Ground loops must be avoided. Therefore, the Bulk and Substrate contacts are shorted to Ground reliably with the above layout and accurate S-parameter measurement results can be obtained. If an individual DC bias is required for the Bulk or Substrate contact, it is recommended to use Ground-Signal probes, and to use a 3rd probe of this type to bias the 3rd contact. This reduces possible device oscillation considerably.
If an individual DC bias is required for the Bulk or Substrate contact, it is recommended to use Ground-Signal probes, and to use a 3rd probe of this type to bias the 3rd contact. This reduces possible device oscillation considerably.

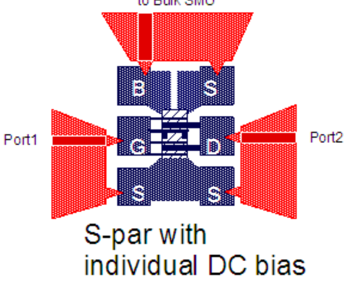


Figure: Layout for S-parameter Measurements and Individual DC bias for Bulk or Substrate, using Ground-Signal Probes

Note
The distance between the center of the pads is called pitch. Typical pitches for on-wafer transistor modeling is about or below 100um.

See also:
T.E.Kolding, A Four-Step Method for De-Embedding Gigahertz on-Wafer CMOS Measurements, IEEE Transactions on Electronic Devices, vol.47, no.4, April 2000

Frequency Limits For Cables and Connectors

Connector	Frequency Limit
3.5mm	< ~ 30GHz
2.4mm	< ~ 50GHz
1.85mm	< ~ 65GHz
1mm	<~110GHz

Web Info

www.agilent.com: search for Application Note 95-1, S Parameter Techniques for Faster, More Accurate Network Designs.

Publications

1. Paul Schmitz, Vector Measurements of High Frequency Networks, Hewlett-Packard, Publication HP5958-0387, April 1989
2. Agilent Technologies Application Notes
   1287-1: Understanding the Fundamental Principles of Vector Network Analyzers, Pub.No. 5965-7710E, 1997
   1287-2: Exploring the Architectures of Network Analyzers, Pub.No 5965-7708E, 1997
   1287-4: Network Analyzer Measurements: Filter and Amplifier Examples, Pub.No5965-7710E, 1997
3. Hewlett-Packard 1997 Back to Basics Seminar, D.Ballo, Network Analyzer Basics, Pub.No. 5965-7917E

4. Hewlett-Packard Application Note 8510-5A: Specifying Calibration Standards for the HP8510 Network Analyzer, Publication HP5956-4352, February 1997
5. Cascade Microtech, Microwave Wafer Probe Calibration Constants, HP8510 Network Analyzer Input, Instruction Manual, 1990
6. Cascade Microtech Application Note: On Wafer Vector Network Analyzer Calibration and Measurements, 1997, Pub No. PYRPROBE-0597
7. Cascade Microtech Application Note: Layout Rules for GHz Probing, 1992, Pub.No LAYOUT19
8. CascadeMicrotech Tecnical Brief: A Guide to Better Vector Network Analyzer Calibrations for Probe-Tip Measurements
9. Cascade Microtech, Winter 2002 European Probing Seminar CD-ROM

Calibration

Contents

- *Calibration Procedures* (iccapmhb)
- *Defining Custom Calkits* (iccapmhb)
- *Introduction to NWA Calibration* (iccapmhb)
- *NWA Calibration With Gating in the Time Domain* (iccapmhb)
- *VNA Cal Kit Manager* (iccapmhb)

Calibration Procedures

In calibration procedures, you will learn about:

- *NWA Calibration in General* (iccapmhb)
- *Entering CalKit Data into the Network Analyzer* (iccapmhb)
- *Entering Calkit Data for On-Wafer Probes* (iccapmhb)
- *Entering Calkit Data for Packaged Devices* (iccapmhb)
- *Entering Calkit Data for Connectorized Devices* (iccapmhb)
- *Entering SOLT Calkit Data - HP Network Analyzer specific* (iccapmhb)
- *Entering SOLT Calkit Data - Agilent PNA Network Analyzer Specific* (iccapmhb)
- *Entering TRL Calkit Data - Agilent Network Analyzer specific* (iccapmhb)
- *Checking Maximum Applicable RF Signal Level* (iccapmhb)
- *NWA Instrument Options in IC-CAP* (iccapmhb)
- *Performing the NWA Calibration from IC-CAP* (iccapmhb)
- *NWA Calibration Verification* (iccapmhb)
- *DC Bias Losses During S-Parameter Measurements* (iccapmhb)
- *NWA Calibration Procedure for At-Temperature Measurements* (iccapmhb)

Checking the max. Applicable RF Signal Level

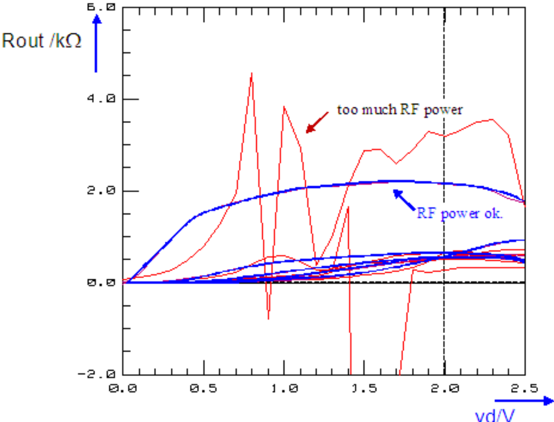
It is very important to not overdrive the test device with too much RF signal. As mentioned, it must be absolutely ensured that particularly the transistor's output signal is not clipping. When this happens, the transistor output signal is no longer a sine function, i.e. the transistor behaves nonlinear. This is a contradiction to a linear NWA ! In this case, the NWA will only measure the fundamental frequency, and ignore the occurring harmonics !

On the other hand, a too small RF signal will result in noisy S-Parameter measurements.

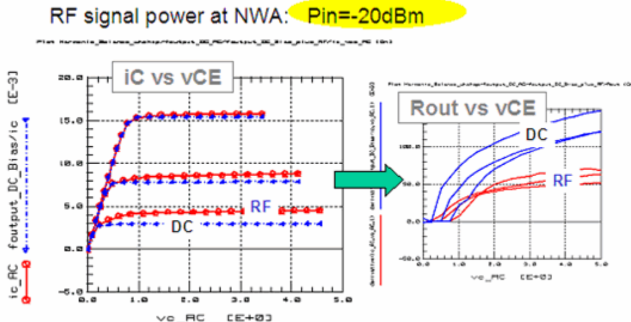
In this context, keep in mind that:
-> a typical RF signal level for modeling transistors is about -40dBm
i.e. the NWA port1 is set to -30dBm, plus additional ~10dB attenuation of cables and connectors.
-> -40dBm corresponds to a RMS voltage of ~2.2mV, what corresponds to ~3mV peak or ~6mV peak-peak !
-> small signal condition for a transistor is usually defined as VT / 10 = 2.7 mV
-> Last not least, when harmonics occur, i.e. the test device is overdriven,
the DC operating point is affected and thus, the DC bias current is affected.

We can use this effect of a shifted DC bias to identify the max. applicable RF signal for NWA measurements with a smart and simple measurement method :
When measuring a DC output characteristics and calculating Rout out of it, the resulting curve is very sensitive. Therefore, we can use this plot to identify possible effects of a too big a AC power applied to the transistor. This means, we measure the DC output characteristics, and let the NWA operate in continuous mode, i.e. unsynchronized to the DC measurement. When this happens, i.e. when the operating point is shifted (ib and ic for bipolar, id for MOS), harmonics do occur! We then reduce the RF power a bit and know the max. allowed RF power for the NWA S-parameter measurements!

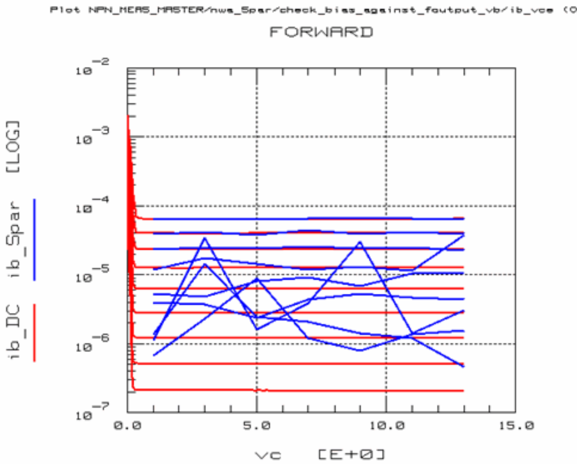
The following plot reflects such a test. The disturbed curves happen when there is too much RF power applied to the transistor.



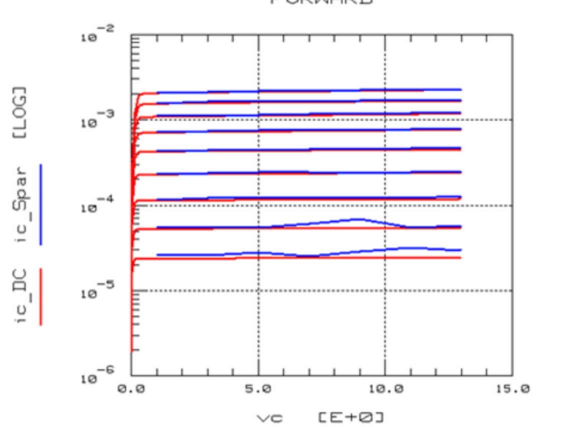
And here another example:



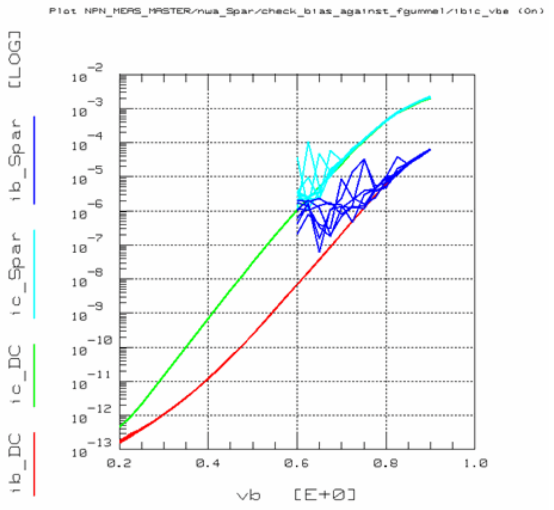
Here some more examples of too big RF signals, causing the transistor to behave nonlinear:
Too big RF signal affects the iB of a bipolar current, when comparing the S-parameter iB with the previously measured DC iB:



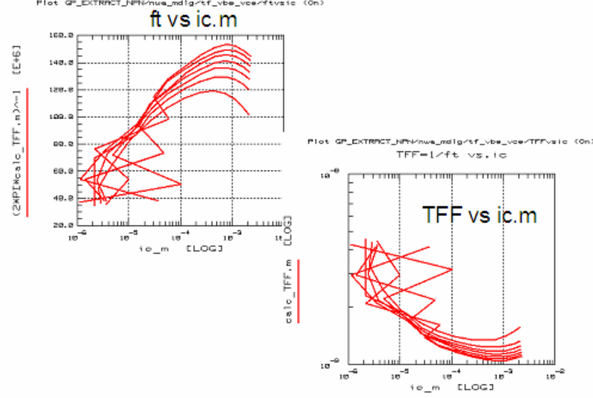
The same, now referring to iC:



Referring to the Gummel plot:



Last not least, a too big RF signal can disturb also the fT curve:



Note
By the way, self heating affects the slope 'after' the maximum fT.

Important Note on the Agilent N5250A 110GHz PNA Network Analyzer

In this network analyzer, the power in the 67-110 GHz range is not controlled by the PNA. It is provided directly by the 67-110 GHz Olesson test head. The power generated by the source in the head is constant and it is around 0dB.
Therefore, the only way to decrease this power is by using the attenuators.
In other words, the power settings in the PNA only controls the power up to 67 GHz. Therefore the trick of setting -30 dB and using the power slope to compensate for the losses between the PNA port and the head port, can only be used in the first band. For a sweep from 10 MHz ... 110 GHz, across the 67 GHz, it is recommended to set the source power in IC-CAP to 0dBm, and to use the attenuatiors on Port1 and Port2 for signal levels which are low enough for your transistor.

Considering the DC Bias Losses During S-Parameter Measurements

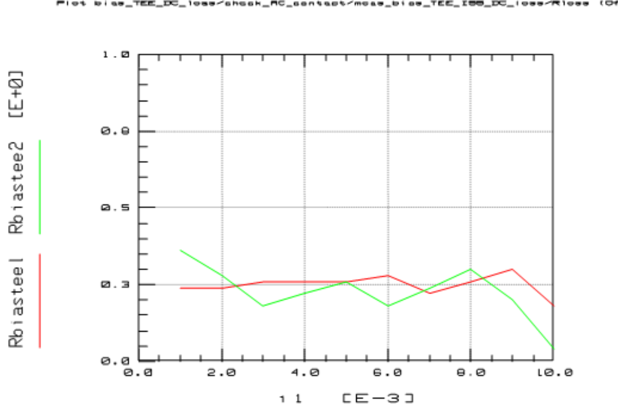
After the network analyzer has been properly calibrated, and this calibration has been verified, we finally have to take the DC bias losses into account, when measuring and then modeling the S-parameters of e.g. diodes or transistors.

These DC losses are **NOT** calibrated out by the VNA calibration !
Therefore, we perform 2 measurements:

- we place the GSG probes on the SHORT standard of the ISS calibration substrate and measure the ohmic loss.
- we perform the same measurement on the SHORT dummy on the wafer.

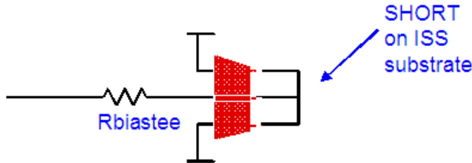
Hint: The following notes refer to the NPN_MEASURE_MASTER.mdl file of the IC-CAP demo_features directory:

When placing the GSG probes on the SHORT standard of the ISS substrate, and when stimulating a current into both probes, we can measure the voltage drop of the contacts. Such a measurement result is depicted below:



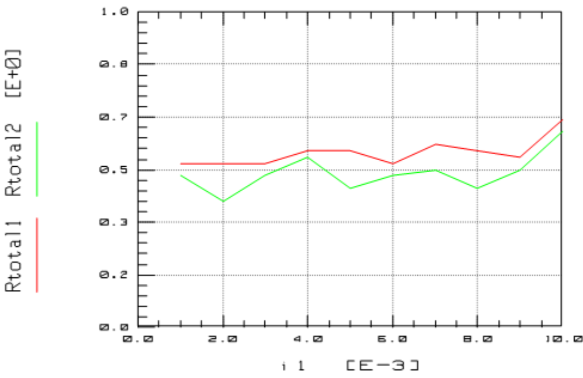
Note
The contact resistance to the gold on the ISS is usually different (much lower) compared to the contact resistance to the aluminum pads on the wafer, see the next measurement.

For the measurement results, we consider these resistors:

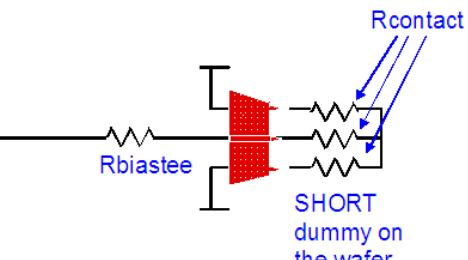


Assuming an ideal contact on the gold SHORT of the ISS substrate, we measure in this setup basically Rbiastee.
Then, we repeat these measurement conditions, but now for the SHORT on the wafer. This will show up as an additional resistor due to the worse GSG-aluminum contact compared to the previous GSG-gold contact.

> **Note**
> You can use the SHORT dummy on the wafer, but you can also place the GSG probes an any aluminum area on the wafer. It only needs to make sure we have a good aluminum short between the GSG probe pins.



For the interpretation of the aluminum SHORT measurements, we consider the following schematic:



Assuming 3x the same contact resistance for the GSG probe, we measure in this setup basically

Rtotal = Rbiastee + 1.5 * Rcontact.

Rbiastee is known from the previous ISS contact measurement. So we get for Rcontact:

Rcontact1 = (Rtotal1 - Rbiastee1) / 1.5
Rcontact2 = (Rtotal2 - Rbiastee2) / 1.5

How to take these contact and bias-TEE resistors into account for your S-Parameter modeling?

For your S-parameter measurements, the DC operating points are affected by these DC losses.
This loss is not taken into account by your VNA calibration. Therefore, the easiest way to perform a correct S-parameter modeling is to to simply take these DC losses into account during your simulations.

This means:
In your S-parameter DUTs, add a Test Circuit which reflects these DC loss resistors, and enter there
Rtotalport1 = Rbiastee1 + 1.5* Rcontact1
Rtotalport2 = Rbiastee2 + 1.5* Rcontact2

Make sure to add a big parallel capacitor to Rtotalport1 and Rtotalport2, in order to not affect the S-parameter simulations by this ohmic loss resistor (which affects the DC bias point, but **NOT** the S-parameters!!!).

For an example, see the Test Circuit in DUT NWA_Spar in
demo_features\3_MEAS_ORGANIZE_n_VERIFY_DATA\0_MASTER_FILES\NPN_MEAS_MASTER_demodata_PELdep.mdl
Here is the corresponding ADS syntax:

```
define emul_bias_TEE (P1   P2)
;The following network emulates the losses in the bias-TEEs
;and the wafer contact losses.
;Since the losses affect only the DC path, they are shorted
;for the AC signals by the capacitors C_short_x
;The values of R_serx are measured in Setup 'check_DC_contact_Spar'
R:R_ser1     P1     P11     R=1m
C:CAC1       P1     P11     C=.1
R:R_par1     P11      0     R=1MEG
R:R_ser2     P2     P22     R=1m
C:CAC2       P2     P22     C=.1
R:R_par2     P22      0     R=1MEG
;call the subcircuit description of the transistor,
;as defined in Tab 'Circuit' of the model file
VBIC_EXTRACT_MASTER:X1 P22 P11 0 0
end emul_bias_TEE
```

And this is the SPICE syntax:

```
.SUBCKT emul_bias_TEE  1=P1   2=P2
R_ser1  1 11 1.5
C_short1 1 11 .1
R_par1   1 0  1MEG
R_ser2  2 22 1.5
C_short2 2 22 .1
R_par2   2 0  1MEG
*call the subcircuit description of the transistor,
*as defined in Tab 'Circuit' of the model file
X1 22 11 0 0  VBIC_EXTRACT_MASTER
.ENDS
```

Entering Calkit Data for Connectorized Devices

When using and Agilent network analyzer with connector calkits, e.g. the 3.5mm calkit, the calkit data are already available in the instrument and need not to be entered manually. For the 26.5GHz 3.5mm calkit, the data are for an 'insertable device', i.e. one cable with a male connector, the other with a female connector (Tdelay_THRU = 0s) , are shown in the table below:

HP85052B 3.5mm calkit (26.5GHz) for an 'insertable device':

STANDARD No.	TYPE	C0 fF / L0 pH	C1/ E-27F/Hz / L1/E-24 H/Hz	C2 E-36F/Hz² / L2/E-33 H/Hz²	C3/ E-45F/Hz³ / L3/E-42 H/Hz³	Fixed or Sliding	OFFSET DELAY psec	OFFSET Z0 Ohm	OFFSET LOSS Ohm/s	FREQ MIN	FREQ MAX	STND LABEL
1	SHORT	2.076	-108.54	2.17	-0.01		31.78p	50	2.366G	0	999	SHORT
2	OPEN	49.4f	-310.13	23.17	159.7m		29.24p	50	2.2G	0	999	OPEN
3	LOAD					fixed	0	50	1.13G	0	999	LOAD
4	THRU						0	50	1.13G	0	999	THRU

Note
For non-insertable devices, see further below.

HP85033D 3.5mm calkit (6GHz) for HP8753 NWAs and again for an 'insertable device':

STANDARD No.	TYPE	C0 fF / L0 pH	C1/ E-27F/Hz / L1/E-24 H/Hz	C2 E-36F/Hz² / L2/E-33 H/Hz²	C3/ E-45F/Hz³ / L3/E-42 H/Hz³	Fixed or Sliding	OFFSET DELAY psec	OFFSET Z0 Ohm	OFFSET LOSS Ohm/s	FREQ MIN	FREQ MAX	STND LABEL
1	SHORT	0					31.8p	50	2.36G	0	999	SHORT
2	OPEN	49.4f	-310.13	23.17	160m		29.24p	50	2.2G	0	999	OPEN
3	LOAD					fixed	0	50	2.3G	0	999	LOAD
4	THRU						0	50	2.3G	0	999	THRU

Note
For non-insertable devices, see further below.

For the 2.4mm 85056A calkit (50GHz) and the 8510C network analyzer, the data are (neglecting the frequency-dependent C's and L's), again for an 'insertable device':

STANDARD No.	TYPE	C0 fF / L0 pH	C1/ E-27F/Hz / L1/E-24 H/Hz	C2 E-36F/Hz² / L2/E-33 H/Hz²	C3/ E-45F/Hz³ / L3/E-42 H/Hz³	Fixed or Sliding	OFFSET DELAY psec	OFFSET Z0 Ohm	OFFSET LOSS Ohm/s	FREQ MIN	FREQ MAX	STNDp LABEL
1	SHORT	2.16p					22.54p	50	0	0	999	SHORT
2	OPEN	29.7f	0	0	0		20.8p	50	0	0	999	OPEN
3	LOAD					broad	0	50	0	0	999	LOAD
4	THRU						0	50	0	0	999	THRU

Insertable and Noninsertable Devices

Note
The THRU cal standard specification with 0 ps is assuming a 'male-female' connection during the THRU measurements. This assumes, on the other hand, a male and a female connector at your DUT.

If your DUT exhibits connectors with the same sex, you need to use a 3.5mm THRU connector for the THRU calibration. Such a THRU connector may not be included in the calkit box. In this case, assume a delay of

TD=measured_ThruLength / C0 with C0 = 3E+8 m/s

for the delay of such a conventional THRU connector. In this case, however, you have to modify the calkit and enter the THRU delay time into your NWA calkit data!

A typical value for HP/Agilent 3.5mm THRU adapter (12 mm length 'w/o connectors') is TD = 95ps .

Below a list of modeled 3.5mm CalKit THRUs.

Delay Times of Agilent 3.5mm CalKit THRUs
(modeled by IC-CAP and ADS)

THRU male-male: 98.9p

THRU male-female: 104.2p

THRU male-male: 42.8p

If you instead have the two cable connectors with the same sex, and need to connect later to a male-female DUT, screw a THRU connector to one of your cables, and leave it connected also during the calibration.

HINT:
Even if you measure on-wafer, the separate connectorized calkit box might be interesting to buy as well, since it allows you to check the network analyzer performance separately.

IMPORTANT HINT:
Under www.vnahelp.com, you can find a wealth of information regarding NWA calibration for connectorized devices.

Entering Calkit Data for On-Wafer Probes

CASCADE MICROTECH

The following is a table of calkit standards definitions for Cascade 100um GSG probes.

Table 1: example of Standard definitions for Cascade G-S-G probes 100um pitch
(use your own proper Cascade standard definitions, referring to your pitch size)

STANDARD No.	TYPE	C0 fF / L0 pH	C1/ E-27F/Hz / L1/E-24 H/Hz	C2 E-36F/Hz² / L2/E-33 H/Hz²	C3/ E-45F/Hz³ / L3/E-42 H/Hz³	Fixed or Sliding	OFFSET DELAY psec	OFFSET Z0 Ohm	OFFSET LOSS Ohm/s	FREQ MIN	FREQ MAX	STND LABEL
1	OPEN	-9.3	0	0	0		0	50	0	0	999	OPEN
2	SHORT	2.4					0	50	0	0	999	SHORT
3	LOAD					fixed	-0,007	500	0	0	999	LOAD
4	THRU						1	50	0	0	999	THRU

Note:NWA standards 1-4 are for SOLT calibration
standards 8-11 are for LRM calibration

From: Cascade Microtech, Microwave Wafer Probe Calibration Constants, HP8510 Network Analyzer Input, Instruction Manual, 1990
See also: Cascade Microtech Application Note: On Wafer Vector Network Analyzer Calibration and Measurements

For the other Cascade ACP (air coplanar probe) G-S-G probes the calkit standards values are:
CalKit constants for Cascade G-S-G probes:

Pitch	C-Open	L-Short	L-Termination
100	-9.3fF	2.4pH	-3.5pH
125	-9.5fF	3.6pH	-2.6pH
150	-9.7fF	4.8pH	-1.7pH

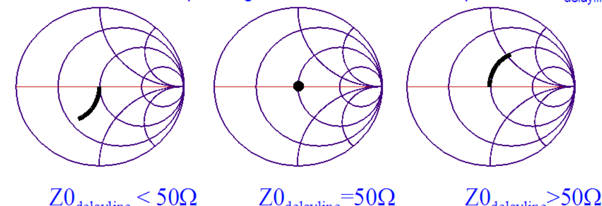
Independent of the pitch, the THRU delay time is always 1ps.
This is the table for the assignment of the cal kit classes:
Table 2:standard class assignment, for all kinds of probes

S11A	1	OPEN
S11B	2	SHORT
S11C	3	LOAD
S22A	1	OPEN
S22B	2	SHORT
S22C	3	LOAD
FORWARD TRANSMISSION	4	THRU
REVERSE TRANSMISSION	4	THRU
FORWARD MATCH	4	THRU
REVERSE MATCH	4	THRU
FORWARD ISOLATION	3	ISOLATION
REVERSE ISOLATION	3	ISOLATION

Note on the 500Ω entry for Z0 of the Cascade LOAD:
or:What to do if I the NWA does not allow to enter a calkit value:

The Cascade LOAD represents a precisely trimmed 50Ω DC characteristics, plus an inevitable inductance. In order to reflect this inductive performance at higher frequencies, it would be necessary to enter the value of that inductance into the NWA calkit. However, most NWAs do not allow to enter an inductance. However, they allow to enter the characteristic impedance Z0 plus the delay time of a lossless delay line. Now, it is important to know that the S11 curve of a delay line of exactly Z0delayline = 50Ω is completely in the center of the Smith chart. Delay lines with a characteristic impedances Z0delayline above 50Ω look like an inductor in series with a resistor of 50Ω, and delay line impedances below Z0delayline = 50Ω look like a capacitance in parallel with a resistor of 50Ω. See the sketches below:

S11 curves of lossless delay lines
depending on their characteristic impedance $Z0_{delayline}$



Therefore, since the Cascade LOAD behaves like a precisely trimmed resistor of 50Ω in series with an inductor, a delay line with Z0delayline = 500Ω (what is the max. entry allowed for Agilent's NWAs) are entered plus a delay time. This delay time value is obtained from the lumped inductance by the following conversion formula:

T_delay = L / Z0delayline = L / 500
On the other hand, the Picoprobe LOAD behaves like a resistor of 50Ω in parallel with a capacitor. Therefore, we have to enter a delay line impedance of Z0delayline = 5Ω (what corresponds to the min. entry allowed for Agilent's NWAs) and a corresponding delay time! See the specification sheet of the Picoprobe GSG probes.

> **Note**
> If The Network Analyzer does not allow L0 FOR THE SHORT, but only a delay, apply the above formula as well. Use Z0delayline = 500Ω for both, Cascade and Picoprobe (since we want to convert an inductance to a delay line, see the notes on a delay line with Z0delayline > 50Ω). For example, to replace the L_short = 2.4pH , a 4.8fs delay and Z0delayline = 500Ω are entered into the NWA calkit table instead.

IF THE NETWORK ANALYZER DOES NOT ALLOW TO ENTER A NEGATIVE DELAY TIME: In this case, the negative series inductance has to be converted into a parallel capacitance:For a load resistance R0 in series with an impedance LSer < 0, there is

$$Z = R0 + j \cdot \omega \cdot L_{Ser} \quad ....(1)$$

$$Z = \left(R0^{-1} + j \cdot \omega \cdot C_{Par} \right)^{-1}$$
$$\approx R0 \cdot (1 - j \cdot \omega \cdot C_{Par}) \qquad \text{for } \omega \cdot R0 \cdot C_{Par} << 1$$

Setting (1)=(2) gives

$$C_{par} = -\frac{L_{Ser}}{R0^2}$$

what finally gives a modified conversion formula

$$T_{delay} = -\frac{L_{Ser} \cdot Z0}{R0^2}$$

Example:
R0 = 50Ω, Z0 = 5Ω, Lser = -3.5pH, and so we enter as calkit data, following the hints on delay lines with Z0 < 50Ω (see above):

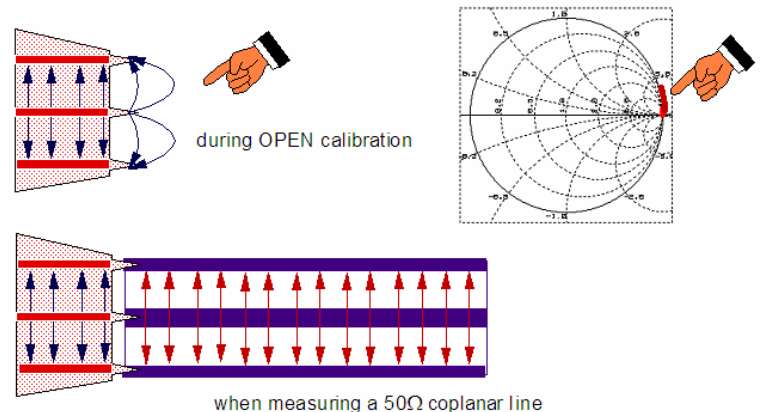
$$T_{delay} = -\frac{-3.5pH \cdot 5\Omega}{2500\Omega^2} = 7f$$
$$Z0 = 5\,\Omega$$

and

$Z0 = 5\,\Omega$

Acknowledgements: thanks to Mr. Meiser of Infineon/Munich for this hint.

Note on the negative capacitance of Cascade probes for an OPEN
The goal of a calibration is to ensure later a correct measurement of the DUT without any parasitics caused by the measurement cables etc. If the DUT was a 50Ω coplanar line, it is expected that the power from port 1 is fed from the probe tip completely into this DUT. Now, an OPEN Cascade probe is also a 50Ω line, yet cut off at the end of the probe tips. When measuring this lift-off probe, the electro-magnetic field at the probe end is distorted and no longer the homogeneous one of a 50Ω line. In other words, it will ~~of course~~ not correspond to a precisely cut-off homogeneous field. It will rather extend a bit over the OPEN end into the air. On the other hand, if the probe is later connected to that 50Ω coplanar line, it is obvious that there should be no discontinuity of the electro-magnetic field any more, see the following figure. Therefore, the distortion, shown with the OPEN probe tip, must correspond to a nonideality which is **not** an ideal infinite Ohm representation.



Note
Another explanation for a negative capacitance with probes in the air is that the wave propagation in air is faster than on the wafer.

For more details, see the publications of Cascade Microtech mentioned at the end of this chapter.

PICOPROBE

Standard definitions for Picoprobe GSG-150 probes

STANDARD No.	TYPE	C0 fF / L0 pH	C1/ E-27F/Hz / L1/E-24 H/Hz	C2 E-36F/Hz² / L2/E-33 H/Hz²	C3/ E-45F/Hz³ / L3/E-42 H/Hz³	Fixed or Sliding	OFFSET DELAY psec	OFFSET Z0 Ohm	OFFSET LOSS Ohm/s	FREQ MIN	STND MAX	LABEL
1	OPEN	6.5	0	0	0		0	50	0	0	999	OPEN
2	SHORT	8.8					0	50	0	0	999	SHORT
3	LOAD					fixed	0,008	5	0	0	999	LOAD
4	THRU						1.13	50	0	0	999	THRU

CalKit constants for Picoprobe G-S-G probes:

Pitch	C-Open	L-Short	LOAD Delay	.
75	6.5fF	3.9pH	19.7fs	and Z0=5Ω
100	6.5fF	5.0pH	15.5fs and Z0=5Ω or 10fs/500Ω	
125	6.5fF	7.1pH	11.5fs and Z0=5Ω	
150	6.5fF	8.8pH	8.0fs and Z0=5Ω or 17.6fs/500Ω	

Independent of the pitch, the THRU delay time is always 1.13ps, because the Picoprobe cal substrate THRU is 200um long, probe contact distance from each end is 25um each, so the effective THRU is 150um. Corresponding to the dielectric constant of the Picoprobe cal substrate, the corresponding delay time is then 1.13ps.

Note
For an explanation of the Z0=5Ω of the LOAD, please see above.

CalKit constants for Picoprobe G-S probes (to be used with the CS-8 Cal.Substrate)

Pitch	C-Open	L-Short	LOAD Delay
100	4.3fF	9.2pH or 18fs/500Ω	0

Independent of the pitch, the THRU delay time is always 1.13ps

Note
The next slides are from Cascade Microtech, so we have decided not to upload them.

Entering Calkit Data for Packaged Devices

Intercontinental Microwave (packaged devices) Test Fixture

Standard definitions for the ICM TestFixture for a SOT538A package.

STANDARD No.	TYPE	C0 fF / L0 pH	C1/ E-27F/Hz / L1/E-24 H/Hz	C2 E-36F/Hz^2 / L2/E-33 H/Hz^2	C3/ E-45F/Hz^3 / L3/E-42 H/Hz^3	Fixed or Sliding	OFFSET DELAY psec	OFFSET Z0 Ohm	OFFSET LOSS Ohm/s	FREQ GHz MIN	FREQ GHz MAX	STND LABEL
1	THRU					fixed	0	50	0	0.395	8.01	THRU
2	LINE1					fixed	145	50	0	0.395	2.51	LINE1
3	LINE2					fixed	50	50	0	2.49	8.01	LINE2
4	SHORT					fixed	0	50	0	0.395	8.01	SHORT

Standard Class Assignment

S11A	**OPEN**	
S11B	SHORT	
S11C	LOAD	
S22A	OPEN	
S22B	SHORT	
S22C	LOAD	
FORWARD TRANSMISSION	THRU	
REVERSE TRANSMISSION	THRU	
FORWARD MATCH	THRU	
REVERSE MATCH	THRU	
FORWARD ISOLATION	ISOLATION	
REVERSE ISOLATION	ISOLATION	
FREQUENCY RESPONSE		
TRL THRU	1	
TRL REFLECT	4	
TRL LINE	2	3

Entering CalKit Data into the Network Analyzer

General Info

About Calibration Kits (a copy of the Agilent PNA Help pages)

A calibration kit is a set of physical devices called standards. Each standard has a precisely known or predictable magnitude and phase response as a function of frequency.
In order for the analyzer to use the standards of a calibration kit, the response of each standard must be mathematically defined and then organized into standard classes that correspond to the error models used by the analyzer.
To be able to use a particular calibration kit, the known characteristics from each standard in the kit must be entered into analyzer memory or recalled from a default list of calibration kits stored in the analyzer.

Calibration Standards

Calibration standards provide the reference for error-corrected measurements in the network analyzer. Each standard has a precisely known definition that includes electrical delay, impedance, and loss. The analyzer stores these definitions and uses them to calculate error correction terms.
During measurement calibration, the analyzer measures standards and mathematically compares the results with "ideal models" of those standards. The differences are separated into error terms that are later removed from device measurements during error correction.

Standard Type

A standard type is one of four basic types that define the form or structure of the model to be used with that standard. The four standard types are shown below:

Standard	Terminal Impedance
SHORT	zero ohms
OPEN	infinite ohms
LOAD	system impedance, Z0
THRU/LINE	no terminal impedance
ARBITRARY	user-defined

Standard Definitions

Standard definitions describe the electrical characteristics of the standards and the frequencies they will be used. Refer to the "standards definition table" included with the calibration kit. Standard definitions include:

- Minimum Frequency Specifies the minimum frequency the standard is used for calibration.
- Maximum Frequency Specifies the maximum frequency the standard is used for calibration.
- Z0 Specifies the characteristic impedance of the standard (not the system characteristic impedance or the terminal impedance of the standard).
- Delay Specifies a uniform length of transmission line between the standard being defined and the actual calibration plane.
- Type Specifies type of standard (SHORT, OPEN, THRU/LINE, LOAD, ARBITRARY).
- Loss Specifies energy loss, due to skin effect, along a one-way length of coaxial cable.

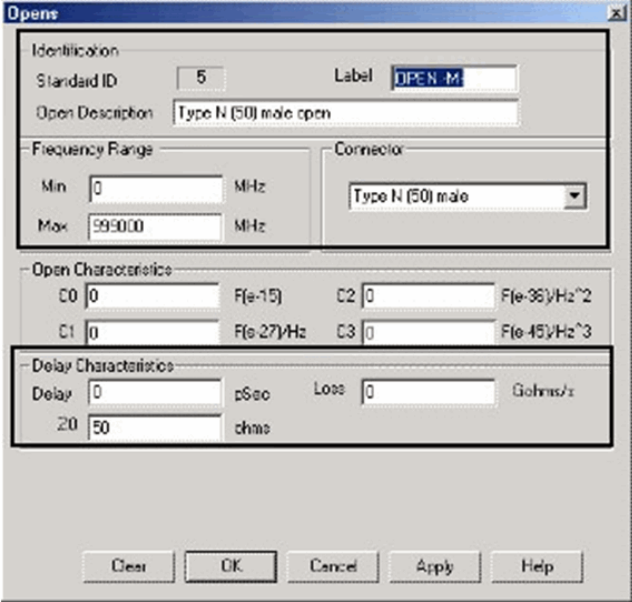
Loss Model Equation

- The value of loss is entered as ohms/second at 1 GHz.
- To compute the loss of the standard, measure the delay in seconds and the loss in dB at 1 GHz. Then use the following formula:

Capacitance Model Equation

C0, C1, C2, C3. Specifies the fringing capacitance for the open standard.

- $C = (C0) + (C1 \times F) + (C2 \times F^2) + (C3 \times F^3)$
- (F is the measurement frequency).
- The terms in the equation are defined when specifying the open as follows:
  - C0 term is the constant term of the third-order polynomial and is expressed in Farads.
  - C1 term is expressed in F/Hz (Farads/Hz).
  - C2 term is expressed in F/Hz^2.
  - C3 term is expressed in F/Hz^3.



Inductance model equation:

L0, L1, L2, L3. Specifies the residual inductance for the short standard.

- $L = (L0) + (L1 \times F) + (L2 \times F^2) + (L3 \times F^3)$
- (F is the measurement frequency).
- The terms in the equation are defined when specifying the short as follows:
  - L0 term is the constant term of the third-order polynomial and is expressed in Henries.
  - L1 term is expressed in H/Hz (Henries/Hz)
  - L2 term is expressed in H/Hz^2.

- L3 term is expressed in H/Hz^3.

Class Assignments

Once a standard is characterized, it must be assigned to a standard "class". A standard class is a group of standards that are organized according to the calibration of the PNA error model.
The number of classes needed for a particular calibration type is equal to the number of error terms being corrected.
A class often consists of a single standard, but may be composed of multiple standards, such as loads or delay lines. Refer to the calibration kit "class assignment" table.
Example: A response calibration requires only one class, and the standards for that class may include an OPEN, or SHORT, or THRU. A 1-port calibration requires three classes. A full 2-port requires 10 classes, not including two for isolation.
The number of standards assigned to a given class may vary from one to seven for unguided calibrations. Guided calibrations allow as many standards as needed.

The different classes used in the PNA:

S11A, S11B, S11C (S22A, S22B, S22C and so forth)
These are the three classes for port 1-reflection calibrations (three classes also for S22 and S33). They are used in the one-port calibrations and the full two-port calibration. They are required in removing the directivity, source match, and reflection tracking errors. Typically, these classes might consist of an open, a short and a load standard for each port.
Transmission and Match (forward and reverse)
These classes are used to perform a full two-port calibration. The transmission class relates primarily to the transmission tracking, while the match class refers to load match. For both of these classes, the typical standard is a thru or delay.

Isolation

The isolation classes are used to perform a full two-port and the TRL two-port calibrations. The isolation classes apply to the forward and reverse crosstalk terms in the PNA error model.

TRL thru

These are used to perform a TRL two-port calibration. The TRL thru class should contain a thru standard or a short line. If it contains a non-zero length thru standard, then the calibration type is called LRL or LRM.

TRL reflect

This class is used to perform a TRL two-port calibration. The TRL reflect class should contain a standard with a high reflection coefficient, typically an open or preferably short. The actual reflection coefficient need not be known, but its phase angle should be specified approximately correctly (± 90 deg).
The exact same reflection standard must be used on both ports in the TRL calibration process.

TRL line (or match)

These are used to perform a TRL two-port calibration. The TRL line or match class should contain line standards, load standards, or both. If a line standard is used, its phase shift must differ from that of the TRL thru standard by 20° to 160°. This limits the useable frequency range to about 8 to 1. Two or more line standards of different lengths may be specified to get broader frequency coverage. It is also common to include a load standard for covering low frequencies, where the line's length would be impractically long. When a load is used, the calibration type is called TRM or LRM.

Note
For more information, read application note 8510-5A, "Specifying Calibration Standards for the Agilent 8510 Network Analyzer". Although the application note is written for the Agilent 8510 series of network analyzers, it applies to the PNA as well. The part number of the publication is 5956-4352.

Entering SOLT Calkit Data (Agilent PNA-NetworkAnalyzer-specific)

Procedure for the older PNA Firmware:

1. Use the mouse, go to the 'Calibration' menu, click 'Advanced Modify Cal Kit'.
2. Select 'Edit Version 1 Kits'.
3. In the new window, select the desired calibration kit ID number, e.g. 7.

   Note
   To restore the selected calibration kit to its predefined values (ID numbers 1-6), click Restore Default Kit.
4. In the 'Cal Kit Name' box, enter your calibration kit name for the new calkit.

Assign Standards to Ports

1. Click the 'Port 1 Label' box and enter your port 1 name, e.g. probe1.

2. In the 'Port 2 Label' box, enter your port 2 name, e.g. enter probe2.
3. In the section 'Choose Standard Nos. for Cal Classes, enter a name for the Cal class and make the selection for the Standard number:

Cal.Class	Cal.Class Label	Standard No.
S11A	'OPEN'	= 2
S11B	'SHORT'	= 1
S11C	'LOAD'	= 3
S21T	'THRU'	= 4
S22A	'OPEN'	= 2
S22B	'SHORT'	= 1
S22C	'LOAD'	= 3
S21T	'THRU'	= 4

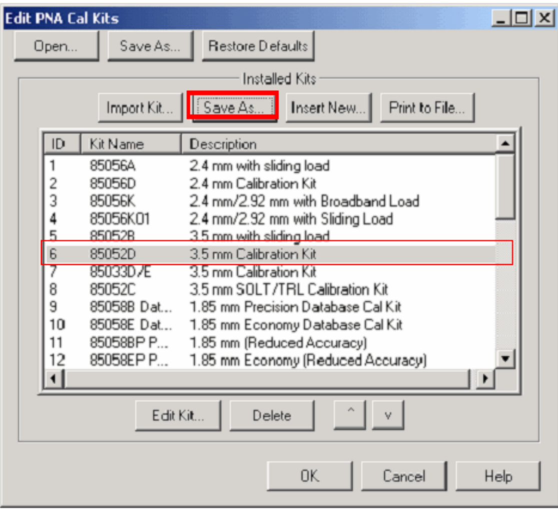
Modify Standard Definition

4. In the Standard No. box in the 'Modify Standard Definition' portion of the dialog box, click the arrow button to select the desired standard number.
5. For the standard number selected in the previous step (during Assign Standards to Ports), enter a label (e.g. Standard #1 is 'SHORT', #2 is 'OPEN', #3 is 'LOAD', #4 is 'THRU'.
6. Then, enter all the definitions of the standard (for example, Minimum Frequency, Maximum Frequency, Z0, Delay, Loss, etc). Refer to the calkit data of your probes, as given ins the tables further below.
7. Repeat steps 1 to 3 for each standard number.
8. Click OK.
   You are back in the starting menu item.
   Click OK.

Procedure for the new PNA Firmware (PNA Type 'B'):

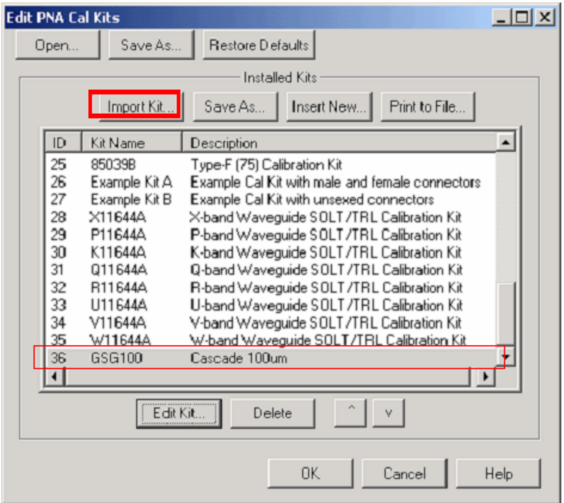
On the PNA, select Calibration/AdvancedModifyCalKit:

Step 1: copy-paste an existing calkit



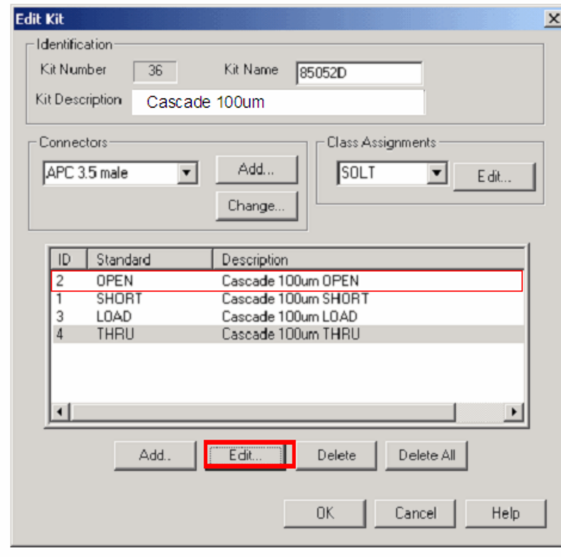
To copy-paste, select an existing calkit (e.g. the 3.5mm one), save it, and then hit 'Import Kit'

Step 2: rename it



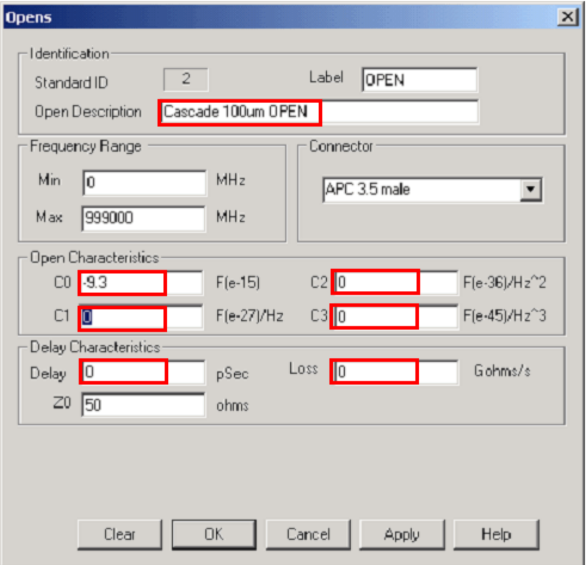
To rename, hit 'Edit Kit'

Step 3: specify the 4 calkit standards



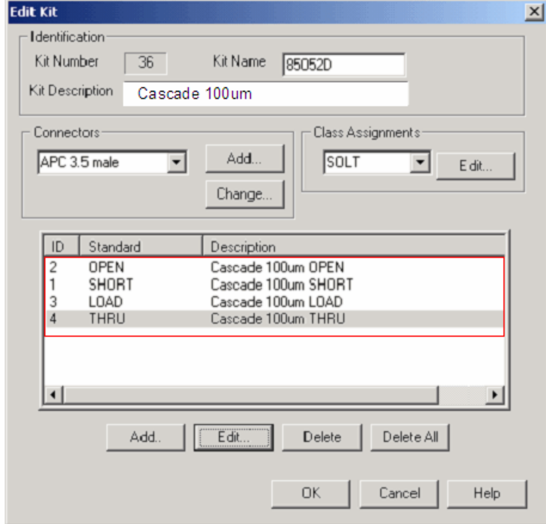
NOTE:
keep the ID
numbers fo
the individu
standards a
shown here
2=OPEN
1=SHORT
3=LOAD
4=THRU

To open the individual calkit specifications, hit 'Edit'

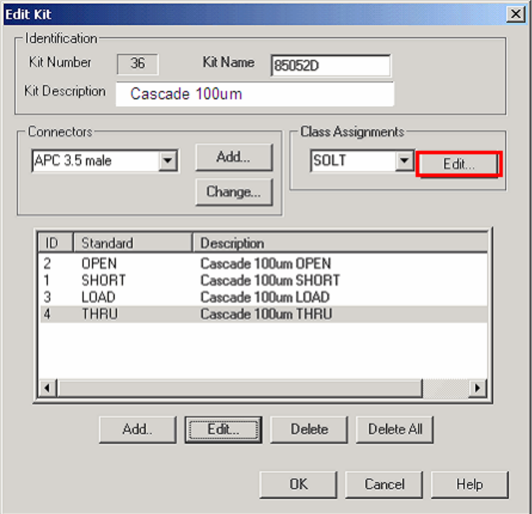


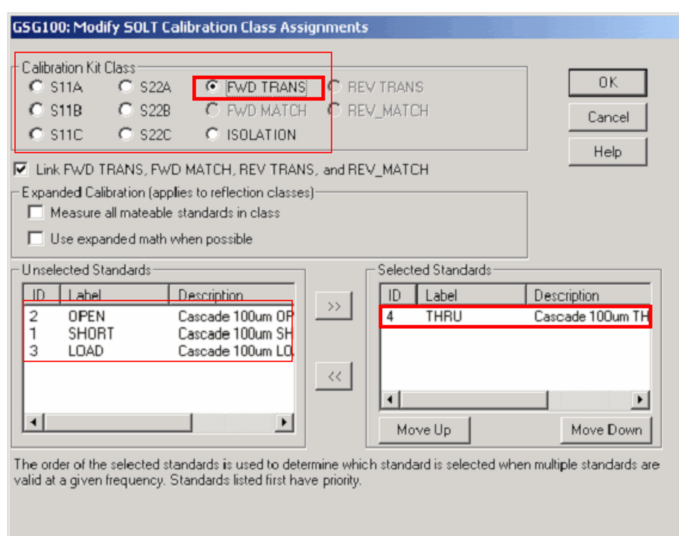
Enter the calkit data for the OPEN, SHORT LOAD and THRU

Step 4: the final result after entering the calkit data



Step 5: check the class assignments





For SOLT calibration:
S11A=2 OPEN
S11B=1 SHORT
S11C=3 LOAD
S22A=2 OPEN
S22B=1 SHORT
S22C=3 LOAD
FWD TRANS = 4 THRU
ISOLATION = 3 LOAD

NOTE: the Class Assignments associate the entered calkit data (OPEN, LOAD, SHORT, THRU) with the fixed arrays (classes) of the SOLT (or other) calibration.

Agilent Technologies

Step 6: save the modified calkit

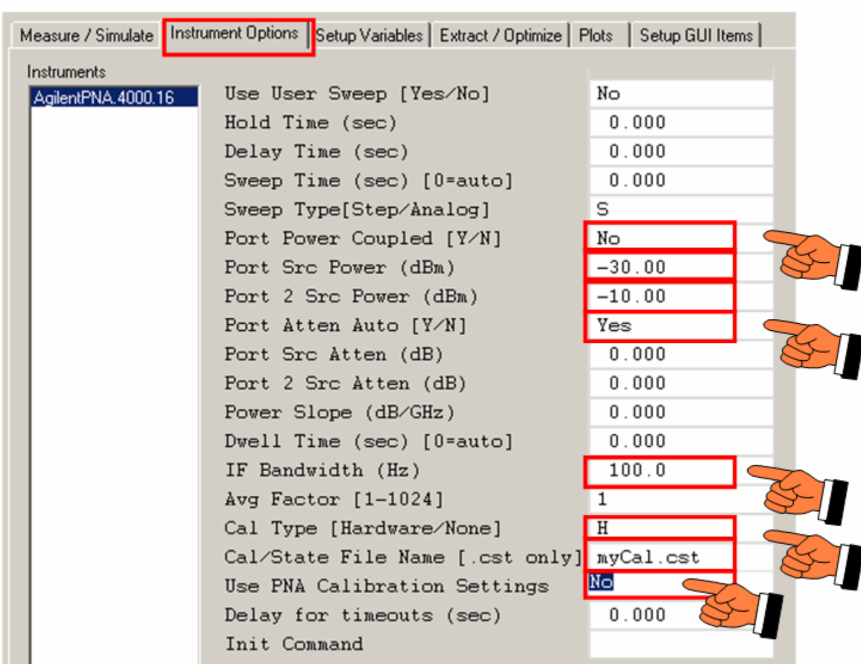
After this, we are back in IC-CAP.

We have to consider two cases:

- A- In Instrument Options, 'Use PNA Calibration Settings'= No
- B - 'Use PNA Calibration Settings'= Yes (This is recommended).

CASE (A): In Instrument Options, 'Use PNA Calibration Settings'= No

When 'Use PNA Calibration Settings'= No



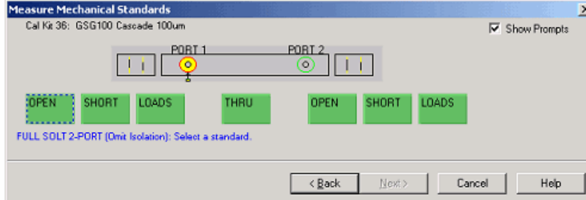
To perform the PNA calibration from IC-CAP, load the file demo_features\3_MEAS_ORGANIZE_n_VERIFY_DATA\0_MASTER_FILES\2_NWA_CAL_VERIFICATION\CAL_VERIFY_MASTER_PELdep.mdl
Begin with Setup 'check_max_RF':

- Connect the transistor
- Check the DC output characteristics sweep conditions
- On the PNA, set the port attenuations to the highest dB values and the PNA signal level to the lowest possible value.
- Set the PNA to continuous measurement mode

- In IC-CAP, hit measure and check the plots.
  You should get an ordinary output characteristics. No distortions.
  Convert these measured data to reference data (in the Outputs of the Setup, change from Type 'B' to 'M', hit enter, and change from 'M' to 'B'. This means that there are now pseudo-simulation data available which represent the output characteristics with no RF distortion (small signal RF excitation condition).
  Then, on the PNA, manually increase the power level (if necessary, adjust the attenuator settings), so that you get as starting RF conditions e.g. -30dBm at port1 and -10dBm at port2.
  For transistors, you should select power-decoupled ports, i.e. you specify for each port an individual power level.
- Make sure the PNA is still in continuous measurement mode (remember, the PNA is **NOT** controlled by IC-CAP at this stage !!)
- Re-perform the DC output measurement with IC-CAP.
  If the power levels are still ok, the new (red) curve is identical to the previous (yellow) reference curve. If there is a distortion visible, namely for low currents, the applied power levels are already too big for the transistor. Repeat this step with reduced PNA power levels.
  If there is no distortion, you may try to increase the power levels.
  After that, proceed in IC-CAP with Setup 'OPEN'.
- Check the frequency range
- In 'Instrument Options', specify
  - Port Power Coupled: No
  - Port Atten Auto: Yes
  - Cal Type: H

- and enter the power levels for each port as evaluated before.
- For 'Cal/State File Name' enter e.g. myCal_1.cst (this will be the calset filename stored after the manual PNA calibration)

Step 7: Calibrate the PNA for that calkit
-> Initialize the calibration in IC-CAP
-> Select the Calibration wizzard, select the previously specified calkit and you end up with the window depicted below:



The next slide explains the individual calibration initialization steps in IC-CAP

Hit 'Calibrate' in IC-CAP.

IC-CAP displays a window stating that you should now manually calibrate the PNA.
DO NOT CLOSE THIS IC-CAP WINDOW UNTIL YOU ARE FINISHED WITH THE CALIBRATION AND HAVE SAVED THE RESULT IN THE PNA myCal_1.cst FILE

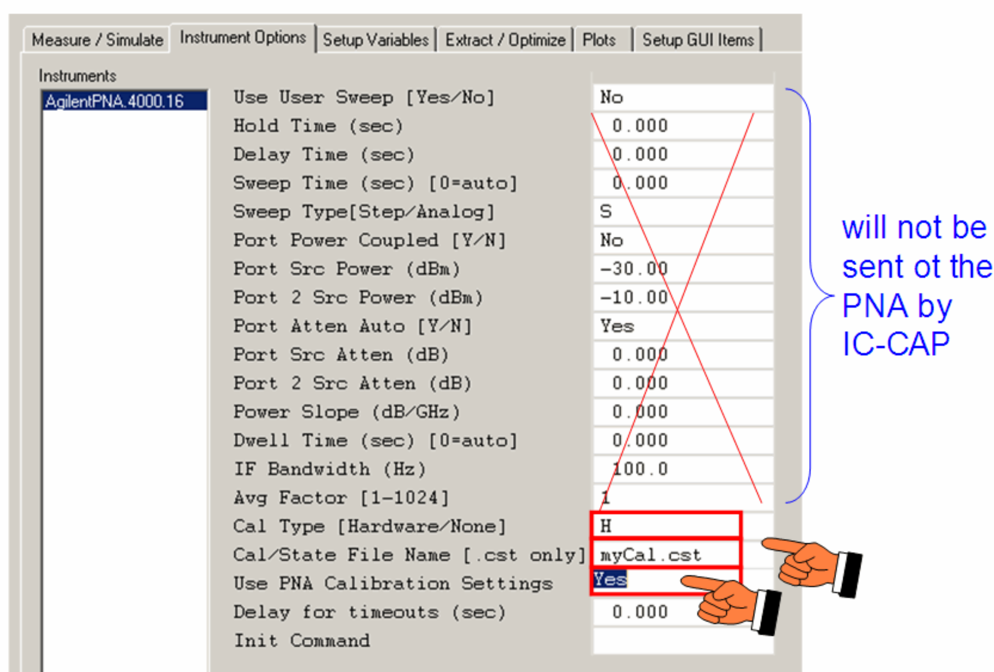
On the PNA, select the 'Calibration Wizzard' menu, and work through until you see the above depicted calibration window.

- connect the cal. standard on the ISS substrate and click on the corresponding green button in the PNA cal window
  Finally, save the calibration **manually** as calset myCal_1.cst in the default directory of the PNA:
- On the PNA, select File/SaveAs, and select the file type *.cst),
- Save the .cst file on the PNA-computer under
  'C:\Program Files\Agilent\Network Analyzer\Documents '.
  This is the directory IC-CAP will ask the PNA to look for the correction data when executing 'Measure' later.

Final Note: make sure that for all subsequent IC-CAP Setups applying this PNA calibration, the Input 'freq' must be identical as well as the 'Instrument Options' !!!
I.e.: copy-paste the 'Instrument Options' to all other PNA Setups in your IC-CAP ModelFile !

Note
On the Agilent PNA instrument mainframe, the default directory for saving and reading calibration (.cal) and state/cal (.cal) files is:
C:\Program Files\Agilent\Network Analyzer\Documents
Calibration files can be saved and read in a different directory by setting the System Variable PNA_CAL_PATH_FILE to another directory.
Example: PNA_CAL_PATH_FILE=C:\my_dir

CASE (B): (RECOMMENDED) In Instrument Options, 'Use PNA Calibration Settings'= Yes



In this case, you perform the PNA calibration without any initialization from IC-CAP !
I.e. applying WinCal, SuessCal or the PNA Cal Menu
After that, on the PNA save the calset manually to the CalFile Name specified later in IC-CAP.

- On the PNA, select File/SaveAs, and select the file type *.cst).
- Save the .cst file to the default directory (IC-CAP will search the file there !).

After that, you only need to set the frequency points in the IC-CAP Setup accordingly to what you had used before on the PNA. And as mentioned just before the InstrumentOptions entry 'Cal/State File Name' must match the .cst file name used before when we had saved th PNA calibration.

Note
Make sure that for all subsequent IC-CAP Setups applying this PNA calibration, the Input 'freq' must be identical as well as the 'Instrument Options', that is, copy-paste the 'Instrument Options' to all other PNA Setups in your IC-CAP ModelFile

Entering SOLT Calkit Data HP NetworkAnalyzer specific

STEP BY STEP PROCEDURE FOR HP8510 and 87xx:

Note
using Picoprobes, you may simply use the provided floppy disc and follow the instructions given in the Picoprobe Calibration Substrate CS-5 Manual.

We start with a calkit list as below.

STANDARD		C0	C1	C2	C3	fixed or sldng	Offset			Freq		Labl
Nr	Type	fF	E-27F/Hz	E-36F/Hz^2	E-45F/Hz^3		Delay psec	Z0 Ω	Loss Ω/s			
		L0 pH	L1 E-24H/Hz	L2 E-33H/Hz^2	L3 E-42H/Hz^3					Min Hz	Max GHz	
1	Open	49.4	-310.1	23.17	-0.159		29.24	50	2.2	0	999	Open
2	Short	2.07	-108.5	2.17	-0.01		31.78	50	2.2	0	999	Short
3	Load					fixed	0	50	0	0	999	Load
4	Thru						97	50	0	0	999	Thru

The calkit entering procedure below refers to the calkit table above. If in your own calkit table the OPEN and SHORT entries are flipped, i.e. Standard Nr.1 = SHORT, Standard Nr.2 = OPEN, either exchange these lines in your calkit table and proceed as described below, or keep that sequence in your calkit table and change the corresponding lines in the procedure below.

Changing calibration constants (modifying a calkit) for the HP 8510

(very similar for Agilent 87xx family):
see also HP8510-5A product note: Specifying Calibration Standards for the HP8510 Network Analyzer

Enter the calkit definitions

On the 8510 network analyzer, press hardkey CAL, the select softkey MORE, MODIFY1 (or 2),
Note: on the 8753E, press hardkey CAL, then select softkey CAL KIT, MODIFY,
In a first step, we define the calibration standards.
Press DEFINE STANDARD
Press 1 and then X1, to enter the 1st line of the table below, i.e. the OPEN calibration standard data
VERY IMPORTANT NOTE: the softkey OPEN has to be underlined now!
If not, repeat the above step: i.e. hit the OPEN softkey, and enter 1 X1,until the underlined softkey and the entered table row match !!
NOTE: if your table's entries have a different order for the OPEN, SHORT, LOAD and THRU,
e.g. SHORT-OPEN-LOAD-THRU, apply this method correspondingly!
Select OPEN
Hit, CO and enter the new C value. Do the same for C1, C2 and C3
Hit SPECIFY OFFSET, and enter the new value.
Do the same with softkeys OFFSET DELAY, OFFSET LOSS, OFFSET Z0,
MINIMUM FREQUENCY, MAXIMUM FREQUENCY
After that, finish this submenu by selecting STD OFFSET DONE
Select LABEL STD (OPEN is displayed in the upper left display corner), and enter the standard
name as given again in the table below, then select TITLE DONE to finish this sub-menu
Select STD DONE (DEFINED)

Standard no.1 has now been defined. To define the remaining standards, refer to the calkit table
above and repeat the steps above from DEFINE STANDARD to STD DONE
-> for the SHORT (press 2 and X1 and select SHORT etc),
-> for the LOAD (press 3 and X1 and select LOAD etc.)
-> and for the THRU (press 4 and X1 and select THRU etc.)

The next, i.e. the 2nd step for entering the calkit data, is the entering of the class assignments (in this step we associate the location of the correction factors in the NWA corr.vectors with the table entries from above)
Proceed as follows:

Select SPECIFY CLASS,
hit S11A, enter 1, x1 (to inform the NWA to use standard no.1 for the S11A class of calibration)
hit S11B, enter 2, x1
hit S11C, enter 3, x1
hit S22A, enter 1, x1
hit S22B, enter 2, x1
hit S22C, enter 3, x1
hit MORE,
hit FWD.TRANS, enter 4, x1
hit REV.TRANS, enter 4, x1
hit FWD.MATCH, enter 4, x1
hit REV.MATCH, enter 4, x1
hit RESPONSE, enter 4, x1
hit FWD.ISOL'N, enter 3, x1
hit REV.ISOL'N, enter 3, x1
Select SPECIFY CLASS DONE

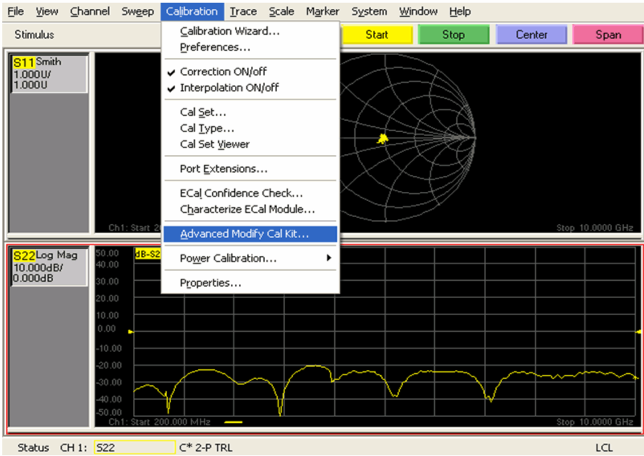
The 3rd step of our calkit entering procedure is to label the class
Select LABEL CLASS,
hit S11A, enter the title 'OPEN', press LABEL DONE
hit S11B, enter 'SHORT' (use the wheel to select the characters!!!)
hit S11C, enter 'LOAD'
hit S22A, enter 'OPEN'
hit S22B, enter 'SHORT'
hit S22C, enter 'LOAD'
hit MORE
hit FWD.TRANS, enter 'THRU'
hit REV.TRANS, enter 'THRU'

hit FWD.MATCH, enter 'THRU'
hit REV.MATCH, enter 'THRU'
hit FWD.ISOL'N, enter 'ISOLATION'
hit REV.ISOL'N, enter 'ISOLATION'
Select LABEL CLASS DONE

Finally, we reach step 4 of our procedure: the labelling of the calkit:
Select LABEL KIT,
hit ERASE TITLE, and enter the new title,
press TITLE DONE,
press KIT DONE (MODIFIED)

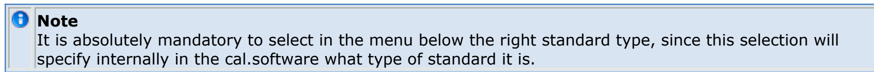
The calkit is now ready to be stored: Hit 'Save'.

Entering TRL Calkit Data Agilent NetworkAnalyzer specific

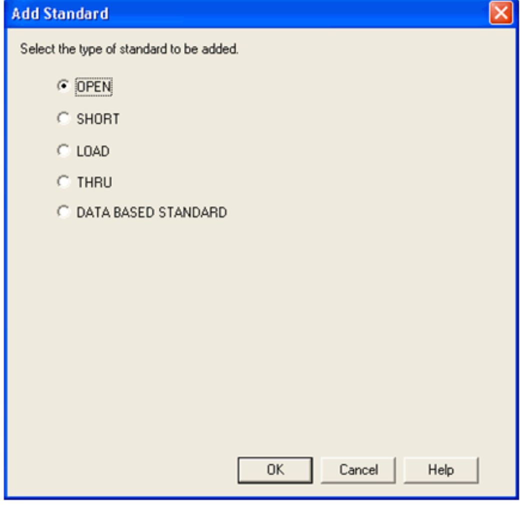


Select the corresponding Standard Type:

- THRU
- REFLECT: use a SHORT (recommended, alternatively use an OPEN)
- LINES (in the Menu depicted below, select THRU for entering the different TRL lines, and then change the name to LINE in the up-popping next menu.)

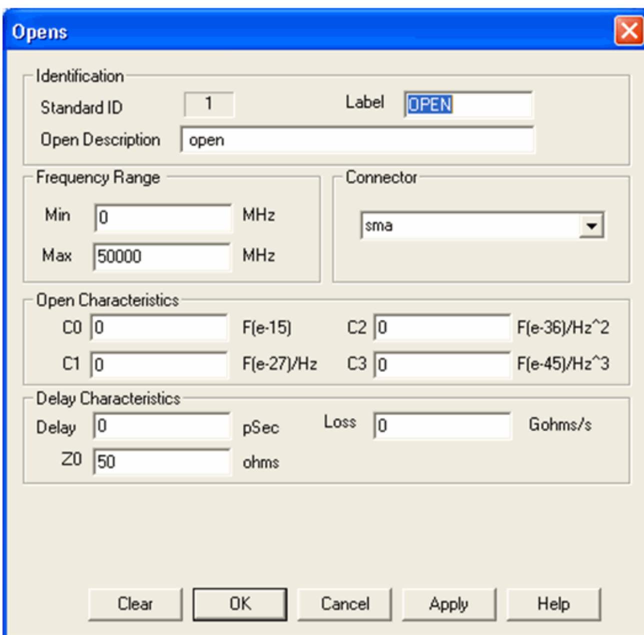


Note
It is absolutely mandatory to select in the menu below the right standard type, since this selection will specify internally in the cal.software what type of standard it is.

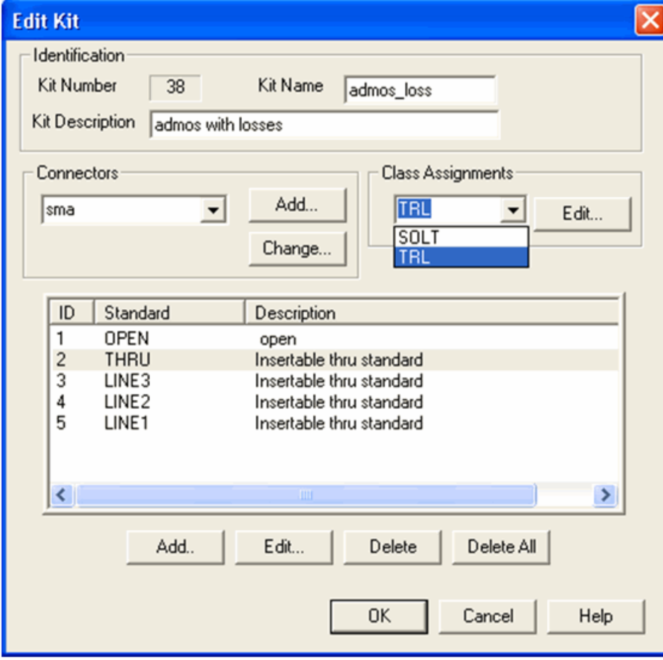


Fill in the calkit data.

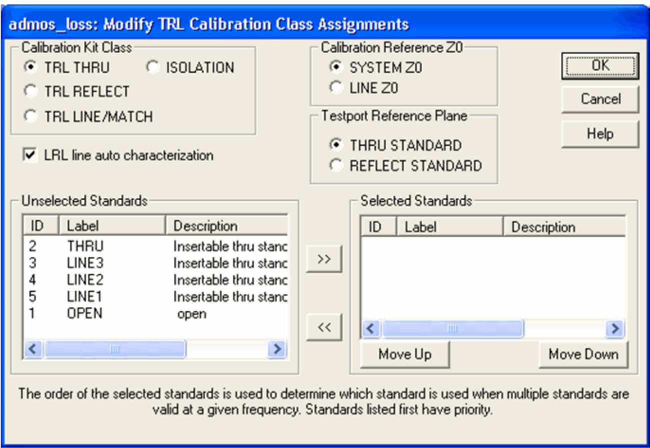
Note
For the LINES, change the Label to LINEx.



After having entered the TRL standards (as mentioned, it is strongly recommended to use a SHORT rather than an OPEN), select 'Class Assignments' as TRL. Then hit 'Edit'.



Click through the 'Calibration Kit Klass' readio boxes and associate the Calkit Standards with these classes.
Then, for the Calibration Reference Z0, select SystemZ0.
For the 'Testport Reference Plane', select THRU standard (this will mike the cal reference plane to the middle of the THRU).

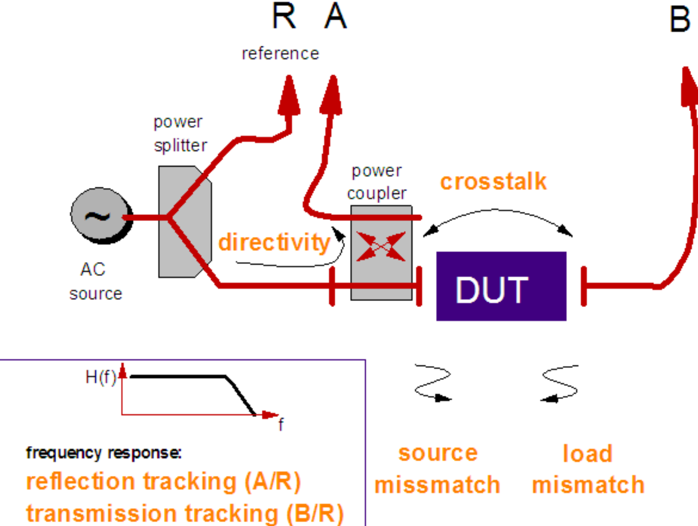


Finally, save the Calkit on the PNA, typically in the default directory.

NWA Calibration in General

VNA Standard Calibration Techniques and Verification

A considerable challenge in S-parameter VNA measurements is to define exactly where the measurement system ends and the DUT begins. This location is called 'reference plane'. This means, all error contributions, inside the VNA and also from the cables up to this reference plane, have to be calibrated out.
The calibration of a NWA is performed by rather complex procedures. Such are Short-Open-Load-Through (SOLT), Through-Reflection-Load (TRL) or Load-Refection-Match (LRM) and the associated error correction model.

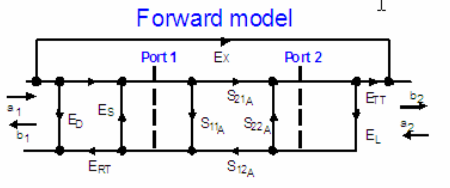


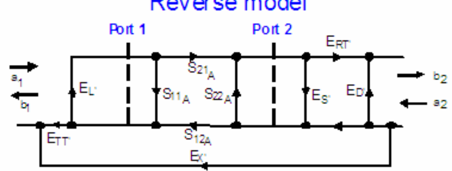
As can be seen in the following figure, there are 6 error contribution terms in forward direction:

Directivity	**Cross-talk of the power coupler**
Crosstalk	Cross-talk inside the S-parameter test set, overlying the DUT
Source Mismatch	Multiple reflections due to non-Z0 input and output impedance of the cables and connectors
Load Mismatch	The same for the opposite port
Reflection Tracking A/R	Frequency dependence of signal path R->A
Transmission Tracking A/R	Same for signal path R->B

For the reverse calibration, another 6 error terms add up to a total of 12 terms. These 12 terms are often referred as the 12-term error correction.
For the different calibration procedures, specific, accurately known standards have to be measured. For connectorized DUTs, with e.g. 3.5mm connectors, usually 4 calibration standards like OPEN, LOAD, SHORT and THRU are applied, with exactly known standard inaccuracies.
It is interesting to note that each corrected S-Parameter is a function of all other (uncorrected) measured S-Parameters, see the following figure:

12 Term Error Correction





E_D = Fwd Directivity
E_S = Fwd Source Match
E_{RT} = Fwd Reflection Tracking
$E_{D'}$ = Rev Directivity
$E_{S'}$ = Rev Source Match
$E_{RT'}$ = Rev Reflection Tracking
E_L = Fwd Load Match
E_{TT} = Fwd Transmission Tracking
E_X = Fwd Isolation
$E_{L'}$ = Rev Load Match
$E_{TT'}$ = Rev Transmission Tracking
$E_{X'}$ = Rev Isolation

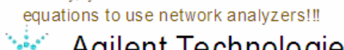
$$S_{11a} = \frac{\left(\frac{S_{11m}-E_D}{E_{RT}}\right)\left(1+\frac{S_{22m}-E_D'}{E_{RT}'}E_S'\right)-E_L\left(\frac{S_{21m}-E_X}{E_{TT}}\right)\left(\frac{S_{12m}-E_X'}{E_{TT}'}\right)}{\left(1+\frac{S_{11m}-E_D'}{E_{RT}}E_S\right)\left(1+\frac{S_{22m}-E_D'}{E_{RT}'}E_S'\right)-E_L'E_L\left(\frac{S_{21m}-E_X}{E_{TT}}\right)\left(\frac{S_{12m}-E_X'}{E_{TT}'}\right)}$$

$$S_{21a} = \frac{\left(\frac{S_{21m}-E_X}{E_{TT}}\right)\left(1+\frac{S_{22m}-E_D'}{E_{RT}'}(E_S'-E_L)\right)}{\left(1+\frac{S_{11m}-E_D}{E_{RT}}E_S\right)\left(1+\frac{S_{22m}-E_D'}{E_{RT}'}E_S'\right)-E_L'E_L\left(\frac{S_{21m}-E_X}{E_{TT}}\right)\left(\frac{S_{12m}-E_X'}{E_{TT}'}\right)}$$

$$S_{12a} = \frac{\left(\frac{S_{12m}-E_X'}{E_{TT}'}\right)\left(1+\frac{S_{11m}-E_D}{E_{RT}}(E_S-E_L')\right)}{\left(1+\frac{S_{11m}-E_D}{E_{RT}}E_S\right)\left(1+\frac{S_{22m}-E_D'}{E_{RT}'}E_S'\right)-E_L'E_L\left(\frac{S_{21m}-E_X}{E_{TT}}\right)\left(\frac{S_{12m}-E_X'}{E_{TT}'}\right)}$$

$$S_{22a} = \frac{\left(\frac{S_{22m}-E_D'}{E_{RT}'}\right)\left(1+\frac{S_{11m}-E_D}{E_{RT}}E_S\right)-E_L'\left(\frac{S_{21m}-E_X}{E_{TT}}\right)\left(\frac{S_{12m}-E_X'}{E_{TT}'}\right)}{\left(1+\frac{S_{11m}-E_D}{E_{RT}}E_S\right)\left(1+\frac{S_{22m}-E_D'}{E_{RT}'}E_S'\right)-E_L'E_L\left(\frac{S_{21m}-E_X}{E_{TT}}\right)\left(\frac{S_{12m}-E_X'}{E_{TT}'}\right)}$$

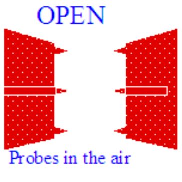
- Notice that each corrected S-parameter is a function of all four measured S-parameters
- Analyzer must make forward *and* reverse sweep to update any one S-parameter
- Luckily, you don't need to know these equations to use network analyzers!!!



For the case of on-wafer measurements, SOLT calibration, and using Ground-Signal-Ground (GSG) coplanar probes, fig. 3 depicts the corresponding test structures, which are usually available on a RF-high-performance ceramic substrate (ISS substrate).

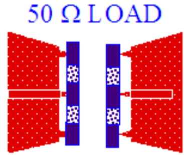
SOLT calibration
for Ground-Signal-Ground probes (G_S_G)

OPEN



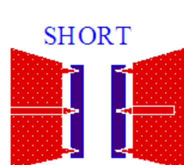
Probes in the air

50 Ω LOAD

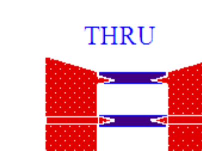


SOLT
calibration

SHORT



THRU



All standard calkit data must be accurately entered into the VNA.
SOLT is sensitive to probe placement (probe tip instance at THRU).

About Calibration Standards

While in the case of the CV meter, the OPEN calibration corrects for an ideal offset capacitor, a NWA calibration uses cal standards (OPEN, SHORT, LOAD, THRU etc.) from a cal kit.

These cal standards do not represent ideal standards. They represent the real, existing standard, including its nonidealities! It means that a SHORT is not an ideal SHORT, but instead represents rather an inductance. The same applies to the THRU, which has a non-ideal delay time. The OPEN corresponds rather to a capacitor than to an ideal OPEN. Therefore, these nonidealities of the G-S-G probes have to be entered into the NWA before calibration. This is called 'modifying the cal kit'.

While this procedure refers to the nonidealities of the calibration standards, the subsequent calibration is related to the selected frequency range, the RF power, the averaging of the NWA etc. After it has been performed, the correction terms are stored in the cal set of the NWA. In other words, the 12-term error vectors are 'filled up'.

Afterwards, when the measurement is performed, the raw measured data arrays will be corrected using a correction technique related to the selected calibration method, and referring to the specified cal set. Finally, this corrected measurement result is transferred into IC-CAP and displayed there (as well as on the NWA monitor).

Therefore, after the calibration, a re-measurement of the OPEN will not represent an ideal

open, but instead exactly those parasitic components as described in the documentation of the OPEN. In the same way, a THROUGH shows up after calibration with its real delay time, and a SHORT represents its inductive behavior!

This re-measuring of the standards together with a simulation of their known values can be used to verify the VNA calibration, and this can be done easily, using IC-CAP.

We simply re-measure the cal.standards, e.g. the OPEN, the SHORT, the THRU and the LOAD. As said above, this measurement will correspond to the nonidealities of the selected cal.standard. In case of Cascade probes, the OPEN, for example, behaves like a negative capacitance of roughly -9fF . Now, after this measurement has been made, we can define a test circuit for that setup in IC-CAP, entering the netlist of the calibration standards. Using SPICE3, a simulator which also permits negative capacitances, we can simulate the expected behavior of the OPEN probes. If the calibration was executed correctly, there is an excellent match between measured and simulated curves.
In a next step, we measure the SHORT, define in IC-CAP the SHORT nonidealities in a SPICE circuit, and simulate. Again, an excellent match between measured SHORT data and simulations has to be achieved. We then continue with the THRU and LOAD measurements and simulations. Only if all 4 standards exhibits an excellent fit, we can assume a correct calibration of the NWA.

> **Note**
> This calibration verification can also be applied to check the quality of an older calibration.

NWA Calibration Procedure For At-Temperature Measurements

Example using WinCal of Cascade Microtech

1. Place the wafer on the thermal chuck and the ISS substrate on the aux chuck
2. Pre-align the wafer in theta and use the WinCal software to align the ISS
3. Set desired temperature
4. Once at temperature, put the probes just on the wafer pad and allow to stabilise at temperature for 15 mins.
5. Move the probes over to the ISS and place on the alignment mark
6. Perform Auto calibration (probe alignment and calibration should be done as fast as possible before probe temperature changes)
7. Move probes back to device and place probe back on device pads
8. Leave for 5 minutes to re-stabilize
9. Use the WinCal verification tool to check calibration stability
10. Measure the device

NWA Calibration Verification

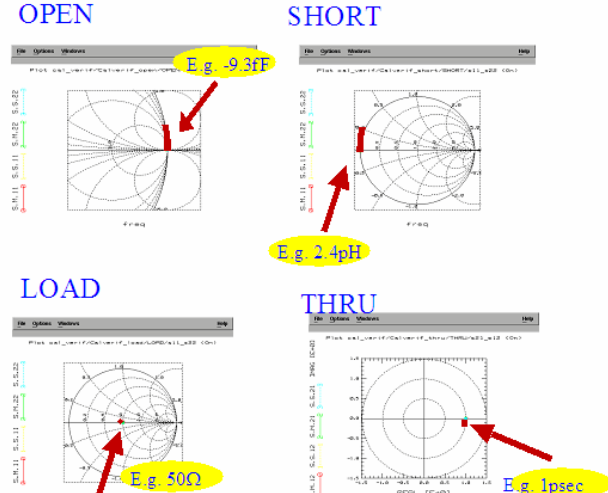
In order to verify the calibration, it is highly recommended to re-measure the calibration standards and to model them, using the calkit data of the GSG probe or the connectorized standards manufacturer.

For this verification, and for an on-wafer SOLT calibration for example, we re-measure the cal.standards, e.g. the OPEN, the SHORT, the THRU and the LOAD. We know that this measurement will correspond to the nonidealities of the selected cal.standard. In case of Cascade probes, the OPEN, for example, behaves like a negative capacitance of roughly -9fF . Now, after this measurement has been made, we can define a test circuit for that Setup in IC-CAP, entering the netlist of this calibration standards. Using SPICE3, a simulator which also permits negative capacitances, we can simulate the expected behavior of the OPEN probes. If the calibration was executed correctly, there is an excellent match between measured and simulated curves.
In a next step, we measure the SHORT, define in IC-CAP the SHORT nonidealities in a SPICE circuit, and simulate. Again, an excellent match between measured SHORT data and simulations has to be achieved. We then continue with the THRU and LOAD measurements and simulations. Only if all 4 standards exhibits an excellent fit, we can assume a correct calibration of the NWA.

> **Note**
> This calibration verification can also be applied to check the quality of an older calibration.

The following figures give an example for Cascade G-S-G probes, 100um pitch.

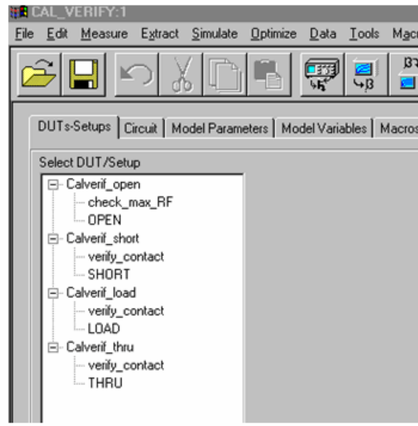
Verifying the calibration by modeling all the standards of the calibration kit



Only if the fitting between simulated and measured data is in the few-percent range, for all 4 re-measured calibration standards, the NWA calibration can be considered as good. If only one fit is bad, re-perform a new NWA calibration.

HINT:
use the IC-CAP CAL_VERIFY_MASTER_demodata.mdl File in the demo_features directory.

The scheme of the IC-CAP CAL_VERIFY_MASTER.mdl file

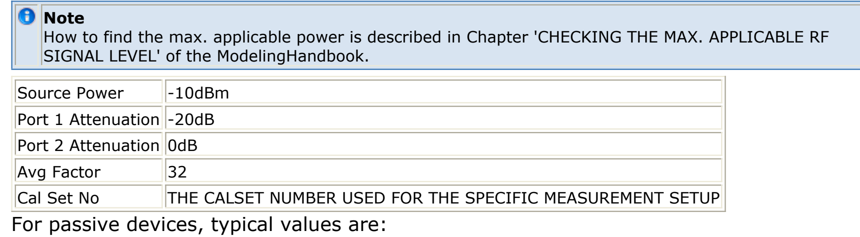


including checking of max. RF power, and the calibration verification of the standards. Additionally, DC Setups are added for each DUT to allow verification of good DC contacts during calibration.

NWA Instrument Options in IC-CAP

Instrument Options for the HP/Agilent 87xx and 8510 Network Analyzers

Usually, for transistors, these settings might be correct:



Note
How to find the max. applicable power is described in Chapter 'CHECKING THE MAX. APPLICABLE RF SIGNAL LEVEL' of the ModelingHandbook.

Source Power	-10dBm
Port 1 Attenuation	-20dB
Port 2 Attenuation	0dB
Avg Factor	32
Cal Set No	THE CALSET NUMBER USED FOR THE SPECIFIC MEASUREMENT SETUP

For passive devices, typical values are:

Source Power	-10dBm
Port 1 Attenuation	0dB
Port 2 Attenuation	0dB

Instrument options for the Agilent PNA Network Analyzers

Setting the Port Power Levels is the main difference between the Instrument Options of the PNA series compared to the older Agilent/HP network analyzers (87xx or 8510):

Set the entry 'Port Power Coupled' to 'No'. This gives an individual, fully decoupled and

independent 'Port Src Power' and 'Port Src Atten' for each port ! I.e. select an individual combination of 'Port Src Power' and 'Port Src Atten' for each port which corresponds to the max. applicable power rate of your device.

Note
How to find the max. applicable power is described in Chapter 'CHECKING THE MAX. APPLICABLE RF SIGNAL LEVEL'
of the ModelingHandbook.

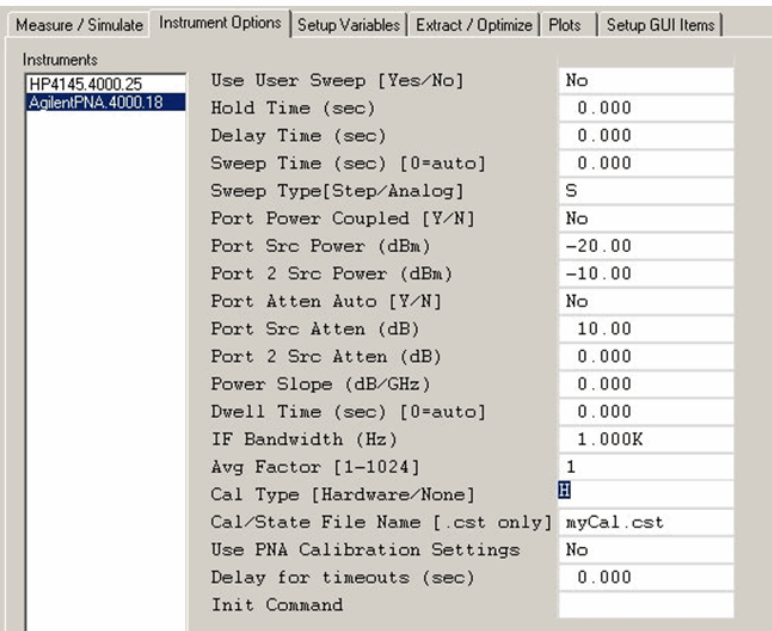
Example of a typical setting for transistors:

- Port 1: Power -20dBm, Attenuation 10dB,
- Port 2: Power -10dBm, Attenuation 0dB

Note
For each port of the PNA, you can achieve any specified, desired power level, however, not all combinations of 'Port Source Power' and Port Source Attenuation' and 'Power Slope' are possible (error message 'Source unleveled')

Therefore, it is recommended to first try manually the possible 'Port Src Power' and 'Port Src Atten' settings directly at the PNA, and to enter these values afterwards into the Instrument Options in IC-CAP, (before clicking on 'Calibrate' in the IC-CAP Setup).

Instead of Cal Set numbers, the PNAs offer to save the Cal Sets with filenames in the memory of the PNA processor. Therefore, save the Cal Set in the default directory of the PNA and give it a filename. Then, enter that filename into the IC-CAP Instrument Options Field 'Cal File Name'.
Available IC-CAP Instrument Options for the PNA series (IC-CAP 2006B Update3)



In details (from the IC-CAP Help File):
YELLOW: must-enter-fields (if Use PNA Calibration Settings=yes).
HINT: leave non-yellow fields default

Entry Field	Note	Default	Typical
Use User Sweep	Yes = use user sweep, i.e. IC-CAP triggers every frequency point. No = use instrument's internal sweep.	No	-
Hold Time	Time, in seconds, the instrument waits before beginning a sweep to allow for dc settling.	0	-
Delay Time	Time the instrument waits before setting each frequency in user sweep mode.	0	-
Sweep Time	Time the instrument takes for each sweep. 0 = Auto	0	-
Sweep Type[SA	S = Stepped mode. A = Analog (ramp) mode	S	-
Port Power Coupled	Yes = Coupled mode. No = Non-Coupled mode. When Ports are coupled, Port Src Power is used for both Port 1 and 2. Port 2 Src Power is ignored. Attenuators are also coupled so that Port Src Atten is used for both ports and Port 2 Src Atten is ignored.	Yes	No
Port Src Power	Define the source Power for Port 1 and 2 when ports are coupled or the source power for Port 1 when ports are uncoupled. The power range depends on the attenuator settings and the PNA model and options.	-	-20 (Transistor) -10 (pass.dev.)
Port 2 Src Power	Define the source power for Port 2 when ports are uncoupled. This option field is ignored when ports are coupled. The power range depends on the attenuator settings and the PNA model and options.	-	-10 (Transistor and pass.dev.)
Port Atten Auto	Yes = Auto mode. No = Non-Auto mode. When attenuators are in auto-mode, the PNA will set the most efficient values for the attenuators to obtain the requested output power at the port. In auto-mode, the full power range is directly available at the output port. In auto-mode, the instrument options Port Src Atten and Port 2 Src Atten are ignored.	No	-
Port Src Atten	Possible Values: 0, 10, 20, 30, 40, 50, 60, 70 dB. The available range depends on the PNA model. For example, the E8364A attenuator range is 0-60 dB. This option is ignored when attenuators are in auto-mode.	0	5 - 10 (Transistor) 0 (pass.dev.)
Port 2 Src Atten	Possible Values: 0, 10, 20, 30, 40, 50, 60, 70 dB . The available range depends on the PNA model. For example, the E8364A attenuator range is 0-60 dB. This option is ignored when attenuators are in auto-mode.	0	0
Power Slope	Can be any value between -2 and +2 dB/GHz	0	-
Dwell Time *	Sets the dwell time, in seconds, between each sweep point. Only available in Stepped sweep type. (Auto - PNA will minimize dwell time)	0	-
IF Bandwidth	Possible Values: 1, 2, 3, 5, 7, 10, 15, 20, 30, 50, 70, 100, 150, 200, 300, 500, 700, 1k, 1.5k, 2k, 3k, 5k, 7k, 10k, 15k, 20k, 30k, 35k, 40k Hz Note: If a non valid value is specified, the PNA will not round it to the nearest available value, rather it will use the lowest valid value close to specified value.	1000	200
Avg Factor	Number of averages per measurement. 1-1024	1	1
Cal TypeHN	H = Hardware calibration. N = No calibration	N	H
Cal File Name	Name of PNA cal file to be used. If no file extension is specified, .cst (cal file & instrument state) is assumed.	myCal.cst	-
Use PNA Calibration Settings	When set to 'yes', IC-CAP will not set the PNA InstrumentOptions, and will accept all user-specified (or WinCal-specified) settings. All InstrumentOptions entries are ignored by IC-CAP, except 'Cal File Name' and 'Cal Type'. As a user, you only need to set the Start-Stop-Step in your 'freq' Input accordingly to the PNA frequency setting.	No	-
Delay for timeouts	For long-running measurements use this option to avoid measurement timeouts.	0	-
Init Command	Command field to set the instrument to a mode not supported by the option table. Command is sent at the end of instrument initialization for each measurement. Normal C escape characters such as \n (new line) are available.	-	-

NOTE on DWELL time:
Dwell time is the time spent at each frequency point before sampling starts. For most applications, you should set dwell time to auto mode. In auto mode, the PNA increases the dwell time as the sweep time increases to comply the total sweep time. If long delays are present in the circuit and additional settling time is needed, set the dwell time to an appropriate value.
Dwell time is not active in analog mode--only in step mode. If the sweep time in analog mode is increased significantly (because of a setting), the PNA can internally switch to step mode and set an optimum value for the dwell time.

NOTE on ATTENUATORS IN AUTO MODE:
When port attenuators are set to auto mode, the PNA automatically chooses the attenuator value that provides the requested power level at the output port. Since accurate S-parameter calibration requires that the attenuator settings do not change during measurements or calibration, auto mode is not recommended.

NOTE on SWEEP TYPE:
Step sweep mode is more accurate than analog (ramp) mode, but analog mode is usually faster than step sweep mode. In step sweep mode, RF phase locking is performed at each frequency, which ensures that the frequency value is very accurate. This results in a longer transition time from one frequency point to the next and a longer total sweep time. In analog mode, the RF frequency is swept across the frequency range and its frequency accuracy depends on the linearity of the VCO (Voltage Controlled Oscillator).

NOTE on Cal TypeHN :
-> **NEVER** use 'N' (no calibration)

NOTE on Use PNA Calibration Settings:
If you prefer to set your PNA manually (or use e.g. WinCal for all PNA settings), then set Use PNA Calibration Settings=Yes
In this case, IC-CAP will accept all settings of the PNA and will no longer be 'the boss' of the PNA instrument (what is the default setting of IC-CAP). You only have to enter the frequency settings and the .cst file correctly.
NOTE: when set to 'Yes', all irrelevant InstrumentOptions fields should commonly be greyed-out (inactive). This is currently not possible for the IC-CAP source code, but this will be changed in ICCAP2008.

IMPORTANT NOTE ABOUT AVERAGING AND PNAs:

It is often preferable to apply some signal-to-noise algorithm to measured data. In the case of the 8510 network analyzer, the general practice is to set averaging to ~16 or ~32. (the added benefit of setting it to higher values is generally negligible).

In the newer PNAs, however, the recommended use paradigm to achieve the same benefit as averaging in the 8510, is to define IF Bandwidth, rather than averaging. The value setting is up to the user but generally values of 200 Hz or less are recommended for modeling applications.

In other words:
reducing IF Bandwidth in PNA is ~ increasing averages in the older 8510...87xx NWAs.

Background: the PNA apply digital signal processing, while the older 8510 and 87xx applied analog signal processing.

Another reason that the 'best practices' recommendation to set IF Bandwidth vs. averaging to reduce the noise, is measurement efficiency. If one sets 'Number of Averages' to 'n', this forces the PNA to take 'n' sweeps (vs. by setting IF Bandwidth, the time for data acquisition at each point may be increased, but there is only 1 sweep).

IMPORTANT NOTE: if Cascade/Wincal is being used to automate the PNA calibration, averaging is not saved with the PNA calibration file, as Cascade recommends the use of IF Bandwidth settings.
So even if one had set averaging > 1 in the IC-CAP Setup Instrument Options, this averaging would not be used or saved in the PNA cal file.

Performing the NWA Calibration from IC-CAP

NWA CALIBRATION UNDER IC-CAP

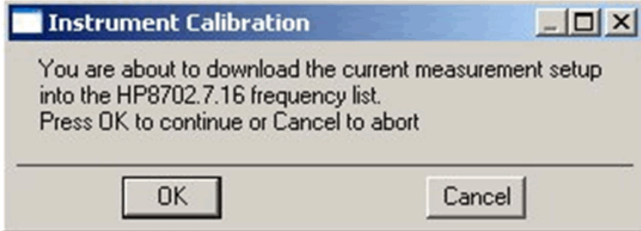
Provided that the Cal Kit Data which is provided with every GSG probes calibration substrate have been entered correctly into the network analyzer (see the other chapter (.pdf file) located in the same directory), and we have determined the maximum applicable RF signal level, the calibration of the NWA under IC-CAP is pretty simple.

The steps are:
After the IC-CAP Setup has been defined with a frequency sweep under 'Inputs' and an S-parameter output, the next step is to define the NWA options under 'Instrument Options'. Enter particularly the port power levels, the averaging and the Cal Set number (for the PNAs the Cal Set Filename).
Make sure to select an appropriate RF power level, in order to not overdrive the DUT.

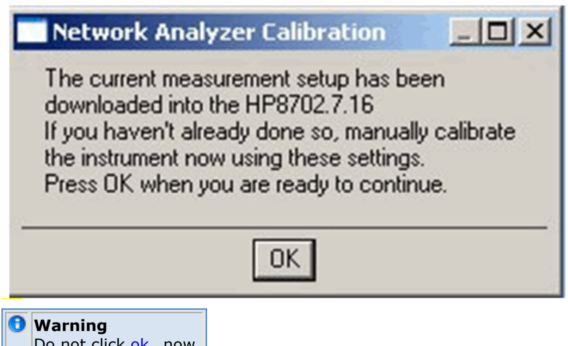
- THE NWA IS A LINEAR MEASUREMENT SYSTEM,
- THEREFORE HARMONICS (OVERDRIVEN DUTs) ARE NOT ALLOWED !
  see the chapter on evaluating the correct RF signal level
  in the other chapter (.pdf file) located in the same directory.

H3. Using Agilent PNAs

1. In the IC-CAP Setup's InstrumentOptions, when you set 'Use PNA Calibration Settings' = Y, then you only need to enter the field for CalType = H, and Cal/StateFileName.cst only . All other InstrumentOptions fields are ignored by IC-CAP.
2. Then, click 'Calibrate' in the pulldown menu of the actual IC-CAP setup.
   IC-CAP transfers all instrument settings to the NWA.
3. Then, a box appears with a note:



4. Press OK.



Warning
Do not click ok now.

5. Perform the calibration manually on the NWA first or use WinCal or any other cal. software.
6. Save your calibration in the Calset specified earlier in the IC-CAP instrument options (see below the infos about older HP NWAs and the Agilent PNA series).
7. After all that, finally click OK in IC-CAP.

Series 8510x and 87xx Network Analyzers of HP / Agilent

1. On the NWA, click 'local', 'Cal', and select the appropriate calkit.
2. Click 'Calibrate Menu', and then click 'Full 2-port'.
3. Perform that full TwoPort calibration manually.
4. Finally, save the calset in the same calset number which was entered into IC-CAP before (under instrument options).

AGILENT PNA NETWORK ANALYZER SERIES

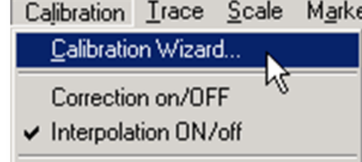
Note

For each port of the PNA, you can achieve any specified, desired power level, however, not all combinations of 'Port Source Power' and Port Source Attenuation' and 'Power Slope' are possible (error message 'Source unleveled'){info)

Therefore, it is recommended to first try manually the possible 'Port Src Power' and 'Port Src Atten' settings directly at the PNA, and to enter these values afterwards into the Instrument Options in IC-CAP, (before clicking on 'Calibrate' in the IC-CAP Setup).

After you have entered the Instrument Options in IC-CAP, click on Calibrate, wait until you get the IC-CAP message window as described above, (do NOT click 'ok' on that IC-CAP window), and go to the PNA. Hit the 'local' button.

With the mouse, select Calibrate/Calibration Wizard,



and you get

Defining Custom Calkits

When measuring packaged devices, the network analyzer calibration is typically performed using 3.5mm connector-based calibration standards.

As an example, the table below gives the HP85033D 3.5mm calkit data for the HP8753 NWAs,for an 'insertable device' (Delay THRU = 0 ps !!) with a male 3.5mm connector at one port and a female 3.5mm connector at the other port.

STANDARD No.	TYPE	C0 fF / L0 pH	C1/ E-27F/Hz / L1/ E-24H/Hz	C2 E-36F/Hz² / L2/ E-33H/Hz²	C3/ E-45F/Hz³ / L3/ E-42H/Hz³	Fixed or Sliding	OFFSET DELAY psec	OFFSET Z0 Ohm	OFFSET LOSS Ohm/s	FREQ MIN	FREQ MAX	STND LABEL
1	SHORT	0					31.8p	50	2.36G	0	999	SHORT
2	OPEN	49.4f	-310.13	23.17	160m		29.24p	50	2.2G	0	999	OPEN
3	LOAD					fixed	0	50	2.3G	0	999	LOAD
4	THRU						0	50	2.3G	0	999	THRU

This means, the calibration plane is ending at the 3.5mm connectors.
It has to be extended to the very location of the packaged device, within the test fixture.

This is usually done by modeling the test fixture consisting of the other 3.5 mm connector and the strip line(s) down to the DUT (device under test). A TDR would be applied in order to evaluate the equivalent circuit along the connections. After that, a NWA would be applied to find out the crosstalk between the connections, and to add the corresponding enhancements to the equivalent circuit. After inserting the DUT, modeling continues, and all model components required from now on for device modeling are part of the 'DUT model'.

Another possibility would be to de-embed the modeled test fixture, to obtain the DUT S-parameters out of the measurements of test fixture and DUT.

On the other hand, if custom calkits, i.e. OPEN, SHORT, LOAD and THRU, were available for the custom test fixture, i.e. and if these custom calkits were accurately known (modeled), we could also use them instead of the 3.5mm connector-based calkits and obtain a calibration including the test fixture. After NWA calibration, the calibration plane would then end where the DUT is connected to, on the fixture !

In other words, we need to specify a table (like the one above) for the custom calkit. This chapter sketches the realization of this idea.

Before we begin, here a copy of parts of Agilent Application Note AN 1287-9:
In-Fixture Measurements Using Vector Network Analyzers

Characterizing calibration standards for SOLT calibration:

Most network analyzers already contain standard calibration kit definition files that describe the characteristics of a variety of calibration standards. These calibration kit definitions usually cover the major types of coaxial connectors used for component and circuit measurements, for example Type-N, 7 mm, 3.5 mm and 2.4 mm connector-based calkits.

Most high-performance network analyzers allow the user to modify the definitions of the calibration standards. This capability is especially important for fixture-based measurements, because the in-fixture calibration standards will not have the same attributes as the coaxial standards. Custom calibration standards, such as those used with fixtures, require the user to characterize the standards and enter the definitions into the network analyzer. The calibration kit definition must match the actual standards for accurate measurements. Definitions of the in-fixture calibration standards can be stored in the analyzer as a custom user-defined calibration kit.

Characterizing a SHORT

The electrical definition of an ideal short is unity reflection with 180 degrees of phase shift. All of the incident energy is reflected back to the source, perfectly out of phase with the reference. A simple short circuit from a single conductor to ground makes a good short standard. For example, the short can be a few vias (plated through holes) to ground at the end of a micro-strip transmission line. If coplanar transmission lines are used, the short should go to both ground planes.
To reduce the inductance of the short, avoid excessive length. A good RF ground should be near the signal trace. If the short is not exactly at the contact plane of the DUT, an offset length can be entered (in terms of electrical delay) as part of the user-defined calibration kit.

Characterizing an OPEN

The open standard is typically realized as an unterminated transmission line. Electrical definition of an ideal open has "unity reflection with no phase shift." The actual model for the open, however, does have some phase shift due to fringing capacitance.
How to determine open capacitance:
The fringing capacitance can be measured as follows:

1. Perform a one-port calibration at the end of the test cable. Use a connector type that is compatible with the fixture. For example, use APC 3.5-mm standards for a fixture using SMA connectors.
2. Connect the fixture and measure the load standard. This data should be stored in memory and the display changed to "data minus memory". This step subtracts out the reflection of the fixture connector (assuming good consistency between connectors), so that we can characterize just the open. An alternative is to use time-domain gating to remove the effect of the connector.
3. Measure the short standard. Set the port extension to get a flat 180 degrees phase response. To fine-tune the value of port extension, set the phase-off set value for the trace to 180 degrees and expand degrees-per-division scale. Mismatch and directivity reflections may cause a slight ripple, so use your best judgment for determining the flattest trace, or use marker statistics (set the mean value to zero).
4. Set the network analyzer display format to Smith chart, the marker function to Smith chart format G+jB (admittance) and then measure the open standard. Markers now read G+jB instead of the R+jX of an impedance Smith chart. Admittance must be used because the fringing capacitance is modeled as a shunt element, not a series element. The fringing capacitance (typically 0.03 to 0.25 pF)can be directly read at the frequency of interest using a trace marker. At RF, a single capacitance value (C0) is generally adequate for the calibration kit definition of the open. In some cases, a single capacitance number may not be adequate, as capacitance can vary with frequency. This is typically true for the measurements that extend well into the

microwave frequency range. Capacitance may vary with frequency, especially at frequencies above 3 GHz.
When measuring the fringing capacitance, a problem can arise if the short standard is electrically longer than the open standard. The measured impedance of the open circuit then appears to be a negative capacitor, indicated by a trace that rotates backwards (counter-clockwise)on the Smith chart. This problem is a result of using an electrically longer short standard as the 180 degrees phase reference. The electrically shorter open will then appear to have positive phase. The remedy for this is to decrease the port extension until the phase is monotonically negative. The model for the open will then have a normal (positive)capacitance value. The value of the negative offset delay that needs to be included in the open standard definition is simply the amount by which port extension was reduced (for instance, the difference in the port extension values between the short and the open).In effect, we have now set the reference plane at the short. Alternatively, the offset delay of the open can be set to zero, and a small positive offset delay can be added to the model of the short standard. This will set an effective reference plane at the open.

Characterizing a LOAD

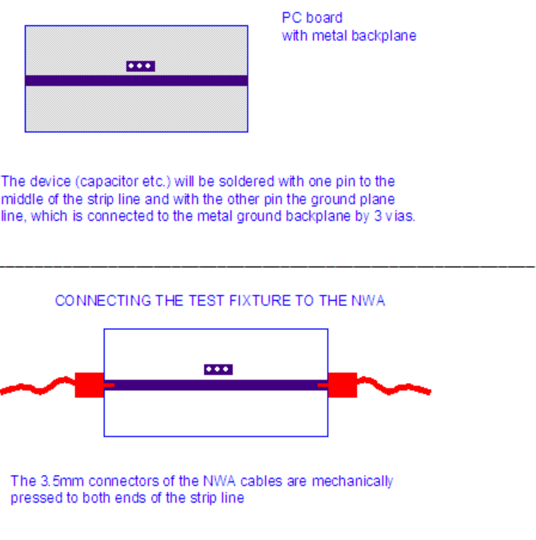
An ideal load reflects none of the incident signal, thereby providing a perfect termination over a broad frequency range. We can only approximate an ideal load with a real termination because some reflection always occurs at some frequency, especially with non-coaxial actual standards.
At RF, we can build a good load using standard surface-mount resistors. Usually, it is better to use two 100-ohm resistors in parallel instead of a single 50-ohm resistor, because the parasitic inductance is cut in half. For example, 0805-size SMT resistors have about 1.2 nH series inductance and 0.2 pF parallel capacitance. Two parallel 100-ohm 0805 resistors have nearly a 20-dB better match than a single 50-ohm resistor. Also note that special RF-performance SMT resistors are available !

Characterizing a THRU

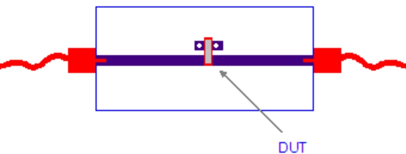
The thru standard is usually a simple transmission line between two coaxial connectors on the fixture. A good thru should have minimal mismatch at the connector launches and maintain a constant impedance over its length (which is generally the case for PCB thrus). The impedance of the thru should match the impedance of the transmission lines used with the other standards (all of which should be 50 ohms).
If the PC board is wider for the transmission line where the DUT will be soldered and since we want the two halves of line to be equal in electrical length to the thru line, the PCB must be widened by the length of the DUT. With a properly designed PC board fixture, the short (or open) defines a calibration plane to be in the center of the fixture. This means the thru will have a length of zero (which is usually not the case for fixtures used in manufacturing applications, where a set of calibration standards is inserted into a single fixture). Since the length is zero, we do not have to worry about characterizing the loss of the thru or its phase shift.

After this prologue of the Application Note, we want to enhance the above mentioned ideas and apply device modeling techniques for obtaining an accurate model of our custom calibration standards.
First of all, this is the test fixture layout:

TEST FIXTURE LAYOUT



DEVICE-UNDER-TEST (DUT) SCENARIO

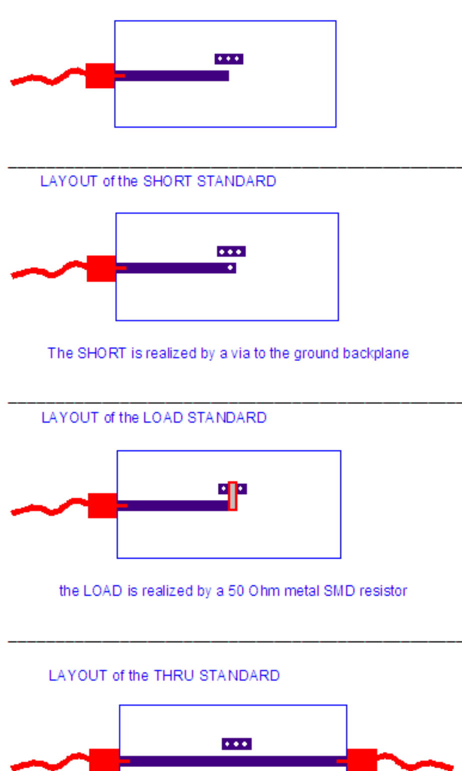


In this scenario, the DUT will be measured in a shunt scenario, i.e. towards Ground.

> **Note**
> Another test fixture layout would be two non-connected striplines, each of the length of the later OPEN, SHORT and LOAD cal. standard, leading to both ends of the DUT. The THRU would be exactly twice the length of the individual OPEN, SHORT and LOAD striplines.

In this case, the DUT would be measured between the NWA ports, and not towards Ground.

The next figures depict the layout of the corresponding calkit standards:



Regarding the calkit standards from above, the modeling has to be performed in such a way that we can enter the model into the calkit definition table of the NWA, taking the limitations of the table entries into account.
This means, we need to specify a shifted calibration plane from the connector ends to the middle of the THRU standard. Therefore, the THRU calkit data will be delay = 0 ps.

For the modeling, we will have to model the two connectors at each side, plus a lossy delay line. From this model, the left connectors and half of the THRU delay line will be used later for de-embedding the OPEN, SHORT and LOAD standards. The remaining performance of these de-embedded standards will be modeled as a frequency-dependent capacitance for the OPEN, plus possibly a fine-tuning delay (taking into account a possible different electrical length of the OPEN compared to half the length of the THRU standard),an inductor for the SHORT standard, plus possibly a fine-tuning delay, a delay and an impedance Z0 in series with the 50 Ohm resistor for the LOAD standard (taking into account the frequency-dependence of the 50 Ohm resistor).

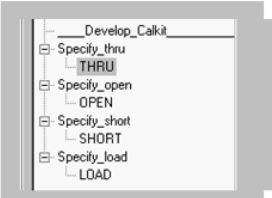
STANDARD No.	TYPE	C0 fF / L0 pH	C1/ E-27F/Hz / L1/ E-24H/Hz	C2 E-36F/Hz² / L2/ E-33H/Hz²	C3/ E-45F/Hz³ / L3/ E-42H/Hz³	Fixed or Sliding	OFFSET DELAY psec	OFFSET Z0 Ohm	OFFSET LOSS Ohm/s	FREQ MIN	FREQ MAX	STND LABEL
1	SHORT	___					___	___	0	0	999	SHORT
2	OPEN	___	___	___	___		___	___	0	0	999	OPEN
3	LOAD					fixed	___	___	0	0	999	LOAD
4	THRU						___	___	0	0	999	THRU

STEP-BY-STEP MODELING PROCEDURE FOR THE CUSTOM CALKITS:

> **Note**
> Pls. contact franz.sischka@agilent.com for a copy of the IC-CAP file shown below.

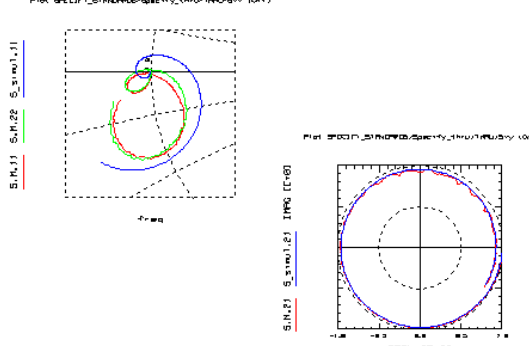
1. a conventional NWA calibration is performed based on a standard 3.5mm connector calkit
2. this calibration is verified by re-measuring all 4 cal standards, and by comparing the 4 measurement results by the corresponding 4 SPICE simulations. See the corresponding calibration verification Model file of IC-CAP, in the demo_features directory.

After that, the custom calkits are connected and measured individually. For each calkit standard, a model will been developed. Since we will have to de-embed the connector and half of the THRU for the OPEN, SHORT and LOAD; what is done by inverse ABC matrix (chain matrix) multiplications, we will use only ABC matrices for the standard modeling, and not a simulator like SPICE.

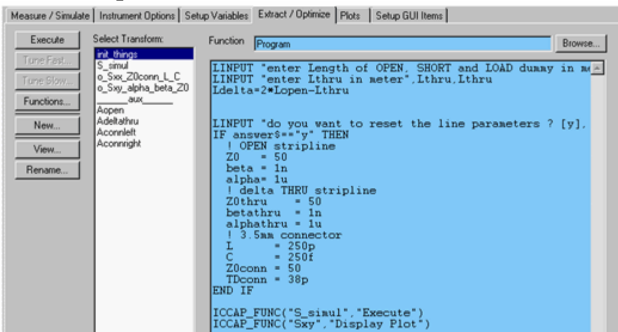


CUSTOM THRU CALKIT: Measurement and SPICE Model
As mentioned, we want the new reference plane in the middle of the THRU standard, so the THRU is modeled first.

MEASUREMENT AND MODELING
OF THE CUSTOM THRU STANDARD



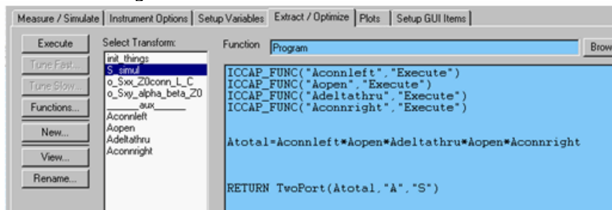
The plots above show the measured data and the fit of the model. Particularly important is the fitting of the Sxy plots, which corresponds to the length of the thru microstrip line and its loss.
The connectors show basically up in the Sxx plot, as a deviation of the measurement from the center of the Smith chart, i.e. Z0 = 50 Ohm.
The modeling transforms in IC-CAP look like this:



Modeling:
we commence with the model parameters of the strip line (plus the possibly existing delta-L of the THRU if the test fixture offers 2 microstrip lines instead of a single one. In this case, the THRU stripline is Ldelta longer than 2*Lopen.)

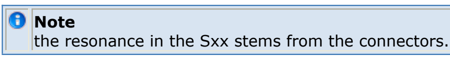
The next detail of IC-CAP depicts the PEL program 'S_simul' which simulates the THRU standard, i.e. the chain of Aconnleft*Aopen*Adeltathru*Aopen*Aconnright

Transform 'Aopen' holds the A matrix data of the microstrip line of the same length as the OPEN standard, Adeltathru the A matrix data of the possibly existing Ldelta strip line (DUT between the ports), Aconnleft and Aconnright model the 3.5mm connector with each a series inductor and a shunt capacitor plus a typ. 38ps delay line of 50 Ohm. Watch out that Aconnleft and Aconnright contain flipped schematics, since A matrices always model from 'left to right' !



Parameter fine-tuning is performed for the Sxy plot with optimizer 'o_Sxy_alpha_beta_Z0', which either tunes or optimizes the delay and the attenuation of the THRU microstrip line. Keep an eye on the starting trace of the Sxx curve: if it ascends from Z0=50 Ohm, then Z0 of the microstrip line is >50 Ohm, if it descends, then it is <50 Ohm !

Then, the Sxx plot is fine-tuned or optimized with 'o_Sxx_Z0conn_L_C'.



Note
the resonance in the Sxx stems from the connectors.

DUTs-Setups | Circuit | Model Parameters | Model Variables | Macros | Model GUI Items

Name	Value
___modeling_the_stripline____	
alpha	14.67u
Z0	46.26
Lopen	10m
beta	4.125n
___modeling_the_connector________	
L	392.6p
C	271.9f
Z0conn	54.62
TDconn	38p

Note
all modeling parameters for this step are Model Variables: stripline: Lopen, alpha, beta, Z0, connector: L, C, Z0conn, TDconn
(because they will be reused in the modeling of the OPEN, SHORT and LOAD)
but the Ldelta microstrip parameters alphathru, betathru and Z0thru are local Setup Variables.

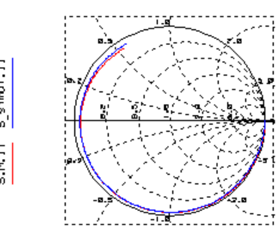
To finish with the THRU, we enter its calkit data into the table:

STANDARD No.	TYPE	C0 fF / L0 pH	C1/ E-27F/Hz / L1/ E-24H/Hz	C2 E-36F/Hz² / L2/ E-33H/Hz²	C3/ E-45F/Hz³ / L3/ E-42H/Hz³	Fixed or Sliding	OFFSET DELAY psec	OFFSET Z0 Ohm	OFFSET LOSS Ohm/s	FREQ MIN	FREQ MAX	STND LABEL
1	SHORT	___					___	___	0	0	999	SHORT
2	OPEN	___	___	___	___		___	___	0	0	999	OPEN
3	LOAD					fixed	___	___	0	0	999	LOAD
4	THRU						0	50	0	0	999	THRU

CUSTOM OPEN CALKIT: Measurement and SPICE Model

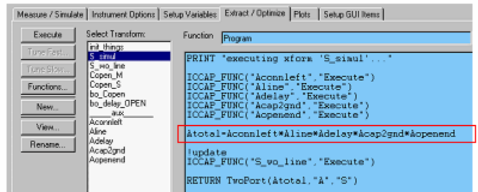
Now that we know the model of the microstrip line on the calibration substrate, as well as the connector model, we can measure the OPEN standard, and de-embed the left connector and the microstrip line from the measurement. After that, we will obtain the measurement of the open ended line.
In the IC-CAP file, this is done in Setup specify_open/OPEN.

MEASUREMENT OF THE CUSTOM OPEN STANDARD



The following IC-CAP screen dump shows the modeling programs:

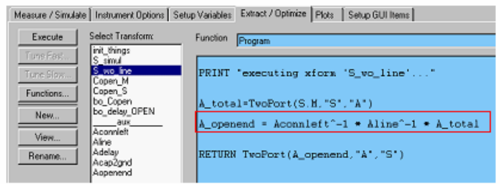
MODELING PROGRAMS FOR THE OPEN STANDARD



The simulation program 'S_simul' consists of a multiplication of A matrices of the left connector, the microstrip of the line (with alpha, beta and Z0 as modeled before in the THRU standard) and a conversion into S-parameters.

For the modeling, we need to de-embed the measurement from the 3.5mm connector and the microstrip line.
This is done in Program 'S_wo_line', see the next slide:

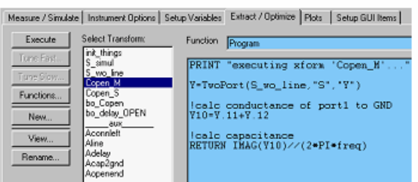
DE-EMBEDDING THE OPEN END
FROM THE CONNECTOR AND THE MICROSTRIP LINE



These de-embedded data are converted into a corresponding OPEN capacitor to Ground,

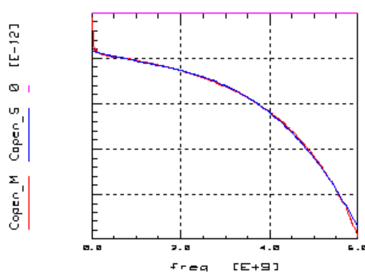
with the following PEL Program:

CALCULATING THE CAPACITANCE OF THE OPEN END



The next plot shows this OPEN capacitance, and its frequency-dependency.

MODELING THE FREQUENCY-DEPENDENT OPEN CAPACITORS



Both, the converted measurement data as well as the fitted curve with

$C_OPEN = C0 + C1 * freq + C2 * freq^2 + C3 * freq^3$

With these capacitance values, we check the fitting in the S-parameters, and obtain:

FINE-TUNING THE DELAY OF THE OPEN STANDARD

IN THE TOTAL MEASUREMENT INCLUDING CONNECTOR AND MICROSTRIP LINE

A little fine-tuning of the OPEN delay fits the S11 data very well, as already shown in the first figure of this chapter.

So, the obtained parameters for the OPEN cal. standard are:

OPEN STANDARD CALKIT DATA

Name	Value
Copen0	-430.0f
Copen1	-1.04459E-022
Copen2	6.60873E-033
Copen3	-4.28848E-042
delay_OPEN	400.0f
Z0_OPEN	50

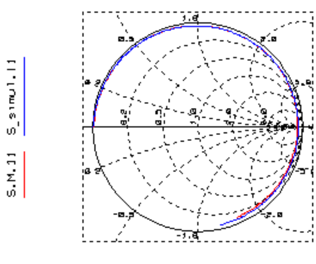
After all, we enter its calkit data into the table:

STANDARD No.	TYPE	C0 fF / L0 pH	C1/ E-27F/Hz / L1/ E-24H/Hz	C2 E-36F/Hz² / L2/ E-33H/Hz²	C3/ E-45F/Hz³ / L3/ E-42H/Hz³	Fixed or Sliding	OFFSET DELAY psec	Z0 Ohm	LOSS Ohm/s	FREQ MIN	MAX	STND LABEL
1	SHORT	___					___	___	0	0	999	SHORT
2	OPEN	**-430**	**-104459**	**6608.7**	**-4288.5**		**0.4**	**50**	0	0	999	OPEN
3	LOAD					fixed	___	___	0	0	999	LOAD
4	THRU						0	50	0	0	999	THRU

Custom SHORT calkit: Measurement and SPICE Model

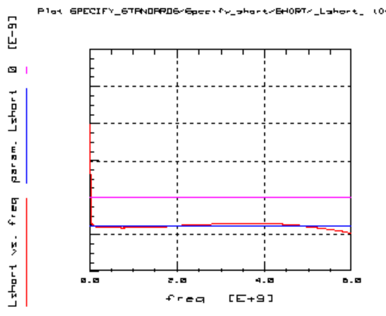
In the same way as with the OPEN standard, we proceed with the SHORT.

MEASUREMENT OF THE CUSTOM SHORT STANDARD



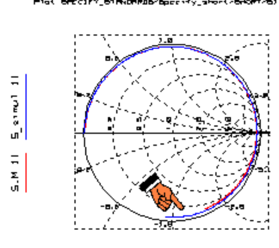
The measured data are again de-embedded from the 3.5mm connector and the microstrip line, in the same way as above for the OPEN.
Then, the shorted-end data are converted into a shunt inductance, what is depicted below:

VISUALIZED SHORTED END AND FITTED PARAMETER



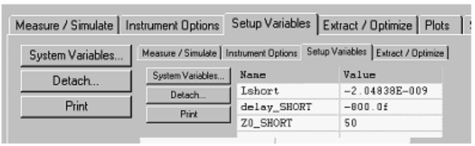
The inductance shows basically no frequency drift, and so the Cal. standard parameter L0 can be derived directly. A little fine-tuning of this inductor together with an additional delay line is performed on the data of the total, non-deembedded data in order to take the electrical difference of the structure with the geometrical structure into account.

FINE-TUNING THE DELAY OF THE SHORT STANDARD
IN THE TOTAL MEASUREMENT INCLUDING CONNECTOR AND MICROSTRIP LINE



So, we obtain these calkit data:

SHORT STANDARD CALKIT DATA



Name	Value
Lshort	-2.04838E-009
delay_SHORT	-800.0f
Z0_SHORT	50

which we enter again into the standards table:

STANDARD No.	TYPE	C0 fF / L0 pH	C1/ E-27F/Hz / L1/ E-24H/Hz	C2 E-36F/Hz² / L2/ E-33H/Hz²	C3/ E-45F/Hz³ / L3/ E-42H/Hz³	Fixed or Sliding	OFFSET DELAY psec	OFFSET Z0 Ohm	OFFSET LOSS Ohm/s	FREQ MIN	FREQ MAX	STND LABEL
1	SHORT	**-2049**					**-0.8**	**50**	0	0	999	SHORT
2	OPEN	-430	-104459	6608.7	-4288.5		0.4	50	0	0	999	OPEN
3	LOAD					fixed	___	___	0	0	999	LOAD
4	THRU						0	50	0	0	999	THRU

Custom LOAD calkit: Measurement and SPICE Model

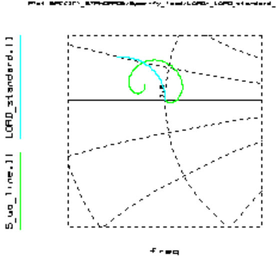
In the same way as with the OPEN and SHORT standards, we proceed with the LOAD.

MEASUREMENT OF THE CUSTOM LOAD STANDARD

The data of the total measurement, as shown above, are de-embedded from the 3.5mm connector and the microstrip line. The modeling goal is now a Z0 = 50 Ohm shunt resistor, in series with again a delay line. Since the ohmic 50 Ohm resistor will exhibit a frequency dependence, this delay line can help modeling this effect.
Also, this resistor should be a real high-frequency resistor, and as mentioned in the application note from above, it is preferable to have 2* 50 Ohm resistors in parallel to reduce their inductance.

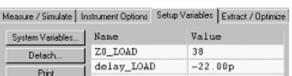
The data after de-embedding look like this:

THE DE-EMBEDDED CUSTOM LOAD STANDARD



In this case, due some non-idealities of the standard, we can only make a compromise. THis is shown as LOAD_Standard data in the same plot above.
These data are:

LOAD STANDARD CALKIT DATA



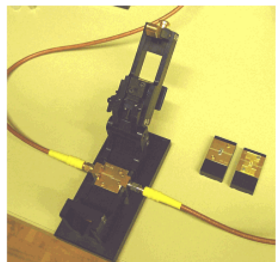
We enter its calkit data into the table and obtain the final calkit table result:

STANDARD No.	TYPE	C0 fF / L0 pH	C1/ E-27F/Hz / L1/ E-24H/Hz	C2 E-36F/Hz² / L2/ E-33H/Hz²	C3/ E-45F/Hz³ / L3/ E-42H/Hz³	Fixed or Sliding	OFFSET DELAY psec	Z0 Ohm	LOSS Ohm/s	FREQ MIN	MAX	STND LABEL
1	SHORT	-1531					-0.8	50	0	0	999	SHORT
2	OPEN	-430	-104459	6608.7	-4288.5		0.4	50	0	0	999	OPEN
3	LOAD					fixed	**-22**	**38**	0	0	999	LOAD
4	THRU						0	50	0	0	999	THRU

NEXT AND FINAL STEP: VERIFICATION OF THE CALKIT

A new NWA calibration is performed with a modified calkit, and the data from the table above.
After that, all cal standards are re-measured and compared to a simulation of SPICE circuits, which consist of the components of the calkit table from above. The calibration was successful only if
ALL 4 RE-MEASUREMENTS MATCH THE SIMULATIONS EXTREMELY WELL !!

A realized custom Calkit for packaged devices



Introduction to NWA Calibration

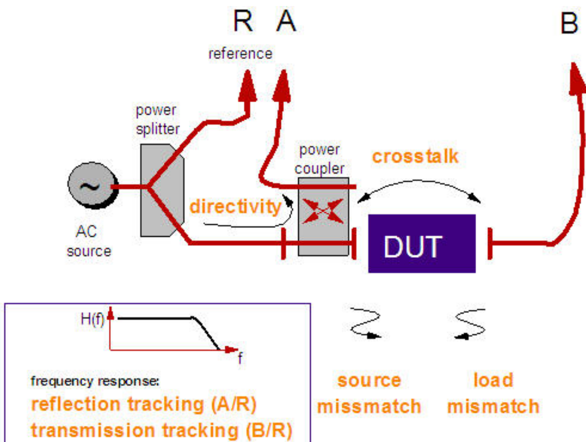
VNA Standard Calibration Techniques And Verification

A considerable challenge in S-parameter VNA measurements is to define exactly where

the measurement system ends and the DUT begins. This location is called 'reference plane'. This means, all error contributions, inside the VNA and also from the cables up to this reference plane, have to be calibrated out.

The calibration of a NWA is performed by rather complex procedures. Such are Short-Open-Load-Through (SOLT), Through-Reflection-Load (TRL) or Load-Refection-Match (LRM) and the associated error correction model.

Figure 1: The 12 error contributions for S-parameter measurements with a VNA



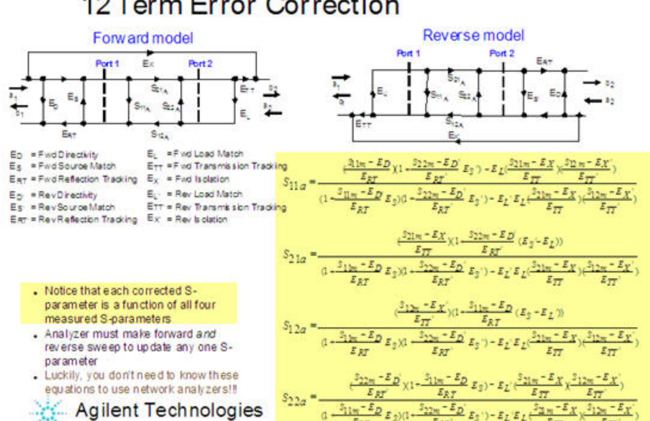
As can be seen in fig.1, there are 6 error contribution terms in forward direction:

- **Directivity**: cross-talk of the power coupler
- **Crosstalk**: cross-talk inside the S-parameter test set, overlying the DUT
- **Source Mismatch**: multiple reflections due to non-Z0 input and output impedance of the cables and connectors
- **Load Mismatch**: the same for the opposite port
- **Reflection Tracking A/R**: frequency dependence of signal path R->A
- **Transmission Tracking A/R**: same for signal path R->B

For the reverse calibration, another 6 error terms add up to a total of 12 terms. These 12 terms are often referred as the 12-term error correction.

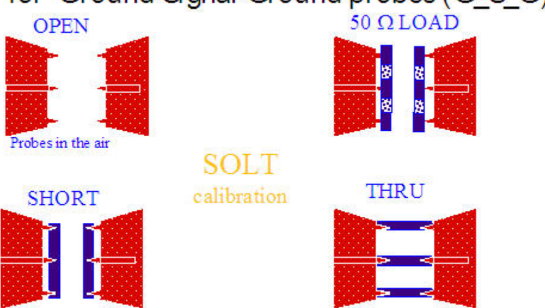
For the different calibration procedures, specific, accurately known standards have to be measured. For connectorized DUTs, with e.g. 3.5mm connectors, usually 4 calibration standards like OPEN, LOAD, SHORT and THRU are applied, with exactly known standard inaccuracies.
It is interesting to note that each corrected S-Parameter is a function of all other (uncorrected) measured S-Parameters, see fig.2 below.

Figure 2: The 12-term error correction formula



For the case of on-wafer measurements, SOLT calibration, and using Ground-Signal-Ground (GSG) coplanar probes, fig. 3 depicts the corresponding test structures, which are usually available on a RF-high-performance ceramic substrate (ISS substrate).

Figure 3: SOLT calibration structures for on-wafer measurements



Some notes about these calibration standards:

while in the case of the CV meter, the OPEN calibration corrects for an ideal offset capacitor, a NWA calibration uses cal standards (OPEN, SHORT, LOAD, THRU etc.) from a cal kit.

These cal standards do not represent ideal standards. They represent the real, existing standard, including its non-idealizations! It means that a SHORT is not an ideal SHORT, but instead represents rather an inductance. The same applies to the THRU, which has a non-ideal delay time. The OPEN corresponds rather to a capacitor than to an ideal OPEN. Therefore, these non-idealizations of the G-S-G probes have to be entered into the NWA before calibration. This is called 'modifying the cal kit'.

While this procedure refers to the non-idealizations of the calibration standards, the subsequent calibration is related to the selected frequency range, the RF power, the averaging of the NWA etc. After it has been performed, the correction terms are stored in the cal set of the NWA. In other words, the 12-term error vectors are 'filled up'.

Afterward when the measurement is performed, the raw measured data arrays will be corrected using a correction technique related to the selected calibration method, and referring to the specified cal set. Finally, this corrected measurement result is transferred into IC-CAP and displayed there (as well as on the NWA monitor).

Therefore, after the calibration, a re-measurement of the OPEN will not represent an ideal open, but instead exactly those parasitic components as described in the documentation of the OPEN. In the same way, a THROUGH shows up after calibration with its real delay time, and a SHORT represents its inductive behavior!

This re-measuring of the standards together with a simulation of their known values can be used to verify the VNA calibration, and this can be done easily, using IC-CAP.

We simply re-measure the cal.standards, e.g. the OPEN, the SHORT, the THRU and the LOAD. As said above, this measurement will correspond to the non-idealizations of the selected cal.standard. In case of Cascade probes, the OPEN, for example, behaves like a negative capacitance of roughly -9fF . Now, after this measurement has been made, we can define a test circuit for that setup in IC-CAP, entering the netlist of the calibration standards. Using SPICE3, a simulator which also permits negative capacitance, we can simulate the expected behavior of the OPEN probes. If the calibration was executed correctly, there is an excellent match between measured and simulated curves.
In a next step, we measure the SHORT, define in IC-CAP the SHORT nonidealities in a SPICE circuit, and simulate. Again, an excellent match between measured SHORT data and simulations has to be achieved. We then continue with the THRU and LOAD measurements and simulations. Only if all 4 standards exhibits an excellent fit, we can assume a correct calibration of the NWA.

Note
This calibration verification can also be applied to check the quality of an older calibration.

NWA Calibration With Gating in the Time Domain

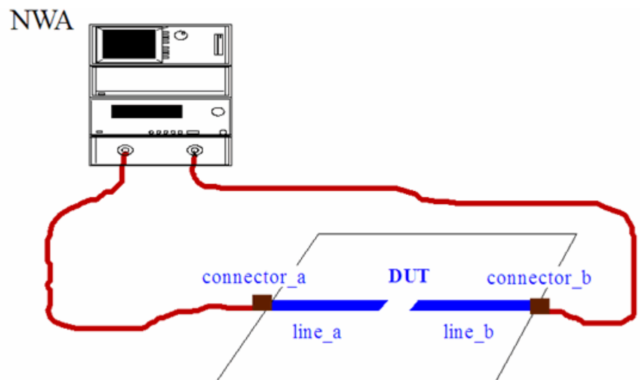
This chapter refers to a de-embedding technique for strip-line test fixtures as described by /Gronau/. Based on this publication, see also /Böhm/ and /Stassen/.

It is assumed for this type of de-embedding that the vector network analyzer (VNA or NWA) has been calibrated already up to the ends of its connection cables. These cables are then connected to a test fixture. The idea is that the DUT is measured on a ceramic substrate (or similar) and is connected by strip lines on this test fixture. If we can correct the VNA calibration for the connector and the strip lines at both sides of the DUT, we would be able to shift our calibration plane to every location on the strip lines. And therefore we can de-embed the effects of the test fixture until the physical limits of the DUT. This is a very important feature and allows to measure devices like dummy packages, passive components like resistors, capacitors, spiral inductors and vias alone

without being affected by parasitic effects from the test fixture.

The measurement setup for the DUT and the test fixture is given in following figure.

The measurement environment for the DUT



The basic idea of this special calibration technique is to describe the effects of the connectors A and B as well as those of the lines A and B with A-matrices each and to de-embed the DUT later with A-matrix manipulations.
Therefore, the A-matrix of the total VNA measurement is:

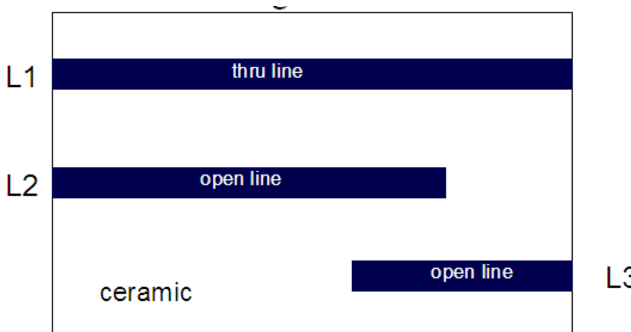
ATOTAL = Aconn_a * Aline_a * ADUT * Aline_b * Aconn_b (1)

Once the matrices Aconn_x and Aline_x are know, we can solve for ADUT :

ADUT = Aline_a-1 * Aconn_a-1 * ATOTAL * Aconn_b-1 * Aline_b-1 (2)

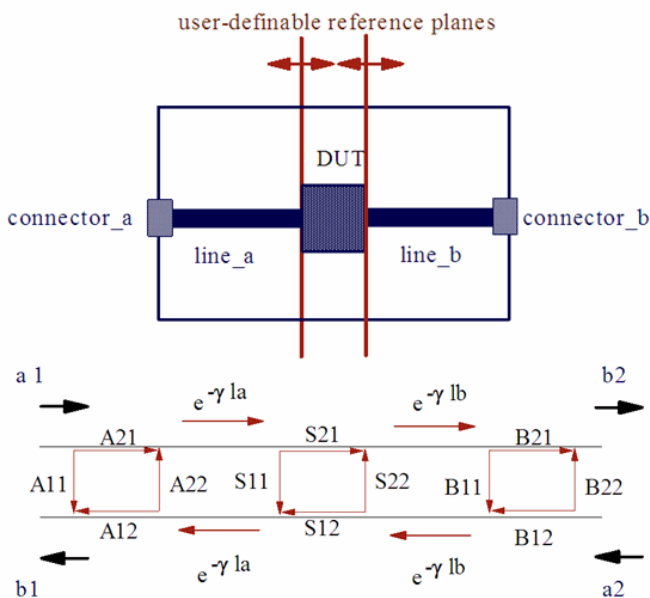
The problem is now to determine the A-matrices of the strip lines and the connectors. This is done with measurements on an auxiliary calibration substrate that holds a THROUGH line and two OPEN strip lines with different length.

Calibration Substrate



The VNA measurements are Fourier transformed to the time domain in the VNA, a gating function is applied to select either the connector reflection or the end reflection, and finally this gated time domain reflectogram is transformed back to the frequency domain. The resulting S-parameters are then transferred back to IC-CAP. These auxiliary measurement results are stored in the setups of the model file and the de-embedding calculations will refer to these data later. That means we will perform a "software calibration" on the later measurement data. In other words, we will obtain partly calibrated data (until the ends of the VNA cables) from the device under test (DUT), and then perform matrix manipulations on this data using the previously measured calibration data. After these manipulations within IC-CAP, we will obtain the S-parameters of the DUT alone and convert them to A-parameter to match equ.(2).
The following figure gives an overview of this concept of user-definable reference planes and the signal flow chart.

Test fixture with user-definable reference planes and the signal flow chart for the fixture.



In order to be able to set a correct gating, i.e. a high time domain resolution, it is necessary to sweep the VNA frequency as high as possible. In the file 'cal_w_td_gating.mdl' , the upper frequency is 26.5GHz. Therefore, an HP8510 or similar with time domain option is recommended. Also, macro "run_cal" (that performs all required measurements of the setups Sxxx) is using HP-IB commands for the HP 8510.

The following figure gives the user interface of the IC-CAP file 'cal_w_td_gating.mdl' :

```
 ____________________________________
|      MACRO                         |
|                                    |
| RUN_DUT_MEAS                       |
| RUN_ALL                            |
| define_things                      |
| run_cal                            |
| run_deemb                          |
| README_equations                   |
| README_theory                      |
| Literature                         |
|                                    |
|                                    |
|____________________________________|

 ____________________________________
|__________DUT-SETUP |
|   ________________________________ |
|  |   cal          S11A             |
|  |                S11B             |
|  |                S22B             |
|  |                S22A             |
|  |                S21L1            |
|  |                S11L2            |
|  |                S11L3            |
|  |                S22L2            |
|  |                S22L3            |
|  |                Calculate        |
|  |                m_thru_line      |
|  |________________README___________|
|   ________________________________ |
|  |____meas_DUT    meas_DUT         |
|  |________________README___________|
|____________________________________|
```

The IC-CAP model file includes:

- setups 'cal / Sxxx' are used to keep the gated S-parameters
- setup 'cal / Calculate' contains all the matrix manipulations
- setup 'cal / m_thru_line' holds the un-gated measurement of L1

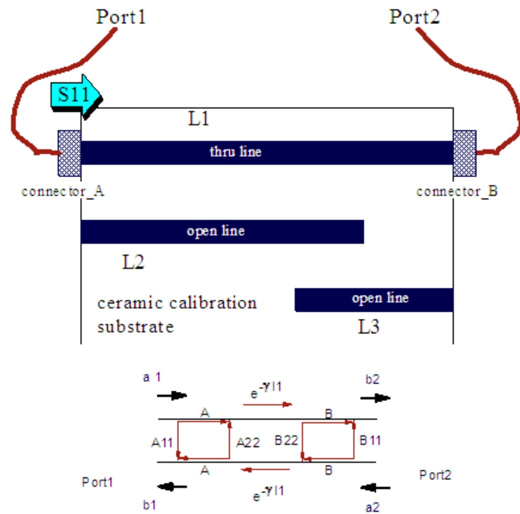
- setup 'meas_DUT / meas_DUT' contains the measurement of the DUT and the test fixture, its transforms are linked to setup 'cal / Calculate' in order to perform the de-embedding.

The individual steps of the procedure are:
The S-matrices of the connectors A and B are called matrix A and matrix B and the strip lines A and B are described with the propagation coefficient γ , i.e. exp(- γ * length).

It is further assumed that
A12 = A21 = A (3)
and B12 = B21 = B (4)

- To begin with, the VNA has to be calibrated to the end of its connection cables, and the calibration has to be stored in a VNA cal set.
  Then we connect the calibration substrate THRU line L1 and have the following situation:

Determination of S11 of connectors A and B with time-domain gating.



> **Note**
> Using IC-CAP, the VNA is not re-setted totally before every measurement. Therefore, we can perform a 'classical' S-parameter measurement from IC-CAP, set the VNA to local afterwards, transform the S-parameters in the VNA to time domain, and set the gating. If we then re-perform a measurement from IC-CAP, the S-parameters transmitted to IC-CAP will reflect the gated S-parameters.

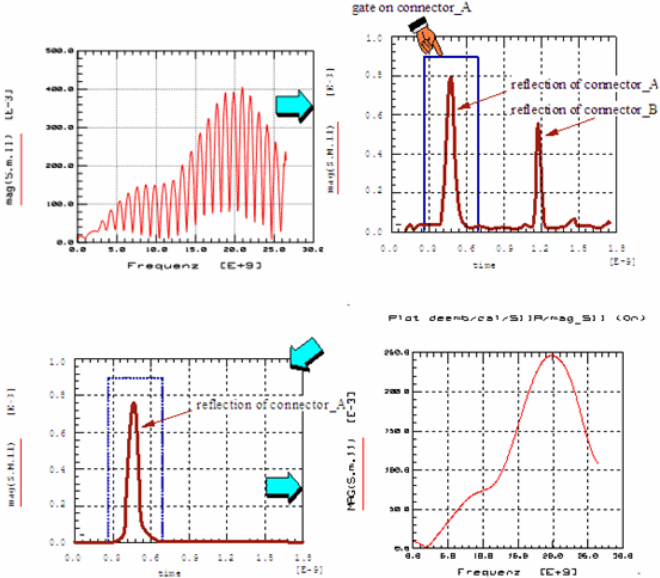
Next,

- The transmission line of length L1 is measured. L1 has to be large enough, such that the individual reflections of connector A and B do not overlap in the time domain.
  i.e.: we set the VNA to local, perform a measurement, display S11, transform it to the time domain and set the gate on the left reflection (connector A). Then, we re-perform a measurement in IC-CAP. After that, we hold in
  - SETUP S11A the gated S-parameter measurement S11 of connector A.

S11A = A11 (5)

This is depicted in the following figure:

The gating of S11 for the separation of S11A, the S11 parameter of connector A



After that, we have to

- Set the VNA to local, and position the gate on the right reflection (connector B). We re-perform a measurement in IC-CAP and hold in

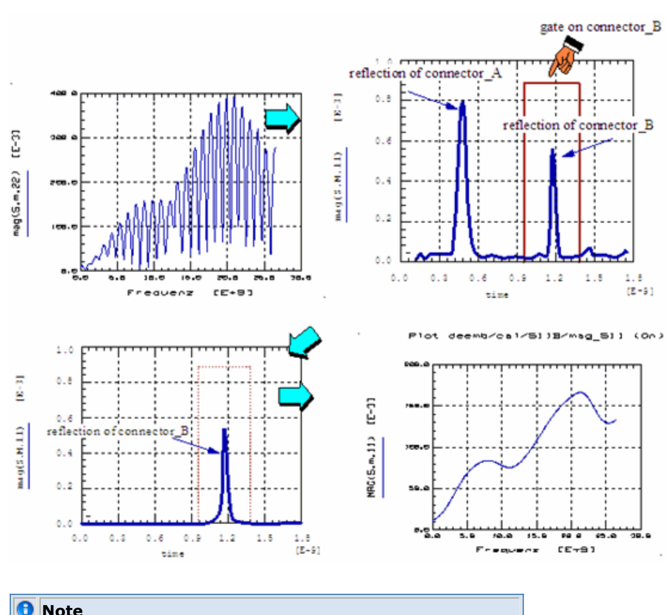
- SETUP S11B the gated S-parameter measurement S11 of connector B, what is mathematically:

$S11B = A21*\exp(- \gamma *L1)*B22*\exp(- \gamma *L1)*A12$

$= A^2*B22*\exp(-2* \gamma *L1)$ (6)

with respect to the signal flow chart.
See the following figure for details.

The gating of S11 for the separation of S11B, the S11 parameter of connector B



> **Note**
> index 1 refers always to port 1, index 2 to port 2 of the VNA.

Now we concentrate on S22:

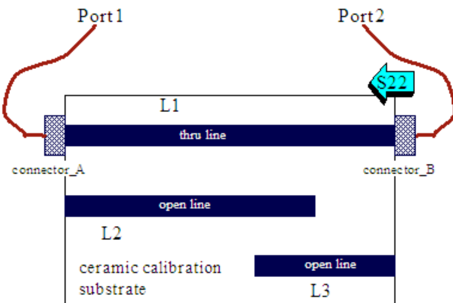
- We set the VNA to local, measure and display S22, transform it to the time domain and set a gate on the left reflection (connector B, the first reflection to be seen in time domain for S22).
  Then, we re-perform a measurement in IC-CAP. After that, we hold in

- SETUP S22B the gated S-parameter measurement S22 of connector B.

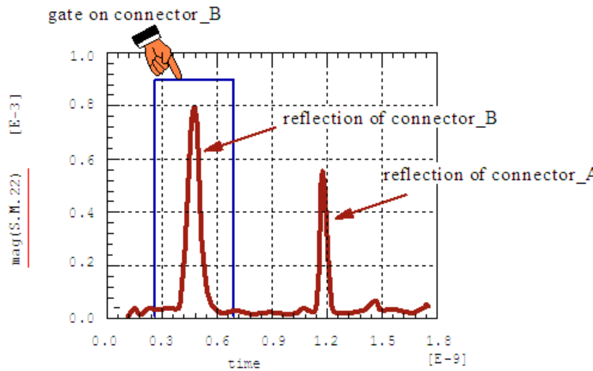
$S22B = B11$ (7)

The following two figures give the details.

The measurement of S22A and S22B with time domain gating.



The gating of S22 for the separation of S22B, the S22 parameter of connector B.



Then,

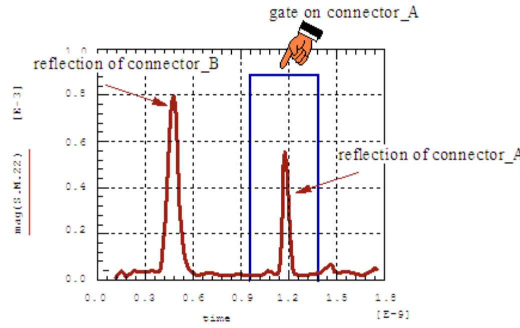
- Measuring again L1 with the VNA set to local, we now position the gate on the right reflection (connector A). We re-perform a measurement in IC-CAP and hold in

- SETUP S22A the gated S-parameter measurement S22 of connector A, what is mathematically:

S22A = B^2*A22*exp(-2* γ *L1) (8)

The following figure shows this procedure:

The gating of S22 for the separation of S22A, the S22 parameter of connector A.



Lastly,

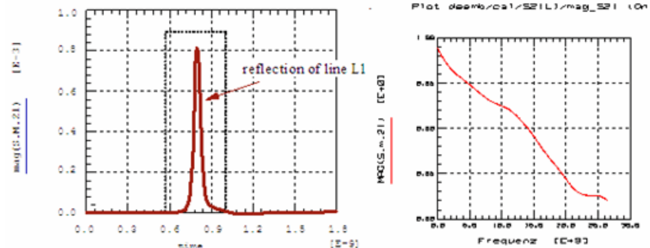
- we position the gate on the strip line in the time-domain equivalent of S21 and yield the measurement result of

- SETUP S21L1,

S21L1 = S12L1 = A*B*exp(- γ *L1) (9)

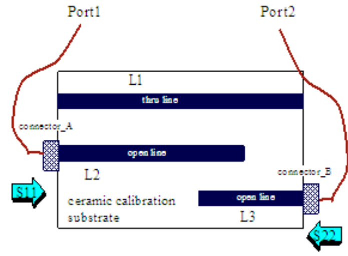
This is shown in following figure:

The gating of S21 and the 'gated S-parameter' S21L1.



After having measured the THRU line L1, we do not know yet all parameters that are required for the de-embedding procedure. We still need to know the value of parameter γ . And that is why we have to perform some more measurements on lines L2 and L3, which have each an open end.
Port1 of the VNA is therefore connected with L2 and port2 with L3.

Measurement setup for the first measurement of lines L2 and L3



We

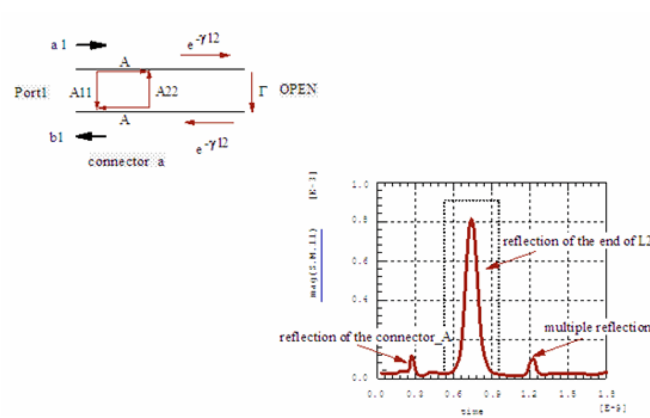
- measure S11 and set a gate on the OPEN reflection of line L2 at port1, and transform back into the frequency domain. After having performed a measurement in IC-CAP,

- SETUP S11L2 contains

S11L2 = A^2* γ **exp(-2** γ *L2) (10)

This is shown in the following figure:

Signal flow chart for line L2 and the gating in the time domain to separate the OPEN of S11L2



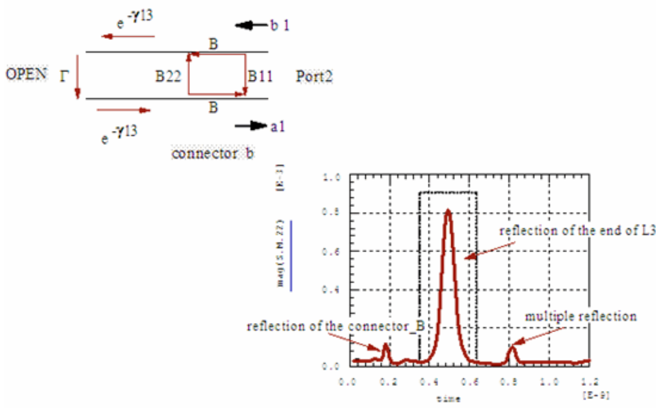
- we do the same gating on the OPEN reflection of line L3 at port2, and transform back into the frequency domain. The following IC-CAP measurement yields

- SETUP S22L3, which contains

S22L3 = B^2* γ **exp(-2** γ *L3) (11)

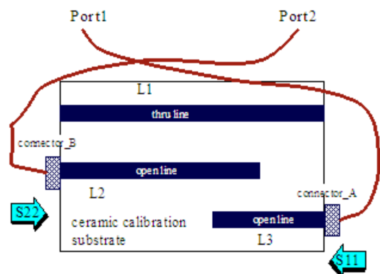
This is depicted in following figure:

Signal flow chart for line L3 and the gating in the time domain to separate the OPEN of S22L3



Next, we have to flip the connections of the calibration substrate. Now, port1 of the VNA is connected with L3 and port2 with L2.

Measurement setup for the second (flipped) measurement of lines L2 and L3



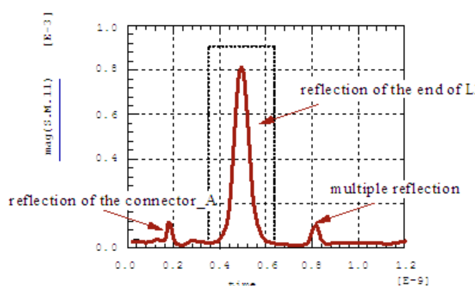
We re-measure with the VNA set to local, then

- set a gate on the OPEN reflection of line L3 at port1, and transform back into the frequency domain. Thus, after having measured with IC-CAP,
- SETUP S11L3 contains

S11L3 = A^2* γ **exp(-2** γ *L3) (12)

See the following figure:

The gating in the time domain to separate the OPEN of S11L3



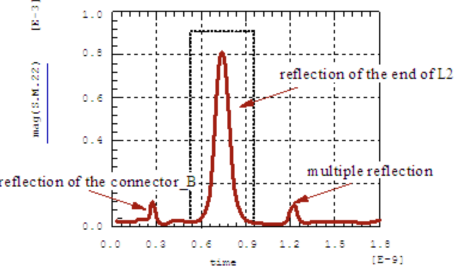
again, we do

- the same gating on the OPEN reflection of line L2 at port2, and transform back into the frequency domain. Thus,
- SETUP S22L2 contains

S22L2=B^2* γ **exp(-2** γ *L2) (13)

given in the following figure:

The gating in the time domain to separate the OPEN of S22L2



After all, we are ready to calculate all unknown S-parameter values:

First, the propagation coefficient of the lines on the substrate is with the assumption of S11L3 / S11L2 = S22L3 / S22L2 (see literature /Gronau 1992/):

$$\gamma = \frac{1}{2*(L2-L3)} * \ln\left(\frac{1}{2} * \left(\frac{S11L3}{S11L2} * \frac{S22L3}{S22L2}\right)\right) \qquad (14)$$

with the log of complex numbers:

$$\underline{z} = r\,e^{j\varphi} \qquad \ln(\underline{z}) = \ln(r) + j\varphi \quad \text{for } -\Pi < \varphi < \Pi$$

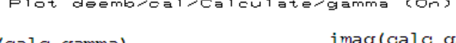
In IC-CAP, the PEL program looks like:

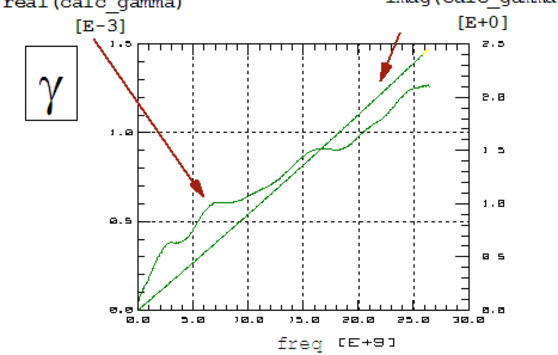
```
tmp1 = MAG(cal/S11L3/S.m.11 / cal/S11L2/S.m.11)
tmp2 = MAG(cal/S22L3/S.m.22 / cal/S22L2/S.m.22)
tmp3 = .5*(tmp1+tmp2)
tmp4 = REAL(cal/S11L3/phs.m.11) - REAL(cal/S11L2/phs.m.11)
tmp5 = REAL(cal/S22L3/phs.m.22) - REAL(cal/S22L2/phs.m.22)
tmp6 = .5*(tmp4+tmp5)

calc_gamma  = 1 / (2 * (L_2 - L_3)) * LOG(tmp3) + j * tmp6
```

where the programs 'phs' contain phase correction algorithms for a continuous phase shift without 2PI steps.

The calculated propagation coefficient γ

Plot deemb/cal/Calculate/gamma (On)





The reflection coefficient is:

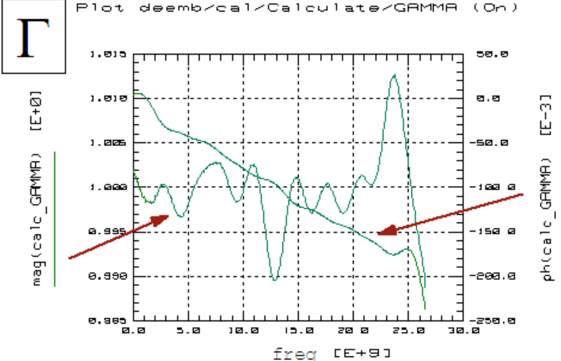
$$\Gamma = \frac{\text{sqrt(S11L2 * S22L2)}}{\text{S21L1}} * \exp(\gamma * (2*L2 - L1)) \quad (15)$$

And in IC-CAP, the PEL program looks like:

```
calc_GAMMA = sqrt(abs(cal/S11L2/S.m.11 * cal/S22L2/S.m.22))
         * exp(j*0.5*(real(cal/S11L2/phs.m.11) + real(cal/S22L2/phs.m.22)))
         / cal/S21L1/S.m.21
         * EXP(calc_gamma*(2*L_2 - L_1))
```

where 'phs' contains again the phase correction algorithms for a continuous phase shift without 2PI steps.

The calculated reflection coefficient Γ



while,

$$A21 = A12 = A = \text{sqrt}\left(\frac{S11L2}{\Gamma}\right) * \exp(\gamma * L2) \quad (16)$$

in IC-CAP:

```
calc_A = SQRT(ABS(cal/S11L2/S.m.11 / calc_GAMMA))
       * EXP(j*0.5*(real(cal/S11L2/phs.m.11)- ph(calc_GAMMA)))
       * EXP(calc_gamma * L_2)
```

and

$$B21 = B12 = B = \text{sqrt}\left(\frac{S22L2}{\Gamma}\right) * \exp(\gamma * L2) \quad (17)$$

in IC-CAP again:

```
calc_B = SQRT(ABS(cal/S22L2/S.m.22 / calc_GAMMA))
       * EXP(j*0.5*(real(cal/S22L2/phs.m.22)- ph(calc_GAMMA)))
       * EXP(calc_gamma * L_2)
```

and

$$A22 = \frac{S22A}{B^2} * \exp(2 * \gamma * L1) \quad (18)$$

implemented in PEL:

```
calc_A22 = cal/S22A/S.m.22 / calc_B^2 * EXP(2*calc_gamma*L_1)
```

and

```
            S11B
B22 =   ------ * exp(2* γ *L1)
            A^2
```
(19)

what looks in IC-CAP like:

```
calc_B22 = ( cal/S11B/S.m.11 / calc_A^2 ) * EXP(2*calc_gamma*L_1)
```

what looks in IC-CAP like:

```
calc_B22 = ( cal/S11B/S.m.11 / calc_A^2 ) * EXP(2*calc_gamma*L_1)
```


Now we are done and can calculate the S matrices of the connectors and the strip lines:

```
S_conn_a.11[i] = cal/S11A/S.m.11[i]
        .12[i] = calc_A[i]
        .21[i] = calc_A[i]
        .22[i] = calc_A22[i]

S_conn_b.11[i] = calc_B22[i]
        .12[i] = calc_B[i]
        .21[i] = calc_B[i]
        .22[i] = cal/S22B/S.m.22[i]

S_line_a.11[i] = 0
        .12[i] = EXP(-calc_gamma[i]* L_P1
        .21[i] = EXP(-calc_gamma[i]* L_P1)
        .22[i] = 0

S_line_b.11[i] = 0
        .12[i] = EXP(-calc_gamma[i]* L_P2)
        .21[i] = EXP(-calc_gamma[i]* L_P2)
        .22[i] = 0
```
(20)

These S-matrices are converted to A-matrices using the IC-CAP TwoPort function.
Finally, these A matrices are manipulated following equation (2), what looks like:

```
tmp1 = (TwoPort(S_lineP1.m,"S","A"))^-1
tmp2 = (TwoPort(S_conn1.m,"S","A"))^-1
tmp3 =  TwoPort(S.m,"S","A")
tmp4 = (TwoPort(S_conn2.m,"S","A"))^-1
tmp5 = (TwoPort(S_lineP2.m,"S","A"))^-1
tmp  = tmp1 * tmp2 * tmp3 * tmp4 * tmp5
RETURN TwoPort(tmp.m,"A","S")
```

with
A_total = A_connA * A_LA * A_DUT * A_LB * AconnB (see 1)

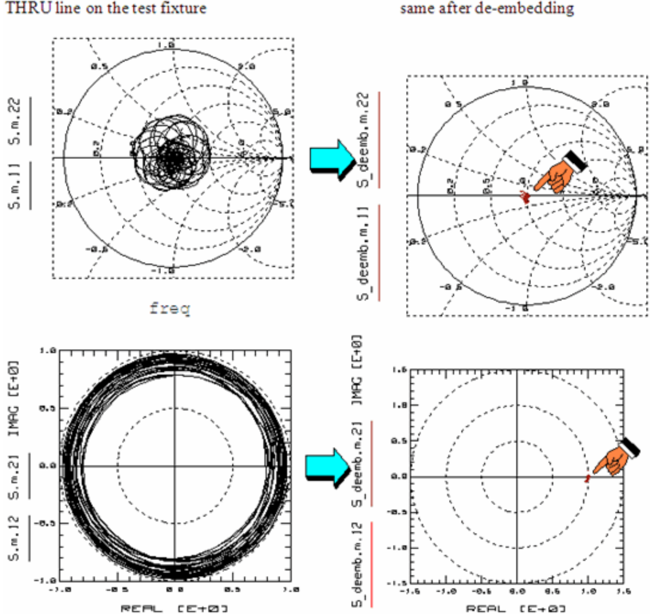
or solved for A_DUT:
-1 -1 -1 -1
A_DUT = A_LA * A_connA * A_total * A_connB * A_LB (see 2)

It should be noted, that a nice feature of this method is to being able to set the calibration plane to every location on the strip line on the ceramic substrate. This means that once the calibration is done, the performance of the DUT alone can be calculated by simply setting the strip line length variables to the length of interest. This feature is especially important for devices like SMT devices, vias etc.

Calibration Verification

In order to check the quality of the calibration, we reconnect LINE_1 of the calibration substrate, set the lengths L_P1 and L_P2 (length of the strip lines of port1 and port2) to 0.5*L1 and 0.5*L2, perform a measurement of strip line L1 and apply the equations (20) to the measurement results. In this case, the calibrated or better de-embedded result should correspond to an ideal THRU: Sxx=0 and Sxy=1 (all power transmitted). Fig.20 shows the result of such a verification measurement for frequencies from 45MHz to 26.5GHz. It can be seen that the procedure works very well.

Verification of the calibration quality



> **Note**
> It is mandatory for this calibration method to exactly determine the calibration substrate strip line lengths up to 1/10 mm.

Publications

G.Gronau, Scattering Parameter Measurement of Microstrip Devices, Microwave Journal, Nov.1992.

G.Gronau, I.Wolff,: A Simple Broad-Band Device De-embedding Method Using an Automatic Network Analyzer with Time-Domain Option, IEEE Trans.on Microwave Theory and Tech., Vol. 37, No.3, March 1989.

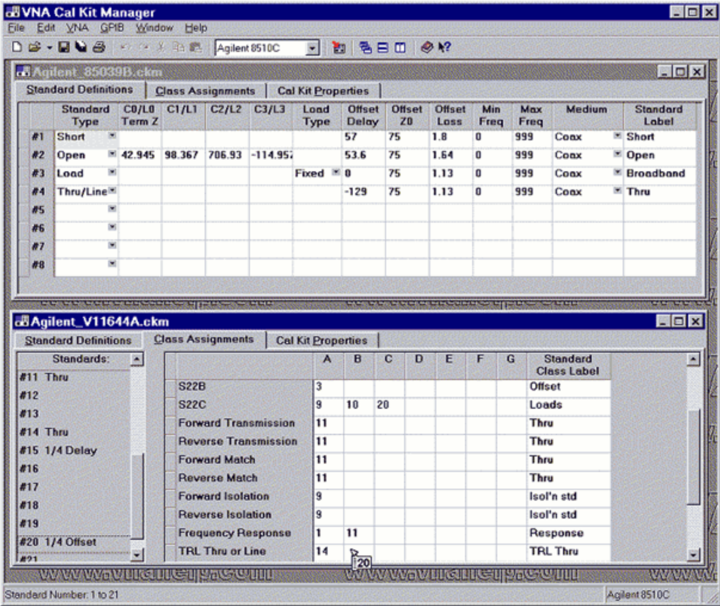
M.Böhm, Diploma Thesis Technical University Ilmenau/Germany: Meßtechnische und rechnerische Untersuchungen von Verbindungsstrukturen für elektronische Schaltungen in einem Frequenzbereich von 45MHZ bis 25GHz, 30.4.1993.

Rudolf Stassen, 'Einsatz eines Mikrowellen-SPitzenmeßplatzes zur Charakterisierung von Transistore direkt auf dem Wafer im Frequenzbereich von 0,045 bis 26,5GHz',
Diplomarbeit am Institut für Schicht- und Ionentechnik am Forschungszentrum Jülich GmbH,
available from Technische Informationsbibliothek, Hannover.

VNA Cal Kit Manager 2.0

Version 2.0 (983 kBytes) for Windows 95 / 98 / NT4 / 2000
from the internet site www.vnahelp.com

This PC-based software allows the easy entering of calkit data into a NWA, without hitting manually the many buttons especially on the elder NWAs.



Installation Hints:

The file setup.exe should be stored in any convenient spot on your hard drive. Then execute the file by double-clicking it in Windows Explorer - this will start the installation program. You can then follow the instructions on the screen.
VNA Cal Kit Manager is absolutely free, provided you abide by the license agreement.

S-Parameters

Contents

- *Definition of S-Parameters* (iccapmhb)
- *Understanding S-Parameter Plots* (iccapmhb)
- *Interpreting S-Parameter Plots* (iccapmhb)
- *S-Parameter Basics for Modeling Engineers* (iccapmhb)
- *S-Parameters Measurements for Modeling* (iccapmhb)
- *Calculating S-parameters from Voltages* (iccapmhb)
- *Small Signal Versus Large Signal S-Parameters* (iccapmhb)
- *Converting N-Port S-Parameters to M-Port* (iccapmhb)
- *Differential S-Parameters* (iccapmhb)
- *Amplifier Characteristics from S-Parameters* (iccapmhb)
- *Q-Factor Calculation* (iccapmhb)
- *S-Parameter Impedance Matching* (iccapmhb)
- *S-Parameter Utilities* (iccapmhb)

Note
Additional Information: Agilent Technologies Application Note AN-95
S-Parameter Techniqueshttp://contact.tm.agilent.com/Agilent/tmo/an-95-1/index.html

Amplifier Characteristics from S-Parameters

We can calculate more information about the 2-port network from the S-parameters. Here some examples. For more details see the publication list below, esp. Hewlett-Packard application note AN95-1.

Input Reflection Coefficient with arbitrary load impedance Z_L

$$S'_{11} = S_{11} + \frac{S_{12} * S_{21} * \Gamma_L}{1 - S_{22} * \Gamma_L} \qquad (1)$$

with the reflection coefficient:

$$\Gamma_L = \frac{a_L}{b_L} \quad \text{of the load impedance } Z_L$$

Note
For $Z_L = Z_0$, it is $S'_{11} = S11$

Output Reflection Coefficient with arbitrary source impedance Z_S

$$S'_{22} = S_{22} + \frac{S_{12} * S_{21} * \Gamma_S}{1 - S_{11} * \Gamma_S} \qquad (2)$$

with the reflection coefficient:

$$\Gamma_S = \frac{a_S}{b_S} \quad \text{of the source impedance } Z_S$$

Voltage Gain with arbitrary Z_S and Z_L

$$A_v = \frac{v_2}{v_1} = \frac{S_{21} * (1 + \Gamma_S)}{(1 - S_{22} * \Gamma_L) * (1 + S'_{11})} \quad \text{using } S'_{11} \text{ from (1)} \qquad (3)$$

Stability:

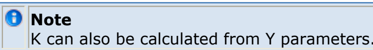
Rollett's stability factor, calculated from the S-parameters, is given by

$$K = \frac{1 + |\underline{D}|^2 - |\underline{S}_{11}|^2 - |\underline{S}_{22}|^2}{2 * |\underline{S}_{12} * \underline{S}_{21}|}$$

where D is the determinant, i.e.

$$\underline{D} = \underline{S}_{11} * \underline{S}_{22} - \underline{S}_{12} * \underline{S}_{21} \qquad (4)$$

Devices with a K-factor K > 1 and a non-negative determinant at a given frequency are said to be unconditionally stable at that frequency, meaning they will not oscillate due to load mismatch. In this case, ANY impedance connected to the ports of the two-port will NOT cause oscillation.
In general, transistors have K < 1 for certain frequencies, particularly lower frequencies, and are therefore only stable for certain impedance values.



Note
K can also be calculated from Y parameters.

$$K = \frac{2 \cdot REAL(Y11) \cdot REAL(Y22) - REAL(Y12 \cdot Y21)}{|Y12| \cdot |Y21|}$$

Power Gain from S-Parameters

With the abbreviations:

$$D = S_{11} * S_{22} - S_{12} * S_{21}$$
$$M = S_{11} - D * S'_{22}$$
$$N = S_{22} - D * S'_{11}$$

where S means conjugate complex S-Parameters, we can calculate the following gain characteristics:

Power Gain

$$\text{power gain} = \frac{\text{power delivered to load}}{\text{power input to network}}$$

$$G = \frac{|S_{21}|^2 \cdot \left(1-|\Gamma_L|^2\right)}{\left(1-|S_{11}|^2\right)+|\Gamma_L|^2 \cdot \left(|S_{22}|^2 - |D|^2\right) - 2 \cdot REAL(\Gamma_L \cdot N)} \qquad (5)$$

with the reflection coefficient:

$$\Gamma_L = \frac{a_L}{b_L} \quad \text{of the load impedance } Z_L$$

of the load impedance Z_L

Available Power Gain

$$\text{available power gain} = \frac{\text{power available from network}}{\text{power available from source}}$$

$$G_A = \frac{|S_{21}|^2 \cdot \left(1-|\Gamma_S|^2\right)}{\left(1-|S_{22}|^2\right)+|\Gamma_S|^2 \cdot \left(|S_{11}|^2 - |D|^2\right) - 2 \cdot REAL(\Gamma_S \cdot M)} \qquad (6)$$

with the reflection coefficient,

$$\Gamma_S = \frac{a_S}{b_S} \quad \text{of the source impedance } Z_S$$

of the source impedance Z_S

Transducer Power Gain

$$\text{transducer power gain} = \frac{\text{power delivered to load}}{\text{power available from source}}$$

$$G_T = \frac{|S_{21}|^2 \cdot \left(1-|\Gamma_S|^2\right) \cdot \left(1-|\Gamma_L|^2\right)}{\left|(1-S_{11} \cdot \Gamma_S) \cdot (1-S_{22} \cdot \Gamma_L) - S_{12} \cdot S_{21} \cdot \Gamma_L \cdot \Gamma_S\right|^2} \qquad (7)$$

in the case of unilateral transducer power gain, i.e. S12=0, this simplifies to:

$$G_{TU} = \frac{|S_{21}|^2 \cdot \left(1-|\Gamma_S|^2\right) \cdot \left(1-|\Gamma_L|^2\right)}{\left|(1-S_{11} \cdot \Gamma_S)\right|^2 \cdot \left|(1-S_{22} \cdot \Gamma_L)\right|^2} = G_S \cdot G_{DUT} \cdot G_L \qquad (8)$$

with

$$G_S = \frac{1-|\Gamma_S|^2}{\left|(1-S_{11} \cdot \Gamma_S)\right|^2} \quad \text{at the source} \qquad (9a)$$

$$G_{DUT} = |S_{21}|^2 \quad \text{power gain of the DUT} \qquad (9b)$$

$$G_L = \frac{1-|\Gamma_L|^2}{\left|(1-S_{22} \cdot \Gamma_L)\right|^2} \quad \text{at the load} \qquad (9c)$$

Maximum Available (stable) power amplification Gain

This maximum gain is obtained when the transistor or amplifier is inserted between matching load/source resistance. I.e. for conjugate complex impedance matching at both sides of the transistor, the amplifier gain is different, usually higher than for a Z0 = 50? environment..

The maximum available gain is:

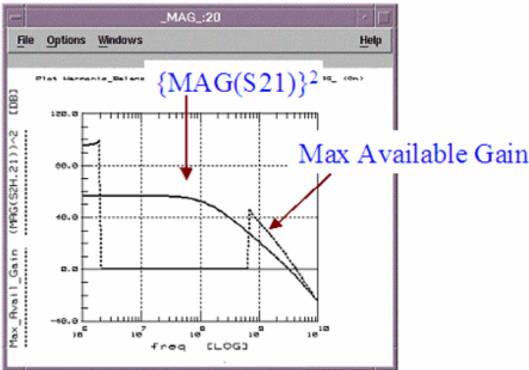
$$MAG = \left|\frac{S_{21}}{S_{12}}\right| \cdot \left(K - \sqrt{K^2 - 1}\right) \qquad (10)$$

with K from (4) and only valid if K>1

MAG calculates the gain of the entire network (DUT plus matching networks). There is still a Z0 termination outside the matching networks.

Note
It is a typical characteristic of RF transistors that they are conditionally unstable for lower frequencies (K<1). Calculating MAG for those frequencies makes therefore no sense. In other words, only for frequencies with K>1, an impedance matching for max. gain MAG can be designed.

It is further helpful to draw |S21|2 and MAG together in the same plot versus log(freq). Usually, MAG is bigger than |S21|2, what means that the power amplification can be improved with impedance matching, compared to the conventional measurement condition Z0=50Ω. See the following figure. Note that for a wide frequency range, the transistor is not unconditionally stable (K<1), and therefore MAG is not defined.



Comparing power amplification |S21|2 to the maximum available gain.

Note
For modeling, it is interesting to compare MAG_measured and MAG_simulated, since basically all S-parameters contribute to the MAG formula. A good fitting for MAG is an indication for a good model fit.

In case that S_{12}=0, the MAG formula from (10) simplifies and we get for the

Maximum Unilateral Transducer Power Gain

$$GU_{max} = \frac{|S_{21}|^2}{\left|\left(1-|S_{11}|^2\right)\cdot\left(1-|S_{22}|^2\right)\right|} = G_{smax} \cdot G_0 \cdot G_{lmax} \qquad (11)$$

with G_{smax}, G_0 and G_{lmax} from (9a)..(9c) (see Agilent App.Note AN95-1.

When the transistor/amplifier being measured by a conventional vector network analyzer, i.e. in a Z_0 environment, the Current amplification for AC-wise shorted output port is

$$\underline{h21} = \frac{-2\,\underline{S21}}{(1-\underline{S11})(1+\underline{S22})+\underline{S12}\,\underline{S21}}$$

This gives the Power Amplification in dB
for this Z_0 environment condition:

$$|\underline{h21}|^2 \text{ in dB} = 20 * \log|\underline{h21}|$$

Mason's power gain

This is another often cited power gain definition:

$$U = \frac{\frac{1}{2}\cdot\left|\frac{S_{21}}{S_{12}}-1\right|^2}{k\cdot\left|\frac{S_{21}}{S_{12}}\right| - \mathrm{Re}\left(\frac{S_{21}}{S_{12}}\right)}$$

with

$$k = \left[\frac{1-|S_{11}|^2-|S_{22}|^2+|\Delta S|^2}{2\cdot|S_{12}S_{21}|}\right] > 1$$

as in equation (4) above, and

$$|\Delta S| = |S_{11}S_{22} - S_{12}S_{21}| < 1$$

Definition of the Transit Frequency

The transit frequency ft is defined as the frequency at which |h21| hits '1'. This frequency ft determines the max. switching frequency and is therefore most interesting for digital applications (output AC-wise shorted).

The definition of the transition frequency assumes, that the transistor behaves like a single pole lowpass with a -20dB decrease per frequency decade. Looking at real measurement data shows, that this assumption is true in most cases, even for very high frequencies.

As we assume a 1-pole low-pass for h21, the gain-bandwidth product is a constant. Therefore it is sufficient to measure a h21 at a fixed frequency higher than the -3dB frequency. In other words, this fixed frequency should be from a -20dB/decade range of h21 . This measurement frequency can be found when transforming the measured S-parameters to H-parameters (using the TwoPort function in IC_CAP). From the dB-plot of ABS(h21(f)) versus log(frequency) we determine a frequency where the slope fits a -20dB/decade roll-off.

This frequency is now used as a fixed frequency f-20dB. After the S-parameters are converted into H-parameters, we get for the constant gain-bandwidth product of this one-pole low-pass filter:

$$1 * f_{T\,1\text{-pole}(DC_bias)} = \left| h_{21(DC_bias)} \right| * f_{-20dB}$$

or

$$f_{T1\text{-pole}(DC_bias)} = \left| h_{21(DC_bias)} \right| * f_{-20dB}$$

Using Y parameters, ft can also be calculated as:

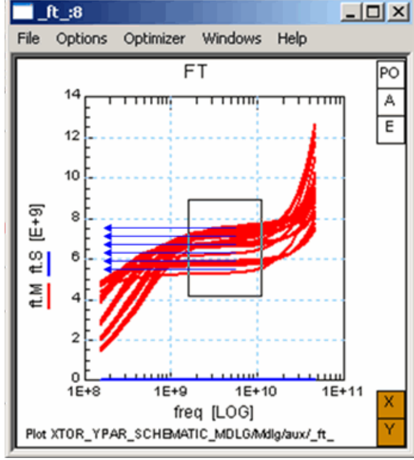
$$f_{T\,1\text{-pole}(DC_bias)} = \left| \frac{Y_{21(DC_bias)}}{Y_{11(DC_bias)}} \right| * f_{-20\,dB}$$

Especially for MOS Transistors:

$$f_T = \frac{gm}{2 \cdot PI \cdot C_{Gate}}$$

(M.T.Yang et.al, On The High-Frequency Characteristics and Model of Bulk Effect, IEICE Trans. Electron., Vol.E88-C, No.5, May 2005)

A short excerpt from the above literature:
At low Vgs in the weak inversion region, the increased electron (hoe) velocity also reduces the total transit time of the carriers. these lead to a significant increase of gm, resulting in a higher fT. For PMOS, the fT increase is less due to less mobility. With increasing Vgs, in strong inversion, fT reaches a maximum due to velocity saturation. Beyond that, fT decreases again, due to mobility reduction. Applying more vB shifts the VTH, and thus, reduces the fT and shifts its maximum to higher vGS .
In practice, the simplest way to see fT is to multiply MAG(H21) by the stimulating frequency array. This gives a plot as shown below:



This visualizes easily where the prerequisite of the formula presented above is valid: where the ft-curve **is flat vs. frequency**. This is marked by the box in the plot above. The different ft-values, represented by the blue arrows, represent the different DC bias conditions.
Yet, this is not the max. frequency for the transistor in analog operation and without the assumption of an AC-wise short at the output of the twoport.

Definition of the Maximum Oscillation Frequency fmax:

This max. frequency is defined by fmax , which refers to the maximum available gain MAG. As long as MAG> 1, the transistor can still generate oscillations. The max. possible oscillation frequency is therefore

$$MAG(f_{max}) = 1$$

or, referring to Manson's power gain:

$$U(f_{max}) = 1$$

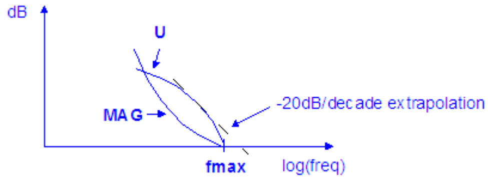
Although neither U nor MAG are representing a 1-pole low-pass filter, and, therefore, a -20dB/decade interpretation does not make real sense for MAG. But many modeling engineers do so, and calculate:

$$f_{max} = \sqrt{MAG} \cdot f \qquad \text{MAG: max. available gain}$$

or

$$f_{max} = \sqrt{U} \cdot f \qquad \text{U: Mason's Gain}$$

where f equals the measurement frequency at the falling slope (where the slope is assumed to drop with -20dB/decade).
In J.Berkner, Kompaktmodelle für Bipolartransistoren, expert Verlag, ISBN 3-8169-2085-3, Fig. 6-45, a nice sketch about the problems which may arise with this method is given. See the next figure:



As can be seen, neither MAG does not follow a -20dB/decade slope for all frequencies, nor does U.
However, theoretically, both curves should hit the 'real fmax' at 0dB,
as described by Gupta, M.S.: "Power Gain in Feedback Amplifiers, a Classic Revisited", IEEE Trans.on Microwave Theory and Techniques, Vol.40, No.5, May 1992. So, applying a -20dB/decade slope interpretation will result in a fmax which is frequency dependent, see the sketched -20dB/decade extrapolation. In many cases, the max(fmax) is then published in data sheets of foundry processes.

> **Note**
> This frequency-dependency of fmax has also been presented during the European IC-CAP workshop 2002 in Berlin by P.Brenner, Infineon, Munich, "Test Structures, Methodology And Specific Problems In Measuring Power Gains And fMAX Characteristics Of High Speed Transistors With fMAX > 50GHz".
> For H21 of a 1-pole low-pass filter (a transistor for example), the MAG(H21) decreases by -20dB per decade in a dB vs. log(freq) plot. In this case, an extrapolation of the -20dB/decade slope by means of: !spar_amp_characteristics_equation32.gif!like described further above makes sense.

There are approximation formulas published which allow fmax to be calculated independent of the max. available gain MAG.
For a bipolar transistor, B.Ardouin proposes:

$$f_{\max} = \sqrt{\frac{f_T}{8\pi \cdot C_{BC} \cdot R_{BB'}}}$$

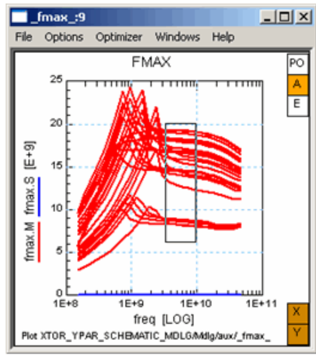
B.Ardouin, Contribution à la modélisation et la caractérisation en hautes fréquences des transistors bipolaires à heterojunction Si/SiGe, PhD Thesis at the University Bordeaux, École doctorale de sciences physiques et de l'ingenieur, Dec.10, 2001
For MOS transistors, G.Knoblinger proposes:

$$f_{\max} = \sqrt{\frac{f_T}{8\pi \cdot C_{GD} \cdot R_{Gate}}}$$

G.Knoblinger, Modellierung des Hochfrequenzverhaltens von MOS-Transistoren, PhD Thesis at the Universität der Bundeswehr, Munich, Fakultät Elektrotechnik, Aug.10, 2001
and M.T.Yang et.al:

$$f_{\max} = \sqrt{\frac{f_T}{4\pi \cdot C_{DB} \cdot \sqrt{(R_{Gate} + R_i + R_{Source}) \cdot R_{Bulk}}}}$$

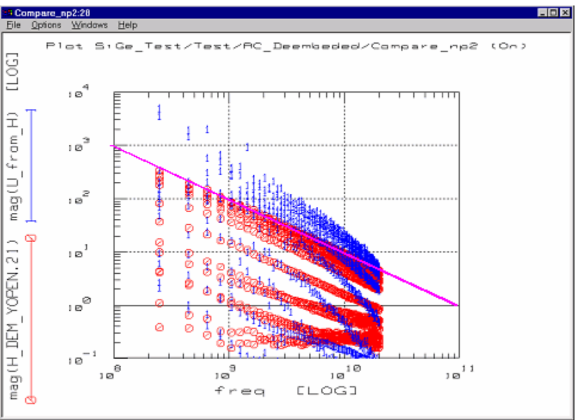
M.T.Yang et.al, On The High-Frequency Characteristics and Model of Bulk Effect, IEICE Trans. Electron., Vol.E88-C, No.5, May 2005



> **Note**
> Like with the ft Plot, the fmax is calculated from the traces inside the box (where the prerequisites are valid), and is a function of the applied DC biases.

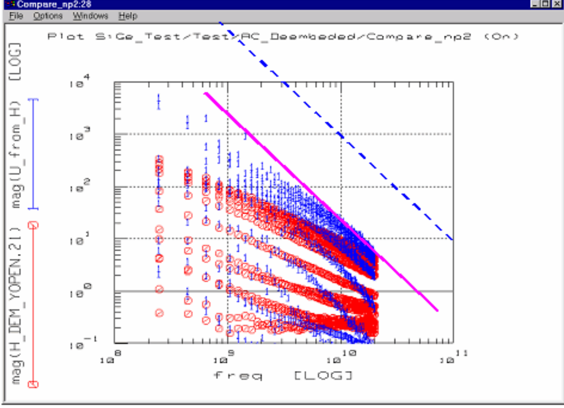
Notes on Estimating fmax vs fT by Extrapolation, and Comparison of fT vs fmax Measurements

Source: Labnotes on calculating fmax and ft, with friendly permission from Nathan Perkins, Agilent Technologies, 2004



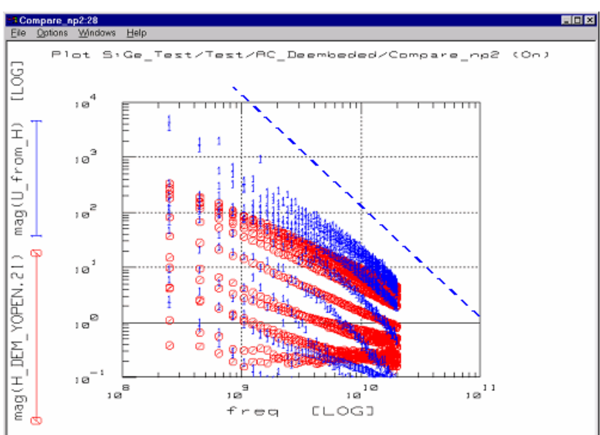
As expected, |H21| rolls off at one decade of magnitude per decade of frequency. Therefore, fT (H21 magnitude=1) is always calculable as fT=mag(H21)*freq for any frequency.

However, |U| rolls off by two decades of magnitude per decade of frequency. This is because U is a power gain, not a current gain. Therefore, the simple approximation of fmax=mag(U)*freq would badly overestimate the fmax value !



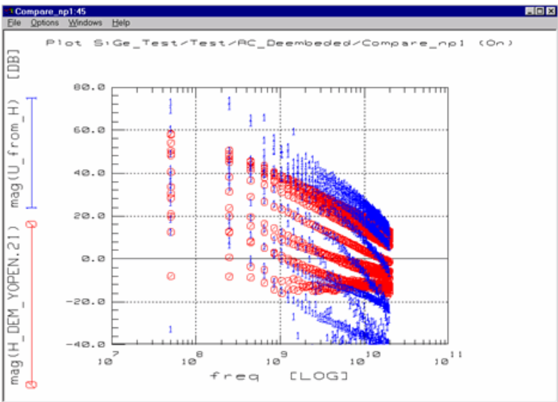
Linearizing this, we get log(U) = -2*log(freq) from U ~ freq-2. Anywhere along a line of a specific DC bias, the product U*freq2 is a constant. If any value of U and freq are measured, and the freq fmax at which U goes to unity is desired, we can calculate:

$$f\,max = \sqrt{U_{meas}} * freq_{meas}$$

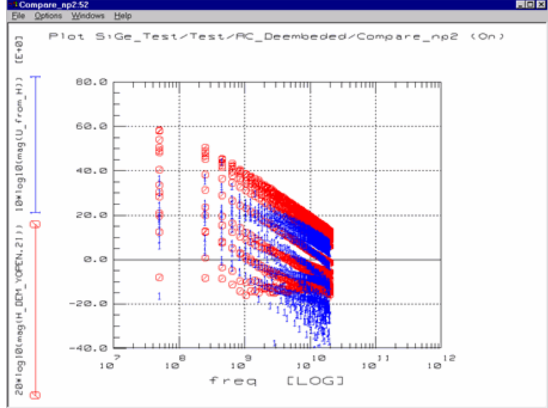


This leads to a solution/description for plotting fT vs fmax on dB scale graphs in IC-CAP.

Plotting raw values will give H21 a rolloff of -20dB/decade (expected value), but U will roll off at -40dB/decade. This is because we are inconsistent in our definition of dB.



This is because an IC-CAP plot always does dB scaling using 20*log(val) and not 10*log(val); e.g. always assumes a measurement of voltage gain rather than a measurement of power gain. Re-plotting with U as a power gain term (e.g., entering UdB as 10*log10(U)), and H21 as a current gain (e.g., 20*log10(MAG(H21)), we see that both now have a roll-off of -20dB/decade of frequency. This is more consistent with many literature publications.



Considerations About the Validity of Unilateral Gain Calculations

All twoport models are bilateral, so both the forward and the reverse signal flow must be considered. If the signal flow in the reverse direction is much smaller than the flow in the forward direction, it is possible to make the simplification that the reverse flow is zero.

The Unilateral Figure of Merit U is a quick calculation that can be used to determine where this simplification can be made without significantly affecting the accuracy of the complete gain formula.

$$u = \frac{|S_{11} * S_{22} * S_{12} * S_{21}|}{\left|\left(1 - |S_{11}|^2\right) * \left(1 - |S_{22}|^2\right)\right|}$$

This gives the error limits on Unilateral Gain Calculations:

$$\frac{1}{1+u^2} < \frac{G_T}{G_{TU}} < \frac{1}{1-u^2}$$

Example:
Suppose that a transistor is to be used in a simple amplifier, and optimized for power gain at a certain frequency by means of lossless input and output matching networks. Since the reverse gain S12 for this transistor is quite small, -50 dB smaller than forward gain S21 , there is a possibility that it can be neglected. If this is so, the design problem will be much simpler, because setting S12 equal to zero will make the design equations much less complicated. In determining how much error will be introduced by assuming S12 = 0, the first step is to calculate the unilateral figure of merit U, using the formula given above.

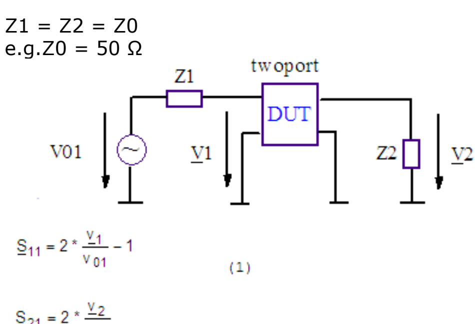
U is then plotted against log(freq). As an example, if the maximum value of U is 0.03, so the maximum error in this case turns out to be about + 0.25 dB at that frequency. This is small enough to justify the assumption that S12 = 0.

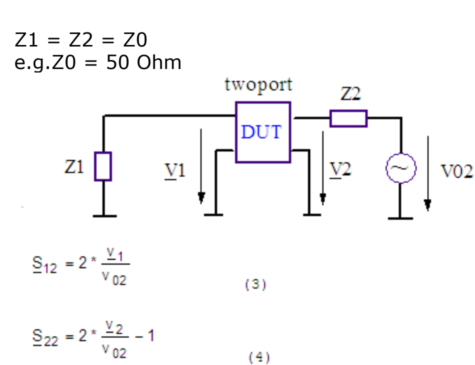
> **Note**
> A small i.e. neglectible reverse gain, or feedback factor, S12 , is an important and desirable property for a transistor to have, for reasons other than it simplifies amplifier design. A small feedback factor means that the input characteristics of the completed amplifier will be independent of the load, and the output will be independent of the source impedance. In most amplifiers, isolation of source and load is an important consideration.

Calculating S-parameters from Voltages

As it was mentioned before, S-parameters can be interpreted in terms of voltage at the DUT in a Z0 environment.

The following sketch gives an explanation about how to calculate them for a given twoport (DUT), imbedded in an external circuit, which itself represents the characteristic impedance.





Converting N-Port S-Parameters to M-Port

Converting 3-port parameters to 2-port and back

In order to transform e.g. Common-Emitter S-Parameters to Common-Base, we will now consider 3-Port S-Parameters. Provided a certain port is connected to ground, we are then able to evaluate the conversion formulas.

> **Note**
> For details, see the publication /Stassen/ mentioned at the end of this chapter.

3-PORT -> 2-PORT

Provided that the device under test is connected with 50Ω to all its 3 ports, the 3-pole network S-parameters are defined as below:

$$b_1 = S_{11} * a_1 + S_{12} * a_2 + S_{13} * a_3$$
$$b_2 = S_{21} * a_1 + S_{22} * a_2 + S_{23} * a_3$$
$$b_3 = S_{31} * a_1 + S_{32} * a_2 + S_{33} * a_3 \qquad (1a..1c)$$

If port 3 is connected to ground, we get for the corresponding reflection coefficient Γ3

$$\Gamma_3 = \frac{a_3}{b_3} = -1$$ -----------------------------------**(2)**

or

$$b_3 = -a_3$$

what changes (1c) to

$$a_3 = -\frac{S_{31} * a_1}{1+S_{33}} - \frac{S_{32} * a_2}{1+S_{33}}$$ -------------------------------------**(3)**

substituting (3) in (1a) and (1b), we get the 2-port network parameters from the 3-port ones as:

$$\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} S_{11} - \frac{S_{13} * S_{31}}{1+S_{33}} & S_{12} - \frac{S_{13} * S_{32}}{1+S_{33}} \\ S_{21} - \frac{S_{23} * S_{31}}{1+S_{33}} & S_{22} - \frac{S_{23} * S_{32}}{1+S_{33}} \end{pmatrix} * \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \qquad (4)$$

Converting from 2-PORT -> 3-PORT

Using the matrix in (4), we can also calculate the 3-port S-parameters out of 2-port parameters.
As described in publication /Stassen/, see below, the sum of each column and row in a 3-port S-parameter matrix equals '1', i.e.

$$\sum_{j=1}^{3} S_{ij} = 1 \quad \text{for } i=1,2,3$$

and

$$\sum_{i=1}^{3} S_{ij} = 1 \quad \text{for } j=1,2,3$$

Referring to the above equations (1) and (2), we can calculate the 3-port S-parameters out of the 2-port ones with Index T (Twoport Measurement) :

$$S_{33} = \frac{\sum_{i=1,2;\, j=1,2} S_{ijT}}{4 - \sum_{i=1,2;\, j=1,2} S_{ijT}}$$

$$S_{32} = \frac{1+S_{33}}{2}(1 - S_{12T} - S_{22T})$$
$$S_{23} = \frac{1+S_{33}}{2}(1 - S_{21T} - S_{22T})$$
$$S_{22} = S_{22T} + \frac{S_{23} + S_{32}}{1+S_{33}}$$

$$S_{31} = 1 - S_{33} - S_{32}$$
$$S_{13} = 1 - S_{23} - S_{33}$$
$$S_{12} = 1 - S_{22} - S_{32}$$
$$S_{11} = 1 - S_{21} - S_{31}$$
$$S_{21} = 1 - S_{22} - S_{23}$$

Once these 3-port S-parameters are known, they can be used to calculate the common-base or common-collector 2-port S-parameters of a bipolar transistor

etc.

Publications:
Rudolf Stassen, 'Einsatz eines Mikrowellen-SPitzenmeßplatzes zur Charakterisierung von Transistore direkt auf dem Wafer im Frequenzbereich von 0,045 bis 26,5GHz', Diplomarbeit am Institut für Schicht- und Ionentechnik am Forschungszentrum Jülich GmbH, available from Technische Informationsbibliothek, Hannover.

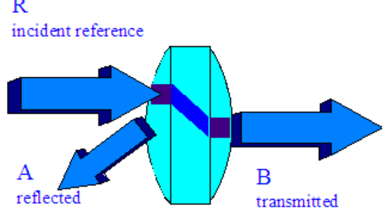
Converting from 2-PORT -> 1-PORT and its Application for the Q factor of RF-passive Components

This is required if a component has been measured in 2-Port mode, and some of its components characteristics might be affected by the characteristic impedance of the opposite VNA port.

We refer to equation. (3) from above and obtain:

$$S_{11_1port} = S_{11} - \frac{S_{12} * S_{21}}{1 + S_{22}}$$

Definition of S-Parameters



Scatter Parameters, also called S-parameters, belong to the group to twoport parameters used in twoport theory. Like the Y or Z parameter, they describe the performance of a twoport completely. Different to Y and Z, however, they relate to the traveling waves that are scattered or reflected when a network is inserted into a transmission line of a certain characteristic impedance ZL. Therefore, S-parameters can be compared to reflection and through-pass of a pair of spectacles.

S-parameters are important in microwave design because they are easier to measure and to work with at high frequencies than other kinds of twoport parameters. They are conceptually simple, analytically convenient and capable of providing detailed insight into a measurement and modeling problem. However, it must kept in mind that -like all other twoport parameters, S-parameters are linear by default. I.e. they represent the linear behavior of the twoport.

This means that S-parameters do relate traveling waves (power) to a twoport's reflection and transmission behavior. Since the twoport is embedded in a characteristic impedance of Z0, these 'waves' can be interpreted in terms of normalized voltage or current amplitudes. This is explained below.

The S-Parameters relate the **b** (out of the device) PowerWaves to the **a** (into the device) PowerWaves

$$b1 = S11 * a1 + S12 * a2$$
$$b2 = S21 * a1 + S22 * a2$$

with ai: PowerWave towards the twoport
bi: PowerWave out of the twoport

Z0

Looking at the S-parameter coefficients individually, we have:

$$S_{11} = \frac{b_1}{a_1} = \frac{\text{PowerWave}_{\text{reflected at port 1}}}{\text{PowerWave}_{\text{towards port 1}}} \Big/ a_2 = 0 \qquad S_{12} = \frac{b_1}{a_2} = \frac{\text{PowerWave}_{\text{out of port 1}}}{\text{PowerWave}_{\text{towards port 2}}} \Big/ a_1 = 0$$

$$S_{21} = \frac{b_2}{a_1} = \frac{\text{PowerWave}_{\text{out of port2}}}{\text{PowerWave}_{\text{towards port1}}} \Big/ a_2 = 0 \qquad S_{22} = \frac{b_2}{a_2} = \frac{\text{PowerWave}_{\text{reflectedat port2}}}{\text{PowerWave}_{\text{towards port2}}} \Big/ a_1 = 0$$

S11 and S21 are determined by measuring the magnitude and phase of the incident, reflected and transmitted signals when the output is terminated in a perfect Zo load. This condition guarantees that a2 is zero. S11 is equivalent to the input complex reflection coefficient or impedance of the DUT, and S21 is the forward complex transmission coefficient.

Likewise, by placing the source at port 2 and terminating port 1 in a perfect load (making a1 zero), S22 and S12 measurements can be made. S22 is equivalent to the output complex reflection coefficient or output impedance of the DUT, and S12 is the reverse complex transmission coefficient.
The accuracy of S-parameter measurements depends greatly on how good a termination we apply to the port not being stimulated. Anything other than a perfect load will result in a1 or a2 not being zero (which violates the definition for S-parameters). When the DUT is connected to the test ports of a network analyzer and we don't account for imperfect test port match, we have not done a very good job satisfying the condition of a perfect termination. For this reason, two-port error correction, which corrects for source and load match, is very important for accurate S-parameter measurements.
Let's now discuss some characteristic S-parameter values.

S11 and s22

Value	Interpretation
-1	All voltage amplitudes towards the twoport are inverted and reflected (0Ω)
0	impedance matching. no reflections at all (5Ω).
1	voltage amplitudes are reflected (infinite Ω).

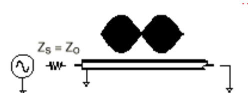
The magnitude of S11 and S22 is always less than 1. Otherwise, it would represent a negative ohmic value. On the other hand, the magnitude of S21 (transfer characteristics) respectively S12 (reverse) can exceed the value of 1 in the case of active amplification. Also, S21 and S12 can be positive an negative. If they are negative, there is a phase shift. Example: S21 of a transistor starts usually at about S21 = -2 -10. This means signal amplification within the Z0 environment and phase shift.

S21 and S12

Magnitude	Interpretation
0	no signal transmission at all.
0...+1	input signal is damped in Z_0 environment.
+1	unity gain signal transmission in the Z_0 environment.
>+1	

The numbering convention for S-parameters is that the first number following the S is the port at which energy emerges, and the second number is the port at which energy enters. So S21 is a measure of power emerging from Port 2 as a result of applying an RF stimulus to Port 1.
In order to better understand the Sxx parameters, let's consider an ideal transmission line, connected to a port of the VNA, and terminated with either a LOAD, an OPEN or a SHORT, or an ideal RESISTOR. (This chapter is from HP App.Note 1287-1).

a transmission line connected to a VNA



At very low frequencies, with wavelengths much larger than the line, the transmission line can be thought of a simple wire. This is adequate for conducting DC or very low frequency power. The resistance of the wire is relatively low and has little effect on low-frequency signals. The voltage and current are the same no matter where a measurement is made on the wire. At higher frequencies, wavelengths are comparable to or smaller than the length of the transmission line (or a conductor in a high-frequency circuit), and power transmission can be thought of in terms of traveling waves. When the transmission line is terminated in its characteristic impedance, maximum power is transferred to the load. When the termination is not equal to the characteristic impedance, that part of the signal that is not absorbed by the load is reflected back to the source.

If a transmission line is terminated in its characteristic impedance Z0 (LOAD condition), no reflected signal occurs since all of the transmitted power is absorbed by the load. Looking at the envelope of the RF signal versus distance along the transmission line shows no standing waves because without reflections, energy flows in only one direction.

When the transmission line is terminated in a SHORT circuit (which can sustain no voltage and therefore dissipates zero power), a reflected wave is launched back along the line toward the source. The reflected voltage wave must be equal in magnitude to the incident voltage wave and be 180 degrees out of phase with it at the plane of the load. The reflected and incident waves are equal in magnitude but traveling in the opposite directions.

If the transmission line is terminated in an OPEN-circuit condition (which can sustain no current), the reflected current wave will be 180 degrees out of phase with the incident current wave, while the reflected voltage wave will be in phase with the incident voltage wave at the plane of the load. This guarantees that the current at the open will be zero. The reflected and incident current waves are equal in magnitude, but traveling in the opposite directions. For both the short and open cases, a standing wave pattern is set up on the transmission line. The voltage valleys will be zero and the voltage peaks will be twice the incident voltage level.

If the transmission line is terminated with say an ideal 25 Ω RESISTOR, resulting in a condition between full absorption and full reflection, part of the incident power is absorbed and part is reflected. The amplitude of the reflected voltage wave will be one-third that of the incident wave, and the two waves will be 180 degrees out of phase at the plane of the load. The valleys of the standing-wave pattern will no longer be zero, and the peaks will be less than those of the short and open cases. The ratio of the peaks to valleys will be 2:1.

Voltage standing wave ratio

The traditional way of determining RF impedance was to measure VSWR using an RF probe/detector, a length of slotted transmission line, and a VSWR meter. As the probe

was moved along the transmission line, the relative position and values of the peaks and valleys were noted on the meter. From these measurements, impedance could be derived. The procedure was repeated at different frequencies. Modern network analyzers measure the incident and reflected waves directly during a frequency sweep, and impedance results can be displayed in any number of formats (including VSWR).

The most general term for ratioed reflection is the complex reflection coefficient, Γ (gamma). The magnitude portion of Γ is called ρ (rho). The reflection coefficient is the ratio of the reflected signal voltage level to the incident signal voltage level. For example, a transmission line terminated in its characteristic impedance Z0, will have all energy transferred to the load so Vrefl = 0 and ρ = 0. When the impedance of the load, ZL is not equal to the characteristic impedance, energy is reflected and ρ is greater than zero. When the load impedance is equal to a short or open circuit, all energy is reflected and ρ = 1. As a result, the range of possible values for ρ is 0 to 1.

Reflection Parameters

Reflection Coefficient $\Gamma = \frac{V_{reflected}}{V_{incident}} = \rho \angle \Phi = \frac{Z_L - Z_O}{Z_L + Z_O}$

Return loss $= -20 \log(\rho), \rho = |\Gamma|$

E_{max} E_{min}

Voltage Standing Wave Ratio

VSWR $= \frac{E_{max}}{E_{min}} = \frac{1 + \rho}{1 - \rho}$

Return loss is a way to express the reflection coefficient in logarithmic terms (dB). Return loss is the number of decibels that the reflected signal is below the incident signal. Return loss is always expressed as a positive number and varies between infinity for a load at the characteristic impedance and 0 dB for an open or short circuit. Another common term used to express reflection is voltage standing wave ratio (VSWR), which is defined as the maximum value of the RF envelope over the minimum value of the RF envelope. It is related to ρ as
(1 + ρ)/(1 - ρ).
VSWR ranges from 1 (no reflection) to infinity (full reflection).

After these thoughts on the Sxx parameters, lets finish by considering the properties of the Sxy parameters. The transmission coefficient is defined as the transmitted voltage divided by the incident voltage. If the absolute value of the transmitted voltage is greater than the absolute value of the incident voltage, a DUT or system is said to have gain. If the absolute value of the transmitted voltage is less than the absolute value of the incident voltage, the DUT or system is said to have attenuation or insertion loss. The phase portion of the transmission coefficient is called insertion phase.

Transmission parameters

$V_{Incident}$ → DUT → $V_{Transmitted}$

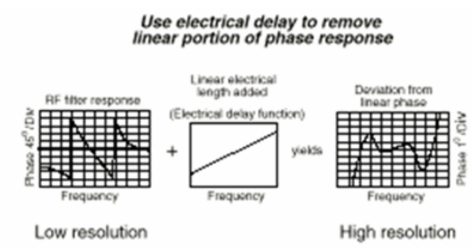
Transmission Coefficient $= T = \frac{V_{Transmitted}}{V_{Incident}} = \tau \angle \phi$

Insertion Loss (dB) $= -20 \, Log \left| \frac{V_{Trans}}{V_{Inc}} \right| = -20 \log \tau$

Gain (dB) $= 20 \, Log \left| \frac{V_{Trans}}{V_{Inc}} \right| = 20 \log \tau$

Direct examination of insertion phase usually does not provide useful information. This is because the insertion phase has a large (negative) slope with respect to frequency due to the electrical length of the DUT. The slope is proportional to the length of the DUT. Since it is only deviation from linear phase that causes distortion in communications systems, it is desirable to remove the linear portion of the phase response to analyze the remaining nonlinear portion. This can be done by using the electrical delay feature of a network analyzer to mathematically cancel the average electrical length of the DUT. The result is a high-resolution display of phase distortion or deviation from linear phase.

Applying electrical delay

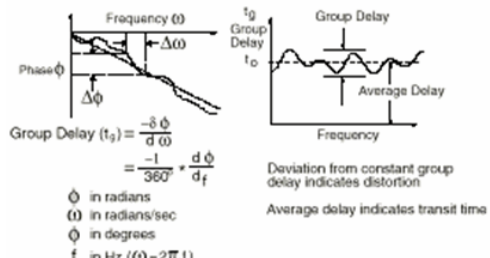


Note
Adding electrical delay should not be used for modeling.

Another useful measure of phase distortion is group delay. This parameter is a measure of the transit time of a signal through a DUT versus frequency. Group delay can be calculated by differentiating the DUT's phase response versus frequency. It reduces the linear portion of the phase response to a constant value, and transforms the deviations from linear phase into deviations from constant group delay, (which causes phase distortion in communications systems). The average delay represents the average signal

transit time through a DUT.

Understanding group delay



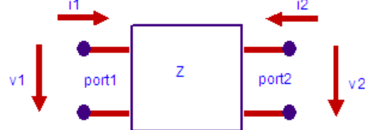
Differential S-Parameters

When measuring S-parameters using a VNA, the DUT is characterized by reflection and thru measurements at both ports. These signals are referring to signal ground.

In some applications, however, the device may be used in differential mode, e.g. in differential amplifier stages etc. In such a case, the characteristics of the device may be different from those measured in the conventional S-parameter measurement environment. This is especially important for passive RF components and their Q factor, i.e. the 'quality factor' of their RF performance. Therefore, it is desirable to also model both, the performance and the Q factor, in the appropriate stimulus condition.

Since the S-parameter are linear twoport parameters, they can be converted into Z-parameters, and the Z-parameters can be converted in special Z-parameters for common mode excitation and differential mode excitation.

Let's begin with the conventional Z matrix definition:



For differential mode condition, we can define:

$v_diff = v_d = v1 - v2$ and $i_diff = i_d = \frac{i1 - i2}{2}$

And for common mode,

$v_common = v_c = \frac{v1 + v2}{2}$ and $i_common = i_c = i1 + i2$

Referring to the Z matrix,

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix} \cdot \begin{bmatrix} i_1 \\ i_2 \end{bmatrix},$$

we can substitute by the common and differential voltages and currents and obtain

$$\begin{bmatrix} v_c \\ v_d \end{bmatrix} = \begin{bmatrix} Z_{cc} & Z_{cd} \\ Zdc & Z_{dd} \end{bmatrix} \cdot \begin{bmatrix} i_c \\ i_d \end{bmatrix}$$

with

$$Z_{cc} = \frac{Z_{11} + Z_{12} + Z_{21} + Z_{22}}{4} \qquad Z_{cd} = \frac{Z_{11} - Z_{12} + Z_{21} - Z_{22}}{2}$$

$$Z_{dc} = \frac{Z_{11} + Z_{12} - Z_{21} - Z_{22}}{2} \qquad Z_{dd} = Z_{11} - Z_{12} - Z_{21} + Z_{22}$$

This conversion scheme can be applied to get differential mode and common mode S-parameters:

- Convert S-parameters to Z
- Apply common/differential conversion
- Convert Z-parameters back to S

Calculating differential and Common-mode S-parameters directly from S-parameters

$$S_{diff} = \frac{S_{11} + S_{22} - S_{21} - S_{21}}{2}$$

$$S_{com} = \frac{S_{11} + S_{22} + S_{21} + S_{21}}{2}$$

Publication: Ch.Inui, M.Fujishima, Characterization of T-Shaped Terminal IMpedances of Differential Short Stubs in Advanced CMOS Technology, 2008 IEEE ICMTS Conference, Edinburgh, UK

Related to the Q-Factor of Spiral Inductors Operated in Differential Mode

In non-differential mode, the Q factor is calculated from the Z-parameters, which have been converted to 1-port Z-parameters as

$$Q = \frac{IMAG(Z_{11_1port})}{REAL(Z_{11_1port})}$$

This refers to a measurement condition where the 1st port is tested, and the 2nd is grounded.
Under differential mode operating conditions, this is not the case. Referring to the sections above, we now have:

$$L_{dd} = \frac{IMAG(Z_{dd})}{2 \cdot PI \cdot freq} \quad \text{and} \quad R_{dd} = REAL(Z_{dd})$$

what gives

$$Q_{dd} = \frac{IMAG(Z_{dd})}{REAL(Z_{dd})} = \frac{L_{dd} \cdot 2 \cdot PI \cdot freq}{R_{dd}}$$

Wrap-up recipe: Under differential excitation of a symmetric inductor, differential impedance, inductance and quality factor are defined as

Zdiff = Z11 + Z22 - Z12 - Z21

Ldiff = 1/(2*PI*f) * IMAG(Zdiff)

Qdiff = IMAG(Zdiff) / REAL(Zdiff)

Differential 1-Port S-Parameters

The differential one-port S-parameter can be calculated from the conventionally measured S-parameters by

$$S_{diff_1port} = \frac{S_{11} + S_{22} - S_{12} - S_{21}}{2}$$

From that, we obtain the differential Z-parameter (with the diff. characteristic impedance Z0):

$$Z_{diff_1port} = \frac{2Z0 \cdot (1 + S_{diff_1port})}{1 - S_{diff_1port}}$$

Acknowledgments:
Special thanks to Laurent Gambus of Philips in Caën for bringing up this topic of differential/common 2-port parameters and also for the basic matrix conversion method.

Publications on Differential/Common Mode Z-parameters

M.Danesh, J.R.Long, R.A.Hadaway, D.L.Harame, "A Q-Factor Enhancement Technique For MMIC Inductors", 1998 IEEE MTT-S Digest, p.183-186

M. Danesh, et al., "Differential driven symmetric microstrip inductors", IEEE TMTT, vol. 50, no. 1, Jan., 2002.
Differential 1-Port S-Parameters

after M.Danesh, J.R.Long, "Differentially Driven Symmetric Microstrip Inducors",
IEEE Trans. Microwave Theory and Techniques, vol.50, no.1, pp.332-341, Jan 2002

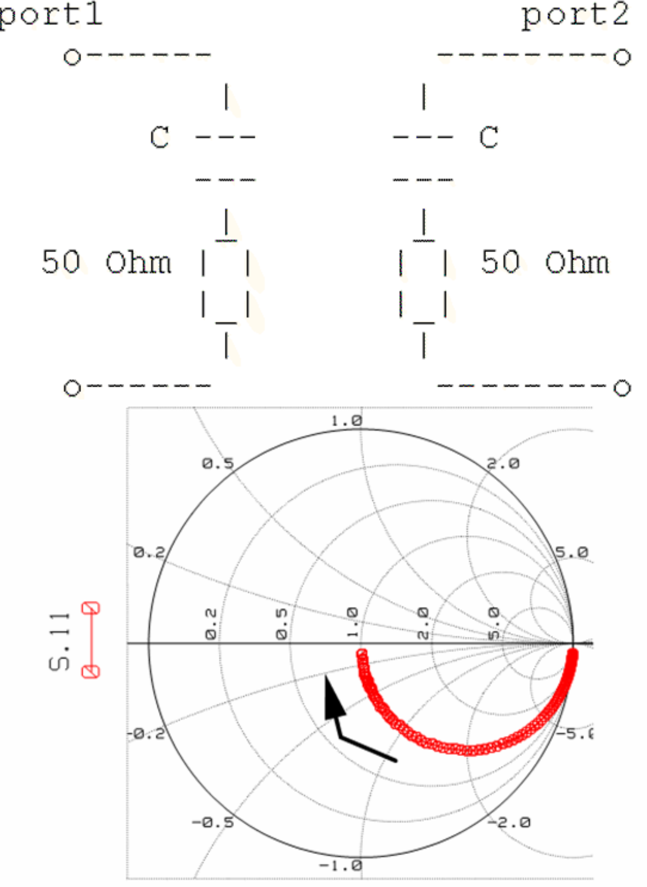
Interpreting S-Parameter Plots

INTRODUCTION TO UNDERSTANDING S-PARAMETER PLOTS

IC_CAP file: 1_Sxy_plots_LCR.mdl

This sub-chapter covers expected s-parameter measurements for typical LCR circuits.

Additional delay lines would shift the phase of the following plots, but not the magnitude. If the delay line is matching the impedance of the NWA ports, the phase shift is linear. Otherwise, it is non-linear. For more details about such phase shifts see the next sub-chapter.

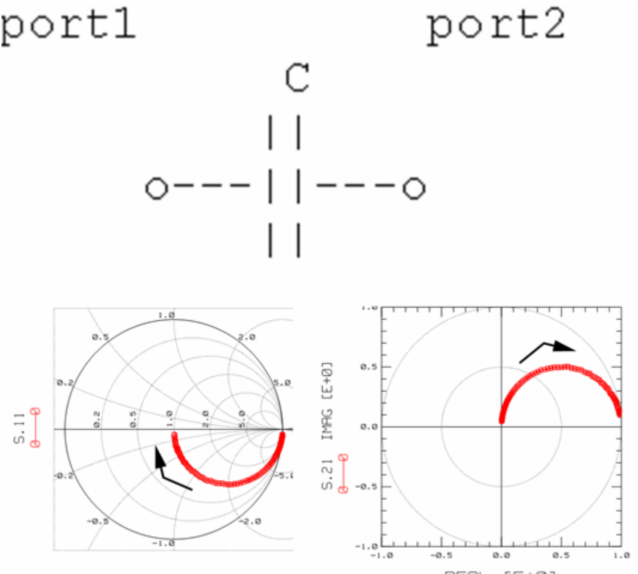


S11

freq	S11	Ohm	Notes
0	+1	infinite	C is an OPEN
infinite	0	50	C is a SHORT, and we see the 50 Ohm resistors of the opposite port

S21

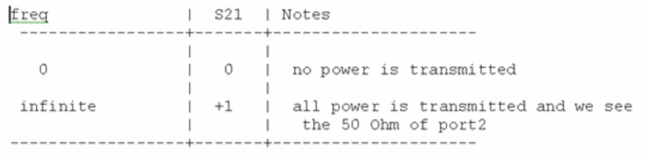
S21 is always 0, because both ports are decoupled. No power can be transmitted between them.



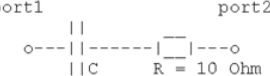
S11:

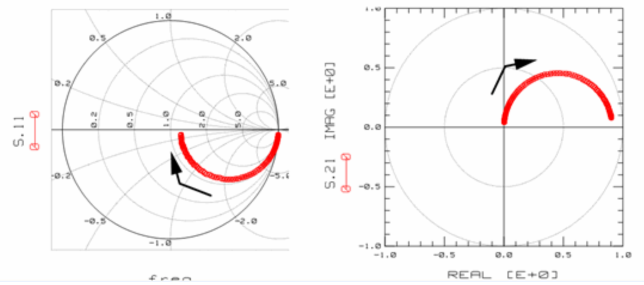
freq	S11	Ohm	Notes
0	+1	infinite	C is an OPEN
infinite	0	50	C is a SHORT, but we see 50 Ohm of Port2 (!!)

S21



freq	S21	Notes
0	0	no power is transmitted
infinite	+1	all power is transmitted and we see the 50 Ohm of port2

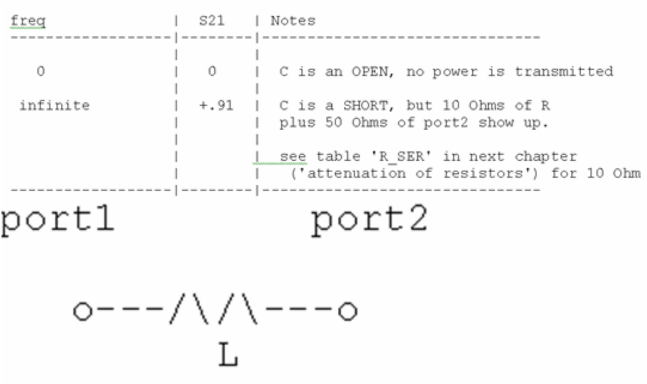




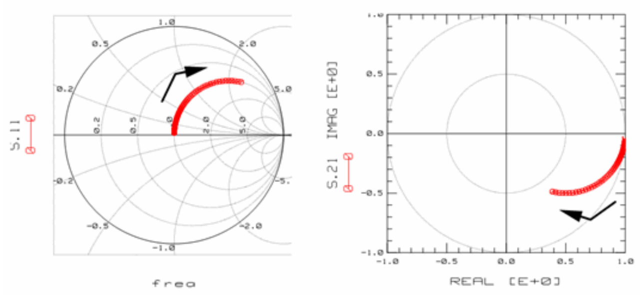
S11

freq	S11	Ohm	Notes
0	+1	infinite	C is an OPEN, all power is reflected back
infinite	+0.1	60	C is a SHORT, but 10 Ohms of R plus 50 Ohms of Port2 show up

S21



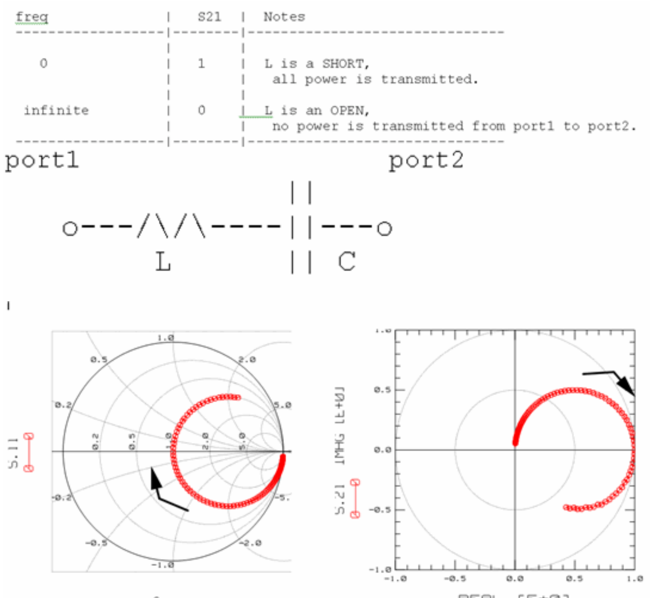
freq	S21	Notes
0	0	C is an OPEN, no power is transmitted
infinite	+.91	C is a SHORT, but 10 Ohms of R plus 50 Ohms of port2 show up. see table 'R_SER' in next chapter ('attenuation of resistors') for 10 Ohm



S11

freq	S11	Ohm	Notes
0	0	50	L is a SHORT, but we see 50 Ohm from port2
infinite	1	infinite	L is an OPEN, therefore is port2 virtually isolated from port1

S21

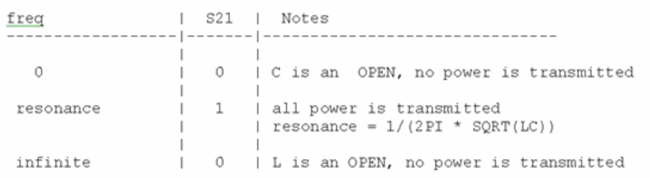


freq	S21	Notes
0	1	L is a SHORT, all power is transmitted.
infinite	0	L is an OPEN, no power is transmitted from port1 to port2.

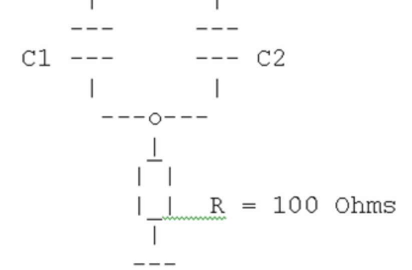
S11

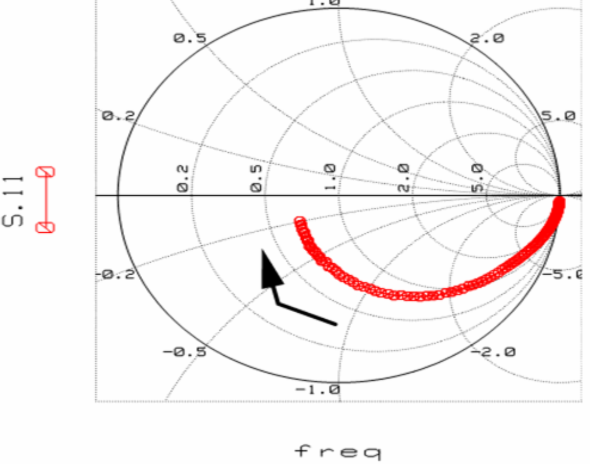
freq	S11	Ohm	Notes
0	+1	infin.	C is an OPEN, all power is reflected back.
resonance	0	50	S11 shows 50 Ohm from port2 resonance = 1/(2PI * SQRT(LC))
infinite	+1	infin.	L is OPEN, all power is reflected back

S21



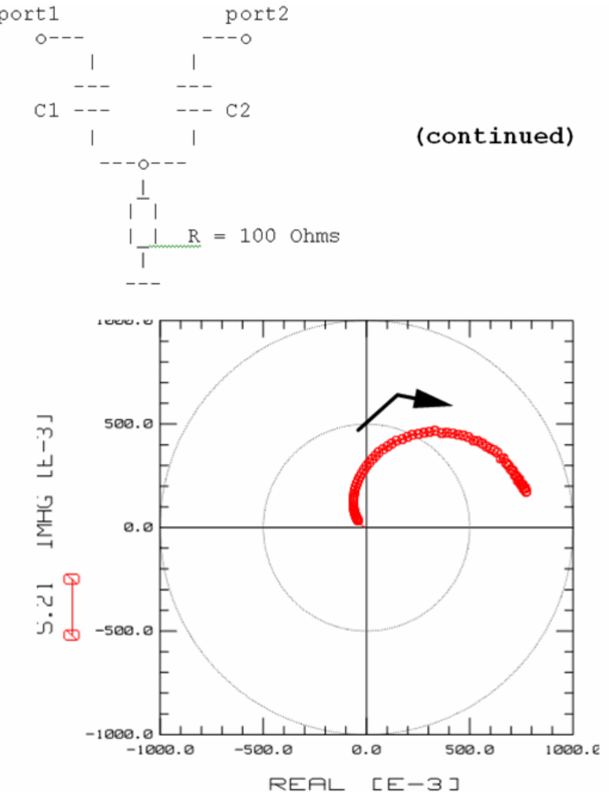
freq	S21	Notes
0	0	C is an OPEN, no power is transmitted
resonance	1	all power is transmitted resonance = 1/(2PI * SQRT(LC))
infinite	0	L is an OPEN, no power is transmitted



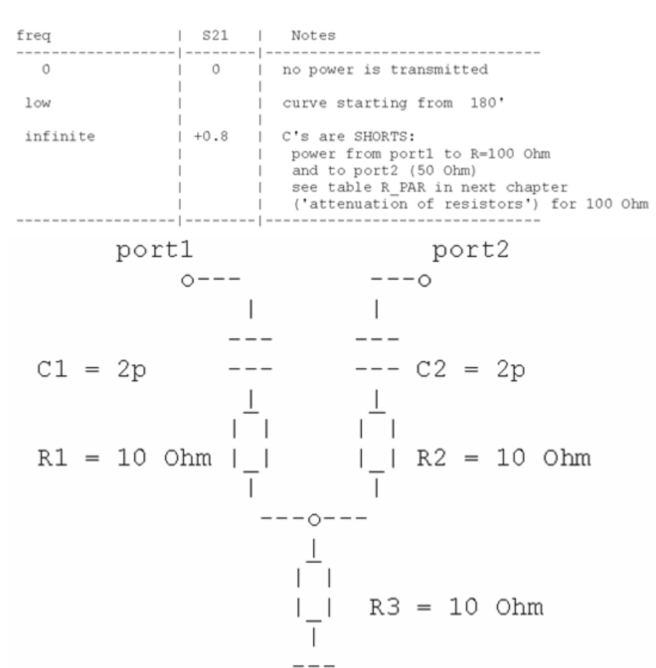


S11

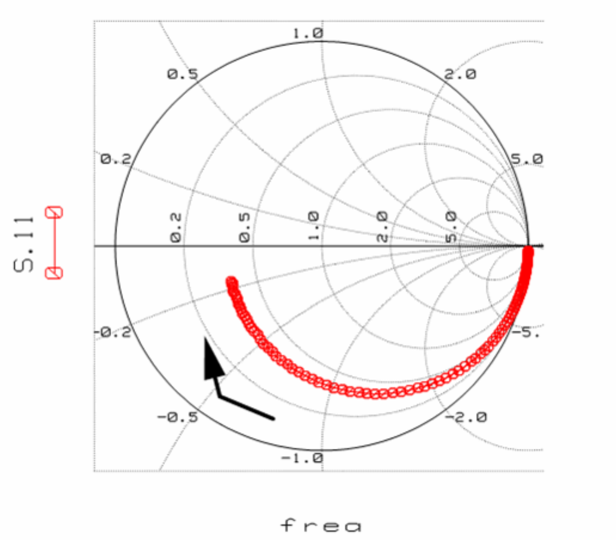
freq	S11	Ohm	Notes
0	+1	infin.	Cs are OPEN, all power is reflected back
low			circle from infinite to R=100 Ohms, Note: port2 (50 Ohms) gets no power yet
high			power hits port2, 50 Ohm of port2 dominate now over R=100 Ohms. Therefore the curve deviates to another circle from from infinite Ohms to 33.3 Ohms (100 // 50 Ohms)
infinite		33.3	100 // 50 Ohms



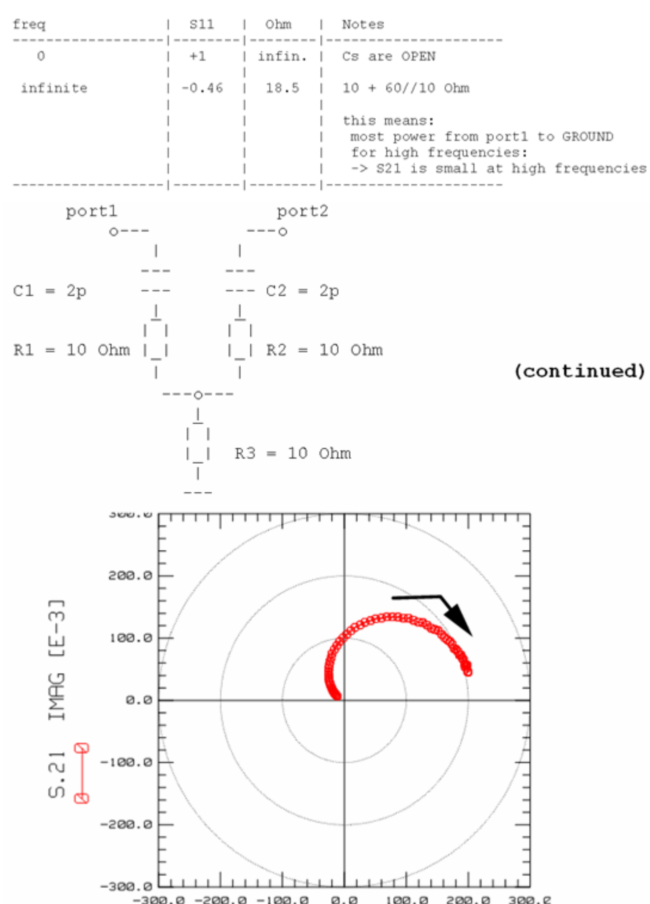
S21



freq	S21	Notes
0	0	no power is transmitted
low		curve starting from 180'
infinite	+0.8	C's are SHORTS: power from port1 to R=100 Ohm and to port2 (50 Ohm) see table R_PAR in next chapter ('attenuation of resistors') for 100 Ohm



S11

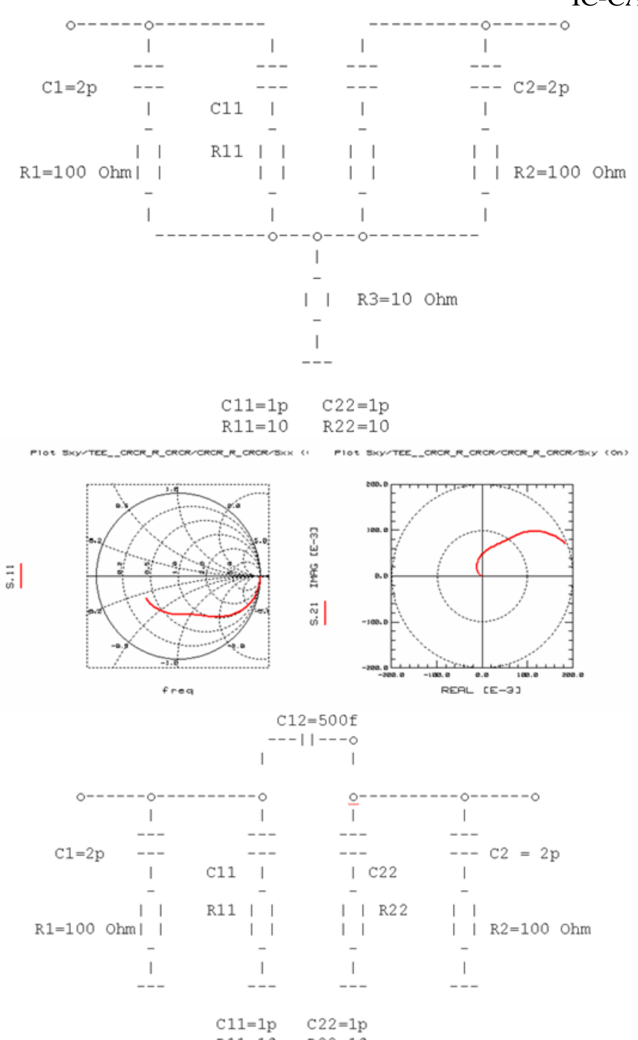


freq	S11	Ohm	Notes
0	+1	infin.	Cs are OPEN
infinite	-0.46	18.5	10 + 60//10 Ohm this means: most power from port1 to GROUND for high frequencies: -> S21 is small at high frequencies

S21

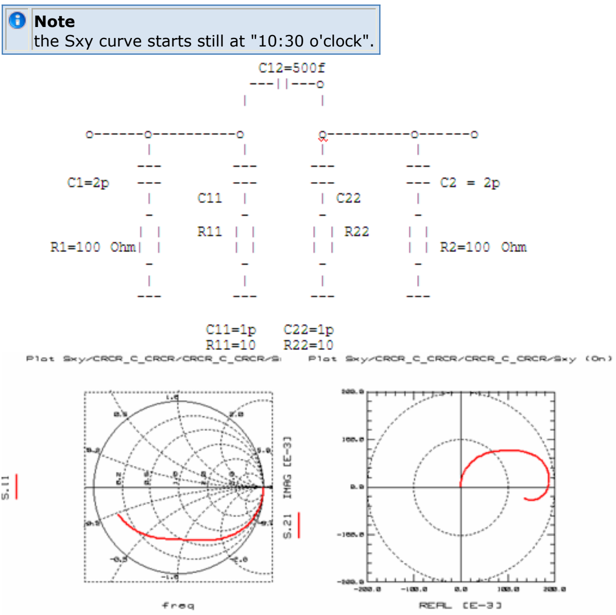
freq	S21	Notes
0	0	no power transmitted
low		curve starting from 180'
infinite	~ 0.2	

NOTE: the Sxy curve starts at "10:30 o'clock". This is a hint for the common resistor to ground!



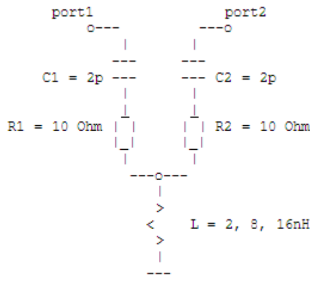
Since the second RC has a higher 3dB frequency than the first, the curve thends for high frequencies to an overlay of both RC branches

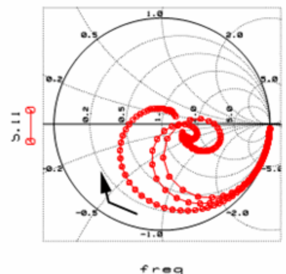
> **Note**
> the Sxy curve starts still at "10:30 o'clock".

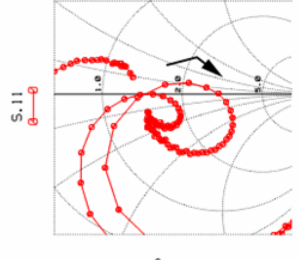


Since the second RC has a higher 3dB frequency than the first, the curve thends for high frequencies to an overlay of both RC branches

> **note**
> The Sxy curve starts upwards, at "12 o'clock". This is a hint for the cross-coupling capacitor.





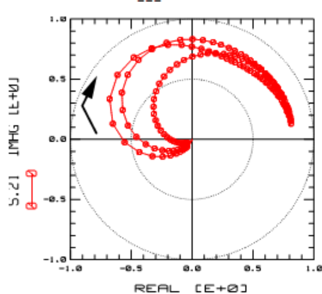


S11

S11:

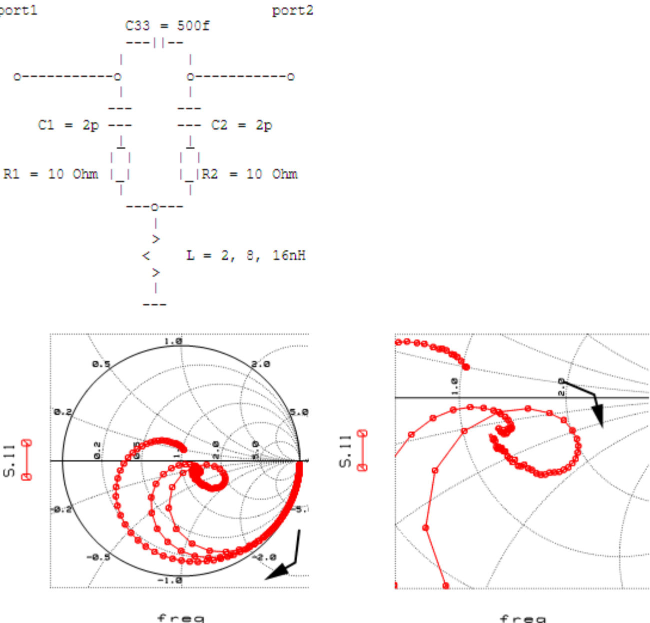
freq	S11	Ohm	Notes
0	1	infin.	Cs are OPEN
low			circle from infinite to 10 Ohm, L is still a SHORT
high			branch to port2 becomes dominant, L opens, PHASE exceeds 180', i.e. L and C's are active!
infinite	>0	70	10+10 + 50 Ohm Cs are SHORTs L is OPEN 50 Ohm from port2 is visible

```
     port1               port2
        o---           ---o
           |           |
          ---         ---
C1 = 2p ---         --- C2 = 2p
           |           |
          _|_         _|_
R1 = 10 Ohm | |       | | R2 = 10 Ohm
          |_|         |_|
           |           |
           ---o---
              |
              >
             <    L = 2, 8, 16nH
              >
              |
             ---
```



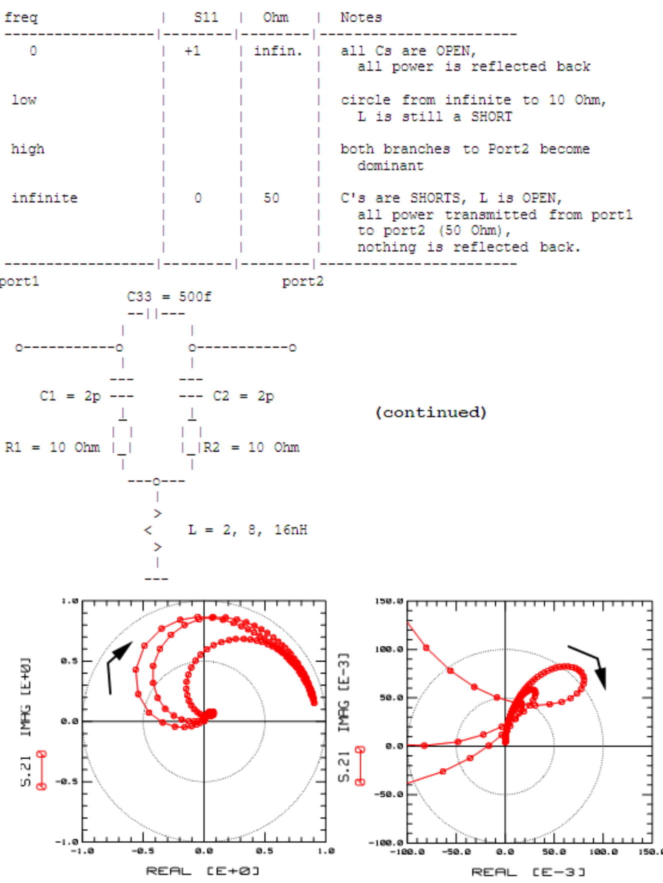
S21

freq	S21	Notes
0	0	no power transmitted
	0	
low		curve starting from 270'
	0	
infinite	+0.83	see table R_SER in next chapter ('attenuation of resistors') (20 Ohm)



Note
the Sxy curve starts at "8:30 o'clock". This is a hint for the inductor to ground

S11

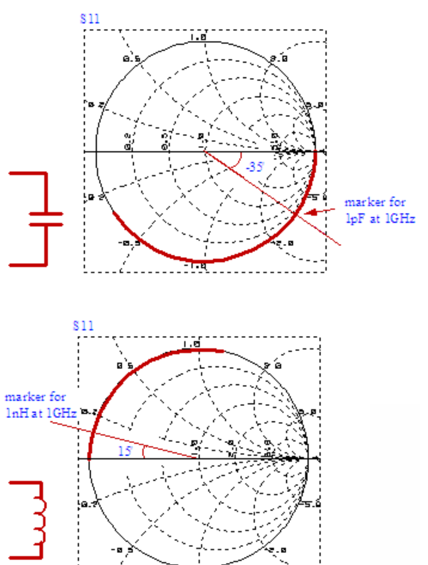


freq	S11	Ohm	Notes
0	+1	infin.	all Cs are OPEN, all power is reflected back
low			circle from infinite to 10 Ohm, L is still a SHORT
high			both branches to Port2 become dominant
infinite	0	50	C's are SHORTS, L is OPEN, all power transmitted from port1 to port2 (50 Ohm), nothing is reflected back.

S21:

freq	S21	Notes
0	0	no power is transmitted
low		curve starting from 90', i.e. C33 dominates over TEE.
medium		looping back as if curve would start from 270' i.e. TEE dominates now over C33.
infinite	+1	C33 is a SHORT

Phase Shift of Inductors and Capacitors



Attenuation of Resistors Related to S-Parameters

S21 VALUES FOR RESISTORS THAT ARE PARALLEL OR IN SERIES WITH THE NWA PORTS

The interpretation of S11 or S22 curves in terms of complex inductances is simple, since we use a Smith chart. Yet, the interpretation of the traces of S12 or S21 seem to be more complex. The following two tables may help in better understanding such Sxy plots. They give the Sxy values for typical series and parallel ohmic resistors. These values will be measured when R_PAR and R_SER are imbedded between 50 Ohm delay lines (Port1 and Port2, respectively):

TABLE R PAR:

```
- - -------o------- - -
           |
          | | R_PAR
          |_|
           |
- - -------o------- - -
```

R PAR	Sxy .
infinite	1,00
1k	0,98
5000	0,95
200	0,89
100	0,80
50	0,66
20	0,44
10	0,28
5	0,17
2	0,07
1 Ohm	0,04

TABLE R SER:

```
            ___
- - ---|___|--- - -
        R_SER

- - ------------ - -
```

R SER	Sxy .
infinite	0,000
100k	0,001
10k	0,010
5k	0,020
2k	0,047
1k	0,091
500	0,167
200	0,33
100	0,500
50	0,67
20	0,83
10	0,91
5	0,95
2	0,98
1 Ohm	0,99

Note: The formula for R_SER is:

$$\frac{2 \cdot Z0}{2 \cdot Z0 + R}$$

How to use these tables

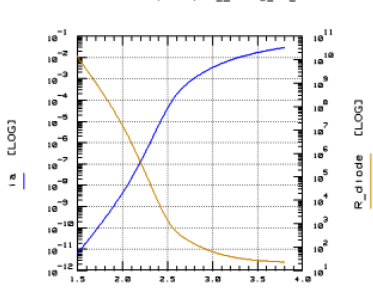
Assuming we have a value of REAL(S21) equal to 0,95 (i.e. 95% power transmission from

Port1 to Port2). As Port2 has 50 Ohms, the series resistor between Port1 and Port2 has 5 Ohms.

If we assume a parallel resistor and again have a value of REAL(S21) equal to 0,95, the parallel resistor to Port2 (50 Ohm) is now 500 Ohm.

Application

An application of the Sxy -> Rseries context is for example for diode modeling.
Let's assume a diode between Port1 and Port2. Its resistance varies a lot with the DC bias, see the plot below:

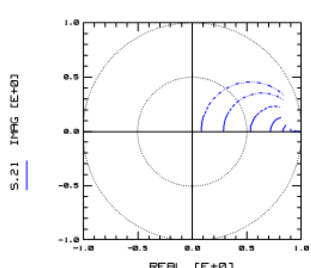


Diode DC characteristics and corresponding diode resistance.

This means. when the diode current changes from pA to mA, the diode resistance drops from G? to ~10?.
Related to S-parameters, for lowest bias, the S21 of a diode is 0, while for highest bias, we expect S21 -> 1. (The end point on the real axis of S21 is determined by the ohmic series resistance of the diode!).

Therefore, the question arises which DC biases to select for diode modeling.
This question can be simply answered by inspecting the R_diode plot and its corresponding S21-value from the table above:
For S21 starting points between S21~0,1 and S21~0,9, R_diode should be between roughly 1k? and 10?. For the diode above, suitable DC bias points would be between 2.5V and 3.5V.

The next plot shows the measurement result based on these DC bias conditions:



S21 characteristics of a diode, biased after the considerations from above

While this S21 plot shows the 'inner diode', i.e. without any series inductance (package or strip lines on the wafer for finger structured devices), the next S21 plots shows the same device, but now including a series inductance.

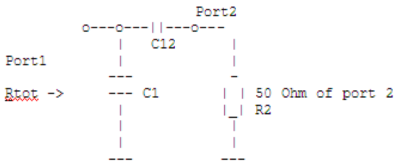
HINT: watch the difference between fig.2 and fig.3:
Fig.2: the diode capacitance shorts the dynamic diode resistance for high frequencies. The end point of all traces is equal to the RS, the ohmic series diode resistance, usually close to '1'.
Fig.3: the shortening effect of the diode capacitance is overlaid by the extra phase shift of the inductance. Therefore it can happen, that the starting angle for inrementing DC bias changes from +90° abruptly to -90° !

Plot led4course/Spar/Spar/Sxy (On)

S21 characteristics including a series inductance(e.g. package)

About the Interpretation of S-Parameters of an Asymmetrical Tee Structure

This tutorial is intended to help with the understanding of a typical small signal modeling problem and the interpretation of its Smith charts. It can be applied to better understand the trace of S11 and S22 curves of small signal transistor models after de-embedding. It explains especially the case where 'elliptic' curves are obtained instead of simple 'half circles'.



Schematic of an asymmetrical TEE schematic

The total resistance as seen at port1 is:

$$R_{tot} = \frac{\frac{1}{pC1} * \left(\frac{1}{pC12} + R2\right)}{\frac{1}{pC1} + \frac{1}{pC12} + R2} = ... = \frac{1}{pC1} * \frac{1}{\frac{1 + C12/C1 + pC12 * R2}{1 + pC12 * R2}}$$

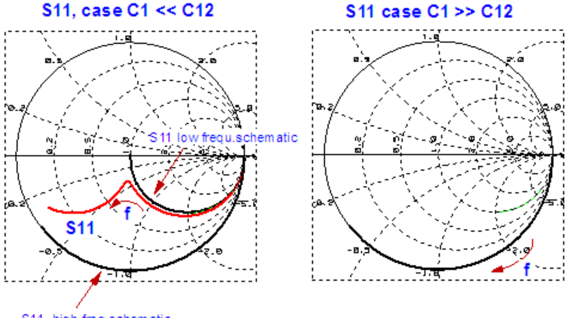
For extreme frequencies, Rtot behaves mainly like a capacitor Ceff with frequency dependent value.

For low frequencies, it behaves like Ceff_low = C1 (1 + C12/C1)

and for high frequencies like Ceff_high = C1

Referring to low frequencies and therefore Ceff_low, we can distiguish between two cases:
C12/C1 >> 1
and C12/C1 << 1

In the first case, Ceff_low ~ C12, and the influence of C1 can be neglected. The expected S11 curve will start at 'infinite' resistance, and lead towards 50 Ohm for higher frequencies (Port2 resistor R2). Yet for even higher ones, C1 will become conductive and will conduct the power flow from Port1 directly to ground. Therefore, R2 will become 'invisible'. Between both ranges, we expect a transitional trace of S11. This situation is depicted in the Smith chart of shown below.



In the case that C1 >> C12, we have Ceff ~ C1 for all frequencies. I.e. the signal flow from Port1 directly into the 'big' C1 is dominating always over the other flow through the 'small' C2 into R2. This is shown in fig.2b.

Fig.3 shows the trace of fig.2a converted to the complex resistance plane. The trace comes for low frequencies from 50? - j ? , and changes its direction to tend to 0?. As can

be seen, this kind of plot is ideally suited to extract the two ohmic resistance values of the underlaying schematic.

Zin from S-Parameter

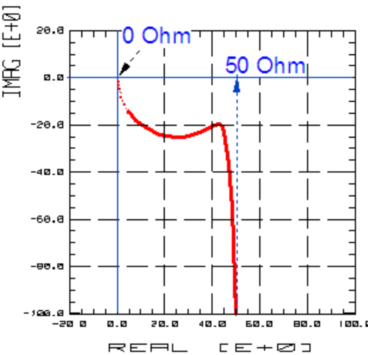
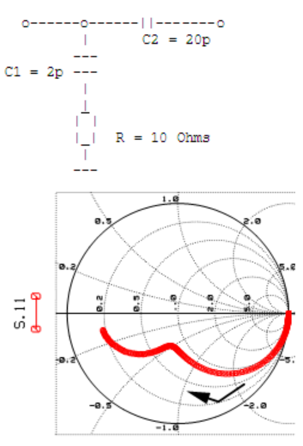


Fig.3: S11 from fig.2a, converted to the complex resistance plane. The knee happens at the same frequency as in fig.2a.

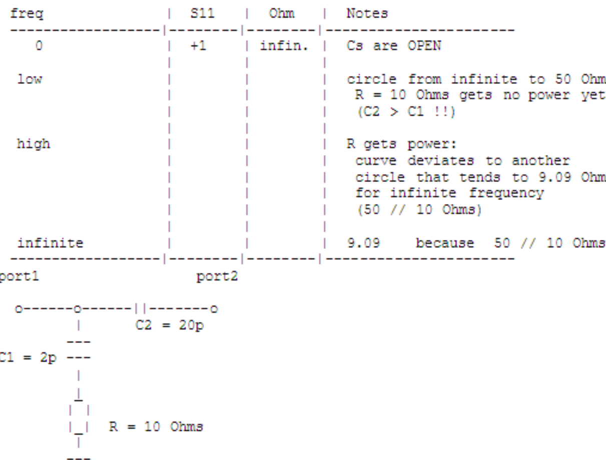
Application for transistors:
Fig.2a is often seen when measuring S22 curves of transistors. From the considerations above, the interpretation might now be easier when understanding the importance of the relative parameter values of CBE and CBC (for a bipolar transistor). This is covered in the last plots of this section.

Another example is the modeling of the output resistance of MOS transistors in the operating point. Here, the measured S-parameters are converted to Y, and Y22-1 is displayed like in fig.3. This allows the easy separation in typically a high-value ohmic output resistance and one or two series RC combination. More details see in the section "Operating Point Modeling of MOS Transistors".

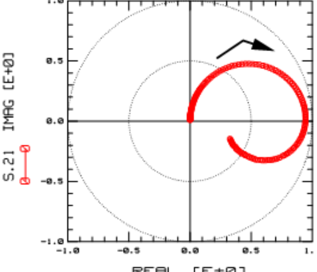
```
port1                    port2

  o-------o------||-------o
          |        C2 = 20p
         ---
C1 = 2p  ---
          |
          |
         | |
         |_|   R = 10 Ohms
          |
         ---
```



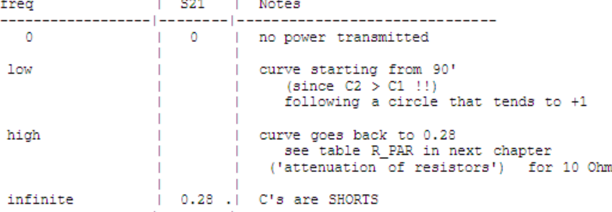
S11



```
freq                 |  S11    |  Ohm    |  Notes
---------------------|---------|---------|-----------------------
   0                 |  +1     | infin.  |  Cs are OPEN
                     |         |         |
 low                 |         |         |  circle from infinite to 50 Ohm
                     |         |         |   R = 10 Ohms gets no power yet
                     |         |         |   (C2 > C1 !!)
                     |         |         |
 high                |         |         |  R gets power:
                     |         |         |   curve deviates to another
                     |         |         |   circle that tends to 9.09 Ohms
                     |         |         |   for infinite frequency
                     |         |         |   (50 // 10 Ohms)
                     |         |         |
 infinite            |         |         |  9.09    because  50 // 10 Ohms
---------------------|---------|---------|-----------------------
port1                    port2

  o-------o------||-------o
          |        C2 = 20p
         ---
C1 = 2p  ---
          |
          |
         | |
         |_|   R = 10 Ohms
          |
         ---
```

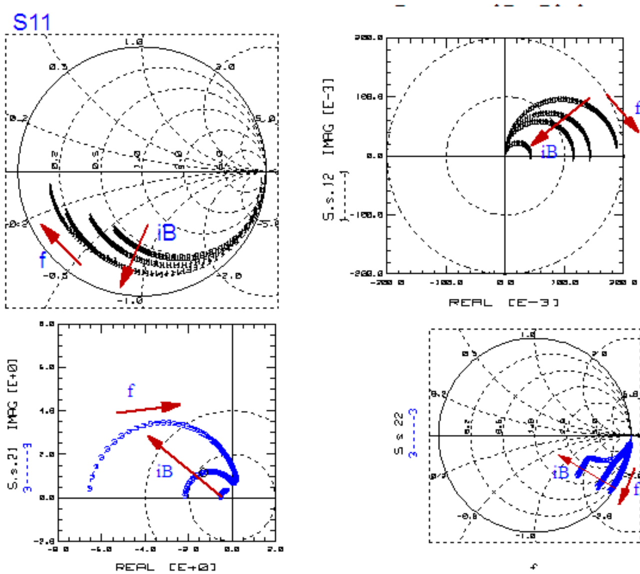


S21



freq	S21	Notes
0	0	no power transmitted
low		curve starting from 90' (since C2 > C1 !!) following a circle that tends to +1
high		curve goes back to 0.28 see table R_PAR in next chapter ('attenuation of resistors') for 10 Ohm
infinite	0.28 .	C's are SHORTS

Interpreting S-Parameter Plots of a Transistor

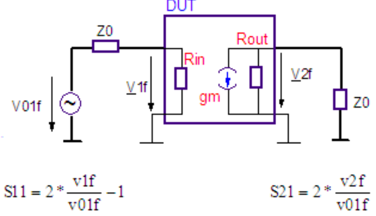


Starting Points of Transistor S-Parameter Plots

example MOS transistor
IC_CAP file: calc_spar_start_from_DC.mdl

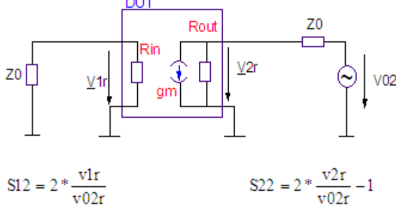
Of particular interest is the starting point of transistor S-parameter traces. These starting ponts are completely determined by the DC performance. This means that, if the starting points do not match for the transistor modeling, there is no chance to match the measured S-parameter traces with the model at all.

The S-parameters of the DUT can be calculated out of these schematics:
Forward Simulation following the formula:



$$S11 = 2 * \frac{v1f}{v01f} - 1 \qquad S21 = 2 * \frac{v2f}{v01f}$$

Reverse Simulation with the formula:



$S12 = 2 * \frac{v1r}{v02r}$ $S22 = 2 * \frac{v2r}{v02r} - 1$

Proveded, Rin, gm and Rout are obtained and known from DC measurements, we can calculate the voltages for Z0=50Ω:

FORWARD: $\frac{v1f}{v01f} = \frac{Rin}{50 + Rin}$ and $v2f = -gm * v1f * \frac{Rout}{50 + Rout}$

REVERSE:

$$\frac{v2r}{v02r} = \frac{Rout}{50 + Rout}$$

These results introduces into the S-parameter equations from above gives finally with V01f = 1 and v02r = 1:

$$S11_DC = 2 * v1f - 1 = 2 * \frac{Rin}{50 + Rin} - 1$$

$$S21_DC = 2 * v2f = -2 * gm * \frac{Rin}{50 + Rin} * \frac{Rout}{50 + Rout}$$

and

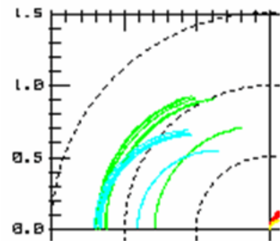
$$S22_DC = 2 * v2r - 1 = 2 * \frac{Rout}{50 + Rout}$$

Note about a possible mismatch between measured and simulated S21 transistor data at lowest frequencies

Although you have achieved an excellent fitting of the DC transfer curve, i.e. the output plot and even the Rout plot match well, you might run into some problems when fitting the S21 curves of a transistor. Since with S-parameter modeling, we can only affect the trend of the simulated curves towards higher frequencies by changing the transit time parameters, the capacitances or some external RF parasitics, we cannot change the starting points at lowest frequencies. These points are completely determined by the DC fit.

This means we might run into a modeling problem like depicted in the following figure:

mismatch between measured and simulated data of S21 at low frequencies



It has been observed that this effect can be due to the ohmic losses in the bias TEEs used for the S-parameter measurements. For example, if we bias the transistor directly from the DC instrument (HP414x, HP415x) into the bias connectors of the S-parameter test set of the VNA, we have to account for a total Rbias of about 2.5 ? at each port. Using the external bias TEEs, the remaining resistance is considerably lower, in the range of 0.5 Ω.

What does this mean for a MOS transistor: no voltage drop at the Gate, since Ig=0, However, a voltage drop for Vd. And, indeed, the biggest S21 mismatch can be observed for DC bias conditions where the output characteristics curves (Id vs. Vd) are most non-linear, i.e. for low Vd.

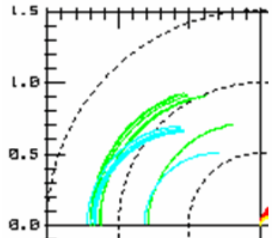
For a bipolar, the Rbias affects again mainly Ic, but also a bit Ib. On the other hand, the output characteristics is usually more linear than that of a MOS.

This means, we define a certain DC bias value for Vd in the S-parameter setup, but this Vd is not appplied to the transistor Drain. In order to compensate for this effect, we firstly measure the Rbias at each port by connecting the probes to the SHORT calibration standard, and by performing a DC current sweep (-10mA .. 10mA) for the SMU, measuring the voltage. Using the derivative function of IC-CAP, we can easily calculate the Rbias. This is done for both ports. After that, we add a Rbias1 and Rbias2 into the circuit deck of our transistor. Rbias1 in series with the Gate, and Rbias2 in series with the Drain. Both resistors are AC-wise shorted by e.g. 1F capacitors. (If we would not short these Rbias, we would influence considerably the RF trace of our S-parameters. By shorting them with these big capacitors, only the DC operating point of the transistor is affected, and not its S-parameter traces).

Example for a enhanced SPICE circuit deck:

```
.SUBCKT bsim_n_DC_bias 1=G 2=D

Rbias1 1 11 2.5
Cbias1 1 11 1
Rbias2 2 22 2.5
Cbias2 2 22 1

Xtransistor 22 11 0 0 bsim

.ENDS
```

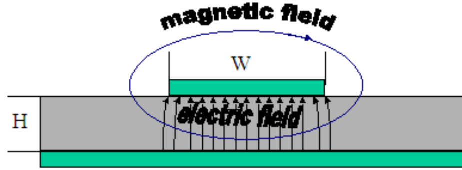
Applying this enhanced schematic, we obtain the curves of fig. 2 instead of those in the following figure:



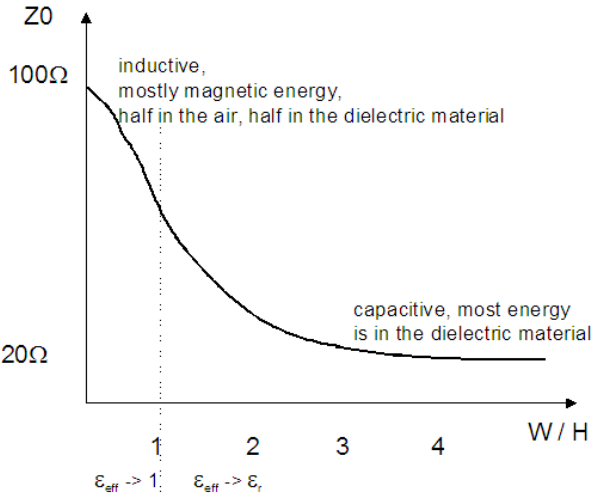
Interpreting S-Parameters of Strip Line

Strip Line Tutorial

A strip line exhibits a TEM field, i.e. a transverse electromagnetic mode field, where the dielectric field is perpendicular to the magnetic field, and both are perpendicular to the direction of wave propagation.



From this sketch, we can conclude the following basic properties of a microstrip line, depending on the geometry factor W/L:



The characteristic impedance, Z0, ranges from about 20Ω to about 100Ω. The limit of 100Ω exists for a very simple reason: the width is much less than the hight, and such a structure cannot be manufactured (under-etching etc).
This sketch allows to make some fundamental considerations:

As a matter of fact, a small microstrip line exhibits less capacitance than a wide one. Inspecting the plot, this concludes that a lower capacitance in a microstrip line comes along with a lower impedance Z0.

This is obvious, when recalling that

$$Z \sim \sqrt{\frac{L}{C}}$$

Referring to crosstalk between lines, we can learn from the sketch above that a low impedance microstrip line is capacitive. I.e. the energy is rather between the metal conductor and the groud. I.e. two low impedance striplines side-by-side, will exhibit less cross-talsk than two high impedance striplines. By the way, this is a key design rule for

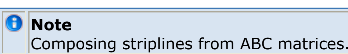
packages and connectors.

As another important outcome, a shielding across a microstrip line will lead to the fact that the impedance of the resulting strip line wil be lower, because more of the electro-magnetic field will now be present in the enlarged electric field consisting of the previous field in the dielectic layer plus the additional space between the active metal layer and the top cover. Therefore, a cover across a microstrip line reduces the resulting impedance, and, thus, reduces cross-talk between adjacent striplines.

To further reduce cross-talk of adjacent lines, i.e. to reduce the impedance of each line (inclrease the electric field, i.e. make the lines more capacitive), reduce the height of the dielectric material.

However, on the other hand, a cover 'kills' the performance of filters designed from strip lines based on electric field coupling!

OPEN ENDED STRIP LINE VERSUS A THRU STRIP LINE



Note
Composing striplines from ABC matrices.

When defining a strip line of an ABC matrix (chain matrix), please note that in this case the line is from Port1 to Port2. If you need to define an open ended strip line with an ABC matrix, make sure to multiply the above ABC strip line matrix with another ABC matrix of a thru with a very big series resistor value !
See the two PEL programs below:

THRU STRIPLINE

```
! A PEL program for a ABC matrix THRU microstip line
COMPLEX Athruline.M.22[SIZE(freq)]
i=0
WHILE i < SIZE(freq)
 tmp=j*2*PI*T1.TD*freq[i]
 Athruline.11[i]=cosh(tmp)
 Athruline.12[i]=T1.Z0*sinh(tmp)
 Athruline.21[i]=T1.Z0^-1*sinh(tmp)
 Athruline.22[i]=cosh(tmp)
 i = i + 1
END WHILE
RETURN Athruline
```

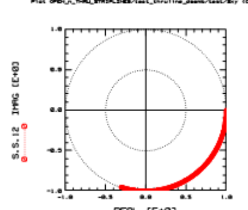
OPEN STRIPLINE

```
! A PEL program for a ABC matrix OPEN microstip line
COMPLEX Aline.M.22[SIZE(freq)]
i=0
WHILE i < SIZE(freq)
 tmp=j*2*PI*T1.TD*freq[i]
 Aline.11[i]=cosh(tmp)
 Aline.12[i]=T1.Z0*sinh(tmp)
 Aline.21[i]=T1.Z0^-1*sinh(tmp)
 Aline.22[i]=cosh(tmp)
 i = i + 1
END WHILE
! ABC matrix of a 1G resistor
! NOTE: the A matrix of an OPEN is NOT defined!
COMPLEX Aopen.M.22[SIZE(freq)]
i=0
WHILE i < SIZE(freq)
 Aopen.11[i]=1
 Aopen.12[i]=1G
 Aopen.21[i]=0
 Aopen.22[i]=1
 i = i + 1
END WHILE
RETURN Aline*Aopen
```

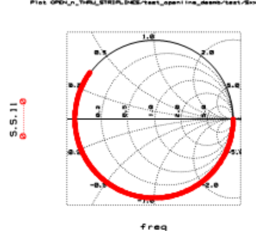
Note
Phase shift of an OPEN and a THRU strip line.

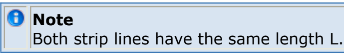
When simulating an OPEN ended strip line, the delay of the strip line shows up in Sxx. Note that due to forward and reverse traveling of the waves along the open ended strip line, the strip line delay shows up multiplied by 2. This becomes evident when comparing such a result with a THRU strip line.

THRU strip line between Port1 and Port2



OPEN ended strip line at Port1



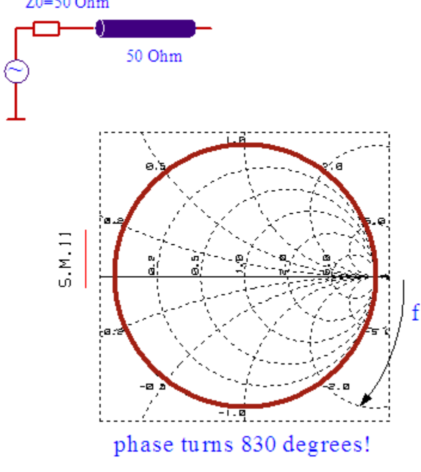

Note
Both strip lines have the same length L.

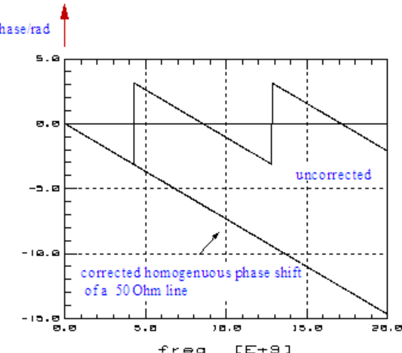
Conclusions

For the same phase shift, a THRU strip line needs a length 2*L compared to an OPEN ended strip line of length L !

S-PARAMETERS OF SERIAL LOSSLESS DELAY LINES WITH DIFFERENT IMPEDANCES

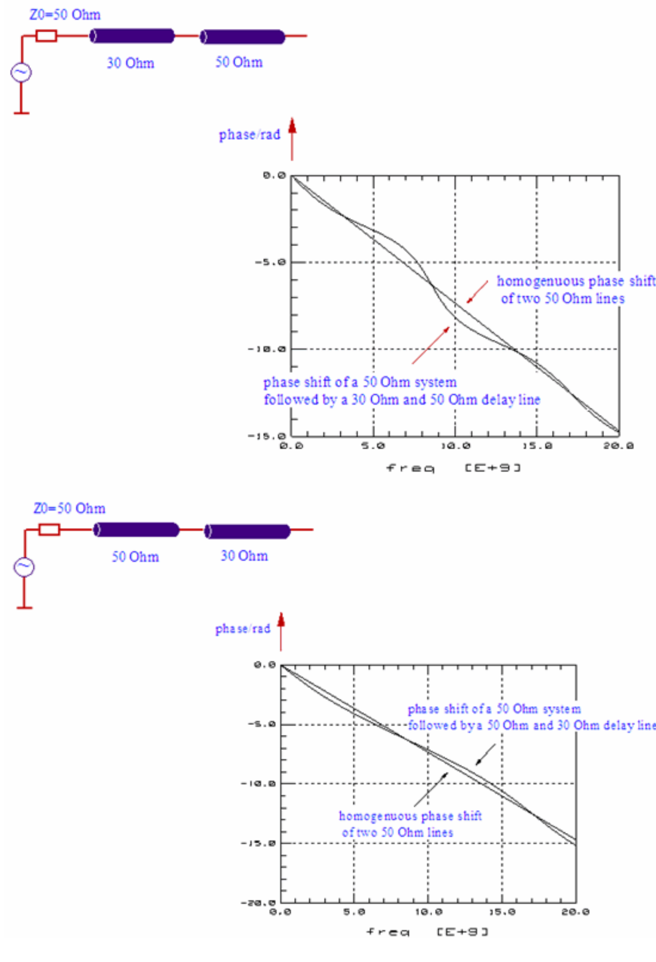
Delay lines show up in S-parameter plots with big phase shift. If they have no loss, ~~like the SPICE models~~ they keep the magnitude. This is depicted below.





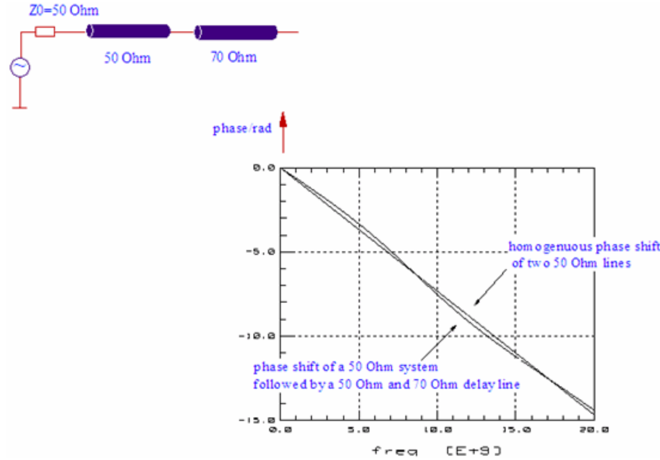
Yet, if there is a mismatch in the impedances, the phase is affected. The following plots give an idea. Both lines have a delay of 30ps. The frequency is swept from 45MHz to 20GHz.

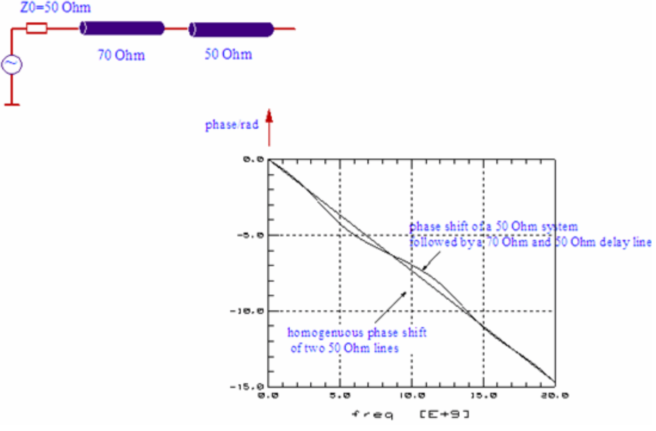
Case 1: impedance steps to lower value



Note
a step to lower impedance: -> more phase shift at low frequencies!

Case 2: impedance steps to higher value





> **Note**
> NOTE: a step to higher impedance:-> less phase shift at low frequencies!

As a final note, we can conclude that for lossless lines, there is no phase ripple, provided there are no multiple reflections (connector mismatch, impedance mismatch). If we encounter phase ripple, we have to take multiple reflections into consideration.

Q-Factor Calculation

The quality factor of a spiral inductor is given by:

$$Q = \frac{IMAG(\underline{Z}_{inductor})}{REAL(\underline{Z}_{inductor})}$$

The best way to get rid of the 50 Ohm terminations of 2-port S-parameters is to calculate 1-port S-parameters:

$$S_{11_1port} = S_{11} - \frac{S_{12} * S_{21}}{1 + S_{22}}$$

From that, we apply the basic S11 <> R conversion, mentioned in the chapter on Smith charts,

$$\underline{R} = Z0 \cdot \frac{1 + \underline{S_{11}}}{1 - \underline{S_{11}}}$$

and obtain for the input impedance at Port 1

$$Z_{11_1port} = Z_0 * \frac{1 + S_{11_1port}}{1 - S_{11_1port}}$$

From this, we can obtain the requested 1-Port characteristic like the Q factor of a spiral inductor:

$$Q = \frac{IMAG(\underline{Z}_{11_1port})}{REAL(\underline{Z}_{11_1port})}$$

Note
The same result is obtained when converting the S-parameters to Y-parameters, and then calculating.

$$Q = \frac{-IMAG(\underline{Y}_{11})}{REAL(\underline{Y}_{11})}$$

Small Signal Versus Large Signal S-Parameters

So far, we introduced the S-parameters and compared them to the other twoport parameters like Y or Z. This means, S-parameters are small signal parameters by definition. For a transistor as an example, the S-parameters do not reflect non-linear amplification phenomena like compression etc.

In general, twoport parameters of non-linear components like transistors or diodes vary as a function of input power. Referring to S-parameters, the parameters |S1l|2 and |S21|2 are defined as a function of power incident at port 1 with no power incident at port 2; whereas |S12|2 and |S22|2 are defined as a function of power incident at port 2. Therefore, in RF simulators like ADS, all S-parameters of nonlinear electrical elements are represented by a S-parameter matrix versus input power at port 1 for forward and port 2 for reverse operation.

This data set has been accepted as a convenient means of characterizing nonlinear devices by their large-signal S-parameters and have been successfully used for designing power amplifiers, oscillators, etc.

However, keep in mind that when the S-parameters become RF-power dependent, harmonics occur. In order to characterize these harmonic frequencies, a spectrum analyzer or a non-linear network analyzer should be applied!

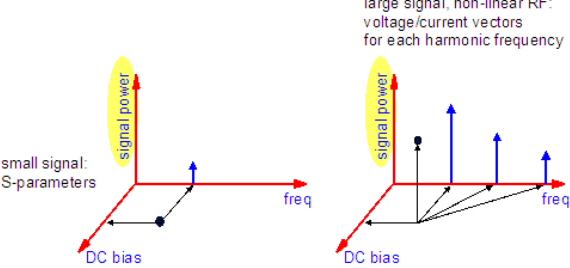
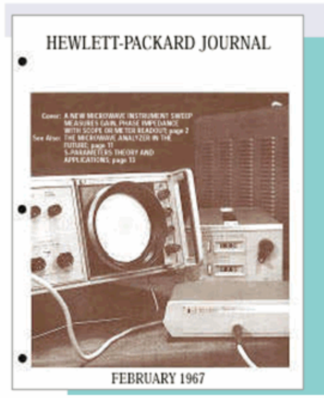


Figure: Small signal S-parameters are a function of frequency and, for transistors, also of bias, while large signal S-parameters also cover the signal power.

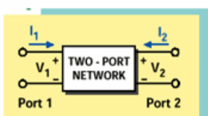
Note: even for the three-dimensional character of large signal S-parameters, they are by definition linear. This means, a single frequency power |al|2 injected into port1 will lead to reflected and transmitted power with exactly the same frequency. No harmonic frequencies!

S-Parameter Basics for Modeling Engineers



This is the cover of the February 1967 issue of the Hewlett-Packard Journal, yet its content remains important today. S-parameters are an essential part of high-frequency design, though much else has changed during the past 30 years. During that time, HP, now Agilent Technologies, has continuously forged ahead to create today's leading test and measurement environment.

Describing Twoport Networks



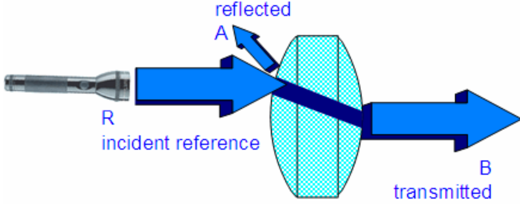
To characterize the performance of such a network, any of several twoport parameter sets can be used, each of which has certain advantages. Each parameter set is related to a set of four variables associated with the two- port model. Two of these variables represent the

excitation of the network (independent variables), and the remaining two represent the response of the network to the excitation (dependent variables).
Example: y parameters

$$\begin{pmatrix} I_1 \\ I_2 \end{pmatrix} = \begin{pmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{pmatrix} * \begin{pmatrix} V_1 \\ V_2 \end{pmatrix}$$

S-Parameter Analogy

"Scattering parameters", which are commonly referred to as S-Parameters, relate to the traveling waves that are scattered or reflected when a network is inserted into a transmission line.

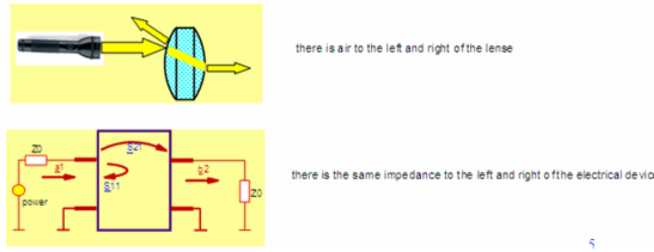


S-parameters are important in microwave design because they are easier to measure and to work with at high frequencies than other kinds of parameters. They are conceptually simple, analytically convenient, and capable of providing a goog insight into a measurement or design problem.

> **Note**
> S-parameters are a mean to characterize an electronic device, in a well-defined environment (i.e. terminated with the characteristic impedance Z0). To better understand this, we consider an example from optics, a lens. If your lens is in a well-defined environment, e.g. air, your optician can characterize its performance from measuring the reflection and throughput characteristics. If your lens would be put on water, and the other side remained in the air, such a measurement would be much more difficult. So, what is true for your spectacles, is true also for S-parameter. They allow to characterize a circuit in a well-defined environment, the characteristic impedance Z0. This means that S-parameters relate traveling waves (power) to a twoport's reflection and transmission behavior.

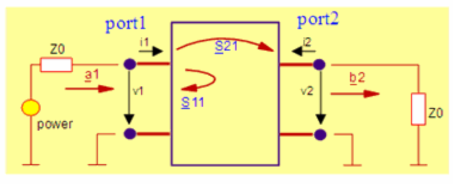
From Light Waves to S-Parameter 'waves'

- at high frequencies, voltages and currents are difficult to measure
- BUT: electrical high-frequency signals are represented by el.power. THEREFORE: when light power works well for optics, why not using electrical power for RF characterization?



> **Note**
> By injecting electrical power to the device, we will get the AC currents and AC voltages at the terminals of the device by simple calculations, as will be seen later.

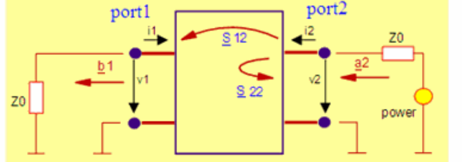
Considering a 2-port: stimulated at port1, then at port 2



The transmitted and reflected PowerWaves are:

$$a1 = \frac{1}{2\sqrt{Z0}}(v1 + i1 \cdot Z0) \quad *$$

$$b2 = \frac{1}{2\sqrt{Z0}}(v2 - i2 \cdot Z0)$$

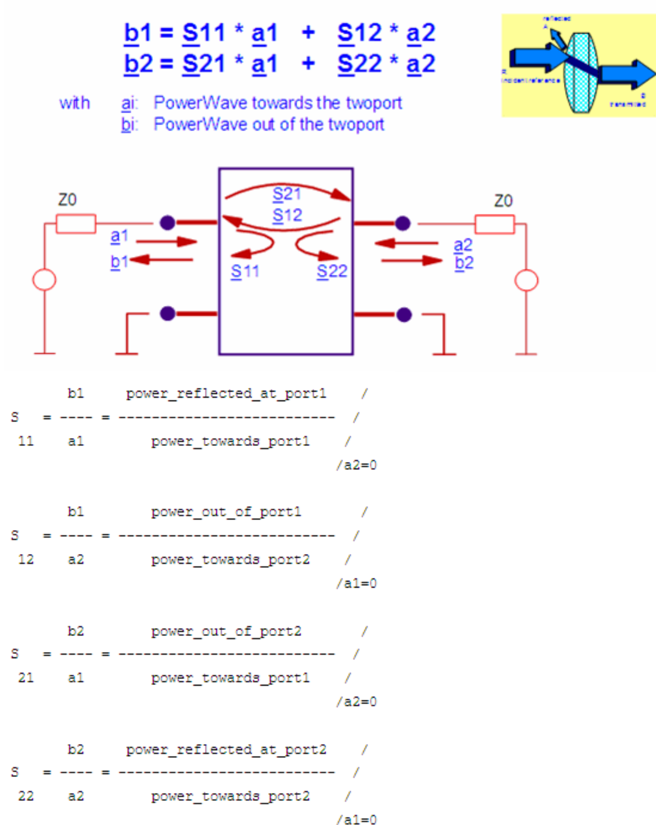


$$a2 = \frac{1}{2\sqrt{Z0}}(v2 + i2 \cdot Z0) \quad *$$

$$b1 = \frac{1}{2\sqrt{Z0}}(v1 - i1 \cdot Z0)$$

> **Note**
> During S-parameter measurements, the power sources provide the RF signal to the twoport. The voltages v1, i1, v2 and i2 are therefore a consequence of that stimulating power source. On the other hand, the above plot explains how they can be calculated out of the measured or simulated S-parameters.

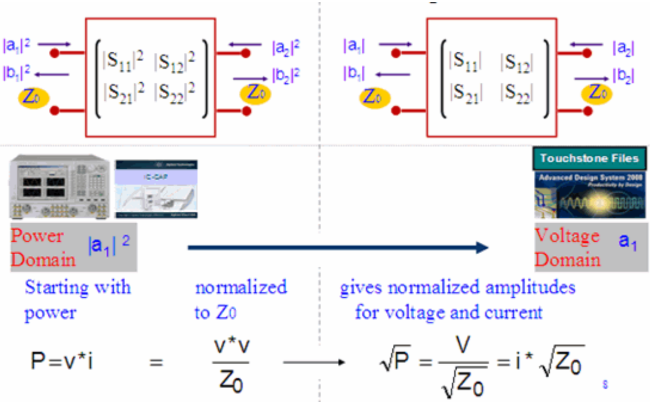
The S-Parameters relate the b(out of the device) PowerWaves to the a (into the device) PowerWaves




```
      b1      power_reflected_at_port1     /
S   = ---- = -------------------------    /
 11   a1         power_towards_port1      /
                                          /a2=0

      b1         power_out_of_port1        /
S   = ---- = -------------------------    /
 12   a2         power_towards_port2      /
                                          /a1=0

      b2         power_out_of_port2        /
S   = ---- = -------------------------    /
 21   a1         power_towards_port1      /
                                          /a2=0

      b2      power_reflected_at_port2     /
S   = ---- = -------------------------    /
 22   a2         power_towards_port2      /
                                          /a1=0
```

The conditions a1=0 and a2=0 are fulfilled with a impedance of Z0 (typ. 50 Ohm) in series with the stimulating voltage source at the input, and with an output termination of again Z0.

S-Parameters in Power or Voltage Domain



Note
S-Parameters are unit-less (they are relations of 'waves').
These 'waves' can be interpreted as 'power waves', or 'voltage waves'. And this makes the comparison between measurement results shown on the instrument's screen and results displayed e.g. in ADS Data Display a bit confusing.

As shown above, instruments display the results as power waves -> you can make a simple test: measure an OPEN standard at the NWA's port1, and then connect a resistor with 150 Ohm. MAG(S11) drops from 1 to ½. On the NWA display, when selecting 'LOG MAG', the trace drops by 3dB.

Simulators (like ADS), and S-parameter data files (e.g. Touchstone), apply the voltge wave interpretation-> the experiment above, performed using an ADS S-parameter simulation, drops by 6 dB !!

Therefore, in IC-CAP Plots, we do not apply the IC_CAP built-in LOG10 scale, which is based on 20*LOG10(MAG()), but specify by ourselves 10*LOG10(MAG()).
This way we get the same Plot result on the NWA display and in IC-CAP (what is important when e.g. verifying the calibration etc).

----
dB Voltage Waves: 20*LOG10(MAG())
dB Power Waves: 10*LOG10(MAG())

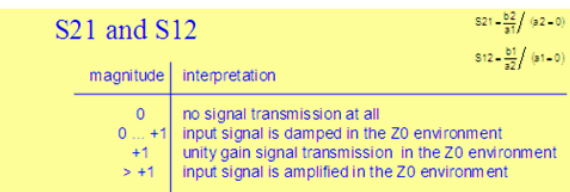
Interpreting Some Important S-Parameter Values

S11 and S22

$S11 = \frac{b1}{a1} / (a2 = 0)$

$S22 = \frac{b2}{a2} / (a1 = 0)$

value	interpretation
-1	all voltage amplitudes towards the twoport are reflected and inverted (0 Ω)
0	impedance matching, no reflections at all (50 Ω)
+1	voltage amplitudes are reflected (infinite Ω)



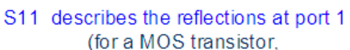
S21 and S12

$S21 = \frac{b2}{a1} / (a2 = 0)$

$S12 = \frac{b1}{a2} / (a1 = 0)$

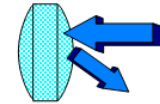
magnitude	interpretation
0	no signal transmission at all
0 ... +1	input signal is damped in the Z0 environment
+1	unity gain signal transmission in the Z0 environment
> +1	input signal is amplified in the Z0 environment

Reflection S-Parameters S11 and S22



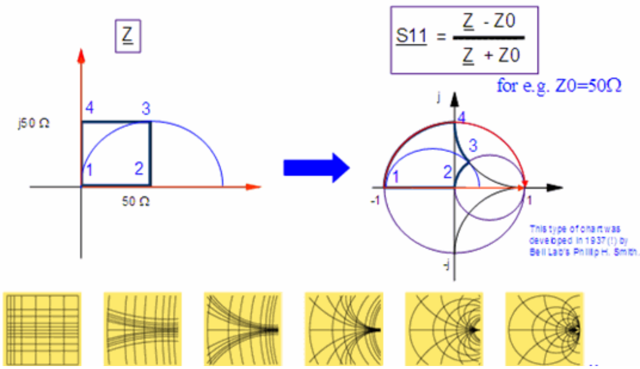
S11 describes the reflections at port 1
(for a MOS transistor,
mostly $C_{Gate\text{-}Source}$, $R_{Gate_ext\text{-}Gate_int}$)

and



S22 the reflections at port 2
(for a MOS transistor,
mostly $C_{Drain\text{-}Source}$, $R_{Drain\text{-}Source}$, $C_{Drain\text{-}Gate}$)

Applying the Smith Chart to S11 and S22



What makes S_{xx}-parameters especially interesting for modeling, is that S11 and S22 can be interpreted as complex input or output resistance of the twoport (including the termination at the opposite side of the twoport with Z0 !!). That's why they are usually plotted in a Smith chart.

The Smith chart is a transformation of the complex impedance plane Z into the complex

reflection coefficient, i.e. Sxx, following:

!spar_figequation5.gif!with the system's reference impedance Z0 = 50 Ω.

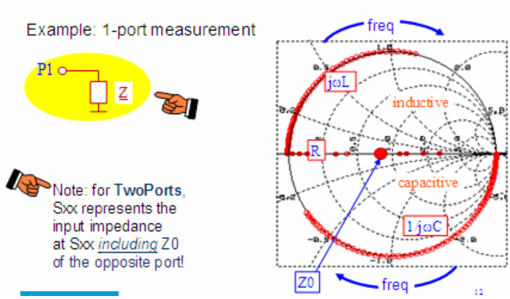
This means that the right half of the complex impedance plane Z is transformed into a circle in the Sxx-domain with radius '1' .
In order to get more familiar with interpreting the Smith chart, let's consider a square with the corners (0/0)Ω, (50/0)Ω, (50/j50)Ω and (0/j50)Ω in the complex impedance plane 'Z' and its equivalent in the Smith chart with Z0=50Ω. Watch the angel-preserving property of this transform (rectangles stay rectangles close to their origins). Also watch how the positive and negative imaginary axis of the Z plane is transformed into the Smith chart, and where (50/j50)Ω is located in the Smith chart. Also verify that the center of the Smith chart represents Z0, i.e. for Z0 = 50Ω, the center of the Smith chart is (50/j0)Ω.

Some more:

Sxx	interpretation
-1	all voltage amplitudes towards the twoport are inverted and reflected (0 Ω)
0	impedance matching, no reflections at all (50 Ω)
+1	voltage amplitudes are reflected (infinite Ω)

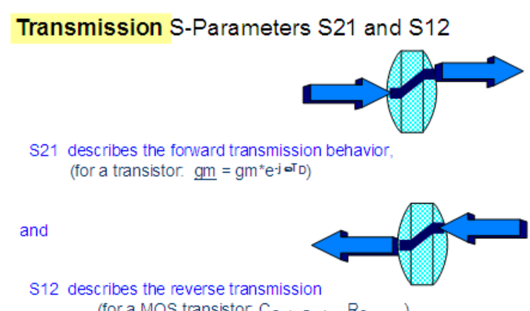
Note
The magnitudes of S11 and S22 are always ≤ 1. Do not forget that included in Sxx is the termination at the opposite side of the two-port, usually Z_0

Interpreting the Smith chart of S11 and S22



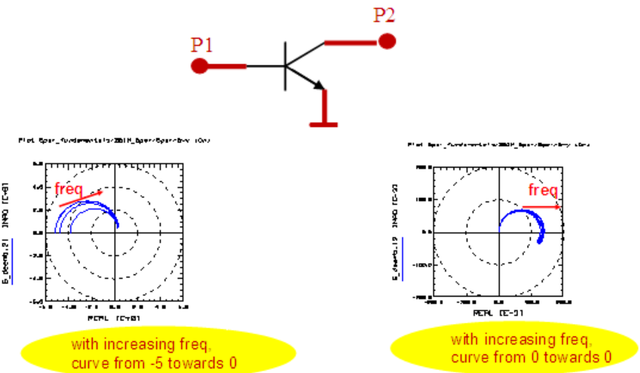
Note
for TwoPorts,Sxx represents the input impedance at Sxx including Z0 of the opposite port. The slide above gives Smith chart plots for different electrical components connected to a single port of a vector network analyzer (VNA). For an ohmic resistor, all measurement results are located on the x-axis, R<Z0 is to the left of the center, R>Z0 to the right. From the scaling of the Smith chart (x = 0 ..1.. infinite), we can calculate the value of the resistor as x*Z0.For inductors and capacitors, however, the measurement result is frequency dependent and follows the sketched traces. Note that the curves always turn to the right with increasing frequency.
We consider the component between one port and ground. For a component between the ports, the 50 Ohm of the opposite port always show up together with the component's impedance Z

Transmission S-Parameters S21 and S12



Interpreting the Polar Plots of S12 and S21 (Sij)

Interpreting the Polar Plots of S12 and S21 (Sij)

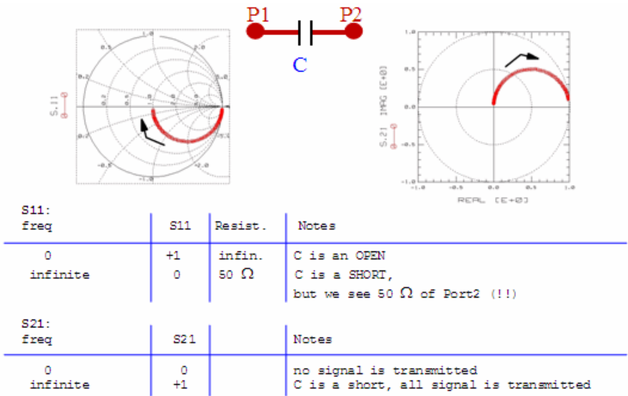


S-Parameter Interpretation at a glance

Sxx	**Input/Output Resistance of the Twoport up to and Including Port 2/Port 1**
Sxy	**Power Transmission of the Twoport in the Z0=50 Ω System**

> **Note**
> Note: when interested in the twoport's input/output resistance alone, without the char.impedance of the opposite VNA port, use the Y parameters.

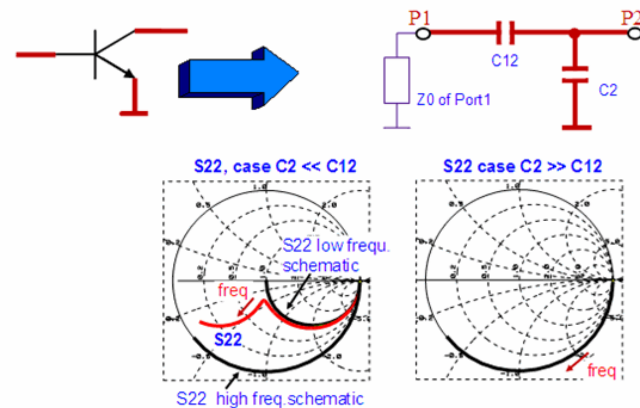
Typical S-Parameter Traces



S11:

freq	S11	Resist.	Notes
0 infinite	+1 0	infin. 50Ω	C is open C is short but we see 50 Ω of Port2
S21			
freq	**S21**	**Notes**	
0 infinite	0 1+	no signal is transmitted C is a short, all signal is transmitted	

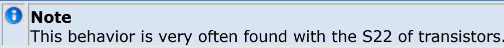
A Specific Sxx Parameter Trace (very often seen with transistors)



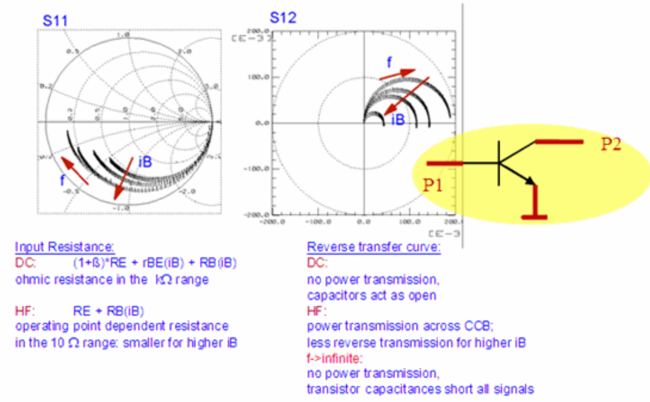
For the schematic above, we can distinguish between two cases: C12 >> C2and C12 << C2

In the first case, we have Ceff~ C12, and the influence of C2 can be neglected. The expected S22 curve will start at 'infinite' resistance, and lead towards 50 Ohm for higher frequencies (Port1 impedance Z0). Yet for higher frequencies, C2 will become more and more conductive and will make the power flowing from Port2 more and more directly to ground. Therefore, Z0 of port1 will become 'invisible'. The S22 curve will change its high frequency end value, and will now tend towards 0Ohm for infinite frequency. Between both ranges, we expect a transitional trace of S22.

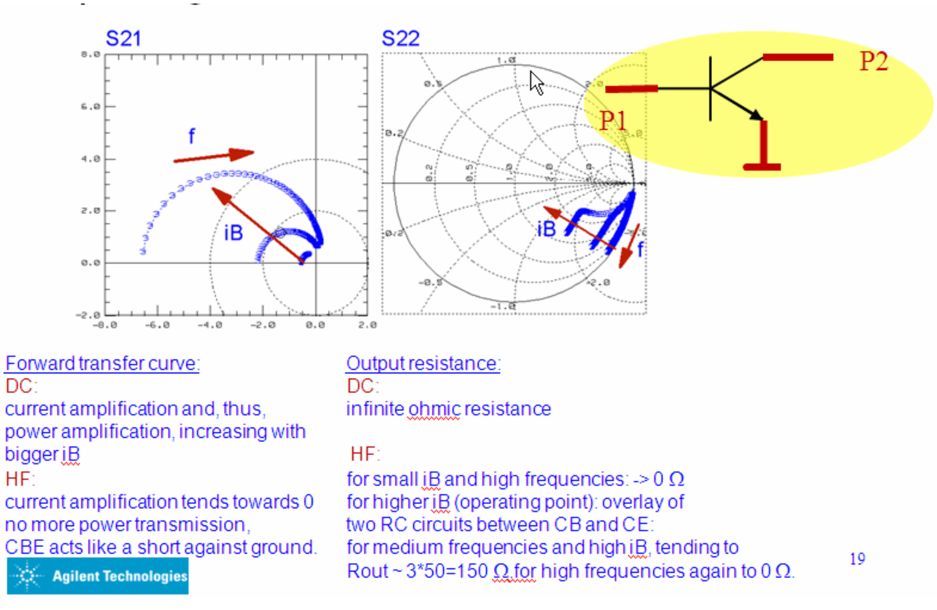
Not to the case C2 >> C12: in this case, the 'easier' path for the signal, for all frequencies, will be from Port2 directly to Ground. C12 and Z0 of port1 will not be visible in the Smith chart.


Note
This behavior is very often found with the S22 of transistors.

Interpreting Transistor S-Parameter Traces



Interpreting Transistor S-Parameter Traces

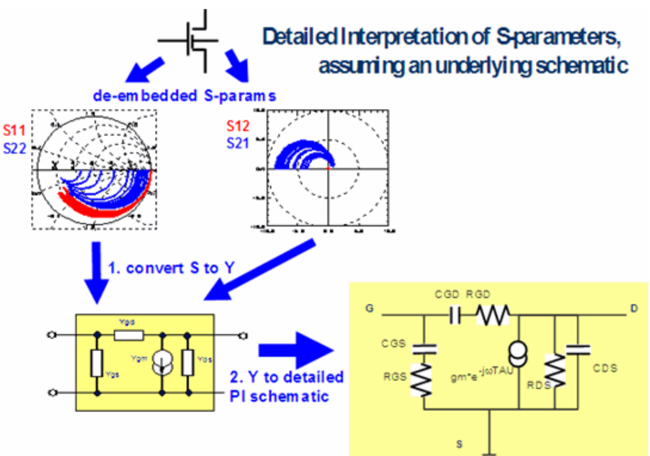


S-Parameter Interpretation Basics

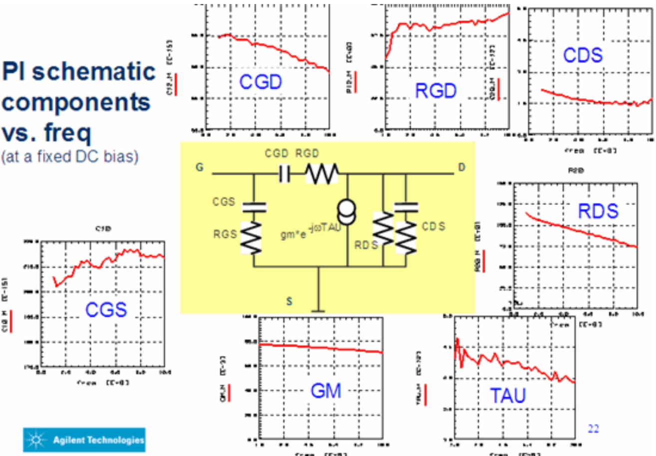
For frequency, 0 -> infinite,
Sxx Parameters of RC or RL circuits are half-circles
Sxx Parameters of CL circuits are full circles (resonance)
Sxy Parameters for active components have a magnitude >1 (power amplification in a 50 W system)

All S-parameter curves turn clock-wise!
Sxx and Sxy Parameters starting points at lowest frequency are identical to the DC measurement.

Detailed Interpretation of S-parameters, assuming an underlying schematic



PI schematic components vs. freq (at a fixed DC bias)



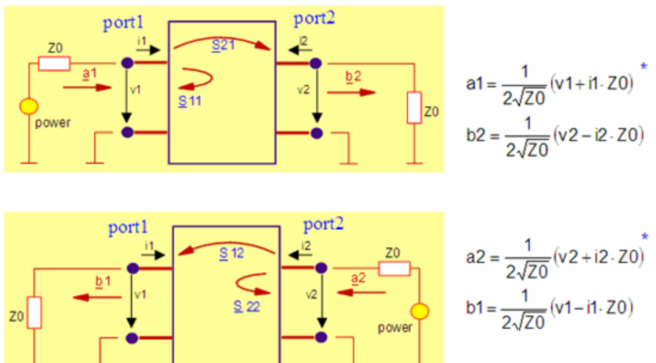
Conclusion

S-parameters are specific twoport parameters,

- for linear circuits
- related to power instead of voltage/current
  As a very important last remark, S-parameters are members of the LINEAR Twoport Theory.
  In other words, a prerequisite of applying them to nonlinear devices like diodes and transistors is to make sure that these nonlinear devices have been measured in small signal excitation. I.e. the applied RF signal level has to be small enough to avoid any non-linearities. This means, only the stimulating frequency is allowed to be present in the measured S-parameters, and no harmonics of this stimulating, fundamental frequency.

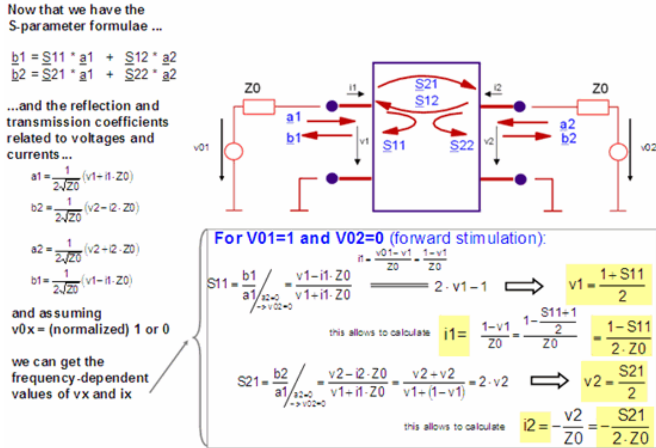
For Experts

Coming back to the transmitted and reflected PowerWaves



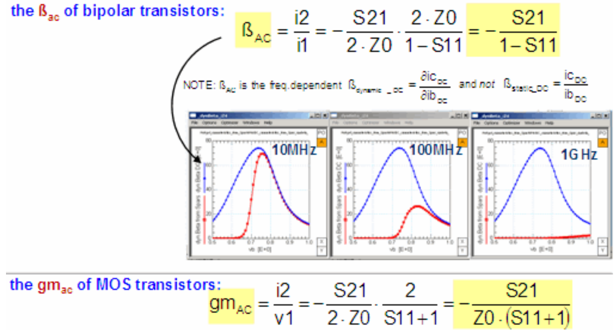
Note
During S-parameter measurements, the power sources provide the RF signal to the twoport. The voltages v1, i1, v2 and i2 are therefore a consequence of that stimulating power source. On the other hand, the above plot explains how they could be calculated out of the measured or simulated S-parameters.

Getting vx and ix from S-parameters



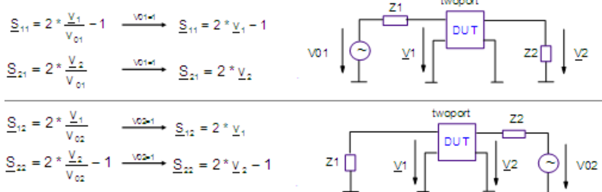
The normalization of v01=1 and v02=1 allows to simplify the above equations. It corresponds to an AC stimulus at Port1 or Port2 of 1V AC signal. This is OK because S-Parameters are linear, and represent the tangent to the DC curve (with 50 Ohm termination load) in the operating point.

and now that the Vx and Ix are known from S-parameters. we can calculate some transistor characteristics which are based on relations like:



Note
Due to the normalization to v01=1 and v02=2, we cannot get the absolute voltages v1, v2 and the currents i1, i2 from the S-Parameters, but we can easily get characteristics like ß and gm which are the quotient of ix and vx.

and inversely, we can also calculate S-Parameters out of a voltage simulation



Again, because S-Parameters are linear, we can scale the applied voltage to any value, and get the corresponding (linearly) scaled other voltage and currents. Therefore, setting V01=1V and V02=1V is pretty practical, because it just simplifies the above equation.

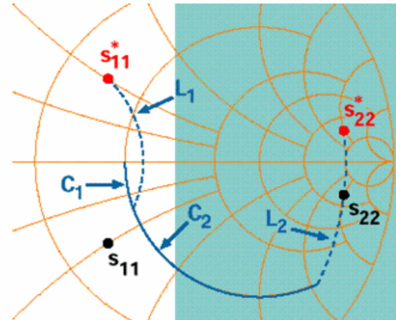
S-Parameter Impedance Matching (for obtaining max. power gain)

This is an excerpt from HP App.Note AN 95-1

In order to amplify the maximum gain, we will now discuss to synthesize matching networks that will transform the source and load impedances Z_0 to the impedances corresponding to reflection coefficients of S11* and S22* , respectively, i.e. the conjugate complex S11 and S22. This can be done for a single frequency or a frequency band.

For simplicity, let's consider a matching for a single frequency. In this case, the matching networks need not be complicated. Simple series capacitor, shunt inductor networks will not only do the job, but will also provide a handy means of biasing the transistor, via the inductor, and of isolating the DC bias from the load and the source.

Values of L and C to be used in the matching networks for the small band amplifier at e.g. 300MHz are determined using the Smith Chart below.



Smith Chart for 300- MHz amplifier design example.

> **Note**
> See AN95-1 for an interactive movie on these manipulations.

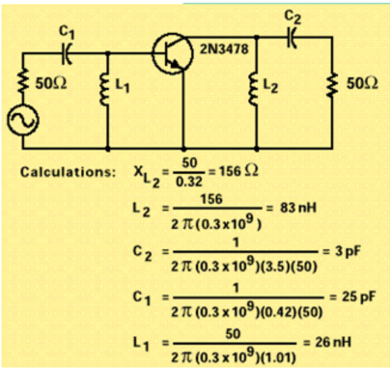
First, points corresponding to S11 ,S11* ,S22 ,and S22* at 300 MHz are plotted in the Smith chart. Each point represents the tip of a vector leading away from the center of the chart, its length equal to the magnitude of the reflection coefficient being plotted, and its angle equal to the phase of the coefficient. Next, a combination of constant- resistance and constant-conductance circles is found, leading from the center of the chart, representing Z_0 , to S11* and S22* . The circles on the Smith Chart are constant-resistance circles; increasing series capacitive reactance moves an impedance point counter-clockwise along these circles.

In this case, the circle to be used for finding series C is the one passing through the center of the chart, as shown by the solid line in the figure above. Increasing shunt inductive susceptance moves impedance points clockwise along constant-conductance circles. These circles are like the constant-resistance circles, but they are on another Smith Chart, one that is just the reverse of the chart shown! (This is a Smith chart for 1/Z !) The constant-conductance circles for shunt L all pass through the leftmost point of the chart rather than the rightmost point. The circles to be used are those passing through S11* and S22* , as shown by the dashed lines in the Smith chart above.

Once these circles have been located, the normalized values of L and C needed for the matching networks are calculated from readings taken from the reactance and susceptance scales of the Smith charts.

Each element's reactance or susceptance is the difference between the scale readings at the two end points of a circular arc. Which arc corresponds to which element is indicated in the Smith chart above.

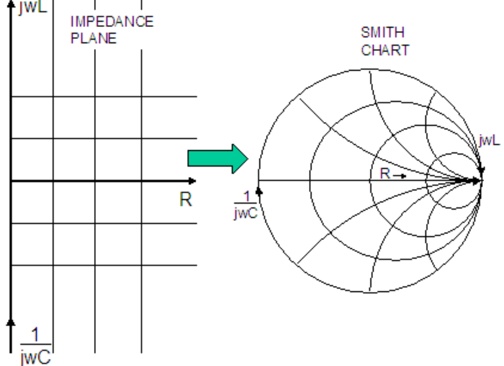
The final network and the element values, normalized and un-normalized, are shown below:



A 300-MHz amplifier with matching networks for maximum power gain.

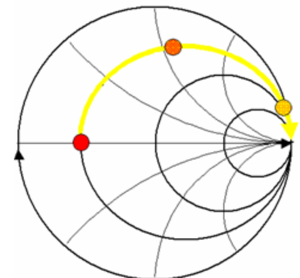
The following small tutorial gives some additional info about impedance matching:

We commence with the already discussed transform of the complex impedance plane into the Smith chart plane, see the figure below.

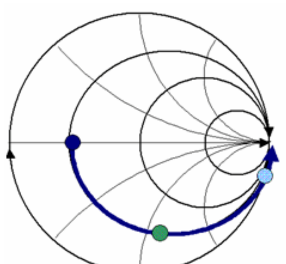


We remember that:

- On the impedance plane, with respect to increasing frequency, all curves turn upwards (for a given, constant L,C,R). Referring to the Smith chart, they therefore always turn clockwise
- For a locus within the lower capacitance half-plane, only a series impedance can bring the locus to the upper, inductive impedance half-plane
- Adding a series capacitance to an impedance within the lower capacitance half-plane, can bring the locus only to the x-axis at the most, even with C=infinite.
  When adding a series inductor:
- a zero inductor in series with a resistor does not affect the locus
- a medium inductor in series with this resistor results in a bigger total impedance
- a L=infinite gives an open circuit.

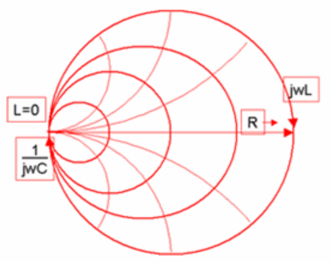


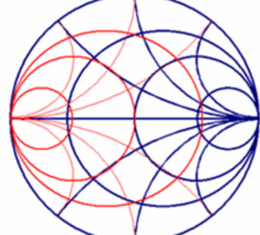
When adding a series capacitor
- an infinite capacitor in series with the resistor does not affect the locus
- a medium capacitor in series results in a bigger total impedance
- a C=0 gives an open circuit



As a next step, we define an auxiliary overlay chart to reflect Smith chart locus shifts of

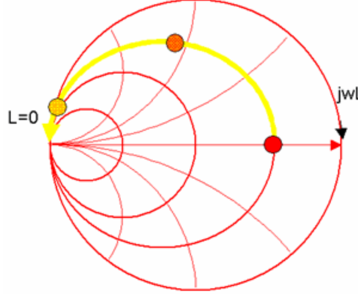
parallel components. After the locus shift has been determined, the resulting impedance is then read from the original impedance Smith chart.
Again, it should be noted that when working with the auxiliary overlay chart, read the resulting impedance from the original impedance Smith chart.



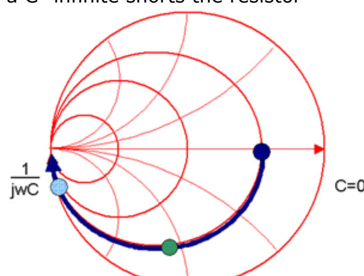


When adding a parallel inductor

- a L=infinite inductor in parallel with a resistor does not affect the locus
- a medium inductor in parallel with this resistor results in a smaller total impedance
- a L=0 shorts the resistor



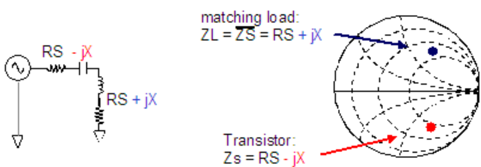
When adding a parallel capacitor
- a C=0 capacitor parallel with a resistor does not affect the locus
- a medium capacitor in parallel with this resistor results in a smaller total impedance
- a C=infinite shorts the resistor



After these pre-remarks, we are now ready to consider the impedance matching procedure:

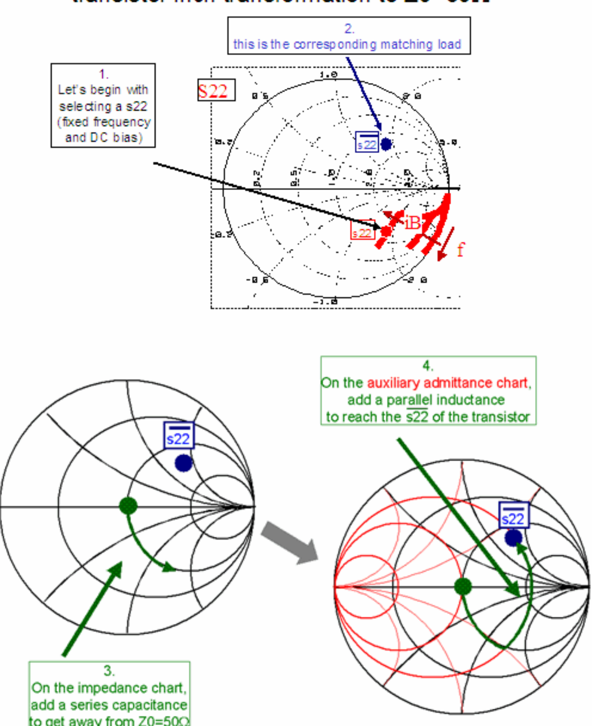
Power Transfer Efficiency

For complex impedances,
maximum power transfer occurs when ZL = ZS* (conjugate match)

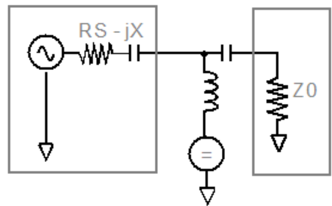


However, we want to stay in the Z0=50Ω environment!
I.e. Z0 has to be transformed to $\overline{ZS}$,
the conjugate complex output impedance of the transistor

Output Matching Network for S22 of a transistor incl. transformation to Z0=50Ω



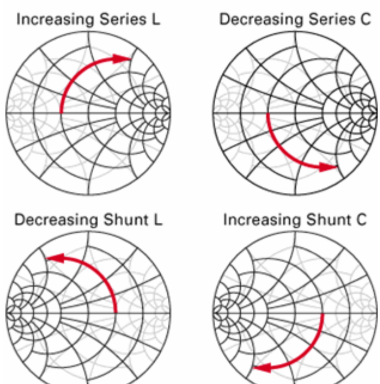
The resulting matching network for the output of the transistor is therefore:



The following is an excerpt of the Agilent web page
http://www.tm.agilent.com/data/static/eng/tmo/Notes/interactive/an-95-1/classes/imatch.html
(as of Jan.2001):

Smith Chart Circuit Elements Paths

You can think of impedance matching using the Smith Chart as driving a car to a specific destination in Smith Town - a city were none of the streets are straight! By adjusting circuit components in appropriate order, we can constrain the circuit reflectance to paths along constant resistance or constant conductance circles. Just like road signs can direct a car along the circular streets of Smith Town, so can we reach the matching impedance condition in a straightforward and deterministic way. The graphs below demonstrate how the various shunt and series L and C components change the circuit reflectance on the Smith Chart. Assuming the given component is the last component in the matching network, the circuit reflectance will move as indicated along constant resistance or constant conductance circles.

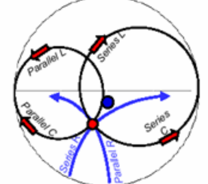


Matching means:

Moving toward the center of the Smith Chart!

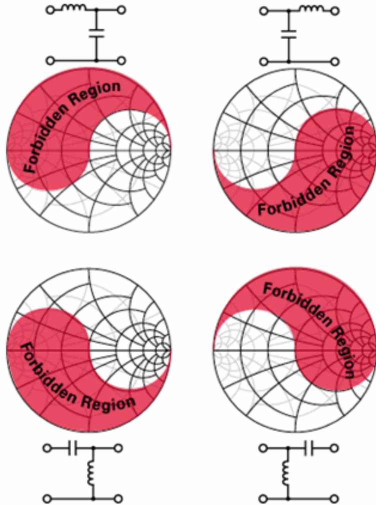
Add Series or Parallel (shunt) components.

Adjust the value to move toward open, short, L, C, or center of chart.



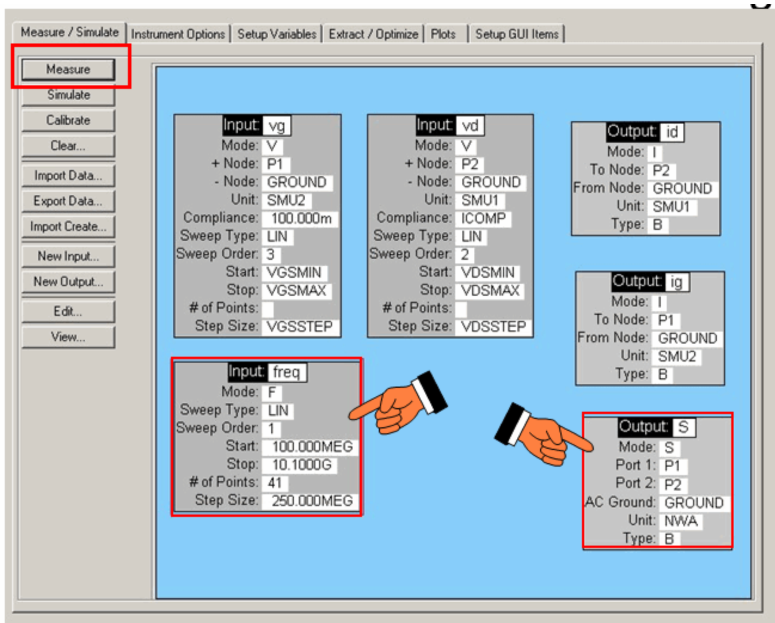
Forbidden Regions of the Smith Chart

For a given load reflectance, only certain L-C matching networks will be capable of transforming the source impedance to the load impedance. In fact, for any load reflectance, exactly two of the four possible L-C matching networks in the Transistor Amplifier-II model above will be able to do the matching job. But which two?
The charts below can be used to determine which matching networks will work in a given load situation. If the load reflectance lies within the forbidden region of the Smith Chart for the indicated matching network, then that network cannot perform the required matching operation. You cannot drive your car into the forbidden neighborhoods of Smith Town! They are unpaved!



Use these charts to determine which matching network should be used. First, visually locate the position of the load reflectance from your inner model on each of the four color Smith Charts. Then, eliminate the two networks whose forbidden regions overlap the reflectance point, and use one of the remaining two networks to perform the impedance match.

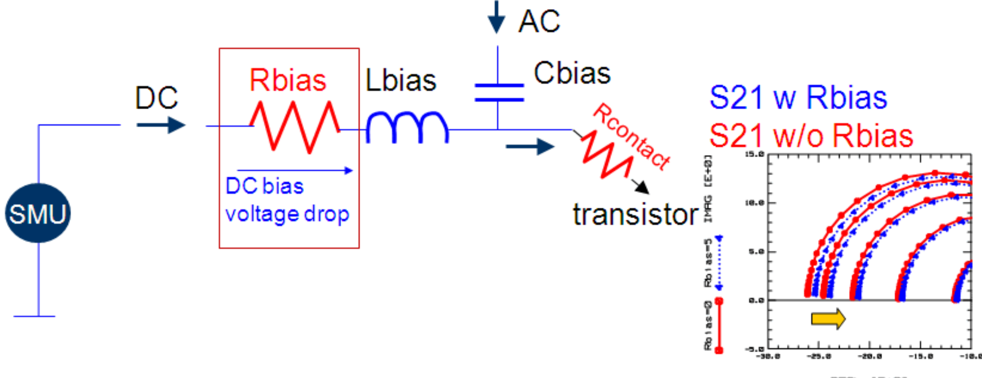
S-Parameters Measurements for Modeling



10 Rules for Successful S-parameter Handling

> **Note**
> This paper is a collection of experiences with making reliable S-parameter measurements. The underlying IC-CAP model files are located at:
> demo_features\3_MEAS_ORGANIZE_n_VERIFY_DATA\0_MASTER_FILES and further
> 20_NWA_CAL_VERIFICATIONselect CAL_VERIFY_PNA_2port_MASTERFILE_demodata.mdl or
> CAL_VERIFY_genl_NWA_2port_MASTERFILE_demodata.mdl for the older HP/Agilent NWAs.
> and 31_DEEMB_VERIFICATION\
> CHECK_OPEN_SHORT_DEEMB_with_a_THRU_demodata__MASTERFILE.mdl

DC Bias Voltage Drop in the VNA Bias TEEs



The NWA bias TEEs, internal or external, represent ohmic losses (Rbias) for the DC bias, and -together with the DC loss of Rcontact-this loss is not calibrated out by the NWA(!)i.e, the bias voltages at the device will be less than provided by the SMUs. This will shift the starting point of the S21 curves towards 0.
During the S-parameter measurements, the transistor is DC-biased by bias TEEs. Typical values for the ohmic losses in the DC path of NWA-built-in bias-TEEs, are ~1Ohm (PNA/ENA), ~2Ohm (older HP/Agilent NWAs). External bias-TEEs may have lower loss, but mostly suffer from a limited frequency range (upper and also (!!) lower frequency*) for the HF. In any case, the DC losses are not calibrated out by the NWA, since the NWA calibration "does not see" them (DC is not included in the NWA signal path).
Additionally, especially for silicon devices, there is a non-negligible DC contact loss, additionally to the losses in the bias-TEE. This contact problem is especially true for aluminum contact pads. Although, AC-wise, this loss is calibrated out (it is included in the AC signal path!), it is shifting the DC operating point.
For device modeling, if there is a problem with the fitting of the simulated S-parameters and the measurement curves for lowest frequencies, especially for bias conditions with higher currents, this total DC bias voltage drop (in the AC/DC bias TEE and the contact loss) must be considered and the simulation circuit must be enhanced with these resistors !Typically, using the Test Circuit of the DUT in the IC-CAP ModelFile. In order to not affect the HF S-parameter simulations, however, the resistors must be shorted AC-wise by big

capacitors. Typically, their value is C=100mF.

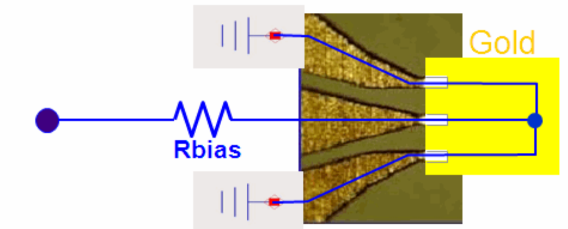
IC-CAP File: In demo_features\3_MEAS_ORGANIZE_n_VERIFY_DATA 0_MASTER_FILES\31_*DEEMB_VERIFICATIONopen the file CHECK_OPEN_SHORT_DEEMB_with_a_THRU_demodata*_MASTERFILE.mdl and see DUT check_DC_contact_Spar.

> **Note**
> For the lowest frequency, the bias-TEE (representing for the RF signal a resonance consisting of its C and L) must have **completely** finished its own 360' low-frequency phase turn.

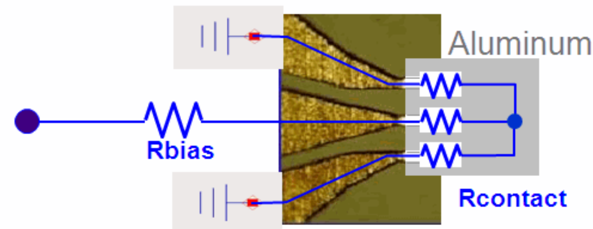
Evaluate the DC bias losses

Bias-TEE resistance Rbias and Probe Contact Resistance Rcontact

1. Contact the SHORT on the ISS calibration substrate (Gold) and measure the resistance Rbias



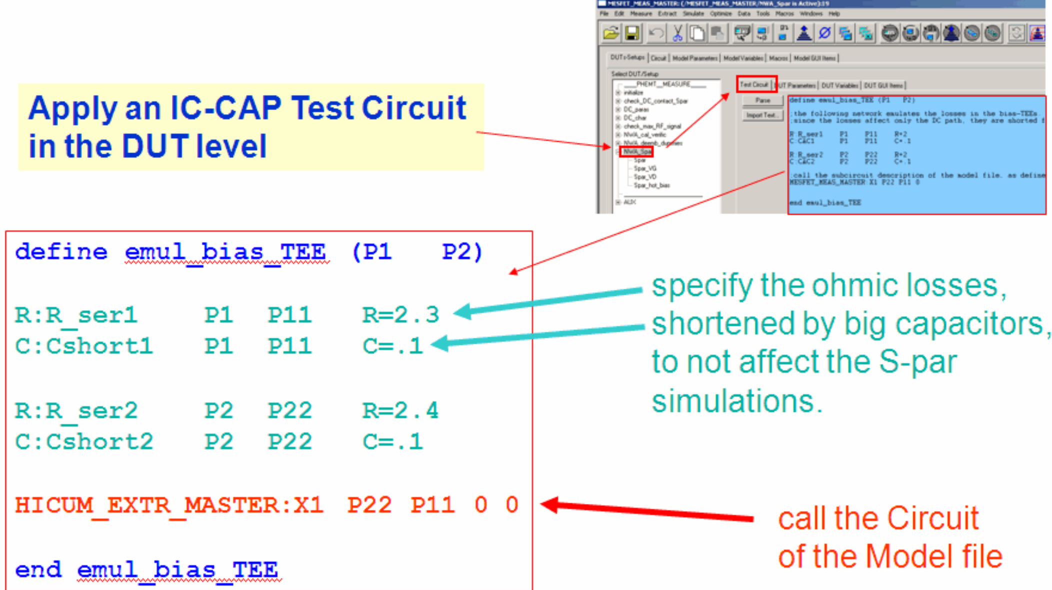
2. Contact the SHORT Dummy on the wafer and measure the resistance Rcontact



How to measure Rbias and Rcontact:

1. In a first measurement setup, we characterize the ohmic losses in the bias TEEs of the S-parameter testset (network analyzer) or in the external bias TEEs. For the measurement, place your GSG probes on the SHORT standard of the ISS substrate (gold contacts). Assuming an ideal contact on the golden SHORT of the ISS substrate, we measure in this setup basically Rbias.
   A typical value for Rbias is ~1Ohm (PNA), ~2Ohm (older NWAs), when using the internal NWA S-Parameter Testset, and ~1Ohm when applying external bias TEEs.
2. In a second step, we place the HF probes on a piece of aluminum on the wafer, or we contact a SHORT on the wafer. Either the 'de-embedding Dummy' on the wafer, or any location on the wafer with enough aluminum area to place the probes for a good short. Assuming 3x the same contact resistance for the GSG probe, we measure in this setup basically Rtotal = Rbias + 1.5 * Rcontact.Rbias is known from the previous ISS contact measurement. So we get for Rcontact:Rcontact = (Rtotal - Rbias) / 1.5
   A typical value for Rcontact is ~1-2 Ohm for silicon, and nearly 0 Ohm for GaAs (gold contacts!).

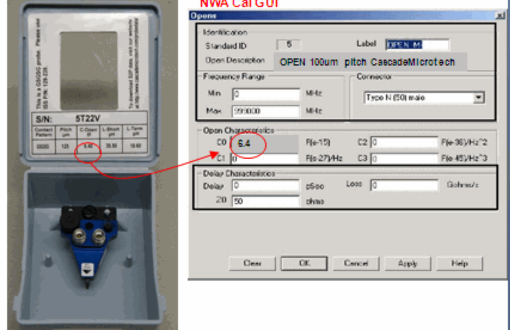
Specify The Probe Contact Resistance In IC-CAP



The network shown above -and specified as a 'DUT Test Circuit' in the IC-CAp ModelFile, emulates the losses in the bias-TEEs and the wafer contact losses. Since the losses affect only the DC path, they are shorted for the AC signals by the capacitors C_short_x
Then comes the line which calls the model, i.e. the subcircuit description of the transistor, as defined in Tab 'Circuit' of the IC-CAP model file
IC-CAP File: In your extraction files, in the DUTs which deal with S-parameters, add the above netlist to the Test Circuit tab.
Example:
demo_features\1_BASIC_MDLG_EXAMPLES\21_Gummel_Poon_demo\
GP_CLASSIC_NPN_Mdlg_Navigator_PELdep.mdland the Test Circuit in DUT 'Spar_mdlg'.

Tell the NWA which calibration standards will be applied "Modifying a Calkit"

Since we will probe on a wafer, and perform a calibration down to the GSG probe tips, we need to 'tell' the NWA about these probe tips and their corresponding ISS calibration substrate. Otherwise, the NWA assumes we calibrate with '3.5mm connector 'cal.standards.



The above screenshot is from the Agilent PNA series of network analyzers. Enter the calkit data, given e.g. on the inside cover of the Cascade_Microtech probes box, or from the spec.sheet of the PicoProbe or Suss probes.

ENTERING THE CALKIT DATA INTO THE NWA

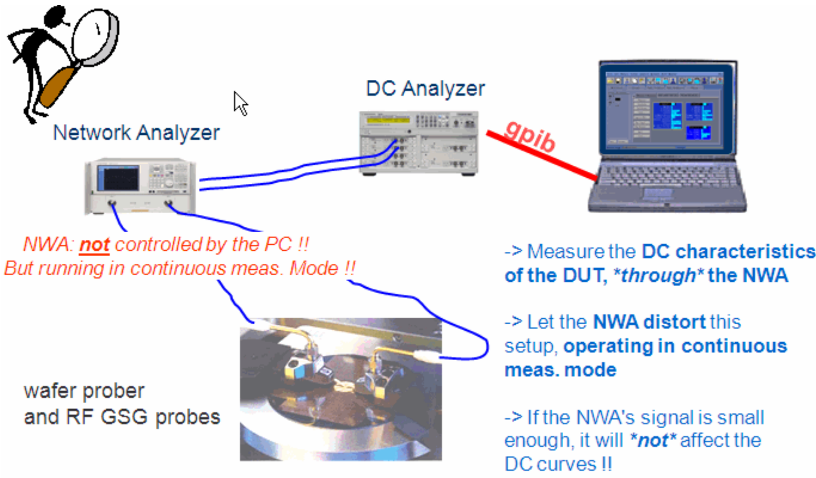
- Table:Example of Standard definitions for Cascade Microtech G-S-G probes, 100um pitch*

STANDARD No.	TYPE	C0/fF or L0/pH	Delay psec	Z0 Ohm	Loss Ohm/s	FREQ min max	STANDARD Label in the calkit
1	OPEN	-9.3	0	50	0	0 999	OPEN
2	SHORT	2.4	0	50	0	0 999	SHORT
3	LOAD		-0,007	500	0	0 999	LOAD
4	THRU		1	50	0	0 999	THRU

These nonidealities have to be entered into the NWA. This is called defining or 'modifying the calkit'. Check the NWA manual about how to do this.
Some notes about entering the calkit data: While in the case of a CV meter, the calibration corrects simply for an ideal offset capacitor, a NWA calibration uses several cal standards (OPEN, SHORT, LOAD, THRU etc.) from a cal kit. These cal standards do not represent ideal standards. They represent the real, existing standard, including its nonidealities! It means that a SHORT is not an ideal SHORT,

but instead represents rather an inductance. The same applies to the THRU, which has a non-zero delay time. The OPEN corresponds rather to a capacitor than to an ideal OPEN. Therefore, these nonidealities of the G-S-G probes have to be entered into the NWA before calibration. This is called 'modifying the cal kit'.
While this procedure refers to the nonidealities of the calibration standards, for all frequencies and independent of the applied power, the subsequent calibration based on these calkit data is related to the selected frequency range, the RF power, the applied IF bandwidth etc. After it has been performed, the correction terms are stored in the cal set of the NWA. In other words, the 12-term error vectors are 'filled up'.
Afterwards, when the measurement is performed, the raw measured data arrays will be corrected using a correction technique related to the selected calibration method, and referring to the specified cal set. Finally, this corrected measurement result is displayed on the NWA monitor. Therefore, after the calibration, a re-measurement of the OPEN will not represent an ideal open, but instead exactly those parasitic components as described in the documentation of the OPEN. In the same way, a THROUGH shows up after calibration with its real delay time, and a SHORT exhibits its specified inductive behavior.

Find Out The Max. Applicable RF Signal For Your Device



CHECKING THE CORRECT NWA PORT POWER SETTINGS
It is very important to not overdrive a nonlinear device like a transistor with too much RF signal. It must be absolutely ensured that particularly the transistor's output signal is not clipping. When this happens, the transistor output signal is no longer a sine function, i.e. the transistor behaves nonlinear. And this is a contradiction to a linear NWA measurements! In this case, the NWA will only measure the fundamental frequency, and ignore the occurring harmonics!
At the same time, when harmonics occur, the DC bias is usually affected and shifted. This is due to Kirchhoff's law: the sum of currents into a node (for all frequencies) must be zero. So, if there appears 'energy' in the harmonics, the 'energy' in the DC is usually affected too.
In this context, keep in mind that:-> a typical RF signal level for modeling transistors is -40dBmi.e. the NWA port1 is set to -30dBm, plus typically ~10dB attenuation of cables and connectors.

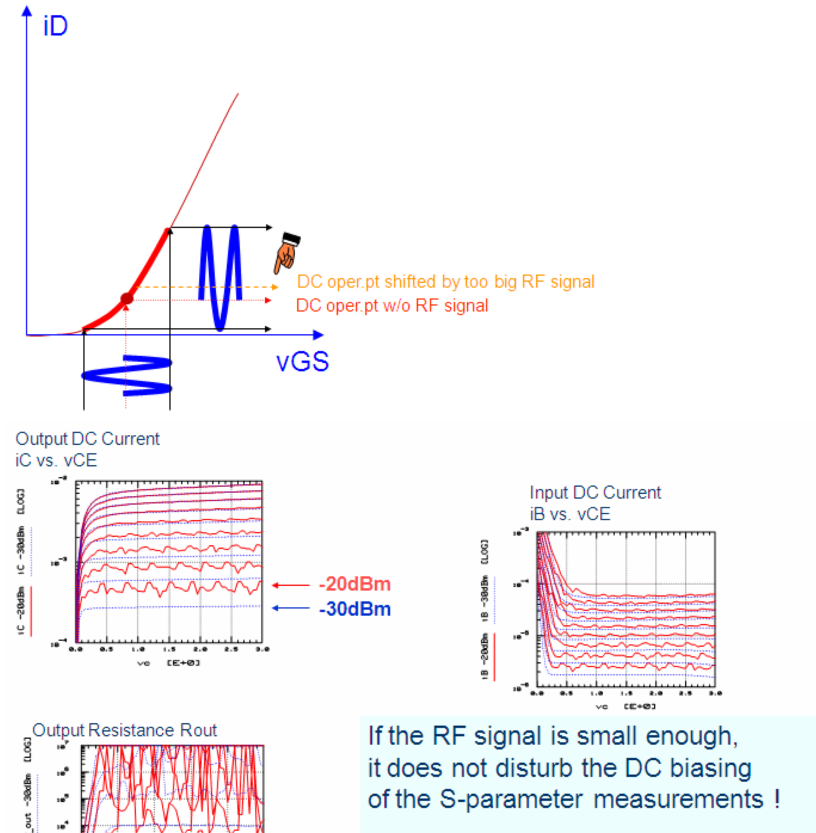
- -40dBm corresponds to a RMS voltage of ~2.2mV, what corresponds to ~3mV peak or ~6mV peak-peak !
- Small signal condition for a transistor is usually defined as VT/10=2.7mV

As mentioned, when harmonics occur, i.e. the test device is overdriven, the DC operating point is affected and thus, the DC bias current is affected.
We can use this effect to identify the max. applicable RF signal for NWA measurements with a smart and simple measurement method, depicted in the slide above.
When measuring a DC output characteristics and calculating Rout out of it, the resulting curve is very sensitive. Therefore, we can use this plot to identify possible effects of a too big an AC power applied to the transistor.
This means, we measure the DC output characteristics, through the NWA test set, or using external bias TEEs, and let the NWA operate in continuous mode, i.e. unsynchronized to the DC measurement. We start with the NWA switched-off, or set to min. RF signal power and max. port attenuations. We keep that output characteristics measurement (and the output resistance plot) as a reference.
Then, we increase the RF signal, and measure once again the output characteristic. When the device is overdriven, i.e. when the operating point is shifted (ib and ic for bipolar, id for MOS), we recognize that easily by a shift of the output characteristics and a shift in the output resistance plot. Now we know that harmonics are present. We then reduce the RF power a bit and know the max. allowed RF power for the NWA S-parameter measurements for that particular transistor device.

Transistor and Diode S-Parameter measurements are performed including a DC bias.

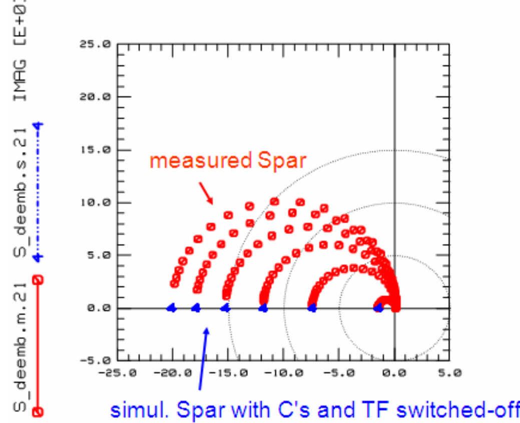
Due to the nonlinear transistor behavior, a too big AC amplitude results in a distorted

output signal. These distortions will shift the DC operating point.



From the plots above, we can determine 2 important things:
the max. applicable RF signal level for the S-parameter measurements for given DC operating points, or, inversely, the min applicable DC bias for a given RF signal. In other words: The starting points of the measured S-par (0Hz)are determined by the DC fitting.

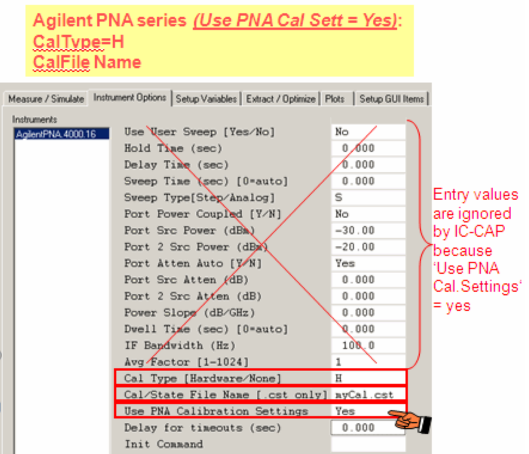
- With RF modeling (capacitors and transit time), only the trace from the starting points towards the end points can be fitted, but not the starting points. If this starting point check fails, verify:
  - too much RF signal
  - self-heating
- Voltage drop in S-par testset> accurate DC contact resistance> DC modeling was performed at different bias range than the S-par modeling.



When using an Agilent PNA/ENA*, specify the NWA power settingsdirectly in the PNA/ENA Cal.Wizzard, or applying Wincal / SussCal

> **Note**
> See the appendix when using the older Agilent/HP NWAs

In this case, thecalibration is not performed by IC-CAP, and the InstrumentOptionsin the IC-CAPSetup are as simpleas shown here.

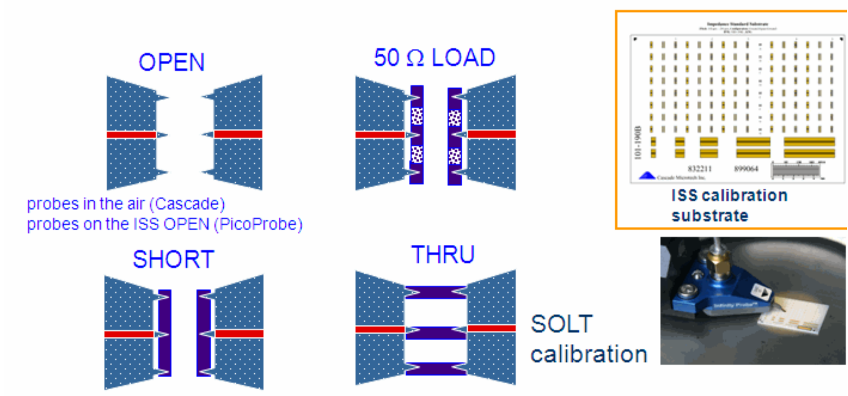


Note
All entry fields, except the red marked ones, are ignored by IC-CAP.

Provided that the Cal Kit Data (which is provided with every GSG probes calibration substrate) have been entered correctly into the network analyzer, and we have determined the maximum applicable RF signal level, the calibration of the Agilent PNA network analyzer is pretty simple. (For the older Agilent/HP NWAs, see the slide at the end of this slide set).
The steps are (for a PNA, and pretty similar for an ENA):

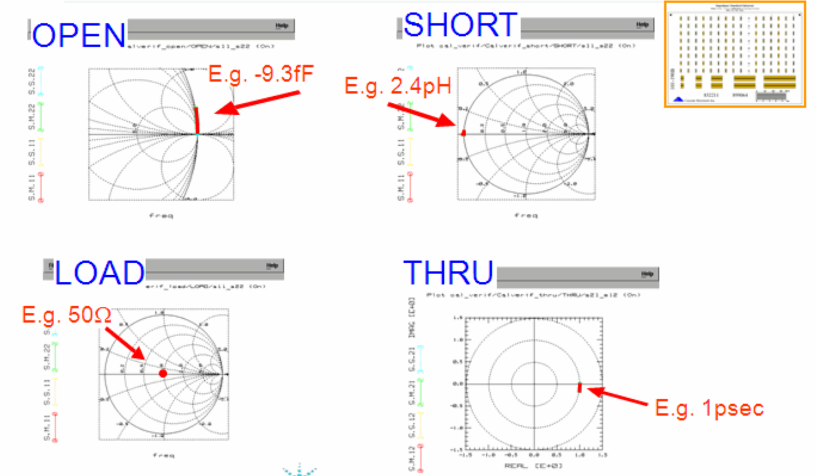
- after the PNA was calibrated (applying SussCal, WinCal, or ECal, PNA's built-in Cal), **manually** save your CalState on the PNA by selecting File/Save_As, select '.cst' file format, and save it to the default directory proposed by the PNA. Don't save it to another directory. IC-CAP would not find it there.
- in IC-CAP, specify the frequency range (IC-CAP Input 'freq') identical to the PNA frequency settings-> in the InstrumentOptions of your IC-CAP Setup, when 'Use PNA Calibration Settings=YES', then simply specify only - CalType = H, - Cal File Name. Note: Port1 and Port2 Source Power, Port Attenuation and IF Bandwidth are now accepted by IC-CAP as they were set by the user (directly on the PNA), or by SussCal/WinCal.
- Don't hit the IC-CAP 'Calibrate' button - it is no longer required in this case
- Simply hit 'Measure'.
  IC-CAP file: The easiest way to verify the calibration of your NWA (PNA) is to use the following file: in demo_features, go to 3_MEAS_ORGANIZE_n_VERIFY_DATA\0_MASTER_FILES\ 20_NWA_CAL_VERIFICATION, and load file CAL_VERIFY_PNA_2port_MASTERFILE_demodata.mdl.

During the calibration, measure the Calibration Standards on the ISS substrate

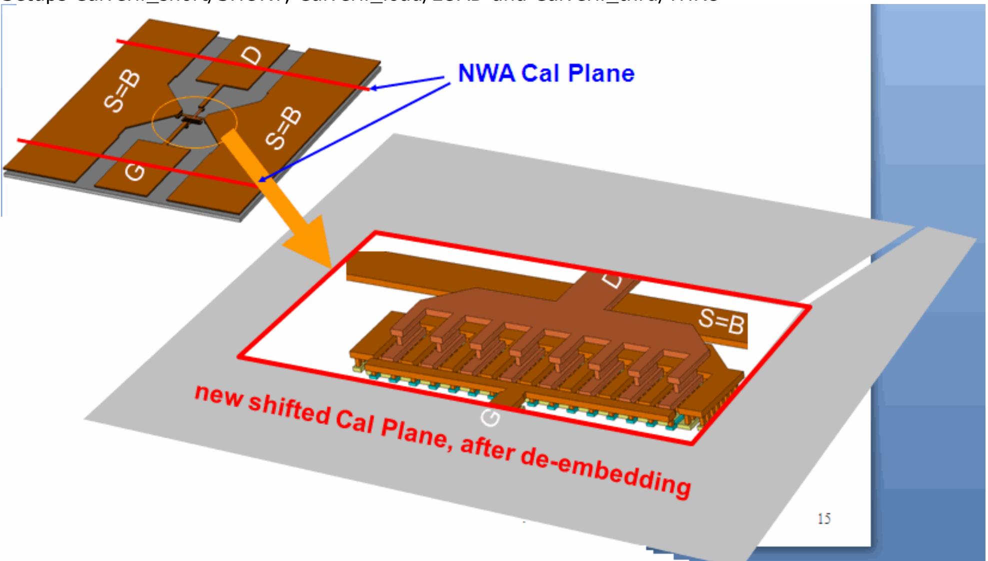


On the NWA's calibration menu, select 'Full 2-Port Calibration'.

Verify the NWA calibration by re-measuring and comparing the result to simulated standards

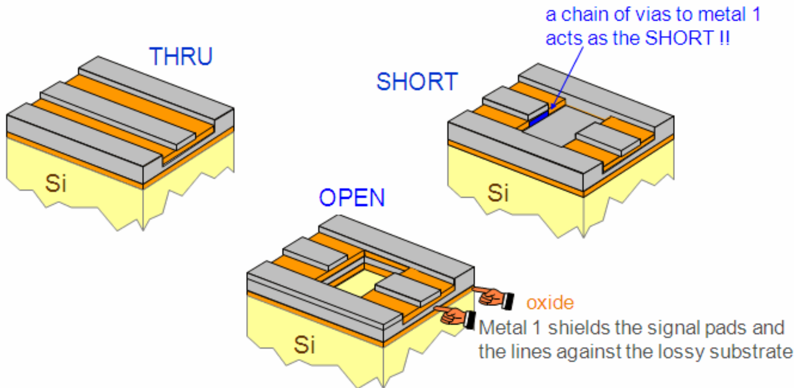


After the calibration is finished, it is important to measure some devices with known performance to verify the calibration quality. But, what is a known, 'golden device'? A set of devices with exactly known characteristics are the calibration standards on the ISS cal substrate. We remeasure all four of them, and compare the measurements individually against the simulation of the calkit data.
A really excellent fitting of measured and simulated data is an absolute MUST for accurate S-parameter measurements!
The goal is a fitting quality for all four measurements in the <2-5% range!
See the IC-CAP file: in demo_features, go to
3_MEAS_ORGANIZE_n_VERIFY_DATA\0_MASTER_FILES\
20_NWA_CAL_VERIFICATIONand load the file
CAL_VERIFY_genl_NWA_2port_MASTERFILE_demodata.mdl, Execute the
Macro'COPY_INSTR_OPTIONS' and then remeasure the standards applying the
Setups Calverif_short/SHORT, Calverif_load/LOAD and Calverif_thru/THRU



In the above slide, in the magnified view, is the inner transistor, which we want to characterize. All the rest has to be de-embedded, i.e. to be stripped-off. In other words, the NWA calibration plane, obtained by the NWA calibration so far, has to be shifted from the GSG probe contact location down to the beginning of the device.

Prepare the De-Embedding: Measure and Model The On-Wafer Dummies

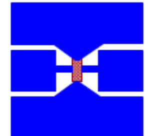


This layout suggestion for Silicon on-wafer Dummy structures uses metal1 to shield the contact pads from the lossy silicon substrate.Due to this metal1 shielding, the measured OPEN Dummy S-parameters represent a much more ideal OPEN Dummy schematic than the more conventional layout without the metal1 shielding. However, the OPEN Dummy capacitance is bigger than without the metal1 shielding. But it is representing much more an ideal capacitance.
For the SHORT Dummy, the short itself is much more ideal with the metal1 shielding compared to without that shielding. This is because of the chain of shorting vias down to the electrically well-behaving metal1 Ground plane. This chain is geometrically concentrated and precisely localized, compared to the conventional SHORT Dummy layout where the device 'hole' is simply filled up with metal1.

> **Note**
> For clarity, the layout details are simplified in the slide above. See publication T.E.Kolding, O.K.Jensen, T.Larsen, "Ground-Shielded Measuring Technique for Accurate On-Wafer Characterization of RF CMOS Devices", IEEE Int.Conf.on Microelectronic Test Structures ICMTS, March 2000, Monterey,CA.

IC-CAP file: go to demo_features, and then 3_MEAS_ORGANIZE_n_VERIFY_DATA\0_MASTER_FILES\31_*DEEMB_VERIFICATION\ and load the file CHECK_OPEN_SHORT_DEEMB_with_a_THRU_demodata*_MASTERFILE.mdl

Layout Comparison:Conventional On-Wafer Dummy Structures



OPEN

the OPEN suffers from the lossy silicon substrate effects, because metal1 shielding is not applied

SHORT

the SHORT suffers from considering the metal short as ideal, while the strip lines at both sides are considered as real, i.e. inductive lossy

This represents the more 'classical' Dummy structure layouts, without using metal1 as a shield against the lossy silicon substrate.

Modeling The On-Wafer OPEN Dummy

OPEN Dummy on Silicon, with metal-1 shielding

OPEN

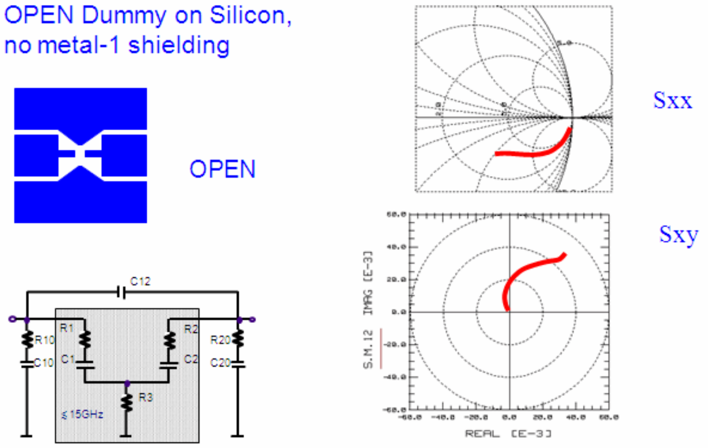
Si

C12 R12 C10 C20 R10 R20

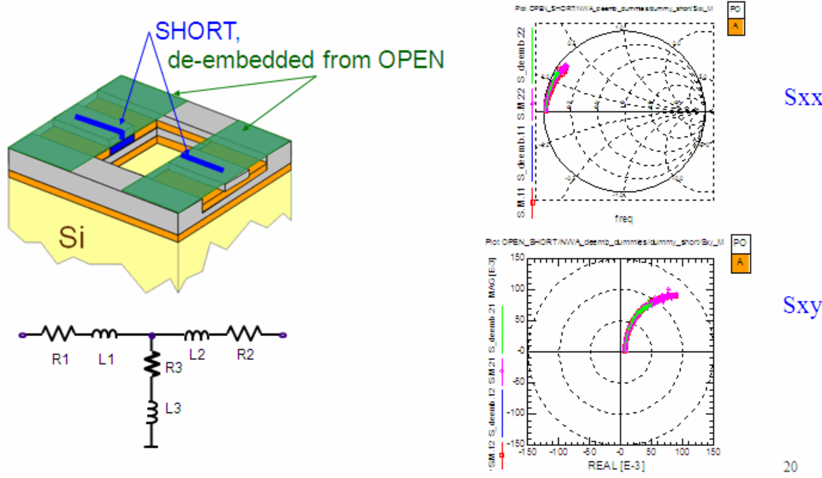
Sxx

Sxy

In this first step, the OPEN Dummy is modelled in order to verify that there are no inductive, serial components present. Such inductors would be a contradiction to the commonly applied OPEN de-embedding, i.e. the OPEN Y-matrix subtraction.
IC-CAP file: Go to demo_features, and then further to
3_MEAS_ORGANIZE_n_VERIFY_DATA\0_MASTER_FILES\31_*DEEMB_VERIFICATION\ and load file CHECK_OPEN_SHORT_DEEMB_with_a_THRU_demodata*_MASTERFILE.mdl.
See the Setup 'OPEN_dummy'.

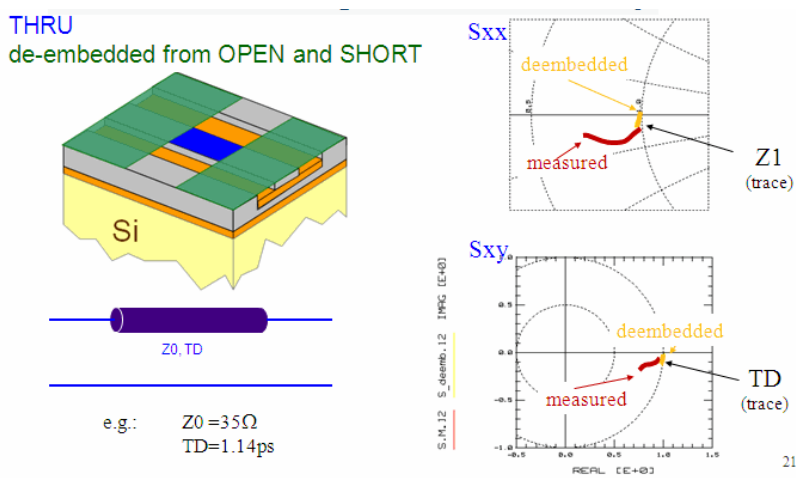


Model The On-Wafer SHORT Dummy



Then, the SHORT Dummy is de-embedded from the OPEN and again modelled. The goal is to verify that there are no parallel components included in the Z matrix of the SHORT. For example, a parallel capacitive component, required to model the SHORT, is a contradiction to the usually applied Z-matrix subtraction.
IC-CAP file: go to demo_features and to directory
3_MEAS_ORGANIZE_n_VERIFY_DATA\0_MASTER_FILES 31__DEEMB_VERIFICATION\
and load the file
CHECK_OPEN_SHORT_DEEMB_with_a_THRU_demodata__MASTERFILE.mdl
Go to Setup 'SHORT_dummy'.

Verify The De-Embedding With Modeling The THRU Dummy



After the OPEN and SHORT Dummy measurements have been verified by modeling, we want to check the de-embedding procedure with a known 'golden device'. In practice, however, like with the NWA cal.verification, the problem is the availability of an 'golden device'. Provided we have a THRU Dummy, we can use this component for a verification of the de-embedding procedure. We do not know the model parameters of the THRU. But we know that the de-embedded THRU should look like a simple delay line, with a certain characteristic impedance Z0, and a delay time TD. Additionally, the delay line may be lossy. In most cases, due to the short geometry of the THRU, the loss is negligible. Anyhow, the model parameters must represent physically meaningful values, and the de-embedded trace must be simple without resonances etc. Just a simple delay line!

Some modeling hints

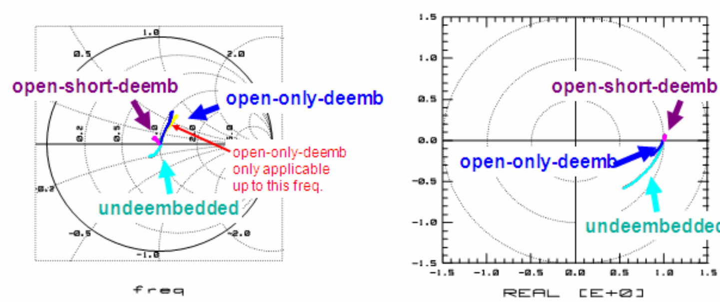
TD: models the phase in Sxy Z0: models the Sxx trace:
for Z0<50W, the curve starts at 50W, turns downwards and then to the left into a circle around the line's Z0. The turning is clock-wise with increasing frequency!
for Z0=50W, the curve is a single point at at 50W in the Smith chart.
For Z0>50W, the curve starts at 50W, turns upwards and then to the right into a circle around the line's Z0. The turning is clock-wise with increasing frequency!

Only if there were no layout bugs or measurement problems with the Dummy structures, we will end up with a physical representation of the small, de-embedded THRU strip line! Therefore, when we obtain a physical strip line characteristic and model, we can assume that the de-embedding, applied later to transistors, diodes etc., will be correct as well!

Note
since the THRU strip line is passive, and not depending on bias or frequency or temperature, it can be used much better for de-embedding verification than checking the de-embedding with a transistor, a diode etc.

IC-CAP file: go to demo features and ´further to 3_MEAS_ORGANIZE_n_VERIFY_DATA\0_MASTER_FILES\31_*DEEMB_VERIFICATION\ and load the file CHECK_OPEN_SHORT_DEEMB_with_a_THRU_demodata*_MASTERFILE.mdl. See Setup 'verify_with_THRU_dummy'.

An Example of bad THRU line de-embedding results



- The OPEN-only-deembedded Sxx is only OK up to 2/3 of the freq. range.
- The OPEN-SHORT-deembedded Sxx turns backwards (unphysical impedance Z0).
- The OPEN-only-deembedded Sxy looks OK
- The OPEN-SHORT-deembedded Sxy turns backwards (neg. delay time)

The slide above represents a de-embedding of Dummy structures with obviously layout problems.
Although the inspection of the measurement results sequence undeembedded >

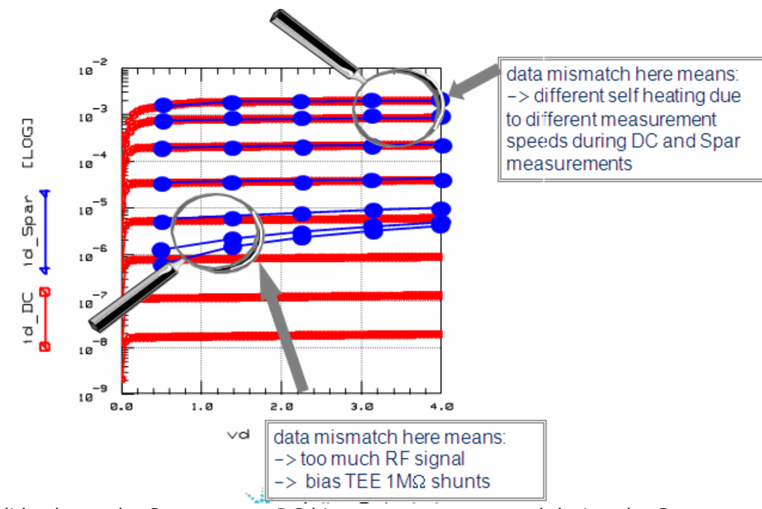
open_deemb > open_short_deemb exhibits the right trend (with each de-embedding step, the resulting THRU becomes more and more the expected simple, short THRU), the open_short_deemb results does not match the expected the modeling target and represents - a strip line with negative delay time (see Sxy: a short delay line with neg. delay time and even with a trend for MAG(Sxy)>1 with increasing frequency) - and an unphysical Z0 (since in Sxx, it turns counter-clock-wise).

Finally, measure the device, but keep an eye on the self-heating

if you have to consider self-heating of your device (typ. >50mW DC power) and if you measure both, first the DC curves like

- transfer curve (id vs vg)
- output characteristics (id vs vd)
  and then the- DC biased S-parameter curves, make sure that you measure the DC curves with the same slow speed as during the S-parameter measurements.
  Verify the identical self-heating by comparing the DC-only curves with the DC bias curves obtained during the S-parameter measurements.

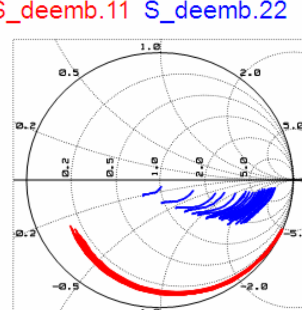
DC Measurements vs. Spar DC Bias



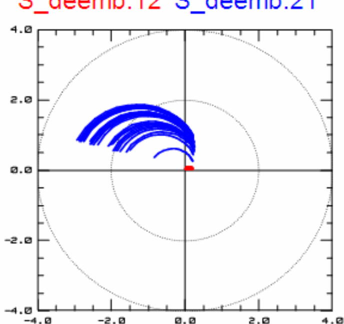
The slide shows the S-parameter DC bias currents, measured during the S-parameter measurements, extracted and overlaid to the conventional DC output characteristics measurement.
IC-CAP file: go to demo_features and then further to
3_MEAS_ORGANIZE_n_VERIFY_DATA\0_MASTER_FILES\1__CHECK_DATA_CONSISTENCY
and load the file FET_MEAS_CHECKTOOLS.mdl or NPN_MEAS_CHECKTOOLS.mdl

**With all prerequisites fulfilled,
we end up with reliable S-parameters**

S_deemb.11 S_deemb.22



S_deemb.12 S_deemb.21



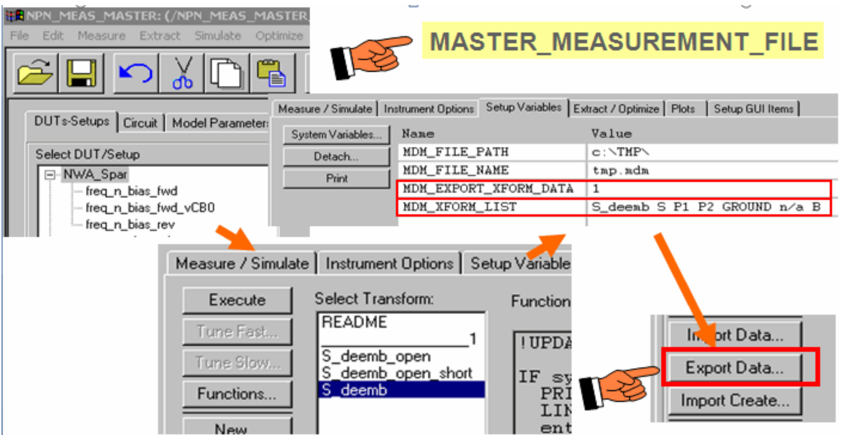
So, How To Deal With S-Parameters?

The 10 Rules At A Glance

1. Measure the DC Bias Voltage Drop in the VNA Bias TEEs
2. Measure The Probe Contact Resistance
3. Tell the NWA which standards will be applied during calibration
4. Identifying the correct AC Power Level for measuring the transistor or diode
5. Perform The NWA Calibration
6. Verify the NWA calibration by remeasuring and comparing the result to simulated standards

7. Prepare the De-Embedding: Measure The On-Wafer Dummies
8. Model the on-wafer OPEN and SHORT dummies
9. Verify the de-embedding with modeling the THRU Dummy
10. Finally, measure the device, but keep an eye on the self-heating

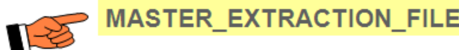
Last not least:Organizing the De-Embedding of S-Parameters in IC-CAP

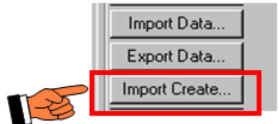


Hint: In the demo_features directory, go to directory 3_MEAS_ORGANIZE_n_VERIFY_DATA\0_MASTER_FILESand load the file NPN_MEAS_MASTERFILE_demodata_PELdep.mdl Check the organization of Setup 'NWA_Spar/freq_n_bias_fwd'.
Particularly, check for:

- Output 'S'
- Transform 'S_deemb'
- Setup Variable 'MDM_EXPORT_XFORM_DATA'=1
- Setup Variable 'MDM_XFORM_LIST' = S_deemb S P1 P2 GROUND n/a B
  This Setup Variable 'MDM_XFORM_LIST' will assure that the Transform 'S_deemb' is added to the .mdm file during 'Export', and that it will be loaded back later as an 'Output' called 'S_deemb' (the default name for de-embedded S-parameter) with Type 'S' Port1 'P1' Port2 'P2' AC Ground 'GROUND' Unit 'n/a' Type 'B'.

Organizing the De-Embedding of S-Parameters in IC-CAP





When importing the .mdm file into the Extraction Master Model File, simply delete the unnecessary Output 'S', and keep only the Output 'S_deemb'.

Wrap-Up Network Analyzer Calibration

with calibration

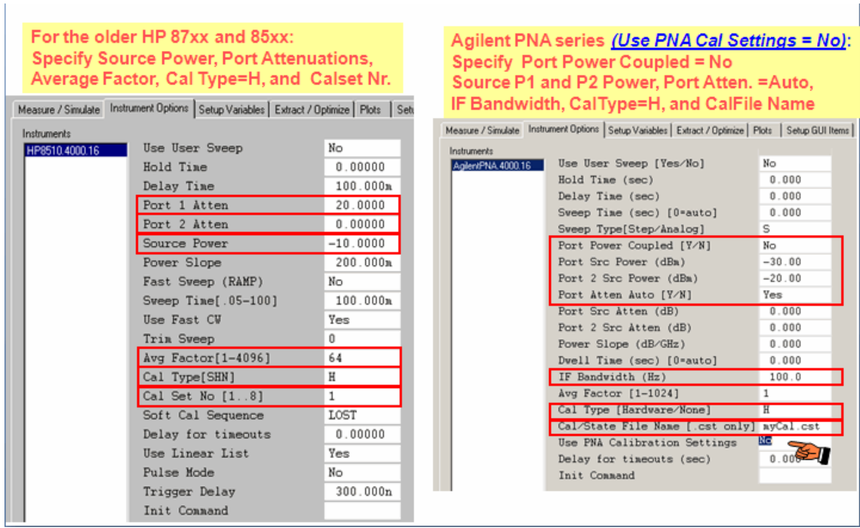
bad calibration

bad calibration and too big a signal

Appendix

Performing the NWA calibration from IC-CAPfor the older HP/Agilent NWAs like HP8510, HP87xx etc.

Specifying the NWA power settings in IC-CAP



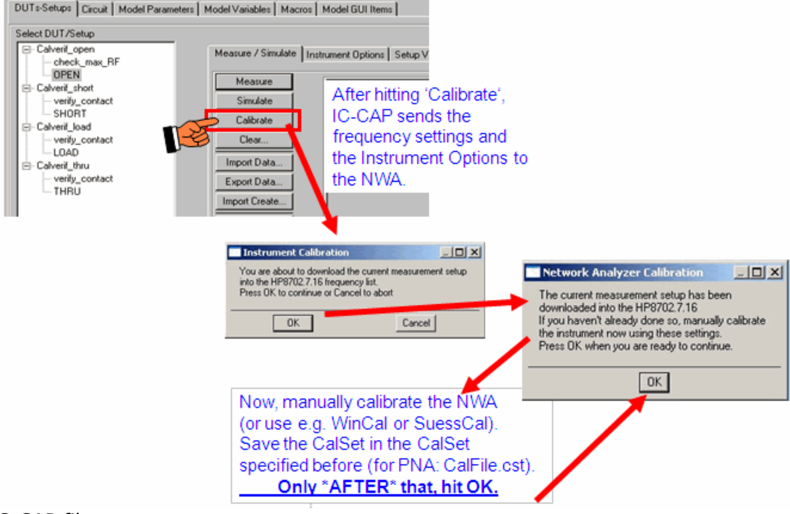
Provided that the Cal Kit Data (which is provided with every GSG probes calibration substrate) have been entered correctly into the network analyzer, and we have determined the maximum applicable RF signal level, the calibration of the NWA is pretty simple.
The steps are:

- First of all, specify the frequency range (IC-CAP Input 'freq')
- In InstrumentOptions (for the older HP/Agilent NWAs like 8510, 87xx), enter the 'Source Power' and the 'port x Atten'
    - Set the 'avg Factor' (typ. 64) - 'Cal Type' = H
    - Specify the 'Cal Set No'.
- In InstrumentOptions (for the Agilent PNA series),
    - When 'Use PNA Calibration Settings'=NO, then specify
- 'Port Power Coupled' = No
- 'Port Src Power' and 'Port 2 Source Power'
  to the value evaluated before
- set 'Port Atten Auto' = Yes
- 'IF Bandwidth' (typ. 100Hz or less)
- 'CalType' = H
- 'Cal/state File Name' as specified in the PNA before.

As mentioned before, make sure to select an appropriate RF power level, in order to not overdrive the DUT.

IC-CAP file: The easiest way to verify the calibration of your NWA (PNA) is to use the following file:in demo_features, go to 3_MEAS_ORGANIZE_n_VERIFY_DATA\0_MASTER_FILES\ 20_NWA_CAL_VERIFICATION, and load the file CAL_VERIFY_genl_NWA_2port_MASTERFILE_demodata.mdl.

For the PNA, use the ModelFIleCAL_VERIFY_PNA_2port_MASTERFILE_demodata.mdl
Then, start The NWA Calibration From IC-CAP

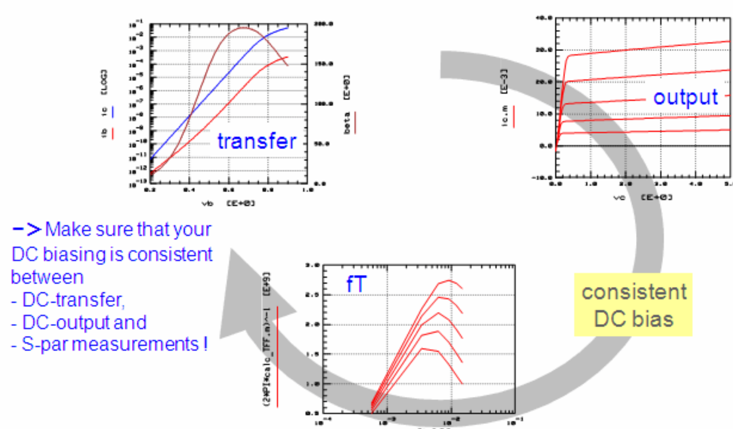


IC-CAP file: go to

demo_features\3_MEAS_ORGANIZE_n_VERIFY_DATA\0_MASTER_FILES\20_NWA_CAL_VERIFICATIONand stay with fileCAL_VERIFY_genl_NWA_2port_MASTERFILE_demodata.mdl Go to Setup Calverif_open/OPEN, and hit the 'Calibrate' button there.

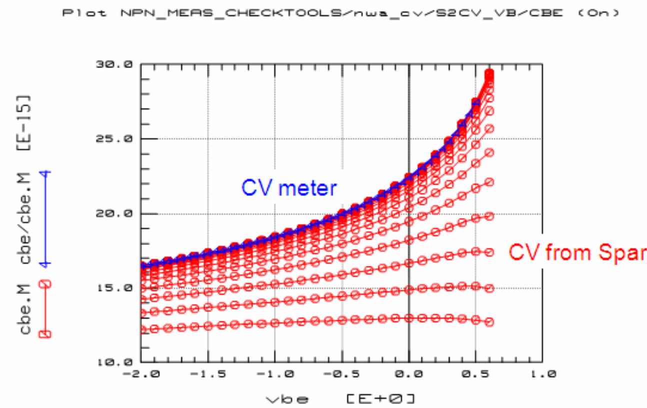
> **Note**
> usually, you need a separate NWA calibration for every different frequency sweep. When using Agilent PNAs, and when using 'Use PNA Calibration Settings'=YES, this step (hitting the 'Calibrate' button) is not required.

A typical modeling mistake: fT modeling at different (higher) iC than DC modeling

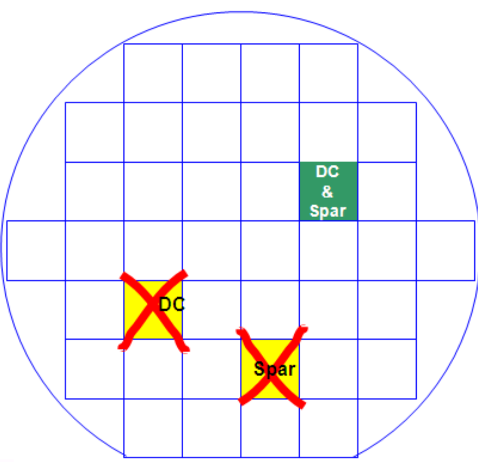


Further Data Consistency Check Checks: CV <> Spar and DC <> Spar Consistency Checks
After we have discussed in details the data consistency checking within DC, CV and S-parameter measurements, we want to close the loop and compare that data consistency also against the measurement domains.

Comparing CV curves from S-par measurements with CV-measurements



When comparing CV curves converted from S-parameters, keep in mind that due to the S_to_Y conversion and the interpretation of these Y-parameters with respect to a PI schematic, the CV curves represent the capacitance in one of the PI branches. With a CV meter, depending on the measurement principle, you may obtain the total capacitance between two nodes (2-pin CV measurement method) or the individual capacitance (4-pin method, guarded wafer chuck).
Last not least: don't merge diff. measurements of diff dies when extracting model parameters.



It is obvious that when developing a model of a component, all types of measurements have to be performed on the same device !

S-Parameter Utilities

Calculation the radiation losses from S-parameters

The radiation losses for passive components can be calculated as:

$$radiation[\%] = 100 * \left[(1 - MAG(S11))^2 - (MAG(S21))^2 \right]$$

Note
If considerable radiation occurs, e.g. with spiral inductors, conventional lumped schematics may no longer be valid.

Understanding S-Parameter Plots

The Smith Chart for Sxx Parameters

What makes Sxx-parameters especially interesting for modeling, is that S11 and S22 can be interpreted as complex input or output resistances of the twoport (including the termination at the opposite side of the twoport with Z0).
That's why they are usually plotted in a Smith chart. This chapter is intended to explain the basics of such Smith charts.

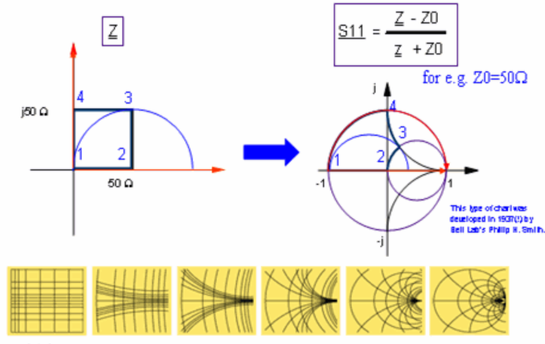
The Smith chart is a transformation of the complex impedance plane R into the complex reflection coefficient Γ (rho) following:

$$\underline{\Gamma} = \frac{\underline{R} - Z0}{\underline{R} + Z0}$$

with the system's reference impedance Z0 = 50 Ω .

This means that the right half of the complex impedance plane R is transformed into a circle in the Γ -domain. The circle radius is '1' (see the following figure).

Relationship between Sxx and the complex impedance of a twoport

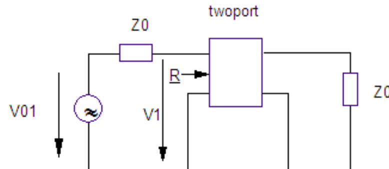


On the other hand, using an network analyzer with a system impedance of Z0, the parameter S11 is equal to:

$$S_{11} = 2 \cdot \frac{v1}{v01} - 1$$

where v1 is the complex voltage at port 1 and v01 the stimulating AC source voltage (typically normalized to '1'). See the following figure and the chapter, *Calculating S-parameters from Voltages* (iccapmhb) for details.

About the definition of S11, (S22 is just analogous)



Under the assumption that R is the complex input resistance at port 1 and Z0 is the system impedance, we get using eq.(2) and the resistive divider formula:

$$\underline{S_{11}} = 2 \cdot \frac{\underline{R}}{\underline{R} + Z0} - 1 = \frac{\underline{R} - Z0}{\underline{R} + Z0}$$

And this is the reflection coefficient Γ from (1) !!

After all, if the reflection coefficient Γ resp. S11 or S22 is known, we get for the complex resistor R:

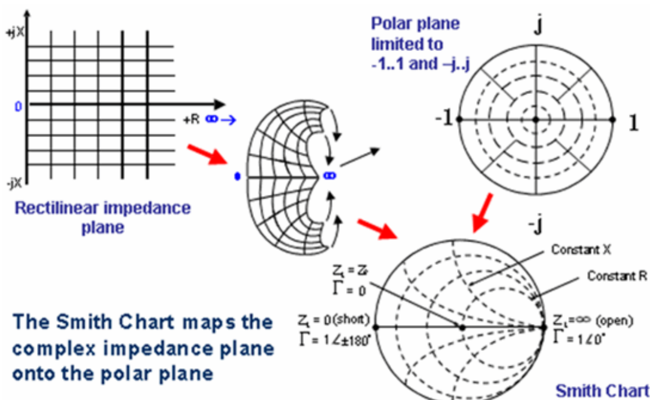
$$\underline{R} = Z0 \cdot \frac{1 + \underline{\Gamma}}{1 - \underline{\Gamma}} = Z0 \cdot \frac{1 + \underline{S_{11}}}{1 - \underline{S_{11}}}$$

, with usually Z0 = 50Ω
This explains how we can get the complex input/output resistance of a twoport directly from S11 or S22, if we plot these S-parameters in a Smith chart.

Let's go back to Figure, "Relationship between Sxx and the complex impedance of a twoport" and consolidate this context a little further:
It shows a square with the corners (0/0)Ω, (50/0)Ω, (50/j50)Ω and (0/j50)Ω in the complex impedance plane and its equivalent in the Smith chart with Z0=50Ω. Please watch the angel-preserving property of this transform (rectangles stay rectangles close to their origins). Also watch how the positive and negative imaginary axis of the R plane is transformed into the Smith chart domain (Ω), and where (50/j50) Ω is located in the Smith chart. Also verify that the center of the Smith chart represents Z0, i.e. for Z0=50Ω,

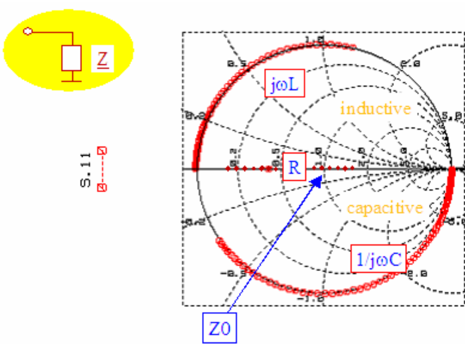
the center of the Smith chart is (50/j0)Ω.
The following explains once again the transformation of the complex ohmic plane to the Smith chart.

How the complex resistance plane is transformed into the Smith chart



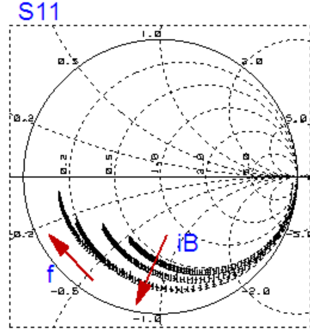
This allows us to make the following statements:

- Sxx on the real axis represent ohmic resistors
- Sxx above the real axis represent inductors
- Sxx below the real axis represent capacitors
- Sxx curves in the Smith chart turn clock-wise with increasing frequency.



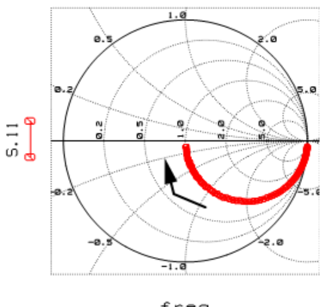
As an example for interpreting Smith charts, The following figure shows the S11 plot of a bipolar transistor. In this case, the locus curve stars with S11≈1 = ∞*Z0 at low frequencies (RBB' + Rdiode + beta*RE). For higher frequencies, the curves then tend towards |S11|{-1...0}
for high frequencies (the CBE shorts Rdiode, and beta = 1. Therefore, the end point of S11 is RBB' + RE). Since RBB' is bias dependent, and decreasing with increasing iB, the end points of the curves represent this bias-dependency. For incrementing frequency, the S11 locus curve turns clockwise-

S11 of a transistor with increasing Base current iB



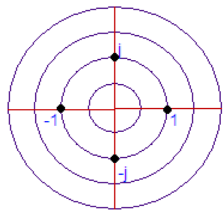
The following figure shows the S11 curve of a capacitor located between the NWA ports. The capacitor represents an OPEN for DC, thus S11 ~ 1 = ∞*Z0. For highest frequencies, it behaves like a SHORT, and we see the 50Ω of the opposite port2 . The transition between the DC point and infinite frequency follows a circle, and the increasing frequency turns the curve again clockwise.

S11 of a capacitor between port 1 and port 2



THE POLAR PLOTS FOR THE S12 AND S21 PARAMETERS

The polar plot for S12 and S21

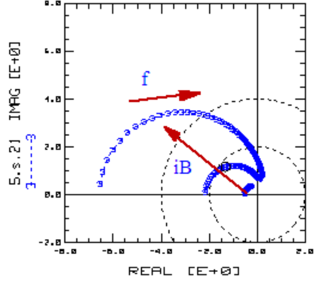


The S21 parameter represents the power transmission from port 1 to port 2, if the twoport is inserted into a matching network with characteristic impedance Z0 of e.g. 50Ω. This means, if no voltage is transmitted, then S21=0 (in the center of the polar plot). If voltage is transmitted, we are on the positive X axis. The curve will be below S21=1 for damping between the port 1 and port 2, and above S21=1 for amplification. If the phase is inverted, we are basically in the left half-plane of the polar plot (REALS21 <0). And, for voltage amplification, but also phase shift of e.g.-180 degrees (a transistor), below S21=-1.

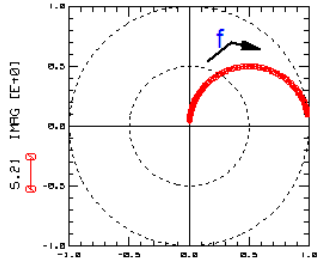
Like with the Smith chart, all S21 and S12 curves turn clock-wise with increasing frequency.

As an example, Figure, "S21 of a transistor with varying Base current iB" shows the S21 plot of a bipolar transistor, and Figure, "S21 of a capacitor between port 1 and port 2" of a capacitor between port 1 and port 2. While the transistor starts with S21 <-1 at low frequencies (voltage amplification in a 50Ω system), its curves tend towards S21=0 for high frequencies (no voltage transmission, the transistor capacitances short all voltage transmission). For the capacitor it's just the opposite: no power transmission for lowest frequencies, but an ideal short (S21=1) for highest frequencies. For more details, see the chapter on understanding S-parameters plots.

: S21 of a transistor with varying Base current iB



S21 of a capacitor between port 1 and port 2


f
S.21 IMAG [E+0]
REAL [E+0]
1.0
0.5
0.0
-0.5
-1.0
-1.0
-0.5
0.0
0.5
1.0

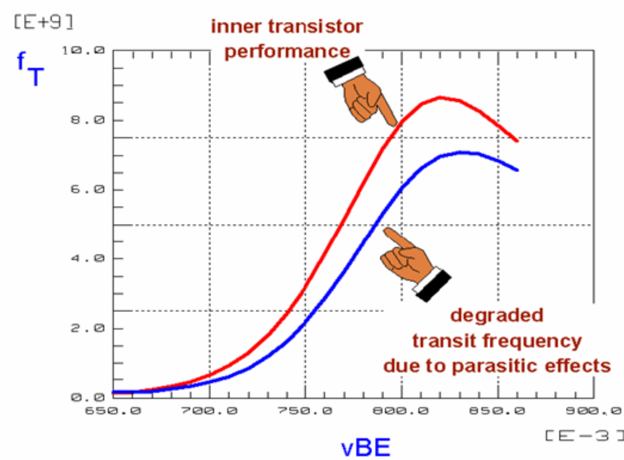
De-embedding Techniques

Contents

- *Basics of De-embedding Techniques* (iccapmhb)
- *De-embedding Techniques Tutorial* (iccapmhb)
- *ft Modeling of a Transistor Affected by Parasitics* (iccapmhb)
- *Interpreting OPEN-Dummy Measurements* (iccapmhb)
- *Twoport Matrices of Basic Schematics* (iccapmhb)
- *Twoport Matrix Conversions* (iccapmhb)
- *Twoport Matrix Definitions* (iccapmhb)
- *Verifying the De-embedding Procedure* (iccapmhb)

Basics of De-embedding Techniques

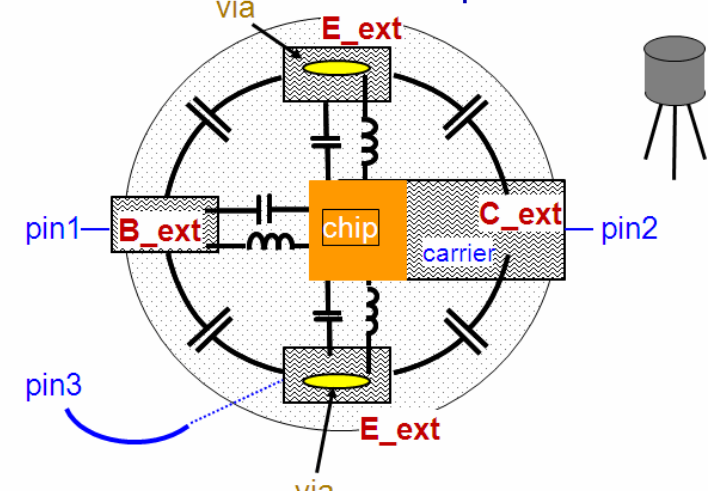
Why De-embedding ?



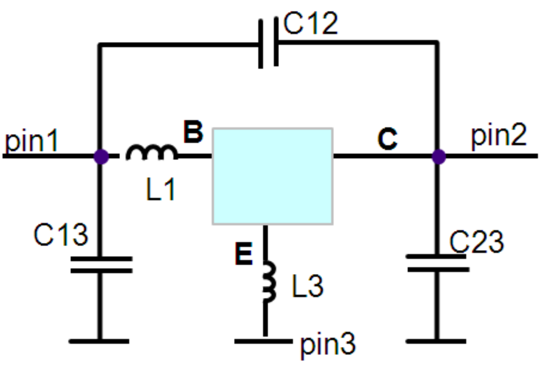
Example:
Chip Carrier of a Packaged Transistor and its Parasitic Components

Chip carrier

The bipolar transistor is mounted with its Collector to the carrier, Base is bonded, Double-Emitter contact (bonded twice) goes to the backside of the carrier.



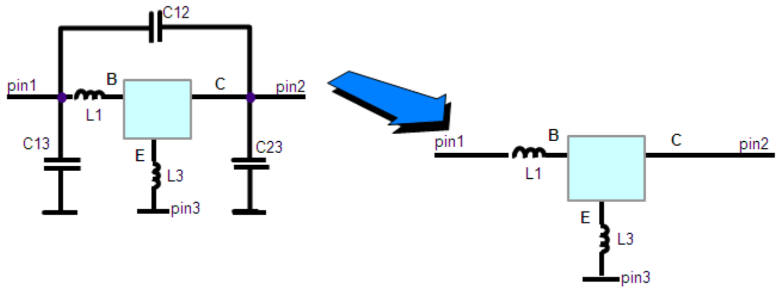
This corresponds to:



$$S_{total} = \begin{pmatrix} S_{11total} & S_{12total} \\ S_{21total} & S_{22total} \end{pmatrix}$$

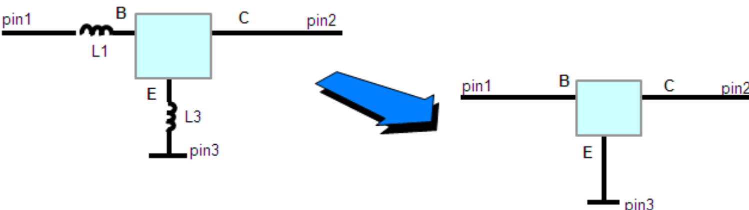
Converting the S-Parameters to Y, we can strip-off

the capacitors



$$\text{Ydeemb} = \begin{pmatrix} Y_{11} - j\Omega(C_{13} + C_{12}) & Y_{12} + j\Omega C_{12} \\ Y_{21} + j\Omega C_{12} & Y_{22} - j\Omega(C_{23} + C_{12}) \end{pmatrix}$$

then, converting to Z parameters

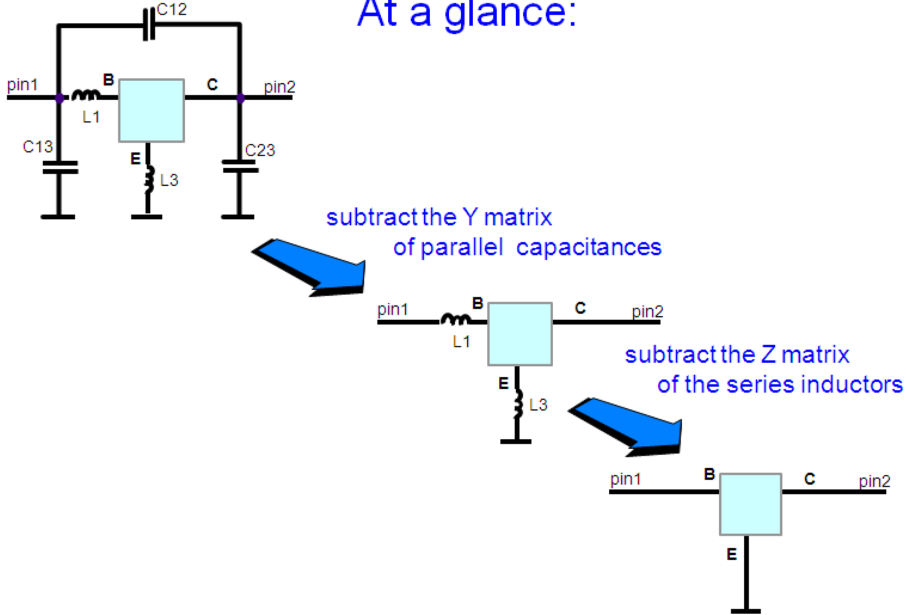


and de-embedding the inductors

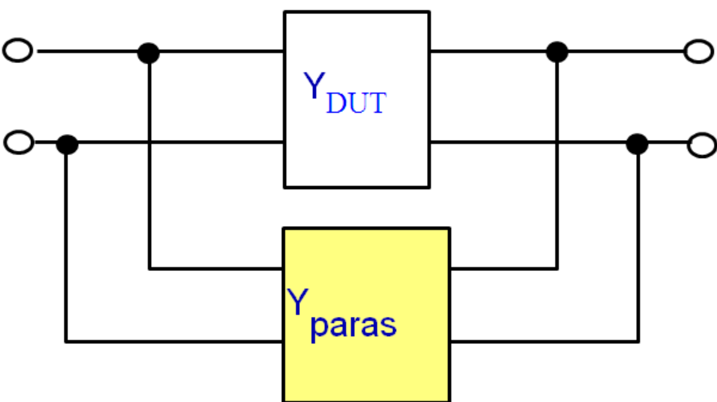
$$\text{Zdeemb} = \begin{pmatrix} Z_{11} - j\Omega(L_1 + L_3) & Z_{12} - j\Omega L_3 \\ Z_{21} - j\Omega L_3 & Z_{22} - j\Omega(L_3)) \end{pmatrix}$$

gives the inner transistor

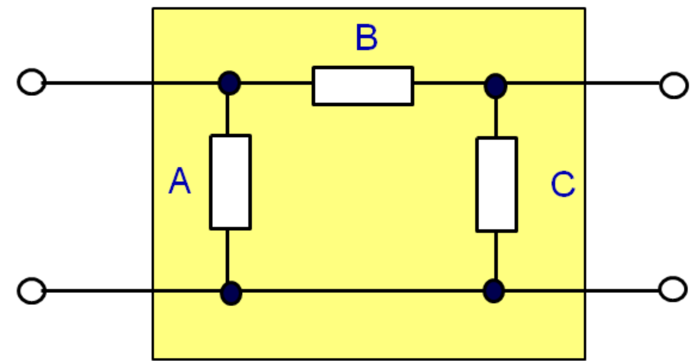
At a glance:



A note on adding of Y matrices



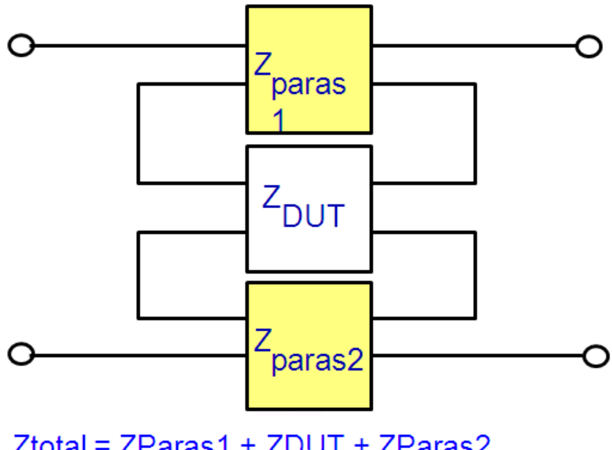
A typical representation of Matrix Yparas



$$Y_{paras} = Y_{PI} = \begin{pmatrix} A+B & -B \\ -B & C+B \end{pmatrix}$$

watch the signs !

A note on adding of Z Matrices



Ztotal = ZParas1 + ZDUT + ZParas2

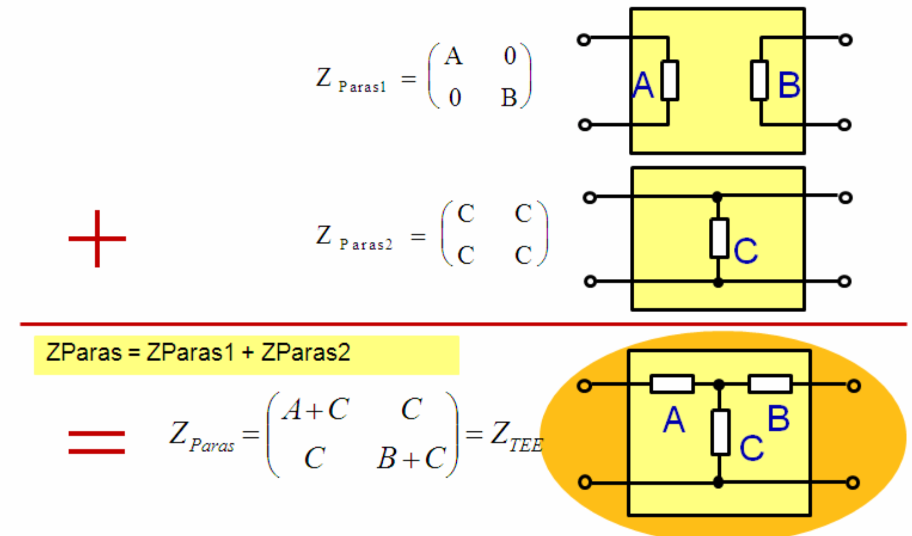
solved for ZDUT gives

ZDUT = Ztotal - (Zparas1 + Zparas2)

or

ZDUT = Ztotal - Zparas

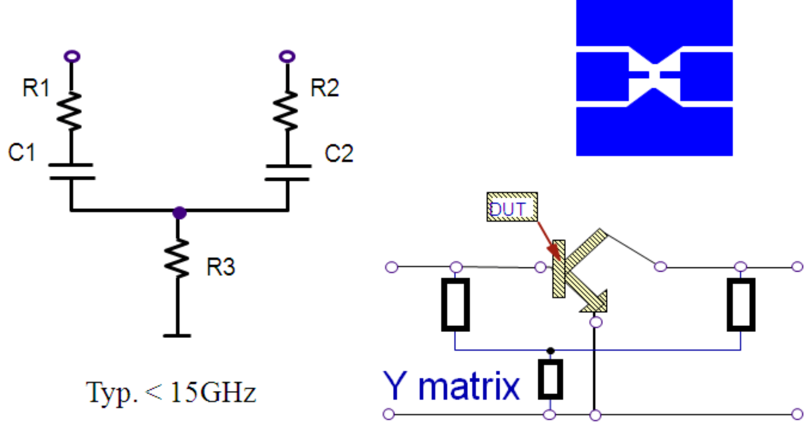
Applied to modeling, matrix Zparas represents a TEE structure:



This means:

- Either model the parasitics by a lumped components circuit and de-embed the individual components
  or, applied to on-wafer measurements:
- de-embed the whole matrix, if it correctly describes all parasitic effects that is, the Y matrix includes only parallel parasitic components, and the Z matrix only series parasitics

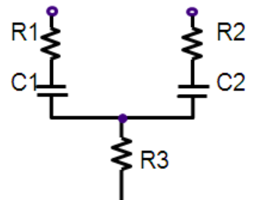
For on-wafer OPEN dummy structures, Yparas looks like this:



The Lumped Components Values are:

Convert Y -> Z and calculate
Zxy = (Z12 + Z21) / 2

A = Z11 - ZxyB = Z22 - ZxyC = Zxy

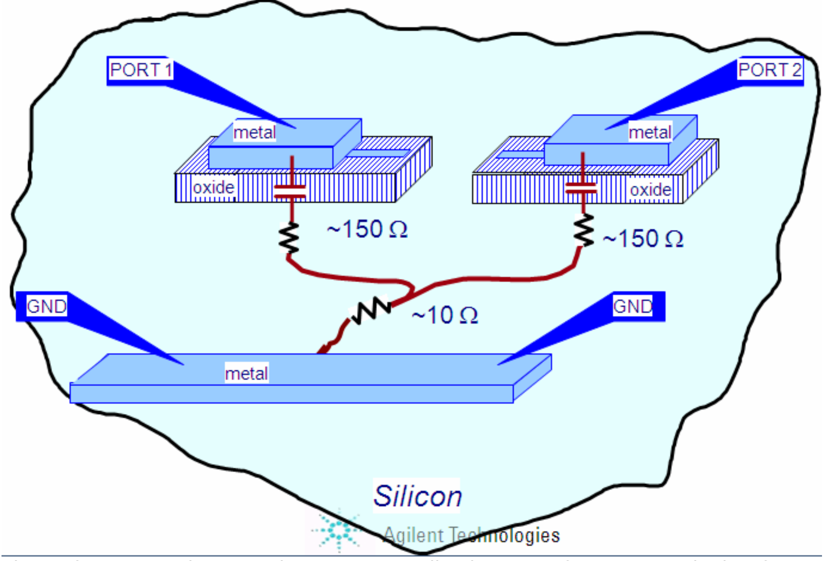


then $R1 = REAL(A)$ $C1 = \frac{-1}{2 * PI * freq * IMAG(A)}$

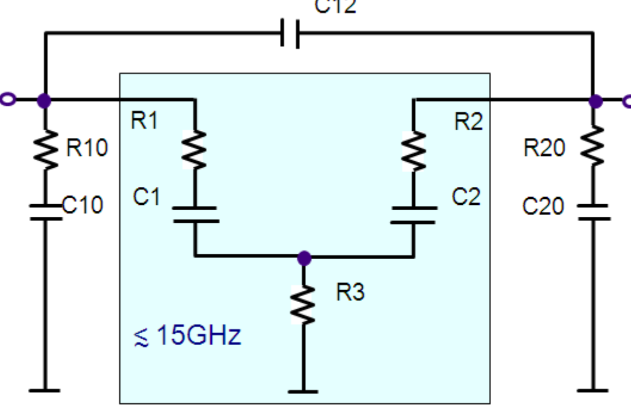
$R2 = REAL(B)$ $C2 = \frac{-1}{2 * PI * freq * IMAG(B)}$

$R3 = REAL(C)$

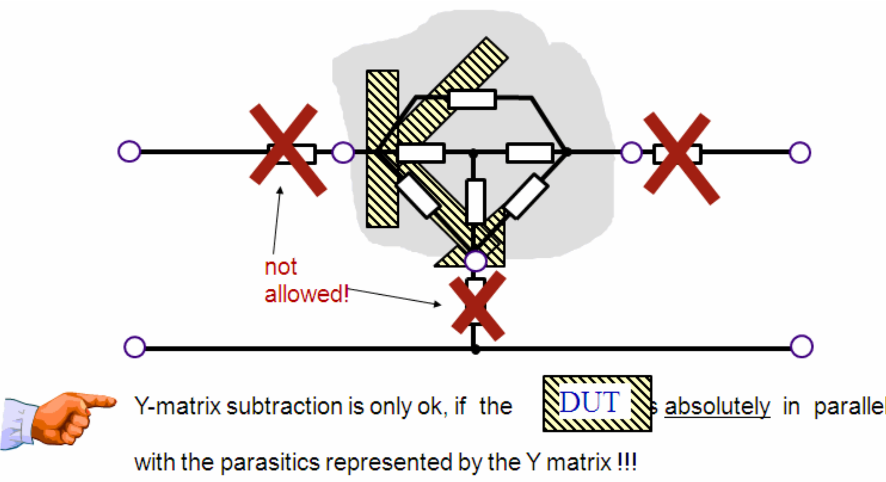
Physical Interpretation of the OPEN dummy



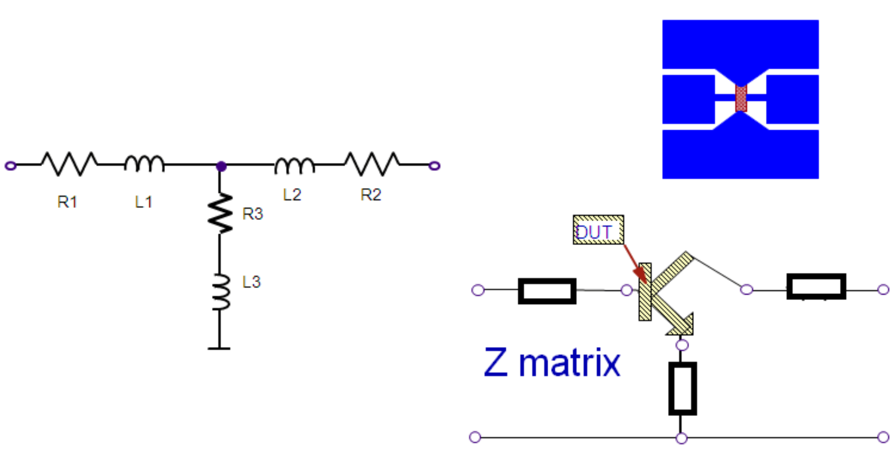
The so far proposed OPEN schematic is usually ok up to a few GHz. For higher frequencies, the following schematic is most often appropriate:



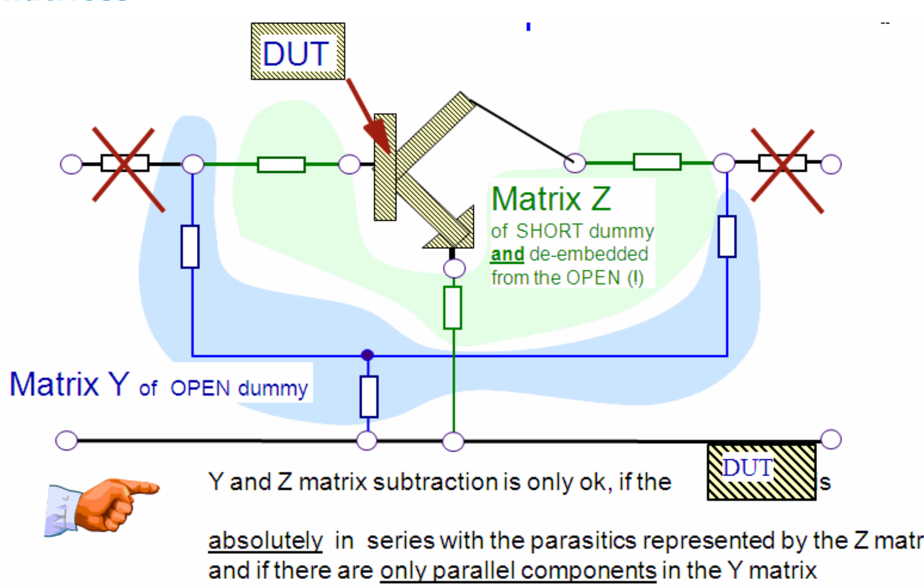
Reminder on the de-embedding of the complete Y matrix



For on-wafer SHORT dummy structures, Zparas looks like this:



about the de-embedding of complete Y and Z matrices



De-embedding of complete Y and Z matrices

Required measurements

'TOTAL' device
'OPEN' dummy device
'SHORT' dummy device.

De-embedding procedure

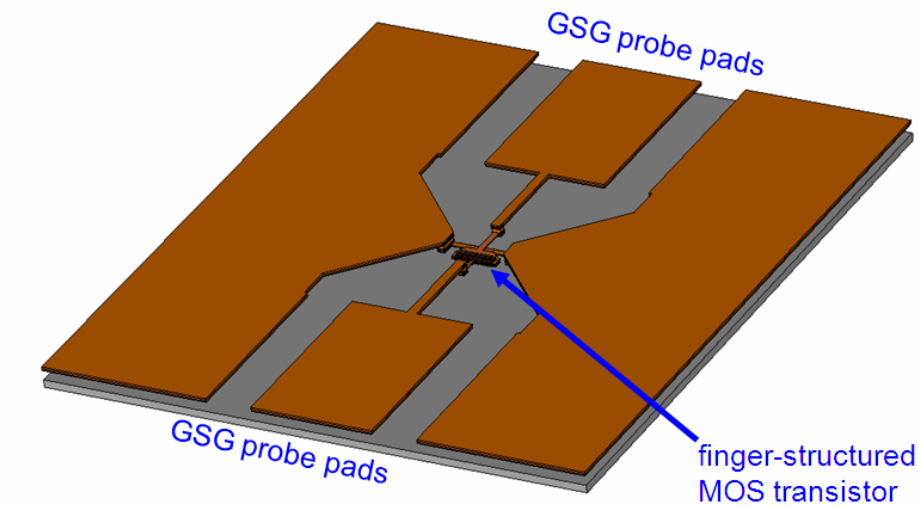
convert S_total -> Y_total,
and calculate Y1 = (Y_total - Y_open)
to de-embed the DUT from the parallel parasitics

convert Y1 -> Z1,
and calculate Z2 = (Z1 - Z_short) to de-embed further the serial parasitics,
finally re-convert Z2 -> S_deemb
to obtain the 'inner' DUT.

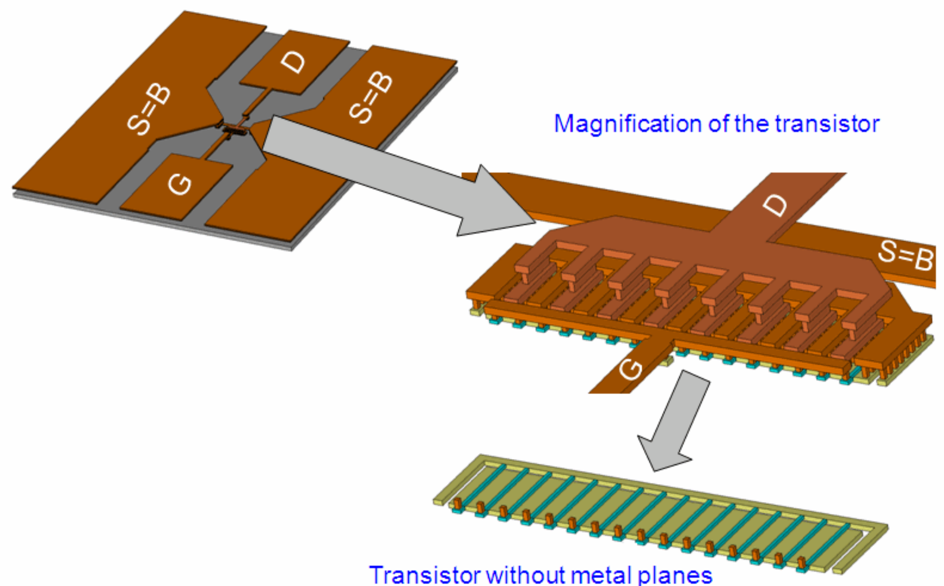
Note:
matrix Z_short has to be de-embedded first from the outer parallel parasitics.
This is done by:
convert
S_short_incl_open -> Y_short_incl_open,

calculate
Y_short = (Y_short_incl_open - Y_open).
and convert
Y_short -> Z_short

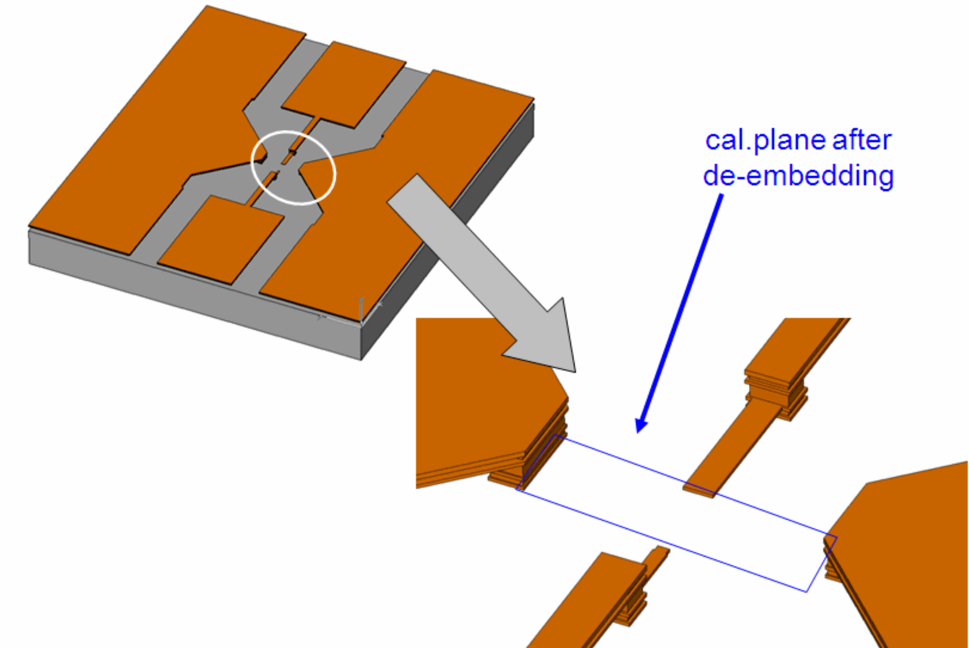
After so much theory, a look at the reality



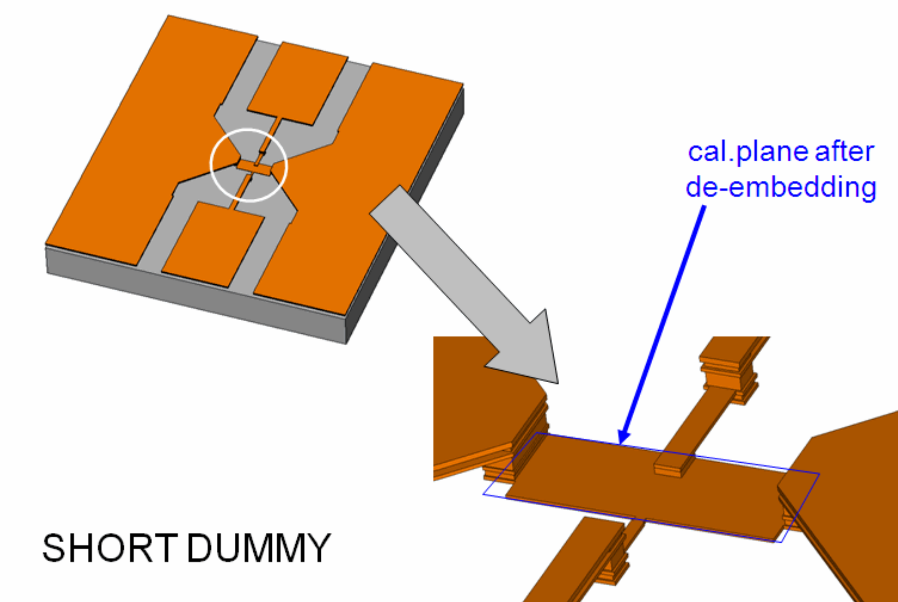
This slide depicts a three-dimensional representation of such a multi-finger MOS transistor. It is interesting to note how 'high' the metal planes are relative to the active silicon area and that the size of the transistor itself is very small compared to the size of the pads.
The goal of de-embedding is to shift the NWA calibration plane, which is (after the NWA calibration) at the end of the probe tips, down to the beginning of the transistor. This beginning is one of the key points for a good de-embedding and a key point for good dummy structures
The next slides depict the corresponding dummy structures.



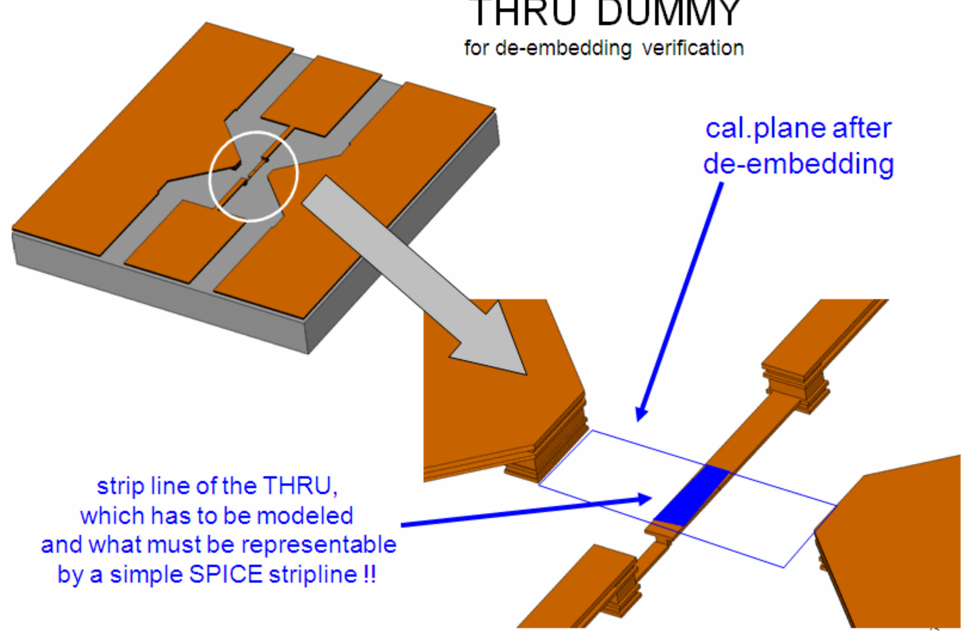
In the above slide, in the section 'Magnification of the transistor' is the inner transistor, which we want to model. All the rest has to be de-embedded, i.e. to be stripped-off. In other words, the NWA calibration plane has to be shifted down here.



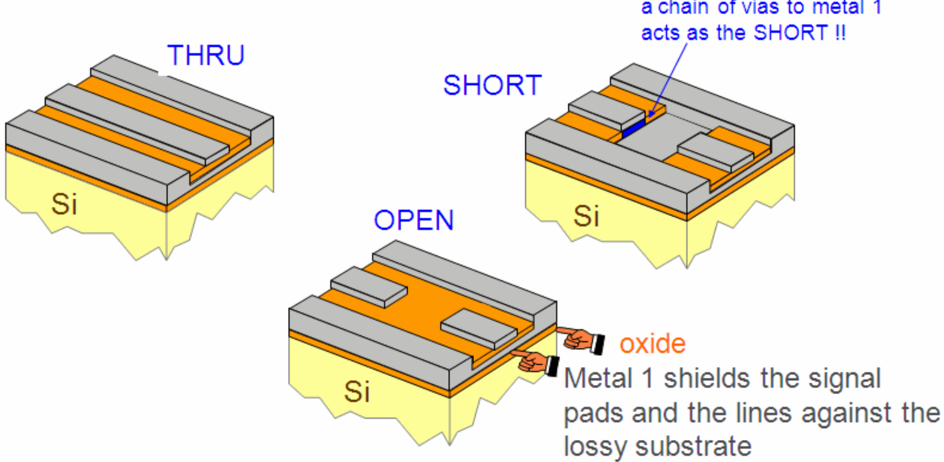
Note that the limits of the blue surrounded area in the OPEN dummy will become the shifted calibration plane. All what is inside will become part of the transistor model!



The SHORT dummy again refers to the desired shifted calibration plane. All what is inside this plane is now filled up with metal, so that we can consider this part to behave ideal, while the striplines from Gate and Drain will behave like inductors.



This is the THRU dummy, for de-embedding verification purpose. See the next slides for a proposed de-embedding verification procedure..



This layout suggestion for Silicon on-wafer dummy structures uses metal1 to shield the contact pads from the lossy silicon substrate.Due to this metal1 shielding, the measured OPEN dummy S-parameters represent a much more ideal OPEN dummy schematic than the more conventional layout without the metal1 shielding.
Also, for the SHORT dummy, the short itself is much more ideal because of the chain of shorting vias. This chain is geometrically concentrated and precisely localized, compared to the conventional SHORT dummy layout there the 'hole' is simply filled up with metal1.

Note
For clarity, the layout details are simplified in the slide above. A correct layout would be composed of conventional contact pads and a stripline each towards the component input and output, like for conventional layouts. So, the main difference between the layout sketch above and 'conventional' layouts is to use the metal1 as a shield fo the signal pads and lines against the lossy silicon substrate.

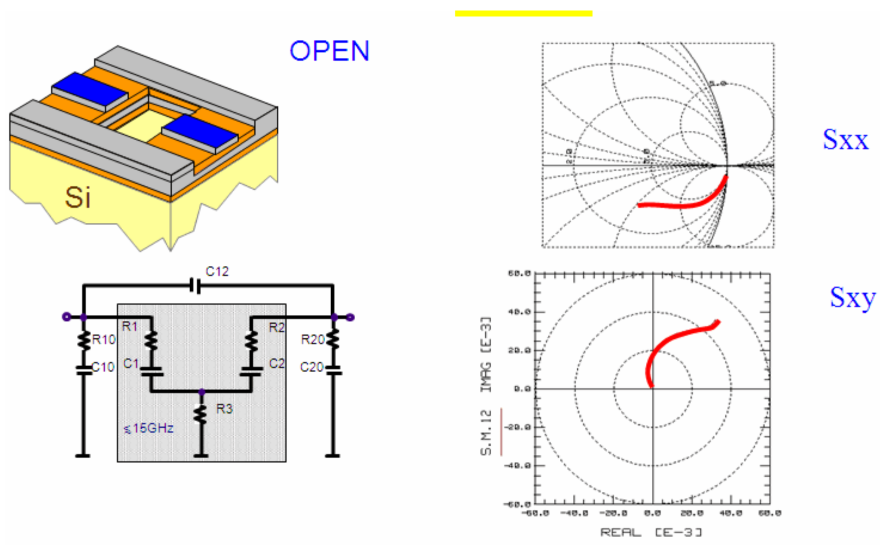
See publication T.E.Kolding, O.K.Jensen, T.Larsen, "Ground-Shielded Measuring Technique for Accurate On-Wafer Characterization of RF CMOS Devices", IEEE Int.Conf.on Microelectronic Test Structures ICMTS, March 2000, Monterey,CA

Verifying the de-embedding for on-Wafer transistor measurements

The Idea:

- Have an OPEN, a SHORT and a THRU dummy measured
- Model the dummies in order to check the previously mentioned prerequisites
- de-embed the THRU from the OPEN and SHORT and model it
- We should get a simple lumped schematic consisting of essentially a delay line which corresponds to the THRU on the wafer

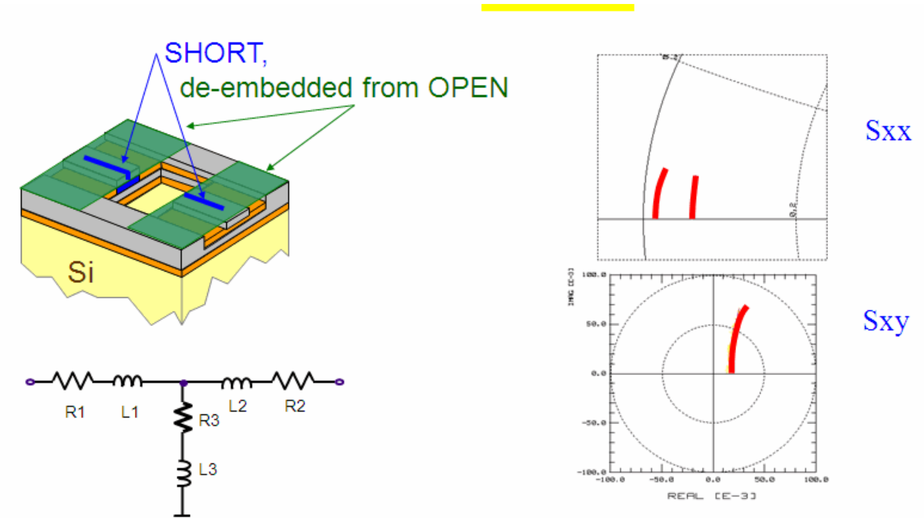
Model The On-Wafer OPEN Dummy



In this first step, the OPEN dummy is modeled in order to verify that there are no serial components included in the Y matrix of the OPEN.

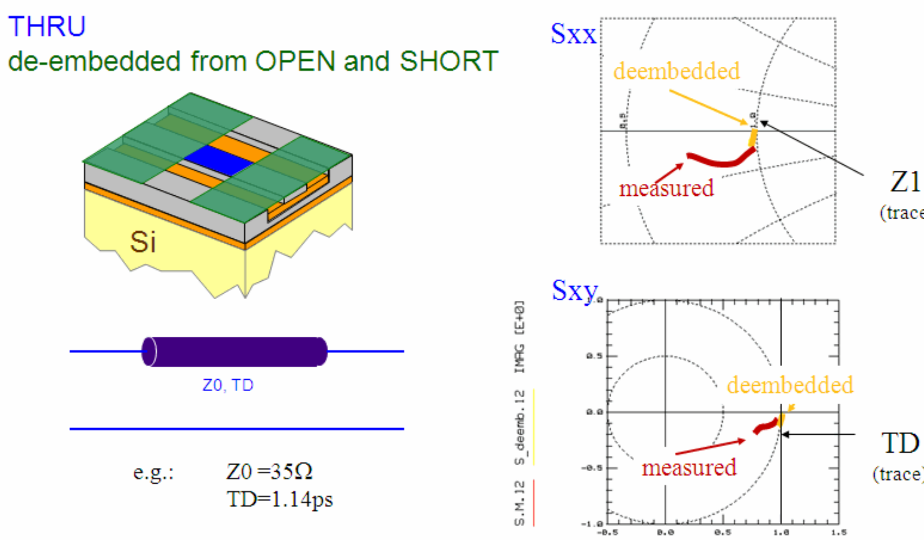
Hint:
In the demo_features directory, load the file 3_MEAS_ORGANIZE_n_VERIFY_DATA\0_MASTER_FILES\3_DEEMB_VERIFICATION\CHECK_DEEMB_OpenShortThru_MASTER_PELdep.mdl

Model The On-Wafer SHORT Dummy



Then, the SHORT is de-embedded from the OPEN and again modeled. THe goal is to verify that there are no parallel components included in the Z matrix of the SHORT.

Verify The De-Embedding With Modeling The THRU Dummy



After the OPEN and (if available/required) SHORT dummy measurements have been modeled, we want to check the de-embedding procedure with a known 'golden device'. In practice, however, like with the NWA cal.verification, the problem is the availability of an 'golden device'. Provided we have a THRU dummy, we can use this measurement for a verification of the de-embedding procedure. The idea is that the THRU should look like a simple delay line, with a certain characteristic impedance Z0, and a delay time TD. Additionally, the delay line may be lossy. But in reality, due to the short geometry of the THRU, this is typically not the case. Anyhow, the model parameters must represent physical meaningful values, and the de-embedded trace must be simple without resonances etc: just a simple delay line!

- TD: models the phase in Sxy
- Z0: models the Sxx trace:
  - for Z0<50W, the curve starts at 50W, turns downwards and then to the left into a circle around the line's Z0. The turning is clock-wise with increasing frequency!
  - for Z0=50W, the curve is a single point at at 50W in the Smith chart.
  - for Z0>50W, the curve starts at 50W, turns upwards and then to the right into a circle around the line's Z0. The turning is clock-wise with increasing frequency!

Hint:

In the demo_features directory, load the file
3_MEAS_ORGANIZE_n_VERIFY_DATA\0_MASTER_FILES\3_DEEMB_VERIFICATION\CHECK_DEEMB_OpenShortThru_MASTER_PELdep.mdl

and model the THRU with the tuners in Setup THRU.

Only if there were no layout bugs or measurement problems with the Dummy structures, we will end up with a physical representation of the small THRU strip line. Therefore, when we obtain that physical strip line, we can assume that the de-embedding, applied later to transistors, diodes etc., will be correct as well!

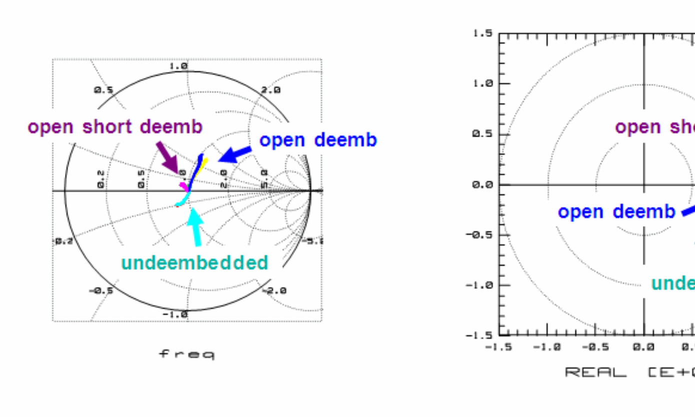
Note
Since the THRU strip line is passive, and not depending on bias or frequency or temperature, it can be used much better for de-embedding verification than checking the de-embedding with a transistor, a diode etc.

As expected, the THRU dummy can be modeled by a delay line. Therefore, we can assume a correct de-embedding procedure, that is, the inner transistor will now be de-embedded correctly.
This means:
After we are sure about the de-embedding, we apply:

$$Z_{DUT} = Z(Y_{TOTAL} - Y_{OPEN}) - Z(Y_{SHORT} - Y_{OPEN})$$

to all measured S-parameters for a correct and verified de-embedding.

An Example of bad THRU deembedding results



The slide above represents a de-embedding of dummy structures with obviously layout problems.
Although the inspection sequence undeembedded -> open_deemb -> open_short_deemb exhibits the right sequence (the resulting THRU becomes more and more the expected simple, short THRU), the open_short_deemb results does not match the expected the modeling target and represents - a strip line with negative delay time (see Sxy: a short delay line with neg. delay time and even with a trend for MAG(Sxy)>1 with increasing frequency) - and an unphysical Z0 (since in Sxx, it turns counter-clock-wise).

De-Embedding Conclusions

De-Embedding is used to strip-off high frequency parasitics from the DUT
In order to determine the required de-embedding steps and to also verify the de-embedding process, it is recommendable to also have a THRU dummy available on the wafer.

Try yourself in IC-CAP

Hit the blue Examples icon,select the demo_features directory and go further to subdirectory

3_MEAS_ORGANIZE_n_VERIFY_DATA/0_MASTER_FILES/31_DEEMB_VERIFICATION

and load the file CHECK_DEEMB_OpenShortThru_MASTER_demodata_PELdep.mdl

Note
You need to load the additional file DEPOTS.mdl, which is located under demo_features directly.

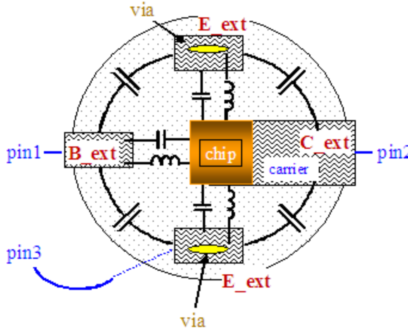
De-Embedding Techniques Tutorial

High frequency measurements are always influenced by the parasitic components (contact pads, packaging, test fixture). This makes the determination of the model parameters of the 'inner device' often pretty complicated. In some cases it is possible to measure and characterize the parasitic components of the contact pads, of the packaging or of the test fixture alone. In other cases we might have to use 'hot/cold' measurement techniques in order to separate the parasitic components from the 'inner device'. Once the values of the parasitic components are known, their effect on the total measurement can be eliminated by matrix operations and we get the s-parameters of the 'inner device'. This is explained in the following tutorial.

De-embedding tutorial

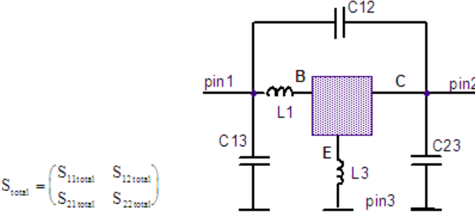
Assumed, we are going to model a transistor chip mounted on a chip carrier of a transistor case. The chip carrier will distort the 'inner' transistor's performance.
Yet, the chip carrier can be described with series inductors of the bond wires and the parallel capacitors of the pads (See the figure below).

Figure1: Chip carrier of a bipolar transistor case and its parasitic components



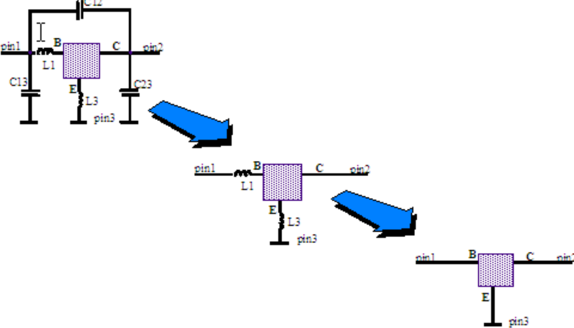
From Figure1, we obtain the simplified equivalent schematic given in Figure2.

Figure2: The equivalent schematic of the chip carrier for high frequencies.



About the de-embedding

The de-embedding procedure is sketched below. It means stripping off the parasitic components from the outside towards the inner DUT.



Stripping off the parasitic components gives the performance of the inner transistor.

De-embedding step by step:
We have measured the S-parameters Stotal of the transistor and its package. We assume that the components have been modeled already. (Note: de-embedding by complete matrix subtraction and not lumped components subtraction is covered further below).

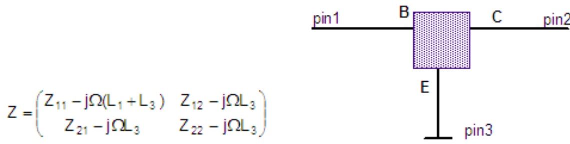
Following our assumption of having parallel capacitors as 'outer' parasitic components, we transform the S-parameters to Y, because a Y matrix represents a PI structure of components. A simple subtraction will de-embed the parasitic capacitor effects.

Figure 3: The equivalent schematic of fig.2 after the de-embedding of the capacitors

$$Y = \begin{pmatrix} Y_{11} - j\Omega(C_{13} + C_{12}) & Y_{12} + j\Omega C_{12} \\ Y_{21} + j\Omega C_{12} & Y_{22} - j\Omega(C_{23} + C_{12}) \end{pmatrix}$$

Now, the 'outer' parasitic components are the two inductors, which are in series with the chip connections. Series parasitics can be easily eliminated by subtracting a Z matrix. Therefore, we transform the Y-parameters from above into Z-parameters and subtract the inductors.

Figure 4: The equivalent schematic of fig.3 after the de-embedding of the series inductors



$$Z = \begin{pmatrix} Z_{11} - j\Omega(L_1 + L_3) & Z_{12} - j\Omega L_3 \\ Z_{21} - j\Omega L_3 & Z_{22} - j\Omega L_3 \end{pmatrix}$$

These Z-parameters are finally transformed back into S-parameters which describe the performance of the 'inner' chip.

$$S_{chip} = \begin{pmatrix} S_{11chip} & S_{12chip} \\ S_{21chip} & S_{22chip} \end{pmatrix}$$

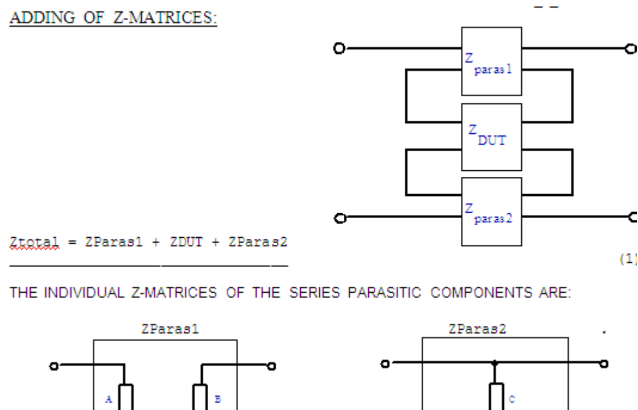
MATRIX MANIPULATIONS IN DETAILS

Note: the mentioned IC-CAP model files can be found in $ICCAP_ROOT\examples

and then
demo_features\3_MEAS_ORGANIZE_n_VERIFY_DATA\0_MASTER_FILES\3_DEEMB_UTILITIES

In the following pages, we will discuss the matrix properties and manipulation schemes relevant for de-embedding.

DE-EMBEDDING OF SERIES PARASITICS (Z-MATRIX)

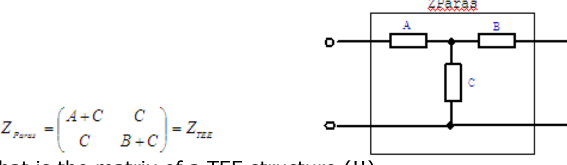


i.e.:

$$Z_{Paras1} = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \qquad Z_{Paras2} = \begin{pmatrix} C & C \\ C & C \end{pmatrix}$$

Adding these matrices of parasitic components gives:
Zparas = ZParas1 + ZParas2 (2)

or:



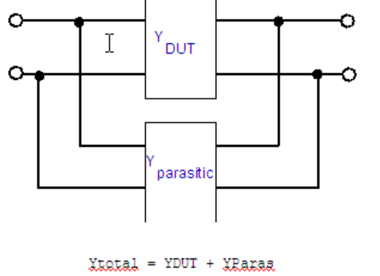
$$Z_{Paras} = \begin{pmatrix} A+C & C \\ C & B+C \end{pmatrix} = Z_{TEE}$$

what is the matrix of a TEE structure (!!).

How to proceed with the de-embedding of series parasitics:
Once the parasitic components A, B and C are known, they can be eliminated by subtracting the Z-matrix Zparas = ZTEE of a TEE structure from the measured data Ztotal. We obtain:
(1) (2)
ZDUT = Ztotal - ZParas1 - ZParas2 = Ztotal - Zparas

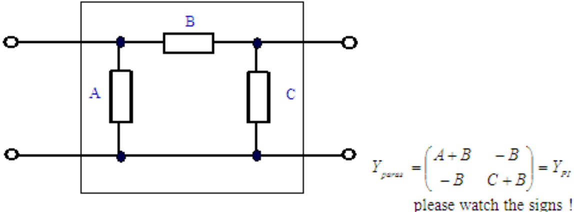
DE-EMBEDDING OF PARALLEL PARASITICS (Y-MATRIX)

ADDING OF Y-MATRICES:



Ytotal = YDUT + YParas (1)

THE Y-MATRIX OF THE PARALLEL PARASITIC COMPONENT IS:



$$Y_{paras} = \begin{pmatrix} A+B & -B \\ -B & C+B \end{pmatrix} = Y_{PI}$$

please watch the signs !

How to proceed with the de-embedding of parallel parasitics:

Once the parallel parasitic components A, B and C are known, they can be eliminated by simply subtracting the Y matrix YPI of a PI structure from the measured data Ytotal and we obtain:

(1)
YDUT = Ytotal - YParas

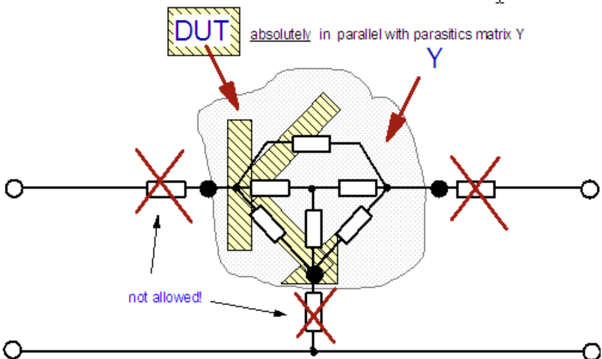
DE-EMBEDDING BY SUBTRACTING THE Z AND Y MATRICES

without modeling their components
IC-CAP files: 3_z_y_deemb.mdl
3_pad_short_open.mdl

In case, we are sure that the connections to the DUT do not suffer from any inductive losses and the 'inner' parasitics are all in parallel with the DUT, we can use a simple Y matrix subtractions for the de-embedding of the DUT. In the other case, when the parasitics are more complex, but still can be separated into a full parallel and another full serial part, a Y and Z matrix subtraction can be applied. In order to determine the components of these matrices, we need a measurement condition where the DUT behaves like a short (either a transistor in full saturation or a dummy device with a SHORT) and another one in which it behaves like an OPEN (a transistor without any bias voltage or an OPEN dummy device).

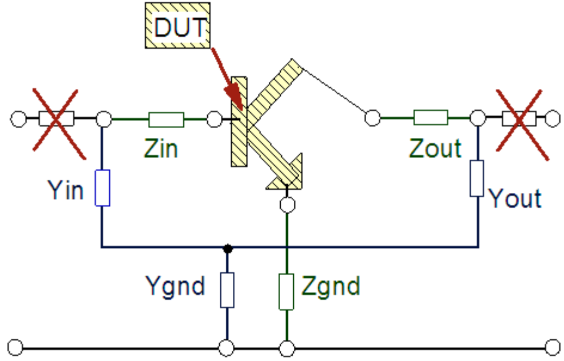
As stated above, it is assumed that the SHORT contains no components in parallel with the DUT (like capacitors) and that the OPEN includes no components that are in series with the connections of the DUT (inductors).

For the case of parasitics only in parallel with the pins of the DUT, we have the following situation:



The de-embedding is performed by simply subtracting the Y matrix of the parasitics from the total measurement. This is usually o.k. up to 10GHz.

For higher frequencies, we must also consider the parasitics which are in series with the DUT. This is illustrated in fig.2. Again, it is assumed that there are only series parasitics included in the Z matrix, and nothing but parallel parasitics in the Y matrix. The depicted situation with the complex resistors is a typical one for on-wafer transistor measurements.



If these conditions are true, the DUT can be embedded like follows, requiring these measurement results:
'total' device
'short' dummy device
'open' dummy device.

The de-embedding of the 'inner' device-under-test (the DUT) from the parasitics is then:

Convert S_total to Y_total and calculate

Y_total_w_short = (Y_total - Y_open)

to de-embed the outer parallel parasitics.

Then, convert Y_total_w_short to Z_total_w_short and calculate

Z_dut = (Z_total_w_short - Z_short_wo_open)

to de-embed the inner serial parasitics.

Finally, re-convert Z_dut to S_dut to obtain the S-parameters of the 'inner' DUT.

Important Note:
matrix Z_short has to be de-embedded first from the outer parallel parasitics, before subtracting it in the second step above.
To do this, convert S_short to Y_short and calculate

Y_short_wo_open = (Y_short - Y_open)

HOW TO DETERMINE THE SEQUENCE OF Y AND Z MATRIX DE-EMBEDDING

In the example above, the serial parasitics are assumed to be close to the DUT, while the parallel parasitics are around them. This is typically the case for on-wafer transistor measurements. Here, usually, the parallel parasitics, the pad capacitors, are the outer ones, while the small lines to the transistors, represented by series inductors are the inner ones.

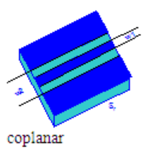
For packaged transistors, this is usually just the other way around.

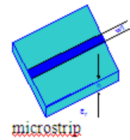
In order to check the sequence of de-embedding or also to determine if an OPEN de-embedding alone is sufficient, it is best to also have a dummy THRU structure. Such a THRU is represented by a delay line. If the THRU after the de-embedding represent a

simple delay line (to compared with a simulated delay line, the de-embedding procedure can be considered o.k.

Another method, however not always successful, is to convert the OPEN dummy S_parameter to Z and to plot MAG(Zxx) and MAG(Zxy) versus the frequency. If there is no rise in magnitude, it can be assumed that there are no hidden inductors in the OPEN structure. For the SHORT, a check of MAG(Yxx) and MAG(Yxy) might give some more insight.

DE-EMBEDDING OF LOSSLESS 50 Ohm DELAY LINES (S-MATRIX)


coplanar

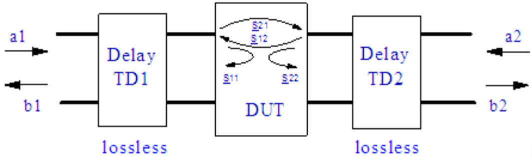

microstrip

Assumption: we are in the S-parameter domain and have the device-under-test (DUT) imbedded between two delay lines:

with a: forward wave
b: reflected wave

In this case, there is b1 = s11total*a1 + s12total*a2

b2 = s21total*a1 + s22total*a2



and the S-parameter matrices of the individual delay lines are:

$$S_{TD1} = \begin{pmatrix} 0 & e^{-j*2PI*f*TD1} \\ e^{-j*2PI*f*TD1} & 0 \end{pmatrix} \qquad S_{TD2} = \begin{pmatrix} 0 & e^{-j*2PI*f*TD2} \\ e^{-j*2PI*f*TD2} & 0 \end{pmatrix}$$

How to proceed with the de-embedding of lossless delay lines:

Provided we know the delay times TD1 and TD2, we can use the 'port shift' properties of S-parameter multiplications (shift of reference planes), and obtain:

$$\begin{pmatrix} S11_{DUT} & S12_{DUT} \\ S21_{DUT} & S22_{DUT} \end{pmatrix} = \begin{pmatrix} S11_{total} * e^{j \cdot 2PI \cdot f \cdot 2 \cdot TD1} & S12_{total} * \exp^{j \cdot 2PI \cdot f \cdot (TD1+TD2)} \\ S21_{total} * \exp^{j \cdot 2PI \cdot f \cdot (TD1+TD2)} & S22_{total} * \exp^{j \cdot 2PI \cdot f \cdot 2 \cdot TD2} \end{pmatrix}$$

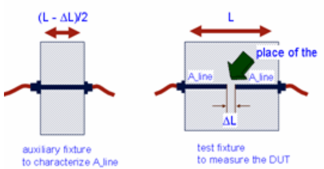
with: f : frequency
TD1: delay time at port 1,
TD2: delay time at port 2

DE - EMBEDDING USING ABC - MATRICES

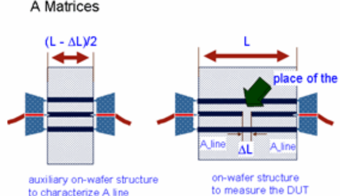
IC-CAP file: 6_a_deemb.mdl

This kind of de-embedding is related to measurement conditions where we have considerably long connection lines to the device. It can be applied for example to insertable devices like a connector etc., but also to special components line spiral inductors. Besides the test fixture for the DUT, which consists basically of a substrate with a strip line for each port of the DUT, we also need a special test structure to characterize the strip line.

This auxiliary fixture consists of a strip line on the same substrate like the test fixture, and the strip line is exactly as long as on the test fixture. In this case, we can apply ABC-matrices (A-matrix, cascade matrix) to perform the de-embedding. The sketch below shows the required substrates for a connectorized situation



and for a on-wafer situation:



With the assumption that the Aline ABC matrix of the auxiliary structure is identical to the connection strip lines on the wafer or on the test fixture (in a connectorized scenario, the connector has been de-embedded !), the total performance of the device-under-test (DUT) including the test fixture can be expressed in A matrices:

$$A_{total} = A_{line} * A_{DUT} * A_{line} \qquad (1)$$

or solved for ADUT (multiplication 'at the left' with , and 'at the right' with:

$$A_{DUT} = A_{line}^{-1} * A_{total} * A_{line}^{-1} \qquad (2)$$

How to proceed:
The measured S-parameters Stotal are transformed to Atotal using IC-CAP's TwoPort function. The same applies to the measured Sline parameters. Then the matrix calculation from (2) is applied in order to obtain the de-embedded ADUT.

IMPORTANT NOTE: when cutting an electro-magnetic effect into a chain of effects, make sure to set the cut where the E-fields have a monotonous trace (where the EM effects of geometrical discontinuities have died out). As an example, when isolating a via from a stripline-via-stripline scenario, the via should be cut off from the strip line not at the location of the via, but rather including a bit of the strip lines.

DE - EMBEDDING OF CHAINS OF ABC - MATRICES

IC-CAP file: 6_chain_of_a_matrices_deemb_verif.mdl

In case of the de-embedding of an 'inner' A-matrix, embedded within a chain of A-matrices, we need to proceed as follows, regarding the inverse multiplication of A-matrices:

Let's assume we have the following situation:

$$\boxed{A_{total} = A_{connleft} * A_{lineleft} * A_{DUT} * A_{lineright} * A_{connright}} \qquad (1)$$

Multiplying both sides of the equation 'at the left' with $A_{connleft}^{-1}$ gives:

$$A_{connleft}^{-1} * A_{total} = \left(A_{connleft}^{-1} * A_{connleft}\right) * A_{lineleft} * A_{DUT} * A_{lineright} * A_{connright}$$

or

$$A_{connleft}^{-1} * A_{total} = A_{lineleft} * A_{DUT} * A_{lineright} * A_{connright} \qquad (2)$$

In the next step, we end up with

$$A_{lineleft}^{-1} * A_{connleft}^{-1} * A_{total} = A_{DUT} * A_{lineright} * A_{connright} \qquad (3)$$

Now, we multiply with the inverse matrices 'at the right', and obtain these intermediate steps:

$$A_{lineleft}^{-1} * A_{connleft}^{-1} * A_{total} * A_{connright}^{-1} = A_{DUT} * A_{lineright} * \left(A_{connright} * A_{connright}^{-1}\right)$$

or

$$A_{lineleft}^{-1} * A_{connleft}^{-1} * A_{total} * A_{connright}^{-1} = A_{DUT} * A_{lineright} \qquad (4)$$

And, after the next step, we finally have isolated A_{DUT} :

$$A_{lineleft}^{-1} * A_{connleft}^{-1} * A_{total} * A_{connright}^{-1} * A_{lineright}^{-1} = A_{DUT}$$

or:

$$\boxed{A_{DUT} = A_{lineleft}^{-1} * A_{connleft}^{-1} * A_{total} * A_{connright}^{-1} * A_{lineright}^{-1}}$$

IMPORTANT NOTE:
watch the flipped sequence (with respect of the original chain of twoports !!) of the inverse A-matrices chain !!!

DE-EMBEDDING OF LOSSY DELAY LINES WITH GENERAL CHARACTERISTIC IMPEDANCE

This is, again, an application of ABC-matrices (A-matrices).

Knowing the de-embedding procedure for A matrices shown above, we now can apply it to the de-embedding of striplines with general impedance Z0. Compared to the S-parameter de-embedding above, the stripline can also be lossy in this case.

First, we have to define the ABC matrix A_{line} of a strip line, with:

Z0	characteristic impedance of the strip line
γ	propagation coefficient of the line
α	loss
ß or TD	phase shift or also delay time of the strip line,

$$\gamma \cdot \omega \cdot L = \left[\sqrt{\omega} \cdot \alpha + j \cdot \omega \cdot ß\right] \cdot L \overset{\text{for lossless line}}{\sim} j \cdot ß \cdot \omega \cdot L = j \cdot \omega \cdot TD$$

Note: roughly TD=10ps per mm strip line can be assumed on a ceramic substrate.

In case of a lossy line, we calculate first:

$$aux[i] = \left(\sqrt{2 * PI * freq[i]} * \alpha + j * 2 * PI * freq[i] * \beta\right) * L \qquad (1)$$

and in case of a lossless line (α=0), we calculate first:

$$aux[i] = j * 2 * PI * freq[i] * \beta * L \quad \text{or} \quad aux[i] = j * 2 * PI * freq[i] * TD$$

which gives for the **ABC matrix of the delay line**:

$$(A_{line}[i]) = \begin{pmatrix} \cosh(aux[i]) & Z0 * \sinh(aux[i]) \\ \frac{1}{Z0}\sinh(aux[i]) & \cosh(aux)[i] \end{pmatrix} \qquad (2)$$

Such an ABC matrix after (2) is then introduced in the ACB matrix de-embedding formula

$A_{DUT} = A_{line}^{-1} * A_{total} * A_{line}^{-1}$,

and we can calculate the de-embed the matrix A_{DUT}.

NOTE: Delay Lines in ADS are specified by length L, so for a lossless delay line, we can substitute TD = L / 3E+8

SPLITTING A THRU INTO 2 SYMMETRICAL MIRROR HALVES

or

OPEN -THRU DE-EMBEDDING

APPLYING DIFFERENTIAL TWOPORT-PARAMETERS

See demo_features\3_MEAS_ORGANIZE_n_VERIFY_DATA\0_MASTER_FILES\31_DEEMB_VERIFICATION\CHECK_OPEN_THRU_DEEMB_demodata_MASTERFILE.mdl

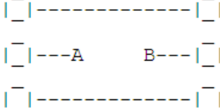
This de-embedding idea is after

Jon Tao, Paul Findley and G. Ali Rezvani, Novel Realistic Short Structure Construction for Parasitic Resistance De-embedding and on-Wafer Inductor Characterization, Proc. IEEE 2005 Int. Conference on Microelectronic Test Structures, Vol. 18, April 2005

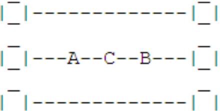
It is especially interesting for the de-embedding of spiral inductors.

The dummy structures should look like this:

- Sketch of the OPEN dummy structure ('A' marks the end of line1, 'B' the end of line2)



- and of the THRU dummy structure ('C' marks the center of the THRU line)



When applying differential TwoPort measurements to a THRU dummy, we obtain a virtual Ground in the middle of the THRU dummy.
The center of the THRU becomes a virtual ground,
i.e. the differential TwoPort parameters of the THRU represent a virtual SHORT.

Measuring a THRU instead of a SHORT avoids

- S-parameter measurement problems due to the near-zero impedance.
- The uncertainty of the impedance measurement of the conventional SHORT dummy structure, where it is assumed to measure from pad1 to A and also from pad 2 to B the
series impedance of the SHORT, while the shortening from A -> Ground and B -> Ground is assumed to be ideal, i.e. 0 Ohm.

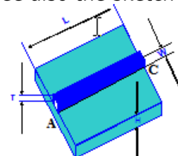
After the conventional OPEN de-embedding, and based on the above layout sketches, the differential impedance of the THRU structure can be calculated as

Zdiff = 1/2 (Z11 + Z22 - Z12 - Z21)

i.e. the impedance of each port to the center virtual ground point (C) is:

Zc = 1/2 Zdiff

We now need to consider the electrical performance between points A and C of the THRU (see also the sketch above):



Assuming the length of the metal stripe between point A (the later connection point of the spiral inductor) and point C (the center point of the THRU) is L, and the metal width is W (L and W can be measured from the layout), and T is the metal thickness, and assuming the specific resistance of the metal layer is rho , the resistance of the line between A and C can be calculated as

and for the special case rhoAluminum ~ 0,0264 um²/um :
R_{AC} ~ 0.0264*L[um] / (W[um]*T[um])

Similarly, the inductance between A and C can be estimated.
For a strip line width W > height of dielectric (H), and line length L, it becomes approximately

$$L_{AC} \sim \frac{1.2\text{pH}}{\text{um}} * L[\text{um}] \cdot \frac{1}{1.4 \cdot \frac{W}{H} + .667 \cdot \ln\left(1.444 + \frac{W}{H}\right)}$$

Therefore, the Z-parameters of the virtual SHORT structure can be calculated as

$$Z11 = Z22 = Zc - R_{AC} - j*2*PI*freq*L_{AC}$$
$$Z12 = Z21 = 0 \qquad (1)$$

By substituting the Z-parameters of the conventional SHORT dummy structure with the Z-parameters of (1), we can apply the well-known OPEN-SHORT de-embedding, as it is described in the chapter 'DE-EMBEDDING BY SUBTRACTING THE Z AND Y MATRICES' further above.

> **Note**
> for strip line width W < height of dielectric (H), and line length L.

$$L_{AC} \sim \frac{0.2\text{pH}}{\text{um}} * L[\text{um}] \cdot \ln\left(\frac{8 \cdot H}{W} + \frac{W}{4 \cdot H}\right)$$

> **Note**
> the conversion of the THRU to a virtual SHORT can also be applied to verify the quality of the de-embedding (by comparing to the real SHORT measurement). Of course, in both cases, the OPEN has to be de-embedded from the THRU and also from the SHORT.

ft Modeling of a Transistor Affected by Parasitics

Tutorial About the f_t Modeling of a Transistor Affected by Parasitics

An introduction to de-embedding

In order to demonstrate the importance of de-embedding for proper s-parameter modeling, this chapter covers the fT modeling of a bipolar transistor with and without de-embedding.

> **Note**
> To demonstrate the de-embedding techniques in a well-defined situation, all data in this chapter come from simulations.

The individual steps of this tutorial are:

1. Inspection of the cutoff frequency curve of a bipolar transistor without any parasitic effects.
2. The same curve, but now distorted by package effects.
3. The de-embedding procedure
4. The de-embedded cutoff frequency curve that matches the one of step 1.

The cutoff frequency of a bipolar transistor chip is depicted in fig_1. It refers to the 'inner' transistor and was obtained from a simulation.

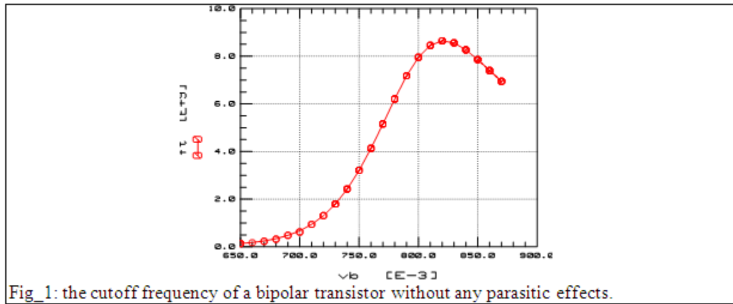


Fig_1: the cutoff frequency of a bipolar transistor without any parasitic effects.

In reality, we would not be able to measure the transistor under test in this manner. Assuming a packaged bipolar transistor, fig_2 shows some typical parasitic elements of such a package.

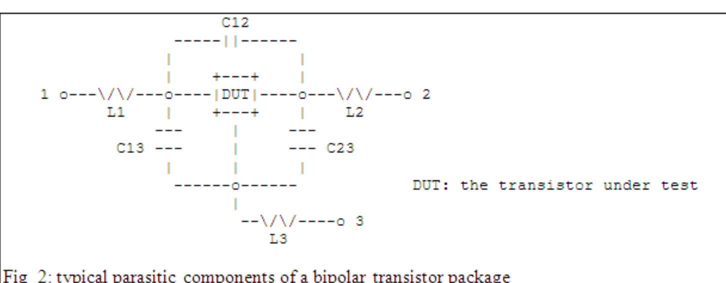


Fig_2: typical parasitic components of a bipolar transistor package

These parasitic elements will degrade the performance of the 'inner' DUT at high frequencies, as given in fig_3.

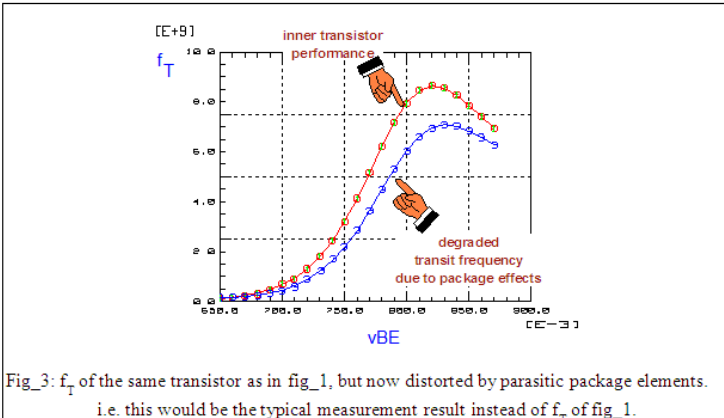


Fig_3: f_T of the same transistor as in fig_1, but now distorted by parasitic package elements. i.e. this would be the typical measurement result instead of f_T of fig_1.

Our goal is to de-embed the curve of fig._3 and to get back that one of fig_1. Provided the parasitic elements are known (what will be explained in the chapters below), we have to perform the following sequence of matrix operations in order to de-embed the 'inner' transistor:

Referring to fig_2, we commence with these pre-considerations:
The inductors are in series with the transistor pins. Therefore they can be represented best by a Z-matrix.
The capacitors are parallel to the DUT. They can be represented best by a Y-matrix.
This leads to the de-embedding procedure:

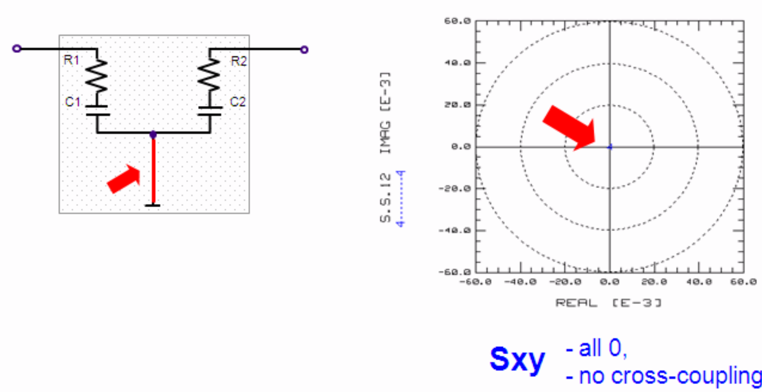
1. convert the measured 'distorted' S-parameters of fig_3 to Z-parameters
2. substract the known parasitic inductors of the SHORT (Z-matrix)
3. convert the Z-parameters to Y-parameters
4. substract the known parasitic capacitors of the OPEN (Y-matrix), which itself is de-embedded from the parasitics of the SHORT!
5. finally convert the Y-parameters back to S-parameters and obtain the data of fig_1
   Table_1: the de-embedding sequence

Hint: see also chapter De-embedding Techniques with Z,Y,S,A Matrices

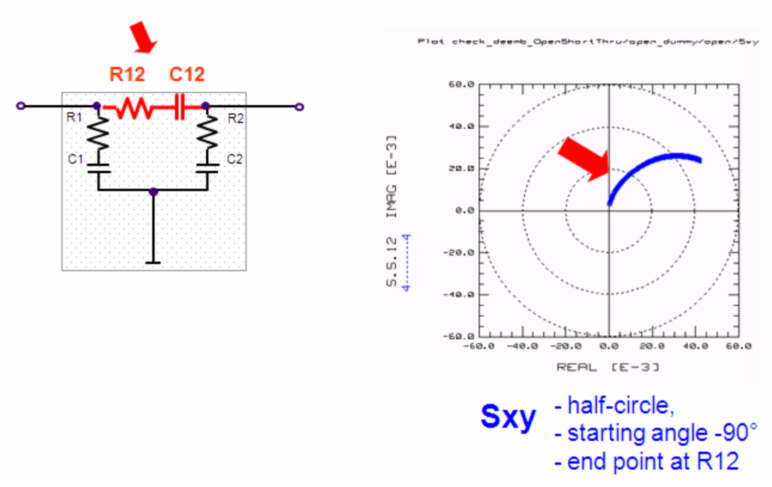
Interpreting OPEN-Dummy Measurements

This slide sequence is about the inspection of measurements of different OPEN dummy structures and about the interpretation of possible underlying physical layout problems.

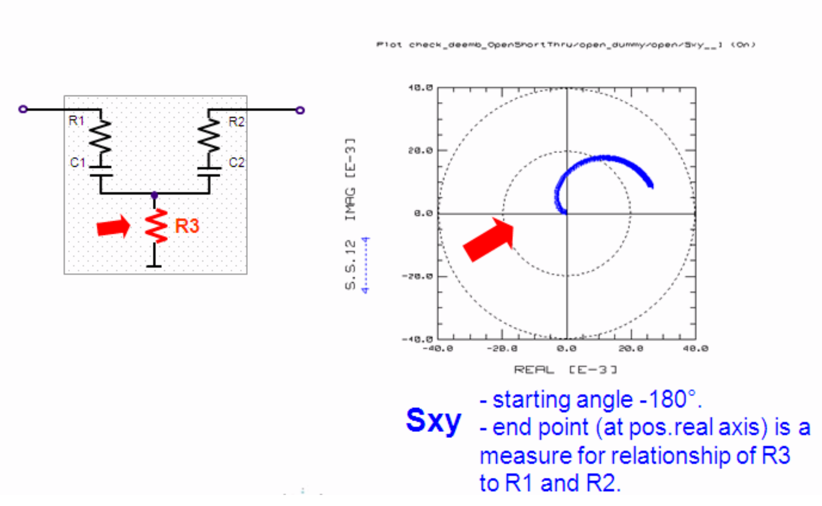
Effect of R1, C1 and R2, C2 on Sxy



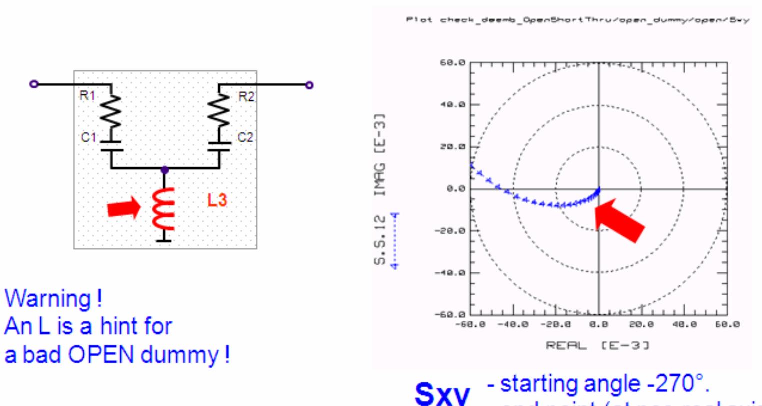
Effect of R12 and C12 on Sxy



Effect of R3 on Sxy



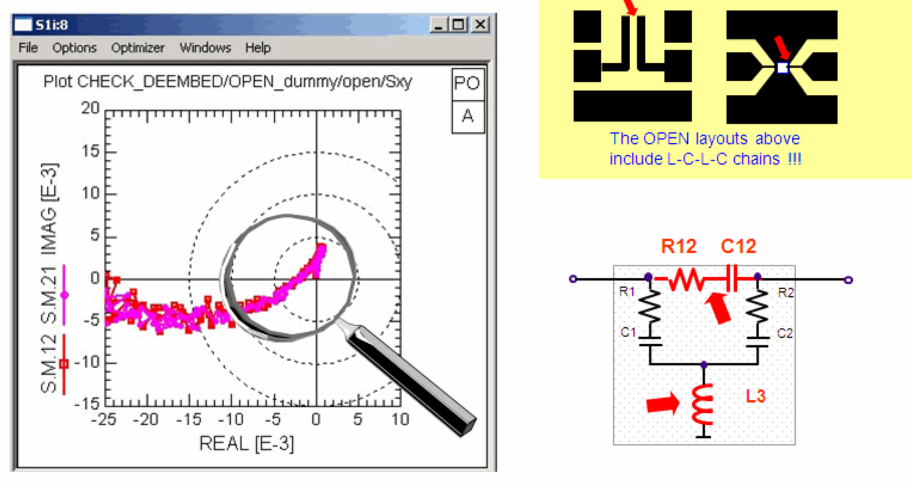
Effect of L3 on Sxy



> **Note**
> The necessity to include an inductor for modeling an OPEN dummy structure is a contradiction of the fact that all components of an OPEN dummy schematic are **in parallel** to the device, and therefore, can be de-embedded by subtracting the Y matrix of the OPEN dummy.

QUIZ:

and what schematic can be assumed behind this OPEN Sxy measurement??



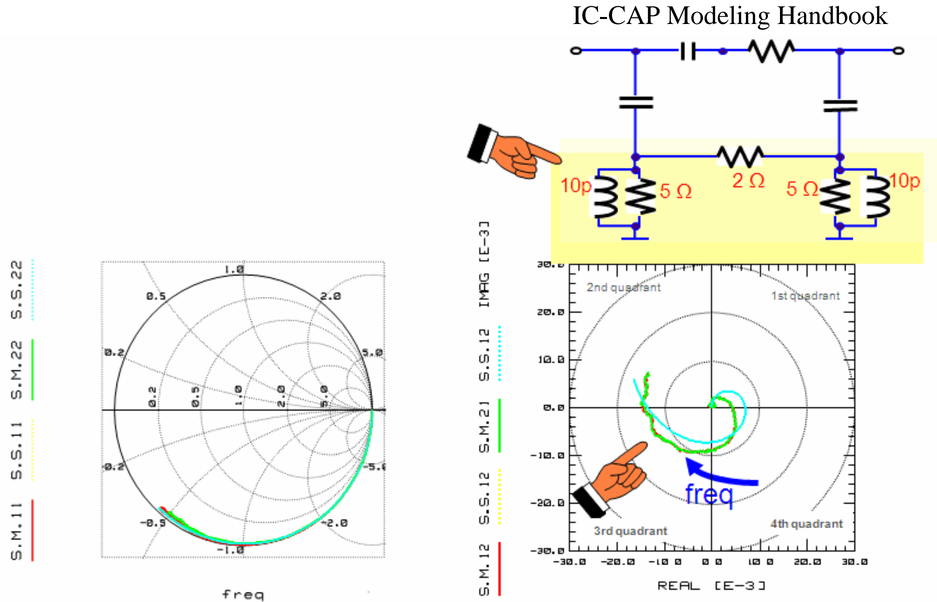
> **Note**
> If your OPEN measurement exhibits a trace as depicted above, it is pretty possible that the THRU de-embedding verification will fail.I.e. that the de-embedded THRU may have a MAG(Sxy) with more than 1, or may turn backwards (to the left), and that the Sxx may also turn backwards.

Open Dummy High Frequency Modeling

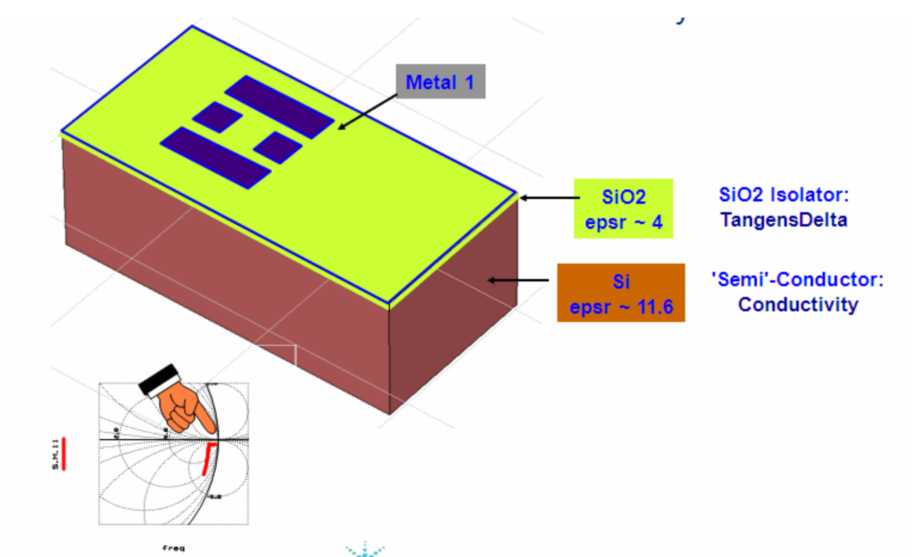
> **Note**
> Such a trace is pretty typically when not applying a metal-1 shield below the dummy layout structures.

> **Note**
> Such a trace is pretty typically when not applying a metal-1 shield below the dummy layout structures.

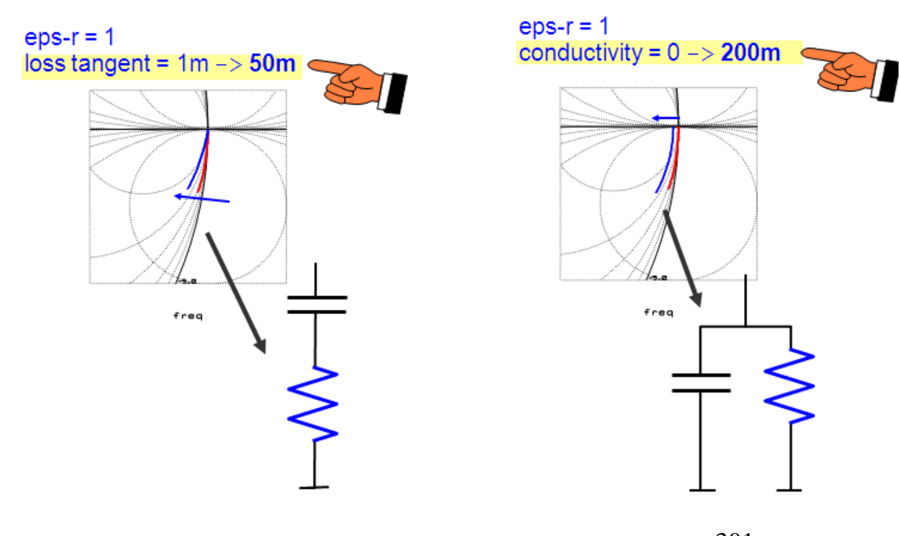
And if Sxy extends into 4th and into 3rd quadrant, its probably the bulk effect.



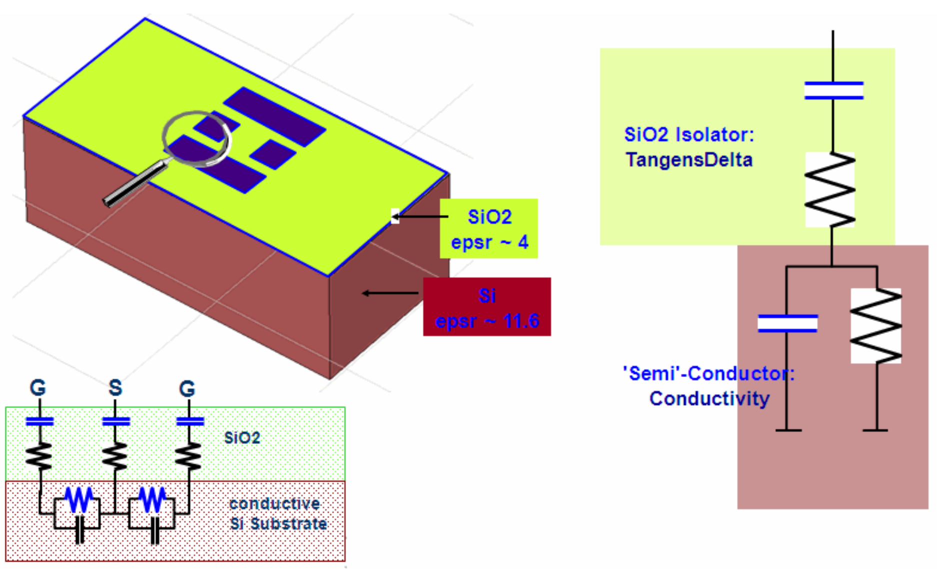
How the Silicon Substrate and the Silicon Oxide parameters affect the 'knee' in the OPEN Dummy Structure S11



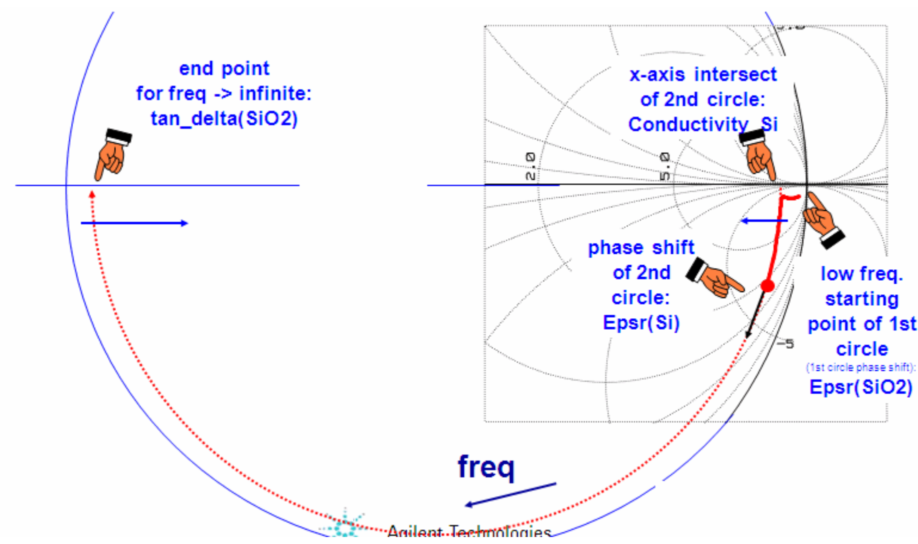
A study on the effect of the oxide loss tangent and the substrate conductivity



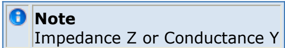
The signal pad of the OPEN sees basically a sequence of SiO2 > silicon substrate > SiO2, i.e. a circuit schematic like below

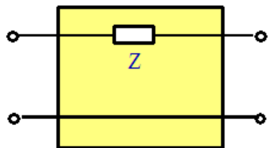


Interpreting the OPEN Sxx related to substrate and oxide params



Twoport Matrices of Basic Schematics

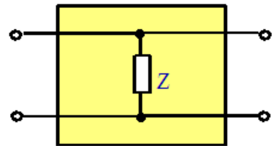




$$\text{Amatrix} = \begin{pmatrix} 1 & Z \\ 0 & 1 \end{pmatrix}$$

Z-matrix not applicable,

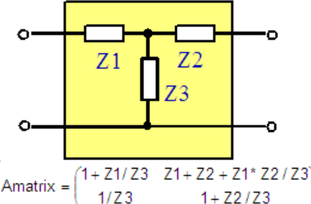
$$\text{Ymatrix} = \begin{pmatrix} 1/Z & -1/Z \\ -1/Z & 1/Z \end{pmatrix}$$



$$\text{Amatrix} = \begin{pmatrix} 1 & 0 \\ 1/Z & 1 \end{pmatrix}$$

$$\text{Zmatrix} = \begin{pmatrix} Z & Z \\ Z & Z \end{pmatrix}$$

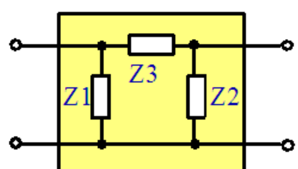
Y-matrix not applicable,



$$\text{Amatrix} = \begin{pmatrix} 1+Z1/Z3 & Z1+Z2+Z1^{*}Z2/Z3 \\ 1/Z3 & 1+Z2/Z3 \end{pmatrix}$$

$$\text{Zmatrix} = \begin{pmatrix} Z1+Z3 & Z3 \\ Z3 & Z2+Z3 \end{pmatrix}$$

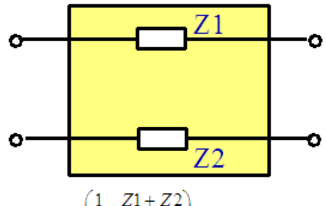
$$\text{Ymatrix} = \frac{1}{Y1+Y2+Y3} \cdot \begin{pmatrix} Y1\cdot(Y2+Y3) & -Y1\cdot Y2 \\ -Y1\cdot Y2 & Y2\cdot(Y1+Y3) \end{pmatrix}$$



$$\text{Amatrix} = \begin{pmatrix} 1+Z3^{*}Y2 & Z3 \\ Y1+Y2+Y1\cdot Y2\cdot Z3 & 1+Y1\cdot Z3 \end{pmatrix}$$

$$\text{Zmatrix} = \frac{1}{Z1+Z2+Z3} \cdot \begin{pmatrix} Z1\cdot(Z2+Z3) & Z1\cdot Z2 \\ Z1\cdot Z2 & Z2\cdot(Z1+Z3) \end{pmatrix}$$

$$\text{Ymatrix} = \begin{pmatrix} Y1+Y3 & -Y3 \\ -Y3 & Y2+Y3 \end{pmatrix}$$



$$Amatrix = \begin{pmatrix} 1 & Z1+Z2 \\ 0 & 1 \end{pmatrix}$$

Z-matrix not applicable,

$$Ymatrix = \frac{1}{Z1 + Z2} \cdot \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

Twoport Matrix Conversions

H to Y Conversion

Since a TwoPort matrix describes the behavior of its circuit completely, the matrices can be converted from one form into another. As an example, to convert a H matrix into its Y equivalent, we just have to rearrange the matrix equations. The H matrix has the form.

$$\begin{pmatrix} v1 \\ i2 \end{pmatrix} = \begin{pmatrix} h11 & h12 \\ h21 & h22 \end{pmatrix} * \begin{pmatrix} i1 \\ v2 \end{pmatrix}$$

and the target Y matrix

$$\begin{pmatrix} i1 \\ i2 \end{pmatrix} = \begin{pmatrix} y11 & y12 \\ y21 & y22 \end{pmatrix} * \begin{pmatrix} v1 \\ v2 \end{pmatrix}$$

The v1 dependency, i.e. the first line in the H matrix, can be solved for i1:

$$i1 = \frac{1}{h11} v1 - \frac{h12}{h11} v2$$

, what is already the first line of the Y matrix.

We insert this result into the second line of the H matrix in order to replace the i1 dependency,

$$i2 = \frac{h21}{h11} v1 + \left(h22 - \frac{h21}{h11} h12 \right) v2$$

with

$$\det(h) = h11 h22 - h12 h21$$

we can rearrange and yield,

$$i2 = \frac{h21}{h11} v1 + \frac{\det(h)}{h11} v2$$

what finally gives the relationship between the y and h parameters:

$$\begin{pmatrix} y11 & y12 \\ y21 & y22 \end{pmatrix} = \frac{1}{h11} \begin{pmatrix} 1 & -h12 \\ h21 & \det(h) \end{pmatrix}$$

Y to Z and Z to Y Conversion

The conversion of the Y to the Z parameters is, related to the special definition of these two matrices, very simple.

Starting with the resistance matrix form,

$$(v) = (z) * (i)$$

we can solve it for (i) by multiplying to the left side with (z)-1, the inverse matrix of (z).

$$(z)^{-1}(v) = (z)^{-1}(z) * (i)$$

$$(z)^{-1}(v) = (i)$$

or Comparing this result with the definition of the y matrix, we get finally,

$$(z)^{-1} = (y)$$

and correspondingly,

$$(y)^{-1} = (z)$$

Note
Remember the matrix inversion scheme.

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - cb} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

with ad - cb = det(matrix),

S to H Conversion

S to H conversion (S2H):

$$h_{11} = \frac{(1+s_{11}) \cdot (1+s_{22}) - s_{12}s_{21}}{(1-s_{11}) \cdot (1+s_{22}) + s_{12}s_{21}} * Z0$$

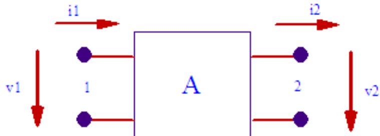
$$h_{12} = \frac{2 \cdot s_{12}}{(1-s_{11}) \cdot (1+s_{22}) + s_{12}s_{21}}$$

$$h_{21} = \frac{-2 \cdot s_{21}}{(1-s_{11}) \cdot (1+s_{22}) + s_{12}s_{21}}$$

$$h_{22} = \frac{(1-s_{11}) \cdot (1-s_{22}) - s_{12}s_{21}}{(1-s_{11}) \cdot (1+s_{22}) + s_{12}s_{21}} \; / \; Z0$$

Matrix Conversions at a Glance

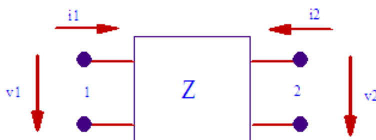
A Matrix



$$v1 = A11*v2 + A12*i2$$
$$i1 = A21*v2 + A22*i2$$

$$(A) = \begin{pmatrix} A11 & A12 \\ A21 & A22 \end{pmatrix} = \frac{1}{Z21}\cdot\begin{pmatrix} Z11 & \Delta Z \\ 1 & Z22 \end{pmatrix} = \frac{1}{Y21}\cdot\begin{pmatrix} -Y22 & -1 \\ -\Delta Y & Z22 \end{pmatrix} - Y11 = \frac{1}{H21}\cdot\begin{pmatrix} -\Delta H & -H11 \\ -H22 & Z22-1 \end{pmatrix}$$

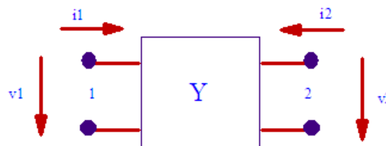
Z Matrix



$$v1 = Z11*i1 + Z12*i2$$
$$v2 = Z21*i1 + Z22*i2$$

$$(Z) = \begin{pmatrix} Z11 & Z12 \\ Z21 & Z22 \end{pmatrix} = \frac{1}{\Delta Y}\cdot\begin{pmatrix} Y22 & -Y12 \\ -Y21 & Y11 \end{pmatrix} = \frac{1}{H22}\cdot\begin{pmatrix} \Delta H & H12 \\ -H21 & 1 \end{pmatrix} - Y11 = \frac{1}{A21}\cdot\begin{pmatrix} A11 & \Delta A \\ 1 & A22 \end{pmatrix}$$

Y Matrix

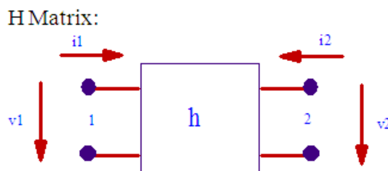


$$i1 = Y11*v1 + Y12*v2$$
$$i2 = Y21*v1 + Y22*v2$$

$$(Y) = \begin{pmatrix} Y11 & Y12 \\ Y21 & Y22 \end{pmatrix} = \frac{1}{\Delta Z}\cdot\begin{pmatrix} Z22 & -Z12 \\ -Z21 & Z11 \end{pmatrix} = \frac{1}{H11}\cdot\begin{pmatrix} 1 & -H12 \\ H21 & \Delta H \end{pmatrix} - Y11 = \frac{1}{A12}\cdot\begin{pmatrix} A22 & -\Delta A \\ -1 & A11 \end{pmatrix}$$

H Matrix

H Matrix:

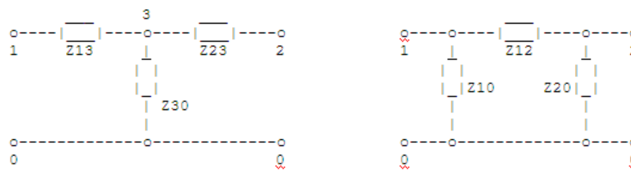


$$v1 = H11*i1 + H12*v2$$
$$i2 = H21*i1 + H22*v2$$

$$(H) = \begin{pmatrix} H11 & H12 \\ H21 & H22 \end{pmatrix} = \frac{1}{Z22}\cdot\begin{pmatrix} \Delta Z & Z12 \\ -Z21 & Y11 \end{pmatrix} = \frac{1}{H22}\cdot\begin{pmatrix} 1 & -Y12 \\ Y21 & \Delta Y \end{pmatrix} - Y11 = \frac{1}{A22}\cdot\begin{pmatrix} A12 & \Delta A \\ -1 & A21 \end{pmatrix}$$

TEE to PI Conversion

The following conversion equations explain how to convert a TEE structure to a PI structure.



With the impedances Z and the admittances Y of the above TEE and PI schematics, the conversion equations are as follows:

$$Z13 = \frac{Z10 * Z12}{\sum Z\mu\nu}$$

$$Y12 = \frac{Y13 * Y23}{\sum Y\nu 0}$$

$$Z23 = \frac{Z20 * Z12}{\sum Z\mu\nu}$$

$$Y13 = \frac{Y13 * Y30}{\sum Y\nu 0}$$

$$Z30 = \frac{Z10 * Z20}{\sum Z\mu\nu}$$

$$Y23 = \frac{Y23 * Y30}{\sum Y\nu 0}$$

with
$Z\mu\nu = Z10 + Z12 + Z20$

with
$Y\nu 0 = Y13 + Y23 + Y30$

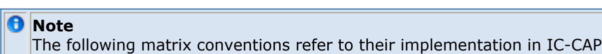
Twoport Matrix Definitions

Terms and Definitions on Series and Parallel Circuits

Series circuit	Parallel circuit
Impedance $Z = \sqrt{R^2 + X^2}$	Admittance $Y = 1/Z = \sqrt{G^2 + B^2}$
Resistance R	Conductance G=1/R
Reactance X	Susceptance B=1/X

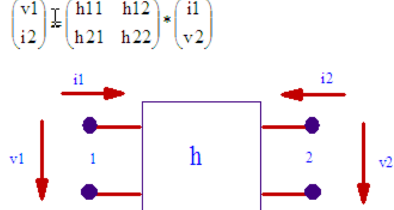
Immittance: A general term for both impedance and admittance, used when the distinction is irrelevant.

Twoport Matrix Signal Flow Definition

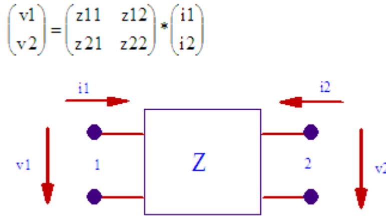


Note
The following matrix conventions refer to their implementation in IC-CAP.

H Matrix



Z Matrix



Y Matrix

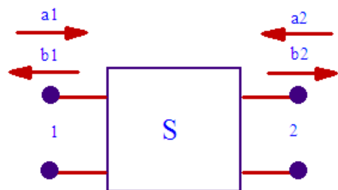
$$\begin{pmatrix} i1 \\ i2 \end{pmatrix} = \begin{pmatrix} y11 & y12 \\ y21 & y22 \end{pmatrix} * \begin{pmatrix} v1 \\ v2 \end{pmatrix}$$

A matrix

$$\begin{pmatrix} v1 \\ i1 \end{pmatrix} = \begin{pmatrix} a11 & a12 \\ a21 & a22 \end{pmatrix} * \begin{pmatrix} v2 \\ i2 \end{pmatrix}$$

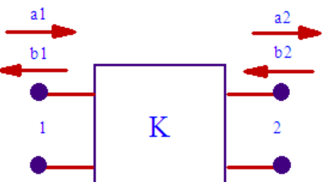
S Matrix

$$\begin{pmatrix} b1 \\ b2 \end{pmatrix} = \begin{pmatrix} s11 & s12 \\ s21 & s22 \end{pmatrix} * \begin{pmatrix} a1 \\ a2 \end{pmatrix}$$



K Matrix

$$\begin{pmatrix} a1 \\ b1 \end{pmatrix} = \begin{pmatrix} k11 & k12 \\ k21 & k22 \end{pmatrix} * \begin{pmatrix} b2 \\ a2 \end{pmatrix}$$



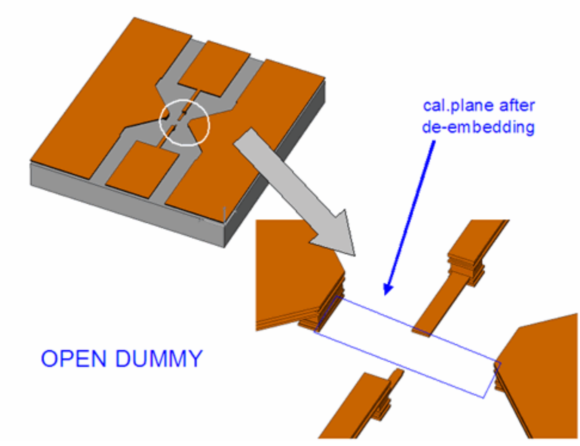
Verifying the De-embedding Procedure

Verifying the de-embedding procedure is very important before applying it to the very DUT, i.e. the transistor or the passive RF component etc.
Without this step, errors or problems associated with the de-embedding will add to the performance of the inner, de-embedded device, and, thus, lead to a wrong device model.
This is especially critical when simply de-embedding (subtracting) the complete Y and Z matrices of the OPEN and SHORT dummy, because, in this case, possible problems may not show up compared to de-embedding of lumped components. In this later case, we would have detected such problems because they would show up with non-physical lumped circuit components of these OPEN and SHORT dummies.
Therefore, it is suggested to

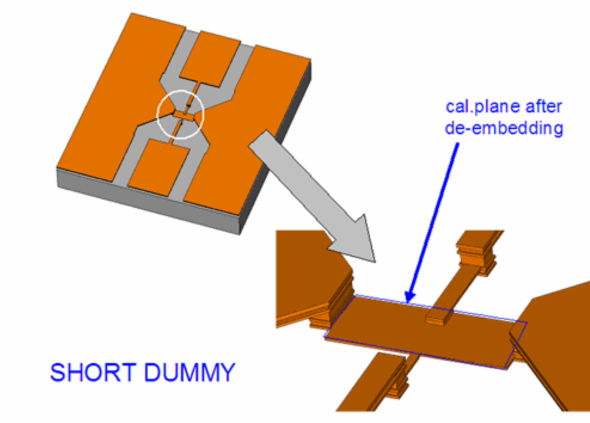
- Model every dummy structure before simply subtracting their total matrices in order to verify its de-embedding prerequisites:
  and to
- Verify the de-embedding procedure with a well-known 'golden device'.
  before applying this procedure to the very DUT.
  The pre requisite for a correct de-embedding is that certain test structures are available on a wafer together with the device under test (DUT) itself. Depending on the selected de-embedding method, an OPEN and SHORT dummy structure is required and must be measured. For the proposed de-embedding verification from above, also a THROUGH dummy structure is necessary.

The principle layouts of these structures are given in the figures shown below. These layouts are for Ground-Signal-Ground Probes (GSG).
Please check the calibration plane as marked in the figures. Every part of the DUT included in this calibration plane will become a part of the DUT model!
In other words, you can think of de-embedding as a shift of the current calibration plane to these new borders on the wafer.

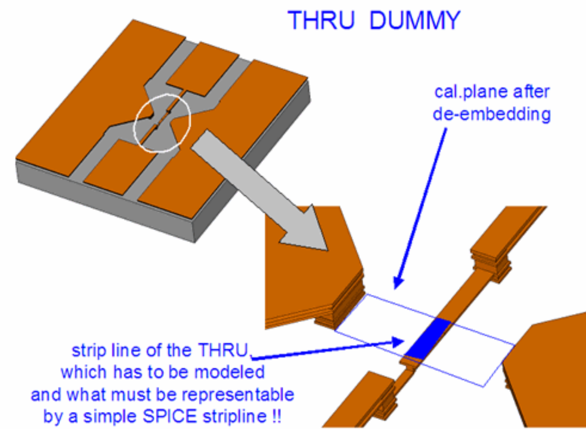
Example of a layout of an OPEN dummy structure.



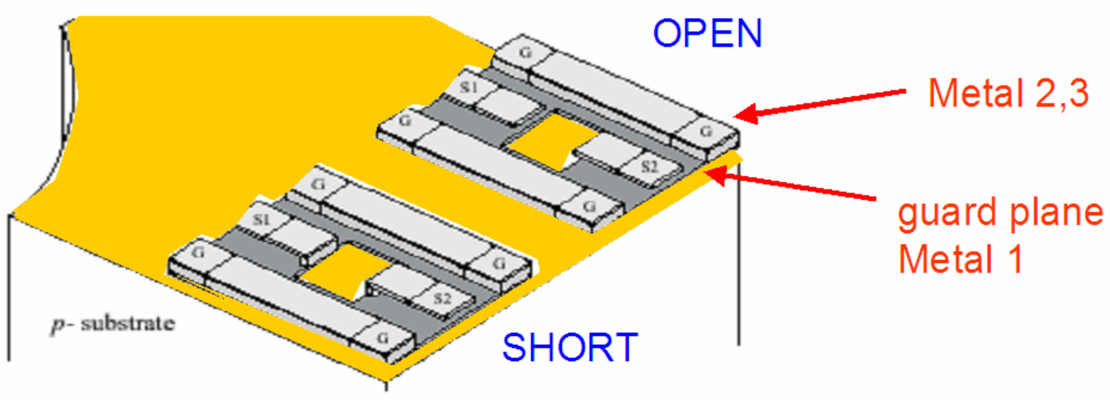
Example of a layout of a SHORT dummy structure



Example of a layout of a THRU dummy structure.



Note
Instead of the 'classical' layout shown above, today's layouts for silicon wafers look more and more often like the one shown in the next figure below. It avoids the effects of the lossy silicon substrate by using the 1st metal plane to shield the contact pads and the lines to the DUT against the lossy silicon substrate. Especially interesting is the drastically improved SHORT dummy performance, since it applies a series of vias to ground (metal 1) at the end of the SHORT, rather than having - as with the SHORT depicted above- a large metal plane which is considered as ideal, while the two microstrip lines of the SHORT dummy are considered as non-ideal. Also, the OPEN is much more ideal with this alternate approach, i.e. its measurement data do not exhibit the second half-circle effect like with the conventional layout from above,i.e., the OPEN can now be modeled up to several 10 GHz with simply the inner schematic of the OPEN circuit schematic as discussed in the next section.



From:
T.E.Kolding, O.K.Jensen, T.Larsen,
Ground-Shielded Measuring Techniques for Accurate On-Wafer Characterization of RF CMOS Devices
IEEE International Conference on Microelectronic Test Structures, March 2000, p.246-251

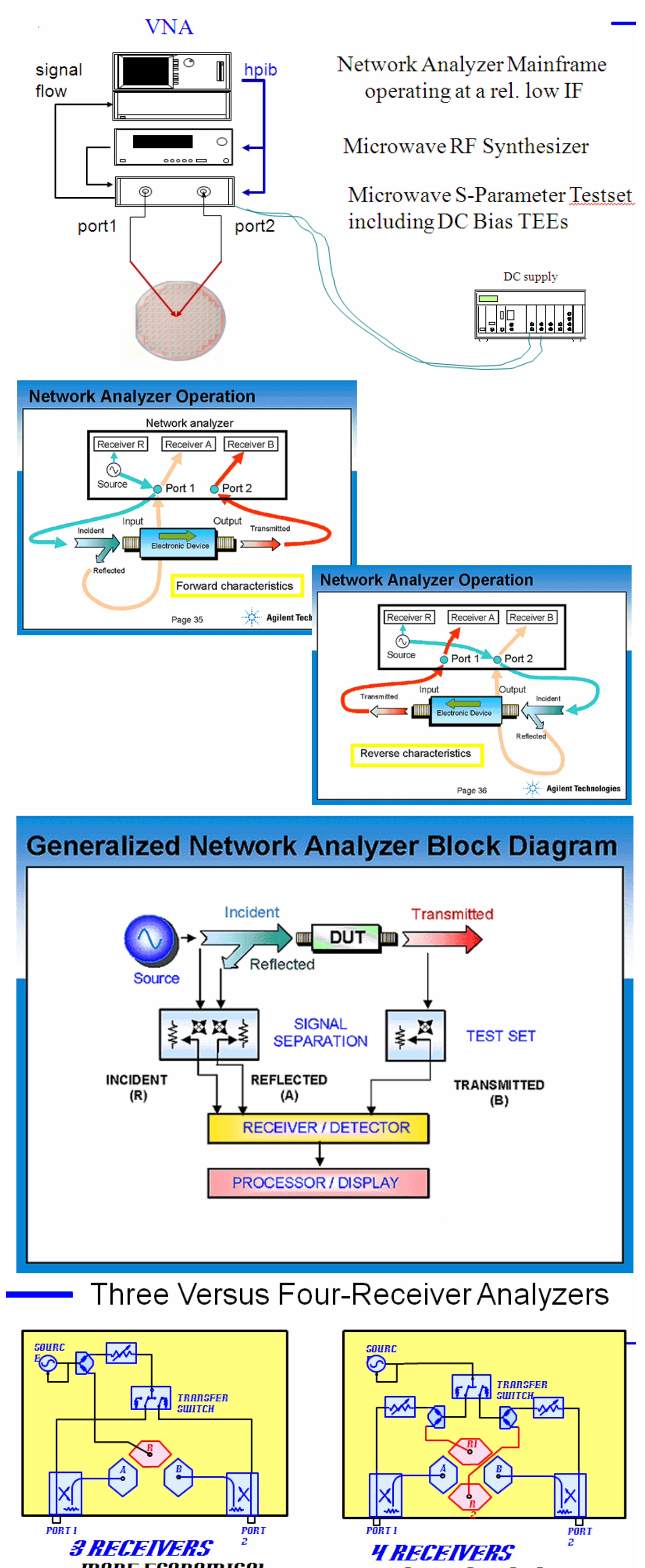
Altogether, this layout suggestion of using the 1st metal as a shield has proofed to give much more robust de-embedding results.

Modeling of the Open Dummy

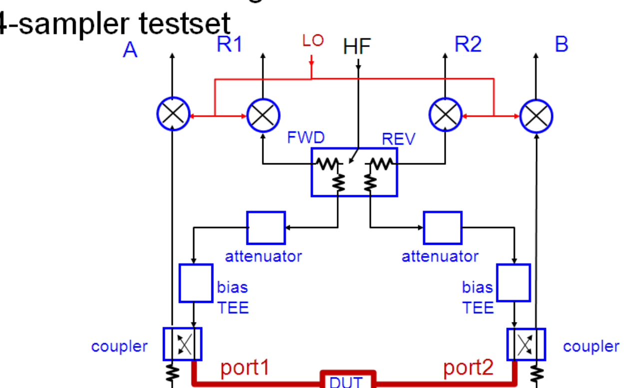
In this first step, the OPEN dummy is modeled in order to verify that there are no serial components included in the Y matrix of the OPEN. See fig.1.

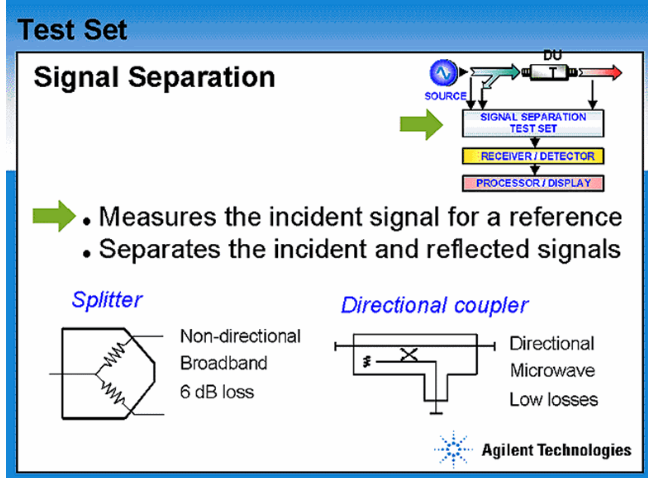
Note
Note: if there were parasitic components in series with the DUT, the subtraction of the Y matrix cannot be applied, and would lead to wrong OPEN-deembedding results.

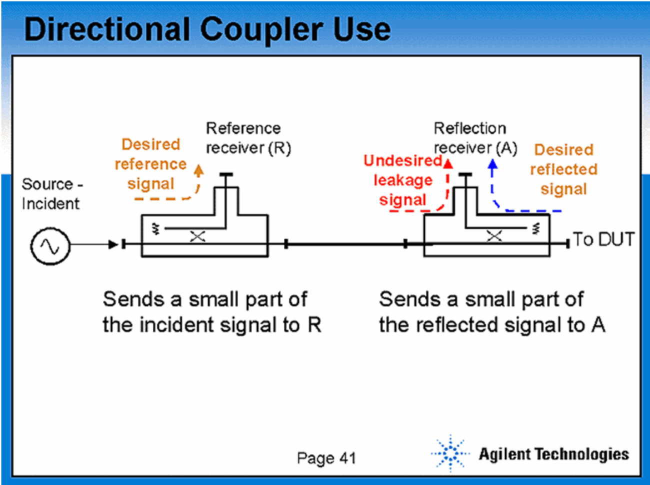
Vector Network Analyzer - Basics for Modeling Engineers

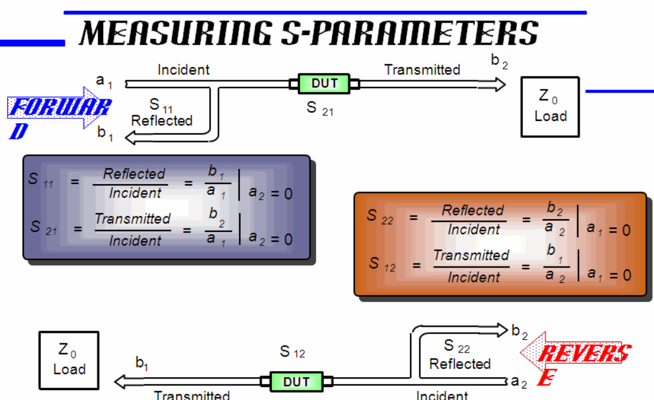


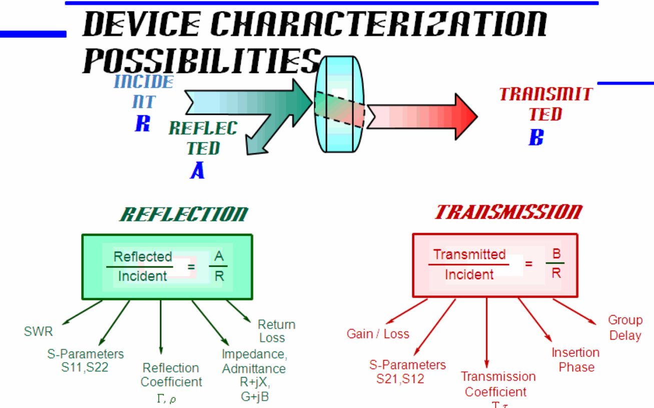
Detailed Block Diagram of the 8510C 4-sampler testset











Measurement Errors

Systematic errors

Due to imperfections in the analyzer and test setup are assumed to be time invariant (predictable)can be characterized (during calibration process) and mathematically removed during measurements.

Random errors

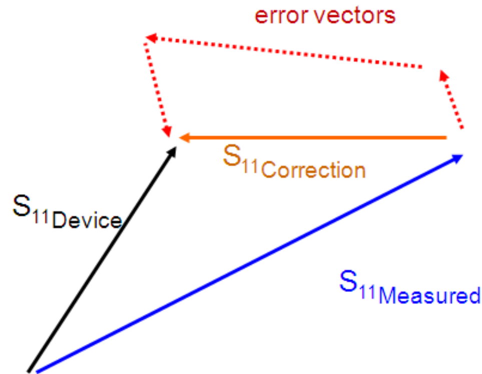
- Vary with time in random fashion (unpredictable)cannot be removed by calibration
- main contributors:
  - instrument noise (source phase noise, IF noise floor, etc.)
  - switch repeatability
  - connector repeatability

Drift errors

These errors are due to instrument or test-system performance changing after a calibration has been done are primarily caused by temperature variation can be removed by further calibration(s).

Vector Error Correction

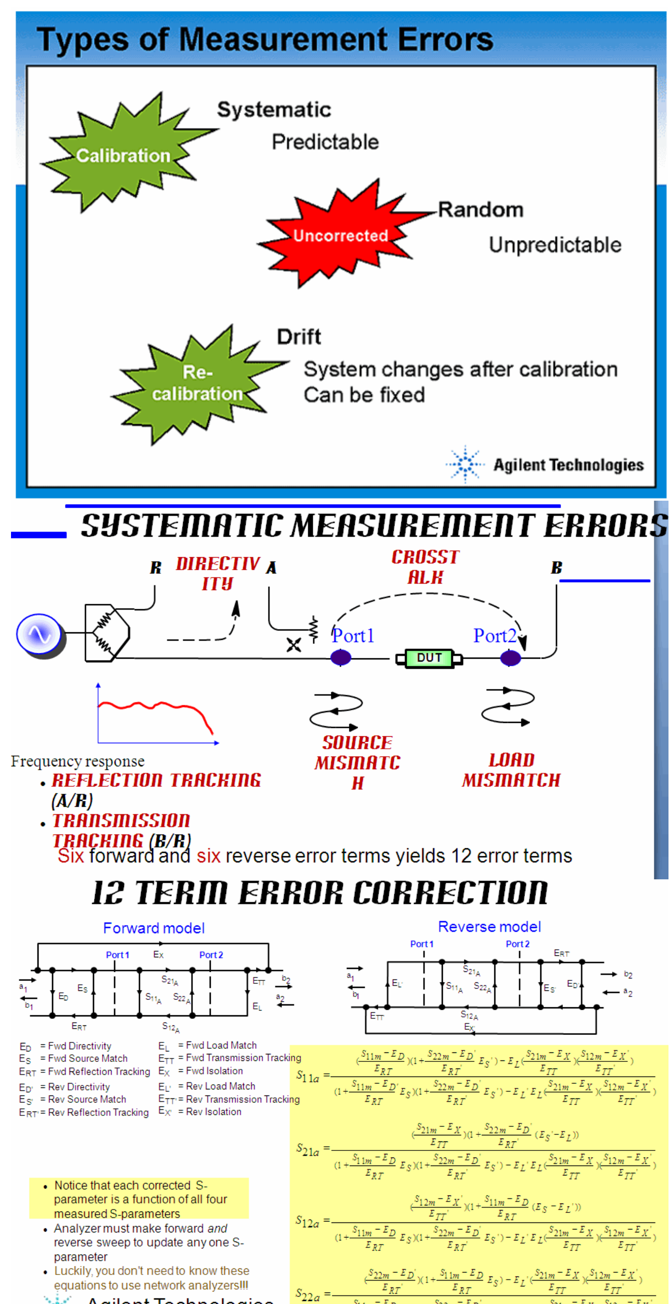
accounts for all major sources of systematic error



What to do for Vector-Error Correction?

- Full 2-port calibration
  (reflection and transmission measurements)
  12 systematic error terms measured usually requires 12 measurements on four known
- Standards (SOLT)
- Standards are defined in cal kit definition table. These cal kit definitions are entered to the network analyzer

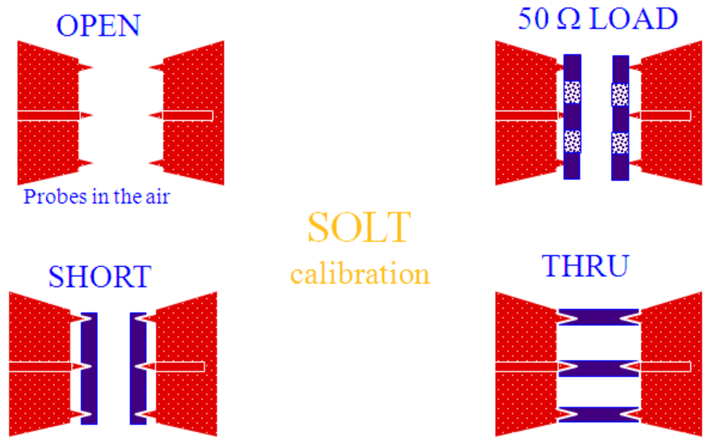
CAL KIT DEFINITION MUST MATCH ACTUAL CAL KIT USED!



Practical Considerations before starting with NWA Measurements

- Enter calkit data (non-idealities of the cal. standards) into the NWA
- Enter your frequency range etc. and Calibrate the NWA and generate the calset for the 12 term error correction (to be used for meas.data correction)
- Verify the calibration
- MEASURE
- if required , de-embed the DUT

SOLT calibration for Ground-Signal-Ground probes (G_S_G)



All standard calkit data must be accurately entered into the VNA. SOLT is sensitive to probe placement (probe tip instance at THRU).

Other Calibration Methods

We know about Short-Open-Load-Thru (SOLT) calibration...

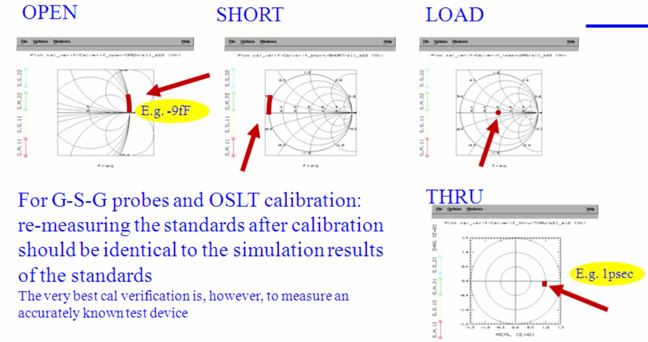
What is TRL, LRM etc ?

- Different two-port calibration techniques
- Use the same 12-term error model as the common SOLT cal
- Use calibration standards that are easier to fabricate and characterize
- Can give better calibration results than SOLT for frequencies >20GHzbut may have limited bandwidth (TRL)
- Two variations:TRL, LRM (require 4 receivers) and TRL*, LRM*(only three receivers needed)

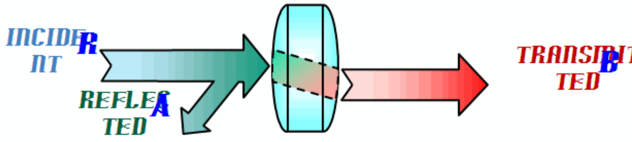
Verify the Calibration

E=m*C² ± 3dB !!?

Verify the Calibration



Summary



Network Analyzer measurements for modeling need

- Good hardware: VNA, probes, calkits
- Good understanding of S-parameters
- Specific chip layout

Nonlinear Network Analyzers

Contents

- *Nonlinear De-Embedding* (iccapmhb)
- *Nonlinear RF Measurements and Device Modeling* (iccapmhb)

Nonlinear De-Embedding

Measurements of RF power devices can best be performed using a non-linear vector network analyzer (NVNA), also called a large signal network analyzer (LSNA). Such instruments offer the unique capability to measure, at the input and output of any two-port device under test, both magnitude and phase of voltages and currents in the frequency domain:

- At the fundamental frequency
- and a specified number of harmonics.

From these results, the corresponding waveforms in the time domain can easily be derived. This large-signal RF information represents the entire device RF behaviour for the given stimulus.

While for packaged devices, when using a test fixture with in-fixture calibration standards, the obtained large signal measurements represent - as desired- exactly the packaged device with its inner circuit plus the package parasitics, on-wafer measurements have to be de-embedded from the probe contact pads and the strip lines to the inner DUT. Generally, de-embedding is associated with S-parameter measurements. Many such de-embedding procedures are published in the literature. However, only little is found on how to de-embed large-signal (nonlinear) vectorial RF (radio frequency) measurements.

And, in fact, the de-embedding of non-linear RF measurements is substantially different from S-parameter de-embedding, as can be seen in the following figure:

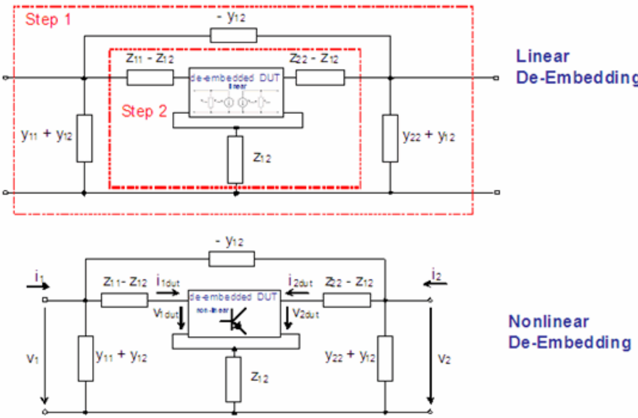


Figure: The different (Open-Short) de-embedding methods:

- De-embedding procedure for linear S-parameter measurements: TwoPort matrix manipulations
- De-embedding procedure for large-signal non-linear RF measurements *Calculation of device voltages and currents out of measured voltages and currents*

> case a) the de-embedded DUT is considered to behave linear (bias-dependent, but linearized in the operating point)
> case b), the de-embedded DUT is still non-linear (may be large-signal-driven in its operating points).

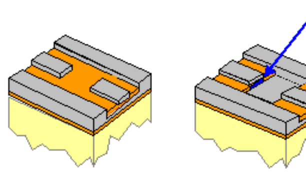
2-Step Nonlinear De-Embedding

In order to do device modeling for on-wafer measurements with the nonlinear network analyzer, the measurement has to be de-embedded

- from the on-wafer contact pads
- and the strip lines from these pads to the device itself.
  Different to the de-embedding process for S-Parameters, we cannot simply apply the conventional Y-parameter or Z-parameter subtractions. The reason is that S-parameters are linear, and that all the twoport matrix theory is also based on linear circuits. For non-linear devices like transistors or diodes, this means a small-signal excitation around a DC operating point, and no harmonic frequencies.

When applying a nonlinear network analyzer, these prerequisites of the mentioned linear twoport theory are not fulfilled. We need to handle the complex (magnitude and phase) voltages and currents. However, since the OPEN and SHORT dummies represent linear circuits, we can use their Y- and Z-matrix elements in the de-embedding calculations. Note that these matrix elements are represented by complex numbers, versus frequency. And these frequency points must match the fundamental frequency and its harmonics from the nonlinear network analyzer measurement.
Also note, that we can apply them 'as they are' and do not need to care about their underlying inner schematic in the de-embedding procedure presented below.

The proposed 2-step de embedding procedure of the nonlinear measurements is based on the conventional OPEN and SHORT dummy structures.



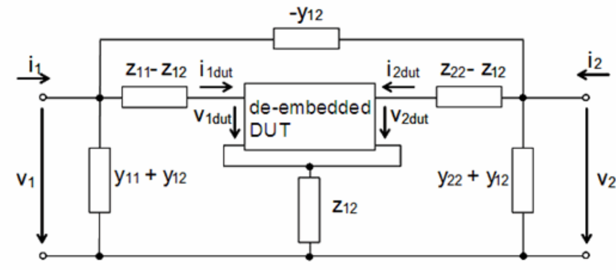


Figure: OPEN-SHORT de-embedding for large signal RF measurements.

The de-embedded large signal currents and voltages (i1dut, i2dut, v1dut, v2dut) are calculated from the measured ones (*i1*, *i2*, *v1*, *v2*) after:

$$i_{1\,dut} = i_1 - v_1(y_{11} + y_{12}) - (v_2 - v_1)y_{12} \quad (1)$$

$$i_{2\,dut} = i_2 - v_2(y_{22} + y_{12}) - (v_1 - v_2)y_{12} \quad (2)$$

$$v_{1\,dut} = v_1 - i_{1\,dut} \cdot z_{11} - i_{2\,dut} \cdot z_{12} \quad (3)$$

$$v_{2\,dut} = v_2 - i_{2\,dut} \cdot z_{22} - i_{1\,dut} \cdot z_{12} \quad (4)$$

and re-arranged:

$$i_{1dut} = i_1 - v_1 \cdot y_{11} - v_2 \cdot y_{12}$$

$$i_{2\,dut} = i_2 - v_2 \cdot y_{22} - v_1 \cdot y_{12}$$

where y_{ij} are the admittance parameters of the OPEN dummy Y-matrix, and z_{ij} are the impedance parameters of the SHORT dummy Z-matrix after the de-embedding of the OPEN dummy effects.

Note
Both, OPEN and SHORT structures, are reciprocal, i.e. y12 = y21 and z12 = z21. Therefore, we can apply the mean values.

and
y12 := (y12+y21)/2
z12 := (z12+z21)/2
to reduce measurement noise.

IC-CAP demo file
demo_features\3_MEAS_ORGANIZE_n_VERIFY_DATA\0_MASTER_FILES\30_DEEMBEDDING\5_nonlinear_deemb_OPEN_SHORT.mdl
Special acknowledgements regarding the de-embedding procedure: Giovanni Crupi, Dipartimento di Fisica della Materia e Ingegneria Elettronica, University of Messina, 98166 Messina, Italy.
Giovanni Crupi, Dominique Schreurs, Electronic Engineering Department, Katholieke Universiteit Leuven, 3001 Leuven, Belgium
Publications
G.Crupi, D.Schreurs, D.Xiao, A.Caddemi, B.Parvais, A.Mercha, S.Decoutere: "Determination and Validation of New Nonlinear FinFET Model Based on Lookup Tables", IEEE MicrowaveAnd Wireless Components Letters, Vol. 17, No. 5, May 2007.

3-Step Nonlinear De-Embedding

Note
This chapter is based on section D of the publication, E. Vandamme, D. Schreurs and C. van Dinther, "Improved three-step de-embedding method to accurately account for the influence of pad parasitics in silicon RF test-structures," IEEE Trans. Electron Devices, vol.48, no.4, pp.737-742, 2001.

The 3-step de-embedding is based on four Dummy structures which are all used in the de-embedding process. The quality of the obtained de-embedding relies on the fact if the individual dummies can be represented by the underlying, assumed schematics. The dummies and their assumed schematics are shown in the following figure.

Note
For silicon, the layout using metal1 as a ground shield against the lossy Silicon substrate is preferred over the conventional layout using a single metal layer. The short circuit of the Short Dummy is achieved by a chain of vias from the upper metal to metal1 (ground).

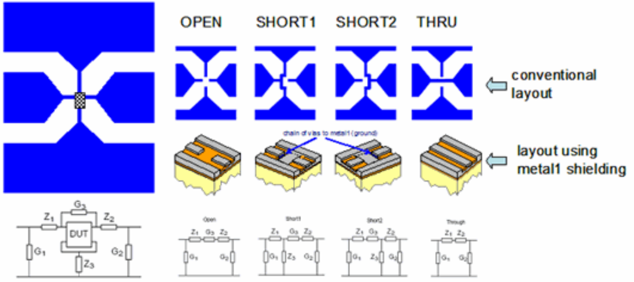


Figure: Layout of the DUT and details of the de-embedding dummy layouts

The non-linear 3-step de-embedding procedure in details:

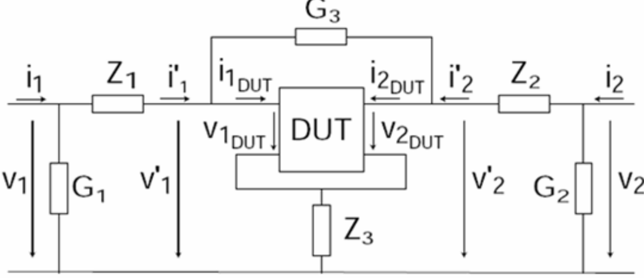


Figure: 3-step de-embedding procedure for large-signal non-linear RF measurements.

The large-signal currents and voltages in the de-embedding scenario shown above in fig.4, are defined in the frequency domain, with the fundamental frequency and a certain number of harmonics. The same frequency stimuli are applied to the S-parameter measurements of the parasitic elements as well. In order to model the large-signal RF behaviour at the level of the d $_{ut}$, currents i_{1dut} and i_{2dut} , and voltages v_{1dut} and v_{2dut} are calculated after:

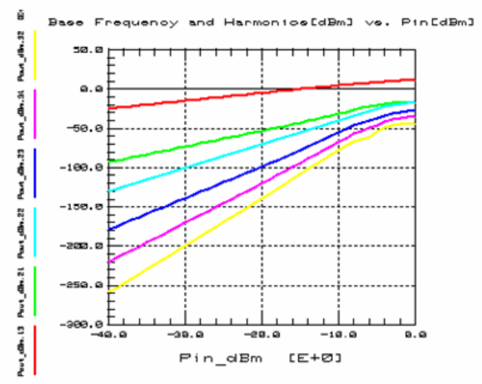
$$i_{1dut} = i_1' - (v_{1dut} - v_{2dut}) \cdot G3$$
$$i_{2dut} = i_2' - (v_{2dut} - v_{1dut}) \cdot G3$$
$$v_{1dut} = v_1' - (i_1' + i_2') \cdot Z3$$
$$v_{2dut} = v_2' - (i_1' + i_2') \cdot Z3$$

where,

$$i'_1 = i_1 - v_1 G_1$$
$$i'_2 = i_2 - v_2 G_2$$
$$v'_1 = v_1 - i'_1 Z_1$$
$$v'_2 = v_2 - i'_2 Z_2$$

and the voltages v_1, v_2 and the currents i_1, i_2 are measured by the nonlinear network analyzer.

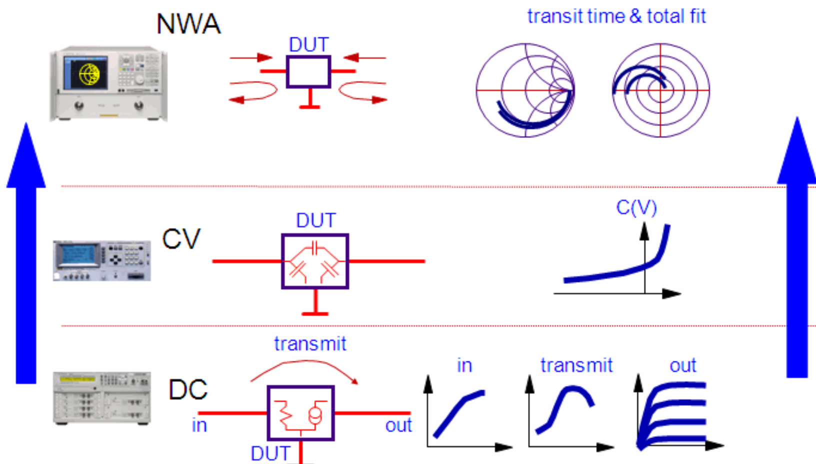
Nonlinear RF Measurements and Device Modeling



Topics

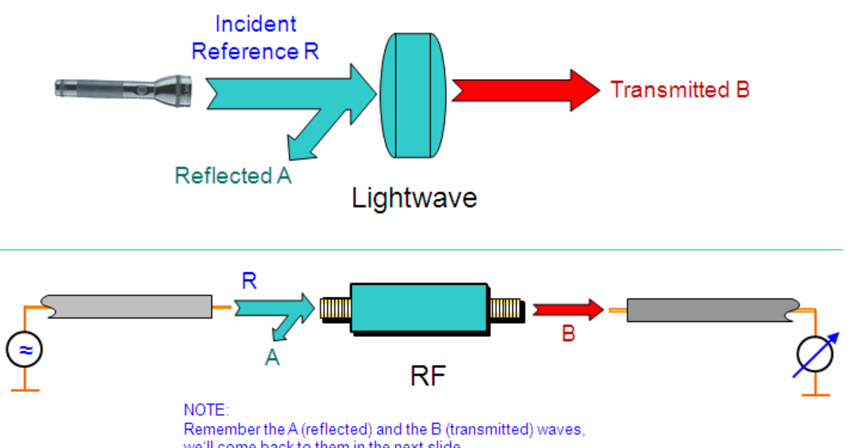
- Introducing Nonlinear RF
  - Linear S-Parameter Measurements
  - Nonlinear VNA Measurements
  - Harmonic Balance (HB) Simulations
  - X-Parameters
  - PHD Model
- Device Modeling Aspects for Nonlinear VNA On-Wafer Measurements

From DC > CV > NWA



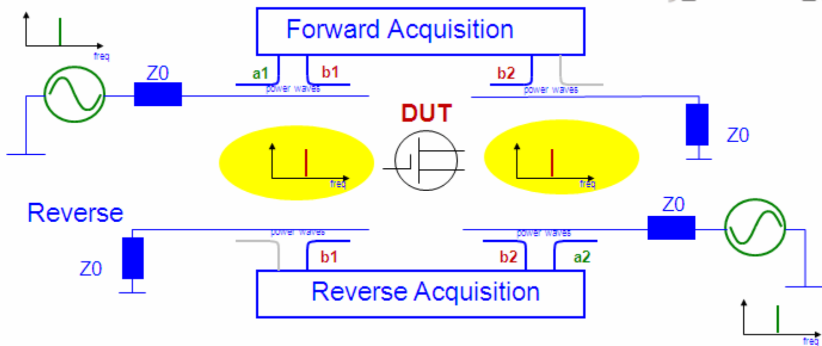
This slide depicts the standard device modeling approached based on DC -> CV -> S-Parameters:The DC modeling is non-linear, while the CV modeling and the S-Parameter modeling are linear in the (many) operating points, i.e. tangents to the measured curves at the operating points.In other words, the modeled 'tangents' to the bias points are a substitute for large-signal non-linear RF excitation (trajectory curves).

Lightwave Analogy to RF Energy



One of the fundamental concepts of RF power transmission involves incident, reflected and transmitted waves traveling along transmission lines. It is helpful to think of traveling waves along a transmission line in terms of a light wave analogy. We can imagine incident light striking some optical component like a clear lens. Some of the light is reflected off the surface of the lens, but most of the light continues on through the lens. If the lens was made of some glossy material, then some of the light could be absorbed within the lens. If the lens had mirrored surfaces, then most of the light would be reflected and little or none would be transmitted.
This concept is valid for RF signals as well, except the electromagnetic energy is in the RF range instead of the optical range, and our components and circuits are electrical devices and networks instead of lenses and mirrors.

Linear Vector Network Analyzer

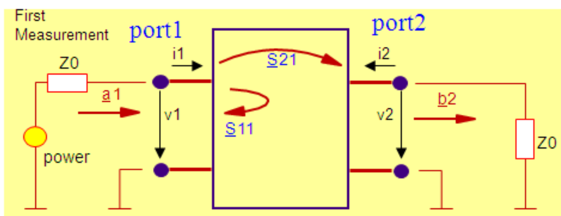


The VNA measures the A and B waves, and then calculates the S-Parameters out of them.
The S-Parameters are a model of the DUT -at its operating point(s)- in a simulator.

In fact a vector network analyzer (VNA) is a complete measurement system > a signal generator, providing sequentially the a1 (forward) and the a2 (reverse) waves:

- a narrow-band S-Parameter testset, measuring (filtering out) the bi waves in forward and reverse,
- and a model extraction capability, calculating the S-Parameters, which can be considered as a model of the DUT.
  In other words, the VNA experiment consists of a forward and reverse excitation and measurement.The model extraction refers to the equation-solving using the measurements, resulting into S-Parameters.

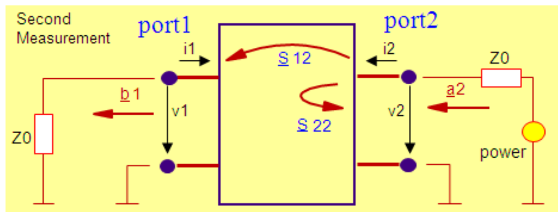
Considering a 2-port: stimulated at port1, then at port 2



The transmitted and reflected PowerWaves are defined as:

$$a1 = \frac{1}{2\sqrt{Z0}}(v1 + i1 \cdot Z0) \quad *)$$

$$b2 = \frac{1}{2\sqrt{Z0}}(v2 - i2 \cdot Z0)$$



$$a2 = \frac{1}{2\sqrt{Z0}}(v2 + i2 \cdot Z0)$$

$$b1 = \frac{1}{2\sqrt{Z0}}(v1 - i1 \cdot Z0) \quad *)$$

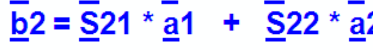
NOTE : a, and b, are in $\sqrt{Watt}$

*) see S.Maas, "The RF and Microwave Circuit Design Handbook", Artech House, Boston, 1998, ISBN 0-89006-973-5, p.37ff.

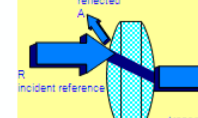
Note
During S-Parameter measurements, the power sources provide the RF signal to the twoport. The voltages v1, i1, v2 and i2 are therefore a consequence of that stimulating power source. On the other hand, the above plot explains how they can be calculated out of the measured or simulated S-Parameters (vi and ii as a function of the ai and bi). In the formula above, voltages vx and the currents ix are 'rms' values, i.e. peak/SQRT(2). The obtained waves ax and bx are 'rms' too. Their unit is SQRT(Watt).
For the final set of S-Parameters, an overlay of forward and reverse measurement is applied. This is possible for the 'linear' S-Parameters. But this is **not** possible for the nonlinear case, the nonlinear network analyzer (NVNA).

The S-Parameters relate the **b** PowerWaves to the **a** PowerWaves

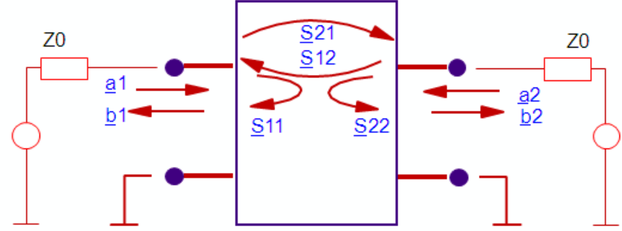
$$b1 = S11 * a1 + S12 * a2$$



$$b2 = S21 * a1 + S22 * a2$$



with ai: PowerWave towards the twoport
bi: PowerWave out of the twoport



Note:
the forward and reverse measurements are *performed sequentially*. They are then *overlaid* to give the four S-Parameters

$$S_{11} = \frac{b1}{a1} = \frac{power_reflected_at_port1}{power_towards_port1}\Bigg|_{a2=0}$$

$$S_{12} = \frac{b1}{a2} = \frac{power_out_of_port1}{power_towards_port2}\Bigg|_{a1=0}$$

$$S_{21} = \frac{b2}{a1} = \frac{power_out_of_port2}{power_towards_port1}\Bigg|_{a2=0}$$

$$S_{22} = \frac{b2}{a2} = \frac{power_reflected_at_port2}{power_towards_port2}\Bigg|_{a1=0}$$

The conditions a1=0 and a2=0 are fulfilled with an impedance of Z0 (typ. 50 Ohm) in series with the stimulating voltage source at the input, and with an output termination of again Z0.

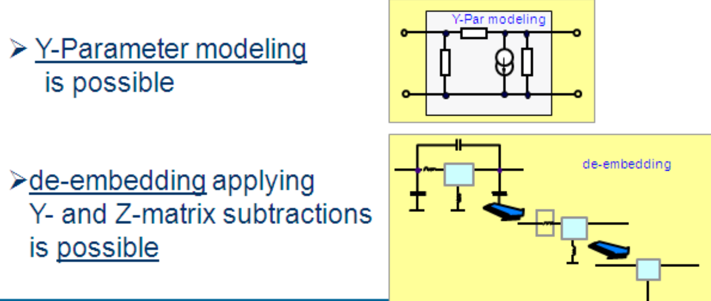
S-Parameters = linear theory

$S11 = \frac{b1}{a1} \Big/ (a2 = 0)$ Input Impedance $S21 = \frac{b2}{a1} \Big/ (a2 = 0)$ Forward Transmission

$S22 = \frac{b2}{a2} \Big/ (a1 = 0)$ Output Impedance $S12 = \frac{b1}{a2} \Big/ (a1 = 0)$ Reverse Transmission

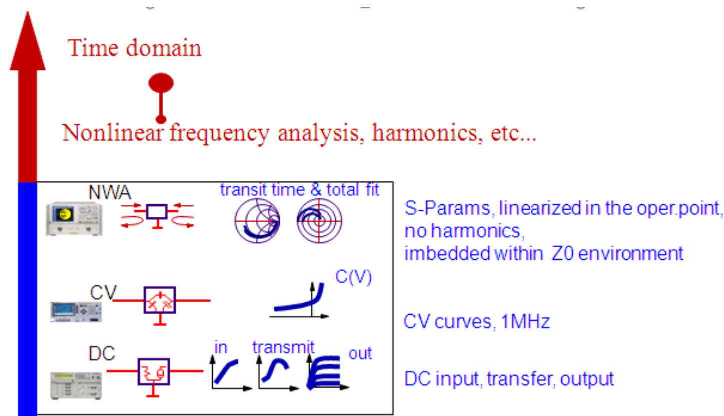
S-Parameters are based on linear circuit theory. Therefore, the forward (a2=0) and reverse (a1=0) measurement of a network analyzer, performed sequentially, can be overlaid to give the 4 S-Parameters.
Also, in linear circuit theory, matrix conversion (from e.g. S-Par to Y-Par, to do PI-schematic-based Y-Parameter modeling, is possible.
Last not least, these matrix conversions also allow the easy de-embedding methods which subtract Y-matrices and Z-matrices.
As will be seen later, this all is **not** possible for nonlinear behavior, and therefore, **other methods** need to be applied there.

S2Y, S2Z etc. matrix conversions are possible



S-Parameters are based on linear circuit theory. Therefore, the forward (a2=0) and reverse (a1=0) measurement of a network analyzer, performed sequentially, can be overlaid to give the 4 S-Parameters.
Also, in linear circuit theory, matrix conversion (from e.g. S-Par to Y-Par, to do PI-schematic-based Y-Parameter modeling, is possible.
Last not least, these matrix conversions also allow the easy de-embedding methods which subtract Y-matrices and Z-matrices.
As will be seen later, this all is **not** possible for nonlinear behavior, and therefore, **other methods** need to be applied there.

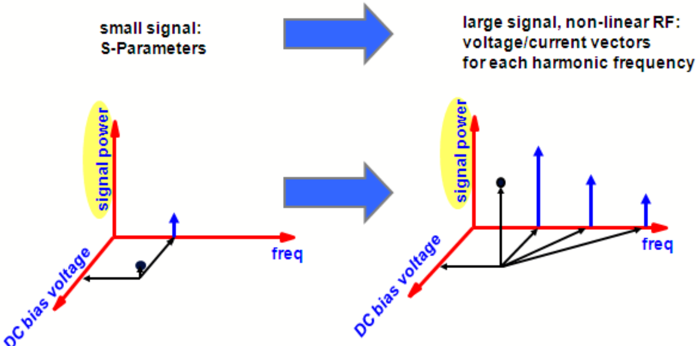
Beyond S-Parameters



Beyond the S-Parameters, we have nonlinear device excitation: instead of a tangent to the operating point, the large-signal excitation leads to a revolving 'satellite' curve around the operating point: the trajectory curve.
This means there are 2 ways of looking at the signals:

- Either in the time domain, or
- Regarding the spectrum (DC, fundamental frequency and harmonics) in magnitude and phase.

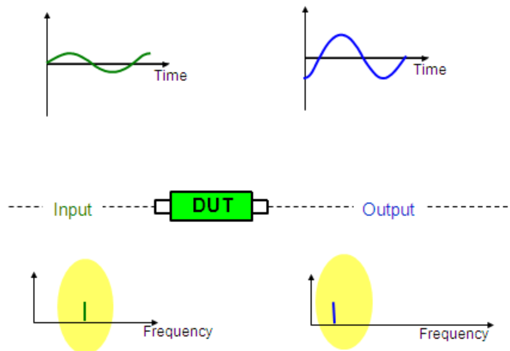
Extending linear to RF-power-dependent signals



Linear versus nonlinear behavior

Linear Behavior

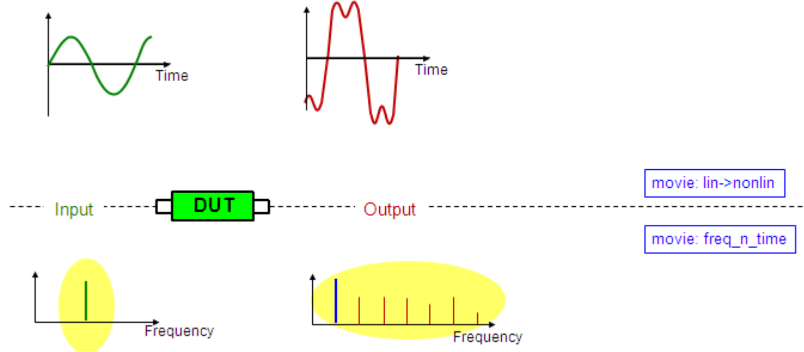
- Input and output frequencies are the same (no additional frequencies created)
- Output frequency only undergoes magnitude and phase change



In order to understand nonlinear measurements, let's review the differences between linear and nonlinear behavior. Devices that behave linearly only impose magnitude and phase changes on input signals. Any sinusoid appearing at the input will also appear as a sinusoid at the output at the same frequency, of course with different magnitude and phase. No new signals are created.

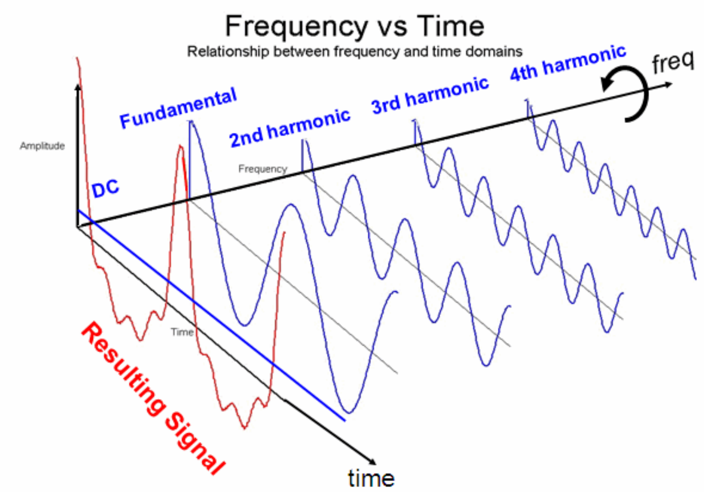
Nonlinear behavior:

- Additional frequencies created (harmonics, inter-modulation)



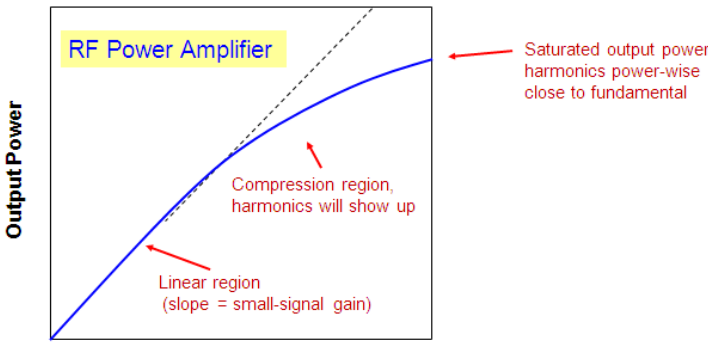
Non-linear devices convert single-frequency stimuli into output signals which exhibit harmonics or intermodulation products. -> New signals are created.

> **Note**
> Many components that behave linearly under most signal conditions can exhibit nonlinear behavior if driven with a large enough input signal.



A typical power-RF example:

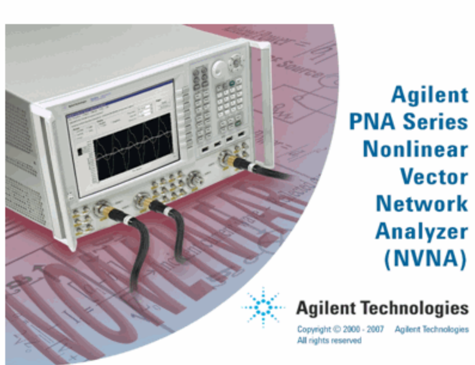
Power Sweep - Compression



Many network analyzers have the ability to perform power sweeps as well as frequency sweeps. Power sweeps help characterize the nonlinear performance of an amplifier. Shown above is a plot of an amplifier's output power versus input power at a single frequency. Amplifier gain at any particular power level is the slope of this curve. Notice that the amplifier has a linear region of operation where gain is constant and is independent of power level. The gain in this region is commonly referred to as small-signal gain. At some point as the input power is increased, the amplifier gain appears to decrease, and the amplifier is said to be in compression. Under this nonlinear condition, the amplifier output signal is no longer sinusoidal – some of the output power is present in harmonics, rather than occurring only at the fundamental frequency.

NVNA Measurements

Component Characterization



The Nonlinear VNA is a new set of Firmware that runs on a standard PNA-X.

The PNA-X can run the normal PNA-X Firmware (S-Parameters) and, also, the NVNA Firmware.

Brief Explanation: Nonlinear device characterization and X-Parameters

The NVNA features two methods for measuring the nonlinear effects of the device under test:

- Nonlinear device characterization (measuring the Ai and Bi waves)
- X-Parameter measurement

Nonlinear device characterization measurements provide calibrated, vector corrected wave forms of the device under test. Measurements such as amplitude and phase of harmonics, group delay, phase between the harmonics and vector corrected time domain waveforms can be made. Nonlinear component characterization measures the incident, transmitted and reflected waves of the DUT. This data can then be displayed in frequency, time or power domains as well as user defined ratios such as I/V to display dynamic loadlines. Additionally, this information can be imported into Agilent's ADS simulator or into IC-CAP for applying conventional models (BSIM, PSP, Hicum, VBIC, Angelov, Statz etc.) to nonlinear RF measurement data. For an amplifier, the designers can measure the full spectral output of the device in both amplitude and phase.

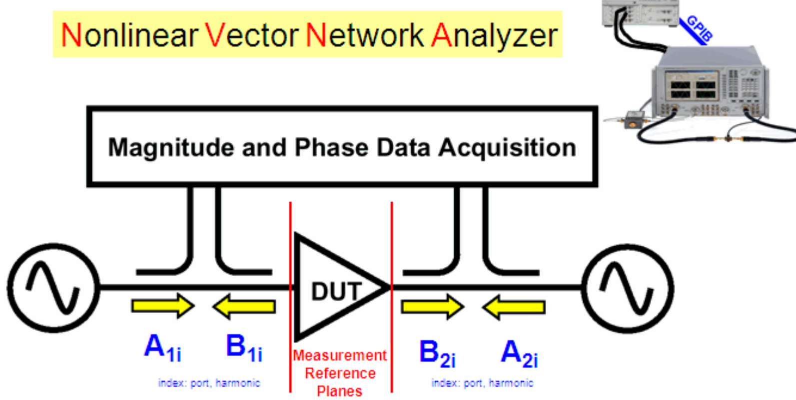
X-Parameters, on the other hand, are the logical, mathematically correct extension of S-Parameters into a nonlinear operating environment. X-Parameters measurements require an additional source which is used to drive the DUT with both a large signal tone and a small signal tone at the same time. With the device information derived from such measurements, the coefficients of the X-Parameter model are calculated. The X-Parameters provide information such as the device gain, match, ... while the device is operating in a linear or nonlinear state. The X-Parameters can then be used like S-Parameters, but now for the non-linear, large-signal case. Since the X-Parameters have power and harmonic dependencies, there are many more X-Parameters than S-Parameters. Possibly one of the strongest benefits of X-Parameters is the ability to accurately cascade the X-Parameters from individual devices using the PHD (Poly-Harmonic-Distortion) model in ADS.

Detailed Explanation

X-Parameters are to nonlinear linear measurements what S-Parameters are to linear measurements. S-Parameters were developed as a method to analyze and model the linear behavior of RF components. They also play a key role in analyzing, modeling, and designing more complex systems which cascade multiple individual components. They are related to familiar measurements: S11 input match, S22 output match, S21 gain/loss, S12 isolation, and can be easily imported into electronic simulation tools such as ADS. While S-Parameters are extremely useful and powerful, they have limitations and are defined only for small signal, linear systems.

The usefulness of X-Parameter information can be seen in the process of designing a power amplifier. The designer is compelled to drive the amplifier into the nonlinear region to get the maximum output power as well as to extract the maximum efficiency. Some form of a feedback circuit is then used to compensate for the nonlinear effects and make the output behave like a high power linear device. A typical approach is to suppress the harmonic outputs of a power amplifier using filters and other components. If the input match of the filtering components does not match the output match of the specific harmonic of interest generated by the amplifier, then the attenuation level of that specific harmonic could be significantly different than anticipated. This leads to a time consuming experience of trail and error at best. With accurate phase and amplitude information from the X-Parameters and using simulation tools such as ADS, designers can design the most robust systems possible in the shortest amount of time, with the highest degree of accuracy.

S-Parameters measure half of the parameters on the forward sweep and half on the reverse sweep. Nonlinear characterization measurements gather all the reflected and transmitted waves at once, and include the measurement of the un-ratioed, power calibrated receiver measurements and the poly harmonic phases. For X-Parameter measurements, both a large signal and small signal at the same frequency is presented to the DUT at the same time. Knowing the phase of each signal and changing this relationship enables the extraction of the measured coefficients.



NVNA Calibration

The NVNA requires, additional to the conventional Vector Calibration known from linear network analyzers, an Ampitude Calibration and a Phase Calibration, using a phase standard.

Phase Reference

Superior in all aspects to existing phase reference technologies.

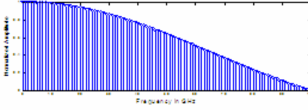
Low temperature sensitivity

Low sensitivity to input power

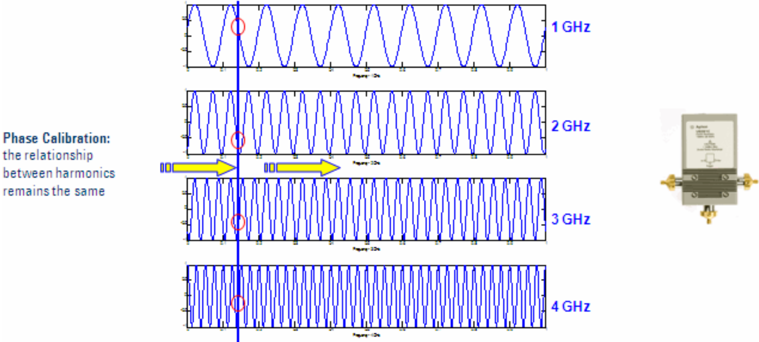
Small minimum tone spacing (< 10 MHz vs 600 MHz)

Low frequency (< 10 MHz vs 600 MHz)

Wide dynamic range due to available energy vs. noise



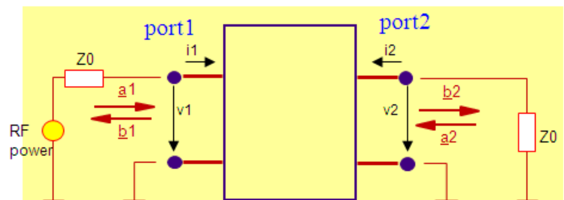
Phase: The critical element



Phase calibration is the important point of NVNA calibration. Due to the down-converision of the signals into the IF domain, the phase information of the harmonics is lost. Phase calibration preserves this information.

NVNA measurement

Similar to a linear VNA,the transmitted and reflected NVNA PowerWaves are defined as:



$$a1 = \frac{1}{2\sqrt{Z0}}(v1 + i1 \cdot Z0)$$

$$b2 = \frac{1}{2\sqrt{Z0}}(v2 - i2 \cdot Z0) \quad {}^{*)}$$

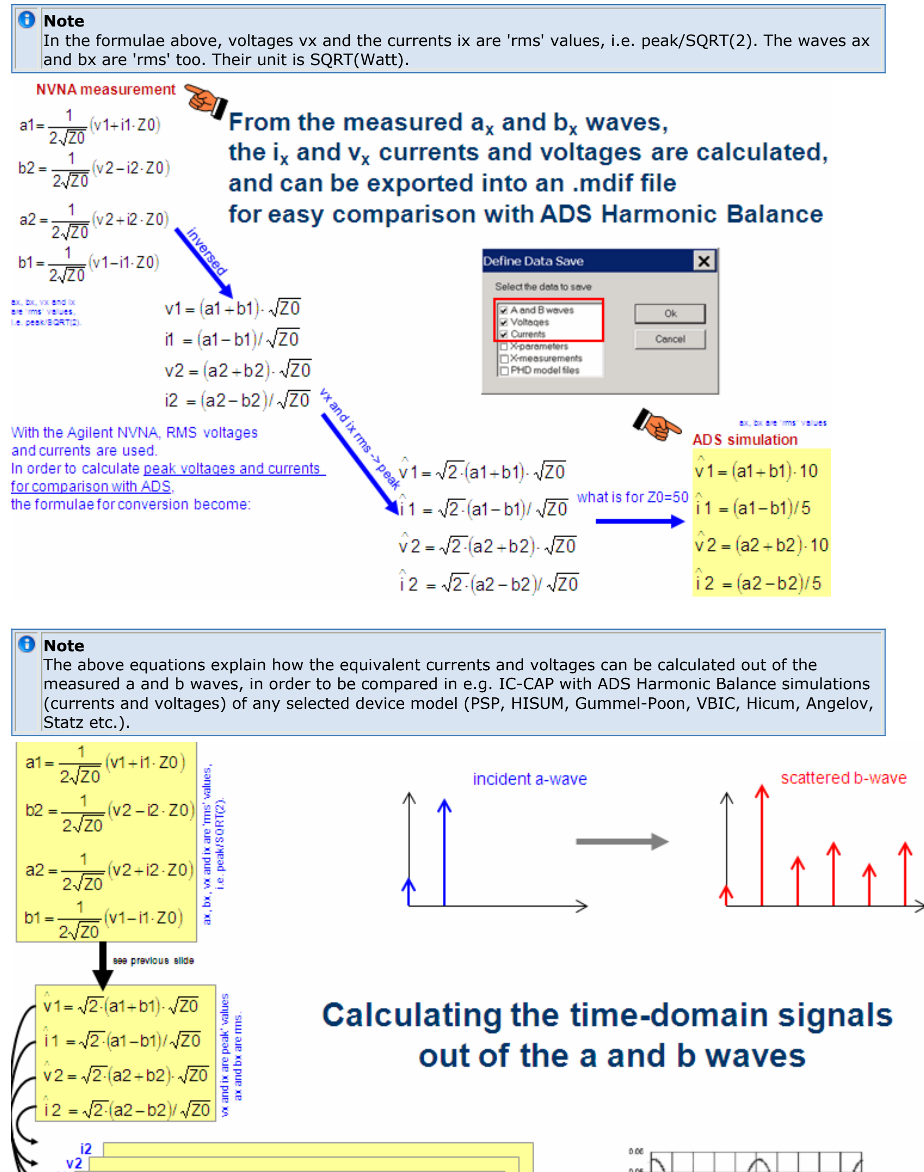
$$a2 = \frac{1}{2\sqrt{Z0}}(v2 + i2 \cdot Z0)$$

$$b1 = \frac{1}{2\sqrt{Z0}}(v1 - i1 \cdot Z0) \quad {}^{*)}$$

NOTE : a, and b, are in $\sqrt{Watt}$

*) see S.Maas, "The RF and Microwave Circuit Design Handbook", Artech House, Boston, 1998, p.37ff, ISBN 0-89006-973-5.

> **Note**
> In the formulae above, voltages vx and the currents ix are 'rms' values, i.e. peak/SQRT(2). The waves ax and bx are 'rms' too. Their unit is SQRT(Watt).



> **Note**
> The above equations explain how the equivalent currents and voltages can be calculated out of the measured a and b waves, in order to be compared in e.g. IC-CAP with ADS Harmonic Balance simulations (currents and voltages) of any selected device model (PSP, HISUM, Gummel-Poon, VBIC, Hicum, Angelov, Statz etc.).

Calculating the time-domain signals out of the a and b waves

$$a1 = \frac{1}{2\sqrt{Z0}}(v1 + i1 \cdot Z0)$$

$$b2 = \frac{1}{2\sqrt{Z0}}(v2 - i2 \cdot Z0)$$

$$a2 = \frac{1}{2\sqrt{Z0}}(v2 + i2 \cdot Z0)$$

$$b1 = \frac{1}{2\sqrt{Z0}}(v1 - i1 \cdot Z0)$$

ax, bx, vx and ix are 'rms' values, i.e. peak/SQRT(2).

see previous slide

$$\hat{v}1 = \sqrt{2} \cdot (a1 + b1) \cdot \sqrt{Z0}$$

$$\hat{i}1 = \sqrt{2} \cdot (a1 - b1) / \sqrt{Z0}$$

$$\hat{v}2 = \sqrt{2} \cdot (a2 + b2) \cdot \sqrt{Z0}$$

$$\hat{i}2 = \sqrt{2} \cdot (a2 - b2) / \sqrt{Z0}$$

vx and ix are peak values, ax and bx are rms.

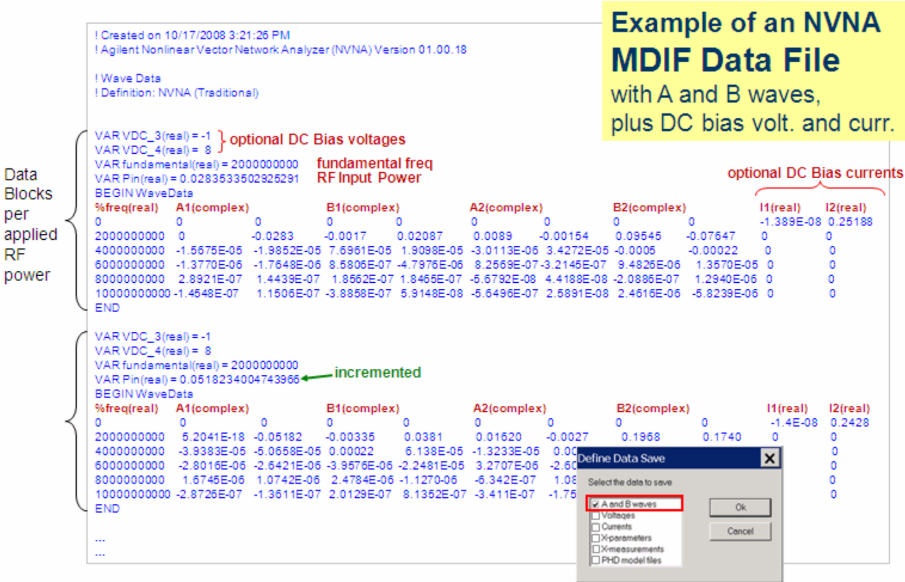
incident a-wave → scattered b-wave

i2, v2, i1, v1:

$$v1(t) = REAL\left(\hat{v}_1[f_0]\right) \cdot \cos(\omega_0 t) - j \cdot IMAG\left(\hat{v}_1[f_0]\right) \cdot \sin(\omega_0 t)$$
$$+ REAL\left(\hat{v}_1[2f_0]\right) \cdot \cos(2\omega_0 t) - j \cdot IMAG\left(\hat{v}_1[2f_0]\right) \cdot \sin(2\omega_0 t)$$
$$+ REAL\left(\hat{v}_1[3f_0]\right) \cdot \cos(3\omega_0 t) - j \cdot IMAG\left(\hat{v}_1[3f_0]\right) \cdot \sin(3\omega_0 t)$$
$$\vdots$$

→ time

The calculated peak voltages and currents at both ports are for DC, fundamental and harmonics. This allows to compose the time-domain signal.

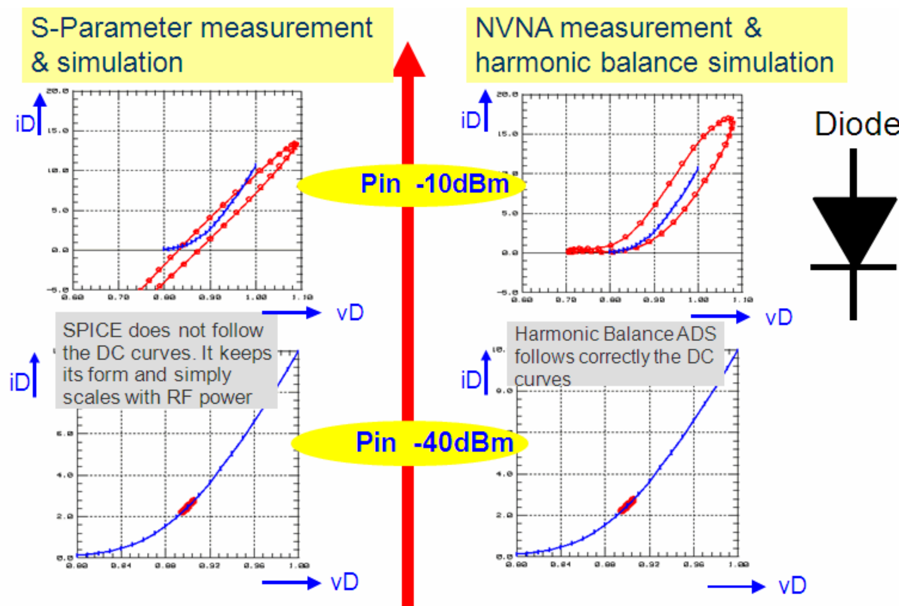


An S-Parameter simulation is a *linear* simulation,
forcing Kirchhoff's law to reflect a *single frequency only*!

Harmonic Balance (HB) simulation is a *non-linear* RF simulation,
with Kirchhoff's law reflecting *all possible harmonics frequencies*!

-> the sum of currents into a node is '0' (for the fundamental freq., all harmonics and DC)

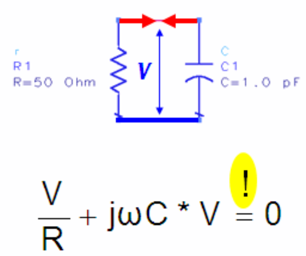
Furthermore, for HB, Z1=Z2=Z0 is an option, not a must !



The slide above depicts the difference between a linear measurement (and its corresponding S-Parameter simulation), and a non-linear measurement (and its corresponding RF simulation).
It shows a diode DC characteristic (blue) and the trajectory curve (red) of a certain DC bias point, vD=0,9V, stimulated with an RF signal of -40dBm (small signal) and of -10dBm (large signal).
The Nonlinear Case (right):The curve represents an ellipse for small RF signal, but for incrementing RF signal levels, this ellipse changes to a more complex trace, due to evolving harmonics (which are a consequence of the nonlinear DC trace!).

I.e. instead of following the measured DC traces in a tangent at a certain operating point, the trajectory curve of the large RF signal begins revolving along the DC characteristic due to the harmonics. It widens up due to the diode's capacitances and the transit time. And the DC operating point is typically also shifted, due to the energy provided by the large-signal RF source!
The Linear Case (left):
Also for large-signal RF stimulation, a linear VNA filters-out (and an S-Parameter simulation only considers) just the fundamental frequency (like it did it also for -40dBm), and thus, the trajectory curve will still exhibit **a tangent to the operating point also for large RF signals**. It will **keep its shape**, just become bigger. And it will **not** represent the real-world complex, satelliting curve following the DC trace.
In other words:due to the linear approach (no harmonics, just the fundamental frequency), the S-Parameter curves can only represent a trajectory curve in circle or ellipse form (Lissajoux curves). It keeps its shape independent of the applied signal power. It will only scale with it. This is equivalent to a linear VNA.
A NVNA measurement or an HB simulation, however, featuring fundamental **and** harmonic frequencies, gives realistic trajectory curves, which can realisticly follow the nonlinear DC curve.

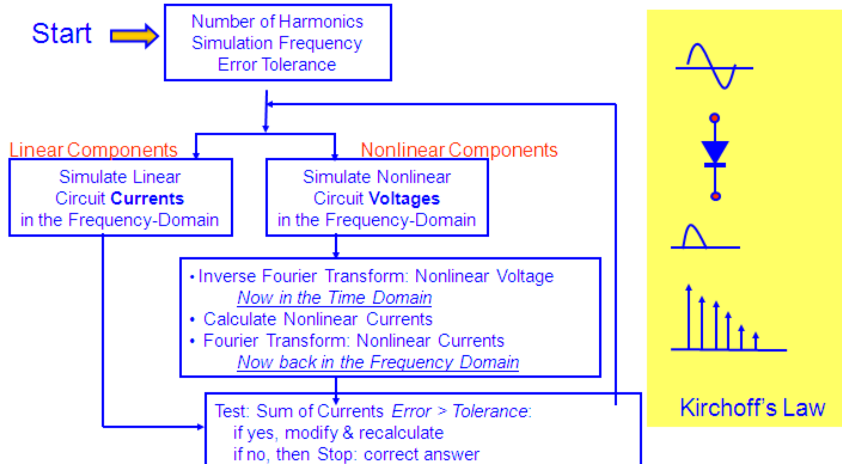
What is Harmonic Balance Simulation?



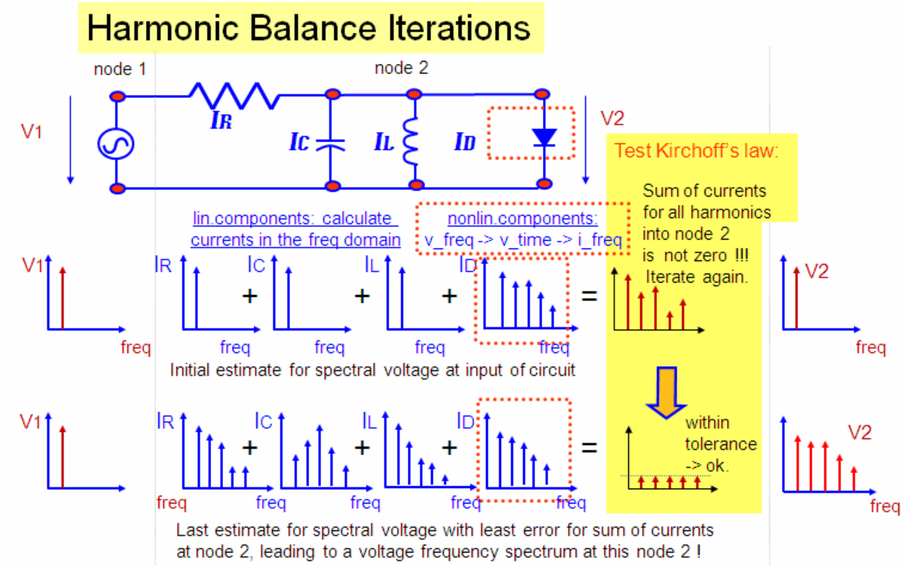
- Kirchoff's current law: sum of currents for all frequencies at each node is zero.
- Such a system of nonlinear equations is solved using e.g. the Newton-Raphson method

> **Note**
> Since most nonlinear devices are described by (Spice) models in the time domain, the simulator has to transform (inverse FFT) the intermediate voltage spectrum into the time domain, evaluate the response of the device and then transform it back to the frequency domain (FFT).

Harmonic Balance Simulation Flow Chart

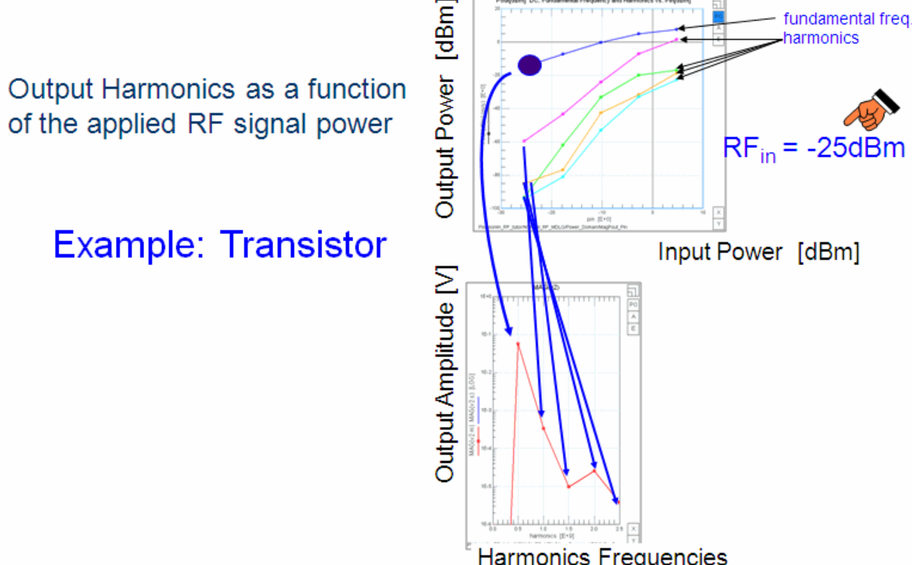


Harmonic Balance Iterations

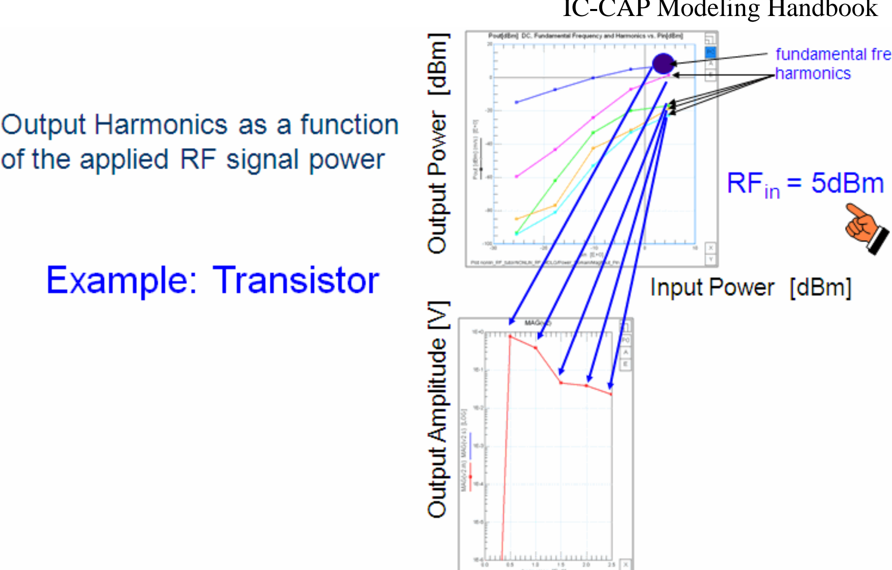


For the circuit given above, the following iterations are performed:

1. First, a DC simulation is executed to obtain the DC operating point of the circuit.
2. A linear AC analysis (Harmonic Balance Fourier Analysis) is applied to analyze the linear and passive components in the frequency domain, stimulated by the fundamental frequency.
3. Nonlinear device models are usually defined as i=f(v). Therefore, the actual voltage frequency spectrum across the nonlinear component is Fourier-transformed into the time domain. This resulting time domain signal is applied to the nonlinear device with a Spice-type simulation. The resulting time domain current is Fourier-converted back to the frequency domain.
4. The sum of the actual current spectra at each circuit node is calculated and it is checked, if this sum is below a certain tolerance limit for all frequencies.
5. Steps 1-4 are iterated until Kirchhoff's current law is satisfied for all frequencies at all nodes. The same is true for fulfilling Kirchhoff's voltage law. Iterative, matrix-driven techniques are used (Newton-Raphson algorithms etc.).
6. The obtained solution is a satisfaction of Kirchhoff's current law (sum of all currents, for all frequencies, is zero for each node) as well as the voltage law. The mathematical analysis has converged. The harmonics are now balanced.

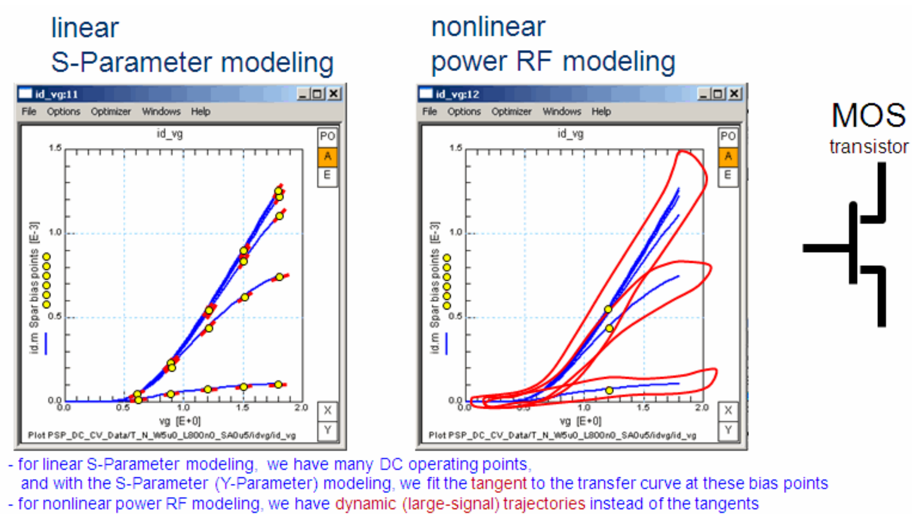


For low RF signal levels, in the linear range of the amplifier characteristics, the harmonic frequencies can be neglected.



For RF signal levels in the nonlinear range of the amplifier characteristics, the harmonic frequencies play a role and can no longer be neglected.

Linear vs. Non-Linear: modeling-wise

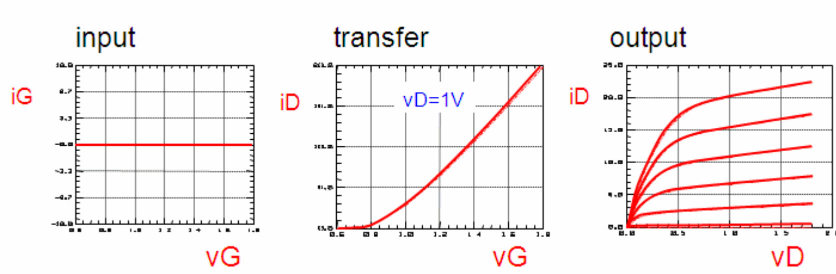


Related to S-Parameters

S-Parameters are a vehicle to describe the performance of a device for high frequencies.
S-Parameters are not the 'real world!'
However,
The real world is: non-linear currents and voltages vs. time!
A quick intermezzo a look at Device Modeling.
Example: MOS-Transistor
what happens to the measurement (and modeling) traces when going from DC -> S-Par -> nonlinear RF ?

From DC to large signal GHz

measuring a MOS transistor's
input, transfer and output characteristics at DC



This slide depicts the input, transfer and output characteristics of a MOS transistor. These plots will be investigated further in the next slides to study the effects of network analyzer measurements and NVNA measurements.

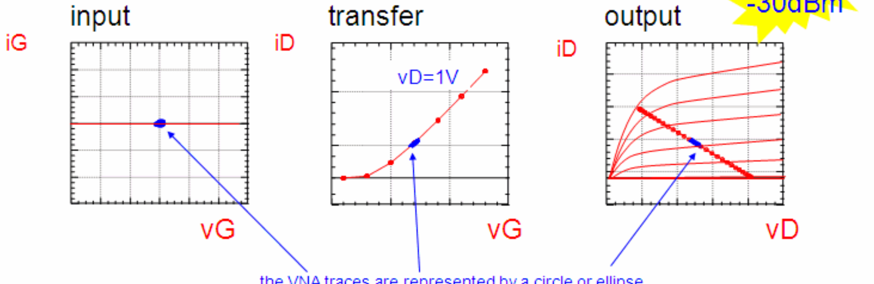
From DC to large signal GHz

Frequency Behavior at small signal levels
is characterized by a Vector Network Analyzer (VNA)

The device is stimulated with small signal levels around the DC bias point.
- If the frequency was low enough, the trajectory curve would exactly follow the DC trace.
- With higher frequency, however, the device can no longer follow the excitation,
i.e. the trajectory curves begin to widen up.

The HF signal has to be small enough to keep the DUT in small signal operation mode,
i.e. the trajectory curve is a circle or an ellipse (no harmonics).

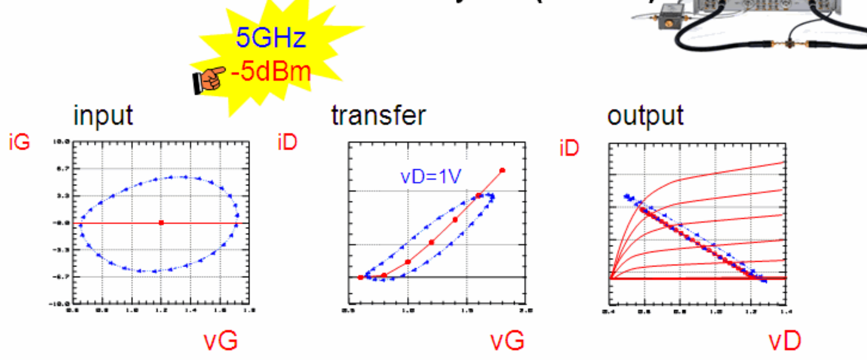
5GHz
-30dBm



Beyond 100MHz, the phase shift between the stimulating voltage and the measured current becomes bigger, and what was a tangent before during the CV measurements, becomes now an ellipse. In the slide above, see especially how the input current of the MOS transistor becomes now an ellipse around the stimulating vG operating point value.

From DC to large signal GHz

Applying larger RF and microwave power levels leads to harmonics and requires a Nonlinear Vector Network Analyzer **(NVNA)**



The NVNA measures the harmonics of an RF signal.
It measures what corresponds to a harmonic balance simulation (ADS).

Applying a large RF signal creates harmonic frequencies. Time-domain-wise, this is represented by complex satelliting curves ("cucumber" or "potato" shapes !).

> Applying the NVNA features realistic self-heating conditions for device modeling (for measurements and simulations).

When performing NVNA measurements, the transistor (or an amplifier etc.) can be operated in the conditions of its later application. The RF signal is satelliting around this operating point. Therefore, the self-heating is automatically covered correctly when measuring with the NVNA.
Besides measuring the Ai and Bi waves, a NVNA can also measure X-Parameters.

X-Parameters

Doing for nonlinear components under large-signal conditions what S-Parameters do for linear systems.

S-Parameters

- Calculated by the VNA out of the measured A1, A2, B1 and B2 power waves (fundamental freq. only)
- They represent a linear black-box-model of the DUT

X-Parameters

- Calculated by the NVNA out of a stimulation at port1 and at port2
- They represent a nonlinear black-box-model of the DUT (fundamental freq. + harmonics + (optionally) DC bias)
  X-Parameters are a rigorous, mathematically correct linearization of device under test (DUT) behavior as represented by a spectral map from incident to scattered pseudo-waves.
  Unlike classic S-Parameters, which capture only linear device behavior and ignore nonlinear behavior such as harmonic/intermod generation and even same-frequency higher order mixing effects, X-Parameters capture linear device behavior and linearize nonlinear behavior about a large signal operating point (LSOP).
  Both, S-Parameters and X-Parameters can be interpreted as a 'model' of the DUT !

> **Note**
> Ai and Bi measurements, as discussed before, are executed extremely fast by the NVNA. Measuring X-Parameters, applying the same stimuli takes much longer. Reason: a (small) stimulus signal is applied also to port 2.

S-Parameters: Nonlinear System X-Parameters

$$b_{ij} = X_{ij}^{(F)}(|a_{11}|)P^{j} + \sum_{k,l\neq(1,1)} \left(X_{ij,kl}^{(S)}(|a_{11}|)P^{j-l} \cdot a_{kl} + X_{ij,kl}^{(T)}(|a_{11}|)\ P^{j+l} \cdot a_{kl}^{*} \right)$$

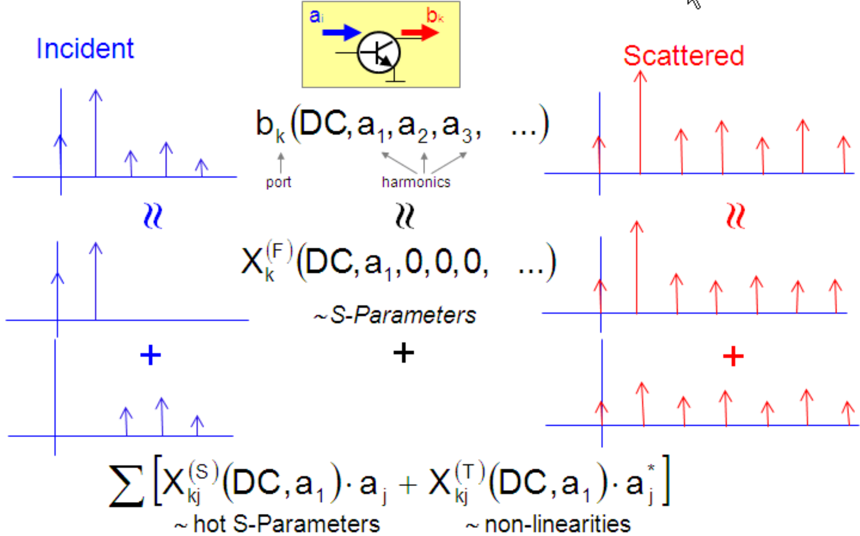
~ S-Parameters ~ hot S-Parameters ~ non-linearities

Indices: i = output port index, j = output frequency index, k = input port index, l = input frequency index

Different to S-Parameters (simple matrix), X-Parameters are represented by a rather non-linear matrix system, split into 3 terms: F-term, S-term and T-term. The F-term can be down-stripped to linear S-Parameters, - the S-term is similar to Hot S22 measurements, a very special application of (linear) S-Parameters to non-linear RF power devices - and the T-term takes into account all the non-linearities.
The terms $X^{(F)}$, $X^{(S)}$ and $X^{(T)}$ depend on the fundamental frequency and its harmonics and the power transfer between port m and port n, and (optionally)

depend also on the DC bias.

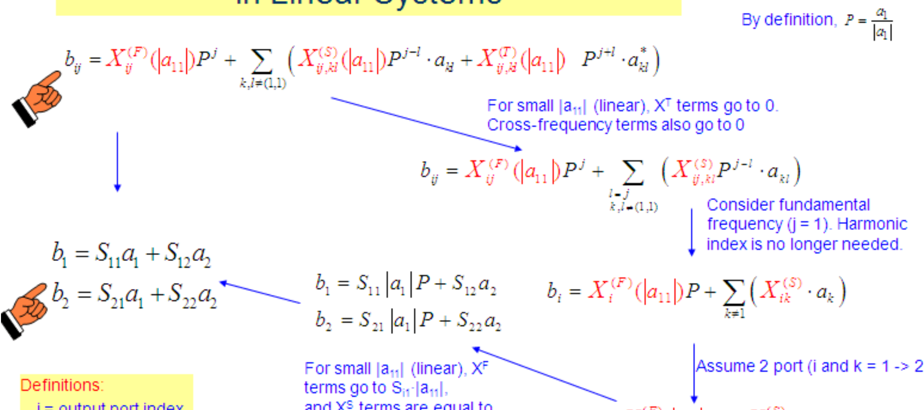
X-Parameters: Systematic Approximations to Nonlinear Mapping



What are HOT S-Parameters?

The output signal of nonlinear devices is input power level dependent. Most amplifiers, however, are operated close to the 1dB compression point. HOT S-Parameter measurements emulate this scenario, using a conventional linear VNA. For HOT S-Parameter measurements, the input signal at port1 is at frequency f1 , and its power level is so that the amplifier's output signal is close to the 1dB compression point. I.e. the amplifier is operated in its later application condition. At the same time, a second small test signal is applied at port2 with f2 = f1+Df, and S22 is measured with this 2nd stimulus. NOTE: The power level of this is typically -30dB smaller than the 1dB compression point output power. NOTE: Df should be bigger than the IF bandwidth of the VNA.

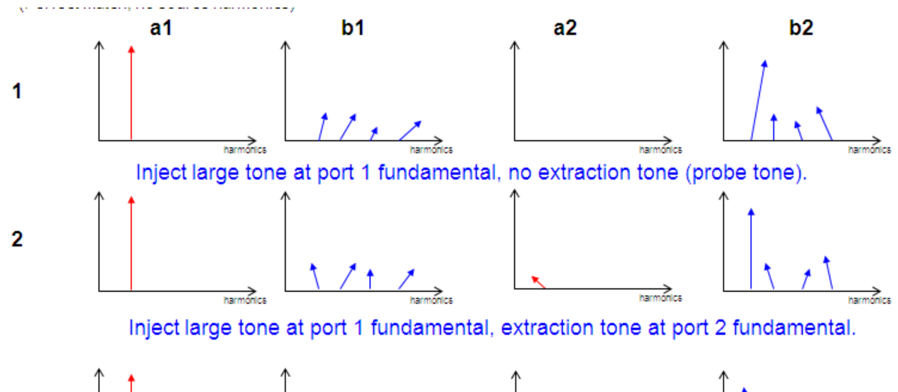
X-Parameters Collapse to S-Parameters in Linear Systems



References
The following references are available for additional information on X-Parameters and the PHD framework:

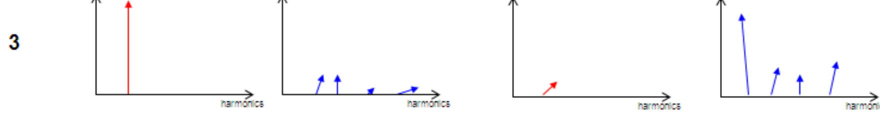
1. D. E. Root et al., "Broad-Band Poly-Harmonic Distortion (PHD) Behavioral Models From Fast Automated Simulations and Large-Signal Vectorial Network Measurements," IEEE Trans. MTT, vol. 53, no. 11, pp. 3656-3664, November 2005
2. J. Verspecht and D. E. Root, "Poly-Harmonic Distortion Modeling," in IEEE Microwave Theory and Techniques Microwave Magazine, June, 2006.
3. J. Verspecht, D. Gunyan, J. Horn, J. Xu, A. Cognata, and D.E. Root, "Multi-tone, Multi-Port, and Dynamic Memory Enhancements to PHD Nonlinear Behavioral Models from Large-Signal Measurements and Simulations," 2007 IEEE MTT-S Int. Microwave Symp. Dig., Honolulu, HI, USA, June 2007
4. J. Horn et al, "X-Parameter Measurement and Simulation of a GSM Handset Amplifier", accepted for publication at European Microwave Conference, 2008

Ideal X-Parameter Extraction Steps (Perfect match, no source harmonics)



Inject large tone at port 1 fundamental, no extraction tone (probe tone).

Inject large tone at port 1 fundamental, extraction tone at port 2 fundamental.



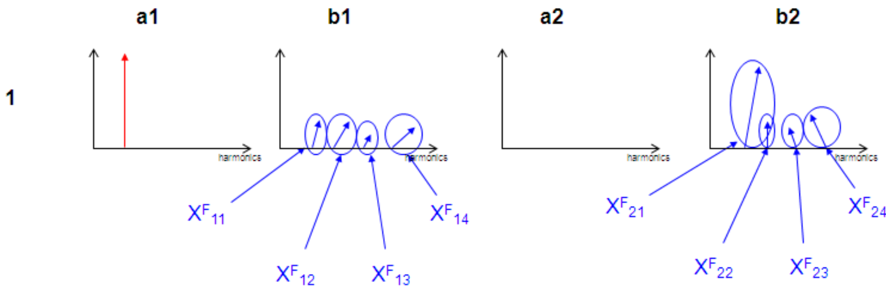
1 - inject large tone at port 1 fundamental
2 - add extraction tone at port 2 fundamental
3 - extraction tone is offset by 90 degrees from previous extraction tone.

47

The slide above describes the measurements necessary to extract X-parameters.
At each operating point (large signal level):

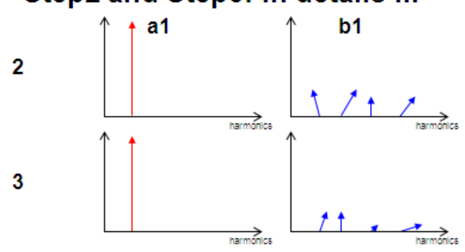
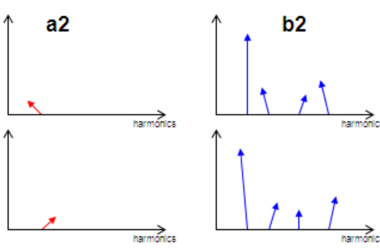
- One measurement is required with no extraction tone to get the XF terms
- Two measurements with extraction tones are required at each port/frequency of interest to get the XS and XT terms. The extraction tones in the two measurements have different phases (ideally 90 degrees offset).

Step1: in details ...



- One measurement is required, with no extraction tone, to get the X^F terms
- These XF terms are the measured scattered waves from measurement Step1

Step2 and Step3: in details ...

Two measurements with extraction tones are required at each port/frequency of interest to get the XS and XT terms.

The extraction tones in the two measurements have different phases (ideally 90° offset).

X^S and X^T terms are extracted by solving the systems of equations

$$b_{21}^{(2)} = X_{21}^{(F)}(|a_{11}|)P + X_{21,21}^{(S)}(|a_{11}|)a_{21}^{(2)} + X_{21,21}^{(T)}(|a_{11}|)P^2 \cdot a_{21}^{2*}$$

$$b_{21}^{(3)} = X_{21}^{(F)}(|a_{11}|)P + X_{21,21}^{(S)}(|a_{11}|)a_{21}^{(3)} + X_{21,21}^{(T)}(|a_{11}|)P^2 \cdot a_{21}^{3*}$$

where P is the phase of a11.

Additional equations at other port-harmonic combinations give other S and T terms related to a_{21}:

$$b_{ij}^{(2)} = X_{ij}^{(F)}(|a_{11}|)P^j + X_{ij,21}^{(S)}(|a_{11}|)P^{j-l} \cdot a_{21}^{(2)} + X_{ij,21}^{(T)}(|a_{11}|)P^{j+l} \cdot a_{21}^{(2)*}$$

$$b_{ij}^{(3)} = X_{ij}^{(F)}(|a_{11}|)P^j + X_{ij,21}^{(S)}(|a_{11}|)P^{j-l} \cdot a_{21}^{(3)} + X_{ij,21}^{(T)}(|a_{11}|)P^{j+l} \cdot a_{21}^{(3)*}$$

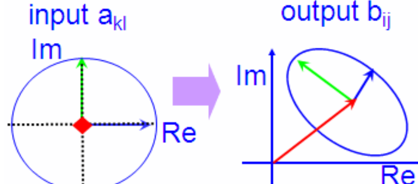
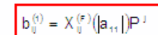
Repeat steps 2 and 3 at all port-harmonic combinations of interest to obtain all S and T terms.

49

X-Parameter Measurement Principle

$$b_{ij} = X_{ij}^{(F)}(|a_{11}|)P^j + \sum_{k,l \neq (1,1)} \left(X_{ij,kl}^{(S)}(|a_{11}|)P^{j-l} \cdot a_{kl} + X_{ij,kl}^{(T)}(|a_{11}|)\ P^{j+l} \cdot a_{kl}^{*} \right)$$

Perform 3 independent experiments with fixed a_{11}, using orthogonal phases of a_{21}

$$b_{ij}^{(1)} = X_{ij}^{(F)}(|a_{11}|)P^j$$

$$b_{ij}^{(2)} = X_{ij}^{(F)}(|a_{11}|)P^j + X_{ij,kl}^{(S)}(|a_{11}|)P^{j-l} \cdot a_{kl}^{(2)} + X_{ij,kl}^{(T)}(|a_{11}|)P^{j+l} \cdot a_{kl}^{(2)*}$$

$$b_{ij}^{(3)} = X_{ij}^{(F)}(|a_{11}|)P^j + X_{ij,kl}^{(S)}(|a_{11}|)P^{j-l} \cdot a_{kl}^{(3)} + X_{ij,kl}^{(T)}(|a_{11}|)P^{j+l} \cdot a_{kl}^{(3)*}$$

Indices: i = output port index, j = output frequency index, k = input port index, l = input frequency index

Wrap-Up:
1st measurement (red underline in the formula, red vector in sketch 'output bij'): -> XF
2nd and 3rd measurement (blue and green underline in the formula, blue and green vector in sketch 'output bij'): -> XS and XT

X-Parameters

The NVNA measures X-Parameters about a single-tone, large signal operating point.

This is achieved by stimulating the DUT with a single large tone at port 1 and, while continuing to apply the large tone, by injecting additional small tones at both ports 1 and 2 at all harmonics of interest. The level of the extraction tone must be small enough to ensure a spectrally linear response, but large enough to ensure that the response is measurable.The extraction tones will be injected into the DUT for the extraction of XS and XT.

At least two phase-offset small tones must be injected at each port/frequency of interest in order to extract the corresponding X-Parameters. Although two phases are sufficient to allow X-Parameter extraction, both the accuracy and robustness of the extraction are improved by using additional phases. A setting of five phases is recommended to provide robust X-Parameter extraction in most measurement configurations.

All resulting waves are measured for each stimulus, and the X-Parameters are calculated directly (if only two phases are used) or using regression (in the case of 3 or more phases).

Intermezzo: NVNA X-Parameter Display

Understanding the NVNA display results

First, let's describe the two X-Parameters: X^S_{pq} and X^T_{pq}.
These two terms relate the incident and scattered waves
at the two selected ports (labeled p and q) as follows:

a → q p → b

$$\Delta b_p = X^S_{pq} \cdot a_q + X^T_{pq} \cdot a_q^*$$

Δb_{pq} represents the change in the DUT's scattered wave b_p due to the phase-normalized incident wave a_q.

As in S-Parameters,
- the first subscript (p in the example above) is associated with the response signal,
- and the second subscript (q in the example above) is associated with the incident signal.

But because the DUT is driven simultaneously with a large signal, the small signal mixes with the large signal to create cross-frequency contributions, making harmonic indexing necessary in addition to port indexing.

So, both the p and q terms are described by a Port and Harmonic combination, as selectable in the NVNA window's "Port, Harm" field:
p (2,1) means port 2, fundamental frequency
q (2,2) means port 2, second harmonic frequency

For the selected harmonics, the X-Parameter display shows four terms for each of the XS and XT X-Parameters.

XSpp and XTpp
XSpq and XTpq
XSqp and XTqp
XSqq and XTqq

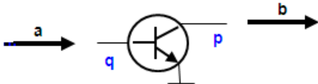
Port,Harmonic indexing in general:

- p=i,k
- q=j,l

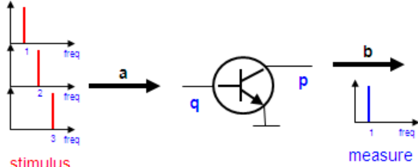
See the indexing in the previous slides.

Intermezzo: Number of X-Parameter measurements

In addition to XS and XT, there is also an XF term.XF is the X-Parameter component of the output due to the large signal input (a11). It is indexed only to the receive (port, harmonic) XFp. For each possible selection of p, there is exactly one XF term.

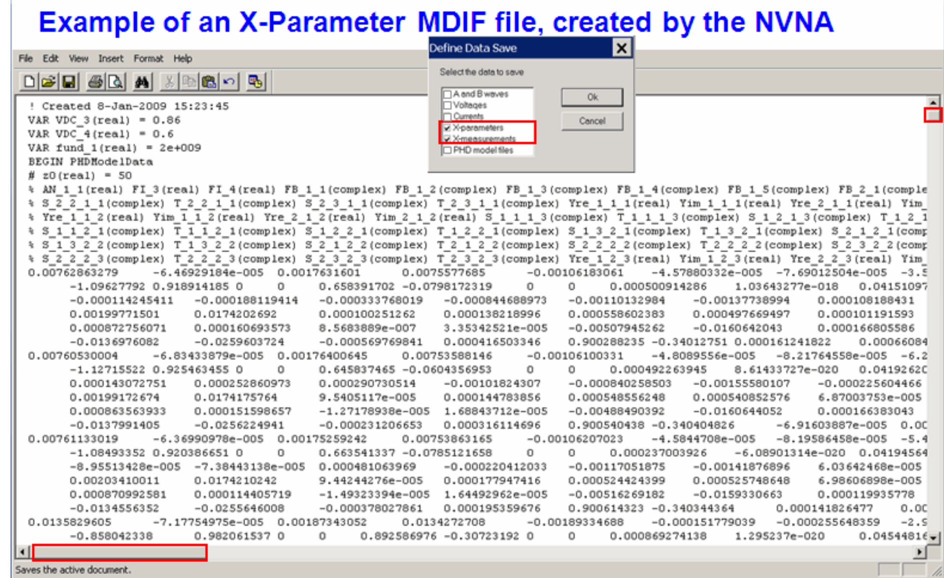


Therefore, in total, there are 13 X-Parameters for each possible (port, harmonic) selection of p. For example, for Port2 and the fundamental frequency, i.e. p=(2,1), and assuming 3 harmonics of interest:
(XF21)
(XS21,11), (XT21,11)
(XS21,12), (XT21,12)
(XS21,13), (XT21,13)
(XS21,21), (XT21,21)
(XS21,22), (XT21,22)
(XS21,23), (XT21,23)



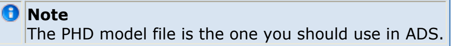
For a 2-port device, with 3 harmonics of interest at each port, there are 6 possible (port, harm) combinations for p. Each (port, harm) combination yields 13 parameters as shown above, and therefore, we end up with a total of 6*13 = 78 X-Parameters to be measured.

Example of an X-Parameter MDIF file, created by the NVNA



The size of such an MDIF file is typically several MByte.
NOTE about file types which can be saved by the NVNA after an X-Parameter measurement:

- The X-Parm measurement file is the raw (still vector corrected) measurements before applied to the X-Parm extraction algorithms.
- The X-Parms file is the extracted X-Parms but without normalization.
- The PHD model file is the normalized X-Parameters with normalization.



Note
The PHD model file is the one you should use in ADS.

Simulating with X-Parameters

Although subsets of X-Parameters can be useful design tools on their own, their true predictive power comes from using the entire set together. Since the X-Parameters include both magnitude and phase information, the computed Δb terms can be combined, along with the XF terms to accurately predict device response to a wide range of input signals.

$$\Delta b_p = X^S_{pq} \cdot a_q + X^T_{pq} \cdot a^*_q$$

This includes properly accounting for upstream harmonics and downstream match.
The complete equation for scattered wave prediction is as follows:

$$b_p = X^F_p + \sum_{q \neq (1,1)} X^S_{pq} \cdot a_q + X^T_{pq} \cdot a^*_q$$

with p=i,k and q=j,l (port,harmonic)

a q p b

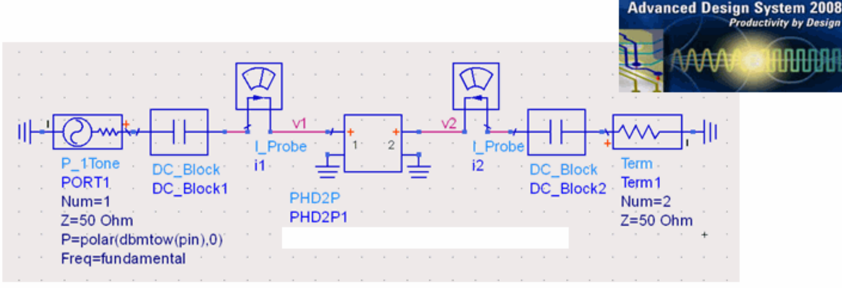
This powerful predictive ability has been combined with:

- Interpolation on frequency and input power as well as any additional independent variables such as DC bias
- Optionally load-pulling,
- Volterra constraints (to ensure physically correct extrapolation below measured power levels in the PHD framework of ADS).

ADS:X-Parameters become the PHD Model

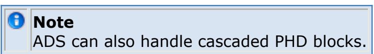
What is the PHD Model ?

PHD Model (Poly Harmonic Distortion Model) means: using the X-Parameters as a 'non-linear blackbox model' in ADS, and applying special interpolation algorithms

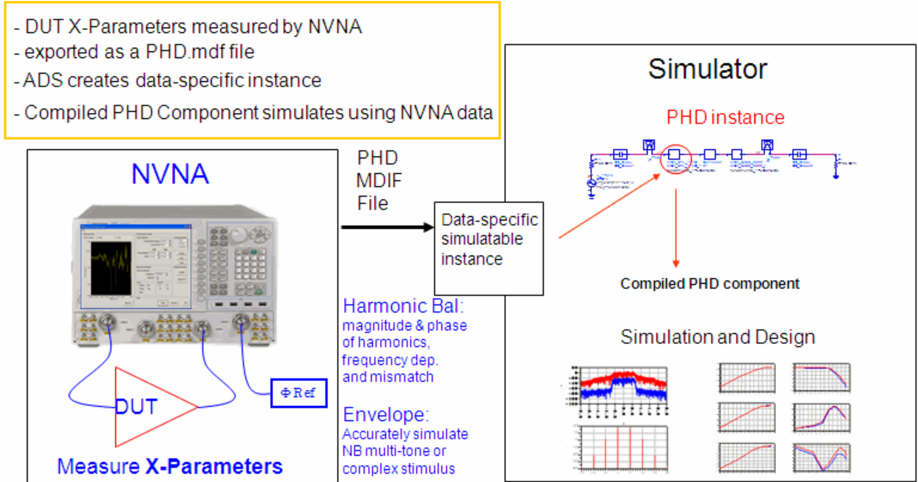


The NVNA can calculate and directly export the nonlin RF measurement into an ADS PHD.mdf data file (save the data as 'mdif file for PHD).Volterra constraints for small signal levels are enforced, ensuring that the harmonic levels go to zero at the appropriate rate.

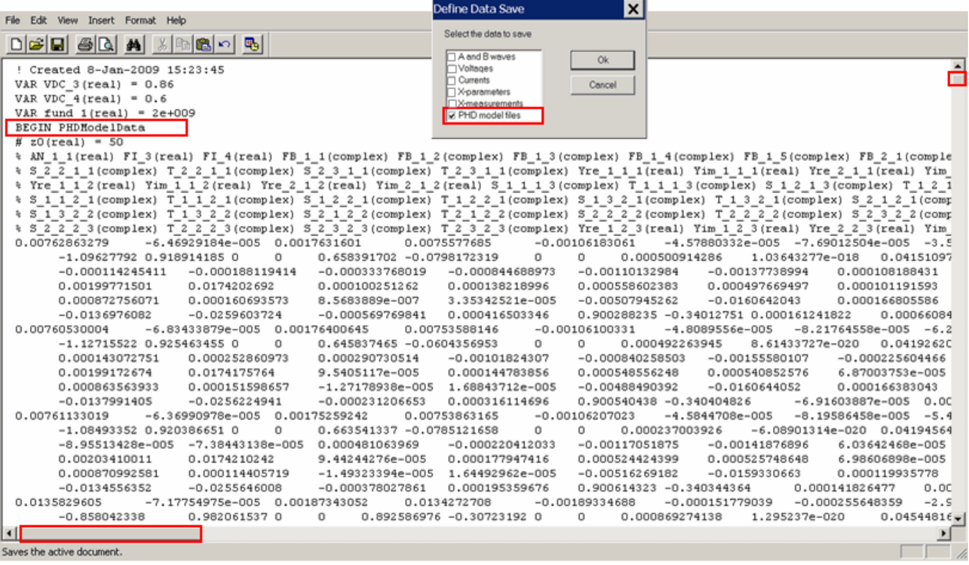
Also the data are re-ordered to be monotonic and fully compatible with ADS.
ADS can directly read this file.



PHD Model

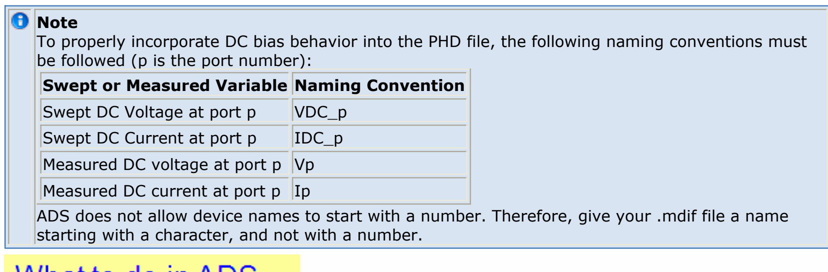


Example of an PHD MDIF file, created by the NVNA



The size of such an MDIF file is typically several MByte.
The file will include- the applied RF signal level (in SQRT(Watt)) in the data column
AN_1_1.
The conversion to dBm is after
Power_dBm = 30 + 10*log10((AN_1_1)^2)

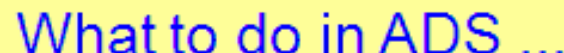
- FB-x-x for the X-Parameter's F-term, T-x-x-x-x for the T-term and S-x-x-x-x for the S-term.
- There may also be DC values present in this file, supposed the user has declared a sequence in the NVNA for setting different VDC levels by GPIB commands for a DC power supply.
  You find these DC bias informations as variables VDC_p where p is the port number on which the DC bias appears. If the bias variables are named VDC_1 and VDC_2, the bias information is covered by shared DC/RF ports 1 and 2. In this case, you will need DC_Feed and RF_Feed components in ADS around your PHD component, to separate DC from AC. If VDC_3 and VDC_4 are used, DC-only ports 3 and 4 will be created on the component. You should consider this when naming the DC bias voltages in the DC instrument driver section of the NVNA. The DC bias currents can be found in the data columns FI_p.

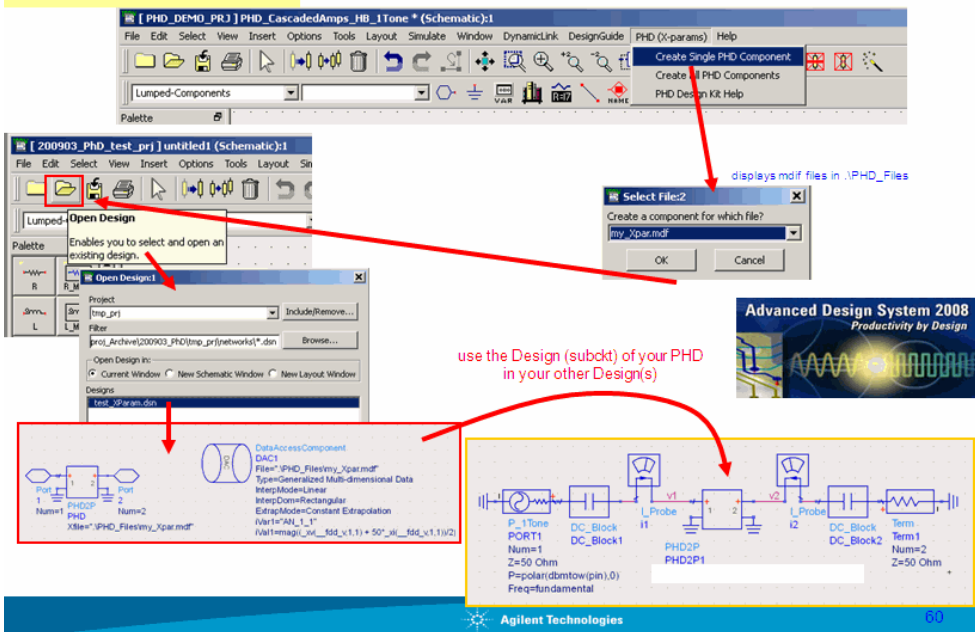


Note
To properly incorporate DC bias behavior into the PHD file, the following naming conventions must be followed (p is the port number):

Swept or Measured Variable	Naming Convention
Swept DC Voltage at port p	VDC_p
Swept DC Current at port p	IDC_p
Measured DC voltage at port p	Vp
Measured DC current at port p	Ip

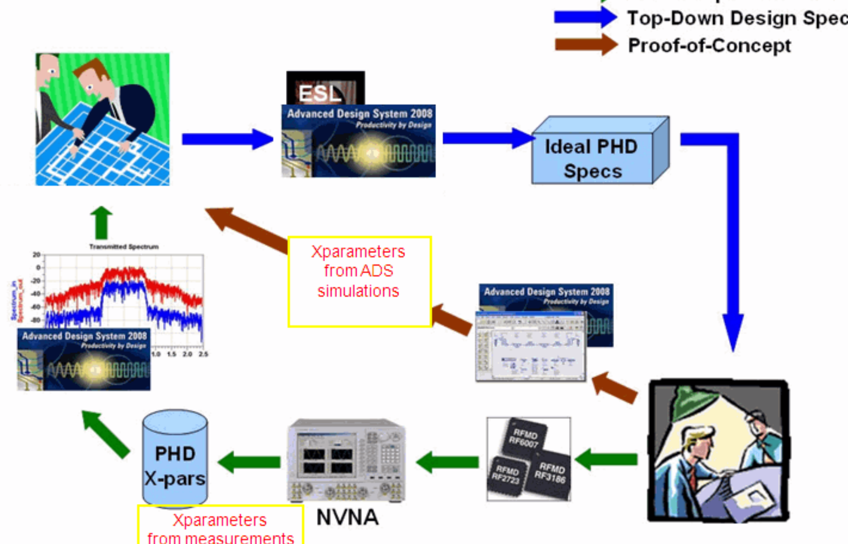
ADS does not allow device names to start with a number. Therefore, give your .mdif file a name starting with a character, and not with a number.





How to load a PHD file into ADS:

1. In the NVNA, save the data as 'mdif file for PHD'.
2. In your ADS project directory, create a new directory called 'PHD_Files'.
3. In your ADS session, select (as shown above) the menu pick 'PHD (X-Params)/Create_single_PHD_Component'. A file browser opens up, and you should be able to select your X-Parameter .mdif file(s), located in the project sub-directory 'PHD_Files'.
4. To see what has happened: In ADS/Main, open a new 'Schematic', and select 'Open Design'. Select the previously created PHD Project. Then close this Design.
5. Open a new Design (or use one of your existing ones), and select the PHD-subcircuit (Design) from 'Component Library/Designs'. Drag your PHD Design into your Schematic.



- Besides creating PHD models from NVNA measurements, PHD models can also be created out of (parts of) a schematic in ADS.
- Using PHD .mdif files, relevant information can be exchanged without disclosing technology information.

Wrap-Up NVNA Measurements

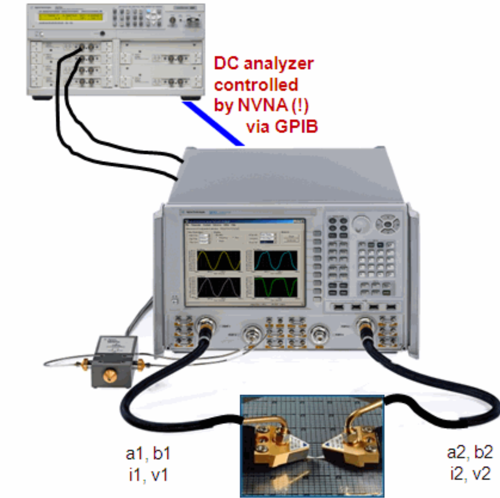
- S-Parameters: a linear matrix, calculated by the VNA out of the measured linear Ai and Bi waves.
- NVNA Harmonics Measurement: The nonlinear Ai and Bi power waves with a RF signal applied to port-1 (corresponding to a HB simul.). Measures Magnitude and Phase of a speficied nr. of harmonics.
- X-Parameters: A highly nonlinear matrix system, calculated by the NVNA out of the nonlinear Ai and Bi waves when applying signals to both ports. Is called PHD Model when used in ADS.

  Agenda
  Introducing Nonlinear RF

  - Linear S-Parameter Measurements
  - Nonlinear VNA Measurements
  - Harmonic Balance (HB) Simulations
  - X-Parameters
  - PHD Model

  Device Modeling Aspects for Nonlinear VNA On-Wafer Measurements

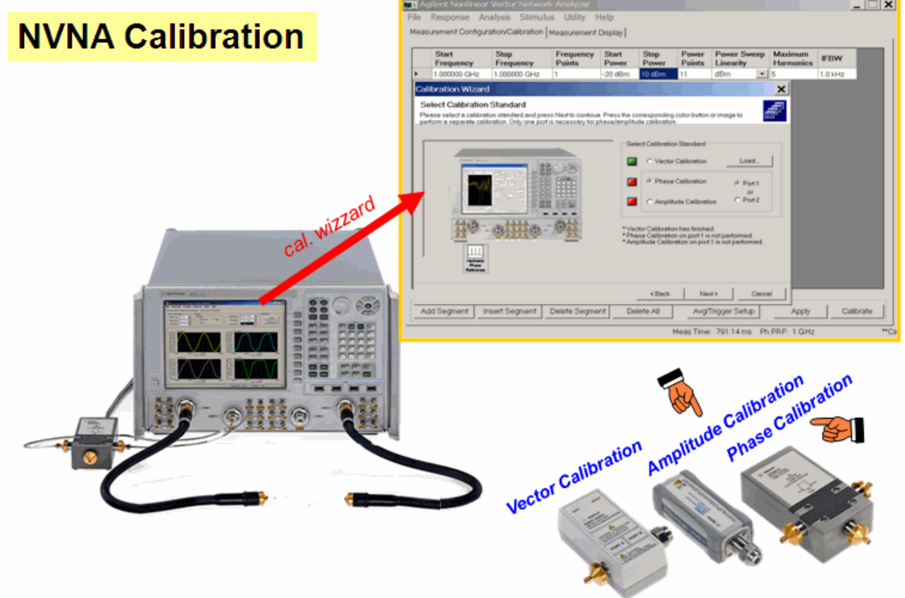
Performing the NVNA on-wafer Measurements



For our application, the NVNA measures the transmitted (b) and reflected (a) PowerWaves and calculates the ix and vx.

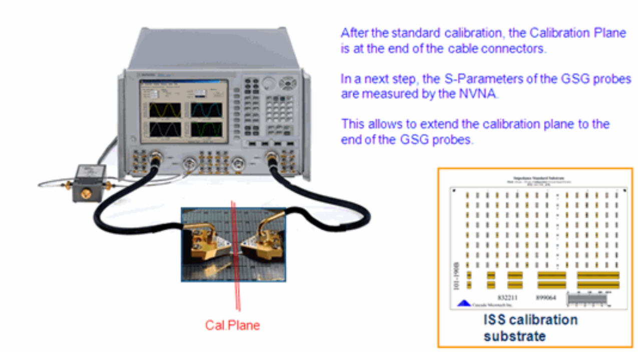
> **Note**
> The NVNA also triggers the DC Analyzer which provides the DC bias for the nonlinear RF measurement.

NVNA Calibration



The NVNA requires, additional to the conventional Vector Calibration known from linear network analyzers, an Ampitude Calibration and a Phase Calibration, using a phase standard.

NVNA Calibration Plane

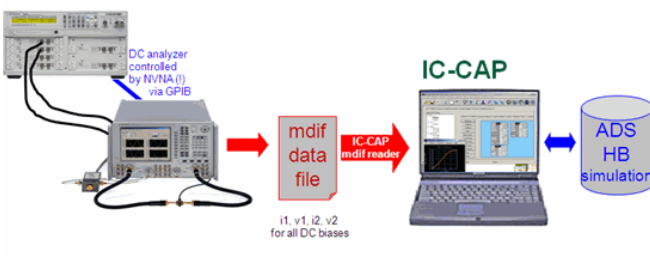


On the NVNA, probe de-embedding is supported either during measurement or during calibration through the NVNA software. The deembedding settings are accessed under the Response->Cal->Fixturing menu.

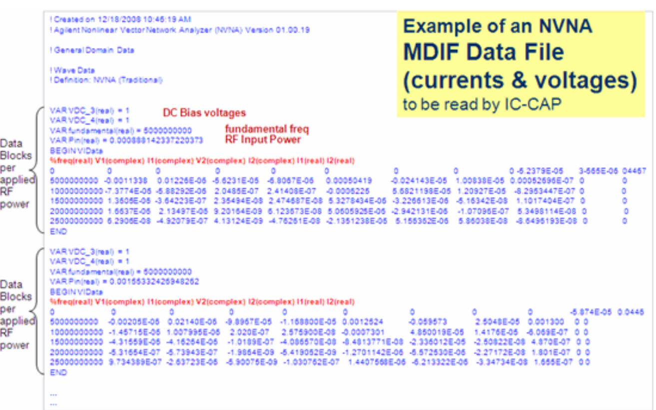
- A connectorized (3.5mm) cal, plus an OPEN, SHORT and LOAD measurement of the ISS standards.

> **Note**
> Measurement of the GSG probe S-parameters can be done with the Adapter Characterization macro in the PNA-X VNA firmware or using 3rd party software such as WinCal.

Importing the NVNA (DC, i1, v1, i2, v2) mdif data file into IC-CAP



Pls. contact the author for the IC-CAP file 'read_NVNA_mdif.mdl'.



As discussed so far, the NVNA features an integrated down-to-the-probe-tips calibration:implemented as:

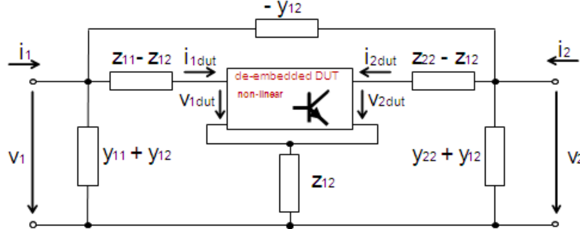
- a connectorized (3.5mm) cal,
- plus an OPEN, SHORT, LOAD measurement of the ISS standards and their de-embedding.

For on-wafer diodes and transistors with contact pads, we need to apply **Nonlinear De-Embedding**

- The measured nonlinear currents and voltages cannot be converted to e.g. S-Parameters or Y-Parameters
- De-Embedding by matrix subtraction is not possible !!!

We have to deal with the voltages and currents (Kirchhoff's Law)

The de-embedded large signal voltages and currents (v1dut, i1dut, v2dut, i2dut) are obtained from the measured ones (v1, i1, v2, i2) after:



NonlinearDe-Embedding

$$i_{1\,dut} = i_1 - v_1(y_{11} + y_{12}) - (v_2 - v_1)y_{12}$$

$$i_{2\,dut} = i_2 - v_2(y_{22} + y_{12}) - (v_1 - v_2)y_{12}$$

$$v_{1\,dut} = v_1 - i_{1\,dut} \cdot z_{11} - i_{2dut} \cdot z_{12}$$

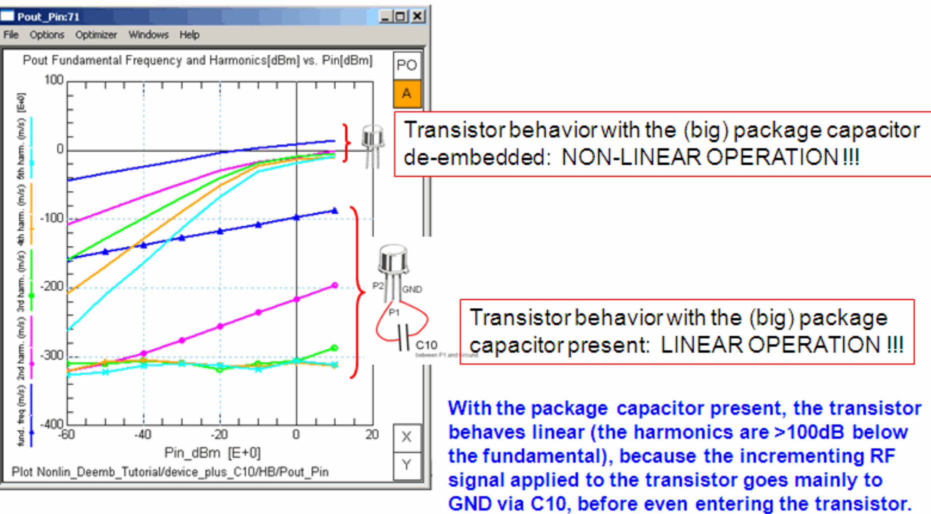
$$v_{2\,dut} = v_2 - i_{2\,dut} \cdot z_{22} - i_{1dut} \cdot z_{12}$$

In the formulae,Yij are the S-to-Y-converted S-Parameters of the OPEN Dummy Zij the S-to-Z-converted S-Parameters of the SHORT Dummy (de-embedded before from the OPEN Dummy)

> Pulication: G.Crupi, D.Schreurs, D.Xiao, A.Caddemi, B.Parvais, A.Mercha, S.Decoutere: "Determination and Validation of New Nonlinear FinFET Model Based on Lookup Tables", IEEE MicrowaveAnd Wireless Components Letters, Vol. 17, No. 5, May 2007
> IC-CAP Demo File: nonlinear_deemb_Tutorial.mdl

Case Study Example

Non-linear de-embedding of a transistor from a big Capacitor



In this case study, a transistor plus an extra package capacitor (with a big capacitance value, to make the effect more obvious) is considered.This extra package capacitor shall be de-embedded.
When we simulate the transistor **plus** the (pretty big) package capacitor, and sweep the RF input signal from -60dBm to +10dBm, and display the fundamental and the harmonics **at the pins** of the transistor, the fundamental follows the input signal, but at a much lower level. This is because most of the stimulating input signal is lost to ground via the (big) capacitor. As a consequence, the harmonics at the very transistor pins are >100dB below the fundamental:

- The transistor behaves linear, although the RF input signal is very big.
  When we simulate the transistor **without** the capacitor, all of the RF input signal goes across the transistor. The fundamental at the output again follows the input signal, with 20dB above the input signal level, and is clipped at an input signal level of ~ -20dBm. And the harmonics are not negligible at all.

- Nonlinear de-embedding affects the operating condition of the DUT: in our example, the transistor changes from linear operation (non-de-embedded) to non-linear (de-embedded).
- Linear deembedding, i.e. subtracting Y- and Z-matrices, keeps the DUT in its operating mode: also after de-embedding, it is linear.

> **Note**
> Therefore, when measuring S-Parameters of devices like transistors, we take so much attention to not overdrive the transistor, and to always keep the transistor in linear operation.

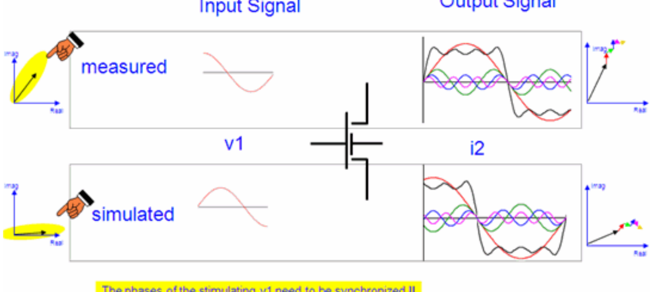
Notes about De-Embedding

- S-Parameter: device was linear before, is linear after de-embedding
- NVNA: device's linearity/non-linearity is affected by the de-embedding

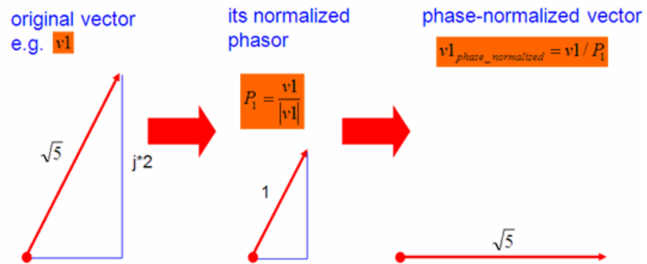
  We have now

  - Shifted the cal. plane to the very DUT,and are ready to
  - Compare these measurements against simulations, to finally do
  - Device modeling

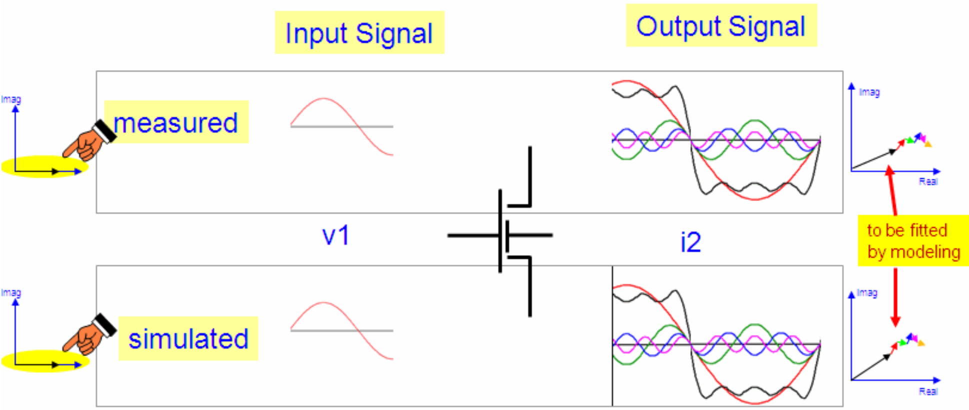
In order to compare the NVNA measurement with an HB simulationin the time domain, the phases of measured and simulated signal need to be normalized.



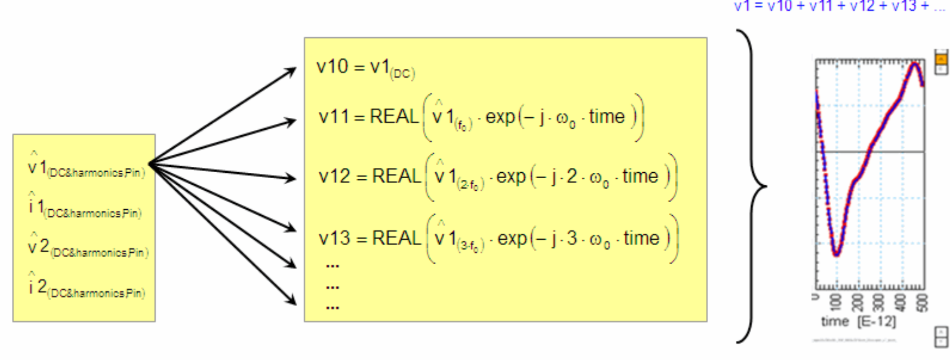
Intermezzo:Phase normalization is achieved by dividing a signal by its phasor



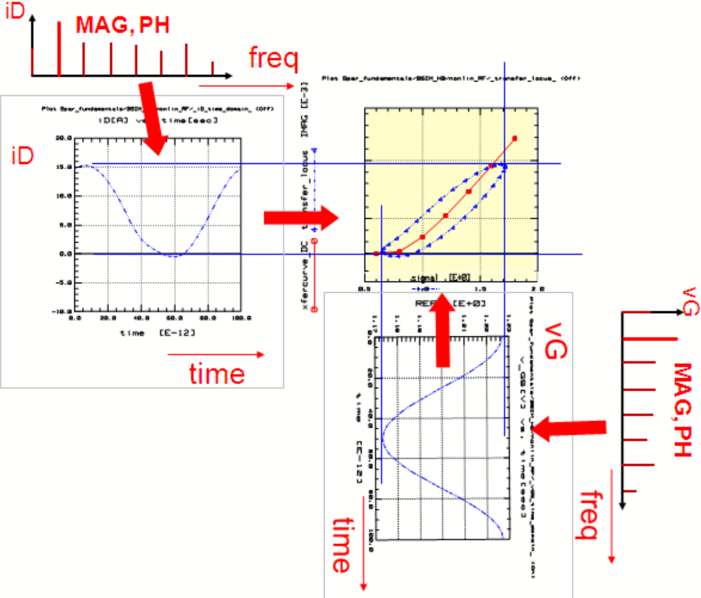
Usually, v_1 is used as reference signal. Therefore, all measured v_x and i_x are divided by the phasor of v_{1meas} and all simulated v_x and ix are divided by the phasor of v_{1simul}.
This means: v_{1meas} has 0 phase, and v_{1simul} has 0 phase as well.



Calculate the Fourier components of the vx and ix

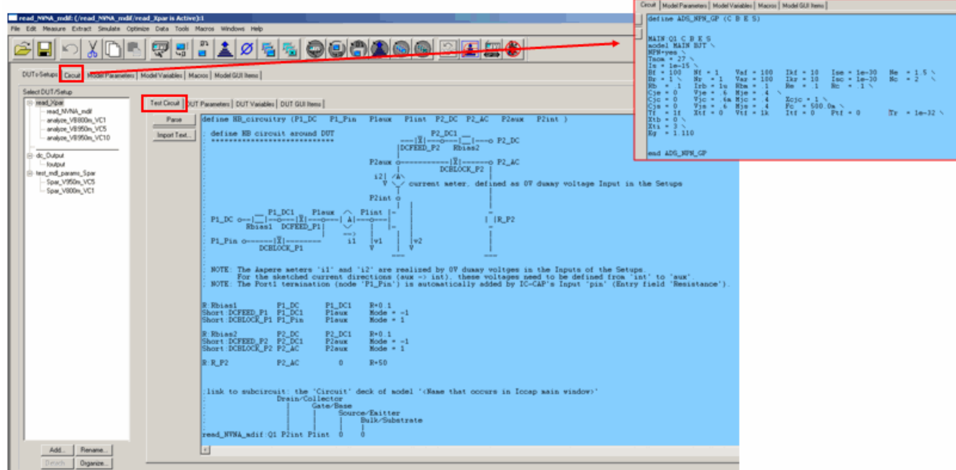


Creating the Dynamic Transfer Trajectory



Now that we have the time-dependent voltages and currents of the fundamental and its harmonics, we can even plot current vs. voltage (trajectories).

Handling NVNA measurements and standard device modelslike PSP or Hicum, Gummel-Poon, Angelov etc. in IC-CAP



Once the NVNA data have been imported into an IC-CAP session, a conventional device modeling can be applied, using a BSIM, PSP, Hicum, Angelov, Curtice model etc. Physics-based device models are a useful application of NVNA measurements, when considering-> scaling rules for compact models that allow you to measure a device of many sizes. -> process variation is another important application for compact models.

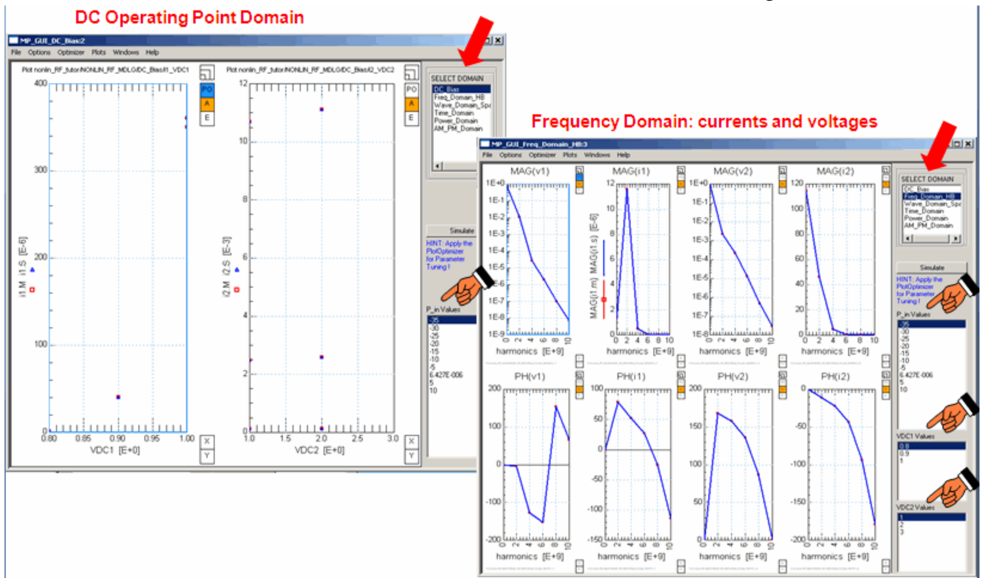
A possible Modeling Sequence:

- DC
- S-Parameters
- Convert to Y-Parameter for PI-schematic modeling
- Extract, verify and fine-tune model parameters using the NVNA data in time and power domain

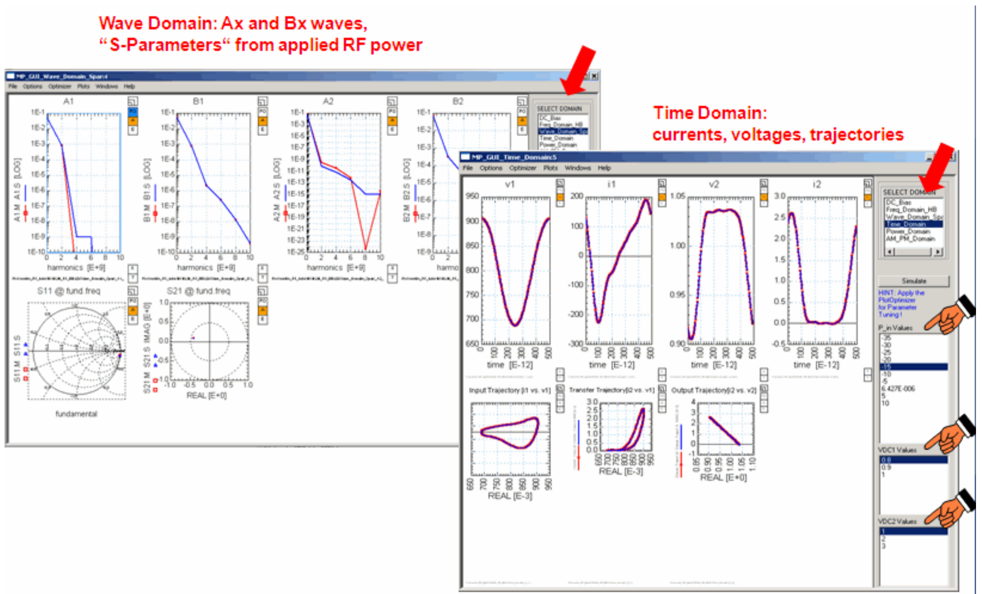
To ease your learning of nonlinear RF measurements, simulations and modeling, an IC-CAP tutorial ModelFile has been developed.

> **Note**
> Pls. contact the author for a copy of the IC-CAP nonlinear RF demo file.

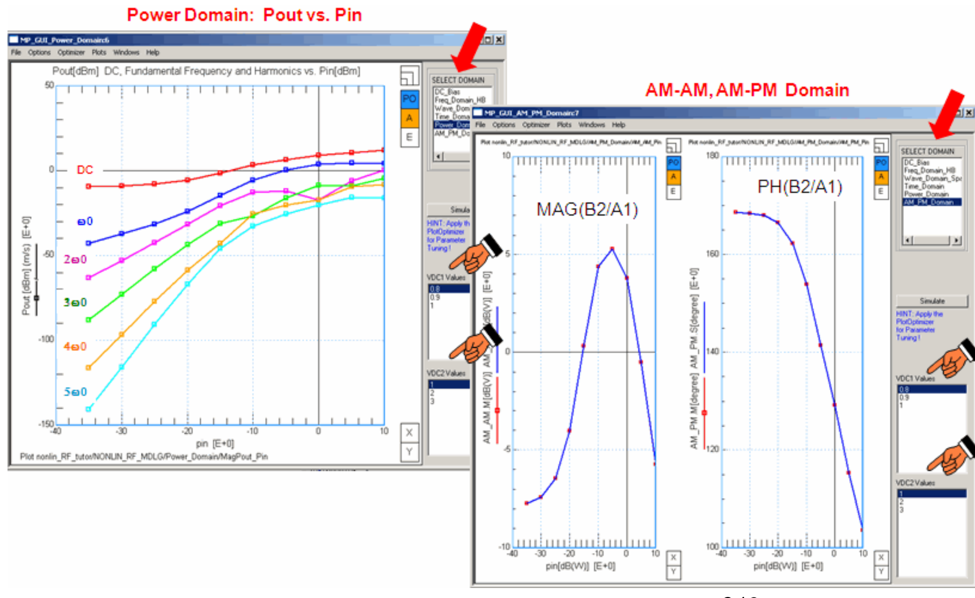
The IC-CAP Nonlinear RF Tutorial ModelFile (1)



The IC-CAP Nonlinear RF Tutorial ModelFile (2)



The IC-CAP Nonlinear RF Tutorial ModelFile (3)



Conclusions

Introducing Non-Linear RF

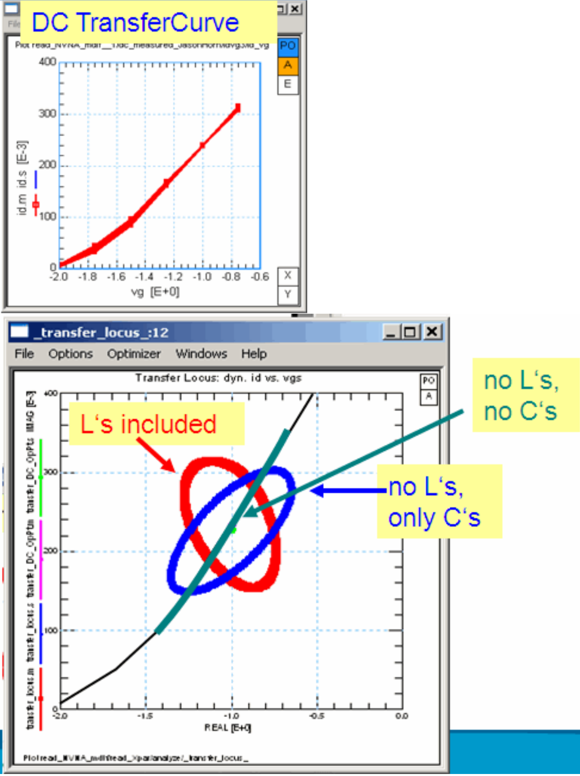
- Linear S-Parameter Measurements
- Nonlinear VNA Measurements
- Harmonic Balance (HB) Simulations
- X-Parameters
- PHD Model

Device Modeling Aspects for Nonlinear VNA On-Wafer Measurements

Modeling Tips and Tricks

Effect of L's and C's on the Dynamic Transfer Curve

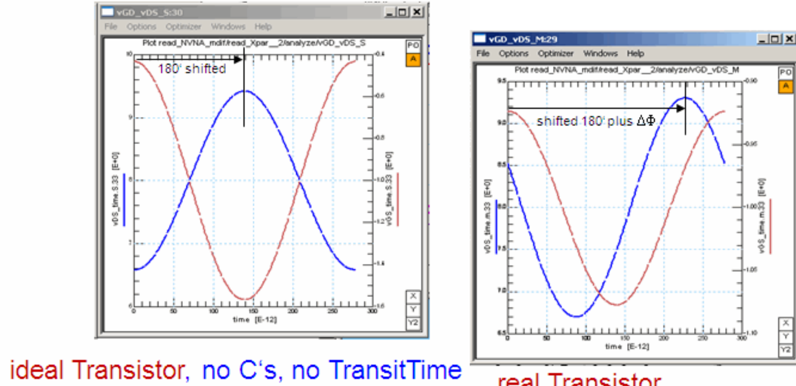
For low RF signal, the dyn. transfer curve follows the DC trace perfectly, if all C's = 0F. When adding the C's with the ADS simulation, the tangent opens up to a 'cucumber'-shaped trace following up and down the DC trace, and widening up with increasing C values. For a trajectory curve perpendicular to the DC trace at low RF signal, we need to apply inductors (package) in addition to the capacitors.



vDS vGS in the time domain

LEFT: The time-domain-converted ADS harmonic balance simulations of a simple transistor model: just a voltage-controlled current source between Drain and Source (gm = 50mS), no capacitors, no transit time.
The data refer to Pin=-30dBm. As it can be expected, vDS is counter-phase to vGS.

RIGHT: Again vGS and vDS, but now including the model capacitors. And again the lowest power applied in the measurements, i.e. no clipping, just the fundamental frequency.
Since the capacitances are not 0, the vDS is **behind** or **comes later** than the ideal vDS in the simulations. (Although may look like vDS comes before vGS).


180' shifted
shifted 180' plus ΔΦ
ideal Transistor, no C's, no TransitTime
real Transistor

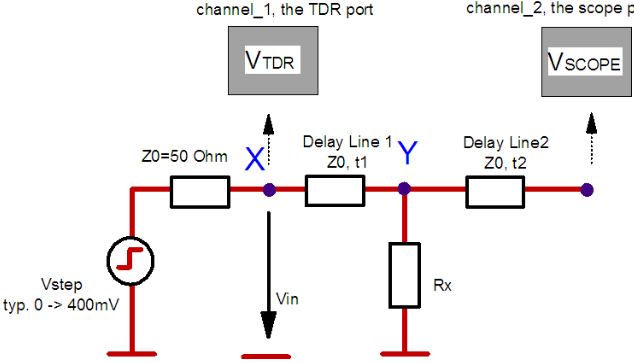
Time Domain

Contents

- *TDR Measurements Basics* (iccapmhb)
- *TDR Calibration Techniques* (iccapmhb)
- *TDR Plots Tutorials* (iccapmhb)
- *Differential TDR Measurements* (iccapmhb)

TDR Measurements Basics

This chapter is intended to make you familiar with TDR measurements and the interpretation of TDR plots. Let's commence with some basic TDR theory: Assuming we have the following measurement, an unknown resistor Zx to ground between two 50 Ohm delay lines.



At any point within the tested device, there is a backreflected and an on-going voltage:

$$V_{refl} = V_{in}\frac{Z - Z_0}{Z + Z_0} \quad (1) \qquad V_{thru} = V_{in}\frac{2 * Z}{Z + Z_0} \quad (2)$$

with Z : impedance seen in forward direction at the actual location.
Example:
Be Vstep = 1 (normalized), then Vin = 0.5 when the step hits the location X at t=0.
At location Y , we have after time t1 and with Rx = 50 Ohm using equ.(1):

$$Vrefl = \frac{1}{2} * \frac{25 - 50}{25 + 50} = -\frac{1}{6} \quad (3)$$

This voltage Vrefl is reflected back to the TDR port and seen there after time 2 * t1 overlaying the on-going Vin=0.5 as

$$VTDR = Vin + Vrefl \overset{(3)}{=} \frac{1}{2} - \frac{1}{6} = \frac{1}{3} \quad (4)$$

The ongoing propagating wave beyond location Y is then,

$$Vthru \overset{(2)}{=} \frac{1}{2} * \frac{2 * 50}{50 + 25} = \frac{1}{3}$$

what is observed with the scope input at the open end at the time t1 + t2 as

$$VSCOPE = Vopen_end = 2 * Vthru = \frac{2}{3}$$

Where factor 2 comes from equation (2) for Z = infinite (open end).
As another example, we can calculate the
impedance Z at any location of the backreflected TDR graph:
From (1), we obtain

$$\frac{V_{refl}}{V_{in}} = \frac{Z - Z_0}{Z + Z_0}$$

or solved for Z:

$$Z = Z_0\frac{V_{in} + V_{refl}}{V_{in} - V_{refl}} \qquad \text{--(5)}$$

It is,

$$V_{step} = (V_{in} + V_{ref}) + (V_{in} - V_{ref}) \overset{(4)}{=} V_{TDR} + (V_{in} - V_{ref})$$

and solved for,

$$(V_{in} - V_{ref}) = V_{step} - V_{TDR} \qquad \text{---(6)}$$

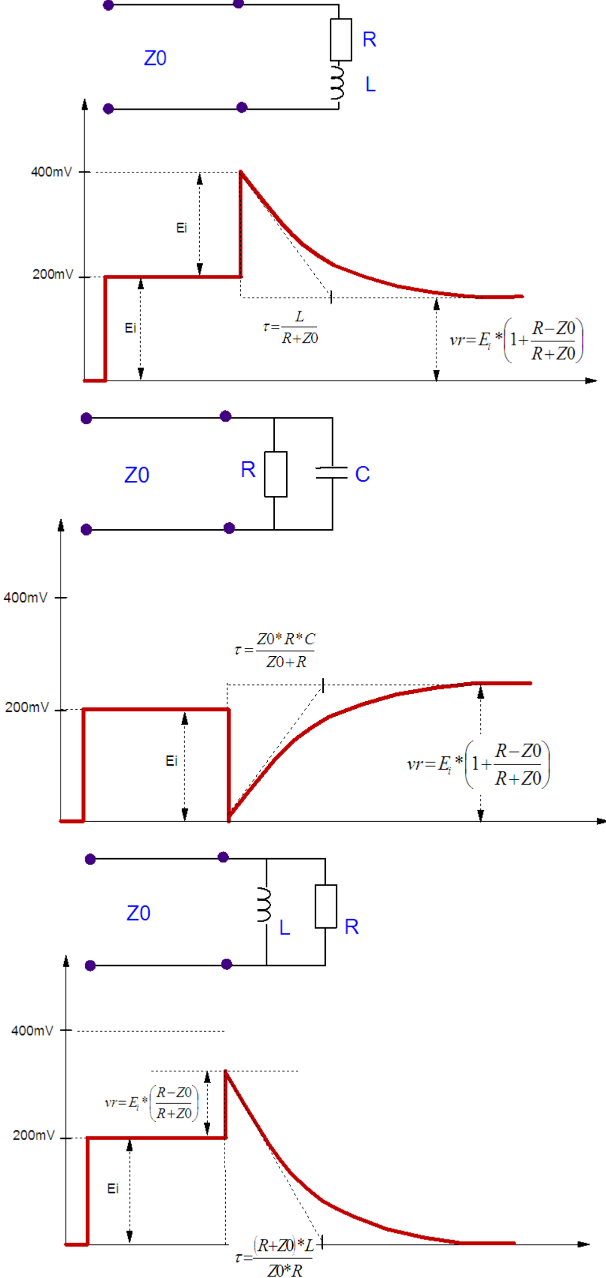
Introducing (6) into (5) gives finally the local impedance along the TDR measurement:

$$Z = Z_0\frac{V_{TDR}}{V_{step} - V_{TDRl}}$$

with Vstep typically 400mV and 0 < VTDR < Vstep

Modeling the Complex Load Out of TDR Plots

Equations for ideal TDR response to complex loads can be derived, allowing to determine R, L, and C values of a circuit. Response time constant and final value, resp. incremental step, are usually the most important factors for such calculations. Some typical TDR responses can be seen in figure 1 below. Its sketches are from the HP application note 62-3 'Advanced TDR techniques'.
For example, with a series RL circuit, with a measured time constant τ of 20ps, and a final value of 150mV from a 200mV input Ei: R=30 Ohm, and L=1.6nH.



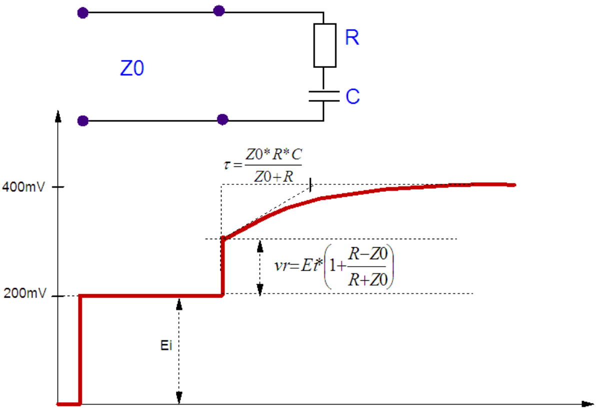


Figure: TDR responses to complex loads

About the Context Between Time and Location

The signal delay TD is given by the length L of the line, and the dielectric constant of the dielectric materiel. The delay is not dependent on the diameter of the conductor.
For a conductor, surrounded homogeneously by a dielectric (εr), it is

$$TD = L * \sqrt{\varepsilon_r * \varepsilon_0 * \mu_0}$$

with
ε_0: 8.854*10-12 F/m

ε_r: dielectric constant of the isolator

μ_0: 1.257*10-6 H/m

L: conductor length
or

$$TD = L[m] * 3.3ns * \sqrt{\varepsilon_0}$$

As a rule of thumb, the speed of a step function on a ceramic substrate strip line (εr~3) is roughly:

```
TDR-time        10ps
--------   =   -----
Distance         1mm

          or:

Distance        10cm
--------   =   -----
TDR-time         1ns
```

Note
On a TDR, the displayed time represents the distance to the location and back.

This gives also an idea about when the connecting components between electronic devices have to be modeled too. This is necessary when the forth and back traveling impulse time is longer than the impulse rise time. This means that the still rising edge of the impulse will be interfered by the back reflection (glitches in digital signals).

TDR Calibration Techniques

TDR (time domain reflectometer) measurements meet ideally the needs of modeling components with geometric extension. This type of devices include packages, connectors, strip lines on PC boards (printed circuit) etc. The advantage of TDR measurements is that the individual components of the DUT appear at separate time slots in a direct relation to their location. The NWA dilemma, namely that all sub-components of the DUT add basically to the 'turning to the right' of the Smith chart, and thus overlay each other's effect, is easily avoided by this method.

Basically, a TDR is a very fast GHz oscilloscope including a fast pulse generator in the picoseconds range. This fast pulse with an amplitude of usually 400mV is applied to the DUT, and the backreflections of that voltage are displayed on the oscilloscope. Therefore, the time axis on that plot corresponds to the physical location of the event on/in the DUT.

For example, when an OPEN ended line is connected to the TDR output, the resulting oscilloscope reading shows a 200mV amplitude (voltage divider: 50? internal resistance and 50Ω characteristic impedance of the line), as long as the impulse is present in the line. Once the open end is reached, there is a jump to 400mV. Otherwise, if the line is SHORTed at its end, the impulse again sees a 50Ω characteristic impedance (while still in the line), yet disintegrates to 0V when hitting the SHORT at the end of the line.

If there is a capacitor between two 50Ω lines, the reflected signal with its level of 200mV (in line 1 in fig.1) will briefly drop down to 0V, 'seeing' the capacitor as a SHORT for the first moment, and then ascend back to 200mV (line 2) with an exponential rise that is proportional to the capacitance value. In case there is an inductor instead of the capacitor, the reflected signal will jump to a full 400mV (because the inductor behaves like an OPEN in the first moment) to once again come back to the 200mV level of line 2.

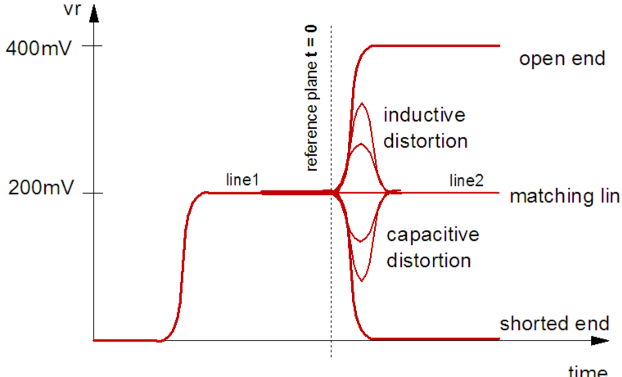


Figure: time domain reflectometry (TDR)

A connector between two lines is commonly seen by a TDR like a short overshoot followed by a short undershoot. We now know what this would mean for modeling: the overshooting represents an inductor L, while the undershooting hints to a capacitor C. The following figure shows such a TDR measurement.

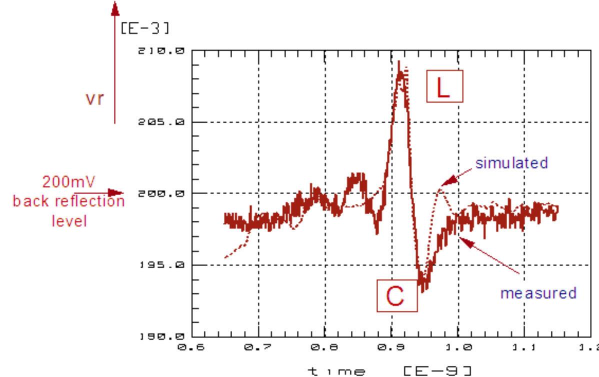


Figure: TDR measurement of a connector.

> **Note**
> The basic principles of understanding TDR measurements are summarized in the chapter entitled 'Interpreting TDR Measurements'.

Calibration Techniques

When measuring and modeling a packaged device in the GHz region, it must be placed in a test fixture. Therefore, when using a TDR, we have to do a calibration of the reflected signal first. This is done by the following two steps:

1. Model the step impulse of the TDR
   (either using the SPICE step function with the parameters slope, start and end value, or more accurately, using a sum of PWL function (piecewise linear) of the simulator)
2. Determine the length of the connection cable
   (modeled using a delay line)
   After that, we connect the cable to the test fixture, model it and, finally, model the component itself. The following figure illustrates these steps.

Because of the sequence of modeling steps, it is always very easy to determine the part of the equivalent schematic related to the test fixture or the DUT itself.

Figure: The individual steps to calibrate a TDR and to perform modeling measurements

TDR Calibration under IC-CAP

Examples about how to calibrate the TDR step function and the test cable in IC-CAP, can be found in the following IC-CAP model file
$ICCAP_ROOT/examples/demo_features/4extraction/specific_PEL_routines/tdr_cal.mdl

Publications

Advanced TDR techniques, Hewlett-Packard Application Note 62-3, Publication HP5952-1141, April 1990.

Evaluating Microstrip with Time Domain Reflectometry, Hewlett-Packard Application Note 1304-1, Publication HP5968-0007E, August 1998

Time Domain Reflectometry Theory, Hewlett-Packard Application Note 1304-2, Publication HP5966-4855E, May 1998

TDR Plots Tutorials

This chapter is intended to give you some ideas on how electronic components look like in the time domain.

Introduction

The characterization of parasitic devices can be done in the time domain or in the frequency domain. But time domain may be smarter, since the parasitic devices show-up with respect to their physical location. In the frequency domain, using s-parameters, we always have an overlay of curves that 'turn to the right' for capacitors, inductors or delay lines.

> **Note**
> For frequencies below 3GHz, a simple parasitic device schematic is often sufficient (two delay lines or capacitors at port1 and port2 and a cross-coupling capacitor, no inductors).

Using time domain, it is best using a time domain reflectometer (TDR) like the hp54xxx series oscilloscopes with a TDR plug-in on channel_1. A voltage step of typically 200mV is triggered by IC-CAP on channel_1 and the back reflections are then measured by channel_1 also. Using other channels of the TDR, one could also measure the cross-coupling from one port of the test object to another.

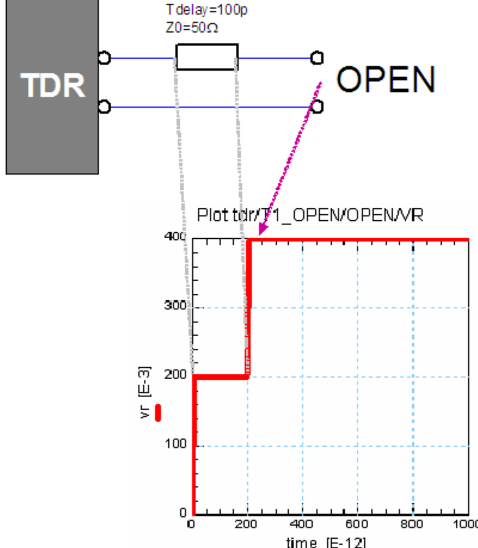
TDR measurements are best for the modeling of geometrically large or spread structures. As an example, we consider the modeling of the test fixture:
This can be done using delay lines, inductors (positive spikes in the back-reflected signal) and/or capacitors (negative spikes). Changes in the amplitude of the reflected signal are modeled using parallel (negative amplitude step) or series resistors (positive amplitude step). After the the measured curve has been fitted, the parameters of the test fixture parasitics are known and the 'inner' transistor can be de-embedded from S-parameter NWA measurements using Y-, Z- and/or S-matrix manipulations.

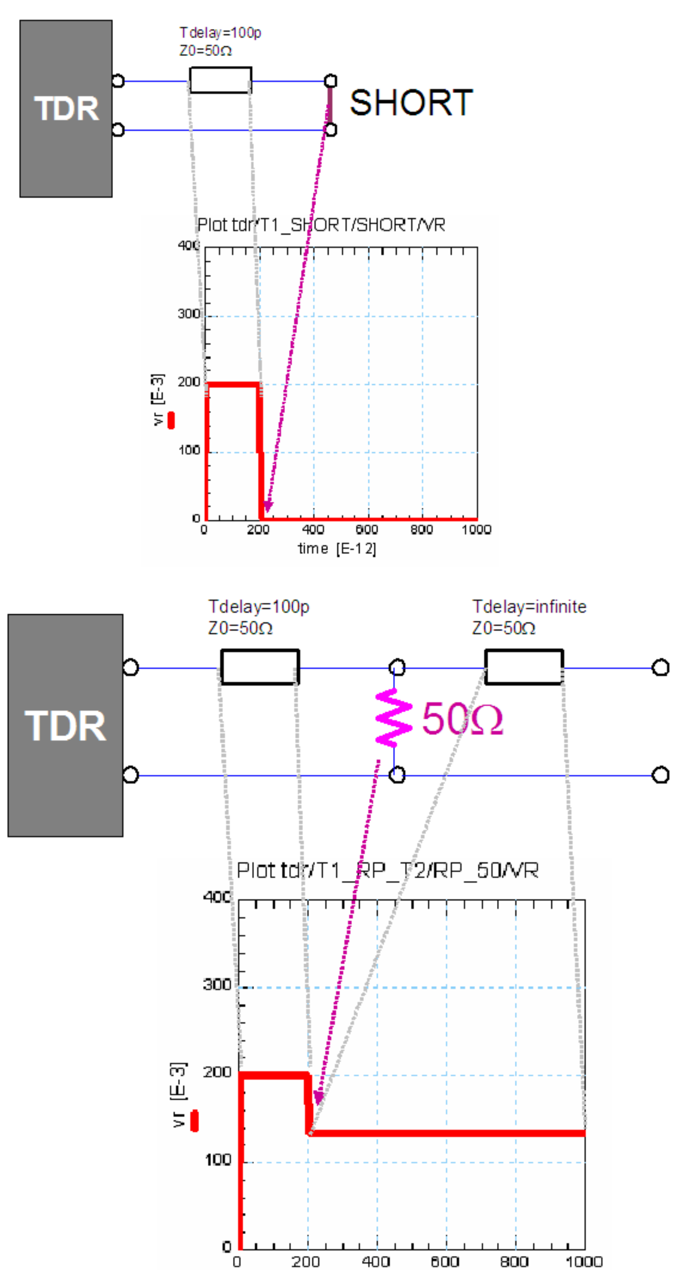
Expected Reflectograms of Typical Parasitic Components

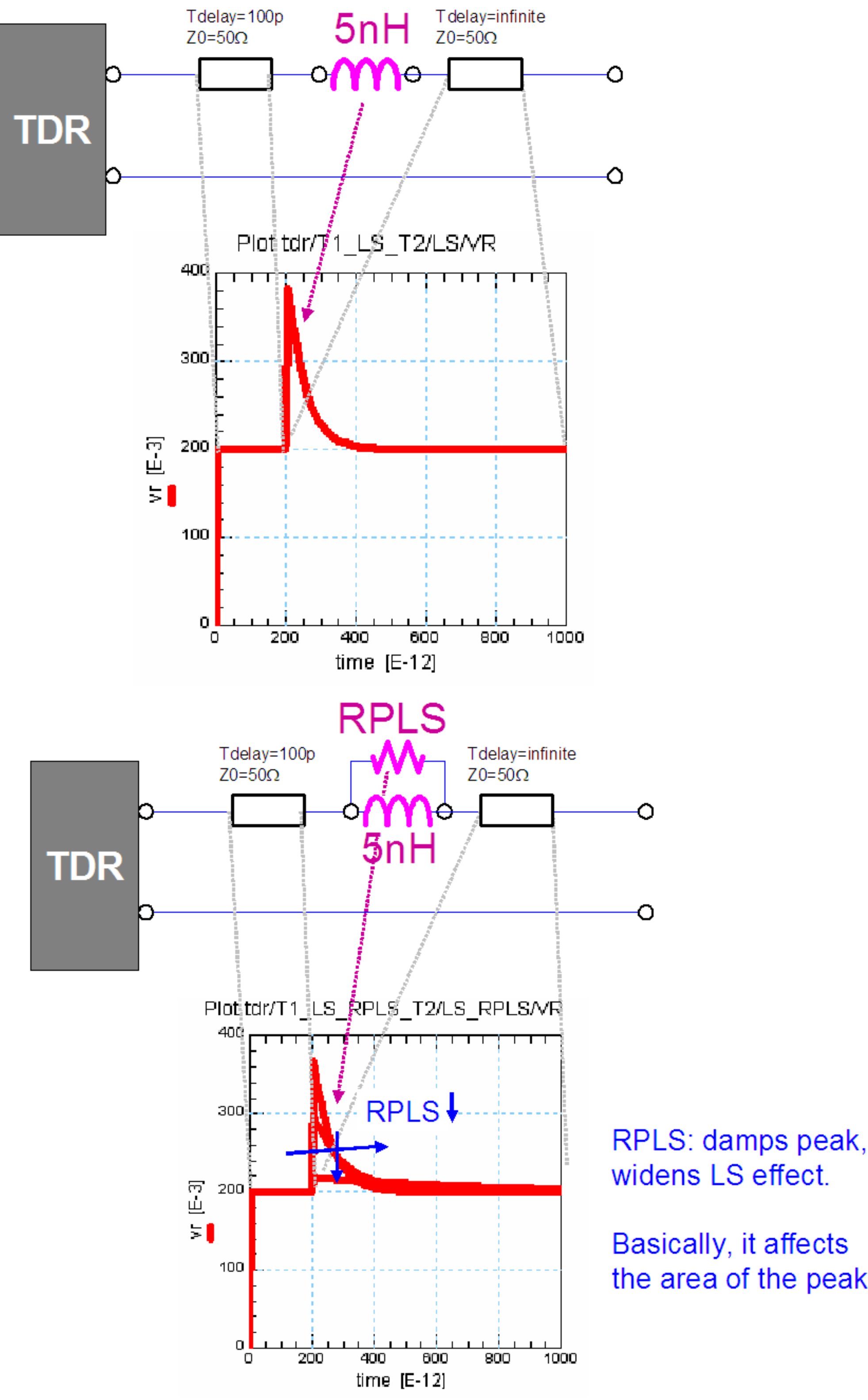
IC_CAP file: tdr_plots.mdl

The following plots are intended to help you with the development of equivalent circuits from reflectograms.

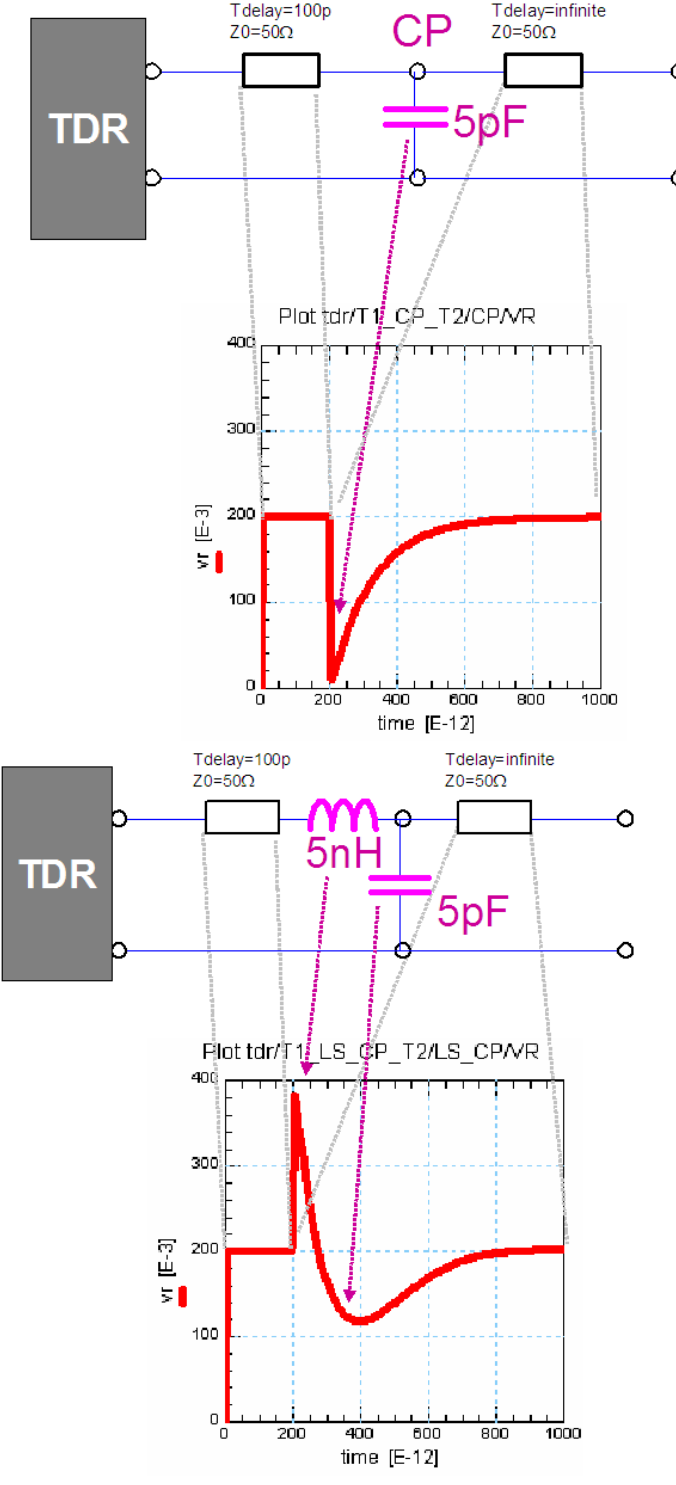
TDR Basics



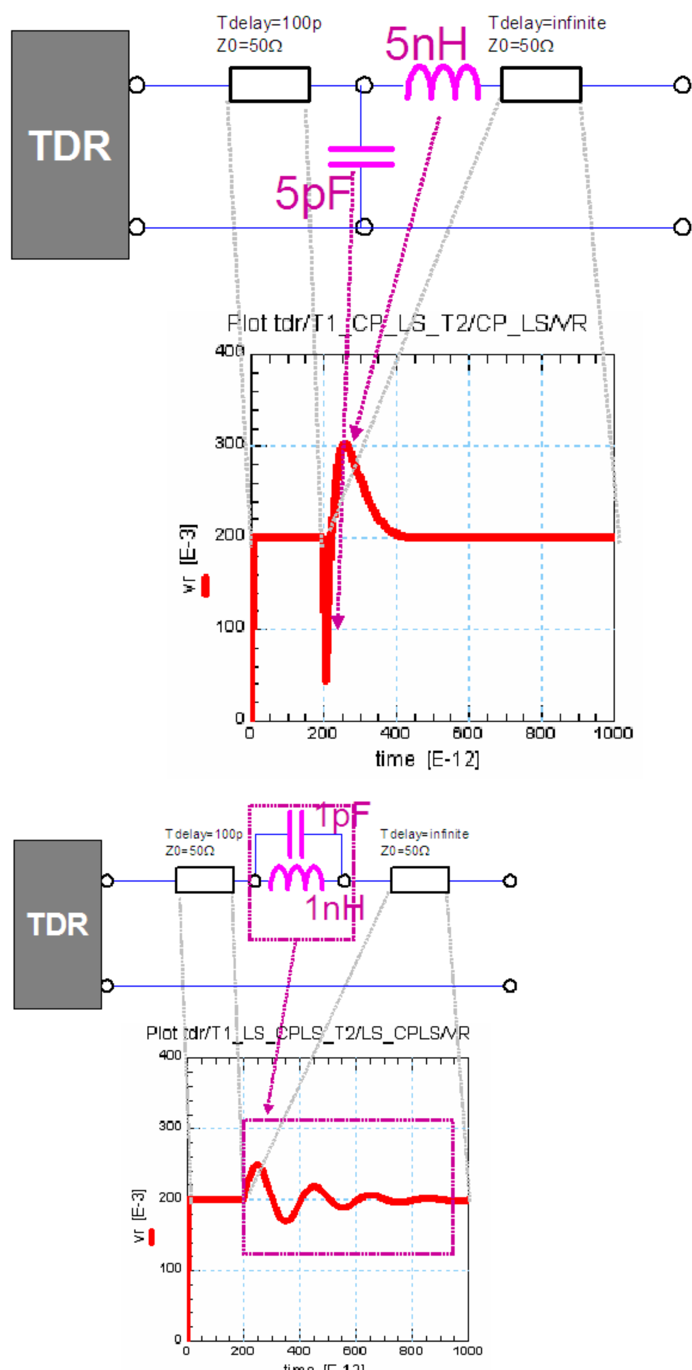

TDR
Tdelay=100p
Z0=50Ω
SHORT
Plot tdr/T1_SHORT/SHORT/VR
vr [E-3]
time [E-12]
TDR
Tdelay=100p
Z0=50Ω
Tdelay=infinite
Z0=50Ω
50Ω
Plot tdr/T1_RP_T2/RP_50/VR
vr [E-3]
time [E-12]



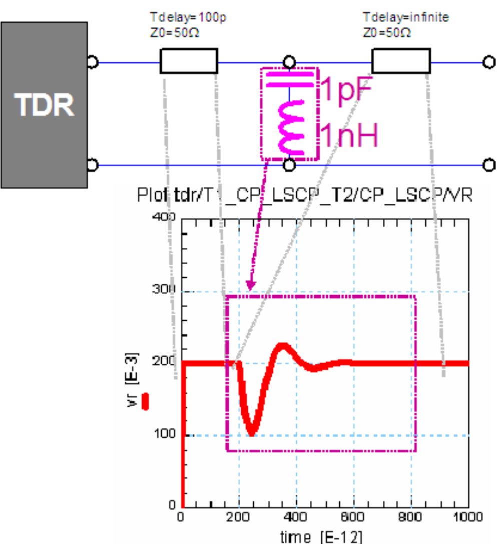
TDR
Tdelay=100p
Z0=50Ω
5nH
Tdelay=infinite
Z0=50Ω
Plot tdr/T1_LS_T2/LS/VR
vr [E-3]
time [E-12]
TDR
Tdelay=100p
Z0=50Ω
RPLS
5nH
Tdelay=infinite
Z0=50Ω
Plot tdr/T1_LS_RPLS_T2/LS_RPLS/VR
RPLS
vr [E-3]
time [E-12]
RPLS: damps peak,
widens LS effect.
Basically, it affects
the area of the peak



TDR
Tdelay=100p
Z0=50Ω
CP
Tdelay=infinite
Z0=50Ω
5pF
Plot tdr/T1_CP_T2/CP/VR
vr [E-3]
time [E-12]
TDR
Tdelay=100p
Z0=50Ω
5nH
Tdelay=infinite
Z0=50Ω
5pF
Plot tdr/T1_LS_CP_T2/LS_CP/VR
vr [E-3]
time [E-12]

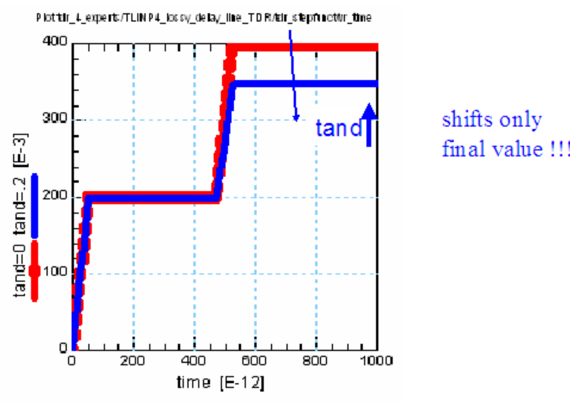


Tdelay=100p
Z0=50Ω
5nH
Tdelay=infinite
Z0=50Ω
TDR
5pF
Plot tdr/T1_CP_LS_T2/CP_LS/VR
vr [E-3]
time [E-12]
Tdelay=100p
Z0=50Ω
1pF
1nH
Tdelay=infinite
Z0=50Ω
TDR
Plot tdr/T1_LS_CPLS_T2/LS_CPLS/VR
vr [E-3]
time [E-12]



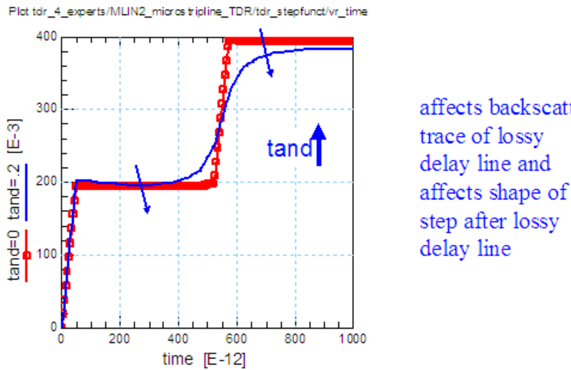
TDR Plots for Experts

ADS TLINP4 Lossy Delay Line (based on telegraph equation)

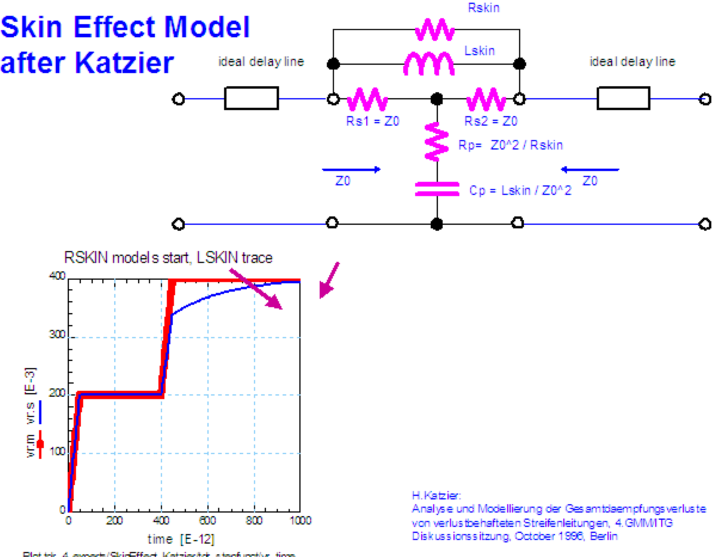


shifts only
final value !!!

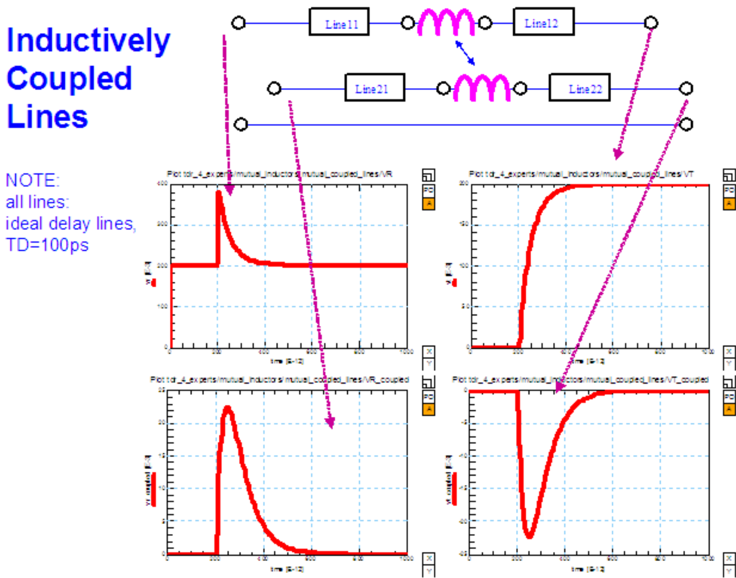
ADS MLIN2 Microstrip Line (with separate substrate model)



affects backscatter
trace of lossy
delay line and
affects shape of
step after lossy
delay line



See also the chapter 'Skin Effect' in the Modeling Handbook.



Note
Parallel strip lines, coupling across with mutual inductors (even with high coupling factors), are usually not much affecting the TDR measurement of the single strip line. But they can affect the corresponding S-parameter measurements heavily! So, for their modeling, either a THRU measurement of the cross-coupled signal using the oscilloscope input of the TDR will be necessary or additional measurements using a network analyzer.

Differential TDR Measurements

Since IC-CAP 5.4, also differential TDR measurement capabilities are available. This follows the growing demand for the modeling of the behavior of connectors, packages etc. when used with differential signals. Therefore, two new entries have been added to the Agilent 54750 Instrument table:

1. **Differential Mode:** Set the instrument in differential mode. Channel 1 and 2 are the TDR channels. The differential stimulus on channel 1 and 2 can be Differential (DIFF) or Common (COMM). Default is no differential stimulus (NONE).
2. **Differential Response Mode:** Once the instrument has been calibrated in differential mode, the response reading can be set Differential (DIFF) or Common (COMM). Note that this field is active only when the Normalize Flag of the response channels is set to yes. Default is DIFF.

To make TDR differential measurements place the Agilent 54754A plug-in in the first two instrument slots (channel 1 and 2). In the IC-CAP measurement page insert one input of type TDR (Unit TDR1 or CH1). Insert one input of Mode T (Type LIN) and set the time interval and the number of points. Insert two outputs of Mode V monitoring channel 1 and 2. In the 54750 Instrument Option Table, set the Differential Mode to DIFF or COMM.

To measure raw data simply set the Normalize flags of CH1 and CH2 to N and run the measurements. To measure normalized data, one needs to perform the TDR normalization before running the measurements. Follow the instructions in the 54754 manual on how to calibrate in differential TDR mode. Once the instrument has been successfully calibrated, set the Normalize Mode to TDR, set Differential Response Mode to DIFF or COMM. To measure the normalized response simply set the Normalize flag of channel 1 and 2 to yes.

Summary Differential TDR

-	Differential Mode	Differential Response Mode	Response Mode	CH1 Norm	CH2 Norm
Raw	DIFF/COMM	Not Relevant	Not Relevant	N	N
Norm	DIFF/COMM	DIFF/COMM	TDR	Y	Y

To make TDT differential measurements place one Agilent 54754A plug-in in the first two instrument slots (channel 1 and 2) and second 54754 plug-in in the third and fourth slots. When measuring differential TDT, the driver assumes that Channel 1 and 2 supply the differential stimulus (input). In the IC-CAP measurement setup page insert one input of type TDR (Unit TDR1 or CH1). Insert one input of Mode T (Type LIN) and set the time interval and the number of points. Insert four outputs of Mode V monitoring channel 1 to 4. In the 54750 Instrument Option Table, set the Differential Mode to DIFF or COMM.

To measure raw data simply set the Normalize flags of CH1,CH2,CH3 and CH4 to N and run the measurements. To measure normalized data, the user needs to perform the TDT normalization before running the measurements. Follow the instructions in the 54754 manual on how to calibrate in differential TDT mode. Once the instrument has been successfully calibrated, set the Normalize Mode to TDT, set Differential Response Mode to DIFF or COMM. To measure the normalized response simply set the normalized flag of channel 3 and 4 to yes.

Summary Differential TDT

-	Differential Mode	Differential Response Mode	Response Mode	CH1 Norm	CH2 Norm	CH3 Norm	CH4 Norm
Raw	DIFF/COMM	Not Relevant	Not Relevant	N	N	N	N
Norm	DIFF/COMM	DIFF/COMM	TDT	**N**	N	Y	Y

Noise

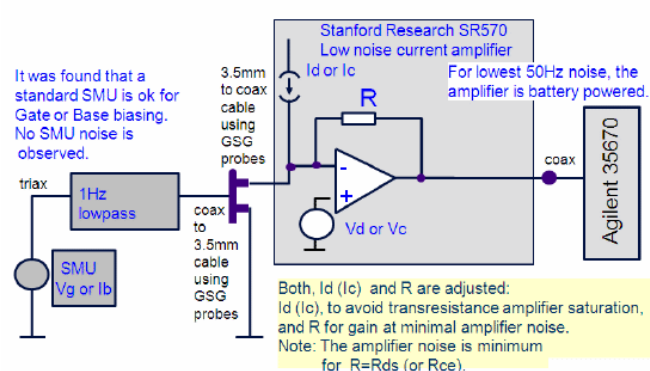
Contents

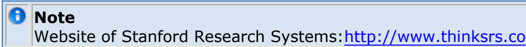
- *Noise Measurement Golden Rules* (iccapmhb)

Noise Measurement Golden Rules

Verifying the Measurement Data of the IC-CAP 1/f Noise Toolkit

The Applied 1/f Measurement Setup

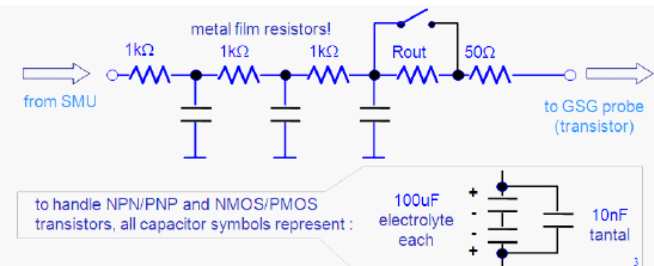



Note
Website of Stanford Research Systems: http://www.thinksrs.com

1Hz Lowpass Filter

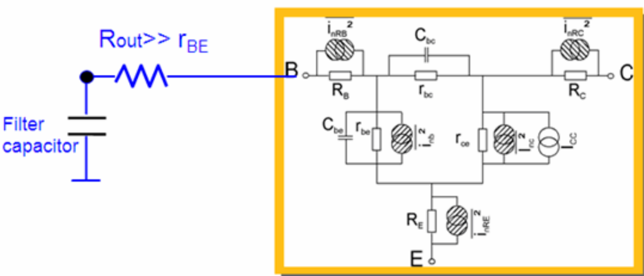
The filter is required for biasing the transistor input. It offers -60dB attenuation at 50Hz.
Required Output Impedance Rout (to avoid shortening the 1/f transistor noise):

- Rout =330 kΩ for bipolar (in any case Rout > Rbe_noise_bias),
- 50Ω for MOS transistors



In a bipolar transistor, the 1/f noise is generated in the B-E region. If the output resistance Rout of the 1Hz filter was smaller than the dynamic, iB-dependent value of rBE, it would AC-wise shorten the rBE (by the big capacitort in the 1Hz filter). Therefore, select Rout >>rBE. For MOS, Rout=50 Ohm is sufficient.

Importance of the Filter Rout for Bipolar Transistors



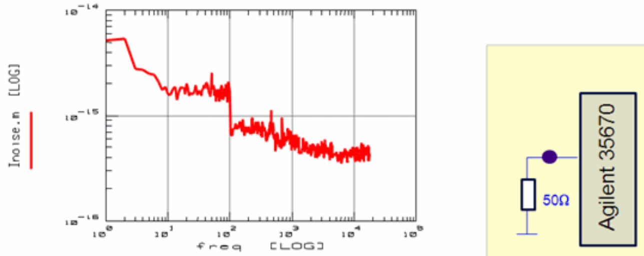
Since the filter capacitors represent a SHORT for the 1/f noisecreated mainly by rbe , Rout of the bias filter must be bigger than any rbe(vbe) during the 1/f noise measurements. Otherwise, the measured 1/f noise will be smaller than the real 1/f noise.

```
.subckt bip_noise 1=C 2=B 5=L 99=AUX
*satisfy SPICE for a noise simulation
Raux 99 0 1
*emulate the bias lowpass filter at the Base
Rfilter1  2 21 1k
Cfilter1 21  0 100u
Rfilter2 21 22 1k
Cfilter2 22  0 100u
*here is the filter output resistor  (high value for bipolar transistors)
```

```
Rout 22 23 330k
*here comes the call to the transistor in the Model file's Circuit tab
X1 11 23 0 noise_1_f_bip
*add a dummy 0V source to sense the transistor noise current
V1    1 11 0
*convert the Collector noise current into an identical voltage for
* further use in the modeling toolkit
H1      5 0 V1 1
R1      5 0 1
.ends
```

Hint
You can study the effect of Rout to a bipolar 1/f measurement also when varying the Rout in the simulation deck.

Checking the resolution of the Dyn. Sign. Analyzer 35670A



In other words, the device noise has to be amplified (by the SR570)above the resolution level of the 35670A.

The Low Noise Amplifier SR570

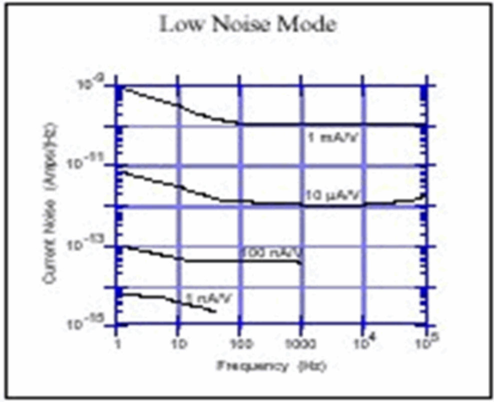
some hints (from the SR570 Manual)

1. The Instrument must be warmed up for about 60 minutes.
2. For best performance, the input current should produce an output voltage of about 1 V or less. This eliminates problems with slew rate limiting in the various amplifier stages.
3. Make sure the source impedance is greater than the inverse of the sensitivity (e.g. with a sensitivity of 1 nA/V use a source impedance greater than 1 GOhm).
4. Use short lengths of high quality coaxial cable to connect to the amplifier input.
5. Keep the amplifier output below 1 VRMS to avoid slew rate limiting at high frequencies.
6. If you want to ground the chassis, use the green connector on the back side, but do not connect the chassis to the amplifier ground (white connector).
7. For noise measurements, disconnect the power cord and use the internal batteries.

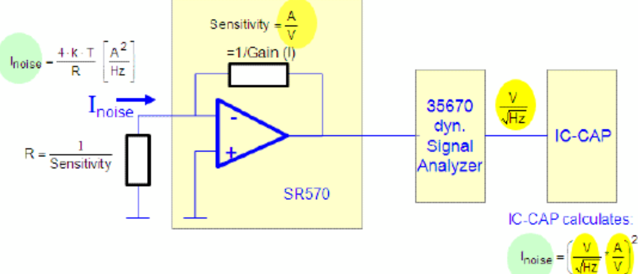
Note
If you experience problems with the 1/f toolkit measurement setup, apply an oscilloscope.
Simply connect the output of the SR570 amplifier to the oscilloscope.
Typical problems with 1/f measurement setup

- Do not use 1Hz filters with unipolar electrolyte-capacitors (instead of cascaded antiparallel electrolyte-capacitors). This may lead to very bad popcorn noise (!!) with frequencies at ~1/10Hz for negative biasing (PMOS), what lead to non-reproducible measurement results.
- Do not connected the voltmeter common (to check the SR570 voltage offset settings) to amplifier common and not to chassis ground of the SR570 either. (These connectors are at the rear side of the SR570). Instead, fully disconnect the voltmeter during the noise measurements.
- The Cascade wafer prober monitor may affect the measurement at several 10kHz: switch it off during 1/f measurements
- Like the 1Hz filter, the SR570 exhibits also a settling time of ~10sec, since it is operated in bandpath mode .
- The SR570 sensitivity (1/gain) should be >= Gout of the transistor. However, a too high sensitivity increases the SR570 noise contribution. For details, see the SR570 manual for the resolution plot "Current Noise as a function of Frequency for several sensitivity settings (typical).
- The offset current settings of the SR570 can add SR570-1/f-noise .
- In the same way, the voltage offset of the SR570 can add SR570-1/f-noise up to 1E-17 A^2/Hz. It is partly cancelled out with an appropriate offset current setting. Note: This was verified with resistor measurements, but may be different with transistor measurements.
- Recommendation: apply the SR570 current offset carefully, and if you are not sure, measure again with a different current offset.

Verifying the SR570 Measurements

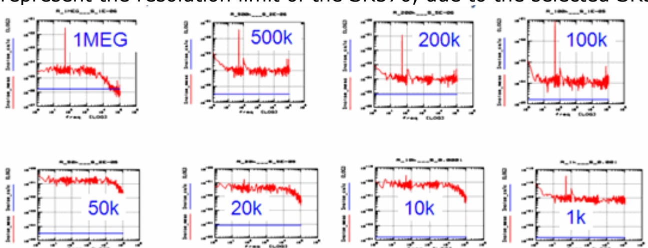


From the SR570 Manual:
Current Noise as a function of Frequency for several sensitivity settings (typical).
In order to verify the SR570 specs, a sequence of measurements
at different Sensitivities (SR570 amplifications) was performed using IC-CAP and the
Agilent 35670A. Both, voltage and current bias were OFF at the SR570 and the source
resistor was set to R=1/Sensitivity.

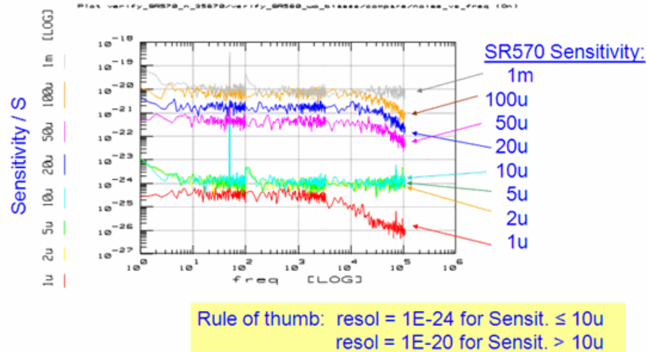


Measured resolutions for a sequence of SR570 Sensitivity Settings.
Sensitivity = 1/R $_{metal_film_resistor}$

In every case, the calculated metal film resistor's white noise level(straight blue line) was
well below the internal SR570 noise level. Therefore, the measured noise levels (red)
represent the resolution limit of the SR570, due to the selected SR570-sensitivity.



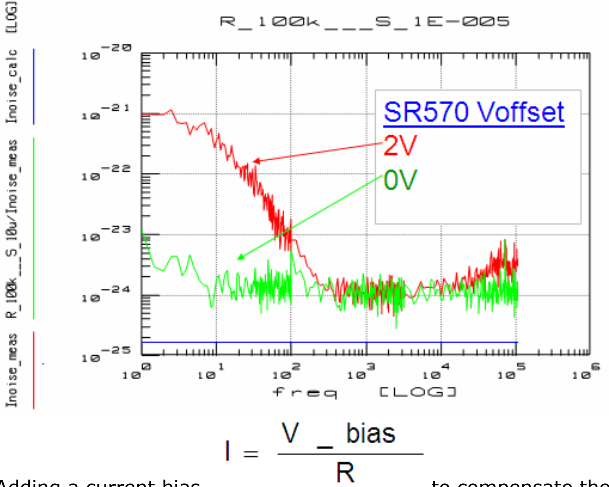
at a glance: SR570 resolution levels as a function of its Sensitivity (Gain)

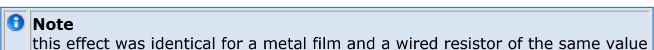


It was observed, that the SR570 bias voltage adds 1/f-like noise to the measurement.
But, that's an effect we have to live with. At Sensitivity=10u, a 100Ω resistor was
measured with and without 2V Voffset

Ioffset = 0

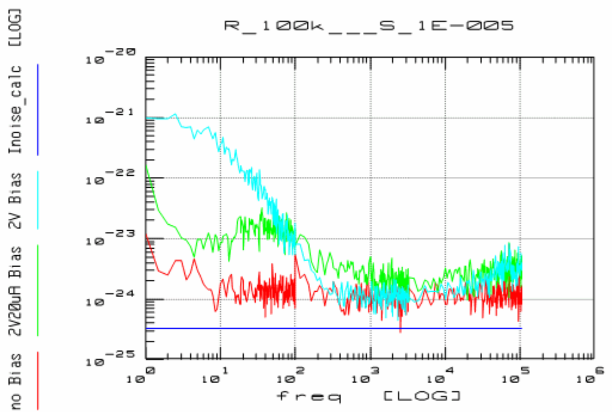
> This effect was identical for a metal film and a wired resistor of the same value. Therefore, the effect stems from the SR570 and not the source resistor.



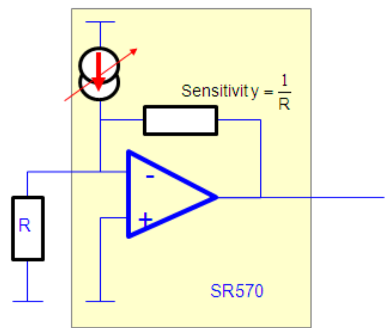
Adding a current bias  to compensate the resistorcurrent reduces the 1/f-like noise addition of the offset voltage,but does not fully compensate it.

At Sensitivity=10u, a 100Ω resistor was measured - without any offset- with 2V 0A Offset- with 2V 5uA Offset.

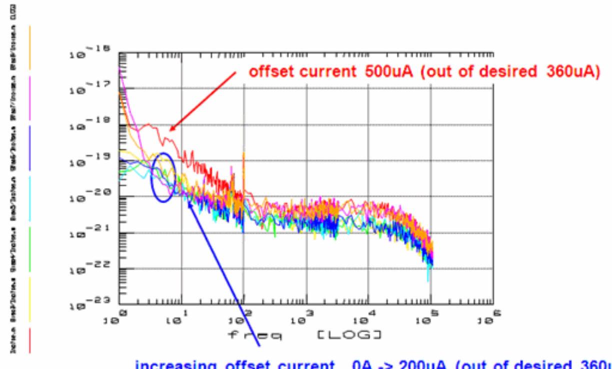
> **Note**
> this effect was identical for a metal film and a wired resistor of the same value



Measurement test sequence for fixed gain (Sensitivity) and varying current offset.

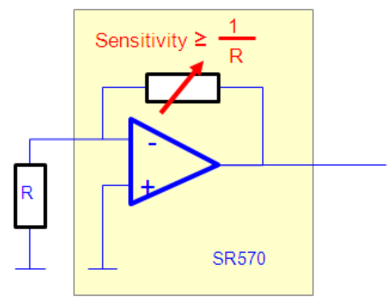


Measurement test sequence (PMOS Transistor)for fixed gain (sensitivity) but varying offset current [required offset current: 360uA].

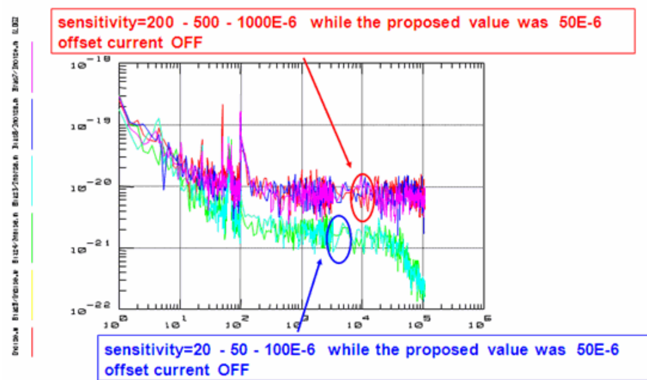


Observation: offset currents below desiredvalue do not add noise to the measurement system.
Measurement test sequence for fixed current offset and varying gain (Sensitivity).

Observation: the SR570 manual states $\text{Sensitivity} \geq \frac{1}{R}$ but when the Sensitivity is increased, the noise level of the SR560 increases too.



Measurement test sequence (PMOS)for offset OFF but varying gain (sensitivity)
[required sensitivity: 50uS = 1/Rout_PMOS]



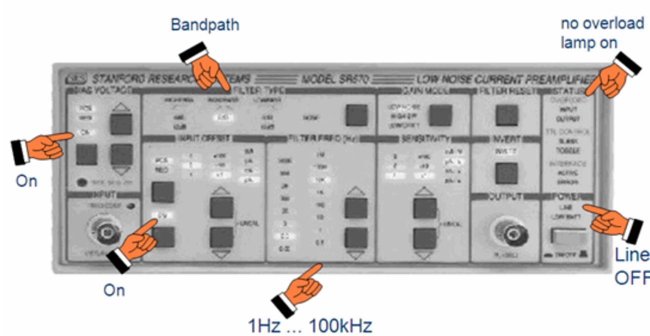
Observation: sensitivity (gain) above desiredvalue shifts the noise level of the measurement system
Conclusions for obtaining best results with the SR570 amplifier:
on the SR570,

1. Set bias offset voltage
2. Set current compensation
3. Set Sensitivity = 1/R4. make sure the SR570 output is above the 35670 noise floor otherwise increase the SR570 amplification (smaller value of the SR570 sensitivity setting)
4. If overflow lights up, adjust the current compensation

Note
If you are not sure whether you measure the DUT's noiseor the SR570's noise, vary the sensitivity, the voltage and current bias and check the effect to your measurement.

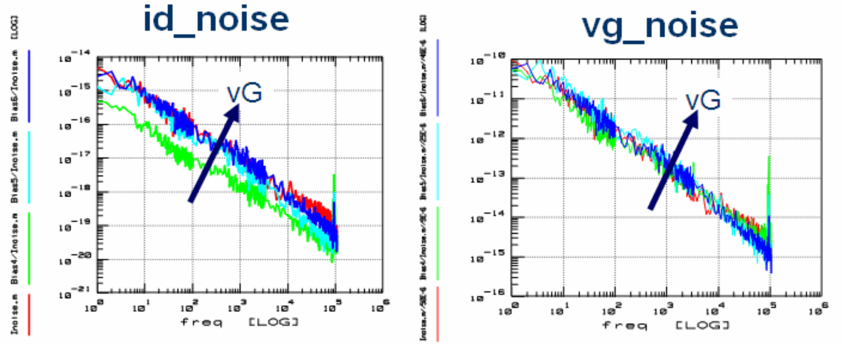
Best Practice Notes, March 2006: Falk Korndörfer

Always apply lowest sensitivity (highest SR570 amplification) for best resolution (1E-23 A^2/Hz). Adjust bias compensation to avoid overload alarm.
Last not least, do not forget to check these SR570 instrument settings:



A Note on Data Consistency Checks for MOS Transistors

For MOS-Transistors, vg_noise is independent of the applied bias. Therefore, calculating vg_noise=id_noise/gm^2 should give overlying curves .



- See: M.T.Yang et.al., Characterization and Model of On-Chip Flicker Noise With Deep N-Well Isolation for 130nm and Beyond,ICMTS 2005, Leuven, Conference Proceedings.
  Acknowledgements:
  special thanks to Edel Griffith of Analog Devices in Limerickand to Falk Korndörfer at IHP-Microelectronics/Frankfurt-Oderfor many interesting discussions !

Appendix

Parts List

Small Parts w/o Supplier

- 1ea 1Hz Filter with switch bipolar-MOS, details see in the slides above
- 1ea test fixture for packaged MOS or bip. transistors (to verify the system performance)

The Following Equipment Is Usually Available At The Customer:

- 1ea triax-triax cable 80cm (SMU->Filter)
- 2ea short Coax->3.5mm cables (Filter->GSG probes, GSG probes ->SR570)
- 1ea short Coax-Coax cable (SR570->HP35670A)
  optional: 2ea Triax(m)->Coax(f) adapters, Trompeter part number ADBJ20-E3-PL75

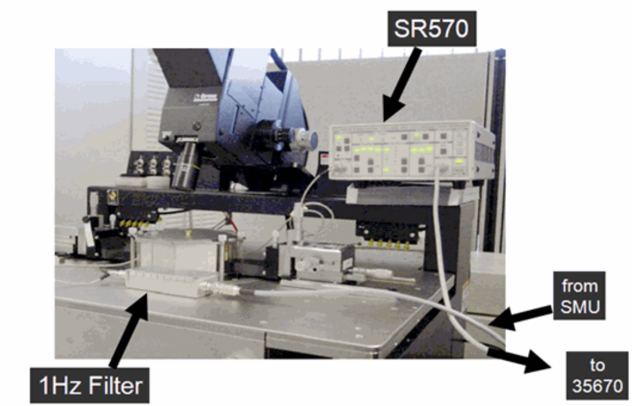
Measurement Instruments

- 1ea Stanford Research Amplifier SR570 (http://www.srsys.com)
- 1ea Agilent 4142 or 415x DC analyzer, other Agilent DC analyzers upon request
- 1ea 35670 dynamic signal analyzer (no special options are required)
  Note: when using a new 35670, make sure that it is NOT set to system controller!

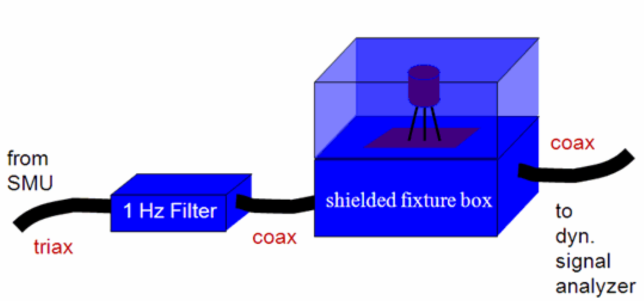
IC-CAP Modules

- 1ea 85199G IC-CAP Noise Measurement Drivers
  1ea 85199A,B,D or 85190A bundle (typically already available at the customer)
- 1ea 85195BL 1/f noise toolkit (includes 1 Hz filters, however, only applicable for NPN and NMOS transistors)

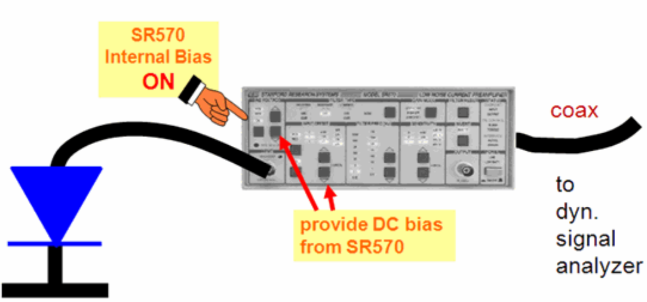
Wafer Prober Setup



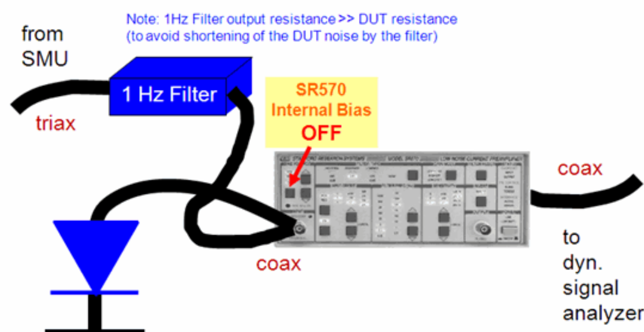
Meas.Setup for Packaged Transistors



Meas.Setup for Resistors, Diodes etc.



> **Note**
> Special acknowledgements to the modeling engineers at ON-Semiconductors in Roznov/Czech Republic, for this smart application.



How to fine-tune the SR570 offset current settings

Important copy from the SR570 Manual

Setting The Input Offset Current

The SR570 can provide a DC current offset to suppress any background currents at the input.
The offset range can be changed from 1 pA to 5mA (both positive and negative) in

discrete increments. Use the up/down arrow keys in the Input Offset section to change the current level.

In addition to these fixed settings, the user may specify arbitrary currents through the UNCAL feature. To set an uncalibrated offset current, the user must press both up and down buttons simultaneously, lighting the UNCAL LED. In this mode, by pressing the up or down pushbuttons, the user may reduce the calibrated current in roughly 0.1% increments from 100% down to 0% of the selected offset value. In contrast to other front-panel functions, when in UNCAL the instrument's key-repeat rate will start slowly and increase to a limit as long as either button is depressed.

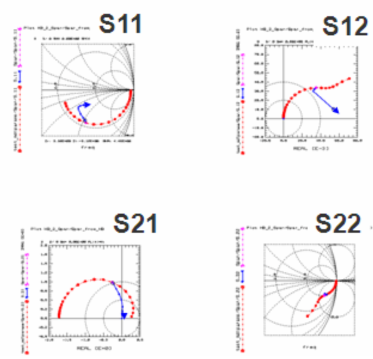
Simultaneously pressing both Offset buttons once again will restore the unit to the previously calibrated current setting, and turn off the UNCAL light

The same applies to the Sensitivity settings.

Spectrum Analyzer

Data Verification

Data Verification and Data Consistency Checks



Prologue

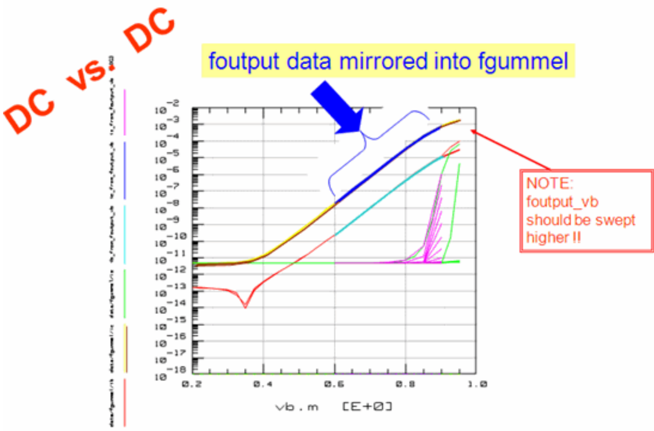
Besides good measurements, the measured data should be consistent from Setup to Setup. Therefore, before performing the first parameter extractions, the data consistency should be proofed.

Possible Checks

- DC data among themselves (e.g. output vs. transfer chacteristics)
- S-parameter bias points vs. DC measurements
- S-parameter starting points vs. simulations of fitted DC parameters
- CV data against S-parameters (which have been converted to CV)

and, using ADS, also linearity of S-parameter measurements

DC vs. DC

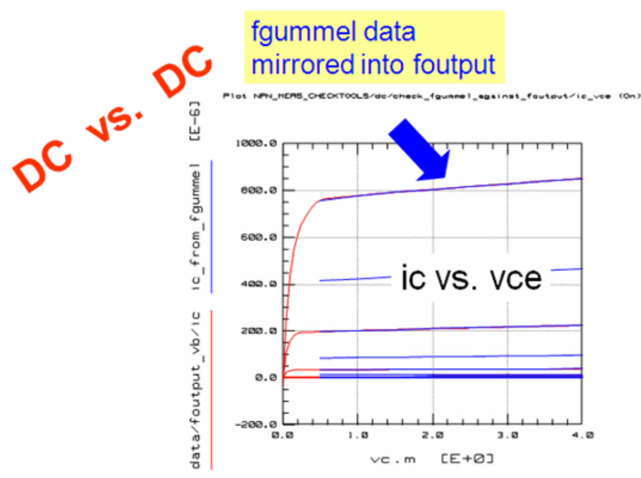


In this example, an IC-CAP macro has converted the currents of the foutput measurement of a bipolar transistor into a pseudo-measured Gummel plot .

IC-CAP Example

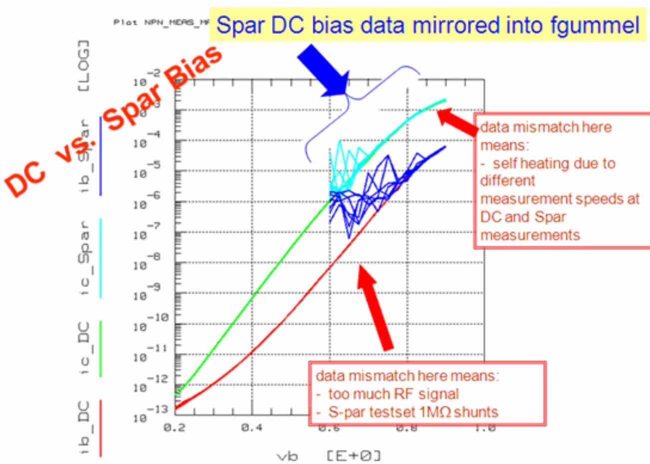
1. Select the demo_features directory
2. Goto sub-directory 3_MEAS_ORGANIZE_n_VERIFY_DATA/0_MASTER_FILES/1_CHECK_DATA_CONSISTENCY
3. and load the model file NPN_MEAS_CHECKTOOLS_demodata.mdl and execute the Transforms 'check_data_consistency' in the different Setups.

Note
The IC-CAP macro applies some data management manipulations like-> save foutput data to an ASCII .mdm file-> read the data back into IC-CAP, but now with inverted sweep orders. This gives a pseudo-measured Gummel plot. Link the data of the original Gummel plot into the same plot.



Of course, DC-DC data checks can also be performed in the other way:In this example, the forward Gummel plot data are converted into pseudo-measured foutput data, and plotted together with the originally measured foutput data.

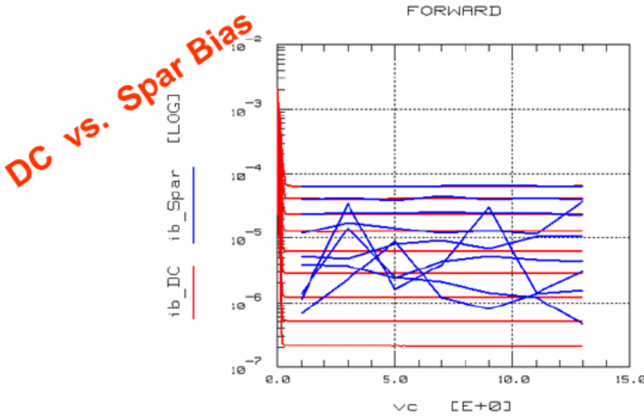
DC vs. Spar Bias



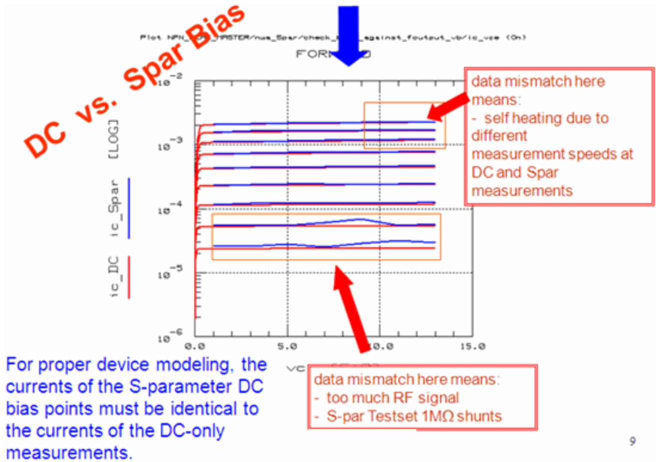
In the next step, we compare the DC bias conditions of the S-parameter measurements with the DC measurements.

> **Note**
> How the IC-CAP macro manipulates the data, export the S-parameter data including the also measured DC bias currents-> import back only the DC bias currents-> compare to the originally measured DC curves.

DC Measurements vs. Spar DC Bias



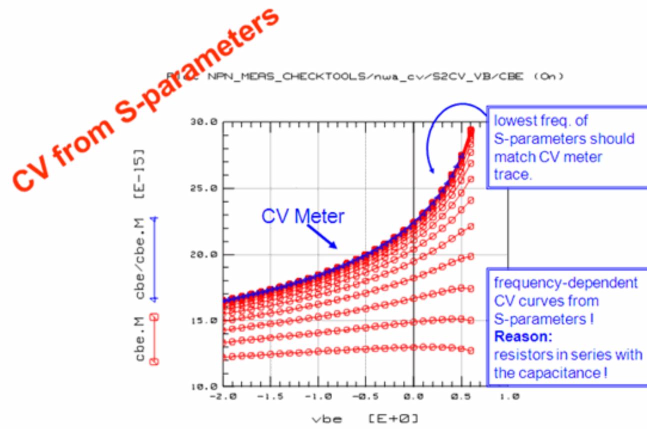
Spar DC bias data mirrored into foutput



To make the test complete, the DC foutput measurement and the DC biases of the S-parameter measurements are compared too. If there is a difference in the DC currents of the S-parameters and the standalone DC measurements:

- Too much RF signal, too small DC bias power (compared to applied RF signal power)
- Self-heating
- Voltage drop in S-par testset, 1MOhm shunt resistors to Ground in the S-par testset
- Unphysically modeled DC performance
- DC modeling was performed at different bias conditions than S-par modeling

CV vs. Spar converted to CV



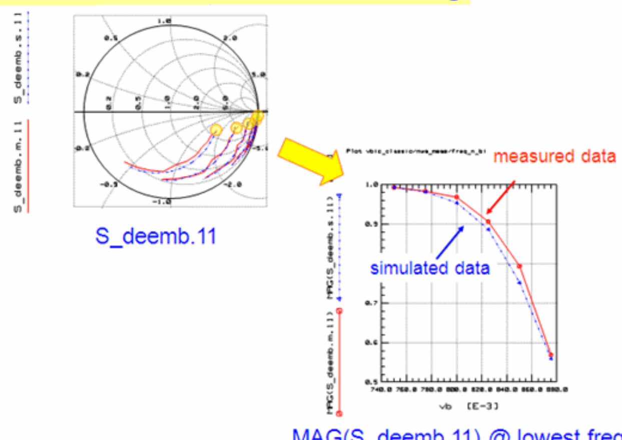
We are now ready to check the data consistency for the frequency range. Commencing with CV, the example above shows the comparison of CV data obtained with a standard LCRZ meter, with pseudo-CV data obtained from converted S-parameters.

Note
Here a quick overview of the IC-CAP macro:

1. De-embed the measured S-parameters
2. Convert the de-embedded S-parameters to Y-parameters
3. Calculate the capacitance out of the imaginary part of the Y-parameters
4. Compare these pseudo-CV data with the original CV meter data.

When comparing CV curves converted from S-parameters, keep in mind that due to the S_to_Y conversion and the interpretation of these Y-parameters with respect to a PI schematic, the CV curves represent the capacitance in one of the PI branches. Measuring with a CV meter, depending on the measurement principle, you may obtain the total capacitance between two nodes (2-pin CV measurement method) or the individual capacitance (4-pin method, guarded wafer chuck).

Spar Starting Points vs. Existing DC Modeling

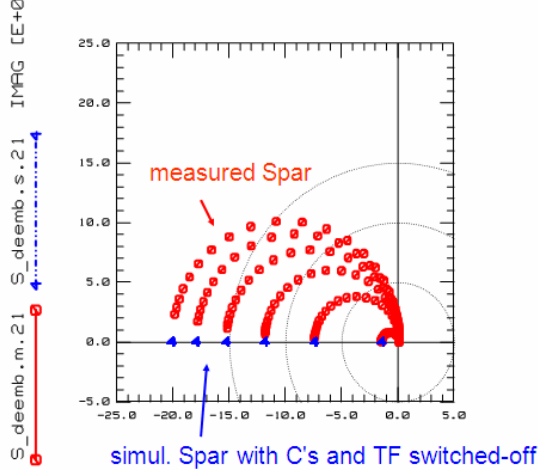


Since the starting points of the simulated S-parameters, or better said, the extrapolated points for 0Hz frequency, are determined by nothing but the DC parameters, we should be able to get a good fit for the low-frequency range even when we haven't yet extracted the high-frequency parameters like the capacitances or the transit time of transistors.

Important Note and Recommendation
In your daily modeling work, you should perform this test every time you are done with fitting the DC (and CV) curves and begin with S-parameter modeling.

The starting points of the measured S-par (0Hz)are determined by the DC fitting

Example Transistor
With RF modeling (capacitors and transit time), only the trace from the starting points towards the end points can be fitted, but not the starting points.
If this starting point check fails, verify this: too much RF signal> self-heating> voltage drop in S-par testset> DC modeling was performed at different bias conditions than S-par modeling.

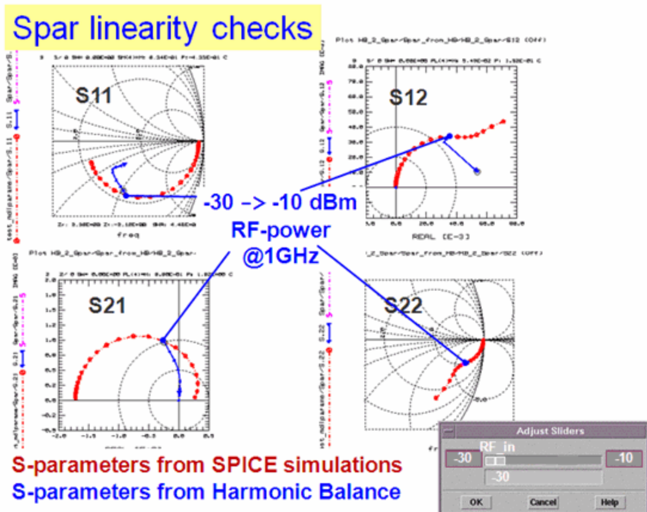


Discussion
If the starting points of the S-parameter curves do not match the simulations with the so far extracted DC parameters:

- Too much RF signal-> self-heating
- Voltage drop in S-par testset
- Unphysically modeled DC performance

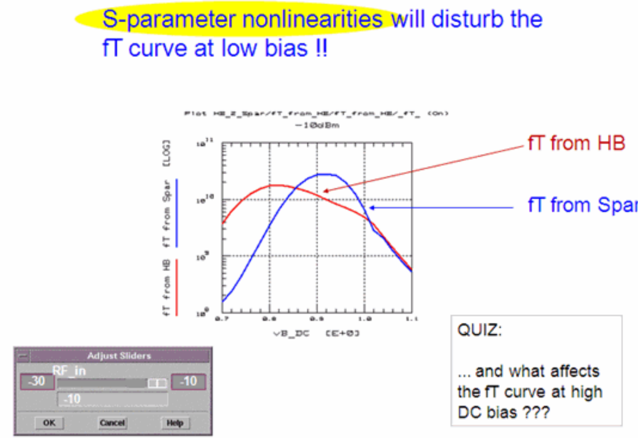
Solve the problem before you continue with S-parameter modeling. The S-parameter points for 0Hz can be extrapolated from the two lowest measured S-parameter frequency points.

Spar linearity checks

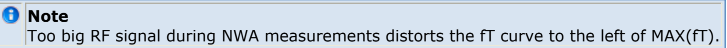


Last not least, regarding our consistency checks, we can even verify the linearity of the performed S-parameter measurements with the NWA.
Provided ADS is available, we can enter the so far extracted model parameters, perform an harmonic balance simulation for a certain fundamental frequency, and convert the simulation result into 'S-parameters as a function of RF power'. (This is done by converting the fundamental frequency simulation results into S-parameters. We compare this result with a linear simulation (what emulates the linear NWA measurement).
We increase the applied RF signal level from e.g. -40dBm up to -10dBm, and check, if the pseudo-S-parameters from the harmonic balance simulation stay constant, or if they begin to become a function of applied RF power. As long as they are independent of the RF power, the device behaves linear. This way, we can backwards check the max. applicable RF signal for the NWA measurement and the NWA calibration.

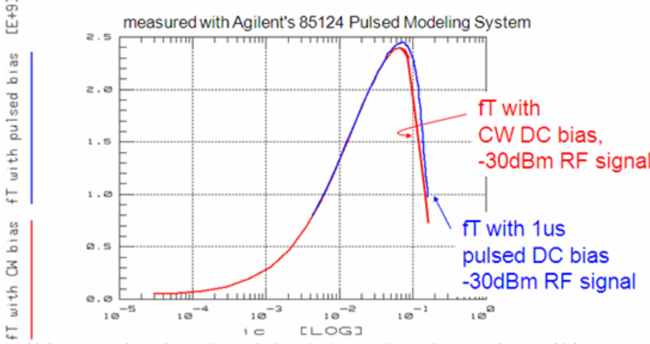
Example in IC-CAP:
load the file
demo_features/3_MEAS_ORGANIZE_n_VERIFY_DATA/0_MASTER_FILES/4_RF_LINEARITY/large_signal_RF_NPN_PELdep.mdl
and execute Transform 'TUNE_ANALYZE_xxx' in the harmonic balance Setups.



When S-parameters are affected by the RF signal, then fT is also affected.

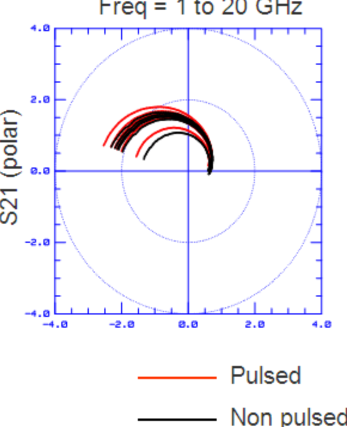

Note
Too big RF signal during NWA measurements distorts the fT curve to the left of MAX(fT).

It is thermal self-heating what affects the fT curve at high DC bias, even FOR low RF signal levels.



Self-heating 'pushes down' the ft for IC bias beyond max(ft).

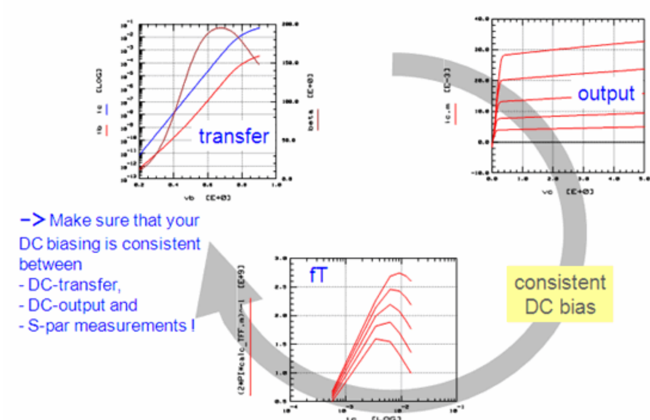
And here's the answer to the previous slide's quiz:
as a general rule, for iC bias conditions beyond max(ft), it is the thermal self-heating what pushes down the ft points towards lower ft values. Depending on the type of transistor, also the max(ft) can be pushed down by self-heating.
Self-heating affects the S-parameters (shift of the operating point).



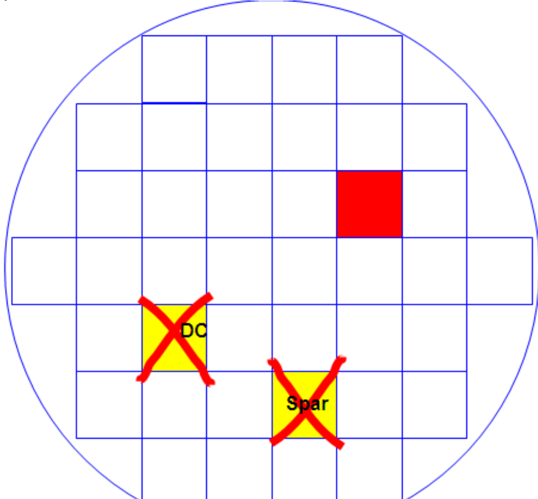
The S21-parameters in this slide depict a transistor measured in pulsed mode (1us pulse for DC and NWA) compared to a conventional CW measurement. This data is the base for the fT curves of the previous slide.

A Typical Modeling Mistake

fT modeling at different (higher) iC than DC modeling



Last not least: don't merge diff. measurements of diff dies when extracting model parameters.

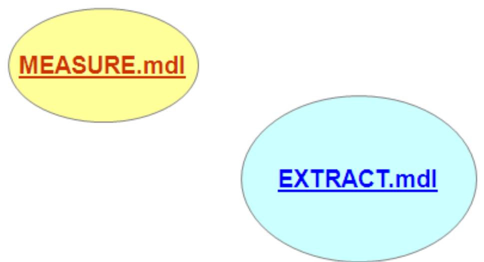


It is obvious that when developing a model of a component, all types of measurements have to be performed on the same device.

Data Management

Measurement Standardization and Data Consistency Checks

The Concept of Separating Measurements From Extractions



Separate measurements from extractions

- 1 MASTER MEASUREMENT file
- 1 MASTER EXTRACTION file

Data Exchange by mdm Files

The separation of measurements from extractions requires as a first step to use 2 model files:- one for the measurements- another for the extractions.
Both files act as strategy files. The measurement file contains the measurement strategy, and the extraction file the modeling and parameter extraction strategy.

Therefore, both files can be saved without measured or simulated data!
The measured data are stored in separate .mdm files instead, and the extracted model parameters in SPICE, ADS, Spectre, ELDO etc. decks.

Required Naming Conventions for the concept of separated measurements and extractions

As a consequencs, it becomes mandatory to standardize on some naming conventions. Otherwise, the data exchange by exported (Measurement_Master.mdl) and then imported (Extraction_Master.mdl) .mdm files will not work.
Diodes: va, vc, ia, ic (anode - cathode), additionally for varactor diodes: vp, vw (poly - well)
Inductors: see under S-parameters
Transistors DC and CV:
MOS and MESFET:

- DC: vg, vs, vd, vb, ig, is, id, ib
  BIP:
- DC: vb, ve, vc, vs, ib, ie, ic, is
- CV: vbe, cbe, etc. with the first index representing the + node, the second the - node

S-PARAMETERS: freq, S (undeembedded data), S_deemb (deembedded data) dummy1, dummy2 for 0V dummy Inputs (to satisfy SPICE)
MEASUREMENT_MASTER <> EXTRACTION_MASTER: naming conventions

With this concept, it becomes mandatory to define some naming conventions. Otherwise, the data exchange by exported (MEASUREMENT_MASTER.mdl) and then imported (EXTRACTION_MASTER.mdl) .mdm files will not work.
For the data import, an ImportCreate is usually performed, which means that the names of the Inputs and Outputs in the mdm files become the names of the Inputs and Outputs of the corresponding master_extraction files. This means, the transforms in this file **have** to use these names.

Suggestion for standardized Input and Output names::
DIODES (DIODE_MEAS_MASTER.mdl) va, vc, ia, ic (anode - cathode)

VARACTOR DIODES (MOS_VARACTOR_MEAS_MASTER.mdl) additionally to diodes: vp, vw (poly - well)

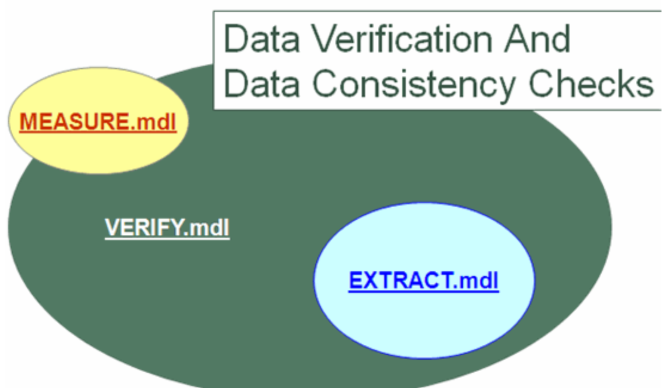
INDUCTORS (SPIRAL_MEAS_MASTER.mdl) see under S-parameters

TRANSISTORS DC and CV

- MOS (BSIM4_DC_CV_MEASURE.mdl) and MESFET: DC: vg, vs, vd, vb, ig, is, id, ib
- BIP (NPN_MEAS_MASTER.mdl):
  - DC: vb, ve, vc, vs, ib, ie, ic, is
  - CV: vbe, cbe, etc. with the first index representing the + node, the second the - node

S-PARAMETERS freq, S (undeembedded data), S_deemb (deembedded data) dummy1,

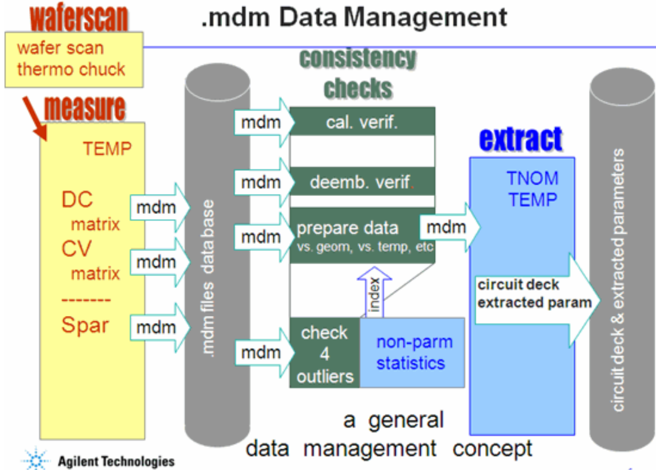
dummy2 for 0V dummy Inputs (to satisfy SPICE)
In order to re-use measurements in different model extraction files, it is important to agree on these naming conventions. So, once the data are measured, the modeling engineer can select the model, without remeasuring the device.



When separating measurementsfrom extraction,we can easilyadd a 3rd stepfor important datainspections and data consistency checks:the VERIFY.mdl

- verify measurement data
- rearrange measurement data (vs. L, W, TEMP, wafer location..)

Taking also the most important data consistency checks into account, requires an intermediate model file layer, i.e. a third model file.



In this VERIFY section, we check for data outliers, scaling effects, network analyzer calibration verification and de-embedding verification etc. We further verify the consistency of the measured DC currents with the DC bias currents of the S-parameter measurements.
It is also in this intermediate layer where non-parametric statistics could be applied to identify the measurement data of the 'golden device' and the 'boundary devices' for the subsequent parameter extraction in the model file EXTRACT.mdl

Going more into details, this slide shows the realization of the previously presented scheme, using

- a master model file for wafer and thermochuck control
- a master model file for performing the measurements
- mdm files as the measurement data base,
- several model files for data verification
- non-parametric statistics for golden device identification and finally
- the master model file for extraction.

Project Flow Suggestion for Reusable Model Files

This is an example of a complete project with measurements separated from extractions, applying data consistency checks, plus re-use of centralized PEL programs (DEPOTS).
With this concept, the measurement data are stored in mdm files,while the model files are all saved without measurement data!

> IC-CAP examples file: NPN_MEAS_MASTER.mdl
> NPN_MEAS_CHECKTOOLS.mdlVBIC_EXTRACT_NPN.mdl
> CAL_VERIFY.mdl
> CHECK_DEEMB_OpenShortThru.mdlcombine_mdms.mdl

In this example, we will discuss the use of a centralized PEL program called 'DEPOTS' plus the separation into a MASTER_MEASUREMENT model file, several MASTER_DATA_CONSISTENCY_CHECK files and a MASTER_EXTRACTION model file.
The project flow for the above sketched scenario is as follows:

1. NPN_MEAS_MASTER (MASTER_MEASUREMENT)
   This is a model-independent, but device specific file for performing all required modeling measurements plus the additional documentation measurements like contact resistance, NWA calibration quality, dummy structure S-parameters etc.During this task, there is an

- Export of the NWA calibration verification measurements, i.e. the (re-)measured OPEN, SHORT, LOAD and THRU calibration standards after the NWA calibration into model file CAL_VERIFY
- export of the S-parameter dummy structure measurements into model file CHECK_DEEMB_OpenShortThru
- call to re-usable central de-embedding PEL routines in model file DEPOTS.

1. Then, we export the measurement data into a mdm file data base
2. We import the mdm files into model file NPN_MEAS_CHECKTOOLS where for example the DC curves are compared to the DC bias curves of the S-parameters (identifying too big RF signal, self-heating, contact problems etc.), or where CV measurements are compared to S-parameter measurements. It is also here where special scaling effects are analyzed. For example, for MOS transistors, .mdm files of the ransconductance vs. L and W are composed here for later VTH extraction. Also, outliers in the .mdm data files are identified and eliminated in this step.
3. Finally, we import the mdm files finally into model file VBIC_EXTRACT_NPN. This VBIC extraction re-uses central diode modeling extractions (like for IS, N, CJO, VJ and M) for VBIC from model files DEPOTS.

> **Note**
> Because of the completeness of the measurements in the NPN_MEAS_MASTER file, we also could import the data to any other bipolar MASTER_EXTRACTION model, Gummel-Poon, VBIC or Hicum. With this concept, the measurement data are stored in mdm files, while all the model files are all saved without measurement or simulated data. This is why there is a small button called 'Save without Measured/Simulated Data' when saving model files. Besides the discussed project flow for the bipolar toolkits for Gummel-Poon*, VBIC*, and Hicum***, this concept is also used in the IC-CAP BSIM and PSP toolkits*, as well as in the toolkits for spiral inductors**, varactor* and diode* modeling.

IC-CAP examples files of the discussed bipolar project flow:

- NPN_MEAS_MASTER.mdl
- DEPOTS.mdl
- cal_verify_MASTER.mdl
- check_deemb_OpenShortThru.mdl
- NPN_MEAS_CHECKTOOLS.mdl
- combine_mdms.mdl
- GP_EXTRACT_NPN.mdl
- VBIC_EXTRACT.mdl

Contacts:

- Franz Sischka, Agilent EEsof, franz.sischka@agilent.com
  - www.xmodtech.com
    - www.admos.de

Examples for Data Consistency Checks

BEFORE MODELING:

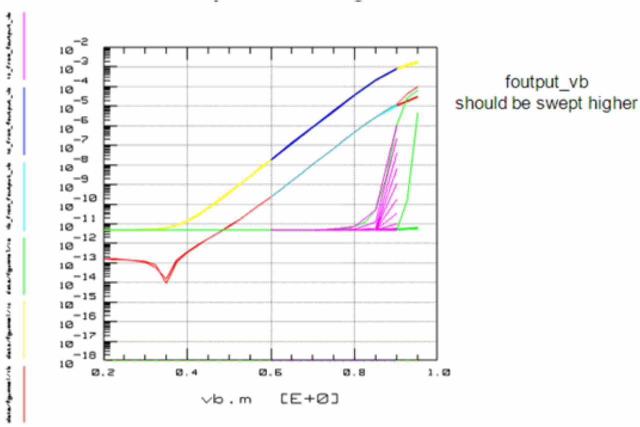
- cross-checking for DC bias conditions
- verification of NWA calibration
- verification of de-embedding
- verification of S-Parameter bias currents vs. DC curves
- comparing CV measurements versus CV from S-Parameters

AFTER MODELING:

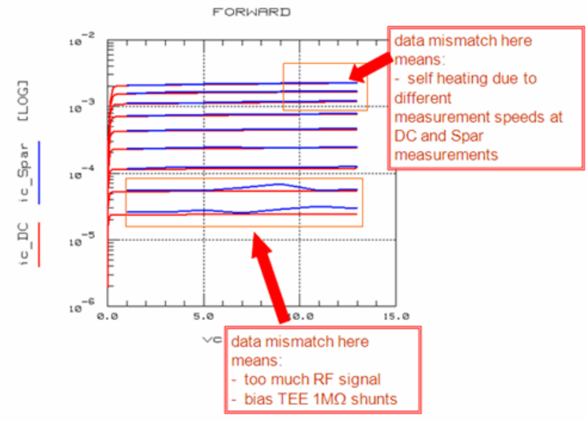
- check nonlinear RF performance of obtained parameter set
  IC-CAP examples file:
  Before modeling:
  demo_features\3_MEAS_ORGANIZE_n_VERIFY_DATA\0_MASTER_FILES\10_CHECK_DATA_CONSISTENCY\NPN_MEAS_CHECKTOOLS.mdl

After modeling:Verification of the linearity of the NWA measurements (after the DC model parameters have been extracted):
demo_features\3_MEAS_ORGANIZE_n_VERIFY_DATA\0_MASTER_FILES\40__RF_LINEARITY\
large_signal_RF_diode.mdl
large_signal_RF_xtors.mdl

Example: DC vs. DC Measurements mirror foutput into fgummel



Example:DC Measurements vs. Spar DC Bias



Advantages of Standardization

- no double PEL programs any more
- ease of use
- clear structure of strategies
  However...
  it takes some effort to convert existing 'hacked' solutions into standardized, reusable modeling schemes

Extraction

Contents

- *Curve Fitting Techniques* (iccapmhb)

Curve Fitting Techniques

Introduction

Contents

- *Synthesizing Measurement Data for Verification of Extraction Methods* (iccapmhb)
- *Regression Analysis* (iccapmhb)
- *Applying Linear Regression Analysis to Modeling* (iccapmhb)
- *Direct, Visual Parameter Extraction* (iccapmhb)
- *General Modeling Techniques* (iccapmhb)
- *Data Interpolation* (iccapmhb)

On the Web

- Articles
- Knowledge Center
- Technical Support
- Training
- Videos

Synthesizing Measurement Data for Verification of Extraction Methods

Using IC-CAP for the extraction of model parameters offers a lot of flexibility in terms of creating user-defined models and implementing the corresponding extraction routines.

But when developing a new extraction strategy, we may run into two major problems:

- do the routines extract the parameters correctly?
- is the model able to fit the measured device at all?

Related to the first problem: are the extraction routines correct ?
Assume that we know the precise parameter values in advance. If we were able to generate quasi-measured data out of these parameters, then the extracted parameters would have to provide exactly these parameters back again!

Using IC-CAP, it is simple to perform such a check. The trick is to 'synthesize' quasi-measured data out of a set of parameters and to apply then the extraction routines to these data. This can be done as follows:

1. Define a measurement Setup in IC-CAP, for which the extraction routines shall be tested. Example: an bipolar output characteristic for an Early-voltage extraction.
2. Select a 'typical' set of parameters (no default values like 'zero' or 'infinite', but instead real realistic values!)
3. Change the 'Output' data type to 'S' (simulated only). The array behind that Output is now simulation data only.
4. Simulate this Setup with these known parameter values.
5. Change the 'Output' data type back to type 'B'.
   IC-CAP doubles now the data field to measurement and simulation data.
   This means: the simulated data of step 4. is now converted to measured data!
6. Reset the model parameters by clicking 'Reset to Defaults' and simulate the Setup using the default parameters.
7. Apply the extraction routine-under-test and check the values of the extracted parameters.

Provided we get the parameter values back within a good tolerance, we can be sure that the extraction works correctly. If we now apply the extraction to real-world measured data, we should obtain the right parameters. This is true if the measured data have the same shape like the model equations! If not, we might have to choose another model or go for subcircuit modeling.

This means, we can now decide whether the model is able to fit the meas. data !

Note
The problem of whether the actual device can be described by the selected model or not, can also be checked by using the method of 'Direct, Visual Parameter Extraction'. If this method does not produce a parameter plateau, the DUT does not behave in the manner described by the model.

Regression Analysis

Regression Analysis is a technique which provides a best curve fit for given data sets.

Mathematical Basics of Curve Fitting

Let's assume we made 'N' measurements yi at the stimulating points xi.
I.e. we obtained the array [xi, yi], which is then shown in a plot.

A curve Y(x) shall be fitted to this array of measured data points using least square curve fitting technique.
Referring to an individual measurement point, the fitting error is:

and for all data points:

$$E = \sum_{i=1}^{N} E_i^2 = \sum_{i=1}^{N} \left[Y(x_i) - y_i \right]^2 \qquad (2)$$

This error shall be minimized.

The fitting will be done by varying the coefficients of the fitting curve of equation (2). The minimum of the total error E depends on the values of these coefficients. This means, we have to differentiate E partially versus the curve coefficients and to set the results to zero. We obtain a system of equations, solve it, and get the values of the coefficients for a best curve fit. This is known as regression analysis.

> **Note**
> This regression analysis is simple for a straight line fit. But in general, measured data is non-linear. Unfortunately, a non-linear regression analysis can be quite complicated. This problem can be solved if we use a suitable transformation on the measured data. This means that the measured data is transformed to a linear context between the yi- and the xi-values. As will be seen in the diode example later, this is a pretty smart way to get the curve fitting parameters easily without much calculations.

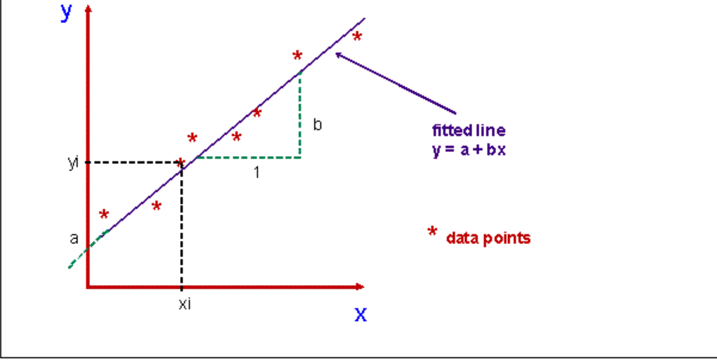
Linear Curve Fitting

IC-CAP File:
demo_features/5_PEL_PROGRAMMING/3_PARAM_EXTRACTIONS_CURVE_FITTING/1_basic_PEL_extractions/1fit_line.mdl

Provided we have got an array of N measured data points of the form [xi,yi].
A linear curve with the equation

$$y(x) = a + b * x \qquad (3)$$

shall be fitted to these points. This situation is depicted below.



Linear regression applied to measurement points

The error of the i-th measurement is:

$$Ei = [a + b * xi] - yi \qquad (4a)$$

Using the least means square method following equ.(2) yields:

$$E = \sum_{i=1}^{N} Ei^2 = \sum_{i=1}^{N} [a + b * xi - yi]^2 \overset{!}{=} \text{Minimum} \qquad (4b)$$

Partial differentiation versus slope 'm' gives:

$$2 \sum_{i=1}^{N} [a + b * xi - yi] xi = 0 \qquad (5)$$

and versus y-intersect 'b':

$$2 \sum_{i=1}^{N} [a + b * xi - yi] = 0 \qquad (6)$$

We obtain from (5) after a re-arrangement:

$$b \sum_{i=1}^{N} xi^2 + a \sum_{i=1}^{N} xi = \sum_{i=1}^{N} yi\ xi \qquad (7)$$

and from (6):

$$b \sum_{i=1}^{N} xi + N\ a = \sum_{i=1}^{N} yi \qquad (8)$$

Multiplying (7) by -N and (8) by xi and adding these two equations allows the elimination of the coefficient 'a', and we can separate the slope 'b':

$$b \left[\left(\sum_{i=1}^{N} xi \right)^2 - N \sum_{i=1}^{N} xi^2 \right] = \sum_{i=1}^{N} xi \sum_{i=1}^{N} yi - N \sum_{i=1}^{N} xi\ yi \qquad (9)$$

or:

$$b = \frac{\sum_{i=1}^{N} x_i \cdot \sum_{i=1}^{N} y_i - N \sum_{i=1}^{N} x_i \cdot y_i}{\left(\sum_{i=1}^{N} x_i \right)^2 - N \sum_{i=1}^{N} x_i^2} \qquad (10)$$

and from (8) for the y-intersect 'a':

$$a = \frac{1}{N} \left(\sum_{i=1}^{N} y_i - m \sum_{i=1}^{N} x_i \right) \qquad (11)$$

with 'b' according to (10).

With equations (10) and (11), we determine the values of the two coefficients of the line which fits best into the 'cloud' of measured data.

Finally, a curve fitting quality factor r2 [regression coefficient] is defined.
Its value ranges from [0 < r2 < 1].
The closer it is to 1, the better is the fit of the linear curve.

$$r^2 = b^2 \frac{\sum_{i=1}^{N} x_i^2 - \frac{1}{N} \left(\sum_{i=1}^{N} x_i \right)^2}{\sum_{i=1}^{N} y_i^2 - \frac{1}{N} \left(\sum_{i=1}^{N} y_i \right)^2} \qquad (12)$$

with 'b' from (10)

General Formula for Regression Analysis

Linear Regression:

$$y = f(x) = a + b \cdot x$$

$$\begin{pmatrix} N & \sum_{i=1}^{N} x_i \\ \sum_{i=1}^{N} x_i & \sum_{i=1}^{N} x_i^2 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{N} y_i \\ \sum_{i=1}^{N} x_i y_i \end{pmatrix} \qquad (13)$$

Quadratic Regression:

$$y = f(x) = a + b \cdot x + c \cdot x^2$$

$$\begin{pmatrix} N & \sum_{i=1}^{N} x_i & \sum_{i=1}^{N} x_i^2 \\ \sum_{i=1}^{N} x_i & \sum_{i=1}^{N} x_i^2 & \sum_{i=1}^{N} x_i^3 \\ \sum_{i=1}^{N} x_i^2 & \sum_{i=1}^{N} x_i^3 & \sum_{i=1}^{N} x_i^4 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{N} y_i \\ \sum_{i=1}^{N} x_i y_i \\ \sum_{i=1}^{N} x_i^2 y_i \end{pmatrix} \qquad (14)$$

Cubic Regression:
continue with the evolution of matrices like above.

Application: Calculating the quadratic polynomial parameters from measurement data

IC-CAP File:
demo_features/5_PEL_PROGRAMMING/3_PARAM_EXTRACTIONS_CURVE_FITTING/1_basic_PEL_extractions/1fit_quadratic.mdl

Starting with equ.(14) from above

$$\begin{pmatrix} N & \sum_{i=1}^{N} x_i & \sum_{i=1}^{N} x_i^2 \\ \sum_{i=1}^{N} x_i & \sum_{i=1}^{N} x_i^2 & \sum_{i=1}^{N} x_i^3 \\ \sum_{i=1}^{N} x_i^2 & \sum_{i=1}^{N} x_i^3 & \sum_{i=1}^{N} x_i^4 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{N} y_i \\ \sum_{i=1}^{N} x_i y_i \\ \sum_{i=1}^{N} x_i^2 y_i \end{pmatrix}$$

we get the following system of 3 equations:

$$\sum_{i=1}^{N} y_i = N \cdot a + \sum_{i=1}^{N} x_i * b + \sum_{i=1}^{N} x_i^2 * c \quad \text{or} \quad \frac{\sum_{i=1}^{N} y_i}{N} = a + \frac{\sum_{i=1}^{N} x_i}{N} * b + \frac{\sum_{i=1}^{N} x_i^2}{N} * c \qquad (15a)$$

$$\sum_{i=1}^{N} x_i y_i = \sum_{i=1}^{N} x_i * a + \sum_{i=1}^{N} x_i^2 * b + \sum_{i=1}^{N} x_i^3 * c \quad \text{or} \quad \frac{\sum_{i=1}^{N} x_i y_i}{\sum_{i=1}^{N} x_i} = a + \frac{\sum_{i=1}^{N} x_i^2}{\sum_{i=1}^{N} x_i} * b + \frac{\sum_{i=1}^{N} x_i^3}{\sum_{i=1}^{N} x_i} * c \qquad (15b)$$

$$\sum_{i=1}^{N} x_i^2 y_i = \sum_{i=1}^{N} x_i^2 * a + \sum_{i=1}^{N} x_i^3 * b + \sum_{i=1}^{N} x_i^4 * c \quad \text{or} \quad \frac{\sum_{i=1}^{N} x_i^2 y_i}{\sum_{i=1}^{N} x_i^2} = a + \frac{\sum_{i=1}^{N} x_i^3}{\sum_{i=1}^{N} x_i^2} * b + \frac{\sum_{i=1}^{N} x_i^4}{\sum_{i=1}^{N} x_i^2} * c \qquad (15c)$$

Equations (15a) .. (15c) are of the form

$$y0 = a + x0 * b + x0^2 * c \qquad (16a)$$

$$y1 = a + x1 * b + x1^2 * c \qquad (16b)$$

$$y2 = a + x2 * b + x2^2 * c \qquad (16c)$$

for which we have already a solution for a, b and c in chapter 6.Data_Interpolation.

Therefore, with the following abbreviations:

$$y0 = \frac{\sum_{i=1}^{N} y_i}{N} \quad x0 = \frac{\sum_{i=1}^{N} x_i}{N} \quad x02 = \frac{\sum_{i=1}^{N} x_i^2}{N}$$

$$y1 = \frac{\sum_{i=1}^{N} x_i y_i}{\sum_{i=1}^{N} x_i} \quad x1 = \frac{\sum_{i=1}^{N} x_i^2}{\sum_{i=1}^{N} x_i} \quad x12 = \frac{\sum_{i=1}^{N} x_i^3}{\sum_{i=1}^{N} x_i}$$

$$y2 = \frac{\sum_{i=1}^{N} x_i^2 y_i}{\sum_{i=1}^{N} x_i} \quad x2 = \frac{\sum_{i=1}^{N} x_i^3}{\sum_{i=1}^{N} x_i} \quad x22 = \frac{\sum_{i=1}^{N} x_i^4}{\sum_{i=1}^{N} x_i}$$

we can re-use the quadratic data interpolation formulae of the chapter Data Interpolation:

$$c = \frac{(y1 - y0) * (x1 - x2) - (x1 - x0) * (y1 - y2)}{(x12 - x02) * (x1 - x2) - (x1 - x0) * (x12 - x22)}$$

$$b = \frac{y1 - y2 + c * (x22 - x12)}{(x1 - x2)}$$

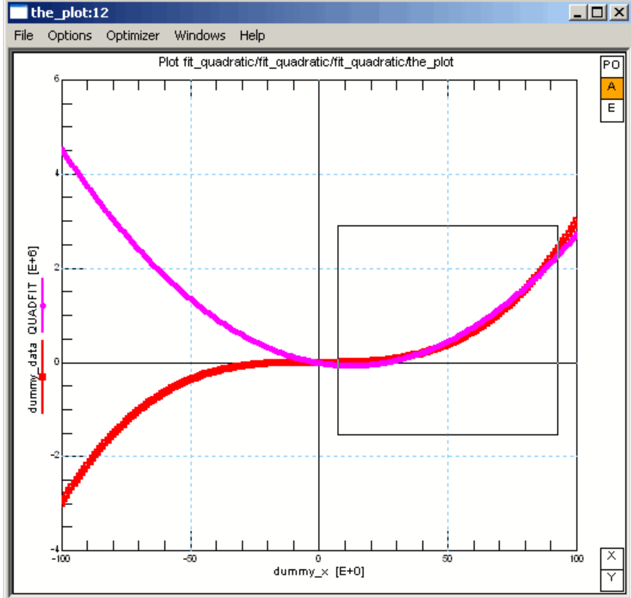
and

$$a = y2 - b * x2 - c * x22$$

and are ready to get the quadratic fitting function y(x)

$$y = a + b * x + c * x^2$$

In the plot below, the red curve (dummy_data) has been fitted by the magenta curve (QUADFIT), inside the marked box.



the_plot:12
File Options Optimizer Windows Help
Plot fit_quadratic/fit_quadratic/fit_quadratic/the_plot
PO
A
E
X
Y
dummy_data QUADFIT [E+0]
dummy_x [E+0]

Applying Linear Regression Analysis to Modeling

For IC-CAP examples, see under directory
$ICCAP_ROOT/examples/demo_features\5_PEL_PROGRAMMING\3_PARAM_EXTRACTIONS__CURVE_FITTING
1_basic_PEL_extractions
and also .../specific_PEL_routines

DC curves:

Applying linear regression analysis to DC modeling data is pretty obvious. Examples include the Early voltage VAF of a bipolar transistor, the threshold voltage of a MOS or the saturation current IS and the exponential coefficient N of a DC diode characteristic.
For more details see the mentioned IC-CAP demo files, or the chapter on diode modeling.

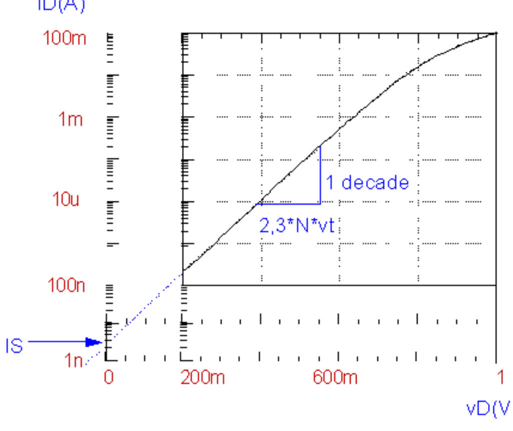
As a brief overview, we consider the diode DC current modeling:

$$i_D = I_S * e^{\frac{v_D}{N * vt}} \qquad (1)$$

with VT: temperature voltage 27mV at 25°C

```
VT = k*T/q = 8.6171 E-5 * (T /'C + 273.15)   (1a)
```

or



DC characteristic of the diode under forward bias and the determination of the DC parameters IS and N

Provided vD > 0, i.e. neglecting the term (-1) in equ.(1), and applying a logarithmic conversion yields:

```
                            vD
log(iD)   =  log(Is)   +   ------ log(e)
                            N VT

          =  log(Is)   +   [1 / (2,3 N VT)]  vD              (2a)
```

This is an equation of the form:

```
 y   =   b    +      m       x            (2b)
```

In order to interpret (2b) linearly, we have to substitute:

```
 y =  log(iD)                                  (2c)
          b =  log(Is)                         (2d)
                       m =  [1 / (2,3 N VT)]   (2e)
                              x =  vD          (2f)
```

And this explains how to manipulate the measured data: after the logarithmic conversion of the measured values of i_D (2c), they are introduced, together with the still linear values of vD (2f) into the regression equations (10) and (11) of the previous chapter as yi- and xi-values. We obtain the y-intersect b and the slope m of the linear regression function. Solving (2d) for Is and (2e) for N we finally are able to calculate these two parameters out of b and m as follows:

$$Is = 10^{b} \qquad (3)$$

```
N  =  1 / (2,3 m VT)
                          with VT from (1a)                (4)
```

CV curves:

When modeling non-linear components such as diodes or transistors, space charge capacitor curve fitting is always an issue. These are voltage dependent capacitors (CV curves). And since this is a central modeling issue, some different methods of applying regression techniques to this problem are discussed in this chapter.

The general space charge capacitance Cs formulation for voltages vD < FC * VJ is:

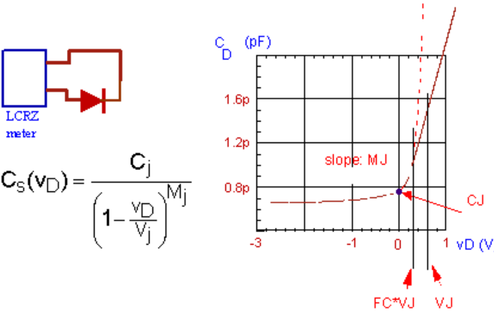
$$C_s = \frac{C_J}{\left(1 - \frac{v_D}{V_J}\right)^{M_J}} \qquad (1a)$$

and else

$$C_s = \frac{C_J}{(1 - F_C)^{(1+M_J)}} * \left[1 - F_C * (1 + M_J) + M_J * \frac{v_D}{V_J}\right] \qquad (1b)$$

with

- CJ: space charge capacitance at vD = 0V
- VJ: diffusion voltage (built-in potential), models the pole of equation (1a) (typ. 0,7V)
- MJ: space charge exponential factor (models the slope of the CV curve)
  (abrupt pn transition (<0,5um): MJ = 1/2)
  (linear pn transition (> 5um): MJ = 1/3)
- FC: switching coefficient in forward mode, default 0,5, see left picture in figure 1.



Parameter FC switches from the hyperbolic model equation to a linear continuation for vD>FC*Vj

For most simulators, the capacitances are rather described by the charges.
This means for vD<FC*Vj

$$qs = \frac{V_j}{(1 - M_j)}\left[1 - \left(1 - \frac{V_D}{V_j}\right)^{(1-M_j)}\right]$$

$$qs = \frac{V_j}{(1 - M_j)}\left[1 - (1 - FC)^{(1-M_j)}\right] + (V_D - FC \cdot V_j) \cdot \left[\frac{1 - FC + \frac{M_j (V_D - FC \cdot V_j)}{2V_j}}{(1 - FC)^{(1+M_j)}}\right]$$

General CV modeling

For simplicity, the capacitor voltage is always less than FC*VJ for parameter extraction.
i.e. we start with equ. (1a), repeated here again for convenience

$$C_s = \frac{C_J}{\left(1 - \frac{v_D}{V_J}\right)^{M_J}} \qquad (1a)$$

A logarithmic conversion of equation.(1a) yields

```
ln(Cs) =   ln(CJ)   -  MJ ln[1 - vD / VJ ]
```

This equation can be linearized following

```
y      =     b    +    m      x
```

when substituting:

$$y = \ln(Cs) \quad (2)$$
$$b = \ln(CJ) \quad (3)$$
$$m = -\ MJ \quad (4)$$
$$x = \ln[1 - vD / VJ] \quad (5)$$

In order to apply regression techniques, we have to transform the stimulus voltage values vD[i] following (5) and the measured capacitance values Cs[i] following (2). Under the assumption that ~0.3 < VJ < ~1 is valid, we use for the transformation of vD[i] a starting value for VJ, e.g.. 0.3V. These two data arrays x[i] and y[i] are applied to a regression analysis, and we get a certain slope m(VJ) and a y-intersect b(VJ). This fitting is also related to a certain fitting factor r^2, the regression coefficient. The values of m, b ad r^2 are stored and then the parameter VJ is incremented to e.g. 0.31V. Then, the same data transform is applied to the stimulating and measured data x[i] and y[i] , and another regression analysis performed. From that we get another triplet of coefficients m[VJ], b[VJ], and r^2. If r^2 is now better than before, we continue by stepping VJ, if it is worse, VJ is stepped in the other direction. When r^2 reached its maximum, the best fitting coefficient r^2 is obtained and the corresponding VJ_opt, m(VJ_opt) and b(VJ_opt) are now known.
This allows to re-substitute:

from (4): MJ = - m[VJ_opt]

and from (3): CJ = exp (b[VJ_opt])

In practice we often have the problem that the measured space charge capacitance is overlaid by an offset capacitor (package, test structure etc.). This leads to the problem of determining that offset capacitance as well. This is covered by the following two suggestions.

CV modeling with MJ=const, but including offset capacitance C_{offs}

In this case, the fitting formula is

$$C_s = \frac{C_J}{\left(1 - \frac{V_D}{V_J}\right)^{M_J}} + C_{OFFS}$$

This equation can also be made linear, yielding
y = m x + b
if the following substitutions are made:
y = Cs
m = CJ

$$x = \frac{1}{\left(1 - \frac{V_D}{V_J}\right)^{M_J}}$$

with

- MJ = constant
- b = C_{offs}

I.e.: we preset MJ to a fixed value:
abrupt PN junction (<0,5um): MJ = 1/2
linear PN junction (> 5um): MJ = 1/3

Using the same procedure as in the previous chapter, VJ is incremented/looped until the best regression factor r^2 is found.
Reverse substitution gives:

CJ = m(VJ_opt)
and
C_{offs} = b(VJ_opt)

where, as noted above, MJ has been specifically predefined.

Total CV modeling including offset capacitance C_{offs}

Assuming that the CV value at vD=0 is known, there is a more 'refined' substitution method that will determine all the CV parameters as well as the offset capacitance Coffs. Again:

$$C_s = \frac{C_J}{\left(1 - \frac{v_D}{V_J}\right)^{M_J}} + C_{OFFS}$$

First, to eliminate the offset capacitance C_{offs}, Cs is derived by vD:

Next, the logarithm is determined:

$$\frac{dC_s}{dv_D} = \frac{C_J * M_J}{V_J} * \frac{1}{\left(1 - \frac{v_D}{V_J}\right)^{M_J+1}}$$

then a linear substitution is again performed:

$$\ln\left(\frac{dC_s}{dv_D}\right) = \ln\left(\frac{C_J * M_J}{V_J}\right) - (M_J + 1) * \ln\left(1 - \frac{v_D}{V_J}\right)$$

in such way that

$$y = b + m \quad x$$

$$y = \ln\left(\frac{dC_s}{dv_D}\right)$$

$$b = \ln\left(\frac{C_J * M_J}{V_J}\right)$$

$$m = -(MJ + 1)$$

$$x = \ln\left(1 - \frac{v_D}{V_J}\right)$$

Here again, VJ is incremented until the maximum regression quality, r^2, is reached. The following reverse substitution is made for this VJ_opt:

$$CJ = \frac{VJ_opt}{MJ} * \exp[b(VJ_opt)]$$

$$MJ = -m(VJ_opt) - 1$$

Finally, Coffs is calculated to:

$$Coffs = Cs(vD=0) - CJ$$

Regression Analysis for S-Parameters

When modeling S-parameters, the problem of fitting curves to circle sectors again is a common problem. Here too, the modeling equation of a circle with its center on the real axis of the Smith chart can be transformed linearly (!) using the following formulation:

Circle equation with the center on the real axis:

$$(x - x0)^2 + y^2 = r^2$$

or:

$$x^2 + y^2 = r^2 - x0^2 + 2\ x0\ x$$

This form can also be linearized (!), into

$$ylin = b + m \quad xlin$$

with the following settings:

$$ylin = x^2 + y^2 \qquad (2a)$$

$$b = r^2 - x0^2 \qquad (2b)$$

$$m = 2\ x0 \qquad (2c)$$

$$xlin = x \qquad (2d)$$

In other words: the measured values x[i] and y[i] are employed in equation (2a) and (2d). Then, the vector ylin[i] = x i [i] 2 + y i [i] 2 is plotted against xlin[i] = x i [i], and a linear regression line fitting is applied.
From the slope m we obtain (2c)

$$x0 = m / 2 \qquad (3a)$$

and from (2b) solved for r

$$r = \sqrt{b + x_0^2} \qquad (3b)$$

Summary

This section illustrated how linear regression can be applied to various modeling problems.

In practice, this method has proven simpler than applying special regression rules for nonlinear curve fitting. This is because device modeling, the equation to be fitted is known. A that remains is to determine a nonlinear transform for that modeling equation to which linear regression can be employed.

Direct, Visual Parameter Extraction

As an additional possibility for parameter extraction, this chapter presents a method for which measured values are directly transformed into parameter values based on a known model equation.

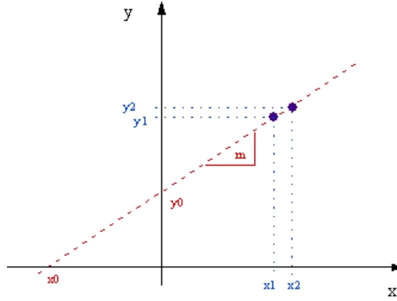
Example, the MOS transistor:
The threshold voltage of different MOS transistor sizes, plotted versus length or width, shows a certain dependency. In old MOS models, the model parameter VTH was the only mean to model VTH. In other words, no dependency of the threshold voltage vs. geometry was included. Applying the visual parameter extraction method, we can develop an improved MOS model by replacing (in the model equations !) the constant parameter VTH by a function of VTH = f(L, W), as obtained by our visualized parameter plots. L and W are the length and width of the MOS transistor.

By the way, this is how company-internal models are developed and improved very often.

Another example: the bipolar transistor:
You know that the x-intersect of lines fitted to an output characteristic of a bipolar transistor should hit always the same point, the Early voltage. If we apply an IC-CAP PEL (parameter extraction language) program to calculate the x-intersect of a line that is fitted to two adjacent measured points, and if we display the result of this operation versus the collector voltage (first order sweep), we will obtain a plot of the 'equivalent' Early-voltages of adjacent measurement points.

The advantage of using this visual extraction method is that we can see clearly, if the model is able to fit the measured data at all. We only have to check if there is a flat region in the transformed data domain or not. If it is there, we can extract the parameter very simply by calculating the mean value of the flat region. And we know at the same time, in which range the parameter is dominant and can therefore be used for fine-tuning with the optimizer. If there is no flat range, the model cannot fit the measured data. We could vary the parameter as much as we like and would not achieve a fit of the simulated to the measured data.

For the application of this method, we start with some basic equations that refer to following figure:



About the definition of some basic equations for the direct visual parameter extraction

Assumed we have
```
y = m*x + y0
```

where m: slope
y0: y-intersect

Then it is:
```
x0 = -b/m
```

and:
```
      y2 - y1
m = ---------
      x2 - x1
```

This finally gives the basic formulas for this modeling method:

starting with

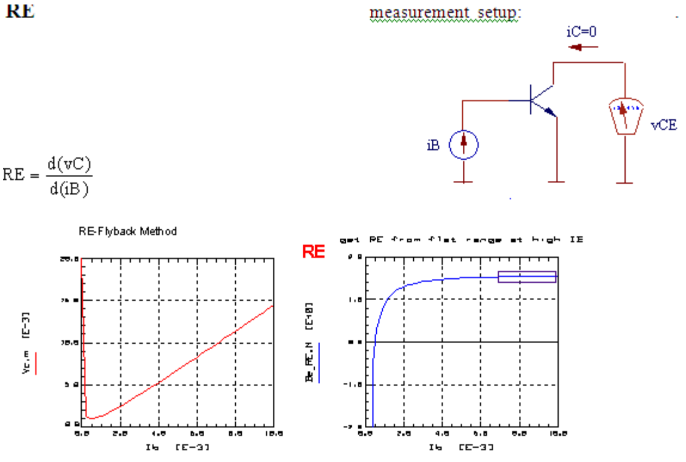
x-intersect x0   y-intersect y0:

$$\frac{y2}{-x0+x2} = \frac{y1}{-x0+x1} \qquad\qquad \frac{y2-y0}{x2} = \frac{y1-y0}{x1}$$

$$x0 = \frac{x1y2 - x2y1}{y2 - y1} \qquad\qquad y0 = \frac{x2y1 - x1y2}{x2 - x1}$$

we get:
The following plots give some examples on how to apply this idea to the parameter extraction of a bipolar transistor using the Gummel-Poon model. It should be mentioned that this method can be applied to all the parameters of this model, as well as to other models like Statz, Curtice, BSIM etc.
IC-CAP demo files about how to apply 'direct visual parameter extraction' to the Gummel-Poon model are available from the author.

The first example is about modeling the parasitic resistors of a bipolar transistor.

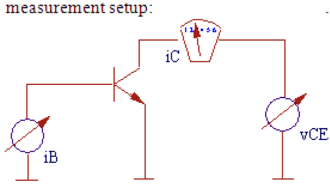


RE is the derivative of the flyback measurement

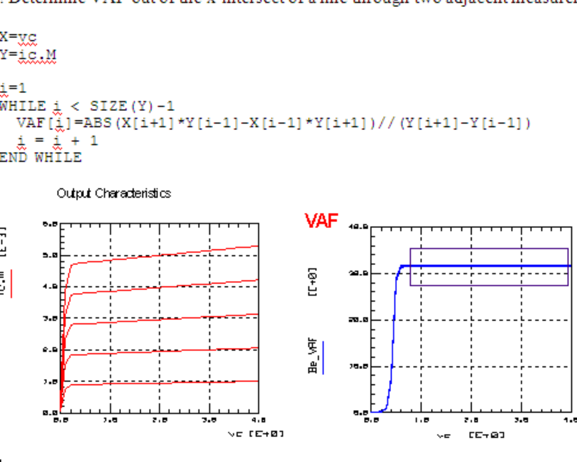
The "visualized" RE is the y-intersect of two adjacent data points from that flyback plot. In this figure, the proposed flyback measurement method and its interpretation to represent the behavior of the Emitter resistor RE (assumed to be an ideal ohmic resistor) seems correct for higher Base currents. Therefore, we extract its value out of the range marked with the box sign.

In the next example, we will see how to apply the visual parameter extraction method to output characteristics of a bipolar transistor.

VAF



! Determine VAF out of the x-intersect of a line through two adjacent measurement points:

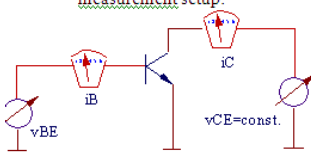


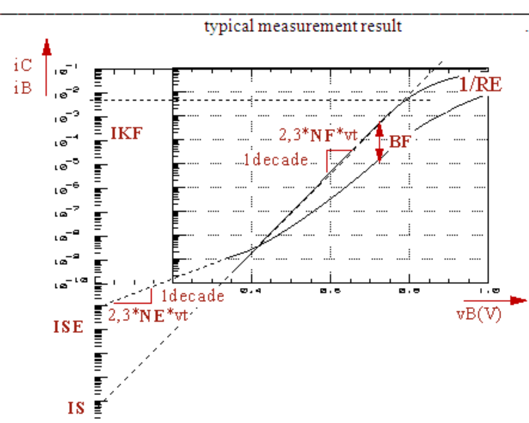
```
X=vc
Y=ic.M

i=1
WHILE i < SIZE(Y)-1
  VAF[i]=ABS(X[i+1]*Y[i-1]-X[i-1]*Y[i+1])//(Y[i+1]-Y[i-1])
  i = i + 1
END WHILE
```


the Early voltage is calculated from the x-intersect of a line through two adjacent data points

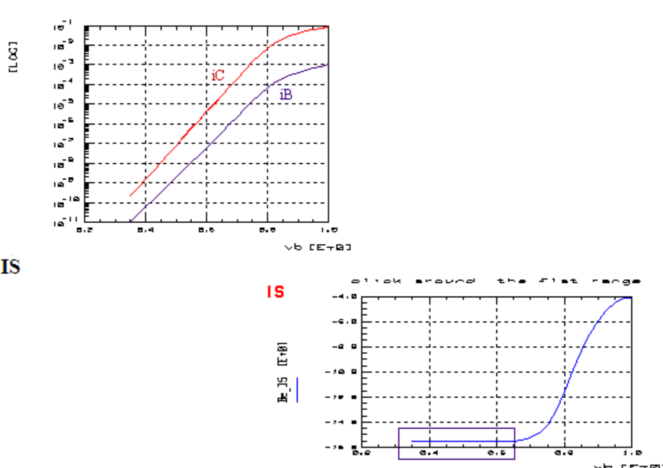
In the example of above figure, we see that the Gummel-Poon model is nicely modeling the increasing slope of the output characteristics with its parameter VAF. Its value is the mean of the data inside the marked box range.

Let's study also the application of the "visual" method for the Gummel-Poon plot:





interpretation of the Gummel-Poon plot for parameter extraction



IS

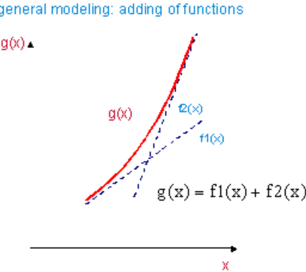
IS

The transformation of measured data directly into the parameter domain works well also for the examples depicted in figures 6 and 7.
In our example, it can be clearly seen that the Base current of that particular transistor does not exhibit the recombination effect (NF in Fig.6 is flat for low vB values).

Note: This is the famous 'knee' in the iB curve visible at lower Base currents.

General Modeling Techniques

When developing an equation for a new model, this equation has to follow the measurement data. In order to develop such an equation from simple partial equations, we can distinguish two situations: Simple adding of partial equations ('summing') or inversely adding ('paralleling').

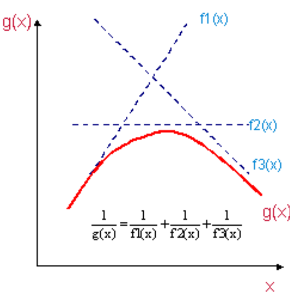


Adding of functions allows to model a curve steepening

While modeling a steepening of a curve with adding of two functions is pretty straight forward (fig.1), the modeling of a declining function is more complex. However, applying 'paralleled' functions, following the equation form

$$\frac{1}{g(x)} = \frac{1}{f1(x)} + \frac{1}{f2(x)} + \frac{1}{f3(x)}$$

allows to also model more complex curve traces, by inversely adding their max. limiting partial functions, see fig.2.



Paralleling of functions allows to model a curve decline

Diode DC modeling

When measuring the DC characteristics of a diode, we have basically 4 regions to model. With increasing bias voltage, this is

1. the recombination effect at low currents
2. the 'conventional', ideal diode region
3. a saturation effect
4. the ohmic region.

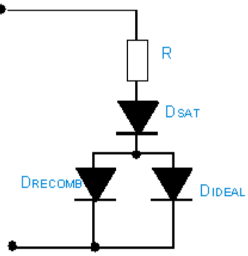
Compared to the conventional, ideal diode region, we typically have another linear trace in a semi-logarithmic plot for the low-current region, i.e. the recombination effect. For a given low voltage vD, there is a higher current flowing than expected from the 'conventional' ideal diode model. This additional current (at the same applied voltage bias) is therefore modeled with a parallel diode DRECOMB.

Then, with increasing vD, comes the 'ideal' diode range, modeled by diode DIDEAL. In other words, with increasing vD, DIDEAL now conducts more and more of the current iD.

Before the ohmic resistor affects the curve, there is often a saturation diode visible in the transition range between the ideal diode and the ohmic range. Since this corresponds to a voltage shift (for a certain, applied current bias), this is a series diode, in series to the so far described parallel diodes. Its effect can be described like this: when the total diode current curve begins to decline, we see an additional voltage drop for a given current. For currents below that value, the voltage drop is zero. Therefore, the IS parameter of the DSAT diode is exactly the 'knee' of the diode DC curve. If, above the knee, the slope is half of the ideal diode, the N parameter of DSAT equals N of DIDEAL. If it is flatter, N of DSAT is even bigger.

After that DSAT region, a series resistor is used to model the ohmic losses.

This means, we end up with a schematic like below:

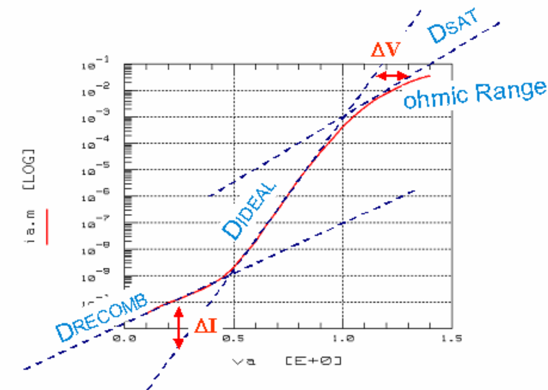


Schematic for modeling the DC performance of a diode

> **Note**
> if you do not want to use DSAT for the CV modeling, treat this diode as an RF short, i.e. set DSAT parameter CJO=100m.

modeling example: DC diode characteristics
higher current at a given vD means a parallel diode

modeling example: DC diode characteristics
higher voltage at a given iD means a series diode



The following measurement result and the modeling steps illustrate the idea for the DC modeling of a typical diode.

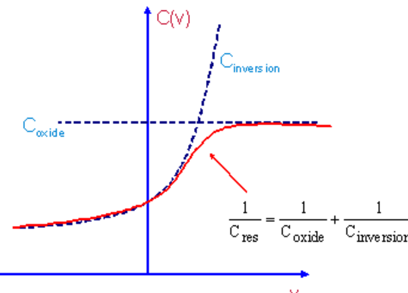
MOS Gate Oxide Capacitance (simplified model)

MOS transistor. In this case, the measured trace (see fig.3) can be interpreted as two series capacitors,

- one of constant capacitance C_{oxide}
- and a second one of varying $C_{inversion}$, after

$$C(v) = \frac{CJO}{\left(1 - \frac{v}{VJ}\right)^{M}}$$

Overlying C_{oxide} with $C_{inversion}$ yields the Gate Oxide Capacitance of a MOS transistor


modeling example: Cox of a MOS transistor
C(v)
C_inversion
C_oxide
1/C_res = 1/C_oxide + 1/C_inversion
v

Data Interpolation

IC-CAP File:
demo_features\3_MEAS_ORGANIZE_n_VERIFY_DATA\
Interpolate_DC_CV_by_QuadrPolynomial____Get_inner_device_volt.mdl

Fitting a quadratic function to three data points:

For three data points [x0, y0], [x1, y1], [x2, y2], we start with a set of 3 quadratic equations

$$y0 = a + b * x0 + c * x0^2 \quad (1)$$
$$y1 = a + b * x1 + c * x1^2 \quad (2)$$
$$y2 = a + b * x2 + c * x2^2 \quad (3)$$

Goal: We need to get the coefficients a, b and c as a function of xi and yi
From (1):

$$a = y0 - b * x0 - c * x0^2 \quad (4)$$

(4) into (2):

$$y1 = y0 - b * x0 - c * x0^2 + b * x1 + c * x1^2 \quad (5)$$

In the same way, from (3):

$$\boxed{a = y2 - b * x2 - c * x2^2} \quad (6)$$

(6) into (2):

$$y1 = y2 - b * x2 - c * x2^2 + b * x1 + c * x1^2$$

$$b * (x1 - x2) = y1 - y2 + c * (x2^2 - x1^2)$$

$$\boxed{b = \frac{y1 - y2 + c * (x2^2 - x1^2)}{(x1 - x2)}} \quad (7)$$

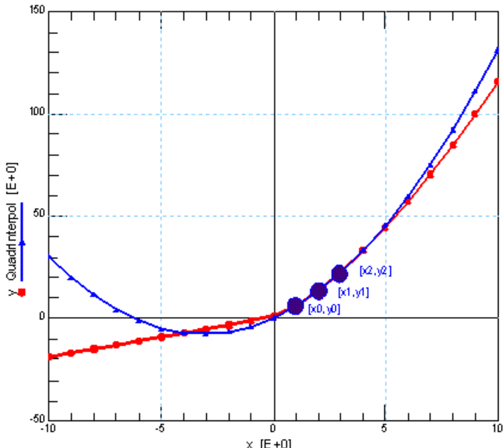
Finally, (7) into (5):

Conclusion:
(6), (7) and (8) represent the equations for the polynomial coefficients a, b and c.

$$y1 = y0 + (x1 - x0) * \frac{y1 - y2 + c * (x2^2 - x1^2)}{(x1 - x2)} + c * (x1^2 - x0^2)$$

$$(y1 - y0) * (x1 - x2) - (x1 - x0) * (y1 - y2) = c * (x1^2 - x0^2) * (x1 - x2) - c * (x1 - x0) * (x1^2 - x2^2)$$

$$\boxed{c = \frac{(y1 - y0) * (x1 - x2) - (x1 - x0) * (y1 - y2)}{(x1^2 - x0^2) * (x1 - x2) - (x1 - x0) * (x1^2 - x2^2)}} \quad (8)$$



Simulations

Introduction

Contents

- *History* (iccapmhb)
- *Simulator Interface* (iccapmhb)
- *Netlist Syntax* (iccapmhb)
- *CV and S-Parameters* (iccapmhb)
- *Harmonic Balance* (iccapmhb)

On the Web

- Articles
- Knowledge Center
- Technical Support
- Training
- Videos

History of Circuit Simulators

The origin of Berkeley SPICE

SPICE (Simulation Program with Integrated Circuit Emphasis) was introduced in May 1972 by Prof. D.O.Pederson at the University of Berkeley, California. It simulates the electrical behavior of circuits based on discrete components In particular, these include:

- Resistors, capacitors, inductors
- Coupled inductors ideal lines and cables
- Independent and controlled current and voltage sources
- Semiconductor components

It soon became a standard for electronic circuit simulation, used by both, the industry and the universities. This fact is based on the technical expertise of the developers of SPICE: especially Prof. Pederson, and Prof. Nagel.

After 17 revisions of SPICE, SPICE2 was introduced in 1975. The main improvements were the dynamical RAM allocation capabilities and the automatic time stepping for time-domain simulations. Especially SPICE version 2g6, was a specially reliable program version, written in Fortran. This version has become the root of many industry in-house simulators and also of many commercially available simulators. This explains why they typically understand the Berkeley syntax (Spectre, Hspice, Pspice, Eldo).

Finally, SPICE3, introduced by the end of the 80s, was the first version written in the C language.

One of the reasons for the success of SPICE is also the implementation of good transistor models. This is especially true for the bipolar model. The Gummel-Poon model has only become popular after it was implemented in SPICE. (However, some simplifications had been made to the original paper of Gummel-Poon.) The same is true for MOS transistor, however, after MOS3, and the introduction of the BSIM series of model, SPICE's model quality had suffered a lot compared to MOS2 and MOS3. Finally, with BSIM3v3, Berkeley SPICE gained back its authority of model definition.

The origin of ADS

The origin of MDS simulator goes back to the late 1960's when Hewlett-Packard introduced the simulator OPNODE which was a linear s-parameter simulator. In 1984 HP wrote a linear simulator called Time-Domain Tutorial which could be used to take a transmission line topology, compute the S-parameters, and use the 8510 to get a time domain reflectometer plot. This time domain tutorial program eventually evolved into the MDS Microwave Linear Simulator MLS in MDS A.01.00.

Ken Kundert worked on MLS until 1985 when HP funded him to get his Ph.D at U.C Berkeley working on nonlinear microwave simulation based upon harmonic balance technology. HP was co-developing harmonic balance in conjunction with Ken Kundert and Berkeley. The internal name for this simulator was "Spectra" which was later named the HP Microwave Nonlinear Simulator or MNS. MNS essentially evolved from Berkeley's code. The Berkeley code was named Spectre (a subtle difference). After Ken Kundert received his doctorate, he left HP to work at Cadence where he built a new time domain simulator based upon the UC Berkeley Spectre code, hence the Cadence Spectre simulator was born.

In 1996, Hewlett-Packard bought EEsof, and MDS and the EEsof simulator products were merged. The new product was called ADS (Advanced Design System).

References

1. Nagel, L.W.; Pederson, D.O.: SPICE. Berkeley, Univ. of California, Electronic Research Laboratory, ERL-M 382, 1973
2. Nagel, L.W.: SPICE2: A Computer Program to Simulate Semiconductor Circuits. Berkeley, Univ. of California, Electronic Research Laboratory. ERL-M520, 1975

Simulator Interface of IC-CAP

IC-CAP always links to separate, standalone simulators. However, to make this concept work independently of the customer's simulator capabilities, three types of simulators are added to the CD-ROM containing IC-CAP:

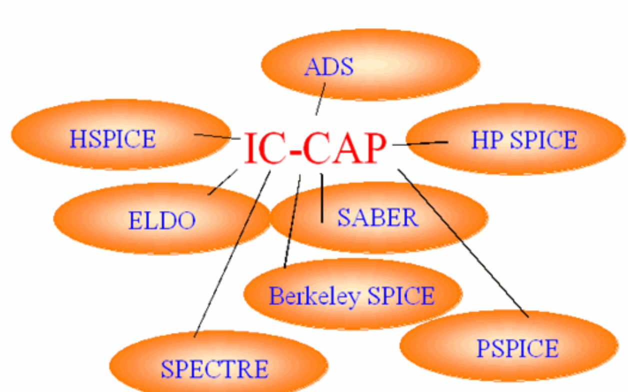
- UCB Spice2g6, (University of California, Berkeley) which is the base of most customer-specific simulators
- UCB Spice3e2
- HPSpice

Since IC-CAP 2006, the license for using ADS linear (DC, CV, S-par) and transient (TimeDomain) simulations is included to IC-CAP. This allows the user to use the latest versions of the industry-standard models like PSP, Hicum, VBIC etc.

- Spice3 is no longer supported by UCB since the early 2000, and therefore, does no longer new models introduced to the marked since then.
- ADS needs to be installed additionally to IC-CAP.

This means that, provided the user has no local simulator available, he can always perform simulations based on ADS or on these 3 types of Spice simulators.

On the other hand, IC-CAP can also interface to other types of simulators. Links, included in the IC-CAP Analysis Module (see IC-CAP file $ICCAP_ROOT/lib/iccap/usersimulators), are depicted in fig.1 below. It should be mentioned that these simulators need not necessarily be available on the local IC-CAP workstation, but that also remote simulation is supported (see again the IC-CAP file $ICCAP_ROOT/lib/iccap/usersimulators).



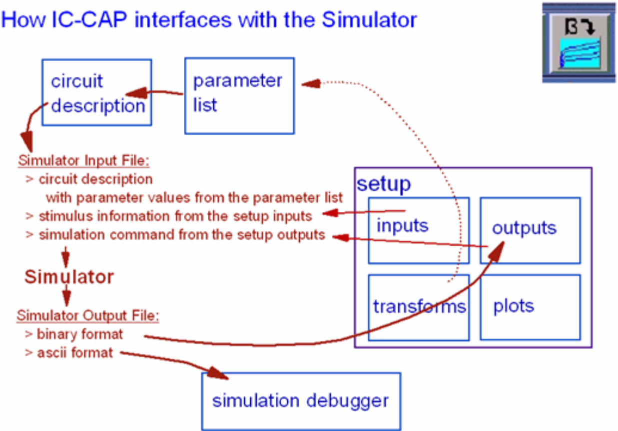
Using IC-CAP's PEL language, or the userC functions, it would be also possible to include the model equations directly into IC-CAP. However, this would imply the following disadvantages:

- if the model equations change, esp. true for user-specific models, this has to be reflected in the IC-CAP model equations.
- if an additional component has to be added to the specified model in IC-CAP, one would have to redefine all model equations. Example: an extra series resistor for a diode model would reduce the voltage drop across the implemented model equations, thus affecting the diode equations.

In order to avoid this dilemma, it is desirable to always use external simulators. In this case, and coming back to the example from above, an additional series resistor is simply added to the simulator sub-circuit. Then, there is no need to modify any model equations inside IC-CAP. The only modification would be to add a series-resistor parameter extraction module in order to determine the value of this additional parasitic resistor.

Because IC-CAP also supports simulators which allow the user to specify own proprietary model formulas like ADS (Verilog-A, sdd), Eldo (Verilog-A, HDLA, ADVance-MS) etc., it is possible to also use user-specific models.
For examples, see
demo_features\6_SIMULATORS\2_ADS\5_NETLIST_TEMPLATES\0_USER_DEF_MODELS

How IC-CAP interfaces with the simulators.



Step-by-step:

1. After the component has been measured, the user specifies a suitable sub-circuit in IC-CAP's Circuit description window as shown in (fig.3)
2. Then, the model parameters are extracted and exported from the IC-CAP Transform hierarchy into the Model Parameter window (which, on the other hand, is a consequence of the user-specific Circuit Window), see fig.3 again.
3. When a simulation is requested by the user (or the optimizer), IC-CAP takes the circuit deck of the Circuit Window, replaces the default values defined there with the actually extracted model parameters from the Parameter List Window, adds the stimulus information of the Setup Inputs and also the requested type of simulation specified with the Setup Output(s) and sends the whole information as a simulator input deck to the simulator.
4. The simulation is executed, locally or remote, using the syntax 'simulator input_deck output_deck'
5. The simulator perform the calculations and exports its simulation results into two files: one ASCII file containing the execution time and the error messages, and a binary file containing the very simulation result.
6. IC-CAP reads the binary file and places the simulation results under the Output icons of the Setup window. The ASCII file is only read if the Simulation Debugger window is open. Otherwise, this file is ignored. (See fig.4)

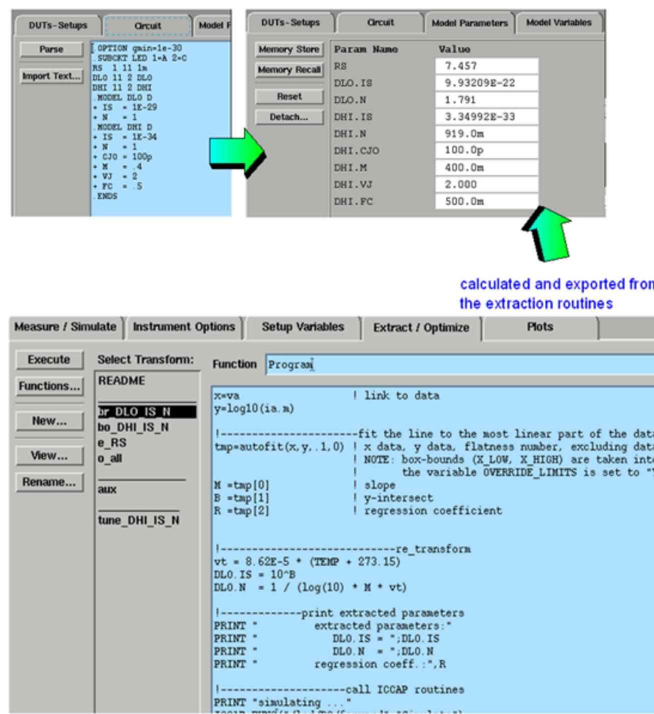
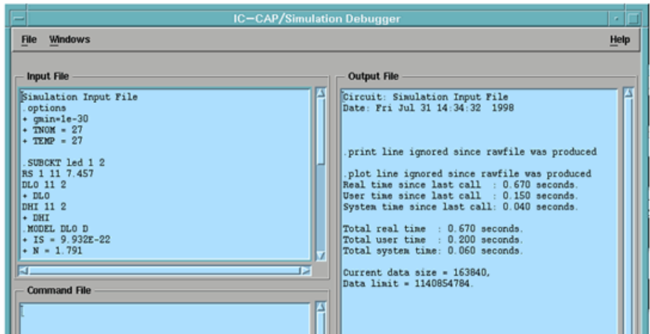


Fig.3: First, a suitable sub-circuit is selected and defined for the simulations. Then, the exacted model parameters are exported into the Model Parameters window. The model parameters showing up in this window are a consequence of the user's sub-circuit definition.


IC−CAP/Simulation Debugger
File
Windows
Help
Input File
Simulation Input File
.options
+ gmin=1e-30
+ TNOM = 27
+ TEMP = 27
.SUBCKT led 1 2
RS 1 11 7.457
DLO 11 2
+ DLO
DHI 11 2
+ DHI
.MODEL DLO D
+ IS = 9.932E-22
+ N = 1.791
Command File
Output File
Circuit: Simulation Input File
Date: Fri Jul 31 14:34:32 1998
.print line ignored since rawfile was produced
.plot line ignored since rawfile was produced
Real time since last call : 0.670 seconds.
User time since last call : 0.150 seconds.
System time since last call: 0.040 seconds.
Total real time : 0.670 seconds.
Total user time : 0.200 seconds.
Total system time: 0.060 seconds.
Current data size = 163840,
Data limit = 1140854784.

Netlist Syntax for ADS and SPICE

Contents

- *Comparing Spice and ADS Syntax* (iccapmhb)
- *Common Parameter List* (iccapmhb)
- *UCB Spice Netlist* (iccapmhb)
- *Spice3 User Defined Model* (iccapmhb)
- *ADS Simulator Netlist* (iccapmhb)
- *ADS User Defined Model* (iccapmhb)
- *Spice ADS Header* (iccapmhb)

Comparing UCB Spice and ADS Netlist Syntax

Table 1: summary of the most commonly used components

Component	**Syntax for Berkeley SPICE**	**Syntax for ADS**
Resistor	R1 1 2 1k	R:R1 1 2 R=1k
Capacitor	C1 1 2 1p	C:C1 1 2 C=1p
Inductor	L1 1 2 1n	L:L1 1 2 L=1n
Coupling between inductors L1 and L2	K1 L1 L2 .8	Mutual:M1 K=0.5 Inductor1="L1" Inductor2="L2"
Ideal Delay Line	T1 1 0 2 0 Z0=50 TD=10n	#uselib "ckt" "TimeDelay" TimeDelay:T1 1 2 ZRef=50 Delay=10p
Independent Voltage Source between node 1 and 2, with 5V	V1 1 2 5	V_Source:V1 1 2 Vdc=5
Diode	D1 1 2 myDiomodl .MODEL myDiomodel D + IS = 1p ...	myDiomodel:d1 A C model myDiomodel Diode Tnom = 27 Is = 1p ...
Gummel-Poon bipolar transistor	Q1 1 2 3 myGPmodel .MODEL myGPmodel NPN + IS = 1p ...	myGPmodel:Q1 C B E model myGPmodel BJT\\NPN=yes Tnom= 27 Is = 1p ...
Subcircuit Call	X1 1 2 3 mySub	mySub:X1 1 2 3
Subcircuit	.SUBCKT mySub 10 20 30 ... .ENDS	define mySub 10 20 30 ... end mySub

Note: most simulators except Berkeley SPICE permit the use of node names instead of numbers

Common Parameter List

How to have a common Parameter List for Spice and ADS

If you plan to do device modeling with exchanging simulators, but maintaining the extractions and optimization steps in your IC-CAP model file, you should set us the netlist(s) so that the Parameter list is the same. In this case, your extraction routines and optimizer settings stay the same, even when exchanging the netlists.

SPICE	ADS	IC-CAP Modelfile Parameter Table
.subckt test 1 2	define test 1 2	
R1.R 1 11 .2	R:R1 1 11 R=.2	R1.R 0.2
L1.L 11 12 50p	L:L1 11 12 L=50p	L1.L 50p
C1.C 12 2 31f	C:C1 12 2 C=31f	C1.C 31f
.ENDS	end test	

And here an example for the corresponding, common parameter extraction PEL code for SPICE and ADS:

R1.R = 4711

UCB SPICE Simulators Netlists (detailed)

For UCB Spice (University of California Berkeley), start every netlist with a comment line, for example
*this is my great circuit

Table1:SPICE Element Component Specifications

Component	General Form	Example
Resistor	RXXXXXXX N1 N2 VALUE <TC=TC1<TC2>>	R1 1 2 1000 TC=0.001,0.015
Capacitor	CXXXXXXX N+ N- VALUE <IC=INCOND>	COSC 15 2 10U IC=3
Inductor	LXXXXXXX N+ N- VALUE <IC=INCOND>	LSHUNT 3 29 10U IC=15.7m
Mutual Inductor	KXXXXXXX LYYYYYYY LZZZZZZZ VALUE	K43 LAA LBB 0.999
Transmission Line	TXXXXXXX N1 N2 N3 N4 Z0=VALUE <TD=VALUE>+ <F=FREQ <NL=NRMLEN>> <IC=V1,I1,V2,I2>	T1 1 0 2 0 Z0=50 TD=10NS
Linear Voltage- Controlled Current Source	GXXXXXXX N+ N- NC+ NC- VALUE	G1 2 0 5 0 0.1M
Linear Voltage- Controlled Voltage Source	EXXXXXXX N+ N- NC+ NC- VALUE	E1 2 3 14 1 2.0
Linear Current- Controlled Current Source	FXXXXXXX N+ N- VNAM VALUE	F1 13 5 VSENS 5
Linear Current- Controlled Voltage Source	HXXXXXXX N+ N- VNAM VALUE	HX 5 17 VZ 0.5K
Independent Voltage Source	VXXXXXXX N+ N- <<DC> DC/TRAN VALUE> + <AC <ACMAG <ACPHASE>>>	VIN 12 0 DC 6
Independent Current Source	IXXXXXXX N+ N- <<DC> DC/TRAN VALUE> + <AC <ACMAG <ACPHASE>>> + SFFM(0 1 10K 5 1K)	ISRC 23 21 AC 0.333 45.0

Table2:SPICE Semiconductor Component Specifications

Component	General Form	Example
Junction Diode	DXXXXXXX N1 N2 MNAME + <AREA><OFF><IC=VD>	DCLAMP 3 7 DMOD 3.0 IC=0.2
BJT	QXXXXXXX NC NB NE <NS> MNAME + <AREA> <OFF> <IC=VBE,VCE>	Q2A 11 26 4 20 MOD1
JFET	JXXXXXXX ND NG NS MNAME + <AREA> <OFF> <IC=VDS,VGS>	J1 7 2 3 JM1 OFF
MOSFET	MXXXXXXX ND NG NS NB MNAME + <L=VAL><W=VAL><AD=VAL><AS=VAL> +<PD=VAL><PS=VAL><NRD=VAL><NRS=VAL> + <OFF> <IC=VDS,VGS,VBS	M1 2 9 3 0 MOD1 L=10U W=5U

Note: see also the UCB Spice Maunal in section Spice/Publications.

User-defined Model Equations in Berkeley SPICE 3

see the ModelFile: demo_features\6_SIMULATORS\1_UCB_Spice\user_def_equat.mdl

Nonlinear resistor in Berkeley SPICE syntax:

$i(v) = 1.5mA \cdot v + 5mA \cdot v^2 + 2m \cdot \sqrt{v}$

The UCB SPICE3 syntax is:
.SUBCKT nonlinR HI=1 LO=2
Bnonlinr 1 2 I = A*(V(1)-V(2)) + B *(V(1)-V(2))^2 + C * SQRT(V(1)-V(2))

NOTE:
(Bxxx is a user-defined SPICE model, which may contain either a voltage (V=...) or a user-defined current (I=...))
.ENDS

However, the IC-CAP circuit parser has some problems with this line, and prints an error.
Therefore, in the spice deck line below, #echo switches the IC-CAP parser off for the whole line.
As a consequence, however, IC-CAP will not detect the parameters
A, B, C etc. and export them into the Model Parameters area.
This means, they cannot be extracted or optimized from within IC-CAP !
Therefore, the $mpar statement tells IC-CAP to export the specified parameter into the Model Parameters table what is between the parenthesis of this command.
.SUBCKT nonlinr HI=1 LO=2
#echo Bnonlinr 1 2 I=$mpar(A=1.5m)*(V(1)-V(2))
$mpar(B=5m)*(V(1)-V(2))^2+$mpar(C=2m)*SQRT(V(1)-V(2))
.ENDS

Gummel-Poon iC characteristic in Berkeley SPICE syntax:

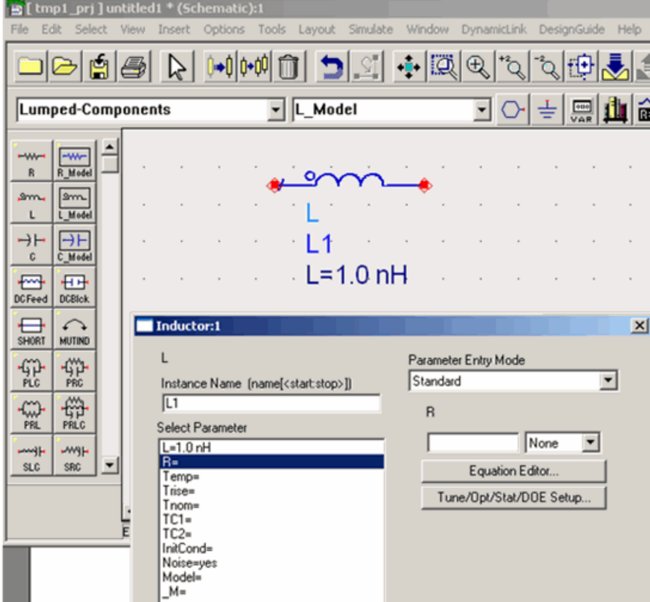
$iC(v_{BE}) = IS \cdot e^{\frac{v_{BE}}{NF \cdot VT}}$

.SUBCKT fgummelC C=1 B=2 E=3

#echo BfgumC 1 3 I=$mpar(IS=1E-25)**exp((V(2)-V(3))/($mpar(NF=1)**
$mpar(VT=1)))
.ENDS

ADS Simulator Netlist Syntax (detailed)

Due to its HF application, ADS uses a somewhat modified circuit deck syntax compared to SPICE.
For example, an inductor in SPICE is a pure, ideal inductor, while in ADS, an inductor for example may include also a resistance, a sophisticated temperature model etc., as visible in the ADS screenshot below:



Therefore, the ADS syntax for such an inductor reads:

L:myInductor node1 node2 L=1n

while it is in Spice simply:

L1 1 2 1n

Getting a list of all ADS components/models:

In order to get an ASCII text file with all ADS model descriptions included, apply the following script
FOR UNIX:

```
#!/bin/sh
####################################################################################
# Author: Rene Stoll, Agilent EEsof
# created: Jan.5, 2000
# Description: generate an ASCII file containing the syntax for ADS components
# How2use:     copy this script to your operating system and run it
####################################################################################
#
# Set path and environment
# Check if OS is SunOs 4.1.3, Solaris 2.5.1, HPUX9.x or HPUX 10.x
# and set the EESOF_DIR variable accordingly so that the proper code is used.
HPEESOF_DIR=/opt/ads
LM_LICENSE_FILE=/opt/licenses/license.dat
#
export HPEESOF_DIR LM_LICENSE_FILE
PATH=$HPEESOF_DIR/bin:/bin:/usr/bin:$MATLAB/bin:/opt/aCC:/opt/aCC/bin
export PATH
SHLIB_PATH=$HPEESOF_DIR/lib/hpux10:$HPEESOF_DIR/hptolemy/lib.hpux10
export SHLIB_PATH
#execute a 'hpeesofsim -h all to see all components available
hpeesofsim -h all  > /tmp/ADS_syntax.txt
```

For WINDOWS:

```
REM ##################################################################################
REM Author: Rene Stoll, Agilent EEsof
REM created: Dec.18, 2007
REM Description: generate an ASCII file containing the syntax for ADS components
REM How2use:     copy this script to your operating system and run it
REM ##################################################################################
set LM_LICENSE_FILE=C:\agileesof\licenses\license.lic
set AGILEESOF_LICENSE_FILE=%LM_LICENSE_FILE%
set HPEESOF_DIR=C:\agileesof\ADS2006A
PATH=%HPEESOF_DIR%\bin;%HPEESOF_DIR%\adsptolemy\lib;%HPEESOF_DIR%\adsptolemy\lib.win32;%PATH%
hpeesofsim.exe -h all >  C:\temp\ADS_ModelsSyntax.txt
```

See also the IC-CAP ModelFile
demo_features\6_SIMULATORS\2_ADS\5_NETLIST_TEMPLATES\00_ADS_Models_and_Netlist_Syntax.mdl

See also the list at
demo_features\6_SIMULATORS\2_ADS\5_NETLIST_TEMPLATES\00_ADS_Models_and_Netlist_Syntax.txt

User-defined Model Equations in ADS

Using ADS sdd's (Symbolic defined devices)

For an IC-CAP example, see
demo_features\6_SIMULATORS\2_ADS\5_NETLIST_TEMPLATES\0_USER_DEF_MODELS\1_sdd\1_Diode_sdd.mdl

Example: a diode, but DC model only:

Custom Modeling of a diode using sdd's

sdd = symbolic defined device

- sdd's allow to specify model equations in a simple basic-like language, which are sent from IC-CAP to ADS, and interpreted by ADS on the fly.
- No compilation is required.
- Therefore, sdd's are an ideal tool for developing models.

sdd syntax for a diode DC model:

```
define diode_SDD_tutor_DC_only (A C)

;defining some constants for later use in the SDD equations
temp=27
vt=(8.62e-5*(temp + 273.15))

;parameters for custom sdd
IS =100a
NF =1

;standard devices as part of this subcircuit
R:RS    A    A1  R=.6

;  → here starts the details of the SDD ←
vdc  =(_v1 - _v2)    ; _v1 is the volt. at node A, _v2 at node C (see define... line)

;define DC current
id = (IS * (exp(vdc / (NF*vt))-1))


;finally, calling the SDD feature in ADS, with the properties defined above
SDD:diodeDConly  A1 0 C 0 I[1,0]=id   I[2,0]=-id

end diode_SDD_tutor_DC_only
```

5/9/2009 Agilent Technologies 4

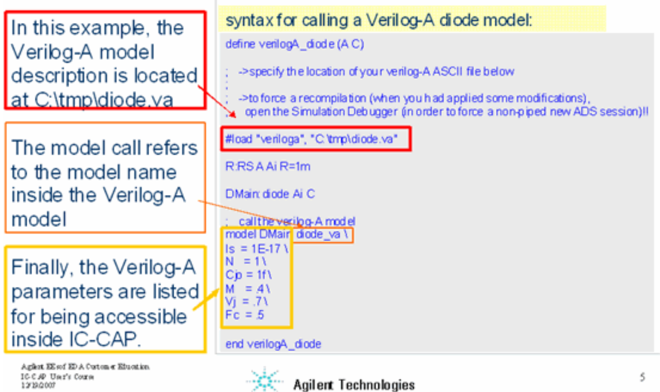
Below the full diode sdd netlist, including charges, as it has to be specified in the IC-CAP ModelFile Circuit tab:

```
define diode_SDD_tutor (A C)
;defining some constants for later use in the SDD equations
temp=27
vt=(8.62e-5*(temp + 273.15))
;parameters for custom sdd
IS =100a
NF =1
CJO=200f
VJ =.7
MJ =400m
FC =500m
TT =2p
;standard devices as part of this subcircuit
R:RS      A         A1   R=.6
L:LS      A1      A11   L=1p
C:CAC     A         C   C=1f
; here starts the details of the SDD
vdc  =(_v1 - _v2)
;define DC current
id = (IS * (exp(vdc / (NF*vt))-1))
;define diode depletion charge (like most simulators, ADS prefers charges instead of capacitances)
depl_charge(cj,v,mj,vj,fc) = (if (v > fc*vj) then depl_charge2(cj,v,mj,vj) else
depl_charge1(cj,v,mj,vj) endif)
depl_charge1(cj,v,mj,vj) = (cj*vj/(1-mj)*(1-(1-v/vj)^(1-mj)))
depl_charge2(cj,v,mj,vj) = (cj*vj/(1-mj)*(1-(1-FC)^(1-mj))+cj/(1-FC)^mj*(v-FC*vj)     \
                     +.5*cj *mj/vj/(1-FC)^(mj+1)*(v-FC*vj)^2)
;define diode diffusion charge
; NOTE: diff.cap = TT* gm = TT * IS / (N * vt) * exp ( v / (N * vt)),
;  but we need the charge for the sdd diff.charge = TT*id(vdc)!
diff_charge(v,is,n,vt,tt) = (tt*is*exp(v/(n*vt)))
charge(cj,v,mj,vj,fc,is,n,vt,tt) = depl_charge(cj,v,mj,vj,fc) + diff_charge(v,is,n,vt,tt)
;finally, calling the SDD feature in ADS, with the properties defined above
SDD:diodeDCnCV  A11 0 C 0 I[1,0]=id   I[2,0]=-id    I[1,1]= charge(CJO,vdc,MJ,VJ,FC,IS,NF,vt,TT)
\
                I[2,1]= -charge(CJO,vdc,MJ,VJ,FC,IS,NF,vt,TT)
end diode_SDD_tutor
```

Using Verilog-A

For an IC-CAP example, see
demo_features\6_SIMULATORS\2_ADS\5_NETLIST_TEMPLATES\0_USER_DEF_MODELS\3_VERILOG_A\diode_VerilogA.mdl

The slide below shows what has to be specified in the IC-CAP ModelFile Circuit tab:



Below the contents of file C:\tmp\diode.va

```
// Make reference to verilog-A discipline 'electrical' (nodes)
`include "disciplines.vams"
module diode_va(anode,cathode);
electrical anode,cathode;
parameter real Is=1e-14 from [0:inf];
parameter real N=1      from [0:10];
parameter real Cjo=0    from [0:inf];
parameter real M=0.5    from [0.1:1];
parameter real Vj=0.7   from [0.1:10];
parameter real Fc=0.5   from [0.4:0.95];
parameter real Tt=1p    from [0:inf];
real Vd, Id, Qd;
real f1, f2, f3, fcp;
analog begin
    f1  = (Vj/(1 - M))*(1 - pow((1 - Fc), (1 - M)));
    f2  = pow((1 - Fc), (1 + M));
    f3  = 1 - Fc*(1 + M);
    fcp = Fc * Vj;
    Vd = V(anode,cathode);
// intrinsic diode.
    if (Vd < 0)
      Id = - Is;
    else
      Id = Is * (exp( Vd / (N * $vt)) - 1);
// capacitance (junction and diffusion).
    if (Vd <= fcp)
      Qd = Tt * Id + Cjo * Vj * (1 - pow((1 - Vd / Vj), (1 - M))) / (1 - M);
    else
      Qd = Tt * Id + Cjo*( f1 + ( 1 / f2 ) * ( f3 * ( Vd - fcp ) + ( 0.5 * M / Vj ) * ( Vd * Vd -
fcp * fcp)));
    I(anode,cathode) <+ Id;
    I(anode,cathode) <+ ddt(Qd);
end
endmodule
```

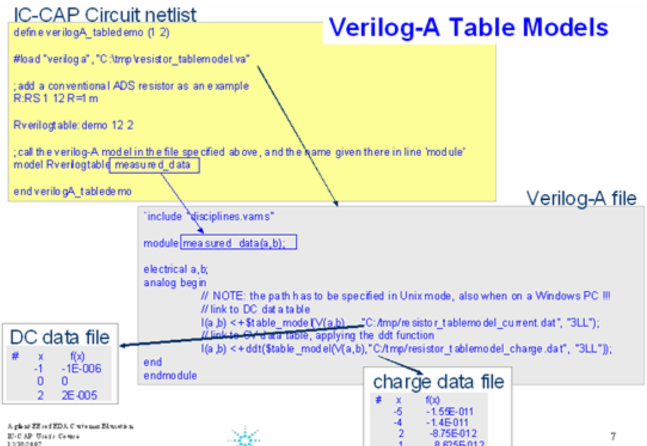
Since ADS2005A,
table models are also supported by Verilog-A

This means that from measured currents and charges, table models can be obtained very easily.

Different methods are selectable for interpolating between and outside the measured data points:
- Interpolation Methods: Linear spline, Quadratic spline, Cubic spline
- Extrapolation Methods: Clamp extrapolation, Linear extrapolation (default)

Agilent Technologies 6

For an IC-CAP examples see:
demo_features\6_SIMULATORS\2_ADS\5_NETLIST_TEMPLATES\0_USER_DEF_MODELS\3_VERILOG_A\table_models\resistor_tablemodel_VerilogA.mdl



Verilog-A Table Models
IC-CAP Circuit netlist
define verilogA_tabledemo (1 2)
#load "veriloga","C:\tmp\resistor_tablemodel.va"
;add a conventional ADS resistor as an example
R:RS1 12 R=1m
Rverilogtable:demo 12 2
;call the verilog-A model in the file specified above, and the name given there in line 'module'
model Rverilogtable measured_data
end verilogA_tabledemo
Verilog-A file
`include "disciplines.vams"
module measured_data(a,b);
electrical a,b;
analog begin
// NOTE: the path has to be specified in Unix mode, also when on a Windows PC !!!
// link to DC data table
I(a,b) <+ $table_model(V(a,b), "C:/tmp/resistor_tablemodel_current.dat", "3LL");
applying the ddt function
I(a,b) <+ ddt($table_model(V(a,b),"C:/tmp/resistor_tablemodel_charge.dat", "3LL"));
end
endmodule
DC data file
x f(x)
-1 -1E-006
0 0
2 2E-005
charge data file
x f(x)
-5 -1.55E-011
-4 -1.4E-011
2 -8.75E-012
1 8.625E-012
7

Netlist Syntax for ADS and SPICE

- Comparing Berkeley UCB SPICE and ADS Netlist Syntax
- How to have a common Parameter List for Spice and ADS
- Berkeley UCB SPICE Simulator Netlists (detailed)
- User-defined Model Equations in Berkeley UCB SPICE 3
- ADS Simulator Netlist Syntax (detailed)
- User-defined Model Equations in ADS

CV and S-Parameters

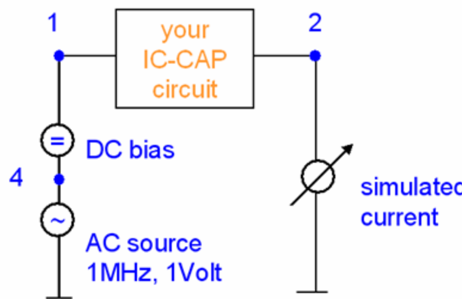
- *CV Curves form SPICE* (iccapmhb)
- *S Parameter Curves form SPICE* (iccapmhb)

How to Obtain CV Curves from SPICE

Calculating CV Curves from Complex Currents

IC-CAP file: demo_features\6_SIMULATORS\0_GENERAL\calc_cap_vs_volt_SPICE.mdl

From an inspection of the Simulation Debugger for spice2, IC-CAP surrounds your Circuit Netlist by the following extra circuitry:



The simulated current result is then interpreted like this:
IC-CAP assumes a situation like with the default LCRZ meter setup (!):
-> a capacitance in parallel with an conductance.

From the simulated current of the schematic above, we obtain the complex admittance

Y = iac / v.

Since v=1, the equation is simply

Y = iac

Therefore, the capacitance value is calculated within IC-CAP from the imaginary part of the current after

C = IMAG(Y) / (2*PI*1MHz) ----------- (1)

and the conductance is

G = REAL(Y) ------------- (2)

IC-CAP applies formula (1) when 'Mode' is set to 'C', and formula (2) when 'Mode' is set to 'G'.

How to Obtain S-Parameter Curves from SPICE

Calculating CV Curves from Complex Voltages

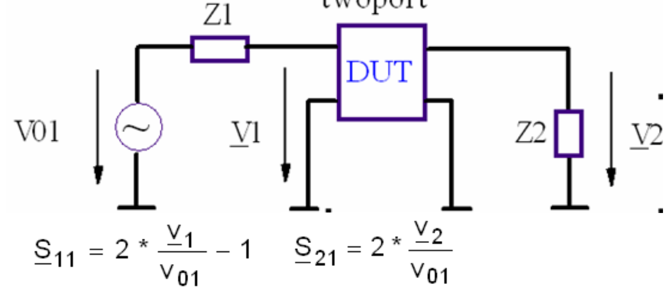
IC-CAP file: demo_features\6_SIMULATORS\0_GENERAL\calc_Spar_vs_freq_SPICE.mdl

As a fact, most SPICE-like, time-based simulators cannot simulate S-parameters. But they can simulate the frequency behavior of circuits in magnitude and phase (or real and imaginary numbers). And IC-CAP can calculate the S-parameters out of these numbers.

How this is done is explained below:

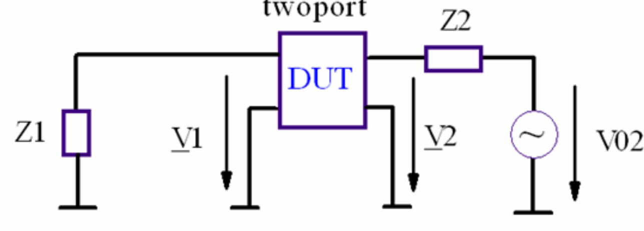
Note: two subsequent simulations are needed in order to obtain one set of S-parameters.

- **1st simulation:** Z1 = Z2 = Z0, e.g.Z0 = 50 Ohm



$$S_{11} = 2 * \frac{V_1}{V_{01}} - 1 \qquad S_{21} = 2 * \frac{V_2}{V_{01}}$$

- **2nd simulation:** Z1 = Z2 = Z0, e.g.Z0 = 50 Ohm



$$S_{12} = 2 * \frac{V_1}{V_{02}} \qquad S_{22} = 2 * \frac{V_2}{V_{02}} - 1$$

This is the reason why IC-CAP has to insert the user-defined circuit (2-port DUT) as a subcircuit into a hyper-circuit and to finally send this total circuit to the simulator. In this manner, the simulator output deck will provide the complex voltages that are required by IC-CAP to calculate the S-parameters.

To speed-up simulations, IC-CAP puts these two schematics together into one big circuit. This 'big' circuit is depicted below in fig.3 using SPICE syntax. Its structure is valid for all other simulator circuit descriptions in IC-CAP, only the syntax may be different. For MDS/ADS, however, these complex calculations are not required, because both simulators support S-parameter simulations directly.
The example shows a S-parameter simulation for a bipolar transistor with inputs VB,VE,VC,VS and frequency. For simplicity, the bias is kept constant and only the frequency is swept.

see IC-CAP file spar_from_volt.mdl

THE USER_DEFINED CIRCUIT DESCRIPTION:

```
Q1 1=C 2=B 3=E 4=S NPN
.MODEL NPN NPN
+ IS  = 2.704E-16
+ BF  = 86.16
+ NF  = 0.979
+ VAF = 86.95
```

Fig.1: The user-defined circuit for the example below (a simple bipolar transistor)
NOTE: no subcircuit description used for simplification.

THE USER_DEFINED SETUP DESCRIPTION LOOKS LIKE THIS:

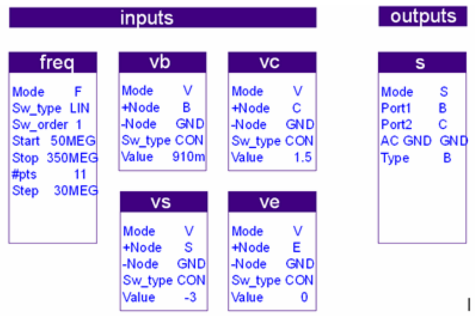


Fig 2: User-defined IC-CAP setup, from which IC-CAP generates the corresponding SPICE deck

Notes on the names of the circuit elements of the following pages:
The user-defined 'Circuit' description consists simply as shown above of a bipolar transistor with 4 connections E,B,C and S. Therefore nodes 1 .. 4 (or 1 .. n in a more general case) as well as 5 ... 8 (or n+1 .. 2n) of the total circuit generated by IC-CAP are used to link the user-defined circuit into the 'overhead circuit'.

Nodes 9 .. 11 (or generally 2n+1 .. 2n+3) as well as 12 .. 14 (or 2n+4 .. 2n+6) of the IC-CAP circuit are used to calculate the voltages required for equ. (1) - (4)
(or in other words to 'emulate' the 'S-parameter test set' of the network analyzer).

The simulator node names as defined by the user in the input fields determine the effective name of the elements in the IC-CAP 'main' circuit, e.g. VCGRO: a voltage from user-defined node C to GROund or LBGRO: an inductor L from user-defined node B to GROund.
node numbers valid for data/bjt_npn.mdl file

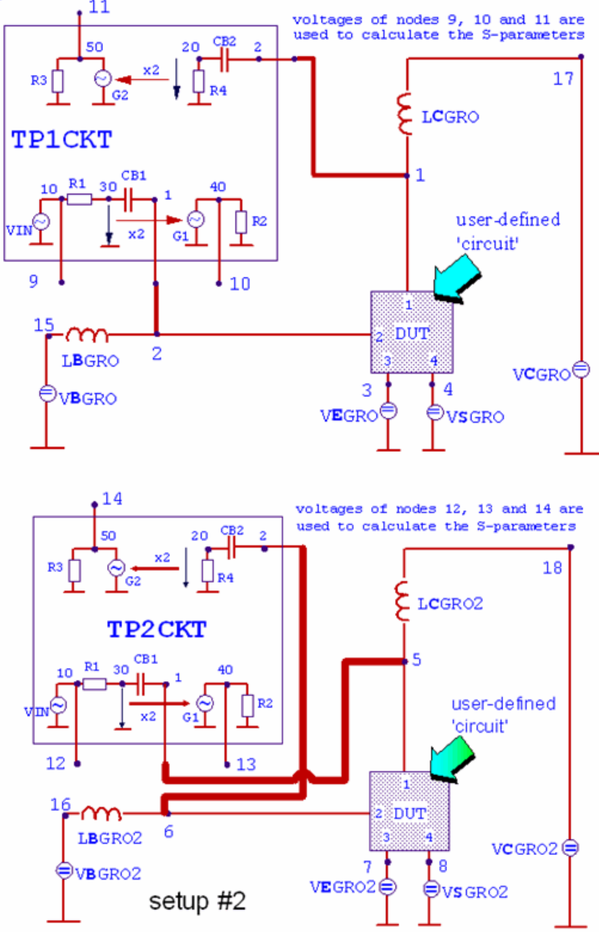


Fig 3: Spice circuit generated by IC-CAP in order to be able to calculate S-parameters out of complex voltages.

The values of LxGROx and CxGROx are defined by the IC-CAP system variables TWOPORT_L and TWOPORT_C, all resistors in TPxCKT that have 50Ohms can be set to a different value when defining the IC-CAP system variale TWOPORT_Z0.

The corresponding UCB SPICE input file generated by IC-CAP and sent to SPICE:

only this first line is the user-defined circuit description!
all the rest has been added by IC-CAP

```
.MODEL npn NPN  IS=2.704E-16  BF=86.16  NF=0.979  VAF=86.95
QCKT  1 2 3 4 npn
QCKT2 5 6 7 8 npn
* START SOURCES
.SUBCKT TP1CKT 1 2 10 40 50
CB1 1 30 1e-05
R1 30 10 50
R2 0 40 1
R3 0 50 1
CB2 2 20 1e-05
R4 20 0 50
G1 0 40 30 0 2
G2 0 50 20 0 2
VIN 10 0 AC 1
.ENDS
.SUBCKT TP2CKT 1 2 10 40 50
CB1 30 2 1e-05
R1 10 30 50
R2 40 0 1
R3 50 0 1
R4 20 0 50
CB2 1 20 1e-05
G1 0 40 30 0 2
G2 0 50 20 0 2
VIN 10 0 AC 1
.ENDS
XTP1CKT 2 1 9 10 11 TP1CKT
XTP2CKT 6 5 12 13 14 TP2CKT
LBGRO 2 15 0.0001
VBGRO 15 0 DC 0.91
LBGRO2 6 16 0.0001
VBGRO2 16 0 DC 0.91
LCGRO 1 17 0.0001
VCGRO 17 0 DC 1.5
LCGRO2 5 18 0.0001
VCGRO2 18 0 DC 1.5
VEGRO 3 0 DC 0
VEGRO2 7 0 DC 0
VSGRO 4 0 DC -3
VSGRO2 8 0 DC -3
* END SOURCES
.AC LIN 11 5e+07 3.5e+08
.PRINT AC V(10) V(9) V(11) V(14) V(13) V(12)
.END
```

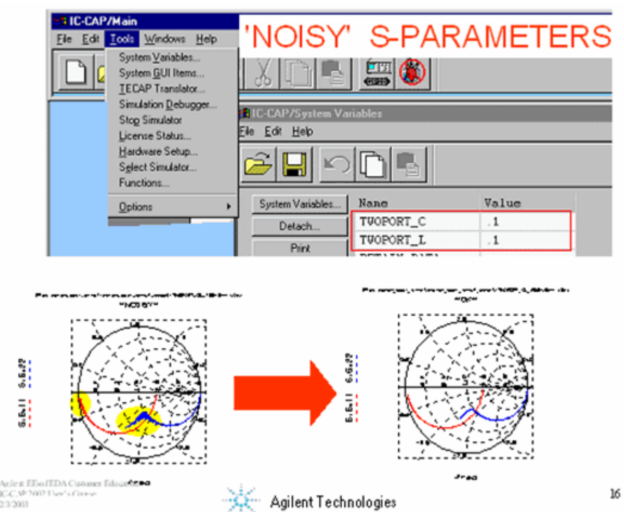
A word on 'Noisy S-parameters' for Spice and spice-like simulators:

As was explained above, IC-CAP needs to emulate a network analyzer and its bias-TEE when requesting complex voltages from spice-like simulators, so that IC-CAP can calculate the S-parameters by itself.

A special problem might now occur with the limited precision of numbers in the computer chip. The default value of Twoport_C and Twoport_L is 100 each in IC-CAP. In the example above, these components became LB_GRO, LC_GRO, and LB_GRO2, LC_GRO2.

Provided the components capacitance is very small, e.g. some 10fF, such a processor rounding effect can happen and as a result, the simulated S-parameters may look like 'noisy', as depicted below. This usually happens in the SPICE simulator solver matrix, where the big values (100F, 100H) together with these small DUT capacitance values lead to numerical rounding effects. To prevent this, specify the two System Variables (under IC-CAP/Main) and set their values to lower ones, e.g. 100m each.

However, do not set them to too small values (!) . Both, TWOPORT_L and TWOPORT_C must be big enough, so that their influence (phase shift due to the resonance of TWOPORT_L and TWOPORT_C) is over for the lowest simulation frequency. This may be critical for low-frequency NWAs who's lowest frequency is in the 30kHz range.



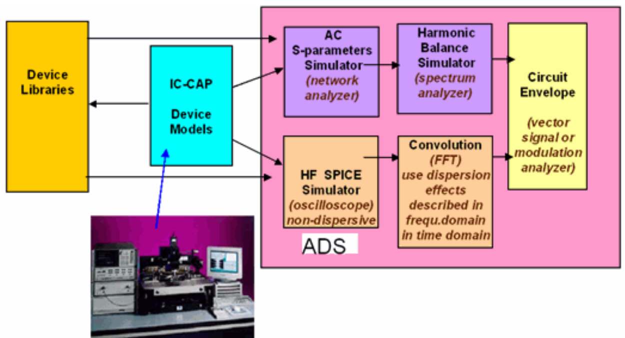
Keep in mind that adding bias-TEEs means a resonance at a few 10 Hz (the L and C of the bias TEE), and that the phase shift of this resonance must be over at the lowes measured/simulated NWA frequency. In other words, what you measure with a NWA and bias TEEs (internal or external ones), represents the S-parameter trace of your DUT **after** a complete 360' phase shift.

Harmonic Balance

ADS

ADS, as an example for a powerful RF simulator, exhibits a group of simulators, which are called automatically according to the simulation task
The total tool chain of ADS plus the IC-CAP part is given below:

The ADS simulator tool chain and the positioning of IC-CAP



Comparing to conventional SPICE-like simulators, ADS' high-frequency SPICE offers these features:

- uses high-frequency models for microstrip lines, bends, gaps and other discontinuities
- used to analyze steady-state response of mixers, oscillators, amplifiers etc.
- frequency dependent components are modeled with approximations that neglect some of the frequency dependent effects, such as dispersion and high frequency loss. This results in faster simulation results, which are nevertheless accurate enough for electrically small components on an IC.
- requires the linear simulator license

The Convolution simulator offers:

- includes dispersion effects and high frequency loss
- results in accurate high frequency results at the expense of longer simulation time
- handles circuits containing distributed elements and S-parameter data used for components
- can accurately analyze circuit start-up and transient conditions at low and high frequencies, where the effects of dispersion and discontinuities are significant (for electrically big components on printed circuit boards

The AC S-parameter simulator allows linear S-parameter simulations. Being part of a RF simulator, enhanced RF models are available.

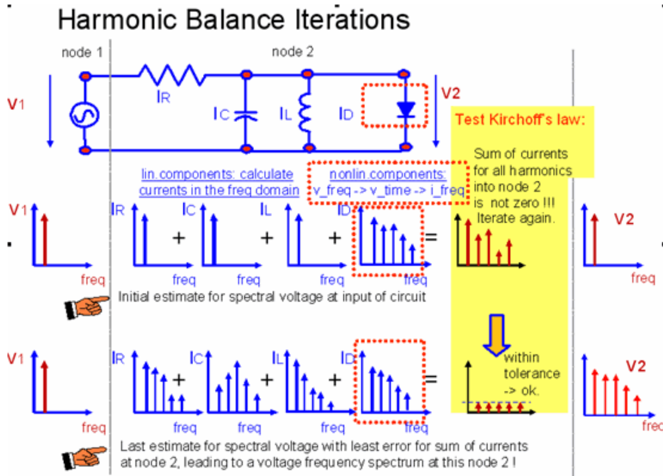
What is Harmonic Balance (HB) Simulation?

Harmonic balance (HB) is a frequency-domain simulation technique. Signals are treated as DC plus summations of finite numbers of sine waves, i.e. as periodic signals. Harmonic balance simulations use one or two sine-wave stimuli to determine (for example) conversion loss vs. drive level, distortion vs. frequency, and gain compression.

Harmonic Balance and Time-Domain Simulation
Harmonic balance calculates voltages and currents in much the same way as time-domain simulators such as SPICE do. Frequency-wise, it solves for the magnitudes and phases of all spectral lines simultaneously. In effect, the entire waveform is being solved for at once. Harmonic balance is extremely efficient if the signals are simple in the frequency domain- as they are, for example, when an amplifier is driven by a sine wave.

Circuits that have non-repeating transient behavior are best analyzed using ADS-Impulse, a time-domain (transient) simulator. In general, harmonic balance is not very useful for analyzing signals that might be analyzed using an arbitrary waveform generator or a wide- bandwidth oscilloscope. ADS-Impulse is also useful for verifying the dynamic performance of a circuit after it has been designed and optimized using harmonic balance.

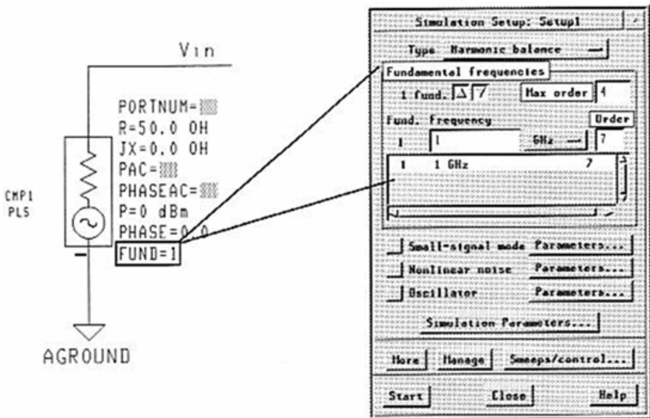
What happens when a nonlinear harmonic balance simulation is executed?



For the circuit given above, the following iterations are performed:

1. First, a DC simulation is executed to obtain the DC operating point of the circuit.
2. A linear AC analysis (Harmonic Balance Fourier Analysis) is applied to analyze the linear and passive components in the frequency domain, stimulated by the fundamental frequency.
3. Nonlinear device models are usually defined as i=f(v). Therefore, the actual voltage frequency spectrum across the nonlinear component is Fourier-transformed into the time domain. This resulting time domain signal is applied to the nonlinear device with a Spice-type simulation. The resulting time domain current is Fourier-converted back to the frequency domain.
4. The sum of the actual current spectra at each circuit node is calculated and it is checked, if this sum is below a certain tolerance limit for all frequencies.
5. Steps 1-4 are iterated until Kirchhoff's current law is satisfied for all frequencies at all nodes. The same is true for fulfilling Kirchhoff's voltage law. Iterative, matrix-driven techniques are used (Newton-Raphson algorithms etc.).
6. The obtained solution is a satisfaction of Kirchhoff's current law (sum of all currents, for all frequencies, is zero for each node) as well as the voltage law. The mathematical analysis has converged. The harmonics are now balanced.

NOTE: Changing the value of an individual component changes the current and the harmonic content of the signal through the other components. It also changes the output voltage. But it does not affect the sum of the currents. This sum is always zero.



In the Agilent Advanced Design System (ADS), harmonic balance simulations assume that all voltage and current waveforms in the circuit are summations of sinusoidal signals. In other words, finite spectrums composed of individual spectral lines.

Energy is not allowed to exist continuously throughout a spectrum. It is discretized into ideal spectral lines that are the harmonics of more basic fundamental frequencies and intermodulation (sum and difference) products of those fundamental frequencies.

This section offers a brief introduction to the parameters that affect the frequencies in the analysis: **FUND**, **ORDER**, and **MAXORDER**.

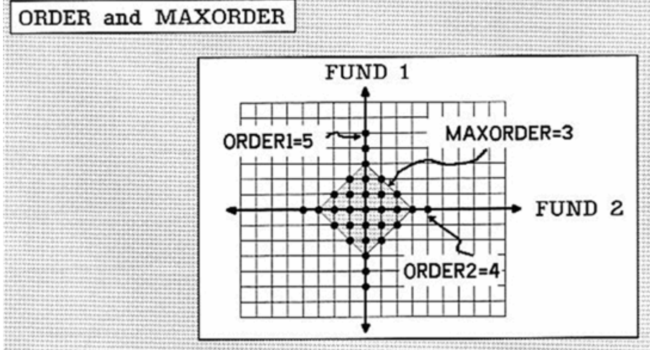
Signal Sources
In harmonic balance simulations, the voltage, current, and power sources in the circuit specify the amplitude of the signal, not the frequency (at least not directly).
In harmonic balance simulation, frequency information is specified in the simulation setup, as part of specifying the number of fundamental frequencies that are being used and their values.
Signal sources for harmonic balance simulation have a parameter FUND that relates the harmonic balance simulation information to the signal source.
FUND is a small integer, and is likely to be 1 or 2.

If FUND is set equal to 1 on a VLS large-signal source, the source produces a sine wave whose frequency is the same as the first fundamental frequency of the harmonic balance simulation component (for example, 1 GHz).

Fundamental Frequencies, or Tones
Spectral lines in harmonic balance simulation results are almost always linear combinations of one or more fundamental frequencies:

FREQ of one spectral line = $\Sigma(\pm m)$*FREQUENCY1 + $(\pm n)$*FREQUENCY2 + ...

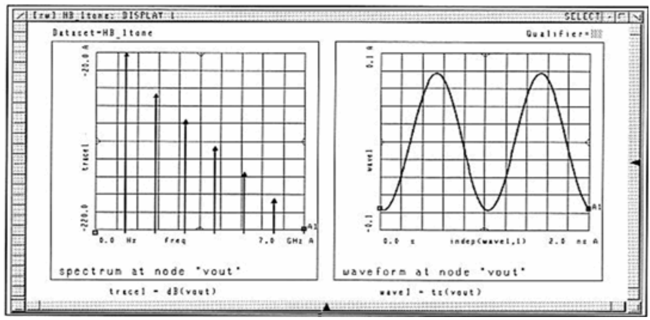
The number of frequencies is important mostly because each additional fundamental frequency means a significant increase in memory usage, simulation time, and dataset storage size. Choose the smallest number that still gives the correct results.
In ADS, fundamental frequencies are also known as tones. Thus, in a one-tone harmonic balance simulation, only one fundamental frequency (FREQ), multiples of that frequency (its harmonics), and DC (the harmonic of order 0) are used.
One-tone harmonic balance is the most common simulation type for amplifiers.



ORDER
In the simulation setup, the parameter ORDER specifies for each fundamental frequency the highest harmonic (largest multiple) of that fundamental frequency that will be considered by the simulator.

The Harmonic Balance Variable MAXORDER
While the parameter **ORDER** affects the number of harmonics for each fundamental, the parameter **MAXORDER** controls the number of intermodulation (mixing) products that will be computed in a two-tone or multi-tone simulation.
For example, if **MAXORDER** is set to 3, only first-order, second-order, and third-order mixing products will be computed.

Presenting the Results of Harmonic Balance Simulations



Dimensionality of Results
In presenting the results of harmonic balance simulations, keep in mind the fact that the node voltages and currents are complex spectrums, not single numbers.

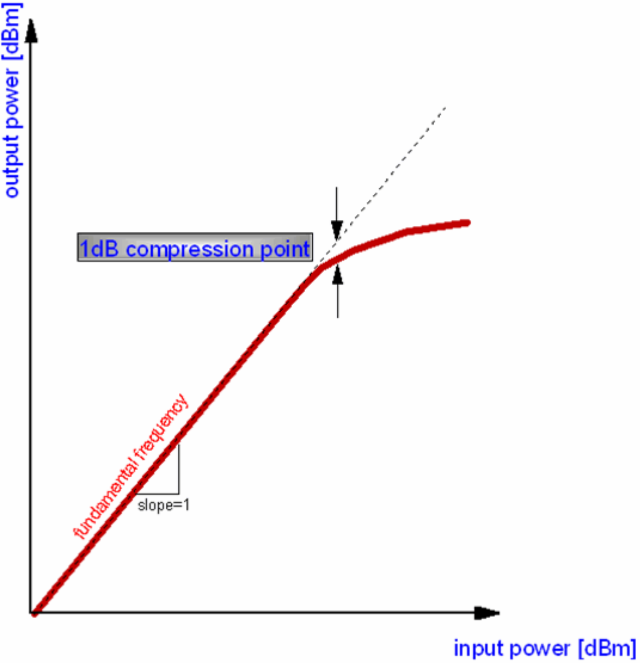
With no swept variables, a simple harmonic balance simulation produces a result with a frequency axis. Therefore, a sweep of (n) variables on the circuit page results in (n + l)-dimensional data.

Nonlinear High Frequency Characteristics

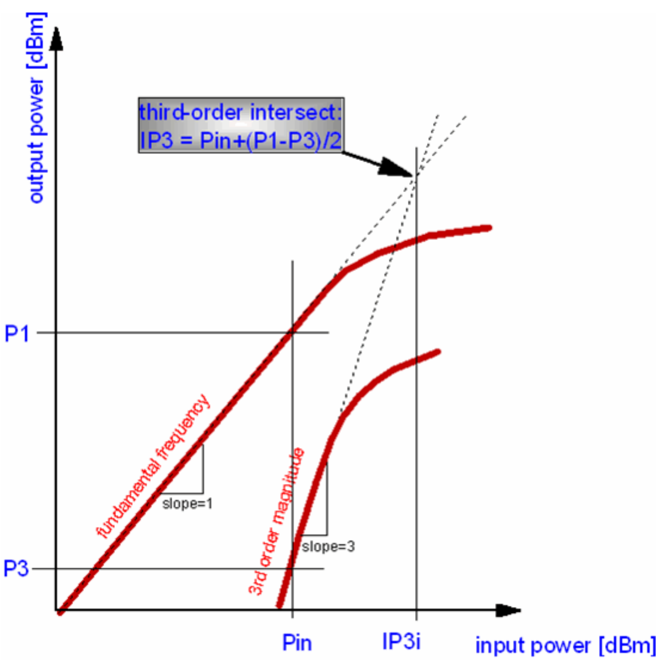
There are 4 main nonlinear high frequency characteristics

- Third Order Intersect (IP3)
- 1-dB Gain Compression (1 DBC)
- Saturated Power (Psat)
- Gain Compression at Saturation (GCS), which are explained with the sketches below:

1dB compression point, a measure for the 1st order harmonic distortion



Definition of the 3rd order intersect, IP3, a measure for 3rd order harmonic distortion



Special Discussion of the Third-Order intersect analysis techniques
The third-order intersect of an element or network is a widely accepted system design parameter because it indicates the degree of nonlinearity of a nonlinear element. The output to input voltage relationship for an electrical component can typically be described as a polynomial relationship like:

$$V_{out} = a0 + a1 * V_{in} + a2 * V_{in}^2 + a3 * V_{in}^3 + a4 * V_{in}^4 \ldots$$

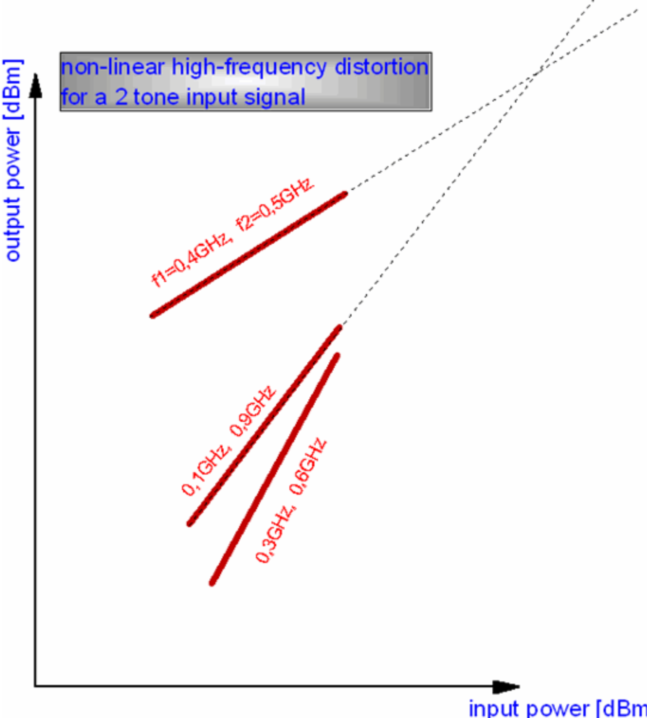
where
a0 output dc offset
al small signal gain
an higher order coefficients; n > 1

Fig.2, above, shows a typical nonlinear element output power versus input power curve for the fundamental and third-order output tones for a single tone input test signal. In the linear operating range of the amplifier, the small signal fundamental curve varies with a 1:1 slope. The small signal third-order curve varies with a 3:1 slope. The third-order intersect is that point where the extrapolated small signal fundamental and third-order curves intersect. At this third-order intersect one may be interested in the input power level, or in the output power level.

These curves are of main interest when performing a distortion analysis on a circuit. Here,

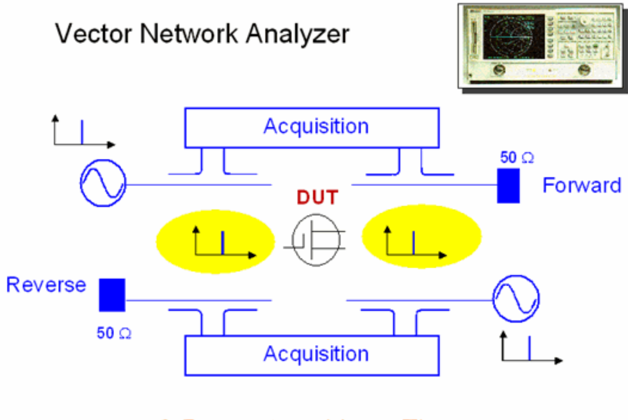
we can distinguish between a 1 tone and a 2 tone signal input. While the 1 tone result looks like the one given in figures 1 and 2, fig.3 shows the result for a 2 tone distortion analysis.

Distortion analysis for a two-tone input

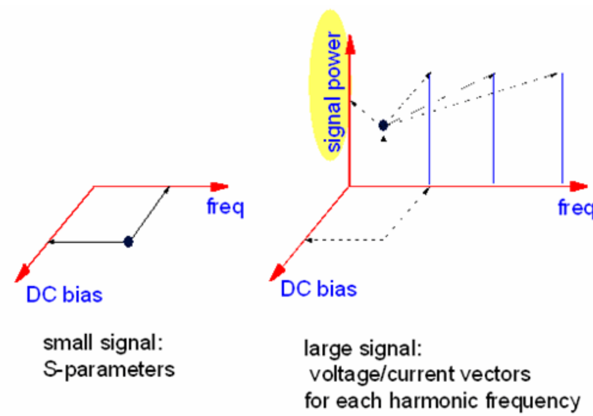


Harmonic Balance Simulations for Modeling

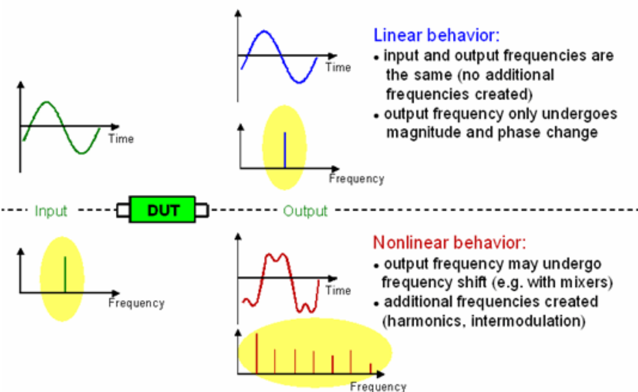
The importance of harmonic balance simulations and its application for modeling can best be seen when considering the measurement principle of a vector network analyzer.



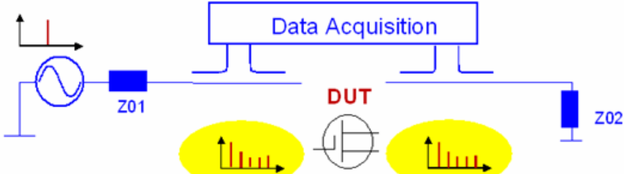
The measurement principle is assuming a linear behavior of the DUT. However, this is only true for passive components. For an active DUT, like a transistor, this is not quite true. Even for very low RF signal levels, i.e. in the -20dBm level, harmonics occur. This is sketched below:



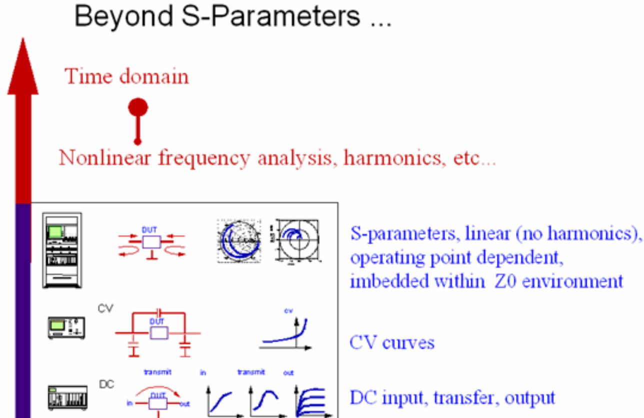
This nonlinear frequency behavior particularly becomes important when considering today's communication systems, where pulsed signals are used rather than CW sinusoidal signals like in the past.



Therefore, a nonlinear vector network analyzer should be used for characterization instead of the conventional S-parameter VNA measurements.

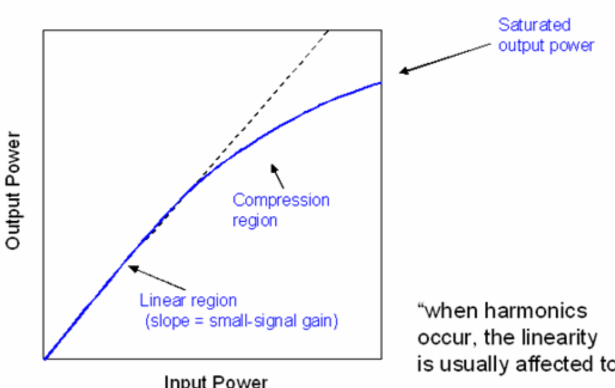


However, such nonlinear network analyzers are very expensive and sophisticated to handle. In many cases, modeling engineers therefore replace such systems by a conventional small-signal VNA measurements (-30dBm RF signal source power) plus additional measurements using a spectrum analyzers. The later, as a disadvantage, allow only to measure the magnitude of the harmonics, and not also the phase. In this case, the true time domain signal cannot be reconstructed out of the harmonics any more, what is possible when using a nonlinear network analyzer.

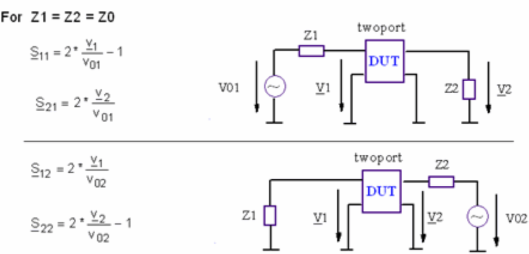


IMPACT OF HARMONIC BALANCE SIMULATIONS FOR CONVENTIONAL S-PARAMETER MODELING

When using a conventional linear vector network analyzer, only the base frequency is measured, and the frequency harmonics are ignored, due to the very small bandwidth of such a VNA. This means, that when harmonics occur, they are not part of the measured S-parameters and therefore can lead to a modeling problem which has not much to do with the real device under test.

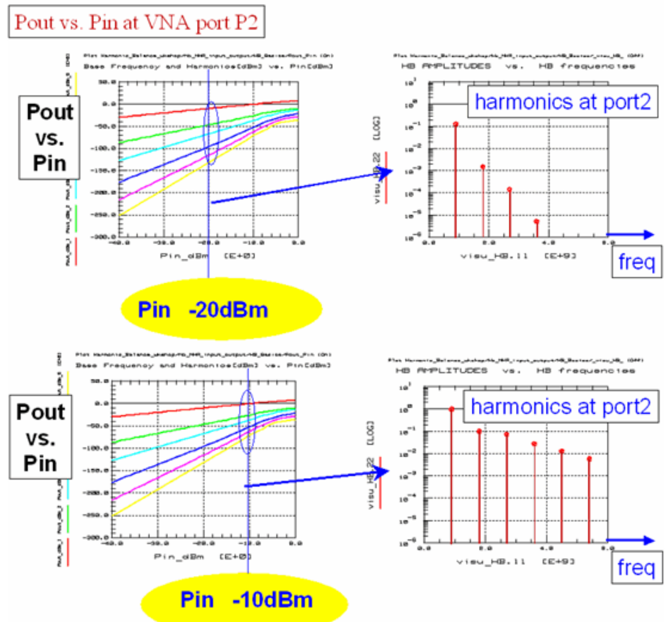


Taking this into account, it has been observed, that the transit frequency modeling of transistors can be performed much more accurately when taking this measurement situation into account. This means that no S-parameter SPICE simulations should be applied for curve fitting, but rather a harmonic balance simulation, from which only the base frequency is considered and compared to the measured VNA S-parameters. Furthermore, the harmonic balance DC operating point conditions can be compared in the same step with the measured DC Collector current or Drain current of the transistor. With these HB simulations applied to S-parameter measurements, the real measurement conditions and their limitations are reflected, and therefore, accurate modeling results for the ft modeling can be achieved.
It must be mentioned, that for such an application, the harmonic balance simulations has to be applied to a frequency sweep which is identical to the VNA measurement frequency steps.
Again, a SPICE simulation is a linearization and does not take the real existing non-linearities into account. It assumes no harmonics, and can therefore lead to wrong simulation results for ft modeling.



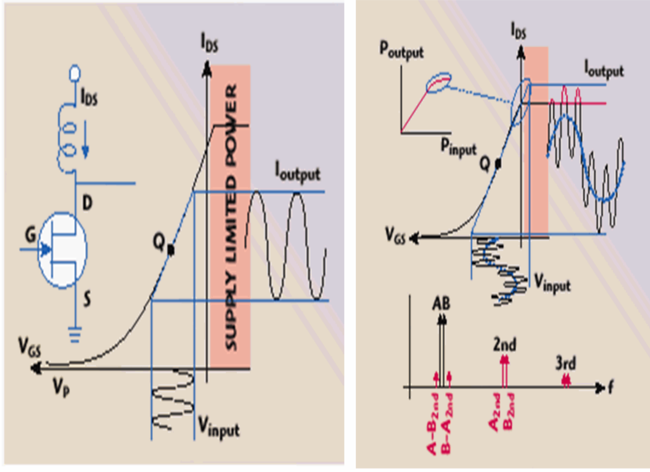
Final remark on using HB simulations for S-parameter modeling:
in this case, no OPEN or OPEN-SHORT de-embedding can be applied, because these twoport operations assume a linear circuit behavior. When applying HB simulations, the OPEN and the SHORT have to be modeled and added to the transistor circuit. Only in this case, the HB simulation will give a correct result.

After the measured S-parameter curves have been modeled using HB simulations, in a next step, the harmonics spectrum is modeled. Based on the mentioned spectrum analyzer measurements, a small fine-tuning of the RF model parameters will help to fit the harmonics vs. RF input power measurements.



With such a complete modeling, from DC, via CV curves, to S-parameters and corresponding HB simulations to finally fitting the harmonics versus RF power, the transistor is modeled as accurately as possible. Its parameter set can now be considered as fulfilling the requirements of its later application, i.e. its operation in the mW RF signal range (~0dBm).

Last not least, the following two sketches depict nonlinearities showing up when a sinusoidal signal or a two-tone sine function is applied to an amplifier, or a single transistor:



History

History of Circuit Simulators

The origins of Berkeley SPICE

SPICE (Simulation Program with Integrated Circuit Emphasis) was introduced in May 1972 by Prof. D. O. Pederson at the University of Berkeley, California.

It simulates the electrical behavior of circuits based on discrete components
In particular, these include:

- resistors, capacitors, inductors and coupled inductors
- ideal lines and cables
- independent and controlled current and voltage sources
- semiconductor components

It soon became a standard for electronic circuit simulation, used by both, the industry and the universities. This fact is based on the technical expertise of the developers of SPICE: especially Prof. Pederson /1/, and Prof. Nagel /2/.

After 17 revisions of SPICE, SPICE2 was introduced in 1975. The main improvements were the dynamical RAM allocation capabilities and the automatic time stepping for time-domain simulations. Especially SPICE version 2g6, was a specially reliable program version, written in Fortran. This version has become the root of many industry in-house simulators and also of many commercially available simulators. This explains why they typically understand the Berkeley syntax (Spectre, Hspice, Pspice, Eldo).

Finally, SPICE3, introduced by the end of the 80s, was the first version written in the C language.

One of the reasons for the success of SPICE is also the implementation of good transistor models. This is especially true for the bipolar model. The Gummel-Poon model has only become popular after it was implemented in SPICE. (However, some simplifications had been made to the original paper of Gummel-Poon.) The same is true for MOS transistor, however, after MOS3, and the introduction of the BSIM series of model, SPICE's model quality had suffered a lot compared to MOS2 and MOS3. Finally, with BSIM3v3, Berkeley SPICE gained back its authority of model definition.

1. Nagel, L.W.; Pederson, D.O.: SPICE. Berkeley, Univ. of California, Electronic Research Laboratory, ERL-M 382, 1973
2. Nagel, L.W.: SPICE2: A Computer Program to Simulate Semiconductor Circuits. Berkeley, Univ. of California, Electronic Research Laboratory. ERL-M520, 1975

The origins of ADS

The origin of MDS simulator goes back to the late 1960's when HP introduced the simulator OPNODE which was a linear s-parameter simulator. In 1984 HP wrote a linear simulator called Time-Domain Tutorial which could be used to take a transmission line topology, compute the S-parameters, and use the 8510 to get a time domain reflectometer plot. This time domain tutorial program eventually evolved into the MDS Microwave Linear Simulator MLS in MDS A.01.00.

Ken Kundert worked on MLS until 1985 when HP funded him to get his Ph. D at U.C Berkeley working on nonlinear microwave simulation based upon harmonic balance technology. HP was co-developing harmonic balance in conjunction with Ken Kundert and Berkeley. The internal name for this simulator was "Spectra" which was later named the HP Microwave Nonlinear Simulator or MNS. MNS essentially evolved from Berkeley's code. The Berkeley code was named Spectre (a subtle difference). After Ken Kundert received his doctorate, he left HP to work at Cadence where he built a new time domain simulator based upon the UC Berkeley Spectre code, hence the Cadence Spectre simulator was born.

In 1996, Hewlett-Packard bought EEsof, and MDS and the EEsof simulator products were merged. The new product was called ADS (Advanced Design System).

Device Modeling

Contents

- *Diode* (iccapmhb)
- *Transistors* (iccapmhb)
- *RF Passive Linear Components* (iccapmhb)
- *Low Frequency Noise* (iccapmhb)
- *Thermal Modeling* (iccapmhb)
- *On Target Modeling* (iccapmhb)
- *Scaled Modeling* (iccapmhb)

On the Web

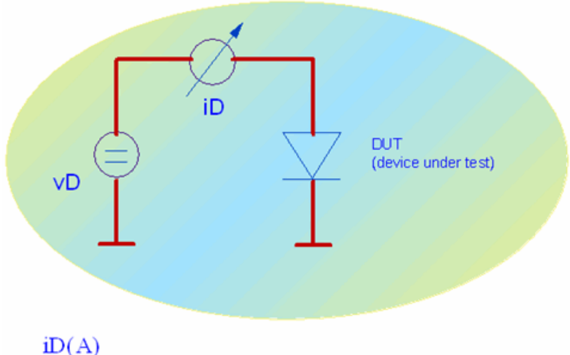
- Articles
- Knowledge Center
- Technical Support
- Training
- Videos

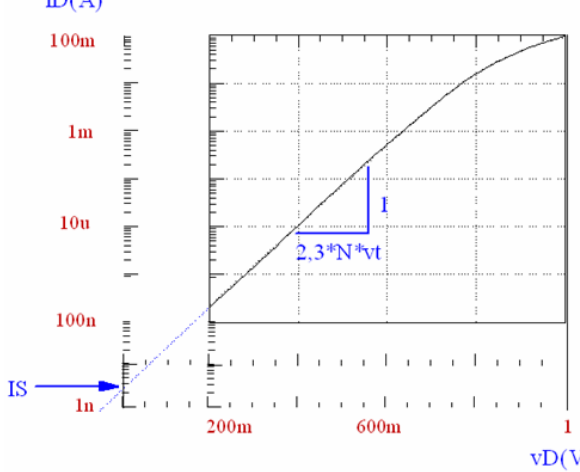
Diode

Contents

- *Modeling a Diode* (iccapmhb)
- *ADS Junction Diode Model* (iccapmhb)
- *Step Recovery Diode* (iccapmhb)

Modeling a Diode





Introduction

Referring to the Modeling Handbook's chapter on curve fitting, regression analysis was introduced as a method for linear curve fitting. It was a pretty simple and straight-forward method. We had to solve the partially differentiated equations for the parameters m and b of the equation

$$y = m*x + b$$

It was mentioned that fitting more complex measurement curves leads to much more complex problems for solving the set of equations for the parameters. And pretty often, this set of non-linear equations cannot be solved without numerical methods.

But there is a way around:
A simple way to fit non-linear curves to measured non-linear data is to transform them to a linear world. But the question is, how to do it. A look at the 'target function' of the model equation gives a hint. An example: in case of an exponential function like

$$i = I_S * e^{\frac{v}{N*vt}},$$

the transform needed is a simple logarithmic conversion. This automatically gives the non-linear transformation for the measured data. Once the measured data are transformed to this linear range, a linear regression analysis is applied. And so we yield the slope and y-intersect of the fitted line. The model parameters are finally calculated out of these two values.

This will become much more transparent in the following diode modeling example.

The SPICE equivalent schematic for a diode is shown in figure 1. It consists of the ideal diode D representing its non-linear DC characteristic plus two voltage dependent capacitors for taking care of the space charge (CS) and delay effects (CD) as well as a series resistor RS for the high-current effects. The series inductor (bonding effect ...) is neglected as well as a parasitic capacitor (housing effects ...).

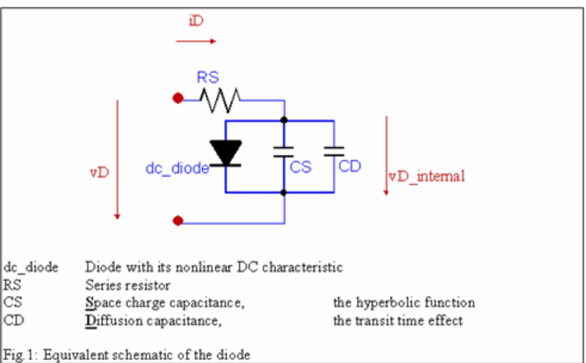


Fig 1: Equivalent schematic of the diode

DC Characterization Parameters IS, N AND RS

Neglecting high current effects, i.e. RS=0 or vD = vD internal, and also neglecting recombination effects for low biasing voltage, the diode current in the forward active region is modeled using:

DC:

$$i_D = I_S \cdot \left(e^{\frac{v_D}{N \cdot vt}} - 1 \right) \qquad (1)$$

with

```
Is :     saturation current   (leakage current, typical fA)
N  :     emission coefficient (ideal diode:     N=1)
VT :     temperature voltage   27mV    at 25°C
         or   VT   =  k*T/q  =  8.6171 E-5 * (T /°C  +  273.15)    (1a)
```

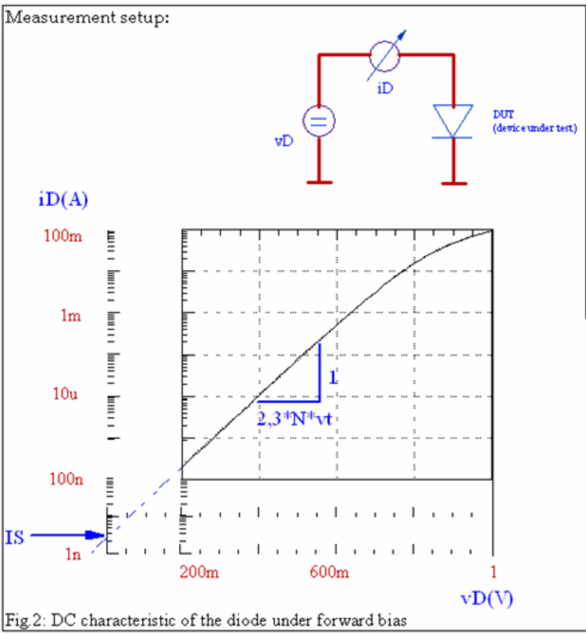


Fig 2: DC characteristic of the diode under forward bias

Determination of the DC parameters IS and N

Provided vD >0, i.e. neglecting the term (-1) in (1), and applying a logarithmic conversion yields:

```
                           vD
log(iD)  =  log(Is)  +  ------ log(e)
                          N VT
         =  log(Is)  +  [1 / (2,3 N VT)]  vD  (2a)
```

This is an equation of the form:

```
y =  b +  mx (2b)
```

In order to interpret (2b) linearly, we have to substitute:

```
y =  log(iD)        (2c)
```

```
b =  log(Is)        (2d)
```

```
m =  [1 / (2,3 N VT)] (2e)


x =  vD            (2f)
```

This explains how to manipulate the measured data: after the logarithmic conversion of the measured values of iD (2c), they are introduced with the still linear values of vD (2f) into the regression equations (10) and (11) of the previous regression analysis chapter 4.1 as yi- and xi-values. We obtain the y-intersect b and the slope m of the linear regression function. Solving (2d) for Is and (2e) for N we finally are able to calculate these two parameters out of b and m as follows:

```
Is  =  10 b  (3)
```

and

```
N  =  1 / (2,3 m VT)
                           with VT from (1a) (4)
```

Validity of this extraction:

The parameter extraction described above is valid only in that range of measured data, where the assumptions are true. This means: eq.(3) and (4) are valid for vD > 0 (data above the measurement resolution (non-noisy data), typ. >0,2 V). The diode current should not be dominated by recombination effects (the weaker slope at low bias voltages) but also below high-current effects (no ohmic effects, the famous knee in the half-logarithmic diode characteristic above typically 0.8 V)

Parameter RS

After the parameters Is and N are extracted, the value of RS can be found from the two highest bias points (biasmax_index and biasmax_index-1 of a sweep from vD_min to vD_max) as follows:

```
vD[max_index] - vD[max_index-1]
RS = ---------------------------------- (5)
      iD[max_index] - iD[max_index-1]
```

This method gives only accurate results for RS if the applied vD was high enough so that the ohmic effect (of RS) dominates over the nonlinear, exponentional diode curve `iD = IS*exp(vD/(N*vt) -1)`

Another method to determine the ohmic part of a diode characteristic, independent of that limitation from above, is to consider the voltage drop along RS, between the ideal diode characteristics and the applied outer voltage. Since we know the applied iD, and since we can calculate the voltage drop across the inner diode (IS and N are known !), we can calculate RS.
This is done by firstly determining the maximum current iD from the sweep by

```
i_RS=ia.m[max_index]
```

This is the current flowing through RS.
Then, we calculate the voltage drop across RS following

```
v_RS =  measured_voltage - ideal_diode_voltage
```

or

```
v_RS =  va.m[max_index] - vt*N*ln(i_RS / IS)
```

what finally gives

```
RS = v_RS / i_RS
```

Of course, as mentioned, the diode DC parameters IS and N have to be determined first.

Pre-requisite for a good RS extraction

the ohmic effect dominates the diode characteristics. Referring to fig.2, the decline for high bias voltage must be clearly visible in the extraction range.

CV Characterization

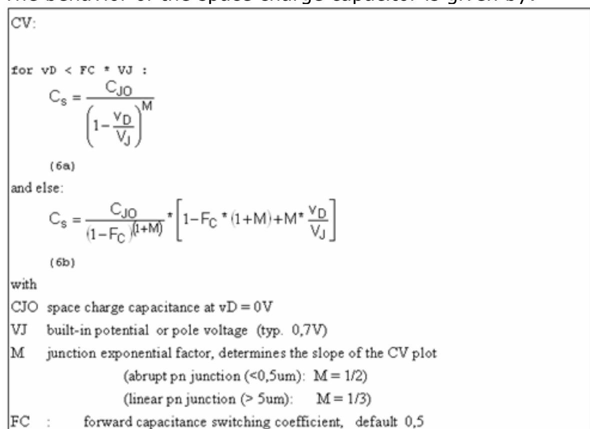
The frequency behavior of a semiconductor can be modeled by a space charge capacitance (dominant at reverse biasing) and a diffusion capacitance (dominant at forward bias) that models the time delay effects. The first capacitance is typically measured with a negative bias using a CV meter (capacitance versus voltage) while the latter is commonly measured

using a network analyzer.

This chapter covers the modeling of the space charge capacitor. Another method is using a network analyzer, measuring the s11 curve with a negative bias. This is not covered here.

Space charge capacitance in general, extracting parameters Cj0, Vj, m

The behavior of the space charge capacitor is given by:



For more details see /Antognetti/.

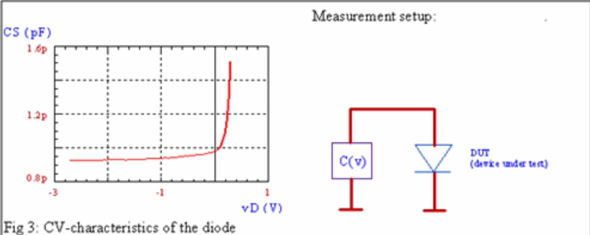


Fig 3: CV-characteristics of the diode

Determination of the CV parameters

We only use the negative bias region for parameter extraction. The logarithmic conversion of (6a) gives:

$$\ln(C_S) = \ln(C_{J0}) - M * \ln\left(1 - \frac{V_D}{V_J}\right) \qquad (7)$$

This equation can be interpreted again as a linear function:

$$y = b + m * x$$

with

y = ln(CS) (8a)

b = ln(CJO) (8b)

m = - M (8c)

and

x = ln[1 - vD / VJ] (8d)

How to proceed

The measured values of CS are converted logarithmically according to (8a). Then, following (8d), the stimulating data of the voltage sweep vD are converted too. Since the parameter VJ has a physical meaning, its value should be in the range of 0,2 .. 1V. Therefore we select 0,2V as a starting value for VJ. These two arrays are now introduced into equations (10) and (11) of the chapter on regression analysis as yi- resp. xi-values. A linear curve is fitted to this transformed 'cloud' of stimulating and measured data and we get the y-intersect b and the slope m for the actual value of VJ. These two values are the best choice for the given VJ. In the next step, this procedure is repeated with an incremented VJ, and we get another pair of m(VJ) and b(VJ) . But now the regression coefficient r2 will be different from the earlier one: depending on the actual value of VJ, the regression line fits the transformed data 'cloud' better or worse. Once the best regression coefficient is found, the iteration loop is stopped and we get VJ_opt as well as

the corresponding b(VJ_opt) and m(VJ_opt).

The final parameter values are then from (8c):

```
M   =  - m(VJ_opt) (9a)
```

and from (8b):

```
CJO  =  exp [ b(VJ_opt) ] (9b)
```

Validity of this extraction

The parameter extraction for the space charge capacitor is valid only for stimulus voltages below `FC * VJ , FC_default = 0,5 .`

In practice, there is always an overlay of this capacitance with some parasitic ones, e.g. packaging or bond pads. If they are not known and therefore cannot be de-embedded (eliminated from the measured data by calculations), the three modeling parameters may have values that have no physical meaning. This is especially true for VJ and M. Nevertheless the fitting of the proposed method is generally very good and pretty easy.

In order to also determine the parasitic offset capacitance, see the examples in the chapter about 'regression analysis applications'.

HF Modeling: Parameter TT

The small signal equivalent schematic of the diode for high frequencies is given in fig.4. When comparing it to fig.1, it can be seen that the element D, representing the non-linear DC transfer curve of the diode has been replaced by the linearized small signal conductance gD.

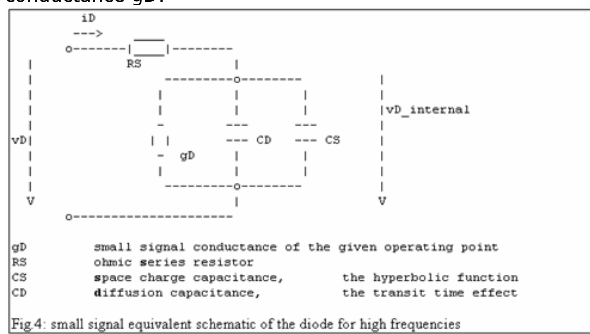


Fig.4: small signal equivalent schematic of the diode for high frequencies

Let's start with gD , the slope to the DC diode characteristics at the operating bias point.

$$g_D = \frac{\partial i_D}{\partial v_D} \overset{(1)}{=} \frac{IS}{N \cdot vt} \cdot \left(\exp\left(\frac{v_D}{N \cdot vt} \right) - 1 \right) = \frac{1}{N \cdot vt} \cdot i_D \qquad (18)$$

The diffusion capacitance in the operating point is given by /Antognetti/:

$$C_D = T_T * g_D = T_T * \frac{1}{N \cdot vt} \cdot i_D \qquad (19)$$

CD is typically overlying the space charge capacitance.
When performing a 2-port measurement of a diode with a network analyzer, we can calculate the total diode capacitance by converting the S-parameters to Y-parameters and calculating
`Cdiode= IMAG(-Y12) / (2 PI freq) .`
The resulting CV curve is depicted below in fig.5:

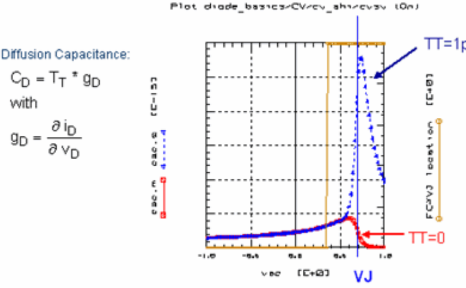


Fig.5: the diffusion capacitance overlays the space charge capacitance for high DC bias.

When applying a NWA, compared to CV meter measurements, there are no restrictions related to positive diode DC biasing and measurement instrument restrictions. Therefore,

we can easily forward bias the diode and study the transition from the space charge capacitance to the diffusion capacitance. This gives the diffusion capacitance.

> NOTE: in practice, especially for packaged devices, the diffusion capacitance is overlaid by the package inductor! See further below!

Determining TT

As can be seen in fig. 6, CD can also be optimized in the S-parameters for medium DC biases, below the influence of RS. In other words, related to S-parameters, TT shifts the Sxx and Sxy traces (adds phase). The effect is dominant for medium and higher DC biases below take-over of RS.

> NOTE:
> When RS begins dominating the diode DC trace, think of the 'inner' diode as a resistor with 1/gD in series with a voltage source of e.g. 0,7V. Therefore, the capacitances CS and CD are shortened by this decreasing diode resistor, and therefore, TT is shortened by this resistor!!

The influence of the diode transit time TT to S-parameters

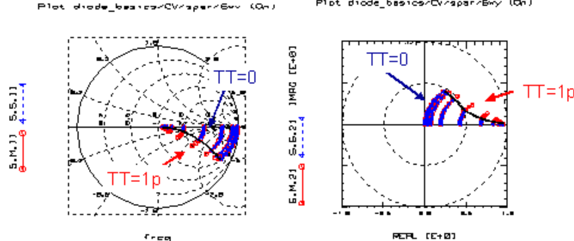


Fig.6: How the transit time TT influences the S-parameters

Finally, when taking also the series inductance into account, which is a typical first-order model of the diode package, we get S-parameters like shown in fig.7:

S-parameter Modeling (package included)

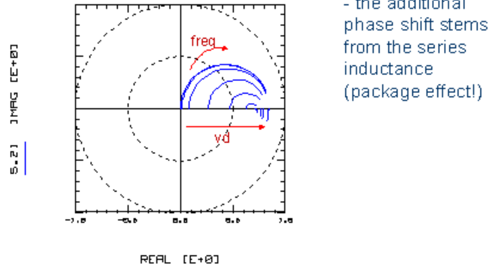


Fig.7: Diode S21-Parameters including the package inductor

The series inductor overlays the so far discussed S-parameters of the inner diode. It basically adds phase to the inner diode S-parameters, and for high DC biasing (where RS domiinates), the inductance affects the diode S21 trace considerable: S21 now turns downwards, tending towards S21->0 for infinite frequency.

Related to modeling, LS can be obtained from optimization of high DC biases.

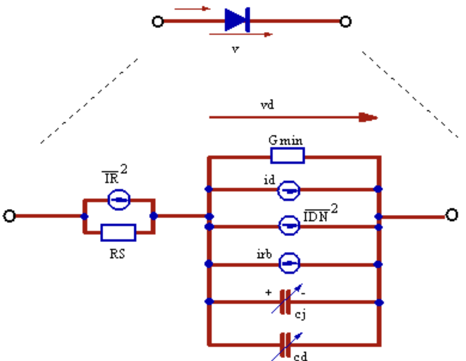
Model Limitations

In order to keep the models simple and usable and to have reasonable simulation times, they might suffer from some limitations:

- **DC**: diodes may show recombination effects at low forward bias voltages. This shows-up as a lower slope on a half-logarithmic scale. In order to cover this effect, the diode model is replaced by a subcircuit, consisting of a diode for the recombination effect, another one in parallel for the upper voltage area and a resistance in series with both diodes. (Both diodes have RS=1e-6 Ohm).
- **CV**: any parasitic capacitance is not included in the diode model. Using again a sub-circuit, a 2nd parasitic capacitance can easily be added.
- **RF**: the parasitic series inductor is not included. Again, a subcircuit could be used for modeling.

Details of the Berkeley SPICE Diode Modeling for Experts

This section gives detailed information about the Berkeley Spice model including thermal and noise modeling.



The complete SPICE parameter list

parameter	description	units	default	example
RS	ohmic resistance	Ohm	0	10
IS	saturation current	A	1.0E-14	1.0E-14
N	emission coefficient		1	1
BV	reverse breakdown voltage	V	infinite	40
IBV	current at breakdown voltage	A	1.E-3	
CJO	zero bias junction capacitance	F	0	2.0E-12
VJ	junction potential	V	1	0.6
M	grading coefficient		0.5	0.5
FC	coeff.for forward-bias deplet.cap.		0.5	0.5
TT	transit time	sec	0	1E-10
EG	activation energy	eV	1.11	1.11 for Si 0.67 for Ge
XTI	saturation current temp.exp.		3.0	3
KF	flicker noise coefficient		0	
AF	flicker noise exponent		1	
TNOM	parameter measurement temp.	'C	27	
GMIN	min.SPICE conductance (a SPICE convergence parameter)		1E-12	

SPICE Diode Model Equations

DC model

forward and reverse leakage:

$$id = IS \cdot \left(e^{\frac{vd}{N * vt}} - 1 \right) + Gmin \cdot vd \qquad \text{with} \qquad vt = \frac{k * TEMP}{q}$$

reverse breakthru current:

$$irb = IBV \cdot \left(e^{-\frac{vd+BV}{vt}} - 1 \right)$$

Note: when implementing the diode model into simulators, consider a linear continuation of the exponential function, in order to improve the simulation convergence and to avoid numeric overflow

$$f(x) = \begin{cases} \exp(x) & x < x0 \\ mx + b & x > x0 \end{cases}$$

with

$$m = \exp(x0)$$
$$b = (1 - x0) \cdot \exp(x0)$$

AC model

Junction capacitance

$$cj = \frac{CJO}{\left(1 - \frac{vd}{VJ}\right)^M} \qquad \text{for} \quad vd<FC*VJ$$

$$cj = \frac{CJO}{(1-FC)^M}\left[1 + \frac{M}{VJ \cdot (1-FC)}(vd - FC \cdot VJ)\right] \qquad \text{for} \quad vd>FC*VJ.$$

For user-defined models, charges have to be specified, $Q = \int C(v)dv$.
Therefore the junction charge is:

for vd<FC*VJ :

$$qj = \frac{CJO \cdot VJ}{1-M} \cdot \left[1 - \left(1 - \frac{vd}{VJ}\right)^{(1-M)}\right]$$

and for vd>FC*VJ :

$$qj = \frac{CJO \cdot VJ}{1-M} \cdot \left[1 - (1-FC)^{(1-M)}\right] + \frac{CJO \cdot (vd - FC \cdot VJ)}{(1-FC)^{M}} + \frac{0{,}5 \cdot CJO \cdot M}{VJ \cdot (1-FC)^{(1+M)}} \cdot (vd - FC \cdot VJ)^2$$

Diffusion capacitance:

$$cd = TT \cdot \left(IS * \frac{1}{vt * N} * e^{\frac{vd}{N * vt}} + Gmin\right)$$

and the corresponding charge (for model implementation):

$$qd = TT \cdot id(vd)$$

Thermal model

$$IS_{TEMP} = IS_{TNOM} * \left(\frac{TEMP}{TNOM}\right)^{\frac{XTI}{N}} * e^{\frac{q * EG}{k * N}\left(\frac{TEMP - TNOM}{TEMP * TNOM}\right)}$$

$$VJ_{TEMP} = VJ_{TNOM} * \left(\frac{TEMP}{TNOM}\right) + 2 * vt * \log \frac{n_i}{n_i_TEMP}$$

with $n_i = 1.45E\text{-}10$

and $$n_i_TEMP = n_i * \left(\frac{TEMP}{TNOM}\right) * e^{\frac{q}{2k}\left(\frac{-EG}{TEMP} + \frac{1.15}{TNOM}\right)}$$

Noise model (used only in AC analysis)

$$\overline{IRS}^2 = \frac{4 * k * TEMP}{RS} * \Delta f$$ thermal noise

$$\overline{IDN}^2 = 2 * q * id * \Delta f + \frac{KF * id^{AF}}{f} \Delta f$$

shot noise flicker noise

For an implementation example, see
demo_features\6_simulators\2_ADS\3_user_def_mdls\1_sdd\basic_sdds\1_diode_sdd.mdl

Publications

Diode modeling and modeling in general:

P.Antognetti, G.Massobrio, Semiconductor Device Modeling with SPICE,
McGraw-Hill, 1988, ISBN 0-07-002107-4

ADS Junction Diode Model

INSTANCE

```
ModelName [:Name] anode cathode <parameter=value> ... ; (device)
```

Parameters	
Area	Junction area factor.
Region	DC operating region, 0=off, 1=on.
Temp (C)	Device operating temperature.
Gd (Siemens)	Small signal conductance.
Cd (F)	Small signal capacitance.
Mode	Nonlinear spectral model on/off.
Noise	Noise generation on/off.

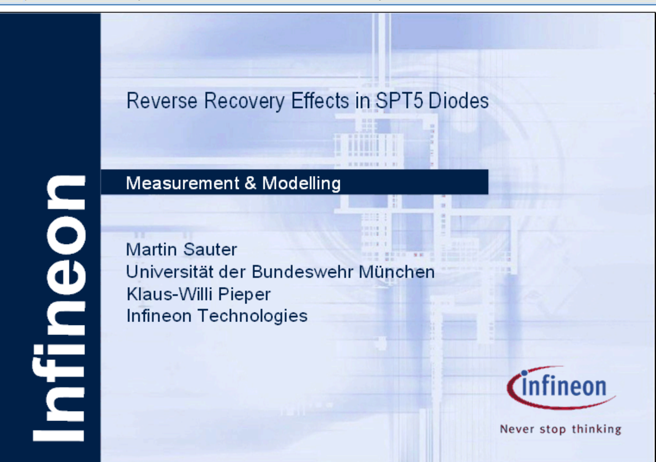
MODEL CARD

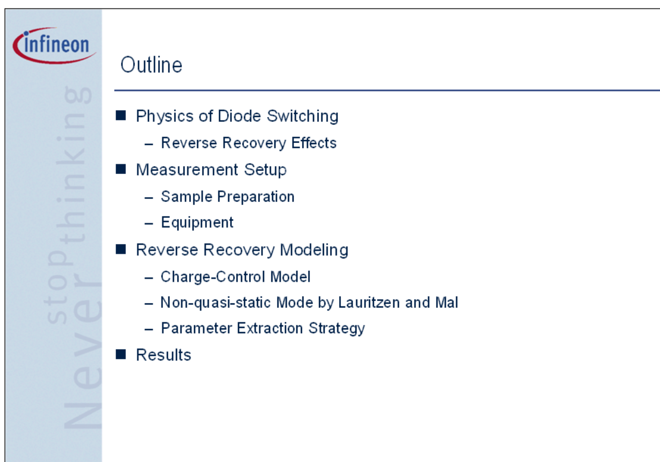
```
model ModelName Diode <Model Parameter=value> ...
```

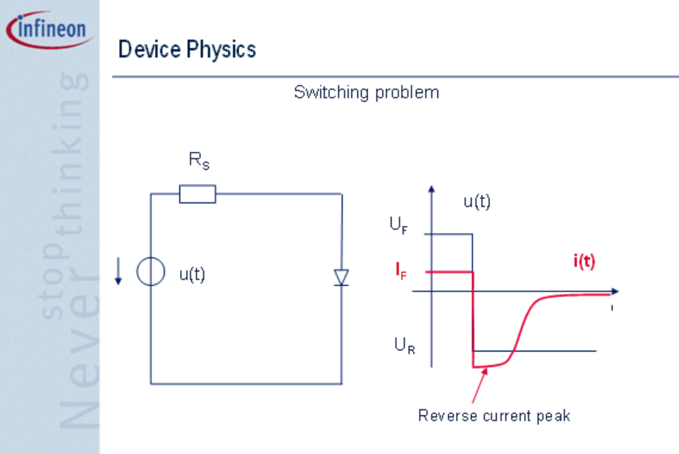
model Parameters	
Isr (A)	recombination current parameter.
Nr	Emission coefficient of recombination current.
Is (A)	Saturation current.
Js (A)	Saturation current.
N	Emission coefficient.
Ikf (A)	high injection knee current.
Rs (Ohms)	Series resistance.
Fc	Forward-bias depletion capacitance threshold.
Ibvl (A)	low-level reverse breakdow knee current.
Nbvl	low-level reverse breakdow ideality factor.
Bv (V)	Reverse breakdown voltage.
Ibv (A)	Current at breakdown voltage.
Nbv	reverse breakdown ideality factor.
Cjo (F)	Zero-bias junction capacitance.
Vj (V)	Junction potential.
Pb (V)	Junction potential.
M	Grading coefficient.
Tt (s)	Transit time.
Ffe	flicker noise frequency exponent.
Eg (eV)	Band gap.
Kf	Flicker-noise coefficient.
Af	Flicker-noise exponent.
Tnom (C)	Parameter measurement temperature.
Xti	Saturation current temperature exponent.
Pt	Saturation current temperature exponent.
Imax (A)	Explosion current.
wBv (V)	Diode reverse breakdown voltage (warning).
wPmax (W)	Maximum power dissipation (warning).

Step Recovery Diode

Note
This chapter has been contributed by Prof. Martin Sauter, University Bundeswehr, Munich, Germany and Klaus-Willi Pieper, Infineon, Munich, Germany.





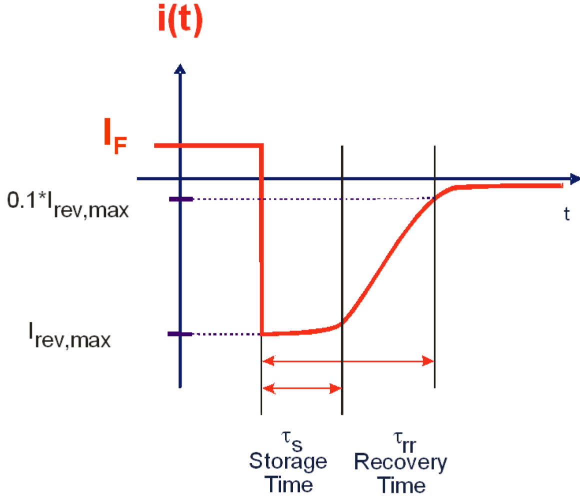


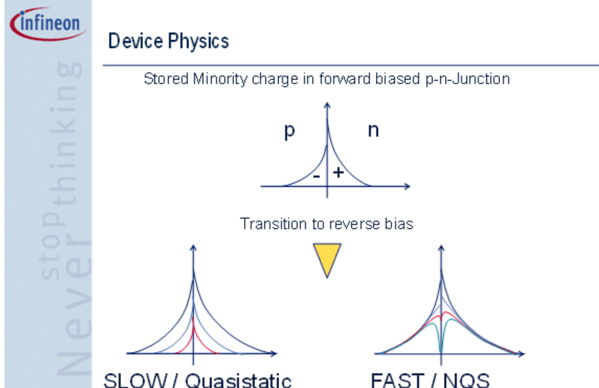
The problem of a diode, "reverse recovery" is a common known problem resulting from many circuits, especially in power electronics. It usually occurs, when a diode switches its state from the forward conducting mode to reverse blocking mode. In this case, the stored minority charge in the diode regions have to be removed from the diode anode and cathode regions, which does not occur instantaneously. The whole process may from a physical point of view be understood as the discharging of the diode diffusion capacitance.

An electrical equivalent circuit and the resulting waveforms are drawn in the Figure above. After the voltage source changes its state from forward voltage UF to reverse voltage UR, a discharge current is flowing, which is mainly limited by the series resistor Rs . This current lasts as long as the diode is in the conducting state and the voltage drop at the junction is about 0.7V. This switching phase is usually caused the "storage phase".

After this time, a more or less exponential decay of the current is observed, although the diode does not anymore operate in the forward region. This phase is usually called the recovery phase. The recovery time now is the time between the switching point of the voltage source and the time until the current has dropped to 10% of its maximum value.

A good comprehensive review of this effect (although in German) is found in the Reference [2].





The physics of the reverse recovery process can in general be described by an examination of the stored minority charge in the diode regions.

From quasistatic point of view, it can be assumed that the minority carrier charge is related to the current by a constant diffusion time, which is in literature called the diode transit time TT. The total diode current is given by the expression:

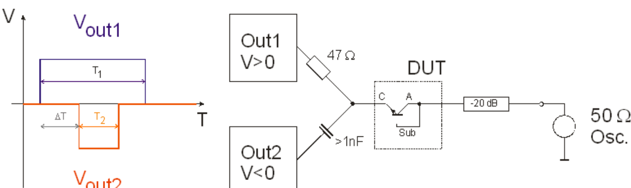
(I) →
$$i(t) = i_D + C_j(u) \cdot \frac{du}{dt} + \frac{dq_T}{dt}$$
$$q_T = T_T \cdot i_D$$

qT describes the stored minority charge in the diode regions and the variable iD is the current which flows under DC conditions when a voltage V is applied to the diode.

If we regard a time dependent voltage u(t), usually a quasistatic assumption is made, which states that the stored minority carrier charge is given by the equation (I). This however is only valid when the stored charge is able to follow the rapidly changing voltage u(t) applied at the diode terminals. A qualitative description of the resulting charge distribution in the diode regions is given in the picture above.

Left: A quasistatic assumption is made, which assumes the minority carriers to follow the diode bias without delay.
Right: A non-quasistatic assumption is made, which shows that the carriers near the junction follow the bias faster than the carriers far away from the junction. The minorities first vanish in the region which is close to the junction. In a later stage of the process, a number of minorities is stored in the diode regions far away from the junction.

Measurement Setup

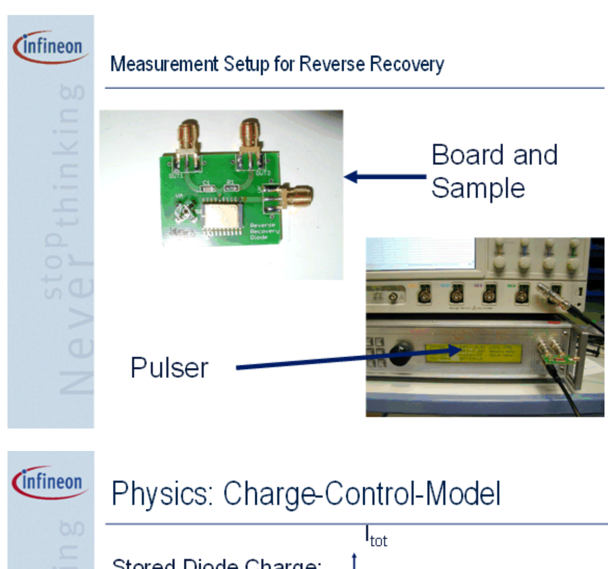


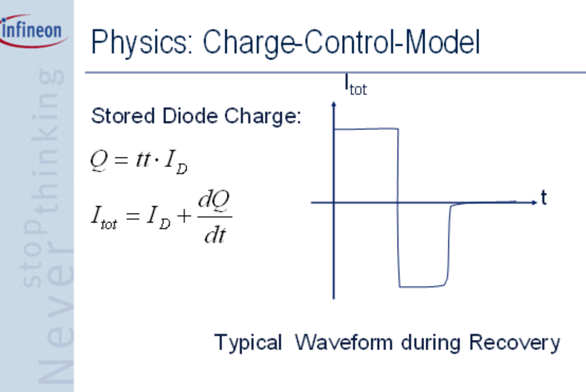
Here we describe how the measurements of the recovery time have been performed. We have used a setup proposed by AVTECH Inc. using a double output pulse generator AVTECH EB4B. This unit has two low-impedance (~2 Ohm) outputs. The output “OUT1” generates a positive pulse, which is sent to the device passing the resistor of 47 Ohm. The width T1 of this pulse can be varied between 100ns and 1us. The pulse rise time is about 1ns.

The output “OUT2” generates a negative blocking pulse with a width T2 between 100 and 300ns, which is coupled to the device by a capacitor with typical values of several nF. The time delay delta_T between the two pulses can be varied as well as the voltages of the two pulses.

A fast Oscilloscope is placed at the output of the device. The input voltage of the Oscilloscope can be used to calculate the current through the device. The frequency characteristics of the oscilloscope had been incorporated into the simulation. Attenuators have to be a added especially to the oscilloscope to measure voltages above 5 V rms.

The samples were prepared into a small CDSO-20 package.
The whole circuit was designed on a system board, which can be seen on the next page.
Data acquisition has been done by the used Tektronix oscilloscope, the ascii data have been transferred afterwards to the UNIX system with ICCAP running.





Standard Diode Model

In the standard diode model, a rather simple equation is implemented. It calculates the stored charge in the diode during forward conduction by one parameter (usually denoted as transit time):

$$Q = tt \cdot I_D$$

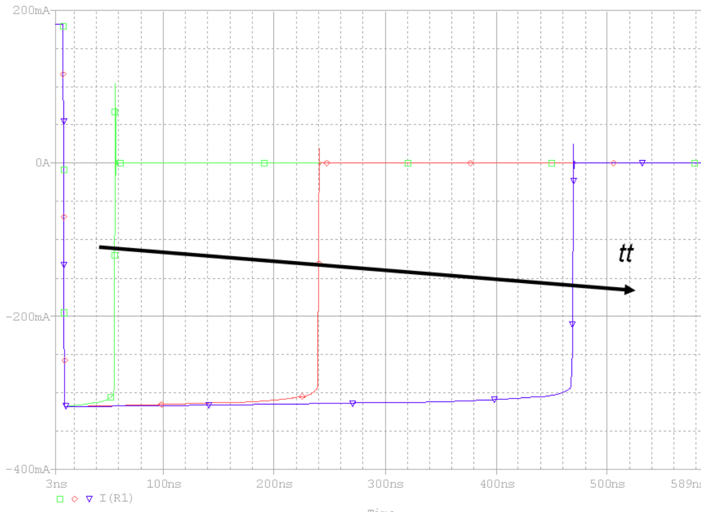
The total current in the diode is then calculated by:

$$I_{tot} = I_D + \frac{dQ}{dt}$$

Here, the junction capacitances are neglected for the reason of simplicity.

If we simulate the diode switching characteristics by this model, we obtain a time dependent current which is rather Step-like with an abrupt change of its value from the recovery peak to zero. The only possibility to change the waveform in the recovery range offered by the standard diode model is by varying the junction capacitance. But this parameter cannot be used as a fit parameter for the recovery waveform.

The following simulation shows the resulting transient waveforms for a rising parameter

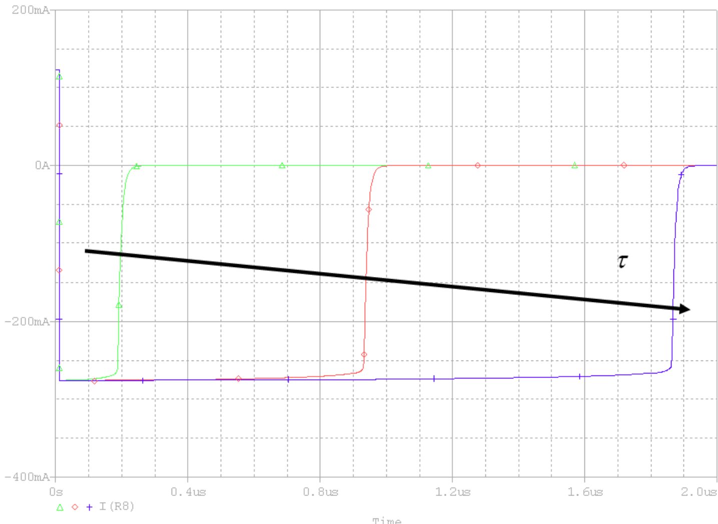


Non Quasi-static Diode Model

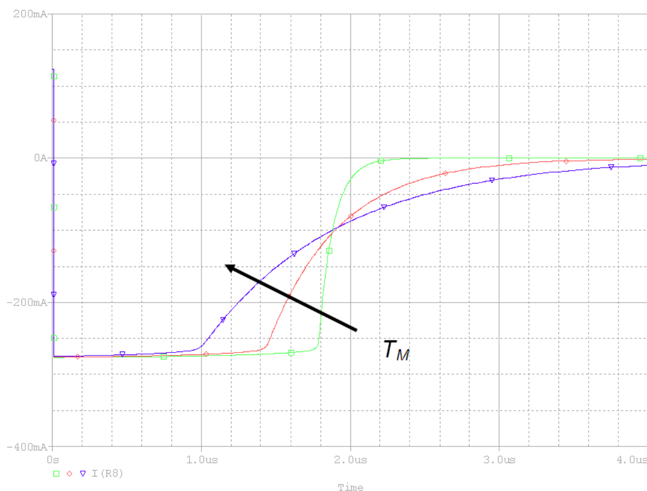
An alternative approach for a model of this recovery phase was developed by Lauritzen and Ma [1]. Originally, it was developed for low-doped high-voltage pin-Diodes.
This model separates the charges in the diode region into two parts Qe and QM .
In the equations ID denotes the static (DC) current of the diode. Itot is the sum of the static current and the current caused be the recovery.

Parameter Extraction

Let us have a look at the resulting transient waveform, if we perform variations of the model parameters TM and Tau.
First we assume TM ~ Tau and perform a variation of Tau . The results look almost identical with the waveforms obtained by the standard model, which is consistent with the mathematical considerations above, where we derived the standard model as a special case of the NQS model.



Variation of TM leads to a much more realistic description of the recovery behaviour by adding an exponential tail to the step characteristic.



Model Implementation

If we take a look at the original equations defined by the Reference 1 , we can perform the following operations with the second equation:

$$\frac{dQ_M}{dt} + Q_M \cdot \left(\frac{1}{\tau} + \frac{1}{T_M} \right) = \frac{Q_e}{T_M}$$

$$\frac{dQ_M}{dt} + Q_M \cdot \left(\frac{\tau + T_M}{\tau \cdot T_M} \right) = \frac{Q_e}{T_M}$$

$$\left(\frac{\tau \cdot T_M}{\tau + T_M} \right) \cdot \frac{dQ_M}{dt} + Q_M = \left(\frac{\tau \cdot T_M}{\tau + T_M} \right) \cdot \frac{Q_e}{T_M}$$

$$\left(\frac{\tau \cdot T_M}{\tau + T_M} \right) \cdot \frac{dQ_M}{dt} + Q_M = \left(\frac{\tau}{\tau + T_M} \right) \cdot Q_e$$

If we apply a Laplace Transform on this equation, we get:

$$\left(\frac{\tau \cdot T_M}{\tau + T_M} \right) \cdot p \cdot Q_M + Q_M = \left(\frac{\tau}{\tau + T_M} \right) \cdot Q_e$$

$$\left[\left(\frac{\tau \cdot T_M}{\tau + T_M} \right) \cdot p + 1 \right] \cdot Q_M = \left(\frac{\tau}{\tau + T_M} \right) \cdot Q_e$$

The term in the square bracket on the left side now is equivalent to the transfer function of a simple RC-circuit if the values of R and C fulfil the condition:

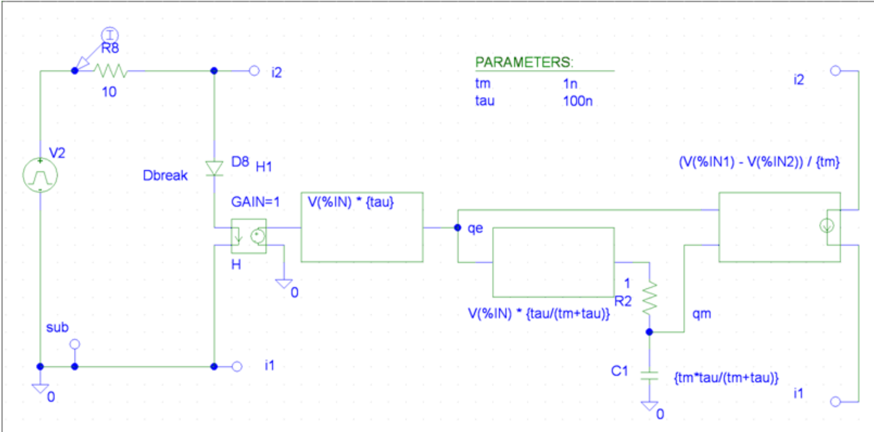
$$RC = \frac{\tau \cdot T_M}{\tau + T_M}$$

This gives an idea of how to implement the differential equations given by Lauritzen into a PSPICE circuit by only using standard elements.

The diode current is measured by a current controlled voltage source. If the transresistance of this element equals the value of Tau, the output voltage is equivalent to the value of Qe

1. The following RC circuit performs the calculation of the transfer function mentioned above. Here, the value of the resistor is set to "1", and the value of the capacitor is given by the two time constants.
2. The third element, which is a voltage controlled current source, performs the calculation of the additional current given in the third equation. The transconductance of this element is equal to Tau^-1

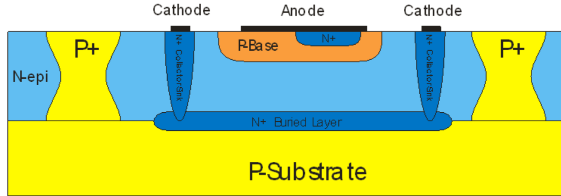
As the resulting current source is in parallel to the original diode, this is an additional contribution given by the equations.
The resulting schematic, which uses PSPICE ABM (Analog behavioural model) elements, is drawn below.



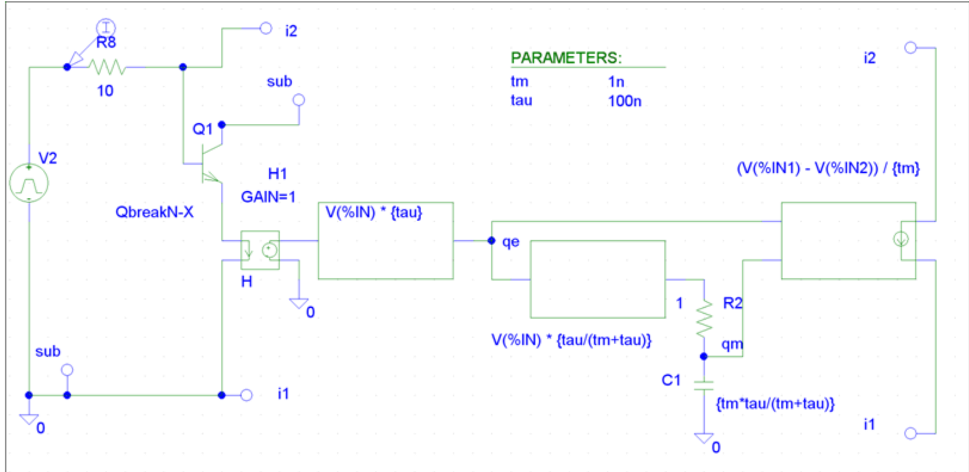
Modelling Results

We apply our model and the parameter extraction strategy to a reverse polarity protection diode, which is part of the Infineon Technologies SPT5 BCD-Process. This diode has a breakdown voltage of about 60V.

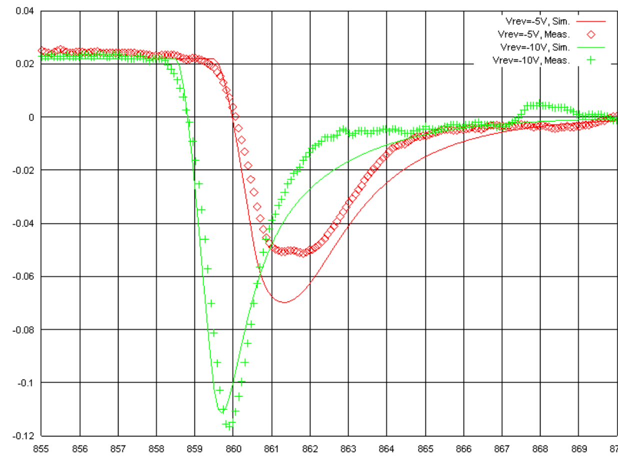
This diode together with the substrate forms a low-gain pnp transistor, where the substrate acts as collector and the diode acts as the emitter-base junction. The diode also includes another doping region on the top, decreasing the number of stored minority carriers in the anode.



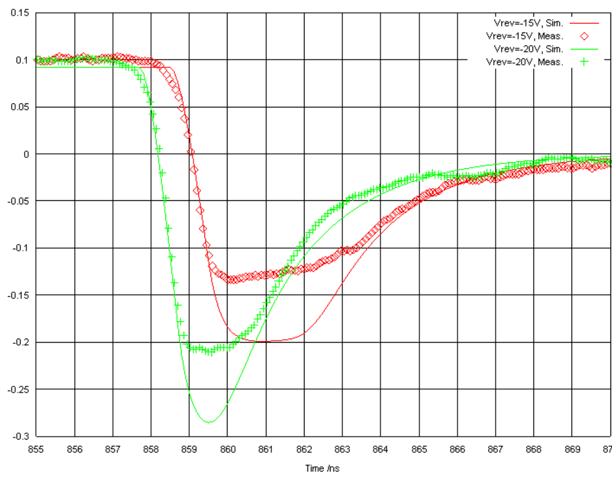
For a proper modelling of the recovery effect, the circuit we showed before has to be modified as drawn below. This has the advantage that the extracted substrate parasitics can be modelled by the pnp transistor and the network, which models reverse recovery behaviour has to be added later.



The results of the modelling of the dynamic characteristics are shown on the next page.



This picture shows the reverse recovery characteristic of the modelled diode. On the x-axis, the time (in ns) is shown, on the y-axis the calculated current (Oscilloscope voltage divided by 50 Ohm) is plotted.
The injection level has been chosen about 25mA on the diode, the reverse voltage was 5 and 10 V, respectively.



This picture shows the reverse recovery characteristic of the modelled diode with a higher injection level (~100mA) of the forward current than in the picture on the page before. Reverse voltage was also higher (15 V and 20V, respectively).
On the x-axis, time (in ns) is shown, on the y-axis the calculated current (Oscilloscope voltage divided by 50 Ohm) is plotted.

Summary

The aim of this paper is to show how the well known model of Lauritzen and Ma can be implemented very easy to model an integrated, very fast switching diode.

The way of implementation of the Lauritzen model is modular and can be added to standard Diode and BJT models as an additional network using only controlled sources and an RC-Circuit
The model is shown to predict the switching time of the measured diodes very well, the simulation was verified at different levels of injection current and for different reverse blocking voltage.

References

1. Lauritzen, P.O.; Ma, C.L.; A simple diode model with reverse recovery, IEEE Transactions on Power Electronics, Volume 6, Issue 2, April 1991 Page(s):188 - 191
2. M. Reisch, Elektronische Bauelemente, Springer Verlag 1998, p. 378-382
3. AVTECH Inc., Manual of Pulse Generator Unit EB4-B

Transistors

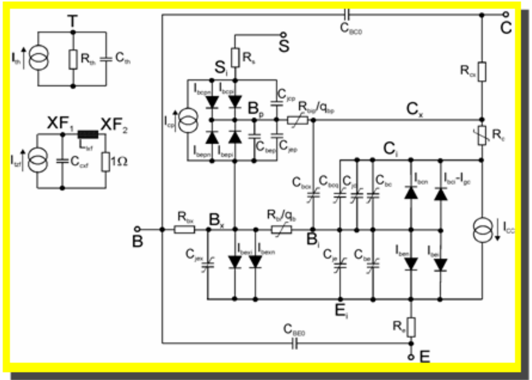
- *Bipolar in General* (iccapmhb)
- *BSIM* (iccapmhb)
- *GaAs Transistor Models in General* (iccapmhb)
- *General Transistor Modeling Strategies* (iccapmhb)
- *Hicum Level 2 model* (iccapmhb)
- *Learning the Angelov Model* (iccapmhb)
- *MOS3* (iccapmhb)
- *MOS Transistors for Dummies* (iccapmhb)
- *PSP* (iccapmhb)
- *The Curtice Mesfet Model* (iccapmhb)
- *The Gummel-Poon model* (iccapmhb)
- *YParameter Modeling* (iccapmhb)
- *VBIC* (iccapmhb)

Bipolar in General

Contents

- *Advanced Bipolar Transistor Modeling* (iccapmhb)
- *Bipolar Transistor Modeling* (iccapmhb)

Advanced Bipolar Transistor Modeling



Why BJT Modeling Is Still Hot?

- BJT is still dominated in RF and analog IC
- SGP model is far not accurate for modern BJT's (L<0.6um) for low risk design
- SiGe HBT for SOC?
- Other advanced HBT process, such as GaAs HBT

What is a Compact Model?

- A set of equations and related parameters implemented in circuit simulators
  - describe the static and dynamic behavior of the device
- Parameter extraction is an issue
  - link people in foundry and design house
- Compromise between physic and empirical
- Compromise between accuracy and complexity
- Multi disciplines / cross sciences

Why Standard Compact Model?

- Standardized model formulations to enhance communication among IC and CAD industries
- Standardized interfaces to simplify and speed up model implementation and test for CAD industries
- Better compact model for the latest technologies to lead the edge design
- Successful story is BSIM3 model from UC Berkeley
- Which model will win as standard model?
  - VBIC
  - Mextram
  - HICUM

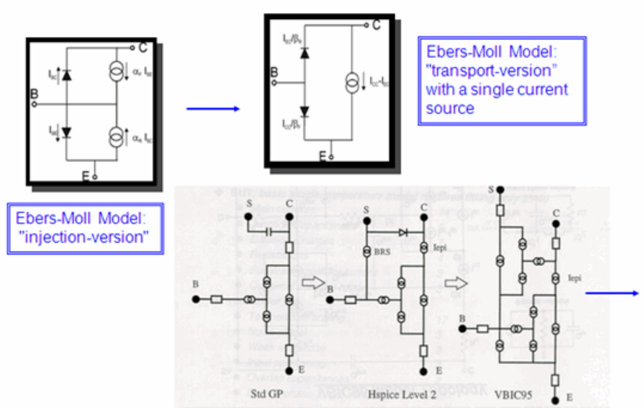
History of BJT Modeling

- 1954: Ebers-Moll model
  Enhanced to EM1,EM2,EM3
- 1970: Gummel-Poon Model
  widely accepted as Spice G-P model (SGP≠GP)
- 1986: Mextram model - Philips
- 1994: public domain
- 1987: HICUM model - Prof. Schroter
- 1999: public domain
- 1995: VBIC model
- 1995: public domain
- 1999: Bipolar model standardization

Most Exquisite TRAnsistor Model current v503.2, new releasev504 in June 2000, available in ADS, IC-CAP. HIgh CUrrent bipolar compact transistor Model 1993 extended to HBT, current v2, used in Rockwell to be available in ADS. HIgh CUrrent bipolar compact transistor Model
1993 extended to HBT, current v2, used in Rockwell to be available in ADS. Vertical Bipolar Inter-Company model current v1.1.5, v1.2 code has been just released available in ADS, IC-CAP.

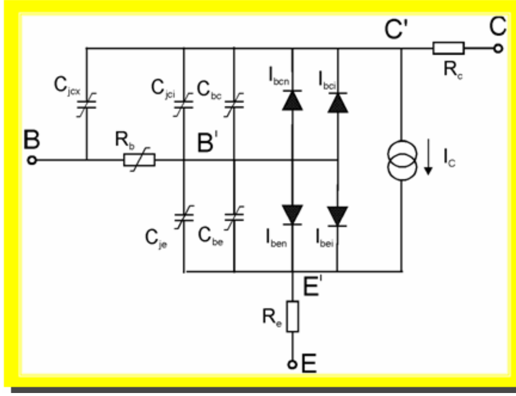
The Gummel-Poon model is not the same as the SPICE Gummel-Poon model which everybody has access to in SPICE-like simulators. It wouyld be unfair to Hermmann Gummel to create the perceiption that these models are the same. Also, the SGPM is not a "new" model, but about 25 years old (it might not fit under the session title).

Example of Model Evolution



The Spice Gummel-Poon Model

- Voltage-controlled Current Source Ic(vbe,vbc)
- Ideal (n≈1) and non-ideal (n≈2) Base Currents in Forward and Reverse Biased Mode
- Space Charge Capacitors
- Diffusion Capacitors
- Bias-dependent Base Resistor
- Fixed Base and Emitter Resistors



SGP-Model: DC Formulation

Current Source I_c:

- Forward and Reverse Operation
- depending from q_b

$$I_C = \frac{I_F - I_R}{q_b} = \frac{IS}{q_b} \cdot \left(e^{\frac{V_{B'E'}}{NF \cdot V_T}} - 1\right) - \frac{IS}{q_b} \cdot \left(e^{\frac{V_{B'C'}}{NR \cdot V_T}} - 1\right)$$

Base Current:

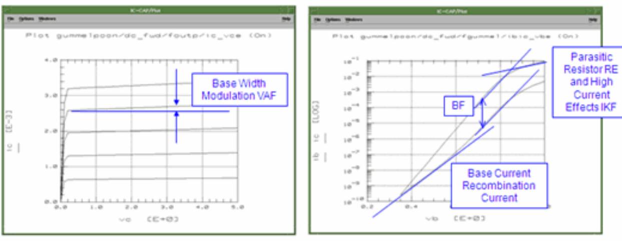
- Forward and Reverse Operation
- ideal and non-ideal Current
- ideal Current coupled to Collector Current Ic

$$I_{be} = \frac{IS}{BF} \cdot \left(e^{\frac{V_{B'E'}}{NF \cdot V_T}} - 1\right) + ISE \cdot \left(e^{\frac{V_{B'E'}}{NE \cdot V_T}} - 1\right)$$

$$I_{bc} = \frac{IS}{BR} \cdot \left(e^{\frac{V_{B'C'}}{NR \cdot V_T}} - 1\right) + ISC \cdot \left(e^{\frac{V_{B'C'}}{NC \cdot V_T}} - 1\right)$$

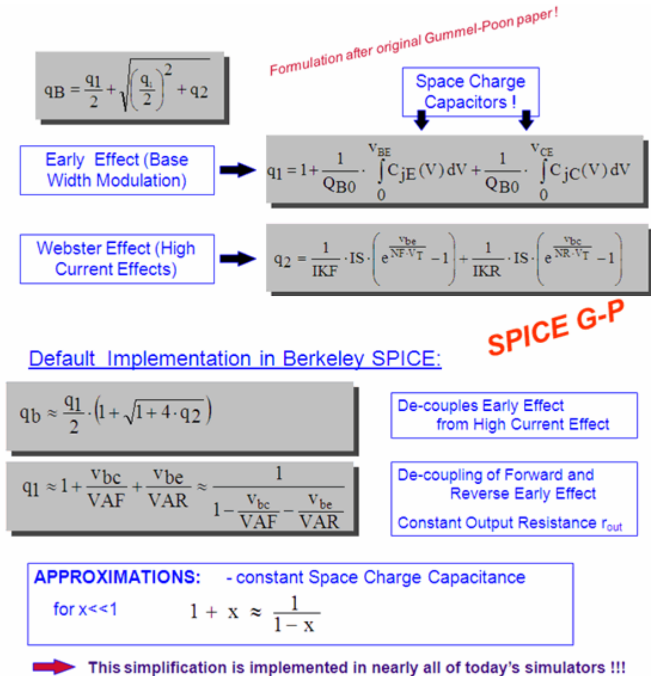
Visualization of the most important DC Effects

Output Characteristics **Gummel Plot**



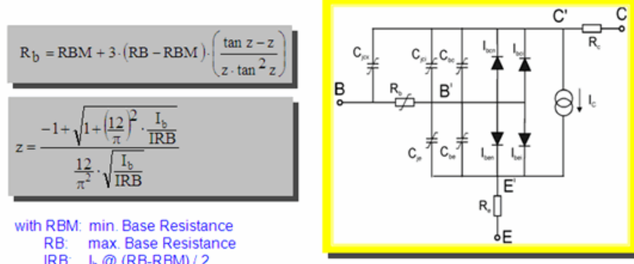
SGP-Model: normalized Base Charge q_b

The Dominant Internal Model Parameter: Normalized Base Charge q_b



SGP-Model: Parasitic Resistors

- Emitter Resistor RE constant
- Collector Resistor RC constant
- Bias-depending Base Resistor



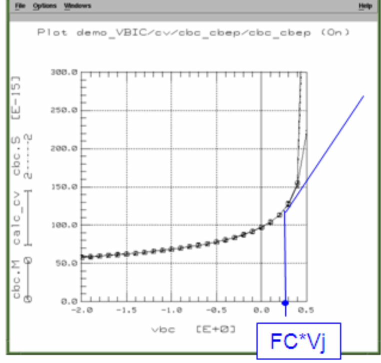
SGP-Model: Capacitors

Space Charge Capacitors:

- Parameter: C_{j0}, V_j, m

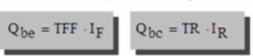
$$C_{ji} = \frac{C_{ji0}}{\left(1 - \frac{V_{ji}}{\Phi_i}\right)^{m_i}}$$

- Linear continuation for v > FC*V_j
- Base-Collector capacitance distributed between inner and outer Base by XCJC
- Substrate capacitance modeled by a constant or using the C_{ji} formula



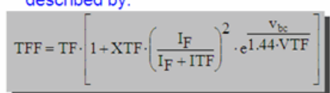
Diffusion Capacitors:

- Determined by the charges

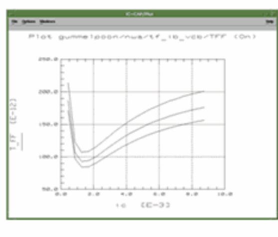


$Q_{be} = TFF \cdot I_F$ $Q_{bc} = TR \cdot I_R$

- Forward transit time empirically described by:



$$TFF = TF \cdot \left[1 + XTF \cdot \left(\frac{I_F}{I_F + ITF}\right)^2 \cdot e^{\frac{V_{bc}}{1.44 \cdot VTF}}\right]$$



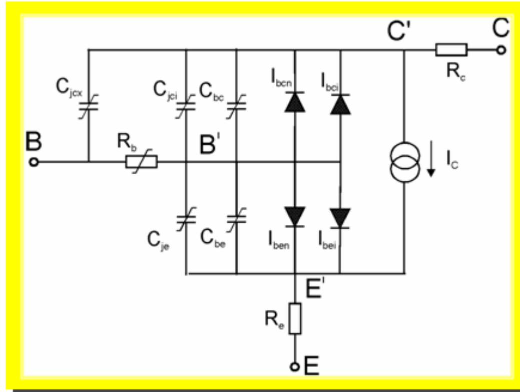
- Reverse transit time is constant

SG-P Model: Additional Phase Shift and Temperature

- Additional Phase Shift at high frequencies: measured phase is bigger than simulated
- Additional phase shift parameter PTF
  - Frequency Domain: added linearly
  - Transient Analysis: second order Bessel function
- Temperature
- Reflects constant ambient temperature
- No self-heating

The Spice Gummel-Poon Model

- Voltage-controlled Current Source Ic(vbe,vbc)
- Ideal (n≈1) and non-ideal (n≈2) Base Currents in Forward and Reverse Biased Mode
- Space Charge Capacitors
- Diffusion Capacitors
- Bias-dependent Base Resistor
- Fixed Base and Emitter Resistors



SGP-Model: Model Parameters

Using Default Parameter Values switches-off the Parameter Effect

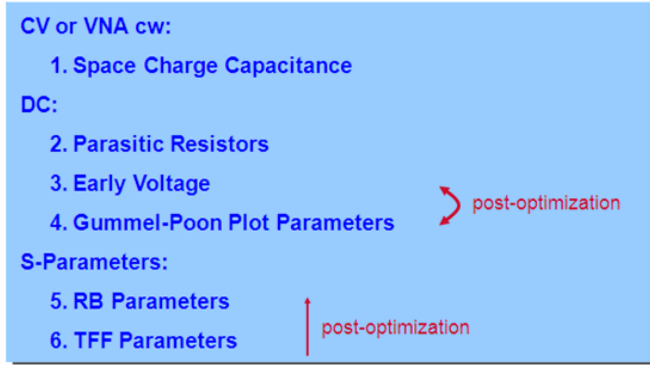
- do not confuse with *Typical Parameters* -

Parameter	Default
DC Forward	
IS	0.1f
NF	1
BF	100
ISE	0
NE	1.5
IKF	∞
DC Reverse	
NR	1
BR	1
ISC	0
NC	2
IKR	∞
Early Modeling	
VAF	∞
VAR	∞

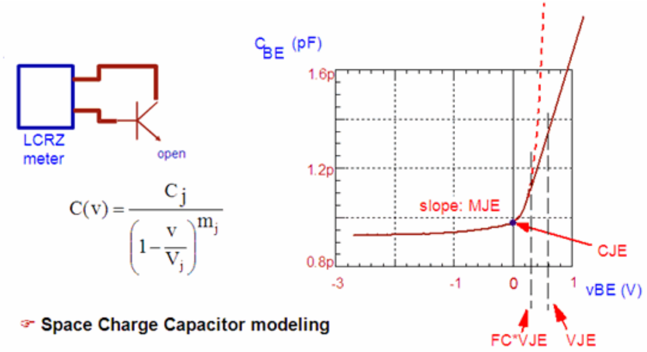
Parameter	Default
Space Charge Capacitances	
CJE	0
VJE	0.75
MJE	0.33
CJC	0
VJC	0.75
MJC	0.33
XCJC	1
CJS	0
VJS	0.75
MJS	0.33
FC	0.5

Parameter	Default
Resistances	
RE	0
RC	0
RBM	RB
RB	0
IRB	∞
Delay Time (Transit Time)	
TF	0
XTF	0
ITF	0
VTF	∞
TR	0
Excess Phase	
PTF	0
Temperature	
TNOM	27
EG	1.11
XTI	3
XTB	0

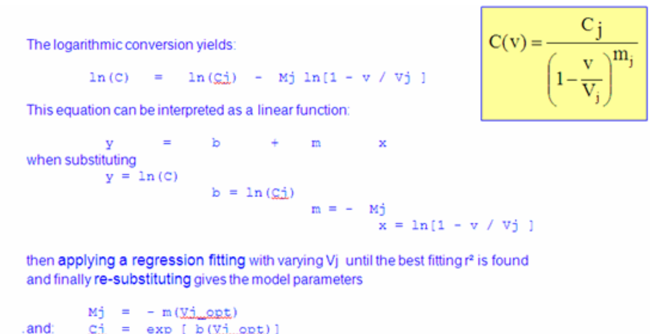
SGP: Parameter Extraction Strategy



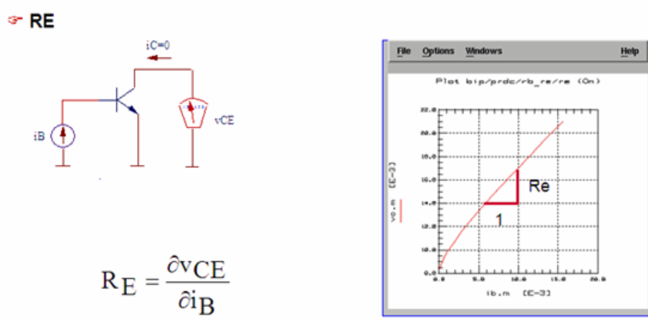
SGP Extraction: Space Charge Capacitors

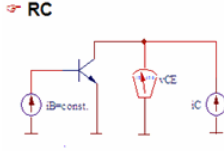


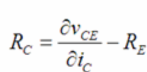
Linearizing the measured C(v) values

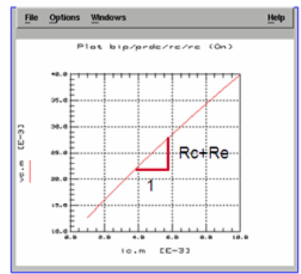


SGP Extraction: Parasitic Resistors: RE



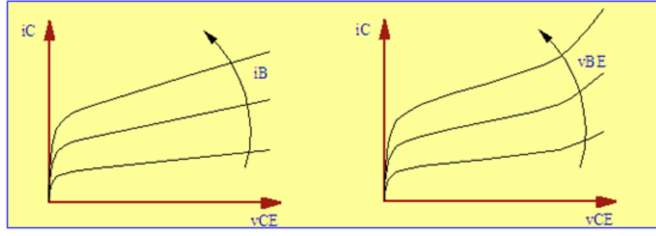




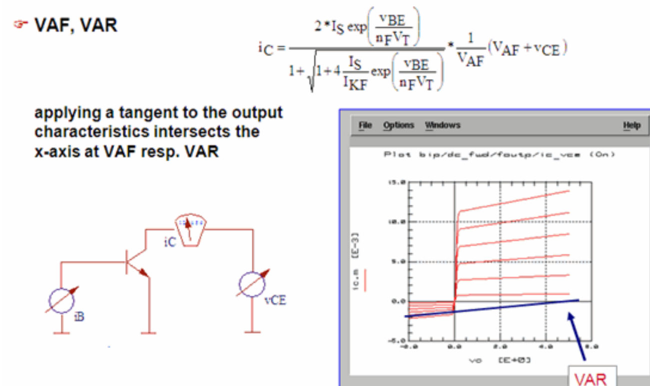


SGP Extraction: DC biasing

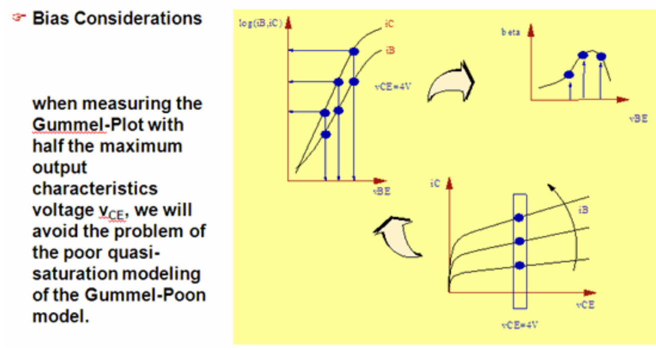
Bias Considerationsapplying a Base current instead a Base voltage prevents from measurements which include the thermal self-heating. Reason: the temperature-varying vBE is not affecting the measurement, because we force the Base current.



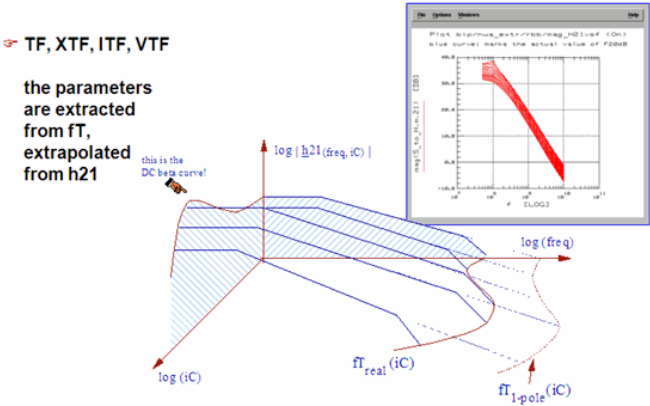
G-P Extraction: Early-Voltages



G-P Extraction: Gummel-Poon Plot Parameters

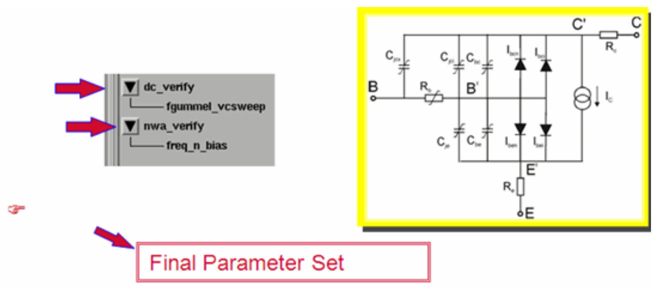


SGP Extraction: TFF Parameters

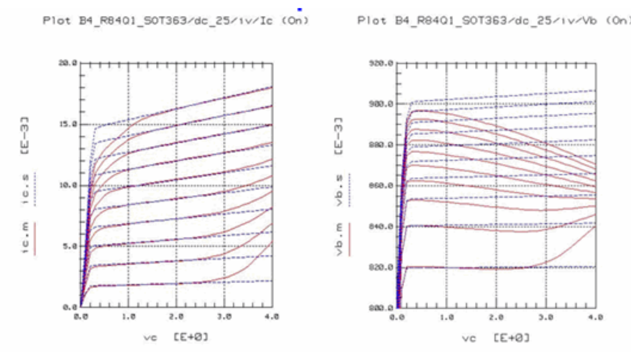


SGP Extraction: Final Parameter Tuning

- Check the modeling result for both, DC and S-parameters
- Fine-tune by optimization, if required



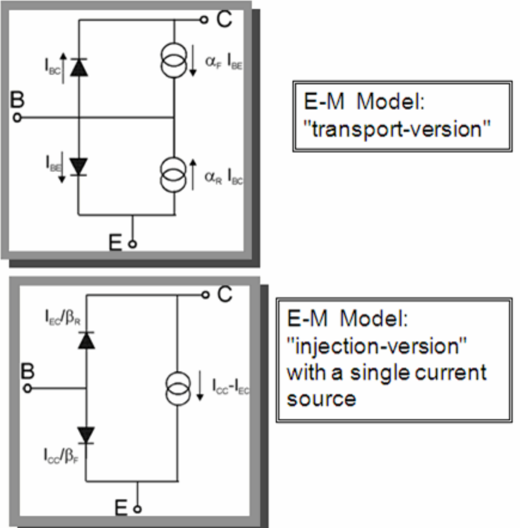
SGP Extraction: Final Result - Example



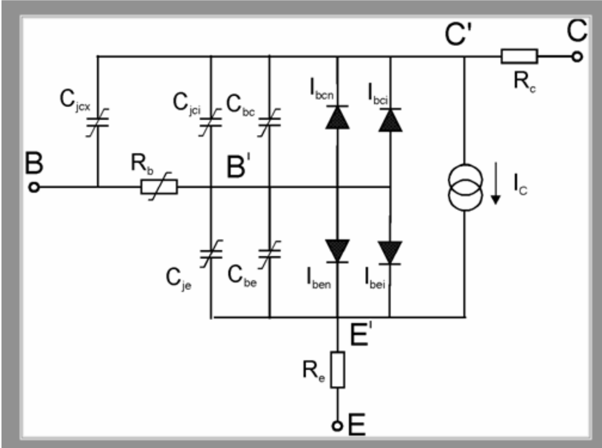
Bipolar Transistor Modeling

History of Bipolar Transistor Modeling

- 1954: Ebers-Moll Model
- Enhanced to EM1, EM2, EM3
- 1970: Gummel-Poon Model
  - Integral charge control relation
  - many improvements
  - widely accepted standard model:SPICE G-P Model
- 1986: MEXTRAM Model
  - Developed by Philips
- 1995: VBIC Model vertical bipolar inter-company model
  - Developed by a consortium ofover 10 US companies



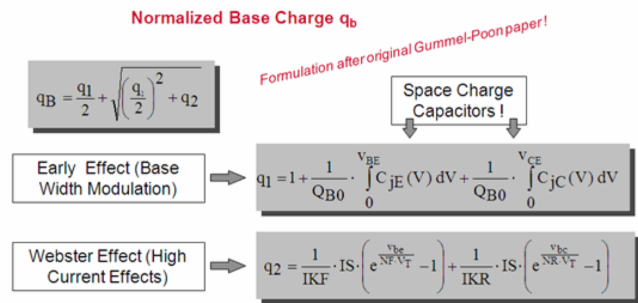
The Gummel-Poon Model



- Voltage-controlled Current Source Ic(vbe,vbc)
- Ideal (n = 1) and non-ideal (n = 2) Base Currents in Forward and Reverse Biased Mode
- Space Charge Capacitors
- Diffusion Capacitors
- Bias-dependent Base Resistor
- Fixed Base and Emitter Resistors

GP-Model: normalized Base Charge q_b

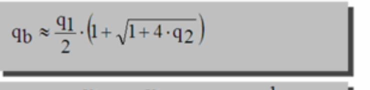
The Dominant Internal Model Parameter:



GP-Model: normalized Base Charge q_b

Default Implementation in Berkeley SPICE:

SPICE G-P

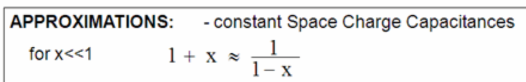


Decouples Early Effect from High Current Effect

$$q_1 \approx 1 + \frac{V_{bc}}{VAF} + \frac{V_{be}}{VAR} \approx \frac{1}{1 - \frac{V_{bc}}{VAF} - \frac{V_{be}}{VAR}}$$

Decoupling of Forward and Reverse Early Effect

Constant Output Resistance r_{out}



This simplification is implemented in nearly all of today's simulators !!!

GP-Model: DC Formulation

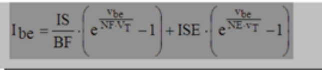
Current Source I_c:

- Forward and Reverse Operation
- depending from q_b

$$I_C = \frac{I_F - I_R}{q_b} = \frac{IS}{q_b} \cdot \left(e^{\frac{V_{be}}{NF \cdot V_T}} - 1\right) - \frac{IS}{q_b} \cdot \left(e^{\frac{V_{bc}}{NR \cdot V_T}} - 1\right)$$

Base Current:

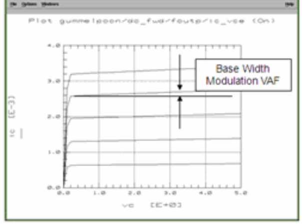
- Forward and Reverse Operation
- ideal and non-ideal Current
- ideal Current coupled to Collector Current Ic



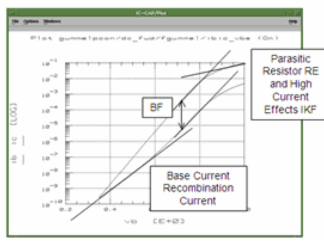
$$I_{bc} = \frac{IS}{BR} \cdot \left(e^{\frac{V_{bc}}{NR \cdot V_T}} - 1\right) + ISC \cdot \left(e^{\frac{V_{bc}}{NC \cdot V_T}} - 1\right)$$

Visualization of the most important DC Effects

Output Characteristics

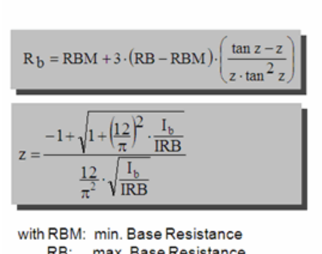


Gummel Plot



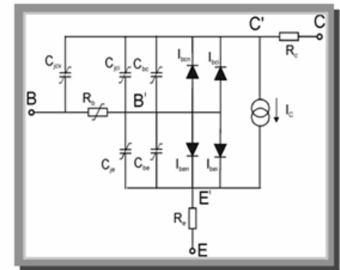
GP-Model: Parasitic Resistors

- Emitter Resistor RE constant
- Collector Resistor RC constant
- Bias-depending Base Resistor



$$R_b = RBM + 3 \cdot (RB - RBM) \cdot \left(\frac{\tan z - z}{z \cdot \tan^2 z} \right)$$

$$z = \frac{-1 + \sqrt{1 + \left(\frac{12}{\pi}\right)^2 \cdot \frac{I_b}{IRB}}}{\frac{12}{\pi^2} \cdot \sqrt{\frac{I_b}{IRB}}}$$



with RBM: min. Base Resistance
RB: max. Base Resistance
IRB: I_b @ (RB-RBM) / 2

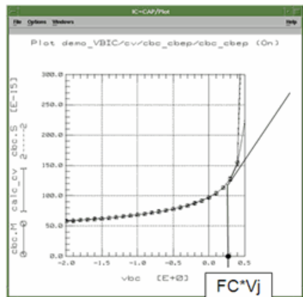
GP-Model: Capacitors

Space Charge Capacitors

- Parameter: C_{j0}, V_j, m

$$C_{ji} = \frac{C_{ji0}}{\left(1 - \frac{V_{ji}}{\Phi_i}\right)^{m_i}}$$

- Linear continuation for v > FC*V_j
- Base-Collector capacitance distributed between inner and outer Base by XCJC
- Substrate capacitance modeled by a constant or using the C_{ji} formula

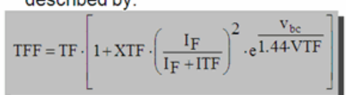


Diffusion Capacitors

- Determined by the charges

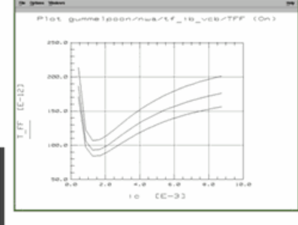
$$Q_{be} = TFF \cdot I_F \qquad Q_{bc} = TR \cdot I_R$$

- Forward transit time empirically described by:



$$TFF = TF \cdot \left[1 + XTF \cdot \left(\frac{I_F}{I_F + ITF} \right)^2 \cdot e^{\frac{V_{bc}}{1.44 VTF}} \right]$$

- Reverse transit time is constant



G-P Model: Additional Phase Shift and Temperature

- Additional Phase Shift at high frequencies: measured phase is bigger than simulated
- Additional phase shift parameter PTF
  - Frequency Domain: added linearly
  - Transient Analysis: second order Bessel function
- Temperature
  - Reflects constant ambient temperature
  - No self-heating

GP-Model: Model Parameters

Using Default Parameter Values switches-off the Parameter Effect (do not confuse with Typical Parameters).

Parameter	Default
DC Forward	
IS	0.1f
NF	1
BF	100
ISE	0
NE	1.5
IKF	∞
DC Reverse	
NR	1
BR	1
ISC	0
NC	2
IKR	∞
Early Modeling	
VAF	∞
VAR	∞

Parameter	Default
Space Charge Capacitances	
CJE	0
VJE	0.75
MJE	0.33
CJC	0
VJC	0.75
MJC	0.33
XCJC	1
CJS	0
VJS	0.75
MJS	0.33
FC	0.5

Parameter	Default
Resistances	
RE	0
RC	0
RBM	RB
RB	0
IRB	∞
Delay Time (Transit Time)	
TF	0
XTF	0
ITF	0
VTF	∞
TR	0
Excess Phase	
PTF	0
Temperature	
TNOM	27
EG	1.11
XTI	3
XTB	0

G-P Extraction: Notes

- For a better overview, we will use these simplifications:
  - no parasitic capacitors
  - no series inductors
  - no temperature modeling
- Simulator: spice2
- IC-CAP example file: gp_classic.mdl

G-P: Parameter Extraction Strategy

Parameters are dependent on each other:

- extraction sequence is important
- alternate approaches are possible

CV

- Space Charge Capacitances

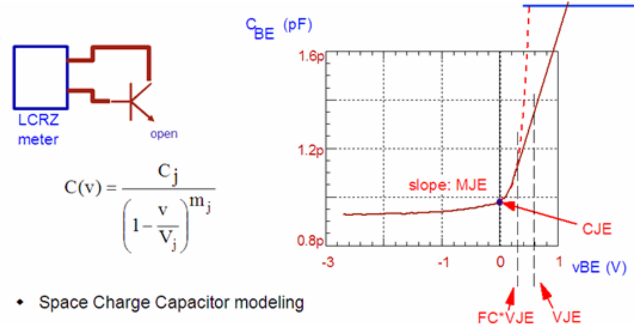
DC

- Parasitic Resistors
- Early Voltage
- Gummel-Poon Plot Parameters

S-Parameters

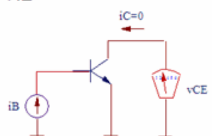
- RB Parameters
- TFF Parameters

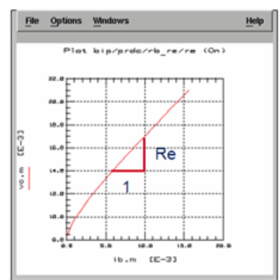
G-P Extraction: Space Charge Capacitors

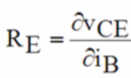


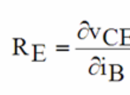
G-P Extraction: Parasitic Resistors: RE

- RE



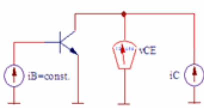


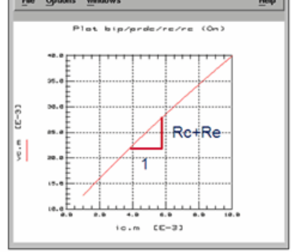




$$R_E = \frac{\partial v_{CE}}{\partial i_B}$$

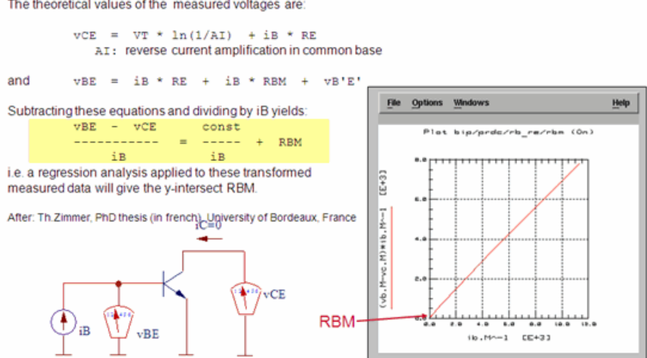
- RC





$$R_C = \frac{\partial v_{CE}}{\partial i_C} - R_E$$

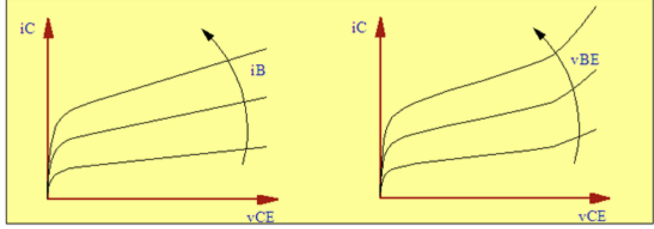
G-P Extraction: Parasitic Resistors: RBM



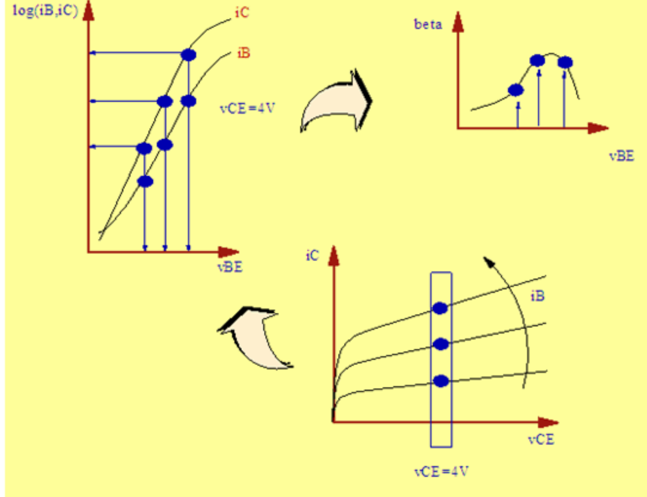
G-P Extraction: DC Biasing

Bias Considerations

Applying a Base current instead a Base voltage prevents from measurements which include the thermal self-heating.
Reason: the temperature-varying vBE is not affecting the measurement, because we force the Base current.

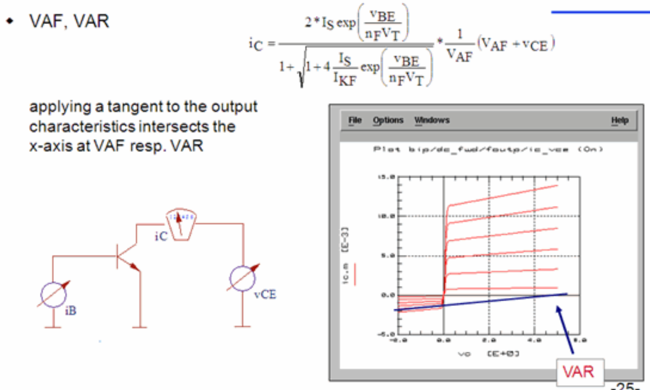


When measuring the Gummel-Plot with half the maximum output characteristics voltage vCE, we will avoid the problem of the poor quasi-saturation modeling of the Gummel-Poon model.

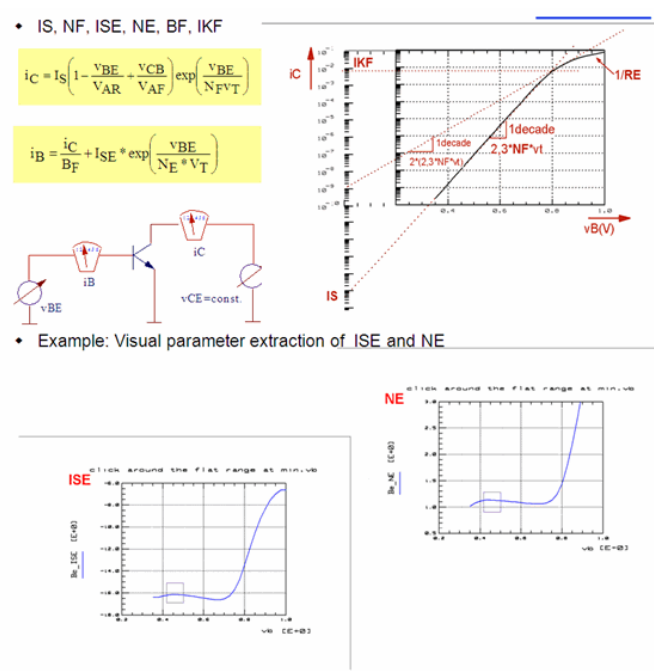


G-P Extraction: Early-Voltages

VAF, VARapplying a tangent to the output characteristics intersects the x-axis at VAF resp. VAR

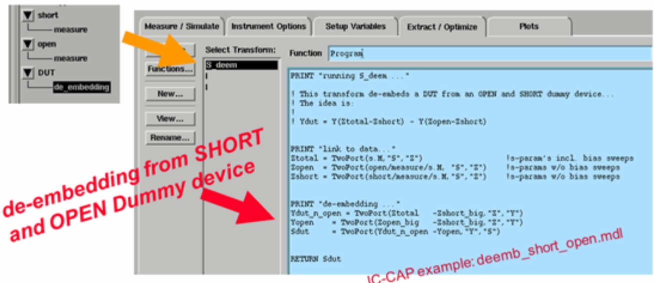


G-P Extraction: Gummel-Poon Plot Parameters



G-P Extraction: S-Parameter De-embedding

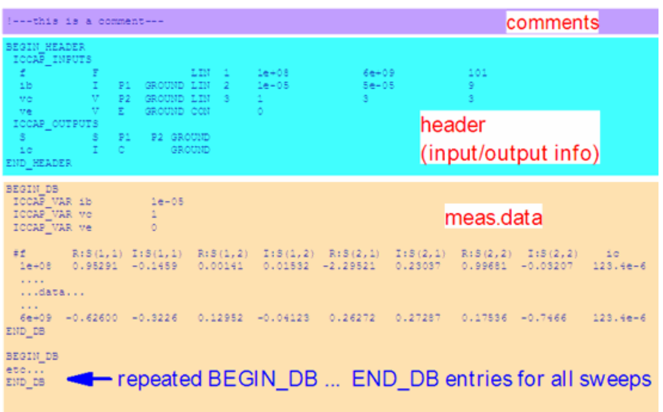
Since the S-parameters of a transistor are always affected by parasitics, de-embedding is required. It is achieved in IC-CAP with a simple PEL program, suited to the specific needs e.g. de-embedding using Y matrix subtraction only: usually ok for <10GHz-> or full OPEN-SHORT de-embedding like below).



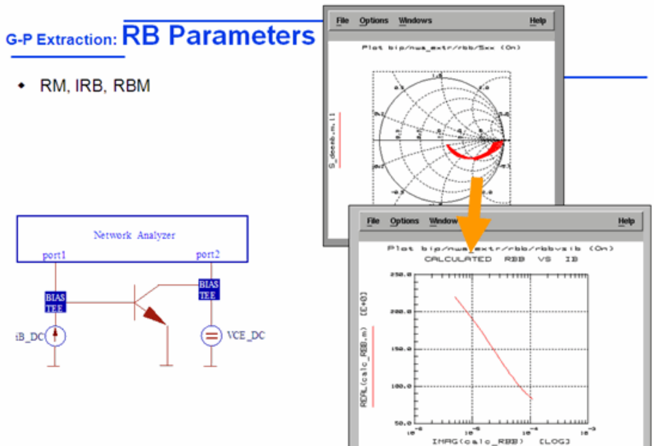
G-P Extraction: S-Parameter: Data Management

In order to reduce NWA calibration time and S-parameter measurement time considerably, we will use the data managementfeature in IC-CAP:i.e.:
the NWA is calibrated only once for a full frequency sweep. Then, the S-parameters are measured for all DC bias conditions, de-embedded once and finally stored to an IC-CAP .mdm file. The parameter extraction setups for rBB (frequency swept, ib swept, vc=constant) and TFF (frequency=constant, ib swept, vc swept) are then re-imported as a subset of de-embedded data out of this .mdm file.

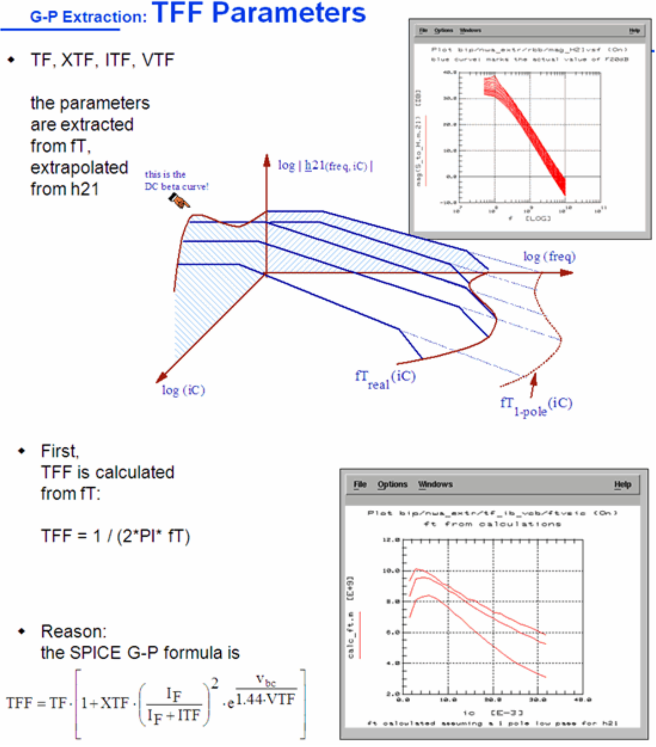
G-P Extraction: Data Management ASCII File



G-P Extraction: RB Parameters



G-P Extraction: TFF Parameters



$$TFF = TF \cdot \left[1 + XTF \cdot \left(\frac{I_F}{I_F + ITF} \right)^2 \cdot e^{\frac{V_{bc}}{1.44 \cdot VTF}} \right]$$

G-P Extraction: TFF Parameters -3-

- Extracting TF, XTF, ITF

$$TFF = TF \cdot \left[1 + XTF \cdot \left(\frac{I_F}{I_F + ITF}\right)^2 \cdot e^{\frac{v_{bc}}{1.44 VTF}}\right]$$

For vBC as small as possible, the exponential term in the TFF formula can be neglected. However, there is usually a limited measurement range, and the effect of XTF and ITF is difficult to see.

$$TFF \sim TF \cdot \left[1 + XTF \cdot \left(\frac{I_F}{I_F + ITF}\right)^2\right]$$

Therefore, a good approach is:

```
TF = min(TFF)
XTF= max(TFF)/TF-1
ITF= MEAN(iC)
```

TFF(psec)

TF(1+XTF)

measured curve

(effect of the space charge capacitors)

ITF

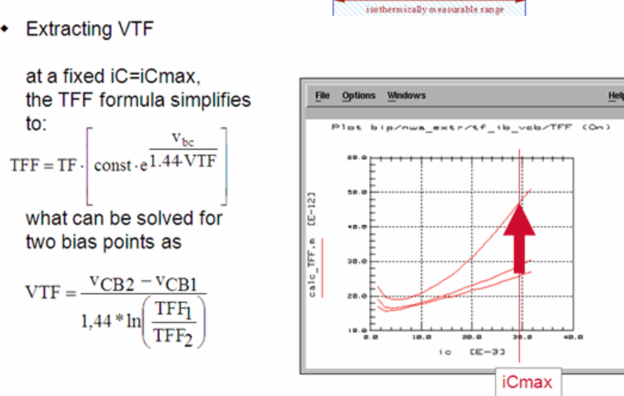
TF

theoretical curve

(from the TFF formula)

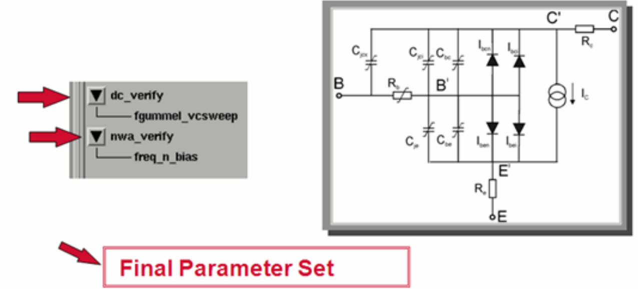
iC

isothermically measurable range



G-P Extraction: Final Parameter Tuning

- Check the modeling result for both, DC and S-parameters
- Fine-tune by optimization, if required

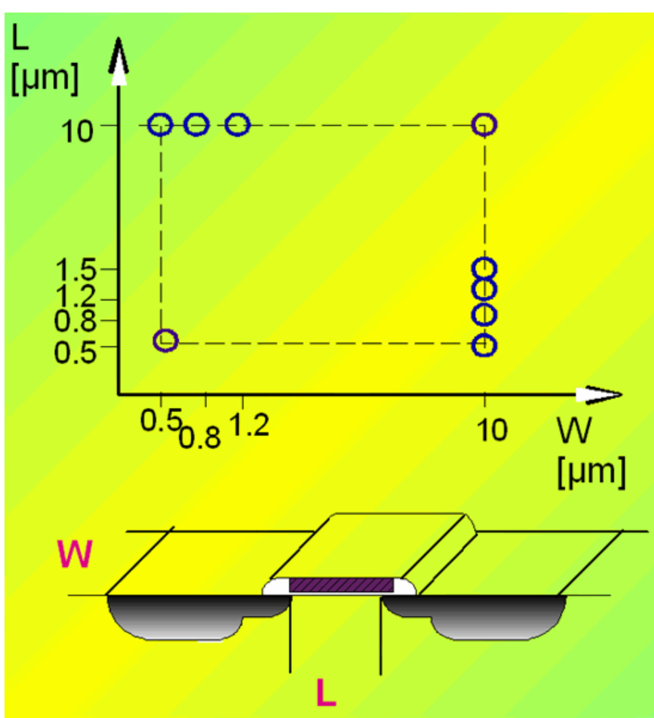


BSIM

Contents

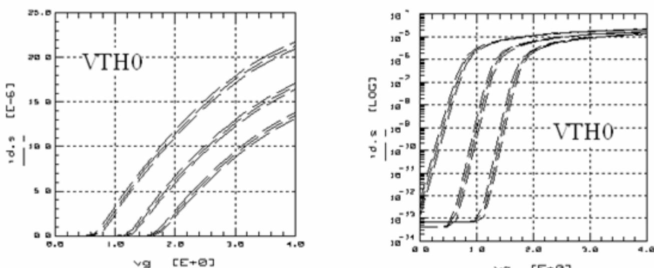
- BSIM3.3v2 - Parameter Extraction Strategy
- Fully Scalable BSIM4 Modeling From DC to RF

BSIM3.3v2 - Parameter Extraction Strategy

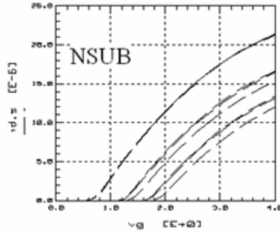


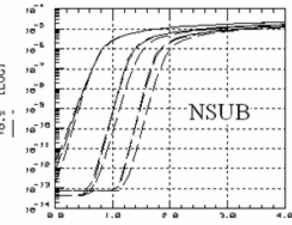
Since IC-CAP version 5.2, a special IC-CAP Toolkit is available for BSIM3 modeling. Therefore, for a detailed technical documentation about the model and the parameter extraction, see the IC-CAP manuals.
The following figures are intended to give an overview, which model parameters affect which fitting problem. The sequence given here is also a possible model parameter extraction strategy.

LARGE: Transconductance id(vg, vb)

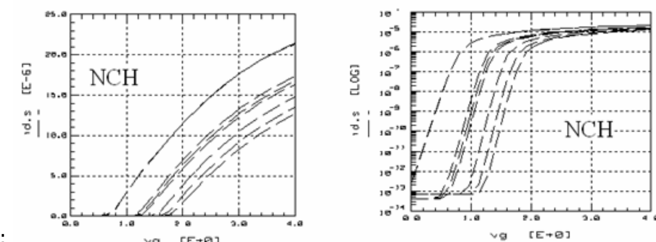


VTH0: equally shifting the transconductance curves (threshold voltage)

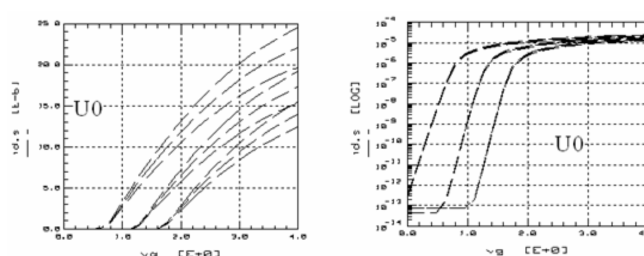




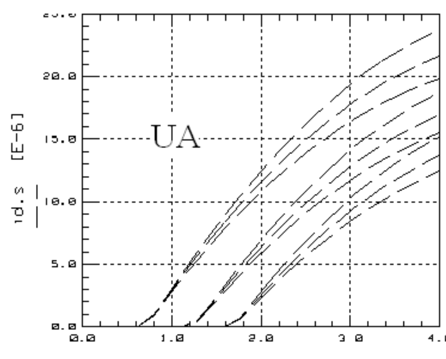
LARGE:
NSUB: shifting the threshold voltage



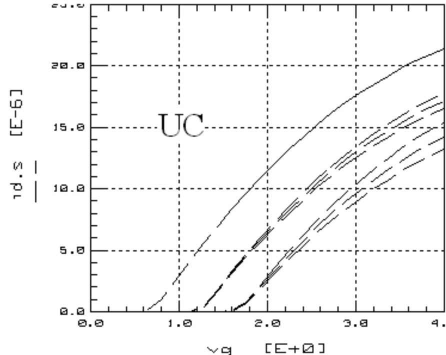
LARGE:
NCH: 2nd order shifting of the threshold voltage



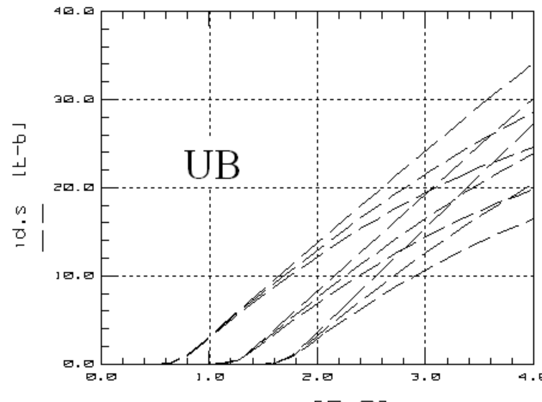
LARGE:
U0: slope of the transconductance curve (Mobility)



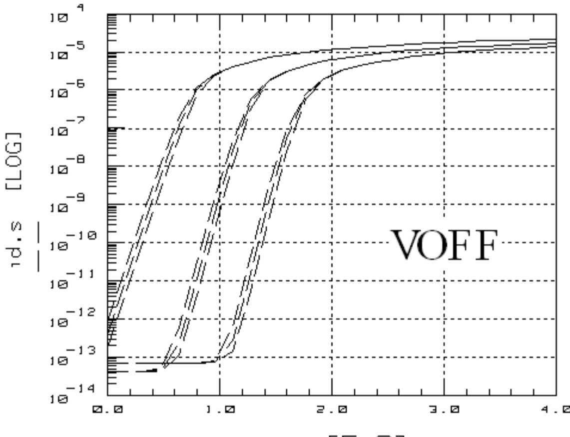
LARGE:
UA: change in slope of the transconductance curve (Reduction of the mobility U0).



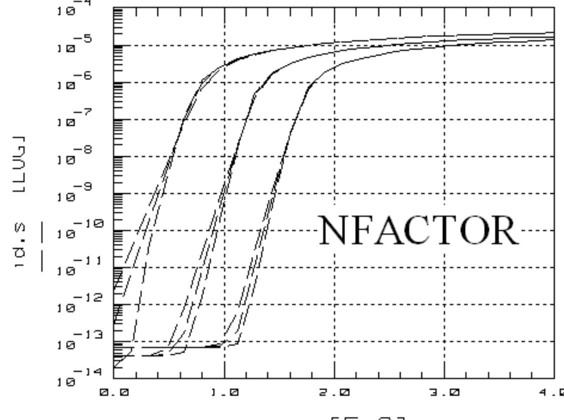
LARGE:
UC: changing the slope of the mobility related to vB



LARGE:
UB: changing the slope for big values of vG

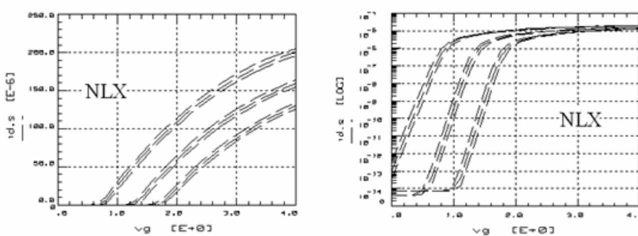


LARGE:
VOFF: y-intersect for a half-logarithmic transconductance curve

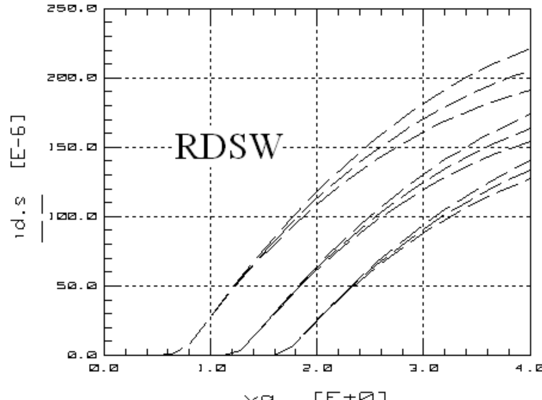


LARGE:
NFACTOR: slope of the half-logarithmically plotted transconductance curve

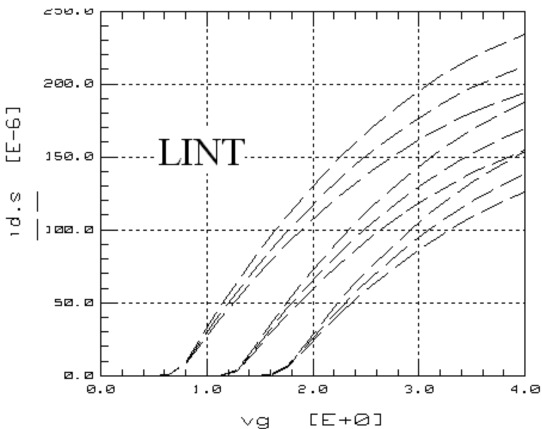
SHORT: Transconductance curve id(vg, vb)



NLX: an equal shift of the transconductance curve for the SHORT

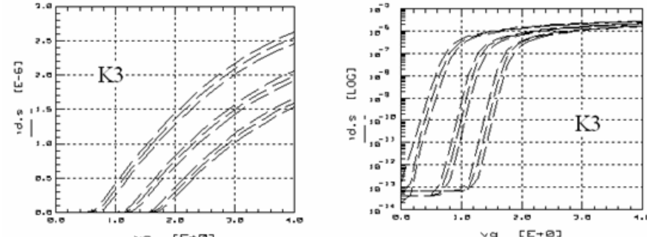


SHORT:
RDSW: change in slope of the transconductance curve for the SHORT

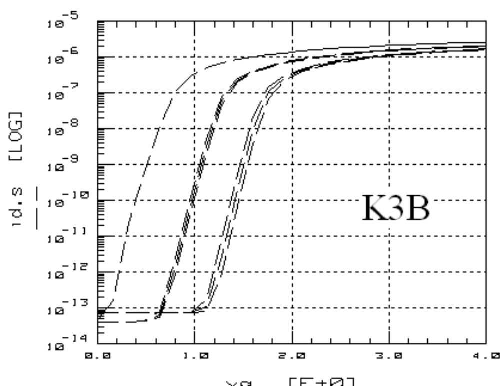


SHORT:
LINT: higher order change in slope for higher Drain currents

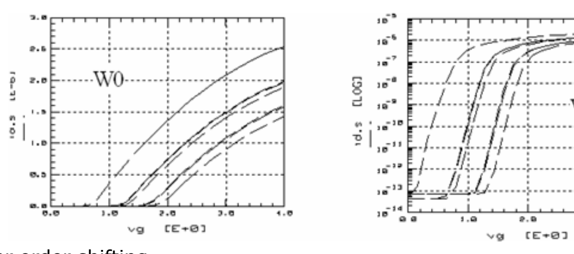
NARROW: Transconductance curve id(vg, vb)



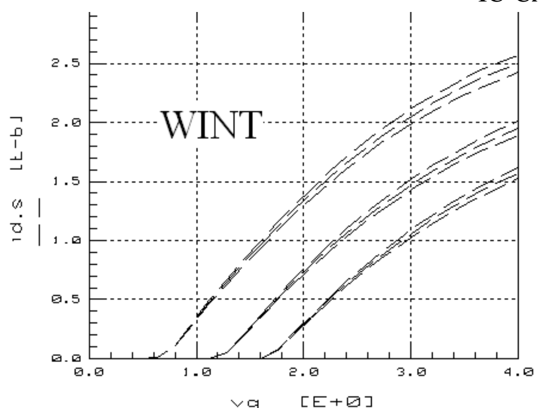
K3: shifting the transconductance curves



NARROW:
K3B: 1st order shifting of the transconductance curves

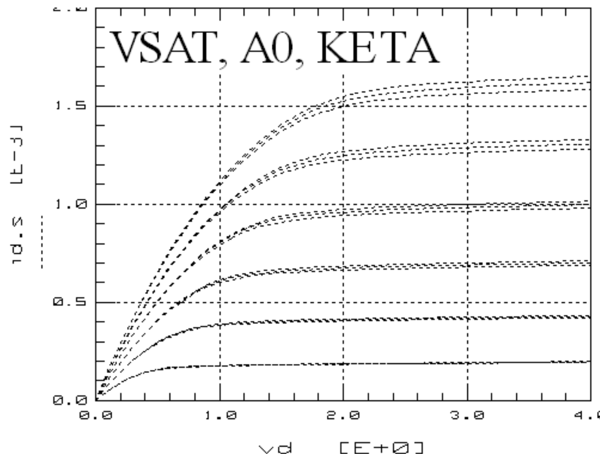


NARROW:
W0: higher order shifting



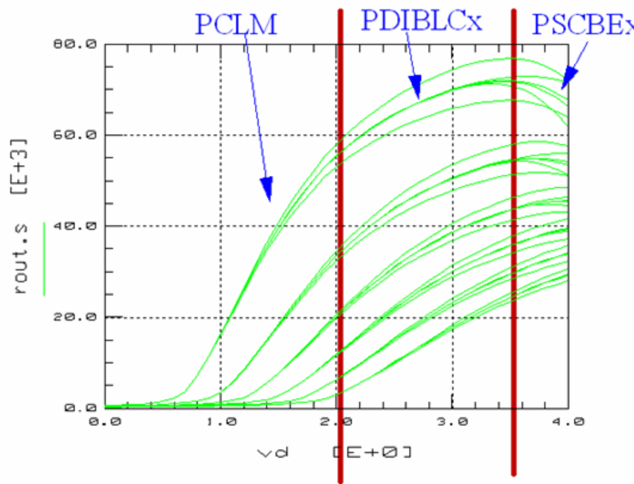
NARROW:
WINT: change in slope

SHORT: Output characteristics id(vd, vg)



VSAT, A0, KETA: amplification of the output characteristics level

SHORT: Output resistance rout(vd, vg)



PCLM: Fitting for low vD channel length modulation
PDIBLC1, PDIBLC2: Fitting for medium vD drain induced barrier lowering
PSCBE1, PSCBE2: Fitting for high vD substrate current induced body effect

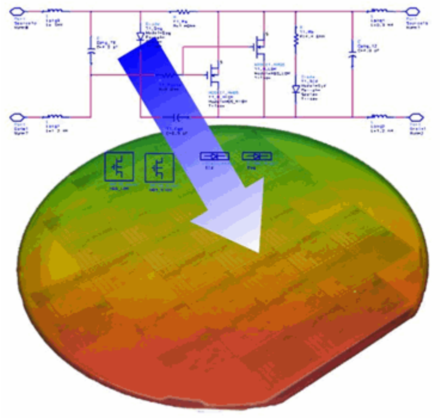
Fully Scalable BSIM4 Modeling From DC to RF

Company Information

AdMOS offers user specific and flexible solutions for a wide variety of applications in the area of modeling and simulation.
Engineering services:

- Modeling service to generate design libraries for semiconductor devices (CMOS down to 90nm, RF CMOS, etc.)

- Simulation models for passive components like on-chip passive, PCB elements for high speed circuits, connectors and packages.

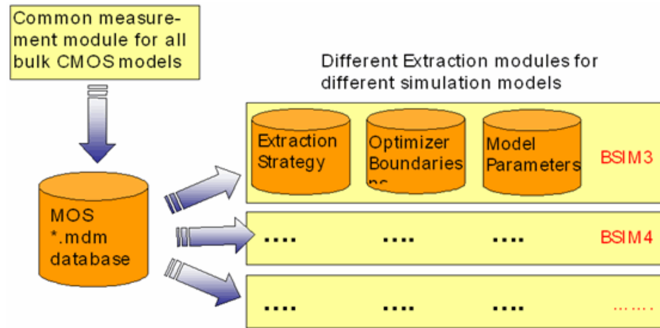


For more info, please see www.admos.com

> The principal services and products of Advanced Modeling Solutions are in the area of device modeling and connector design.

Standardized and reusable measured data sets for MOS modeling

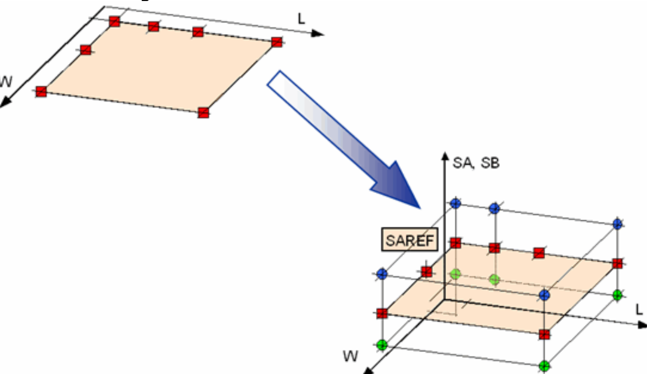
Data reuse for different models



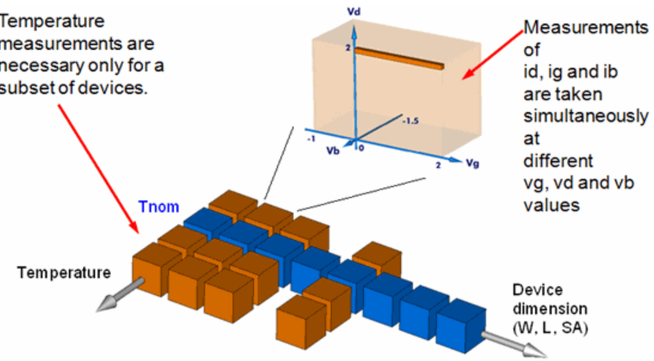
> The second part of the database strategy is to add additional information for each derived type of model to the common set of measured data.
> A user can define a certain extraction strategy (order and type of extraction functions) for a certain model, e.g. BSIM4 which is used to determine the model parameters.
> This extraction strategy is stored together with the final results (model parameters) and the boundaries of the parameters in the project data base.
> This makes all parameter extractions repeatable and self consistent.

Test devices for STI stress effects

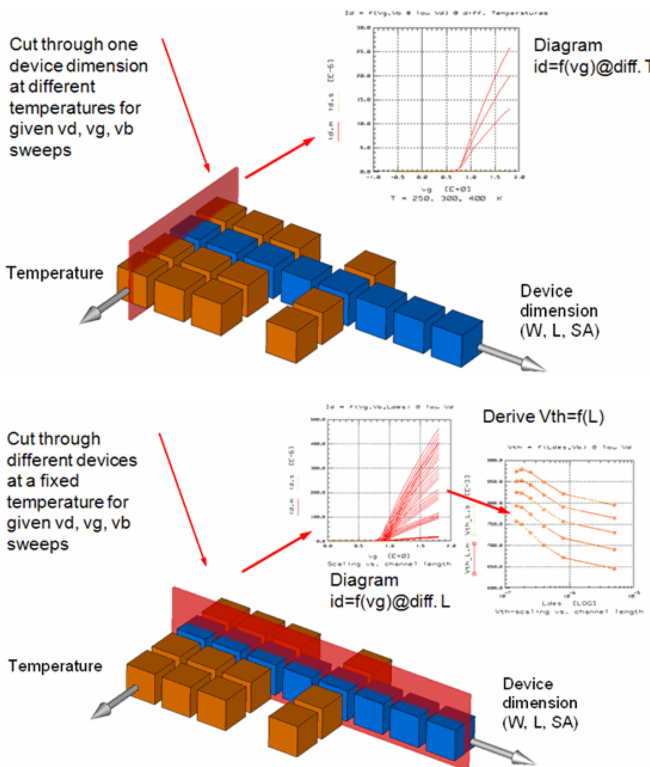
A third dimension is added by BSIM4.3.0, with SA, SB (SD) as parameters. In this diagram, the Temperature as a 4th dimension is neglected.
Until BSIM4.2.1, gate length (L) and gate width (W) have been the major device geometry parameters. The layout of test devices was focused on these two instance parameters to account for length- and width scaled effects.



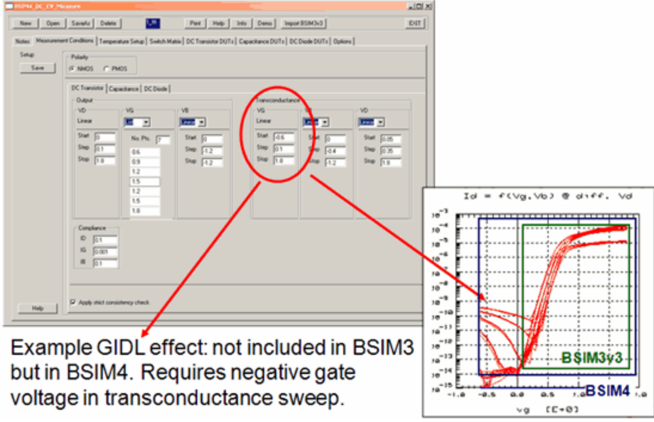
Data concept



Data operation

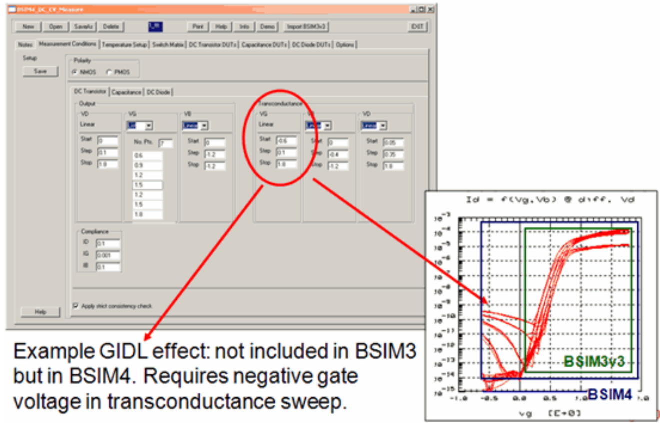


Measurement concept
The measurement environment is in independent from any particular simulation model.



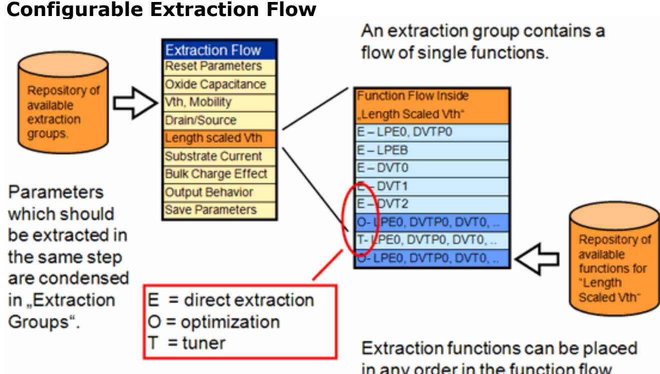
This slide describes the key features of the new BSIM3 Modeling Package. The following slides will show each of these features in detail.

The resulting database of .mdm files can easily be generated by other programs than IC-CAP, e.g. by high volume production measurement environment. All files are in ASCII format.



General Parameter Extraction features in the BSIM4 Modeling Package

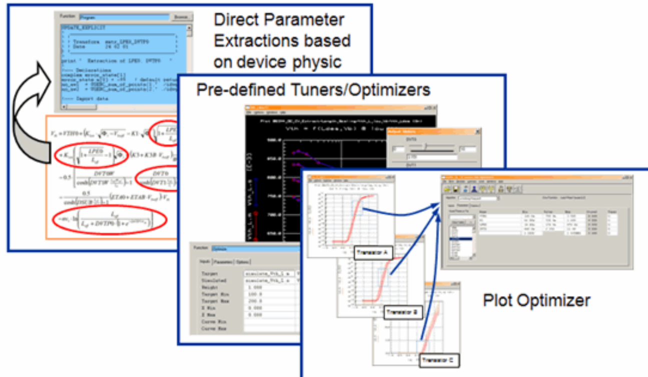
Configurable Extraction Flow



This feature can be configured.

The extraction methodology in the BSIM3/4 Modeling Package is open and can be configured very flexible to adopt the software to a certain CMOS process.
In many cases, model parameters are grouped together when they describe the same physical effects. These extraction groups can be arranged in the first hierarchical level, the extraction flow, in any order. It is also possible to invoke a certain group more than once if necessary.
Inside an extraction group, the flow of the functions (second hierarchical level) to determine a certain model parameter can be set by the user. For each model parameter, either the physically based direct extraction method, the manual tuner or the optimizer can be invoked to determine this parameter. In many cases, the tuner or the optimizer are modifying different parameters in one run, which have been extracted separately.
Using this method, the complete extraction strategy can be easily modified and adopted to the special need of a certain CMOS process.
Once a certain extraction strategy is established, the fully automatic run can be used. In this mode, all extractions are called in the given sequence inside a macro and the intermediate results are logged either in the failure log file or the output window log file. This method is ideal to generate a large number of model parameter sets from different measurements for a further use in a statistical modeling approach.

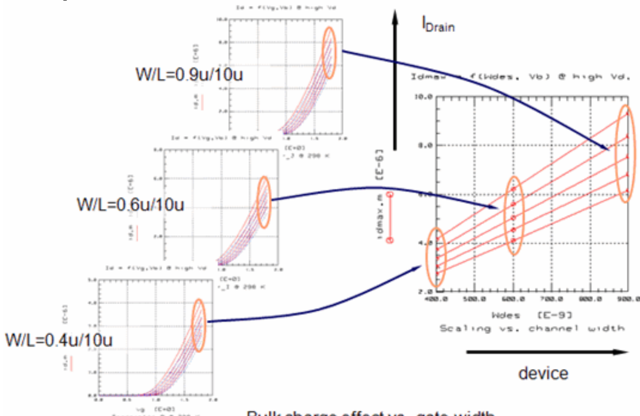
Extraction Method



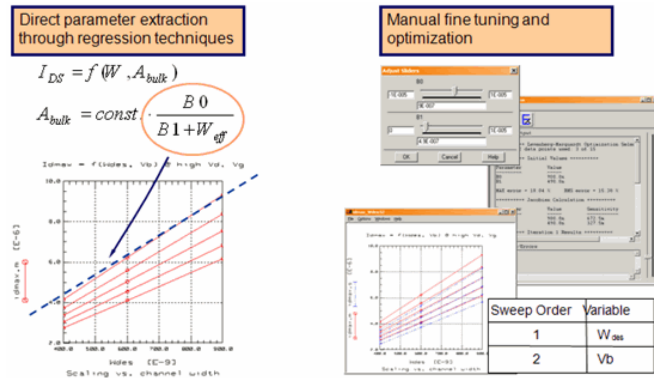
As already mentioned in the last slide, the BSIM3/4 Modeling Packages support different methods of the determination of model parameters.
The first one is the physically based direct parameter extraction. Here, the parameters are

calculated according to their meaning in the model equation directly from measured data. These functions carefully check whether certain physical effects can be seen in the actual data set and they isolate the valid range of measured data to determine a certain effect. An example for such a parameter extraction for different width scaled devices is given in a later slide.
The pre defined tuners and optimizers automatically select reasonable optimizing targets and ranges. They contain a dedicated parameter check and can be customized by the user as well.
Moreover, the BSIM3/4 Modeling Packages support the usage of the new plot optimizer feature. In a special setup, certain diagrams from different devices at different temperatures can be grouped together to build the base for applying the plot optimizer feature. Once such an optimization is set up, it can be included in the extraction flow.

Example: Direct Extraction for BSIM4



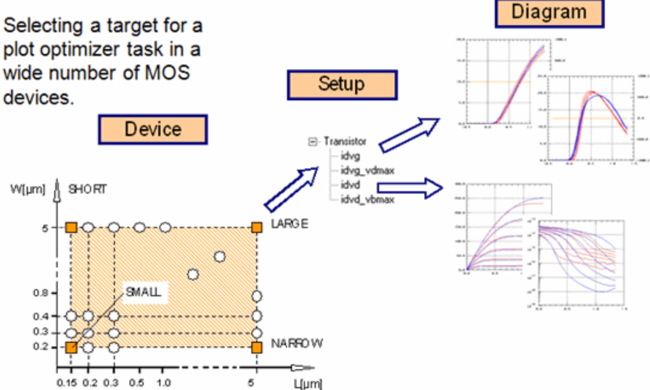
This examples demonstrates the generation of such a condensed data array to extract the width dependent bulk charge effect.
Here different devices with different gate widths and a constant large gate length are taken to isolate this gate width dependent effect. From each device, the drain currents at the maximum gate voltage and at different bulk voltages are extracted and all values are arranged in a new diagram on the right hand.
Now, this diagram has the gate width as an x-axis while different curves accounts for different bulk voltages Vb.
This diagram is the base for further extractions or optimizations as shown in the next slide.



In this slide, the new diagram is taken as a base for either direct extractions or optimizations.
On the left hand, the two equations demonstrate the principle of applying a linear regression method to determine the both parameters B0, B1. This can be done by using the pre-assumption that the gate length is constant while the gate width is variable. Then, after a view transformations, which are not shown in this slide, B0 and B1 can be determined from the slope and intersection of the applied tangent.

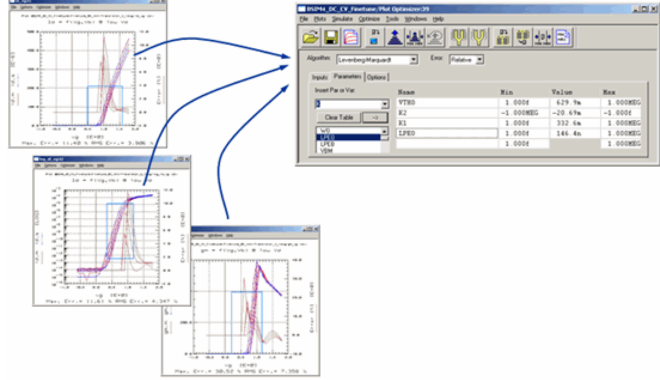
Moreover, the condensed diagram can be easily used to apply an optimizer or a tuner on the gate width dependent data set (see the right hand example). While the measured data is combined by data base operations, the simulated data can be directly generated. This is achieved by defining the gate width as a parametric input with sweep order 1 (equal to the x-axis of the diagram) and the bulk voltage as an input with sweep order 2. Using this setup, any optimization runs very quickly invoking the built-in spice3e2 simulator or any of the supported commercial simulators.

Example: Plot Optimizer



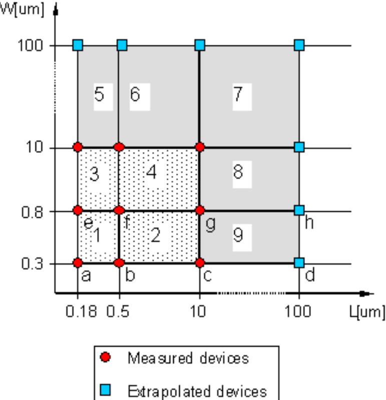
The plot optimizer allows an easy fine tuning of model parameters for different:

- devices
- setups
- regions of operation at once



Binning

- Binned models for BSIM3 and BSIM4 can be generated. They can be extended by extrapolation to cover a wider range of device dimensions than available through measurements.
- It can be configured, whether a certain effect which depends on the device geometry, e.g. the threshold voltage, is modeled by the built-in scalable model equations or by the binning feature.
- Taking only a few parameters as binning parameters can enhance the overall fitting quality while preserving the mostly physical nature of the model.
- Binned model sets are generated ready to use in the ADS, HSPICE and Spectre format.



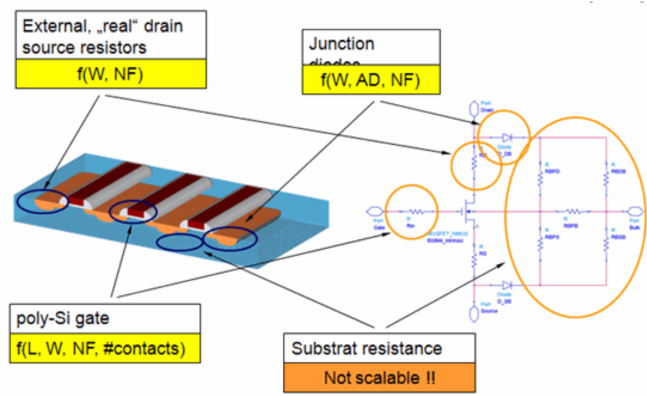
The BSIM3/4 Modeling Packages support the generation of binned models. They can be extended by extrapolation to cover a wider range of device dimensions than available through measurements. Without this extensions, many simulators would not be able to simulate devices with W=10um or L=10um ! The condition for selecting a certain binned area is:
Lmin <= L < Lmax.
That means, a device with L=10um cannot use the parameter set of bin #2 even but must select the parameter set of the extended bin #9.

Moreover, it can be configured, whether a certain effect which depends on the device geometry, e.g. the threshold voltage, is modeled by the built-in scalable model equations

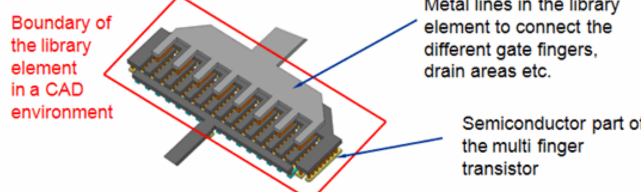
or by the binning feature.
Taking only a few parameters as binning parameters can enhance the overall fitting quality while preserving the mostly physical nature of the model.
Binned model sets are generated ready to use in the ADS, HSPICE and Spectre format.

Advanced BSIM4 RF models

Built-in BSIM4 RF Features

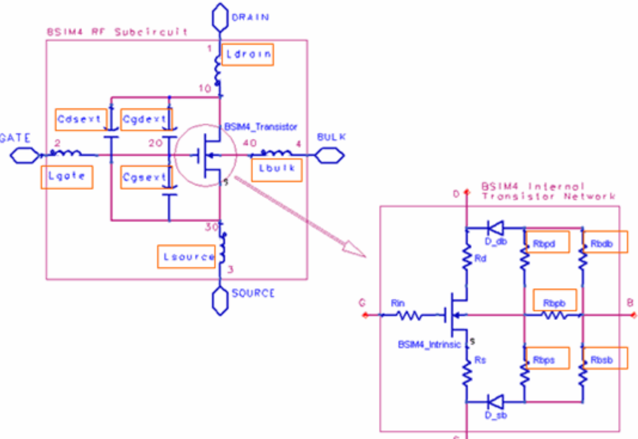


Missing BSIM4 RF Features



- The BSIM4 RF model takes only into account the behavior of the „semiconductor" part of the multi-finger transistor. The influence of the metal layers (crosstalk, delay) is not considered
- Gate length variations (see one of the next slides) inside a multi finger device are not taken into account
- The substrate resistance network is not scalable.

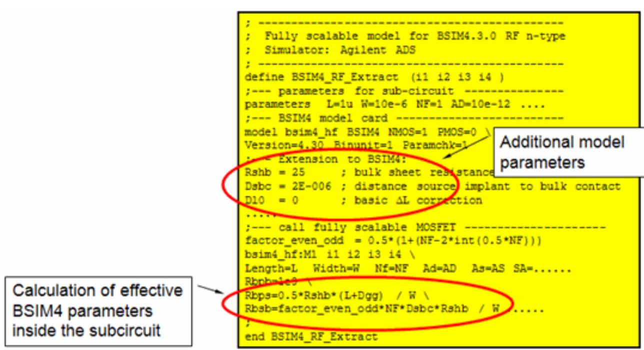
Structure of Scalable BSIM4 RF Macro Model
A sub circuit contains the core BSIM4 model surrounded by external elements representing crosstalk and delay behavior of metal layers.



The values the marked elements are calculated as functions of the device dimensions W, L and NF rather than setting them to fixed numbers.

Scalable BSIM4 RF Macro Model

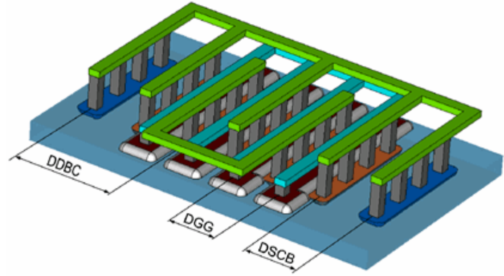
- Extensions to BSIM4
- Scalable substrate resistance model
- Channel length reduction variation inside multi finger devices is included
- Crosstalk and delay effects due to metal interconnection in multi finger device is included.



Scalable transistor models cover a certain range of major device dimensions. In case of a RF MOS transistor, these are the gate length and gate width of a single transistor finger and the number of gate fingers. These models have a structure, such that a design engineer can change the parameters to get an optimum transistor behavior for a certain application.

The fully scalable model has in principle the same structure than the single transistor model. However, it uses equations to determine the values for the substrate resistance parameters and the modified channel length reduction. These equations are derived from simple assumptions. Please see the following sketches and the model equations for more details.

Example: Scalable Substrate Resistance
Standard BSIM4.3 model:
RBPS, RBPD, RBSB, RBDB are fixed numbers



Scalable RF model:
Implementation of a set of equations according to [1] and own developments to describe a scalable substrate resistance behavior:

$$RBPS = RBPD = \frac{R_{SHB}(L + D_{gg})}{2 \cdot W} \qquad RBDB = \frac{factor_even_odd \cdot NF \cdot D_{dbc} \cdot R_{SHB}}{W}$$

$$RBSB = \frac{factor_even_odd \cdot NF \cdot D_{sbc} \cdot R_{SHB}}{W} \qquad factor_even_odd = \begin{cases} 0.5 & \text{for even NF} \\ 1 & \text{for odd NF} \end{cases}$$

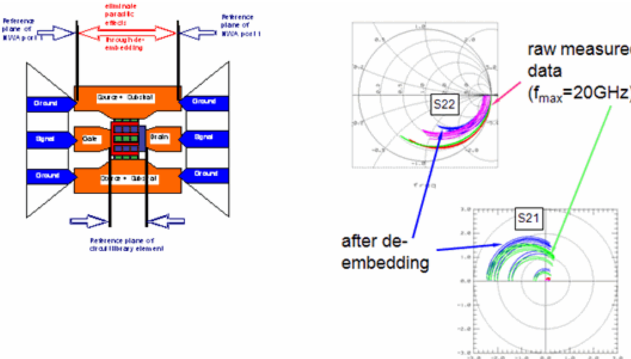
In the RF single device model, the values of the resistors RBPB, RBPD, RBPS, RBDB and RBSB in the substrate resistance network parameters are set directly to fixed values.

In contrast, for the RF scalable model, the substrate resistance network parameters RBPB, RBPD, RBPS, RBDB and RBSB are derived using 4 new model parameters:

RSHB sheet resistance of the substrate
DSBC distance between the source contact and the outer source area
DDBC distance between the drain contact and the outer drain area
DGG distance between two gate stripes

The equations to determine the values for the substrate resistance parameters contains in addition the gate length, gate width of one finger and the number of fingers.

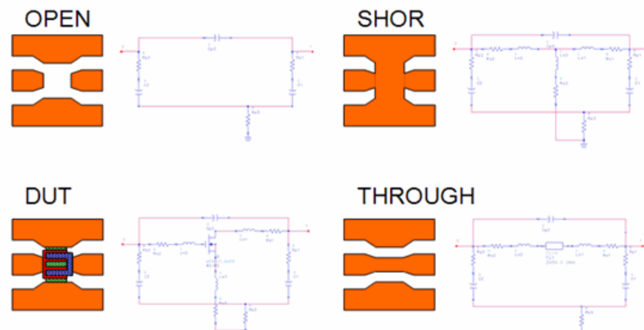
Measurement and De-embedding
The calibration plane of the NWA is located at the end of the probe tips.
The influence of the parasitic elements between the calibration plane of the NWA and the transistor must be eliminated through de-embedding.



The goal of a S-parameter measurement of an integrated RF transistor is to determine the frequency behavior of the transistor itself without taking into account the parasitic influence of the cables, probes, pads and the interconnection to the transistor. The first step to achieve this goal is to calibrate the network analyzer using a special probe substrate. This will move the calibration plane to the end of the probe tips. However, the parasitic influence of the pads and the interconnections to the transistor still remains. It must be eliminated by using de-embedding methods. After all these procedures, the remaining S-parameters describe the real behavior of the transistor as it will be used in a design library.

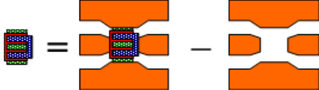
The two diagrams on the right hand demonstrate the effect of the de-embedding. The lupper one shows the input and output reflection S11 and S22 before and after the de-embedding. Looking at S22, the difference in both, phase and magnitude can be seen very clearly. A similar behavior can be observed in the transmission behavior. The raw measured data of the forward transmission S21 clearly has an increased phase shift and a decreased magnitude compared to the de-embedded data. As a summary, the raw measured data underestimates the real behavior of the RF MOS transistor. It is absolutely necessary to perform a correct de-embedding to get the correct transistor data as a base for accurate modeling.

De-embedding and verification test structures



The pre requisite for a correct de-embedding is that certain test structures are available on a wafer together with the device under test (DUT) itself. Depending on the selected de-embedding method (see the next pages) an OPEN, SHORT or THROUGH dummy pad structure must be available and must be measured. The principle layouts of these structures are given above. However, in real layouts, the transistor itself can be much smaller compared to the size of the dummy pads.
Together with the principal layouts, the ideal equivalent electrical circuits for these devices are given. The described de-embedding methods are only valid, if the electrical behavior of the dummy pad structures can be represented by these equivalent circuits.

De-embedding Procedures
OPEN
This method is limited to lower frequency applications (<3GHz) because it takes only into account parasitic effects parallel to the DUT.



SHORT-OPEN
Should be applied for data measured at frequencies roughly > 3GHz.

De-embed from Open:
$Y_{DUT/Open} = Y_{Total} - Y_{Open}$
$Y_{Short/Open} = Y_{Short} - Y_{Open}$

Convert to Z:
$Z_{DUT/Open} = Z(Y_{DUT/Open})$
$Z_{Short/Open} = Z(Y_{Short/Open})$

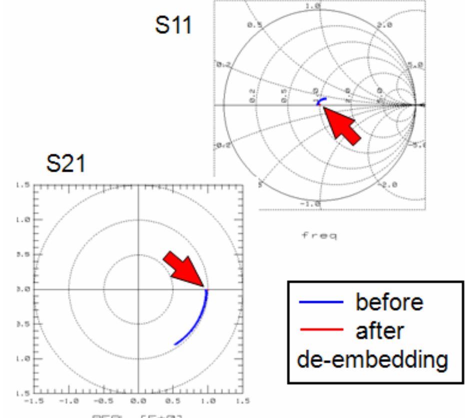
De-embed from Short:
$Z_{DUT} = Z_{DUT/Open} - Z_{Short/Open}$

Convert to S:
$S_{DUT} = S(Z_{DUT})$

This slide describe the two standard de-embedding methods which are implemented in the BSIM3/4 Modeling Package.
A common and very often used de-embedding method is the de-embedding from an OPEN device. The pre requisite for using this method is, that the behavior of the OPEN dummy pad structure can be represented only by parallel elements to the DUT (see the equivalent circuits). The transformed Y-parameters from the OPEN can simply be subtracted from the DUT's Y-parameters. The advantage of this method is that only one device in addition to the DUT is needed. However, this method is limited to lower frequencies roughly smaller than 3-5 GHz. This frequency limit is not fixed and depends on the layout and size of the whole test structure.A more advanced method to remove the parasitic influence from the measured data is the de-embedding using an OPEN and SHORT dummy pad device. The idea behind this method is, that the electrical behavior of the pads around the DUT can be described by a combination of parallel and serial circuit elements. The implementation is shown in the right part above. In a first step, the influence of the parallel parasitic elements is removed from both, the SHORT and the device under test (DUT). In the following step, the Y-parameters are converted to Z-parameters and in this representation, the influence of the serial parasitic elements is removed by subtracting the partially de-embedded Z-parameters of the SHORT from the partially de-embedded Z-parameters of the DUT. After a further conversion, the S-parameters of the DUT are available as a base for the device modeling.
Besides these both standard methods, the BSIM4 Modeling Package allows the implementation of user defined de-embedding methods.

De-embedding verification with THROUGH
After a correct de-embedding of the parasitic components, the S-parameters of the THROUGH should show the behavior of an ideal transmission line:

- Delay time TD representing the distance between the two pads.
- S11, S22 starting at characteristic impedance.



To ensure that the de-embedding is done correctly, the verification using a THROUGH device is an ideal method. The verification is very important because if not all parasitic effect are removed from the transistors measured data, the simulation model must fit unknown parasitic effects and does not represent the pure transistor, as it will be used in a design library.
The idea behind the verification is, that a THROUGH should behave like a transmission line after the de-embedding of the parasitic pad influence. The diagrams show the typical behavior of such a through before and after the de-embedding procedure. The blue curve shows this ideal transmission line behavior which is given by a phase shift in S21 due the delay time of the line and S11, S22 starting at the characteristic impedance in the Smith chart.

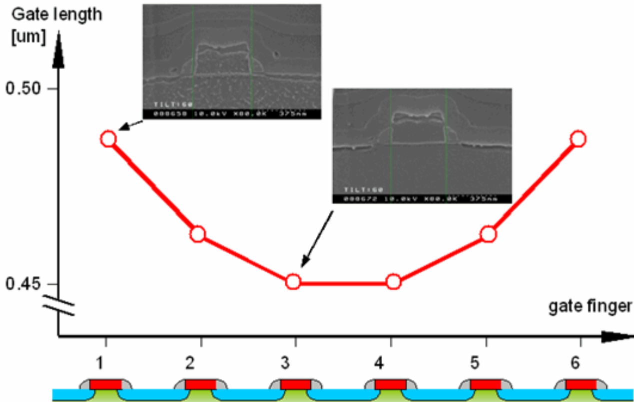
DC model modifications
The first step in achieving a precise RF model is to modify the DC model to match both, output resistance Rout together with the S-parameters at the lowest frequency. The following effects have to be considered:

- Output resistance: Has to be readjusted to match S22 and S21 at lowest frequency.
- Channel length reduction DL :Due to lithographic effects, the channel length

reduction can show a significant variation inside a multi-finger transistor.
- If the target simulator does not provide a multiplier for several gate fingers, the following effects have to be adjusted:
  - drain source resistance Rds
  - channel width reduction ΔW

After preparing the measured S-parameters, the real modeling task starts looking at the DC model parameters. DC model parameters are normally derived from measurements of single finger transistors of different gate lengths and gate widths. However, the typical RF MOS transistors are implemented as multi finger transistors which may show a different behavior than a singe finger transistor. The first step is, that the output resistance parameters mostly have to be re-adjusted to match both, the output resistance Rout = f(Vd) and the starting point of the S-parameters S21 and S22. This behavior will be shown in detail in the next slide. Another effect, which can often be observed within multi-finger transistors is, that the channel length reduction DL is different compared to single finger transistors. A research work of one of our customers showed, that DL can vary up to 30% within a multi finger transistor. Without adjusting DL in the simulation model, an accurate device behavior cannot be simulated. At least, depending on the target simulator, the drain source resistance and the channel width reduction have to be re-adjusted in the case that the target simulator does not have a multiplier for n numbers of gate fingers.

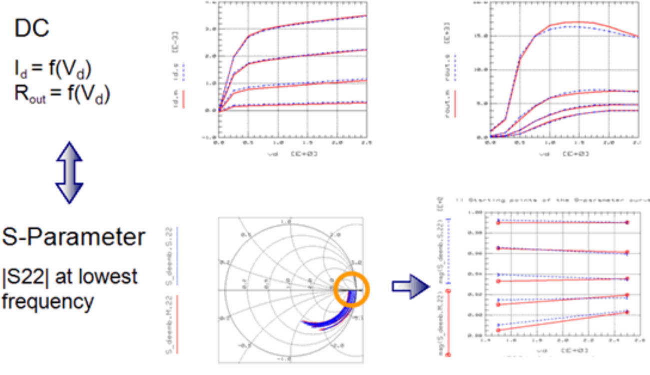
Gate length variation in multi-finger transistors



This diagram shows the effect of gate length variation inside a multi finger device, which can be observed in many CMOS processes. The lower sketch is a principle cross section through a typical multi finger devices with 6 gate fingers. It could be shown through SEM (Scanning Electron Microscope) pictures, that the length of the poly-Si gate is different for the outer gate fingers than for the inner gate fingers. However, that means, that the channel length reduction, which is normally derived from single finger transistors does not correspond to the effective channel length in such multi finger transistors.

When doing modeling, it is naturally not always possible to make this SEM analysis and it is mostly necessary to rely only on measured electrical data. The impact of this effect can be seen very good in a shift of the threshold voltage Vth, which is through the short channel effect heavily dependant on the effective gate length. In addition, it can be also seen in a difference between measurement and simulation of the drain current at high drain and gate voltages. However, if such an effect can be identified, it is necessary to compensate and to readjust the parameter LINT in the BSIM simulation models. The disadvantage of such a procedure is, that now the RF simulation model contains different values for the model parameters than the model for the conventional single finger devices.

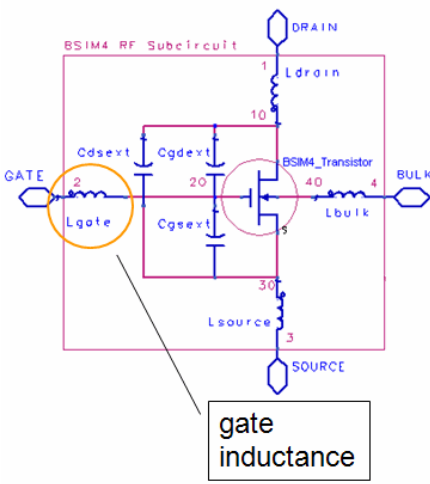
DC output resistance and S-parameters



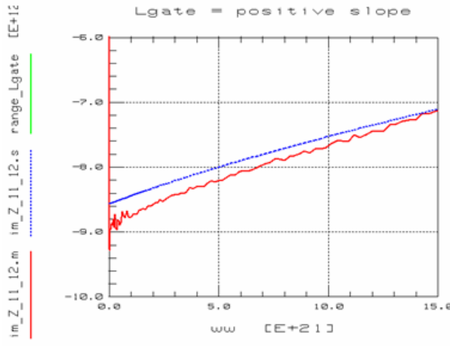
This slide shows the correlation between the DC behavior and the S-parameters of the DUT. The starting point of the S-parameters at lowest frequency (here 100MHz) are only determined by the DC model parameters of the transistor model. The lower right diagram shows the magnitude of S22 at different DC operating conditions at the lowest frequency. The very good fit between measured data (red) and simulated curves (blue) is only

achieved by a perfect fit of the DC behavior, mainly the output drain current and the output resistance Rout. Without correct starting points of the S-parameters at the lowest frequency, no good overall fit can be achieved. This is again an indication, that only physically based DC model parameters can be used for further RF modeling.

Extraction of external circuit elements

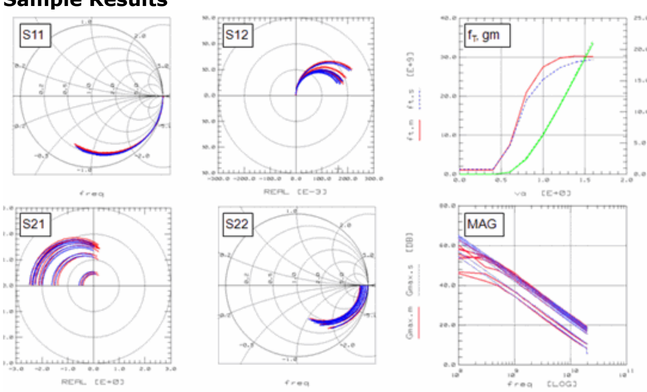


Example:
Extraction of the gate inductance from S-parameters measured at zero bias voltage.



After fitting the starting points of the S-parameters, the model parameter values of the external elements are determined from S-parameter measurements over the full frequency range. In the example shown above, the value of the gate inductance is determined from a measurement on the cold transistor, where the gate and drain voltage is set to zero to switch of the transistor influence.

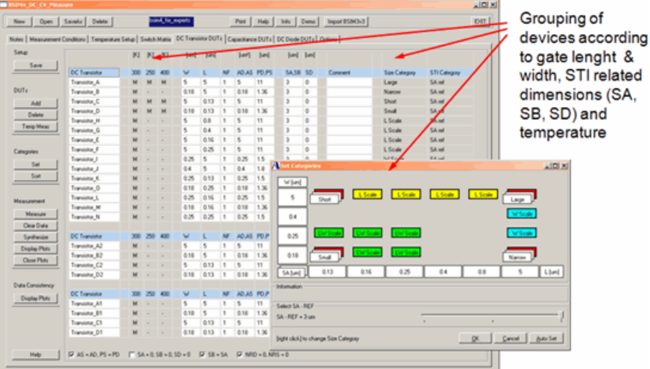
Sample Results



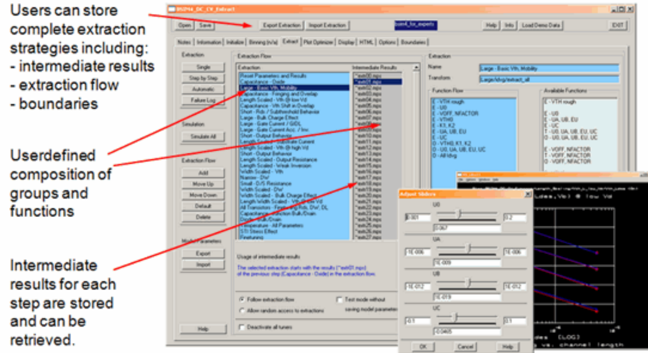
Frequency range: 0.1 .. 20GHz, L=0.24um, meas. sim.
Shown above are the final results of the model generation using the extended BSIM4 RF model. The red curves are representing the measured behavior of the MOS transistor at different DC operating conditions while the blue ones are the corresponding simulation results. It can be seen, that the fit between measured and simulated is very accurate. This could only be achieved by using physically based extractions which results in realistic values for resistance, channel length reduction etc. It can also be seen that the modifications of the extended BSIM3/4 RF model have been necessary to achieve this simulation results.

Advanced usability features

Test & Measurement

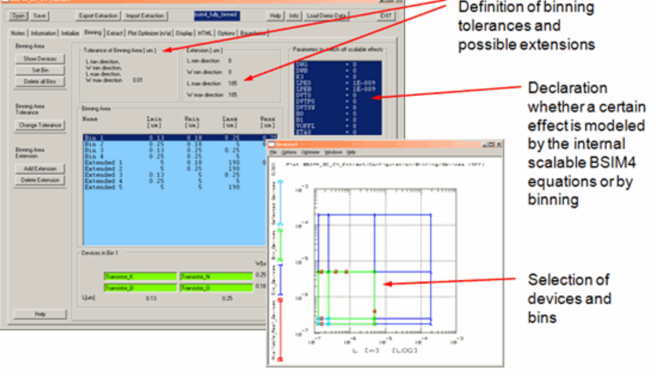


Configurable Extractions

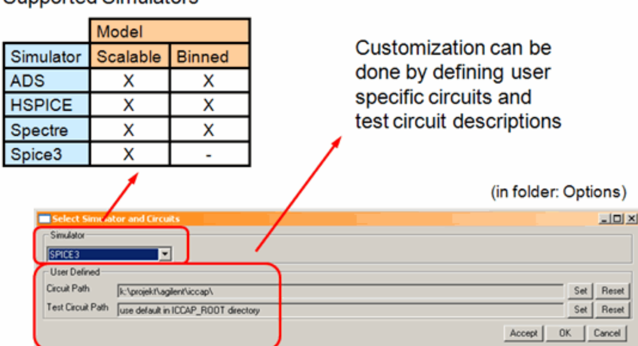


Features have been developed according to experience from AdMOS modeling service.

Define binned models

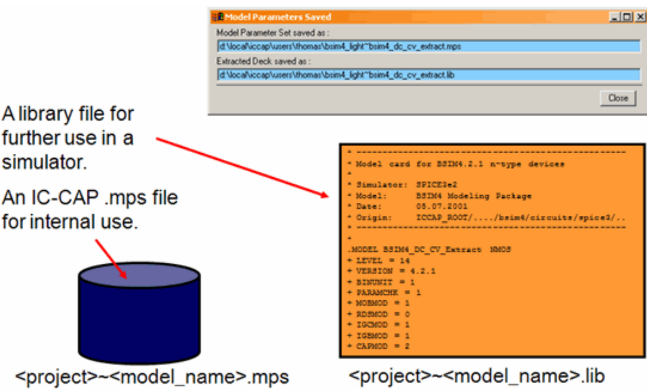


Simulator Support


Supported Simulators

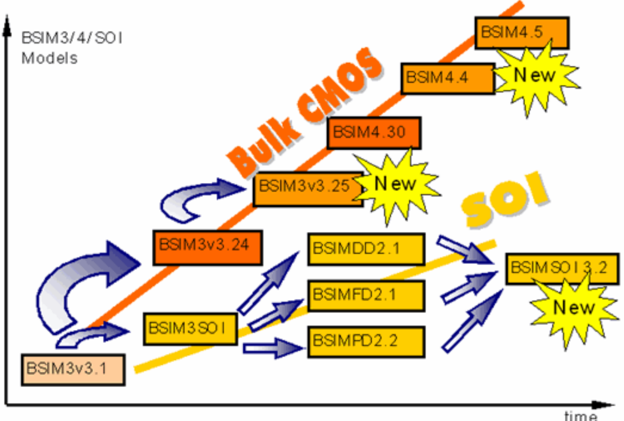
	Model	
Simulator	Scalable	Binned
ADS	X	X
HSPICE	X	X
Spectre	X	X
Spice3	X	-

The extraction module generates two types of outputs:



Outlook

Roadmap of BSIM Simulation Models



The diagram shows the roadmap of simulation models which are based on the BSIM3v3 model. Two principal branches can be identified which cover the mainstream of "conventional" bulk CMOS devices and the upcoming Silicon On Insulator (SOI) technology. All the models in both branches are based on the BSIM3v3.1 model. BSIM3v3 is a quasi standard model in the semiconductor industry. However, to cover all the physical effects of the steadily scaled MOS devices, UC Berkeley published the new BSIM4 model in March 2000.
In the second branch, the new BSIMSOI 3.0 models now unites the fully depletion and partially depletion mode of SOI transistors which was represented in the past by different models.

Outlook

- The BSIM3/4/SOI Modeling Packages will be kept up to date with the current model versions released from UC Berkeley (see anouncement in the CMC minutes).
- In a further step, AdMOS and Agilent Technologies are looking into the Next Generation MOS models (BSIM5, PSP, HiSIM2, EKV3.0), which are currently in the evaluation of the Compact Model Council. Our goal is it to provide supperior MOS modeling support beyond the 90nm CMOS node.

GaAs Transistor Models in General

Different to Silicon Transistor models, GaAs models have a relatively small number of model parameters. Another differentiator to silicon is that there have always been 'many' models for GaAs transistors, compared to the pretty clear structure for MOS (BSIM, PSP etc.) and for bipolar (Gummel-Poon, VBIC, Hicum etc.).

Also, in many cases, modeling engineers apply PI-schematic-based models for individual DC bias points, and then apply scaling vs. DC bias to obtain the general model. (See the chapter about Y-parameter modeling.)
This method is quite similar to the approach with the PSP MOS model.

Last not least, GaAs transistors exhibit in many cases a frequency trapping in the 10kHz range. Therefore, some GaAs models exhibit model parameters for modeling the DC-gm, and fully decoupled RF parameters for modeling the RF-gm.

General Transistor Modeling Strategies

The guidelines in this slide set present a general-purpose transistor modeling strategy, independent of the selected model, but focusing on modeling single transistors.

In this slide set, we consider a bipolar transistor as an example. The same strategy can also be applied to MESFETs, single MOSFETs etc.

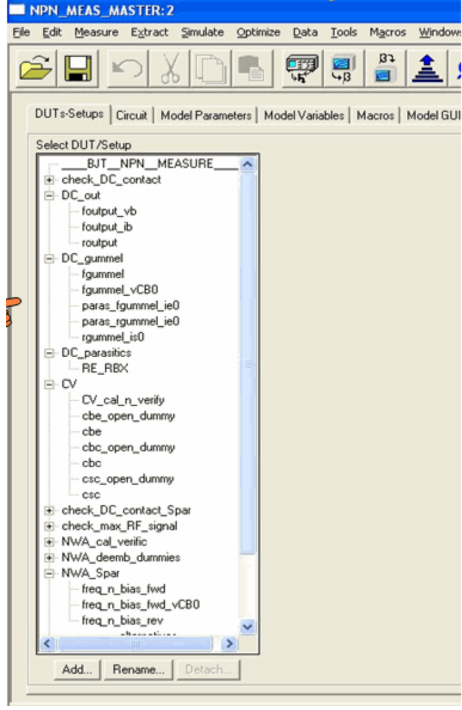
> **Note**
> The proposed sequence of measurements, data verification and final device modeling can also be applied to scaling transistor modeling like for MOS (BSIM, PSP etc). Of course, the modeling strategy for these scaling models is then different from the here proposed one.

Measurements

from the demo_features directory
demo_features/3_MEAS_ORGANIZE_n_VERIFY_DATA\0_MASTER_FILES
load for example the file
NPN_MEAS_MASTERFILE_PELdep.mdl
and perform the predefined standard measurements.
Export the measured data into mdm files (Macro EXPORT_DATA)

> **Note**
> besides the conventional transistor measurements, you find also in this MEAS_MASTERFILE Setups for the measurement of the probe contact resistances, the NWA cal.verification and the de-embedding verification.

Screenshot of NPN_MEAS_MASTERFILE_PELdep.mdl



Data Consistency Checks DC

After the measurements have been performed, it is absolutely important to check the consistency of the data:Concerning the DC data, you should check the following:
does a transfer curve, converted to a pseudo-output curve, match the original output curve?
do the sweep limits of the transfer and the output characteristics match each other, or are the transfer sweeps in another voltage range than the output sweeps?
are the DC contact resistances known?
From the demo_features directory, load the file
demo_features\3_MEAS_ORGANIZE_n_VERIFY_DATA\0_MASTER_FILES\
1__CHECK_DATA_CONSISTENCY\NPN_MEAS_CHECKTOOLS_PELdep.mdl
import the corresponding subset of the previously exported meas.data and follow the READMEs in the model file.

Data Consistency Checks S-Parameters

1. DC Bias Conditions of the S-parameters:

- Are the DC bias currents of the S-parameter measurements identical to those of the DC-only measurements?

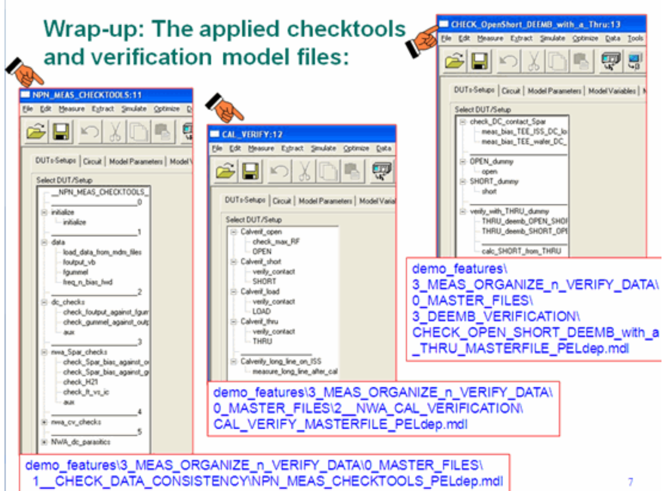
- Are the DC voltage sweep ranges of the S-parameters identical to the DC-only measurements?
- Are the DC contact resistances known?

1. Continue applying the file NPN_MEAS_CHECKTOOLS_PELdep.mdl

NWA calibration verification:
demo_features\3_MEAS_ORGANIZE_n_VERIFY_DATA\0_MASTER_FILES\2__NWA_CAL_VERIFICATION\CAL_VERIFY_MASTERFILE_PELdep.mdl

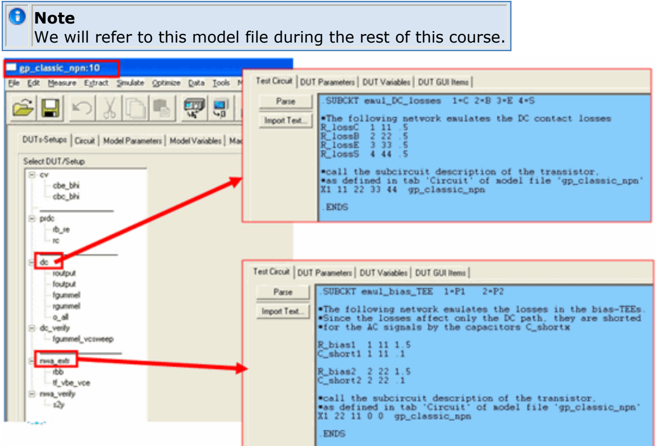
De-embedding Verification:
demo_features\3_MEAS_ORGANIZE_n_VERIFY_DATA\0_MASTER_FILES\3_DEEMB_VERIFICATION\CHECK_OPEN_SHORT_DEEMB_with_a_THRU_MASTERFILE_PELdep.mdl

Wrap-up: The applied checktools and verification model files:



Enter the DC contact resistances and also the measured ohmic losses of the NWA's S-parameter Testset into your MASTER_EXTRACTION.mdl file
for example:
demo_features\1_BASIC_MDLG_EXAMPLES\21_Gummel_Poon_demo\GP_CLASSIC_NPN_1998.mdl

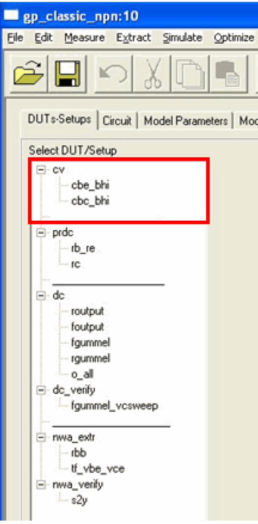
> **Note**
> We will refer to this model file during the rest of this course.



After all the data have been checked, we can finally start the modeling

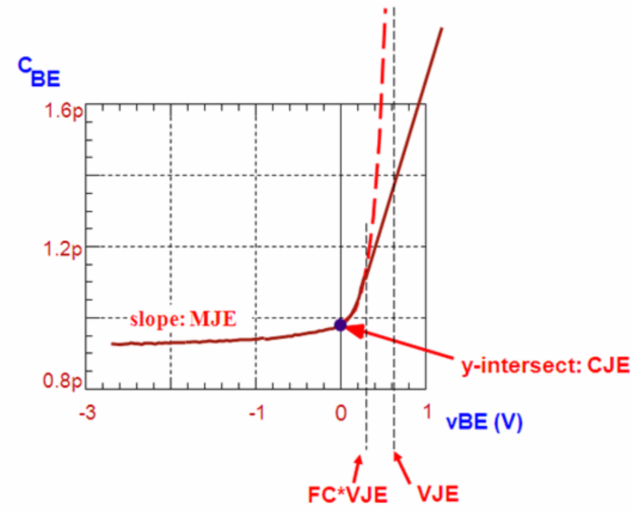
usually, the CV modeling is performed first, since in many models (especially the bipolar models except Gummel-Poon) the capacitance parameters affect the DC modeling.

Additionally, the CV model parameters are not overlaid by any other model parameter, what makes them the ideal parameter to start our modeling work.



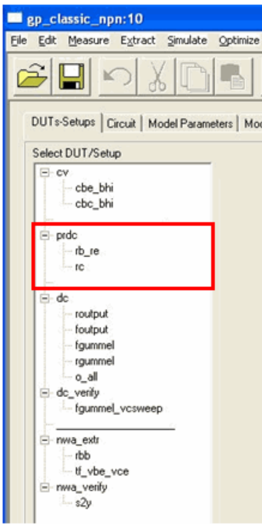
Hicum2 is an exception to all other models, here you need to start with modeling the transit time ft.

Example: CV Modeling Gummel-Poon

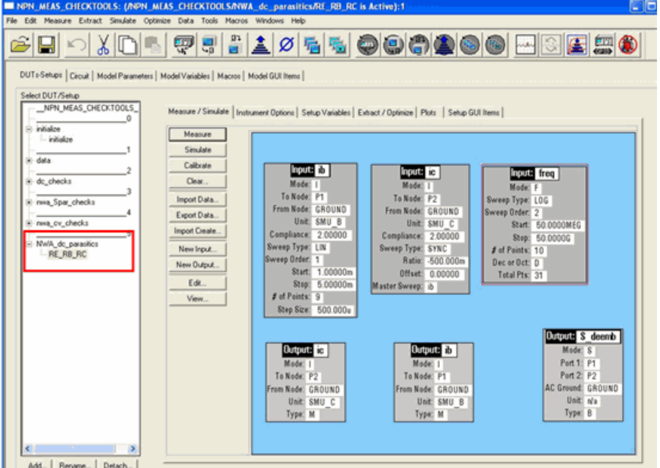


In a second modeling step, the parasitic ohmic resistances are extracted from special measurements. Either from DC measurements or from hot S-parameters (strong bias current into the Base or Gate (MÉSFETS), and half the current suck out of the Collector or Drain. However, no simulation or optimization is performed at this step, since the other DC parameters are not yet known.

> **Note**
> You can also skip this step and extract the ohmic parameters from the transfer curves (Gummel-Poon plot) and the output characteristics for low vCE or vDS.

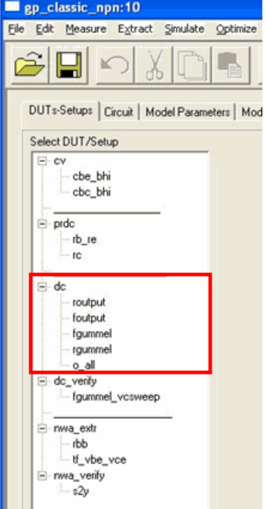


Note
Extracting the parasitic resistors and inductors from Spar measurements and the NPN_MEAS_CHECKTOOLS file.



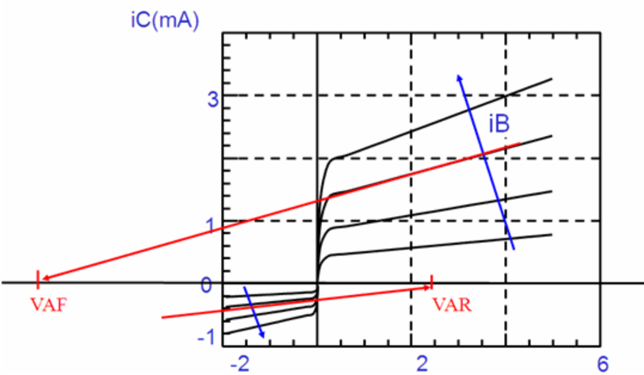
Now we are ready to perform the DC modeling. This DC modeling is the fundamental step for all transistor modeling, since the DC curves determine with 100% the starting point of the S-parameters at lowest frequency (as will be seen later). Therefore, an as good as possible DC modeling fit is mandatory for a good model.

Note
For bipolar models, we first need to fit the slopes of the foutput and routput curves (Early voltage parameters), but still ignore the fitting of the curves themselves. This is necessary since the Early voltage parameter affects the Gummel-Poon plot iC(vBE). The final curve fitting will be possible after the transfer curves (Gummel-Poon plot) are fitted. For all other types of transistors, we usually skip this extra step and commence directly with the Transfer Curve modeling.

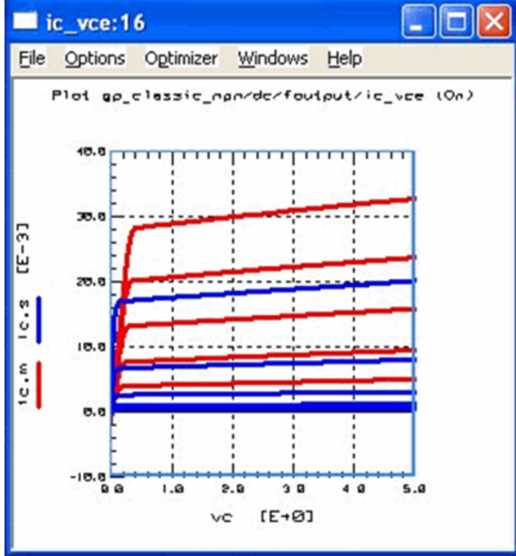


The bipolar intermediate step: fitting the Early Voltage parameter for foutput and routput,

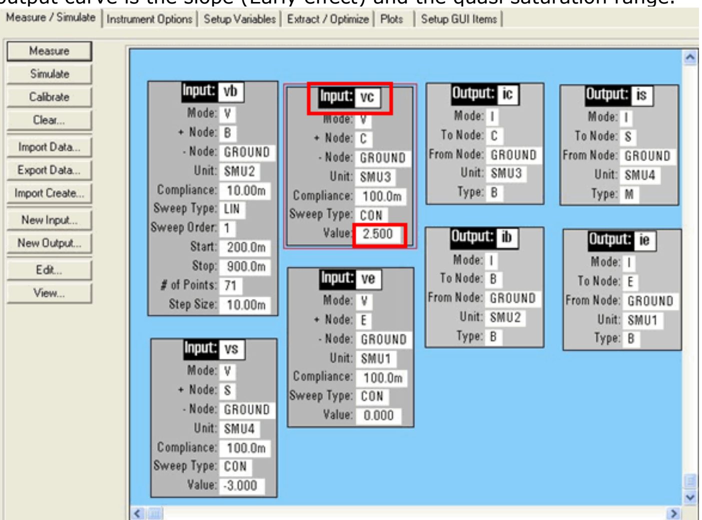
but ignoring the matching of the curves. Only the slopes are considered for the moment.



This is the intermediate result of the Early modeling for bipolar transistors. As you can see, only the slopes are fitted (Early Voltage Parameters), but the curves themselves do not fit yet. As mentioned, for all other other transistor types, this step is skipped.



We are now back with a common, transistor-type independent modeling strategy. Usually, we bias a transfer curve Setup such that the Drain or Collector voltage is half the max. value of the output characteristics. This means that once we have fitted the transfer curve (Gummel-Poon curve for bipolar), we have automatically also fitted the output characteristics at that selected Drain/Collector voltage. All what remains is to adjust in the output curve is the slope (Early effect) and the quasi saturation range.



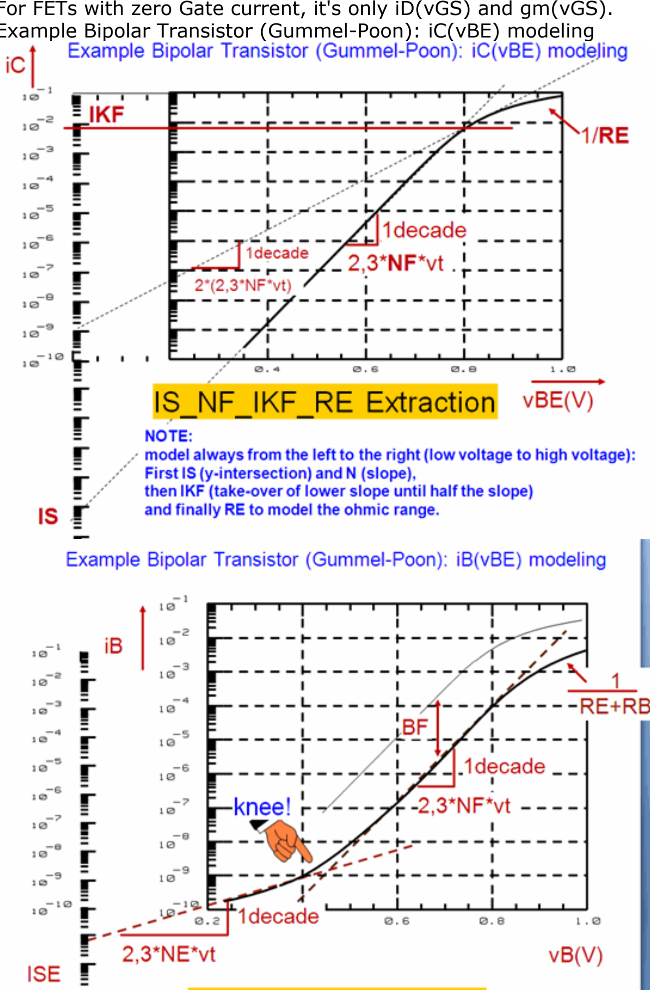
Transfer Curve(s) Modeling

For bipolar transistors, modeling the transfer curve (Gummel-Poon Plot) means to model the input characteristics iB(vBE), and the two transfer characteristics iC(vBE) and beta(vBE).

For FETs with non-zero Gate current (e.g. Junction-FETs etc, we model the input characteristics iG(vGS) with vDS=0

and the transfer characteristics iD(vGS) and gm(vGS).

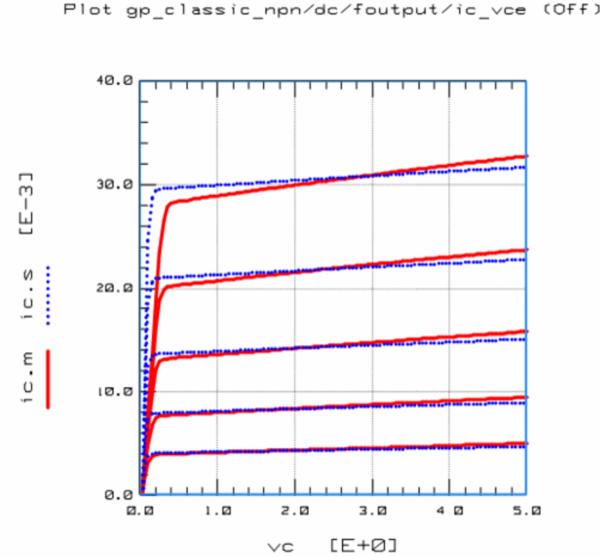
For FETs with zero Gate current, it's only iD(vGS) and gm(vGS).
Example Bipolar Transistor (Gummel-Poon): iC(vBE) modeling



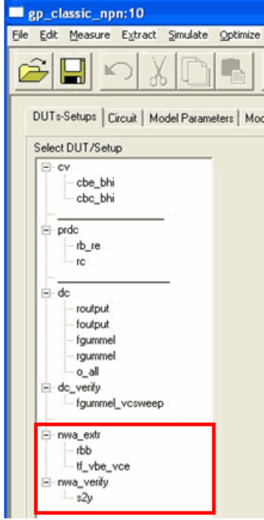
Output Characteristics Modeling

After the transfer characteristics has been modeled, we finally come to the output characteristics modeling. Since you can think of the transfer characteristics as a cut thru the output characteristics for a given vCE or vDS, the fitting we have applied so far to the transfer characteristics must match the output curves at exactly that vCE or vDS.

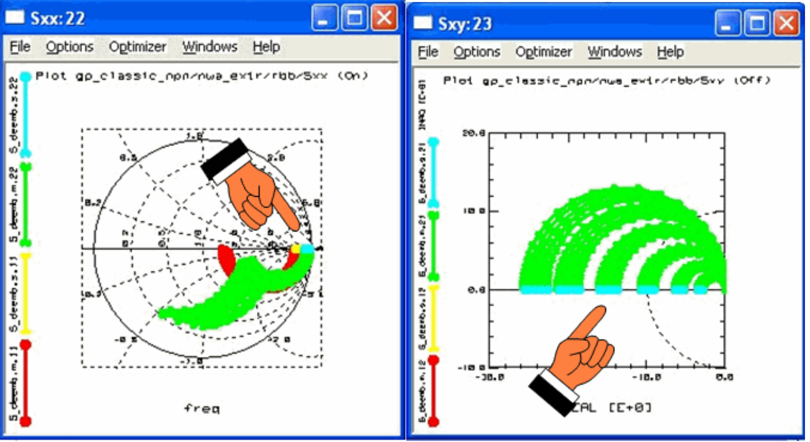
With the Transfer Curve(s) fitted, the Output curves match perfectly at the Collector/Drain voltage used for the transfer curves.
Typically, only a fine-tuning of the Early voltage and the Collector (Drain) resistance is required for obtaining the final fit.

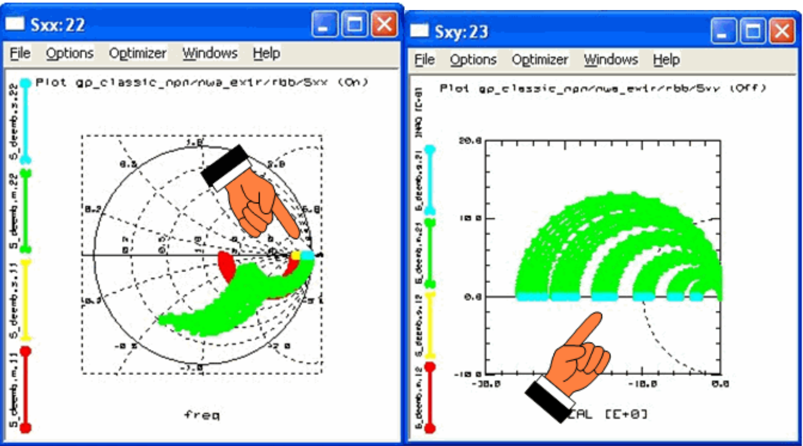


S-parameter Fitting



We are now ready to commence with the S-parameter modeling
In the first step, we switch off all extracted or non-zero capacitances (i.e. we divide all extracted capacitance parameters CJx by 1E6, what means we do not loose their extracted values, but only switch-off their effect).
The same with transit time parameters.

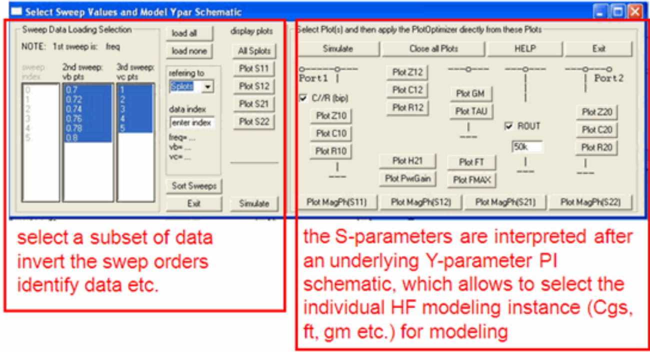




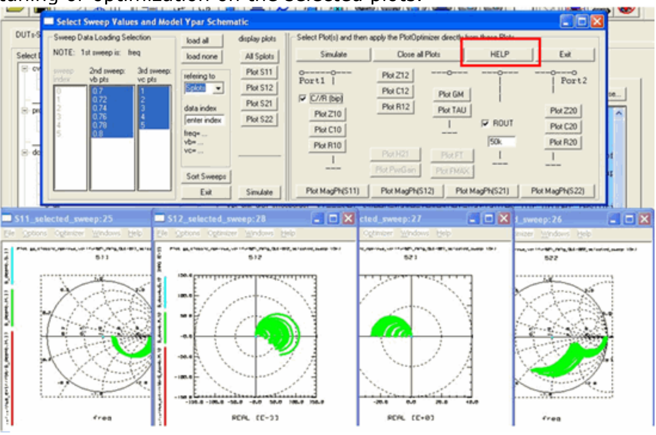
If the data consistency check was o.k., if the de-embedding was perfect and if the DC modeling was perfect too, then the simulated S-parameters should all be located on the x-axis, and exactly at the extrapolations (freq -> 0Hz) of the measured S-parameters (whicih are now representing an ideal, frequency-independent transistor).

If this is not the case, then consider modeling the RB parameters of the bipolar transistor (cannot be modeled correctly in DC, only in S-parameters), consider the bias TEE losses, consider the DC contact losses etc.

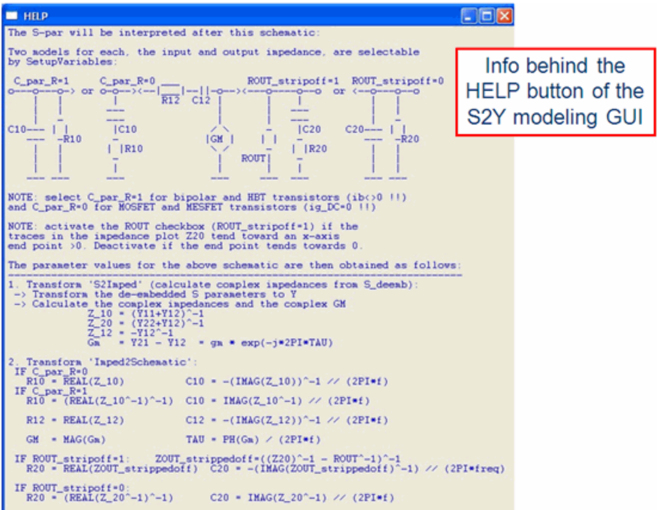
In any case, consider a DC parameter.
Finally, switch on again the capacitance and the transit time parameters and copy from the DEPOTS.mdl file the GUI. CopyPaste_ITEMS/S2Y_Mdlg_GUI into your MASTER_EXTRACTION model file.



Apply the GUI for separating individual curves (CV curves, ft curves etc) out of the S-parameter. From the up-popping plots, select PlotOptimizer and perform a parameter tuning or optimization on the selected plots.



Note
Hit the HELP button in the GUI for detailed explanations and help.



Conclusions

The presented general transistor modeling strategy should assist you in finding the ‚golden modeling path' for your special selected transistor model.

Overview of the Hicum Level 2 model

Forewords

Compact models are only a simplified representation of real life:
There's always a bias, frequency, temperature or geometry range for which a given model is not accurate enough!

One should keep this in mind when doing modeling in order to achieve the best possible compromise.

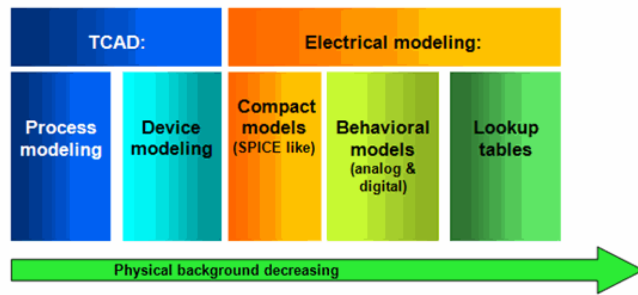
The best compromise should match the following requirements:

- describe the complete process behavior (not just dumb fitting)
- avoid unphysical behavior
- focus on bias, frequency, temperature and geometry ranges that are relevant to circuit design

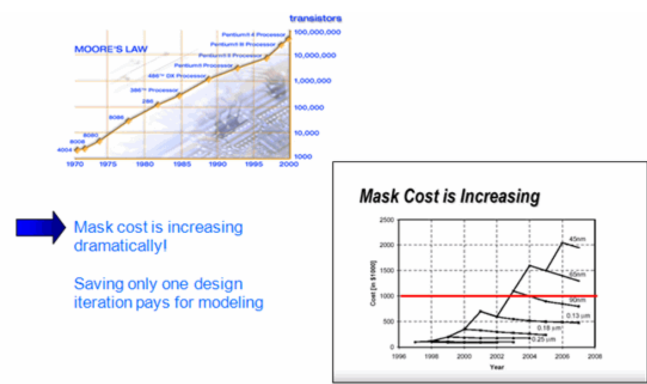
What is compact modeling?

Many significations behind the word modeling:

> **Note**
> Behavioral languages are more and more used for compact models definition (see CMC web site)



Need for accurate models



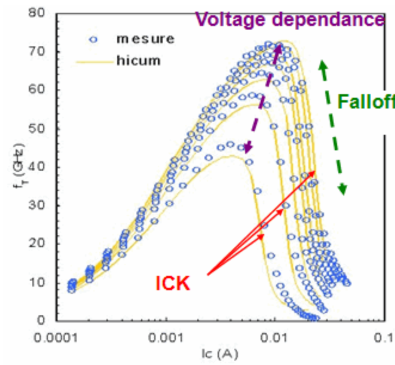
HiCUM can help circuit design

Better RF modeling:

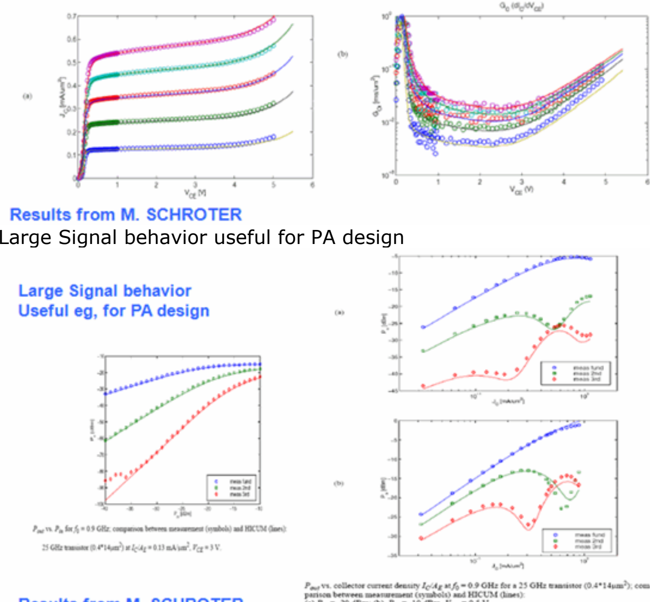
1. Better ICK modeling
2. Better description of abrupt fT falloff
3. Better maximum fT voltage dependence modeling

This is important for many applications:

- Selecting Circuit bias
- for PAs where current can go beyond peak fT

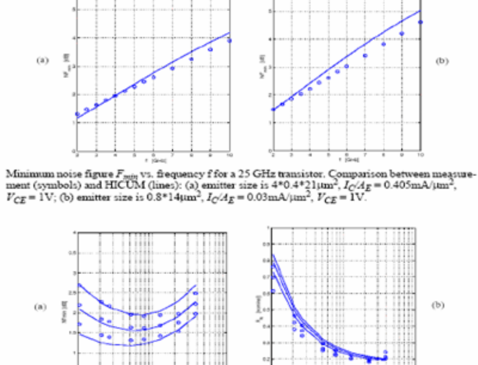


Output conductance modeling (good derivatives are required for large signal simulations)



Large Signal behavior useful for PA design

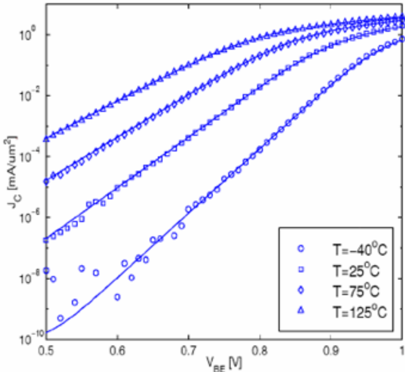
LNA design -> associated to a scalable model, designers can choose optimal transistor



Minimum noise figure F_{min} vs. frequency f for a 25 GHz transistor. Comparison between measurement (symbols) and HICUM (lines): (a) emitter size is 4*0.4*21µm², I_C/A_E = 0.405mA/µm², V_{CE} = 1V; (b) emitter size is 0.8*14µm², I_C/A_E = 0.03mA/µm², V_{CE} = 1V.

Comparison between measurement (symbols) and HICUM (lines) for a 25 GHz transistor (4*0.4*21µm²): (a) Minimum noise figure F_{min} vs. collector current density I_C/A_E; (b) equivalent noise resistance R_n vs. collector current density. f/GHz = 1, 2, 3; V_{CE} = 1V.

Bandgap references -> Very difficult to obtain correct trends. Not only requires "fitting" but physical results: non physical XTI gives wrong results!



HiCUM is also better for Scaling and Statistics

The model physics based definition allows a better geometry scaling

- Note that there are no scaling equations in the model!
- But it's easier to describe model parameters variations with geometry
- Scaling is usually achieved via an external program (able to deliver static libraries) or using simulator macro language

In the same way, it's easier to describe statistical variations with HiCUM

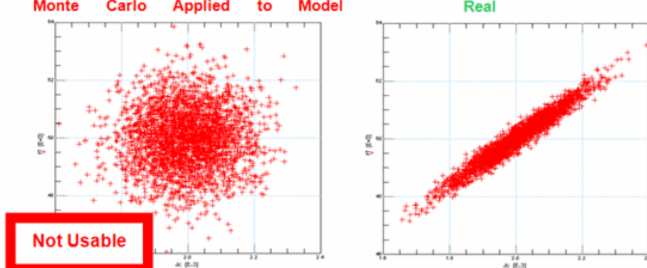
- Corner models (be careful of definition)
- True Monte-Carlo (process)
- Matching

Additionally HiCUM can be used to realize predictive models

- Obtain realistic trends for next process generation
- Start designs early then switch to finalized models close to tape out

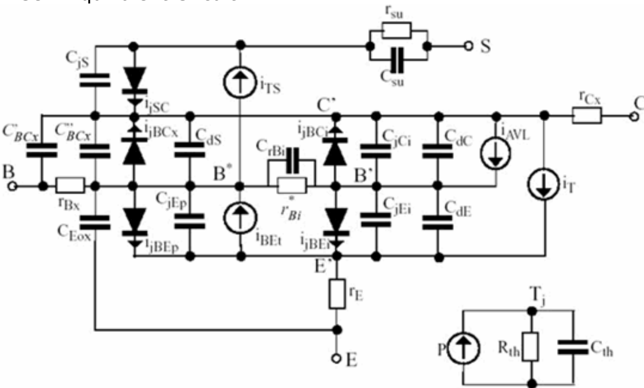
Statistics can help designers

Designers are sometimes skeptical about Monte-Carlo Simulations: they are right!
If one uses model parameters directly as Monte-Carlo simulation input : correlation is lost (left figure)
HiCUM help delivering TRUE statistical models which give the right trend (on the right)



HiCUM Core Model

HiCUM Equivalent Circuit



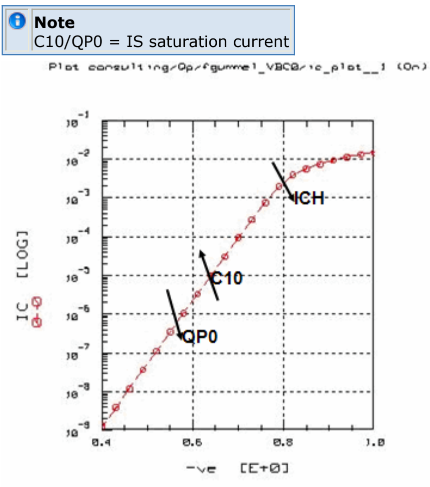
Transfer current

Infinite ICH leads to no effect,

Reducing ICH increases IC at high current densities (only for 2D and 3D effects)

The curve depends also on Capacitances, weighting factors and transit time

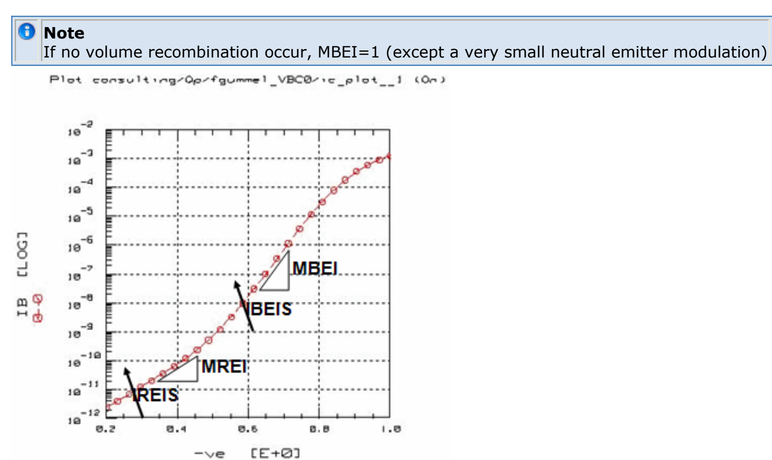
Note
C10/QP0 = IS saturation current



Base Current

BE base current
No Beta unlike SGPM. Identical behaviour for internal and peripheral currents

Note
If no volume recombination occur, MBEI=1 (except a very small neutral emitter modulation)

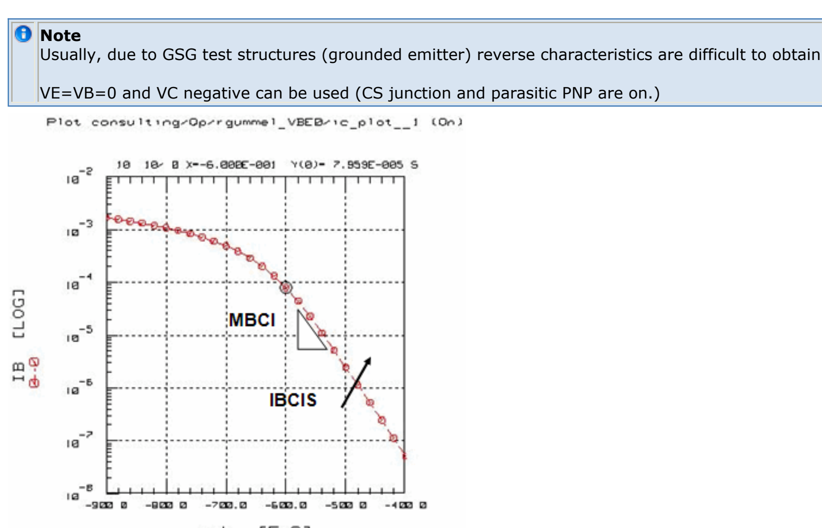


BC base current
No Reverse Beta unlike SGPM

Identical behavior for internal and external current

Note
Usually, due to GSG test structures (grounded emitter) reverse characteristics are difficult to obtain

VE=VB=0 and VC negative can be used (CS junction and parasitic PNP are on.)



Depletion charges and capacitances

The well known capacitance formulation is:

$$C_j = \frac{C_{j0}}{\left(1 - \frac{V}{Vd}\right)^z}$$

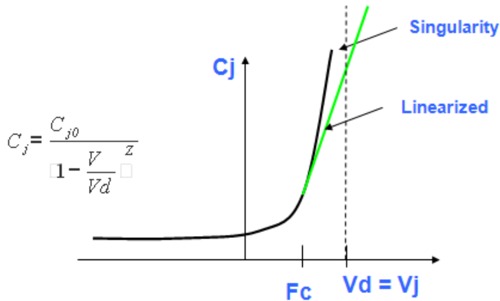
With 1/3 < z < 1/2

With C_{j0} the zero bias capacitance, Vd the junction potential and z the grading coefficient.

This formulation is not satisfying from a numerical point of view:
In SGPM the function is made linear above FC
The capacitance is still increasing, but negligible compared to the diffusion capacitance

In Hicum, the capacitance value at forward bias plays a role:
Used to compute qjEi and qjCi used in the GICCR (Early effect)

A different behaviour has to be used !

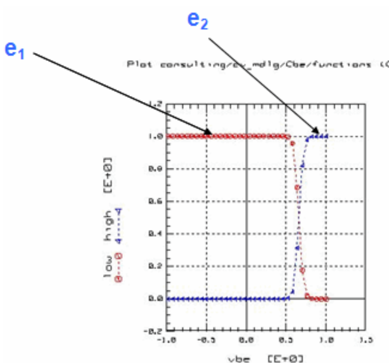


BE Capacitance

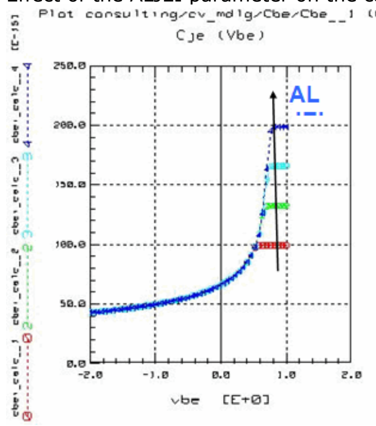
The BE capacitance formulation is based on weighting functions e1 and e2:

$C_{TOTAL} = e_1\ C_J + e_2\ A_{LJEI}\ C_{J0}$

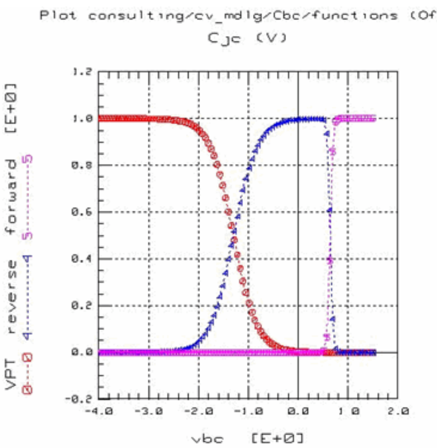
The capacitance is split in two parts using e_1 and e_2



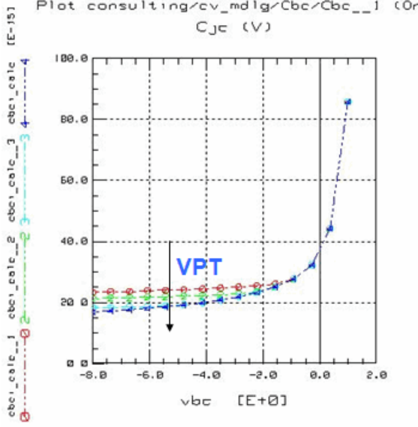
Effect of the ALJEI parameter on the capacitance:



3 Weighting functions:
The capacitance is divided in 3 parts (High reverse bias, reverse and low forward bias, high forward bias) With weighting functions: The punch through function (large reverse) depends on parameter VPTCI. Here VPTCI = 2 V.



Effect of the VPTCI parameter on the capacitance:
This effect is more pronounced for devices with low collector doping profiles (No SIC). Plot of CjCi for VPT= 30, 6, 3 and 2 V.



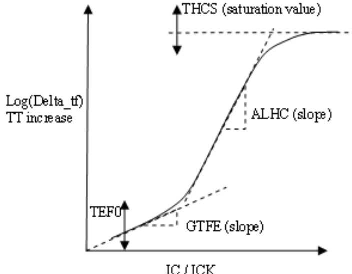
Transit Time

The transit frequency fT is a consequence of stored charges which are described via the transit time:

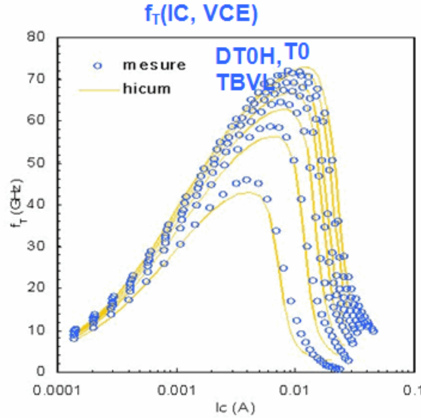
$$\tau_F = \frac{1}{2\pi f_T} - (C_{JE} + C_{JC})\frac{V_T}{I_C} - (RCX + RE)C_{JC}$$

The transit time model in HiCUM is separated into 3 parts:

- At low current density: the transit time (TF0) has a voltage dependence (VBC):
- At medium and high current densities: the critical current ICK has a voltage dependence (VCE)
- At high current densities: the transit time increase has a current dependence (IC) and a voltage dependence via ICK.
  Below is presented the transit time increase Delta_tf versus collector current (normalized with respect to the critical current ICK) and the influence of the HiCUM model parameters THCS and ALHC on the curve.

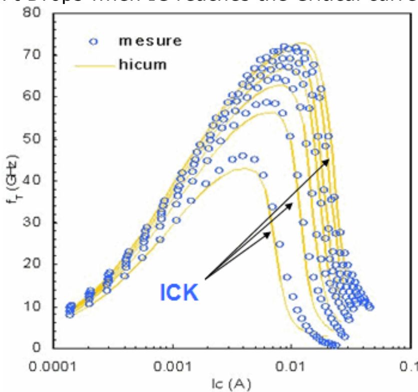


Transit Frequency



Transit Frequency and Critical current

Ft Drops when IC reaches the Critical current ICK.



fT(IC, VCE)

Critical current ICK

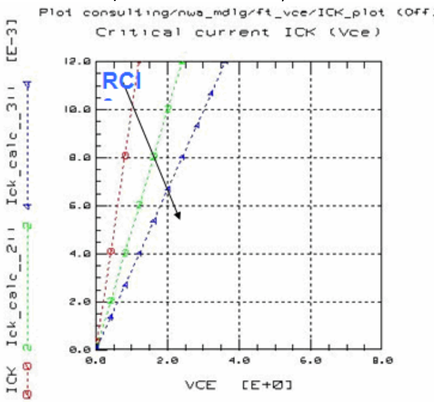
ICK versus VCE plot:
VPT= 1000
VLIM = 100

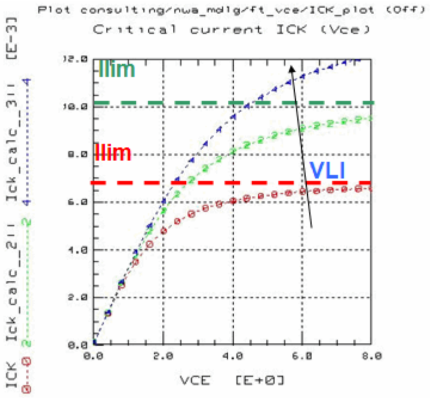
VPT and VLIM have no effect

The slope at VCE = 0 is 1/RCI0

Plots for RCI0=100, 200 and 300 Ohms
Critical current ICK
ICK versus VCE plot:
VPT= 1000
RCI0 = 300
VLIM = 2, 3 and 4 V

The saturation value is:
Ilim= VLIM / RCI0 =6.66 mA, 10 mA and 13.33 mA

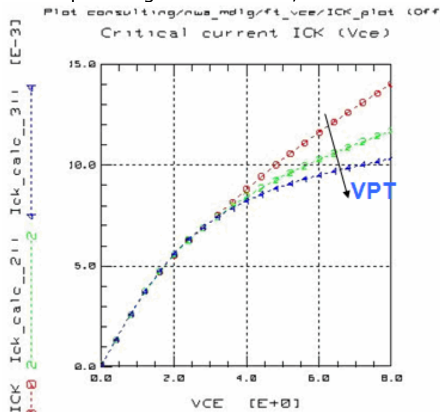




ICK versus VCE plot:
RCI0 = 300
VLIM = 3 V

VPT= 10, 20 and 50 V

The slope at high VCE is: Ilim / VPT



Scalable Modeling

Scalable models are mandatory for Bipolar devices to:

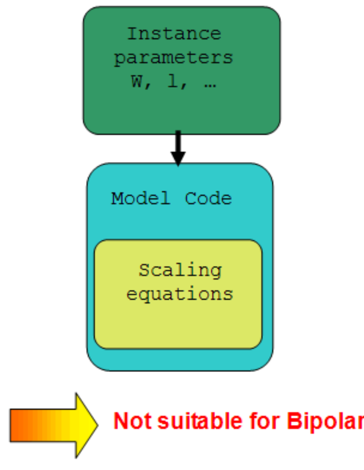
- Reduce parameter extraction and measurement effort
  - Work on reduced transistor sets
- Reduce test chip area
- Provide more transistor configurations to designers
  - Allow modeling of non measured devices
- Help designers to optimize circuits wrt device sizes
  - Scaling equations must be continuous and differentiable
- Improve statistical modeling
  - Model geometry induced statistical effects

> **Note**
> Scaling methodologies also help to understand and capture physical effects and their geometry dependence which could simply not be done with single transistor approach

Method #1: MOS like, eg bsim3

Scaling equation included in model code.

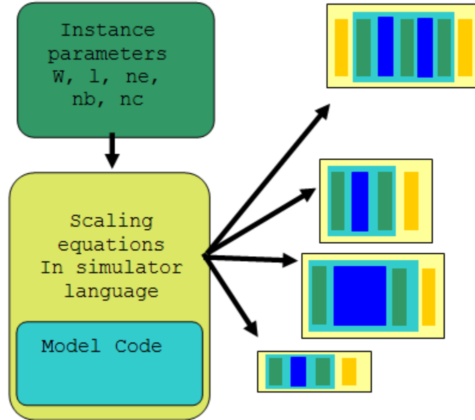
- In Bipolar scaling equations are too complex: would need a new model every year
- Models need to apply to a wide variety of bipolar processes plethora of scaling equations in model code
- Model code never stable for most advanced technologies
- Burden for model developers, less support, less research
- Pressure quick fix of deficiencies with fitting parameters = bad model for a long time



Method #2: SMACH Extraction Toolkit

HiCUM with simulator macro language
Scaling equation included in macro language

- Scaling equations and model core decoupled
- Very flexible solution
- Complex equations difficult to handle with simulator macro language
- Need efficient tracking of scaling equation and rigorous update of related extraction methods
- No scaling rules IP protection

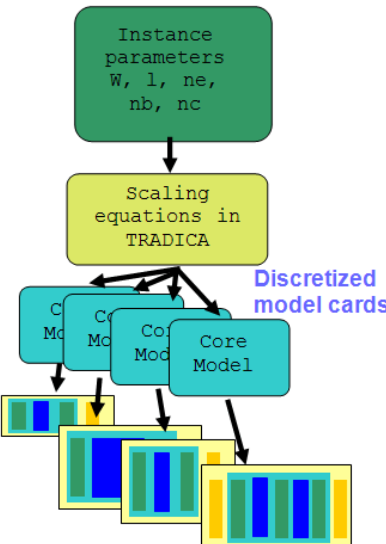


Solution used in SMACH: Very efficient when flexibility is the priority

Method #3: TRADICA /HMT EXTRACTION TOOLKIT

TRADICA approach
Scaling equation in an external program

- Advanced scaling equations can be used
- Model library has very simple syntax (simulator support)
- Scaling equations and model core decoupled
- Models can also be produced dynamically
- Scaling rules IP protected



Method #2 and #3 can be used to apply "external" scaling to HiCUM, but no scaling equations are in the model code ...

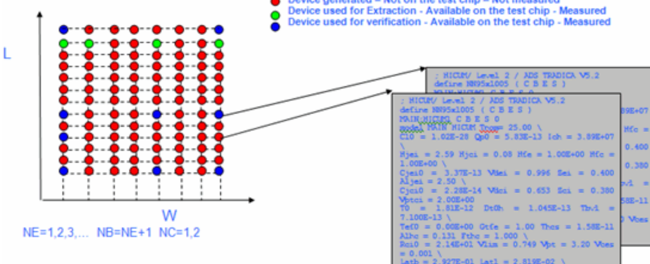
So, why HiCUM is considered as a SCALABLE MODEL?

- Every HiCUM model parameter can be expressed wrt physics based quantities and geometrical quantities relying on physics based approach
- Note this is the key of successful modelling activities. This is also the most powerful path to capture proper statistical variations

The ultimate goal is to model the process, not a single die!

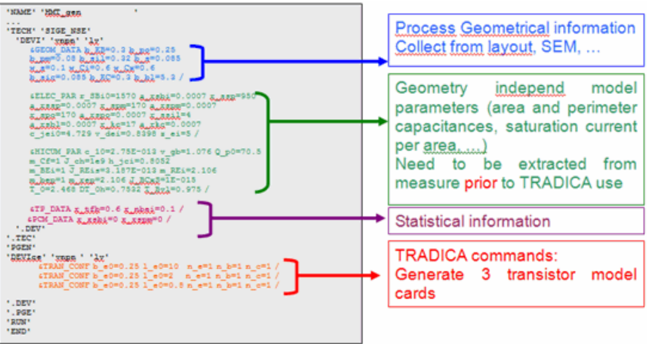
Scalable Modeling: the TRADICA/HMT example

TRADICA produces scalable model parameters (according to scaling rules and geometry independent model parameters)
Parameters are mapped in single device model cards for any configuration: width, length, number of emitter stripes, number of base contacts, collector contacts.



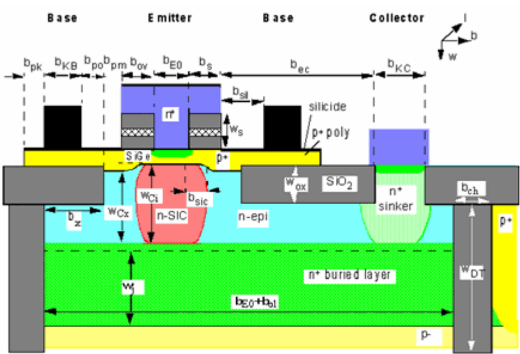
- Can be done statically or dynamically
- Geometry independent parameters are extracted from a subset
- Virtually W and L continuous models

TRADICA Input File Example:

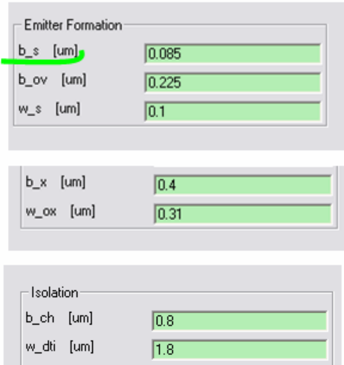


Process and geometry data collection

1st step: determine geometry parameters from SEM cross section



HMT Window manual entry

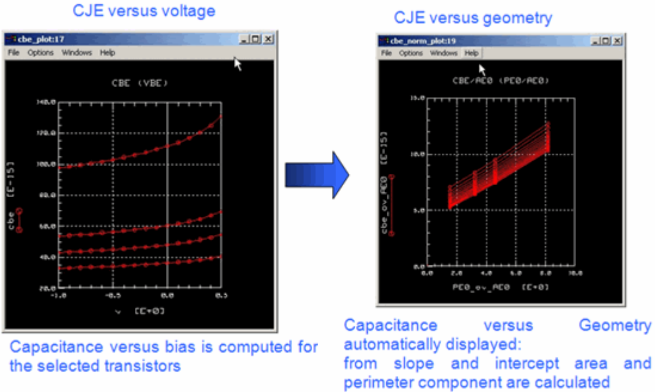


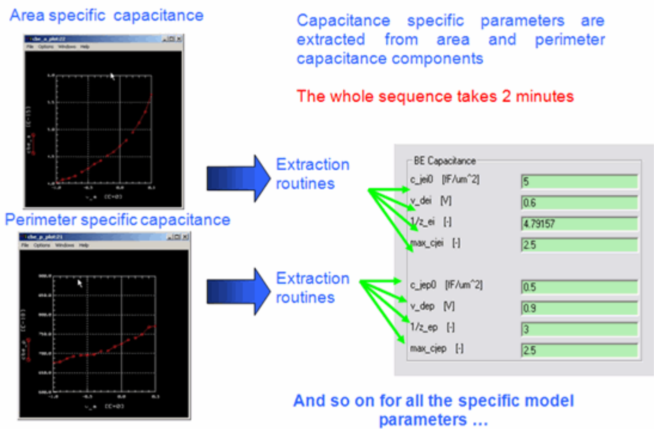
Process specific parameters extraction using HMT/TRADICA:

- Select available devices
- Import data (automatic procedure)
- Navigate through data if necessary
- Run extraction
- Go to the next extraction step

> **Note**
> Process specific parameters should be extracted from multi-geometry data in order to preserve consistency,
> one should avoid extracting model parameters on individual transistors and to fit scaling rules to model parameters
> The latter method (although proposed as standard method for PSP mos model) often provides usually poor or/and unphysical results.

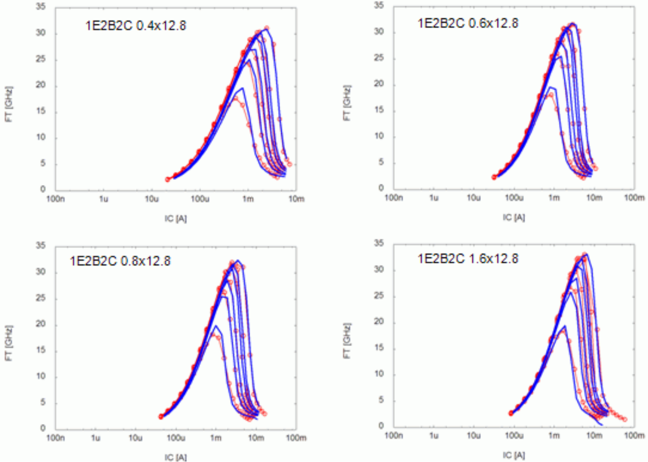
Extraction Example: Junction Capacitance





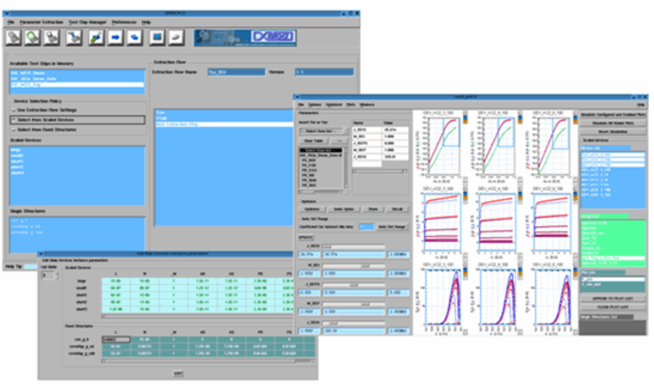
Example results: f_T (30GHz Process)

Comparable results are routinely obtained for processes up to 200GHz.



Statistical Modeling: The SMACH Example

SMACH User Interface



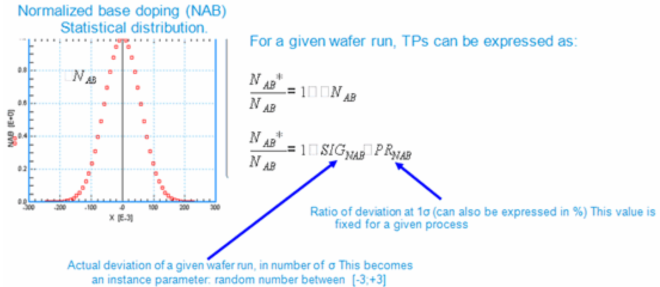
Strategy for Statistical Modeling

Which parameters should be used as independent input variables?
In SMACH Technological Parameters (TPs) are used:

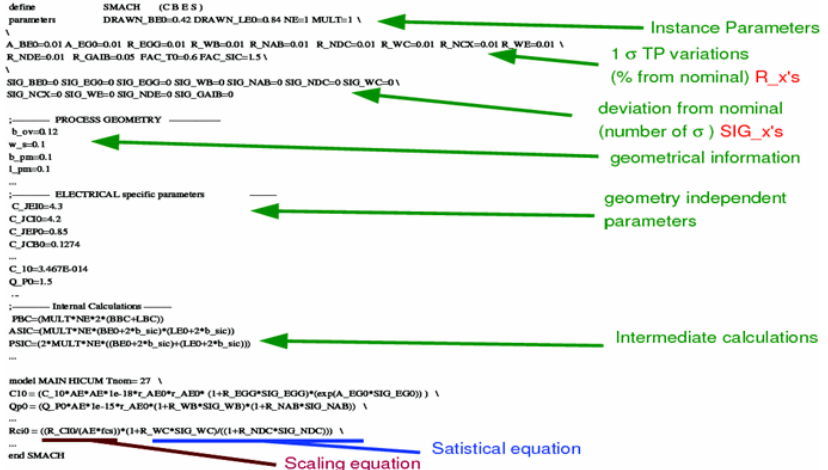
- TPs are Base doping, Emitter width, Base thickness, Collector doping, etc.
- One can find relationships between model parameters and TP
- TPs are uncorrelated
- Reduces measurement and extraction effort
- But TPs statistical variations are typically unknown!
  SMACH allows determining TPs (unknown) statistical variations from (known) PCM measured variations
  In order to do this, PCM are re-simulated using the scalable and statistical model
  Optimization is used: target is RMS error between measured and simulated PCMs, parameters are TPs statistical variations

This is done automatically within the SMACH statistical PLUGIN

Definition



Example of (ADS based) HiCUM scalable & statistical macro model



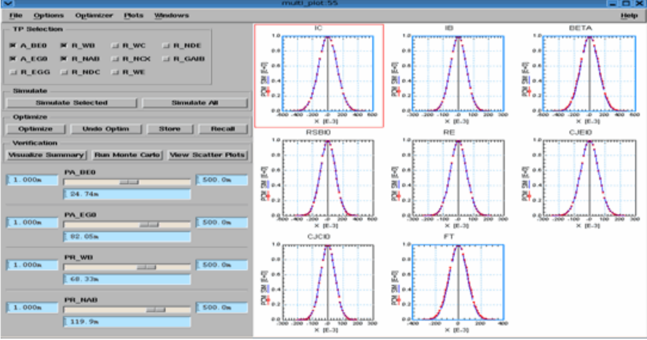
Statistical Modeling Procedure

Need to define a list of used PCMs:

- bias conditions
- Transistor size
- Get back a proposed list of TPs related to these PCMs

Based on the previous step a list of PCMs Gaussian distributions appear

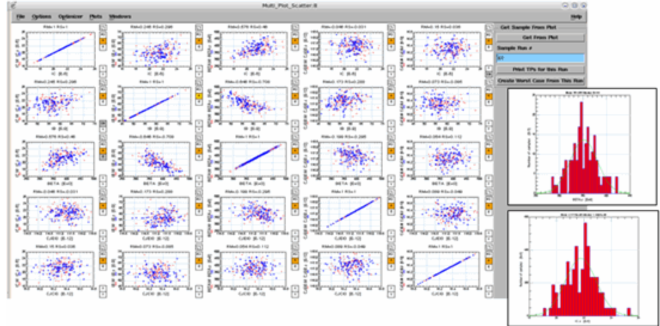
- Measured and Simulated PCMs are compared in these plots
- TPs deviations are the variables to be optimized
- Once finished, TPs are known!
- One can run full Monte-Carlo analysis to check correlation between PCMs



Check Correlation between PCMs

The following type of multi-plot is automatically generated:

- check correlation between PCMs
- compare measured and re-simulated PCMs based on full Monte-Carlo simulation (here 100 samples but can go up to several thousands)



Statistical Modeling Procedure

The Statistical modeling procedure has the following advantages:

- It's fast: the whole procedure takes about 30 minutes
- It's reliable: Measured PCMs are "guaranted" to be reproduced by the model
- it's physics based: model parameters and Technological Parameters (TPs) are linked via physics based equations
- Only PCM measurements are required (usually available since routinely done all along the foundry process manufacturing)
- The procedure is automatically adapted depending on which PCMs are available
- Since the model is a scalable model, Statistical variations are also "natively" geometry dependent

Learning the Angelov Model

The following sequence of plots gives an overview about the influence of the model parameters and their fitting properties. It represents also a possible sequence for the model parameter extraction.

IC-CAP File: Learning_the_Angelov_Model.mdl

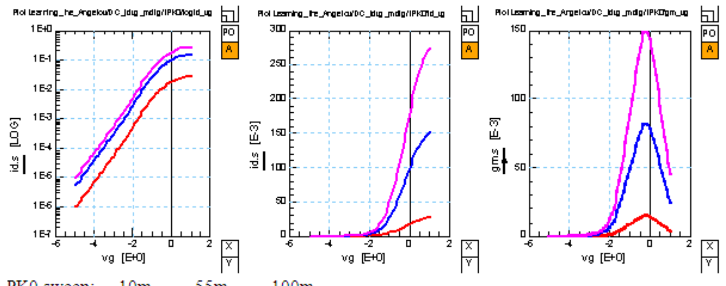
The default parameter values for the plots below are:

Idsmod = 0 Igmod = 0 Capmod = 1 Selft = 1 Noimod = 1
Ipk0 = 0.1
Vpks = -0.2
P1 = 2.0
P2 = 0.0
P3 = 0.0
Alphar = 0.1
Alphas = 1.5
Lambda = 0.01
Lambda1 = 0.0
The parameter values specified above are the recommended 'starting values' of I.Angelov

Bias conditions:

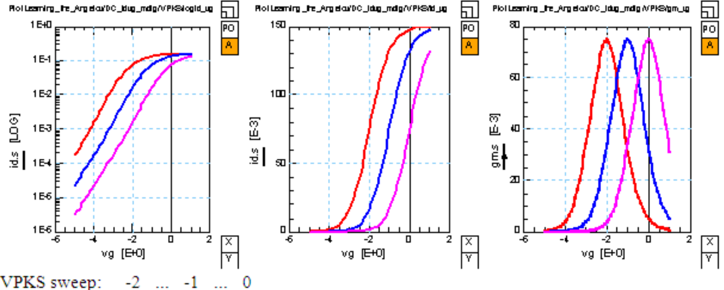
vGS: -5V .. 1V
vDS: 0V .. 10V

IPK0 Current for maximum transconductance



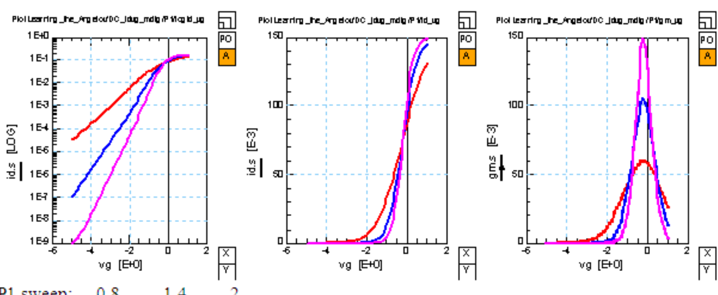
PK0 sweep: 10m ... 55m ... 100m

VPKS Gate voltage for maximum transconductance



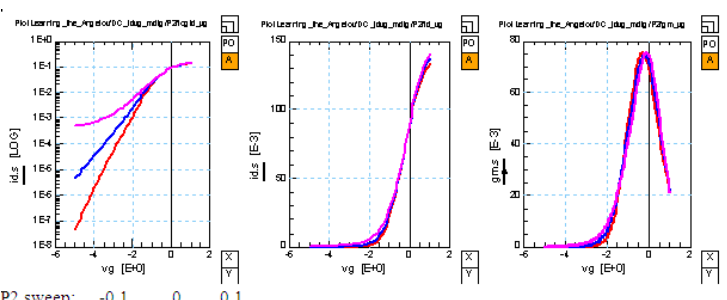
VPKS sweep: -2 ... -1 ... 0

P1 Polynomial coefficient for channel current



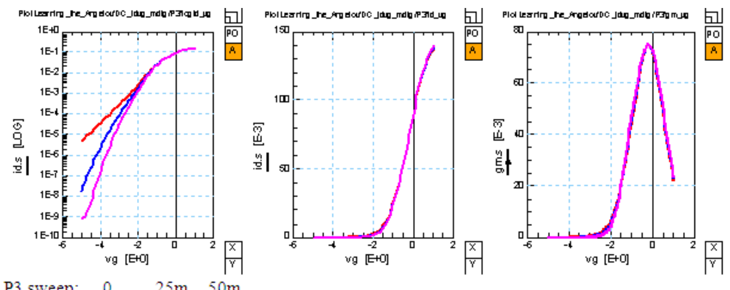
P1 sweep: 0.8 ... 1.4 ... 2

P2 Polynomial coefficient for channel current



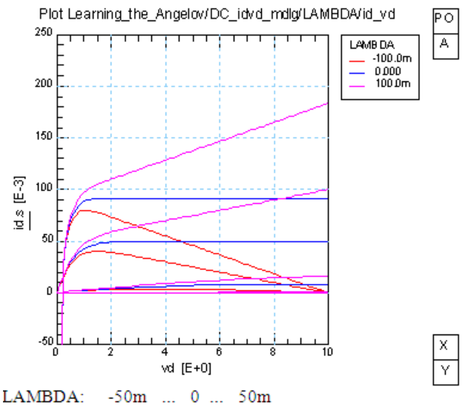
P2 sweep: -0.1 ... 0 ... 0.1

P3 Polynomial coefficient for channel current



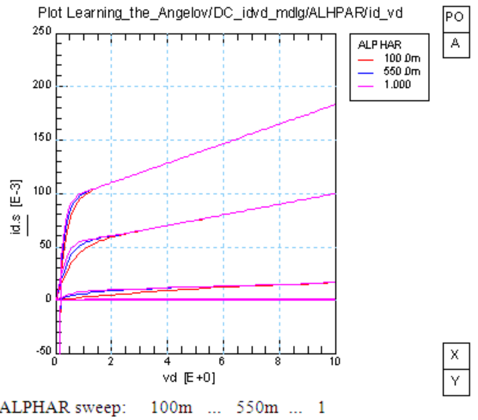
P3 sweep: 0 ... 25m. 50m

LAMBDA Channel length modulation parameter



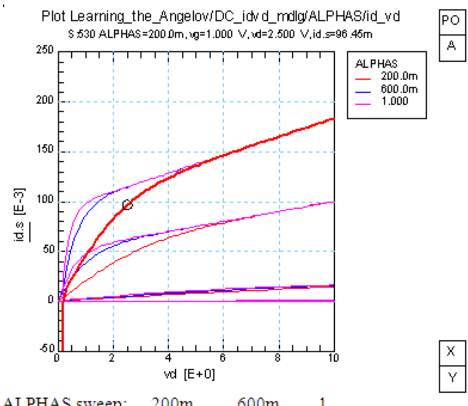
LAMBDA: -50m ... 0 ... 50m

ALPHAR Saturation parameter



ALPHAR sweep: 100m ... 550m ... 1

ALPHAS Saturation parameter

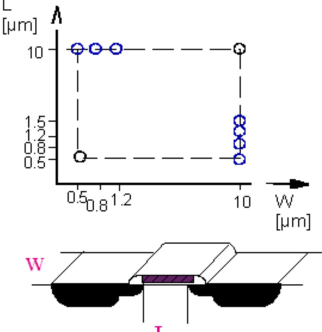


ALPHAS sweep: 200m ... 600m ... 1

After all, optimize the above parameters altogether, on idvg and idvd simultaneously, and include also the ohmic resistor parameters RG, RD and RS in the optimization.

UCB MOS Level-3 MODEL

The main application of MOS transistors is in integrated circuits and digital circuits. For this purpose, they are used within the same circuit with different layout sizes. Therefore, they have to be modeled including their geometry information, what is essentially the Gate length and width. This means, MOS models and also the MOS3 model are scalable and contain always the model parameters length L and width W.



Cross-section of a MOS transistor and the different geometry devices needed for modeling

For the parameter modeling, these measurements have to be performed:

- transfer (transconductance) curve iD(vG,vB) for a constant and low value of vD,
- output characteristics iD(vD, vG)
- and for BSIM3 also the output characteristics RD(vD,vG)

This output characteristics is of special importance for MOS modeling in analog applications. Hereby, the output resistance of a transistor stage is degrading the input characteristics of the following one by composing a lowpass filter consisting of Rout and Cin.

For detailed information about the model equations, see /Vladimirescou/ . Table 1 puts the most important MOS3 model parameters together.

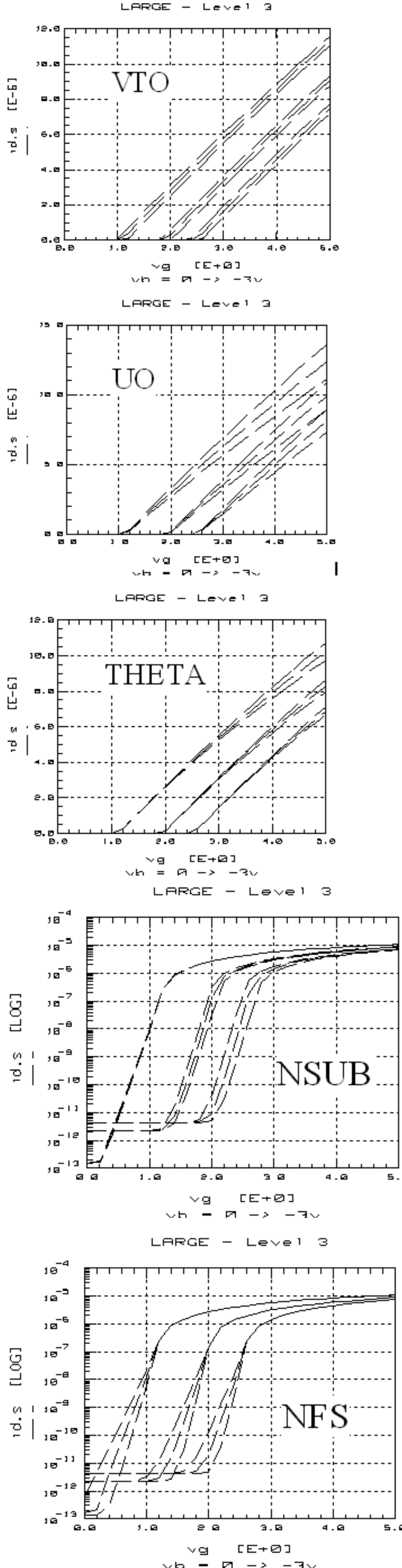
Parameter	Description	Unit	Default	Example
process parameters				
TOX	oxide thickness	meter	100n	100n
extracted from the linear region				
VTO	extrapolated zero bias threshold voltage	V	0	1
NSUB	substrate doping concentration	1/cm³	0	4.00E+15
UO	surface mobility at low gate voltage	cm²/Vs	600	700
THETA	mobility reduction parameter	1/V	0	0.1
extracted from the subthreshold region				
NFS	effective fast surface state density	1/cm²	0	1.00E+10
ETA	static feedback	-	0	1
extracted from the saturation region				
VMAX	maximum drift velocity of carriers	meter/s	0	5.00E+04
KAPPA	saturation field factor	-	0.2	0.5
short channel effect				
XJ	metalurgical junction depth	meter	0	1u
LD	lateral diffusion coefficient (eff.channel length)	meter	0	0.8u
narrow channel effect				
DELTA	width effect on threshold voltage	-	0	1
WD	channel width reduction (eff.channel width)	meter	0	0
parasitic elements				
RS	source ohmic resistance	Ohm	0	1
RD	drain ohmic resistance	Ohm	0	1
gate capacitance				
CGDO	G-D overlap capacitance per meter channel width	F/meter	0	40p
CGSO	G-S overlap capacitance per meter channel width	F/meter	0	40p
junction capacitances				
large area, small perimeter (a square test device)				
CJ	zero-bias bulk capac. per square meter of junct.area	F/m²	0	2.00E-04
MJ	bulk junction grading coefficient	-	0.5	0.5
PB	bulk junction potential	V	0.8	0.87
large perimeter, small area (a meander test device)				
CJSW	zero-bias sidewall capac. per meter of perimeter	F/m	0	1n
MJSW	perimeter capacitance grading coefficient	-	0.33	0.33

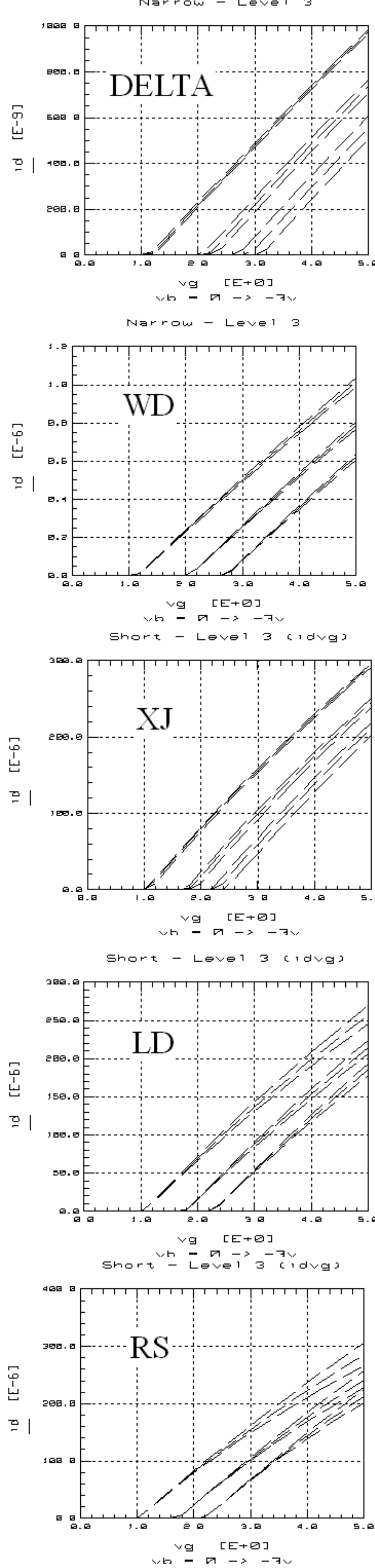
Literature:

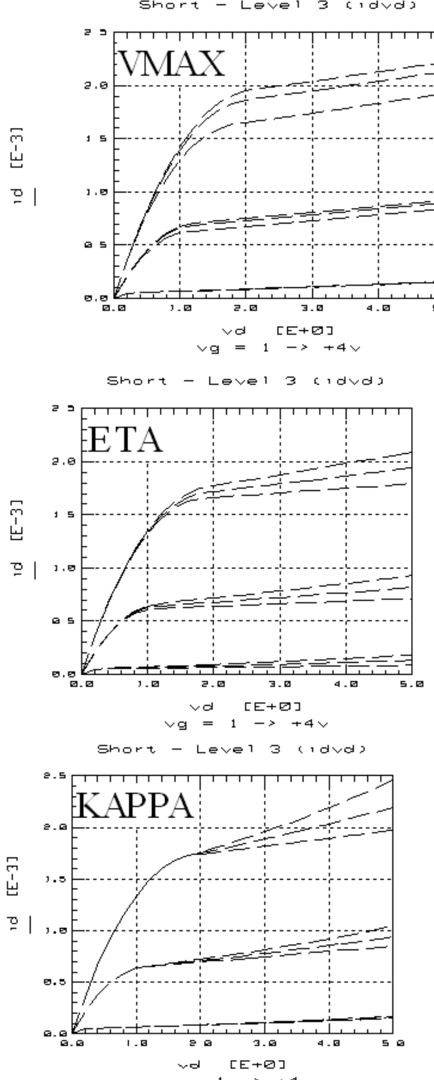
HP94452A UCB MOSFET Model and Parameter Extraction Manual

A.Vladimirescou, S.Liu, The Simulation of MOS Integrated Circuits Using Spice2, Memorandum UCB/ERL M80/7, University of California, Berkeley

The following sequence of plots gives an overview about the influence of the model parameters and their fitting properties. It represents also a possible sequence for the model parameter extraction.

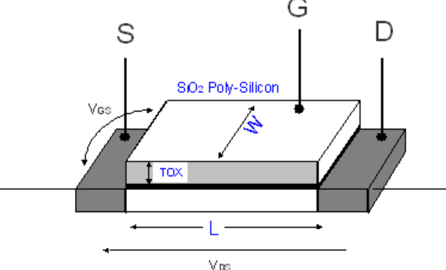

LARGE - Level 3
VTO
id.s [E-6]
vg [E+0]
vb = 0 -> -3v
LARGE - Level 3
UO
id.s [E-6]
vg [E+0]
vb = 0 -> -3v
LARGE - Level 3
THETA
id.s [E-6]
vg [E+0]
vb = 0 -> -3v
LARGE - Level 3
NSUB
id.s [LOG]
vg [E+0]
vb = 0 -> -3v
LARGE - Level 3
NFS
id.s [LOG]
vg [E+0]
vb = 0 -> -3v


Narrow - Level 3
DELTA
id [E-9]
vg [E+0]
vb = 0 -> -7v
Narrow - Level 3
WD
id [E-6]
vg [E+0]
vb = 0 -> -7v
Short - Level 3 (idvg)
XJ
id [E-6]
vg [E+0]
vb = 0 -> -7v
Short - Level 3 (idvg)
LD
id [E-6]
vg [E+0]
vb = 0 -> -7v
Short - Level 3 (idvg)
RS
id [E-6]
vg [E+0]
vb = 0 -> -7v


Short - Level 3 (idvd)
VMAX
id [E-3]
vd [E+0]
vg = 1 -> +4v
Short - Level 3 (idvd)
ETA
id [E-3]
vd [E+0]
vg = 1 -> +4v
Short - Level 3 (idvd)
KAPPA
id [E-3]
vd [E+0]
vg = 1 -> +4v

MOS Transistors for Dummies

or: The MOS Drain current equation on one page (after B.Höfflinger, IMS, Stuttgart)



We start with neglecting the influence of V_{DS} on the electric field between Gate and Source,i.e. $V_{DS} << V_{GS}$, and consider the charge in the Gate capacitor.

$$C_{OX} = \frac{\varepsilon_{ox} \cdot L \cdot W}{T_{ox}}$$

and its charge:

$$Q = C_{OX} \cdot V_{GS}$$

The current from Drain to Source is

$$I_{DS} = \frac{Q}{\tau}$$

with the charge velocity

$$v = \frac{L}{\tau}$$

On the other hand, the charge velocity is

$$v = \mu \cdot E_{DS}$$

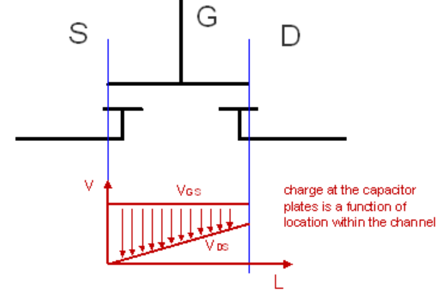
and the Drain Source voltage is

$$V_{DS} = E_{DS} \cdot L$$

Inserting these equations into each other gives

$$I_{DS} = \frac{Q}{\tau} = C_{OX} \cdot V_{GS} \cdot \frac{\mu \cdot V_{DS}}{L^2} = \frac{\varepsilon_{ox}}{T_{ox}} \cdot \frac{W}{L} \cdot \mu \cdot V_{GS} \cdot V_{DS}$$

Now, let's extend this idea to also include the effect of V_{DS} on V_{GS} vs. the Gate length L i.e. $V_{DS} > V_{GS}$.



The charge at $y = 0$ is

$$Q_{(y=0)} = C_{OX} \cdot V_{GS}$$

while it is at $y = L$

$$Q_{(y=L)} = C_{OX} \cdot (V_{GS} - V_{DS})$$

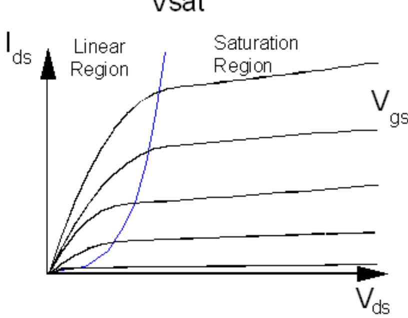
To keep things simple, let's consider the mean charge:

$$Q_{(mean)} = C_{OX} \cdot \frac{V_{GS} + V_{GS} - V_{DS}}{2}$$

what finally gives:

$$I_{DS} = \frac{Q_{(mean)}}{\tau} = \frac{\varepsilon_{ox}}{T_{ox}} \cdot \frac{W}{L} \cdot \mu \cdot \left(V_{GS} - \frac{V_{DS}}{2}\right) \cdot V_{DS}$$

This equation is valid until I_{ds} gets into saturation.



Compare this result with the MOS Level 1, where is:

$$I_{DS} = KP \cdot \frac{W}{l} \cdot \left(v_{GS} - VTH - \frac{v_{DS}}{2} \right) \cdot v_{DS} \cdot (1 + LAMBDA \cdot v_{DS})$$

with KP=U0*COX U0 mobility (600cm²/V/s)
COX = ε_{SiO2} / TOX
$\varepsilon_{SiO2} = \varepsilon_0 \, \varepsilon_r$ = 34.5E-12 As/V/m

LAMBDA channel length modulation,
corresponding to early voltage for bipolar transistors

Let's now interpret these basic formulas a bit and learn some basics about gm and fT.

We commence with:

$$gm = \frac{\partial i_D}{\partial v_{GS}}$$

Derivation gives:

$$gm = \left(\frac{\varepsilon_{ox}}{T_{ox}} \cdot \frac{W}{L} \cdot \mu \right) \cdot v_{DS} = \left(\frac{\varepsilon_{ox}}{T_{ox}} \cdot \frac{W}{L} \cdot \mu \right) \cdot E_{DS} \cdot L$$

Inserting the charge velocity v = µ *E $_{DS}$ ~ 107cm/sec, gives

$$gm = \frac{\varepsilon_{ox}}{T_{ox}} \cdot W \cdot v$$

from what we learn that wide transistors exhibit more gm:

$$W \uparrow \;\rightarrow\; gm \uparrow$$

Considering now the cutoff or transit frequency of MOS transistors:

$$\omega_T = \frac{gm}{Cox} = \left[\frac{\varepsilon_{OX}}{T_{OX}} \cdot W \cdot v \right] \cdot \left[\frac{T_{OX}}{\varepsilon_{OX} \cdot W \cdot L} \right] = \frac{1}{TAU}$$

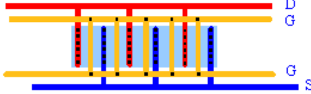
Like stated above, it is TAU = L / v with v ~ 107cm/sec, what gives

$$f_T = \frac{1}{2 \cdot PI \cdot TAU} = \frac{v}{2 \cdot PI \cdot L}$$

from what we learn that short transistors exhibit more fT:

$$L \downarrow \;\rightarrow\; f_T \uparrow$$

This means that for high-frequency MOS transistors and high gain gm, we need short and wide transistors. Practically, this is a finger-structured transistor layout !



As an example, assuming L = 0,1um, we get f $_T$ ~ 100GHz.

MOS Modeling for Dummies

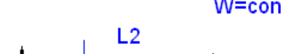
In the MOS model Level 1, I $_{DS}$ is including the lateral underdiffusion LD:

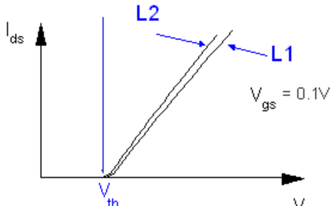
$$I_{DS} = KP \cdot \frac{W}{L - 2 \cdot LD} \cdot \left(v_{GS} - VTH - \frac{v_{DS}}{2} \right) \cdot v_{DS} \cdot (1 + LAMBDA \cdot v_{DS})$$

or, simplified by neglecting the output characteristics slope in saturation, i.e. LAMBDA,

$$I_{DS} = KP \cdot \frac{W}{L - 2 \cdot LD} \cdot \left(v_{GS} - VTH - \frac{v_{DS}}{2} \right) \cdot v_{DS}$$

The following sketch depicts two I $_{DS}$ curves of two transistors with same W, but different lengths L1 and L2.





It is obvious that the threshold voltage, V_{TH}, can be estimated from the x-intersection of this plot.
In order to extract the model parameters KP and LD, we consider the slope gm of I_{DS} :

$$\frac{\partial I_{DS}}{\partial v_{GD}} = KP \cdot \frac{W}{L - 2 \cdot LD} \cdot v_{GS} = m \cdot v_{GS}$$

Reading the two slopes m1 and m2 from the above plot, and from the starting point

$$KP \cdot W = m2 \cdot L2 - 2 \cdot m2 \cdot LD$$

$$KP \cdot W = m1 \cdot L1 - 2 \cdot m1 \cdot LD$$

Solving this set of equation gives:

$$LD = \frac{m2 \cdot L2 - m1 \cdot L1}{2 \cdot (m2 - m1)}$$

and

$$KP = \frac{m1 \cdot m2}{m2 - m1} \cdot \frac{L1 - L2}{W}$$

PSP

Contents

- *Background of the PSP Model Extraction Package* (iccapmhb)
- *Global PSP Model* (iccapmhb)
- *PSP Model Extraction Package* (iccapmhb)
- *PSP Parameter Extraction Strategy* (iccapmhb)

Background of the PSP Model Extraction Package

Introduction to PSP

- The PSP model is a compact MOSFET model intended for digital, analogue, and RF design,
- It has been jointly developed by Philips Research (now NXP) and The Pennsylvania State University (the PSP co-author moved to Arizona State University).
- The roots of PSP lie in both MOS Model 11 (Philips Research) and the SP model (Penn State University).
- In December 2005, the Compact Model Council (CMC) has elected PSP as the new industrial standard model for compact MOSFET modeling.

Physical principles in PSP

- PSP is a surface-potential based MOS Model, containing all relevant physical effects (mobility reduction, velocity saturation, DIBL, gate current, lateral doping gradient effects, STI stress, etc.) to model present-day and upcoming bulk CMOS technologies in the nanometer regime. The source/drain junction model, c.q. the JUNCAP2 model, is fully integrated in PSP.
- PSP not only gives an accurate description of currents, charges, and their first order derivatives (i.e. transconductance, conductance and capacitances), but also of the higher order derivatives, resulting in an accurate description of electrical distortion behavior. The latter is especially important for analog and RF circuit design.
- The model furthermore gives an accurate description of the noise behavior of MOSFETs. Finally, PSP has an option for simulation of non-quasi-static (NQS) effects.

Internal model structure of PSP

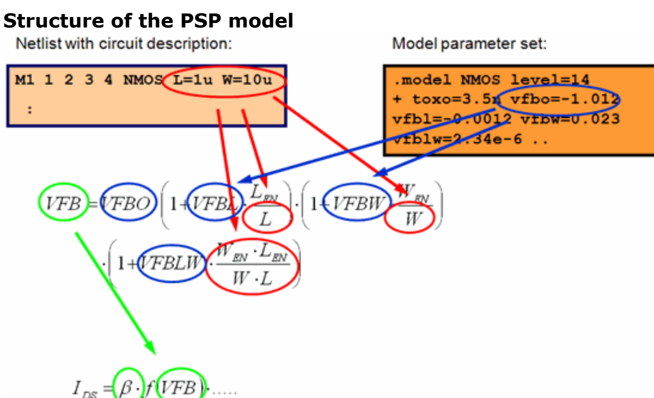
BSIM4 Simulator

Structure of the BSIM4 model

Calculation of the electrical behavior of the MOSFET in a complex set of equations which contain both, device dimensions (L, W, NF, ..) and model parameters (TOXE, DVT0, VTH0,..)

PSP Simulator

Structure of the PSP model



In a first step, a scaling of parameters is performed. From a so called Global Model which contains both, device dimensions and model parameters, an effective parameter is calculated which is called Local parameter.
The effective transistor properties are calculated from Local parameters only.

Types of parameter scaling

Local parameter Equation in the **Global Model**

No scaling at all:

$$TOX = TOXO$$

Scaling vs. L or W only:

$$RS = RSW1 \cdot \frac{W_{EN}}{W} \cdot \left(1 + RSW2 \cdot \frac{W_{EN}}{W}\right)$$

Complex scaling vs. L, W and combinations of L,W:

$$DPHIB = \left(DPHIBO + DPHIBL \cdot \left[\frac{L_{EN}}{L}\right]^{DPHIBLEXP}\right) \cdot \left(1 + DPHIBW \cdot \frac{W_{EN}}{W}\right) \cdot \left(1 + DPHIBW \cdot \frac{W_{EN} \cdot L_{EN}}{W \cdot L}\right)$$

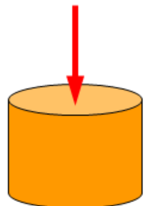
Parameter extraction principle for scalable models

Taking into account the structure of the PSP model with different types of scaling, we can identify the following kinds of parameter extractions:

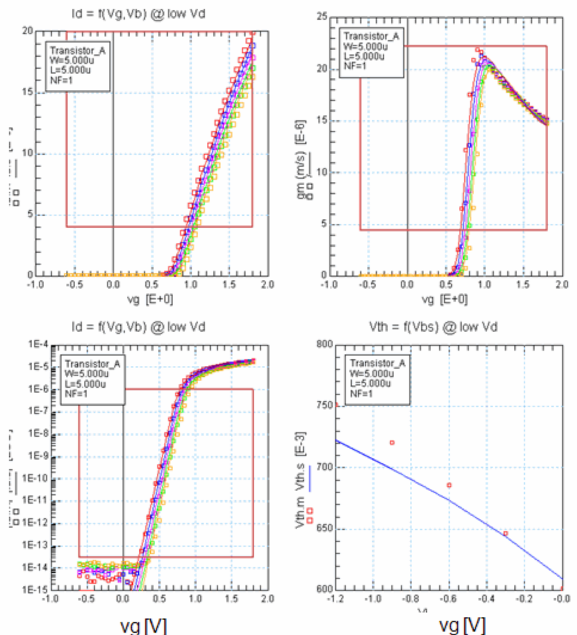
- Parameters, which do not scale and which are identical in the local and in the global model like the very fundamental parameters like TOXO.
- Parameters not affecting the „transistor" part in the PSP model (junction capacitance or the diode).
- Parameters which scale with the transistor dimensions (L, W). These parameters are mostly describing the „transistor" behavior and influence the core of the PSP model. The following pages will highlight the extraction methodology for those devices in detail.

Local extraction

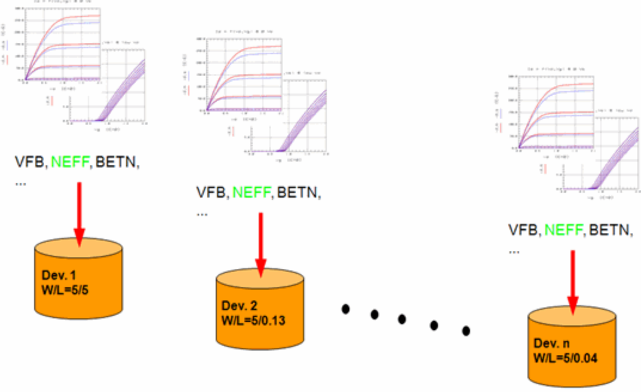
The goal is to extract a set of local parameters from one device only without taking into account the device dimensions:
VFB, BETN, NEFF, ...



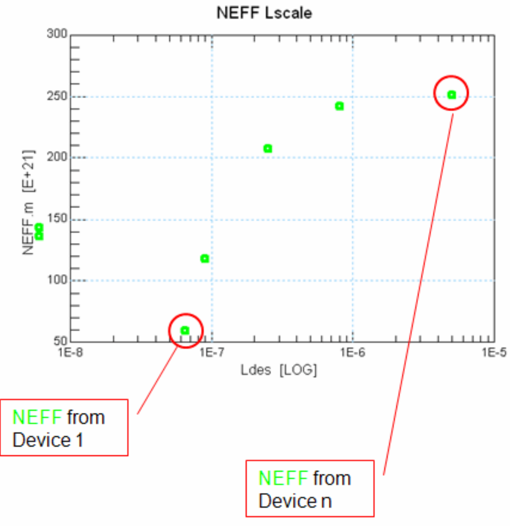
A set of Local model parameters, which is valid for one dedicated transistor

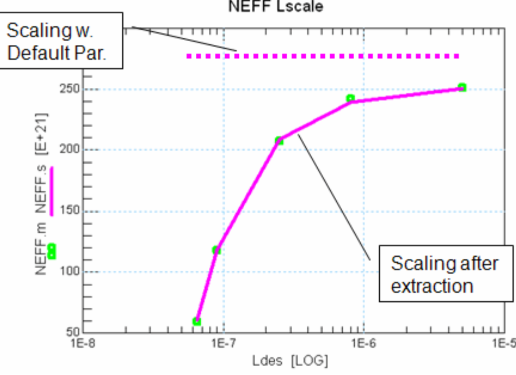


Repeat this procedure for different devices with different Dimension



Local Parameter vs. Dimension
The local parameter values are plotted vs. device dimensions





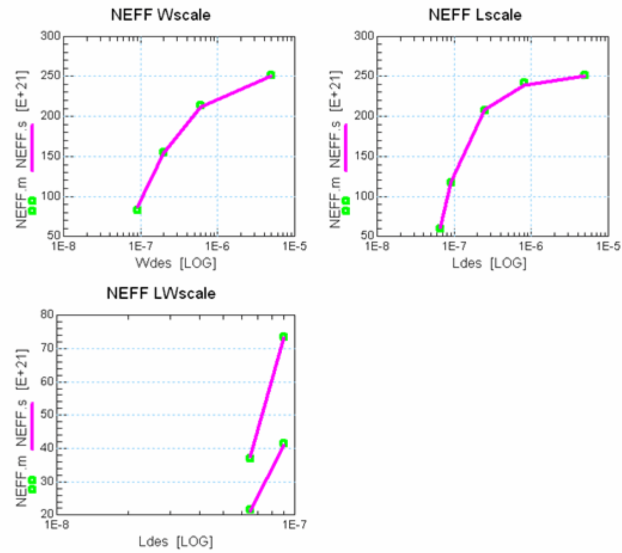
Determine the scaling parameters from this curve according to the built in scaling equation:

$[A^2/Hz]$

Global Scaling

With a more complex scaling equation, parameters are plotted vs.

- L @wide W,
- W @ wide L,
- and L,W simultaneously



Example for a typical scaling equation:
$0.2nA/\sqrt{Hz}$

Binning in PSP

- The PSP model provides in addition a binning feature.
- Instead of the global scaling, a scaling of local parameters inside defined bin boundaries can be applied.
- However, the difference to former MOS models with binning features is, that PSP provides 3 types of scaling equations which are shown on the following page.
- This approach allows to make the binning a little bit more physical compared to the approach in the BSIMxx models.

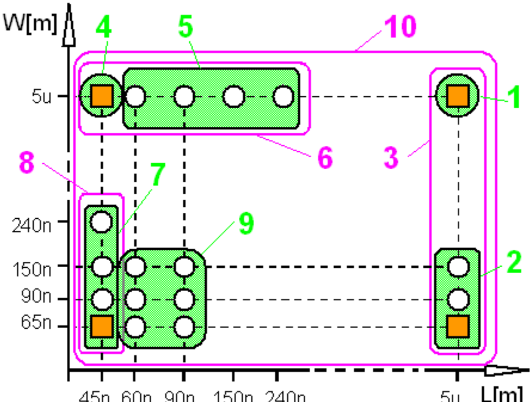
Type 1: $[V^2/Hz]$
Type 2: $E_{nv}[V/\sqrt{Hz}]$
Type 3: A^2/Hz

Parameter	Explanation	Equivalent BSIM4 Expression
LE	effective channel length L-	Leff
WE	effective channel width W	Weff
LEN	normalized channel length = 1e-6	-
WEN	normalized channel width = 1e-6	-

Example extraction flow

Extraction Steps

The extraction procedure is clearly defined by the model definition itself.
Due to the different scaling equations, the stepping sequence (1..10) shown in the figure below must be followed.

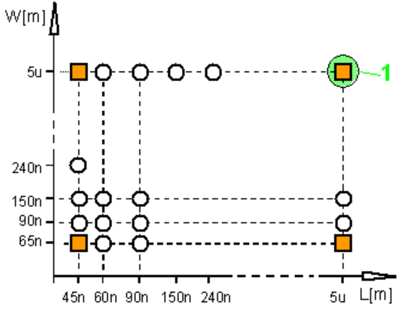


A few typical steps are described below:

- Extraction of Local parameter sets of a similar group of devices.
- Adjustment of Global model parameters to a group of extracted Local parameters.

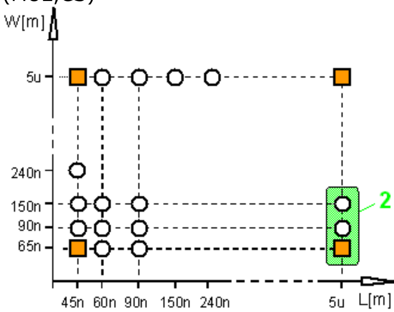
1. **Long-Wide**

- Determination of a basic set of Local parameters (NEFF, BETN,..)
- Extraction of some Local parameters, which will be fixed for all other devices (VP, THEMU, ..). These parameters are immediately transformed into appropriate global parameters (VPO, THEMUO,..)



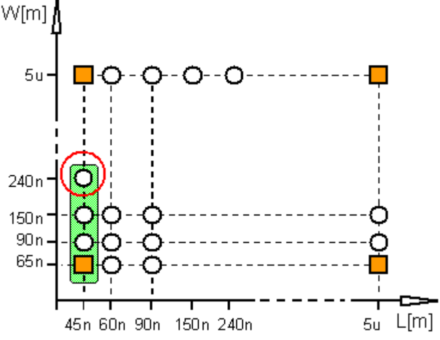
1. **Long-Width Dependence**

- Determination of a basic set of Local parameters (NEFF, BETN,..)
- Extraction of some Local parameters, which will describe narrow channel effects only (MUE,CS)



1. **Scaling of Width Dependence**
   In this step, some global parameters, which are only width dependent are already determined from the appropriate local parameters.
   In this case these are:
   MUEO, MUEW from MUE from steps 1,2
   CSO, CSW from CS from steps 1,2

Impact on further extraction steps
When the parameter set of the highlighted device is extracted, some local parameters must be set and calculated from the global parameters before the extraction starts.
These are exactly those parameters, which have been extracted only in steps 1 and 2 and which are describing narrow channel effects.
In addition, there is no way to simply "transfer" these parameters from a certain other device, because such a device with an appropriate gate width does not exist.



$$MUE = MUEO \cdot \left(1 + MUEW \cdot \frac{W_{EN}}{W}\right)$$

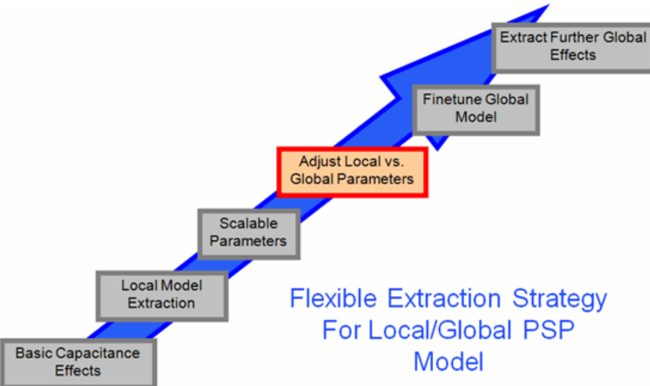
$$CS = CSO \cdot \left(1 + CSW \cdot \frac{W_{EN}}{W}\right)$$

Summary

- We introduced the internal structure of the PSP model and highlighted the approach

in separating the model equations into a so called local and a global model.
- Starting from this model structure the principle of the two step parameter extraction method going from local to global models is demonstrated.
- Finally, the example for the extraction flow shows how to do such a parameter extraction step by step in detail taking into account the complex relation between different parameter groups.
- As a conclusion, this approach is implemented in the brand new PSP Extraction Package in IC-CAP from Agilent Technologies.

PSP Extraction Flow



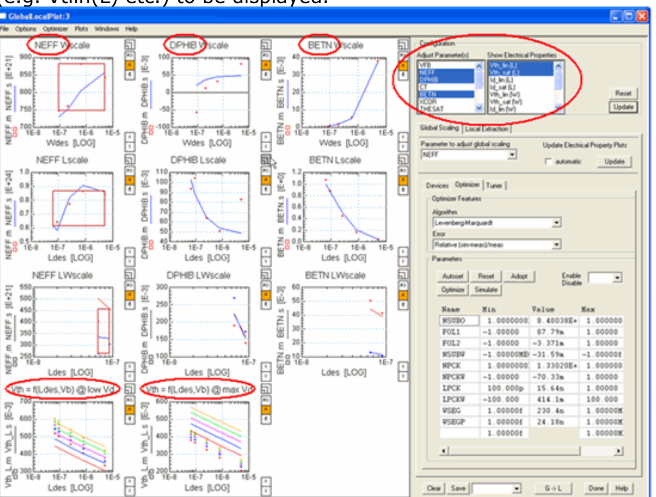
Requirements from PSP modeling teams

High level description of necessary functionality of a simultaneousadjustment of both, global and local model parameters:

- Show more than one group of global parameters simultaneously (e.g. DPHIB and NEFF)
- Display in addition scaling behavior of the global model for several devices (e.g. Vtlin vs. L, Idlin vs. L, etc.)
- Select one device and invoke the local parameter tuning/optimization for this device. Show the change of the parameters in all the following diagram groups: - local device behavior (I-V curves)- "measured" data of global parameters (Parameters vs. L, W)
- Invoke the optimization/tuning of global parameters. Show the effect of changing a parameter in the "simulated" part of the global parameters group and in the selected diagrams for electrical scaling (e.g. Vtlin, Idsat, ..).

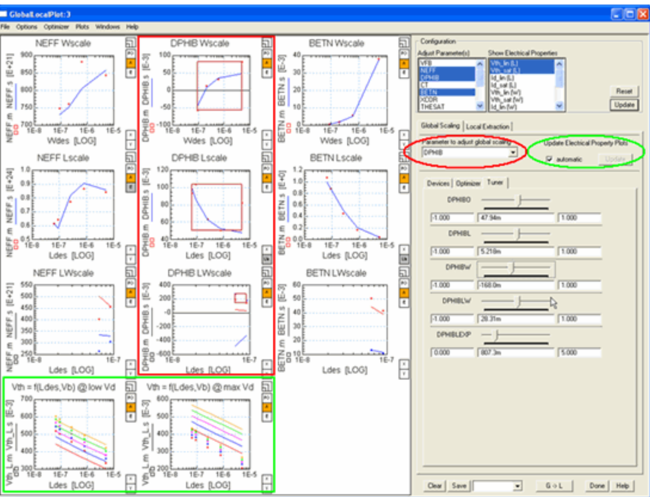
Plot Configuration

Select the parameters to display the scaling as well as some electrical scaling diagrams (e.g. Vtlin(L) etc.) to be displayed.



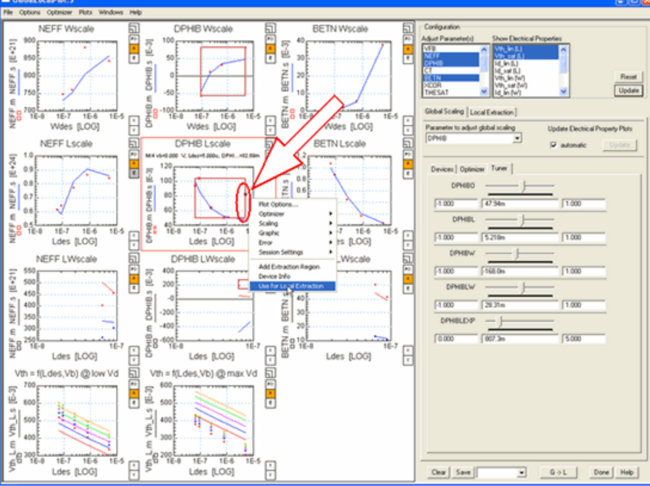
Adjust Scaling

Select the parameter to be adjusted by tuning or optimization.
The electrical scaling diagrams can be updated simultaneously or manually.
Due to the use of ADS as a simulator, this is done in real time.



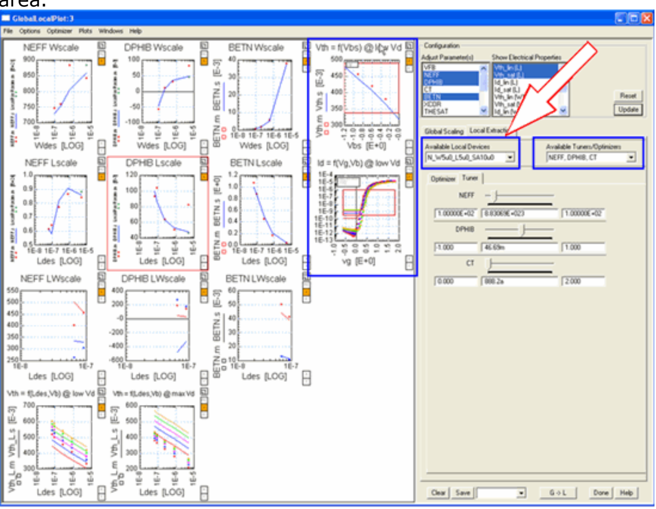
Switch to local extraction

Select a device, which is far off the scaling. Using the right mouse button, a menu can be invoked which allows the switching to the Local Extraction mode.

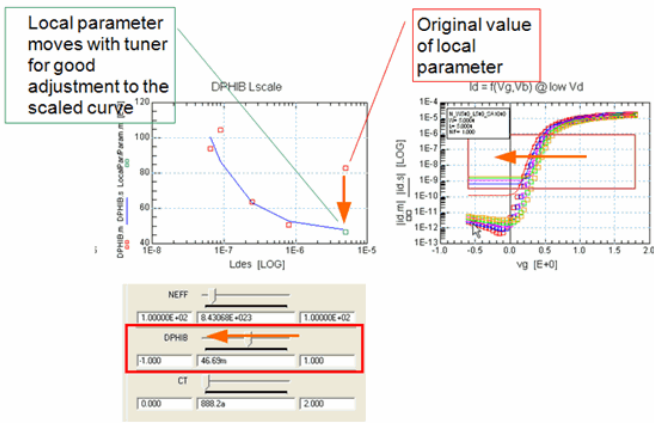


Adjust local parameter

In addition, the I-V curves of a local device are shown
All available tuners / optimizers for this device can be selected.
After selecting such an optimizer / tuner, the appropriate diagrams will appear in the plot area.

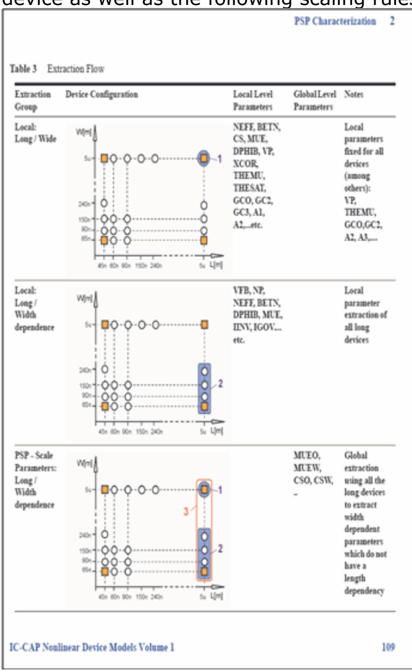


Simultaneous update of I-V curves of a local device and the local parameter in the scaling diagrams while tuning a local parameter.



Documentation

The integrated documentation in IC-CAP describes the extraction steps in detail.
The different kinds of extracted local parameters depending on the dimensions of the device as well as the following scaling rules are outlined in a easy to follow flow.



PSP Characterization 2

Table 3 Extraction Flow

Extraction Group	Device Configuration	Local Level Parameters	Global Level Parameters	Notes
Local: Long / Wide		NEFF, BETN, CS, MUE, DPHIB, VP, XCOR, THEMU, THESAT, GCO, GC2, GC3, A1, A2,...etc.		Local parameters fixed for all devices (among others): VP, THEMU, GCO,GC2, A2, A3,...
Local: Long / Width dependence		VFB, NP, NEFF, BETN, DPHIB, MUE, IINV, IGOV... etc.		Local parameter extraction of all long devices
PSP - Scale Parameters: Long / Width dependence			MUEO, MUEW, CSO, CSW, ..	Global extraction using all the long devices to extract width dependent parameters which do not have a length dependency

IC-CAP Nonlinear Device Models Volume 1 109

Technical Details

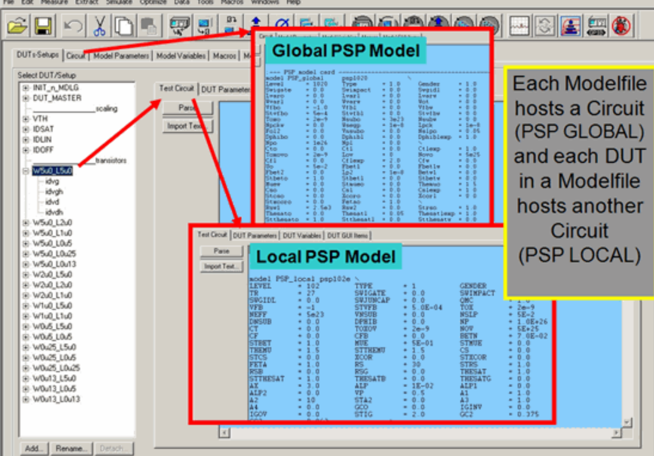
- Simulators supported: ADS*), Spectre (others like HSPICE will be added depending on the availability of PSP1020 in those simulators.
- Supports the generation of scalable PSP1020 models as well as of the binned PSP1021 model.

Global PSP Model

The IC-CAP Model File structure with its organization in DUTs and according Setups (Measurement, Simulation and Extraction/Optimization environment) makes it an ideal tool for handling the PSP model.

The local model is hosted in the Circuit on DUT level, while the global model is located in the Circuit on Model File level.

Scaling e.g. vs. L, W etc, is done by using the LSYNC sweep in any Setup.

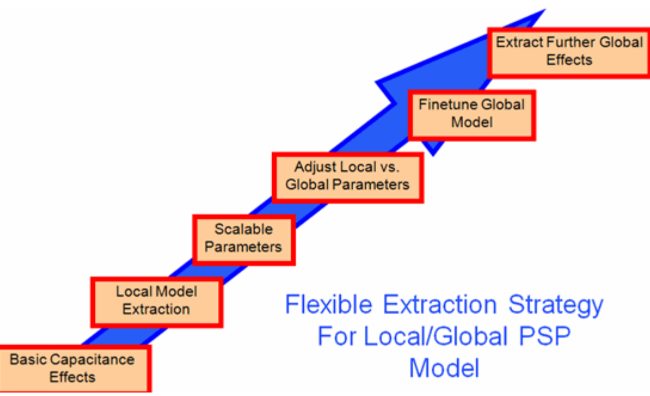


PSP Model Extraction Package

Key Features

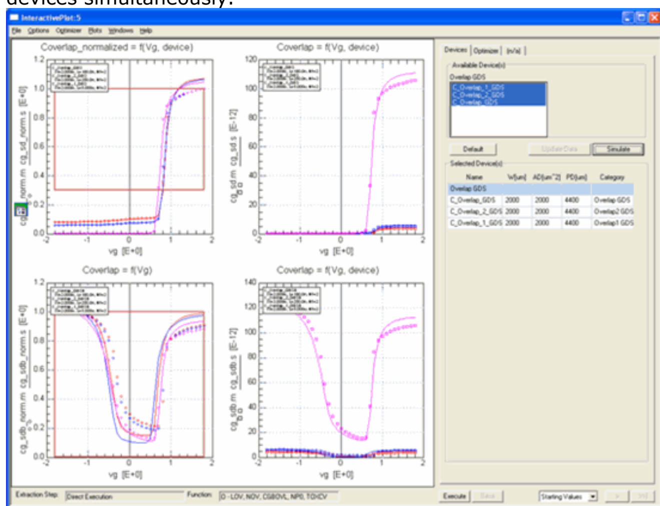
- The IC-CAP PSP Modeling Package features modules for both, measurements and extractions, of DC, CV and S-parameter behavior.
- It has been developed based on the recommendations of the model developers (NXP, Philips Research and Penn State University) as well as from modeling teams already using the PSP model.
- The extraction procedure was successfully tested on devices with a minimum feature size down to the 45nm technology node.

PSP Extraction Flow

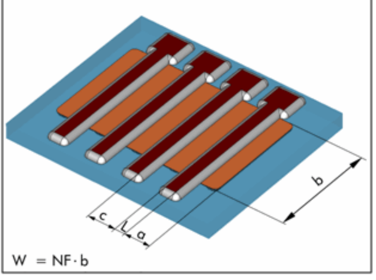


Basic Capacitance

The oxide capacitance as well as the overlap capacitance effects are derived from the well known test structures.
For this purpose, the global model of PSP is invoked and the fitting is done on multiple devices simultaneously.

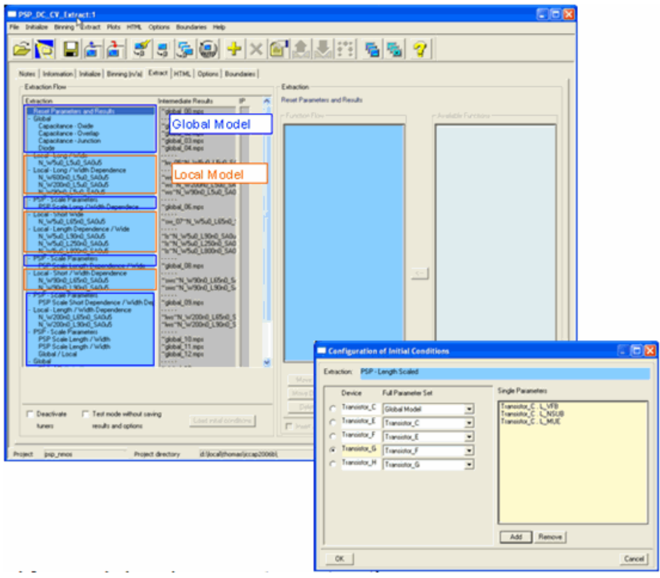


This extraction must be performed prior to the generation of local models of single transistor devices



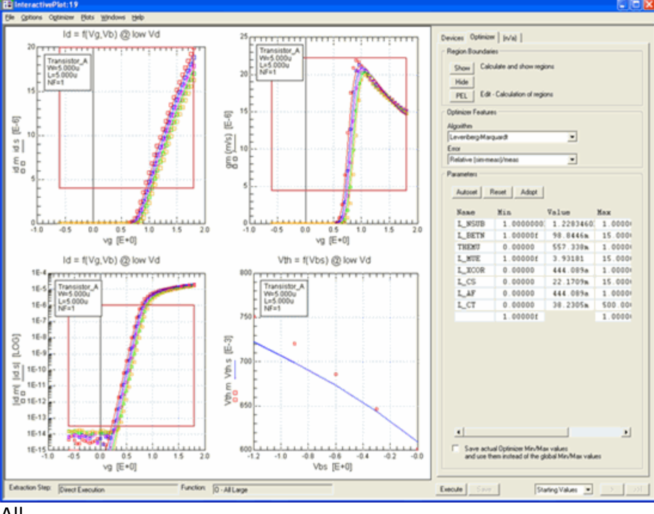
Local Model Extraction

In general, the extraction flow control of the PSP Modeling Package can combine both, local and global parameter extraction.
This is necessary due to the dependency of local model parameters on basic parameters from Cox and Coverlap extractions.



- Initial conditions can be specified for each local parameter extraction.
- Complete start parameter sets as well as single parameters can be taken from previous results.

Extractions of local parameters are defined according to the recommendations of the model developers.



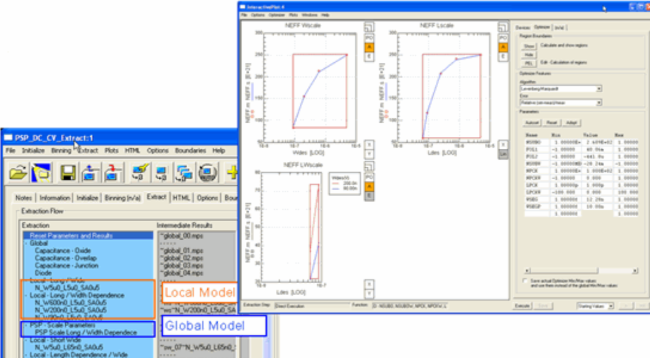
All

- extractions
- optimizations
- tuners

  can easily be customized with respect to the selection of parameters, optimization algorithms, optimization region etc.

Scalable Parameters

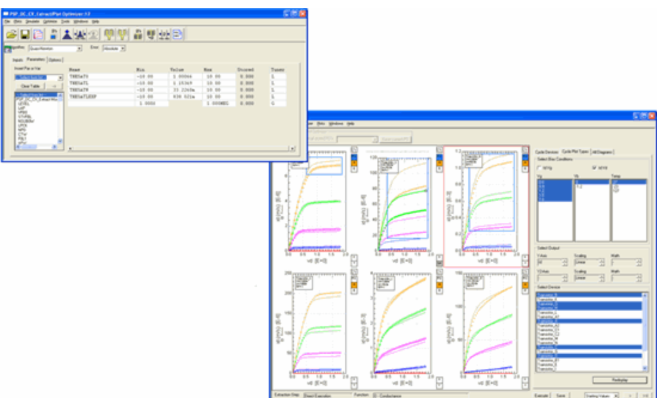
Plot local parameters vs. gate length or gate width to extract and adjust global model parameters.



The IC-CAP PSP package extraction flow provides all groups for global model parameter

fitting according to the categories of the different device dimensions.

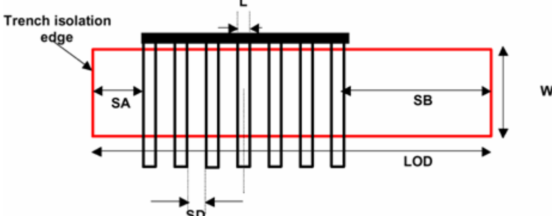
Verification and Fine tuning of Global Model



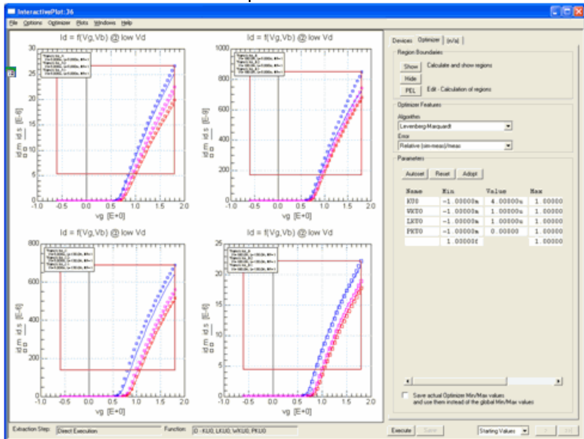
It is very likely, that it is necessary to fine tune a global model after extracting the global model parameters from local models.
This functionality is provided by the combined data display and the Plot Optimizer inside the PSP Modeling Package.
The user can select a huge variety of characteristic diagrams and can apply own, graphically defined optimizers/tuners to those diagrams.
Once such an optimizer is defined, it can be reused for later extractions.

Extract STI Parameters

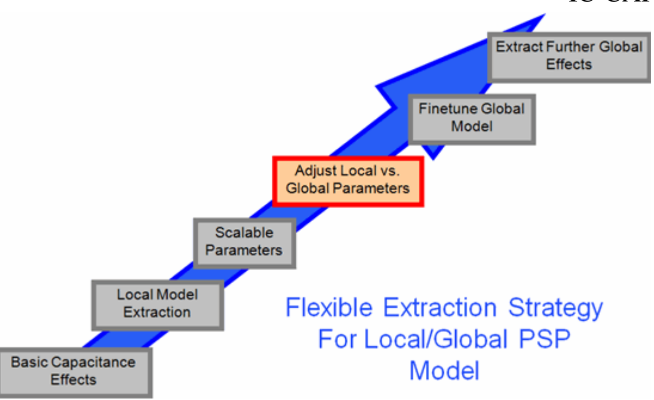
The STI (Shallow Trench Isolation) stress parameters can only be extracted after having a well fitted global model available.



This is done as a final step in the extraction flow.



Extensions to the PSP Model Extraction Package from the Modeling Bench



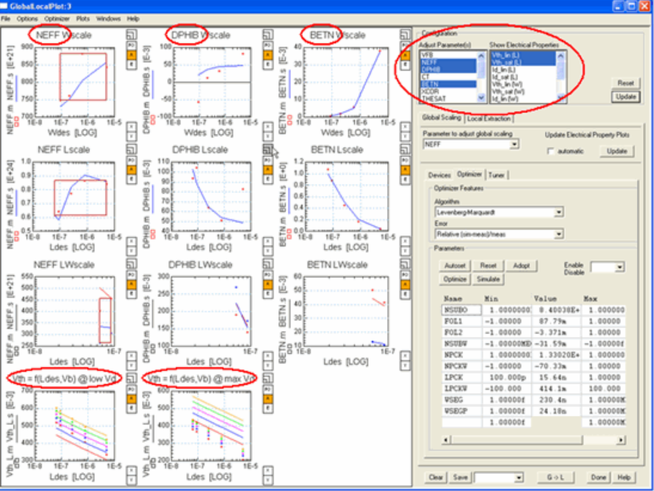
Requirements from PSP modeling teams

High level description of necessary functionality of a simultaneous adjustment of both, global and local model parameters:

- Show more than one group of global parameters simultaneously (e.g. DPHIB and NEFF)
- Display in addition scaling behavior of the global model for several devices (e.g. Vtlin vs. L, Idlin vs. L, etc.)
- Select one device and invoke the local parameter tuning/optimization for this device. Show the change of the parameters in all the following diagram groups: - local device behavior (I-V curves)- "measured" data of global parameters (Parameters vs. L, W)
- Invoke the optimization/tuning of global parameters. Show the effect of changing a parameter in the "simulated" part of the global parameters group and in the selected diagrams for electrical scaling (e.g. Vtlin, Idsat, ..).

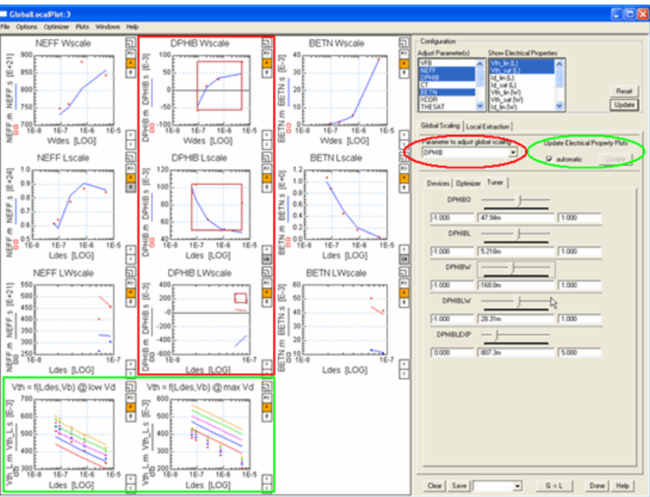
Plot Configuration

Select the parameters to display the scaling as well as some electrical scaling diagrams (e.g. Vtlin(L) etc.)



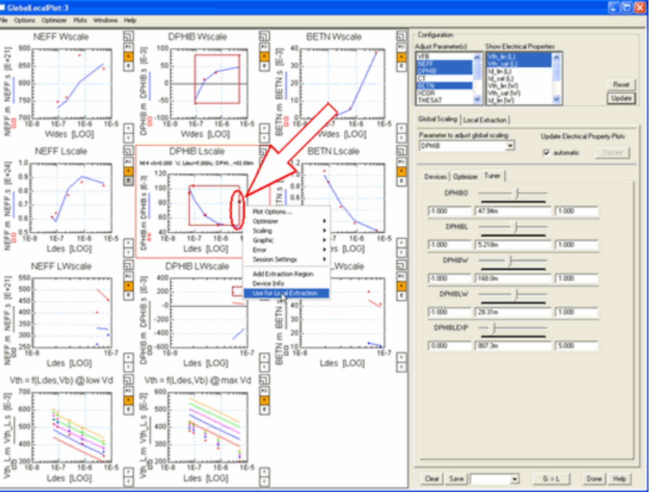
Adjust Scaling

Select the parameter to be adjusted by tuning or optimization.
The electrical scaling diagrams can be updated simultaneously or manually.
Due to the fast link to ADS as a simulator, this is done in real time.



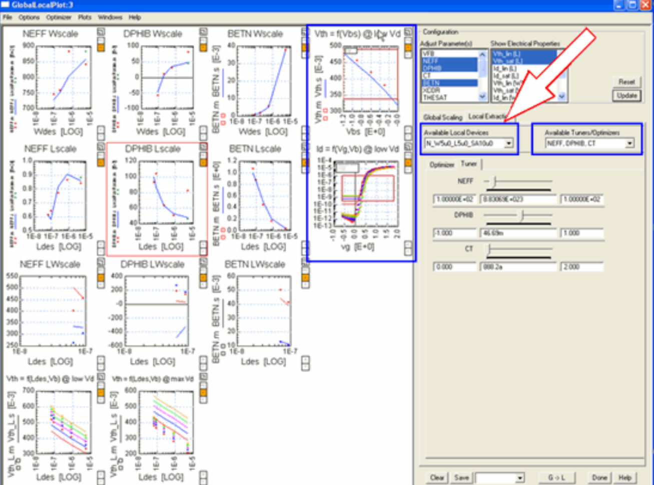
Switch to local extraction

Select a device, which is far off the scaling.
Using the right mouse button, a menu can be invoked which allows the switching to the Local Extraction mode.

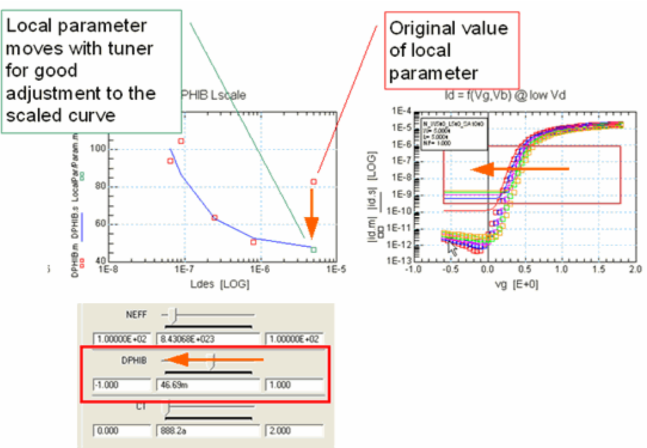


Adjust local parameter

In addition, the I-V curves of a local device are shown.
All available tuners / optimizers for this device can be selected.
After selecting an optimizer / tuner, the appropriate diagrams will appear in the plot area.

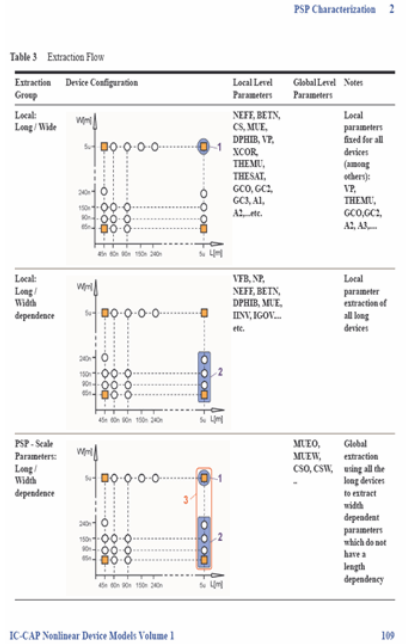


Simultaneous update of I-V curves of a local device and the local parameter in the scaling diagrams, while tuning a local parameter.



Documentation

The integrated documentation in IC-CAP describes the extraction steps in detail.
The different kinds of extracted local parameters depending on the dimensions of the device as well as the following scaling rules are clearly outlined and can be easily followed.



PSP Characterization 2

Table 3 Extraction Flow

Extraction Group	Device Configuration	Local Level Parameters	Global Level Parameters	Notes
Local: Long / Wide		NEFF, BETN, CS, MUE, DPHIB, VP, XCOR, THEMU, THESAT, GCO, GC2, GC3, A1, A2,...etc.		Local parameters fixed for all devices (among others): VP, THEMU, GCO,GC2, A2, A3,...
Local: Long / Width dependence		VFB, NP, NEFF, BETN, DPHIB, MUE, IINV, IGOV... etc.		Local parameter extraction of all long devices
PSP - Scale Parameters: Long / Width dependence			MUEO, MUEW, CSO, CSW, ...	Global extraction using all the long devices to extract width dependent parameters which do not have a length dependency

IC-CAP Nonlinear Device Models Volume 1 109

PSP Modeling

Don't worry about the size and readability of a single plot. Using 2 or more FlatScreens, the PSP Modeling Package can display all required information at once.

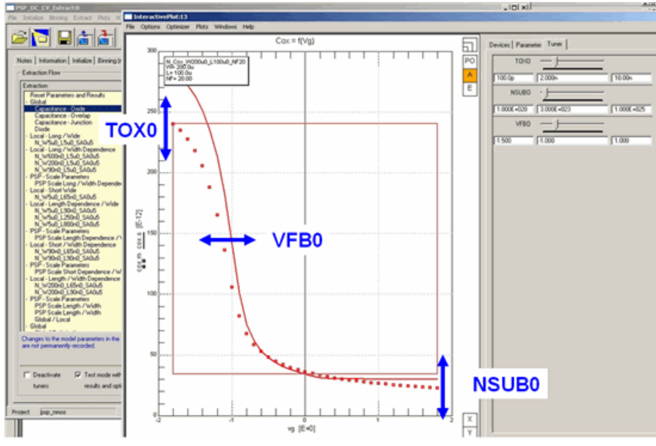
Technical Details

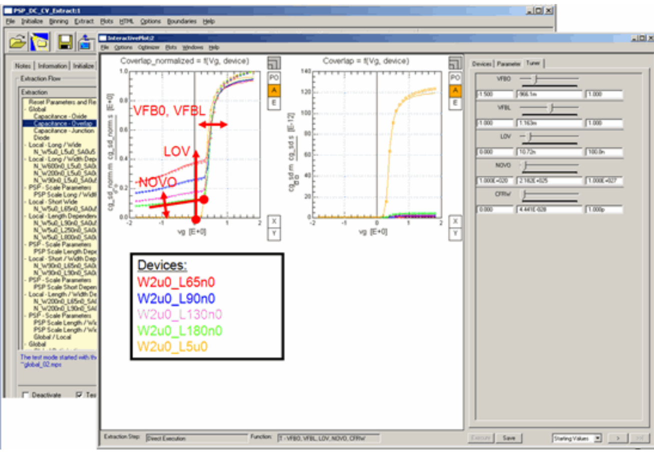
- Simulators supported: ADS*), Spectre (others like HSPICE will be added depending on the availability of PSP1020 in those simulators).
- Supports the generation of scalable PSP1020 models as well as of the binned PSP1021 model.
- The PSP Modeling Package is delivered starting December 7, 2006 through an update to IC-CAP2006 (available on the Agilent EEsof Knowledge Center)
- For more information, please visit the Agilent website: http://eesof.tm.agilent.com/products/iccap_psp.html

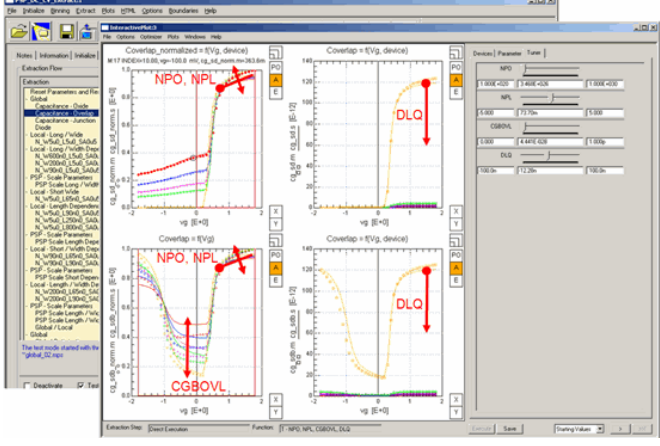
PSP Parameter Extraction Strategy

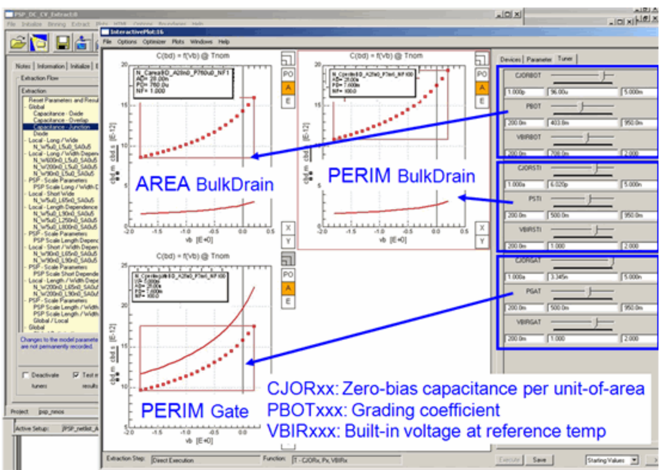
This information below gives an overview about the impact of the PS model parameters and a proposed parameter extraction sequence.

CV Modeling

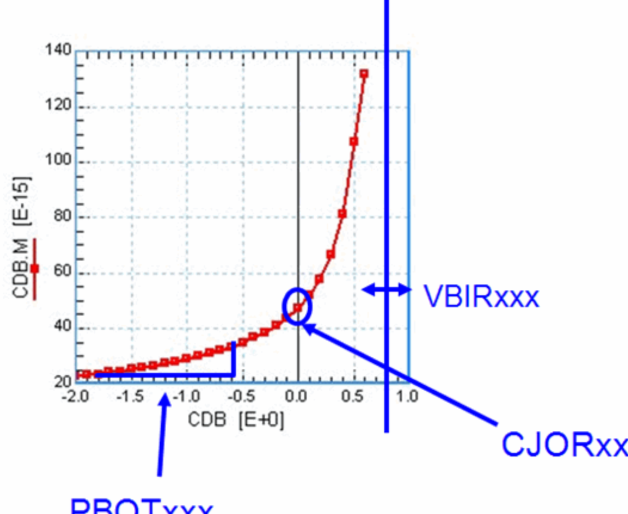






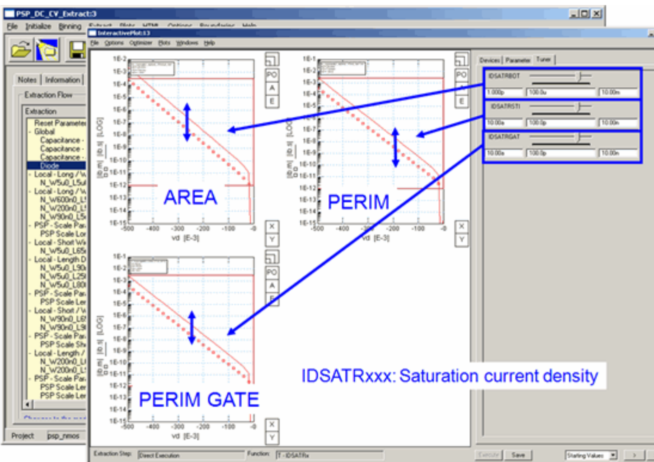


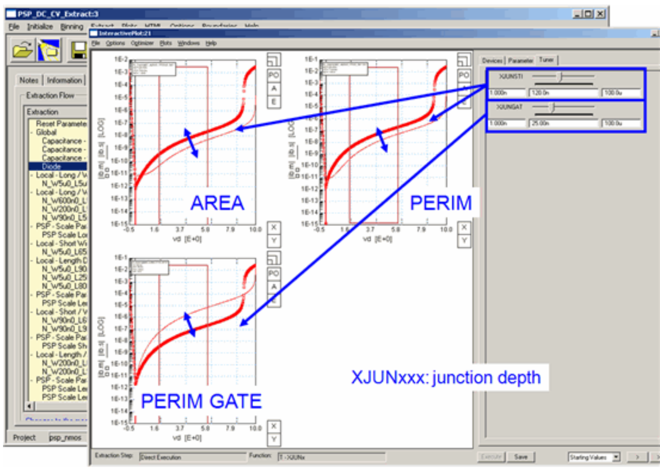
Modeling of junction capacitance

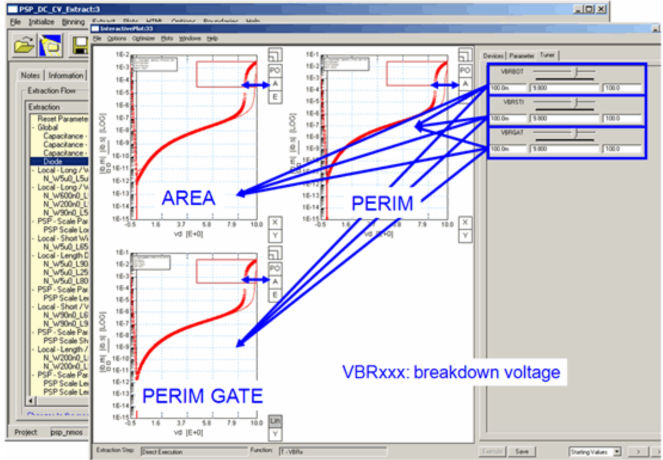


VBIRxxx: Built-in voltage at reference temp
PBOTxxx: Grading coefficient
CJORxx: Zero-bias capacitance per unit-of-area

Diode

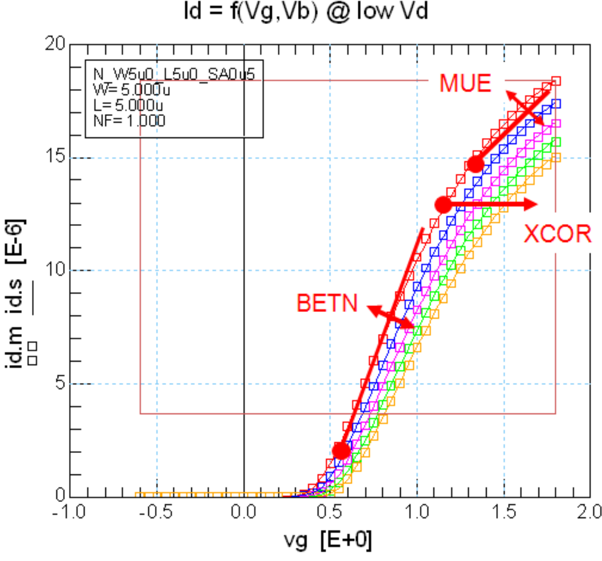






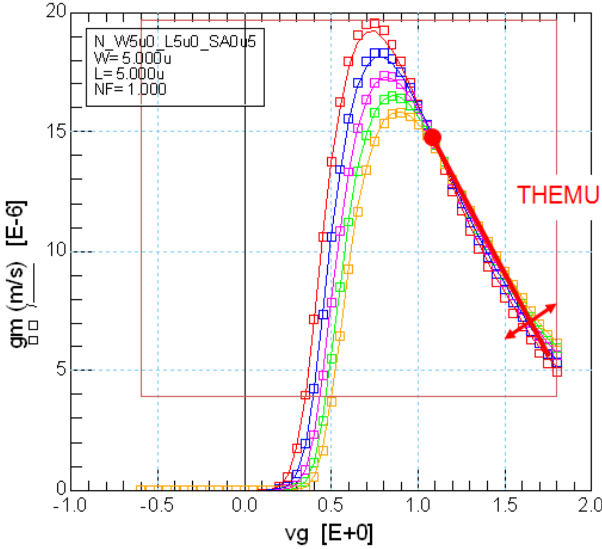
Individual Transistor DC Modeling

id_vg



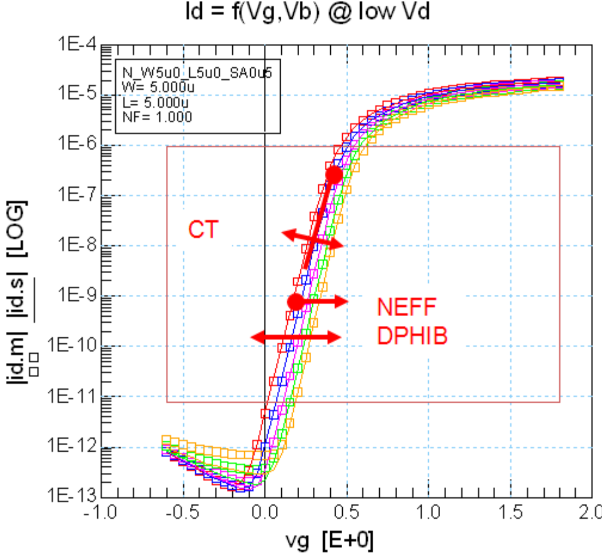
BETN: (0..70m..inf) Channel aspect ratio times zero-field mobility
MUE: (0..0.5..inf) Mobility reduction coefficient at nominal reference temperature TR
XCOR: (0..0..inf) Non-universality factor

gm



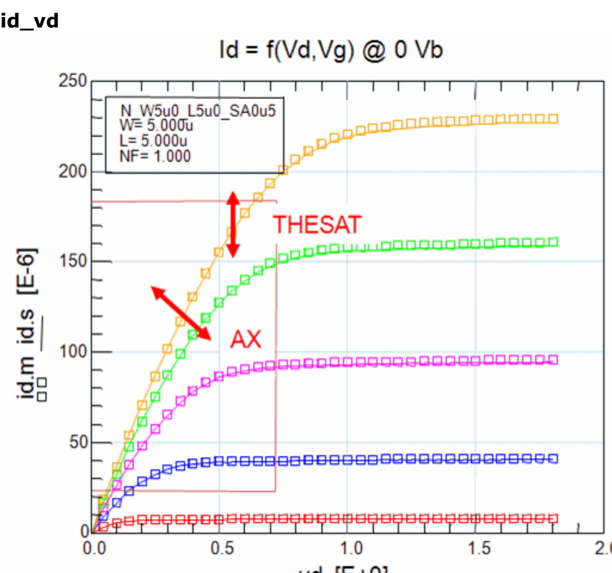
THEMU: (0..1.5..inf) Mobility reduction exponent at nominal reference temperature TR

id_vg Subthreshold



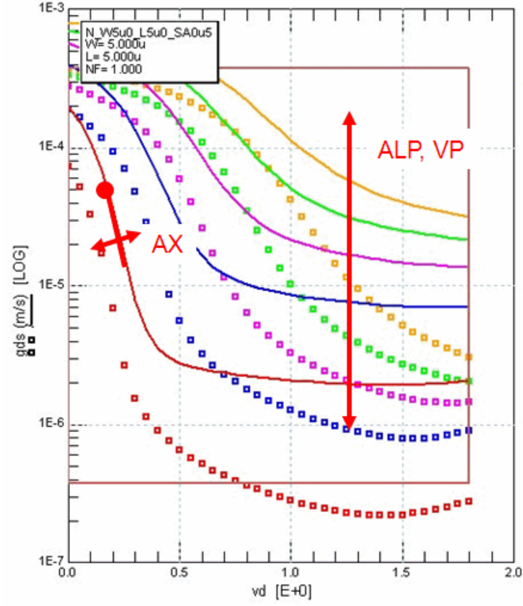
NEFF: (1e20..5E23..1e26) Effective substrate doping
DPHIB: (-inf..0..inf) Offset parameter for Surface Potential PHIB
CT: (0..0..inf) Interface states factor

id_vd



THESAT: (0..1..inf) Velocity saturation parameter at nominal reference temperature TR
AX: (2..3..inf) Linear/saturation transition factor

gds_vd

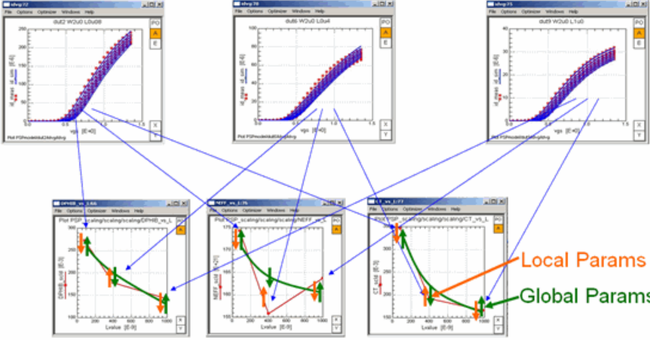


ALP: (0..10m..inf) channel-length-modulation (CLM) pre-factor
ALP1: (0..0..inf) CLM enhancement factor above threshold
ALP2: (0..0..inf) CLM enhancement factor below threshold
VP: (1e-10..50m..inf) CLM logarithm dependence factor
AX: (2..3..inf) Linear/saturation transition factor

PSP Parameter Extraction Flow for a single transistor (Large/Wide):

1. id_vg: NEFF, BETN, THEMU, MUE
2. id_vg substhreshold: NEFF, DPHIB, CT
3. id_vg: XCOR, CS, BETN, THEMU, MUE
4. fine-tuning: All above parameters on all id_vg
5. id_vd: THESAT, AX
6. gds: ALP, ALP1, ALP2, VP, AX
7. fine-tuning: All above parameters on id_vd, gds

Local <-> Global Parameters



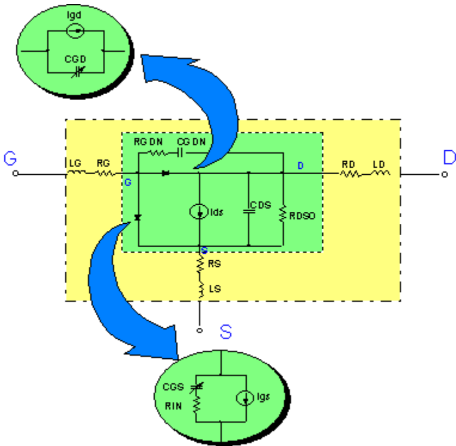
- For each geometry, local params are extracted
- The geometry scaling of the local params is plotted, and the global params are fitted to this curve.
- Simulate devices with the final global model.
- Finally, each device is fine tuned to simultaneously fit both, - the device behavior (local params) as well as - the scaling equation from the global model.

> While the modeling of the individual transistor (parameter mini-set) seems easy, the modeling of all different geometries (mini-sets) is much more complex because of model-specific dependencies between the miniset parameters.
>
> Additionally, the global PSP model (parameter maxi-set) has to handle measurement outliers and mini-set parameter outliers.

The Curtice Mesfet Model

An important model for GaAs-MESFETs is the Curtice model.
Walter R.Curtice developed it around 1980 with the RCA Labs in Princeton, New Jersey, USA.

Fig.1 depicts the equivalent schematic, consisting of the Curtice model itself, as well as additional enhancements for high frequency modeling using Lx and Rx, x = G,S,D.



Curtice model

The following table compiles the model parameters:

Common parameters for all Curtice model levels:

Parameter	Description	default	unit
RD, RG, RS	external Drain, Gate and Source resistance	1m	Ohm
VBR	reverse GD resp. GS breakdown voltage	0	V
RIN	series resist.to CGS (models HF input impedance)	1m	Ohm
CGSO	Gate-Drain junction capacitance with zero bias	0	F
CGDO	Gate-Drain junction capacitance with zero bias	0	F
VBI	built-in voltage (same as PB in Statz model)	1	V
CGS	interelectrode Gate-Source bias-independent cap	300f	F
CGD	interelectrode Gate-Drain bias-independent cap	30f	F
CDS	interelectrode Drain-Source bias-independent cap	300f	F
LD, LG, LS	external inductances	0	H
TAU	internal time delay from drain to source	10p	sec
A5	dependency of TAU versus Vds	1sec	/V

Diode model parameters:

Parameter	Description	default	unit
IS	diode saturation current	10f	A
N	diode emission coefficient	1	V
XTI	diode saturation current temp.coefficient	0	Ohm
EG	diode energy gap	1.11	eV

Piecewise linear model parameters:

Parameter	Description	default	unit
R1	approx. D-G breakdown resistance	0	Ohm
R1	resistance relating breakdown volt.to channel currents	0	Ohm
RF	eff.value of forw.bias G-S resistance	0	Ohm
MODEL	switch between level 1 and level 2 model	1 or 2	

Level 1 parameters:

Parameter	Description	default	unit
ALPHA	coeff. of Vds in tanh function for quadr.model	0.8	1/V
BETA	transcond.param. (same as BETA in Statz)	3m	A/V
LAMBDA	channel length modulation param for quadr.model	40m	1/V
VTO	quadr.model gate threshold volt.(same as VTO in Statz)	-6	V

Level 2 parameters:

Parameter	Description	default	unit
GAMMA	coeff. of Vds in tanh function for quadr.model	.5	1/V
BETA	transcond.param. (same as BETA in Statz)	100u	A/V
A0	0-order coeff. of v1 in Ids cubic equation	83.6m	A
A1	1-order coeff. of v1 in Ids cubic equation	15.28m	A/V
A2	2-order coeff. of v1 in Ids cubic equation	-1.293m	A/V²
A3	3-order coeff. of v1 in Ids cubic equation	-233u	A/V³
RDSO	Drain-Source leakage path resistance	1m	Ohm
VDSO	value of Vds at which A0 .. A3 are determined	6	V
VDSDC	the Vds bias at which RDSO, CDG and CGS were measured	0.8	V

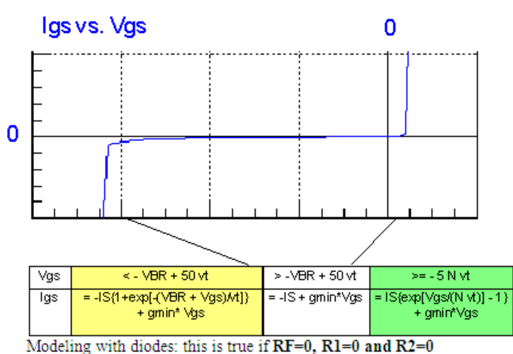
A special of this model is the fact that it behaves differently, depending on parameter values, which act as model switches. This means, the input characteristic can be modeled by using either diodes or simply a piece-wise linear function. Also, for the transfer curve, id = f(vg), a so-called 'quadratic' (Level_1) or 'cubic' (Level_2) modeling can be selected. The parameter BETA has a double meaning (see table above). Also, for AC modeling, the user can select between different modeling features. Fig. 2 lists these modeling specials.

Model level		
Parameter	Value	Result
MODEL	1	Quadratic equation used to model drain-source current.
	2	Cubic equation used to model drain-source current.
DC		
Parameter	Value	Result
RF	0	Gate-source junction DC current modeled with a diode.
	>0	Gate-source junction DC current modeled with piecewise linear currents..
R1	0	Gate-drain junction DC current modeled with a diode.
	>0	Gate-drain junction DC current modeled with piecewise linear currents.
AC		
Parameter	Value	Result
CGS	0	A nonlinear pn-junction capacitance model is used. CGSO and VBI are used to calculate CGS.
	>0	A constant CGS value is used.
CGD	0	A nonlinear pn-junction capacitance model is used. CGDO and VBI are used to calculate CGD.
	>0	A constant CGD value is used.
TAU	0	A5 * Vds is used as Tau.
	>0	A constant TAU value is used.

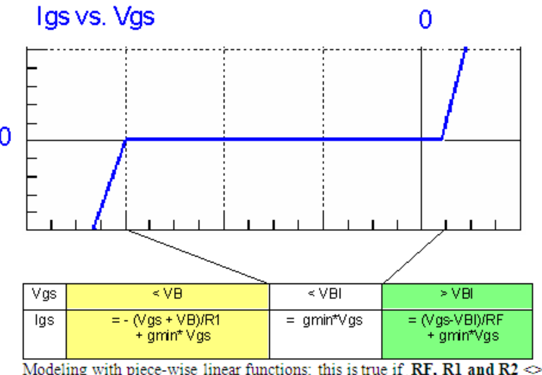
Overview about the different modeling possibilities of the Curtice model

DC Modeling

The two modeling possibilities for the input characteristics are given in fig.3.
Parameter determination is done following the common diode modeling techniques.

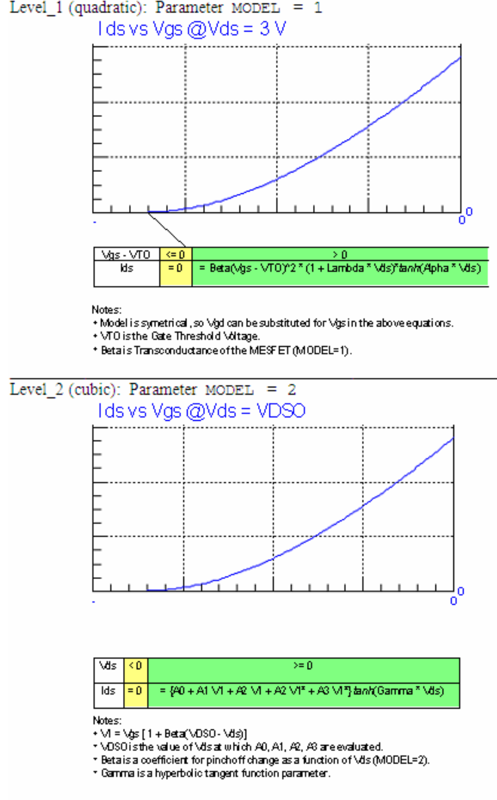


Modeling with diodes: this is true if **RF=0, R1=0 and R2=0**



Modeling with piece-wise linear functions: this is true if **RF, R1 and R2** <> 0

Modeling with diodes: this is true if RF=0, R1=0 and R2=0



Possibilities to model the DC semiconductor input characteristics

Modeling with piece-wise linear functions: this is true if RF, R1 and R2 <> 0

The transfer curve can either be modeled using Level_1 (quadratic model) or using Level_2 (cubic model). Fig. 4 puts both possibilities together.

Level_1 (quadratic): Parameter MODEL = 1

The determination of the parameters for Level_1 is similarly to the Statz model: first the modeling of the transfer curve with BETA , and then the output characteristics using the parameters ALPHA (low vDS) and LAMBDA (saturated output characteristics, 'Early effect'). For Level_2, the parameters A0 .. A3 can be obtained by deriving the transfer curve several times versus V1, GAMMA fits the output characteristics for low vDS and BETA models the slope in the saturation region.

Dynamic performance

Electronic charge is modeled by a constant charge between Drain and Source as well as a delay time TAU:

Qds = CDS * Vds + TAU * Ids

TAU stands for the internal time delay (carrier lifetime). This delay time is modeled with a constant value, if TAU is specified, i.e. non-zero. However, if TAU=0 and A5 is specified, SPICE calculates the delay TAU using:

TAU = A5 * Vds

Also, with the modeling of the Drain-Gate or Source-Gate capacitances, the Curtice model offers two possibilities. Either, when specifying CGS and CGD, what means they are non-zero, the charges are modeled linearly with

Qgs = CGS * Vgc where Vgc is the part of Vgs across CGS.
Qgd = CGD * Vgd
Qds = CDS * Vds

Otherwise, if CGS or CGD are set to zero each, the known space charge formulation is used:

$$CGS = \frac{CGSO}{\sqrt{1 - \frac{Vgs}{VBI}}} \qquad CGD = \frac{CGDO}{\sqrt{1 - \frac{Vds}{VBI}}}$$

with

$$CGx = \frac{dQGx}{dVgx}$$

x = S,D

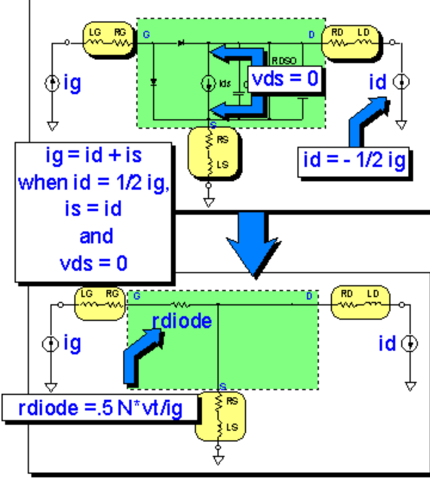
For positive bias voltages above Vgx > FC*VBI is:

$$CGx = \left(CGxO * \sqrt{1-FC}\right) * \left[1 + \frac{Vgx - FC*VBI}{2*VBI*(1-FC)}\right]$$

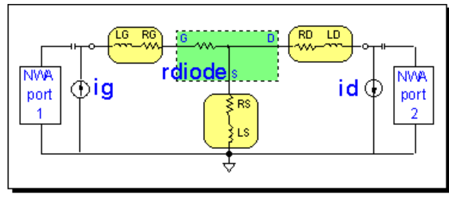
x = S,D

When interested only in the modeling around the operating point, the linear model of the capacitances and thus the determination of the parameters CGS, CGD and CDS is sufficient. This can be achieved by a S-parameter measurement at the operating point in question. Otherwise, the modeling is done using CV measurements of the capacitors CGS and CGD following the common CV parameter extraction rules (see chapter diode or bipolar transistor).

To determine the external parasitic inductor values Lx and Rx, the transistor is measured bias-overdriven (id = -ig/2). This means, the transistor is fully saturated and behaves like a good SHORT. Fig. 5a visualizes the measurement considerations and fig. 5b the parameter extraction.



Measurement technique to bias-overdrive the GaAs transistor. This allows the measurement of the series inductances and resistances



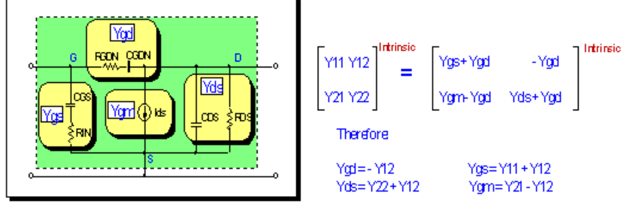
$$\begin{bmatrix} Z11 & Z12 \\ Z21 & Z22 \end{bmatrix} = \begin{bmatrix} RG + rdiode + RS + j\omega(LS + LG) & RS + j\omega LS \\ RS + j\omega LS & RD + RS + j\omega(LS + LD) \end{bmatrix}$$

Therefore:

RG = Real (Z11) - RS - rd   LG = Imag (Z11) - LS
RS = Real (Z12)   LS = Imag (Z12)
RD = Real (Z22) - RS   LD = Imag (Z22) - LS

how to determine the parasitic component values in bias-overdriven measurement mode

Finally, the remaining capacitors of the Curtice model are determined. This is done by de-embedding the measured S-parameters from the now known series components (Z matrix de-embedding). Afterwards, the result is converted into Y matrix parameters. For the linear modeling of the capacitors with Qij = Cij * Vij , fig.6 gives the details.



- CGS and RIN:

  $CGS = -1 / [\omega * imag(1/Ygs)]$

  $RIN = real(1/Ygs)$

- CGDN and RGDN:

  $CGDN = -1/[\omega * imag(1/Ygd)]$

  $RGDN = real(1/Ygd)$

- CDS and RDS:

  $CDS = imag(Yds)/\omega$

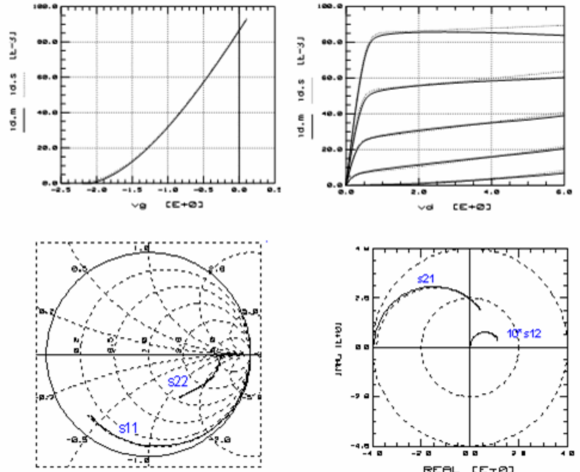
  $RDS = 1/real(Yds)$

- TAU or A5:

  $$TAU = \frac{phase[Ygm*(1+j\omega RIN*CGS)]}{\omega}$$

  $A5 = TAU / Vds$

TO conclude, fig. 7 shows the measurement and simulation result.



The modeling result: measurement (lines) and simulation (dotted)

Publications for the Curtice model

W.R.Curtice, GaAs MESFET Modeling and Nonlinear CAD, IEE Trans.on Microwave Theory and Techniques, vol.36, no.2, Feb.1988

W.R.Curtice, A MESFET Model for use in the Design of GaAs Integrated Circuits, IEEE Trans.Microwave Theory and Techniques, vol. MTT-28, pp.448-456, May 1980.

The Gummel-Poon model

Contents

- *AC Small Signal Modeling* (iccapmhb)
- *Appendices* (iccapmhb)
- *Gummel-Poon Bipolar Model* (iccapmhb)
- *Gummel-Poon CV Modeling* (iccapmhb)
- *Introduction to Gummel-Poon* (iccapmhb)
- *Limitations of Gummen Poon Model* (iccapmhb)
- *Modeling the Base Resistor* (iccapmhb)
- *Modeling the Parameter XCJC* (iccapmhb)
- *Modeling the Resistors* (iccapmhb)
- *Modeling the Temperature Effects* (iccapmhb)
- *Non Linear DC Modeling* (iccapmhb)
- *Preface* (iccapmhb)
- *Publications* (iccapmhb)
- *Transit Time Modeling* (iccapmhb)

AC Small Signal Modeling

Overview

We are now ready to consider the basics of modeling for frequencies higher than 100MHz. It is assumed that the measurements have been made on the pure semiconductor device without being affected by packaging parasitics, bond pads or other parasitics. If this is not possible, de-embedding techniques have to be applied. This means to find the proper semiconductor behavior out of the distorted measurement by de-embedding.

It should also be mentioned that the probe pins have to have an excellent frequency performance within the transistor measurement frequency range. Once again the key to meaningful AC measurements and thus modeling is a good network analyzer calibration with excellent standards and a correctly defined calibration kit data in the network analyzer.

If you need additional support for de-embedding, calibration and for better understanding S-parameters, please refer to the additional toolkits for IC-CAP. Please contact the author.

Let us go first for the AC small signal equivalent schematic. It can be derived from fig.2b of the introduction chapter as a linearization at each bias point of the transistor.

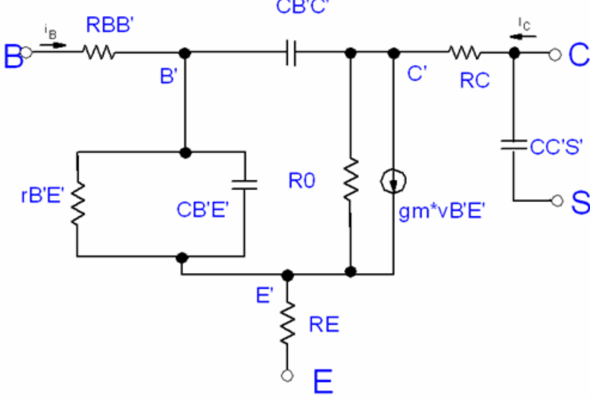


Fig.AC-1: AC small signal schematic of the bipolar transistor
NOTE: XCJC effect neglected.

The following equations give the values of the internal elements in fig.AC-1.
They represent the linearized DC- and CV-equations at the DC operating point.

gB'E'
From equ.(B) in the introductory chapter we get from the derivative of iB versus vB'E':
$[A^2/Hz]$
where the second term can most often be neglected for operating points of iC above 1 uA.

g0 The output conductance is:
$0.2nA/\sqrt{Hz}$

CB'E' or CPI

including the delay time effect modeled by TFF is given in equ.(P) and (R3) of the introductory chapter for the particular operating point voltages.
As a first order estimation, CB'E' simplifies to
$[V^2/Hz]$

while
CB'C' or CMU (MU or µ stands for 'mutual'

simplifies because of vBC < 0 at forward operation to:
$E_{nv}[V/\sqrt{Hz}]$

gm
The transconductance gm finally is using equ. (H)
A^2/Hz

Appendices

Linear Curve Fitting: Regression Analysis

IC-CAP File:
$ICCAP_ROOT/examples/demo_features/4extraction/basic_PEL_routines/1fit_lin.mdl

Let's assume we made 'N' measurements y_i at the stimulating points x_i. I.e. we obtained the array (x_i , y_i). Subsequently, these measured values were plotted.

A curve Y (x) shall be fitted to this array of measured data points using least square curve fitting technique. Referring to an individual measurement point, the fitting error is:

$$E_i = Y(x_i) - y_i$$

and for all data points:

$$E = \sum_{i=1}^{N} E_i^2 = \sum_{i=1}^{N} \left[Y(x_i) - y_i \right]^2$$

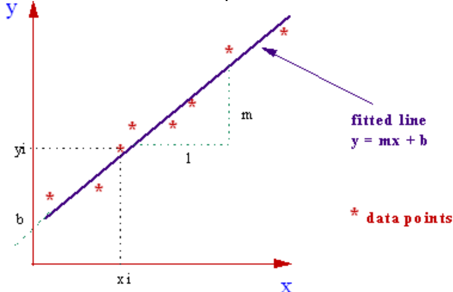
This error shall be minimized.

The fitting will be done by varying the coefficients of the fitting curve of equation above. The minimum of the total error E depends on the values of these coefficients. This means, we have to differentiate E partially versus the curve coefficients and to set the results to zero. We obtain a system of equations, solve it, and get the values of the coefficients for a best curve fit. This is known as regression analysis.

> **Note**
> This regression analysis is simple for a straight line fit. But in general, measured data is non-linear. Unfortunately, a non-linear regression analysis can be quite complicated. This problem can be solved if we use a suitable transformation on the measured data. This means that the measured data is transformed to a linear context between the yi- and the xi-values. As will be seen in the diode example later, this is a pretty smart way to get the curve fitting parameters easily without much calculations.

Provided we have got an array of N measured data points of the form (x_i , y_i).

A linear curve with the equation:
y (x) = m x + b
shall be fitted to these points. This situation is depicted below.



Linear regression applied to measurement points

The error of the i-th measurement is:
Ei = [m xi + b] - yi
Using the least means square method yields:

$$E = \sum_{i=1}^{N} Ei^2 = \sum_{i=1}^{N} [m\,xi + b - yi]^2 \overset{!}{=} \text{Minimum}$$

Partial differentiation versus slope 'm' gives:

$$2 \sum_{i=1}^{N} [m\,xi + b - yi]\, xi = 0$$

and versus y-intersect 'b':

$$2 \sum_{i=1}^{N} [m\,xi + b - yi] = 0$$

We obtain (from previous equations) after a re-arrangement:

$$m \sum_{i=1}^{N} xi^2 + b \sum_{i=1}^{N} xi = \sum_{i=1}^{N} yi\, xi$$

and from previous equations:

$$m \sum_{i=1}^{N} xi + N\,b = \sum_{i=1}^{N} yi$$

Multiplying by -N and xi and adding these two equations allows the elimination of the coefficient 'b', and we can separate the slope 'm':

$$m \left[\left(\sum_{i=1}^{N} xi \right)^2 - N \sum_{i=1}^{N} xi^2 \right] = \sum_{i=1}^{N} xi \sum_{i=1}^{N} yi - N \sum_{i=1}^{N} xi\, yi$$

or

and from previous equations for the y-intersect 'b':

$$b = \left[\sum_{i=1}^{N} yi - m \sum_{i=1}^{N} xi \right] / N$$

with 'm' according to previous equations:

With previous equations, we determined the values of the two coefficients of the linear curve which fits best into the 'cloud' of measured data.
Finally, a curve fitting quality factor r^2 is defined. Its value ranges from ($0 < r^2 < 1$). The closer it is to 1, the better is the fit of the linear curve.

Verifying the quality of extraction routines

Let's start with the first problem: assumed, we would know the parameters to-be-extracted in advance, it would be easy to check the validity of the extraction routines!

Using IC-CAP, it is simple to perform such a check. The trick is to 'synthesize' quasi-measured data out of a set of parameters and to apply then the extraction routines to these data. This can be done as follows:

1. Define a measurement setup in IC-CAP, for which the extraction routines shall be tested. Example: an output characteristic for an Early-voltage extraction.
2. Select a 'typical' set of parameters (no default values like 'zero' or 'infinite', but instead real realistic values!)
3. Change the 'Output' data type to 'S' (simulated only). The array behind that output is no longer one-dimensional, i.e. no measurement data any more, but only simulation data.
4. Simulate this setups using these parameter values.
5. Change the 'Output' data type back to type 'B'.
   IC-CAP doubles now the data field to measurement and simulation data. Thus the simulated data of step 4. is now converted to measured data!
6. Reset the model parameters by clicking 'Reset to Defaults' and simulate the setup using the default parameters.
7. Apply the extraction routine-under-test and check the quality of the extracted parameters.

Provided we get the parameter values back within a good tolerance, we can be sure that the extraction works correctly. If we now apply the extraction to real-world measured data, we should obtain the right parameters. This is true if the measured data have the same shape like the model equations! If not, we might have to choose another model or go for subcircuit modeling. And this leads us to the second part of this paper:

Direct visual parameter extraction of BF, ISE and NE

or: Checking if the model can fit the measured data at all

Using IC-CAP for the extraction of model parameters offers a lot of flexibility in terms of creating user-defined models and implementing the corresponding extraction routines.

But when developing a new extraction strategy, we may run into two major problems:
do the routines extract the parameters correctly?
and is the model able to fit the measured device at all?

This appendix proposes a method that allows us to

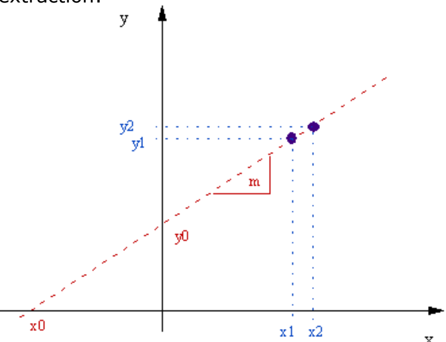
- verify the quality of the extraction routines
- check the fitting of our model to the measured data
  and to perform the parameter extraction simultaneously.

This model-fit-check method is also called 'Direct Visual Parameter Extraction'

The point is to transform the measured data to a domain where the parameter itself can be displayed against the measurement stimulus. As an example, we know that the x-intersect of lines fitted to an output characteristic of a bipolar transistor should hit always the same point, the Early voltage. If we apply an IC-CAP PEL (parameter extraction language) program to calculate the x-intersect of a line that is fitted to two adjacent measured points, and if we display the result of this operation versus the collector voltage (first order sweep), we will obtain a plot of the 'equivalent' Early-voltages of adjacent measurement points.

The advantage of using this method is that we can see clearly, if the model is able to fit the measured data at all. We only have to check if there is a flat region in the transformed data domain or not. If it is there, we can extract the parameter very simply by calculating the mean value of the flat region. And we know at the same time, in which range the parameter is dominant and can therefore be used for fine-tuning with the optimizer. If there is no flat range, the model cannot fit the measured data. We could vary the parameter as much as we like and would not achieve a fit of the simulated to the measured data!

For the application of this method, we start with some basic equations that refer to figure below which defines some of the basic equations for the direct visual parameter extraction:



Assumed we have:
y = m*x + y0
where m: slope & y0: y-intersect
Then it is:
x0 = -b/m
and:

$$m = \frac{y2 - y1}{x2 - x1}$$

Some more useful formulas to calculate:

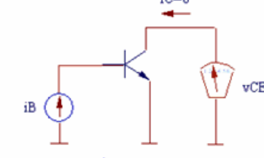
	x-intersect x0	y-intersect y0:
starting with	$\frac{y2}{-x0+x2} = \frac{y1}{-x0+x1}$	$\frac{y2-y0}{x2} = \frac{y1-y0}{x1}$
we get:	$x0 = \frac{x1y2 - x2y1}{y2 - y1}$	$y0 = \frac{x2y1 - x1y2}{x2 - x1}$

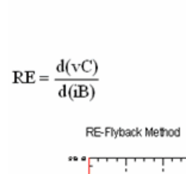
These equations are implemented to the model files of directory "visu_n_extr".

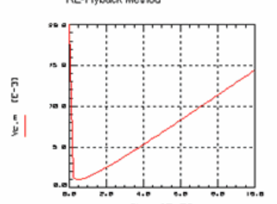
The following plots give some examples on how to apply this idea to the parameter extraction of a bipolar transistor using the Gummel-Poon model. It should be mentioned that this method can be applied to all the parameters of this model, as well as to other models like Statz, Curtice, BSIM etc.

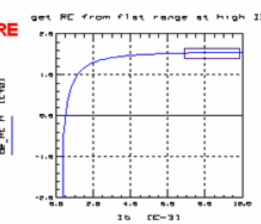
RE

measurement setup:



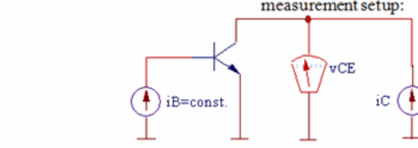


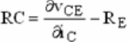


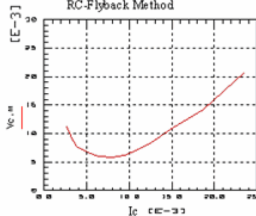


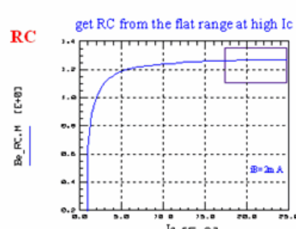
RE is the derivative of the flyback measurement

the "visualized" RE is the y-intersect of two adjacent data points from that flyback plot.



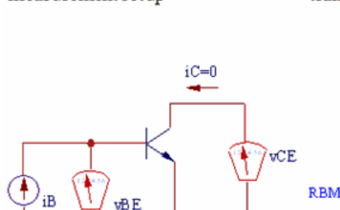


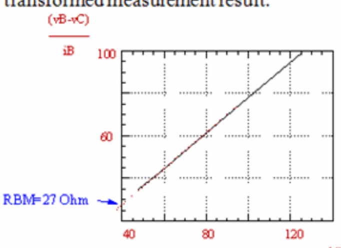




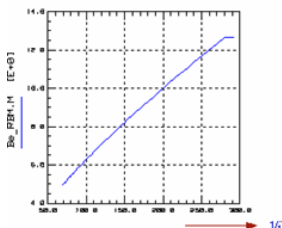
RC can be calculated from measured vC at stimulated iB and iC
the "visualized" RC is the y-intersect of two adjacent data points

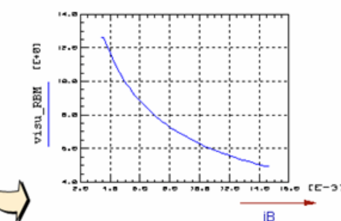
RBM





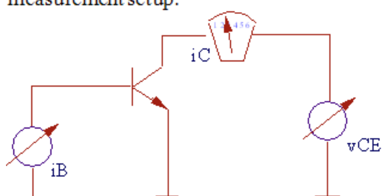
Next, the y-intersects of two adjacent data points in this transformed data plot are drawn versus 1/iB again. From the curve above, we might expect a flat range, but in reality we get a curve like the left one below: In our example, RBM ranges from 5 to 12 Ohm. In this case, we might display the visualized data versus iB, and interpret the result similar to figures in the rBB chapter in order to estimate not only RBM but also IRB and RB (see plot on the right).





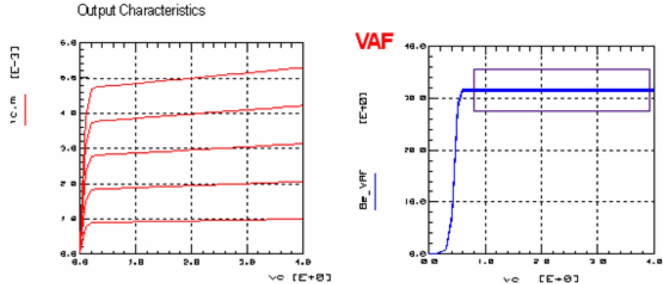
The figure above represents the visual extraction of RBM and to estimate IRB and RB possibly.

VAF



Determine VAF out of the x-intersect of a line through two adjacent measurement points:

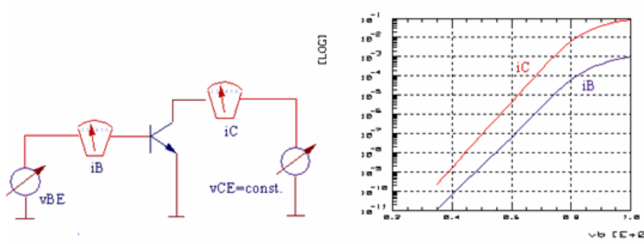
```
X=vc
Y=ic.M
i=1
WHILE i < SIZE(Y)-1
 VAF[i]=ABS(X[i+1]*Y[i-1]-X[i-1]*Y[i+1])//(Y[i+1]-Y[i-1])
 i = i + 1
END WHILE
```



Early voltage is calculated from the x-intersect of a line through two adjacent data points in the above figures.

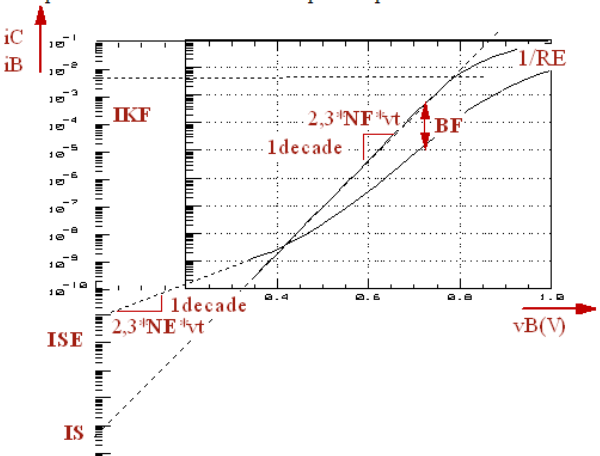
Let's study also the application of the "visual" method for the Gummel-Poon plot:

measurement setup: measurement result:



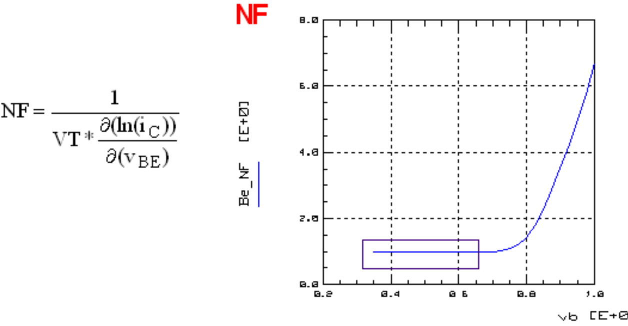
interpretation of the Gummel-Poon plot for parameter extraction:

interpretation of the Gummel-Poon plot for parameter extraction:



NF can be obtained from the inverse derivative of the iC-Gummel-plot:

click around the flat range

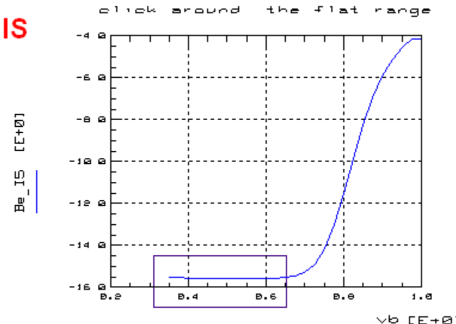


$$NF = \frac{1}{VT * \frac{\partial(\ln(i_C))}{\partial(v_{BE})}}$$

IS

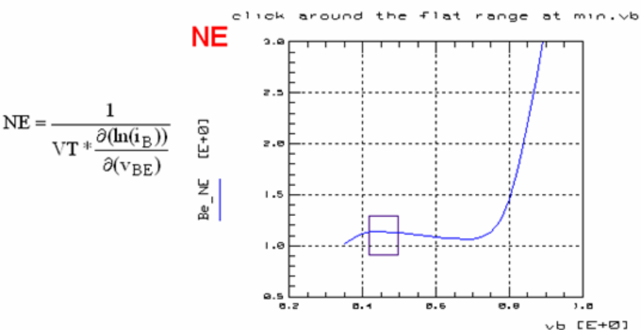
```
!Determine IS out of the y-intersect of a line through two adjacent measurement points:
X=vb
Y=log10(ic.M)
i=1
WHILE i < SIZE(Y)-1
  ISE[i]=(X[i+1]*Y[i-1]-X[i-1]*Y[i+1])//(X[i+1]-X[i-1])
  i = i + 1
END WHILE
```

IS from the y-intersect of a line through two adjacent measurement points:



NE

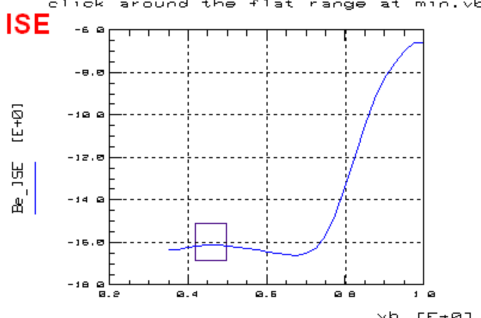
The following figure shows NE as obtained from the inverse derivative of the iB-Gummel-plot



ISE

Determine ISE out of the y-intersect of a line through two adjacent measurement points:

```
X=vb
Y=log10(ib.M)
i=1
WHILE i < SIZE(Y)-1
  ISE[i]=(X[i+1]*Y[i-1]-X[i-1]*Y[i+1])//(X[i+1]-X[i-1])
  i = i + 1
END WHILE
```



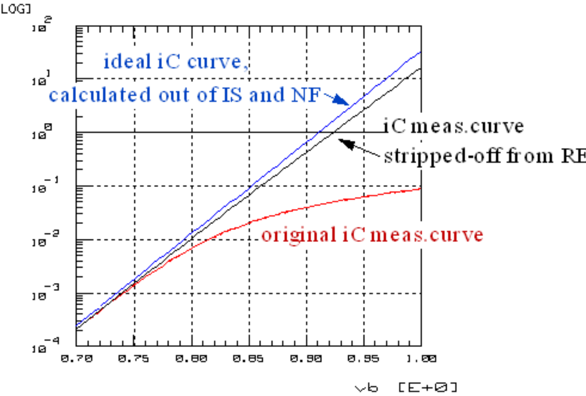
As can be expected from an inspection of fig_6, the measured data does not show a big recombination effect on the ib(vbe) curve. This means that the parameters ISE and NE will not contribute a lot to the curve fitting and may be difficult to extract. This is exactly the meaning of the transformation results in above figures.

IKF

In order to be able to visualize the effect of IKF (see fig.6 in the introduction chapter), we have to 'strip-off' the effect of RE on the Gummel-Poon iC curve:

```
! this extraction assumes that RE is already extracted properly. It eliminates the effect of RE on
the Gummel-Plot, such that IKF can be extracted from the 'knee' of the 'stripped-off' ic-Gummel-
plot!
!strip-off the effect of RE on the Gummel-plot
!iCmeas = IS*exp(vBEint/(vt*NF) = IS*exp((vBEext-(iC+iB)*RE-iB*RB)/(vt*NF))
! i.e. multiplying iCmeas by exp((vBEext-(iC+iB)*RE-iB*RB)/(vt*NF))
! will give iC without the influence of RE and RB !!
X  =ABS(vb)
Yic=ABS(ic.M)
Yib=ABS(ib.M)
! calculate the stripped-off Gummel Plot
```

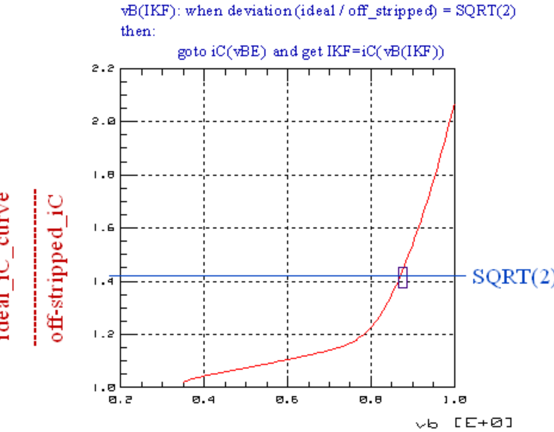
```
tmp = Yic*(exp((MAIN.RE*(Yic+Yib)+MAIN.RBM*Yib)//(MAIN.NF*VT)))

! calculate the ideal ic Gummel curve
!  as a reference
tmp1= MAIN.IS*(exp(X//(MAIN.NF*VT)))

RETURN tmp+j*tmp1
```



Calculating IKF

```
X = ABS(vb)
Y = IMAG(strip_off_RE)*(REAL(strip_off_RE))^-1
Y1 = REAL(strip_off_RE)

i=SIZE(Y)-1
index = 0
WHILE i > 0
  IF Y[i] < SQRT(2) THEN
    index = i
    i = 0
  END IF
  i = i - 1
END WHILE

PRINT index
! calculuate IKF out of iC(vBE) at that index
MAIN.IKF = Y1[index]
```

Calculating IKF from the ratio ideal_iC_curve / stripped-off iC:

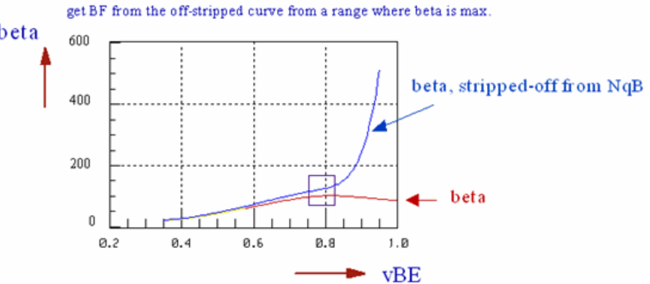


BF

The idea is to strip-off the base charge effect (NqB) from the beta plot and then apply the often cited "max. beta" extraction for BF:

```
!Calculation  of  ß = iC/iB
vbe  = vb-ve
vbc  = vb-vc
beta = ABS(ic) / ABS(ib)
! with some simplification is:
!     beta ~ BF/nqb
!          nqb  = q1//2*(1+SQRT(1+4*q2))
!    with   q1  ~ (1-vbe//VAR-vbc/VAF))^-1
!           q2  ~ IS//IKF*exp(vbe//(NF*VT))
!           VT :  temp.voltage (a model variable)

q1 = (1-vbe//MAIN.VAR-vbc//MAIN.VAF)^-1
q2 = MAIN.IS//MAIN.IKF*exp(vbe//(MAIN.NF*VT))
```

```
nqb=q1//2*(1+SQRT(1+4*q2))

BF=ic//ib*nqb
RETURN BF
```



This method of direct visual parameter extraction can be applied to all Gummel-Poon parameters. For more examples see the IC-CAP files in directory "visu_n_extr".

Let us conclude with an example of a PEL program used for this method. It is the VAF transform. The program performs either the data transformation (SWITCH=1) or performs the parameter extraction (SWITCH=-1):

```
X=vc                                            ! link to stimulus data
Y=ic.M                                          ! link to measured data
Y = SMOOTH3(Y)                                  ! smooth measured data
SWITCH = -SWITCH                                ! a model variable
LINPUT "data transform (1) or extraction(-1) ?",SWITCH,dummy
SWITCH = dummy
PRINT "calculating the x_intersect ..."
tmp = Y ! a lousy array declaration
i=1
WHILE i < SIZE(Y)-1
 tmp[i]=ABS(X[i+1]*Y[i-1]-X[i-1]*Y[i+1])//(Y[i+1]-Y[i-1])
 i = i + 1
END WHILE
tmp[0] = tmp[1]                             ! watch-out for the array bounds
tmp[SIZE(tmp)-1] = tmp[SIZE(tmp)-2]
IF SWITCH == -1 THEN                  ! sum-up all parameters within the box
 i = 0
 N = 0
 result = 0
 WHILE i < SIZE(X)
    IF X[i] > X_LOW AND X[i] < X_HIGH THEN
      result = result + tmp.M[i]
      N = N+1
    END IF
      i = i+1
 END WHILE
 MAIN.VAF = result//N                              ! export parameter value
 PRINT "MAIN.VAF = ",MAIN.VAF
ELSE
 MENU_FUNC("Visu_va","Display Plot")
 LINPUT "click a box and re-run this transform to extract VAF",dummy
 RETURN tmp
END IF
```

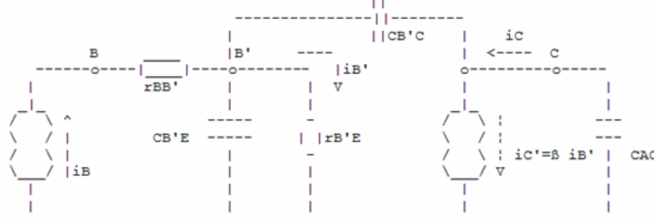
Note
Included in your "bipolar toolkit" is a directory called "visu_n_extract" that contains
IC-CAP model files with suggestions on direct visual extraction for most of the Gummel-Poon parameters.

Literature:

HP-EESOF, Characterization Solutions Journal, Spring 1996.

Calculation of h21 of the Gummel-Poon Model

We start from the small signal equival.schematic of the following figure, using some small simplifications (Collector is AC-wise shorted, since h21 = iC / iB for RL = 0):



The equations:

Input:
vB = iB rBB' + vB' _ iB = vB / rBB' - vB' / rBB'

Output:
0 = (iC - ß iB') / pCB'C + vB'

Internal 1:
vB' = (iC - ß iB' + iB) rB'E
using rB'E = rB'E // CB'E
Internal 2:
iB' = vB' / rB'E

This gives the matrix:

```
    iC          vB           iB'           vB'      |
-------------------------------------------------------- ---------
                                                    |
    0       1 / rBB'          0        -  1 / rBB'  |  iB
                                                    |
 1 / pCB'C       0       - ß / pCB'C        1       |   0
                                                    |
  -rB'E          0          ß rB'E          1       | iB rB'E
                                                    |
    0            0           -1          1 / rB'E   |   0
```

Solved for iC:

```
          | iB      1 / rBB'         0         -  1 / rBB'|
          |                                               |
          |  0          0        - ß / pCB'C        1      |
          |                                               |
          |iB rB'E      0          ß rB'E           1      |
          |                                               |
          |  0          0           -1          1 / rB'E  |
iC   = ------------------------------------------------------
         |   0       1 / rBB'         0         -  1 / rBB'|
         |                                               |
         |1 / pCB'C      0       - ß / pCB'C        1      |
         |                                               |
         | -rB'E         0          ß rB'E          1      |
         |                                               |
         |   0           0           -1          1 / rB'E  |
```

Solving for the 2nd column of the nominator and the 2nd column of the denominator:

```
                   |  0       - ß / pCB'C       1      |
                   |                                   |
        -  1 / rBB' |iB rB'E      ß rB'E          1      |
                   |                                   |
                   |  0           -1          1 / rB'E |
iC   = ----------------------------------------------------
                  |1 / pCB'C     - ß / pCB'C       1      |
                  |                                      |
       -  1 / rBB' | -rB'E          ß rB'E          1      |
                  |                                      |
                  |   0             -1          1 / rB'E |
```

Now solving for the 1st column of the nominator and the 3rd row of the denominator:

```
                      |  - ß / pCB'C       1      |
          - iB rB'E   |                           |
                      |      -1          1 / rB'E |
iC   = ------------------------------------------------------------
            |1 / pCB'C    1|              |1 / pCB'C   - ß / pCB'C|
     -(-1)  |              |  + 1 / rB'E  |                       |
            | -rB'E       1|              | -rB'E         ß rB'E  |
```

gives for h21:

```
                                      ß
                 - rB'E { -  -------------- - (-1)}
         iC                    pCB'C rB'E
 h21 = ----  =  ---------------------------------------------------
         iB         1                     1      ß rB'E    ß rB'E
                 ------  + rB'E  + ------- { -------- - -------- }
                  pCB'C                  rB'E     pCB'C      pCB'C

                 ß / rB'E  -  pCB'C
             =  ---------------------
                 1 / rB'E  +  pCB'C
```

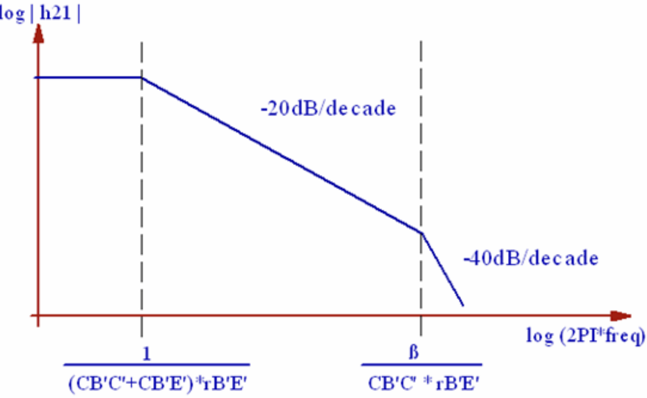
Finally re-substituting
1 / rB'E = 1 / rB'E + pCB'E
yields:

```
        ß /rB'E    -       pCB'C
 h21 =  -----------------------------
        1 / rB'E  +  p(CB'C + CB'E)
```

or

```
        ß  -      pCB'C rB'E
h21 =  ------------------------------
        1  + p(CB'C + CB'E) rB'E
```

what is depicted below:



Now we are ready to calculate the transit frequency fT:

from previous equations, follows for h21 = 1 :

$$1 + 4\,PI^2\, fT^2\, (CB'C + CB'E)^2\, r^2B'E \;=\; ß^2 \;-\; 4\,PI^2\, fT^2\, c^2B'C\, r^2B'E$$

or:

$$f_T = \frac{g_{B'E}}{2PI}\sqrt{\frac{ß^2 - 1}{(C_{B'C} + C_{B'E})^2 - C_{B'C}^2}}$$

with:

$$gB'E' \;=\; \frac{diB'}{dvB'E} \;\sim\; \frac{iB'}{NF\ VT}$$

$$CB'C \;\sim\; CSBC(vBC)$$

and:

$$CB'E \;\sim\; \frac{TFF}{NF\ \ VT}\; iC$$

Simplification

In order to keep things a little simpler for parameter extraction, previous equation is modified a bit, neglecting the zero (at high frequencies) against the pole(low frequencies):

$$h21 \sim \frac{ß}{1 \;+\; p(CB'C + CB'E)\; rB'E}$$

Calculating again the transit frequency for this simplified h21 yields:

$$ß^2 \;=\; 1 + 4\, PI^2\, f^2T1\text{-}pole\, (CB'C + CB'E)^2\, r^2B'E$$

$$f_{T1-pole} = \frac{1}{2PI}\sqrt{\frac{ß^2 - 1}{(C_{B'C} + C_{B'E})^2\, r_{B'E}^2}}$$

$$\begin{matrix} CB'E > CB'C \\ \sim \\ ß > 1 \end{matrix} \qquad \frac{ß}{2\ PI\ CB'E\ rB'E}$$

$$\begin{matrix} (9) \\ \sim \\ (7) \end{matrix} \qquad \frac{ß}{2\ PI\ \frac{TFF}{NF\ VT}\ iC\ \ \frac{NF\ VT}{iB'}}$$

since ß = iC / iB', we get the pretty simple form

$$fT1\text{-}pole \quad \sim \quad \frac{1}{2\ PI\ TFF}$$

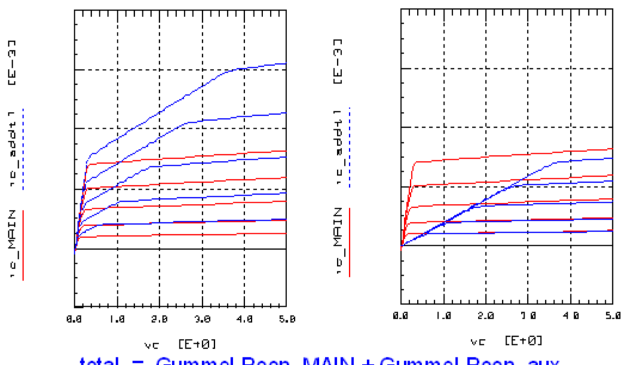
Or solved for TFF:

$$TFF \;=\; \frac{1}{2\ PI\ fT1\text{-}pole}$$

what is the well-known formula for the TFF parameter extraction.

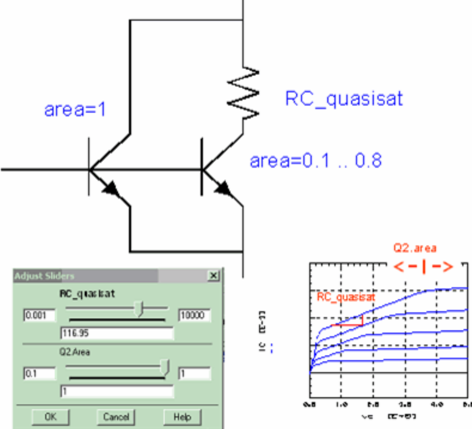
Modeling the output characteristics quasi-saturation effect with Gummel-Poon

Modeling Example:
quasisaturation modeling with Gummel-Poon



The Gummel-Poon model is known for its poor fitting capabilities in the quasi-saturation range. However, an output characteristics of a bipolar transistor can be split into two overlying Collector currents: a first one (MAIN transistor) responsible for the output slope in saturation, and another one, added, for modeling the quasi-saturation.

In this case, the trick lies in the fact to use the same model parameter set for both transistors. However, the two transistors are in parallel with a series Collector resistor for the quasisaturation range. See the next slide for the schematic:



In this example: increasing RC_quasisat affects the slope in quasisaturation, while Q2.area determines the transition to the non-saturated range of iC.

See the previous slide to better understand the effect of these 2 parameters.

> **Note**
> Of course, during the Gummel-Poon plot modeling,, this scenario has to be considered as well.
> Therefore, it is smart to set the area of the quasisaturation transistor to a variable , e.g. 'area', ranging from 10m to 1, and to set the area of the MAIN transistor to (1-area).
> This way, the Gummel-Poon plot modeling is not affected outside the quasisaturation range !

Gummel-Poon Bipolar Model

Structure of this Manual

- Introduction
  - Operating Modes of the Bipolar Transistor
  - The Equivalent Schematic and the Formulas of the SPICE Gummel-Poon Model
  - A Listing of the Gummel-Poon Parameters
  - A Quick Tutorial on the Gummel-Poon Parameter Extractions
  - Proposed Extraction Strategy
- CV Modeling
  - Extraction of CJE, VJE, MJE, as well as CJC, VJC, MJC
- Parasitic Resistor Mmodeling
  - Extraction of RE
  - Extraction of RC
  - Extraction of RBM
- Nonlinear DC Modeling
  - Extraction of VAR and VAF
  - Extraction of IS and NF
  - Extraction of BF , ISE and NE
  - Extraction of IKF
  - Reverse Parameters NR, BR, ISC, NC and IKR
- AC Small Signal Modeling, Parameter Extraction
  - Extraction of RB, IRB and RBM
  - Extraction of TF, ITF, and XTF
  - Extraction of VTF
  - Extraction of PTF
  - Extraction of TR
  - Modeling of XCJC
- Temperature Effects
- Model Limitations
- Appendices
  - Linear Curve Fitting: Regression Analysis
  - About the Modeling Dilemma
  - Verifying the Quality of Extraction Routines
  - Direct Visual Parameter Extraction of BF , ISE and NE
  - Calculation of h21 of the Gummel-Poon Model
- Publications

This product has been developed to meet the local demands of European IC-CAP users for more technical background information on extraction techniques and for the availability of extraction source code.

Published for the first time in 1990, it has been updated since then several times.

It is part of a series of supplementary modeling toolkits for the IC-CAP users. These products feature source code and detailed technical description of the extraction routines. Please contact the author for further information.

The author would like to thank the many users for valuable inputs, and is hoping for fruitful discussions also in the future.

About This Manual

This manual is intended to explain the basics of modeling a bipolar transistor using the Gummel-Poon model as it is implemented in the simulation program SPICE of the University of California Berkeley (UCB) /see publication list/.
It is part of the Gummel-Poon Bipolar Model Parameter Extraction Toolkit.

This toolkit includes the IC-CAP model file
GP_CLASSIC_NPN.mdl the MASTER model file
which is described in this manual

and featuring the data management features of IC-CAP 5.x, i.e. separating measurements from extractions:

NPN_MEAS_MASTER.mdl a master file for measurement
GP_EXTRACT_NPN.mdl a master file for modeling

As well as many other IC-CAP model files covering topics like:

- Model parameter extraction using the tuner feature
- Direct visual parameter extractions
- Alternate modeling methods for DC- CV- and RF-parameters.
- Bipolar transistor modeling including the parasitic transistor.

> After you have become familiar with the modeling procedure itself, i.e. file GP_CLASSIC_NPN.mdl, you are encouraged to split the modeling into 2 parts: separate measurements and separate extraction strategy. In this case, all measurements are performed using the file NPN_MEAS_MASTER.mdl. Then, the data are exported into IC-CAP mdm files (ASCII files) and imported into the master extraction file GP_EXTRACT_NPN.mdl for extraction. This method allows to improve continuously the extraction strategy file, independent of the measurement data.

> NOTE:
> This manual and the underlying IC-CAP model file GP_CLASSIC_NPN.mdl are intended to explain the basics of the Gummel-Poon modeling. Therefore, it covers the classical Gummel-Poon model without enhancements for also modeling the parasitic transistor.
> However, as stated above, such model files are included in the file sets of this toolkit. Please see the README macros in these IC-CAP model files for more details.
>
> You are also invited to get in contact with the author for assistance with such modeling problems.

The IC-CAP model file "GP_CLASSIC_NPN.mdl" features
The extractions are written using PEL (parameter extraction language) and are open to the user. They can be easily modified to meet specific user needs.

Subcircuit model description, open for user enhancements (HF modeling, parasitic pnp etc.).

All transistor pins are connected to SMUs for flexible measurements

The transistor output characteristic and S-parameter measurements use a Base current stimulus rather than a Base-Emitter voltage in order to avoid 1st order thermal effects being visible. However, self-heating might be present and affect the Gummel plots in the ohmic range.
See also the file GP_MEAS_MASTER.mdl

Organization of the chapters in this manual
There are 5 main chapters, which explain how to determine the model parameters from CV (capacitance versus voltage), then parasitic ohmic resistors, and DC, to finally high frequency measurements using network analyzers.

More chapters cover side aspects of bipolar modeling.

The individual chapters follow always this scheme:

- Explanation of the parameter-dependent measurement setup
- Explanation of the mathematical basics for the parameter extraction
- Explanation of the parameter extraction
- Explanation about how to use the IC-CAP file.

Gummel-Poon CV Modeling

Since the CV parameters are like for most bipolar models, independent of the other model parameters, they are usually extracted first. We follow this idea and begin with the CV modeling, followed then by the parasitic resistor modeling and the non-linear DC curves. Finally, the S-parameter measurements are modeled.

The Gummel-Poon Capacitor Equations

Provided that: vBE = vB'E' and vBC = vB'C', the capacitors in the Gummel-Poon model given in the introduction chapter with equations (O)... (S) are:

$$CBC = \underbrace{CSBC}_{\text{space charge}} + \underbrace{CDBC}_{\text{diffusion cap.}}$$

$$= \frac{CjC}{[1 - vBC / VJC]^{M_{JC}}} + TR \frac{diC}{dvBC}$$

$$\overset{(H)}{=} \frac{CjC}{[1 - vBC / VJC]^{M_{JC}}} + \frac{TR}{NR\ VT} \frac{IS}{NqB} \exp\left(\frac{vBC}{NR\ VT}\right) \qquad (CV\text{-}1)$$

with equation (H) from the introduction chapter.

$$CBE = \underbrace{CSBE}_{\text{space charge}} + \underbrace{CDBE}_{\text{diffusion cap.}}$$

$$= CSBE + \frac{dQ_F}{dv_{BE}} \qquad \text{with QF equals the total forward charge}$$

$$\sim \frac{CjE}{[1 - vBE / VJE]^{M_{JE}}} + TFF \frac{diC}{dvBE}$$

$$\overset{(H)}{=} \frac{CjE}{[1 - vBE / VJE]^{M_{JE}}} + \frac{TFF}{NF\ VT} \frac{IS}{NqB} \exp\left(\frac{vBE}{NF\ VT}\right) \qquad (CV\text{-}2)$$

again with equation (H1) from the introduction chapter and additionally with,

$$TFF = TF \left\{ 1 + XTF \left[\frac{if}{if + ITF} \right]^2 \exp\left[\frac{vBC}{1{,}44\ VTF} \right] \right\} \qquad (CV\text{-}3)$$

and the ideal Collector current if from (C).
CSBi models the space charge and CDBi the diffusion capacitance between Base and Emitter or Base and Collector respectively. vBE and vBC are the stimulating voltages.

Note
For CBE, and its diffusion capacitor part CDBE, the exact definition is (private communication with D.Celi, ST Crolles, France)

$$C_{DBE} = \frac{dQ_F}{dv_{BE}} = \frac{dQ_F}{di_C} \cdot \frac{di_C}{dv_{BE}} = TF_{ac} \cdot gm$$

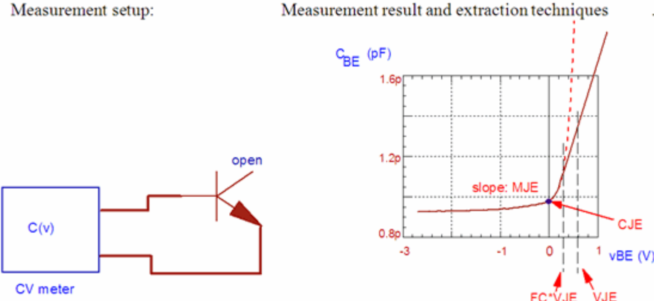
where TFac is the small signal transit time which is different from TFF. However, in this toolkit, we keep the simplified equation of CV-2 and a corresponding raw parameter extraction, and apply a quick optimizer run to get the true TF parameter.

Modeling the Space Charge Capacitors

Extraction of CJE, VJE, MJE, (CJC, VJC, MJC and CJS, VJS, MJS is the same)

- CJE B-E zero-bias deplet.capacitance
- VJE B-E built-in potential
- MJE B-E junction exponential factor

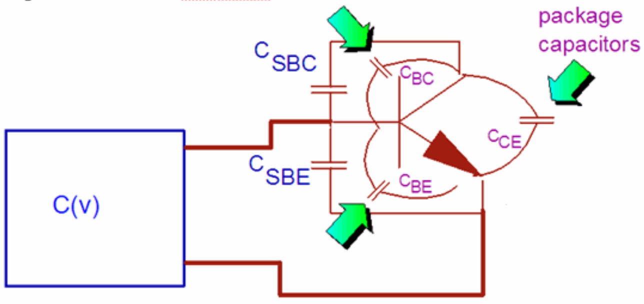
These parameters model the Base-Emitter and the Base-Collector space charge capacitance, i.e. the first term in (CV-1) and (CV-2). The second terms with the TFF and TR parameters will be modeled later by S-parameter mesurements. For the measurement of the Base-Emitter capacitance, the Collector is left open while the Emitter is open during the measurement of the Base-Collector capacitance. In both cases, the modeling formula is the same. Therefore this chapter covers only the modeling of the Base-Emitter capacitance.



Measuring and modeling the Base-Emitter capacitance.

Note on the Influence of the Remaining Capacitances of the Open Pin:

as one of the transistor pins is left open, the measurements of CSBC and CSBE are always an overlay of the other capacitances. CSBi (i = E,C) and, when measuring packaged devices, the small parasitic package capacitors CBC , CBE.and CCE.



CV meter

The total measured capacitance is therefore CSBi in parallel with the parasitic ones. This means that the measurement results are always too big. When using a capacitance meter like the Agilent4284, that eliminates by its measurement principle parasitic capacitances to ground, this effect can be avoided by applying an AC short to the open transistor pin versus ground (big capacitor).

The Equation

The behavior of the space charge capacitor is given by equations (CV-4a) and (CV-4b):

For vBE < FC * VJE:

$$C_{SBE} = \frac{C_{JE}}{\left(1 - \frac{v_{BE}}{V_{JE}}\right)^{M_{JE}}} \qquad \text{(CV-4a)}$$

and else:

$$C_{SBE} = \frac{C_{JE}}{(1 - F_C)^{(1+M_{JE})}} * \left[1 - F_C * (1 + M_{JE}) + M_{JE} * \frac{v_{BE}}{V_{JE}}\right] \qquad \text{(CV-4b)}$$

with

CJE : space charge capacitance at vBE = 0V

VJE : built-in potential or pole voltage (typ. 0,7V)

M_{JE} : junction exponential factor, determines the slope of the cv plot
(abrupt pn junction (<0,5um) : M_{JE} = 1/2)
(linear pn junction (> 5um) : M_{JE} = 1/3)

FC : forward capacitance switching coefficient, default 0,5

Determination of the CV Parameters

For simplicity, we only use the measurement data from the negative bias. The logarithmic conversion of (CV-4a) yields:

$$\ln(CSBE) = \ln(CJE) - M_{JE} \ln[1 - vBE / VJE] \qquad \text{(CV-5)}$$

This equation can be interpreted as a linear function according to the ideas of linear regression analysis:

$$y = b + m\,x$$

with

$$y = \ln(CSBE) \quad \text{(CV-6a)}$$

and

$$b = \ln(CJE) \quad \text{(CV-6b)}$$

$$m = -\,MJE \quad \text{(CV-6c)}$$

$$x = \ln[1 - vBE / VJE] \quad \text{(CV-6d)}$$

Linear regression means to fit a line to given measurement points. Therfore, the three main equations of a linear regression are b=f(xi,yi) and m=f(xi,yi), together with a fitting quality factor r^2=f(xi,yi,m,b). For a good fit, r^2~0.9...0.9999. See also the appendix.

How to Proceed

the measured values of CSBC are logarithmically converted according to (CV-6a). Following (CV-6d), the stimuli data of the forcing voltage vBE are nonlinearily converted too. This is done using a starting value for the unknown parameter VJE (e.g. 0,2V). These two arrays are now introduced into the regression equations (see appendix) as corresponding yi- resp. xi-values. A linear curve is fitted to this transformed 'cloud' of stimulating and measured data. Thus we get the y-intersect b(VJE) and the slope m(VJE) for the actual value of VJE. In the next step, this procedure is repeated with an incremented VJE, and we get another pair of m(VJE) and b(VJEC). But now the regression coefficient r^2 will be different from the earlier one. i.e. depending on the actual value of VJE, the regression line fits better or worse the transformed data 'cloud'. Once the best regression coefficient is found, the iteration loop is exited and we finally get VJE_opt as well as the corresponding b(VJE_opt) and m(VJE_opt).

Thus we get from (CV-6c):

$$M_JE = -\,m(VJE_opt)$$

and from (CV-6b):

$$CJE = \exp[\,b(VJE_opt)\,]$$

Validity of this extraction: The parameter extraction for the space charge capacitor is valid only for stimulus voltages vBE below FC * VJE, FC_default = 0,5.

What to do in IC-CAP

Since this is our first parameter extraction step, we first reset all parameter values to default, see IC-CAP Window: 'Model Parameters'
Otherwise, we might end up with a mix of parameter values obtained during our last transistor modeling and today. Therefore, execute Macro 'INITIALIZE' for this task.

Then, for our intended manual parameter extraction, make sure to set the Model Variables

MACRO_CONTROLLED=0
HELP_ENABLED=1

Extract the CV Parameters

open setup "/gp_classic_npn/cv/cbe_bhi"
(means modeling of CBE, with Base contact at high voltage pin),
import the data from NPN_MEAS_MASTER_demodata_PELdep.mdl
(located in demo_features\3_MEAS_ORGANIZE_n_VERIFY_DATA\0_MASTER_FILES)
click a box into plot "*cbe*" (capacitance vs. voltage) to select the measurement data used later for extractions. Click 'Copy to Variables' under 'Options' in that plot. This will cause IC-CAP to save the box corners in the 'cbe_bhi' Setup Variables
X_LOW, X_HIGH, Y_LOW, Y_HIGH
perform transform "br_CJE_VJE_MJE" (box regression). This transform applies a data transformation and regression analysis to the data inside the box.
Then simulate with the extracted parameter values, using simulation or the substitute transform calc_cv. Do the same for the capacitor CBC in setup 'cbc_bhi'.

{info:try also macro 'extract_n_opt_CV', or execute Transform 'OPEN_MDLG_GUI'

Some comments on CV-modeling

In practice there is always an overlay of this capacitance with some parasitic ones, e.g. package or pad capacitances. If they are not known and therefore cannot be de-embedded (calculated out of the measured data), the extracted CV parameter values may have no physical meaning. This may happen especially to VJC and MJC.

If there are resolution problems with fF-capacitances and CV meters, a network analyzer can be used instead of the CV meter as well. In this case, the Base is biased and Emitter and Collector are grounded. The measured S-parameters are deembedded, converted to Y parameters and the CV traces can be calculated out of their imaginary parts.

Introduction

Contents

- Operating modes of the bipolar transistor
- The Gummel-Poon equivalent schematic
- The Gummel-Poon model equations
- List of the SPICE Gummel-Poon parameters
- A quick tutorial on the Gummel-Poon parameters
- Proposed global extraction and optimization strategy

This manual describes the modeling of a bipolar transistor using the Gummel-Poon model as implemented in the simulator SPICE. It should be mentioned that the Gummel-Poon model itself covers only the internal part of a real-transistor. Therefore, on-wafer parasitics like a parasitic pnp transistor are not covered. Also, packaging parasitics and other non-ideal effects are not part of the model. However, they can be added by using a sub-circuit rather than just the stand-alone model.

Please check the example files included in the file directory of this toolkit for examples. Parasitic effects is specially important for network analyzer (NWA) measurements. The modeling procedures presented in this manual refer to already de-embedded measurements.

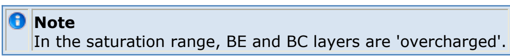
For on-wafer measurements, test probes that allow NWA calibrations down to the chip (like Cascade or Picoprobe probes) are commonly used. De-embedding means here to eliminate on-wafer parasitics, which are due to the test pads (OPEN dummy) and the lines from the test pads to the transistor itself (SHORT dummy). This is done by subtracting the Y matrix of the OPEN from the total measurement, followed ~~if required~~ by the subtraction of the Z matrix of the SHORT. It should be mentioned that in this case the SHORT itself has to be de-embedded first from the OPEN parasitics!

For packaged devices, we need to use a test fixture. In this case, the NWA has to be calibrated down to the ends of its cables using the calibration standards (SOLT) of the actual connector type. As a next step, the test fixture has to be modeled (OPEN, SHORT, THRU). Finally, the DUT (device under test) is inserted into the test fixture and measured. The now known test fixture parasitics can be de-embedded and the extraction techniques of this manual can be applied to the down-stripped inner device. A file including such a procedure is included in the toolkit filesets. See the example more_files/packaged_xtor_in_testfixture.mdl

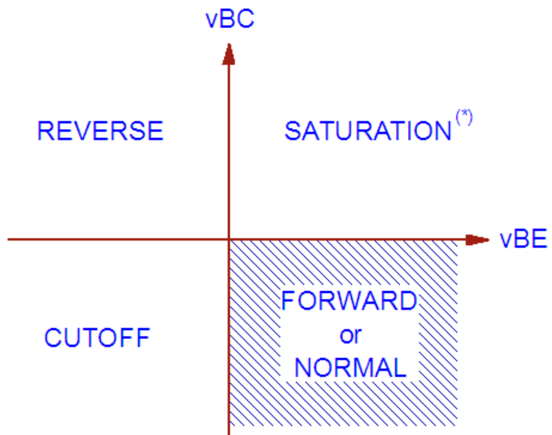
Please contact the author if you wish more info or help on de-embedding.

Operating Modes of the Bipolar Transistor

There are four operating modes of a bipolar transistor as illustrated in figure 1. The saturation region, for example, the region vCE<0.3V in the DC output characteristics, is described by the ohmic resistors. The DC and AC extraction procedures that are proposed in this manual cover mainly the forward region. Since the model is symmetrical, the reverse parameters can be extracted following the same ideas, but applied to the reverse measurements.



Operating Modes of a Bipolar npn Transistor

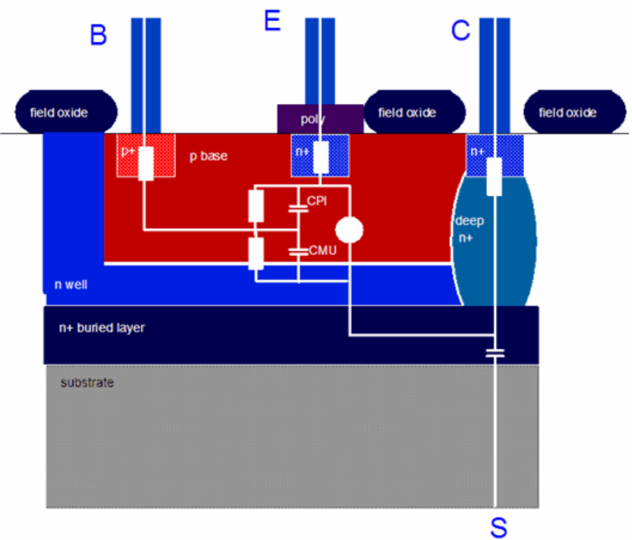


The Gummel-Poon Equivalent Schematic

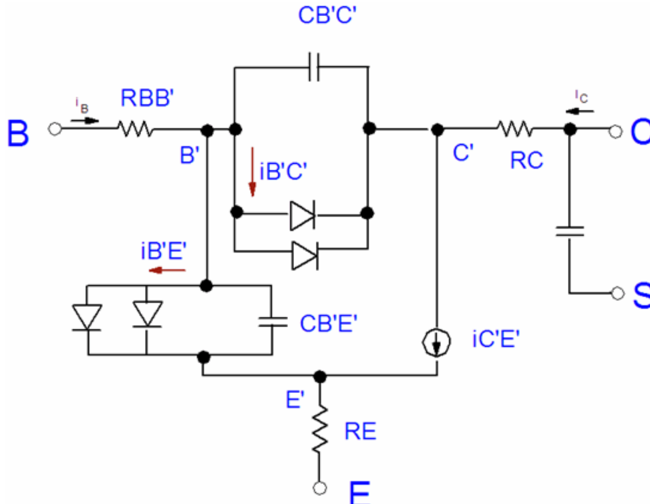
The figure, "Gummel-Poon large signal schematic of the bipolar transistor" shows the large signal schematic of the Gummel-Poon model. It represents the physical transistor: a

current-controlled output current sink, and two diode structures including their capacitors. This structure represents pretty much the physical situation of a bipolar transistor, see "physical situation for a bipolar transistor, neglecting the parasitic pnp transistor".

physical situation for a bipolar transistor, neglecting the parasitic pnp transistor

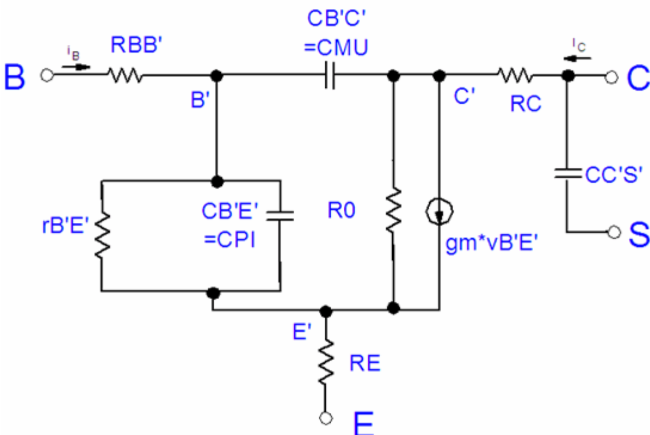


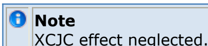
Gummel-Poon large signal schematic of the bipolar transistor



From figure "Gummel-Poon large signal schematic of the bipolar transistor" the small signal schematic for high frequency simulations can be derived. This means, for a given operating point, the DC currents are calculated and the model is linearized in this point (fig.AC small signal schematic of the bipolar transistor). Such a schematic is used later for SPICE S-parameter simulations.
It should be noted that the schematic after figure "AC small signal schematic of the bipolar transistor" is a pure linear model. It cannot be used to predict non-linear high-frequency behavior of the transistor. For this case, RF simulators like ADS perform high-frequency simulations using the large signal model (harmonic balance simulations).

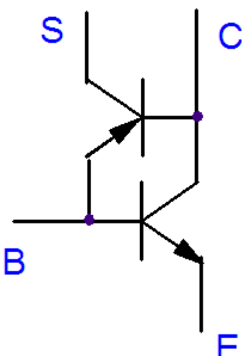
AC small signal schematic of the bipolar transistor





In order to make the presentations of the schematics complete, fig."subcircuit schematic when including the parasitic pnp" depicts the subciruit used for modeling a npn transistor including the parasitic pnp. As said above, IC-CAP files for this type of modeling are included in the file sets of this toolkit. However, the description of this manual does not cover this. See the macros in the model files instead.

Sub-circuit schematic when including the parasitic pnp



The Gummel-Poon Model Equations

For the reader's convenience all the Gummel-Poon equations are presented at a glance. In order to make them better understandable, we assume no voltage drops at RB, RE and RC, i.e. vB'E'=vBE and vB'C'=vBC.

TEMPERATURE VOLTAGE

$$vt = \frac{k\ T}{q} = 8.6171\ E\text{-}5 * (T\ /\text{'C} + 273.15)$$

Base Current

$$iB = iBE + iBC \tag{A}$$

$$iB = if/BF + iBErec + ir/BR + iBCrec \tag{B}$$

rec: recombination effect

with

ideal forward diffusion current: $$if = IS \left\{ \exp\left[\frac{vBE}{NF\ vt} \right] - 1 \right\} \tag{C}$$

B-E recombination effect: $$iBErec = ISE \left\{ \exp\left[\frac{vBE}{NE\ vt} \right] - 1 \right\} \tag{D}$$

ideal reverse diffusion current: $$ir = IS \left\{ \exp\left[\frac{vBC}{NR\ vt} \right] - 1 \right\} \tag{E}$$

B-C recombination effect: $$iBCrec = ISC \left\{ \exp\left[\frac{vBC}{NC\ vt} \right] - 1 \right\} \tag{F}$$

(see equiv. schematic in fig_2)

this gives:

$$iB = \frac{IS}{BF} \left\{ \exp\left[\frac{vBE}{NF\ vt} \right] - 1 \right\} + ISE \left\{ \exp\left[\frac{vBE}{NE\ vt} \right] - 1 \right\} + \frac{IS}{BR} \left\{ \exp\left[\frac{vBC}{NR\ vt} \right] - 1 \right\} + ISC \left\{ \exp\left[\frac{vBC}{NC\ vt} \right] - 1 \right\} \tag{G}$$

Collector Current

$$iC = 1/NqB\ (if - ir) - ir/BR - iBCrec \tag{H}$$

(see definition of iB above and equiv.schematic in fig_2)

or:

$$i_c = \frac{IS}{Nqb} * \left[\left(\exp\frac{vBE}{NF * vt} - 1 \right) - \left(\exp\frac{vBC}{NR * vt} - 1 \right) .... \right] - \frac{IS}{Nqb} * \left[\exp\frac{vBC}{NR * vt} - 1 \right] - ISC * \left[\exp\frac{vBC}{NC * vt} - 1 \right] \tag{I}$$

with the base charge equation

$$NqB = \frac{q_{1S}}{2} * \left(1 + \sqrt{1 + 4 q_{2S}} \right) \tag{J}$$

for the modeling of non-idealities like the base-width modulation:

$$q1s = \frac{1}{1 - \frac{vBE}{VAR} - \frac{vBC}{VAF}} \tag{K}$$

and the hi-level injection effect:

$$q2s = \frac{IS}{IKF} \left[\exp\left(\frac{vBE}{NF * vt} \right) - 1 \right] + \frac{IS}{IKR} \left[\exp\left(\frac{vBC}{NR * vt} \right) - 1 \right] \tag{L}$$

Base Resistor

$$RBB = RBM + 3(RB - RBM) \frac{\tan(z) - z}{z * \tan^2(z)} \tag{M}$$

with

$$z = \frac{\sqrt{1 + \left(\frac{12}{PI} \right)^2 \frac{i_B}{i_{RB}}} - 1}{\frac{24}{PI^2} \sqrt{\frac{i_B}{i_{RB}}}} \qquad PI = 3,14159 \tag{N}$$

Space Charge and Diffusion Capacitors

$$CBC = \underbrace{CSBC}_{\text{space charge}} + \underbrace{CDBC}_{\text{diffusion cap.}} \quad (O)$$

$$= \frac{CJC}{[1 - vBC / VJC]^{MJC}} + \frac{TR}{NR\ vt} \frac{IS}{NqB} \exp\left[\frac{vBC}{NR\ vt}\right] \quad (P)$$

and

$$CBE = \underbrace{CSBE}_{\text{space charge}} + \underbrace{CDBE}_{\text{diffusion cap.}} = \quad (Q)$$

$$= \frac{CJE}{[1 - vBE / VJE]^{MJE}} + \frac{TFF}{NF\ vt} \frac{IS}{NqB} \exp\left[\frac{vBE}{NF\ vt}\right] \quad (R)$$

with the transit time

$$TFF = TF \left\{ 1 + XTF \left[\frac{if}{if + ITF}\right]^2 \exp\left[\frac{vBC}{1{,}44\ VTF}\right] \right\} \quad (S)$$

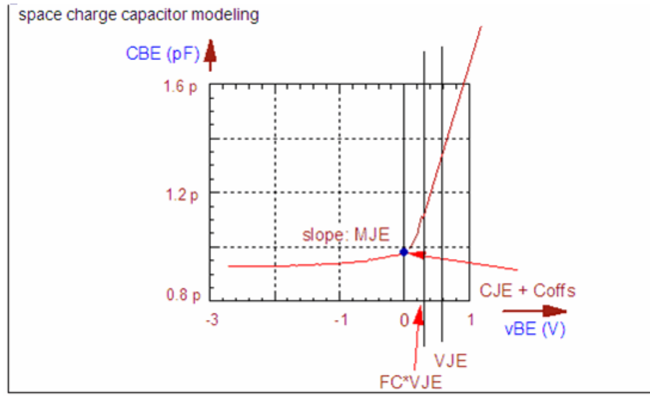
and the ideal forward base current if from the definition of iB, i.e. equation (C).

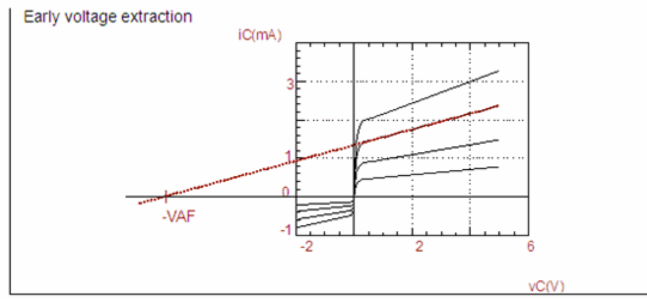
List of the SPICE Gummel-Poon Parameters

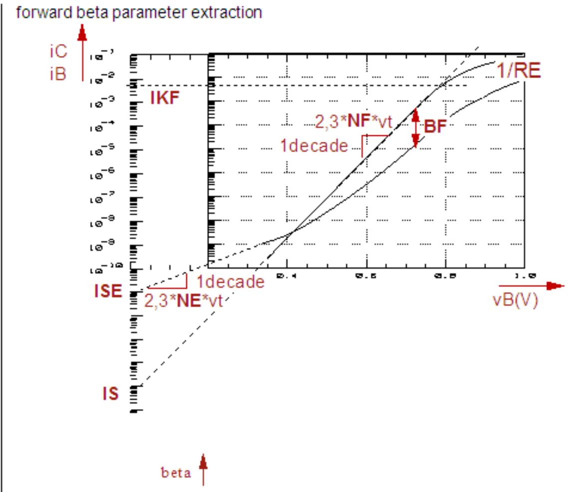
Name	Parameter explanation	SPICE default	typ.value	Unit
DC:				
IS	transport saturation current	.1E-15	1.E-15	A
XTI	temperature exponent for effect on IS	3	3	
EG	energy gap for temperature effect on IS	1.11	1.11	eV
BF	ideal forward maximum beta	100	150	
BR	ideal reverse maximum beta	1	.5	
XTB	forward & reverse beta temp.coeff.	0	2.5	
VAF	forward Early voltage	infinite	100	V
VAR	reverse Early voltage	infinite	50	V
NF	forward current emission coeff.	1	1.0	
NR	reverse current emission coeff.	1	1.0	
NE	B-E leakage emission coeff.	1.5	1.7	
NC	B-C leakage emission coeff.	2	1.3	
ISE	B-E leakage saturation current	0	.1E-12	A
ISC	B-C leakage saturation current	0	1.E-13	A
IKF	forward beta hi current roll-off	infinite	.05	A
IKR	reverse beta hi current roll-off	infinite	.3	A
OHMIC PARASITICS:				
RB	zero bias base resistance	0	100	Ohm
IRB	current at medium base resistance	infinite	.0001	A
RBM	min.base resistance at hi current	RB	25	Ohm
RE	emitter resistance	0	5	Ohm
RC	collector resistance	0	10	Ohm
CBE:				
CJE	B-E zero-bias deplet.capacitance	0	1.E-12	F
VJE	B-E built-in potential	.75	.6	V
MJE	B-E junction exponential factor	.33	.4	
CBC:				
CJC	B-C zero-bias deplet.capacitance	0	.5E-12	F
VJC	B-C built-in potential	.75	.6	V
MJC	B-C junction exponential factor	.33	.4	
XCJC	fraction of B-C capacitor connected to int.base	1	1	
CCS:				
CJS	zero-bias collector-substrate capacitance	0	0	F
VJS	substrate junction built-in potential	.75	0	V
MJS	substrate junction exponential factor	0	0	
CAPACITOR FORWARD CHARACTERISTICS:				
FC	forward bias depletion cap.coeff.	.5	.5	
TRANSIT TIME:				
TF	ideal forward transit time	0	1.E-12	sec
XTF	coeff.for bias dependence of TF	0	10	
VTF	voltage describing VBC dependence of TF	infinite	5	V
ITF	hi-current parameter for effect on TF	0	20.E-3	A
PTF	excess phase at frequency 1/(TF*2PI)	0	0	deg
TR	ideal reverse transit time	0	50.E-12	sec
NOISE:				
KF	flicker noise coeff.	0		
AF	flicker noise exponent	1		
TEMPERATURE EFFECTS				
.TEMP	device temperature for simulation /'C	27	27	'C
.OPTIONS TNOM	device meas. and param. extraction temp	27	27	'C

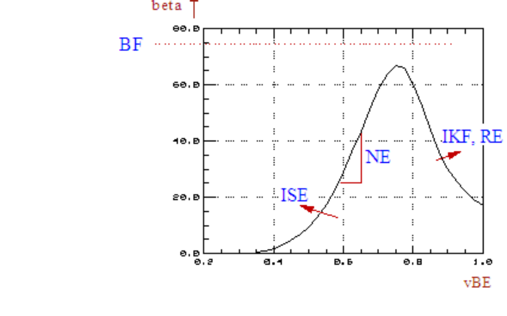
A quick Tutorial on the Gummel-Poon Parameters

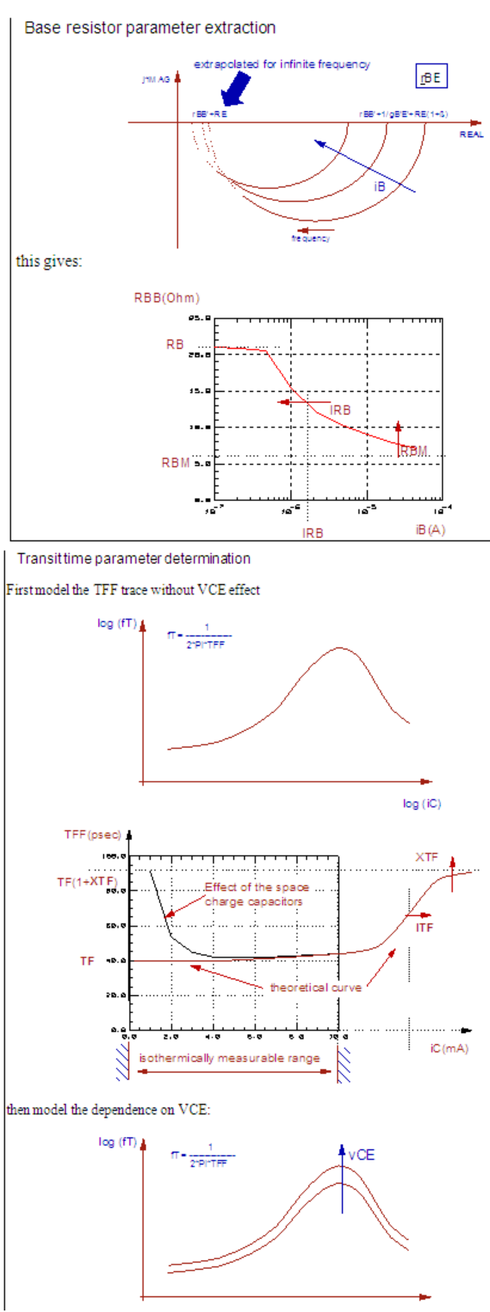
Although it is recommended to go through the individual extraction steps of the corresponding sections of this manual, this chapter puts together the graphical equivalents of the parameter extraction techniques.


space charge capacitor modeling
CBE (pF)
1.6 p
1.2 p
0.8 p
slope: MJE
CJE + Coffs
-3
-1
0
1
vBE (V)
VJE
FC*VJE


Early voltage extraction
iC(mA)
3
1
0
-1
-VAF
-2
2
6
vC(V)


forward beta parameter extraction
iC
iB
IKF
1/RE
2,3*NF*vt
BF
1decade
ISE
1decade
2,3*NE*vt
vB(V)
IS


beta
BF
NE
IKF, RE
ISE
vBE



Proposed Global Extraction and Optimization Strategy

First reset the model parameters to default (Window Model Parameters). This will firstly get rid of old parameter values which belong to the last modeling and not to our actual, current one, and secondly, the default parameters are those which reduce the complexity of a model completely. For example, VAF=1000 means: no Early effect, IKF=1000 no knee current, all resistors Rx=1m means no ohmic effects and so on. During the extraction process, we get more and more parameter values, and thus, the model becomes more and more complex and accurate.

CV

- Extract the CV parameters CJx, MJx and VJx for the BE- and BC-capacitance, optional also for CCS.
- optimize the CV parameters.

Ohmic Parasitics

- Extract the parasitic resistors RE, RBM and RC from flyback-measurements.
  or:
- Extract them from overdriven S-parameter measurements: high current at Base, half the current out of the Collector, Emitter grounded, frequency swept.

DC

- Extract VAR and VAF from the output characteristics.
- Extract IS, NF, ISE, NE and BF from the forward Gummel-Poon plots.
- Optimize the Gummel-Poon plot for IS, NF, ISE and NE, well below ohmic effects show up.
- Extract IKF from the ß-curve.
- Optimize RE in the upper region of the Gummel-Poon plots (iB and iC).
- Optimize BF and IKF in the ß-curve at high bias.
- Fine-optimize VAR, BR, VAF and BF in the output characteristics setup.
- Fine-tune all DC parameters in all DC setups.

S-parameters:

- De-embed the measurement data.
- Extract the base resistor parameters RB, IRB and RBM from S11 measurements with swept frequency and base current as a secondary sweep.
- Transform S- to H-parameters and get a frequency f-20dB from the -20dB/decade of h21.
- Measure again S-parameters, but now with the constant frequency f-20dB and swept iB and swept vCE and extract TF, XTF and ITF, as well as VTF.
- Optimize S-parameter fitting of TF, XTF, ITF (lowest vCE).
- Optimize the S-parameter fitting of VTF (all vCE).
- Go back to the rBB' setup and optimize the S-parameter fitting of the RB, IRB and RBM.
- Then, again in the rBB' setup, optimize the TF, XTF, ITF and VTF parameters.

Finally,

- Re-simulate all setups and check the fitting quality in the verify setups.
- If required, perform optimizer fine-tuning.

Macro 'extract_n_opt_ALL' in IC-CAP file gp_classic_npn.mdl contains an example for such a modeling strategy based on the measurement data included in the file.
This strategy may vary a bit depending on the actual data. In this case, simply modify the macro to meet your local requirements.

> **Note**
> A smart way of defining or verifying the most appropriate extraction strategy is to synthesize quasi-measured data from simulation results, and to check the extraction routines on these data. This means to simulate all setups using a given parameter set, to transform these simulated data into measured ones and to try to get the (known) parameters back again.
>
> In this way you are sure that your extraction strategy works well for a perfect Gummel-Poon transistor. If you have afterwards problems during the curve fitting, you might consider that your physical device under test may not be so well represented by the Gummel-Poon model!
>
> To 'synthesize' such pseudo-measured data in IC-CAP, make sure the parameter values in the IC-CAP parameter list are all set to typical values that you will expect later for your parameter extraction, perform a simulation for every setup in your model file, change the setup output data type to 'S', hit <RETURN>, change it back to 'B' again and hit again <RETURN>. Now you have identical data in both measured and simulated arrays. Then reset the parameter values to default and try your extraction strategy.

Last but not the least, before performing your measurements, i.e. before defining the measurement ranges, contact tour design engineer colleagues and ask them about the specific operating range.

As a general rule, modeling should be done in those regions where the transistor will be operated later.

Limitations of the Gummel-Poon Model

Ohmic Effects

The Collector and Emitter resistance parameters are constant and not functions of current or voltage. They have no temperature coefficients.

Forward DC Modeling

The parameter IKF models the begin of the decrease in beta. As a limitation of the model the slope of ß above the knee current IKF has a fixed value of "-1" on a log-log scale. However, this is most often overlaid by RE.

The modeling of the saturated region in the output characteristics (VCE < 0.5V) lacks of specific parameters. Therefore the model cannot cover modern transistors in this range (quasi-saturation).

No reverse breakdown effects are included in both Base-Collector and Base-Emitter diode.

Reverse DC Modeling

The reverse DC modeling suffers from a separate parameter IS. Thus NR sometimes has to be mis-used for better fitting the reverse iE versus vBC plot.

Like in the forward region, the slope of ß above the knee current IKR has a fixed value of "-1" and also the output characteristics saturation range is modeled inflexible.

AC Modeling

The TF modeling, especially versus vCE, is not physical and often not accurate
The TR parameter is not a function of current or voltage like TFF.

Temperature Modeling

The TNOM value of VJE, VJC and VJS must be greater than 0,4V to insure convergence for temperature analysis up to 200'C.

Applications in Integrated Circuits

There is no parasitic transistor included in the model.

Conclusion

Disregarding these limitations, the Gummel-Poon model is a good compromise between accurate modeling and a limited amount of parameters. It is still very useful especially when enhancing it with external parasitics like inductors, parasitic diodes or lateral pnp transistors.

Modeling the Base Resistor rBB'

Extraction of RB, IRB and RBM

RB zero bias Base resistance
IRB curr. at medium Base resistance
RBM min.Base resistance at hi current

It is assumed that XCJC = 1, and that ß is the DC current amplification.

This chapter explains how to model the Base resistor from S11 data.
It is organized like this:

- Derivation of the small signal schematic for the parameter extraction
- Short introduction into the basics of the Smith chart
- Discussing the expected frequency dependence of rBE, considering rBB' constant
- Enhancing the schematic for the bias-dependent omic resistor rBB'

As an approximation to keep the equations simpler, we further assume: vBE = vB'E' and vBC = vB'C'. Simulations and optimizer runs after the parameter estimation will eliminate this simplification.

The measurement setup for the rBB' characterization is given below in the following figure (RBB-1):

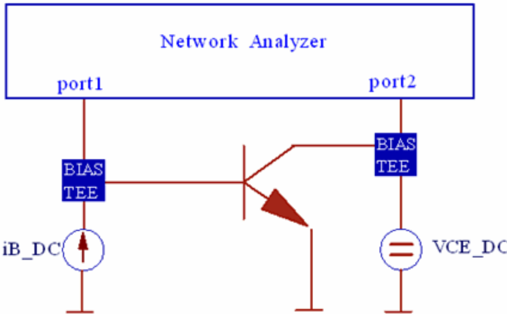


Fig.RBB-1: Measurement setup for the rBB measurement

To begin with, we refer to fig.AC-1 from the previous chapter. We simplify it to cover mainly the imput impedance. This leads to the schematic of fig.RBB-2. This figure explains the two cases: frequency -> 0 Hz and frequency -> ∞Hz

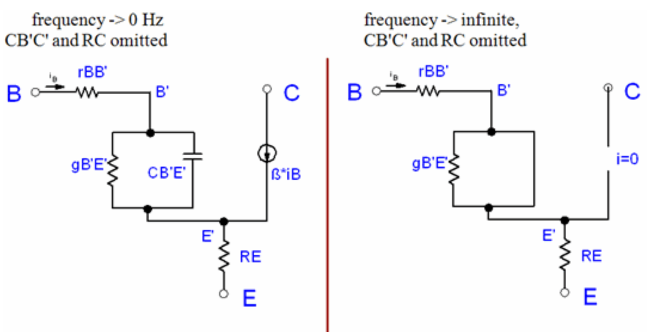


Fig.RBB-2:
Simplified AC Base-Emitter input impedance scheme for low and high frequencies

In order to evaluate the schematic and the device parameters of fig.RBB-2, we have to consider the measured S11 data. This is best done by displaying them in a Smith chart.

As a reminder, a Smith chart transforms the right side of the complex resistor plane R into the area of a circle of radius '1' using the transform
$[V^2/Hz]$

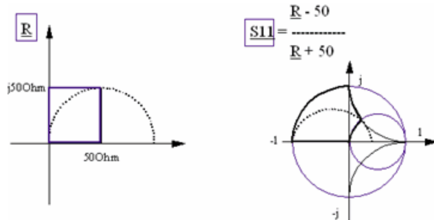


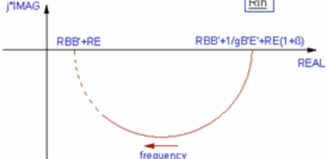
Fig.RBB-3: Note on the Smith chart transformation
(S11 = 1 for R = infinite , S11 = 0 for R = 50 Ohm , S11 = -1 for R = 0 Ohm)

Therefore, we can use S11 instead of H11 for the RIN modeling as well and our

measurement result should look like figure RBB-4 .

> Note: In order to get familiar with the problem, we consider first the hypothetical case that rBB' is no function of bias. In other words, the Base resistor is considered as a constant, ohmic resistor RBB'.

Complex resistor plane for the interpretation of RIN



and the same measurement result displayed in a Smith chart:

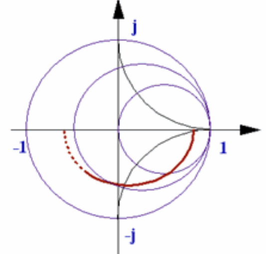


Fig.RBB-4: idealized input resistor curve, assuming an ohmic, bias-independent RBB'

Ideally, RIN should look like a circle.
The starting point at DC is RBB'+1/gB'E'+RE(1+ß). For higher frequencies, CB'E' will act more and more like a short and eliminate the influence of resistor 1/gB'E'. For infinite frequencies, RIN should hit the x-axis at RBB' + RE (effects of CB'C' and RC omitted!). Now the Base-Emitter capacitance has completely shorted 1/gB'E' and thus the transconductance gm became 0 as well. This means that the transistor has no beta any more.
$E_{nv}[V/\sqrt{Hz}]$

As RE is known from DC measurements, the value of RBB' can be estimated quite accurately.

This method is advantageous because the estimation of the Base resistor is affected only by the parameter RE. Moreover, there is mostly RBB' >> RE , so that the influence of a uncertain value of RE is minimized using this method.

So far we considered rBB' to be simply ohmic, i.e.constant. In reality, rBB' is modeled more complexly. One separate resistor from the outer Base contact to the inner Base contact (ohmic RBM) and a bias-dependent part from the inner Base contact to inside the inner Base. This means, the higher iB, the more iC is extending its flow area closer to the internal Base contact due to current crowding. This means, we expect a lower Base resistor for higher bias. The sketch below depicts that. The Gummel-Poon model combines both resistors into a single, bias dependent one.

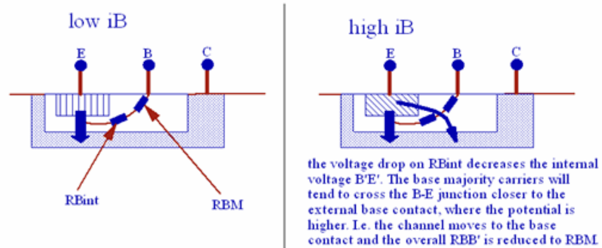


Fig.22a: current crowding leads to a bias dependent Base resistor

Now, overlying this DC bias dependency with the frequency dependence from above, we end up with S11 curves like sketched in fig.RBB-6.

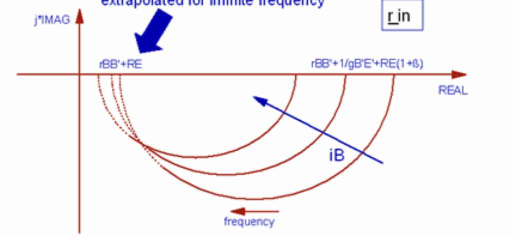


Fig.RBB-6: idealized input resistor curve with a real, bias-dependent Base resistor rBB', effects of CB'C' and RC omitted.

Real-world measurement curves will look like these curves at low frequencies only. This is

due to the overlay of more second order effects. In order to separate rBB' with the proposed method, we must fit circles to the S11 curves at low frequencies and then calculate the x-intersect from an extrapolation of the circle for infinite frequency, which is then assumed to be equal to rBB'+RE. This is shown in fig.RBB-7.

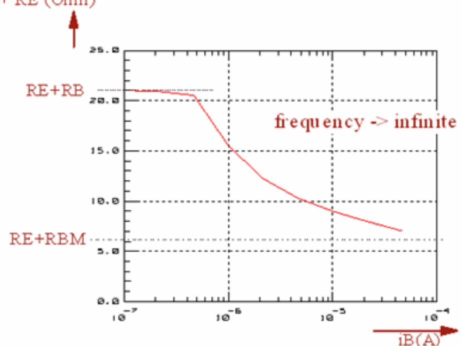


Fig.RBB-7: Idealized ohmic Base- plus Emitter-resistor versus iB

Notes on some limitations of this extraction strategy

Because of the influence of 1/gB'E' on the range where the circle must be fitted to, iB should be as high as possible to not dominate the r BB' - effect by 1/gB'E' . Also, to keep the rBB' influence dominant over RE (1+ß), iB should also be as high as possible, so that iC > IKF, and therefore ß, is as small as possible.
Unfortunately, the rBB' -measurement could now be dominated by thermal effects. Moreover, this range of iB typically is also not the operating one. This contradicts general rule to always concentrate on meaningful measurements close to the operating range for good parameter extractions.
Finally, the trace of Fig.RBB-7 is often overlaid by the parameters TF, ITF, XTF and VTF, which will be extracted next. Therefore, an optimization (of the S11 parameters) of this setup should only be applied after the fitting of these transit time parameters.

The extraction strategy

Circles must be fitted to the low-frequency sections of interest. They are centered to the x-axis. The suitable circle formula is:
A^2/Hz

This again can be considered as a linear form (!) with

$$C(f) = C_0 \cdot \frac{1}{1+\left(\frac{f}{f_g}\right)^2} \qquad (8)$$

This means: The measured data xi and yi are introduced into equ.(RBB-3a). Next the ylin are plotted versus the xlin and a straight line regression is applied. From the slope m, using (RBB-3c), we get:

$$C(f) = KB \cdot \frac{I^{AB}}{1+\left(\frac{f}{FB}\right)^2} \qquad (9)$$

and from the y-intersect b using (RBB-3b):

$$C_{1/f} = \frac{\alpha}{N_{tot}} \cdot \frac{1}{f} \qquad (10)$$

Finally the left circle intersection with the x-axis (for the frequency -> infinite) for our rBB'-extraction is:

$$C_{1/f} = KF \cdot \frac{I^{AF}}{f^{B}} \qquad (11)$$

After all these pre-considerations, we are now able to generate the trace of RBB out of the measured S-parameters. This means we are now ready to consider the formula for RBB in the Gummel-Poon model.

Equation

The nonlinear Base resistor is described in the Gummel-Poon model as:

$$\overline{i_r^{\,2}} = 4 \cdot k \cdot T \cdot \frac{1}{R} \cdot \Delta f$$

Fig.RBB-8 shows the plot of this equation:

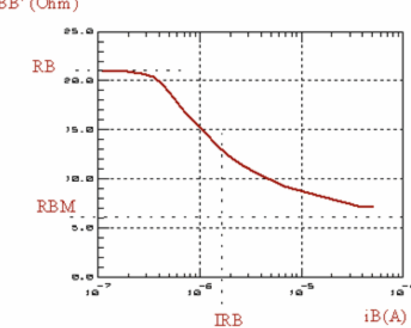


Fig.RBB-8: Base resistor curve as implemented in the model

This means: after we got rBB' from the measurement, we now have to fit the model curve from fig.RBB-8 to the measured data of fig.RBB-7 (after subtraction of RE).

Procedure

When iB -> 0 then z -> 0 and therefore

$$\overline{v_r^2} = 4 \cdot k \cdot T \cdot R \cdot \Delta f$$

We get from (53) solved for RB:
RB = rBB' [iB -> 0] , zero bias Base resistance
When iB -> infinite then z -> PI/2 and

$$\overline{i_D^2} = 2 \cdot e \cdot I_d \cdot \Delta f + KF \cdot \frac{I_d^{AF}}{f} \cdot \Delta f \qquad (12)$$

and

$$\overline{i_R^2} = 4 \cdot k \cdot T \cdot \frac{1}{R_{par}} \cdot \Delta f \qquad (13)$$

what gives from equation (RBB-4) solved for RBM:
RBM = rBB' [iB -> infinite] , min.Base resistance at hi current
Finally when iB = IRB then z = 1,21. Thus

$$\overline{i_{R,i}^2} = 4 \cdot k \cdot T \cdot \frac{1}{R_i} \cdot \Delta f, \quad i = b, c, e \qquad (14)$$

That's why
IRB is the current where the Base resistor is half its max.value
RBM + (RB - RBM) / 2

What to do in IC-CAP

After importing the de-embedded data, the modeling steps are:

- Open setup "/gp_classic_npn/nwa_extr/rbb",
- Import the data from the .mdm file,
- Simulate the setup with the so far determined DC and CV parameters
- Execute transform 'initialize'
- Perform transform "calc_RBB" to convert the measured S-parameters to rBB'
  (set model variable MACRO_CONTROLLED=0 in order to obtain the
  Check of the upper frequency limit for the rBB extraction)
- Check the plot "rbb_vbe" (rbb versus vbe)
- Perform transform "e_RB_IRB_RBM"
- Simulate with the extracted parameters.
- Optimize after you are finished with the TFF parameter extractions.
  (see proposed optimization sequence given there).

See also transform READ_ME

> NOTE: As you might experience, it can be quite complex to obtain a reasonable S11 plot from which a rBB' curve like that one in fig.RBB-8 can be derived. If despite all of these efforts the transformed measured data do not match the curve of fig.RBB-8, set RB=RBM and model the Base resistor bias independent.

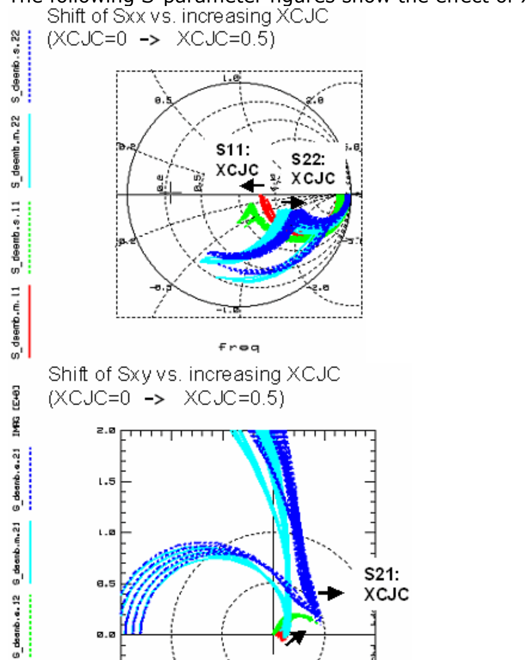
> Note: avoid thermal self-heating (esp.when measuring packaged devices). This will show up if the fitting of the forward Gummel plot of iB for high vBE becomes worse when the fitting of the S11 plot is improving during fine-tuning of RE and RB. If this occurs, reduce the bias for both the rBB and forward Gummel setup. If you need these high bias values, consider using pulsed measurements (DC bias pulse width around 1us).

Modeling the Parameter XCJC

This parameter distributes the CCB junction capacitance between the inner and the outer Base contact. Its default value is XCJC=1, meaning the CCB capacitance is tied completely to the inner Base. For XCJC=0, the capacitance is between the outer Base and the Collector.

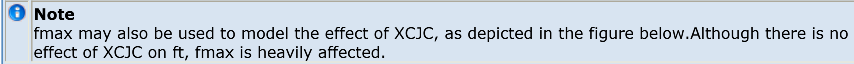
This parameter is difficult to determine from CV measurements. However, if the geometry of the device is known, it can be calculated pretty easily after:
XCJC = 1 - AE/AB,
where AE is the Emitter area, and AB the total area of the Base, including the Emitter area.

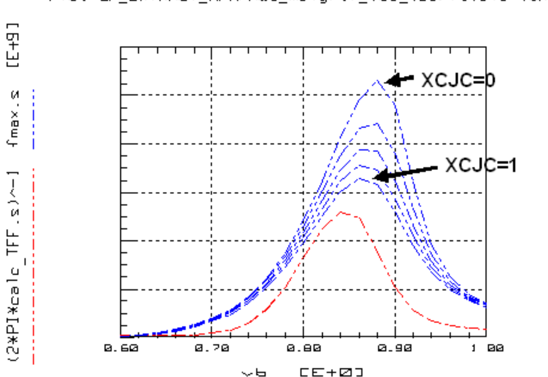
Usually, this parameter is fine-tuned after the other HF parameters have been determined.
The following S-parameter figures show the effect of XCJC for device modeling.



Generally speaking, if S12 becomes 'big' for high frequency, it is either Re or XCJC!

Note
fmax may also be used to model the effect of XCJC, as depicted in the figure below.Although there is no effect of XCJC on ft, fmax is heavily affected.





A final remark on the S-Parameter Extraction and Optimization Strategy

When you run into problems when fitting both, the transformed rBB' and TF curves, you should try to optimize what you have measured, i.e. S-parameters rather than the rBB' or the ft and TF plots. This is the real world and the fitting there might be more important than the fitting of the rBB' plot with all its limitations (extrapolated S11 at infinite frequency) or the fT plot (again extrapolated from H21 from S-to-H parameter transformation).

Of course, the best modeling result is a good fit in all domains, the S-parameters and the transformed rBB' and TFF curves.

Modeling the Resistors

Contents

- Extraction of RE
- Extraction of RC
- Extraction of RBM from DC measurements

An alternate method to calculate the ohmic parasitic resistors from s-parameter measurements.
The methods given below are considered as standard extractions. But the parameter values are pretty often merely a 'first guess'. Also, the other model parameters are still not yet knwon. Therefore, no simulation or optimization is performed in the setups of DUT prdc in file gp_classic_npn.mdl

Instead, these parasitic resistor parameters are fine tuned in the setups dc/fgummel and dc/rgummel. In the Gummel plots, they are tuned in order to fit the ohmic regions: RE in the forward Gummel plot (iC and iB vs vBE) and RC in the reverse plot (iE and iB vs vBC).

Modeling the Emitter Resistor

Extraction of RE

Measurement setup:

iC=0

iB

vCE

Measurement result:

vCE(mV)

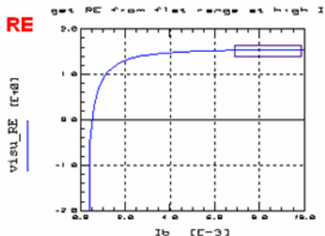
RE

1

iB(mA)

$$R_E = \frac{\partial v_{CE}}{\partial i_B}$$

transformed measured data:



Measurement of the open Collector voltage ('flyback method')
and the transformed measurement data in the RE domain (delta(vCE) / delta(iB))

Extracting the parameters:

The ohmic emitter resistor is physically located between the internal Emitter E' and the external Emitter pin E. When we apply a Base current and have the Emitter pin grounded, we get a voltage at the open Collector that is proportional to the Base current through this Emitter resistor. If we derivate vCE with respect to iB, we get the equivalent RE for each operating point. The value of RE is then the mean value of the flat range in this plot.

What To Do in IC-CAP:

- Import the data of setup rb_re
- Run transform visu_RE and select 'data transform'
  this will derive the measured data and display the
  calculated effective RE against the stimulus iB.

- Click a box around the most constant range of measured data and click 'Copy to Variables'
- Re-execute transform visu_RE to extract the RE value (select 'extract parameter' for this operation mode).

Do not simulate or optimize this setup, since

- The other DC model parameters are not known yet
- The Gummel-Poon model cannot represent 'unconventional' measurement conditions like the actual flyback method. The values of the ohmic parasitics will be fine-tuned later in the Gummel-Poon plots

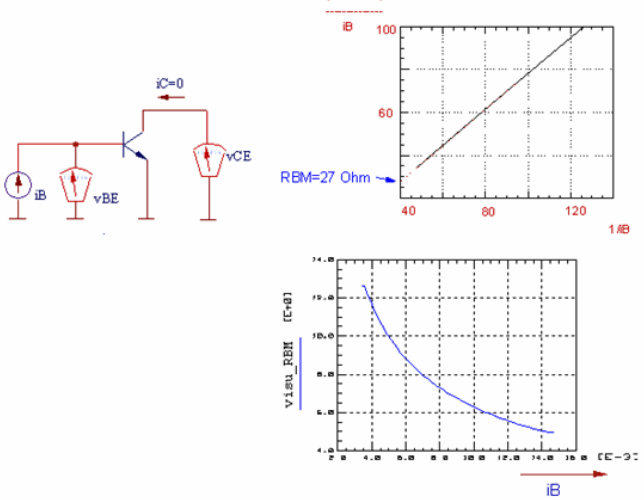
Extraction of RBM from DC measurements

RBM min.Base resistance at high current

There are several methods to determine the Base resistor: either the constant part of it (RBM) from pure DC measurements, or the non-linear RBB' = f(RB,IRB,RBM) from a s11 plot or from noise figure measurements.

Applying these three methods to the same transistor 'will generate typically three different values' for the Base resistor.

An interesting method to determine RBM is to use the RE-flyback method, with additionally measuring vBE /T.Zimmer/. This method is applied now.



Measurement setup and determination of RBM out of transformed measured data.

The theoretical values of the measured voltages are:

$[A^2/Hz]$

Subtracting these equations and dividing by iB yields:

$0.2nA/\sqrt{Hz}$

i.e. a regression analysis applied to these transformed measured data will give the y-intersect RBM.

In a final step, we then apply a loop to these data, in which a line is fitted to two adjacent points, and the y-intersect is calculated. The incremental y-intersect is then displayed against the stimulus iB.

> When RB becomes measurable DC-wise (the 'ohmic' range in the Gummel-Poon plot), its value is typically already lowered to the value of RBM. This means, parameter RB (the higher Base resistor value for lower Base bias), cannot be determined by this method. Therefore, we simply set RB=RBM.

> NOTE: See also the appendix chapter 'direct visual parameter extraction'.

What To Do in IC-CAP:

- The measurement of setup rb_re is re-used
- Run transform visu_RBM and select 'data transform'
  this will calculate the local Base resistor for each bias point, as described above, and display the RBM value against the stimulus iB.

- Click a box around the most constant range of measured data and click 'Copy to Variables'
- Re-execute transform visu_RBM to extract the RBM value

Again, do not simulate or optimize this setup, since the other DC model parameters are not known yet.

> If a sensitivity analysis for a Gummel-Plot shows a reasonable impact of the Base resistor to the forward and reverse Base current, an optimizer run on these two curves simultaneously might make sense to obtain a guess on the actual value of RBM.
> However, this is usually not the case.

Modeling the Collector Resistor

For the extraction of RC, the same flyback method like for RE is applied. The only difference is that the Collector pin is grounded, and the Emitter pin is left open and its voltage is measured.

What To Do in IC-CAP:

- Import the data of setup rc
- Run transform visu_RC and select 'data transform'
  this will derivate the measured data and display the calculated effective RC against the stimulus iB.

- Click a box around the most constant range of measured data and click 'Copy to

Variables'
- Re-execute transform visu_RC to extract the RC value.

Again, do not simulate or optimize this setup

Note: Try also and study macro 'extract_resistors".

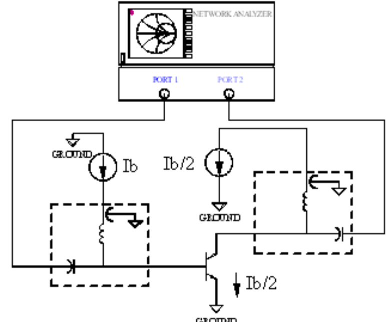
Notes
As mentioned above, these 'classical' extractions of the ohmic model parameters are used to get a good estimation about the parameter values. The values will be fine-tuned later in the setups fgummel and rgummel.

For details on alternate DC modeling methods of the parasitic resistors, see also the publications of /Berkner/ and /MacSweeny/.

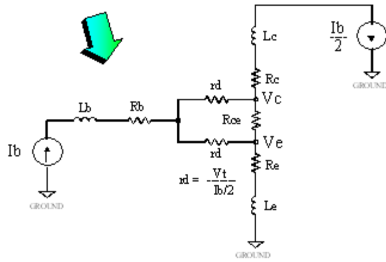
If there is a parasitic pnp transistor present, this method will not give accurate RC values. See the corresponding model file of this toolkit.

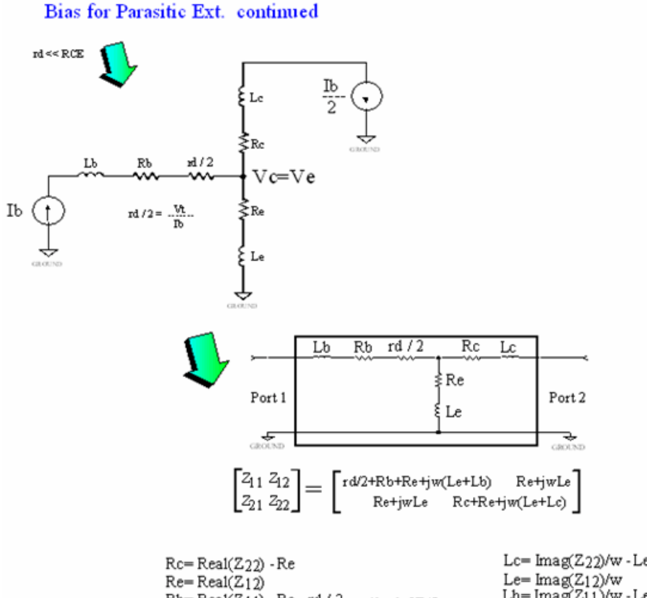
An alternate method to calculate the ohmic parasitic resistors from s-parameter measurements

Since the fitting of the S-parameters is the goal of a good transistor modeling, it makes sense to think about extracting the ohmic parameters from S-parameter measurements also. The following figures sketch a reliable way to do that. The basic idea is to overdrive the transistor and to reduce its effect to simple diode characteristics ('hot' measurement). With the known value of the Base current, the remaining resistor values can be calculated easily.

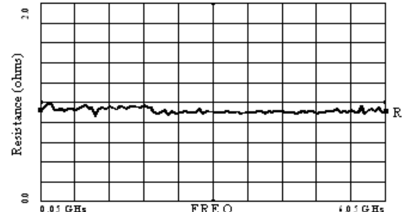


Bias for Parasitic Extraction





The resistor values are finally displayed versus frequency and their values are obtained as a simple mean value. If there is a frequency drift, take the mean value from the lowest frequency. The plot below gives an example:



Modeling the Temperature Effects

The parameters given below are modified when the selected simulation temperature TEMP is different from the extraction temperature TNOM. (Temperatures in 'K).

Used auxiliary variables:

$$VT = \frac{k\ TEMP}{q}$$

$$EG = 1.16 - \frac{7.02\ E\text{-}4\ TEMP^2}{TEMP + 1108}$$

$$Ni = 1.45\ E10 \left(\frac{TEMP}{TNOM} \right)^{1.5} \exp\left[\frac{q}{2k} \left(-\frac{EG}{TEMP} + \frac{1.1151}{TNOM} \right) \right]$$

Temperature dependent modeling parameters:

$$IS(TEMP) = IS(TNOM) \left(\frac{TEMP}{TNOM} \right)^{X_{TI}} \exp\left[\frac{EG}{VT} \left(\frac{TEMP}{TNOM} - 1 \right) \right]$$

$$BF(TEMP) = BF(TNOM) \left(\frac{TEMP}{TNOM} \right)^{X_{TB}}$$

$$BR(TEMP) = BR(TNOM) \left(\frac{TEMP}{TNOM} \right)^{X_{TB}}$$

$$ISE(TEMP) = ISE(TNOM) \left(\frac{TEMP}{TNOM} \right)^{-X_{TB}} \left[IS(TEMP) / IS(TNOM) \right]^{1/NE}$$

$$ISC(TEMP) = ISC(TNOM) \left(\frac{TEMP}{TNOM} \right)^{-X_{TB}} \left[IS(TEMP) / IS(TNOM) \right]^{1/NC}$$

$$VJE(TEMP) = VJE(TNOM) \left(\frac{TEMP}{TNOM} \right) + 2\ VT\ \ln \left[\frac{1.45\ E10}{Ni} \right]$$

$$VJC(TEMP) = VJC(TNOM) \left(\frac{TEMP}{TNOM} \right) + 2\ VT\ \ln \left[\frac{1.45\ E10}{Ni} \right]$$

Non Linear DC Modeling

Contents

- Extraction of VAR and VAF
- Extraction of IS and NF
- Extraction of BF , ISE and NE
- Extraction of IKF
- Extraction of the remaining reverse parameters NR, BR, ISC, NC and IKR

Three measurements are required in order to extract the DC parameters:

- An output plot including both, forward and reverse operation,
- And two so-called Gummel plots, one for forward and another for reverse mode.

These three plots have a certain context between each other. Neglecting this context can easily lead to one of the famous, so-called 'infinite modeling loops'.

This can be explained as follows:
Let's consider the forward Gummel plot. It is based on a measurement of i_B and i_C simultaneously, versus v_{BE} and is typically plotted half-logarithmically. Most often, the applied Collector-Base voltage is set to v_{BC} = 0V. The reason for this is that it simplifies the modeling equations (H)...(L) drastically. However, this approach can easily lead to the 'infinite loop' mentioned above.
IC-CAP does not need the simplified equations. The optimizer in IC-CAP always uses a true simulator like SPICE in the background that includes the complete Gummel-Poon equations. Therefore, while extracting the DC parameters or other parameters, there is no reason for having v_{BC} = 0 for the Gummel plot.

We can take a smarter approach. We first measure the forward output characteristic and extract VAR and VAF. Then, we leave this setup for the moment, and measure the forward Gummel plot. Differently from the commonly used method mentioned above, we apply a v_{CE} that is not zero, but between 2V and half the value of the maximum v_{CE} of the output plot. We do this for the following reason. Once the Gummel plot is fitted for this special voltage, the following output plot simulation already hits the measured curves exactly in the middle of the output characteristic. A final fine-tuning is then easily achieved by adjusting VAF and BF. Otherwise, if we use v_{BC} = 0 for the Gummel plot, it can easily happen that if the Gummel plot itself is nicely fitted, the output characteristics doesn't match and so on. Because, in this case, if the Gummel plot fits, this means that the output characteristic fits in the *saturation* range (v_{CE} ~ 0.2 ... 0.9V) and not int the desired linear range (v_{CE} ~ 0.5 to v_{CEmax}).

An illustration of this idea is presented below in fig.DC-1. First, the output characteristic is measured and VAR and VAF are extracted. Then, considering a cut through this plot for a fixed v_{CE} (4V in the example), and using this value of v_{CE} when measuring the Gummel-Poon plot, we have data points that refer directly to our previous output characteristics measurement with the corresponding v_{BE} . The relationship between i_C and i_B leads to the beta plot, also plotted against linear v_{BE} instead of the usual logarithmic i_C (which is the same for i_C below the ohmic effects in the Gummel-Poon plot) and again highlighting the corresponding output data points by buttons. Therefore, if beta fits, so does the output characteristic, which we were starting from.

Therefore, if we extract the DC forward parameters from a Gummel-Poon measurement that is biased like this, all measurements fit together.

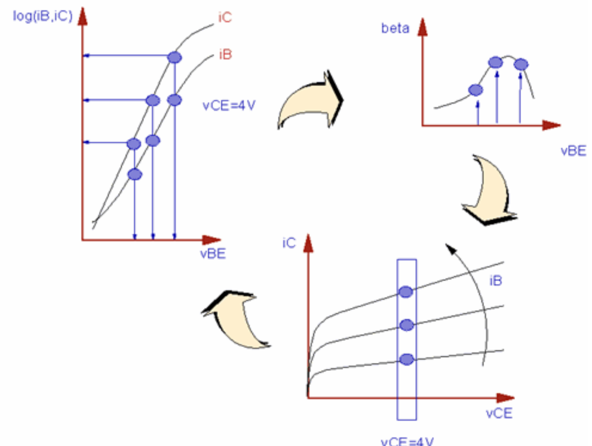


Fig.DC-1: Proposed context of the DC forward measurement setups.

For the output characteristic, by forcing i_B instead of v_{BE}, we also prevent from measuring thermal self-heating effects, which are not included in the standard Gummel-Poon model. However, we should also measure the same Collector currents values with a corresponding v_{BE} as well. This is sketched below in fig.DC-2.

Note:Such a check is implemented in file data_mgmt/BIP_MEAS_MASTER.mdl

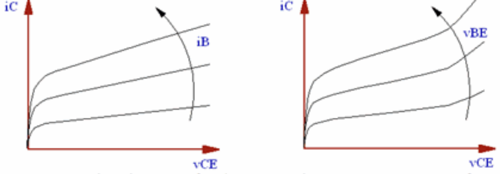


Fig.DC-2: Forcing iB rather than vBE for the output characteristics prevents from obtaining measurement curves including the self-heating effect.

Note: i_C is scaled identically for both plots.

Because, if the output characteristic drifts off when forcing v_{BE}, we should be careful when measuring the Gummel plot, because it could be affected by self-heating as well. The ohmic effects are in this case *overlaid* by the thermal self-heating, and we will either get wrong model parameters for RE and IKF, or no good fitting at all.

It is recommended in this case to apply a v_{CE} as low as possible for the Gummel plot (below the thermal runaway), but well above the saturation region of the output plot.

Note: See also IC-CAP file bip_output_char_i_or_v.mdl in the more_files directory

Modeling the Output Characteristic

Extraction of VAR and VAF

VAR reverse Early voltage
VAF forward Early voltage

Modern fast bipolar transistors exhibit small values of VAR. Due to the simplifications of the G-P space charge model implementation in SPICE, this may affect the other model parameters. This typically happens for VAR<5.

Therefore, the extraction of the nonlinear DC parameters is best started with the extraction of VAR, followed by VAF. As will be shown in this chapter, VAR/VAF can be determined with only little overlay of the other (actually still unknown) parameters.

After the Early voltages are extracted, and before optimizing the fit of this setup, we need to go ahead and extract the remaining DC forward model parameters from the Gummel plot. Only then, with the correct BF etc., the simulation of the output characteristic can fit the measured data.

Therefore, we come back to this setup and fine-tune the VAR/VAF values by optimization later.
Now, let's discuss the theoretical background of the Early voltage extraction. For an easier understanding, we consider VAF.

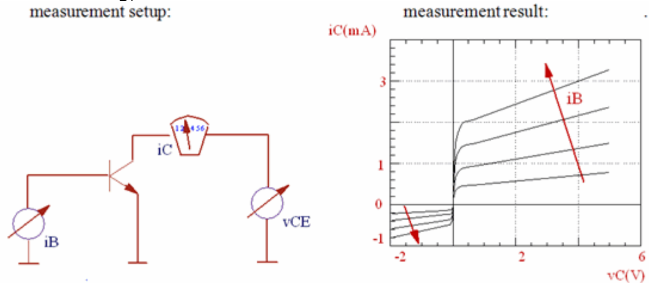


Fig.DC-3: Measurement of the output characteristic

The equation:
Provided that: vB'E'=vBE and vB'C'=vBC , the Gummel-Poon model describes iC by
$E_{nv}[V/\sqrt{Hz}]$
with the Base charge equation
A^2/Hz
for the modeling of non-idealities like the Base-width modulation:
$[V^2/Hz]$
and the hi-level injection effect:

$$C(f) = C_0 \cdot \frac{1}{1+\left(\frac{f}{f_g}\right)^2} \quad (8)$$

In order to handle this complex formula, we have to start with some simplifications:

We consider only the forward active region. Here, the Base Collector voltage is vBC < 0V; therefore the terms

$$C(f) = KB \cdot \frac{I^{AB}}{1+\left(\frac{f}{FB}\right)^2} \quad (9)$$

in equ.(DC-4) and (DC-4) may be neglected.

Thus (DC-4) becomes:

$$C_{1/f} = \frac{\alpha}{N_{tot}} \cdot \frac{1}{f} \quad (10)$$

Equ.(DC-6) in (DC-5) yields:

$$C_{1/f} = KF \cdot \frac{I^{AF}}{f^B} \quad (11)$$

Fig.DC-3 showed iC versus vCE. Therefore, (DC-8) has to be re-arranged using

$$\overline{i_r^2} = 4 \cdot k \cdot T \cdot \frac{1}{R} \cdot \Delta f$$

(DC-9) in (DC-8):

$$\overline{v_r^2} = 4 \cdot k \cdot T \cdot R \cdot \Delta f$$

with typically vBE << VAR and vBE << VAF we get:

$$\overline{i_D^2} = 2 \cdot e \cdot I_d \cdot \Delta f + KF \cdot \frac{I_d^{AF}}{f} \cdot \Delta f \quad (12)$$

and $$\overline{i_R^2} = 4 \cdot k \cdot T \cdot \frac{1}{R_{par}} \cdot \Delta f \quad (13)$$

or

$$\overline{i_{R,i}^2} = 4 \cdot k \cdot T \cdot \frac{1}{R_i} \cdot \Delta f, \quad i = b, c, e \quad (14)$$

Thus we got iC = f(vCE,iB) as shown in fig.DC-3, with vBE = f(iB).

Extracting the parameter:

We consider all assumptions from above valid. This means that we should be sure that the output characteristics measurement has been taken in the linear range of the Gummel-Poon plot, i.e. with a maximum iC well below the ohmic effects. Then, the Collector current of the output characteristics measurement (equation DC-11), becomes zero for vCE = -VAF.

How to proceed:

VAF is the x-axis intersect of the tangent fitted to the linear region of the output characteristics.

> NOTE: As you will find out with your own measurements, VAF is rather a function of the bias current than a constant. The standard deviation of the values of VAF found by applying tangents to all slopes in the output plot is most often very big. Depending on the type of transistor, sigma(VAF) can range up to VAF/2. The reason is that the assumptions in equations (DC-5)..(DC-11) are pretty straight forward. Therefore an estimation of VAF by using only 1 tangent may be sufficient, when an optimizer run is performed later (after the extraction of the remaining DC forward parameters). Please note again that the IC_CAP optimizer calls the simulator which includes the full set of model equations and therefore finds the correct final value of VAF.

An alternate method could also be to determine VAF out of the delta of two Gummel plot curves iC(vBE) for two different Collector-Emitter bias voltages. See equation (DC-18) of the next chapter.

What to do in IC-CAP:

open setup "/gp_classic_npn/dc/routput",
import the measured data
in plot 'ie_vec', specify a box for parameter extraction
perform transform "br_VAR" (extract VAR)
simulate with the extracted value of VAR.
Then,
open setup "/gp_classic_npn/dc/foutput",
import the measured data
in plot 'ic_vce', specify a box for parameter extraction

perform transform "be_VAF" (extract VAF)
simulate with the extracted value of VAF.

Do not be confused about the simulation result, and that the curves do not match.
Because all other DC parameters are still set to default, it is only important that the slopes of simulated and measured curves match!
We will have a much better fitting after the extraction of the other DC forward parameters.

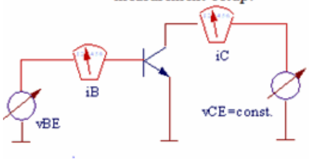
Have also a look into "/gp_classic_npn/dc/routput/READ_ME".

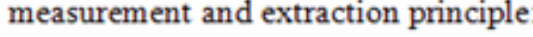
Modeling the Collector Current

Extraction of IS and NF

IS transport saturation current
NF forward current emission coefficient

These 2 parameters, together with the already known Early voltages, are the only ones that are dominant in the measurement setup given in the figure DC-4 below. NF determines the slope and IS the y-intersect of the half-logarithmically plotted iC(vBE).





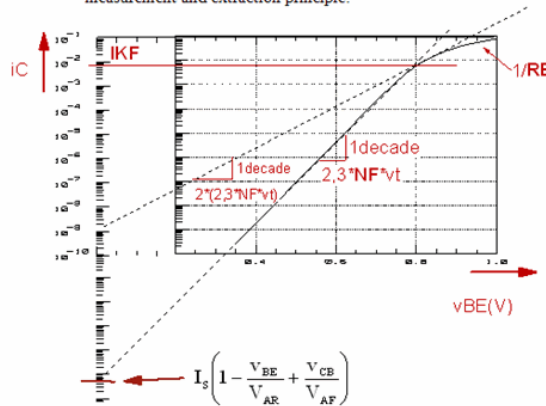


Fig.DC-4: Measurement of the Collector current vs. B-E voltage

The equation:
Provided that vB'E'=vBE and vB'C'=vBC , we start again with the iC formula (H) ... (L) from the introduction chapter:

$$\overline{i_{b,S}^{2}} = 2 \cdot e \cdot I_b \cdot \Delta f \qquad (15)$$

In order to determine IS from (DC-12) for small vBE, i.e. no ohmic and no IKF effects, we get for forward biasing

vBE ~ 0,7V << |VAR|
vBC < 0 V
and |vBC| << |VAF|.
Therefore (DC-12) simplifies to:

$$\overline{i_{c,S}^{2}} = \overline{i_{nc}^{2}} = 2 \cdot e \cdot I_c \cdot \Delta f \qquad (16)$$

Let's have a closer look to equ. (DC-13). Firstly, the formula reminds to apply the following series approach for small values of x :

$$\overline{i_{nb}^{2}} = 2 \cdot e \cdot I_b \cdot \Delta f + KF \cdot \frac{I_b^{AF}}{f} \cdot \Delta f \qquad (17)$$

what means for our case:

$$\left(\overline{i_{nb}^{2}}\right)^{ADS,BJT} = 2 \cdot e \cdot I_b \cdot \Delta f + KF \cdot \frac{I_b^{AF}}{f} \cdot \Delta f + KB \cdot \frac{I_b^{AB}}{1+\left(\frac{f}{FB}\right)^2} \cdot \Delta f \qquad (18)$$

NqB from (DC-17) is split into two parts: q1S represents a lowering of the Collector current for increasing Early voltages (DC-14). This can be seen in the iC Gummel plot as a curve shift to lower Collector currents. On the other hand, the other coefficient q2S begins to contribute for high Collector currents above IKF in forward operation resp. IKR in reverse (DC-15a), and reduces the Collector current as well.

For the modeling of IS and NF, we consider the lower and medium current ranges well below the Effect of IKF or the influence of the ohmic Resistor RE . Therefore, (DC-16) simplifies to:

$$\left(\overline{i_{nc}^{\,2}}\right)^{PSPICE} = 2 \cdot e \cdot I_c \cdot \Delta f + KF \cdot \frac{I_c^{\,AF}}{f} \cdot \Delta f \qquad (19)$$

For bigger values of the Early voltages, the terms |vBx| < |VAx| can be neglected and we obtain:

$$\left(\overline{i_{nb}^{\,2}}\right)^{VBIC} = 2 \cdot e \cdot I_b \cdot \Delta f + KF \cdot \frac{I_b^{\,AF}}{f^{BF}} \cdot \Delta f \qquad (20)$$

NOTE: Compare (DC-19) with the measurement result given in the figure DC-4 above.

Extracting the parameters:

Following the curve fitting techniques given in the chapter on regression analysis in the appendix, a transformation can be applied to the measured data in order to obtain a linear context between the measured values of iC and the stimulating values of vBE in (DC-19):

A log10 conversion of (DC-19) gives:

$$\overline{i_{RD}^{\,2}} = 4 \cdot k \cdot T \cdot \frac{1}{R_D} \cdot \Delta f \qquad (21)$$

and

$$\overline{i_{RS}^{\,2}} = 4 \cdot k \cdot T \cdot \frac{1}{R_S} \cdot \Delta f \qquad (22)$$

This can be considered as a linear form:

$$\overline{i_{nD,th}^{\,2}} = \frac{8}{3} \cdot k \cdot T \cdot g_m \cdot \Delta f \qquad (23)$$

How to proceed:

We select a sub-range of the measured data, where the half-logarithmicly plotted data represent a straight line. Then the logarithmically converted iCi of this sub-range are interpreted as y- and the linear vBEi values as x-data for the regression formula. Applying these formulas, we obtain y-intersect 'b' and the slope 'm' of the straight fitted line.

From comparing (DC-20a) with (DC-20b) we know how to re-substitute the parameters out of 'b' and 'm':

$$\overline{i_{nD}^{\,2}} = \frac{8}{3} \cdot k \cdot T \cdot g_m \cdot \Delta f + KF \cdot \frac{I_D^{\,AF}}{f} \cdot \Delta f \qquad (24)$$

and

With

$$\left(\overline{i_{nD,th}^{\,2}}\right)^{HSPICE} = \frac{8}{3} \cdot k \cdot T \cdot \beta \cdot (v_{gs} - V_{T0}) \cdot \frac{1+\alpha+\alpha^2}{1+\alpha} \cdot GDSNOI \cdot \Delta f \qquad (25)$$

and

$$\alpha = \begin{cases} 1 - \dfrac{v_{ds}}{v_{gs} - V_{T0}} & \text{linear region} \\ 0 & \text{saturationregion} \end{cases} \qquad (26)$$

Validity of the extraction:
vBE between 0,2V no noise and 0,7V [no high current effects]

What to do in IC-CAP:

open setup "/gp_classic_npn/dc/fgummel",
import the data,
click a box into plot "ibic_vbe" around a linear range for the IS/NF extraction
click 'Copy to Variables' (check how the box bounds are exported into the
setup variables X_LOW, X_HIGH, Y_LOW, Y_HIGH)
perform transform "br_IS_NF" (box regression IS, NF), which refers to X_LOW etc.
simulate with the extracted parameter values.
optimize with transform "bo_IS_NF"

Also have a look into "/gp_classic_npn/dc/fgummel/READ_ME".

HINT:
Transforming the measured data such that the model parameter can be displayed directly against the stimulating voltage or current is another smart way to determine model parameters. In the case of NF this would mean to start with

$$\overline{i_{Rd}^{\,2}} = 4 \cdot k \cdot T \cdot \frac{1}{R_d} \cdot \Delta f \qquad (27)$$

and

$$\overline{i_{Rs}^{\,2}} = 4 \cdot k \cdot T \cdot \frac{1}{R_s} \cdot \Delta f \qquad (28)$$

to convert it logarithmically in order to obtain

$$\overline{i_{nD}^{2}} = \frac{8}{3} \cdot k \cdot T \cdot g_m \cdot \Delta f + KF \cdot \frac{I_D^{AF}}{f \cdot C_{ox} \cdot L_{eff}^{2}} \cdot \Delta f \qquad (29)$$

This is the mathematical representation of the half-logarithmic Gummel plot for iC. The parameter NF is proportional to the slope and we have therefore to differentiate ln(iC) with respect to v_{BE} and obtain:

$$\alpha = \begin{cases} 1 - \frac{V_{ds}}{V_{D,sat}} & \text{in the linear region} \\ 0 & \text{in the saturation region} \end{cases} \qquad (30)$$

Solved for NF gives

$$\overline{i_{1/f}^{2}} = KF \cdot \frac{I_D^{AF}}{f^{EF} \cdot C_{ox} \cdot L_{eff}^{2}} \cdot \Delta f \qquad (31)$$

Therefore, if we display the calculated NF (what is the 'effective NF' for every measured data point) versus v_{BE}, we get

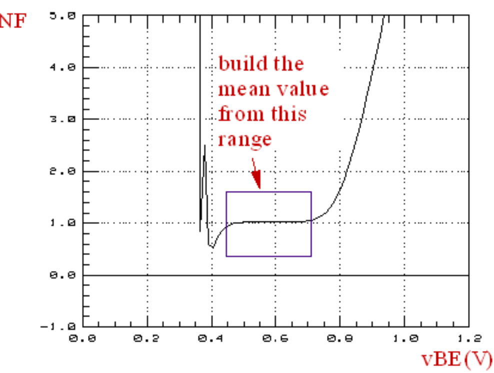


Fig.DC-5: Direct visual extraction for parameter NF

This allows us to check, if the model is able to fit the measured data at all (if there is a constantly flat range) and then to easily extract the parameter as the mean value of that flat range.

In directory "visu_n_extr" of this toolkit you will find more IC-CAP model files that follow the idea of direct visual extraction. See also appendix A for more infos.

Comments on ic_measured above IKF:
Referring to many graphical parameter extraction methods and also to fig.DC-4 above, some more background information is given for the model curve of ic above IKF.

From (DC-17) we get neglecting the Early-effect:

$$\overline{i_{th}^{2}} = \frac{8}{3} \cdot k \cdot T \cdot (g_m + g_{ds} + g_{mb}) \cdot \Delta f \qquad (32)$$

or solved for NqB:

$$\overline{i_{1/f}^{2}} = f(NOIA, NOIB, NOIC, EF, EM) \qquad (33)$$

Let us consider the two cases:

$$\overline{i_{th}^{2}} = \frac{4 \cdot k \cdot T \cdot \mu_{eff}}{L_{eff}^{2}} \cdot |Q_{inv}| \cdot \Delta f \qquad (34)$$

with $$Q_{inv} = -W_{eff} \cdot L_{eff} \cdot C_{ox} \cdot V_{gsteff} \cdot \left(1 - \frac{A_{bulk}}{2 \cdot (V_{gsteff} + 2 \cdot v_t)} \cdot V_{dseff}\right) \qquad (35)$$

Interpreting the result:
From (DC-22b) we learn that the ic curve has half the slope for currents above IKF (see fig.DC-4). In practice, however, there is always an overlay with the ohmic resistor RE, and therefore (DC-22b) is not so well suitable for extracting IKF.
However, the overlaying parameter RE is affecting basically both, the i_C and the i_B curve in the same way. This means that the effect of RE cancels out for the beta curve $ß=i_C/i_B$.
On the other hand, parameter IKF affects only i_C. Therefore, IKF is commonly extracted from the beta curve of the transistor

Modeling the Base Current

Extraction of BF, ISE and NE

BF ideal forward maximum beta
ISE B-E leakage saturation current
NE B-E leakage emission coefficient

In the literature, the three parameters of this chapter are most often introduced with their corresponding influence on the different ranges of the iB curve in DC-6.

In practice, there is most sometimes an overlay of the influences. This is especially true for BF in the beta plot (overlaid from IKF and NE). Also, modern transistors have pretty low recombination effects for the B-E diode: the 'famous knee' (see finger pointer in fig.DC-6) is not visible. Therefore we will not follow the graphical extraction method, but develop another method instead. We will derive a formula for the 3 parameters directly from measured data that has been taken from the range around the 'knee'.

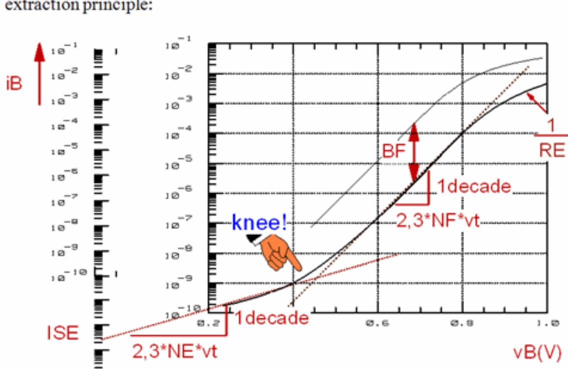


Fig.DC-6: measurement of the Base current vs. B-E voltage

The equation:
Provided that vB'E'=vBE and vB'C'=vBC , then

$$\overline{i_{nD,th}}^2 = \frac{8}{3} \cdot k \cdot T \cdot g_m \cdot \Delta f \qquad (23)$$

We assume once again that:

vBE ~ 0,7V << |VAR|
and vBC < 0 V

This simplifies equ.(DC-23) to:

$$\overline{i_{nB\ 1/f}}^2 = KF \cdot \frac{I_{B_DC}^{AF}}{f} \cdot \Delta f \quad \left[A^2\right] \qquad (37)$$

Introducing (DC-19) -i.e. the Collector current iC with neglected high current effects into (DC-24) yields the pretty simple form:

$$S_{iB} = \frac{\overline{i_{nB\ 1/f}}^2}{1Hz} = KF \cdot \frac{I_{B_DC}^{AF}}{f} \quad \left[\frac{A^2}{Hz}\right] \qquad (38)$$

We will use both iC = f(vBE) and iB = f(vBE) from the simultaneously measured currents of the Gummel-Poon measurement of fig.DC-4.

We now have iB as a function of vBE as desired.

Extracting the parameters:
This equation (DC-25) is one of the few cases during the bipolar modeling, where a non-linear transform applied to the measured data doesn't give a straight line. (At least, the author had not sufficient intuition!). Therefore the partial derivations of the curve fitting error in (DC-25) versus BF, ISE and NE have to be calculated and then set to zero. The solution of this system of equations finally gives these 3 parameters.

As iB ranges from pico- to milli-Ampere, we will have to minimize the relative error between measured and fitted curve. Thus we get from (DC-25) after dividing by iB:

$$S_{iB} \cdot f = KF \cdot I_{B_DC}^{AF} \qquad (39)$$

Equation (DC-26) is only approximately true for the real measured data iBi, iCi and vBEi. Therefore it is expanded by the individual error Ereli for every data point of index i:

$$S_{iB@1Hz} = KF \cdot I_{B_DC}^{AF} \qquad (40)$$

Using least means square techniques we now have:

I_{B_DC}

It can be shown that the parameters BF and ISE can be separated out of the partial derivations with respect to BF and ISE with a reasonable effort. This is unfortunately not

possible for NE . This parameter has to be iterated - similar to VJ of the space charge capacitor - until the sum of individual errors according to (DC-29) is minimized.

Step by step:
The partial derivation of (DC-29) versus BF is:

$$\log_{10}\left(S_{iB@1Hz}\right) = \log_{10}(KF) + AF \cdot \log_{10}\left(I_{B_DC}\right) \quad (41)$$

and versus ISE:

$$y = a + b \cdot x \quad (42)$$

(DC-30) is expanded by

$$y = \log_{10}\left(S_{iB@1Hz}\right)$$

$$a = \log_{10}(KF)$$

$$b = AF$$

and (DC-31) by

$$x = \log_{10}\left(I_{B_DC}\right)$$

These two new equations are added and their sum is solved for ISE:

$$AF = b \quad (43)$$

Now we can also separate BF from (DC-30):

$$KF = 10^{a} \quad (44)$$

Validity of the extraction:
vBE above measurement resolution and below high current effects.

How to proceed:

A subset of the measured data iBi and iCi , i.e. the range around the 'KNEE' (see fig.DC-6) are selected and introduced into equations (DC-32) and (DC-33). Next a suitable starting value for NE is selected (e.g. NE = 1) and the error according to (DC-29) is calculated. NE is then incremented until this error becomes a minimum. The triplet of NE, BF and ISE of this minimized error is the final parameter extraction result.

> NOTE: the complexity of (DC-32) and (DC-33) illustrates that transforming measured data to a linear context and applying linear regression techniques is often a much smarter approach for parameter extraction.

What to do in IC-CAP:

in setup "/gp_classic_npn/dc/fgummel",
click a box into plot "ibic_vbe" around the 'knee' at low vb,
click "Copy to Variables",
perform transform "br_ISE_NE_BF" (box regression ISE, NE, BF),
simulate with the extracted parameter values.
perform transform "bo_ISE_NE_BF" (box optimization ISE, NE, BF),

If there is no 'knee' with your measured transistor, the Base current recombination effect does not occur. In this case, switch off the Base current recombination effect in the G-P model. This can be done by setting ISE to a very small value (ISE=1E-30) and the slope parameter NE to a flat slope (NE=2).

Have also a look into "/gp_classic_npn/dc/fgummel/READ_ME".

> Note:
> For low values of VAR, the Collector current formula of (DC-19) inserted into (DC-25) is not quite correct. It would lead to a too low extracted value of BF, due to the shift of the iC Gummel plot. Equation (DC-18), without the assumption of big Early voltages, is better in this case. Therefore, correct the measured Collector current values to $r_{BE} = \frac{\partial v_{BE_DC}}{\partial i_{B_DC}}$.

> VAR<<VAF, so this correction is sufficient before inserting them into equations (DC-32) and (DC-33).

Modeling the Current Amplification at High Current

Extraction of IKF

IKF forward beta high current roll-off

Referring to fig. DC-4, IKF models the Webster push-out effect. This effect describes a decrease of the proportionality of log(iC) versus vBE. Unfortunately, as already mentioned, this effect is also overlaid by RE. However, while RE affects mainly both, iB and iC, IKF only affects iC. Therefore, it can be best extracted out of ß=iC/iB.

From a modeling standpoint, the beta plot should not be considered isolated from its origin, the iB and iC curve. Therefore, we display it always versus the same stimulus vBE, together with iB and iC. This helps a lot in better understanding the influence of the parameters ISE and NE on the increase of beta, and of IKF for the decrease.

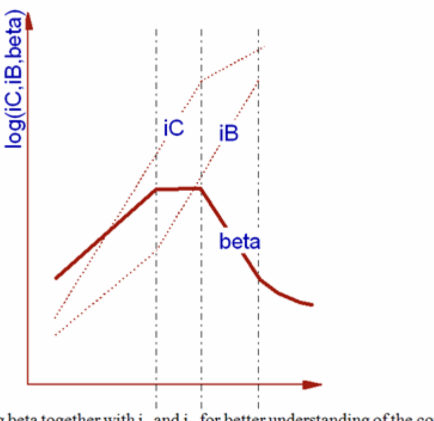


Fig.DC-7: Plotting beta together with i_C and i_B for better understanding of the context of the model parameters

It becomes also clear why for some transistors, BF seems not to affect the maximum of the beta trace at all: there is no parallel region between log(iC) and log(iB) in the Gummel plot, or referring to the beta plot, IKF reduces already beta, before it can reach the value of the BF for increasing bias. This is shown in fig. DC-8.

same measurement setup as for fig.DC-4

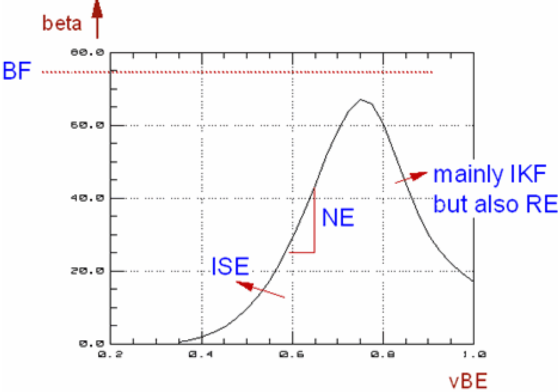


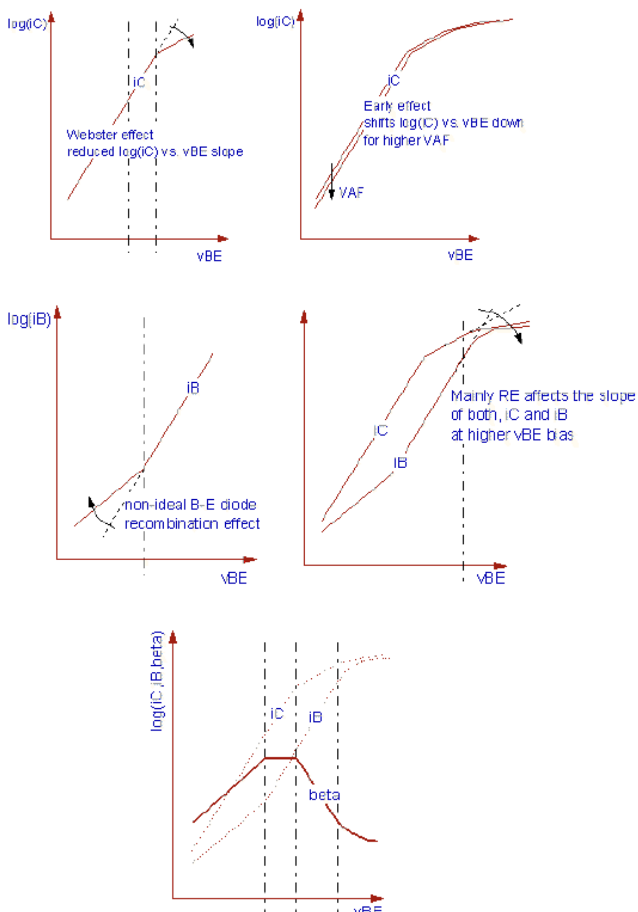
Fig.DC-8: Beta from the measurements of fig.DC-4 and fig.DC-3.

> Note: log(ß) is plotted versus vBE . The conventional plot using log(ß) versus log(iC) is identical to it, provided we neglect high-current effects. But this way of plotting log(ß) is more straightforward, because it displays a measurement result versus a stimulus and not another measurement result.

NOTE: for best results, estimate or extract RE first.

Hint
Avoid thermal self-heating effects. For Collector currents >10...50mA, thermal self-heating has to be taken into account. This becomes visible if the beta-plot for a *forward and reverse vBE sweep looks different* at high vBE. To avoid this, DC pulsed measurements with pulse widths about 1us should be used in this case (the HP4142 offers only pulse widths ≥100us).

Before we start with the extraction of IKF, we are now ready to understand the following schematized Gummel plots. They characterize at a glance the different effects for the Base and Collector current in the Gummel-Poon model.



The equation:
The current amplification is defined as:
i_{nB}^2
Provided that: vB'E'=vBE and vB'C'=vBC and further

$$\overline{i_{nD\ 1/f}}^2 = \mathbf{KF} \cdot \frac{I_{D_DC}^{AF}}{\mathbf{f}^{EF} \cdot \mathbf{COX} * \mathbf{Leff}^2} \cdot \Delta\mathbf{f} \qquad (45)$$

We introduce (DC-24) for iB and(DC-16) for iC into (DC-34):

$$COX = \frac{\varepsilon_0\ \varepsilon_{Si}}{TOX} = \frac{3.45E-11}{TOX}$$

$$Leff = L - 2 * \left(LINT + \frac{LL}{L^{LLN}} + \frac{LW}{W^{LWN}} + \frac{LWL}{L^{LLN} * W^{LWN}}\right)$$

We further introduce an approximation for NqB, see equ.DC-18. This gives:
$Leff = L - 2 * LINT$
Divided by

$$S_{iD} = \frac{\overline{i_{nD\ 1/f}}^2}{1Hz} = KF \cdot \frac{I_{D_DC}^{AF}}{f^{EF} \cdot COX * Leff^2} \qquad (46)$$

we get:

$$\log_{10}(S_{iD}) = const. - EF \cdot \log_{10}(f) \qquad (47)$$

Extracting the parameters:

As we want to consider again relative errors, we proceed as in the chapter of the determination of the iB parameters:
$1/f^{EF}$

This formula is again more ore less true for the measured data iCi and iBi with the stimulating voltage vBE. Thus we have to introduce again an individual error Ereli for each measured data point of index i:
'f^{EF}'

Finally the total error for all measured data (1 .. N) is (least means square):

$$S_{iD} \cdot \mathbf{f}^{EF} = \mathbf{KF} \cdot \frac{I_{D_DC}^{AF}}{\mathbf{COX} * \mathbf{Leff}^2} \qquad (48)$$

How to proceed:

In order to keep things simple, (DC-39) is solved for a best IKF by iteration. Thus IKF is set to a starting value, e.g. 10A, and then divided by 2 in every iteration, until the total error given in (DC-39) is minimized. Fine-tuning is then done by the optimizer.

What to do in IC-CAP:

open setup "/gp_classic_npn/dc/fgummel" and plot "ibic_vbe"
(beta is the right-axis data), then
perform transform "e_IKF" (extract IKF) and
check the simulation result.
run transform 'o_BF_IKF_RE' for fine-tuning the parameters of this setup.

Have also a look into "/gp_classic_npn/dc/fgummel/READ_ME"
and also in some alternate methods on the IKF-extraction in the 'direct visual parameter extraction' model file.

Extraction of the remaining Reverse Parameters

Extraction of NR, ISC, NC, BR, IKR

The reverse modeling can be performed like the forward modeling. Simply exchange Emitter and Collector.

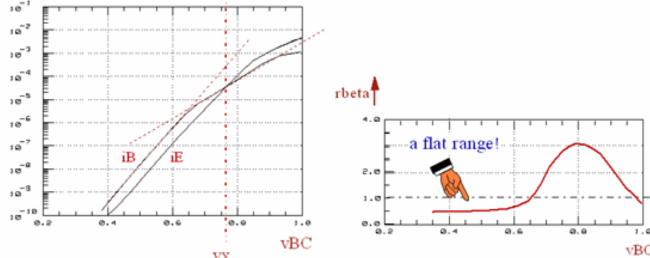
What to do in IC-CAP:

open transform README in setup "/gp_classic_npn/dc/rgummel" and follow the modeling sequence given there

Last not least, macro 'extract_n_opt_DC' includes a suitable automated modeling strategy for both DC forward and reverse.
Included in this example is also the interesting and pretty often recognizable effect, that the reverse Early voltage is affecting the forward modeling, due to its low value.
The strategy used in this macro covers that effect by looping a bit between forward and reverse extraction and optimization.

This sequence may be different for your actual transistor. Just correct the macro if required.

Note on reverse Gummel-Poon modeling
If your reverse beta curve and reverse Gummel plot look like below,



(a steeper slope of log(iB) versus vBC for vBC < vx, something that is not included in the Gummel-Poon model), you might consider replacing the Gummel-Poon recombination modeling (parameters ISC and NC) by a external diode with its parameters IS, N and RS.

For more details, refer to file rgummel_special.mdl under directory 'more_files' in this toolkit file collection.
See also the chapter on the limitations of the Gummel-Poon model at the end of this manual.

Preface

This product has been developed to meet the local demands of European IC-CAP users for more technical background information on extraction techniques and for the availability of extraction source code. Published for the first time in 1990, it has been updated since then several times. It is part of a series of supplementary modeling tool-kits for the IC-CAP users. These products feature source code and detailed technical description of the extraction routines. Please contact the author for further information.

About Gummel Poon User Information

This information is intended to explain the basics of modeling a bipolar transistor using the Gummel-Poon model as it is implemented in the simulation program SPICE of the University of California Berkeley (UCB) /see publication list/. It is part of the Gummel-Poon Bipolar Model Parameter Extraction Toolkit.

This toolkit includes the IC-CAP model file GP_CLASSIC_NPN.mdl, the MASTER model file which is described in this manual and featuring the data management features of IC-CAP 5.x, i.e. separating measurements from extractions:

- NPN_MEAS_MASTER.mdl a master file for measurement
- GP_EXTRACT_NPN.mdl a master file for modeling

as well as manyother IC-CAP model files covering topics like:

- model parameter extraction using the tuner feature
- direct visual parameter extractions
- alternate modeling methods for DC- CV- and RF-parameters.
- bipolar transistor modeling including the parasitic transistor.

Note
After you have become familiar with the modeling procedure itself, i.e. file GP_CLASSIC_NPN.mdl, you are encouraged to split the modeling into 2 parts: separate measurements and separate extraction strategy. In this case, all measurements are performed using the file NPN_MEAS_MASTER.mdl. Then, the data are exported into IC-CAP mdm files (ASCII files) and imported into the master extraction file GP_EXTRACT_NPN.mdl for extraction. This method allows to improve continuously the extraction strategy file, independent of the measurement data.

IMPORTANT NOTE:
This manual and the underlying IC-CAP model file GP_CLASSIC_NPN.mdl are intended to explain the basics of the Gummel-Poon modeling. Therefore, it covers the classical Gummel-Poon model without enhancements for also modeling the parasitic transistor. However, as stated above, such model files are included in the file sets of this toolkit. Please see the README macros in these IC-CAP model files for more details. You are also invited to get in contact with the author for assistance with such modeling problems.

The IC-CAP model file "GP_CLASSIC_NPN.mdl" features

The extractions are written using PEL (parameter extraction language) and are open to the user. They can be easily modified to meet specific user needs. Sub-circuit model description, open for user enhancements (HF modeling, parasitic pnp etc.). All transistor pins are connected to SMUs for flexible measurements

The transistor output characteristic and S-parameter measurements use a Base current stimulus rather than a Base-Emitter voltage in order to avoid 1st order thermal effects being visible. However, self-heating might be present and affect the Gummel plots in the ohmic range. See also the file GP_MEAS_MASTER.mdl.

Organization of Information

There are 5 main chapters, which explain how to determine the model parameters from CV (capacitance versus voltage), then parasitic ohmic resistors, and DC, to finally high frequency measurements using network analyzers.

More chapters cover side aspects of bipolar modeling.

The individual chapters follow always this scheme:

1. explanation of the parameter-dependent measurement setup
2. explanation of the mathematical basics for the parameter extraction
3. explanation of the parameter extraction
4. explanation about how to use the IC-CAP file.

Publications

- P.Antognetti, G.Massobrio: Semiconductor Device Modeling with SPICE, McGraw-Hill, 1988, ISBN 0-07-002107-4
- F.van der Wiele, W.L.Engl, P.G.Jespers: Process and Device Modeling for Integrated Circuit Design, NATO Advanced Study Institute, Louvain-la-Neuve, Belgium, 1977
- D.A.Hodges, H.G.Jackson: Analysis and Design of Digital Integrated Circuits, McGraw-Hill, New York, 1983, ISBN 0-07-029153-5
- P.R.Gray, R.G.Meyer: Analysis and Design of Analog Integrated Circuits, John Wiley&Sons, New York, 1984, ISBN 0-471-87493-0
- I.Getreu: Modeling the Bipolar Transistor,Tektronix Publication No. 062-2841-00, 1976
- H.K.Gummel, H.C.Poon: An Integral Charge Control Model of Bipolar Transistors, Bell Syst.Techn.J., 49 1970 , p.827
- Simplification of DC characterization and Analysis of Semiconductor Devices, HP application note 383-1
- EEsof, Xtract Supplement Manual, March 1994, HP-EEsof.
- Application Note AN 1202-4: Advanced Bipolar Transistor Modeling Techniques, IC_CAP marketing group
- T.Zimmer: Contribution à la modélisation des transistors haute fréquence, Thèse à l'Université de Bordeaux I, 17.7.1992.
- Zimmer, Meresse, Cazenave, Dom, 'Simple Determination of BJT Extrinsic Base Resistance', Electron.Letters, 10.10.91, vol.27, no.21, p.1895
- Paul Schmitz, Vector Measurements of High Frequency Networks, Hewlett-Packard, Publication HP5958-0387, April 1989
- P. van Wijnen, On the Characterization and Optimization of High-Speed Silicon Bipolar Transistors, PhD Thesis University of Delft, 1992, publshed 1995 by Cascade Microtech, Inc., Beaverton, Oregon
- F.X.Sinnesbichler, Großsignalmodellierung von SiGe-Heterobipolartransistoren für den Entwurf von Millimeterwellenschaltungen, Dissertation (PhD Thesis) Universität München, VDI-Verlag, Düsseldorf, ISBN 3-18-332709-0
- B.Ardouin, Contribution à la modélisation et à la caractérisation en hautes fréquences des transistors bipolaires a heterojunction Si/SiGe, PhD Thesis University Bordeaux, École doctorale de sciences physiques et de l'ingenieur, December 2001

Paras.resistor modeling

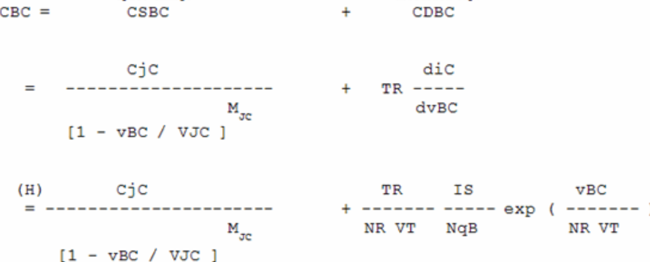
- J.Berkner: A Survey of CD-Methods for Determining the Series Resistances of Bipolar Transistors Including the new ΔI_{Sub} Method, SMI GmbH, Frankfurt/Oder, published in the proceedings of the European IC-CAP User's Group Meeting 1994, October 10-11, Colmar, France
- DC and paras.resistor modeling:
  D.MacSweeny et.al.: Modeling and Parameter Extraction for Lateral Bipolar Devices Using IC-CAP Sub-Circuits, Proceedings of the 1996 Europ.IC-CAP User Meeting in The Haague, Netherlands
- T.Ning, D.Tang: Method for Determining the Emitter and Base Series Resistance of Bipolar Transistors, IEEE Transactions on Electron Devices, vol.ED-31, pp.409-412, April 1984
- CV and TF modeling:
  J.Berkner: Parasitäre Effekte bei der SPICE-Modellparameterbestimmung für integrierte Bipolartransistoren, Halbleiter Elektronik Frankfurt/Oder GmbH, published in the proceedings of the European IC-CAP User's Group Meeting 1993, June 22-23, Esslingen, Germany
- TF modeling:
  Kendall et al.: Direct Extraction of Bipolar SPICE Transit Time Parameters Without Optimization, published in the proceedings of the US IC-CAP User's Group Meeting 1993.

For providing valuable feedback on the extraction methods presented in this document, the author would like to especially acknowledge:
Mr. Jörg Berkner, Infineon, Munich, Germany
Mr.Dermot MacSweeny from the NMRC in Cork, Ireland,

Transit Time Modeling

For a better overview, the Gummel-Poon basic capacitor equations are repeated first from the CV chapter:

Provided that: vBE = vB'E' and vBC = vB'C', the capacitors in the Gummel-Poon model given in the introduction chapter with equations (O) ... (S) are:



And

$$CBE = \underbrace{CSBE}_{\text{space charge}} + \underbrace{CDBE}_{\text{diffusion cap.}}$$

$$= CSBE + \frac{dQ_F}{dv_{BE}} \quad \text{with QF equals the total forward charge}$$

$$\sim \frac{CjE}{[1 - vBE / VJE]^{M_{JE}}} + TFF \frac{diC}{dvBE}$$

$$\text{(H)} = \frac{CjE}{[1 - vBE / VJE]^{M_{JE}}} + \frac{TFF}{NF\ VT} \frac{IS}{NqB} \exp\left(\frac{vBE}{NF\ VT}\right)$$

again from the introduction chapter and additionally with:

$$TFF = TF \left\{ 1 + XTF \left[\frac{if}{if + ITF} \right]^2 \exp\left[\frac{vBC}{1{,}44\ VTF} \right] \right\}$$

and the ideal Collector current If from C
CSBi models the space charge and CDBi the diffusion capacitance between Base and Emitter or Base and Collector respectively. vBE and vBC are the stimulating voltages.

For CBE, and its diffusion capacitor part CDBE, the exact definition is (private communication with D.Celi, ST Crolles, France):

$$C_{DBE} = \frac{dQ_F}{dv_{BE}} = \frac{dQ_F}{di_C} \cdot \frac{di_C}{dv_{BE}} = TF_{ac} \cdot gm$$

where TF_{ac} is the small signal transit time which is different from TFF. However, in this toolkit, we keep the simplified equation of CV-2 and a corresponding raw parameter extraction, and apply a quick optimizer run to get the true TF parameter.

Modeling the Diffusion Capacitor CDBC

Extraction of TF, ITF and XTF

TF - ideal forward transit time
XTF - coefficient for bias dependence of TF
ITF - high-current parameter for effect on TF

For forward active operation of the transistor, the AC behavior is modeled by CBC and CBE. In this operating mode, the already modeled CV-capacitance CSBC dominates over the diffusion capacitance CDBC , while in another equation, the more important term is CDBE. This chapter covers the modeling of CDBE.

CDBE is described by the bias-dependent transit time T_{FF}.

TFF is modeled with the formula:

$$TFF = TF \left\{ 1 + XTF \left[\frac{if}{if + ITF} \right]^2 \exp\left(\frac{vBC}{1{,}44\ VTF} \right) \right\}$$

with the ideal forward Base current i_f

TF is the ideal forward transit time modeling the 'excess charge'. The parameters XTF, and ITF cover the operating point dependence from the DC bias iC ~ if , while VTF describes the dependence from vCB ~ vCE.

Pre-considerations concerning the measurement

Like in the previous chapter, the parameter estimation is again performed using a

simplified model, whereas the parameter fine-tuning is finally done during an optimizer run using the full set of SPICE model equations.
Referring to appendix B, it can be shown that the transistor's h21(f)-parameter behaves frequency wise like a low-pass filter with the transfer function

```
                1 - p / p01(iC,vCE)
h21(f)  =  ß  ---------------------
                1 + p / pp1(iC,vCE) with   p = j * 2PI * f
```

or inverted

$$f_{T\ 1-pole}(i_C, v_{CE}) = \frac{1}{2*PI*T_{FF}(i_C, v_{CE})}$$

where $f_{T1-pole}$ is a function of the bias current i_C and the bias voltage v_{CE}

In many publications, like e.g. /Sinnesbichler p.106/, it is mentioned that the transit time after the above equation is:

$$T_{FF}(i_C, v_{CE}) = \frac{1}{2*PI*f_{T\ 1-pole}(i_C, v_{CE})} - RC*CBC$$

In this case, the TFF used for modeling is RC*CBE smaller than the value converted from f_T.

In some other publications, this formula is extended to

$$T_{FF}(i_C, v_{CE}) = \frac{1}{2*PI*f_{T\ 1-pole}(i_C, v_{CE})} - (RC + RE + RB/ß)*CBC$$

or after /B.Ardouin, p.198/

$$T_{FF}(i_C, v_{CE}) = \frac{1}{2*PI*f_{T\ 1-pole}(i_C, v_{CE})} - (CBE + CBC)\cdot\frac{v_T}{i_C} - (RCX + RE)\cdot CBC$$

In practice, however, with the goal of a direct extraction of the TFF parameters followed by a post-optimization, the additional terms can be neglected and the simple equation (TFF-1) is sufficiently correct.

Now back to the parameter extraction

We will first consider the extraction of the parameters TF, ITF and XTF. VTF will be covered later. This means, a special measurement for fT(iC) and later for fT(iC, vCE) is needed. As we assume a 1-pole low-pass for h21, the gain-bandwidth product is a constant. Therefore it is sufficient to measure a h21(iC, vCE) at a fixed frequency higher than the the -3dB frequency.

In other words, this fixed frequency should be from a -20dB/decade range of MAG[h21(freq, iC, vCE)] . This measurement frequency can be found when transforming the measured rBB' S-parameters to H-parameters (using the TwoPort function). From the dB-plot of ABS(h21(f)) versus log(frequency) we determine a frequency where the slope fits a -20dB/decade roll-off.

> **Note**
> If your MAG[h21] does not follow the -20dB/decade law, there is probably a so-called over-deembedding. This means more parasitics are subtracted than present in reality. The opposite de-embedding problem, under-deembedding, does not affect the slope, but it can show up like a second -20dB/decade slope shifted in frequency.

This frequency is now used as a fixed frequency f-20dB for the S-parameter measurements of this setup. The underlying DC bias values are a swept iC and a constant and small value of vCE (to neglect the VTF effects). Then, these S-parameters are converted into H-parameters and we get for the constant gain-bandwidth product of this assumed one-pole low-pass filter:

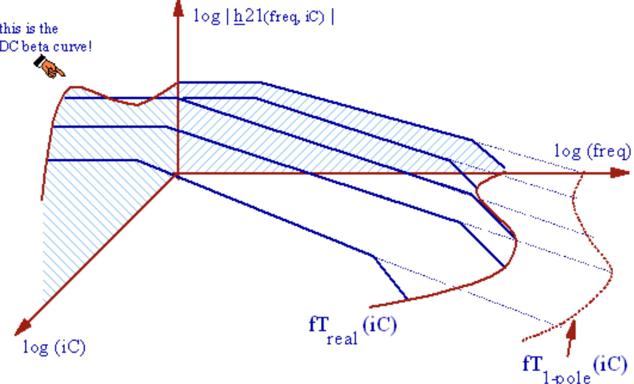
$$1*f_{T\,1-pole}(i_C, v_{CE}) = |h_{21}(i_C, v_{CE})|*f_{-20dB}$$

or

$$f_{T1-pole}(i_C, v_{CE}) = |h_{21}(i_C, v_{CE})|*f_{-20dB}$$

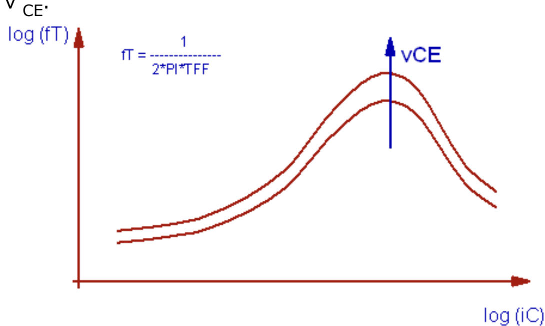
$f_{T1-pole}$ after equ.TFF-2 is valid for all DC bias conditions, i.e. for the whole bias-dependent array of h21. This new array fT1-pole is then introduced into (TFF-1), what gives the bias-dependent array of TFF to be fitted.

Figure below shows log|h21| as a two-dimensional function of the Collector current i_C and the frequency freq. It shows the transit frequency with and without simplification (Appendix B). The dependence of v_{CE} is neglected for simplification.



Log |h21| as a function of log (iC) and log (f) for v_{CE}=const.

Usually, f_T is plotted against i_C . This is the typical diagram published in many data sheets. the following figure shows such a curve, also including the dependence of f_T from v_{CE}.



Note
For a correct modeling, check the fT curve at low iC for so-called self-biasing! This effect occurs if the RF signal power at the Base is in the range of the DC bias power. Under this condition and considering the non-linear diode characteristic at the Base of the transistor, the rectified AC signal will contribute to the DC bias! A flat trace of the fT curve at low Collector current is an indicator for that effect.

For more detailed examples about how the RF power level might affect the fT curve, see literature P.v.Wijnen, chapters 3 and 4, and the IC-CAP examples on non-linear RF modeling, available from the author.

Pre-considerations concerning the model equation:

As cited at the beginning of this section, we start with:

$$TFF = TF \{ 1 + XTF \left[\frac{if}{if + ITF} \right]^2 \exp \left[\frac{vBC}{1,44\ VTF} \right] \}$$

with the ideal Collector current i_f v_{BC} as well as v_{BE} are the DC bias voltages at the operating point.

In this equation, i_f is the ideal Collector current. If we consider currents below IKF ,we can set if = ic . After the extraction of the parameters of this section, we will use a final optimization on the S-parameter curves, which will eliminate this small error.

Therefore the curve-to-be-fitted is:

$$TFF = TF \{ 1 + XTF \left[\frac{iC}{iC + ITF} \right]^2 \exp \left[\frac{-\ vCB}{1,44\ VTF} \right] \}$$

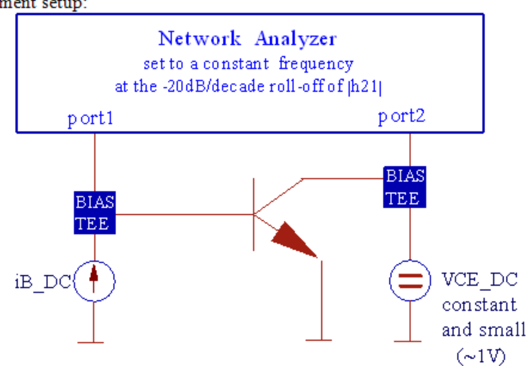
If we choose vCB ~ 0, we can further simplify and get finally:

$$TFF = TF \{ 1 + XTF \left[\frac{iC}{iC + ITF} \right]^2$$

Validity of (TFF-4): iC<IKF and linear forward operating point with vCB~0 or as small as possible.

Performing the measurement

Measurement setup:



Measurement setup for the TFF parameter measurement

First the network analyzer is set to a constant frequency on the -20dB/decade roll-off of |h21(f)|. The used test frequency had been estimated from the S_to_H parameters of the RBB measurement. More details were given above.

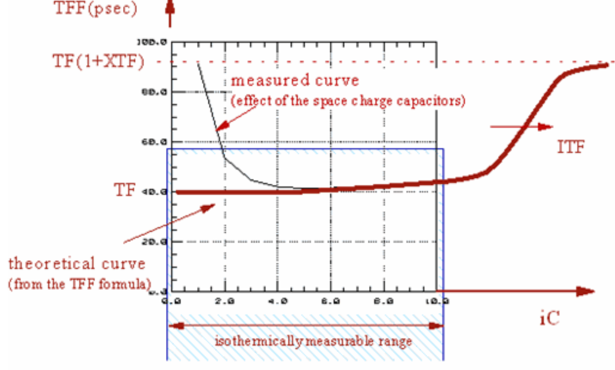
Next the transit frequency for a 1-pole low-pass model is calculated as given in (TFF-2):

```
fT1-pole(iC,vCE)  =  |h21(iC,vCE)| * f-20dB
```

Then we calculate TFF = 1/(2 PI fT1-pole) as the bias-depended total transit time.

Extracting the parameters

The following figure highlights the theoretical trace of TFF of equation. This figure shows the theoretical curve in addition to the measured one to make things more clear. The real measured curve is overlaid by the space charge capacitor effects for low collector currents. This can be seen also from equation in the chapter on CV modeling. For a detailed description of this effect, see /Berkner 1993/



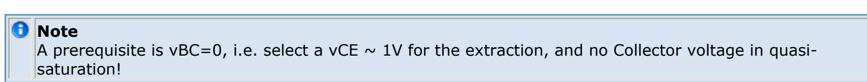
The theoretical transit time TFF (bold) as a function of iC for vCB = 0 compared to the theoretical trace.

Due to these overlay and measurement problems, it had been found that a pretty simple and straight-forward extraction technique can be applied that gives nevertheless quite reasonable results. This method is explained below. There exist some more complex strategies, but the extraction results may be not much better. As sketched in the above figure, this is mainly because it is not possible to force such a high Collector current that the trace can be obtained from measurements without being overlied and distorted by thermal self-heating effects.

How to proceed

TF

TF is extracted as the minimum value of TFF.



Note
A prerequisite is vBC=0, i.e. select a vCE ~ 1V for the extraction, and no Collector voltage in quasi-saturation!

XTF

The behavior of TFF is given in the above figure4. It is difficult to measure for a higher Collector current due to thermal limitations. So XTF is estimated from the trace of TF at max. applicable Collector bias current under the assumption that it would be TFF at infinite current:

MAX(TFF) = TF (1 + XTF)

or

```
         MAX(TFF)
XTF =  ----------    - 1
            TF
```

ITF

Referring to the same measurement restrictions as above, a good first-order estimation of ITF is related to the max. Collector current measured:
ITF = MAX(iC_meas) / 2

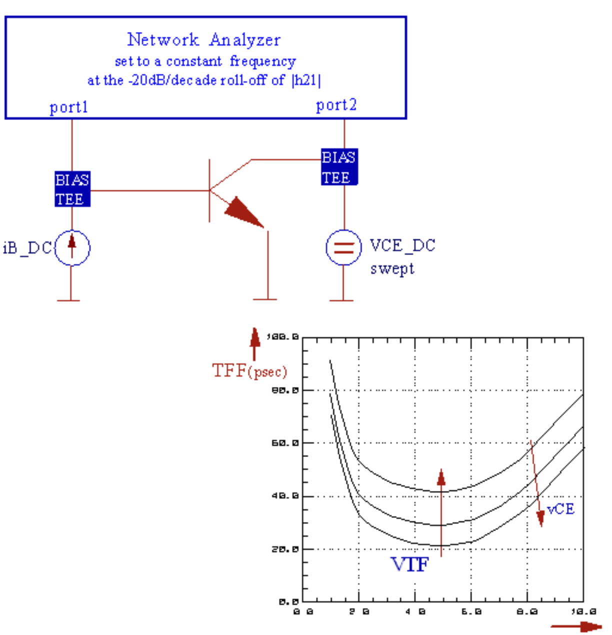
Again, since the end of the TFF trace is often not measurable, correct this estimation by ITF=5*ITF(equ.TFF-6).

This ITF extraction method follows the idea of the Base resistor IRB parameter extraction!

Note
The extraction of the parameter TF, ITF and XTF is performed simultaneously with parameter VTF in a single extraction routine.

Extraction of VTF

VTF Voltage describing the vBC dependence of TF
Finally, we consider also the vCE sweep.
Measurement setup:



The transit time TFF as a function of iC and vCE

VTF can be obtained from for a fixed value of iC:

$$TFF = const \cdot e^{\frac{-vCB}{1.44 \cdot VTF}}$$

Or

$$\frac{TFF1}{TFF2} = \frac{\exp[-v_{CB1} / 1{,}44\ VTF]}{\exp[-v_{CB2} / 1{,}44\ VTF]} = \exp\left[\frac{v_{CB2} - v_{CB1}}{1{,}44\ VTF}\right]$$

This gives:

$$\ln\left[\frac{TFF1}{TFF2}\right] = \frac{v_{CB2} - v_{CB1}}{1{,}44\ VTF}$$

And finally:

$$VTF = \frac{v_{CB2} - v_{CB1}}{1{,}44 * \ln\left(\frac{TFF_1}{TFF_2}\right)}$$

After the extraction of these four parameters for CDBC , we will next run an optimization to improve the fitting of the fT plot. However, it is very important, that we do not forget to also optimize the S-parameter fittings for all bias conditions after that (setup rbb).

Note
Notes:
From the pre-considerations given above, the plots RBB vs. iB and TFF vs. iC represent curves which had been extrapolated from the S-parameters. So, the S-parameter measurement is "the real world" and the fitting should be optimized in this domain.

Again, all the AC extraction methods need absolute clean measurements and elimination of parasitics by de-embedding techniques. Otherwise no curve fitting might be possible or the parameters obtained might make no physical sense.

What to do in IC-CAP

1. open setup "/gp_classic_npn/nwa_extr/tf_vbe_vce",
2. select a -20dB/decade frequency from the plot 'mag_h21vsf' of setup 'rbb' (middle

mouse click)and enter it to input 'freq'
3. import the de-embedded data of this -20dB/decade frequency into setup 'tf_vbe_vce' from the .mdm file
4. perform transform "calc_TFF"
5. check the plots "ft_vbe" and "*TFF*"
6. perform transform "e_TF_ITF_XTF_VTF"
7. simulate with the extracted parameters

How to optimize the S-parameter Setup in IC-CAP

For the data supplied with the file 'gp_classic_npn.mdl', this S-parameter sequence is best:

1. In setup "/gp_classic_npn/nwa/tf_vbe_vce", run the optimizer transforms "to_TF_ITF_XTF" to optimize for the smallest vCB bias., then, execute "o_VTF" to opotimize the fitting for the biggest vCB bias finally, call "o_TF_ITF_XTF_VTF to optimize all bias points"

1. Go back to setup "rbb" and optimize the S-parameter fitting of the RB, IRB and RBM setup by running "o_RB_IRB_RBM"
2. In the same setup, execute finally "o_TFF", what optimizes all TFF parameters at all bias conditions of setup "rbb".

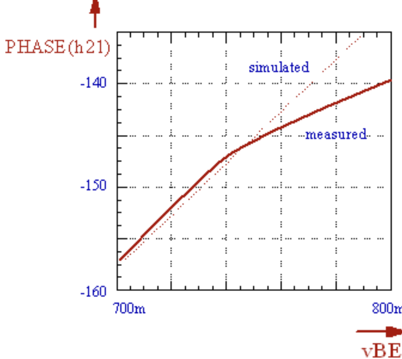
See also macro "extract_n_opt_NWA"

Note
Depending on the device type, it has been observed that the fT-fitting affects also the rBB-fitting.

Extraction of PTF

PTF excess phase at frequency 1/(2PI * TF)

Implemented into the model as a 2nd order all-pass Bessel-function, this parameter can be used to add some extra phase to the RF simulation curves. It can be obtained when plotting the phase of h21 of the TFF measurement from above (fig.TFF-5)versus vBE /T.Zimmer/. As a limitation of the method, the measurable range is again only covering a small part of the desired curve. Therefore it is advisable to use an optimizer run in order to get the value of PTF.



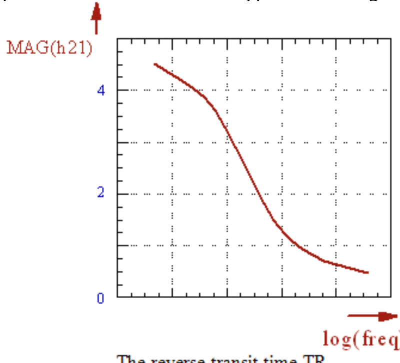
Measured and simulated phase (h21) before optimization.

Extraction of TR

TR ideal reverse transit time

The reverse transit time of the Gummel-Poon model is modeled by one parameter only, TR.

TR can be obtained from an h21 measurement versus frequency at a typical reverse operating point. An optimizer run on the S-parameters of this setup is used to obtain the parameter value from a typical starting value, e.g. TF * 100.



The reverse transit time TR.

Note
After /Sinnesbichler/, TR can be optimized in the S22 and S12 plot of the reverse biased

S-parameters. In this publication, the reverse biasing of the S-parameter is for example:
1st order DC sweep: vCE = -1 ... -5V
2nd order DC sweep: vBE = -0,7 ... -1V

Note
TR can also be obtained from pulse measurements using an oscilloscope. In many cases, TR is the dominant parameter in such a setup, often more important than TFF and the junction capacitances. Therefore, for the modeling of transistors in digital applications, the TR modeling is a must.

YParameter Modeling

Contents

- *Modeling of Transistors in the Operating Point* (iccapmhb)
- *S2Y-Parameter Transistor Modeling GUI* (iccapmhb)
- *Testing S-Parameters for Small Signal Transistors* (iccapmhb)

Modeling of Transistors in the Operating Point

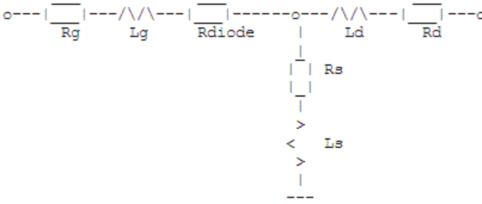
The Modeling of Transistors in The Operating Point

FIRST EXAMPLE: GaAs TRANSISTOR

If a transistor is operated at a fixed bias point only, it is not necessary to model it also for large signal behavior. In this case, a small signal model is used instead. This chapter gives a examples on how to do this.

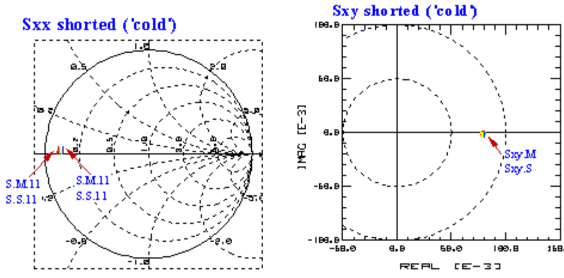
It is assumed that a packaged GaAs transistor shall be modeled. The package contributes with series components like inductors and resistors. This will be modeled in a first step. The 'inner' small signal transistor is modeled afterwards.

In order to separate the serial components of the package, the transistor is over-biased and behaves like a good SHORT (so-called cold measurement). This is achieved by biasing a strong current iG into the Gate and drawing half of its value out of the Drain with the Source connected to Ground. The assumed underlying equivalent schematic is sketched in the following figure:



Schematic of the serial components of the transistor package under the condition of a shorted inner transistor (Figure 1)

The following figures show the measurement (and simulation for later reference) of the transistor under this bias condition (COLD measurement).



Measurement and simulation result of the packaged transistor in COLD measurement mode (Figure 2)

We transform the measured S-parameters to Z, since the assumed schematic of Figure 1 is a TEE structure.

$$\begin{pmatrix} Z11 & Z12 \\ Z21 & Z22 \end{pmatrix} = \begin{pmatrix} Rg + rdiode + Rs + j\omega(Ls + Lg) & Rs + j\omega Ls \\ Rs + j\omega Ls & Rd + Rs + j\omega(Ls + Ld) \end{pmatrix}$$

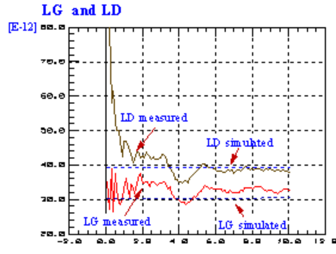
Solved for the components of the schematic, we get:

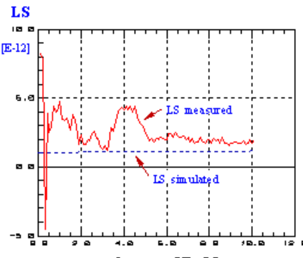
```
Rs = REAL(Z12)                       Ls = IMAG(Z12)
Rg = REAL(Z11)  - Rs - rdiode        Lg = IMAG(Z11)  - Ls
Rd = REAL(Z22)  - Rs                 Ld = IMAG(Z22)  - Ls
```

After having plotted these parameters versus the frequency (fig.3 and fig.4), we can easily check the applicability of our assumed underlying schematic. If the parameters have no major frequency dependence, then the assumption is correct. In our case, this is true, and we calculate the mean parameter values out of these plots. A simulation with these parameters is also illustrated in figures 3,4 and 2. The slight difference stems from the fact that the measured data obviously represents not quite simply a TEE schematic.

!opr_pt3_resistor values.gif !

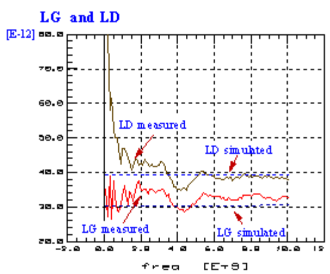
Resistor values of the schematic of fig.1 and the simulation result using the extracted mean parameter values (Figure 3)

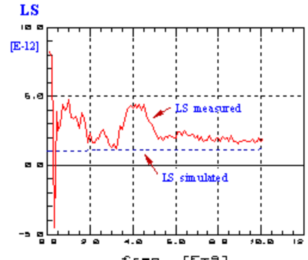




Inductor values of the schematic of fig.1 and the simulation result using the extracted mean parameter values (Figure 4)

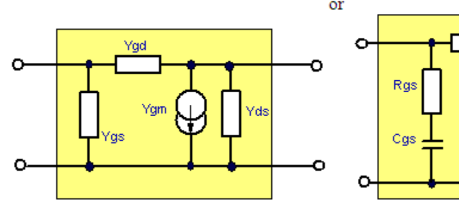
In the second step, we bias the transistor in the operating point. The measured S parameters (fig.6) are transformed to Z, and the previously obtained inductors and resistors are de-embedded. The result is depicted in fig.5.



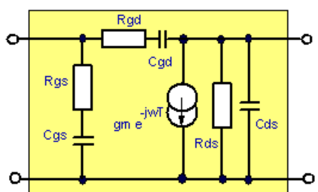


Measurement (and simulation results for later reference) of the transistor biased in operating point (Figure 5)

Then, the result is converted to Y parameters, since we now assume the 'inner transistor' schematic of Figure 6, which represents a PI structure:



or



Schematic of the Inner Transistor (Figure 6)

The S to Y converted parameters are now interpreted following the assumptions of fig.6a:

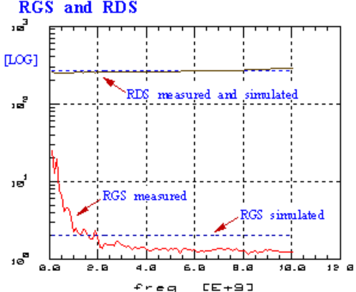
$$\begin{pmatrix} Y11 & Y12 \\ Y21 & Y22 \end{pmatrix} = \begin{pmatrix} Ygs + Ygd & -Ygd \\ Ygm - Ygd & Yds + Ygd \end{pmatrix}$$

$Ygd = -Y12$
$Ygs = Y11 + Y12$
$Yds = Y22 + Y12$
$Ygm = Y21 - Y12 = gm * \exp(-j\omega T)$

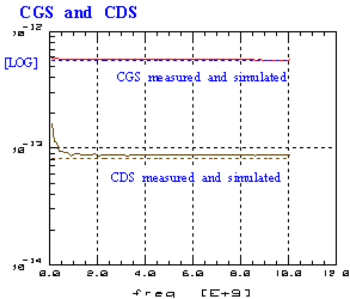
and solved for the component values as given in fig.6b:

Cgs and Rgs: Cgs = -1 / (2*PI*freq*IMAG(1 / Ygs))
Rgs = REAL(1 / Ygs)

Cds and Rds: Cds = IMAG(Yds) / (2*PI*freq)
Rds = 1 / REAL(Yds)

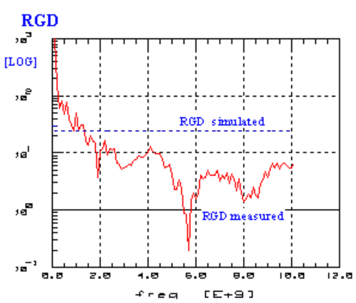


Trace of Rgs and Rds versus frequency (Figure 7a)

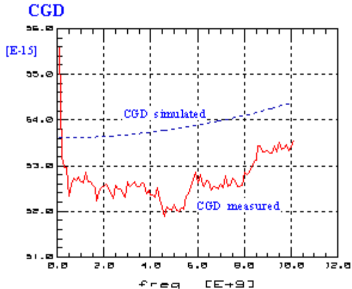


Trace of Cgs and Cds versus frequency (Figure 7b)

Cgd and Rgd: Cgd = -1 / (2*PI*freq*IMAG(1 / Ygd))
Rgd = REAL(1 / Ygd)

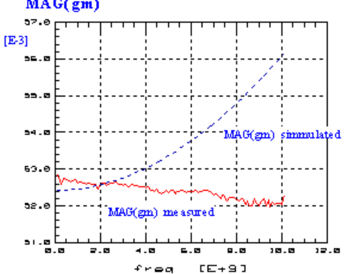


Trace of Rgd versus frequency (Figure 8a)

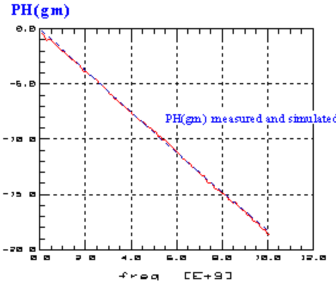


Trace of Cgd versus frequency (Figure 8b)

Ygm: Ygm = Y21 - Y12



Trace of MAG(gm) versus frequency (Figure 9a)



Trace of PH(gm) versus frequency (Figure 9b)

After this modeling, we can check the quality of the developed equivalent schematic. The simulation result, which matches properly the measured data, is depicted in fig. 5.

Modeling with Parameter Tuning

Parameter tuning is another method to determine the values of the circuit components of small signal schematics. Again, a GaAs transistor is assumed. The underlying schematic is given in fig.1.

Small signal GaAs transistor schematic

The S-parameter modeling can be done in the following order:

1. All components of the equivalent schematic are 'switched-off'. I.e. all resistors set to 1mOhm, inductors and capacitors to 1 Atofarad resp. -henry (10-18).
2. Modeling of S21 using Rout for f->0.
   gm is known from DC-measurements and should not be changed. Therefore, R3 is used for the fitting.
   (Attention, if id>5mA, self-heating of the transistor chip may have occurred. In such a case, gm_DC will be different from gm_AC !!)
3. Modeling of S11:
   R1 is the ohmic value for infinite frequency (may be overlaid by R3)
   C1 fits the phase shift

   > Note: if Sxx shows a notch, we should consider that the power flows for low frequencies through C12 to the port 2 of the NWA and passes through C1 or C2 only for higher frequencies. This is the case for C12 > C2 resp. C1.
   > In such a case, we should remember that the opposite NWA port represents 50 Ohm and that these 50 Ohm will show up in the S11-curves!

Modeling of **S22**:
Rout is the ohmic value for low frequencies
R2 is the ohmic value for infinite frequencies (may be overlaid by R3)
C2 turns the phase

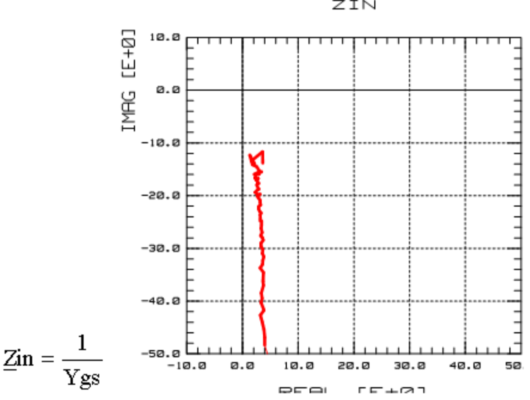
1. Fine-tuning of **S21**:
   First, the influence of L3 should be checked
   Then, fine-tuning C12 and R3, we improve the upper frequency fitting.
2. Finally, L1 and L2 are tuned.

!opr_pt_small_signal1.gif!

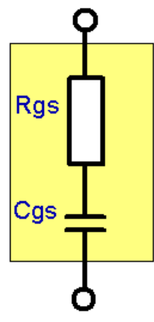
Solved for the conductance of the Y schematic branches, we obtain

$$
\begin{aligned}
Ygd &= -Y12 \\
Ygs &= Y11 + Y12 \\
Yds &= Y22 + Y12 \\
Ygm &= Y21 - Y12 = gm \cdot \exp(-j\ \omega\ TAU)
\end{aligned}
$$

ZIN



For ZINlike given above, we assume the following schematic:



what leads to the component values:

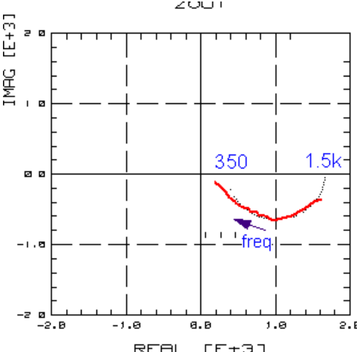
Cgs and Rgs: Cgs = -1 / (2*PI*freq*IMAG(1 / Ygs))

ZTRANS

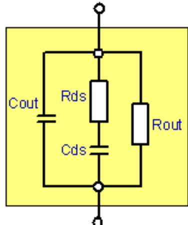
$$\underline{Z}\text{trans} = \frac{1}{\text{Ygd}}$$

Cgd and Rgd: Cgd = -1 / (2*PI*freq*IMAG(1 / Ygd))
Rgd = REAL(1 / Ygd)

ZOUT



$$\underline{Z}out = \frac{1}{Yds}$$



Neglecting Cout for the moment, we can separate for
Rout, Cds and Rds, Cout:
Rout = 1 / (REAL(Yds)), extrapolated to 0Hz
then we can calculate
Zds = (Yds - 1/Rout)^-1
what finally gives Rds = REAL(Zds)
and Cds = -IMAG(Zds)^-1 / (2*PI*freq)

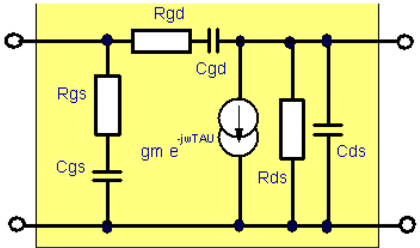
C_{out} is then finetuned for high frequencies in S22

gm

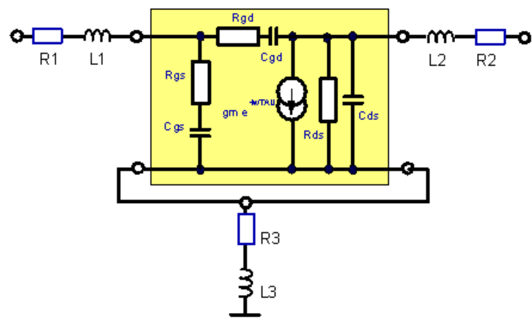
Ygm: Ygm = Y21 - Y12

gm = MAG(Ygm)

TAU = -PH(Ygm) / (2PI * freq)



Developing A Small Signal Model For Mos Transistors



$$Z_{TEE_PI} = \left[Z_{TEE} + Y_{PI}^{-1}\right] \quad (1)$$

$$Z_{TEE} = \begin{pmatrix} Rg+Rs & Rs \\ Rs & Rd+Rs \end{pmatrix} + j\omega * \begin{pmatrix} Lg+Ls & Ls \\ Ls & Ld+Ls \end{pmatrix} \quad (2)$$

$$Y_{PI} = \begin{pmatrix} j\omega(Cgs+Cgd) & -j\omega Cgd \\ gm * e^{-j\omega T} - j\omega Cgd & Gds + j\omega(Cds+Cgd) \end{pmatrix} \quad (3)$$

For frequencies below 1/10 fT, analytical expressions of the Z_{TEE_PI} parameters can be derived which represent the specific frequency behavior of (1):

$$REAL(Z_{TEE_PI_ij}) = REAL(Z_{TEE_ij}) + \frac{A_{ij}}{\omega^2 + B} \quad \text{for } i,j \in 1,2 \quad (4)$$

$$\frac{1}{\omega} IMAG(Z_{TEE_PI_ij}) = \frac{1}{\omega} IMAG(Z_{TEE_ij}) - \frac{E_{ij}}{\omega^2 + B} - \frac{F_{ij}}{\omega^2(\omega^2 + B)} \quad \text{for } i,j \in 1,2 \quad (5)$$

$$y_intersect = REAL(Z_{TEE_kl}) - \frac{A_{kl}}{A_{ij}} * REAL(Z_{TEE_ij})$$

$$slope = \frac{A_{kl}}{A_{ij}}$$

Publications:

G.Dambrine, F.Heliodore, E.playez, A NEw Method for Determining the FET Small-Signal Equivalent Circuit, IEEE Trans. on MIcrowave Theory and Techniques, vol.36, no.7, July 1988

M.Berroth, R.Bosch, High Frequency Equivalent Circuit of GaAs FETs for LArge Signal Applications, IEEE Trans. on Microwave Theory Tech., Vol.MTT-39, No.2, 1991, p.224

D.Lovelance, J.Costa, N.Camilleri, Extracting Small-Signal Model Parameters of Silicon MOSFET Transistors, IEEE, MTT-S intern.Microwave Symp.Digest, 1994, pp.865-868

S.Lee, H.K.Yu, C.S.Kim, J.G.Koo, K.S.Nam, A Novel Approach to Extracting Small-Signal Model Parameters of Silicon MOSFETS's, IEEE Microwave and Guided Wave Letters, vol.7, no.3, March 1997, pp.75-77

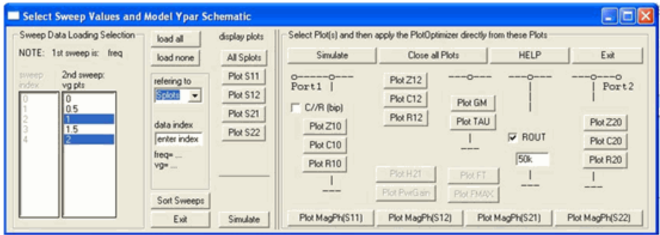
J.P.Raskin, G.Dambrine, R.Gillon, Direct Extraction of the Series Equivalent Circuit Parameters for the Small-Signal Model of SOI MOSFET's, IEEE Microwave and Guided Wave Letters, vol.7, no.12, Dec.1997, pp408-410

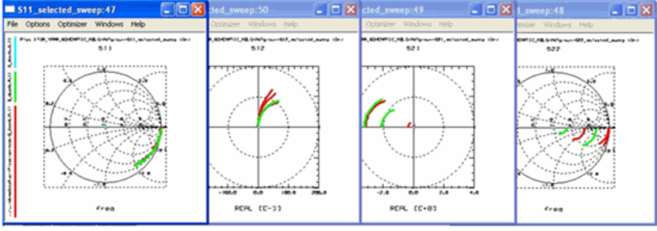
C.C.Meng, G.-H.Huang, Unique Determination of the Equivalent Circuit Parameters for Uniform-Doped GaAs MESFET, The Fifth Symposium on Nano Device Technology, SNDT'98, National Nano Devices Laboratories, National Science Council, Chungli, Taiwan R.O.C.

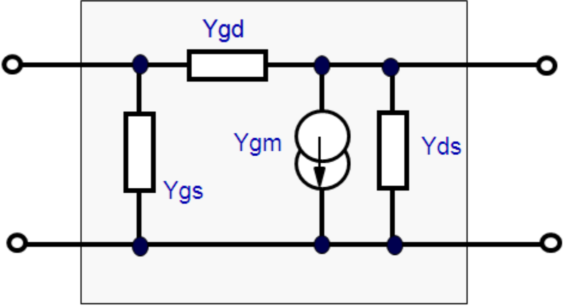
S2Y-Parameter Transistor Modeling GUI

This chapter is a cookbook about how to use the file

demo_features\1_BASIC_MDLG_EXAMPLES\
24_general_Transistor_DC_n_S2Ypar_Schematic_Mdlg\
Xtor_S2Ypar_schematic_Mdlg_PELdep.mdl







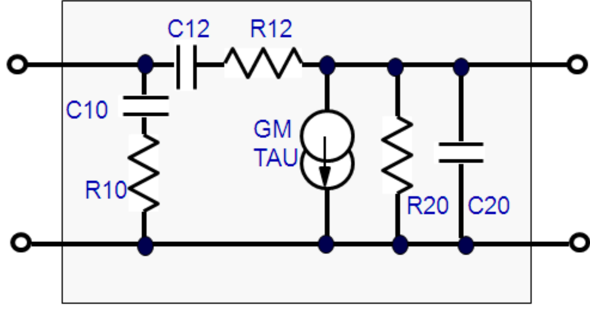
The underlying schematic of Y parameters is a PI schematic

Starting with the Y-parameters of the MOS transistor small-signal schematic

$$\begin{pmatrix} Y11 & Y12 \\ Y21 & Y22 \end{pmatrix} = \begin{pmatrix} Ygs + Ygd & -Ygd \\ Ygm - Ygd & Yds + Ygd \end{pmatrix}$$

Ygd = -Y12Ygs = Y11 + Y12Yds = Y22 + Y12Ygm = Y21 - Y12 = gm * exp(-jωTAU)

How to Proceed



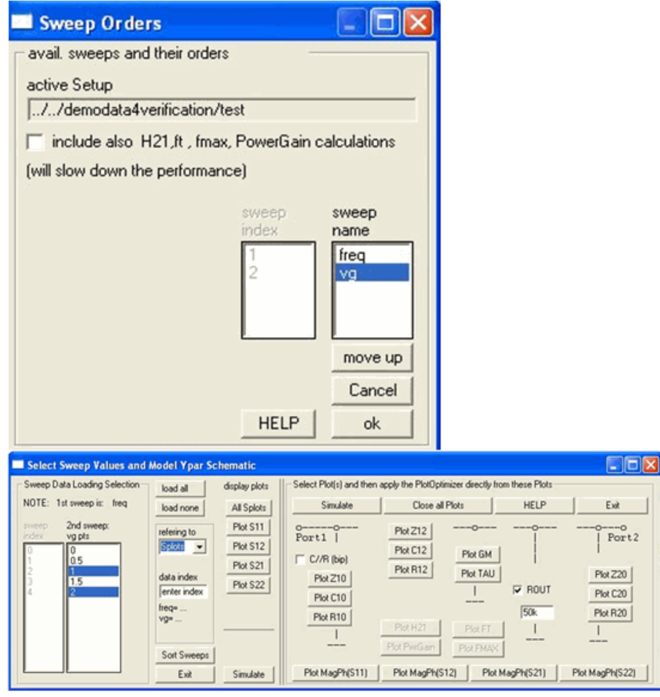
The following individual steps explain how to get the MOS transistor schematics out of the de-embedded S-parameters.
In the IC-CAP GUI, check boxes allow to distinguish between MOS transistors (Cgs in series with Rgs) and bipolar transistors (Cbe in parallel to Rbe).
Also, two models for the output impedance are selectable in the IC-CAP GUI.

1. Convert de-embedded S-parameters to Y
2. Calculate complex impedances and Gm
   Z_10 = (Y11+Y12)^-1

Z_20 = (Y22+Y12)^-1
Z_12 = -Y12^-1
Gm = Y21 - Y12 = gm * exp(-j*2PI*TAU)

3. Calculate finally
   R10 = REAL(Z_10) C10 = -(IMAG(Z_10))^-1 // (2PI*f)
   R12 = REAL(Z_12) C12 = -(IMAG(Z_12))^-1 // (2PI*f)
   GM = MAG(Gm) TAU = PH(Gm) / (2PI*f)
   R20 = (REAL(Z_20^-1)$^{-1)}$ C20 = IMAG(Z_20-1) // (2PI*f)

How to use the GUI



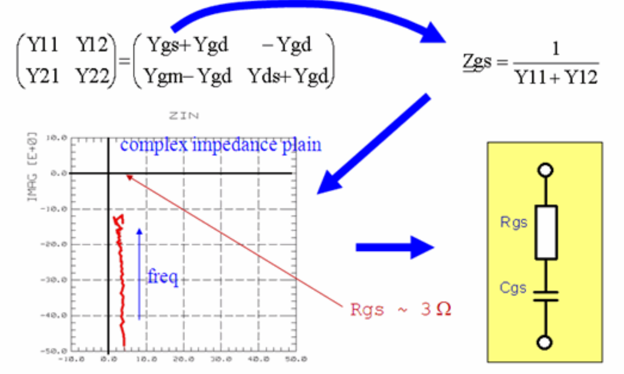
After starting the GUI, you can change the sweep order at your convenience (step 1. above).

Then, a 2nd GUI pops up. On its left, you can select a subset of DC biases for data inspection, but also for your later modeling using the right side of this GUI.

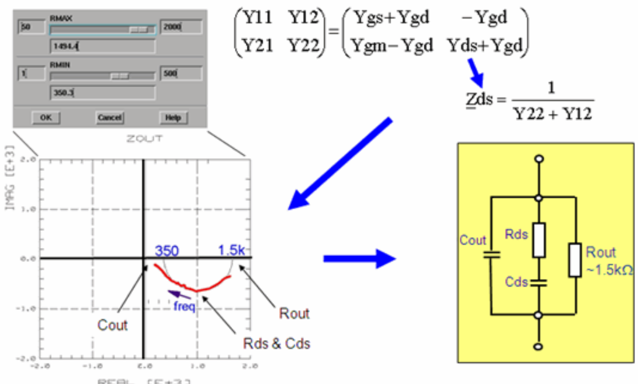
In this right side (step 3. in the slide above), you can specify which output impedance model you want the plots to apply to, and which input inpedance (depending on MOS or bipolar transistors) you want to use.

The following sections can assist you in better understanding the Impedance Plots of the S2Y modeling procedure.

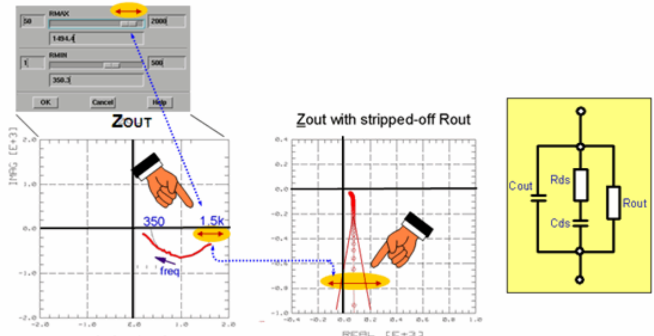
Interpreting the input impedance schematic



From the output impedance to its schematic

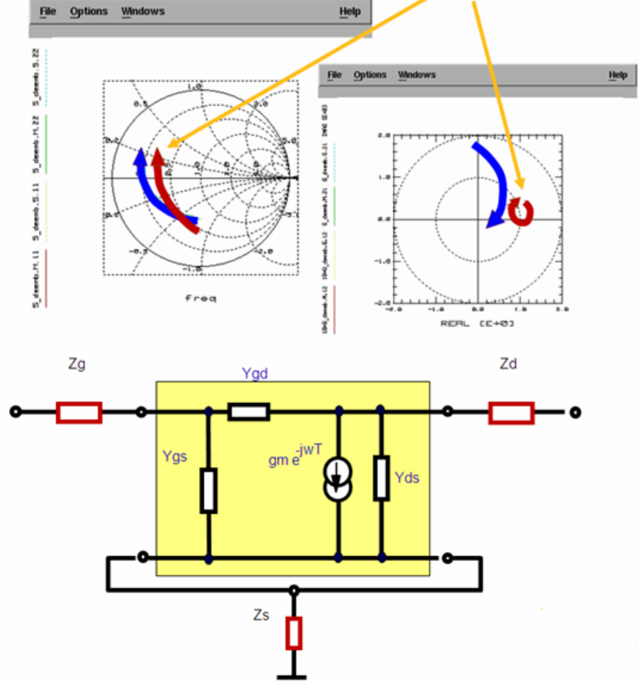


Note: While tuning Rout, keep an eye on the off-stripped impedance plot Zout, representing the rest of the schematic !

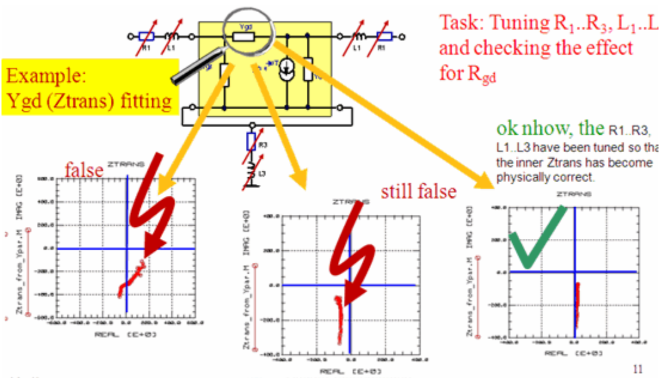


What if the locus curve becomes inductive !!

If the impedance plots of de-embedded S-parameter measurements exhibit inductive behavior (Sxx extend into the upper ohmic part of the Smith Chart, S12 shows a resonance), you should first be suspicious about your de-embedding, and also about the de-embedding layout. However, if a de-embedding verification gives good results (using the THRU dummy), your inner transistor may indeed include inductances (the lines of longer finger structures etc.).



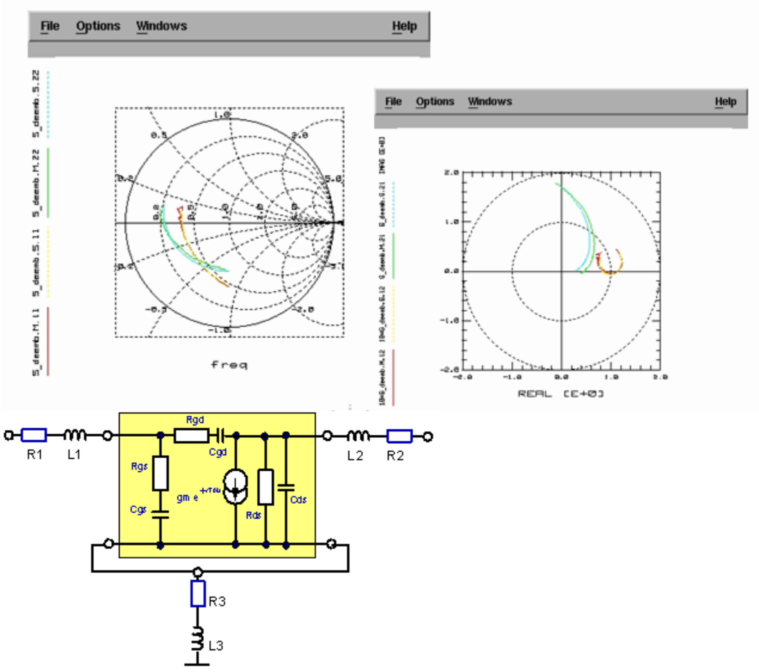
Fine-tuning the series parameters so that the remaining inner Y parameters become physically correct.



You can find the schematic and the parameter values of the series impedances by tuning their values and inspecting the traces of the locus curves of the inner PI schematic.

Note: this method is not implemented in the IC-CAP GUI applied in this chapter.

Modeling Result Examples



Conclusions:

Y-parameter (PI schematic) modeling
is much more sensitive than S-parameter modeling
offers a nice separation of model parameters in the different Yxy plots
gives much inside view to the effects of model parameters

Y-parameter modeling is easy when applying the model-independent Y-Parameter Modeling GUI in IC-CAP

In the demo_features directory, go to \1_BASIC_MDLG_EXAMPLES\24_genl_Xtor_DC_n_S2Ypar_Schematic_Mdlgand load the file Xtor_S2Ypar_schematic_Mdlg_PELdep.mdl

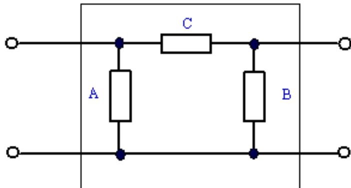
Try the S2y modeling yourself with the demo data included there

Of course it is most interesting when you apply it to your own, measured data.
Try it !!!

Testing S-Parameters for Small Signal Transistors

IC_CAP file:
demo_features\1_BASIC_MDLG_EXAMPLES\29_TestSpars4Circuits\3transistor_PI_mdlg.mdl

This chapter is intended to present some matrix manipulations for the interpretation of S-parameters of MOS and bipolar transistors. This is done by transforming the S-parameters to Y-parameters.



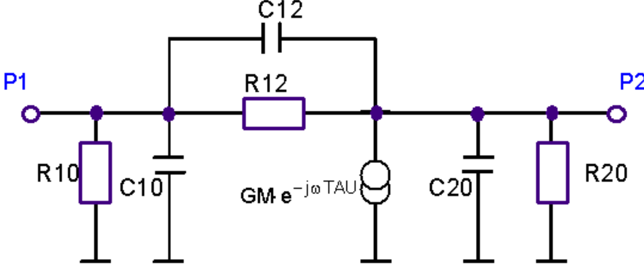
PI circuit interpreted as a Y matrix

$$Y_{PI} = \begin{pmatrix} A+C & -C \\ -C & B+C \end{pmatrix}$$

watch the signs!

Converting S-parameters into a small signal schematic of a bipolar transistor:

Bipolar:



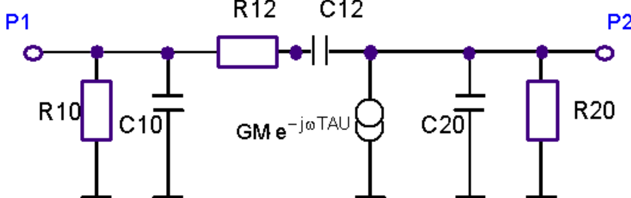
How to obtain the components values out of the S-parameters:

1. Transform the S parameters to Y
2. Calculate the complex conductance and the complex GM
   Y_10 = Y11+Y12
   Y_20 = Y22+Y12
   Y_12 = -Y12
   GM = Y21 - Y12 = gm * exp(-j*ω*TAU)
3. Separate the results
   R10 = REAL(Y_10)^-1 C10 = IMAG(Y_10) / (2PI*f)
   R20 = REAL(Y_20)^-1 C20 = IMAG(Y_20) / (2PI*f)
   R12 = REAL(Y_12)^-1 C12 = IMAG(Y_12) / (2PI*f)
   gm = MAG(GM) Tau = PH(GM) / (2PI*f)

Converting S-parameters into a small signal schematic of a bipolar transistor,

ZBC representing the off-state diode:

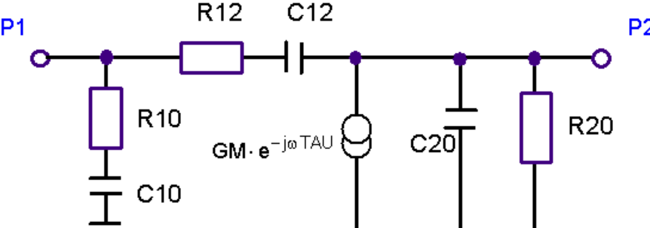
Bipolar:



How to obtain the components values out of the S-parameters:

1. Transform the S parameters to Y
2. Calculate the complex conductances and the complex GM
   Y_10 = Y11+Y12
   Y_20 = Y22+Y12
   Y_12 = -Y12
   GM = Y21 - Y12 = gm * exp(-j*ω*TAU)
3. Separate the results
   R10 = REAL(Y_10)^-1 C10 = IMAG(Y_10) / (2PI*f)
   R20 = REAL(Y_20)^-1 C20 = IMAG(Y_20) / (2PI*f)
   R12 = REAL(Y_12^-1) C12 = -(IMAG(Y_12^-1) * (2PI*f))^-1
   gm = MAG(GM) Tau = PH(GM) / (2PI*f)

Converting S-parameters into a small signal schematic of a MOS transistor:



How to obtain the components values out of the S-parameters:

1. Transform the S parameters to Y
2. calculate the complex conductance and the complex GM
   Y_10 = Y11+Y12
   Y_20 = Y22+Y12
   Y_12 = -Y12
   GM = Y21 - Y12 = gm * exp(-j*ω*TAU)
3. separate the results
   R10 = REAL(Y_10^-1) C10 = -(IMAG(Y_10^-1) * (2PI*f))^-1
   R20 = REAL(Y_20)^-1 C20 = IMAG(Y_20) / (2PI*f)
   R12 = REAL(Y_12^-1) C12 = -(IMAG(Y_12^-1) * (2PI*f))^-1
   gm = MAG(GM) Tau = PH(GM) / (2PI*f)

VBIC

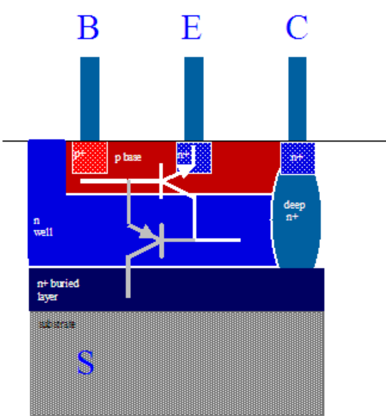
Contents

- *Introduction to VBIC* (iccapmhb)
- *VBIC Background Information* (iccapmhb)
- *Recapitulating the Gummel-Poon Model* (iccapmhb)
- *The VBIC Model* (iccapmhb)
- *Comparing the VBIC and Gummel-Poon Parameters* (iccapmhb)
- *VBIC Modeling Strategy* (iccapmhb)
- *Measuring and Extracting the VBIC Parameters* (iccapmhb)
- *Regression Analysis* (iccapmhb)
- *Acknowledgements, Web info, and Publications* (iccapmhb)

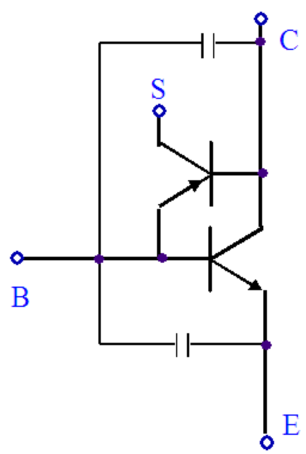
Introduction to VBIC

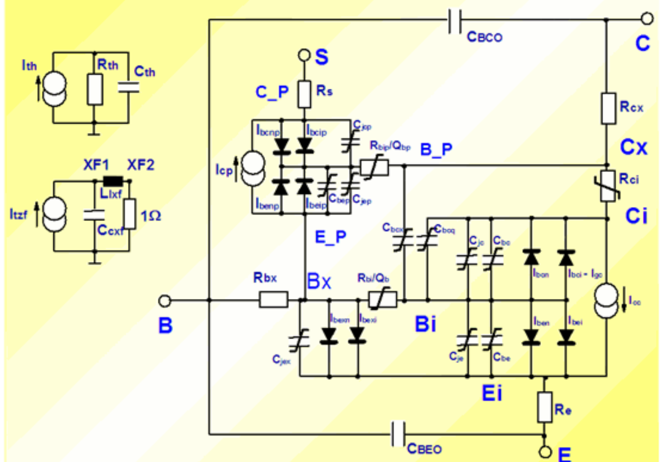
VBIC - The Vertical Bipolar Inter-Company Bipolar Model

VBIC Model Physically



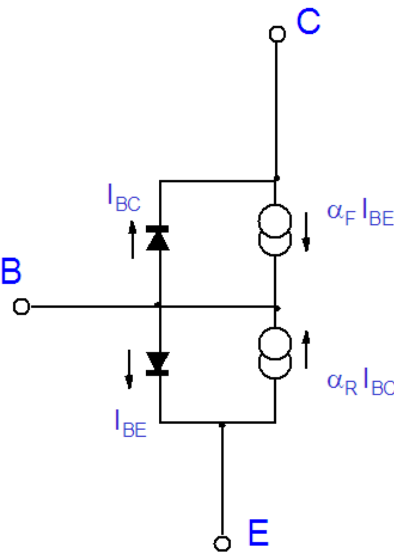
VBIC model Electronically





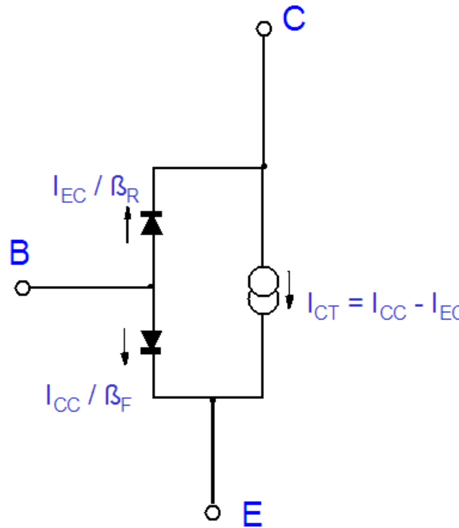
In 1954, Ebers and Moll have proposed a large signal model for bipolar transistors. This model is still the background of today's bipolar transistor models. It describes the fundamentals of the DC behavior. However, low and high current effects, as well as parasitic resistors and dynamic behavior are not yet covered. Fig.1 depicts the topology of the underlying equivalent schematic.

Equivalent Schematic for a Bipolar Transistor after Ebers-Moll



Based on the schematic of fig.1, an alternate, yet mathematically identical formulation has been introduced. Instead of injection currents, it is based on a transport current. This means, the two current sources of the forward and reverse current are combined into a single current source (Fig. 2). The main difference between both models is the different reference currents. The injection version model is based on both diode currents, while the transport version makes use of the currents ICC and IEC, which give the current ICT.

Hybrid Equivalent Schematic of the Ebers-Moll Model



This large signal model has been enhanced and improved many times. Pedersen introduced a classification into three Ebers-Moll model versions EM1, EM2 and EM3, see Getreu.

The EM3 model covers already all essential effects, which are then included in the Gummel-Poon model, published in 1970 see publications . The important advantage of modeling the bipolar transistor with the Gummel-Poon model is especially the very clear and standardized description of many effects by introducing the "integral charge control relation". Therefore, with the years, the Gummel-Poon (G-P) model has become a standard for the modeling of bipolar transistors.

It should be mentioned, however, that this modeling standard is usually a special version of the G-P model, which has been implemented into the simulation program SPICE of the University of Berkeley UCB, California. In some details, this implementation differs from the original G-P formulation. This is especially true for the Early effect. See e.g. the Agilent ADS implementation which allows to select either the simplified UCB SPICE version, or the original Gummel-Poonm paper model.

For modern transistors with the continuous trend to smaller geometries, second order effects become more and more important. Due to higher integration and the necessity to improve the design yield, the need for more precise simulation results and thus to better models has increased. Many companies have therefore developed in-house models, and in some cases made them public. Such a model is the Philips MEXTRAM model. It was developed in 1986 by de Graaff, Klostermann and Jansen.

Later, in 1995, an US industry consortium has proposed a new bipolar model, called

VBIC95. Its goal was to become an accepted standard for today's bipolar transistors. Besides an improved modeling, also including the parasitic PNP transistor of integrated NPN transistors, the VBIC95 is aimed to be as much as possible similar to the standard G-P model. Today, it has changed its name to VBIC.

The following list gives the improvements of the VBIC model compared to the SPICE G-P model:

- Precise implementation of the Base width modulation
- Parasitic Substrate transistor
- Improved Kull model for quasi-saturation
- Enhanced delay time modeling
- Approximating a distributed Base region
- weak avalanche current effects
- consistent treatment of the additional phase factor for small signal and time domain analysis
- improved capacitance model
- self-heating: improved temperature modeling

Since the VBIC is based on the Gummel-Poon model, and since we want to sketch the similarities between both models, we will now recap the details of the standard SPICE G-P model.

A Note on PNP Transistors

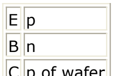
The VBIC model is focusing by its own name on a vertical bipolar transistor, i.e. per definition to a NPN transistor. If applied to a PNP transistor, first check its physical structure. Then decide, what kind of sub-circuit you may want to use for the modeling. See /Getreu/.

Lateral PNP

Lateral PNP has a wide Base, therefore, its 'ß' is low, and as a consequence, its 'ft' is low too.

Substrate PNP

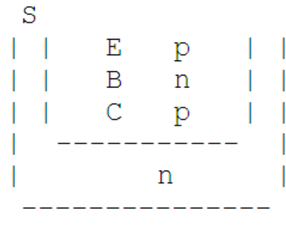
Substrate PNP has usually the following sequence of layers:



E	p
B	n
C	p of wafer

This means, that its Collector is identical to the substrate, i.e. at negative supply voltage. Such transistors usually feature a bad 'ß'.

Vertical PNP (modeling possible using VBIC)



Its Collector current iC flows vertical, 'ß' is high (because the Base can be made very thin), and as a consequence, its 'ft' is high.

VBIC Background Information

VBIC Model History Notes

Rev.1.2 updates (Sept. 24, 1999)

- 3-terminal version defined
- Base-emitter breakdown model added
- Reach-through model added for B-C depletion capacitance
- Homotopy version of code added
- Limited exponential version added
- Completely new code generation added
- C, FORTRAN, Verilog-A, Perl, and MAST code provided
- Bug in psibi mapping with temperature fixed
- Bugs in electrothermal derivatives and solver stamp fixed
- DTEMP local temperature difference parameter added
- VERS and VREV (version revision) parameters added
- NKF high current beta rolloff parameter added
- Temperature dependence added to IKF
- Ability to select SGP qb formulation added (QBM)
- Ability to separate IS for fwd and rev added (ISRR)
- Fixed collector-substrate capacitance added (CCSO)
- Separate temperature coeffs added for RCX, RBX, RBP
- tl node eliminated
- POLARITY OF SOME BRANCHES REVERSED FOR VERILOG-A COMPATIBILITY
- Ith flows from dt to ground and so is negative
- Ixzf flows from xf1 to ground as so is negative of Itzf
- Igc component moved into Ibc
- Icc broken into forward and reverse components, Itxf or Itzf and Itzr

Release 1.2 Sept. 18, 1999

1.1.5 updates July 28, 1996

- Dependence of Irbp on Vbci added to "branch currents"
- Itf/Itr renamed Itfi/Itri to avoid name conflicts
- Resistor collapse and code bypass condition changed from par = 0 to par <= 0
- Branch current and charge dependencies separated for self-heating and no self-heating
- Depletion charge and avalanche routines that provide derivatives for self-heating added
- Self-heating solver and examples added (HBT)
- Extra external node added for self-heating to allow coupling of thermal models between devices

1.1.4 updates

- Qbe diffusion term made equivalent to SGP (divide by qb)
- Solver example including excess phase added (Icc separated into Itzf|Itxf and Itzr for this)
- Error in sgp_to_vbic in PTF to TD translation fixed

1.1.3 updates

- Ith bug fixed and Igc term added
- BFN exponent added to 1/f noise
- RTH default changed to zero
- parameter aliases added

1.1.2 updates

- EAI bug fixed in PE/PC/PStemperature mapping
- Single->double precision in decomp/solve/vbict/QCDEPL
- Scale changed to vscale in solver to avoid name conflict
- Avalanche model added, element Igc
- Initialization changed in solver
- AC solver and AC and temperature tests added
- Missing term in derf_Vrci added
- Potential numerical problem in Irci fixed

1.1.1 updates

- VJ->V bug fixed in qj definition
- Potential numerical problems with ITF fixed
- Typo derf_Vcci fixed to derf_Vrci in FORTRAN code

SDD - Definition of the VBIC 1.1.4 Model in MDS

Although the toolkit is aimed to use the VBIC implementation in ADS (optional HSpice, Eldo, Spectre etc.), the following MDS definition of the VBIC model (rev.4. version1.1.4),

implemented as "Symbolically Defined Device" (SDD), may be interesting to study.
This example is provided with the toolkit as vbicsdd.inc.

Note
The '\' sign means continuation with the next line.

```
sdd:CMP1 B 0 C 0 E 0 S 0 \
                V_bi 0 V_ci 0 V_ei 0 V_si 0 V_cx 0 V_bx 0 V_bp 0 V_th 0 V_cxf 0 V_rxf 0 \
       i[5,0]=ibi  \
       i[5,1]=qbe+qbc+qbcx \
       i[6,0]=ici  \
       i[6,1]=-qbc  \
       i[7,0]=iei-ire  \
       i[7,1]=-qbe-qbex  \
       i[8,0]=isi-irs  \
       i[8,1]=qbcp  \
       i[9,0]=icx-ircx  \
       i[9,1]=-qbcx  \
       i[10,0]=ibx-irbx  \
       i[10,1]=qbex+qbep  \
       i[11,0]=ibp  \
       i[11,1]=-qbep-qbcp  \
       i[12,0]=irth-Ptot  \
       i[12,1]=qcth  \
       f[1,0]=ib-irbx  \
       f[1,1]=-qbeo-qbco \
       f[2,0]=ic-ircx  \
       f[2,1]=qbco  \
       f[3,0]=ie-ire  \
       f[3,1]=qbeo  \
       f[4,0]=vs - irs*RS_T - vsi  \
       f[13,0]=itzf - itxf  \
       f[13,1]=-qcxf  \
       f[14,0]=vcxf - vrxf  \
       f[14,1]=-flxf
ics:CMP255 V_cx 0 idc=0
ics:CMP256 V_bp 0 idc=0
ics:CMP257 V_bx 0 idc=0
ics:CMP252 V_si 0 idc=0
ics:CMP253 V_ci 0 idc=0
ics:CMP251 V_ei 0 idc=0
ics:CMP250 V_bi 0 idc=0
ics:CMP254 V_th 0 idc=0
ics:CMP270 V_cxf 0 idc=0
ics:CMP271 V_rxf 0 idc=0
 ie=_i3
 irs=_i4
 ic=_i2
 ib=_i1
 vci = _V6
 vbi = _V5
 vs = _V4
 ve = _V3
 vc = _V2
 vb = _V1
 delt = _V12
 vrxf = _V14
 vcxf = _V13
 PE_T = psibi(PE,EAIE)
 PC_T = psibi(PC,EAIC)
 PS_T = psibi(PS,EAIS)
 GAMM_T = GAMM * (rT ^ XIS * exp(-EA * (1.0 - rT) / vtv))
 VO_T = VO * rT ^ XVO
 CJCP_T = CJCP * (PS / PS_T) ^ MS
 CJEP_T = CJEP * (PC / PC_T) ^ MC
 CJC_T = CJC * (PC / PC_T) ^ MC
 CJE_T = CJE * (PE / PE_T) ^ ME
 NR_T = NR * (1.0 + TNF * (tdev - tini))
 NF_T = NF * (1.0 + TNF * (tdev - tini))
 IBCNP_T = IBCNP * (rT ^ XIN * exp(-EANS * (1.0- rT) / vtv)) ^ (1.0 / NCNP)
 IBCIP_T = IBCIP * (rT ^ XII * exp(-EAIS * (1.0- rT) / vtv)) ^ (1.0 / NCIP)
 IBCN_T = IBCN * (rT ^ XIN * exp(-EANC * (1.0- rT) / vtv)) ^ (1.0 / NCN)
 IBENP_T = IBENP * (rT ^ XIN * exp(-EANC * (1.0- rT) / vtv)) ^ (1.0 / NCN)
 IBEIP_T = IBEIP * (rT ^ XII * exp(-EAIC * (1.0- rT) / vtv)) ^ (1.0 / NCI)
 IBCI_T = IBCI * (rT ^ XII * exp(-EAIC * (1.0- rT) / vtv)) ^ (1.0 / NCI)
 IBEN_T = IBEN * (rT ^ XIN * exp(-EANE * (1.0- rT) / vtv)) ^ (1.0 / NEN)
 IBEI_T = IBEI * (rT ^ XII * exp(-EAIE * (1.0- rT) / vtv)) ^ (1.0 / NEI)
 ISP_T = ISP * (rT ^ XIS * exp(-EA * (1.0- rT) / vtv)) ^ (1.0 / NFP)
 IS_T = IS * (rT ^ XIS * exp(-EA * (1.0- rT) / vtv)) ^ (1.0 / NF)
 RBP_T = RBP * rT^XRC
 RS_T = RS * rT^XRS
 RE_T = RE * rT^XRE
 RBI_T = RBI * rT^XRB
 RBX_T = RBX * rT^XRB
 RCI_T = RCI * rT^XRC
 RCX_T = RCX * rT^XRC
 AVC2_T = AVC2 * (1.0 + TAVC * (ftdev - tini))
 iei = -icc-ibe-ibex
 irbp = vrbp * qbp / RBP_T
 irbi = vrbi * qb / RBI_T
 irci = iohm / sqrt(1.0 + derf * derf)
 derf = IVO * RCI_T * iohm / (1.0 + 0.5 * IVO * IHRCF * (sqrt(vrci * vrci + 0.01)))
 iohm = (vrci + vtv * (kbci - kbcx - ln(rkp1))) / RCI_T
 rkp1 = (kbci + 1.0) / (kbcx + 1.0)
 kbcx = if (RCI==0.0) then 0.0 else sqrt(1.0 + GAMM_T * exp(vbcx / vtv)) endif
 kbci = if (RCI==0.0) then 0.0 else sqrt(1.0 + GAMM_T * exp(vbci / vtv)) endif
 vbcx = vbci - vrci
 qbco = CBCO*vbc
```

```
qbeo = CBEO*vbe
qbcp = CJCP_T*qdbcp
qbep = TR*itfp+CJEP_T*qdbep
qbcx = QCO*kbcx
qbc = TR*itr+CJC_T*qdbc+QCO*kbci
qbex = CJE_T*qdbex*(1.0-WBE)
qbe = tff*it_f/qb+CJE_T*qdbe*WBE
slTF = if (ITF<=0.0) then 1.0 else 0.0 endif
IHRCF = if (HRCF<=0.0) then 0.0 else 1.0/HRCF endif
IITF = if (ITF<=0.0) then 0.0 else 1.0/ITF endif
IVTF = if (VTF<=0.0) then 0.0 else 1.0/VTF endif
IVO = if (VO_T<=0.0) then 0.0 else 1.0/VO_T endif
IIKP = if (IKP<=0.0) then 0.0 else 1.0/IKP endif
IIKF = if (IKF<=0.0) then 0.0 else 1.0/IKF endif
IIKR = if (IKR<=0.0) then 0.0 else 1.0/IKR endif
IVER = if (VER<=0.0) then 0.0 else 1.0/VER endif
IVEF = if (VEF<=0.0) then 0.0 else 1.0/VEF endif
tff= TF*(1.0+QTF*q1)*(1.0+XTF*exp(vbci*IVTF/1.44)*(slTF+(ritf/(ritf+1.0))^2)*sgitf)
ritf= it_f*sgitf*IITF
sgitf= if (it_f>0.0) then 1.0 else 0.0 endif
ibc = diode(vbci,IBCI_T,NCI)+diode(vbci,IBCN_T,NCN)
ibex = (1.0-WBE)*(diode(vbex,IBEI_T,NEI)+diode(vbex,IBEN_T,NEN))
ibe = WBE*(diode(vbei,IBEI_T,NEI)+diode(vbei,IBEN_T,NEN))
iccp = (itfp-itrp)/qbp
qbp = 0.5*(1.0+sqrt(1.0+4.0*q2p))
q2p = itfp*IIKP
itrp = diode(vbcp, ISP_T, NFP)
itfp = ISP_T*(WSP*exp(vbep/(NFP*vtv))+(1.0-WSP)*exp(vbci/(NFP*vtv))-1.0)
CEP = if (TD<=0.0) then 0.0 else TD endif
LEP = if (TD<=0.0) then 0.0 else TD/3.0 endif
icc= if (TD>0.0) then itxf-itzr else itzf-itzr endif
itxf = vrxf
qcxf = CEP*vcxf
flxf = LEP*itxf
itzr = itr/qb
itzf = it_f/qb
qb = 0.5*(q1+sqrt(q1*q1+4.0*q2))
q2 = it_f*IIKF+itr*IIKR
q1 = 0.5*(sqrt((q1z-0.0001)^2+0.0001*0.0001)+q1z-0.0001)+0.0001
q1z = 1.0+qdbe*IVER+qdbc*IVEF
itr = diode(vbci, IS_T, NR_T)
it_f = diode(vbei, IS_T, NF_T)
rT = tdev / tini
tini = TNOM + tabs
qdbcp = qj(vbcp,PS_T,MS,FC,AJS)
qdbc = qj(vbci,PC_T,MC,FC,AJC)
qdbep = qj(vbep,PC_T,MC,FC,AJC)
qdbex = qj(vbex,PE_T,ME,FC,AJE)
qdbe = qj(vbei,PE_T,ME,FC,AJE)
icx = irci+irbp
vrbp = vcx-vbp
vrs = vs-vsi
vbcp = vsi-vbp
vbep = vbx-vbp
vre = ve-vei
vrbi = vbx-vbi
vrbx = vb-vbx
vbex = vbx-vei
vrci = vcx-vci
vrcx = vc-vcx
vei = _V7
vbp = _V11
vbx = _V10
vcx = _V9
vsi = _V8
tabs = 273.15
qq = 1.602189E-19
kb = 1.380662E-23
qsingle(v,p,f,a,m) = ql0(v,p,f,a,m)+(1.0-f)^(-m)*(v-vl(v,p,f,a)+vl0(p,f,a))-q0(p,f,a,m)
vl(v,p,f,a) = 0.5* (dvh(v,p,f)-mv(v,p,f,a))+p*f
mv(v,p,f,a) = sqrt(dvh(v,p,f)*dvh(v,p,f)+a)
ql0(v,p,f,a,m) = -p*(1.0-vl(v,p,f,a)/p)^(1.0-m)/(1.0-m)
q0(p,f,a,m) = -p*(1.0-vl0(p,f,a)/p)^(1.0-m)/(1.0-m)
vl0(p,f,a) = 0.5* (dv0(p,f)-mv0(p,f,a))+p*f
mv0(p,f,a) = sqrt(dv0(p,f)*dv0(p,f)+a)
dv0(p,f) = dvh(0,p,f)
qspice(v,p,m,f) = qlo(v,p,f,m)+qhi(v,p,f,m)
qhi(v,p,f,m) = if (dvh(v,p,f)>0.0) then dvh(v,p,f)*(1.0-f+0.5*m*dvh(v,p,f)/p)/((1.0-f)^(1.0+m))
\
          else 0.0 endif
qlo(v,p,f,m) = if (dvh(v,p,f)>0.0) then p*(1.0-(1.0-f)^(1.0-m))/(1.0-m)  \
          else p*(1.0-(1.0-v/p)^(1.0-m))/(1.0-m) endif
dvh(v,p,f) = v-f*p
qj(v,p,m,f,a) = if (a<=0.0) then qspice(v,p,m,f) else qsingle(v,p,f,a,m) endif
vtv = kb * tdev / qq
diode(v,is,n) = is*(exp(v/(n*vtv)) - 1)
vbci = vbi-vci
vbc = vb-vc
vbe = vb-ve
vbei = vbi-vei
isi = ibcp-iccp
ibp = -ibep-irbp-ibcp
ibx = ibex+irbi+ibep+iccp
psiio(p) = 2.0 * vtv * ln(exp(.5 * p / vtv)-exp(-0.5 * p / vtv))
psiin(p,e_a) = psiio(p) * rT - 3.0 * vtv * ln(rT) - e_a * (rT - 1.0)
psibi(p,e_a) = psiin(p,e_a) + 2.0 * vtv * ln(0.5 * (1.0 + sqrt(1.0 + 4.0 * exp(-psiin(p,e_a) /
vtv))))
tdev = if (RTH>0) then TAMB + tabs + delt else TAMB + tabs endif
Ptot = if (RTH>0) then ibe*vbei+(ibc-igc)*vbci+icc*(vbei-vbci)+ibep*vbep+ibcp*vbcp+iccp* \
```

```
        (vbep-vbcp)+ircx*vrcx+irci*vrci+irbx*vrbx+irbi*vrbi+ire*vre+irbp*vrbp+irs*vrs+ibex*vbex
\
        else 0.0 endif
 irth = if (RTH>0) then delt / RTH else 0.0 endif
 qcth = if (CTH>0) then CTH * delt else 0.0 endif
 igc = if (AVC1==0.0) then 0 else (itzf-itzr-ibc)*avalm(vbci,PC_T,MC,AVC1,AVC2_T) endif
 ibi = ibe+ibc-igc-irbi
 ici = icc-ibc+igc-irci
 vla(v,p) = 0.5 * (sqrt((p-v)^2 + 0.01) + (p-v))
 avalm(v,p,m,av1,av2) = av1 * vla(v,p) * exp(-av2 * vla(v,p)^(m - 1.0))
 ibep= if (IBEIP==0.0 && IBENP==0.0) then 0.0 else    /
           diode(vbep,IBEIP_T,NCI)+diode(vbep,IBENP_T,NCN) endif
 ibcp= if (IBCIP==0.0 && IBCNP==0.0) then 0.0 else    /
           diode(vbcp,IBCIP_T,NCIP)+diode(vbcp,IBCNP_T,NCNP) endif
 irbx = if (RBX_T <= 0.0) then 1E-15 else (vb-vbx)/RBX_T endif
 ircx = if (RCX_T <= 0.0) then 1E-15 else (vc-vcx)/RCX_T endif
 ire = if (RE_T <= 0.0) then 1E-15 else (ve-vei)/RE_T endif
R:CMP272 0 V_cxf r=1E20 OH
R:CMP273 0 V_bx r=1E20 OH
R:CMP274 0 V_bi r=1E20 OH
R:CMP275 0 V_rxf r=1E20 OH
R:CMP276 0 V_th r=1E20 OH
R:CMP277 0 V_cx r=1E20 OH
R:CMP278 0 V_bp r=1E20 OH
R:CMP279 0 V_ei r=1E20 OH
R:CMP280 0 V_ci r=1E20 OH
R:CMP281 0 V_si r=1E20 OH
```

SIMPLIFIED VBIC EQUATIONS, rev. 1.1.4:

Main transistor Collector current without the case check for parameter=0, and without temperature effects

```
ici = icc  -  ibc  +  igc  -  irci
>>      icc= itxf - itzr
                itxf = vrxf, with     vrxf = _V14
                       itzr = itr/qb
                               itr = diode(vbci, IS_T, NR_T)
                                      diode(v,is,n) = is*(exp(v/(n*vtv)) - 1)
                                  qb = 0.5*(q1+sqrt(q1*q1+4.0*q2))
>>               ibc = diode(vbci,IBCI_T,NCI)+diode(vbci,IBCN_T,NCN)
                           igc = itzf - itzr - ibc) * avalm(vbci,PC_T,MC,AVC1,AVC2_T)
                                   itzf = it_f/qb
                                           it_f = diode(vbei, IS_T, NF_T)
                                           itzr = itr/qb
                                                  itr = diode(vbci, IS_T, NR_T)
                                                  ibc =
diode(vbci,IBCI_T,NCI)+diode(vbci,IBCN_T,NCN)
>>                                     irci = iohm / sqrt(1.0 + derf ^2)
                                                iohm = (vrci + vtv * (kbci - kbcx - ln(rkp1))) /
RCI_T
                                                            vrci = vcx-vci
```

Main transistor Base current without the case check for parameter=0, and without temperature effects

Different to the SGP model, the Base-Emitter current is defined independent from the Collector-Emitter current. No beta parameter is used (i.e. no BF wit the VBIC model). The Base Emitter current consists of an ideal and an nonideal part, both devided into an intrinsic (Ibe) and an extrinsic (Ibex) part by the geometrical parameter WBE:

For version 1.1.4, there is:

and

The Base-Collector current is defined independent of the inverse transport a current as well:

Paras transistor Collector current without the case check for parameter=0, and without temperature effects

Ifp = Isp * (WSP*exp(vbep/(NFP*vt))+(1-WSP)*exp(vbci/(NFP*vt))-1)

Information on the VBIC Code, Release 1.2.

For a short synopsis of the changes see the chapter "History".
Here is a list of the major changes in version 1.2:

1. The name is now VBIC and not VBIC95.
2. The thermal network has been returned to its original form, which was how it was implemented in all simulators anyway. The "tl" node was incorrect, the Ith current had to circulate from dt to tl and so could not allow tl to funtion as a coupling node. Ith has to have one end grounded.

   Note
   This means that the value of RTH used for single device self-heating differs from that used when a thermal network couples more than one device.

3. All of the model additions agreed to at the BCTM meetings have been implemented

- temperature dependence of IKF
- separate temperature coefficients for intrinsic and extrinsic resistances
- a 3 terminal version
- base-emitter breakdown model (simple exponential)
- reach-through model to limit base-collector depletion capacitance
- VERS version parameter added (also VREV for version revision)
- separate activation energy added for ISP

1. Additional changes were made based on feedback from many sources

- errors in solvers and derivatives for electrothermal model fixed
- simple continutation added to improve solver convergence
- QBM parameter add to switch to SGP qb formulation
- NKF added to parameterize beta(Ic) high-current rolloff
- fixed collector-substrate capacitance added (CCSO)
- for HBTs, ISRR added to allow separate IS for reverse operation
- an error in the built-in potential temperature mapping was fixed
- code bypass for efficiency, if some parameters are zero
- limited exponential version provided
- the transport current Icc was explicitly separated into forward and reverse components

1. The automated code generation has been completely rewritten. All code, including solvers, is now generated. Solvers exist for all combinations of the code.
2. IMPORTANT: note that the polarities of some of the current branches have changed. This was necessary because Verilog-A supports (or appears to support) branches to ground referenced from a node to ground, and not from ground to a node.
   The Ith and Itzf branches in the thermal and excess phase networks are now defined as the negative of what they were, but the connection polarity is switched. Ith is now negative, but flows from dt to ground. This must be taken into account when setting up the matrix stamp properly.

Equivalent Circuit Network for VBIC 1.2

Default Parameters rev.1.2

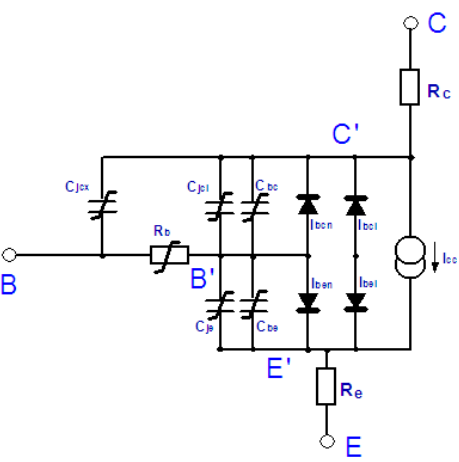
```
VERS 1.2
VREV 0.0
Parameter Extraction Temperature
TNOM 27.0
Local Temp. Dependence
DTEMP 0.0
CV
FC 0.9
CBEO 0.0
CJE 0.0
PE 0.75
ME 0.33
AJE -0.5
CBCO 0.0
CJC 0.0
QCO 0.0
CJEP 0.0
PC 0.75
MC 0.33
AJC -0.5
CJCP 0.0
PS 0.75
MS 0.33
AJS -0.5
CCSO 0.0
Resistors
RCX 0.0
RCI 0.0
VO 0.0
GAMM 0.0
HRCF 0.0
RBX 0.0
RBI 0.0
RE 0.0
RS 0.0
RBP 0.0
Early Voltage
VEF 0.0
VER 0.0
Main Transistor
IS 1.0e-16
NF 1.0
NR 1.0
IBEI 1.0e-18
WBE 1.0
NEI 1.0
IBEN 0.0
NEN 2.0
IBCI 1.0e-16
```

```
NCI 1.0
IBCN 0.0
NCN 2.0
AVC1 0.0
AVC2 0.0
IKF 0.0
NKF 0.5
IKR 0.0
ISRR 1.0
Parasitic Transistor
ISP 0.0
WSP 1.0
NFP 1.0
IBEIP 0.0
IBENP 0.0
IBCIP 0.0
NCIP 1.0
IBCNP 0.0
NCNP 2.0
IKP 0.0
Thermal Model
RTH 0.0
CTH 0.0
Transit Time
TF 0.0
QTF 0.0
XTF 0.0
VTF 0.0
ITF 0.0
TR 0.0
TD 0.0
Flicker Noise
KFN 0.0
AFN 1.0
BFN 1.0
Select SGP qB Formulation
QBM 0.0
Temperature  & Misc.
EA 1.12
EAIE 1.12
EAIC 1.12
EAIS 1.12
EANE 1.12
EANC 1.12
EANS 1.12
XIS 3.0
XII 3.0
XIN 3.0
TNF 0.0
TAVC 0.0
VRT 0.0
ART 0.1
XRE 0
XRBI 0
XRCI 0
XRS 0
XRCX 0
XRBX 0
XRBP 0
XIKF 0
XVO 0
XISR 0.0
DEAR 0.0
EAP 1.12
VBBE 0.0
NBBE 1.0
IBBE 1.0e-6
TVBBE1 0.0
TVBBE2 0.0
TNBBE 0.0
EBBE 0.0
```

Recapitulating the Gummel-Poon Model

The figure GP-1 depicts the equivalent schematic of the Gummel-Poon large signal model.

GP-1: Equivalent Schematic of the Gummel-Poon Large Signal Model



Definitions

Currents are always considered to flow into the device. This means, for example, that IC is a current into the Collector of the transistor. Voltages are indexed with their reference nodes. In order to have a clear notation, voltage drops across parasitic resistors are only considered, if this is required to better understand a specific detail or to prevent from confusion. This means, that vBE stands usually for a voltage between inner Base and inner Emitter vB'E'. Similarly, the parameter dependencies of temperature are neglected. Finally, all notation is based on a NPN transistor.

The Normalized Base Charge Qb

An essential detail for the calculation of the DC, as well as the AC performance, is the majority carrier Base charge normalized to its value without bias.

$$q_B = \frac{Q_B}{Q_{B0}} = \frac{\int_{x_{E'}}^{x_{C'}} e \cdot A_j \cdot p(x)\,dx}{\int_{x_{E0}}^{x_{C0}} e \cdot A_j \cdot N_A(x)\,dx} \qquad \text{(GP-1)}$$

q_b can be calculated as:

$$q_B = \frac{q_1}{2} + \sqrt{\left(\frac{q_1}{2}\right)^2 + q_2} \qquad \text{(GP-2)}$$

with

$$q_1 = 1 + \frac{1}{Q_{B0}} \cdot \int_0^{V_{BE}} C_{jE}(V)\,dV + \frac{1}{Q_{B0}} \cdot \int_0^{V_{CE}} C_{jC}(V)\,dV \qquad \text{(GP-3)}$$

covering the Early effect (Base width modulation) and

$$q_2 = \frac{1}{IKF} \cdot IS \cdot \left(e^{\frac{V_{be}}{NF \cdot VT}} - 1\right) + \frac{1}{IKR} \cdot IS \cdot \left(e^{\frac{V_{bc}}{NR \cdot VT}} - 1\right) \qquad \text{(GP-4)}$$

covering the Webster effect (high current behavior).
The following simplifications are applied to all common implementations of the G-P model, see e.g. the UCB SPICE version.
Equation (GP-2) is approximated by:

$$q_b \approx \frac{q_1}{2} \cdot \left(1 + \sqrt{1 + 4 \cdot q_2}\right) \qquad \text{(GP-5)}$$

and charge q_1 is approximated by:

$$q_1 \approx 1 + \frac{V_{bc}}{VAF} + \frac{V_{be}}{VAR} \approx \frac{1}{1 - \frac{V_{bc}}{VAF} - \frac{V_{be}}{VAR}} \qquad \text{(GP-6)}$$

in order to obtain, similarly to the Ebers-Moll model, a constant output conductance. Besides the approximation 1+x ~ 1/(1-x) for x<<1, equation (GP-6) contains especially the assumption of invariable and constant space charge capacitors. The second approximation in equation (GP-6) can lead to rather big modeling errors at low Early voltages. Therefore, as mentioned earlier, the simulator ADS of Agilent Technologies allows the user to select between one of these approximations.

DC Performance

The G-P Collector current IC comprises of a forward and reverse component following

$$I_C = \frac{I_F - I_R}{q_b} = \frac{IS}{q_b}\cdot\left(e^{\frac{v_{be}}{NF\cdot V_T}} - 1\right) - \frac{IS}{q_b}\cdot\left(e^{\frac{v_{bc}}{NR\cdot V_T}} - 1\right) \quad \text{(GP-7)}$$

Base width modulation and high current effects are modeled by the bias-depending Base charge q_b.

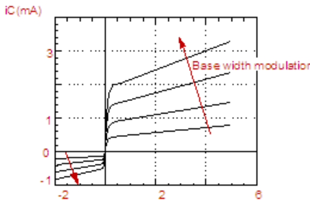
The Base current is composed of an ideal (I_{bei} , I_{bci}) and an non-ideal part (I_{ben} , I_{bcn}). The latter covers the recombination current, a low-bias effect of I_b.

$$I_{be} = I_{bei} + I_{ben} = \frac{IS}{BF}\cdot\left(e^{\frac{v_{be}}{NF\cdot V_T}} - 1\right) + ISE\cdot\left(e^{\frac{v_{be}}{NE\cdot v_T}} - 1\right) \quad \text{(GP-8)}$$

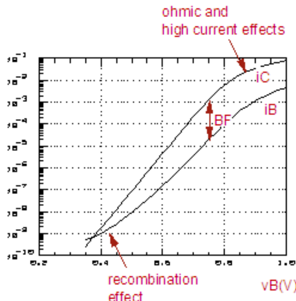
In reverse mode, the Base current is described by

$$I_{bc} = I_{bci} + I_{bcn} = \frac{IS}{BR}\cdot\left(e^{\frac{v_{bc}}{NR\cdot V_T}} - 1\right) + ISC\cdot\left(e^{\frac{v_{bc}}{NC\cdot v_T}} - 1\right) \quad \text{(GP-9)}$$

Figures GP-2 and GP-3 visualize the most important DC effects covered by the SPICE G-P model.



GP-2: Typical output characteristic of a bipolar transistor



GP-3: Gummel plot of a bipolar transistor

Resistors

The parasitic resistors Re and Rc are assumed to be constant. R_b, however, is modeled bias-dependent. In the SPICE G-P model, it is implemented as:

$$R_b = RBM + 3\cdot(RB - RBM)\cdot\left(\frac{\tan z - z}{z\cdot\tan^2 z}\right) \quad \text{(GP-10)}$$

with

$$z = \frac{-1 + \sqrt{1 + \left(\frac{12}{\pi}\right)^2\cdot\frac{I_b}{IRB}}}{\frac{12}{\pi^2}\cdot\sqrt{\frac{I_b}{IRB}}}$$

Dynamic Behavior

The capacitors between Base and Emitter, as well as between Base and Collector, represent each the space charge (depletion) and diffusion capacitor.

Space Charge Capacitors

The space charge capacitors are described by:

$$CBE_i = \frac{CJE}{\left(1 - \frac{v_i}{VJE}\right)^{MJE}} \quad \text{(GP-11)}$$

To prevent from the pole in equation (GP-11) at vi = VJE , a linear continuation of (GP-11) is used for voltages vi > FC*VJE. The Base-Collector capacitor Cjc is distributed between inner and outer Base node of the model by:

$$CBC_i = XCJC\cdot CBC \quad \text{(GP-12)}$$

and

$$CBC_x = (1 - XCJC)\cdot CBC \quad \text{(GP-13)}$$

In addition to the capacitors depicted in fig. GP-1, an additional capacitor between Collector and Substrate is added to the SPICE G-P model. This substrate capacitor is either considered to be a constant or it is also modeled after equation (GP-11).

Diffusion Charge Capacitors:

The other charge, stored with the diffusion capacitors, is calculated by:

$$Q_{be} = TFF \cdot I_F \qquad \text{(GP-14)}$$

and for B-C:

$$Q_{bc} = TR \cdot I_R, \qquad \text{(GP-15)}$$

where IF stands for the current through Collector-Emitter, and IR for Emitter-Collector, see equation (GP-7). This means for the capacitors

$$C_{be} = TFF \cdot \frac{\partial iC}{\partial vBE} = TFF \cdot \frac{IS}{q_b} \cdot \frac{1}{NF \cdot V_T} \cdot e^{\frac{v_{be}}{NF \cdot V_T}} \qquad \text{(GP-16)}$$

and

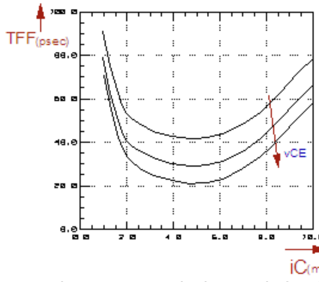
$$C_{bc} = TR \cdot \frac{\partial iE}{\partial vBC} = TR \cdot \frac{IS}{q_b} \cdot \frac{1}{NR \cdot V_T} \cdot e^{\frac{v_{bc}}{NR \cdot V_T}} \qquad \text{(GP-17)}$$

The transit time TFF itself is modeled by the empirical equation

$$TFF = TF \cdot \left[1 + XTF \cdot \left(\frac{I_F}{I_F + ITF}\right)^2 \cdot e^{\frac{v_{bc}}{1.44 VTF}}\right] \qquad \text{(GP-18)}$$

The figure GP-4 shows the trace of TFF vs. the Collector current Ic with vbc as secondary sweep

Fig. GP-4: Trace of TFF ($I_{c, vbc}$)

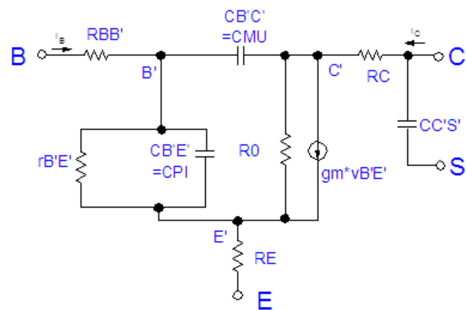


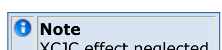
Since the measured phase shift of Ic is usually bigger than what is covered by the model equations, the UCB SPICE G-P model has an additional parameter PTF. When simulating in the frequency range, the additional phase shift is added to the phase of Ic. For the transient simulation in the time domain, however, a different implementation with Bessel function is used in UCB SPICE.

From fig. GP-1, the small signal schematic for high frequency simulations can be derived. This means, for a given operating point, the DC currents are calculated and the model is linearized in this point, see the figure below. Such a schematic is used later for SPICE S-parameter simulations.

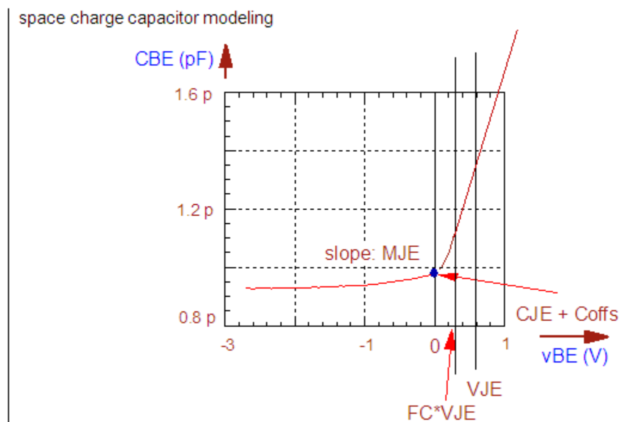
It must be noted that this schematic is a pure linear model. It cannot be used to predict non-linear high-frequency behavior of the transistor. In order to do this, RF simulators like ADS apply harmonic-balance techniques to perform nonlinear RF large signal simulations.

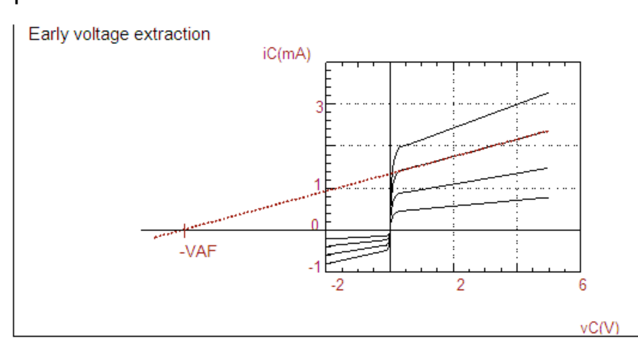
GP-5: AC small signal schematic of the bipolar transistor

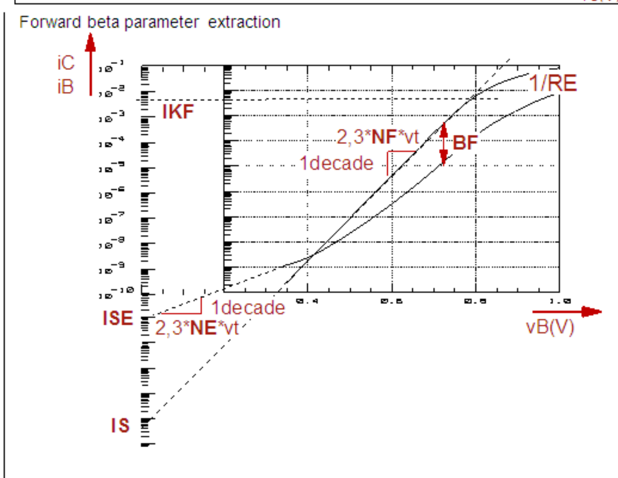


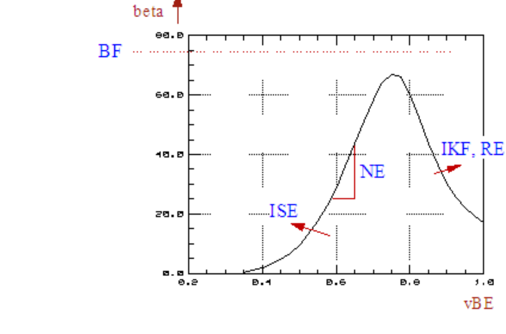

Note
XCJC effect neglected.

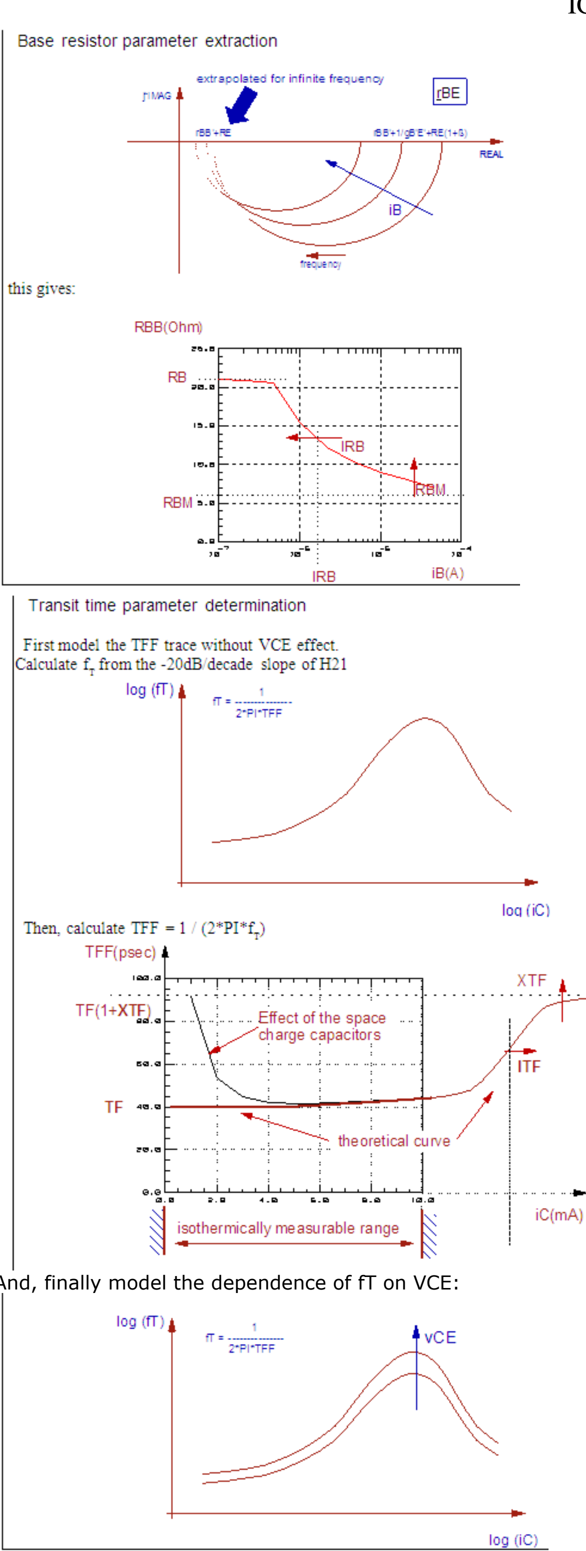
A Quick Tutorial on the Gummel-Poon Parameters


space charge capacitor modeling
CBE (pF)
1.6 p
1.2 p
0.8 p
slope: MJE
CJE + Coffs
-3
-1
0
1
vBE (V)
VJE
FC*VJE


Early voltage extraction
iC(mA)
3
1
0
-1
-VAF
-2
2
6
vC(V)


Forward beta parameter extraction
iC
iB
IKF
1/RE
2,3*NF*vt
BF
1decade
ISE
1decade
2,3*NE*vt
vB(V)
IS


beta
BF
NE
IKF, RE
ISE
vBE



And, finally model the dependence of fT on VCE:

Gummel-Poon Model Parameter List

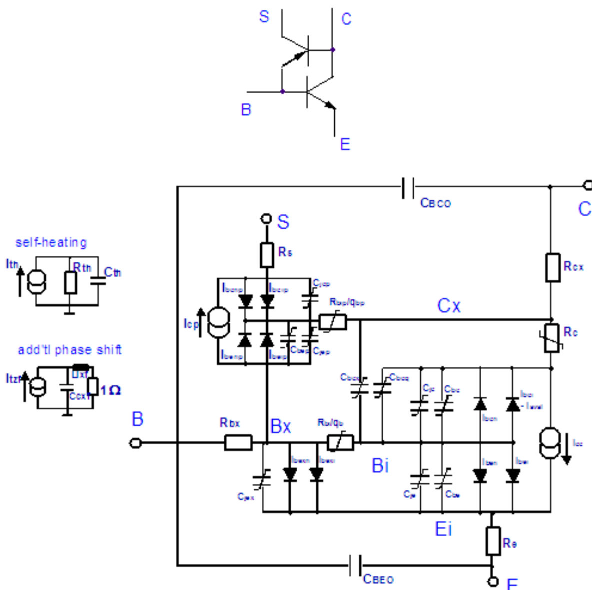
The following table compiles the G-P model parameters and their SPICE default values. These default values basically switch off the effect covered by these parameters. This enables to simulate a transistor behavior also in the case of not knowing all the model parameters.
However, these default values should not be confused with typical parameters!

Name	Parameter explanation	SPICE default	typ.value	Unit
DC:				
IS	transport saturation current	.1E-15	1.E-15	A
XTI	temperature exponent for effect on IS	3	3	
EG	energy gap for temperature effect on IS	1.11	1.11	eV
BF	ideal forward maximum beta	100	150	
BR	ideal reverse maximum beta	1	.5	
XTB	forward & reverse beta temp.coeff.	0	2.5	
VAF	forward Early voltage	infinite	100	V
VAR	reverse Early voltage	infinite	50	V
NF	forward current emission coeff.	1	1.0	
NR	reverse current emission coeff.	1	1.0	
NE	B-E leakage emission coeff.	1.5	1.7	
NC	B-C leakage emission coeff.	2	1.3	
ISE	B-E leakage saturation current	0	.1E-12	A
ISC	B-C leakage saturation current	0	1.E-13	A
IKF	forward beta hi current roll-off	infinite	.05	A
IKR	reverse beta hi current roll-off	infinite	.3	A
OHMIC PARASITICS:				
RB	zero bias base resistance	0	100	Ohm
IRB	current at medium base resistance	infinite	.0001	A
RBM	min.base resistance at hi current	RB	25	Ohm
RE	emitter resistance	0	5	Ohm
RC	collector resistance	0	10	Ohm
CBE:				
CJE	B-E zero-bias deplet.capacitance	0	1.E-12	F
VJE	B-E built-in potential	.75	.6	V
MJE	B-E junction exponential factor	.33	.4	
CBC:				
CJC	B-C zero-bias deplet.capacitance	0	.5E-12	F
VJC	B-C built-in potential	.75	.6	V
MJC	B-C junction exponential factor	.33	.4	
XCJC	fraction of B-C capacitor connected to int.base	1	1	
CCS:				
CJS	zero-bias collector-substrate capacacitance	0	0	F
VJS	substrate junction built-in potential	.75	0	V
MJS	substrate junction exponential factor	0	0	
CAPACITOR FORWARD CHARACTERISTICS:				
FC	forward bias depletion cap.coeff.	.5	.5	
TRANSIT TIME:				
TF	ideal forward transit time	0	1.E-12	sec
XTF	coeff.for bias dependence of TF	0	10	
VTF	voltage describing VBC dependence of TF	infinite	5	V
ITF	hi-current parameter for effect on TF	0	20.E-3	A
PTF	excess phase at frequency 1/(TF*2PI)	0	0	deg
TR	ideal reverse transit time	0	50.E-12	sec
NOISE:				
KF	flicker noise coeff.	0		
AF	flicker noise exponent	1		

The VBIC Model

The following figure (VBIC-1) shows the complete equivalent schematic of the VBIC model. Besides the main NPN transistor between Base, Collector and Emitter, there is an additional parasitic PNP transistor between Base, Collector and Substrate. An additional circuit is added to cover self-heating (Rth,Cth). Also, an extra circuit for adding phase shift to Icc is available. The resistances Rc, Rbi/qb and Rbip/qbp are modeled non-linearly. Therefore, they should rather be considered as nonlinear voltage dependent current sources.

VBIC-1: Equivalent schematic of the VBIC model



A note on the current flow direction of the ICP source of the parasitic PNP. As can be seen from the main NPN transistor, the transport current is always in the same direction like its driving currents, i.e. the diode currents Ibei and Iben. For the parasitic PNP, this has to be true also: flowing from its Emitter to the Collector (what is the common forward operating condition for a PNP!).

We will now look into the details of the VBIC formulation. Thermal effects will be neglected. Currents will be indexed with their corresponding node names and flowing into the nodes. Voltages, indexed by their two node names, will additionally be named by abbreviations like 'i' for internal, 'x' for external, 'o' for outer and 'p' for parasitic.
The following formulas refer to VBIC rev.1.1.4

The Normalized Base Charges q_b and q_{bp}

Since the VBIC has one of its major roots in the Gummel-Poon model, it is again the Base charge qb, which is one of the most important internal model parameters. However, with the VBIC, approximations like in the SPICE G-P model are not used at all.
As a consequence, this allows a bias dependent modeling of the output conductance g_{CE} !
In the Gummel-Poon model, with its UCB-SPICE definition of the Early voltages, g_{CE} is a constant !
Back to q_b , this charge is implemented as:

$$q_b = \frac{q_1}{2} + \sqrt{\left(\frac{q_1}{2}\right)^2 + q_2} \qquad \text{(VBIC-1)}$$

with q1 as a function of the space charge capacitors

$$q_1 = 1 + \frac{q_{je}}{VER} + \frac{q_{jc}}{VEF} \qquad \text{(VBIC-2)}$$

(different from the SPICE G-P model, see eqn. (GP-6))

and

$$q_2 = \frac{1}{IKF} \cdot IS \cdot \left(e^{\frac{V_{bei}}{NF \cdot VT}} - 1 \right) + \frac{1}{IKR} \cdot IS \cdot \left(e^{\frac{V_{bci}}{NR \cdot VT}} - 1 \right) \qquad \text{(VBIC-3)}$$

with the temperature voltage VT after (VBIC-29). Here, q_{je} and q_{jc} are the normalized charges of the space charge (depletion) capacitors Cje and Cjc. This implies that the space charge capacitors have to be modeled before the Early voltages are extracted!

For the parasitic PNP transistor, its Base charge q_{bp} is given by:

$$q_{bp} = \frac{1}{2} \cdot \sqrt{1 + 4 \cdot q_{2p}} \qquad \text{(VBIC-4)}$$

yet neglecting the Early effect!
The charge

$$q_{2p} = \frac{I_{tfp}}{IKP} \qquad \text{(VBIC-5)}$$

covers the Webster effect of the parasitic PNP transistor, but only in forward bias mode.
Coming back to the space charge capacitors, they can be modeled either after the SPICE G P model, or, depending on the switch parameter AJx, after a VBIC-specific new approach. For the case of the SGP model, and in reverse bias (Vx < FC * Px) the normalized depletion charge function is defined as:

$$qj = \frac{P_x}{(1-M_x)}\left[1-\left(1-\frac{V_x}{P_x}\right)^{(1-M_x)}\right] \qquad \text{(VBIC-6)}$$

where Px represents the built-in potential, Mx the junction grading coefficient and Vx the applied voltage of the appropriate junction.
For a forward biased junction (Vx > FC * Px) the depletion charge function is given by:

$$qj = \frac{P_x}{(1-M_x)}\left[1-(1-FC)^{(1-M_x)}\right]+(V_x-FC\cdot P_x)\cdot\left[\frac{1-FC+\frac{M_x(V_x-FC\cdot P_x)}{2P_x}}{(1-FC)^{(1+M_x)}}\right] \qquad \text{(VBIC-7)}$$

Space Charge Capacitors

Integrating the charges from above, the Base-Emitter space charge capacitance is given for vBE < FC*PE by

$$C_{SBE} = \frac{C_{JE}}{\left(1-\frac{v_{BE}}{P_E}\right)^{M_E}} \qquad \text{(VBIC-8)}$$

and else,

$$C_{SBE} = \frac{C_{JE}}{(1-F_C)^{(1+M_E)}} * \left[1-F_C*(1+M_E)+M_E*\frac{v_{BE}}{P_E}\right] \qquad \text{(VBIC-9)}$$

The same formula applies also to the other two space charges capacitors.

For these inner space charge capacitors, there is, besides this SPICE G-P formulation, also an alternate continuous formula implemented in VBIC. This alternate method requires no linear continuation in order to avoid the pole in the SPICE-like equation. The model parameters AJx with x=C, E or S, select the preferred model (-0.5, the default value, selects the G-P formula). The Base-Emitter space charge capacitance is, in analogy to the partitioning of the Base-Emitter current by the DC parameter WBE (see further below), distributed between internal and external Base.

The Base-Collector space charge capacitance is distributed between the capacitances C_{jc} (Base and Collector of the main transistor) and Cjep (Base and Emitter of the parasitic one). Another space charge capacitance, C_{jcp}, is located between Base and Collector of the parasitic PNP. It corresponds to the substrate capacitance in the SPICE G-P model.

DC Performance of the Main NPN Transistor

NPN Collector-Emitter Current

The Collector current source I_{CC} is determined, similar to the SPICE G-P model, by the two transport current terms I_{TF} and I_{TR}

$$I_{cc} = \frac{I_{TF}-I_{TR}}{q_b} = \frac{IS}{q_b}\cdot\left(e^{\frac{Vbei}{NF\cdot VT}}-1\right)-\frac{IS}{q_b}\cdot\left(e^{\frac{Vbci}{NR\cdot VT}}-1\right) \qquad \text{(VBIC-10)}$$

with the temperature voltage VT after (VBIC-29).
This means that the Collector-Emitter current in forward operation is given by:

$$I_{ce} = \frac{IS}{q_b}\cdot\left(e^{\frac{Vbei}{NF\cdot VT}}-1\right) \qquad \text{(VBIC-11)}$$

and in reverse operation by:

$$I_{ec} = \frac{IS}{q_b}\cdot\left(e^{\frac{Vbci}{NR\cdot VT}}-1\right) \qquad \text{(VBIC-12)}$$

NPN Base-Emitter Current

Like with the SPICE Gummel-Poon model, the Base-Emitter current covers the ideal and non-ideal (recombination) behavior. However, with the VBIC model, it is not coupled to the Collector current ICC like with the Gummel-Poon model, where we have the model parameters BF and BR. Here, two separate, parallel diode currents are used instead, one for the 'ideal' part (index I) and the other for the 'non-ideal' part (index N) of the Base-Emitter current:

$$I_{be} = (I_{ben} + I_{bexn}) + (I_{bei} + I_{bexi}) = IBEN \cdot \left(e^{\frac{V_{bei}}{NEN \cdot VT}} - 1 \right) + IBEI \cdot \left(e^{\frac{V_{bei}}{NEI \cdot VT}} - 1 \right) \qquad \text{(VBIC-13)}$$

Note
The VBIC parameter naming is very mnemonic.

I	ideal part of the Base current
N	non-ideal part of the Base current
X	external part of the Base current

The VBIC model features a split of the Base-Emitter current 'to the right and to the left' of the internal Base resistor RBI. This means that the very inner Base-Emitter current from node 'Bi' to 'Ei' is given by:

$$I_{be_internal} = I_{ben} + I_{bei} = WBE \cdot I_{be} \qquad \text{(VBIC-14)}$$

and the outer Base-Emitter current from the node 'Bx' to 'Ei' is given by:

$$I_{be_external} = I_{bexn} + I_{bexi} = (1 - WBE) \cdot I_{be} \qquad \text{(VBIC-15)}$$

Note
WBE, does not affect the DC fitting, but the S-parameter fitting instead.

NPN Base-Collector Current

The Base-Collector current is calculated in analogy to (VBIC-13):

$$I_{bc} = I_{bcn} + I_{bci} = IBCN \cdot \left(e^{\frac{V_{bci}}{NCN \cdot VT}} - 1 \right) + IBCI \cdot \left(e^{\frac{V_{bci}}{NCI \cdot VT}} - 1 \right) \qquad \text{(VBIC-16)}$$

With the VBIC, a weak avalanche current is described by

$$I_{bc_avalanche} = (I_{cc} - I_{bc}) \cdot AVC1 \cdot (PC - V_{bci}) \cdot \exp\left[-AVC2 \cdot (PC - V_{bci})^{MC-1} \right] \qquad \text{(VBIC-17)}$$

It overlays the current iB (see iB vs. vCE in the foutput setup), what then shows up in the forward output characteristics of iC vs. vCE.

DC Performance of the Parasitic PNP Transistor

PNP Collector-Emitter Current

The current Icp of the parasitic PNP transistor, again, is the difference between the forward and reverse transport currents Itfp and Itrp.

$$I_{cp} = \frac{I_{tfp} - I_{trp}}{q_{bp}} \qquad \text{(VBIC-18)}$$

However, the control of the forward transport current Itfp is split into the voltages vBCI of the NPN and vBEP of the parasitic PNP transistor:

$$I_{tfp} = ISP \cdot \left(WSP \cdot e^{\frac{V_{bep}}{NFP \cdot VT}} + (1 - WSP) \cdot e^{\frac{V_{bci}}{NFP \cdot VT}} - 1 \right) \qquad \text{(VBIC-19)}$$

The parameter WSP controls that splitting. Its default value is 1.
WIth WSP=0, a second knee in the log(iB) vs. vCE and log(iS) vs. vCE is modeled, with WSP=1, the log(iB) and log(iS)drop in a single step, see file VBIC_TUTOR.mdl for details.

The reverse transport current is defined by:

$$I_{trp} = ISP \cdot \left(e^{\frac{V_{bcp}}{NFP \cdot VT}} - 1 \right) \qquad \text{(VBIC-20)}$$

Note once again the VBIC mnemonic:

P	parasitic PNP transistor

PNP Base-Emitter Current

The Base-Emitter current is, like with the main NPN, split into an ideal and non-ideal part in forward and reverse operating mode.

$$I_{bep} = IBENP \cdot \left(e^{\frac{V_{bep}}{NCN \cdot VT}} - 1 \right) + IBEIP \cdot \left(e^{\frac{V_{bep}}{NCI \cdot VT}} - 1 \right) \qquad \text{(VBIC-21)}$$

Note
Since the parasitic Base-Emitter is identical to the Base-Collector of the main NPN transistor, the reverse emission coefficients of the 'MAIN NPN', i.e. NCN and NCI, are also used for the parasitic forward Base current formula.

PNP Base-Collector Current

Finally, the Base-Collector current for reverse operation of the PNP is

$$I_{bcp} = IBCIP \cdot \left(e^{\frac{v_{bcp}}{NCIP \cdot VT}} - 1 \right) + IBCNP \cdot \left(e^{\frac{v_{bcp}}{NCNP \cdot VT}} - 1 \right) \qquad \text{(VBIC-22)}$$

Resistors

The parasitic resistors RE of the Emitter and RS of the Substrate contact are modeled with a constant value. The Base resistance, however, is comprised of a constant part RBX and a variable part RBI/qb

$$Rb = RBX + RBI/qb \qquad \text{(VBIC-23)}$$

with qb after equations (VBIC-1) ..(VBIC-3)
The parasitic PNP Base resistance, RBIP is modeled by:

$$Rbi,eff = RBIP/qbp \qquad \text{(VBIC-24)}$$

with qbp after equations (VBIC-4) ..(VBIC-5)

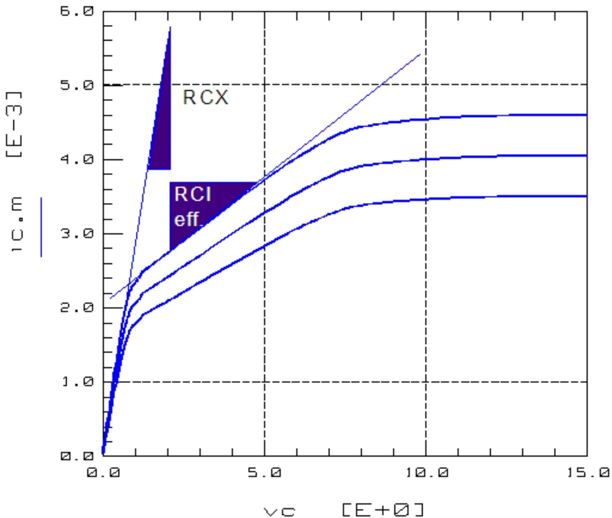
The Collector resistance, different to the G-P model, now also consists of a constant part RCX and a current-dependent, non-linear part RCI. The later models the quasi-saturation of the output characteristics, see below.

Quasi-Saturation

Quasi-saturation, as implemented in VBIC, is essentially based on the Kull model. This means that the current through the resistor RCI is depending on the inner and outer Collector voltage. With some smaller modifications to the Kull model, which essentially refer to the modeling of the velocity saturation at high voltages, the current Irci through resistor RCI is given by

$$I_{rci} = \frac{I_{ohm}}{\sqrt{1 + \left(\frac{I_{ohm} \cdot RCI}{V0 \cdot \left(1 + \frac{0.5 \sqrt{V_{rci}^2 + 0.01}}{V0 \cdot HRCF} \right)} \right)^2}} \qquad \text{(VBIC-25)}$$

In order to better understand this complex formulation, we consider its terms individually after private communication with J.Berkner, Infineon, Munich/:



First of all, the slope for low vBE is determined, like with the Gummel-Poon model, by the external Collector resistor RCX. For higher bias levels, the internal resistor RCI comes into play. This resistor RCI is bias dependent, while RCX is not. The bias dependency of RCI is related to both, the voltage drop across it, plus the current thru it.

We commence with the dependency versus voltage. To simplify things, we start with the ohmic law, applied to RCI and considering $v_{rci} = v_{Cx} - v_{Ci}$.

$$I_{ohm} = \frac{V_{rci} + V_{corr}}{RCI} \qquad \text{(VBIC-26)}$$

This means, the voltage drop across RCI is depending on a correction voltage. For non-saturated bias, Vcorr = 0, and the internal Collector resistor is constant and equal to RCI. When quasi-saturation occurs: Vcorr > 0. This means that Iohm increases and the effective internal Collector resistor is therefore reduced.
This correction voltage is given by:

$$V_{corr} = VT \cdot \left(K_{bci} - K_{bcx} - \ln \frac{1 + K_{bci}}{1 + K_{bcx}} \right)$$

with the coefficients

$$K_{bci} = \sqrt{1 + GAMM \cdot e^{\frac{v_{bci}}{VT}}} \quad \text{(VBIC-27)}$$

$$K_{bcx} = \sqrt{1 + GAMM \cdot e^{\frac{v_{bcx}}{VT}}} . \quad \text{(VBIC-28)}$$

and

$$VT = \frac{k * TNOM}{q} \quad \text{(VBIC-29)}$$

with

$$\frac{k}{q} = 8.6171E - 5$$

In addition to this reduction of the effective, internal Collector resistor, its value can further be reduced due charge carrier velocity saturation, and we obtain equation VBIC-25 from above.

Simplified for HRCF -> infinite, VBIC-25 reduces to

$$I_{rci} = \frac{I_{ohm}}{\sqrt{1 + \left(\frac{I_{ohm} \cdot RCI}{V0}\right)^2}}$$

From this simplification, we see that Irci is further reduced if Iohm * RCI ~ VO, i.e. for increasing values of Iohm.

Dynamic Behavior

The additional charge caused by quasi-saturation (Base widening), also influences the dynamic behavior of the transistor (S-parameter). This is covered by parameter QCO which models the charges stored with the capacitances Cbcx and Cbcq.

Therefore, QCO affects the fT modeling for low vCE, i.e. in quasi-saturation.

$$Q_{bcx} = QC0 \cdot K_{bcx} \quad \text{(VBIC-30)}$$

$$Q_{bcxq} = QC0 \cdot K_{bci} \quad \text{(VBIC-31)}$$

with K_{bcx} and K_{bci} from (VBIC-27) and (VBIC-28).

Besides the quasi-saturation charges of (VBIC-30) and (VBIC-31) and the constant parasitic capacitors CBC0 and CBE0, the model features for both, the main NPN and the parasitic PNP transistor, individual space charge and diffusion capacitors.

Space Charge Capacitors

Since this capacitor contributes to the Early effect modeling, it has been already discussed in the very first section of this chapter.

Diffusion Charge Capacitors of the NPN Transistor:

The diffusion capacitance of the main transistor are modeled basically in analogy to the G-P model. The charge, stored in the capacitors Cbe and Cbc, is calculated by

$$Q_{be} = TFF \cdot I_F \quad \text{(VBIC-32)}$$

and for B-C:

$$Q_{bc} = TR \cdot I_R . \quad \text{(VBIC-33)}$$

The forward transit time is calculated by

$$TFF = TF \cdot (1 + QTF \cdot q_1) \cdot \left(1 + XTF \cdot \left(\frac{I_F}{I_F + ITF}\right)^2 \cdot e^{\frac{v_{bci}}{1.44 \cdot VTF}}\right) \quad \text{(VBIC-34)}$$

very similar to the SPICE G-P model.
The term (1+QTF•q1) covers additionally a dependency of the Base width modulation.

Diffusion Charge Capacitors of the PNP Transistor

Related to the parasitic PNP transistor, the model features only one diffusion capacitance Cbep, which is described by

$$Q_{bep} = TR \cdot I_{fp} \quad \text{(VBIC-35)}$$

Additional Phase Shift

The additional phase shift is calculated with the VBIC model by a separate network. From the parameter TD, the phase is generated by a Bessel function of second order. The advantage of this method is essentially the consistent description of additional phase in the small signal and transient analysis.

Self-Heating Modeling

The most important difference to the Spice Gummel-Poon is that with VBIC, the temperature is no longer a static 'environment parameter'. Based on the additional temperature network consisting of the thermal resistor RTH and the temperature time constant 1/(RTH*CTH) , the actual, bias-dependent temperature conditions of the transistor are reflected.

Temperature Modeling

The temperature modeling is comparable to SPICE G-P. However, some enhancements have been included. See the sdd-Definition further below.

VBIC Model Parameter List

Compared to the SPICE Gummel-Poon model, we have to handle a lot more parameters with the VBIC. However, please keep in mind that the naming of the VBIC parameters is really very mnemonic:

I	Ideal part of the Base current
N	Non-ideal part of the Base current, recombination effect, below the 'knee'
P	A model parameter of the parasitic PNP

With this in mind, the VBIC can be understood very easily with its similarity to the G-P model.

VBIC Parameter	Default Value
Parasitic Capacitors	
CBE0	0
CBC0	0
Space Charge Capacitors	
CJE	0
PE	0.75
ME	0.33
AJE	-0.5
CJC	0
PC	0
MC	0.75
AJC	-0.5
CJEP	0
CJCP	0
PS	0.75
MS	0.33
AJS	-0.5
FC	0.9
Early Modeling	
VEF	1k
VER	1k
DC Forward	
IS	0.1f
NF	1
IBEI	1a
NEI	1
IBEN	0
NEN	2
IKF	0
Quasi -Saturation	
RCI	1m
GAMM	0
VO	1k
HRCF	1k
QCO	0
Time Delay Modeling	
TF	0
QTF	0
XTF	0
ITF	0
VTF	0
TR	0
Excess Phase Delay Time	
TD	0
Flicker Noise	
AFN	1
BFN	1
KFN	0

Parameter	Default Val
DC Reverse	
NR	1
IBCI	0.1f
NCI	1
IBCN	0
NCN	2
IKR	0
Distributed Base	
WBE	1
Parasitic Transistor	
ISP	0
WSP	1
NFP	1
IBEIP	0
IBENP	0
IKP	0
IBCIP	0
NCIP	1
IBCNP	0
NCNP	2
Avalanche Effect	
AVC1	0
AVC2	0
Resistances	
RE	1m
RBX	1m
RBI	1m
RS	1m
RBP	1m
RCX	1m
Temperature Modeling	
CTH	1*
RTH	0
TAMB	27
TNOM	27
EA	1.12
EAIE	1.12
EAIC	1.12
EAIS	1.12
EANE	1.12
EANC	1.12
EANS	1.12
XRE	0
XRBI, XRBX	0
XRCI, XRCX	0
XRS	0
XVO	0
XIS	3
XII	3
XIN	3
TNF	0
TAVC	0

Note
If RTH is specified, CTH=0 (the common default value) is not allowed. Otherwise, the transistor operated e.g. in pulsed mode, would react with self-heating, faster than its electrical delay.

Comparing the VBIC and Gummel-Poon Parameters

The following table gives a comparison between the VBIC and the G-P model parameters.

Table: Comparison between VBIC and Gummel-Poon Parameters

VBIC	G-P	Remarks
Parasitic Capacitors	-	-
CBE0	-	External capacitors, not included in the G-P model
CBC0	-	External capacitors, not included in the G-P model
Space Charge Capacitors	-	-
AJE	-	AJE selects one of the space charge models (-0.5 for G-P version)
CJE	CJE	-
PE	VJE	-
ME	MJE	-
AJC	-	AJC selects one of the space charge models (-0.5 for G-P version)
CJC	CJC * XCJC	-
PC	VJC	-
MC	MJC	-
CJEP	CJC*(1-XCJC)	-
CJCP	CJS	GP only includes the CV path from C->S, however not the DC C->S diode
PS	VJS	-
MS	MJS	-
AJS	-	AJS selects one of the space charge models (-0.5 for G-P version)
FC	FC	Default for VBIC: 0.9; at G-P: 0.5
WBE	-	WBE distributes the Base-Emitter current space charge capacitor between inner and outer Base
Early Modeling	-	-
VEF	VAF	The Early effect is modeled differently
VER	VAR	With the VBIC model
DC Forward Main Transistor		
IS	IS	-
NF	NF	-
Forward Base:		
IBEI	IS / BF	With the VBIC, the forward Base current is not coupled to the Collector current. Instead, two parallel diodes, one to model the ideal (I) Base current and another to cover the non-ideal or recombination (N) effect, are used.
NEI	NF	
IBEN	ISE	
NEN	NE	
IKF	IKF	
VBIC	G-P	Remarks
DC	reverse Main Transistor	Note: IS from the forward Collector current model is used.
NR	NR	-
IBCI	-	IS / BF With the VBIC model, the reverse Base current is not coupled to the Emitter current. Instead, two parallel diodes, one to model the ideal (I) Base current and another to cover the non-ideal or recombination effect (N) effect, are used.
NCI	NR	-
IBCN	ISC	-
NCN	NC	-
IKR	IKR	-
Parasitic Transistor		
ISP	-	G-P does not cover a parasitic transistor
NFP	-	-
WSP	-	Distributes parasitic collector current control to vbci of the main transistor and vbep of the parasitic transistor
Forward Base	-	-
BEIP	-	For the exponential coefficient, NCI of the main transistor is used
IBENP	-	For the exponential coefficient, NCN of the main transistor is used
Reverse Base		
IBCIP	-	With the VBIC model, the reverse Base current is not coupled to the Collector current. Instead, two parallel diodes, one to model the ideal (I) Base current and another to cover the non-ideal or recombination effect (N), are used.
NCIP	-	-
IBCNP	-	-
NCNP	-	-
IKP	-	-

Avalanche Effect	-	-
AVC1	-	-
AVC2	-	-
Resistances	-	-
RE	RE	-
RBX	RBM	The bias-dependent Base resistance is modeled differently in both models.
RBI	RB - RBM	-
-	RB	-
-	IRB	-
RS	-	-
RBP	-	-
RCX	RC	Constant, external Collector resistance
VBIC	G-P	Remarks
Quasi-Saturation		
CI	-	VBIC has a modified Kull model implemented
GAMM	-	-
VO	-	-
HRCF	-	-
QCO	-	-
Transit Time Modeling		
TF	TF	-
QTF	-	Describes the additional dependency of the transit time from qb.
XTF	XTF	-
ITF	ITF	-
VTF	VTF	-
TR	TR	-
Excess Phase	-	-
TD	π*TF*PTF/180	The VBIC implementation is consistent between small signal and transient analysis
Temperature Dependence	-	-
CTH	-	VBIC includes self-heating effects
RTH	-	-
TAMB	-	Environmental temperature
TNOM	TNOM	Measurement temperature for parameter extraction
EA	EG	The G-P model only contains one energy gap
EAIE	-	-
EAIC	-	-
EAIS	-	-
EANE	-	-
EANC	-	-
EANS	-	-
XRE	-	Temperature coefficients of the resistors are not covered with G-P
XRB	-	-
XRC	-	-
XRS	-	-
XVO	-	-
XIS	XTI	-
XII	XTB	-
XIN	XTB	-
TNF	-	-
TAVC	-	-

Since the VBIC model is based essentially upon the G-P model, most of the G-P parameters can be converted to VBIC. However, the following details have to be kept in mind strictly:

- **Space Charge Capacitances**
  For all space charge capacitances, the parameter Ajx with x = E, C and S must be set to a value less than or equal zero, in order to obtain the same formulation in both models. In this case, the values of the Base-Emitter capacitance can be transferred. With the CJC parameter of the Base-Collector capacitance, it must be reflected that and If the actual SPICE implementation of the G-P model only contains a constant value for the Substrate capacitance, CJCP will hold that value and MS is set to 0.

Finally, FC, modeling the transition between the hyperbolic formulation and the linear continuation, has different default values in both models.

- **Diode Currents**
  The forward and reverse parameters IS, NF and NR can be transferred directly. The ideal Base current sections are not coupled to the transport currents. For VBIC, there is IBEI = IS/BF and IBCI = IS/BR . The G-P parameters of the non-ideal or recombination section of the Base current can, however, be transferred directly to VBIC. Finally, setting WBE =1 , the Base current distribution (inner and outer Base) is switched off.

- **Early Modeling**
  The implementation of the Early effect in the Gummel-Poon and the VBIC model is so different, that the parameters cannot be converted (different modeling of the normalized Base charge qb). Especially with small G-P Early voltages, the resulting error can be considerably big. Yet, as a general rule, the G-P Early parameters are usually bigger than those of the VBIC model. For rather big values of the G-P Early voltage, only slight modifications should be required.
- **Parasitic Transistor and Avalanche effect**
  Using the mentioned default values in VBIC, the parasitic transistor and the Avalanche effects are switched off. Exception: Base-Collector space charge capacitance.
- **Resistances**
  The constant resistors of Emitter and Collector can be overtaken. From the parameters RBM, RB and IRB of the G-P model, suitable values have to be generated for the VBIC parameters RBI and RBX, since the models differ here.
- **Quasi-Saturation**
  Using suitable parameter values, no quasi-saturation effects are taken into account with the VBIC. Setting GAMM = 0, Rc is reduced to an ohmic resistance. QC0 = 0 eliminates the influence of the additional capacitances Cbcx and Cbcq.
- **Transit Time Parameters**
  Setting QTF = 0, the G-P parameters of the transit time TFF as well as the excess phase can be transferred without any change.
- **Temperature Modeling**
  Setting Rth = 0, the VBIC temperature model including self-heating is reduced to G-P, which only covers a constant ambient temperature.

VBIC Modeling Strategy

For a good modeling result, it is essential to follow a good parameter extraction strategy. It is especially important to follow a certain sequence of extractions, since most model parameters depend on each other. Usually, the first parameters to be extracted are those, which do not or only lightly depend on others. Then, when proceeding through the extraction strategy, the more nested parameters are extracted subsequently, and the model fits more and more accurately.

Related to the VBIC model, the following parameter extraction sequence is proposed. See also the VBIC toolkit VBIC_EXTRACT.mdl file and its macros.

CV

Since the Base charge is the basic relationship of the VBIC Early effect description, it is the **space charge capacitors** which have to be modeled first.

DC

First, the **ohmic parasitics** are extracted from specific measurement setups.
Then, the **Early voltages** are extracted from the DC output characteristics for non-quasisaturation and no avalanche effect. The parameters, however, are not yet optimized. This is due to the fact that the other DC parameters are not yet know.

We will optimize the Early parameters after the fitting of the Gummel plots. The diode parameters ISx and Nx as well as the knee currents IKx of the **main NPN transistor** are extracted from forward and reverse **Gummel-Poon** measurements. The transistor should not be in quasi-saturation.

From measurements with either an open Emitter contact, or vEC=0 for the main NPN, the diode parameters ISx and Nx and the knee current IKP of the **parasitic PNP transistor** are extracted. The **quasi-saturation** parameters, except QC0, are calculated from the output characteristics iC(vCE, vBE), as well as the **avalanche** parameters, and the selfheating parameter **RTH**

DC Finetuning

Finally, the output characteristics, especially the quasi-saturation region fitting, is fine-tuned by optimization.

S-parameters

The transit time parameters **TF, XTF, ITF, VTF** are extracted from S-parameter measurements in setup 'tf_vbe_vce' with non-quasi-saturated bias conditions. Since the dynamic model description for this bias conditions are identical to the SPICE Gummel-Poon model, the same extraction strategy is applied here too. For this condition, we set **QTF=0**

In VBIC, the Base resistance is modeled by RBB'=RBX + RBI/qb. Since qb represents a quite complex formula (see equ.(VBIC-1 ff.), the 'input-impedance-circle method' from the Gummel-Poon model cannot applied easily to get a starting value for the inner Base resistance **RBI**. Therefore, RBI is obtained by optimizing the S11 plot of setup 'biased_Spar'.

The S-parameter quasi-saturation parameter **QCO**, which affects the high frequency performance, is determined from S-parameter measurements under quasi-saturation DC bias condition. This is done in setup 'biased_Spar', referring only to the highest vBE, including the saturated vCE bias condition.

S-Parameter Finetuning

The S-parameter fitting for all DC bias conditions is fine-tuned using optimization.

Measuring and Extracting the VBIC Parameters

Measurements

For the measurements, please refer to IC-CAP file BJT_MEAS_MASTER.mdl.
Perform the measurements of all setups, and execute finally macro 'EXPORT_DATA' in order to generate mdm files for import into file VBIC_EXTRACT.mdl.

Parameter Extractions

The IC-CAP model file VBIC_EXTRACT.mdl contains the steps described in this chapter.

Space Charge Capacitances

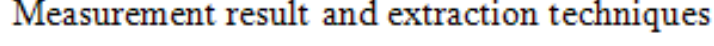
CJx, Mx, Px, AJx, FC

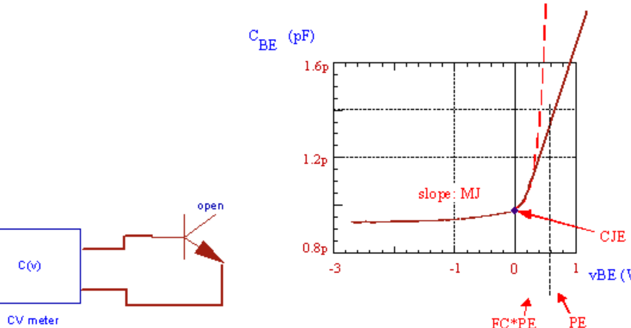
The space charge capacitances are determined from CV measurements between Base-Emitter, Base-Collector and Base-Substrate. Since we will use the G-P description, the parameters AJE, AJC and AJS have to be set to -0.5 (default!).

> **Note**
> it has been observed that the trace of C(v) above Pj does not influence the S-parameter fitting, because the diffusion capacitance (parameters TF, ITF, XTF and VTF) usually overlay the space charge capacitance at this bias condition.

We will refer to the modeling of the BE capacitance. The BC and CS capacitors are modeled correspondingly.

For the measurement of the Base-Emitter capacitance, the Collector and Substrate are left open.



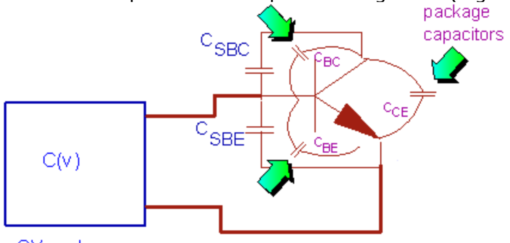


> **Note**
> On the DC bias for the CV measurements:
> A reasonable rule of thumb is to include biases up to where the capacitance is 2 to 3 times the zero bias capacitance. You want to be less than FC*VJ.
> The default FC for VBIC is 0.9, rather than the 0.5 of Spice Gummel-Poon.

> **Note**
> On the influence of the remaining capacitances of the open pin:
> as one of the transistor pins is left open, the measurements of the actual C(v) trace is overlaid by the other capacitances CSBi (i = E,C)

The total measured capacitance is therefore CSBi in parallel with the parasitic ones. This means that the measurement results are always a bit too big.

When using a capacitance meter like the HP4284, that eliminates by its measurement principle parasitic capacitances to ground, this effect can be avoided by applying an AC short to the open transistor pin versus ground (big capacitor, 1nF).



The behavior of the space charge capacitor is given by equations (VBIC-8) and (VBIC-9):
For vBE < FC*PE, we have:

$$C_{SBE} = \frac{C_{JE}}{\left(1 - \frac{v_{BE}}{P_E}\right)^{M_E}}$$ (CV-1)

and else

$$C_{SBE} = \frac{C_{JE}}{(1 - F_C)^{(1+M_E)}} * \left[1 - F_C * (1 + M_E) + M_E * \frac{v_{BE}}{P_E}\right]$$ (CV-2)

with

CJE : space charge capacitance at vBE = 0V

PE : built-in potential or pole voltage (typ. 0,7V)

ME : junction exponential factor, determines the slope of the cv plot
(abrupt pn junction (<0,5um): ME = 1/2)
(linear pn junction (> 5um): ME = 1/3)

FC : forward capacitance switching coefficient, default 0,9

Determination of the CV parameters

For simplicity, we only use the measurement data from the negative bias, i.e. we begin with the equation (CV-1):

$$C_{SBE} = \frac{C_{JE}}{\left(1 - \frac{v_{BE}}{P_E}\right)^{M_E}}$$

A logarithmic conversion yields:

ln(CSBE) = ln(CJE) - ME ln(1 - vBE / PE) (CV-3)

This equation can be interpreted as a linear function according to the ideas of linear regression analysis:
y = b + m x
with
y = ln(CSBE) (CV-4)

b = ln(CJE) (CV-5)
and
m = - ME (CV-6)

x = ln(1 - vBE / PE) (CV-7)

Linear regression means to fit a line to given measurement points. Therfore, the three main equations of a linear regression are b=f(xi,yi) and m=f(xi,yi), together with a fitting quality factor r^2=f(xi,yi,m,b). For a good fit, r^2~0.9...0.9999. See also the appendix on linear regression.

How to proceed

The measured values of CSBC are logarithmically converted according to (CV-4). Following (CV-7), the stimuli data of the forcing voltage vBE are nonlinearily converted too.

This is done using a starting value for the unknown parameter PE (e.g. 0,2V). These two arrays are now introduced into the regression equations as corresponding yi- and xi-values. A linear curve is fitted to this transformed 'cloud' of stimulating and measured data. Thus we get the y-intersect b(PE) and the slope m(PE) for the actual value of PE. In the next step, this procedure is repeated with an incremented PE, and we get another pair of m(PE) and b(PE). But now the regression coefficient r^2 will be different from the earlier one. I.e. depending on the actual value of PE, the regression line fits better or worse the transformed data 'cloud'. Once the best regression coefficient is found, the iteration loop is exited and we finally get PE_opt as well as the corresponding b(PE_opt) and m(PE_opt).

Thus we get from (CV-6):
ME = - m(PE_opt)

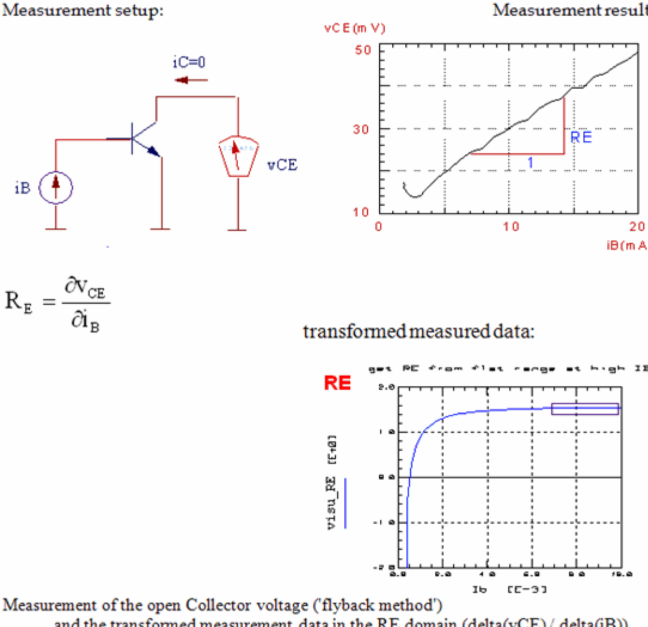
and from (CV-5):
CJE = exp [b(PE_opt)]

After that, we apply the same methodology to the other two CV curves.

> **Note**
> The Base-Collector capacitance is distributed between intrinsic and extrinsic transistor with the program 'partition_CBC'. The main part of this capacitance is usually associated with the parasitic transistor (default settings CJC=0.05*CBC_total and CJEP=0.95*CBC_total).
> From the C(v) measurements, however, the partitioning cannot be fine tuned. This is done with the S-parameters. Bsically, the partitioning affects the knee in S22, but also the magnitude of S21 for higher frequencies. See file VBIC_TUTOR.mdl.

Parasitic Resistors From DC Measurements

RE



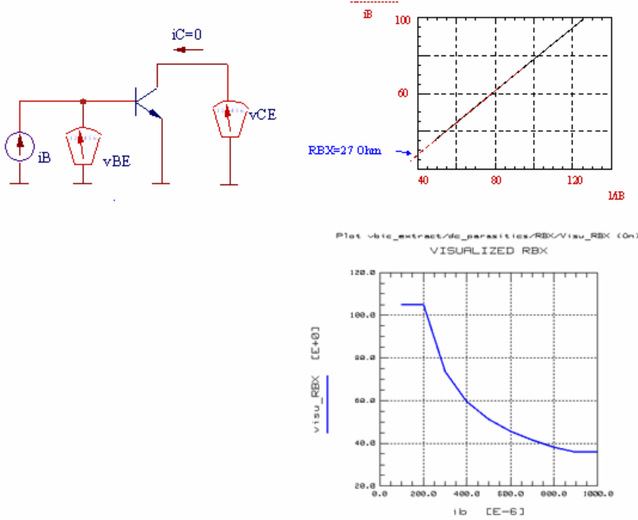
Measurement of the open Collector voltage ('flyback method')
and the transformed measurement data in the RE domain (delta(vCE) / delta(iB))

Extracting the parameters

The ohmic emitter resistor is physically located between the internal Emitter E' and the external Emitter pin E. When we apply a Base current and have the Emitter pin grounded, we get a voltage at the open Collector that is proportional to the Base current through this Emitter resistor. For this mesurement, we leave the substrate contact open. We then derivate vCE with respect to iB, we get the equivalent RE for each operating point. The result is displayed in a separate plot. The value of RE is then the mean value of the flat range in this plot.

RBX

An interesting method to determine RBX is to use the RE-flyback method, with additionally measuring vBE /T.Zimmer/. This method is applied now.



Measurement setup and determination of RBX out of transformed measured data.
The theoretical values of the measured voltages are:
vCE = VT * ln(1/AI) + iB * RE
with AI: reverse current amplification in common Base

and vBE = iB * RE + iB * RBX + vB'E'
Subtracting these equations and dividing by iB yields:

```
 vBE  -  vCE       const
-----------    =   -----  +  RBX
      iB            iB
```

i.e. after plotting the measured data accordingly, we get RBX as the y-intersect.

In a parameter visualization step, we apply a loop to these data, in which a line is fitted to two adjacent points, and the local y-intersect is calculated. The incremental y-intersects are then displayed against the stimulus iB, and represent RBX vs. iB. RBX is extracted

from the most constant range in this plot.

Note
RBX may also be obtained form a flyback measurement on the parasitic PNP with iE_MAIN = 0, i.e Emitter pin of the MAIN transistor left open.

RBI

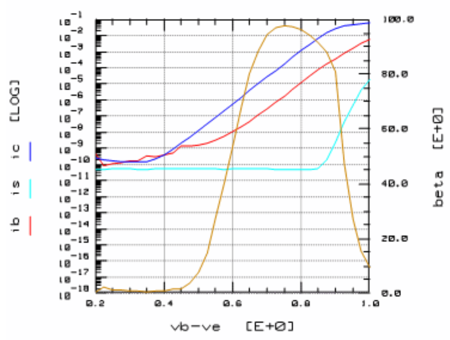
Note
RBI is modeled using S-parameters. See further down.

RCX

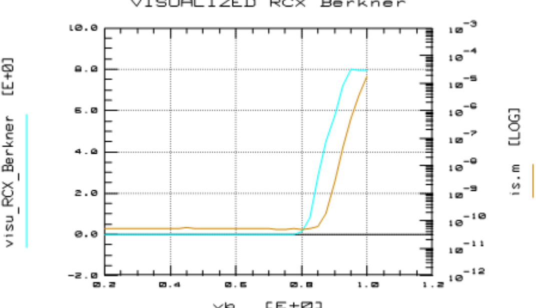
RCX is extracted from the slope of 'iS' of the Gummel-Poon plot.
The method is after the Δ Isub-method in:
J.Berkner, A Survey of DC-Methods for Determining the Series Resistances of Bipolar Transistors including the New delta-Isub Method, Proceedings of the European IC-CAP user meeting 1994, Colmar, France
Referring to two alues of iS and iC at two voltages vBE for which iS rises linearly on a log(iS)_vs_vBE plot, it is

$$RCX = \frac{VT \cdot \ln\left(\frac{I_{S1}}{I_{S2}}\right)}{I_{C1} - I_{C2}}$$

As an example, here a Gummel-Poon plot including iS:



As explained, RCX can be extracted from the slope of iS, see next plot:



RCI

A first-guess RCI is extracted from the slope of the quasi-saturated output characteristics at highest iB

Note
RCI is typically bigger than RCX (5-10 fold), and its value is typically not representing a physical ohmic resistance value.

RS

From the reverse Gummel plot of the parasitic transistor, the substrate resistor Rs can be extracted. It is visible as a decline of the slope for high currents.
RS is tuned-in and optimized in the parasitic reverse Gummel setup.

Early Voltages

VEF, VER

When extracting the VBIC Early voltage parameters, their interaction with the space charge capacitances must be considered. Referring to publication
C.C.McAndrew, L.W.Nagel, "Early Effect Modeling in SPICE",
IEEE Journal of Solid-State Circuits, vol. 31, Nr. 1, Januar 1996, equ.9
we start with

$$\frac{g_{of}}{I_C} = \frac{\frac{c_{jc}}{VEF}}{\left[1 + \frac{q_{je}}{VER} + \frac{q_{jc}}{VEF}\right]} \quad \text{and} \quad \frac{g_{or}}{I_E} = \frac{\frac{c_{je}}{VEF}}{\left[1 + \frac{q_{je}}{VER} + \frac{q_{jc}}{VEF}\right]}$$

with
gof, gor output conductance in forward and reverse mode
qje, qjc normalized charges
cje, cjc normalized capacitances
VEF, VER Early voltages

These two equations can be rearranged into:

$$a \cdot \frac{1}{VEF} + b \cdot \frac{1}{VER} = 1$$
$$c \cdot \frac{1}{VEF} + d \cdot \frac{1}{VER} = 1 \quad \text{with}$$

$$a = q_{jc} - \frac{I_C \cdot c_{jc}}{g_{of}}$$
$$b = q_{je}$$
$$c = q_{jc}$$
$$d = q_{je} - \frac{I_E \cdot c_{je}}{g_{or}}$$

They can be solved for the Early voltages and we get:

$$VEF = \frac{c - \frac{a*d}{b}}{1 - \frac{d}{b}} \qquad VER = \frac{b - \frac{a*d}{c}}{1 - \frac{a}{c}}$$

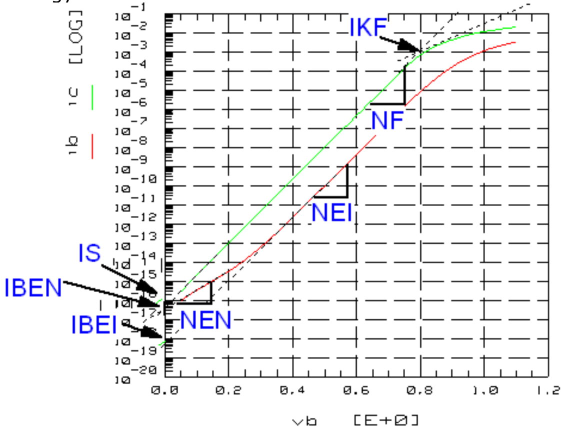
The operating point has to be selected so that no quasi-saturation or high-current effects (avalanche or thermal) disturb the measured data.

See the DUT 'VEF_VER' in file VBIC_EXTR.mdl for details.

Note
VER is often quite low. It is approximately VER ~ VEF * CJC / CJE and CJC is a lot less than CJE in "normal" BJTs because the B-C doping is a lot less than the B-E doping.

Diode Parameters

The DC parameters of the Base and Collector current of the main NPN transistor are determined from Gummel-Poon measurements. Referring to the figure below, the parameters are extracted from their dominant bias sweep ranges using regression techniques or visual extraction techniques. The following graphic depicts the basic strategy:



Determining the forward Gummel-Poon parameters of the main transistor.

The extraction of the IS, NF, IBEx and NEx parameters of the VBIC model is probably the most tedious task in the whole modeling process. Because this 'diode fitting' has to be applied to both, the iC and iB in both, forward and reverse opoeration, and additionally also once again to the parasitic PNP transistor, it can become a bit confusing. Therefore, it is most important to always remember the VBIC nomenclature:

I ideal part of the Base current
N non-ideal part of the Base current
P parasitic PNP

With this in mind, the modeling of the different diode currents becomes much more transparent.

Because of this repetitive task, we refer here only to the iC(vBE) modeling of the main NPN transistor. And this methology can then be applied to all the other parameters like:

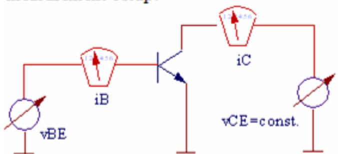
NPN forward: IS, NF, IBEI, NEI, IBEN, NEN, IKF
NPN reverse: NR, IBCI, NCI, IBCN, NCN, IKR

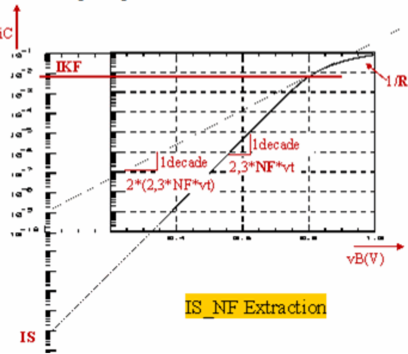
PNP forward: ISP, NFP, IBEIP, IBENP, IKP
PNP reverse: IBCIP, NCIP, IBNCNP, NCNP

IS, NF

IS transport saturation current
NF forward current emission coefficient

NF determines the slope and IS the y-intersect of the half-logarithmically plotted iC(vBE).





Provided that vB'E'=vBE and vB'C'=vBC , we start with the ice formula (VBIC-11):

$$i_{ce} = \frac{IS}{q_b} \cdot \left(e^{\frac{v_{be}}{NF \cdot VT}} - 1 \right)$$

We simplify this equation by setting the normalized Base charge qb=1. In other words, we neglect the Early effect for this extraction. This assumption, which is in most cases no big simplification, will be corrected later by applying a quick optimization to the extracted parameters.

Also, we select a low vCE for the IS and NF, IBEx and NEx parameter extractions. Otherwise, it might hgappen that due to the reverse voltage vCB, a leakage current from the Collector into the Base overlays the positive Base current and, therefore, does not allow to accurately extract the parameters IBEN and NEN.

Extracting the parameters

We begin with

$$i_{ce} = IS \cdot \left(e^{\frac{v_{be}}{NF \cdot VT}} - 1 \right) \quad \text{with} \quad VT = \frac{k*TNOM}{q}$$

For v $_{be}$ > 0.2, a very typical condition to obtain noise-free measurement data, this simplifies further to

$$i_{ce} = IS \cdot e^{\frac{v_{be}}{NF \cdot VT}}$$

We first apply a non-linear transformation to this equation, i.e. the measured data, in order to obtain a linear context between the measured values of iC and the stimulating values of vBE:

A log10 conversion gives:

$$\log_{10}(i_{ce}) = \log_{10}(IS) + \frac{v_{be}}{NF \cdot VT} \log_{10}(e)$$

or

$$\log_{10}(i_{ce}) = \log_{10}(IS) + \frac{1}{2{,}3026 \cdot NF \cdot VT} \cdot v_{be}$$

This can be considered as a linear form:

with substituting:

$$y = \log_{10}(i_{ce})$$

$$b = \log_{10}(IS)$$

$$m = \frac{1}{2{,}3026 \cdot NF \cdot VT}$$

$$x = v_{be}$$

How to proceed

We select a sub-range of the measured data, where the half-logarithmicly plotted data represent a straight line. Then the logarithmically converted icei of this sub-range are interpreted as y- and the linear vBEi values as x-data for the regression formula. Applying these formulas, (see the appendix), we obtain y-intersect 'b' and the slope 'm' of the straight fitted line.

A final re-substitution gives the parameters I_S and N_F out of 'b' and 'm':

$IS = 10^b$
and

$$NF = \frac{1}{2{,}3026 \cdot m \cdot VT}$$

Validity of the extraction

vBE between 0,2V (no noise) and 0,7V (no high current effects), low vCE (to keep the Early effect low).

> **Note**
> many modeling engineers do not extract NF, but keep it rather NF=1. The reason is that for NF ‡ NR, the power balance of the transistor is violated (it generates power instead of behaving like a controlled resistor). They instead extract TNOM from the fitting of the slope.

In this case, the above formula for NF changes to

$$VT = \frac{1}{2{,}3026 \cdot m \cdot 1}$$

what is solved for TNOM

$$TNOM = \frac{q}{k} \cdot VT = \frac{q}{k} \cdot \frac{1}{2.3026 \cdot m} = \frac{5{,}04E3}{m} - 273{,}15$$

A hint on visualized parameter extractions:

Transforming the measured data such that the model parameter can be displayed directly against the stimulating voltage or current is another smart way to determine model parameters. In the case of NF this would mean to start with

$$i_{CE} = IS \exp\left(\frac{v_{BE}}{NF \cdot v_T}\right)$$

to convert it logarithmically in order to obtain

$$\ln(i_{CE}) = \ln(IS) + \frac{1}{NF \cdot v_T} * v_{BE}$$

This is the mathematical representation of the half-logarithmic Gummel plot for i_C.

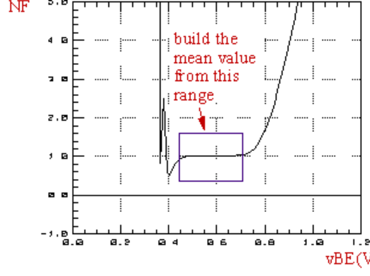
The parameter NF is proportional to the slope and we have therefore to differentiate $\ln(i_C)$ with respect to v_{BE} and obtain:

$$\frac{\partial \ln(i_{CE})}{\partial v_{BE}} = \frac{1}{NF v_T}$$

Solved for NF gives

$$NF = \frac{1}{V_T * \frac{\partial(\ln(i_{CE}))}{\partial(v_{BE})}}$$

Therefore, if we display the calculated NF (what is the 'effective NF' for every measured data point) versus v_{BE}, we get



Note
Applied to modeling iB(vBE), this method allows to determine the exact sub-range of data from which to extract NEN and NEI as the mean value of that flat range. The same principle cal also be applied to extract IBEN and IBEI.

IKF

Referring to the log(iCE)_vs_vBE curve from the top of this chapter, for which we had set vBC ~ 0, we extract IKF from vBC > 0, typically a vBC which corresponds to vCE = vCE_{max} of the foutput setup. This is because for vBC ~ 0, quasi-saturation and Webster effect (IKF) are overlaying each other.

Therefore, IKF is extracted from a separate Gummel-Poon measurement, with a sufficiently high vCE. Before we do so, we first perform a quick fine-optimization of the so far determined Gummel-Poon parameters IS, NF and IBEx, NEx for the actual DC bias conditions.

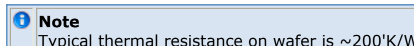
Then, for the determination of the parameter value, we follow the simple idea that IKF represents that iC value where beta=ic/ib has fallen to half its maximum value. Although, there is always an overlay of IKF and RE, this method gives usually quite reliable results. Again, like with the IS, NF, IBEx and NEx parameters, a quick optimizer finetuning is applied to match both, the RE and the IKF value above the knee in the log(iC)_vs_vBE plot.

Applying the same methodology, we extract the main transistor reverse parameters as well as the forward and reverse Gummel parameters of the parasitic transistor. See the setups in DUT 'dc_gummel' for details.

Output Characteristics

As a quick remark on the measurement setup, the question is to either stimulate the output characteristic with a Base current or a Base voltage. It was found that the quasi-static behavior of the output characteristics shows up much better when forcing a Base voltage than a current.

However, with a forcing vB, thermal effects show up much more. While this is a problem for Gummel-Poon, the VBIC model features the parameter RTH.



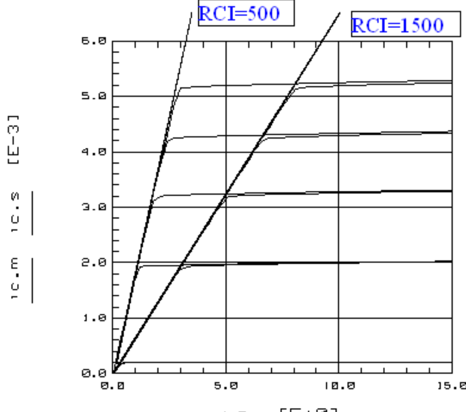
Note
Typical thermal resistance on wafer is ~200'K/W

VEF, VER

These parameters were already extracted at the beginning, and therefore, need now only be fine-optimized. With all the Gummel-parameters (IS, NF, IBEx, NEx) extracted, the output characteristics should fit now well for medium vCE and no avalanche or thermal runaway

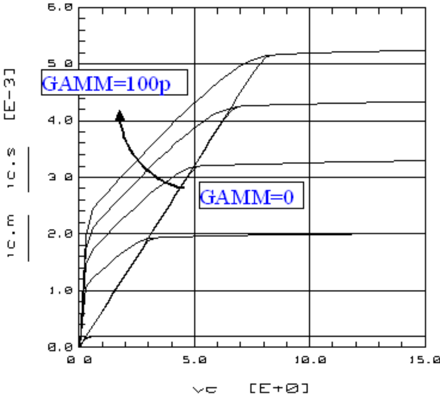
RCI, V0 and GAMM (the Quasi Saturation Parameters)

In setup foutput, the quasi-saturation parameters RCI, V0 and GAMM are determined.

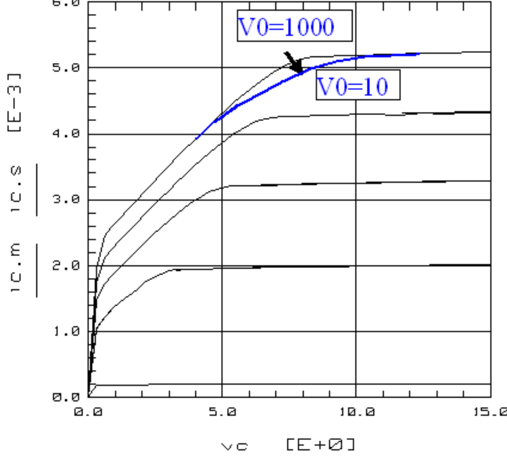


The figure above visualizes, how the parameter RCI affects the output characterization fitting. It basically determines the slope of the saturated range. Therefore, its value can be determined from the transition to quasi-saturation. It should be mentioned that RCI does not necessarily relate to a physical value.

With the parameter GAMM, the effect of quasi-saturation is deferred to higher currents, see the figure below. However note that GAMM also affects through the capacitors Cbcq and Cbcx the dynamic behavior of the transistor.



VO determines the begin of velocity saturation. This means a smoothing at the high-end of the quasi-saturation. For big values of VCO, its influence on the curve vanishes.



Finally, HRCF was added empirically to the quasi-saturation model in VBIC, to reflect an increase of iC with higher vCE. The influence of HRCF is increasing with smaller parameter values. Over a wide range, therefore, HRCF conflicts somehow with VO. We therefore set it to its default value of 1000.

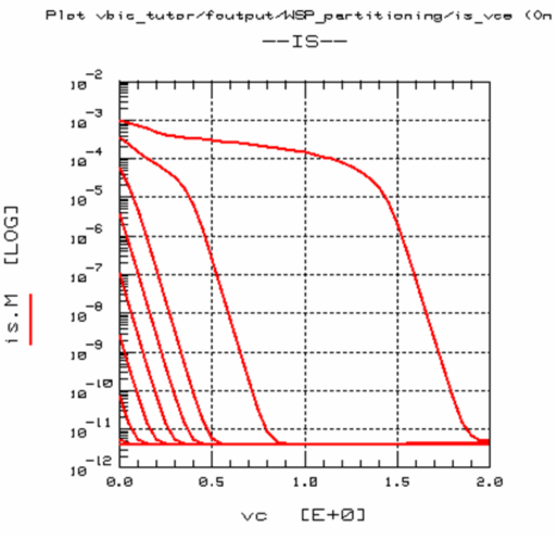
The high-frequency quasisaturation parameter

QC0

not affecting the DC performance, will be determined later from S-parameter measurements.

WSP (DC Current Distribution Between Main NPN and Parasitic NPN)

Parameter WSP, which partitions the control of Itfp between intrinsic main and parasitic transistor, is used to model the 2nd knee in plots log(ib)_vs_vce and log(is)_vs_vce. It can be used for fine-tuning. Its default value '1' represents no such 2nd knee, while the other limit '0' includes it. See setup 'foutput_aval_therm_is_ib' for details.

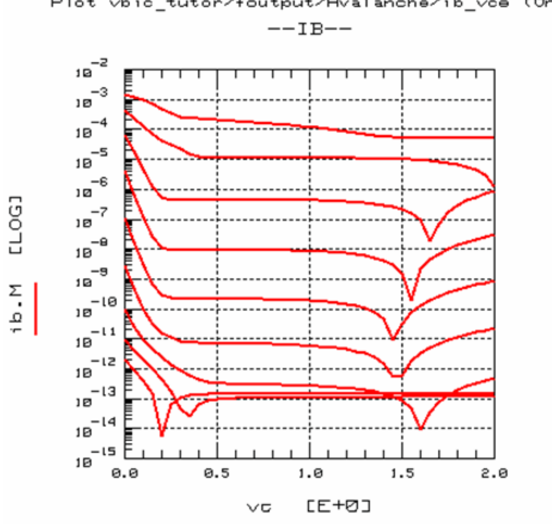


AVC1 and AVC2 (Avalanche Effect Modeling)

The weak avalanche current is modeled as a special current from Collector to the Base following:

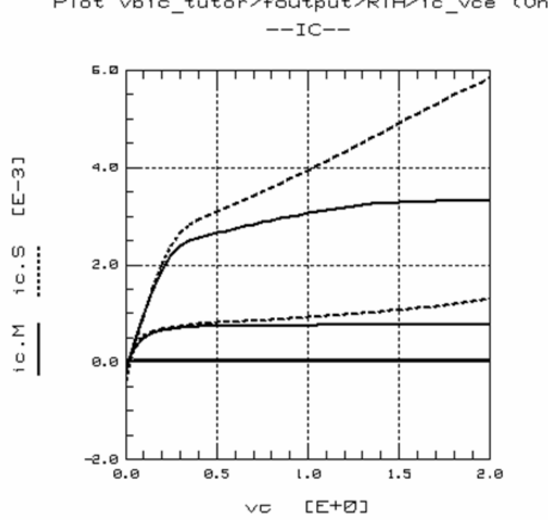
$$i_{bc_avalanche} = (I_{cc} - I_{bc}) \cdot AVC1 \cdot (PC - V_{bci}) \cdot \exp\left[-AVC2 \cdot (PC - V_{bci})^{MC-1} \right]$$

Here, AVC1 and AVC2 are the avalanche model parameters, PC and MC are the Base-Collector space-charge capacitor model parameters.
See the transforms in setup 'foutput_aval_therm_is_ib' for details.



RTH (Output Characteristics Thermal Effect Modeling)

See the transforms in setup 'foutput_aval_therm_is_ib' for details.



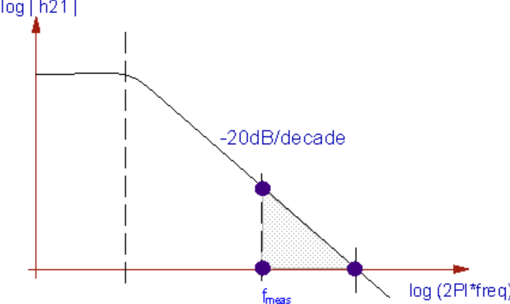
Note
In Spectre, RTH is only active if parameter selft=1

S-Parameter Modeling

The transit time T_{FF} is calculated from transit (cutoff) frequency measurements following the known formula

$$T_{FF} = \frac{1}{2 \cdot \pi \cdot f_T}$$

fT is the cutoff frequency where H21 = 1 (0dB). This is usually extrapolated from H21 measurements by fitting a -20dB/decade slope, see below.



This means, we get

$$f_{T\ 1\text{-pole}}(i_C, v_{CE}) = \frac{1}{2 * PI * T_{FF}(i_C, v_{CE})}$$

or solved for the parameter of interest:

$$T_{FF}(i_C, v_{CE}) = \frac{1}{2 * PI * f_{T\,1\text{-pole}}(i_C, v_{CE})}$$

where fT1-pole is a function of the bias current iC and the bias voltage vCE .
The modeling equation for TFF (VBIC-34) is

$$TFF = TF \cdot (1 + QTF \cdot q_1) \cdot \left(1 + XTF \cdot \left(\frac{I_F}{I_F + ITF} \right)^2 \cdot e^{\frac{Vbci}{1.44 \cdot VTF}} \right)$$

with

$$I_F = \frac{IS}{qb} \cdot \left(e^{\frac{Vbei}{NF \cdot VT}} - 1 \right) \qquad \text{after (VBIC-11)}$$

and

$$q_1 = 1 + \frac{q_{je}}{VER} + \frac{q_{jc}}{VEF} \qquad \text{after (VBIC-2)}$$

This means that, except for QTF, we can apply the known Gummel-Poon extraction methods also to the VBIC model.

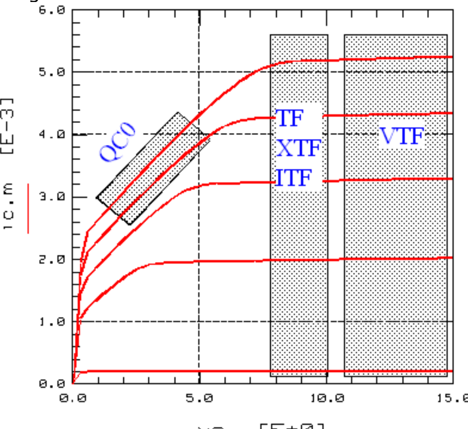
For small values of the forward transport current IF, the above equation simplifies to

$$TFF \approx TF \cdot (1 + QTF \cdot q_1)$$

This would allow to model TF and QTF. However, the effects are difficult to separate. Therefore, we start with QTF=0. If required, we obtain its final value from S-parameter fine-tuning.

Since the VBIC model includes the quasi-saturation effect, we need to watch out for the right DC bias conditions to extract the HF parameters of TFF. Therefore, we first extract TF, XTF, ITF and VTF from values of vCE where there is no quasi-saturation and no avalanche or thermal effect. Then, we select an S-parameter measurement in DC quasisaturation, and model QCO by optimization.

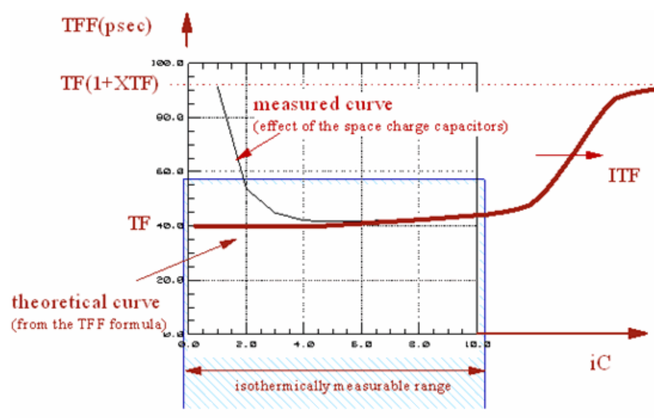
The graphic below gives detailed information about the best parameter extraction bias ranges.



DC bias condition for extraction of the transit time parameters

Parameter extraction in detail

We proceed like with the Gummel-Poon model, and consider first the obtained trace of TFF for as low as possible vCE, but not in quasi-saturation !



The theoretical transit time TFF (bold) as a function of iC for low vCB compared to the theoretical trace.

The above figure shows the theoretical curve in addition to a typically measured one. For low frequencies, the real measured curve is overlaid by the space charge capacitor effects for low collector currents.
On the other hand, the DC bias conditions for which the TFF parameter ITF and XTF show up, cannot be measured without self heating effects, because the required bias current iC is usually well above ~50mA. However, because the VBIC includes RTH, and if this RTH modeling was performed carefully in the foutput setup, this should not influence the TF, ITF, XTF extraction.

Due to these overlay and measurement problems, it had been found that a pretty simple and straight-forward extraction technique can be applied that gives nevertheless quite reasonable results. This method is explained below. There exist some more complex strategies, but the extraction results may be not much better.

How to proceed

TF

is extracted as the minimum value of TFF.

XTF

The behavior of TFF was given in the above figure. In many cases, measurement data for a higher Collector current are not available due to compliance. So XTF is estimated from the trace of TF at max. available Collector bias current under the assumption that it would be TFF at infinite current:

$$MAX(TFF) = TF\,(1 + XTF)$$

or

$$XTF = \frac{MAX(TFF)}{TF} - 1$$

This usually gives a pretty good first-order estimation. Due to the Collector current limitations, an estimation correction like XTF = 5..10 * XTF_extracted can improve the starting conditions for the optimizer.

ITF

Referring to the same measurement restrictions as above, a good first-order estimation of ITF is to use the max. Collector current measured:
ITF = MAX(iC_meas)

Again, since the end of the TFF trace is often not measurable, correct this estimation by ITF = ~5*ITF_extracted.

> **Note**
> This XTF and ITF extraction method follows the idea of the Gummel-Poon Base resistor parameter extraction!

> **Note**
> In the TFF equation VBIC-34, when TFF = TF ? (1 _ XTF / 2), i.e. TFF is in the middle between its minimum and maximum value, the corresponding IF bias current is iC_meas = 2,41 ? ITF.

VTF

Finally, we consider also the vCE sweep, but, again, not in quasi-saturation:

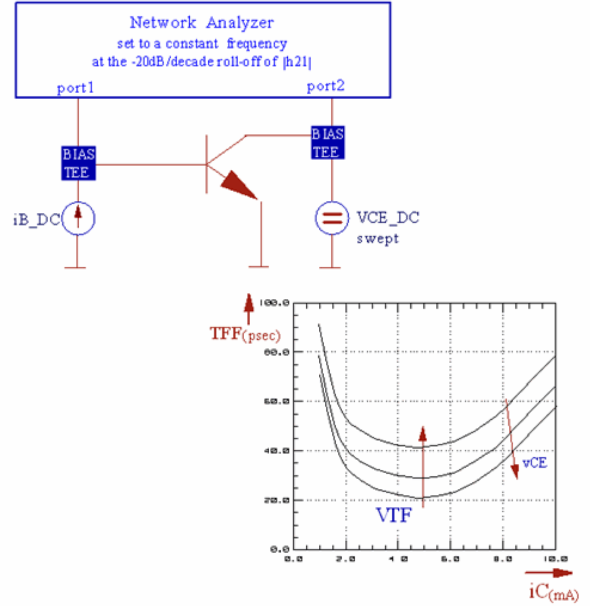
$$MAX(TFF) = TF\,(1 + XTF)$$

or

$$XTF = \frac{MAX(TFF)}{TF} - 1$$

VTF can be obtained from (VBIC-34) for a fixed value of iC:

Measurement setup:



The transit time TFF as a function of iC and vCE

This gives:

$$TFF = \cdot const \cdot e^{\frac{-v_{CB}}{1.44 \cdot VTF}}$$

or

$$\frac{TFF1}{TFF2} = \frac{\exp[-v_{CB1} / 1{,}44\ VTF]}{\exp[-v_{CB2} / 1{,}44\ VTF]} = \exp\left[\frac{v_{CB2} - v_{CB1}}{1{,}44\ VTF}\right]$$

$$\ln\left[\frac{TFF1}{TFF2}\right] = \frac{v_{CB2} - v_{CB1}}{1{,}44\ VTF}$$

and finally:

$$VTF = \frac{v_{CB2} - v_{CB1}}{1{,}44 * \ln\left(\frac{TFF_1}{TFF_2}\right)}$$

QCO

this parameter is very dominant in the gm plot, obtained from S-to-Y parameter conversions. We select a bias condition in quasi-saturation and apply an optimization.

TD

The delay time parameter affects mostly the ttrace of S21 in the 1st quadrant. It is well visible in the phase of gm. Its value is optimized for S21 in setup 'biased_Spar'.

RBI

In VBIC, the Base resistance is modeled by RBB'=RBX + RBI/qb. Since qb represents a quite complex formula (see equations (VBIC-1 ... VBIC-3), the 'input-impedance-circle method' from the Gummel-Poon model cannot applied easily to get a starting value for the inner Base resistance RBI. Therefore, RBI is obtained by optimizing the S11 plot of setup 'biased_Spar'.

Final Parameter Tuning

Since the high-frequency model parameters can affect back the DC model results, all modeling setups should be re-simulated and fine-tuned by optimization. Using the measurements of setups verify_DC and verify_S, the error between measured and simulated curves can be analyzed easily for all relevant bias points.

Parametrized Modeling vs. Transistor Geometry

The following info is a copy of the VBIC committee extraction suggestions:

CV

From devices of different base-emitter area to perimeter ratios CJE can be modeled as function of emitter-base area and perimeter, by extracting area, perimeter, and if necessary constant (corner) components from the different area/perimeter structures. CJE is then calculated as the sum of area, perimeter, and constant components, based on specific device geometry.

If no layout information is known, the CJC/CJEP splitting should be done so that ft is modeled well (note that typically most of the capacitance should be in CJEP).
If the relative areas of the b-c junction under the emitter and not under the emitter (e.g. intrinsic and extrinsic b-c junction areas) are known, partition the extracted CJC between CJC and CJEP accordingly.

Preferred approach

If devices of different base-collector intrinsic/extrinsic areas and perimeters are available, area and perimeter components of CJC+CJEP can be easily determined. CJC is then calculated from the area of the base-emitter, and CJEP from the base-collector perimeter and the excess of the base-collector area over the base-emitter area, plus a constant (corner) component if so modeled.

DC

Ibc components can be split into intrinsic (IBCI/IBCN) and extrinsic (IBEIP/IBENP) in manner analogous to the split for CJC/CJEP. IS, IBEI, IBEN, etc. can all be related to geometry (area and perimeter) by determining them for two or more area/perimeter ratios and then calculating the area and perimeter components.

Thermal Modeling

The VBIC model includes many parameters which allow the modeling of the transistor behavior at different operating temperatures.

Here some hints on this kind of modeling from the VBIC 1.1.4 extraction recommendations of the VBIC committee:

The temperature dependence of the junction built-in potentials and capacitances is determined by the activation energies, which are determined from the temperature dependence of saturation currents.

From low-bias FG Ic data over temperature, determine EA by optimization. XIS can also be included in the optimization, but from my experience the optimization is relatively insensitive to XIS, and EA is by far the major controlling factor. Therefore, XIS should be left at default (XIS=3).

From low-bias FG Ib data over temperature, determine EAIE and EANE by optimization. Again, XII and XIN do not affect this optimization much, and so should be set to 3 rather than being included in the optimization. Different temperature dependences for IS, IBEI and IBEN are necessary to model the variation of beta with temperature properly, the beta roll-off at low Vbe, caused by the non-ideal component of Ibe, has a different temperature variation than the variation of the peak/flat beta with temperature.

There are no "free" parameters for the temperature variation of Ie in reverse mode operation, they are fixed by EA and XIS, determined from the temperature variation of Ic in forward mode operation. EAIC and EANC are determined by optimizing the fit to low bias RG Ib data. Ibc has separate temperature parameters from Ibe, as it has a slightly different variation with temperature.

From low-bias measurements of the parasitic b-c current (over temperature) determine IBCIP/NCIP/IBCNP/NCNP (EAIS/EANS) as done for intrinsic device Ibe/Ibc parameters. These parameters are really only used to flag improper biasing of a device, so reasonable estimated could be used instead of making measurements and extracting parameter values.

APPENDIX

Plots of transistor characteristics like MAG and FMAX can be used to check the final S-parameter fitting. Since they include all the S-parameter, i.e. S11 ... S22, they can be used to show the general S-parameter fitting at a glance.

MAG

The maximum amplification gain is obtained when the transistor or amplifier is inserted between matching load/source resistance. I.e. for conjugate complex impedance matching at both sides of the transistor, the amplifier gain is different, usually higher than for a Z0 = 50Ω environment..

The maximum available gain is

$$MAG = \left|\frac{\underline{S}_{21}}{\underline{S}_{12}}\right| * \left(K - \sqrt{K^2 - 1}\right)$$

with K

$$K = \frac{1 + |\underline{D}|^2 - |\underline{S}_{11}|^2 - |\underline{S}_{22}|^2}{2 * \left|\underline{S}_{12} * \underline{S}_{21}\right|}$$

where D is the determinant, i.e. $\underline{D} = \underline{S}_{11} * \underline{S}_{22} - \underline{S}_{12} * \underline{S}_{21}$

MAG is only valid if K>1

MAG calculates the gain of the entire network (DUT plus matching networks).

Note
it is a typical characteristic of RF transistors that they are instable for lower frequencies (K<1). Calculating MAG for those frequencies makes therefore no sense. In other words, only for frequencies with K>1, an impedance matching for max. gain MAG can be designed.

It is further helpful to draw $|S21|^2$ and MAG together in the same plot versus log(freq). Usually, MAG is bigger than $|S21|^2$, what means that the power amplification can be improved with impedance matching, compared to the conventional modeling measurement condition Z0=50Ω. See the following figure. Note that for a wide frequency range, the transistor is not unconditionally stable (K<1), and therefore MAG is not defined.

Comparing power amplification |S21|2 to the maximum available gain

FMAX

The transit frequency ft is defined as the frequency at which |h21| hits '1'. This frequency ft determines the max. switching frequency and is therefore most interesting for digital applications (output AC-wise shorted).

Yet, this is not the max. frequency for the transistor in analog operation and without the assumption of an AC-wise short at the output of the twoport.

This max. frequency is defined by FMAX , which refers to the maximum available gain MAG. As long as MAG> 1, the transistor can still generate oscillations. The max. possible oscillation frequency is therefore

Regression Analysis

Regression Analysis is a technique which provides a best curve fit for given data sets.

Mathematical Basics of Curve Fitting

Let's assume we made 'N' measurements yi at the stimulating points xi.
I.e. we obtained the array [xi, yi], which is then shown in a plot.

A curve Y(x) shall be fitted to this array of measured data points using least square curve fitting technique.
Referring to an individual measurement point, the fitting error is:

and for all data points:

$$E = \sum_{i=1}^{N} E_i^2 = \sum_{i=1}^{N} \left[Y(x_i) - y_i \right]^2 \qquad (2)$$

This error shall be minimized.

The fitting will be done by varying the coefficients of the fitting curve of equation (2). The minimum of the total error E depends on the values of these coefficients. This means, we have to differentiate E partially versus the curve coefficients and to set the results to zero. We obtain a system of equations, solve it, and get the values of the coefficients for a best curve fit. This is known as regression analysis.

> **Note**
> This regression analysis is simple for a straight line fit. But in general, measured data is non-linear. Unfortunately, a non-linear regression analysis can be quite complicated. This problem can be solved if we use a suitable transformation on the measured data. This means that the measured data is transformed to a linear context between the yi- and the xi-values. As will be seen in the diode example later, this is a pretty smart way to get the curve fitting parameters easily without much calculations.

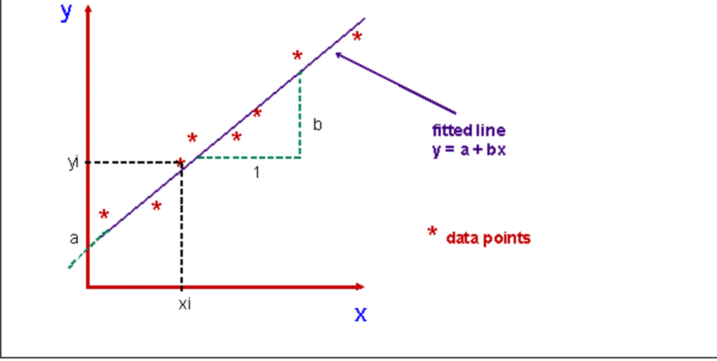
Linear Curve Fitting

IC-CAP File:
demo_features/5_PEL_PROGRAMMING/3_PARAM_EXTRACTIONS_CURVE_FITTING/1_basic_PEL_extractions/1fit_line.mdl

Provided we have got an array of N measured data points of the form [xi,yi].
A linear curve with the equation

$$y(x) = a + b*x \qquad (3)$$

shall be fitted to these points. This situation is depicted below.



Linear regression applied to measurement points

The error of the i-th measurement is:

$$Ei = [\ a + b * xi\ ] - yi \qquad (4a)$$

Using the least means square method following equ.(2) yields:

$$E = \sum_{i=1}^{N} Ei^2 = \sum_{i=1}^{N} [\ a + b * xi - yi\ ]^2 \overset{!}{=} \text{Minimum} \qquad (4b)$$

Partial differentiation versus slope 'm' gives:

$$2 \sum_{i=1}^{N} [a + b * xi - yi\ ]\ xi = 0 \qquad (5)$$

and versus y-intersect 'b':

$$2 \sum_{i=1}^{N} [a + b * xi - yi\ ] = 0 \qquad (6)$$

We obtain from (5) after a re-arrangement:

$$b \sum_{i=1}^{N} xi^2 + a \sum_{i=1}^{N} xi = \sum_{i=1}^{N} yi\ xi \qquad (7)$$

and from (6):

$$b \sum_{i=1}^{N} xi + N\ a = \sum_{i=1}^{N} yi \qquad (8)$$

Multiplying (7) by -N and (8) by xi and adding these two equations allows the elimination of the coefficient 'a', and we can separate the slope 'b':

$$b \left[\left(\sum_{i=1}^{N} xi \right)^2 - N \sum_{i=1}^{N} xi^2 \right] = \sum_{i=1}^{N} xi \sum_{i=1}^{N} yi - N \sum_{i=1}^{N} xi\ yi \qquad (9)$$

or:

$$b = \frac{\sum_{i=1}^{N} x_i \cdot \sum_{i=1}^{N} y_i - N \sum_{i=1}^{N} x_i \cdot y_i}{\left(\sum_{i=1}^{N} x_i \right)^2 - N \sum_{i=1}^{N} x_i^2} \qquad (10)$$

and from (8) for the y-intersect 'a':

$$a = \frac{1}{N} \left(\sum_{i=1}^{N} y_i - m \sum_{i=1}^{N} x_i \right) \qquad (11)$$

with 'b' according to (10).

With equations (10) and (11), we determine the values of the two coefficients of the line which fits best into the 'cloud' of measured data.

Finally, a curve fitting quality factor r2 [regression coefficient] is defined.
Its value ranges from [0 < r2 < 1].
The closer it is to 1, the better is the fit of the linear curve.

$$r^2 = b^2 \frac{\sum_{i=1}^{N} x_i^2 - \frac{1}{N} \left(\sum_{i=1}^{N} x_i \right)^2}{\sum_{i=1}^{N} y_i^2 - \frac{1}{N} \left(\sum_{i=1}^{N} y_i \right)^2} \qquad (12)$$

with 'b' from (10)

General Formula for Regression Analysis

Linear Regression:

$$y = f(x) = a + b \cdot x$$

$$\begin{pmatrix} N & \sum_{i=1}^{N} x_i \\ \sum_{i=1}^{N} x_i & \sum_{i=1}^{N} x_i^2 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{N} y_i \\ \sum_{i=1}^{N} x_i y_i \end{pmatrix} \qquad (13)$$

Quadratic Regression:

$$y = f(x) = a + b \cdot x + c \cdot x^2$$

$$\begin{pmatrix} N & \sum_{i=1}^{N} x_i & \sum_{i=1}^{N} x_i^2 \\ \sum_{i=1}^{N} x_i & \sum_{i=1}^{N} x_i^2 & \sum_{i=1}^{N} x_i^3 \\ \sum_{i=1}^{N} x_i^2 & \sum_{i=1}^{N} x_i^3 & \sum_{i=1}^{N} x_i^4 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{N} y_i \\ \sum_{i=1}^{N} x_i y_i \\ \sum_{i=1}^{N} x_i^2 y_i \end{pmatrix} \qquad (14)$$

Cubic Regression:
continue with the evolution of matrices like above.

Application: Calculating the quadratic polynomial parameters from measurement data

IC-CAP File:
demo_features/5_PEL_PROGRAMMING/3_PARAM_EXTRACTIONS_CURVE_FITTING/1_basic_PEL_extractions/1fit_quadratic.mdl

Starting with equ.(14) from above

$$\begin{pmatrix} N & \sum_{i=1}^{N} x_i & \sum_{i=1}^{N} x_i^2 \\ \sum_{i=1}^{N} x_i & \sum_{i=1}^{N} x_i^2 & \sum_{i=1}^{N} x_i^3 \\ \sum_{i=1}^{N} x_i^2 & \sum_{i=1}^{N} x_i^3 & \sum_{i=1}^{N} x_i^4 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{N} y_i \\ \sum_{i=1}^{N} x_i y_i \\ \sum_{i=1}^{N} x_i^2 y_i \end{pmatrix}$$

we get the following system of 3 equations:

$$\sum_{i=1}^{N} y_i = N \cdot a + \sum_{i=1}^{N} x_i * b + \sum_{i=1}^{N} x_i^2 * c \quad \text{or} \quad \frac{\sum_{i=1}^{N} y_i}{N} = a + \frac{\sum_{i=1}^{N} x_i}{N} * b + \frac{\sum_{i=1}^{N} x_i^2}{N} * c \qquad (15a)$$

$$\sum_{i=1}^{N} x_i y_i = \sum_{i=1}^{N} x_i * a + \sum_{i=1}^{N} x_i^2 * b + \sum_{i=1}^{N} x_i^3 * c \quad \text{or} \quad \frac{\sum_{i=1}^{N} x_i y_i}{\sum_{i=1}^{N} x_i} = a + \frac{\sum_{i=1}^{N} x_i^2}{\sum_{i=1}^{N} x_i} * b + \frac{\sum_{i=1}^{N} x_i^3}{\sum_{i=1}^{N} x_i} * c \qquad (15b)$$

$$\sum_{i=1}^{N} x_i^2 y_i = \sum_{i=1}^{N} x_i^2 * a + \sum_{i=1}^{N} x_i^3 * b + \sum_{i=1}^{N} x_i^4 * c \quad \text{or} \quad \frac{\sum_{i=1}^{N} x_i^2 y_i}{\sum_{i=1}^{N} x_i^2} = a + \frac{\sum_{i=1}^{N} x_i 3}{\sum_{i=1}^{N} x_i^2} * b + \frac{\sum_{i=1}^{N} x_i^4}{\sum_{i=1}^{N} x_i^2} * c \qquad (15c)$$

Equations (15a) .. (15c) are of the form

$$y0 = a + x0 * b + x0^2 * c \qquad (16a)$$

$$y1 = a + x1 * b + x1^2 * c \qquad (16b)$$

$$y2 = a + x2 * b + x2^2 * c \qquad (16c)$$

for which we have already a solution for a, b and c in chapter 6.Data_Interpolation.

Therefore, with the following abbreviations:

$$y0 = \frac{\sum_{i=1}^{N} y_i}{N} \quad x0 = \frac{\sum_{i=1}^{N} x_i}{N} \quad x02 = \frac{\sum_{i=1}^{N} x_i^2}{N}$$

$$y1 = \frac{\sum_{i=1}^{N} x_i y_i}{\sum_{i=1}^{N} x_i} \quad x1 = \frac{\sum_{i=1}^{N} x_i^2}{\sum_{i=1}^{N} x_i} \quad x12 = \frac{\sum_{i=1}^{N} x_i^3}{\sum_{i=1}^{N} x_i}$$

$$y2 = \frac{\sum_{i=1}^{N} x_i^2 y_i}{\sum_{i=1}^{N} x_i} \quad x2 = \frac{\sum_{i=1}^{N} x_i^3}{\sum_{i=1}^{N} x_i} \quad x22 = \frac{\sum_{i=1}^{N} x_i^4}{\sum_{i=1}^{N} x_i}$$

we can re-use the quadratic data interpolation formulae of the chapter Data Interpolation:

$$c = \frac{(y1 - y0) * (x1 - x2) - (x1 - x0) * (y1 - y2)}{(x12 - x02) * (x1 - x2) - (x1 - x0) * (x12 - x22)}$$

$$b = \frac{y1 - y2 + c * (x22 - x12)}{(x1 - x2)}$$

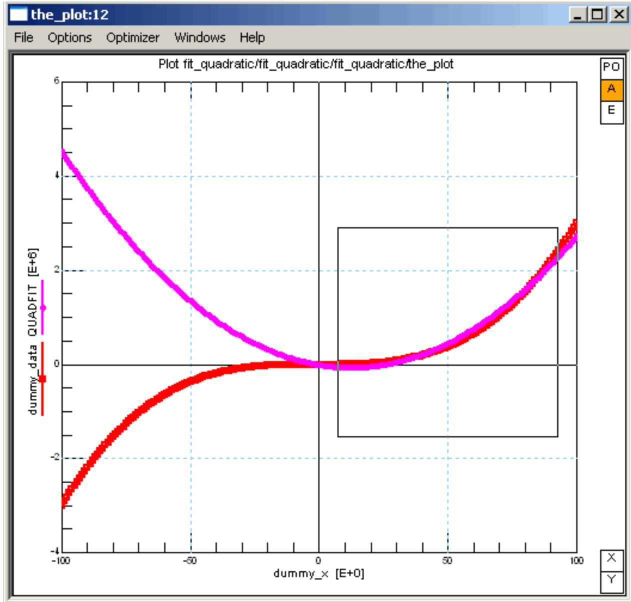
and

$$a = y2 - b * x2 - c * x22$$

and are ready to get the quadratic fitting function y(x)

$$y = a + b * x + c * x^2$$

In the plot below, the red curve (dummy_data) has been fitted by the magenta curve (QUADFIT), inside the marked box.


the_plot:12
File Options Optimizer Windows Help
Plot fit_quadratic/fit_quadratic/fit_quadratic/the_plot
PO
A
E
dummy_data QUADFIT [E+0]
dummy_x [E+0]
X
Y

Acknowledgements, Web info, and Publications

Acknowledgements

The author would like to thank especially F. X. Sinnesbichler of the Technical University Munich, Germany, Institute of High Frequency Techniques. His IC-CAP workshop about VBIC in Munich, Jan.1998, in which he emphasized on introducing the VBIC model with its similarities to the Gummel-Poon toolkit, inspired much the beginning of this VBIC toolkit.

Very special acknowledgments also to J.Berkner of Infineon Technologies, Munich, Germany and E.Gerhardt of Alcatel SEL, Stuttgart, Germany, for many helpful discussions.

Also, many thanks to Fujiang Lin, formerly working at the Agilent Singapore Modeling Center, for informations about the new version 1.2 model and special parameter extractions methods.

Web Info

Info on the web since July, 2000:http://www.fht-esslingen.de/institute/iafgp/neu/VBIC/index.html

Publications

- J.J. Ebers, J.L. Moll, "Large-Signal Behaviour of Junction Transistors", Proceedings of the IRE, pp. 1761-1772, Dezember 1994.
- I. E. Getreu, "Modelling the Bipolar Transistor", Elsevier, Amsterdam, 1978, also available as Tektronix book, Tektronix part no. 062-2841-00.
- B. K. Gummel, H. C. Poon, "An Integral Charge Control Model of Bipolar Transistors", Bell Systems Technical Journal, vol. 49, pp827-853, 1970.
- H. C. de Graaff, F.M. Klaassen, "Compact Transistor Modelling for Circuit Design", Springer Verlag, Wien, 1990.
- C. McAndrew et al, "VBIC95: An Improved Vertical, IC Bipolar Transistor Model", Proceedings of the 1995 BiCMOS Circuits and Technology Meeting, pp.170-177, 1995, Minneapolis.
- C. McAndrew et al, "VBIC95, The Verticl Bipolar Inter-Company Model", IEEE Journal of Solid-State Circuits, vol. 31, Nr. 10, Oktober 1996.
- C.C.McAndrew, L.W.Nagel, "Early Effect Modeling in SPICE", IEEE Journal of Solid-State Circuits, vol. 31, Nr. 1, Januar 1996.
- F. Najim (ed.), "VBIC95: An Improved Bipolar Transistor Model", IEEE Circuits and Devices Magazine, vol. 12, pp. 11-15, March 1996.
- J. Parker, M. Dunn, "VBIC95 Bipolar Transistor Model andAssociated Parameter Extraction", Proceedings of the 1995 HP EEsof US ICCAP Users' Meeting, Washington, Dezember 1995.
- P. Antognetti, G. Massobrio, "Semiconductor Device Modeling with SPICE", McGraw-Hill, New York, 1988.
- G.M. Kull et al, "A Unified Circuit Model for Bipolar Transistors Including Qasi-Saturation Effects", IEEE Transactions on Electron Devices, vol. 32, Nr. 6, pp. 1103-1113, Juni 1985.
- P.B. Weil, L.P. McNamee, "Simulation of Excess Phase in Bipolar Transistors", IEEE Transactions on Circuits and Systems, vol. 25, Nr. 2, pp. 114-116, Februar 1978.
- F.van der Wiele, W.L.Engl, P.G.Jespers, "Process and Device Modeling for Integrated Circuit Design", NATO Advanced Study Institute, Louvain-la-Neuve, Belgium, 1977
- J. Berkner, "Parameter Extraktion for BJT Quasisaturation Models", Proceedings of the 4th European IC-CAP User Meeting, Berlin, Oktober 1997.
- F. X. Sinnesbichler, G. R. Olbrich, "Early Effect Modelling of Bipolar Junction Transistors with Low Early Voltages", Proceedings of the 4th European IC-CAP User Meeting, Berlin, Oktober 1997.
- G. Neaves et al, "Characterization of Low VAR Bipolar Transistors Using a Revised SPICE Simulator", IEEE Proceedings of the 1992 Bipolar BICMOS Circuits and Technology Meeting, pp 229-232, 1992, Minneapolis.
- T.Zimmer: Contribution à la modélisation des transistors haute fréquence, Thèse à l'Université de Bordeaux I, 17.7.1992.
- Zimmer, Meresse, Cazenave, Dom, 'Simple Determination of BJT Extrinsic Base Resistance', Electron.Letters, 10.10.91, vol.27, no.21, p.1895
- F. Sischka, "Gummel-Poon Bipolar Model Parameter Extraction Toolkit", Agilent Technologies GmbH, München, Germany

RF Passive Linear Components

Contents

- Test Fixture
  - *Time Domain Modeling* (iccapmhb)
  - *Frequency Domain Modeling* (iccapmhb)
- RLC
  - *RLC Modeling* (iccapmhb)
  - *Resistor* (iccapmhb)
  - *Capacitor* (iccapmhb)
  - *Inductor* (iccapmhb)
- *Strip Lines* (iccapmhb)
- *Modeling Vias* (iccapmhb)
- *Dummy Chip* (iccapmhb)
- Packages
  - *Package Modeling* (iccapmhb)
  - *SOT23 Package* (iccapmhb)
  - *TSOP44 Package* (iccapmhb)
- *Modeling Connector* (iccapmhb)
- *Frequency Dependent Losses* (iccapmhb)
- *Test for Passive Circuits* (iccapmhb)

Modeling of a Chip Capacitor up to 26 GHz

First, the NWA is calibrated with the method of /Gronau 1992/ , see chapter on network analyzer measurements and calibration. This method allows a shift of the calibration plane up to the pins of the capacitor chip. The following figure shows the measurement setup, the theoretical curve for an ideal capacitor from DC to infinite frequency as well as the s21 measurement curve of the SMT capacitor up to 26GHz. The chip capacitor was pressed to two 50 Ohm strip lines.

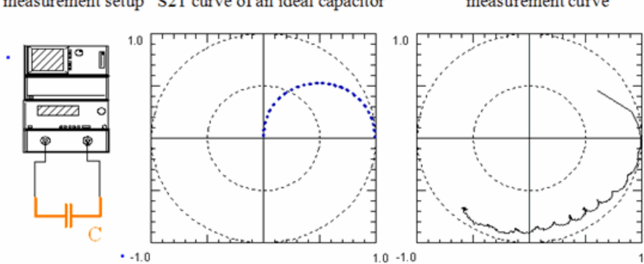


Fig.1: Measurement setup, theoretical S21 curve and measured curve of a 47pF SMU capacitor from 45MHz to 26GHz.

Some pre-considerations:
For the frequency '0', the capacitor blocks the power flow. In the above figure, the theoretical curve starts at (0 + j*0). At infinite frequency, all power is passing the capacitor, and that's why Sxy should end at (1 + j*0). Between these points, a semi-circle should be expected. Yet, the real measurement gives a different picture: the lower measurement frequency of 45MHz causes a certain phase shift of the first measurement point. At higher frequencies, parasitic inductors show up, since the curve exceeds the point (1 + j*0).

Therefore, the modeling can be performed in the following sequence:
The equivalent schematic starts simply with a capacitor (CMAIN). Its value is the specified manufacturer value (see the following figure). A simulation should show a good agreement with the measured curve for low frequencies. Next, we are concentrating on the fit at higher frequencies. We determine the value of RMAIN (since the S21 curve does not hit exactly (1 + j0), then the value of LMAIN (for the raw trace beyond (1 + j*0). We also evaluate the parasitic connectivity parameters L1,C1 and L2,C2.

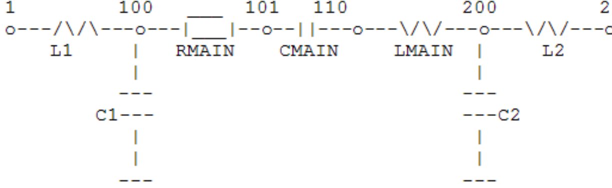


Fig.2: The equivalent schematic of the 47pF capacitor up to 26GHz.

The fitting quality of the model is depicted in the following figure for the magnitude and phase of S21.

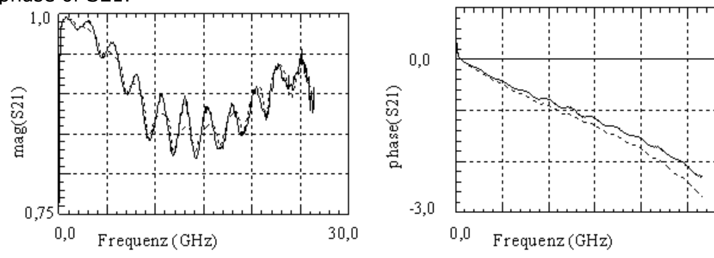


Fig. 3: Magnitude and phase of measured and simulated (dotted) S21

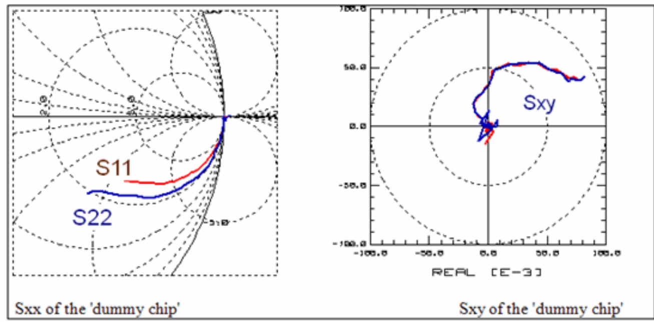
Modeling of a Dummy Chip on a Wafer

Using on-wafer calibration techniques allows to correct for all measurement errors down to the chip pads. Nevertheless is it important for frequencies above some 100MHz to accurately model the parasitic effects like bond pad capacitances on the chip itself, because they affect and distort the correct measurement of the 'inner device' on the chip. Once they are known, they can be eliminated from the measurement of the whole chip using de-embedding techniques, see chapter on de-embedding in section Vector Network Analyzers.

The following shows the layout of a so-called 'open dummy chip'. To the left are the 3 pads for the ground-signal-ground connection of the network analyzer's port 1, and to the right those of port 2. The transistor itself is left 'empty', so that we measure nothing but the parasitics of these pads.

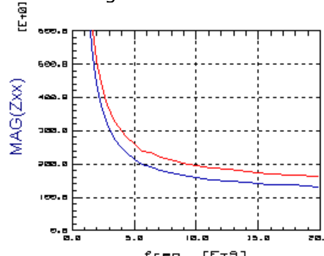
The following figure shows layout of an 'open dummy chip' device on a wafer (the active part of the bipolar transistor is 'left empty').

Following figures show the S-parameter measurement result of a structure like in above figure. The frequency range is from 50 MHz to 20 GHz. The curves can be considered as typical, and as will be seen shortly can be interpreted quite conveniently.



Sxx of the 'dummy chip' Sxy of the 'dummy chip'

To get a first idea about the equivalent schematic, the S-parameters are first transformed to Z-parameters. Fig.4 gives the magnitude of Z11 and Z22. There is no increase in value visible for high frequencies. This leads to the first modeling result: we will probably be able to neglect inductive elements for the equivalent schematic to be developed.



Next, we will test the applicability of a PI structure. So the S-parameters are converted to Y-parameters.

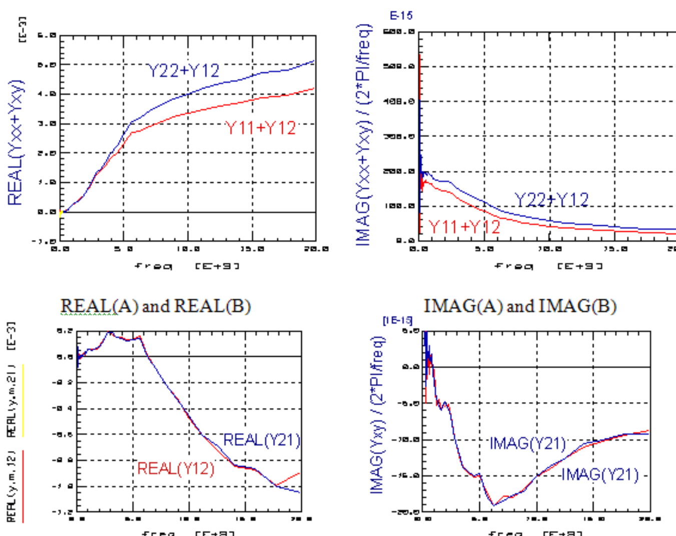
$$Y_{PI} = \begin{pmatrix} A+C & -C \\ -C & B+C \end{pmatrix}$$

PI-structure as Y-matrix

Assuming the PI equivalent schematic (fig.5) and expecting capacitors and dielectric losses as its elements (C in parallel to G), we plot

```
REAL(A) = REAL(Y11+Y12),
IMAG(A) = IMAG(Y11+Y12) / (2*PI*freq)
REAL(B) = REAL(Y22+Y12),
IMAG(B) = IMAG(Y22+Y12) / (2*PI*freq)
REAL(C) = REAL(-Y12)
IMAG(C) = IMAG(-Y12) / (2*PI*freq)
```

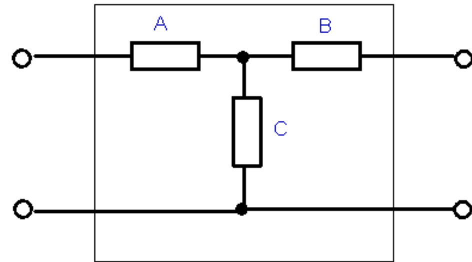
versus frequency. The results are given in following figure.



Unfortunately we cannot get the expected curves like REAL(A) and REAL(B) to be a frequency-independent resistor and IMAG(A) and IMAG(B) to be a capacitor with a constant value over the frequency. Therefore we have to reject the PI equivalent structure.

We will now test the applicability of a T structure, consisting of capacitors in series with resistors. Such a capacitor is typical for lossy silicon capacitor parasitics. The simplest representation of such a structure is a Z matrix, see figure below.

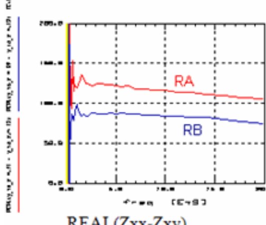
$$Z = \begin{pmatrix} A+C & C \\ C & B+C \end{pmatrix}$$



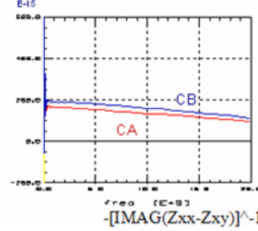
So we convert the measured S-parameters to Z, and plot

REAL(A) = REAL(Z11-Z12)
REAL(B) = REAL(Z22-Z12).

This corresponds to the ohmic resistors RA resp. RB of components A resp. B in fig.10. As given in fig.11, there is no frequency dependency. Therefore, the assumption of a TEE structure for these resistors is valid and we get RA = 110 Ohm as a mean value for frequencies above 3GHz. Using the same strategy, RB is calculated out of the plot REAL(Z22-Z12) as 80 Ohm.

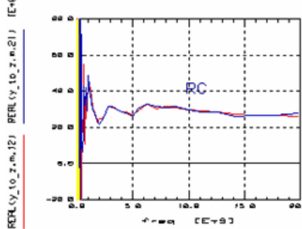


REAL(Zxx-Zxy)

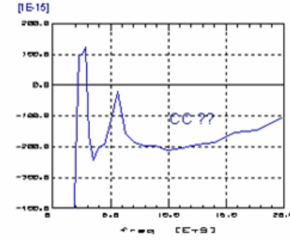


-[IMAG(Zxx-Zxy)]^-1

Then, we plot the inverse of the imaginary parts divided by 2*PI*freq as depicted in above figure. Also this type of transformed data looks quite constant, what can be interpreted with respect to fig.10 as capacitors CA and CB in series with the already identified resistor RA and RB. Their values are CA = 140fF and CBC = 160fF from above figure.



REAL(Zxy)

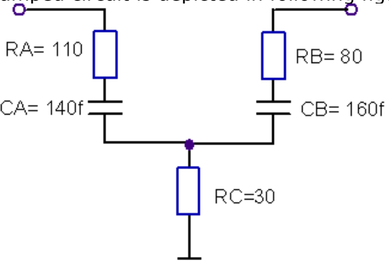


IMAG(Zxy)

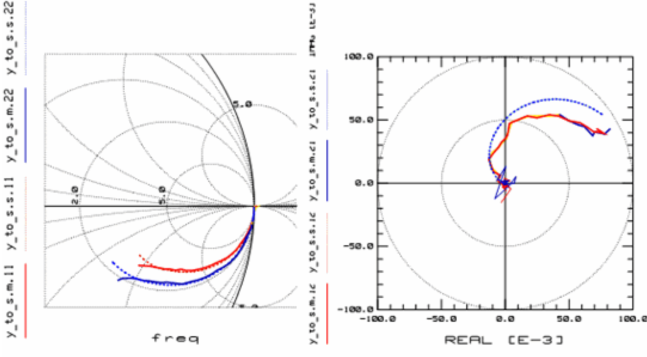
Finally in above figure, we see the real part of Z12 and Z21. Also here, we have an ohmic

behavior for frequencies above 3 GHz, and the mean value is RC = 30 Ohm. On the other hand, the imaginary part of Zxy gives no clear picture for our purpose and is therefore neglected.

We are now finished with the development of the equivalent schematic. The resulting lumped circuit is depicted in following figure:



Finally, we define a netlist using the appropriate simulator syntax like UCB-Spice, MDS, Eldo, Saber, Spectre, Hspice, or Pspice. The simulation results are then plotted in following figures together with the measurement results. Using the optimizer will give us the fine-tuned parameter values.

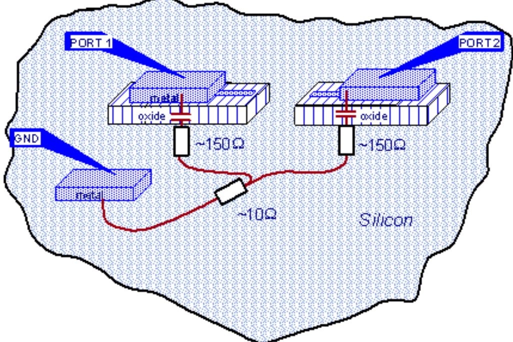


Sxx measured and simulated　　Sxy measured and simulated

The final rms fitting error using the proposed equivalent schematic is about 15%. With respect to our goal of simply subtracting the Y matrix of our 'open dummy structure' for de-embedding the DUT on the wafer, we can state that the developed equivalent schematic is totally in parallel with the 'inner device'. Therefore a simple subtraction of the S-to-Y transformed 'dummy device' data from the also S-to-Y transformed 'total device' data will be sufficient for the de-embedding of the 'inner device'.

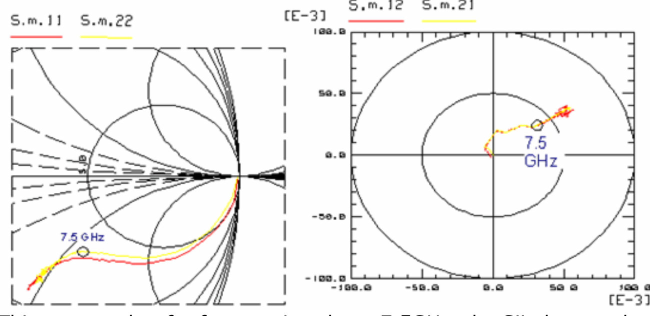
Physical Interpretation of the Equivalent Schematic:

The resistors RA and RB can be interpreted as the loss in the oxide capacitor, while RC is basically the resistance in the bulk of the substrate. The figure below shows the physical interpretation of the OPEN dummy schematic. Shown are the test pads, the probes and the silicon substrate.



Modeling the Second Order Effects of an Open Dummy Structure

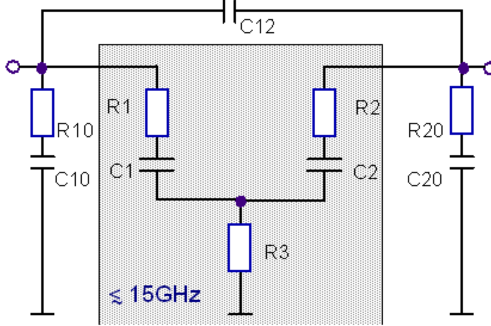
Like already a bit visible in above figures, the measured traces may deviate from the half-circle for high frequencies. This indicates that the underlying schematic is too simple. Fig.18 depicts such a situation for a 50GHz open dummy measurement.



This means, that for frequencies above 7.5GHz, the Sii plots tend to another half circle, which will hit the X axis for infinite frequencies at a much lower ohmic value. This effect, also explainable with HF power finding a lower-resistance path to ground for higher frequencies than with the simple equivalent schematic, is now modeled with the components C10, R10 resp. C20, R20. As stated, their capacitance values are much smaller than those of C1 and C2. (See also chapter 3.3.4., section 'About the interpretation of S-parameters of an asymmetrical TEE structure).

Physically, this new HF power path is cross-coupling between the pads of the dummy structure.

For the Sij plot in above figure, the trace showing up at frequencies above 7.5GHz, means more cross-over power from port1 to port2 and vice versa than for frequencies below 7.5GHz. Since the so far added C10, R10 do definitively not lead to more HF power flowing between the ports, this effect can only be described with an additional cross-over capacitance C12. The Figure below explains the improved schematic:



The parameter extraction is done as follows:
Firstly, the model parameters of R1, R2, R3 and C1, C2 are extracted as mean values from the transformed S-parameter curves, for a frequency range up to ~10..15GHz and following the procedures mentioned in the section above.

Note
Since the underlying simple TEE structure is not valid for the data, the traces will not be flat, but frequency dependent!

Then, the additional components C10, R10, C20, and R20 are added to the circuit and tuned with the tuning feature in IC-CAP, mainly fitting the Sxx traces. Then, C12 is added too and fine-tuned basically to fit the Sxy traces, and not distorting the Sxx plots.

Finally, an optimizer run will find the best parameter values.

PUBLICATIONS:

an interesting publication on OPEN dummies on different silicon substrates is:
L.P.Chen et.al., Extraction of Parasitic Parameters of Dummy Devices on Different Silicon Substrates, 1998 IEEE Radio Frequency Integrated Circuits Symposium, June 7-9, 1998, Baltimore, pp. 321-324, ISSN 1097-2633

Modeling Frequency Dependent Losses

IC_CAP demo files: skin_tdr.mdl, skin_spar.mdl

Time Domain:

Frequency dependent losses of lines show up with TDR measurements as a decline of the slope of the TDR step function.

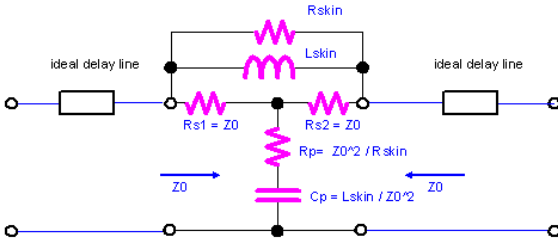
Frequency Domain:

Regarding S-parameter measurements, the effect is represented by an S-shaped decrease in the MAG(Sxy).

Modeling

Conventional SPICE models, however, cannot cover this effect. They only allow the modeling of ideal delay lines without a decrease in magnitude. Modern high-frequency simulators like ADS feature several models to cover different non-ideal effects of lines.

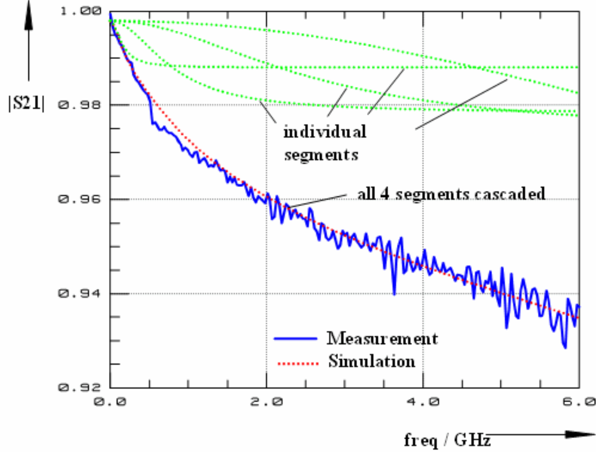
/Katzier/ proposes a special SPICE subcircuit for modeling this effect by using conventional SPICE components.



Lacross and Racross are modeling the damping of |S21|. Here, Lacross is the most important fitting parameter and Racrossis used for fine-tuning. Note that the main effect is the variation in the magnitude of Sxy, while the phase of Sxy is nearly not affected. There are no analytic methods to extract the values of Lacross and Racross and therefore, their values are either tuned or optimized.

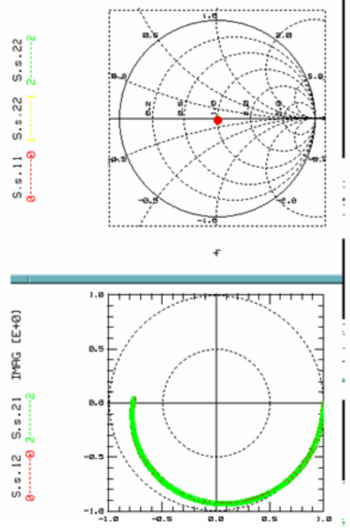
On the other hand, the proposed schematic can easily be cascaded, due to its nice feature of exhibiting Z0 at its input and output independently of any parameter settings.

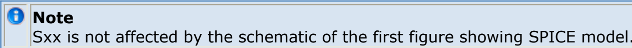
The following figure depicts such a scenario:
the modeling of a 25cm long line using a cascade of 4 schematics. the model of above figure can easily be cascaded for an improved fit of geometrically long devices, e.g. strip lines etc. (plot by courtesy of H.Katzier)



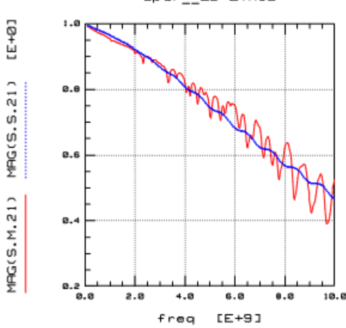

```
.SUBCKT mdlfreqlosses 1=P1 20=P2
Rs1 1 3  T1.Z0
Rs2 3 2  T1.Z0
Lacross  1 2  1n
Racross  1 2  10
Rp   3 33  T1.Z0^2//Racross
Cp  33  0  Lacross//T1.Z0^2
T1 2 0 20 0 Z0=50 TD=200p
.ENDS
```

The above table shows the SPICE input deck for the circuit of the figure shown above. The performance of this model including the ideal delay line is given with the next 3 plots. Note that the model affects mainly the decrease in magnitude of S21.





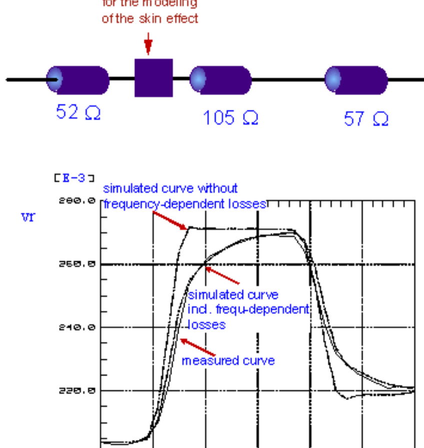
Note
Sxx is not affected by the schematic of the first figure showing SPICE model.



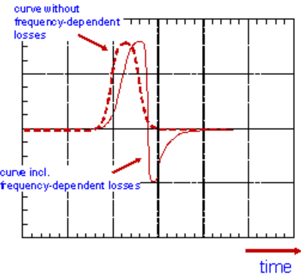
The above figure shows Frequency-dependent losses with S-parameters: the s-shaped decrease of MAG(Sxy) is modeled by the proposed schematic.

Time Domain

In the time domain, the frequency-dependent losses effect is visible as a quasi-'lowpass' slope of the back-reflected step response. The above figure shows a modeling of a chain of delay lines with and without this effect. Again, LACROSS and RACROSS were used to tune the slope.

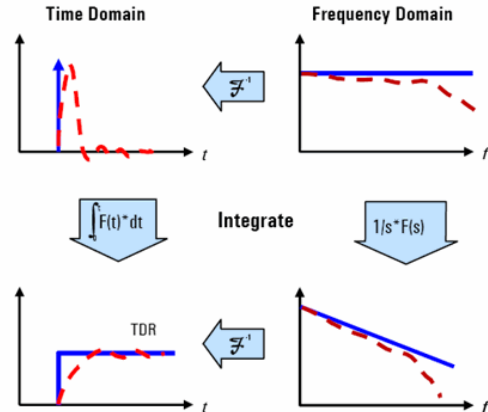


The above figure shows measuring and modeling the end of an open delay line: frequency-dependent losses effect with TDR measurements.
Applied to a Gaussian impulse, the effect becomes visible with a unsymmetrical impulse response with an undershoot. This proofs nicely the plot of fig.4, since a step function is the integral of a pulse.



Note
It should be noted that the proposed schematic features an excellent fitting capability in both, TDR and S-parameter domain. However, when keeping the length parameter of the preceding strip line unaffected, it requires to modify the strip line's delay time, or when applying a more physical strip line model like ADS's MLIN2, to modify the ?r of the substrate and the epsr_losstangent. In other words, the combination of delay line plus the proposed TEE schematic should be as a blackbox which then in total fits the delay line.

The figure below shows the effect in the time-domain and the frequency-domain.



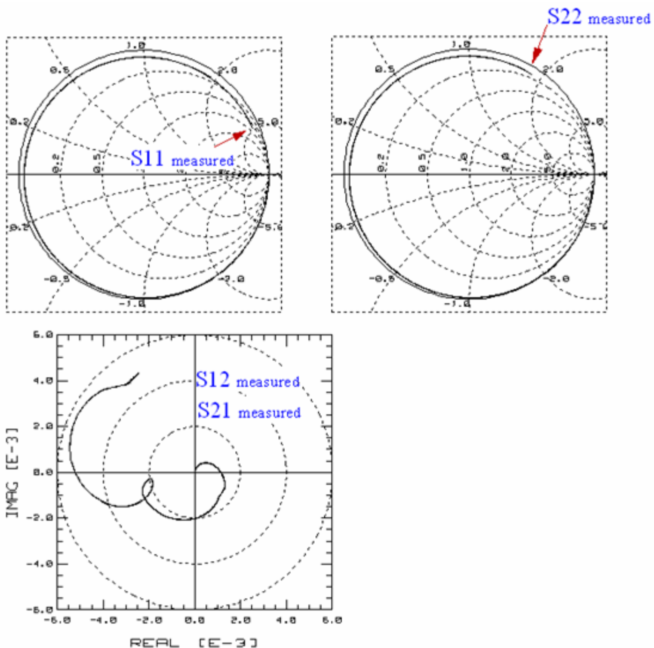
Publications:

H.Katzier: Analyse und Modellierung der Gesamtdaempfungsverluste von verlustbehafteten Streifenleitungen, 4.GMM/ITG Diskussionssitzung, October 1996, Berlin

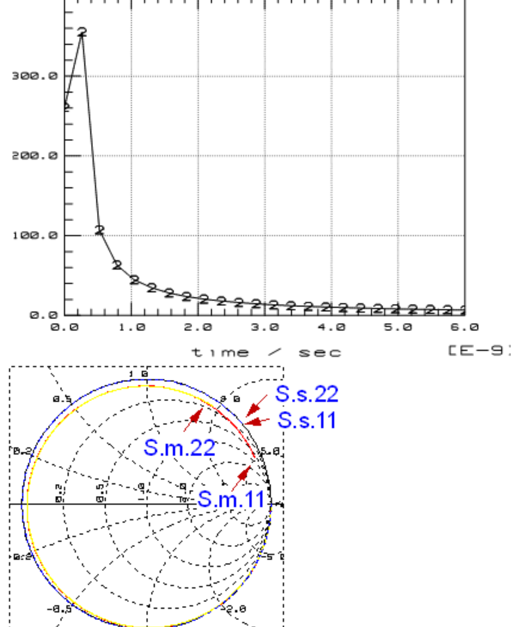
Test Fixture Modeling in Frequency Domain

This chapter gives an example on the modeling of a test fixture for packaged transistors.

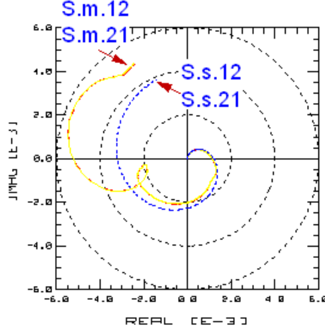
In a first step, the HP8753D NWA has been calibrated up to the ends of its 3.5mm connectors. Then the test fixture has been connected and measured. Figure 1 shows the result.



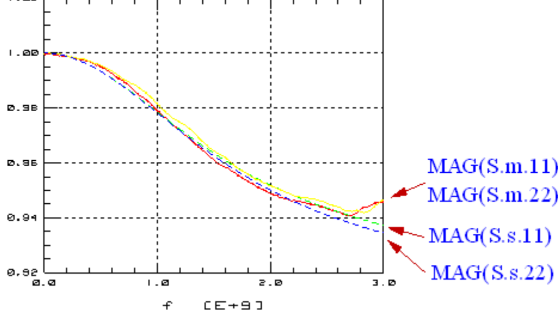
As the next step, we will check the applicability of delay lines to the equivalent schematic. Therefore, FFT transform is applied to get an idea on the time delays of the microstrip lines on the fixture. From fig.2 we get an estimation for the delay lines of our equivalent schematic. We assume TD=250ps. An optimizer run is used to overcome the uncertainties with this method that stem from the limited bandwidth of 3GHz of the HP8753D, and we get a delay of ~290ps. The phase fitting after this step is depicted in fig.3 below:



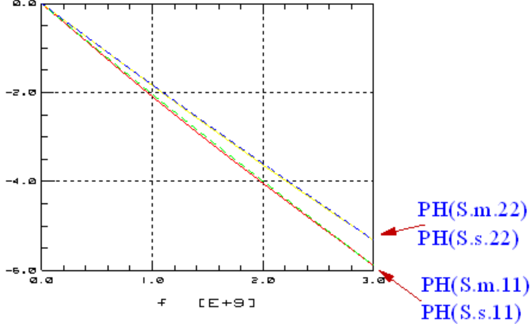
Next, the schematic for the simulation is developed further. From Sxy, the cross-coupling can be modeled with a capacitor C and a resistor R.



Then, back to Sxx and the estimation of 28fF for the non-idealities of the open strip lines (which will be contacted by the SOT23 package later), Cx and Rx have been added to help in the modeling of the decrease in MAG_Sxx. During this procedure, the delay times of the strip lines at port1 and port2 were re-optimized manually.

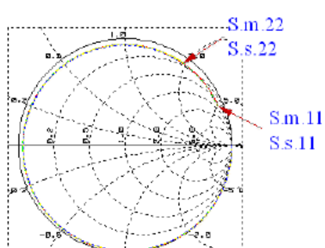


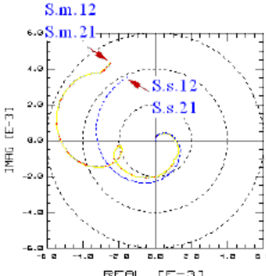
From the phase plots in ph_Sxx, it can be found that a split of each delay line into two parts with different Z0 improves the fit. This tiny improvement affects very much the de-embedding quality later.




```
.SUBCKT MEAS  2=P2  1=P1  3=E
T1   1  0  11 0 Z0=50 TD=76.6p
T11  11 0 111 0 Z0=47 TD=76.6p
C1  111  1111 85f
R1  1111 0 1k
Copen1 111 0 28f
T2   2 0  22 0 Z0=50 TD=69p
T22 22 0 222 0 Z0=48 TD=69p
C2  222  2222 85f
R2  2222 0 1k
Copen2 222 0 28f
C   111 12 2.6f
R   12 222 6.9k
.ENDS
```

The final modeling results are given below.





Modeling of Spiral Inductors

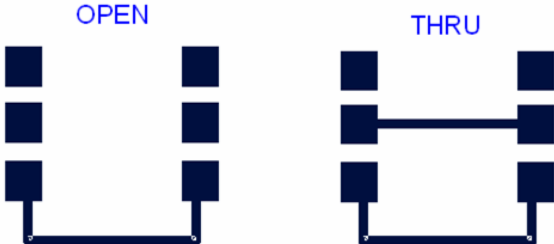
On-wafer spiral inductors are a commonly used component in integrated high frequency amplifiers. Their performance contributes a lot to the overall performance. Additionally, they occupy a big space on the wafer, what makes them ext4remely expensive components. Therefore, a thorough modeling, usually for both, the S-parameter and the Q factor, has to be applied.

De-embedding the measured data before modeling

It must be emphasized that one of the most important issues for accurate spiral inductor modeling is a preceding, very thorough de-embedding. Since basically all passive components exhibit a pretty large geometry compared to the transistor area, the connection lines from the test pads to the component itself cannot be neglected. Therefore, a dummy THRU has to be measured besides the conventional OPEN dummy. Such a THRU dummy can be considered as an equivalent to the SHORT dummy for transistors.

Altogether, the de-embedding procedure is like this:

1. de-embed the total measurement from the OPEN dummy first,
2. then, the THRU dummy is de-embedded from the OPEN dummy and its phase shift and if required, the loss is modeled with respect to the length of the THRU strip line. After that, we can calculate the equivalent ABC matrices of stub at port 1 and of port 2.
3. finally, de-embed the result of step (1) from the de-embedded THRU of step (2).



While the OPEN de-embedding is performed applied to Y matrices, the THRU is best represented by either an S matrix or an ABC matrix. If the THRU can be modeled by a lossless delay line, we calculate the de-embedded performance of the spiral inductor directly out of the total S-parameters by

$$\begin{pmatrix} S11_{DUT} & S12_{DUT} \\ S21_{DUT} & S22_{DUT} \end{pmatrix} = \begin{pmatrix} S11_{total}*\exp(j*2PI*freq*(2*TD1)) & S12_{total}*\exp(j*2PI*freq*(TD1+TD2)) \\ S21_{total}*\exp(j*2PI*freq*(TD1+TD2)) & S22_{total}*\exp(j*2PI*freq*(2*TD2)) \end{pmatrix}$$

with TD1= delay time of strip line at port1
TD2= delay time of strip line at port2

If the THRU is represented by a lossy delay line, we apply ABC matrix de-embedding. This is done by firstly modeling the lossy delay line with an ABC matrix each like

$$(A_{line}[freq]) = \begin{pmatrix} \cosh(aux[freq]) & Z0 * \sinh(aux[freq]) \\ \frac{1}{Z0}\sinh(aux[freq]) & \cosh(aux[freq]) \end{pmatrix}$$

$$aux[freq] = \left(\sqrt{2*PI*freq}*\alpha + j*2*PI*freq*\beta\right)*L$$

where
L = length of the strip line
α = damping factor
β = phase shift factor
and then, secondly, applying the de-embedding for both strip lines at both connections of the spiral inductor as

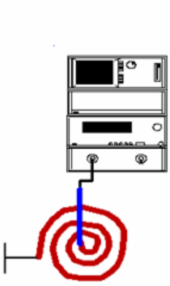
$$A_{DUT} = A_{line\,1}^{-1} * A_{total} * A_{line\,2}^{-1}$$

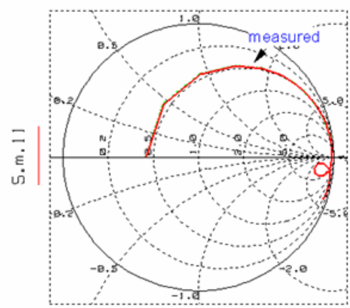
See the chapter 3.3.3.3 on De-embedding for more details.

1-port modeling of a spiral inductor

This chapter presents a tutorial about how to use IC-Cap's PEL language for the flexible and accurate modeling of spiral inductors, measured as a 1-port component (See fig.1). In this example, PEL is used for extractions and also for the implementation of equations into the SPICE netlist (e.g. resonance formula).

Fig.1 shows the test setup and the measurement result of the spiral inductor. The measurement was taken from 300kHZ to 3GHz using an HP8753 network analyzer under IC-CAP.





The modeling was performed in the following order:
Looking at the lowest frequencies in fig.1, we assume a resistive behavior at 300kHz. The real part of S11 is
REAL(S11) = -0,387
or

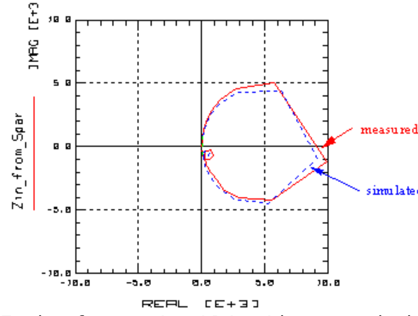
$$RL1 = 50 * \frac{1+S_{11}}{1-S_{11}} = 22{,}1Ohm$$

This is our starting point for the equivalent schematic of fig.5 below.

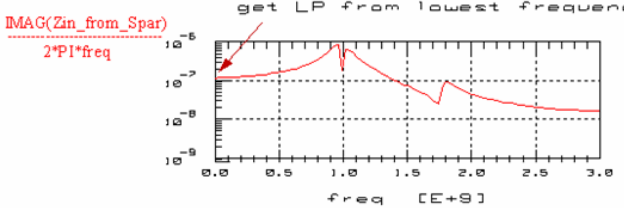
Then, in order to get the value of the designed inductor, we convert the measured S11 parameters to the complex resistance plane following

$$\underline{Zin_from_Spar} = 50 * \frac{1+\underline{S_{11}}}{1-\underline{S_{11}}}$$

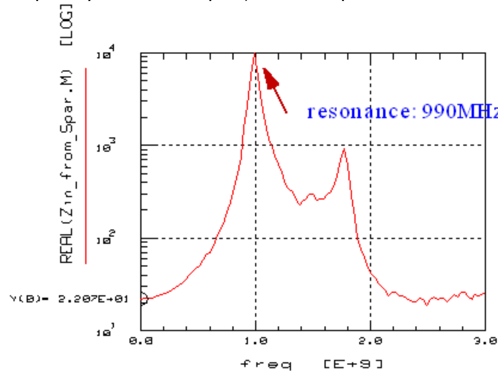
The result is given in fig.2.



For low frequencies, LP is without any doubt affecting the curve the most. Therefore, we divide the imaginary part of the complex resistance of equation (2) by (2*PI*freq). The result is shown in fig.3. From the lowest frequencies, we obtain Lp=120nH.



It can be seen that our DUT exhibits a resonance at high and another one at even higher frequencies. Next, fig.4 shows the real part of the complex resistance (fig.2) versus frequency. From this plot, both frequencies can be determined easily.



Considering first the resonance at 990MHz, we assume a capacitor Cp. in parallel to LP. Furthermore, we neglect the second resonance for the moment. This allows to apply the resonance frequency formula equ.3:

$$Res_freq = \frac{1}{2 * PI * \sqrt{Lp * Cp}}$$

Solved for Cp., and introducing Res_freq=990MHz , we obtain for Cp.

$$Cp = \frac{1}{(2 * PI * 990MEG)^2 * Lp}$$

with LP=120nH from fig.3.

> **Note**
> The formulas can be entered directly into IC-Cap's parameter list, allowing optimization with nested parameters.

Since the S11 curve of fig.1 does not go exactly through S11=1 at the first resonance, we assume a parallel resistor to LP and Cp. Its value can be obtained using equation (1) at the resonance frequency of 990MHz. We obtain:
Rp=10 Ohm.

Now, we are ready for the modeling of the second resonance.

It has to be a circuit, which behaves neutral for low and high frequencies. This means a series resonance circuit, in parallel to the so far developed schematic. Therefore, we add the three more components RS, Cs and Ls to the schematic in fig.5. They represent the resonance at 1.7GHz, see fig.4. From fig.1, we further assume a bandwidth of 150MHz.

	series resonance	parallel resonance
resonance frequency	$f_0 = \frac{1}{2PI\sqrt{L * C}}$	$f_0 = \frac{1}{2PI\sqrt{L * C}}$
bandwidth	$BW = \frac{R}{2PI * L}$	$BW = \frac{R}{2PI * L}$

Referring to table 1, the parameter values for Cs and RS can be calculated dependent on a given Ls as follows (f0=1.7GHz, BW=150MHz):

$$Cs = \frac{1}{(2 * PI * 1{,}7\,G)^2 * Ls}$$

$$R_S = 2 * PI * Ls * 150MEG$$

Again, these formulas can be entered as 'values' into IC-CAP's parameter list. Therefore, by tuning Ls or using the optimizer, we can improve the total fitting for fig.1 at the second resonance (set XMIN and XMAX in the optimizer).

Of course, both resonance circuits overlay each other, and fine-tuning the second resonance will distort the fitting of the first one and vice versa. Therefore, we introduce a tuning parameter DELTA for Cp. as shown in equation (7) for a final optimization step.

$$Cp = \frac{1}{(2 * PI * 990MEG)^2 * Lp} + DELTA$$

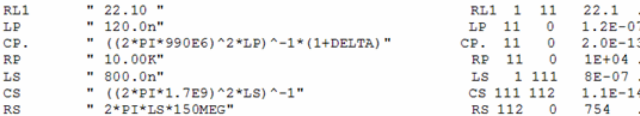
This DELTA parameter is optimized with respect to the phase of S11.

Figure 5 shows the final result: the SPICE circuit deck and the schematic of our modeling approach. Fig.6 depicts the IC-CAP parameter table with the used formulas and also the final parameter values.

The curve fitting is given in fig.7. As can be seen, there is a very good agreement between the measured and simulated curve.

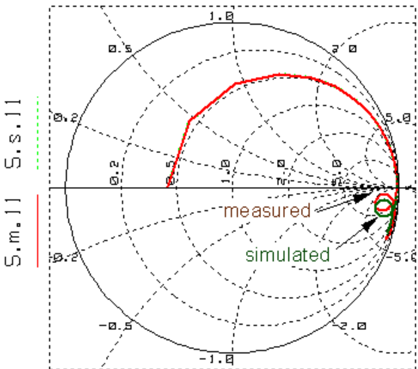
```
.SUBCKT MEAS 2=P2 1=P1

RL1  1  11 23
LP  11   0 100p
CP  11   0 100f
RP  11   0 1k

LS   1 111 10n
CS 111 112 10p
RS 112   0 200

Rdummy 2 0 1T
.ENDS
```

IC-CAP parameter table including the formulas and the final parameter values is:

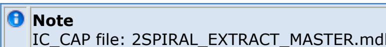


```
RL1      " 22.10 "                            RL1   1   11   22.1  .
LP       " 120.0n"                             LP  11    0   1.2E-07
CP.      " ((2*PI*990E6)^2*LP)^-1*(1+DELTA)"  CP.  11    0   2.0E-13
RP       " 10.00K"                             RP  11    0   1E+04 .
LS       " 800.0n"                             LS   1  111   8E-07 .
CS       " ((2*PI*1.7E9)^2*LS)^-1"             CS 111  112   1.1E-14
RS       " 2*PI*LS*150MEG"                     RS 112    0   754   .
```

Model fit for the 1-port spiral inductor is:



2-port modeling of a spiral inductor



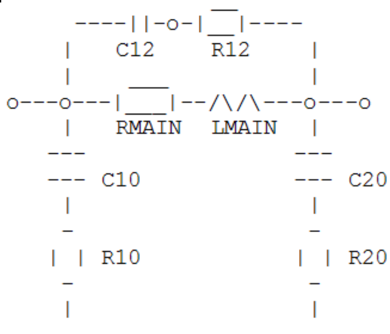
Note
IC_CAP file: 2SPIRAL_EXTRACT_MASTER.mdl

In this example, the spiral inductor has been measured between the ports of the network analyzer. In a first approach, we will try to find the parameters of a typical, given circuit for such type of inductors. In most cases, however, this will not lead to the desired frequency independent model parameters. Therefore, in a second approach, we will develop a suitable subcircuit out of the measured S-parameter data and calculate the model parameter values for the lumped components. This method will follow two approaches: a first approach which assumes a inner PI structure, possibly overlaid by outer parasitics, and a second one which assumes a PI for then inner part of the schematic.

Let's first go for the standard method, a given, predefined circuit.
Referring to the mentioned IC-CAP file, we consider

DUT spiral_ind_quicky

This is the underlying schematic for the standard interpretation of the Spiral S-parameters:



The parameter values can be obtained as follows:

1. Transform the S parameters to Y
   Interprete the Y-pars after a simple PI schematic
```
 o-o-|__|-o-o
   | | C    |
   -        -
  |A|      |B|
   -        -
   |        |
  ---      ---
```
2. calculate Yxy = (Y12 + Y21) / 2
   and then A = Y11 + Yxy
   B = Y22 + Yxy
   C = -Yxy
   Z10 = 1/A
   R10 = REAL(Z10)
   C10 = (-IMAG(Z10)*2*PI*freq)^-1

Z20 = 1/B
R20 = REAL(Z20)
C20 = (-IMAG(Z20)*2*PI*freq)^-1

Z12 = 1/C
RMAIN = REAL(Z12) !RMAIN: typ. before resonance, at low frequ.
LMAIN = IMAG(Z12)//(2*PI*freq) !LMAIN: typ. before resonance, at low frequ.
C12 = -1/(2*PI*freq*IMAG(Z12)) !C12: typ. after resonance, at high frequ.

The underlying schematic is too simple for today's spirals up to 26GHz.
This is indicated by the frequency dependent component values. An other indication is that the schematic isn't able to produce an inductor Sxy simulation trace **around** the zero. It is rather tending **towards** the zero.. Such a trace is a hint for another overlying, more complex schematic.

Therefore, the next DUTs in the mentioned IC-CAP file 2SPIRAL_EXTRACT_MASTER.mdl refer to publications of more complex schematics.

Finally, the IC-CAP file features a more general passive components modeling tool. this is described below:

DUT check_4_PI

The setups of this chapter develop an equivalent schematic based on a PI structure. Both, the details of the schematic, and the parameter values are investigated. It is checked, if there is an overlying TEE structure, and if yes, the model parameters of this overlying structure are extracted as well.

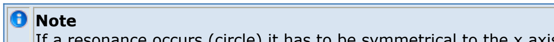
A little bit of theory:
This setup is used to develop a PI based schematic out of the deembedded data. Usually, the data are simply converted to Y parameters (see spiral_ind_quicky). In many cases, however, this is too simple, and an overlying more complex circuit must be assumed.

Therefore, the input, the transfer and the output resistance of a PI structure (Y matrix) are calculated and displayed, under the assumption of overlay series parasitics (TEE structure, Z matrix), and a cross-over capacitor (another Y matrix).
See the header of transform S_2_Y_2_Z_2_Y.

Purpose of transform S_2_Y_2_Z_2_Y:
The de-embedded S-parameters are converted first to Y, and a possibly overlying cross-capacitor C12_ is subtracted. Next, the resulting Y-parameters are converted to Z-parameters, and possibly overlying series components (L1_, R1_, L2_, R2_, R3_) are subtracted. After that, the remaining data are converted to Y parameter and finally checked for the inner PI based schematic.

However, the parameter values of these overlying parasitics are not yet known. Transform check_4_TEE_overlay includes TUNER function which calls transform S_2_Y_2_Z_2_Y. Starting with some reasonable default values for the assumed overlay parasitics, (middle position of the tuner), and including also parameter values which effectively switch-off the effect of the corresponding parasitic (left position of the tuner), the tuner bars are adjusted top-down in order to get as physical as possible remaining Y-parameters. In order to check the physical meaning of these Y-parameters, they are converted to Rin, Rtrans and Rout, and these complex resistances are plotted on a real/imaginary scale. Therefore, the transforms Zin_from_Ypar, Ztrans_from_Ypar, and Zout_from_Ypar calculate complex resistanc from the resulting S_2_Y_2_Z_2_Y Y-parameters. This is displayed in the plots ZIN_Y, ZTRANS_Y, and ZOUT_Y.

How to proceed:
As a first step, transform check_4_TEE_overlay includes a tuner, which allows the tuning of the assumed overlaying TEE parasitics. The goal is to tune these parasitics until we obtain physical plots ZIN_Y, ZTRANS_Y, and ZOUT_Y. This is true if all curves turn clock-wise with frequency, and the left half-plane of the complex resistance plane is not hit.


Note
If a resonance occurs (circle) it has to be symmetrical to the x axis.

After we obtain as physical as possible results for the plots ZIN_Y, ZTRANS_Y, and ZOUT_Y, we have automatically found the parameter values of the overlying parasitics.

Therefore, we now copy the adjusted parameter values, stored temporarily in the setup variable table, into the parameter list of the DUT test circuit.

Having obtained the input, the transfer and the output resistance of the inner, PI based schematic, we have to interprete these plots in the next step. Transform calc_L_or_C_utility supports this task.
This means, we middle-click a data point in one of these plots (beginning at the lowest frequency), and can calculate the corresponding inductor. Then, these values are entered into the parameter list. If required, the schematic is also enhanced for more components.

In order to further evaluate the schematic behind ZIN_Y, ZTRANS_Y, and ZOUT_Y, the setups Zin_strip_RC, Ztrans_strip_RL and Zout_strip_RC are used to strip off step-by-step the so far determined main components of the partial schematic and to determine the second order components. However, this method is very sensitive, and it might be faster to develop the schematic of the inner resistances by applying the calc_L_or_C_ulility to the plots ZIN_Y, ZTRANS_Y, and ZOUT_Y.

After the ZIN, ZTRANS and ZOUT plots are modeled satisfyingly, we go back to fit the S-parameter plots of this setup.

In the next step, all the parameters are finetuned, and optimized.

Note: in the tuner transforms, the tune bars are programmed so that their middle value is

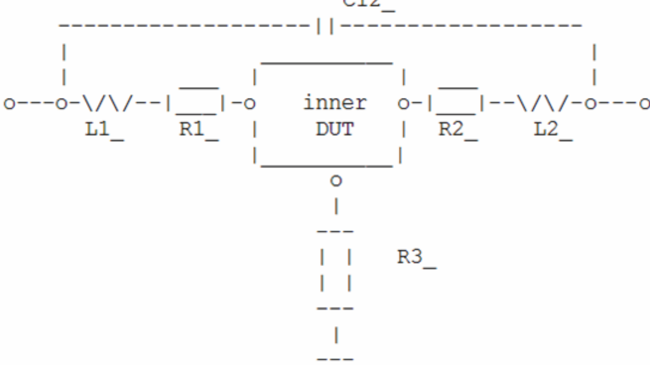
the starting value. This allows you to find your way back if you are lost!

Finally, Program fit_qual calculates the vector error between the measured de-embedded data and the simulation result

Program S11_1port calculates the 1-port S parameter out of the de-embedded data, and program Q calculates out of it the quality factor Q.

Contents of Transform S_2_Y_2_Z_2_Y

This transform performs a de-embedding of the following parasitic structure:

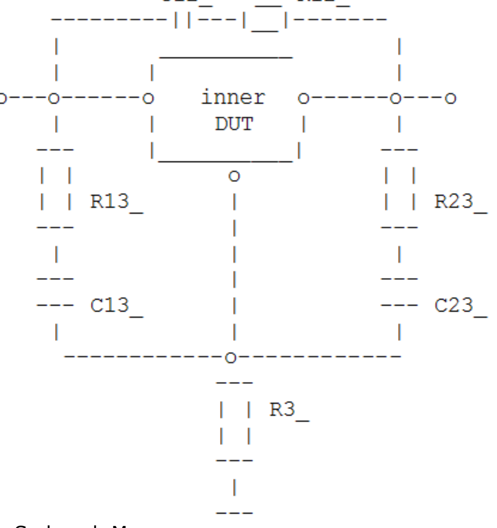



```
s=S_deemb.M
y = TwoPort(s,"S","Y")
!de-embed possibly overlaying C12 cross-couple capacitor
i=0
WHILE i < SIZE(y)
y.M.11[i]=y.M.11[i] -j*2*PI*freq[i]*C12_
y.M.12[i]=y.M.12[i] +j*2*PI*freq[i]*C12_
y.M.21[i]=y.M.21[i] +j*2*PI*freq[i]*C12_
y.M.22[i]=y.M.22[i] -j*2*PI*freq[i]*C12_
i = i + 1
END WHILE
z = TwoPort(y,"Y","Z")
!de-embed possibly overlaying TEE R_series_L-structure
i=0
WHILE i < SIZE(z)
z.M.11[i]=z.M.11[i] -R1_-j*2*PI*freq[i]*L1_ -R3_
z.M.12[i]=z.M.12[i] -R3_
z.M.21[i]=z.M.21[i] -R3_
z.M.22[i]=z.M.22[i] -R2_-j*2*PI*freq[i]*L2_ -R3_
i = i + 1
END WHILE
RETURN TwoPort(z,"Z","Y")
```

DUT check_4_TEE
same idea as above, but now investigating in a TEE schematic, possibly overlaid by a PI structure.

Contents of Transform S_2_Z_2_Y_2_Z
PRINT "running transform 'S_2_Z_2_Y_2_Z'..."

This transform performs a de-embedding of the following parasitic structure:



s=S_deemb.M

z = TwoPort(s,"S","Z")

de-embed possibly overlaying common ground resistor

```
i=0
WHILE i < SIZE (z)
z.M.11[i]=z.M.11[i] -R3_
z.M.12[i]=z.M.12[i] -R3_
z.M.21[i]=z.M.21[i] -R3_
z.M.22[i]=z.M.22[i] -R3_
i = i + 1
END WHILE
y = TwoPort(z,"Z","Y")
de-embed possibly overlaying PI structure (with R_series_C branches)
i=0
WHILE i < SIZE(y)
y.M.11[i]=y.M.11[i] -(R13_+(j*2*PI*freq[i]*C13_)^-1)^-1 -(R12_+(j*2*PI*freq[i]*C12_)^-1)^-1
y.M.12[i]=y.M.12[i] +(R12_+(j*2*PI*freq[i]*C12_)^-1)^-1
y.M.21[i]=y.M.21[i] +(R12_+(j*2*PI*freq[i]*C12_)^-1)^-1
y.M.22[i]=y.M.22[i] -(R23_+(j*2*PI*freq[i]*C23_)^-1)^-1 -(R12_+(j*2*PI*freq[i]*C12_)^-1)^-1
i = i + 1
END WHILE

RETURN TwoPort(y,"Y","Z")
```

Publications:

Y.Sun, J.L.Tauritz, R.G.F.Baets, Micromachined RF Passive Components ant Their Applications in MMICs, Int'l Journal of RF and Microwave Computer-Aided Engineering, vol.9, no.4, July 1999, Wiley&Sons, ISSN 1096-4290

Ali.M.Niknejad, Analysis, Design and Optimization of Spiral Inductors and Transformers for Si RF ICs, master thesis at the University of Berkeley, California, is on the web at http://kabuki.eecs.berkeley.edu/~niknejad/

Calculating the Q-factor of spiral inductors

The quality factor of a spiral inductor is represented by:

$$Q = \frac{IMAG(\underline{Z}_{inductor})}{REAL(\underline{Z}_{inductor})}$$

It cannot be calculated correctly out of the 2-Port S-Parameters because of the Z0=50 ? of Port 2.

In this case, we convert the 2-port S-parameters to 1-port S-parameters and obtain:

$$S_{11_1port} = S_{11} - \frac{S_{12} * S_{21}}{1 + S_{22}}$$

From that, we apply the basic S11 <> R conversion (Smith chart),

$$\underline{R} = Z0 \cdot \frac{1 + \underline{S_{11}}}{1 - \underline{S_{11}}}$$

and obtain for the input impedance at Port 1

$$Z_{11_1port} = Z_0 * \frac{1 + S_{11_1port}}{1 - S_{11_1port}}$$

This finally gives the requested 1-Port Q factor of the spiral inductor

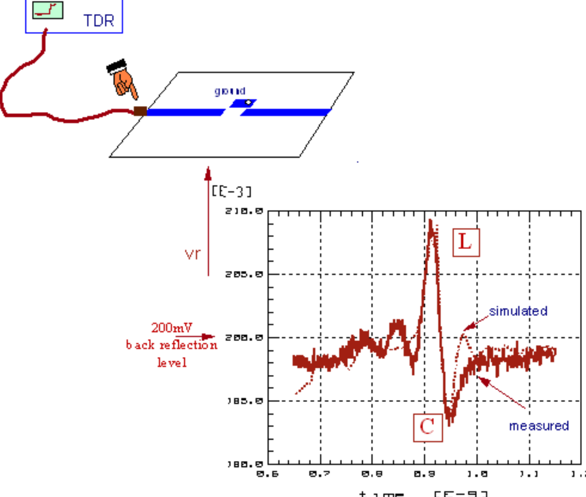
$$Q = \frac{IMAG(\underline{Z}_{11_1port})}{REAL(\underline{Z}_{11_1port})}$$

Note
The same result is obtained when converting the 2-port S-parameters to 2-port Y-parameters, and then calculating

$$Q = \frac{-IMAG(\underline{Y}_{11})}{REAL(\underline{Y}_{11})}$$

Modeling of a 3.5mm Connector

The connector is soldered to a board, made from FR4 material (up to 3GHz) or from Rogers (for higher frequencies).
The TDR cable is connected with the connector-board combination. Since a TDR allows an easy modeling 'from the left to the right', we can focus on the modeling of the connector.



On the TDR trace in the above figure, we see first an inductor effect, followed by a capacitor effect. This gives the first netlist approach for the connector model. The final parameter values of the connector capacitance cp1, the connector inductance ls1 as well as the connector delay with the delay line tconn are fine-tuned either manually (TUNER) or by an optimizer.

```
.subckt mdlconnector 101=tinp 105=toutp
v0   1  0    PWL(289p 1.375m  0.40404n 0)
...
* TDR step function into 50Ohm,
* plus delay line for the path TDR->Connector...
v99 100 99 PULSE(0 -0.625m  0.79596n  4.04p  4.04p 0)
ri    100  101   51.1
t1    101   0   102  0   TD=150p Z0=50
* now comes the connector
tconn 102   0   103  0   TD=38p Z0=50
ls1   103  104  460p
cp1   104   0   140f
* finally the line on the board (not to be modeled)
tmic1 104   0   105   0 TD=500p Z0=50
rend  105   0   1G
.ends
```

The above figure also gives the simulation result: the inductor models the overshoot, and the capacitor the undershoot.

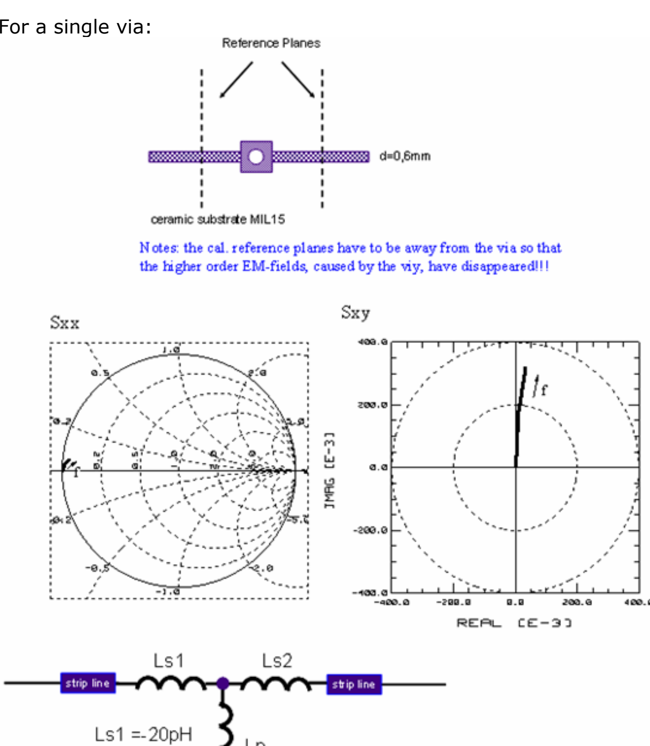
Delay time tconn.TD reflects the delay of the connector (33ps). Delay line tmic1.TD (for the modeling of the delay time of the strip line on the board) is set to 500ps and is not modeled here in this example.

Modeling Vias

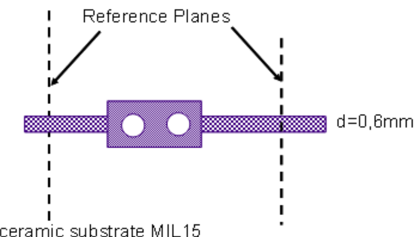
This chapter is about the modeling of vias on ceramic substrates. The test structure is a 50 Ohm strip line between the NWA ports, with different vias to ground (backside of the ceramic substrate) in the middle. This means, we will expect more Sxy for higher frequencies, since the via will behave worse and worse.

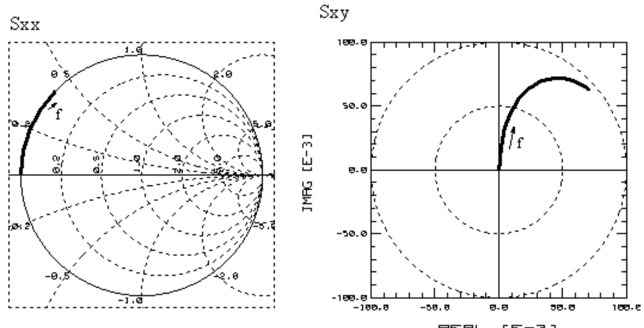
The following examples are cited from literature /Böhm/.

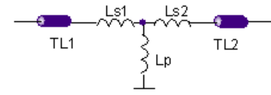
For a single via:



For a serial twin-via:



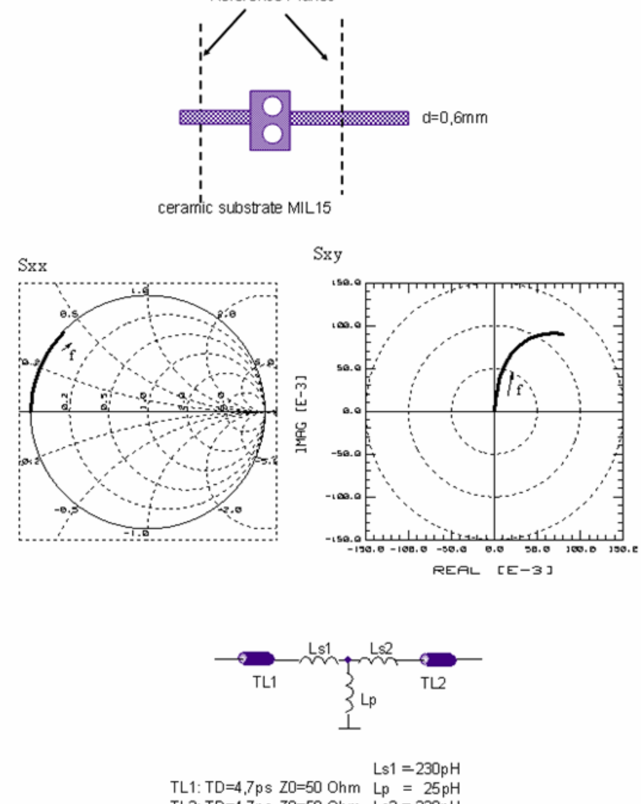




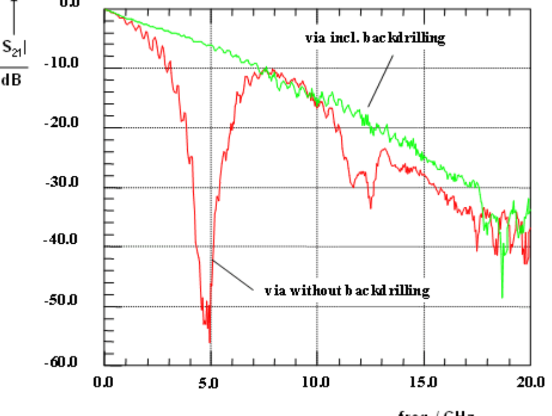
TL1: TD=5,8ps Z0=60 Ohm Ls1 =-290pH
TL2: TD=5,8ps Z0=60 Ohm Lp = 23pH
Ls2 =-290pH

Note
The higher phase shift stems from the broader reference planes!

And for a parallel twin-via:



Note
When vias are used to connect more than 2 layers, a special antenna effect may occur. For example, if a via in a 12-layer board only interconnects layer1 with layer2, the remaining via down to the last 12. layer may act as a resonance. This is depicted below.

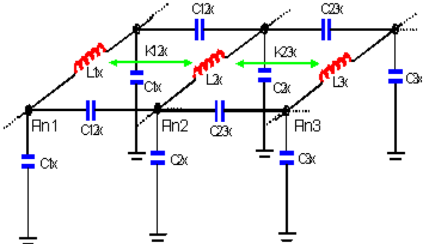


Publications:
M.Böhm, Diploma Thesis: Meßtechnische und rechnerische Untersuchungen von Verbindungsstrukturen für elektronische Schaltungen in einem Frequenzbereich von 45MHz bis 25GHz, TU Ilmenau/Thüringen, Germany, Oktober 1992

Package Modeling Workshop

Introduction

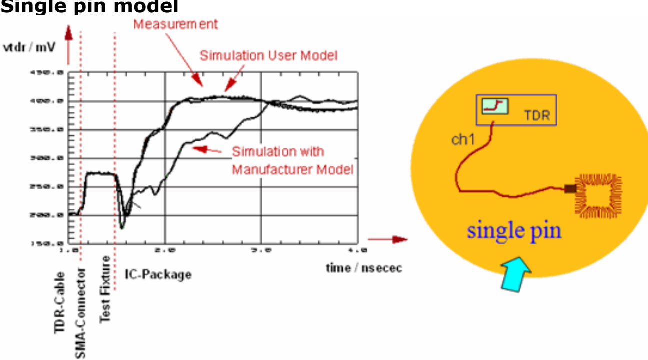
With today's clock rates,where chips become faster than the package performance, the package cannot be neglected any more and needs to be modeled very accurately.



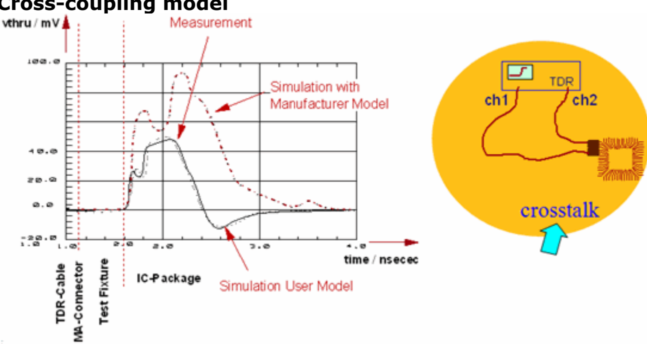
Limitations with Standard Models

Some problems with manufacturer's models such as:

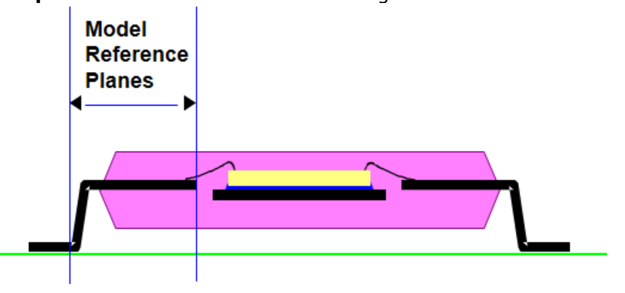
Single pin model



Cross-coupling model



Step1. The Reference Plane of a Package Model has to be defined



Note
It is very important to define the beginning and the end of the package model.

Two Types of Characterization Measurements Required

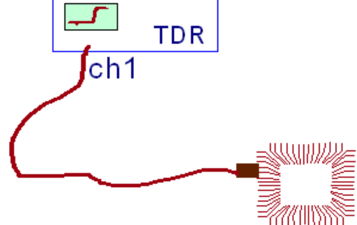
When compared to a conventional transistor modeling, packages represent a big geometry. This implies a much more complex equivalent schematic. It contains Ls, Rs, Cs and Delay Lines which are directly related to the geometries of the package.

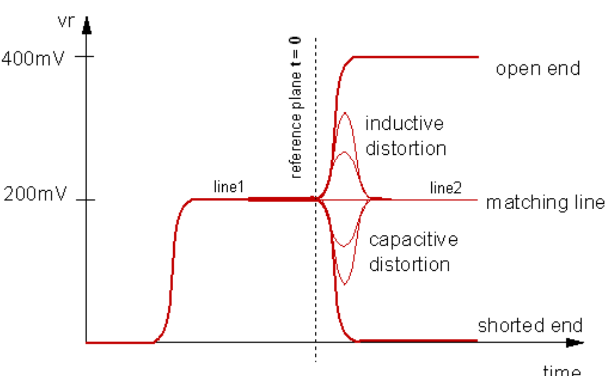
- Time Domain Reflectometer (TDR) measurements allow the identification of schematic components related to geometries, but are poor in crosstalk resolution between the package pins.

- Vector Network Analyzer (VNA) measurements have excellent resolution related to crosstalk, but S-parameter measurements are difficult to interpret for big delay times.

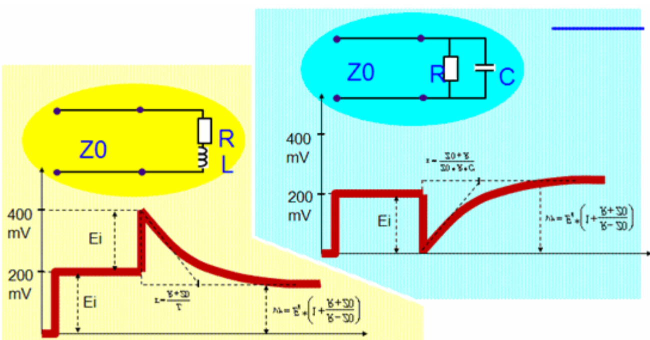
Time Domain Reflectometer (TDR) basics

For Identifying the Lumped Components of a Single Line, Time Domain Reflectometer (TDR) is commonly used. A TDR stimulates a very fast step voltage and measures its back reflections. The measurement trace allows then to identify capacitors, inductors and delay lines with respect to their physical location.

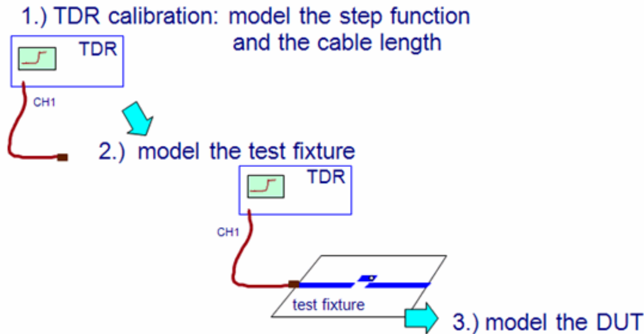




Interpreting TDR Traces

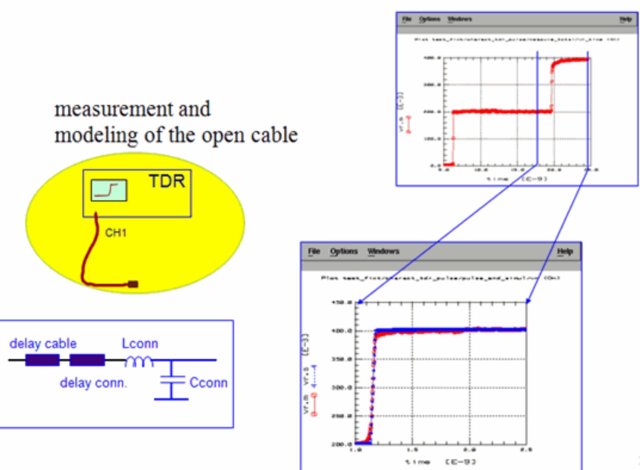


Time Domain Reflectometry Basics



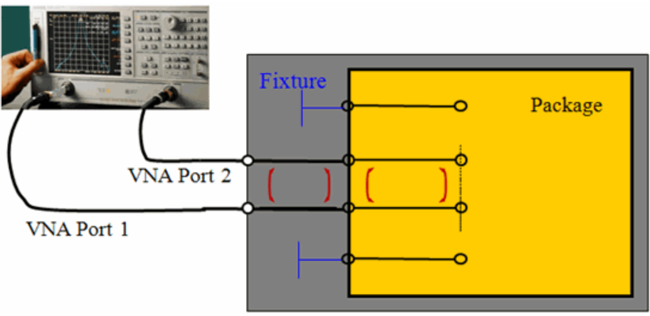
TDR Applied to Package Modeling

TDR Calibration

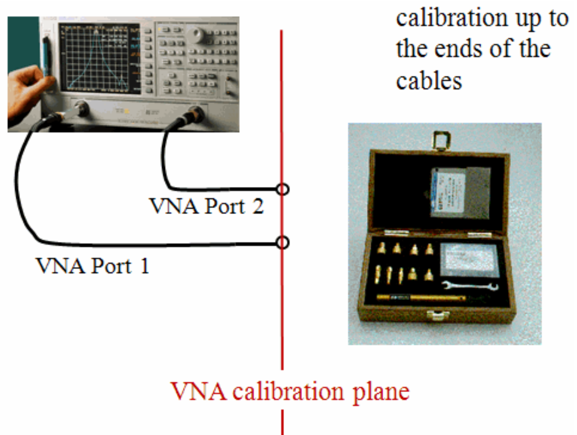


Vector Network Analyzer (VNA) basics

VNA Usage for Package Modeling for accuracy and cross-talk



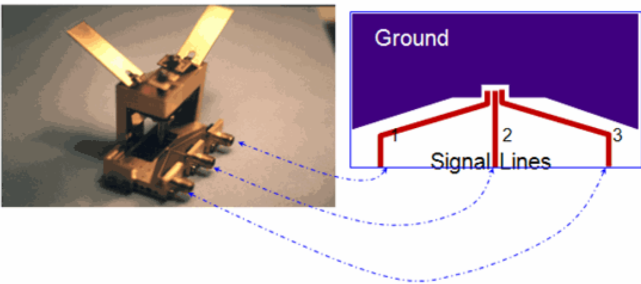
Vector Network Analyzer Calibration



After both, the TDR and VNA have been calibrated, we now need to model the test fixture.

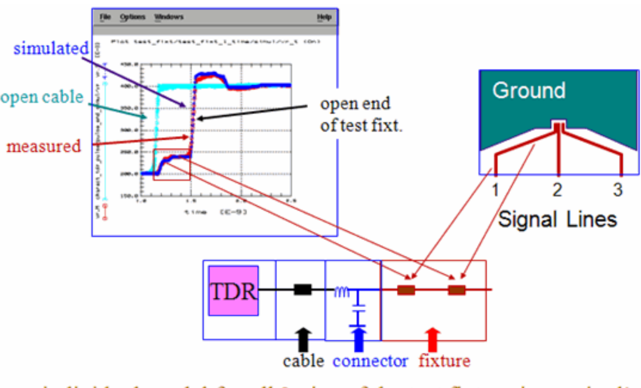
Test Fixture for Package Modeling

Test fixture for TSOP44 for DRAMs

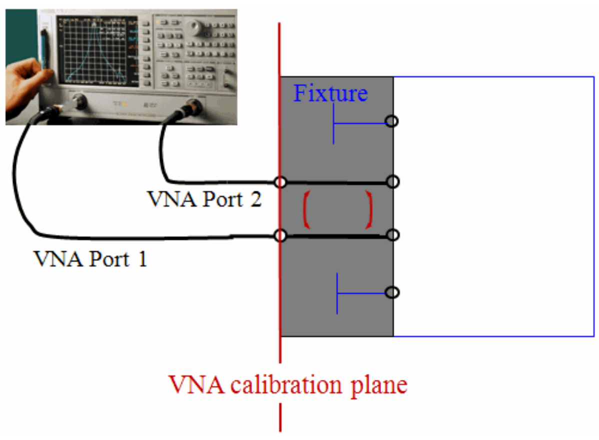


Before modeling the package, the test fixture has to be modeled standalone. Then, after inserting the package, the additional model components belong to the package.

Test Fixture Time Domain Modeling Result

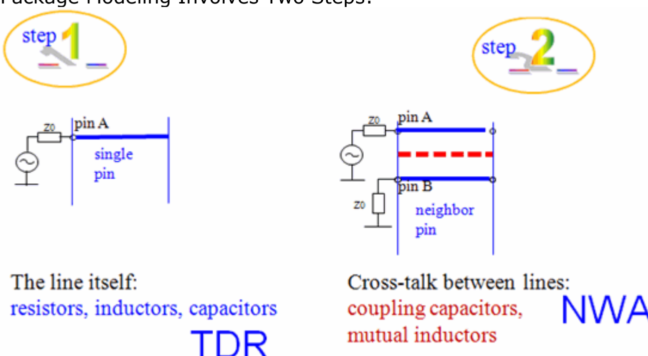


Network Analyzer test fixture model fine-tuning

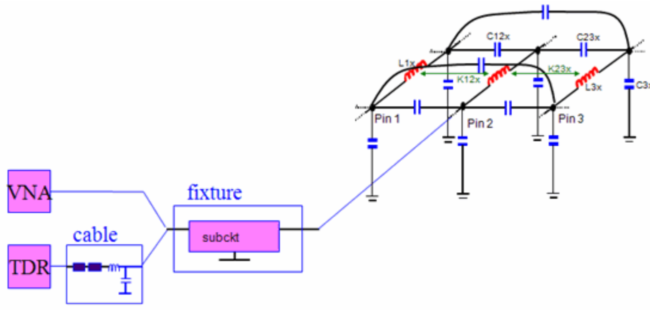


Package Modeling

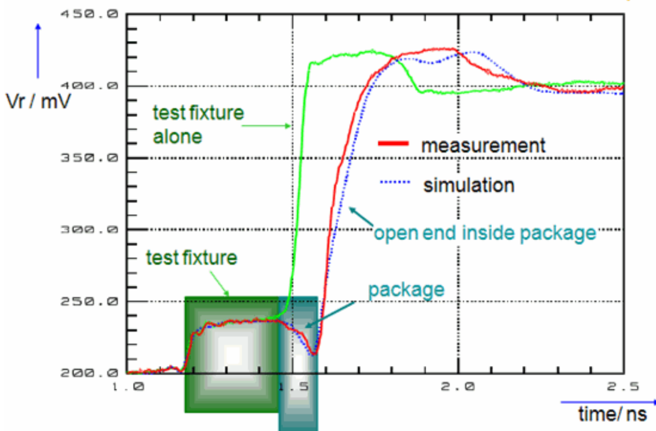
After the instruments are calibrated, and the fixture is modeled, Package modeling is done.
Package Modeling Involves Two Steps:



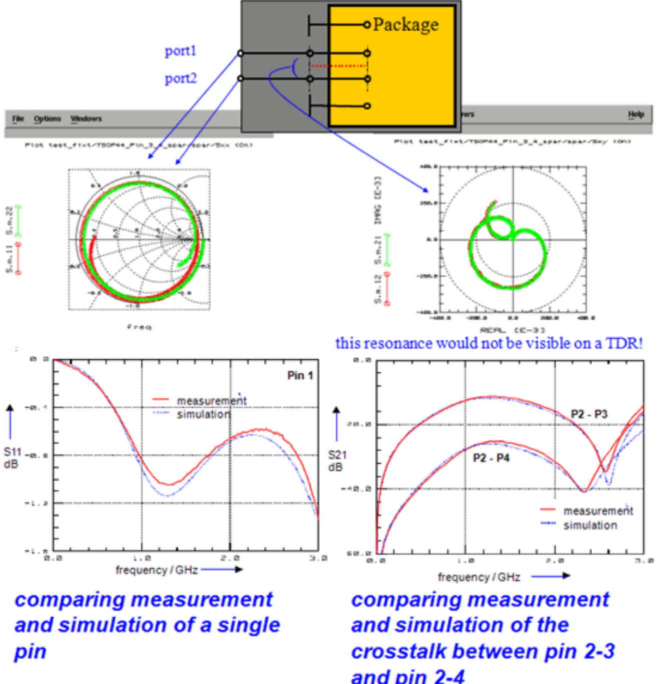
The Basic Equivalent Schematic



TDR Modeling Result (incl.fixture)- Step1



TDR Modeling Result (incl.fixture)- Step2

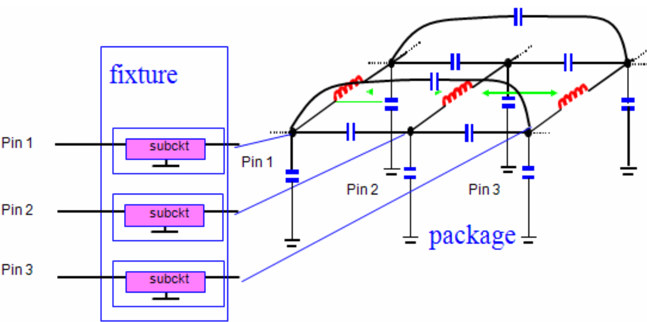


comparing measurement and simulation of a single pin

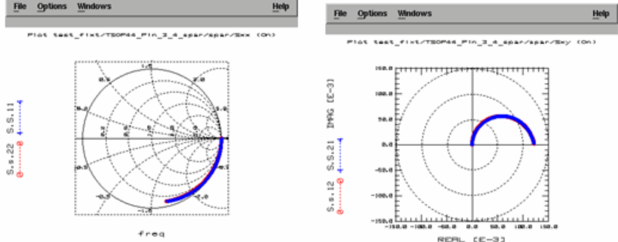
comparing measurement and simulation of the crosstalk between pin 2-3 and pin 2-4

The Total Model

After both, the test fixture and the package have been carefully modeled step-by-step, the package model can be clearly identified out of the total schematic.



The Final Package S-Parameters



Package Modeling Conclusions

A Method to Develop an Equivalent SPICE Schematic for packages has been proposed. The adaptive modeling procedure assures accurate and reliable models, and thus, successful system designs.

Modeling of a SMT Resistor

In this case, the NWA has to be calibrated with a suitable method in order to shift the calibration plane right up to the pins of the DUT. (see /Gronau 1992/ and also the chapter 'network analyzer measurements' for details)

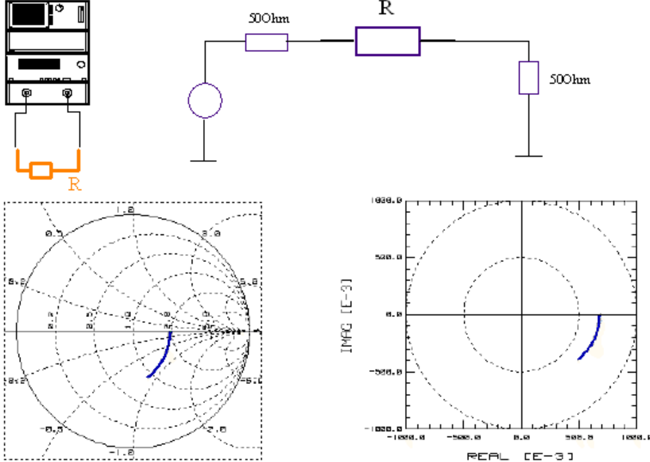


Fig.2 shows the measurement result of a 47 Ohm resistor. The Smith chart shows a real part of normalized 1.95*50 Ohm = 97,5 Ohm for low frequencies. I.e. the DUT plus the 50 Ohms of port 2 of the NWA are in series with each other. For higher frequencies, the resistor shows a capacitive behavior.

The polar diagram of S21 shows power transmission from port1 to port2 with capacitive effects at higher frequencies. This leads to the equivalent schematic below.

```
              L1                               L2
o---o---/\/\---o---|___|---o---/\/\---o---o
    |          |     R0    |          |
   ---        ---         ---        ---
   ---        ---         ---        ---
C11 |          | C12  C22 |           | C21
   ---        ---         ---        ---
```

with

```
R0      = 47Ohm        L1=L2   = 60pH
C11=C21 = 70fF         C12=C22 = 90fF
```

PUBLICATIONS about the NWA calibration:

G.Gronau, Scattering Parameter Measurement of Microstrip Devices, Microwave Journal, Nov.1992.

G.Gronau, I.Wolff,: A Simple Broad-Band Device De-embedding Method Using an Automatic Network Analyzer with Time-Domain Option, IEEE Trans.on Microwave Theory and Tech., Vol. 37, No.3, March 1989.

M.Böhm, Diploma Thesis Technical University Ilmenau/Germany: Meßtechnische und rechnerische Untersuchungen von Verbindungsstrukturen für elektronische Schaltungen in einem Frequenzbereich von 45MHZ bis 25GHz, 30.4.1993.

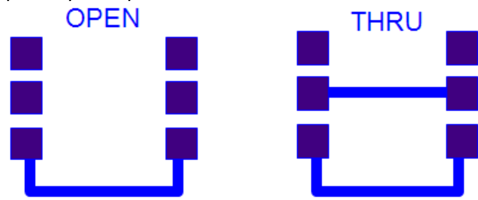
Rudolf Stassen, 'Einsatz eines Mikrowellen-Spitzenmeßplatzes zur Charakterisierung von Transistore direkt auf dem Wafer im Frequenzbereich von 0,045 bis 26,5GHz', Diplomarbeit am Institut für Schicht- und Ionentechnik am Forschungszentrum Jülich GmbH,
available from Technische Informationsbibliothek, Hannover.

Developing Models

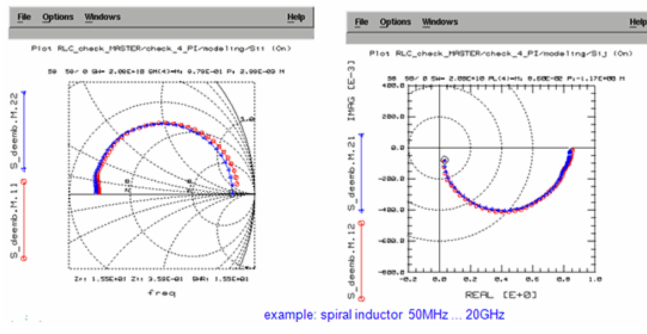
This topic explains developing models for Passive On-Wafer Components & Small Signal Behavior of Transistors.

Introduction

The RF modeling of passive components is commonly based on a given equivalent schematic. However, this may lead to frequency-dependent lumped components. The following procedure proposes a new method which allows to develop both, the schematic and the corresponding model parameters.
Step1. First of all, we need special dummy structures for accurate de-embedding, especially for spiral inductors



It is recommended to have OPEN and THRU dummies available,and to model them first. With the modeling, we check if there
are really no series components hidden in the OPEN dummy, no parallel components in the THRU, and if we can de-embed
the DUT from the THRU by ABC matrix de-embedding.
Step2. After we have verified the dummy device performance,the device is carefully de-embedded:

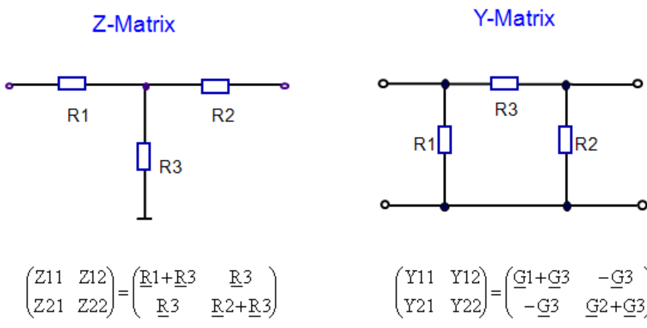


example: spiral inductor 50MHz ... 20GHz

> **Note**
> it cannot be overemphasized enough, that an accurate, verified de-embedding is the prerequisite for accurate passive device models!

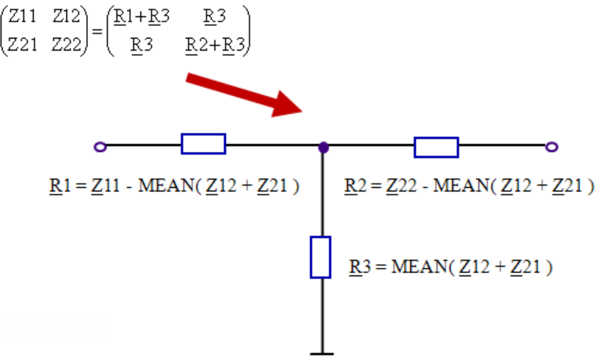
Developing the Equivalent Schematic

Important schematics for free RF modeling

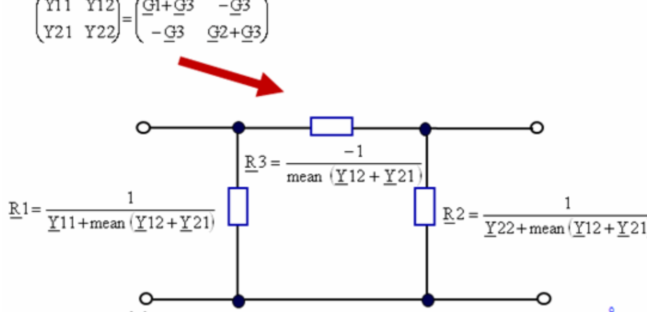


$$\begin{pmatrix} Z11 & Z12 \\ Z21 & Z22 \end{pmatrix} = \begin{pmatrix} \underline{R}1+\underline{R}3 & \underline{R}3 \\ \underline{R}3 & \underline{R}2+\underline{R}3 \end{pmatrix} \qquad \begin{pmatrix} Y11 & Y12 \\ Y21 & Y22 \end{pmatrix} = \begin{pmatrix} \underline{G}1+\underline{G}3 & -\underline{G}3 \\ -\underline{G}3 & \underline{G}2+\underline{G}3 \end{pmatrix}$$

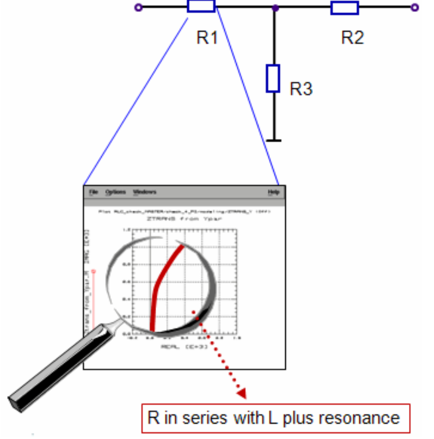
Measurement -> De-embedding -> Z-Matrix



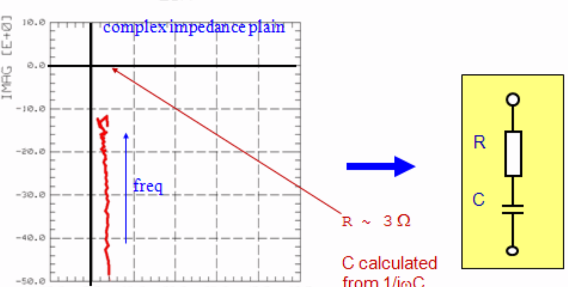
Measurement -> De-embedding -> Y-Matrix

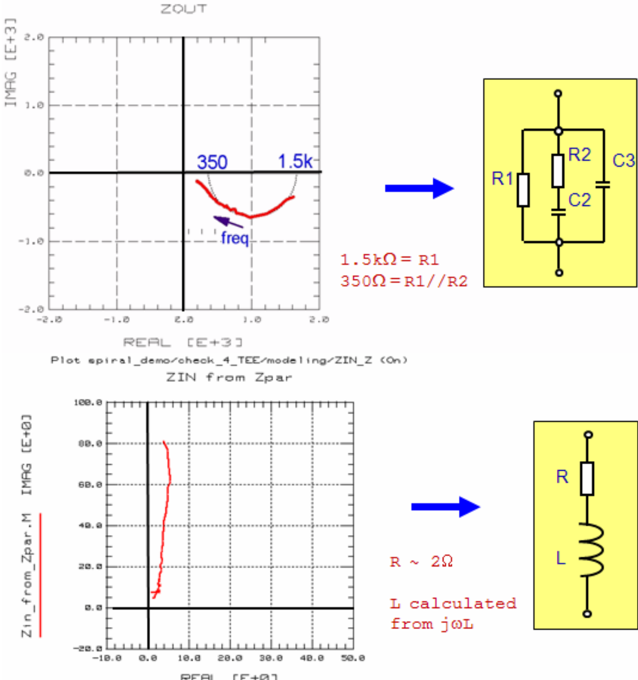


To develop the right subcircuit behind the matrix components
From the de-embedded S-parameters, we calculate the complex impedances of the branches of the Y or Z matrix, and analyze
their locus curve on the complex impedance plane.



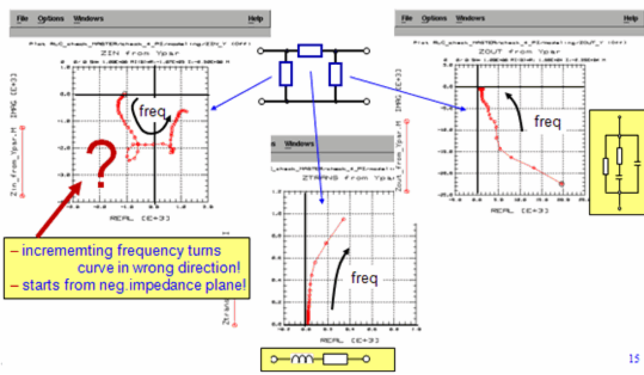
Locus curves of typical impedances



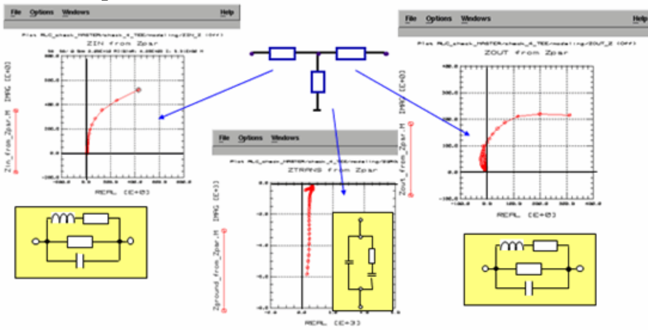


Selecting a PI or a TEE schematic
The de-embedded S-parameters are first converted to e.g. Y-parameters, and the locus of the individual complex impedance of each PI structure branch is roughly analyzed and checked if the locus curve has a physical meaning.
Also, a rough schematic of that branch is developed. Then, the same is done for the opposite matrix, e.g. Z-matrix.

Spiral Inductor
Schematic developed from the locus curves based on a PI structure

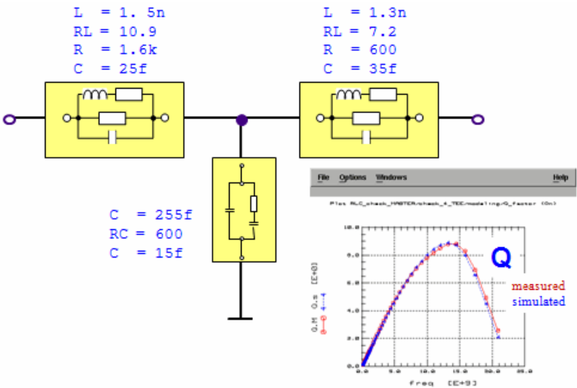


Spiral Inductor
Checked against the locus curve of a TEE structure

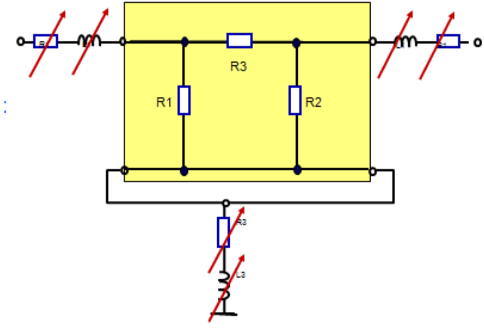


The TEE structure seems to represent the actual device more physically than the PI structure.
In the next step, we do the fine-work and develop accurately the branch schematic and fine-tune the model parameters.

TEE structure schematic selected



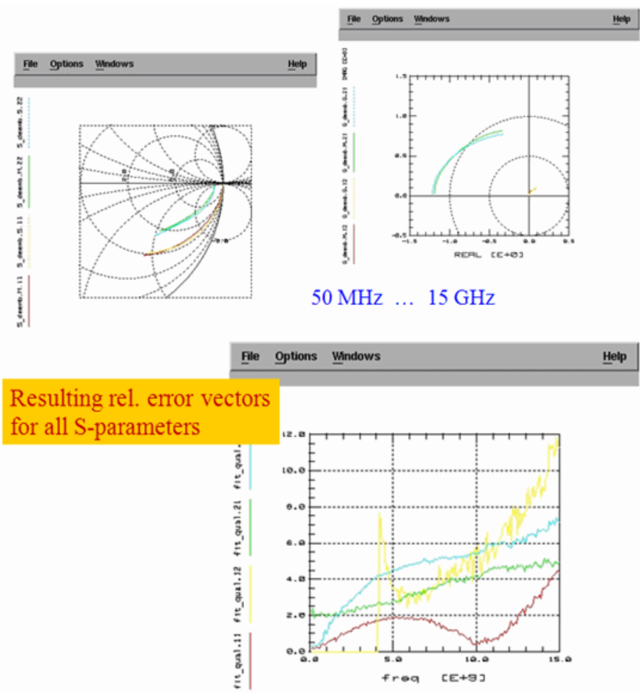
If a TEE or a PI alone does not make it then the locus curve inspection method can also be applied to overlying PI and TEE structures. For example, In the case of an assumed PI, overlaid by a TEE, we set the TEE inductors and resistors first to 0, and then 'tune their values in', until the locus curves of the inner PI components correspond to physical locus curves. Then, the so found TEE components are stripped-off, and the remaining inner PI is again modeled with the method given before.



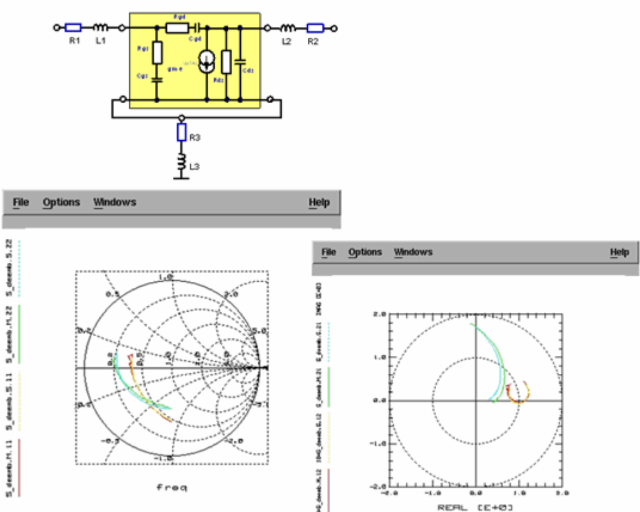
Application of this method to other RF modeling problems
Besides spiral inductors, it can be applied to:

- MIM capacitors,
- varactor diodes
- resistors
- transistors small signal model, quasistatic and non-quasistatic

Small Signal Modeling Result



Modeling result for a SOI transistor up to 50GHz



Summary

A method has been proposed:

- To adaptively develop an equivalent schematic for passive components
- To precisely extract their parameter values

This method:

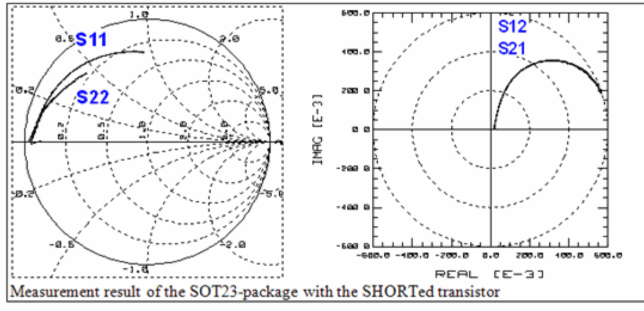
- This method has been implemented into IC-CAP and shows a very good match between measured and simulated data.
- It can be combined with data management for parameter scaling

Modeling of a SOT-23 Package

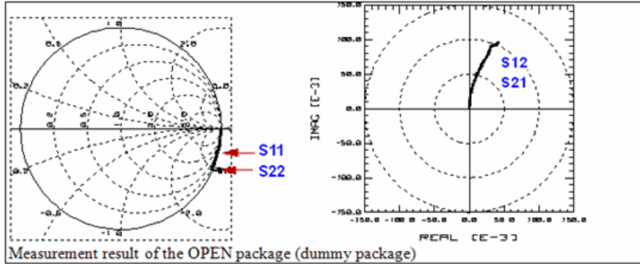
IC-CAP file: demo_features/.../.../sot23.mdl

The SOT23 package

For the modeling of packages, two types of measurements are required: a SHORTed package and an OPEN package. In most cases, an OPEN or 'dummy' package is available. It will be used for the evaluation of the package capacitances. For the inductors and resistors, the SHORTed package will be used. If it is not available, a 'cold' measurement of the transistor may help. In this case, a strong current is applied to port1, while half of that current is sunk out of port2. With such a bias, the transistor behaves like a good SHORT, and we are able to measure the parasitic inductances and resistors. The following figures show the results of these two measurements. The frequency range was from 30kHZ to 3GHz (HP8753D).



Measurement result of the SOT23-package with the SHORTed transistor



Measurement result of the OPEN package (dummy package)

Since we assume an equivalent schematic with LC structures, we start with the modeling of the SHORT measurements (inductors and resistors in series with the ports). As a first attempt, the figure below sketches the starting point for our schematic.

> **Note**
> The resistors Rcoldxy represent the behavior of the shorted transistor in the 'cold' measurement mode.

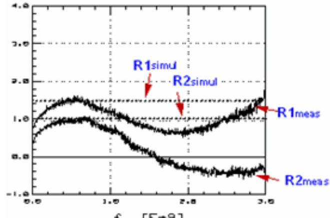
Beginning the modeling with the lowest frequencies, the resistors R1 and R2 model the ohmic losses of the connection between the test fixture and the package and also in the package itself. They represent the 'non-zero Ohm' starting points of the Sxx curves. On the other hand, Sxy is not '0' for low frequencies. Therefore, we have to model some power crosstalk from Port1 to Port2, what is reflected by resistor R3. In order to evaluate the correct values of these three resistors, the S-parameters are converted to Z, and the real part of it is displayed against the frequency. The PEL program that performs that transformation (and also for the imaginary parts for the modeling of the inductors later) is printed out.

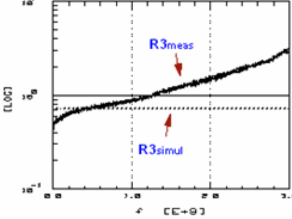
```
            --------|__|---------
            |       Rcold13      |
1     ___   10   L1 |11   1mOhm      21|   L2   20    ___     2
o---|___|---o---\/\/---o                o---/\/\---o---|___|---o
    R1                 |                |              R2
                       -                -
                      | |              | |
                      | |Rcold12  Rcold23| |
                       -  1mOhm    1mOhm -
                       |                |
                       ----------o-----------
                               31|                 |
                                 --/\/\---o---|___|---|
                                   L3     30   R3     |
```

Schematic for first attempt to model the SHORTed SOT-23 package.

```
!PEL program in IC-CAP to calculate the resistors and inductors out of the S-parameters of the
'cold' measurement.
!PRINT "running transform 'Z_to_RLx' ..
tmpx=freq !import stimulus data
tmpy=TwoPort(S,"S","Z") !import measured/simulated data
i=0
WHILE i < SIZE(tmpy)
tmpxyM=(tmpy.M.12[i]+tmpy.M.21[i])//2 ! calculate the
tmpxyS=(tmpy.S.12[i]+tmpy.S.21[i])//2 ! values of the
tmpy.M.11[i]=(tmpy.M.11[i]-tmpxyM) ! underlying
tmpy.S.11[i]=(tmpy.S.11[i]-tmpxyS) ! TEE structure devices
tmpy.M.22[i]=(tmpy.M.22[i]-tmpxyM) ! Rx and j*2*PI*Lx
```

```
tmpy.S.22[i]=(tmpy.S.22[i]-tmpxyS)
!make the imag.parts contain the inductors only
tmpy.M.11[i]=REAL(tmpy.M.11[i])+j*IMAG(tmpy.M.11[i])*(2PI*tmpx[i])^-1
tmpy.S.11[i]=REAL(tmpy.S.11[i])+j*IMAG(tmpy.S.11[i])*(2PI*tmpx[i])^-1
tmpy.M.12[i]=REAL(tmpy.M.12[i])+j*IMAG(tmpy.M.12[i])*(2PI*tmpx[i])^-1
tmpy.S.12[i]=REAL(tmpy.S.12[i])+j*IMAG(tmpy.S.12[i])*(2PI*tmpx[i])^-1
tmpy.M.21[i]=REAL(tmpy.M.21[i])+j*IMAG(tmpy.M.21[i])*(2PI*tmpx[i])^-1
tmpy.S.21[i]=REAL(tmpy.S.21[i])+j*IMAG(tmpy.S.21[i])*(2PI*tmpx[i])^-1
tmpy.M.22[i]=REAL(tmpy.M.22[i])+j*IMAG(tmpy.M.22[i])*(2PI*tmpx[i])^-1
tmpy.S.22[i]=REAL(tmpy.S.22[i])+j*IMAG(tmpy.S.22[i])*(2PI*tmpx[i])^-1
i = i + 1
END WHILE
RETURN tmpy
```

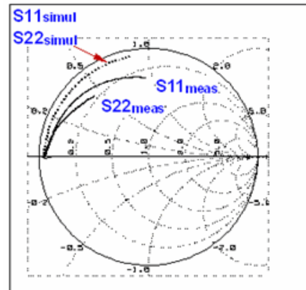


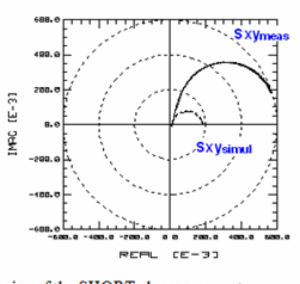


After transforming the S-parameters to Z and calculating Zxx-Zxy, REAL(Zxx) represent R1 and R2, while REAL(Zxy) represents R3.

Ideally, the resistors should show-up with frequency-independent values. In reality, as can be seen above, this is not the case due to the fact that the supposed schematic (a simple TEE structure) is too simple. Therefore, it would fit the S-parameter only if its resistors would have exactly this frequency dependence!

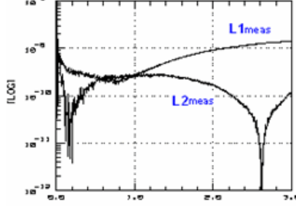
Nevertheless, we want to stay with 'ordinary' frequency independent resistors. So we extract their values from a frequency range where we are well above the lower frequency limitations of the network analyzer and well below 2nd order effects at higher frequencies. In our example, we use a maximum flat region of the curves and therefore the frequency range between 500MHz and 700MHz. The simulation result is given in figure below. We see a nice fitting for low frequencies, as expected. In the next step, we will discuss the fitting for higher frequencies and the determination of the inductor values.

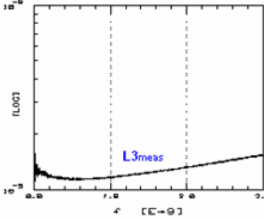




The first attempt to model the low-frequency behavior of the SHORTed measurement

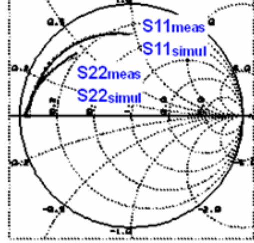
A second result from the little program above are the values of the inductors. They are shown in the following figure. Unfortunately, these plots give no clear picture about where to extract the values of Lx (there are no flat regions visible). Therefore, the extraction is done manually, with respect to the S-parameters, but using the values of the plots in figure below (for low frequencies) as a first guess.

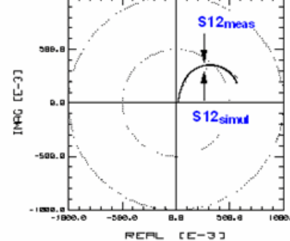




The values of the inductances Lx obtained from the S-to-Y transformation

After all, an optimizer run fine-tunes the values of the inductors Lx with respect to the S-parameters. The result is given in the figure below:





The final result of the modeling of the 'cold' or SHORTed SOT-23 package

The next step is to determine the values of the package capacitors. They show up with the 'open' (dummy package) measurement. Yet, that 'open' measurement is overlaid by the already known series inductors and resistors. Therefore, the measured S-parameters have

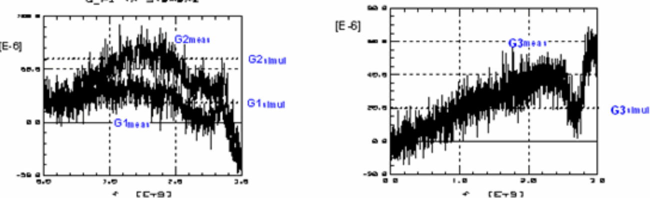
to be de-embedded from these parasitic components first. This is done with a subtraction of the corresponding Z-matrices.
We further assume a PI structure for the parasitic capacitors, and therefore, we have to convert the resulting Z-parameters to Y-parameters. See the little program below which contains printout of the de-embedding program that strips-off the dummy device data from the already known parasitic inductors and resistors.

```
PRINT "running xform Ydeem ..."
            ! Z_open         -             Z_short
tmp = TwoPort(S_open,"S","Z") - TwoPort(S_short, "S","Z")
            ! transform to Y parameters
tmp = TwoPort(tmp,"Z","Y")
RETURN tmp
```

As already mentioned, the resulting Y matrix is then interpreted with respect to a PI structure. This is done in transform 'Y_to_GCx', see the following code sample which contains printout of the transformation of the Y-matrix elements into the capacitors and parallel conductors:

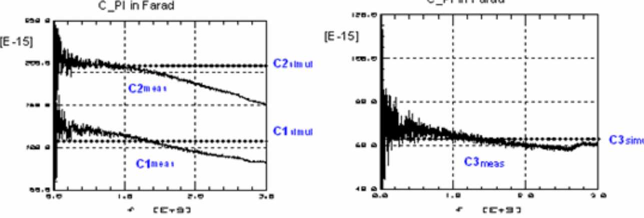
```
tmpx=freq !import stimulus data
tmpy=Ydeem !import measured/simulated data
i=0
WHILE i < SIZE(tmpy)
tmpxyM=(tmpy.M.12[i]+tmpy.M.21[i])//2 ! calculate the
tmpxyS=(tmpy.S.12[i]+tmpy.S.21[i])//2 ! values of the
tmpy.M.11[i]=(tmpy.M.11[i]+tmpxyM) ! underlying
tmpy.S.11[i]=(tmpy.S.11[i]+tmpxyS) ! TEE structure devices
tmpy.M.22[i]=(tmpy.M.22[i]+tmpxyM) ! Rx and j*2*PI*Lx
tmpy.S.22[i]=(tmpy.S.22[i]+tmpxyS)
!change sign of Yxy
tmpy.M.12[i]=-tmpy.M.12[i]
tmpy.S.12[i]=-tmpy.S.12[i]
tmpy.M.21[i]=-tmpy.M.21[i]
tmpy.S.21[i]=-tmpy.S.21[i]
!make the imag.parts contain the inductors only
tmpy.M.11[i]=REAL(tmpy.M.11[i])+j*IMAG(tmpy.M.11[i])//(2PI*tmpx[i])
tmpy.S.11[i]=REAL(tmpy.S.11[i])+j*IMAG(tmpy.S.11[i])//(2PI*tmpx[i])
tmpy.M.12[i]=REAL(tmpy.M.12[i])+j*IMAG(tmpy.M.12[i])//(2PI*tmpx[i])
tmpy.S.12[i]=REAL(tmpy.S.12[i])+j*IMAG(tmpy.S.12[i])//(2PI*tmpx[i])
tmpy.M.21[i]=REAL(tmpy.M.21[i])+j*IMAG(tmpy.M.21[i])//(2PI*tmpx[i])
tmpy.S.21[i]=REAL(tmpy.S.21[i])+j*IMAG(tmpy.S.21[i])//(2PI*tmpx[i])
tmpy.M.22[i]=REAL(tmpy.M.22[i])+j*IMAG(tmpy.M.22[i])//(2PI*tmpx[i])
tmpy.S.22[i]=REAL(tmpy.S.22[i])+j*IMAG(tmpy.S.22[i])//(2PI*tmpx[i])
i = i + 1
END WHILE
RETURN tmpy
```

The REAL parts of the returned array of the program shown above represent the conductors that are in parallel with the capacitors of the PI structure. This is depicted in following figure. Also plotted are the extracted conductor values. Again, they represent the mean value of a maximum flat region of these curves.



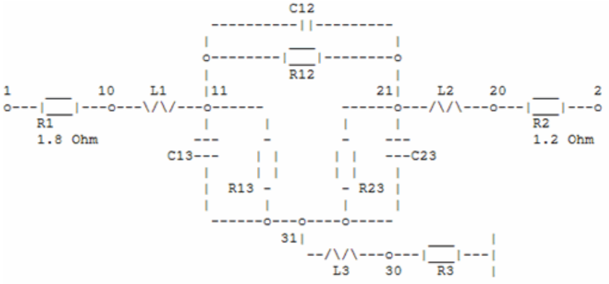
Measured and extracted conductor values for the PI schematic of the 'open' or dummy device measurement

On the other hand, the IMAG parts of the program in above figure represent the capacitors. This is displayed in the figure below. As we end-up with rather flat plots, it seems that our assumption of a PI structure is quite correct. Therefore, the extraction of the capacitor values is pretty straight-forward, and simply the mean value of a maximum flat range of that plots. The resulting values are shown in the plots also.



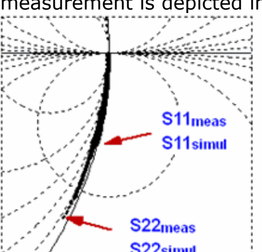
Measured and extracted capacitor values for the PI schematic of the 'open' or dummy device measurement

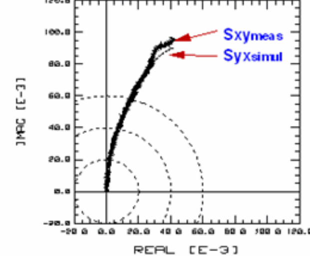
This leads finally to the total equivalent schematic, as shown in the figure below:



The total equivalent schematic for the SOT-23 package as the overlay of the schematics of the shorted and open measurement modeling.

The fitting of the final equivalent schematic of above figure to the SOT-23 OPEN measurement is depicted in the figure below:

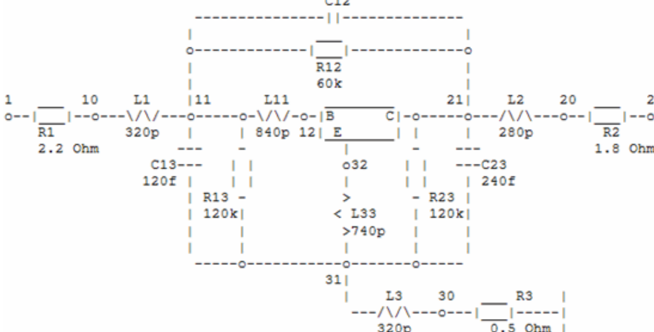




The curve fitting of the 'open' or dummy device modeling with the final equivalent schematic of the SOT-23 package.

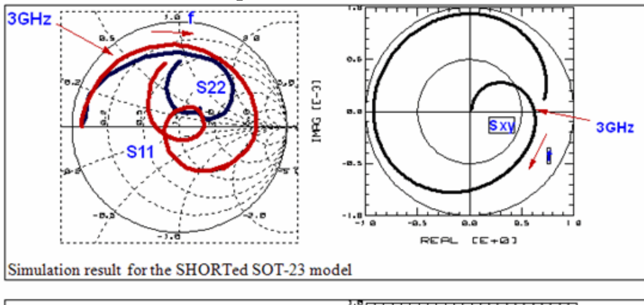
Another model for the SOT-23 package is given in the figure below. We want to check its performance against the model from above, which was valid up to 3GHz.

This model represents more the actual physical facts in the package: The transistor is fixed with its Collector to the package, while Emitter and Base are contacted by bond wires. Therefore, the nodes 11 and 31 are now 'inside' the inductors L1 and L3:

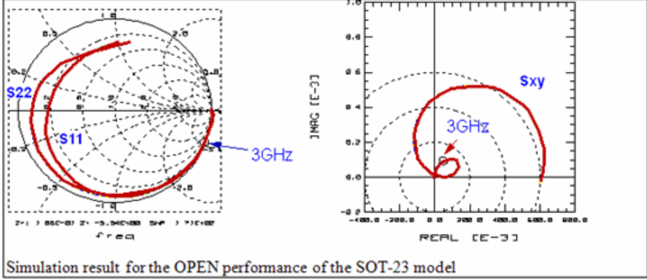


The performance of the model of the above figure was simulated up to 26GHz.

The goal is to compare it with the model from above for the frequency range up to 3GHz, see the markers in the figures below:



Simulation result for the SHORTed SOT-23 model



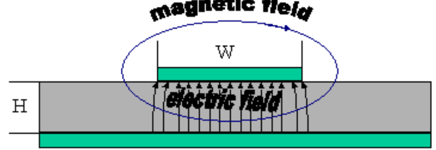
Simulation result for the OPEN performance of the SOT-23 model

Result: The developed (simpler) schematic is very much applicable for frequencies up to 3GHz. The additional parasitic components are required for the performance above 5GHz.

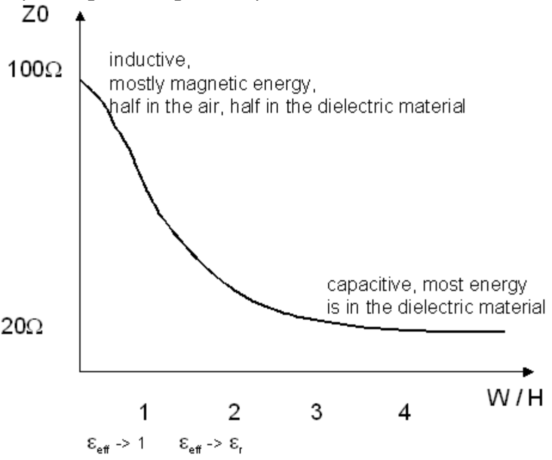
S-Parameters of Strip Lines

Strip Lines Tutorial

A strip line exhibits a TEM field, i.e. a transverse electromagnetic mode field, where the dielectric field is perpendicular to the magnetic field, and both are perpendicular to the direction of wave propagation.



From this sketch, we can conclude the following basic properties of a microstrip line, depending on the geometry factor W/L:



The characteristic impedance, Z0, ranges from about 20Ω to about 100Ω. The limit of 100Ω exists for a very simple reason: the width is much less than the height, and such a structure cannot be manufactured (under-etching etc).

This sketch allows to make some fundamental considerations:
As a matter of fact, a small microstrip line exhibits less capacitance than a wide one. Inspecting the plot, this concludes that a lower capacitance in a microstrip line comes along with a lower impedance Z0.

This is obvious, when recalling that

$$Z \sim \sqrt{\frac{L}{C}}$$

Referring to crosstalk between lines, we can learn from the sketch above that a low impedance microstrip line is capacitive. I.e. the energy is rather between the metal conductor and the ground. I.e. two low impedance striplines side-by-side, will exhibit less cross-talk than two high impedance striplines. By the way, this is a key design rule for packages and connectors.

As another important outcome, a shielding across a microstrip line will lead to the fact that the impedance of the resulting strip line will be lower, because more of the electro-magnetic field will now be present in the enlarged electric field consisting of the previous field in the dielectric layer plus the additional space between the active metal layer and the top cover. Therefore, a cover across a microstrip line reduces the resulting impedance, and, thus, reduces cross-talk between adjacent striplines.

To further reduce cross-talk of adjacent lines, i.e. to reduce the impedance of each line (increase the electric field, i.e. make the lines more capacitive), reduce the height of the dielectric material.

However, on the other hand, a cover 'kills' the performance of filters designed from strip lines based on electric field coupling!

Modeling a lossy strip line from S-parameters

Tutorial

The S-parameters of a strip line of length L and a characteristic impedance Z0, measured with a network analyzer, can be considered like a transition of a 50Ω system into a Z0 system with a long delay.

Therefore, the Sxx parameters in a Smith chart start in the center, i.e. at 50Ω, and turn then to a curve around the value of Z0 of the strip line. For a lossless line, this curve is represented by circles around the line's Z0. For a lossy line, this looks like a looping towards Z0. Like with all S-parameters, this turning is always clock-wise.

- For a Z0 < 50Ω, the curve starts at 50Ω, and turns with -90° clockwise, i.e. downwards. If the line is lossless, we have a circling curve centered around Z0, and touching again and again the center of the Smith chart, i.e. 50ΩFor a lossy line, we have a looping around Z0 with the end point at Z0.
- For the special case of a line with a characteristic impedance of Z0=50Ω, lossless or lossy, we have all S11 and S22 parameters in the center of the Smith chart: at 50Ω.
- For a Z0 > 50Ω, the curve starts at 50Ω, and turns with +90° straight upwards. If the line is lossless, we have again a circling curve centered around the line's Z0, and touching the center of the Smith chart, i.e. 50Ω again and again (if it is long enough!)? If the curve is lossy, we will again have a looping towards the end point Z0.

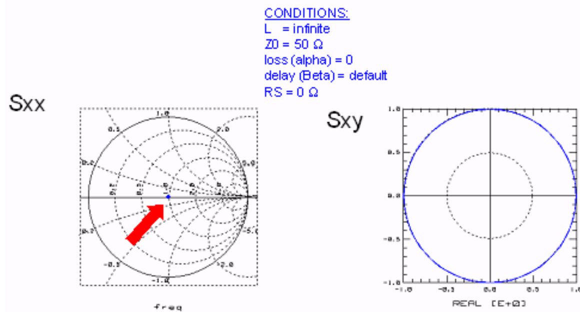
The Sxy parameters in the polar plot start at '+1', and also turn clock-wise.

- For a lossless line of Z0=50Ω, we have circles with magnitude '1' . For a lossy 50Ω line, the circles are replaced by a looping towards '0', because for a long, lossy line, there is no signal reflected back any more.
- If the Z0 of the line is <>50Ω, we have a change in magnitude (see further below in the next chapter), which is represented in the Sxy plot as ellipses for a lossless line, or as an elliptic looping towards '0' for a lossy line.

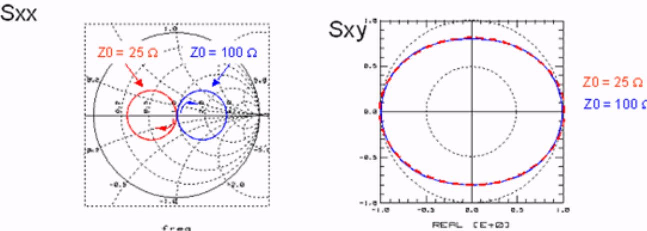
For the modeling

we commence with setting all line parameters to default, and the length to infinite.
The loss parameter is set to '0'.
In this case, we can determine Z0 from fitting the magnitude of the simulated curve to the trace of the measured data in the Smith chart. We ignore the phase.

For Z0 = 50Ω we have this situation:
Step1. Set length to infinite and adjust Z0 in smith chart

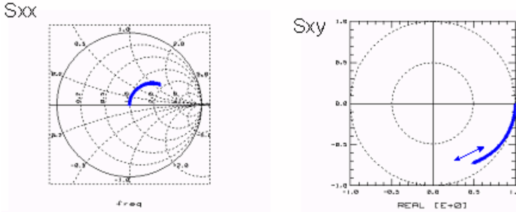


And for Z0 less or bigger than 50Ω we have:

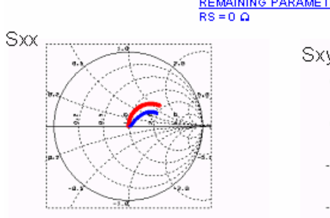


> **Note**
> For Sxx, all curves start for freq=0 in the center of the smith chart i.e. at 50Ω. Their end point at infinite length is their Z0!

After having determined Z0, we set the physical length of the strip line to its real, physical value and adjust the strip line's delay by fitting Sxy:
Step2. Set length to its physical value and adjust delay (beta) in Sxy

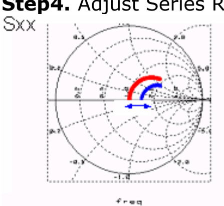


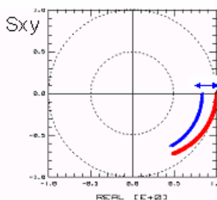
Then, we adjust the loss, again by fitting Sxy:
Step3. Adjust loss (alpha) in Sxy



If the starting point of Sxx is not at 50Ω, and the starting point of the Sxy is not at '1', we need to consider a series resistor (contact resistance) in series with the delay line model. This is, last not least, the last remaining parameter of typical strip lines on the wafer:

Step4. Adjust Series Resistor Rs





> **Note**
> If a shift to the left is required in Sxx, split the delay line into 2 and insert a resistor to ground between them!

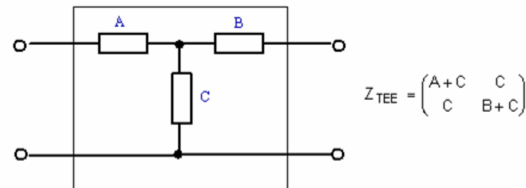
Modeling a Strip Line Transition

When the width of a strip line changes, current crowding occurs. This is modeled by the L, while the change in the step contour is modeled by the capacitance C.

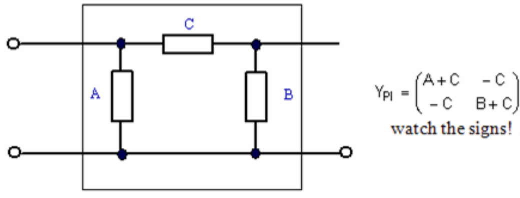
Testing S-parameters For Typical Passive Circuits

IC_CAP file: pass_RLC_TEE_or_PI.mdl

This chapter is intended to present some matrix manipulations for the interpretation of S-parameters of typical passive LRC networks. This is done by transforming the S-parameters to Y or Z-parameters. For these types of matrices exist simple interpretation capabilities, provided their underlying circuits have a structure like in the following figure.



a TEE circuit interpreted as a Z matrix.



a PI circuit interpreted as a Y matrix.

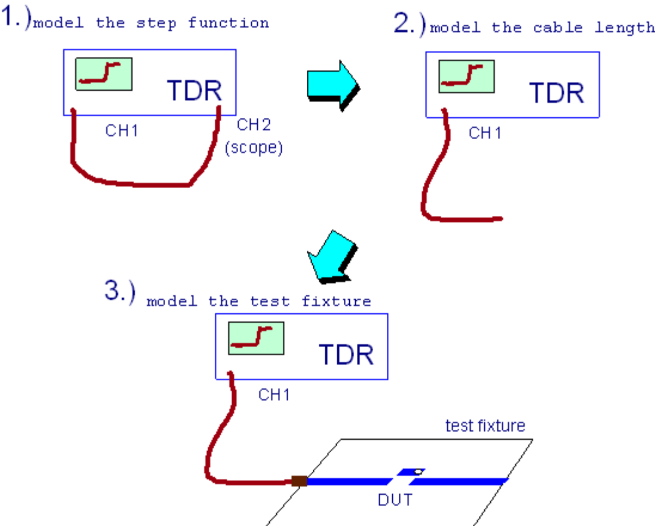
Test Fixture Modeling in Time Domain

When measuring a SMT (surface mount technology) or packaged device, the test fixture contributes to the total measurement. Sometimes it is not possible to use dummy devices like SHORT, OPEN, THRU with known cal-kit data.

Hint
a NWA calibration will only work correctly, if the 'non-idealities' of the cal-set components are known. This is the so-called cal-kit data. This data has to be entered into the NWA before the instrument is calibrated and the calibration result is stored in a cal-set of the NWA.

A nice work-around is to use a TDR (time domain reflectometer) in order to characterize the effects of the test fixture in the time domain and to perform a de-embedding later in the frequency domain. What has been a full 2-port calibration for a NWA, is now:

- first the modeling of the TDR's step function in a THROUGH measurement,
- second the modeling of the TDR cable length, an OPEN measurement.
  and finally the modeling of the structure on the test fixture itself.



WHAT CAN BE MODELED USING A TDR?

- Resistors: the back-reflected curve changes its level.
- Inductors: the back-reflected curve overshoots.
- Capacitors: the back-reflected curve undershoots.
- Delay lines: the separation in time between reflections.

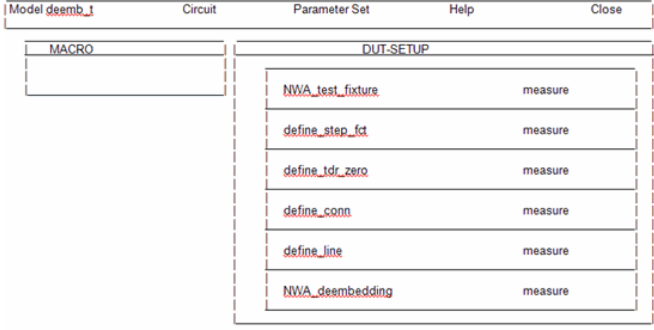
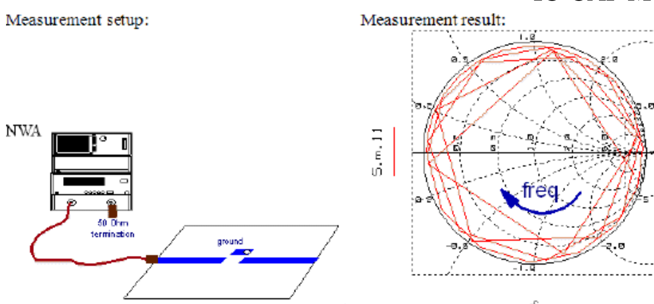
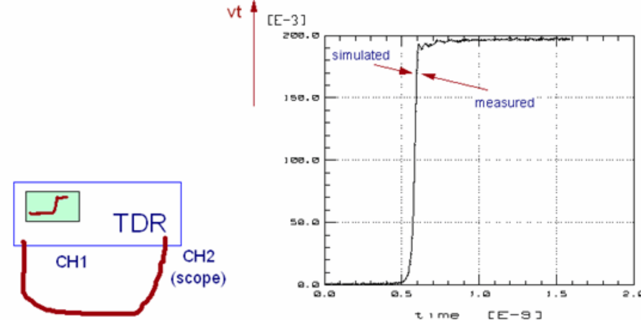


Fig.2: the IC-CAP structure: measure the test fixture with the NWA
model the TDR step response
model the 'zero' of the TDR
model the test fixture connector
model the strip line on the test fixture
perform a NWA measurement of the DUT and de-embed it from the parasitic effects of the test fixture.

The starting point is again the S-parameter measurement of the parasitic structure alone, see fig.3. As our test fixture contains a strip line, we have a lot of phase shift. Our task is to 'turn back' this phase shift and to obtain the S-parameters of the open. To keep things simple, we only consider S11.



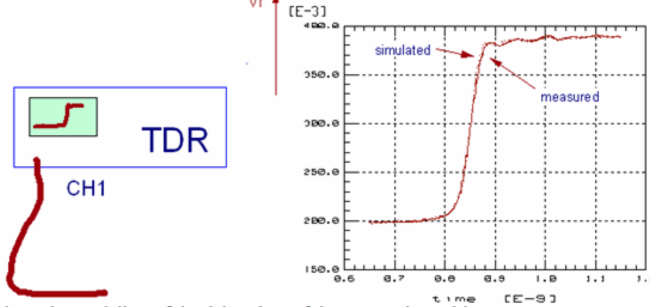
Setup 'define_step_fct' of the model file in fig.2 is used to model the TDR step function. Therefore, we connect the channel CH1 of the TDR (the very TDR channel) with an TDR scope input (e.g. CH2), see fig.1. The measurement result is given below:



The model corresponding to the measurement of fig. 4 is defined in the test circuit of this setup: several SPICE pulse functions are used to model the step function, plus a delay line. It looks like this:

```
.subckt mdltdr 101=tinp 105=toutp
* a series of step functions in the time*
* domain                               *
*****************************************
v0   1  0    PWL(289p 1.375m  0.40404n 0)
v1   2  1  PULSE(0  1.5625m      0.4n  4.04p  4.04p 0)
v2   3  2  PULSE(0 -0.1875m 0.40404n  4.04p  4.04p 0)
v3   4  3  PULSE(0  0.375m  0.40808n  4.04p  4.04p 0)
...
...
v99 101 99 PULSE(0 -0.625m  0.79596n  4.04p  4.04p 0)
ri   101  100   50
t1 100 0 105 0 TD=1n Z0=50
ra   105    0   50
.ends
```

The delay time t1.TD=1n reflects the length of the actual connection cable.
In the next step, we disconnect the cable from CH2, and use another cable which we will use later to connect CH1 with the test fixture. Then, we measure the back-reflection on CH1 of the open cable. Setup 'define_tdr_zero' holds the result, what is depicted in fig.4:

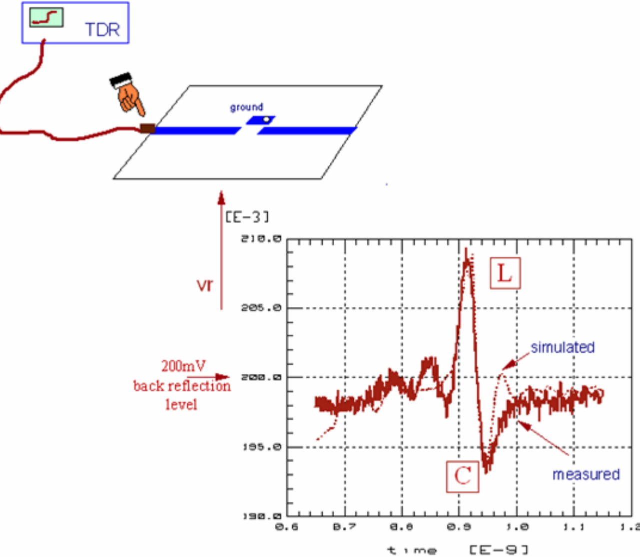


The SPICE circuit of this setup is the one from above, but now enhanced with the correct cable delay time and components to model the open end.

```
.subckt mdltdr 101=tinp 105=toutp
v0   1  0    PWL(289p 1.375m  0.40404n 0)
...
...
v99 100 99 PULSE(0 -0.625m  0.79596n  4.04p  4.04p 0)
ri 100  101   50
r1 101  102   2
t1 102  0   103    0   TD=150p Z0=50
ra 103    0  43k
c1 103    0  1f
.ends
```

After these two pre-measurements, we are ready to go after the modeling of the test fixture itself.

So, the cable is connected with the test fixture. As we model 'from the left to the right', we first concentrate on the modeling of the connector of the test fixture. Setup 'define_conn' holds the result, what is given in fig. 6.



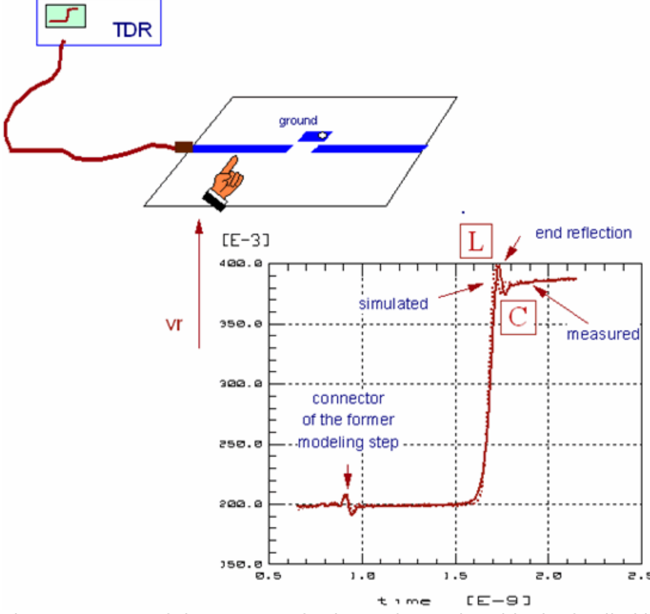
And again, the test circuit of this setup is enhanced to fit the measured data:

```
.subckt mdltdr 101=tinp 105=toutp
v0   1  0    PWL(289p 1.375m  0.40404n 0)
...
...
v99 100 99 PULSE(0 -0.625m  0.79596n  4.04p  4.04p 0)
ri    100  101   50
r1    101  102   1.1
t1    102   0   103  0   TD=188p Z0=50
ls1   103  104  460p
cp1   104   0   140f
r2    104  105  .01
tmic1 105   0   106   0 TD=500p Z0=50
rend  106   0   1G
.ends
```

Fig.6 also gives the simulation result: the inductor models the overshoot, and the capacitor the undershoot.

Delay time t1.TD reflects now the delay of the cable (150ps) plus the connector (33ps). Delay line tmic1.TD (for the modeling of the delay time of the strip line on the test fixture) is set to 500ps and will be modeled in detail in the next step.

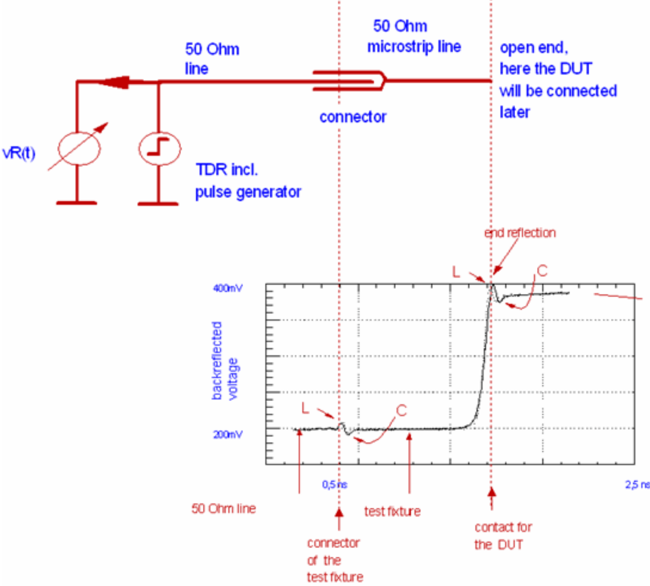
Finally, we consider the strip line and its open end on the test fixture. The TDR measurement is kept in setup 'define_line' and is given in fig 7:



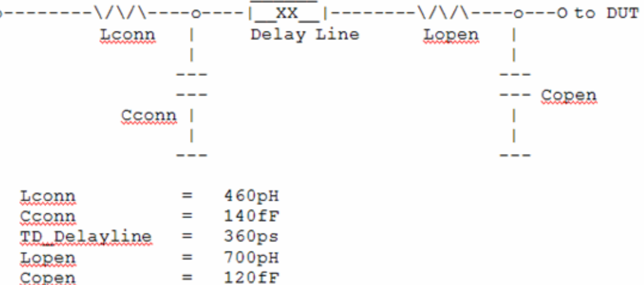
The test circuit of this setup is further enhanced and looks finally like:

```
.subckt mdltdr 101=tinp 105=toutp
v0   1  0    PWL(289p 1.375m  0.40404n 0)
...
...
v99 100 99 PULSE(0 -0.625m  0.79596n  4.04p  4.04p 0)
ri    100  101  50
r1    101  102  1.1
t1    102   0   103   0 TD=188p Z0=50
ls1   103  104  460p
cp1   104   0   140f
r2    104  105  .01
tmic1 105   0   106   0 TD=366p Z0=50
lend  106  107  700p
cend  107   0   120f
rend  107   0   4.5k
.ends
```

After all, we have the following model:



The final SPICE schematic is:



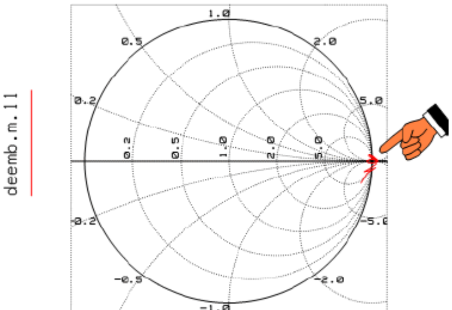
Now, we are done and can re-connect our test fixture with the NWA to perform a measurement. With the S-parameters, we perform a de-embedding using the equivalent schematic of the test fixture as developed in the time domain.

Therefore, the de-embedding follows this scheme:

```
transform S_to_Z:   de-embed   r1
transform Z_to_S:   de-embed the delay of the connector (33ps)
transform S_to_Z:   de-embed   ls1
transform Z_to_Y:   de-embed   cp1
transform Y_to_Z:   de-embed   r2
transform Z_to_S:   de-embed the delay of the strip line tmic1 (366ps)
transform S_to_Z:   de-embed   lend
transform Z_to_Y:   de-embed   cend
transform Y_to_S:   to obtain the de-embedded S-parameters
```

To check the quality of this procedure, we apply the de-embedding to the measurement of fig.3, i.e. the S11 parameters of the strip line at port 1 of the NWA.
Fig.10 shows the de-embedded result: the OPEN of the strip line on the test fixture.

frequency: 600MHz ... 6GHz



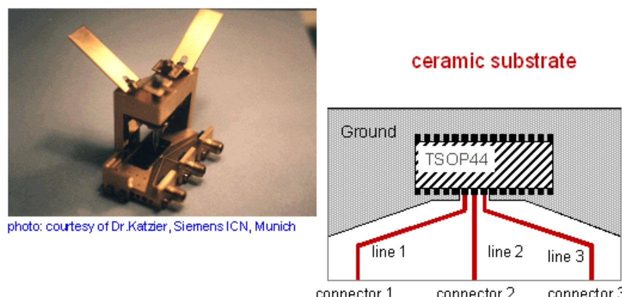
For frequencies up to 3GHz, the de-embedded curve represents an ideal OPEN, while for higher frequencies we might include in our SPICE schematic a special circuit the is able to reflect dispersion effects. See chapter 'Modeling the skin effect' in the appendix.

Modeling of a TSOP44 Packagae

A step-by-step tutorial, based on the IC-CAP model file IC-CAP file: tsop44.mdl

The task is to model a TSOP44 package, using the following test fixture:

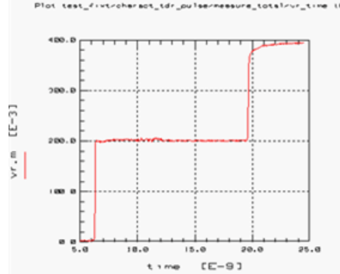
Test fixture for TSOP44



The test board, on the right side, shows that 3 pins of the TSOP44 will me modeled, while the other pins are connected to ground.

TDR-related issues

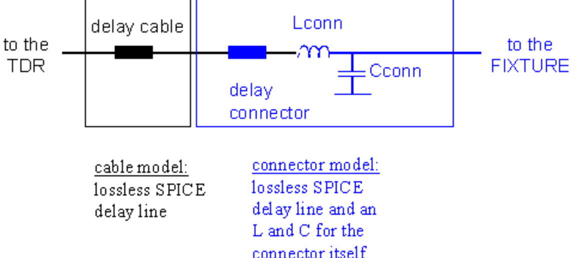
Before we can start measuring the package with the TDR, we have to model the step function of the TDR.



In this plot, we can see that the signal splitter in the TDR detects the emitted TDR step function after 6ns. Until roughly 19ns, we have 200mV, what corresponds to a lossless 50Ω delay line. At 19ns, we see an open end and some power radiation (the 'creeping' slope at the end reflection).
Note: This 'creeping' slope does not need to be modeled for our project, since the connector at the end of the cable will be connected later to the test fixture, i.e. it will no longer radiate a signal. However, for didactical reasons, Copen and Ropen are added to the SPICE circuit of the section 'charact_TDR_pulse', and removed in all other sections of the model file.

I.e. we can start with this subcircuit model:

Modeling the cable from the TDR to the test fixture



As a SPICE circuit for the TDR pulse, with default parameter values, we define:

```
.subckt tdr_pulse 95
*
Tcable     95 0   96 0 Z0=50 TD=500p
Tconn_half 96 0   97 0 Z0=50 TD=22p
Lconn      97 100  1f
Cconn     100   0  1a
Copen     100 102  1f
Ropen     102   0  1G
*
.ends
```

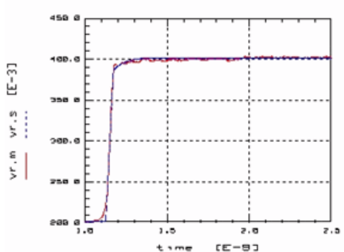
Note
22psec for the delay of half a 3.5mm connector is a typical value.

> **Note**
> For the data shown in this tutorial, i.e. the data in file TSOP44.mdl, Lconn and Cconn are not required due to the good performance of the connector.

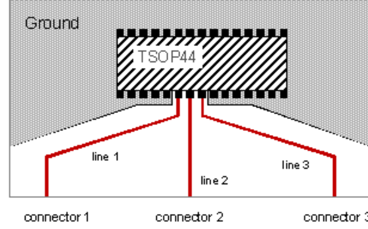
For our example, the extracted parameters are:
Tcable.Z0 = 50
Tcable.TD = 542p
Tconn_half = 22p
with a pulse risetime = 44p

> **Note**
> Copen = 75f and Ropen = 780Ω for the intermediate, didactical modeling of the open cable radiation.

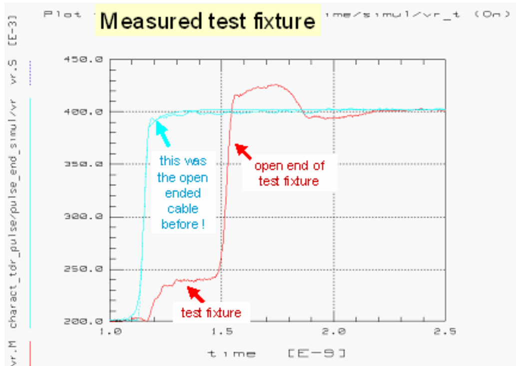
Here is the fitted result, referring to the cable end, i.e. the step from 200mV -> 400mV :



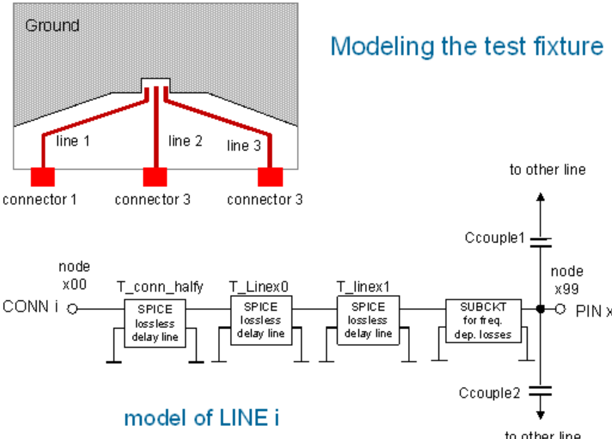
In the next step, we have to model the test fixture. It's layout is given below:



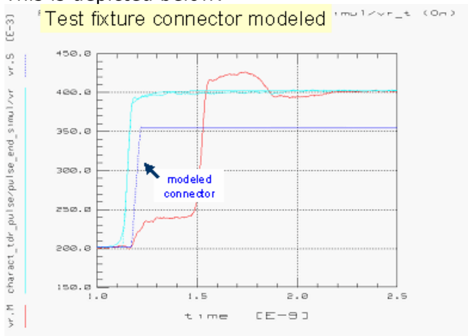
The TDR cable is connected to line 1 of the test fixture and we obtain the following measurement result:



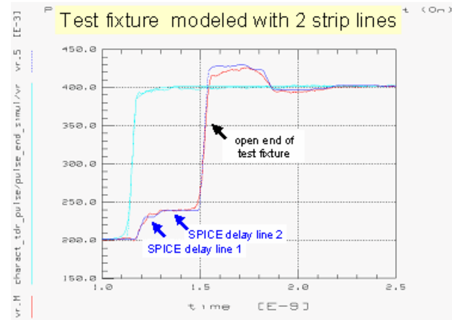
For the modeling, we will model all three lines, one after the other.
As can be seen below, the first SPICE model in this test fixture subcircuit is a delay line for modeling the half of the connector. For the strip line on the test fixture, since there is a knee in line 1, we will use two strip lines in series in the SPICE deck. Finally, a special SPICE circuit will be used to model frequency-dependent effects which are not covered by the ideal SPICE delay line model.



Let's commence with modeling the half of the connector, associated with the test fixture. This is depicted below:

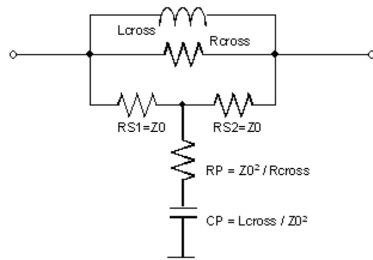


In the next step, we need to model the strip line on the test fixture. We will apply one ideal SPICE delay line before the knee on the layout, and another one after the knee. The modeled measurement is given in the next figure

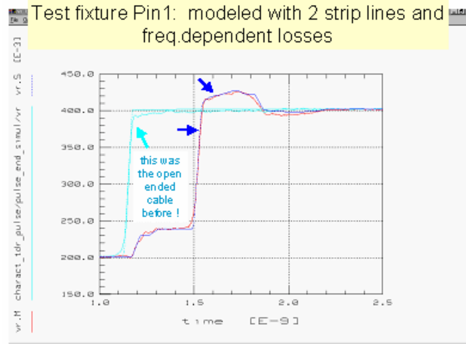


There is still a mismatch for the slope of the OPEN end as well as for the overshoot. This can be modeled with frequency-dependent losses of the strip lines. However, SPICE only offers lossless strip lines. However, using the following subcircuit /after Katzier/, we can also fit this region.

Transmission line section
with frequency dependent losses



Applied to modeling the test fixture line 1 in the time domain, we finally obtain the following modeled TDR performance:

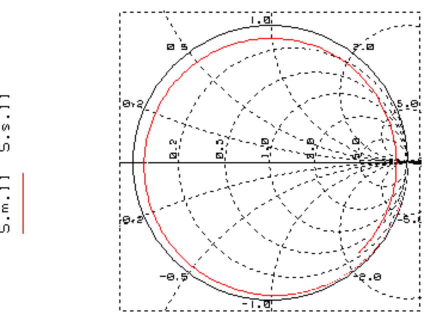


What is left is the modeling of the crosstalk between the lines on the test fixture. This is best measured using a network analyzer. Compared to TDRs (which are broadband measurement instruments), a NWA has a better resolution, and also allows an easy separation of crosstalk (Sxy) and main strip lines (~Sxx). On the other hand, using a TDR offers the advantage to have a clear picture of the geometric sequence of the SPICE elements, while in NWA plots, i.e. the Smich chart or the polar plot, we only see turning curves.

Therefore, before modeling the test fixture lines cross talk, we verify the fitting of the Sxx plot of the NWA. This is given in the next slide (Test Fixture Pin1: modeled with 2 strip

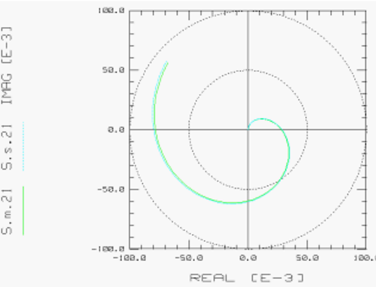
lines and freq. dependent losses)

Plot test_fixt/test_fixt_1_2_spar/spar/S11 (On)

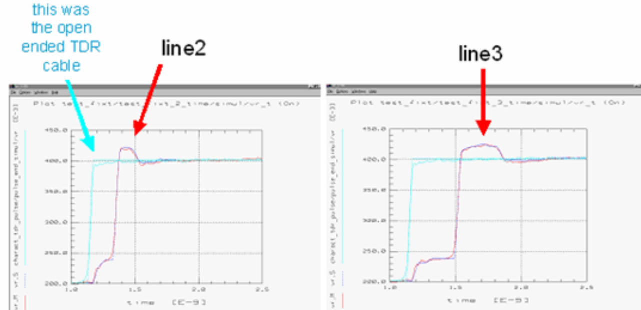


freq

We are now ready to model the crosstalk, referring to plot S21 (crosstalk of Pin1 (NWA Port1) to Pin2 (NWA Port2)):
Test Fixture Pin1: modeled crosstalk to line2



In the same way, lines 2 and 3 of the test fixture are modeled. Here the result:

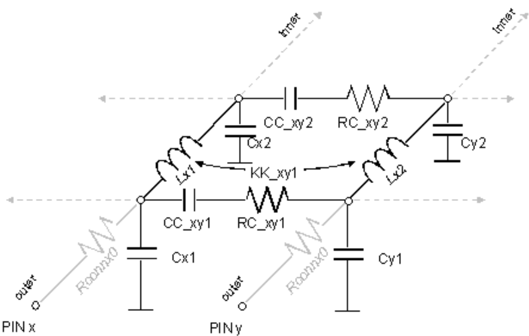


Please note that the test fixture modeling has only to be performed once for a given test fixture!
After the accurate modeling of all the lines of the test fixture, we are finally ready to cover the modeling of the TSOP44 package.

Since a package model, like a connector model, is represented by a big sub-circuit, it is essential to agree first on some standardized SPICE circuit node names.

The following figure depicts a segment of a package model:

TSOP44 - one section of a 2-pin model



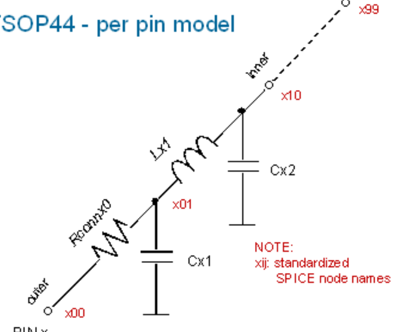
Such a segment is usually able to model a package performance up to ~500MHz. For higher frequencies, more segments have to be cascaded. For package models up to 3GHz, typically 3 segments are cascaded.

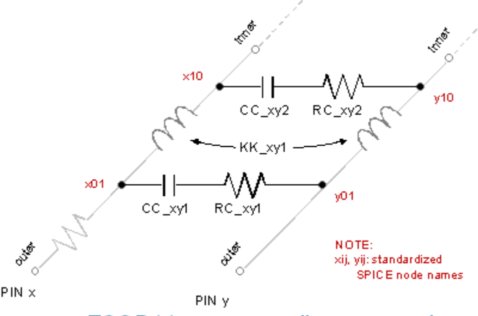
Standardization of SPICE node names:

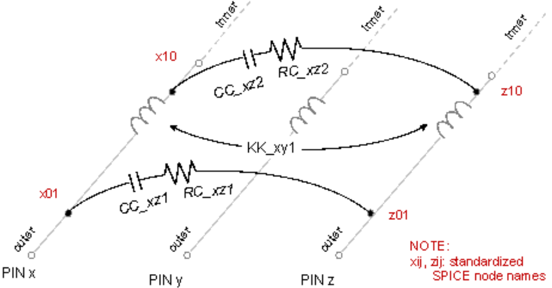
we use 3-digit node names:
package pin 1: the outer package node is 100,

the first segment is from 101 to 110
the next from 110 to 120
the last from iii to 199
package pin 2: the outer package node is 200,
the first segment is from 201 to 210
etc.

This leads to the following sub-models for the TSOP44 package:





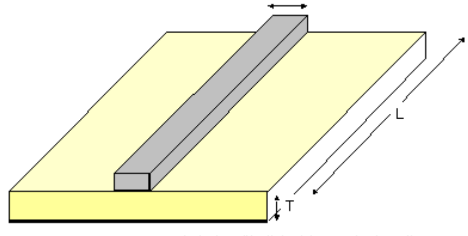


Modeling considerations for our example:

- We will only model pin4
- For simplicity of this demo, we assume that all pins of the TSOP44 package have an identical model

We commence with modeling the SPICE components along the pin. The start values of the model parameters could either come form 3-D field simulations, or, like in this example, from some raw hand calculation formulae.

Line capacitance towards ground plane and line inductance:

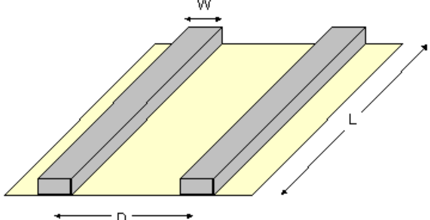


epsilon_r = 2

$$C = 88.5E-6 \cdot \text{epsilonr} \cdot \left(\frac{W \cdot L}{T}\right) \quad \text{in fF}$$

$$L \sim 2 \cdot L \cdot \left(\ln\left(\frac{L}{W}\right) + 0.8\right) \quad \text{in nH}$$

Taking 2 lines inter considerations, we get for the capacitance between them:
Equations from : Nührmann, Das grosse Werkbuch Elektronik, Franzis-Verlag, Munich, 1984, ISBN 3-7723-6544-2
This gives the first modeling result for the TDR measurement of the package, pressed with an electrically neutral stamp onto our test fixture:

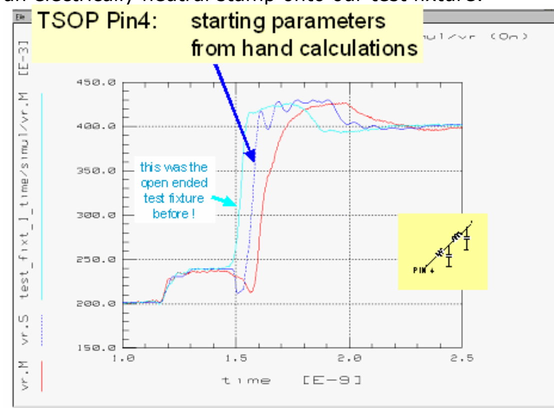


substrate with dielectric constant epsilonr

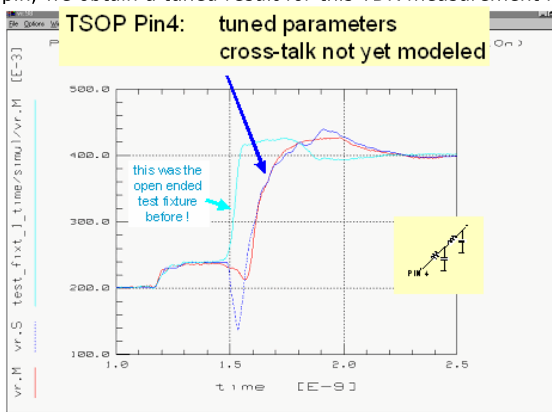
$$C = \frac{88.5E-6 \cdot epsilonr}{1.2} \cdot \left(\frac{W \cdot L}{D}\right) \text{ in fF}$$

Equations from : Nührmann, Das grosse Werkbuch Elektronik, Franzis-Verlag, Munich, 1984, ISBN 3-7723-6544-2

This gives the first modeling result for the TDR measurement of the package, pressed with an electrically neutral stamp onto our test fixture:

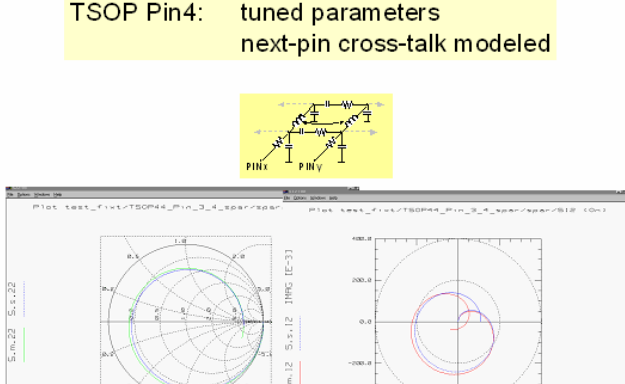


Fine-tuning the line capacitances towards ground and the inductance along the package pin, we obtain a tuned result for this TDR measurement like in the plot below:

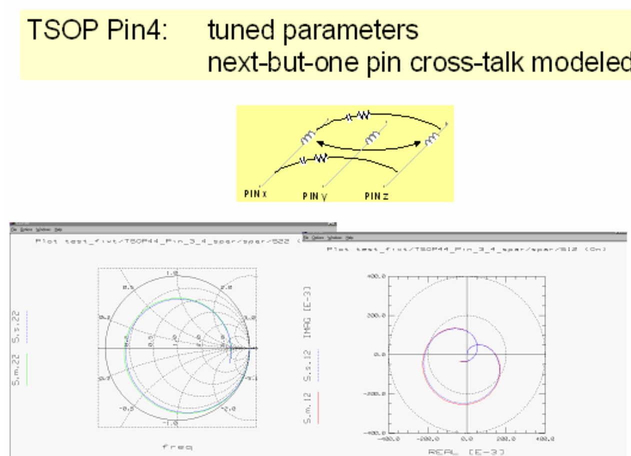


The remaining fitting problems are assumed to be related to cross-coupling between the lines. For resolution reasons, as mentioned already further above, we therefore switch to NWA measurements.

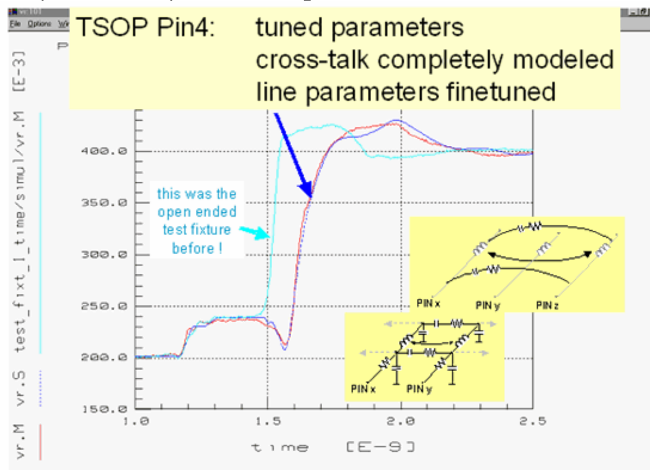
With the NWA calibrated to the ends of its two cables, and with the test fixture already modeled, as well as with a first estimation of the SPICE components along the package pin, we can fine-tune the crosstalk PICE components between our package pin and its two direct neighbor pins, and obtain:



In a second step, we now also tune the SPICE parameters representing the cross-talk to the next-but-one neighbor pins of the package. The plot below shows the result:



With this fitting result, we go back to the TDR measurement and check the fitting there. If required, a small parameter fine-tuning is performed here and back in the S-parameter domain as well.
The plot below depicts the fitting result for the TDR measurement:



Publications

Modeling and Simulation of high pincount connector systems, T. Gneiting, H. Khakzar, Proc. 4. European IC-CAP User Meeting, Berlin 1997

H.Wolter, H.Katzier, O.Podebrad und T.Weiland: Field-Based Analysis of a High Pincount Board Connector; IEEE Transaction on Components, Packaging, and Manufacturing Technology, Part B. Advanced Packaging, Feb. 1996, Vol.: 19, No. 1, 23-31.

H.Katzier, R.Reischl und P.Pagnin: SPICE-Models for High Pincount Board Connectors; IEEE Transaction on Components, Packaging, and Manufacturing Technology, Part B. Advanced Packaging, Feb. 1996, Vol.: 19, No. 1, 3-6.

H.Katzier, R.Reischl: Elektrische Simulationsmodelle für hochpolige Leiterplattenverbinder;

Elektronik, Nov. 1997, No. 24, 104-112.

Mark S. Mirotznik and Dennis Prather: How to choose EM software. IEEE Spectrum, 12, 1997, 53-58.

H.Katzier, R.Reischl, Siemens AG, Munich, Reliable SPICE Models for Lossy Coupled Transmission Lines, , 2000 High-Performance System Design Conference

R.Reischl; Modeling of BGA-IC-Packages; European ICCAP-User-Meeting, 17-18. June 1999, Marseile, France.

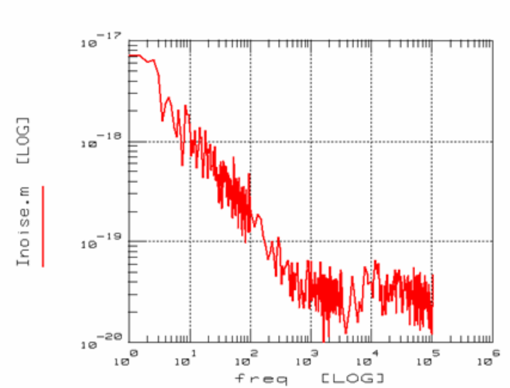
Low Frequency (1/f) Noise Modeling for Semiconductors

Contents

- *Introduction* (iccapmhb)
- *Types of Noise in Semiconductor* (iccapmhb)
- *Noise Models in Semiconductors* (iccapmhb)
- *Low Noise Measurement Setup* (iccapmhb)
- *Low Noise Measurements and Modeling* (iccapmhb)
- *Noise Parameter Extraction and Verification* (iccapmhb)
- *Appendix* (iccapmhb)

Introduction

Phase noise of oscillator circuits is among the key parameters of today's communication systems. It limits the modulation quality of the information signal, and the cross-talk to adjacent channels. In the receiver, on the other hand, it can reduce the selectivity and the demodulation quality. With today's trend to even more complex modulation schemes, low phase noise becomes even more critical. Therefore, minimizing noise with the design of high frequency communications circuits is a must for the cost-effective exploitation of limited bandwidth, and to obtain low bit error rates.
In order to predict the noise behavior of such systems correctly, accurate noise models are required. Without them, the design and optimization of an amplifier's noise figure or the phase noise of oscillators cannot be successful.
A key element is the transistor. While for linear circuits the modeling of the noise at the operating frequency f0 is sufficient, non-linear circuits do also convert the low-frequency noise up to the operating frequency range. For oscillators as an example, the 1/f low frequency noise at fm will be mixed upwards to contribute to the phase noise of the total circuit at the frequencies f0_fm and f0-fm . Therefore, noise modeling of transistors can be split into a high and a low frequency segment. The graph below depicts the Collector low-frequency noise spectral density $[A^2/Hz]$ of a bipolar transistor.
It can be seen that the resolution of the toolkit measurement setup is at about $0.2nA/\sqrt{Hz}$.



Broadband noise in bipolar or FET models is determined essentially by thermal noise and shot noise. These noise sources are automatically determined within the model from the large signal model parameters. An important prerequisite for this is, however, that the noise determining parameters, e.g. the Base resistor of bipolar transistors, has been determined carefully and that their values have a physical meaning. The low frequency noise, on the other hand, dominated by the 1/f noise, is modeled by some specific model parameters. In general, these are the AF and KF and sometimes BF/EF parameters, covered in this chapter. With some models, like the BSIM3v3, an alternate 1/f noise model is available too [10,15].

Noise Term Definitions

In order to keeps things simple, we consider white noise of a resistor. This frequency-independent and bias-independent noise is described in terms of a voltage by the well-known formula

$$\overline{v_{nR}^2} = 4 \cdot k \cdot T \cdot \Delta f * R$$

where,

$\overline{v_{nR}^2}$	*effective noise voltage* (root means square value), in [V²]
$4 \cdot k \cdot T \cdot \Delta f$	this product represents a power, with dimension [A*V]
R	resistance value

Since the term 4kTΔf represents a power, and if we normalize this power to Δf, we end up with a power density. Usually, this normalization is done for Δf=1Hz.

$$\frac{\overline{v_{nR}^2}}{\Delta f} = 4 \cdot k \cdot T * R$$

This result can be plotted against frequency. In case of our resistor, it is a constant value. For semiconductors, such a plot can also exhibit frequency dependencies, e.g. 1/f noise. In any case, such a plot shows a spectrum.

Therefore, this term $\frac{\overline{v_{nR}^2}}{\Delta f}$ is called *power noise spectral density*.
However, this is not exact, referring to the dimensions. A more appropriate term is therefore *voltage noise spectral density*. Its symbol is usually Snv, and its dimension is

$[V^2/Hz]$

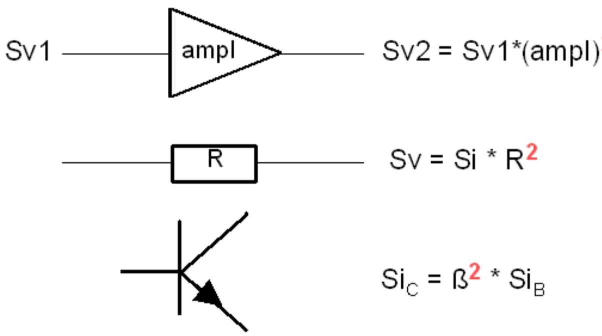
$$S_{nv} = \frac{\overline{v_{nR}^{\,2}}}{\Delta f} = 4 \cdot k \cdot T * R$$

In some cases, the square root of Snv is used, called the *equivalent noise voltage*, $E_{nv}[V/\sqrt{Hz}]$

$$E_{nv} = \sqrt{\frac{\overline{v_{nR}^{\,2}}}{\Delta f}}$$

> Note: Using Env instead of Snv does not require to calculate noise transmissions/amplifications using squares of R, amplification etc.
> When expressing noise in terms of currents, e.g. $S_{ni} = \frac{\overline{i_{nR}^{\,2}}}{\Delta f} = 4 \cdot k \cdot T * \frac{1}{R}$, the same terminology applies.

> Note: When performing circuit calculations based on noise spectral densities, keep in mind that you have to calculate with squares instead of the commonly used expressions.



Types of Noise

This chapter deals with the description of the different noise mechanisms and their noise density spectrum in semiconductor components. It is basically aimed to highlight the main fundamentals rather than to explain the individual physical effects. We start with the thermal noise of the ohmic parasitics, and lead then over to the semiconductor-specific noise sources.

Thermal Noise

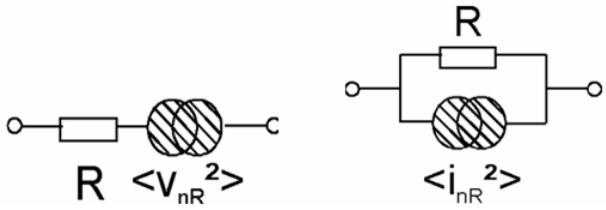
Related to the thermal oscillation of electrons in a resistor, we can measure a current at its contacts, without applying an external voltage. This is the thermal noise current. The effective current is described by

$$\overline{i_{nR}^2} = 4 \cdot k \cdot T \cdot \frac{1}{R} \cdot \Delta f \qquad (1)$$

An equivalent schematic for this condition is a noise-free resistor with a current noise source in parallel.
In analogy to this circuit, a noise-free resistor and a serial noise voltage source can be used as well. In this case, the effective voltage is given by

$$\overline{v_{nR}^2} = \overline{i_{nR}^2} * R^2 \overset{(1)}{=} 4 \cdot k \cdot T \cdot R \cdot \Delta f \qquad (2)$$

Equivalent schematic of a noisy resistor



Note: since noise sources have a mean value of '0', but an effective value of usually non-zero, effective voltages and currents are used to calculate the circuit performance with respect to the noise sources. This means that such an effective voltage is amplified by the square of the amplifier's gain. As another example, two parallel current noise sources can be represented by a single noise current source with a value of

$$\overline{i_{nR_total}^2} = \overline{i_{nR1}^2} + \overline{i_{nR2}^2} \overset{(1)}{=} 4 \cdot k \cdot T \cdot \left(\frac{1}{R_1} + \frac{1}{R_2}\right) \cdot \Delta f$$

Also, the ohmic law is now referring to the square of voltages and currents, see Equation(2).

Back to the thermal resistor noise, the equivalent noise power density spectra are

$$C_i(f) = 4 \cdot k \cdot T \cdot \frac{1}{R} \qquad (3)$$

and

$$C_V(f) = 4 \cdot k \cdot T \cdot R \qquad (4)$$

These frequency independent noise spectra represent a simplification. An accurate calculation based on a quantum mechanic model gives

$$C_i(f) = 4 \cdot \frac{1}{R} \cdot h \cdot f \left(\frac{1}{2} + \frac{1}{e^{\frac{h \cdot f}{k \cdot T}} - 1} \right) \qquad (5)$$

so that equations (1) to (4) are basically only valid for h*f<<k*T , i.e. for 'low' frequencies and high temperatures. However, the quantum noise for h*f>>k*T has to be considered basically only for frequencies very much higher than in RF and microwave applications, i.e. the >1013Hz range.

Shot Noise

The current through the space charge area of a diode is composed of many individual current impulses, due to the transport of individual charge carriers. Since this motion of electrons/wholes is statistical, we always have, besides the expected DC current, also a noise component. With the assumption of individual, rectangular current impulses of the width τ for every charge component, we can calculate a power density spectrum after [1,2]

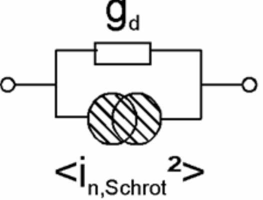
$$C_i(f) = 2 \cdot e \cdot I \cdot \frac{\sin^2(\pi \cdot f \cdot \tau)}{(\pi \cdot f \cdot \tau)^2} \qquad (6)$$

For low frequencies, i.e. small values of τ * f , there is sin(x) /x ~ 1 and we get the commonly used equation

$$C_i^{NF}(f) = 2 \cdot e \cdot I \qquad (7)$$

for the shot noise. The figure below shows the equivalent schematic of an ideal, but noisy

diode. The noise current is described by a current source in parallel with the small signal conductance.



Generation-Recombination Noise

In semiconductors, there is a static charge carrier generation/recombination process. This refers to a statistic change of the carrier states. Referring to the energy band model, transitions between conduction and valence band as well as various trap levels are possible, leading to a noise spectrum following

$$C(f) = C_0 \cdot \frac{1}{1+\left(\frac{f}{f_g}\right)^2} \qquad (8)$$

The current dependency of the combination/recombination noise, also called burst-noise or popcorn-noise [3], is generally reflected by a modeling equation like

$$C(f) = KB \cdot \frac{I^{AB}}{1+\left(\frac{f}{FB}\right)^2} \qquad (9)$$

with the modeling parameters KB, AB and FB.

1/f Noise

A pretty often measurable phenomenon is noise with a spectrum proportional to 1/f. This leads to the name 1/f noise. Another name is flicker noise. It is caused essentially by recombination effects at defects in the semiconductor volume, the borders of diffusion areas or the material surface.
An empirical description after Hooge [4, 5] is a spectrum with

$$C_{1/f} = \frac{\alpha}{N_{tot}} \cdot \frac{1}{f} \qquad (10)$$

where Ntot means the total number of moving charges in the device. The Hooge-Parameter is a material characteristic.
In most simulation programs, the current dependency of the 1/f noise is covered in analogy to the generation/recombination noise by an exponential form like

$$C_{1/f} = KF \cdot \frac{I^{AF}}{f^{B}} \qquad (11)$$

with the model parameters AF, KF and B. B is commonly set to '1'.

Noise Models in Simulation Programs

When simulating the noise behavior of circuits, all circuit components need to have noise models included. All lossy components will exhibit thermal noise, corresponding to the simulation temperature TEMP. Semiconductor devices will additionally also exhibit 1/f noise.

Resistors

Their noise distribution is modeled by white noise, as described above i.e., the effective current is described by

$$\overline{i_r^2} = 4 \cdot k \cdot T \cdot \frac{1}{R} \cdot \Delta f$$

representing a noise-free resistor with a current noise source in parallel.
Alternatively, a noise-free resistor and a serial noise voltage source can be used as well.

$$\overline{v_r^2} = 4 \cdot k \cdot T \cdot R \cdot \Delta f$$

Inductors, Capacitors, and Lines

Inductors, Capacitors, and Lines are considered noise-free. This is valid for ideal components. For real RF components, including parasitic losses (SPICE sub-circuits), the noise contributions stem basically from the parasitic resistors.

Diodes

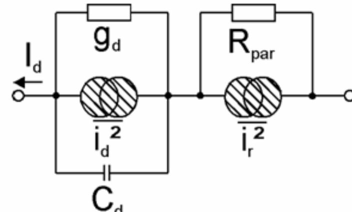
The main noise source of diodes is shot noise. Besides this, 1/f noise as well as thermal noise due to the parasitic resistor is observed. The capacitance CDiode, covering the space charge and the diffusion capacitance, does not contribute to the noise.
This leads to a small signal schematic with the noise sources.

$$\overline{i_D^2} = 2 \cdot e \cdot I_d \cdot \Delta f + KF \cdot \frac{I_d^{AF}}{f} \cdot \Delta f \qquad (12)$$

and

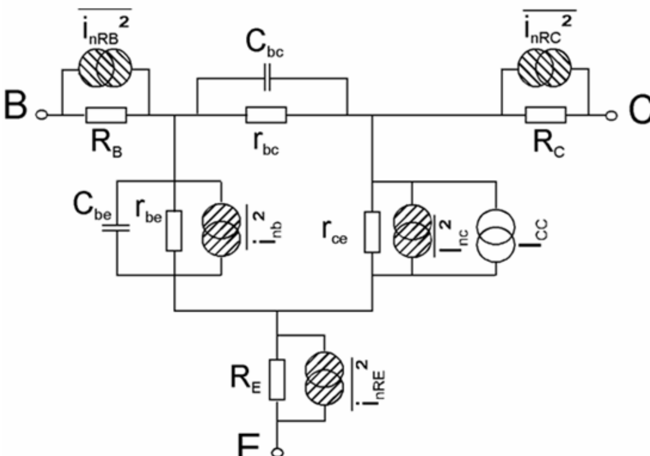
$$\overline{i_R^2} = 4 \cdot k \cdot T \cdot \frac{1}{R_{par}} \cdot \Delta f \qquad (13)$$

Since the shot noise formula refers directly to the diode current Id and the thermal noise to the parasitic resistor Rpar , it is only the parameters AF and KF which have to be modeled additionally.



Bipolar Transistors

SPICE and similar simulators feature a noise schematic for bipolar transistors.



Associated with each physical resistor (Rb at the Base, Rc at the Collector and Re at the Emitter), is a thermal noise source

$$\overline{i_{R,i}^{2}} = 4 \cdot k \cdot T \cdot \frac{1}{R_i} \cdot \Delta f, \quad i = b, c, e \qquad (14)$$

Base and Collector currents are considered to be independent. Therefore, the shot noise can be described by a source

$$\overline{i_{b,S}^{2}} = 2 \cdot e \cdot I_b \cdot \Delta f \qquad (15)$$

at the Base and

$$\overline{i_{c,S}^{2}} = \overline{i_{nc}^{2}} = 2 \cdot e \cdot I_c \cdot \Delta f \qquad (16)$$

at the Collector. The total 1/f noise of the transistors is described by a noise source in parallel with the Base-Emitter contact. This leads to the noise source at the Base following

$$\overline{i_{nb}^{2}} = 2 \cdot e \cdot I_b \cdot \Delta f + KF \cdot \frac{I_b^{AF}}{f} \cdot \Delta f \qquad (17)$$

with, again, the noise model parameters AF and KF.
Agilent's Advanced Design System (ADS) include with the BJT transistor model additionally a generation/recombination noise effect in the noise spectrum of the Base. Referring to equ. (17), this means

$$\left(\overline{i_{nb}^{2}}\right)^{ADS,BJT} = 2 \cdot e \cdot I_b \cdot \Delta f + KF \cdot \frac{I_b^{AF}}{f} \cdot \Delta f + KB \cdot \frac{I_b^{AB}}{1 + \left(\frac{f}{FB}\right)^2} \cdot \Delta f \qquad (18)$$

with the new model parameters KB, AB and FB.
With PSPICE, the Collector current ic includes an additional 1/f noise source. This gives

$$\left(\overline{i_{nc}^{2}}\right)^{PSPICE} = 2 \cdot e \cdot I_c \cdot \Delta f + KF \cdot \frac{I_c^{AF}}{f} \cdot \Delta f \qquad (19)$$

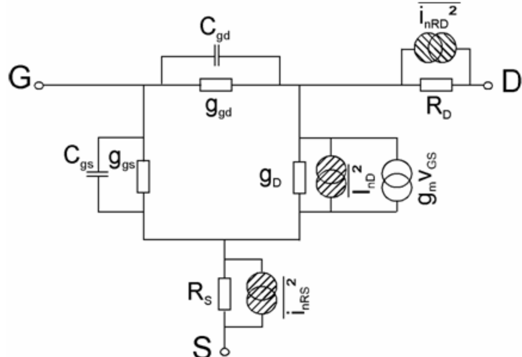
using the same parameters KF and AF of the Base contact.
The VBIC model for bipolar transistors [6] includes, similar to the UCB SPICE BJT model, thermal noise sources as well as shot noise for the Base-Emitter current and the transport current. The 1/f noise is described again by a noise source between Base and Emitter with

$$\left(\overline{i_{nb}^{2}}\right)^{VBIC} = 2 \cdot e \cdot I_b \cdot \Delta f + KF \cdot \frac{I_b^{AF}}{f^{BF}} \cdot \Delta f \qquad (20)$$

However, the frequency dependence is additionally modeled by the parameter BF. Besides the main NPN transistor, VBIC includes also a parasitic PNP transistor. Its noise contribution is modeled in analogy to the NPN, sharing the same model parameters for the 1/f noise. More details can be found in [7,8].
Noise in the MEXTRAM model is described like with the SPICE BJT model [9]. Besides thermal noise, there are shot noise sources associated with the Base-Emitter current as well as the transport current. 1/f noise is described after (17).

Junction FETs

The noise schematic of field effect transistors is most often as shown below



The noise behavior is described by four noise sources. The resistors RD and RS are associated with thermal noise after

$$\overline{i_{RD}^{2}} = 4 \cdot k \cdot T \cdot \frac{1}{R_D} \cdot \Delta f \qquad (21)$$

and

$$\overline{i_{RS}^{2}} = 4 \cdot k \cdot T \cdot \frac{1}{R_S} \cdot \Delta f \qquad (22)$$

The amplifier effect of JFETs is based on changes in the channel resistor. This leads to describing the noise of the channel current ID also by thermal noise

$$\overline{i_{nD,th}^{2}} = \frac{8}{3} \cdot k \cdot T \cdot g_m \cdot \Delta f \qquad (23)$$

Additionally, an 1/f noise source is assumed with the channel after

$$\overline{i_{nD}^{2}} = \frac{8}{3} \cdot k \cdot T \cdot g_m \cdot \Delta f + KF \cdot \frac{I_D^{AF}}{f} \cdot \Delta f \qquad (24)$$

with the model parameters AF and KF. The shot noise of the Gate current is usually neglected.
This model is valid for the simulators SPICE2, SPICE3, PSPICE and ADS. With HSPICE, an additional series resistor at the Gate is assumed with corresponding thermal noise. For the

thermal noise of the channel, there is a choice to replace equ.(23) by a more accurate model [3].

With $$\left(\overline{i_{nD,th}^{2}}\right)^{HSPICE} = \frac{8}{3}\cdot k\cdot T\cdot\beta\cdot\left(V_{gs}-V_{T0}\right)\cdot\frac{1+\alpha+\alpha^{2}}{1+\alpha}\cdot GDSNOI\cdot\Delta f \qquad (25)$$

and $$\alpha = \begin{cases} 1-\dfrac{V_{ds}}{V_{gs}-V_{T0}} & \text{linear region} \\ 0 & \text{saturation region} \end{cases} \qquad (26)$$

it differentiates between linear and saturated region. And, with GDSNOI , an additional noise parameter is introduced.

MOSFETs

The noise formulation of MOSFET models in most of the commonly used simulation programs is based on the SPICE2 model of the UCB University of California, Berkeley, described in [3]. It is known, however, that the thermal channel noise included in this model is essentially valid only in the saturated region of the output characteristics [11]. Therefore, the latest UCB model, the BSIM3v3 model [10], models the noise differently, see further below.
The figure depicts the noise equivalent schematic of the SPICE2 MOS2 and MOS3 model. Like with the JFET transistor, the resistors RD and RS are associated with thermal noise after

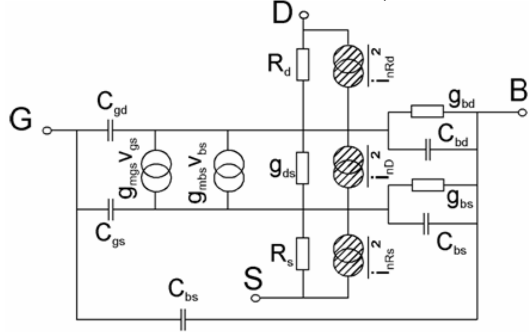
$$\overline{i_{Rd}^{2}} = 4\cdot k\cdot T\cdot\frac{1}{R_d}\cdot\Delta f \qquad (27)$$

and $$\overline{i_{Rs}^{2}} = 4\cdot k\cdot T\cdot\frac{1}{R_s}\cdot\Delta f \qquad (28)$$

The channel noise is described by

$$\overline{i_{nD}^{2}} = \frac{8}{3}\cdot k\cdot T\cdot g_m\cdot\Delta f + KF\cdot\frac{I_D^{AF}}{f\cdot C_{ox}\cdot L_{eff}^{2}}\cdot\Delta f \qquad (29)$$

using again the parameters AF and KF for the flicker noise. Cox and Leff are calculated inside the model from the other model parameters.



In PSPICE, an additional thermal noise of the Gate and of the Bulk resistor is included. With ADS, the models MOSFET Level 1-3 and EEMOS1 are also based on figure above. With EEMOS1, however, no 1/f noise is included.. With HSPICE, like with the mentioned JFETs from above, the user has the choice between this and an alternate noise description. The thermal channel noise is described again after equ.(25), with the exception

$$\alpha = \begin{cases} 1-\dfrac{V_{ds}}{V_{D,sat}} & \text{in the linear region} \\ 0 & \text{in the saturation region} \end{cases} \qquad (30)$$

With the BSIM3v3 model of UCB SPICE3, the user can select for both, the flicker noise as well as for the thermal channel noise, one of the following models. This is done by setting the model parameter noimod accordingly. Either, a slightly modified SPICE2 model describes the 1/f noise by

$$\overline{i_{1/f}^{2}} = KF\cdot\frac{I_D^{AF}}{f^{EF}\cdot C_{ox}\cdot L_{eff}^{2}}\cdot\Delta f \qquad (31)$$

Here, with the same functionality like in the VBIC bipolar model, an additional parameter EF has been introduced, modeling the frequency dependence of the noise spectrum. The thermal channel noise is calculated in BSIM3v3 after

$$\overline{i_{th}^{2}} = \frac{8}{3}\cdot k\cdot T\cdot\left(g_m+g_{ds}+g_{mb}\right)\cdot\Delta f \qquad (32)$$

Or, depending on the parameter NOIMOD, the BSIM3v3 noise model calculates the 1/f noise after

$$\overline{i_{1/f}^{2}} = f\left(NOIA, NOIB, NOIC, EF, EM\right) \qquad (33)$$

using a relatively complex formula, and involving the five new noise parameters NOIA, NOIB, NOIC, EF and EM as well as large signal model parameters [34].
For the channel noise, it is

$$\overline{i_{th}}^2 = \frac{4 \cdot k \cdot T \cdot \mu_{eff}}{L_{eff}^2} \cdot |Q_{inv}| \cdot \Delta f \quad (34)$$

with $$Q_{inv} = -W_{eff} \cdot L_{eff} \cdot C_{ox} \cdot V_{gsteff} \cdot \left(1 - \frac{A_{bulk}}{2 \cdot (V_{gsteff} + 2 \cdot v_t)} \cdot V_{dseff}\right) \quad (35)$$

The procedure described further below in this paper, covering the 1/f noise parameter extraction, can therefore be applied to the BSIM3v3 model only if the SPICE2 model for the 1/f noise has been selected specifically by the parameter NOIMOD. For extracting the parameters NOIA, NOIB, NOIC, EF and EM, see the BSIM3v3 toolkit for IC-CAP.

MESFETS

The noise behavior of MESFETs is modeled in the simulation programs PSPICE and HSPICE essentially analogous to the corresponding JFET -models [3]. See [12] for more details.

As a final remark, the noise bandwidth Δf is set to Δf=1Hz in most simulators.

Low (1/f) Noise Measurement Setup

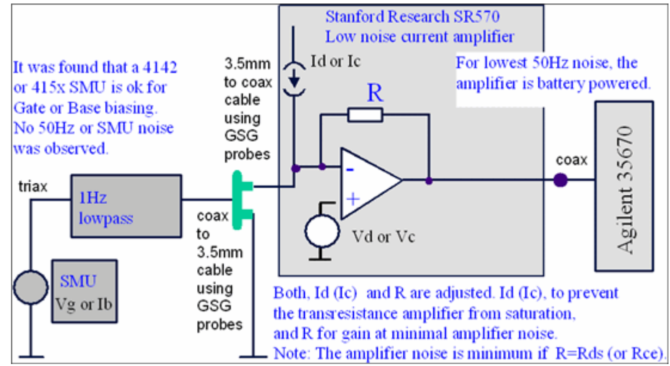
When measuring noise of high frequency transistors, we continue to distinguish between a low frequency range up to several MHz, followed then by a second range up to the maximum operating frequency of the device (GHz). In the first range, the noise power density spectrum is measured in order to characterize the 1/f noise and the generation/recombination noise. In the upper range, the noise figure Fmin, and the parameters Rn , Yopt are measured.

We will now discuss a possible measurement setup for the first range, the 1/f noise characteristics. It is depicted in figure below. It consists of an HP4142 or HP415x SMU, which, however, is operated in 'power supply' mode. This means that it is triggered to output either a current or voltage (depending on the type of transistor), and to keep this current or voltage switched on, until the 1/f noise measurement is over and a new 'stop' command is sent to the SMU. The SMU is connected to the Gate or Base of the transistor by a 1Hz filter. The SMU is connected to the Gate or Base of the transistor by a 1Hz filter. Its output resistance is switchable. For MOS transistors, a low output impedance of the filter is required (e.g. 50Ω), while a high output impedance (e.g. 500kΩ) is used for bipolar transistors.
Note: for a bipolar transistor, the 1/f noise is generated in the Base region. If the 1Hz filter had a low output impedance, it would shorten this noise source and we would measure/simulate a too low 1/f noise characteristic. The higher the 1Hz filter's output impedance is, the less shortening occurs, see also [14]. In the toolkit, a value of ~330kΩ has been found to work best.

The output of the transistor is connected to a special low-noise current amplifier. Besides amplifying the 1/f noise, this amplifier also serves to bias the Drain or Collector of the transistor, and also to compensate the DC Drain or Collector current with an ultra-low noise DC current source. This allows to amplify only the noise current of the transistor:

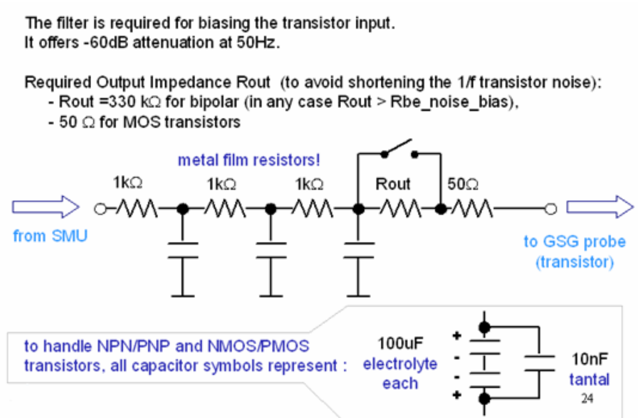
$$\overline{V^2_{output_amplifier}} = R^2 \cdot \overline{I^2_{noise_transistor}}$$

With this setup, the cabling is minimized. Only 2 short coax cables are used to bias the transistor and to measure the noise. Therefore, extremely reliable and extremely clean noise measurements are obtained.
The output of the noise amplifier is further connected to an HP35670 dynamic signal analyzer, which then converts the 1/f noise time signals into a frequency spectrum.
The whole setup is controlled by special IC-CAP macros for bipolar transistors, and for MOS transistors. Ready-to-use extraction routines for the noise parameters AF, KF and EF (for MOS) are included as well.

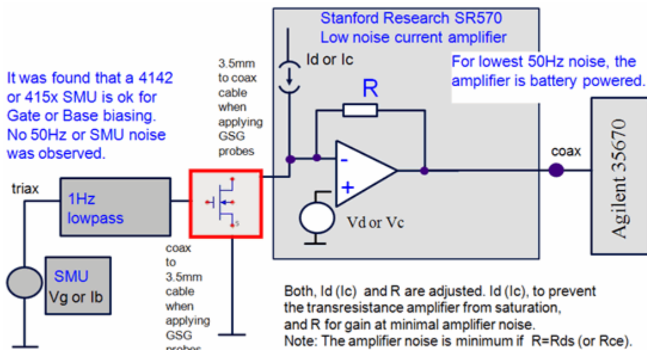


Note: For on-wafer measurements, special attention has to be paid regarding shielding. It is recommended to use a prober with a special 1/f shielding. However, most important, it is recommended to use the G-S-G RF probes instead of using DC probes

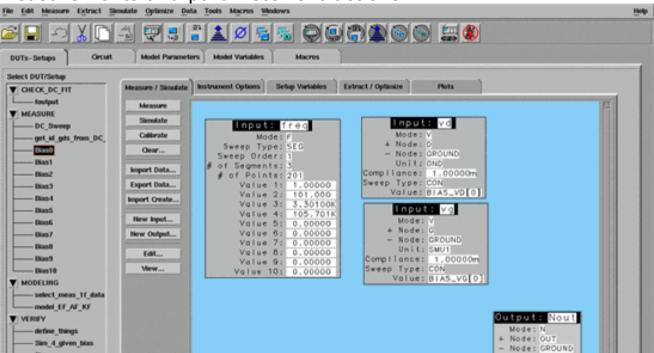
1Hz Low Pass Filter

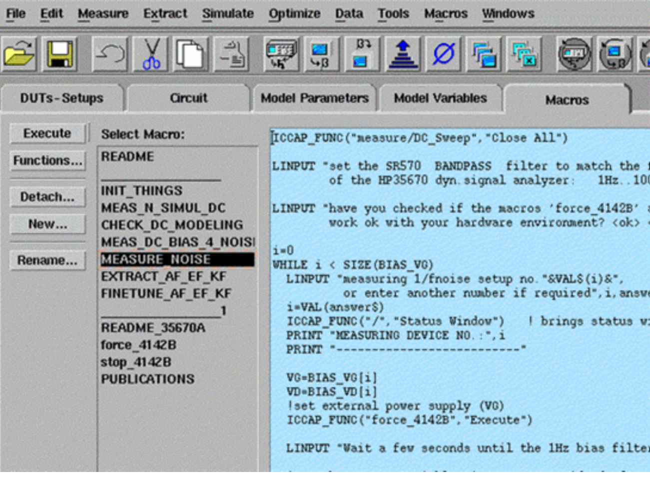


Measurement Setup



The 1/f noise modeling toolkit contains all required IC-CAP measurement setups and driver macros for measuring the bipolar and MOS transistor 1/f noise and to extract the model parameters. Figure below shows its structure, followed by the macros for measurements and parameter extractions.

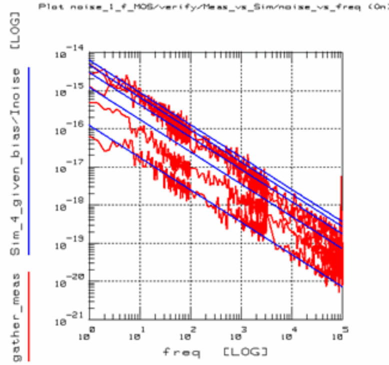




Measurement resolution of the setup:
The current noise spectral density resolution obtainable with this system has been measured as ~5E-21 A^2/Hz.

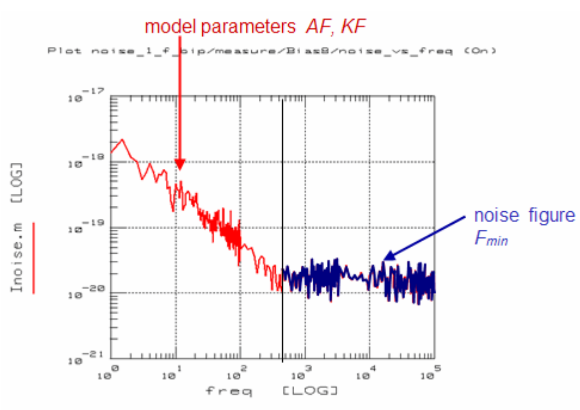
1/f Noise Measurements and Modeling

This presentation is about the bipolar and MOS transistor 1/f noise modeling toolkit before its implementation into IC-CAP as Agilent product number 85195B.



Introduction

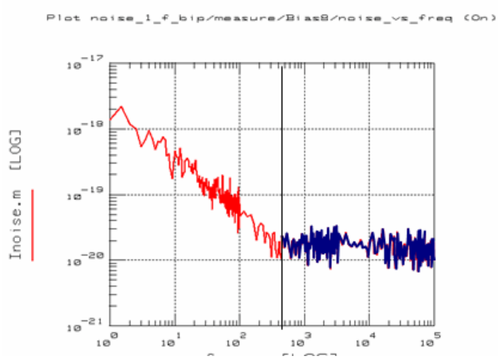
Design Goals For Minimum Noise Circuits in RF Applications



- Low frequency noise:'classical' model parameters AF, KF
- High frequency noise,above 1/f noise:noise figure Fmin

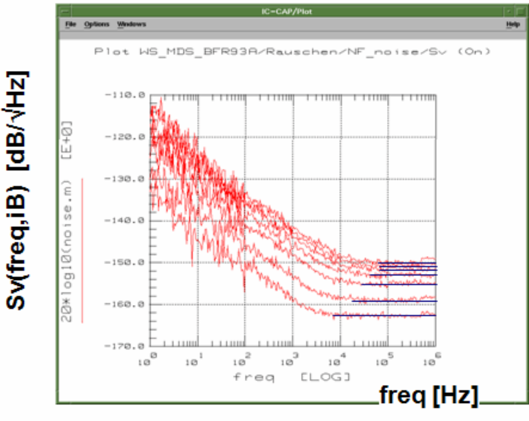
Low Frequency 1/f Noise

As shown in the figure below an effect of low frequencies, this type of noise is upconverted in mixers and VCOs, and can affect the overall circuit performance considerably.



Noise dependencies in Semiconductors

1/f Noise: Operating Point and Frequency Dependent
White Noise: Operating Point Dependent



Description of Noise Signals

Noise signals exhibit power, and therefore, their quadratic mean value is non-zero

$$\overline{v_{noise}^{2}(t)} = \lim_{T\to\infty} \frac{1}{T} \int_{0}^{T} v_{noise}^{2}(t)dt \neq 0$$

This means that noise signals can be described with effective values

$$\overline{v_{noise}^{2}(t)} = v_{noise\,eff}^{2}$$

However, since the time function of noise is unknown, the conventional formula for effective signals in the time domain cannot be applied.
The Wiener-Kintchine theorem permits to associate the auto-correlation function

$$R(\tau) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} v_{noise}(t) \cdot v_{noise}(t+\tau)\, dt$$

with the **spectral** power distribution

$$R(0) = \int_{-\infty}^{\infty} S(f)df = S(f)\Delta f = \overline{v_{noise}^{2}(t)}$$

6

This means, noise is best described by its spectral distribution function
Therefore, we measure the noise spectral density per 1Hz:

Voltage: $S_V(f) = \frac{\overline{v_{noise}^{2}}}{\Delta f}$ Unit: $[V^2/Hz]$

Current: $S_i(f) = \frac{\overline{i_{noise}^{2}}}{\Delta f}$ Unit: $[A^2/Hz]$

The resistor effective noise voltage (root means square value), in $[V^2]$, is

$$\overline{v_{nR}^{2}} = 4 \cdot k \cdot T \cdot \Delta f * R$$

The term 4kTΔf represents a power. We normalize this power to f, what gives a power density.
Usually, this normalization is done for Δf=1Hz.

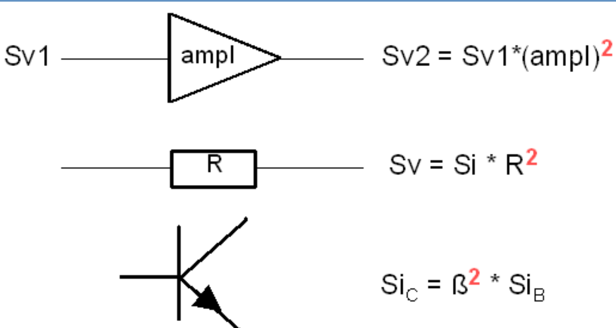
$$\frac{\overline{v_{nR}^{2}}}{\Delta f} = 4 \cdot k \cdot T * R$$

This result can be plotted against frequency. In case of our resistor,it is a constant value.
For semiconductors, such a plot can also exhibit frequency dependencies, e.g. 1/f noise.
In any case, it shows a spectrum.
Therefore, this term is called **voltage noise spectral density**. Its symbol is Snv, and its dimension is $[V^2/Hz]$

$$S_{nv} = \frac{\overline{v_{nR}^{2}}}{\Delta f} = 4 \cdot k \cdot T * R$$

Circuit Calculations with Noise Spectral Densities

When performing circuit calculationson noise spectral densities, or effective noise voltages/currents, keep in mind the unit is a square: $[V^2]$, $[V^2/Hz]$, I^2, or I^2/Hz etc.

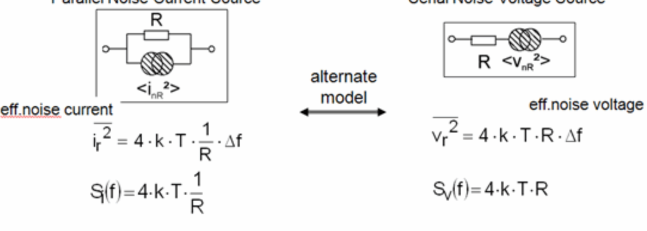


Therefore:

Noise Mechanism

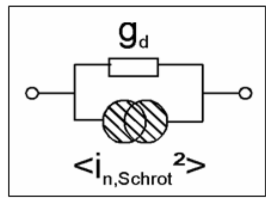
Thermal Noise

Thermal Noise is current through a resistor
Lossy passive components are known to produce temperature dependent and bias independent white noise due to the thermal induced motion of charge carriers.



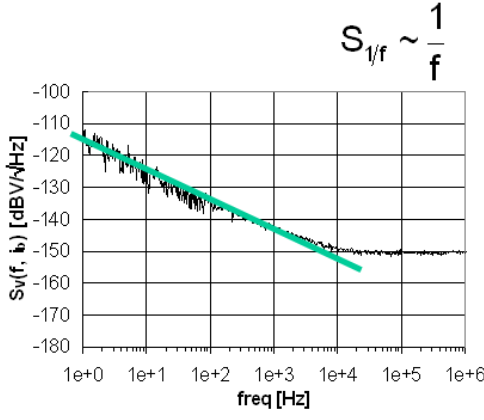
Shot Noise

Shot Noise is current through a semiconductor junction
Semiconductor junctions are known to generate noise caused by the quantized and random current flow across them. This (typically) white noise is bias-dependent but independent of temperature



1/f Flicker Noise

This low-frequency, bias-dependent noise effect is typically attributed to traps associated with contamination and crystal defects.

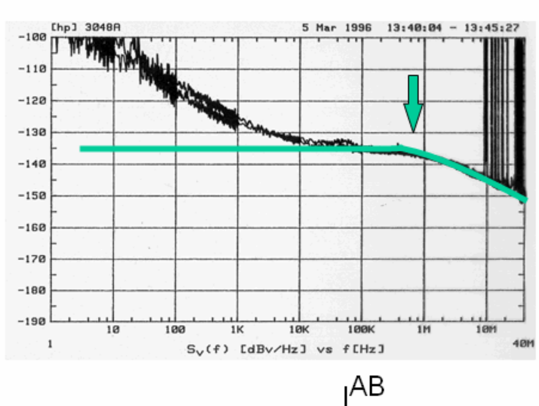


$$S_{1/f} = KF \cdot \frac{I^{AF}}{f^{B}}$$

Generation/Recombination Noise (also called: burst or popcorn noise)

This noise is due to generation/recombination effects and trapping and is observed in some semiconductor devices when other noise effects are minimal. It is bias and frequency dependent.

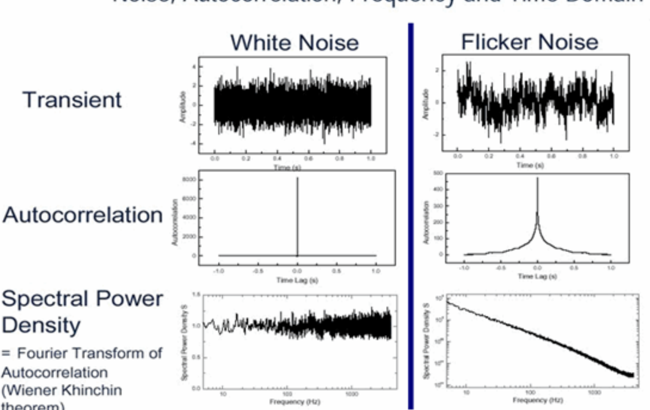
1/f Noise overlaid by Generation/Recombination Noise



$$S_{gr} = KB \cdot \frac{I^{AB}}{1 + \left(\frac{f}{FB}\right)^2}$$

The Shape of White Noise and Flicker Noise

Noise, Autocorrelation, Frequency and Time Domain



Noise Models

Diodes

1/f noise contributions

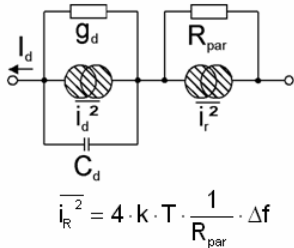
- Thermal Noise: Rpar

$$\overline{i_r^2} = 4 \cdot k \cdot T \cdot \frac{1}{R} \cdot \Delta f$$

- Shot Noise and 1/f Noise: Id

$$\overline{i_D^2} = 2 \cdot e \cdot I_d \cdot \Delta f + KF \cdot \frac{I_d^{AF}}{f} \cdot \Delta f$$

Noise Models in Simulators



Bipolar Transistors

1/f noise contributions

- Thermal Noise: RB, RC, RE

$$\overline{i_{nR\,i}^2} = 4 \cdot k \cdot T \cdot \frac{1}{R_i} \cdot \Delta f$$

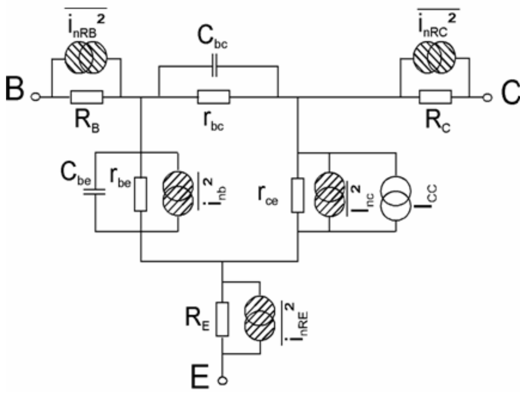
- Shot Noise and 1/f Noise: Ibe

$$\overline{i_{nb}^2} = 2 \cdot e \cdot I_b \cdot \Delta f + KF \cdot \frac{I_b^{AF}}{f} \cdot \Delta f$$

- Shot Noise: Ice

$$\overline{i_{nc}^2} = 2 \cdot e \cdot I_c \cdot \Delta f$$

Noise Models in Simulators



Junction FETs

1/f noise contributions

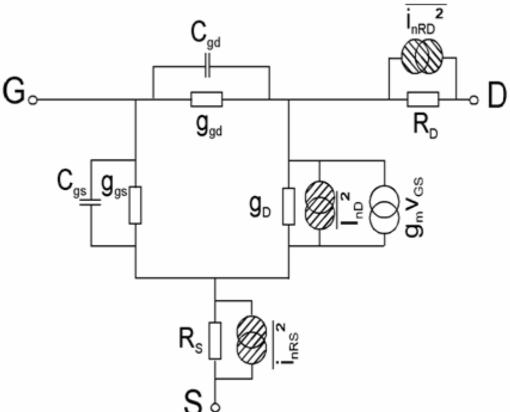
- Thermal Noise: RD, RS

$$\overline{i_{RD}^2} = 4 \cdot k \cdot T \cdot \frac{1}{R_D} \cdot \Delta f$$

- Thermal Channel Noise and 1/f Noise: ID

$$\overline{i_{nD}^2} = \frac{8}{3} \cdot k \cdot T \cdot g_m \cdot \Delta f + KF \cdot \frac{I_D^{AF}}{f} \cdot \Delta f$$

Noise Models in Simulators



MOSFETs

1/f noise contributions

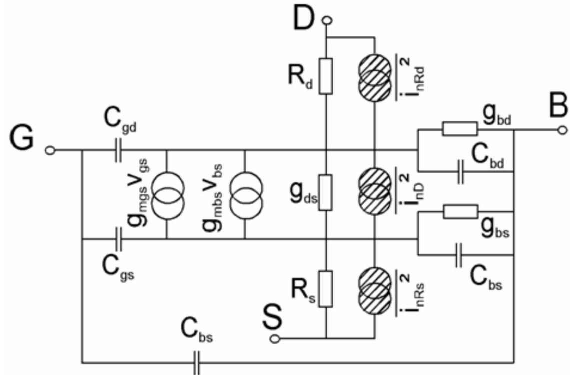
- Thermal Noise: RD, RS

$$\overline{i_{RD}}^2 = 4 \cdot k \cdot T \cdot \frac{1}{R_D} \cdot \Delta f$$

- Thermal Channel Noise and 1/f Noise: ID

$$\overline{i_{nD}}^2 = \frac{8}{3} \cdot k \cdot T \cdot g_m \cdot \Delta f + KF \cdot \frac{I_D{}^{AF}}{f \cdot C_{ox} \cdot L_{eff}^2} \cdot \Delta f$$

Noise Models in Simulators



MOSFETs: BSIM3v3.2

1/f Noise and Channel Noise:
conventional, like in MOS3:

$$\overline{i_{1/f}}^2 = KF \cdot \frac{I_D{}^{AF}}{f^{EF} \cdot C_{ox} \cdot L_{eff}{}^2} \cdot \Delta f$$

$$\overline{i_{th}}^2 = \frac{8}{3} \cdot k \cdot T \cdot (g_m + g_{ds} + g_{mb}) \cdot \Delta f$$

BSIM3v3 specific:

$$\overline{i_{1/f}}^2 = f(NOIA, NOIB, NOIC, EF, EM)$$

$$\overline{i_{th}}^2 = \frac{4 \cdot k \cdot T \cdot \mu_{eff}}{L_{eff}{}^2} \cdot |Q_{inv}| \cdot \Delta f$$

Model Choice:
Parameter ***noimod*** selects a combination of the four modeling possibilities:

NOIMOD		1/f Model conventional	1/f Model BSIM3v3 specific
channel noise model	conventional	1	3
channel noise model	BSIM3v3 specific	4	2

The 1/f Noise and its Model Parameters

The thermal and shot noise is determined inside the model from the other model variable values.

NOTE: This implies that only physical meaningful values for the resistor and DC large signal parameters will finally give an accurate modeling of these types of noise sources.

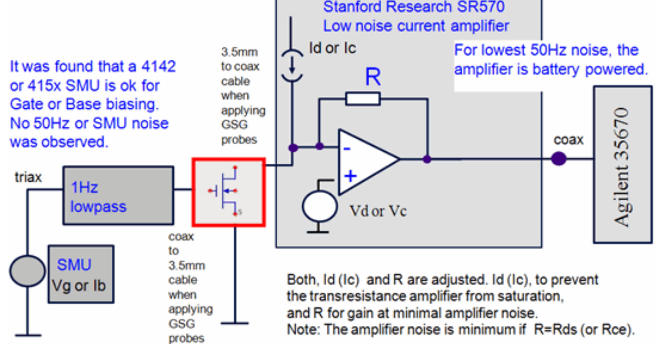
This means, only the 1/f noise can be modeled individually.
The modeling of the 1/f Noise follows for:

- All Diode Models,
- Bipolar Models like the Gummel-Poon, VBIC or MEXTRAM,
- JFET Models,
- MOSFET Models
- BSIM3v3/BSIM4 when selecting the conventional Noise 1/f-Model

Extraction Method for the Parameters AF, KF and B is required.

1/f Noise Measurement Toolkit in IC-CAP

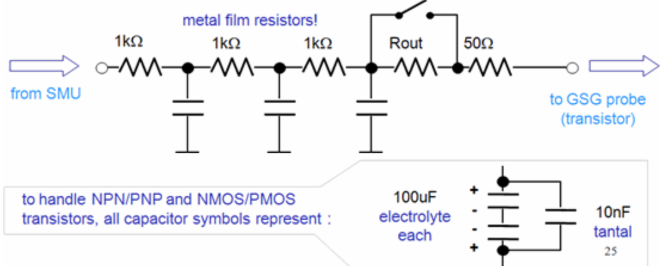
Applied 1/f Measurement Setup



1Hz Lowpass Filter

The filter is required for biasing the transistor input.
It offers -60dB attenuation at 50Hz.

Required Output Impedance Rout (to avoid shortening the 1/f transistor noise):- Rout =330Ω k for bipolar (in any case Rout > Rbe_noise_bias),- 50Ω for MOS transistors



Agilent 4142B or 415xB
Operated as a power supply with two special macros in IC-CAP:

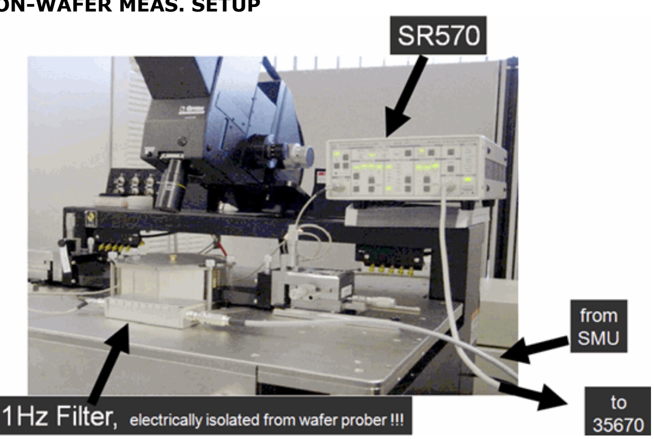
- Voltage or current ON
- Performing the 1/f measurement, typ. 30 sec., with dyn. signal analyzer
- Voltage or current OFF

Stanford Research SR570 Current Preamplifier
1 pA/V maximum gain
Adjustable bias voltage with test point +- 5V max
Variable input offset current +-1pA .. +-5mA
Low Noise, High BW, and Low Drift modes 1 MHz maximum bandwidth

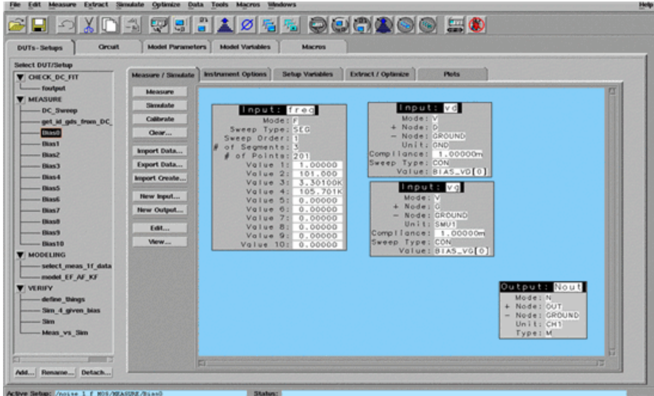
Agilent 35670A
The 35670A is a portable two- or four-channel dynamic signal analyzer

- 102.4 kHz Bandwidth
- Measurements: Linear Spectrum, Cross Spectrum, Power Spectrum, Coherence, Power, Spectral Density, Frequency Response, Time Waveform, Auto-correlation, Cross-Correlation, Histogram etc.

ON-WAFER MEAS. SETUP

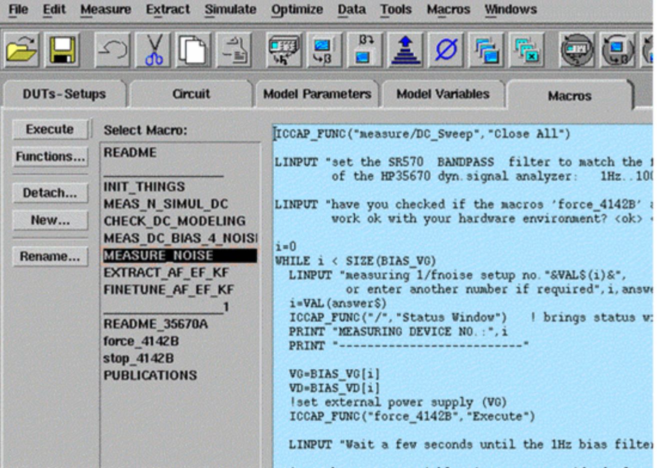


IC-CAP Toolkit for Bipolar and MOS

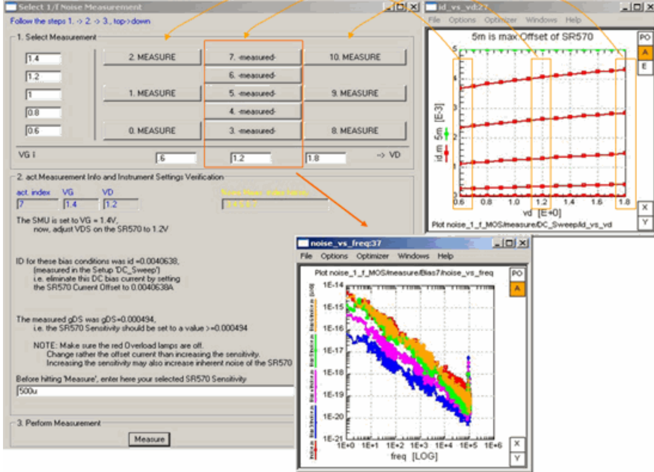


The slide above and the following IC-CAP screenshot slides describes the 1/f bipolar and MOS transistor toolkit before its implementation into IC-CAP as Agilent product number 85195B.

Ready2go, open measurements and extraction routines (user modifiable)



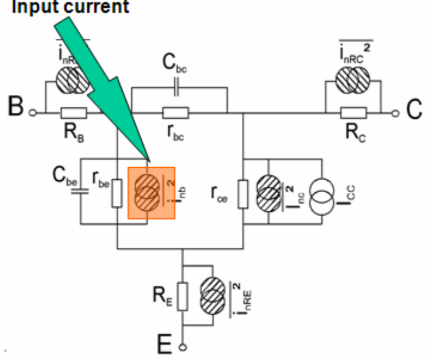
Intuitive User Interface



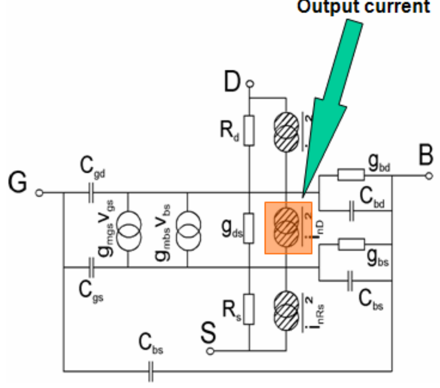
By default, 5 DC bias points at mean_vd are measured for 1/f noise, plus 3 a low and 3 at high vd for verification.
(can be changed by user if required).

1/f Parameter Extraction

Dominant 1/f noise source location:
BIPOLAR input current Ib



MOS output current Id



Bipolar Transistors vs. MOS Transistors

Since the 1/f modeling formula refers to ib for bipolar transistors, a separate measurement of ß = Ic / Ib is required in order to transform the noise, measured at the Collector, to the Base.

$$\overline{i_{1/f}^{\,2}} = KF \cdot \frac{I_D^{\,AF}}{f^{EF} \cdot C_{ox} \cdot L_{eff}^{\,2}} \cdot \Delta f$$

MOS

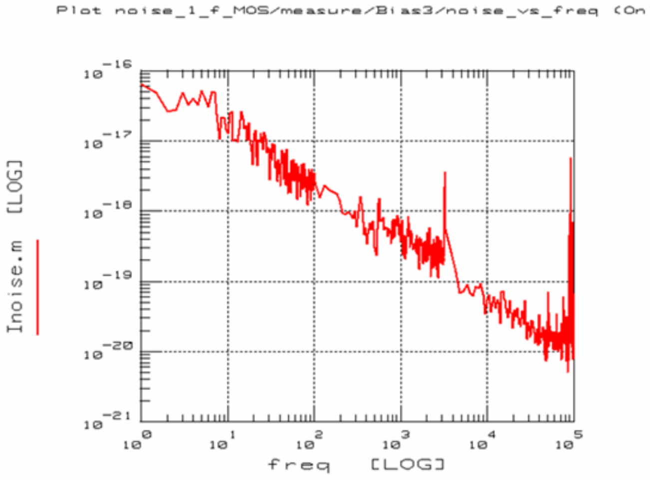
$$\overline{i_{nb}^{\,2}} = KF \cdot \frac{I_b^{\,AF}}{f} \cdot \Delta f$$

Bipolar

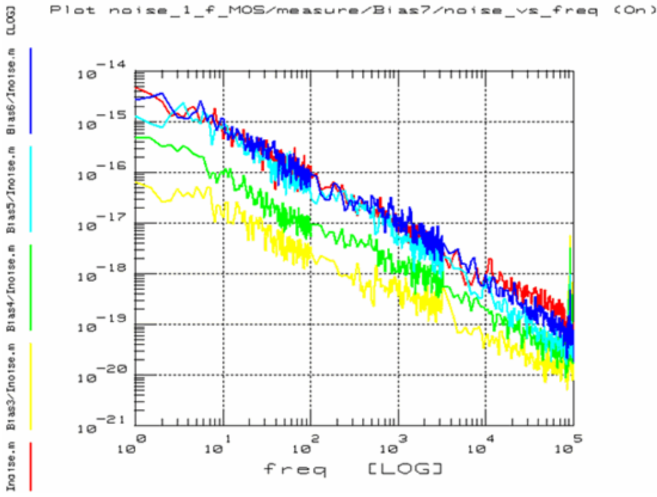
Examples

MOS Transistor

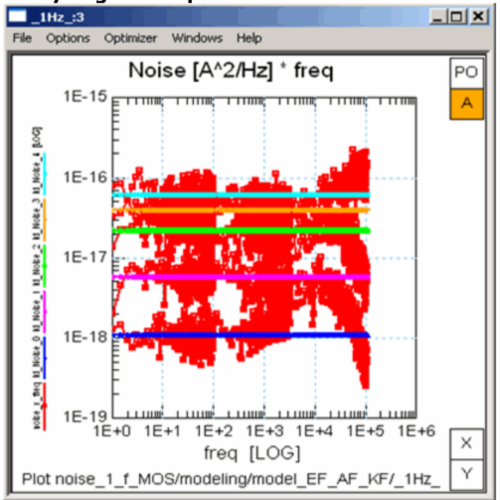
1st DC condition



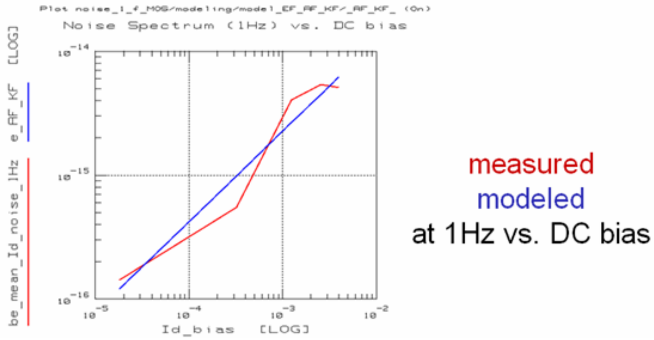
All DC conditions



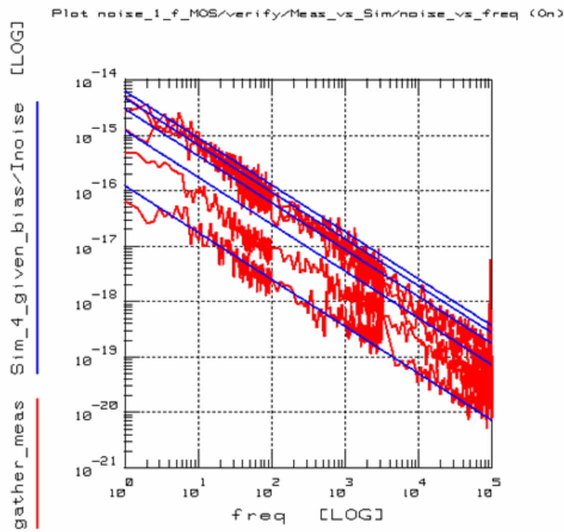
Analyzing the slopes



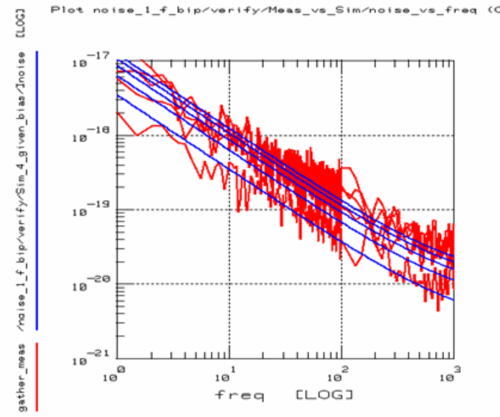
Measured and modeled 1Hz points



Fitted model

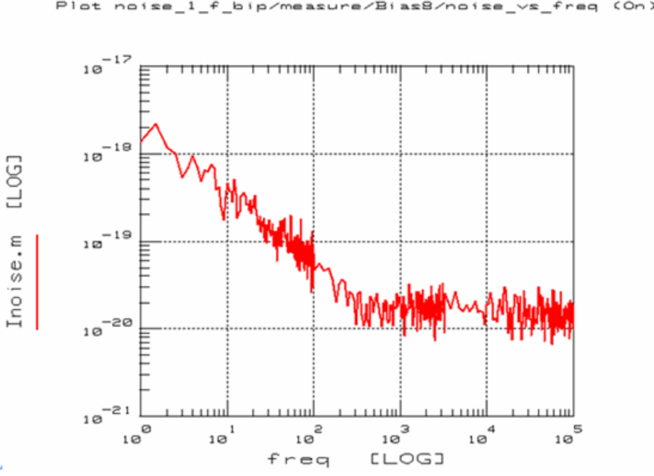


1/f modeling at a glance: multiply Noise Spectrum by freq

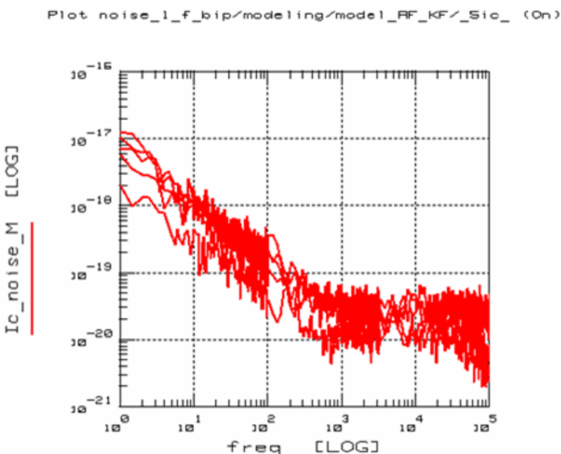


Bipolar transistor

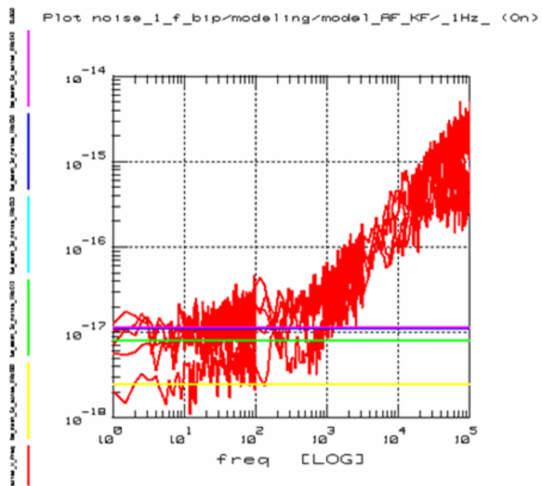
1st DC condition



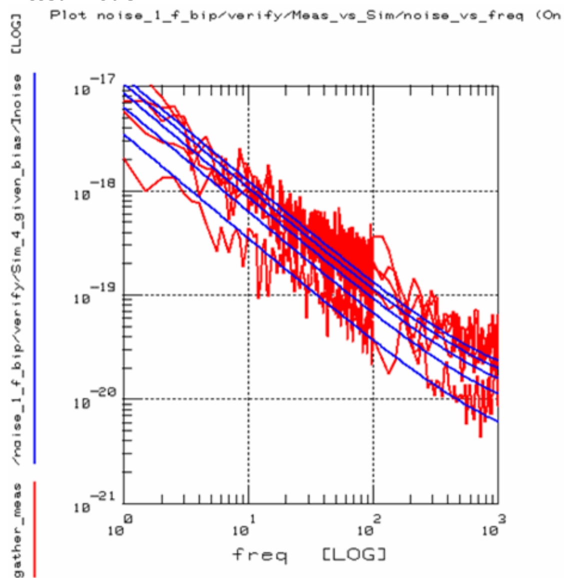
All DC conditions



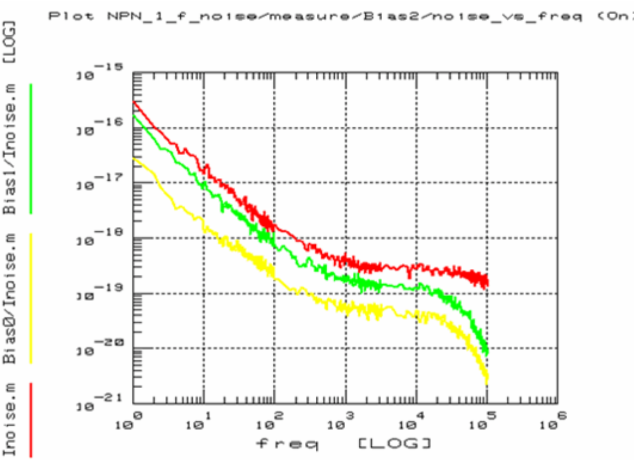
Analyzing slopes



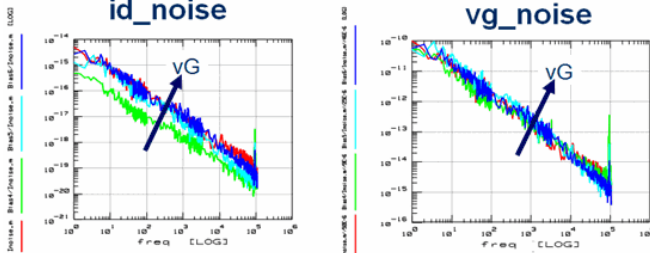
Fitted model



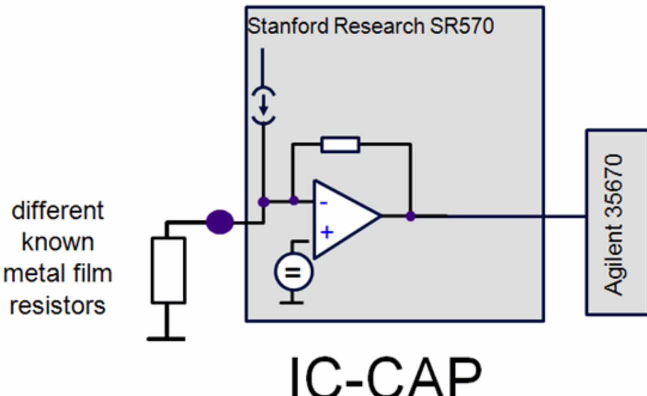
Bipolar measurement result with 50 averages (instead only 10) of the dyn.signal analyzer



Verification Issues
A note on data consistency checks for MOS transistors.
For MOS-Transistors, vg_noise is independent of the applied bias. Therefore, calculating vg_noise=id_noise/gm^2 should give overlying curves .



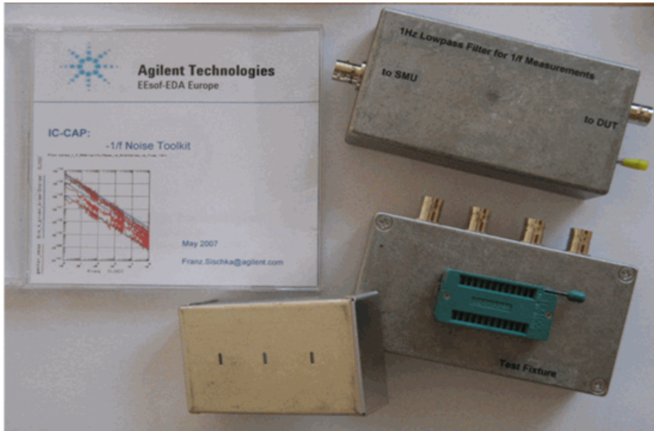
*See: M.T.Yang et.al., Characterization and Model of On-Chip Flicker Noise With Deep N-Well Isolation for 130nm and Beyond,ICMTS 2005, Leuven, Conference Proceedings



Noise measurements of different metal film resistors (low 1/f contribution), were executed and compared with noise calculations. Measurements agreed very well with calculations.

Conclusions
Modeling of 1/f noise is important for accurate design of tomorrow's communication systems.
A precise modeling strategy for most transistor models, using AF and KF, has been demonstrated.
A ready2go IC-CAP toolkit is available.

Toolkit Parts incl. TestFixture for verification with packaged standard device


Agilent Technologies
EEsof-EDA Europe
IC-CAP:
-1/f Noise Toolkit
May 2007
Franz.Sischka@agilent.com
1Hz Lowpass Filter for 1/f Measurements
to SMU
to DUT
Test Fixture

AF, KF and BF/EF Noise Parameter Extraction and Verification

Considering the parameter extraction, we refer to the method introduced in [13, see also 14]. Following this approach, there is no need to exactly determine the corner frequency fc like with other proposed methods. For the corner frequency fc, the 1/f noise equals the white noise. For frequencies below that, the 1/f noise is dominant, allowing to neglect all other noise sources for the further analysis. The current noise source of the transistor can now be calculated out of the measured noise at the output of the trans-resistance amplifier.

Bipolar Transistors:

For the bipolar transistor models, the origin of the 1/f noise is the Base region, see equation (17). However, the effective 1/f current noise spectral density $[A^2/Hz]$ is measured at the Collector of the transistor. Therefore, the 1/f noise at the Base has to be calculated first after

$$S_{iB} = \frac{1}{\beta^2_{(I_{B_DC})}} \cdot S_{iC} \quad \left[\frac{A^2}{Hz}\right] \tag{36}$$

To begin with the parameter extraction, we first repeat the formula of the 1/f effective noise current generated at the Base (equ.17)

$$\overline{i_{nB\ 1/f}^2} = KF \cdot \frac{I_{B_DC}^{AF}}{f} \cdot \Delta f \quad \left[A^2\right] \tag{37}$$

Note: the VBIC model features an additional 1/f noise parameter, BF, see equ.(20). It acts like the EF parameter in the BSIM3v3 model: to fit the -10dB/decade slope of the measured 1/f noise. For details about its extraction, please refer to the next section on MOS transistors.

In order to match equ.(37) to equ.(36), we normalize to Δf and set Δf = 1Hz. This gives the Base current noise spectral density.

$$S_{iB} = \frac{\overline{i_{nB\ 1/f}^2}}{1Hz} = KF \cdot \frac{I_{B_DC}^{AF}}{f} \quad \left[\frac{A^2}{Hz}\right] \tag{38}$$

Since the 1/f slope is a 'given' for our actual modeling problem, our next step is to get rid of it by multiplying the measured curve with the frequency points 'f'. This results in a flat trace where we had the 1/f slope before.

$$S_{iB} \cdot f = KF \cdot I_{B_DC}^{AF} \tag{39}$$

The advantage of this method is that we are now easily able to identify the value of the 1/f noise at 1Hz, what will be used in the next step. The 1Hz noise value is simply calculated as the mean value of a maximum flat sub-range of these so transformed data. Considering the extrapolated measurement result at 1Hz, the above formula simplifies to

$$S_{iB@1Hz} = KF \cdot I_{B_DC}^{AF} \tag{40}$$

This means, we are now ready to obtain an 1Hz value of our 1/f noise for each bias condition I_{B_DC}.

In the next step, we draw these values against the bias current. We apply a logarithmic conversion to the above formula and obtain

$$\log_{10}\left(S_{iB@1Hz}\right) = \log_{10}(KF) + AF \cdot \log_{10}\left(I_{B_DC}\right) \tag{41}$$

what can be interpreted as a linear function like

$$y = a + b \cdot x \tag{42}$$

where

$$y = \log_{10}\left(S_{iB@1Hz}\right)$$

$$a = \log_{10}(KF)$$

$$b = AF$$

and the DC bias current at the Base is transformed by $x = \log_{10}\left(I_{B_DC}\right)$

A linear regression is applied, which returns the y-intersect 'a' and the slope 'b' of a best fitting line for equ.(42).

The noise parameters AF and KF are then calculated after

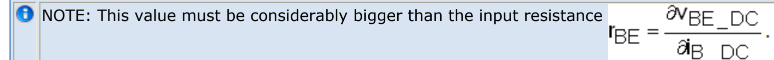
$$AF = b \tag{43}$$

and

$$KF = 10^a \tag{44}$$

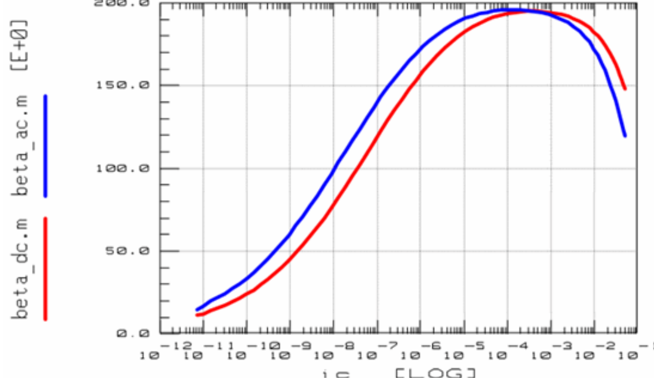
Step-by-step modeling procedure for bipolar transistors

1. For a first DC bias condition (e.g. iB_DC = 1uA, vCE = 2V), the 1/f noise spectral density is measured. As stated above, for the bipolar transistor, the 1Hz Base filter's output impedance is set to a high value, e.g. 330kΩ

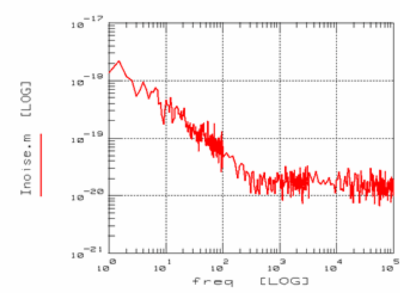


NOTE: This value must be considerably bigger than the input resistance $r_{BE} = \frac{\partial v_{BE_DC}}{\partial i_{B_DC}}$.

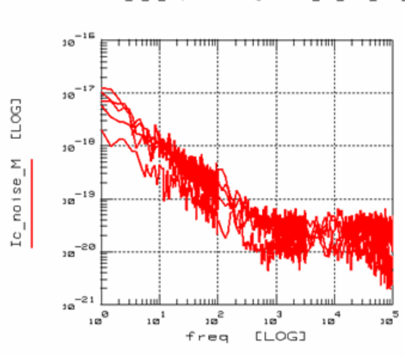
Otherwise, it would shorten the 1/f noise current source i_{nB}^2 at the Base.



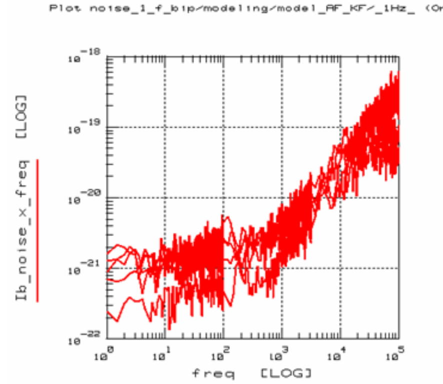
The following plot shows the measured noise current at the Collector:



2. This is repeated for e.g. 5 more different Base currents, but the same vCE. And we get five 1/f noise curves for each iB_DC bias condition:



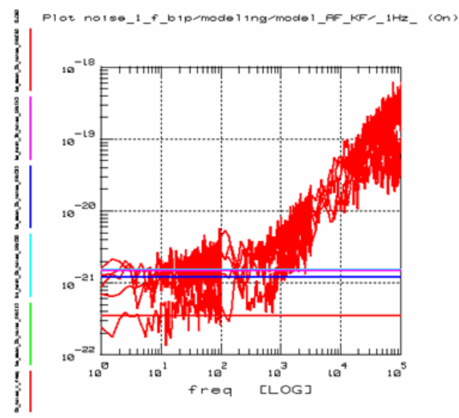
3. We have so far considered the noise current data measured at the Collector of the transistor. The 1/f noise source of a bipolar transistor is, however, located and modeled in the Base region. Therefore, we have to divide the above obtained Collector current noise spectral density SiC by beta2, and obtain SiB, see equ.(36) above.
4. After we have obtained SiB at the Base, we multiply it by 'freq' , see (39), and obtain:



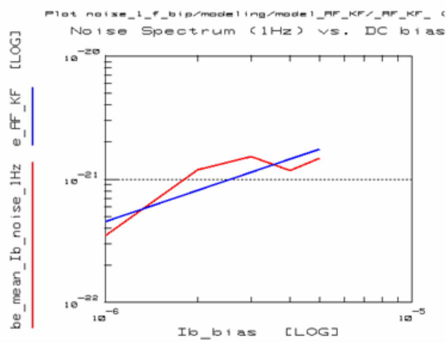
Thus, the 1/f noise appears as flat traces and can easily be identified.

5. We calculate the mean value of each transformed 1/f curve for each Base bias

condition from the max. flat traces. These mean values represent the 1Hz values of the 1/f current noise spectral density, as a function of the iB_DC bias.



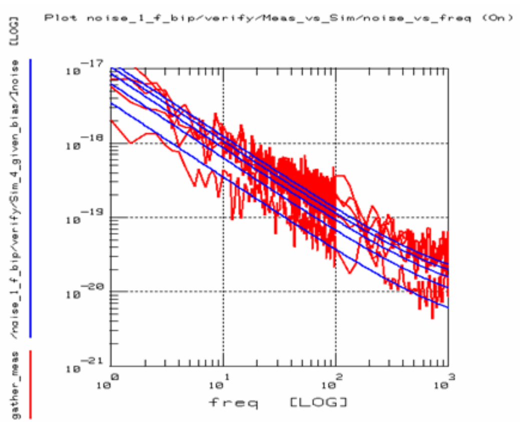
6. Finally, we are ready to draw the 1Hz Base noise data points against the DC bias i_{B_DC}, and fit a line to these data points, see equations (41) and (42).



From the y-intersection and the slope of the fitted curve, we calculate AF and KF after equations (43) and (44).

7. After the model parameters have been obtained, the simulation result of the Collector current noise spectral density is compared with the original measured data, and the AF and KF model parameters are fine-tuned.

Measured and simulated Collector noise spectral density $[A^2/Hz]$ for the bipolar transistor.



MOS Transistors

The 1/f noise description for MOS transistors in this chapter refers to the AF, KF, EF formulation. For details about how to model the 1/f noise in BSIM3v3 using the parameters NOIA, NOIB etc., see publication [15].
What makes the modeling easier compared to the bipolar case, is that the dominant 1/f noise source is located at the output of the transistor, i.e. the Drain-Source channel. On the other hand, the transistor Gate impedance is high. Related to the measurement setup and our 1Hz filter at the Gate, this corresponds to a low output impedance. We choose 50Ω.
On the other hand, a 3rd model parameter is used to reflect slight differences from the theoretical 1/f slope: the parameter EF.
Also, different to the bipolar case, the 1/f noise formulation of the BSIM3v3 model includes some geometry parameters.
As with the bipolar case, we first refer to the formula for the 1/f Drain-Source effective

noise current as

$$\overline{i_{nD\ 1/f}}^2 = KF \cdot \frac{I_{D_DC}^{AF}}{f^{EF} \cdot COX \cdot Leff^2} \cdot \Delta f \qquad (45)$$

with

$$COX = \frac{\varepsilon_0\,\varepsilon_{Si}}{TOX} = \frac{3.45E-11}{TOX}$$

$$Leff = L - 2 \cdot \left(LINT + \frac{LL}{L^{LLN}} + \frac{LW}{W^{LWN}} + \frac{LWL}{L^{LLN} \cdot W^{LWN}} \right)$$

or simplified: $Leff = L - 2 \cdot LINT$

Again, we normalize to Δf, then set Δf = 1Hz and obtain the Drain-Source current noise spectral density in $[A^2/Hz]$.

$$S_{iD} = \frac{\overline{i_{nD\ 1/f}}^2}{1Hz} = KF \cdot \frac{I_{D_DC}^{AF}}{f^{EF} \cdot COX \cdot Leff^2} \qquad (46)$$

As the first modeling step, we extract the parameter EF, the 1/f slope correction. A log conversion of equation (45) gives:

$$\log_{10}(S_{iD}) = const. - EF \cdot \log_{10}(f) \qquad (47)$$

i.e. we convert both the measured noise data and the frequency points logarithmically and apply a regression curve fitting. The parameter EF is then the '-slope' of the fitted line.
Since the $1/f^{EF}$ slope is now already modeled, we can get rid of it by multiplying the measured curve with the frequency points 'f^{EF}'. This results in a flat trace where we had the $1/f^{EF}$ slope in the measurements before.

$$S_{iD} \cdot f^{EF} = KF \cdot \frac{I_{D_DC}^{AF}}{COX \cdot Leff^2} \qquad (48)$$

After this step, we are again easily able to identify the value of the $1/f^{EF}$ noise at 1Hz: it is simply calculated as the mean value of a maximum flat sub-range of these so transformed data.
Considering the extrapolated measurement result at 1Hz, the above formula simplifies to

$$S_{iD@1Hz} = KF \cdot \frac{I_{D_DC}^{AF}}{COX \cdot Leff^2} \qquad (49)$$

This means, we are now ready to obtain an 1Hz value of our $1/f^{EF}$ noise for each bias condition i_{D_DC}.
In the next step, we draw these values against the i_{D_DC} bias current. We apply a logarithmic conversion to the above formula and obtain

$$\log_{10}(S_{iD@1Hz}) = \log_{10}\left(\frac{KF}{COX \cdot Leff^2}\right) + AF \cdot \log_{10}(I_{D_DC}) \qquad (50)$$

what can be interpreted as a linear function like

$$y = a + b \cdot x \qquad (51)$$

where

$$y = \log_{10}(S_{iD@1Hz})$$

$$a = \log_{10}\left(\frac{KF}{COX \cdot Leff^2}\right)$$

$$b = AF$$

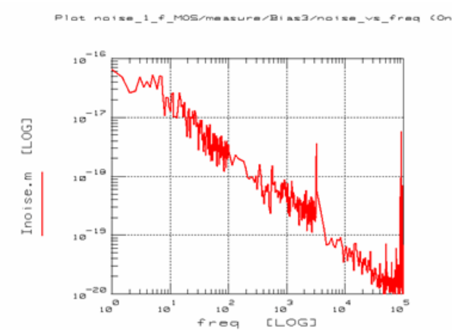
and the DC bias current at the Base is transformed by $x = \log_{10}(I_{D_DC})$
A linear regression is applied, which returns the y-intersect 'a' and the slope 'b' of a best fitting line.
The noise parameters AF and KF are then calculated after

$$AF = b \qquad (52)$$

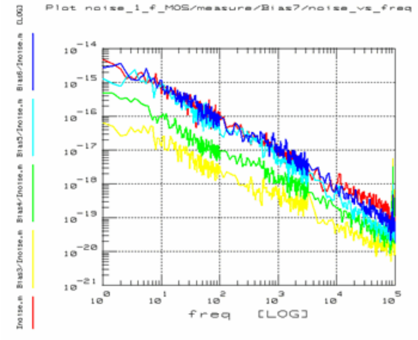
and $KF = COX \cdot Leff^2 \cdot 10^a$ (53)

Step-by-step modeling procedure for BSIM3v3 CMOS transistors

1. For a first DC bias condition (e.g. vG = 0.6V, vDS = 1V), the 1/f Drain current noise spectral density $[A^2/Hz]$ is measured. The 1Hz Base filter's output impedance is set to 50Ω. The following plot shows the measurement result:

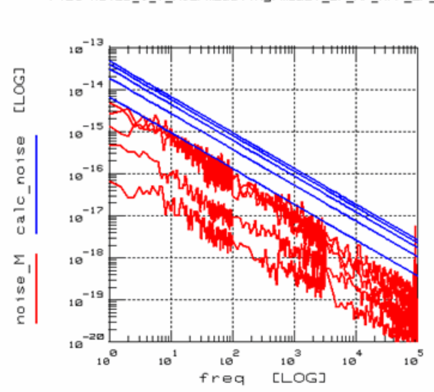


2. This is repeated for e.g. 5 more different Gate voltages, but the same vDS. And we get five 1/f noise curves for each vG DC bias condition:

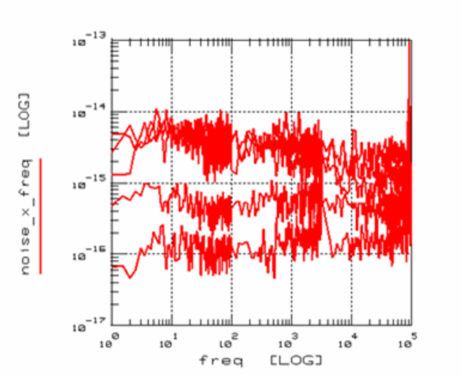


3. In the next step, we extract the EF parameter after (47), and check if the slopes of the intermediate simulations match.

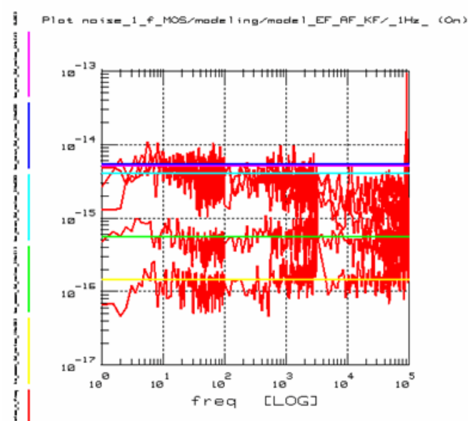
> Note: the AF and KF have not yet been extracted, therefore only the slopes are important to compare.



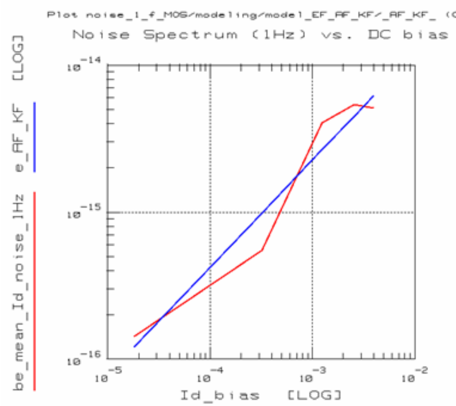
4. Now, we are ready to multiply by 'f^{EF}' in order to easier extract the 1Hz value of the noise.



5. We calculate the mean value of each transformed 1/f curve for each vG bias condition from the max. flat regions. These mean values represent the 1Hz values of the 1/f Drain current noise spectral density, as a function of the iD_DC bias.



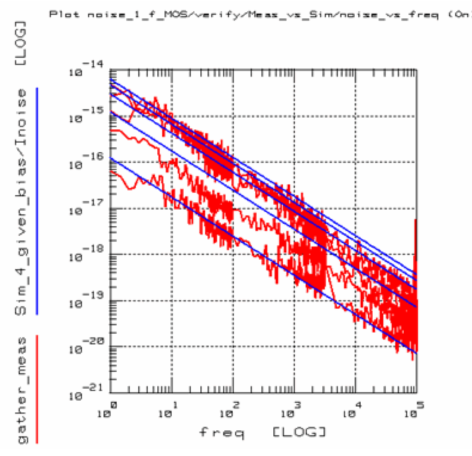
6. Finally, we are ready to draw the 1Hz Drain current noise data points against the DC bias current iD_DC, and to fit a line to these data points, see equations (50) and (51).



From the y-intersection and the slope of the fitted curve, we calculate AF and KF after equations (52) and (53)

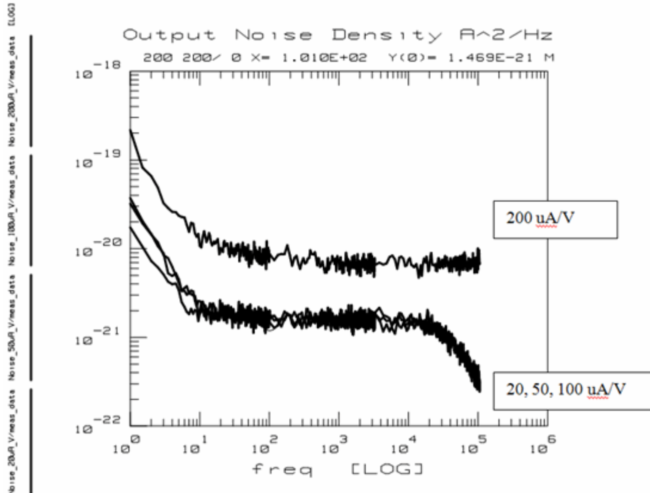
7. After the model parameters have been obtained, the simulation result is compared with the original measured data, and the AF, KF and EF model parameters are fine-tuned.

Measured and simulated Drain current noise spectral density $[A^2/Hz]$ for the MOS transistor



> The noise data returned from hspice and spice3 differs from the noise data returned from hpeesofsim, mns, and spectre.
> hspice and spice3 have units of V^2/Hz and the other simulators have V/Sqrt(Hz).

Noise floor information on the toolkit measurement setup



The figure above shows the current noise density of the system measured at the device output (at the low noise current amplifier LNA).
The noise is expressed in A^2/Hz and varies with the LNA sensitivity. The figure shows the noise floor for the 4 most commonly used values of the LNA sensitivity: 20 uA/V, 50 uA/V, 100 uA/V and 200 uA/V.

Comments:

- The 1/f noise observed at the beginning of the trace is due to the internal 1/f noise of the LNA.
- The noise drop observed at low sensitivity values (high gain) is due to the bandwidth limitation of the LNA. Note that this is not in the frequency band used for the extraction (typically between 10 Hz and 1 kHz). Using a sensitivity of 200 uA/V or greater does not have this limitation but on the other hand, increases the noise floor.
- This noise floor should be compared to the output current noise density of the DUT . If the DUT is a CMOS, this would be the drain current noise density. If the DUT is a bipolar, this should be the collector current noise density. Note that in the bipolar case, the current noise source is actually modeled at the input (base current). The Ib noise density is determined by dividing the Ic noise density by the DUT current gain (squared).

Parts list for 1/f noise toolkit

Instruments:

- 1 Wafer Probe Station with GSG probes (alternatively shielded DC probes)
- 1 Agilent DC Analyzer 414x or 415x (input biasing for transistor)
- 1 Custom made Lowpass Filter (1Hz), input triax, output coax or 3.5mm connector following the recommended circuit description
- 1 Stanford Research Amp SR570
- 1 35670A Dynamic Signal Analyzer
- 1 Multimeter for controlling the DC bias provided by the SR570

Cables:

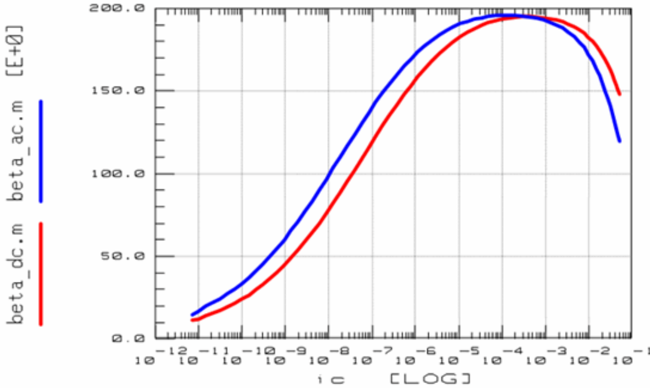
- 1 Cable: filter output (coax or 3.5mm) to GSG probe (3.5mm)
- 1 Cable: GSG probe (3.5mm) to SR570 (coax)
- 1 Cable: SR570 (coax) to 35670A (coax)

IC-CAP:

- 85199A (framework), 85199B(analysis module: simulator interface)
- 85199D (DC analyzer driver)
- 85199G (35670 driver),
- 85195B (1/f noise toolkit for AF, KF, EF parameter extraction for NPN-BJTs and MOS)

For PNP/PMOS transistors, the user has to make the changes related to the inverse voltages by himself.

For bipolar transistors, the noise measured at the collector has to be transformed to the base: i.e. dividing by ß.
This ß, however, is the AC-beta d(ic)/d(ib), and not the DC-beta ic/ib.



Explanation of Spice circuit deck for 1/f noise simulations

SPICE Circuit Deck in the DUT Test Circuit:

```
.subckt _noise 1=C 2=B 5=L 99=AUX
*satisfy SPICE for a noise simulation
Raux 99 0 1
*emulate the bias lowpass filter at the Base/Gate
RB1  2 21 10k
CB1 21  0 100u
RB2 21 22 10k
CB2 22  0 100u
*here is the filter output resistor
* -high value for bipolar transistors: bigger than biggest rBE
* - 50 Ohm for MOS
RB3 22 23 330k
*here comes the call to the transistor subcircuit
X1 11 23 0 noise_1_f_bip
*add a dummy 0V source to sense the transistor noise current
V1    1 11 0
*convert the Collector noise current into an identical voltage
H1     5 0 V1 1
Rspice 5 0 1
.ends
```

The netlist implements a workaround for probing the current noise density flowing through the collector node because the simulator cannot output current noise densities directly. Here is the detailed explanation of the circuit:

- The external nodes of the circuit are 1=C, 2=B, 5=L , 99=AUX
- The AUX=99 node is just an auxiliary nodes used only to run an AC simulation. In the IC-CAP setup verify/simulate, look at the input sweeps: the node AUX is connected to an AC generator. We need to run an AC simulation since we want to get noise data of the device, on the other hand we do not want to apply any AC signal to the device so we use this dummy node connected to a 1 Ohm resistor. Note that the value of the resistor is not relevant.
- The line: V1 1 11 0 connects a voltage generator between node 1 and 11 with zero voltage. In the spice syntax this acts as a current probe. Note that we want the noise contribution at the Collector as a current.
- In order to get the noise part of the current through V1, we setup another independent sub-circuit.
  The line: H1 5 0 V1 1 represents a current-controlled-voltage-generator. Therefore the voltage between node 5 and 0 is controlled by the current through V1, multiplied by '1'.
- The resistance called Rspice is connected between node 5 and 0. Note that the value of this resistance is not relevant.
- Finally, in the IC-CAP setup, we output the noise at node 5, which in turns represents the noise of the current flowing in the collector.

The syntax for ADS:

```
define ads_noise (C B AUX)
; note: ADS circuit syntax is case sensitive!!!!!
R:RB1  B 21 R=10k    Noise=no
C:CB1 21  0 C=100u
R:RB2 21 22 R=10k    Noise=no
C:CB2 22  0 C=100u
R:RB3 22 23 R=1MEG   Noise=no
;here comes the transistor
noise_1_f_bip:X1 11 23 0
;add a dummy resistor to sense the noise current  (current-controlled voltage source in MDS cannot
be controlled by a V=0 source)
R:Rdummy C 11 R=1m   Noise=no
;convert the Collector noise current into an identical voltage
cdvs:Haux 11 C AUX 0 M=1 ANG=0 T=0
R:Raux   AUX 0 R=1k   Noise=no
end ads_noise
;define a current dependent voltage source
```

```
#echo define cdvs(senssnk senssrc pos neg)
#echo   parameters M=1 ANG=0 T=0
#echo   gain=M*polar(1,-ANG)*exp(-j*omega*T)
#echo   Z_Port:v senssnk senssrc pos neg Z[2,1]=gain
#echo end
```

Please note that you have to calculate the square of the ADS noise in order to get the spectral density as calculated by SPICE. In the demo_features directory of IC-CAP, see the file
6_SIMULATORS\0_general\noise_1_f_using_Spice_MDS_ADS.mdl

For the **Gummel-Poon model**, there is a special GP BJT model parameter 'rbnoi' , describing the effective base noise resistance. This parameter is available in Spectre, and also available in ADS. Below are the governing (thermal) Rb noise equations from the Spectre (4.4.6) and ADS (2002C) documentation showing identical implementation (Default for ADS Rbnoi=Rb)
Spectre:

if RBNOI is specified: $$\overline{i_{RB}^2} = \frac{4kT}{RB} \cdot \frac{RBNOI}{RB} \Delta f$$

otherwise: $$\overline{i_{RB}^2} = \frac{4kT}{RB} \Delta f$$

ADS: $$\overline{i_{RB}^2} = \frac{4kT}{RB} \cdot \frac{RBNOI}{RB} \Delta f$$

Rbnoi was added only to the Gummel-Poon model by Cadence, so it was added to ADS's Gummel-Poon model too for compatibility. VBIC, Hicum and Mextram have more sophisticated models for the base resistance, so this parameter wasn't added to these other models by Cadence, Agilent or anyone else.

Acknowledgments and Publications
I would like to thank specially F. X. Sinnesbichler of the Technical University Munich, Institute of High Frequency Techniques, for having laid the base for this manual. A big part of the chapters stem from this work. I would also especially acknowledge Mr.Sinnesbichler's IC-CAP workshop about Noise modeling in Munich, July 1998.

In the same way, I have to acknowledge Mr.Knoblinger of Infineon Technologies AG in Munich, for his hint on using the special low-noise current amplifier as an excellent and straight-forward way to measure 1/f noise accurately and without big influences from power line frequency harmonics.

For detailed verification tests and measurements of the proposed measurement setup, I'd like to mention A.Blaum, O.Pilloud, G.Scalea, J.Victory, Motorola, Geneva

For important discussions on the noise terminology, I'd like to mention Mr. Berkner, also Infineon Technologies AG in Munich.

For detailed verification tests and measurements of the proposed measurement setup I am grateful to A.Blaum, O.Pilloud, G.Scalea, J.Victory, Motorola, Geneva

Web Info
Stanford Research Web Page: http://www.thinksrs.com//

Publications
[1] R.Müller, „Rauschen", Springer-Verlag, Berlin, 1990.
[2] B.Schiek, H.-J.Siweris, „Rauschen in Hochfrequenzschaltungen", Hüthig Buch Verlag, Heidelberg, 1990.
[3] G. Massobrio, P. Antognietti, „Semiconductor Device Modeling with SPICE", McGraw-Hill, New York, 1993.
[4] F. N. Hooge, „1/f Noise is No Surface Effect", Physics Letters, Vol. 29A, Nr. 3, pp. 139-194, April 1969.
[5] F. N. Hooge, „1/f Noise Sources", IEEE Transactions on Electron Devices, Vol. 41, Nr.11, pp. 1926-1935, Nov. 1994.
[6] C. McAndrew at al, „VBIC95: An Improved Vertical, IC Bipolar Transistor Model", Proceedings of the 1995 BiCMOS Circuits and Technology Meeting, , Minneapolis, pp. 170-177, 1995
[7] C. McAndrew et al, „VBIC95, The Vertical Bipolar Inter-Company Model", IEEE Journal of Solid-State Circuits, vol. 31, No. 10, pp. 1476-1483, October 1996.
[8] F. X. Sinnesbichler, G. R. Olbrich, „VBIC - The Vertical Bipolar Inter-Company Model. Ein Überblick über das Modell und über zugehörige Parameterextraktionen", Hewlett-Packard IC-CAP Workshop-Reihe 1997/98, München, Januar 1998.
[9] H. C. deGraaff, F. M. Klaassen, „Compact Transistor Modelling for Circuit Design", Springer-Verlag, Wien, 1990.
[10] Y. Cheng et al, „BSIM3v3 Manual", University of California, Berkeley, 1996.
[11] C.C. McAndrew, „Practical Modeling for Circuit Simulation", IEEE Journal of Solid-State Circuits, vol. 33, no. 3, pp. 439-448, March 1998.
[12] A. Cappy, „Noise Modeling and Measurement Techniques", IEEE Transactions on Microwave Theory and Techniques, vol. 36, no. 1pp., 1-10, Jan. 1988.
[13] F. X. Sinnesbichler, M. Fischer, G. R. Olbrich, „Accurate Extraction Method for 1/f-Noise Parameters Used in Gummel-Poon Type Bipolar Junction Transistor Models", IEEE MTT-S Symposium, Baltimore, 1998, pp. 1345-1348.
[14] J. C. Costa, D. Ngo, R. Jackson, N. Camilleri, J. Jaffee, „Extraction of 1/f Noise

Coefficients for BJTs", IEEE Transactions on Electron Devices, vol. 40, no. 11, pp.1992-1999, Nov. 1994.
[15] J.C.Vildeuil, M.Valenza, D.Rigaud, Universite de MontpellierII, France, CMOS 1/f Noise Modeling and Extraction of BSIM3 Parameters Using a New Extraction Procedure ICMTS1999 Conference Gothenburg, March 15-18,1999, ISBN 0-7803-5270-X
[16] C.G.Jakobson, I.Bloom, Y.Nemirovsky, 1/f Noise in CMOS Transistors for Analog Applications, 19. Convention of Electrical and Electronics Engineers in Israel, 1996, pp.557-560
[17] A.Blaum, O.Pilloud, G.Scalea, J.Victory, F.Sischka, A New Robust On-Wafer 1/f Noise Measurement and Characterization System, ICMTS conference 2001,
Kobe, Japan, March 19-22,2001

The use of the SR560 low-noise voltage amplifier (lower system noise, but no internal DC bias available) is published in:
Hardy et al., Low-Frequency Noise in Proton Damaged LDD MOSFET's,
IEEE Trans. on el. devices, vol.46, no.7, July 1999, P.1341

Thermal Modeling

Note: See also the chapter on Pulsed Measurements in the Measurement chapter.

When specifying the thermal model parameter RTH, make sure to set CTH to a non-zero value. Otherwise, your device is 'thermally faster' than electrically i.e. if you simulate pulses, you do not get the expected fast risetime of the electrical performance, followed by the slower thermal time constant.
If you are not sure about your CTH value, set it to a big value, e.g. several milli-Farad.

When simulating the temperature dependence of circuits, two global simulator variables come into play:
TEMP: the temperature for the simulation result
TNOM: the temperature at which the device was measured and modeled.

This assumes that there is a constant simulation temperature, and no dependence of the simulation result from self-heating effects.

This can cause a major problem to modeling, because as a general rule, we have to account for self-heating effects at currents above ~10mA (for typical vmax = 5V). While this is true when measuring packaged devices, for on-chip measurements using a thermochuck, the critical current is about 50mA.

Therefore, for modeling, we can consider three cases:

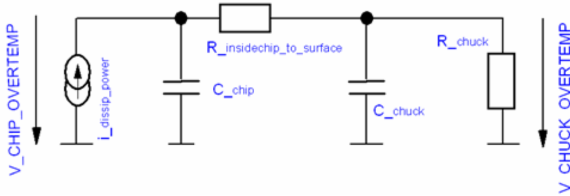
- Stay well below a few mA for all device characterization measurements
- For currents above some mA, measure isothermally (pulsed), and avoid self-heating,
- Measure both, pulsed and CW at the higher currents, change the device models and enhance them for self-heating.

Modeling Procedure

For the last case, the modeling procedure is as follows:

1. Perform a full device measurement characterization at e.g. 25°C in pulsed mode
2. Extract all temperature-independent model parameters (conventionally)
3. Perform another pulsed measurement at the other temperatures, e.g. -25°C and +75°C
4. Extract now the temperature parameters (e.g. XTI and XTB for the Gummel-Poon model)
5. Perform a *conventional, non-pulsed* measurement. The deviation from the pulsed measurement of above is then due to self-heating, modeled by the extra thermal network of the model (see fig. 1). This means, the fitting is obtained by tuning the thermal network model parameters.

Electro Thermal Analogy



A simple thermal model is depicted above. A current, called i_dissip_power, is calculated (inside the custom thermal model) from the dissipated power inside the model.

In the thermal model, this current feeds the thermal RC equivalent schematic. Here, C_chip models the thermal memory of the inner, active part of the chip. R_insidechip_to_surface models the thermal resistor from this inner chip region to its surface. The thermal storage capabilities of the chuck is represented by capacitor C_chuck. Finally, R_chuck models the thermal resistance of the chuck.

There are two voltages available with this simple thermal model: V_chip_overtemp models the inner chip overtemperature, and V_chuck_overtemp represents the overtemperature of the chuck.

Back to the custom thermal model of the DUT itself, it senses the voltage V_chip_overtemp and adds it, interpreted as a temperature, to the actual value of the parameter TNOM (or the model specific temperature parameters).

Static and Dynamic Thermal Modeling

For the thermal modeling, we have to distinguish between static and dynamic thermal modeling. Static dynamic modeling means that the steady-state overtemperature was reached for every individual measurement point. For the corresponding measurements, this means that the DC analyzer for example is set to the absolutely slowest measurement mode possible

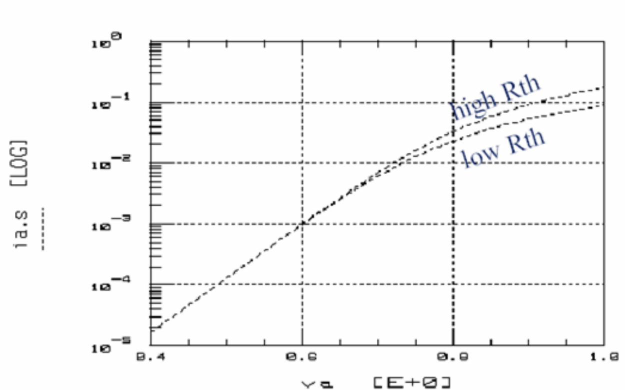
In the dynamic case, the end temperature may not have been reached for every measurement point, and also the actual temperature of the current point may be an

overlay of the self-heating right before and the current one. In this case, for a DC measurement using an HP 414x or 415x instrument, we will measure the time required for that measurement, calculate the corresponding time for each measurement step, and replace later the pure DC simulation by a time domain simulation with the same number of measurement points like the measurements, and each measurement point lasting as long as in the measurement. This means, the kind of the simulation is a copy of the measurement process.

Example: Static and Dynamic Thermal Modeling of the DC characteristic of a Temperature Dependent Resistor

IC_CAP file: diode_selfheat_static.mdl
This example for modeling the static thermal DC characteristics of a diode is after:
Patrick Halloran and Lu Ke, University College Dublin, Ireland, 'Using the Symbolically Defined Device to Simulate Diode Self-Heating', published in the Hewlett-Packard High-Frequency Horizons Journal June 1993.

This model consists of a diode DC model, enhanced for static thermal self-heating. By varying the thermal resistor Rth, the dependence of the diode characteristic in the ohmic range can be studied to also depend on self-heating, described by this parameter Rth. It can be seen that for low Rth, i.e. self-heating, the trace in the ohmic region is below the one without self-heating.

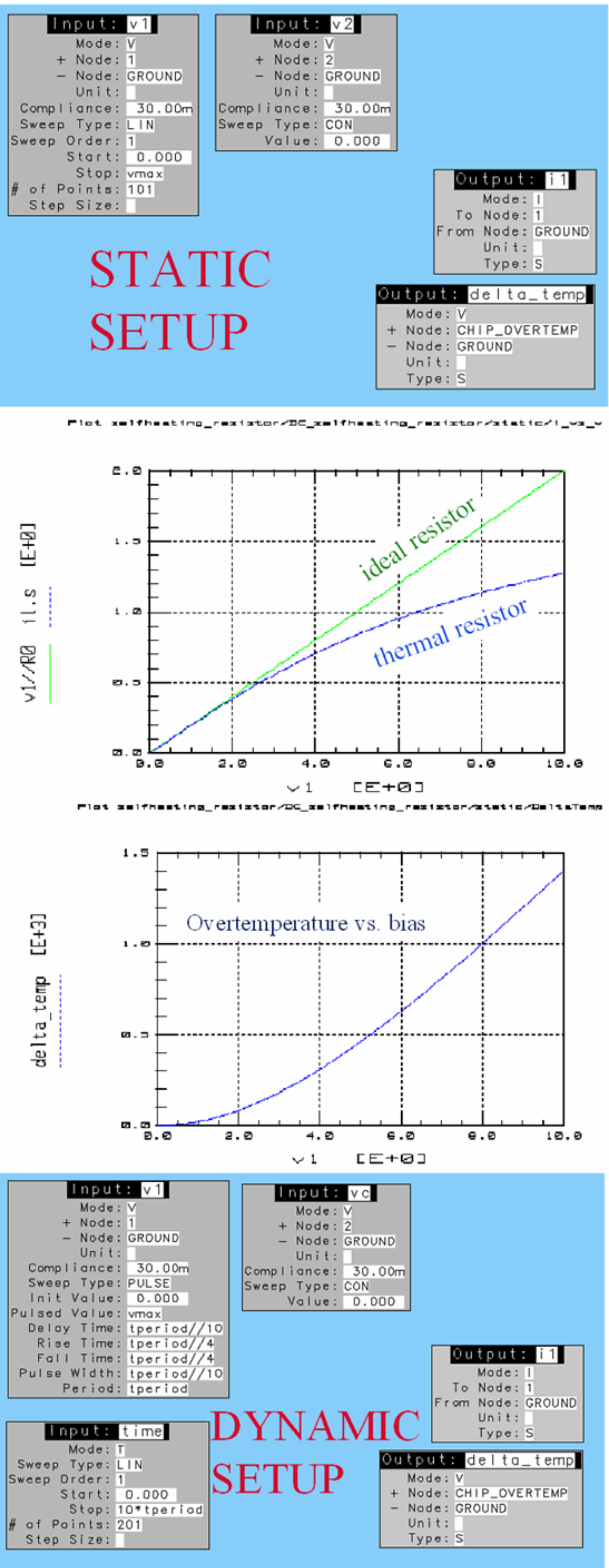


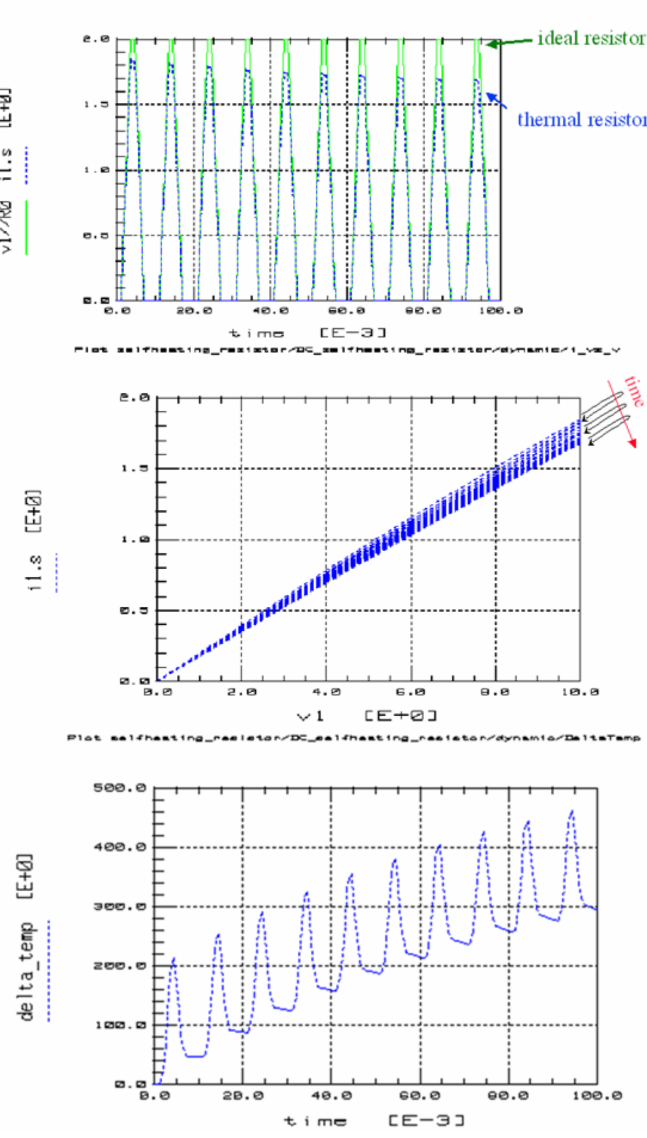
Self-heating reduces the measured diode DC currents and thus affects the measured effective ohmic resistance of the diode.
The user-defined sdd circuit description is given below:

```
define SelfHeatingStaticDiode (A C)
is= 1e-17
n = 1
rs= 5
Pt=3
Eg=1.11
Tamb=300
Rth= 1k
k=BOLTZMANN
q=QELECTRON
  Tjunct = (Tamb + deltaT)
  exparg = (-q * Eg/(n*k*Tjunct) * (1.0 - Tjunct/Tamb))
  isT = (is * (Tjunct/Tamb)**(Pt/n)*exp(exparg))
  id = (isT*(exp(q*vd/(n*k*Tjunct))-1.0))
  deltaT = (_v1 * _i1 * Rth)
  vd=(_v1-(_i1)*rs)
sdd:selfheating_diode A C  f[1,0]=(_i1-id)
end SelfHeatingDiode
```

Example: Static and Dynamic Thermal Modeling of the DC charcteristic of a Temperature Dependent Resistor

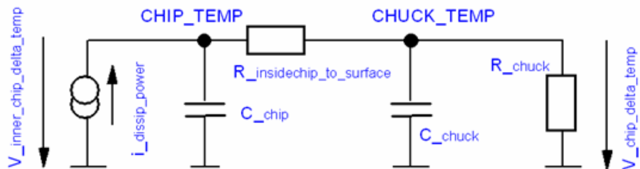
IC_CAP file: resistor_selfheat_static_n_dynamic


Input: v1
Mode: V
+ Node: 1
− Node: GROUND
Unit:
Compliance: 30.00m
Sweep Type: LIN
Sweep Order: 1
Start: 0.000
Stop: vmax
# of Points: 101
Step Size:
Input: v2
Mode: V
+ Node: 2
− Node: GROUND
Unit:
Compliance: 30.00m
Sweep Type: CON
Value: 0.000
Output: i1
Mode: I
To Node: 1
From Node: GROUND
Unit:
Type: S
STATIC
SETUP
Output: delta_temp
Mode: V
+ Node: CHIP_OVERTEMP
− Node: GROUND
Unit:
Type: S
ideal resistor
thermal resistor
v1//R0
i1.s
[E+0]
v1 [E+0]
Overtemperature vs. bias
delta_temp [E+3]
v1 [E+0]
Input: v1
Mode: V
+ Node: 1
− Node: GROUND
Unit:
Compliance: 30.00m
Sweep Type: PULSE
Init Value: 0.000
Pulsed Value: vmax
Delay Time: tperiod//10
Rise Time: tperiod//4
Fall Time: tperiod//4
Pulse Width: tperiod//10
Period: tperiod
Input: vc
Mode: V
+ Node: 2
− Node: GROUND
Unit:
Compliance: 30.00m
Sweep Type: CON
Value: 0.000
Output: i1
Mode: I
To Node: 1
From Node: GROUND
Unit:
Type: S
Input: time
Mode: T
Sweep Type: LIN
Sweep Order: 1
Start: 0.000
Stop: 10*tperiod
# of Points: 201
Step Size:
DYNAMIC
SETUP
Output: delta_temp
Mode: V
+ Node: CHIP_OVERTEMP
− Node: GROUND
Unit:
Type: S



ADS Netlist

```
;ADS simul input file for temperature-dependent resistor
define selfheating_resistor (1 2 CHIP_OVERTEMP)
;define some variables
temp=25
tnom=25
teffK=((temp+273.15)+_v3)
tnomK=(tnom+273.15)
R0=5
a=10m
;calc. temp.dependent resistor current.
; _v1 and _v2 are the potentials at the sdd ports 1 and 2
; HINT: the 3rd sdd port represents the CHIP_OVERTEMP
i_res= ((_v1 - _v2) / (R0 * (1 + a * teffK/tnomK)))
;calc. dissipated power
p_res= (i_resistor*_v1)
SDD:resistor 1 0 2 0 CHIP_OVERTEMP 0  I[1,0]=i_res I[2,0]=-i_res I[3,0]=-p_res
;-> this the thermal capacitor of the chip region around the device
C:C_chip    CHIP_OVERTEMP   0 c=.01m
;-> this is the thermal resistor from inside the chip to its surface
R:R_inside_chip_to_surface  CHIP_OVERTEMP  CHUCK_OVERTEMP r=10
;-> this the thermal capacitor of the chip surface and the chuck
C:C_chuck   CHUCK_OVERTEMP  0 c=1m
;-> this the thermal resistor between the chip surface and the chuck
R:R_chuck   CHUCK_OVERTEMP  0 r=100
end sdd_selfheating_resistor
```



HINTS:
If device models exhibit a thermal node, they include usually a thermal resistor and a thermal capacitor.
The default values are usually:
RTH = 0
CTH = 0

When specifying RTH, make sure to set CTH to a non-zero value. Otherwise, your device is 'thermally faster' than electrically! I.e. if you simulate pulses, you do not get the expected fast risetime of the electrical performance, followed by the slower thermal time constant.

If you are not sure about your CTH value, set it to a big value, e.g. several ms.

Temperature Modeling

If a model includes temperature effects, proceed as follows:

1. Perform a full device measurement characterization at e.g. 27°C
2. Set TNOM=TEMP=27, and extract all temperature-independent model parameters (conventionally).

For a diode, extract e.g. IS N, RS, CJO etc.

Then:

1. Perform another measurement at the other temperatures, e.g. -25°C and +75°C
2. Keep the previously extracted parameters fixed,

- Keep TNOM=27 (this is the parameter extraction temperature),
- Set TEMP (the simulation 'prediction' temperature to the one of the measurement, and extract/optimize now (only)the temperature parameters (e.g. XTI for the classical SPICE diode model).

However,if XTI cannot fit the thermal behavior of your diode, you need to extract a full set of model parameters at each measured temperature (TNOM=TEMP=meas.temp).

And you obtain N full parameter sets for N measurement temperatures

NOTE: set the temperature parameters to default, i.e. switch-off their effect.

For the diode, set XTI=0 i.e., you will end up with N individual IC-CAP .mps parameter files (In the ModelEile, perform a 'Save Model Parameters' for each modeled temperature).

In the next step, we need to interpolate these N parameter values vs. the temperature i.e. to find a fitting curve for each parameter and its N temperature-dependent values.

Procedure

From demo_features/DEPOTS.mdl, apply Transform RUN_ALL in GUI_DEPOTS/PARAMS_TO_MDM.
This will convert the N .mps files into a single .mdm file, which has the temperature as the Input (sweep) with N temperature values, and as many Outputs as you had modeling parameters in your .mps files before.

Then, ImportCreate the .mdm file into a new Modelfile Setup, and display each model parameter (Output) vs. the temperature (Input).

Write little PEL Programs for finding the fitting functions, e.g. Program 'IS_fit':
RETURN IS_0 + IS_1*TEMP + IS_2*TEMP^2
or any other fitting function.

NOTE: the Variables IS_0, IS_1 etc. will be SetupVariables.

Display the result of IS_fit in the same plot as the IS (from Output) vs. temperature.
Apply a PlotOptimizer to tune/optimize your fitting parameters (IS_0, IS_1 etc.).

Last not least, combine all your diode DC curves, CV curves etc. into a single, big .mdm file (applying GUI_UTILITIES/COMBINE_MDMS from demo_features/DEPOTS.mdl), and load these big files into a new ModelFile. Copy-Paste your Circuit from the individual-temperature-modeling Modelfile into the new all-temperature Modelfile, replace the model parameters by the previously identified temperature fitting equations either in the ModelParameter Tab list of the Modelfile, or if your simulator (e.g. ADS, Spectre, Eldo) permits into the Circuit Tab of the Modelfile.
Simulate.
Apply a final PlotOptimizer for fine-tuning of
IS_0, IS_1, IS_2 and N_0, N_1, N_2 etc. etc.

NOTE: A demo about scaled parameter modeling can be found under
demo_features_2008xxxx\1_BASIC_MDLG_EXAMPLES\24_Scaled_Parameters_Mdlg\scaled_parameterized_modeling.mdl

Although it is about geometry scaled modeling, the example given there can be applied to temperature modeling as well.

Publications

R.Bouchakour, C.Lallement, T.Maurel: Thermal and Electrical Modeling and Characterization of Bipolar and Power MOSFET Transistors With IC-CAP and Saber, Hewlett-Packard Characterization Solutions Revue, Winter/Spring 1994, S.6-13
and also published in the Proceedings of the 1993 European IC-CAP User Meeting, Esslingen, Germany

T.Maurel, R.Bouchakour, C.Lallement, Modeling and Simulation of the Electrothermal Behavior of the Power Bipolar Transistor, Proceedings of the 1994 European IC-CAP User Meeting, Colmar, France

M.Dunn, B.Schaefer, Link measurements to nonlinear bipolar device modeling, published in Microwaves&RF, Feb 1996, reprint as HP publication number 5964-9170E

Patrick Halloran and Lu Ke, University College Dublin, Ireland, 'Using the Symbolically Defined Device to Simulate Diode Self-Heating', published in the Hewlett-Packard High-Frequency Horizons Journal June 1993

K.Lu, P.Halloran, T.J.Brazil, Simple Method to Simulate Diode Selfheating Using SPICE, Electronics Letters, August 13, 1992, Vol.28, No.17

On Target Modeling

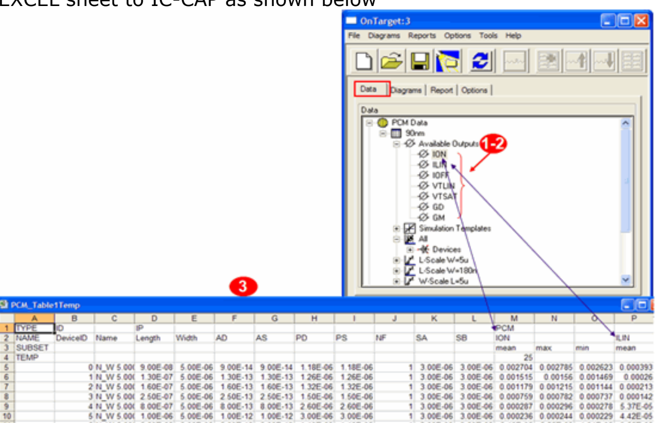
Contents

- *Using PCM Data for Device Modeling* (iccapmhb)
- *Working of On Target Modeling* (iccapmhb)
- *Tutorial to use On Target Modeling* (iccapmhb)

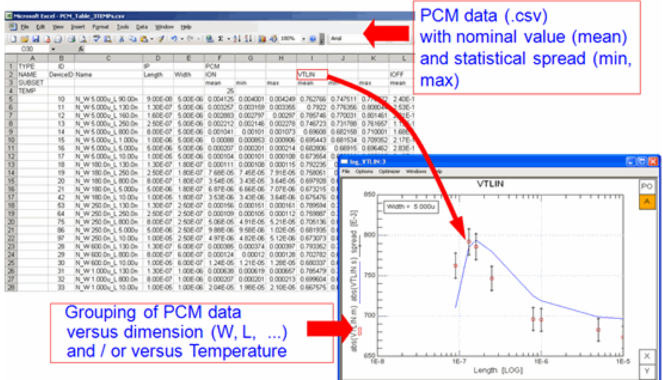
Using PCM Data for Device Modeling

Introduction

The key feature to use PCM Data for Device Modeling is to import PCM Data from an EXCEL sheet to IC-CAP as shown below



PCM Data Representation



Application Scenarios

Foundry
Production measures and publishes key parameters like (Vtlin, Vtsat, ...).

- Applying the On-Target Tool, an existing model can be compared to these PCM data. If required, the model is adjusted to changes in the process.

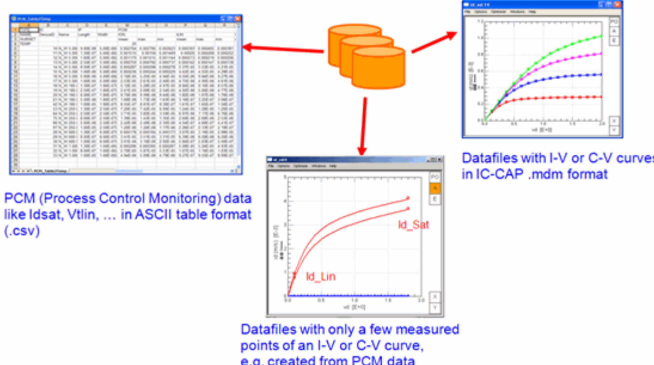
Process Development
PCM Data are available from TCAD Simulations. Maybe, additionally a few I-V Curves, with a few data points.

- The On-Target Tool allows to establish a first Model, without having Silicon available.

Modeling and Model Centering
A 'classic' modeling engineer extracts model parameters based on I-V curves.

- Provided there are PCM data available, the model can be verified against the statistical spread of the PCM data.
- Instead of measuring several wafers with complete measurement sweeps.

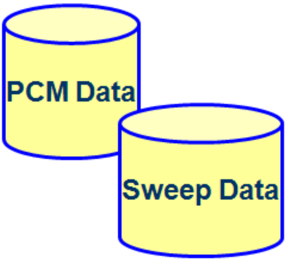
Data to be used for Application


PCM (Process Control Monitoring) data like Idsat, Vtlin, … in ASCII table format (.csv)
Datafiles with I-V or C-V curves in IC-CAP .mdm format
Datafiles with only a few measured points of an I-V or C-V curve, e.g. created from PCM data
Id_Sat
Id_Lin

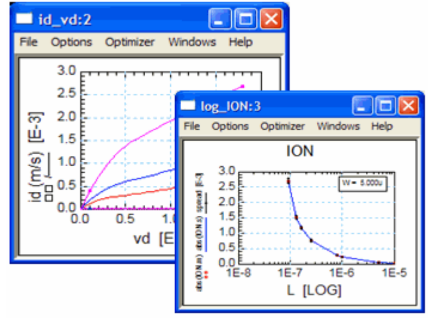
On Target Modeling working

Three fundamental steps are as:

1. **Data Import**

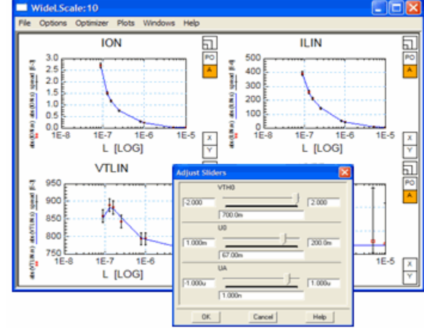


2. **Data Organization**

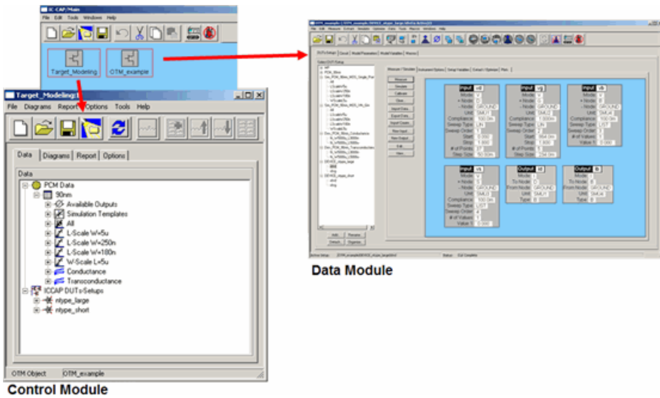


Create scaling
and I-V Diagrams

3. **Model Validation**



IC-CAP Structure



Control Module

- IC-CAP .mdl file with GUI
- Manages the Data Module .mdl file
- Handles templates for PCM simulation, plots, reports, etc.

Data Module

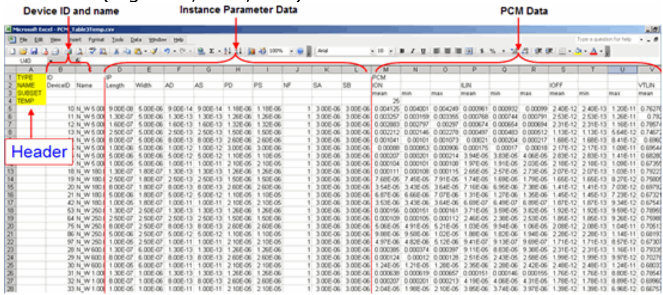
- Implemented as a standard IC-CAP .mdl file

- Holds the data for the On-Target modeling tasks
  - Setups with meas. and simul. data
- User specific extensions

> The OnTarget Toolkit consists of a Control Module 'Target_Modeling', and a Data Module 'OTM_xxxx'. The 'Target_Modeling' is the master, the 'ITM_xxx' the slave.

IC-CAP has defined a Simple .csv Data Format for storing

- Device ID and name,
- Instance Parameter data,
- PCM data (e.g. ION, ILIN, etc.)



For each PCM value, the spreadsheet above may include a spread (Mean, Min and Max values),and also different temperature values.

THE IC-CAP .csv DATA FORMAT IN A NUTSHELL:
The Header keywords 'TYPE', 'NAME', 'SUBSET' and 'TEMP' are a MUST in the .csv file.
Also, in the row 'TYPE', the keywords 'ID', 'IP' and 'PCM' are a MUST.(quote}

Any PCM data names are accepted.

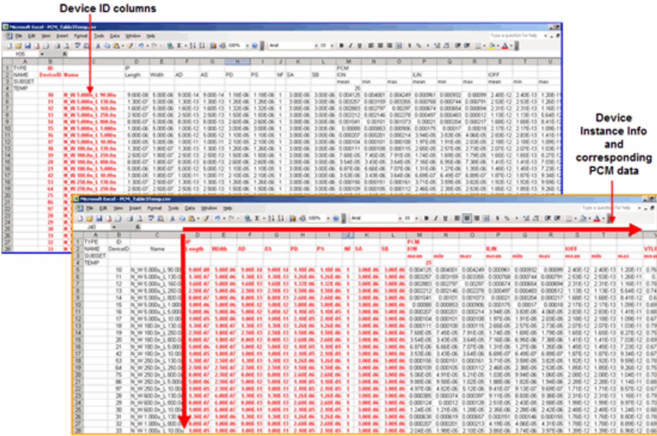
> NOTE:
> If you don't have min and max values, neither mean values,
> but only pure measured data, you can delete the columns 'min' and 'max' in the PCM columns section.Enter your meas. data to the column 'mean'. Important: the keyword 'mean' in the row SUBSET has to remain.

Further info from the IC-CAP Manual:
The first n rows of the table are the KEYWORD rows.
Keywords are:

- TYPE with
  - ID=Identification
  - IP=Instance parameters
  - PCM=PCM parameters
- NAME with
  - DeviceID
  - Name
  - followed by any number and combination of Instance Parameters (IP) and PCM values (PCM)
- SUBSET: specifies in detail the PCM values: Allowed names are mean, min, max. These identifiers are used to add statistical data.
- TEMP: Temperature of the PCM value

In the file, each column is separated by a comma (,). This enables you to import the file into table calculation programs like MS Excel.

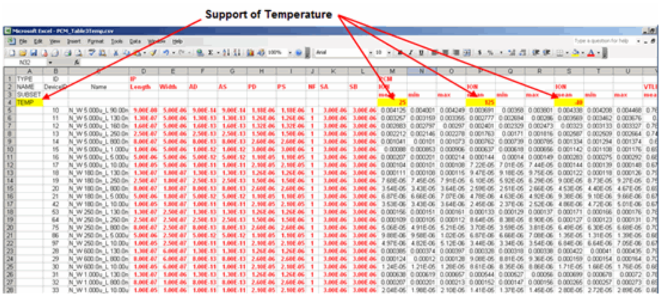
- A further rule for the rows NAME and TEMP is that if a cell does not contain a name, the value of the previous one is used. For example, T=25°C is used as a temperature for all PCM values as long as no other value is specified.
- An empty cell inside the NAME row leads to an error message.
  For example, if the cell following AD is empty, you will get the error message Instance Parameter Name AD is defined twice. Reason: In this case, the name of the previous cell is used and double definitions are not allowed.
  An empty data value cell (PCM data) is assumed to have the value of zero.
  Important, when you don't have a fully populated PCM data table.



IC-CAP .csv data format:

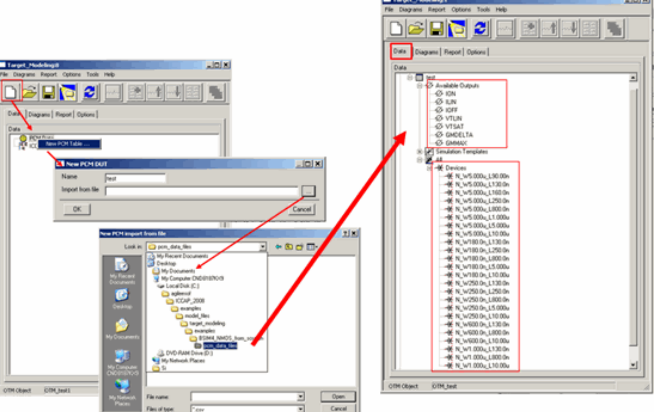
- The 2.d and 3rd column is reserved for the DeviceID and the device Name.
- Then indicated by the keyword 'IP', follows a unlimited number of columns, describing the instance parameters (L, W, SA, etc)
- Finally, indicated by the keyword 'PCM', follows a unlimited number of columns, describing the measured PCM values.

Blank entries in the PCM values are supported.

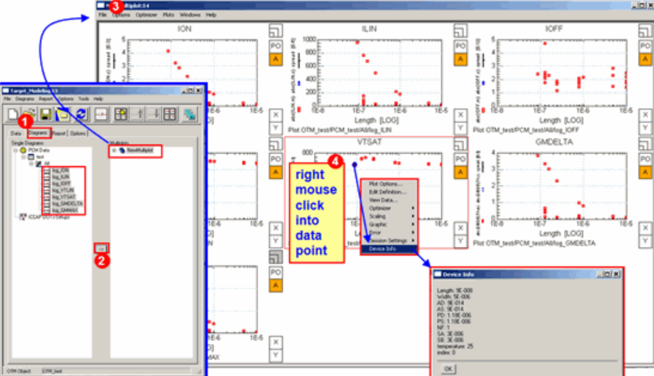


When reading a row, the value stays identical until another value is read. Therefore, you can specify e.g. ION or the temperature values in every column, or simply in the first.

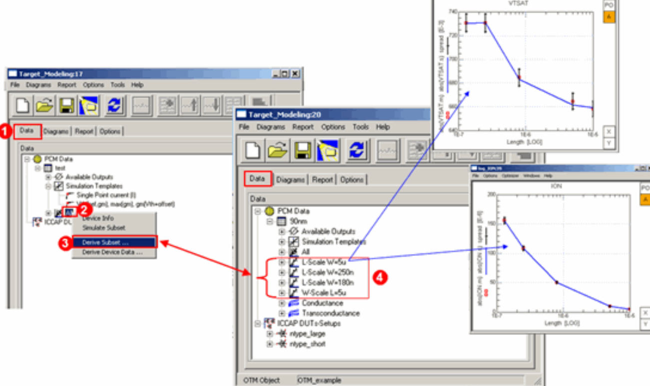
Loading the .csv file into Target Modeling GUI



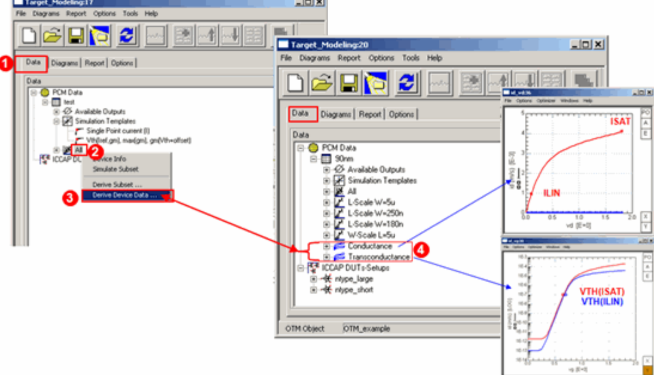
1. First, inspect the loaded PCM data



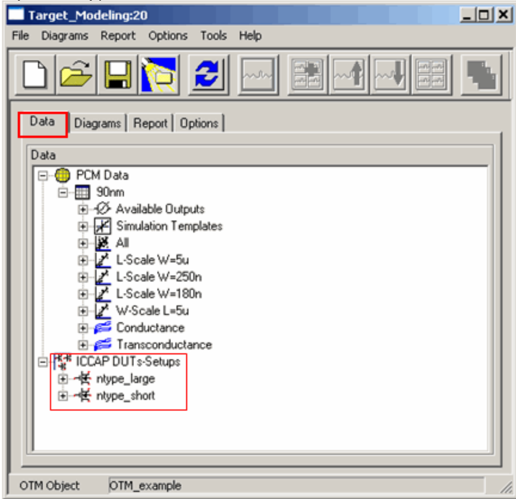
2. Group your PCM data (Derive Subset)



3. Optionally, convert your PCM data to pseudo-sweep data (Derive Device Data)



4. Optionally, add conventional I-V measurements

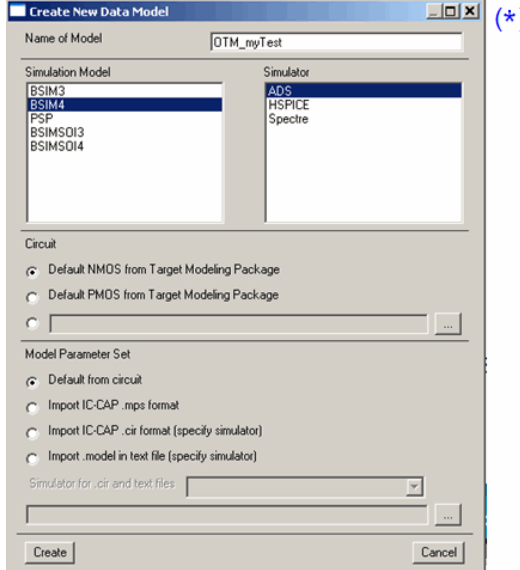


Till now the data are imported and organized.
Now you have to set up the simulations for the imported .csv data.

The PCM data have been imported before, with no detailed information about how they were measured. E.g. for ION, we imported the meas. value, but for making simulations, we need to tell IC-CAP what specific vd and vg values have to be applied in the simulations we are going to perform now.

Supported simulators and models

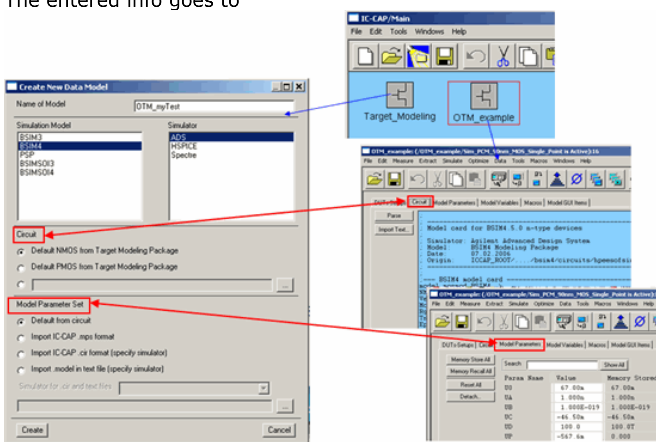
- Support for ADS, Spectre and HSPICE
- Direct support for BSIM3, BSIM4, PSP, BSIMSOI
- User defined models can be added e.g. HV-MOS, macro-models etc.



(*)

* This GUI pops up when you set up a new .csv data set,as the very first window to be filled out.

The entered info goes to

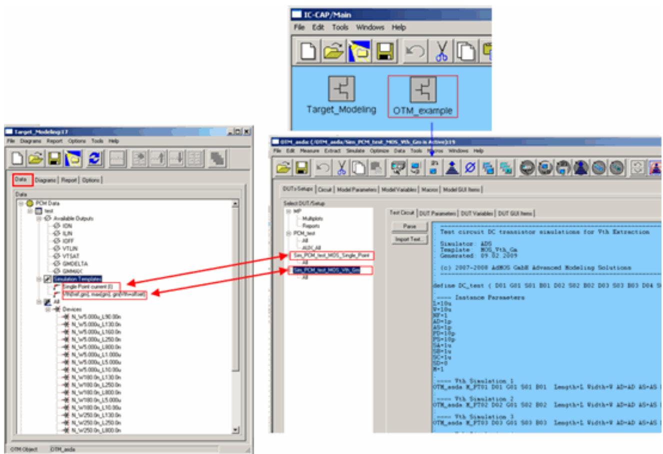


Simulation Templates

So far, in the Data tab below, we have discussed the

- Available Outputs
- Devices

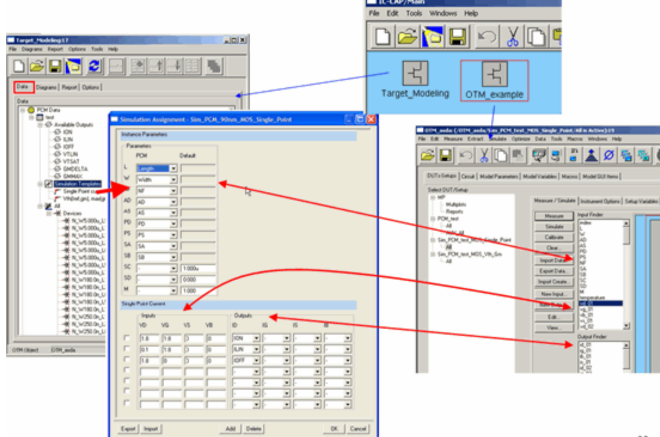
  Now: Simulation Templates



After the model and the simulation details have been defined, we need to specify how to simulate e.g. ION, VTH etc.,i.e. the bias voltage values.

> While the Simulator, and the model were already specified right at the beginning, even before importing the .csv data file, it is now time to specify what *kind of simulation* need to be performed on the PCM data. In the Data ModelFile, is basically a DUT TestCircuit for the ION, ILIN, IOFF, ISAT etc., and another DUT TestCircuit for the VTxx and the GMxx data.
>
> Right-click on 'Simulation Templates' to add the simulation circuits (DUT TestCircuits in the Data-ModelFile).

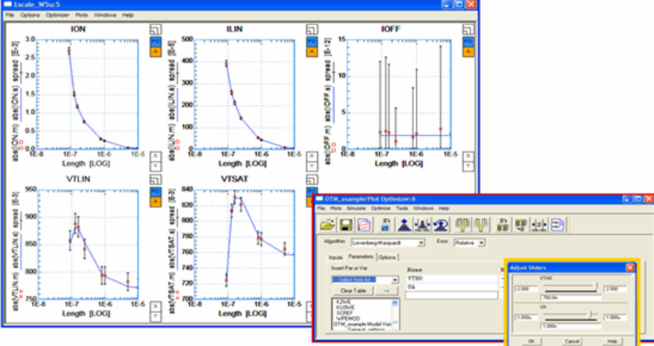
Simulation bias conditions



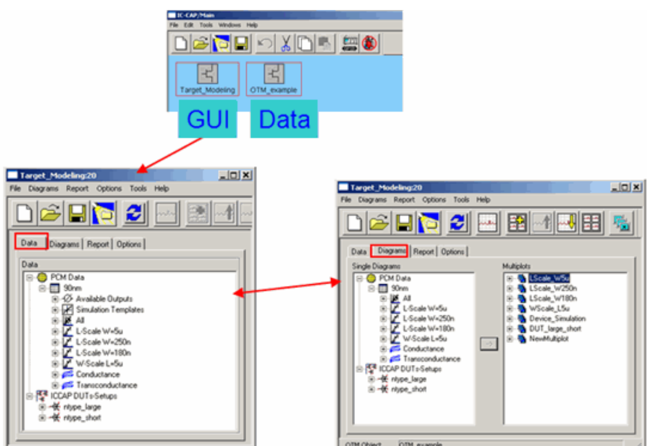
The PCM data were meas. with specific stimulus conditions. The simulator needs to apply them *too*.
So far:

- Data are imported and organized
- Corresponding simulation settings have been made
  Now you will apply the PlotOptimizer to fine-tune, adjust, update etc. your Model Parameters.

Applying PlotOptimizer



- A tandem of manager and data base ModelFile.
- A GUI provided by the manager ModelFile to perform all requested tasks.

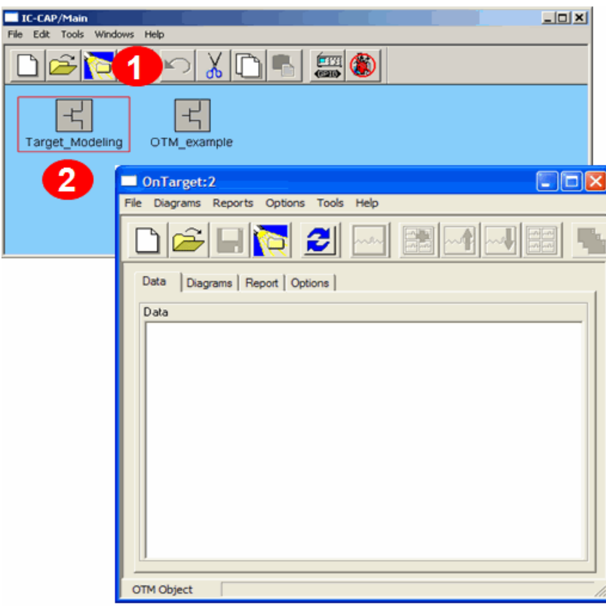

GUI
Data
Target_Modeling
OTM_example
Target_Modeling:20
Data
Diagrams
Report
Options
Single Diagrams
Multiplots
PCM Data
90nm
ICCAP DUTs-Setups
OTM Object
OTM_example

Tutorial

How Target Modeling Works: Opening and Understanding the Example
In this tutorial we will cover:

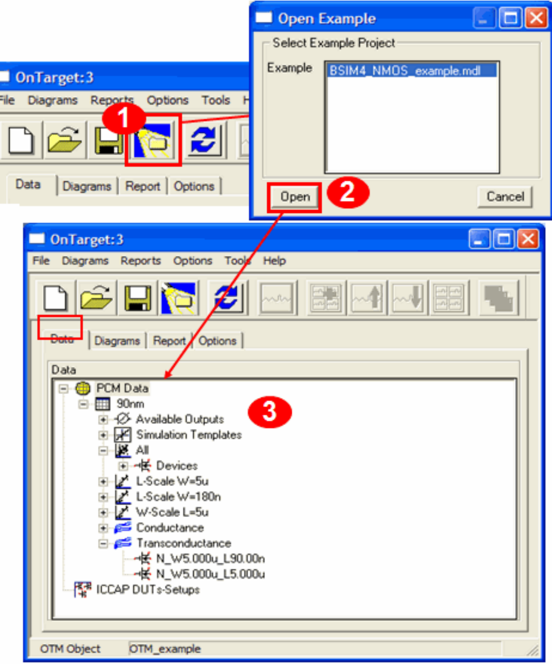
- How to start the Target Modeling Package
- How to load the predefined example
- Viewing the device table information
- Viewing the available PCM data
- Viewing the simulation templates
- Viewing the subsets of devices for scaling diagrams
- Defining I-V curves from PCM data
- Handling the optional conventional I-V curves
- Display diagrams
- Using the PlotOptimizer to extract the model
- Generating an HTML report

How to start the Target Modeling Tool



1. Under IC-CAP Examples, load the file: Target_Modeling.mdlfrom Examples/model_files/target_modeling
2. Open the new dedicated OnTarget window by double-clicking on the "Target_Modeling" icon.

Loading a Predefined Example



1. Click on the "Examples" icon to load a predefined example
2. Select the BSIM4_NMOS_example file and click "Open"
3. The Target Modeling GUI gets populated with data.

This example already includes
in **Tab 'Data':**

- Device PCM data of all loaded devices
- A subset of predefined scaling PCM data
- I-V Setups obtained from PCM data
- Conventional I-V Setups

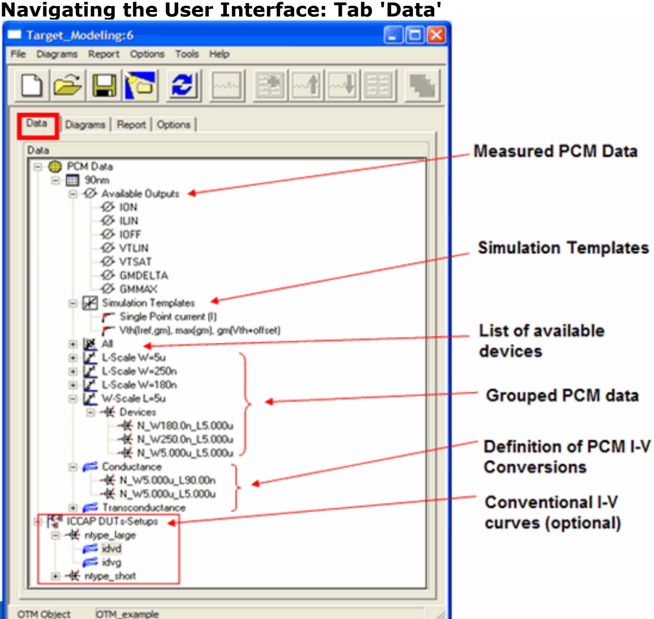
in **Tab 'Diagrams':**

- A set of predefined multiplots diagrams

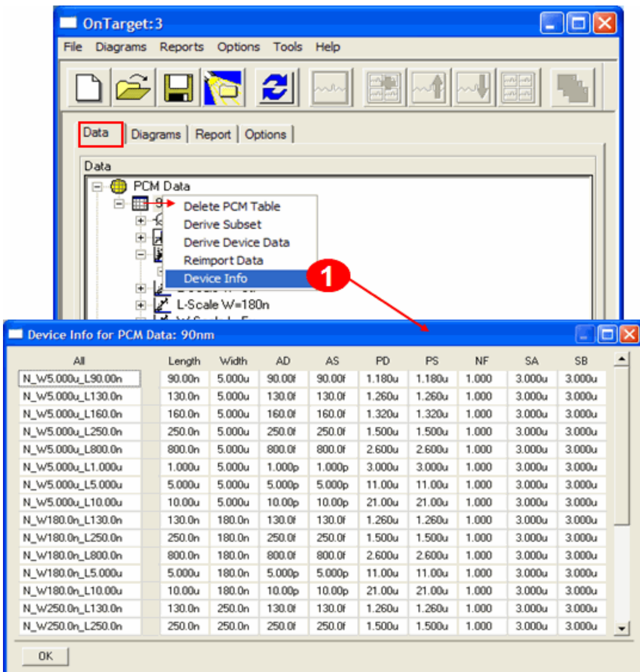
in **Tab 'Options':**

- ADS is selected as simulator

Navigating the User Interface: Tab 'Data'



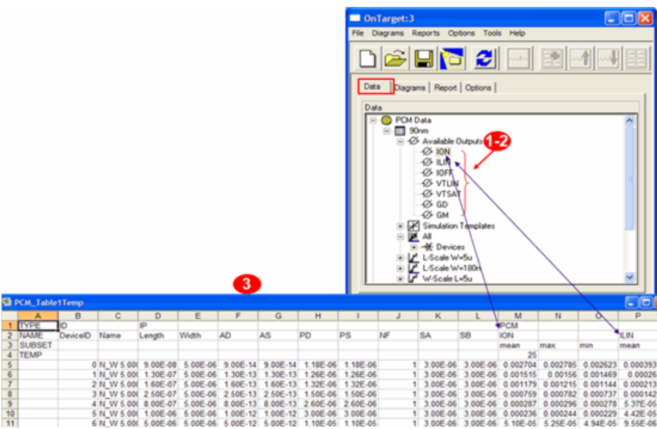
Exploring the Example:Device Table



All	Length	Width	AD	AS	PD	PS	NF	SA	SB
N_W5.000u_L90.00n	90.00n	5.000u	90.00f	90.00f	1.180u	1.180u	1.000	3.000u	3.000u
N_W5.000u_L130.0n	130.0n	5.000u	130.0f	130.0f	1.260u	1.260u	1.000	3.000u	3.000u
N_W5.000u_L160.0n	160.0n	5.000u	160.0f	160.0f	1.320u	1.320u	1.000	3.000u	3.000u
N_W5.000u_L250.0n	250.0n	5.000u	250.0f	250.0f	1.500u	1.500u	1.000	3.000u	3.000u
N_W5.000u_L800.0n	800.0n	5.000u	800.0f	800.0f	2.600u	2.600u	1.000	3.000u	3.000u
N_W5.000u_L1.000u	1.000u	5.000u	1.000p	1.000p	3.000u	3.000u	1.000	3.000u	3.000u
N_W5.000u_L5.000u	5.000u	5.000u	5.000p	5.000p	11.00u	11.00u	1.000	3.000u	3.000u
N_W5.000u_L10.00u	10.00u	5.000u	10.00p	10.00p	21.00u	21.00u	1.000	3.000u	3.000u
N_W180.0n_L130.0n	130.0n	180.0n	130.0f	130.0f	1.260u	1.260u	1.000	3.000u	3.000u
N_W180.0n_L250.0n	250.0n	180.0n	250.0f	250.0f	1.500u	1.500u	1.000	3.000u	3.000u
N_W180.0n_L800.0n	800.0n	180.0n	800.0f	800.0f	2.600u	2.600u	1.000	3.000u	3.000u
N_W180.0n_L5.000u	5.000u	180.0n	5.000p	5.000p	11.00u	11.00u	1.000	3.000u	3.000u
N_W180.0n_L10.00u	10.00u	180.0n	10.00p	10.00p	21.00u	21.00u	1.000	3.000u	3.000u
N_W250.0n_L130.0n	130.0n	250.0n	130.0f	130.0f	1.260u	1.260u	1.000	3.000u	3.000u
N_W250.0n_L250.0n	250.0n	250.0n	250.0f	250.0f	1.500u	1.500u	1.000	3.000u	3.000u

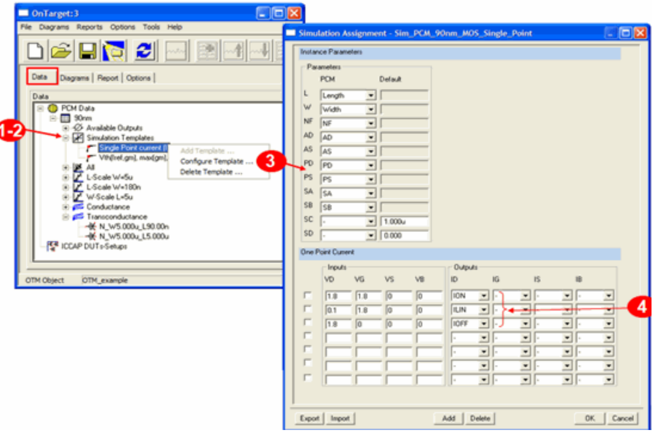
1. In the "PCM Data" tree, right click on "90nm" (this is the name of the set of loaded devices) and select "Device Info"
2. This will show the "Device Table" for the selected PCM Data.
3. Device information was read from the PCM data Excel spreadsheet

Exploring the Example:Available PCM data



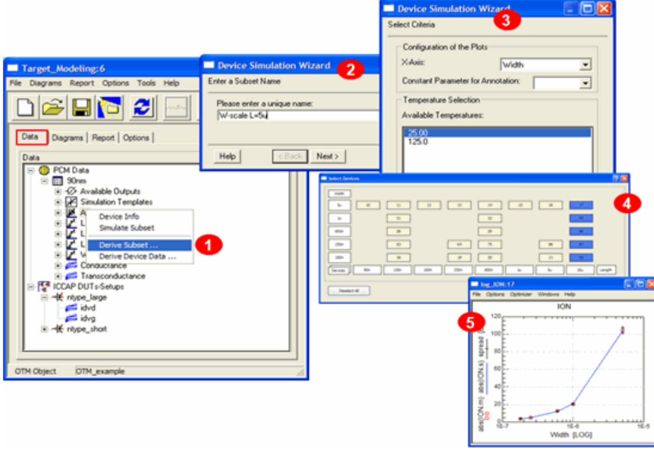
1. In the "PCM Data/90nm" tree, open the "Available Outputs" tree
2. This will show the available PCM data. This information was also read from the PCM data Excel spreadsheet.
3. For each PCM value, the spreadsheet may include a spread (Min and Max values)

Exploring the Example:Simulation Templates for PCM



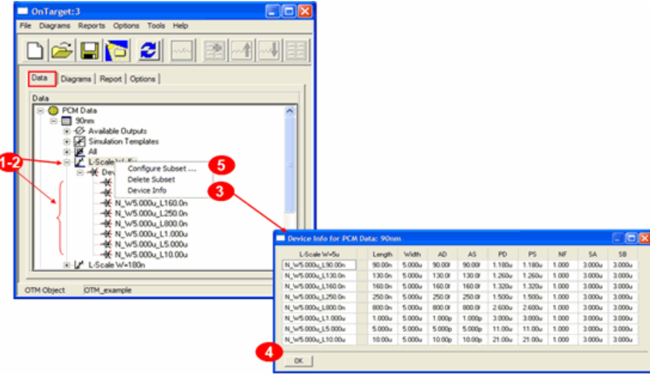
1. In the "PCM Data/90nm" tree node, open the "Simulation Templates" tree
2. This shows the two predefined templates for simulating the PCM data: single-point Id_Vd and VTH.
3. Right click on one of the two templates and select "Configure Templates..." to open the bias definitions for I or V simulations.
4. Each PCM data has its own bias definition to be used by the simulations later

Understanding the Example:Creating a subset of devices



1. In the "All" devices tree node, right click and select "Derive Subset"
2. Name the subset and click "Next"
3. Select X-Axis and Temperature
4. Select Devices
5. Closing the dialog will create the subset of devices and automatically create scaling diagrams

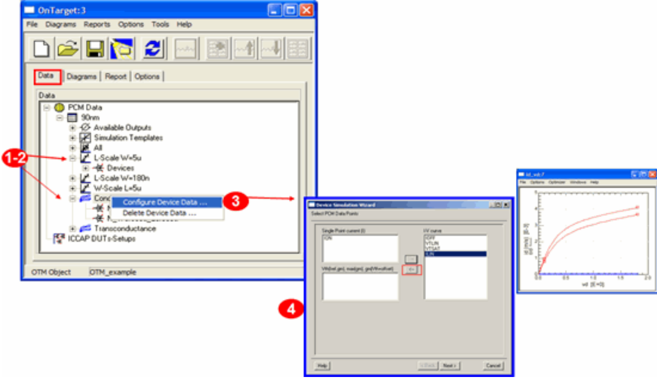
Understanding the Example:Viewing subset of devices



1. In the "PCM Data/90nm" tree node, open the "L-Scale W=5u" and "Devices" sub-tree
2. This is a predefined, customer-selected scaling, as described in the slide before
3. Right click on "L-Scale W=5u" and select "Device Info...".

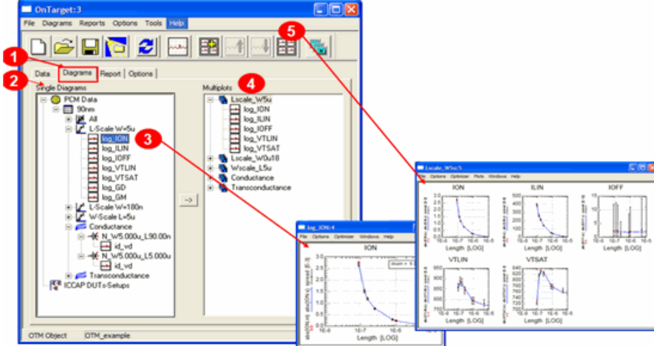
4. This will show the information for the devices which have been grouped into this subset
5. You can select "Configure Subset..." to modify the device selection

Understanding the Example:Defining I-V Setups from PCM data



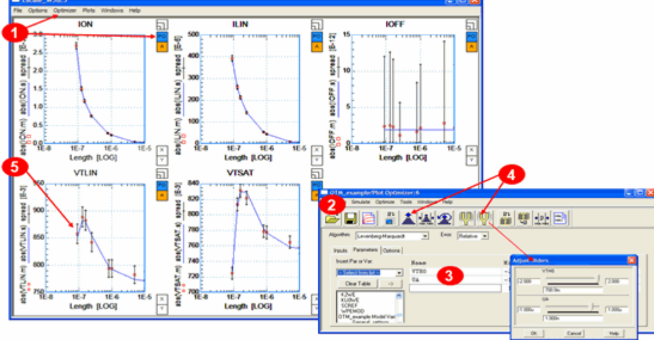
1. In the "PCM Data/90nm" tree node, open the "Conductance" tree
2. "Conductance" is a predefined I-V diagram (id_vd)
3. Right click on "Conductance" and select "Configure Device Data...".
4. This opens a wizard that helps the user to select PCM data, configure an I-V curve and select the devices to be used for simulation.

Understanding the Example:THE PLOTS



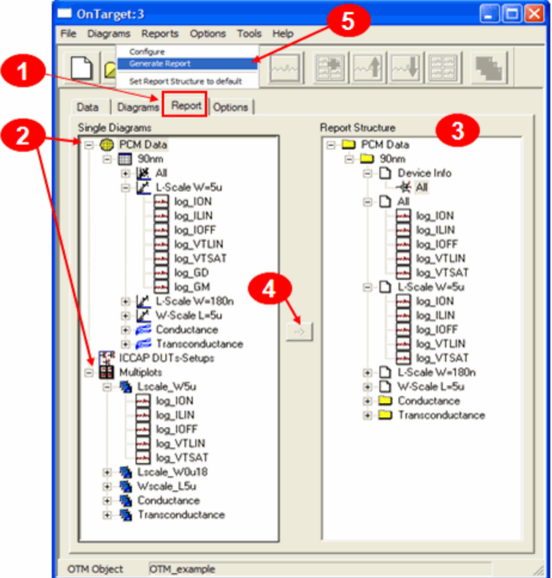
1. Select the Tab "Diagrams"
2. All the available plots are defined in the "Single Diagrams" tree (on the left).
3. Display single plots by double-clicking on their names.
4. Multiple plots diagrams are defined in the "Multiplots" tree (on the right).
5. Display MultiPlots by double-clicking on their names.

Understanding the Example:Using PlotOptimizer to tune the parameters



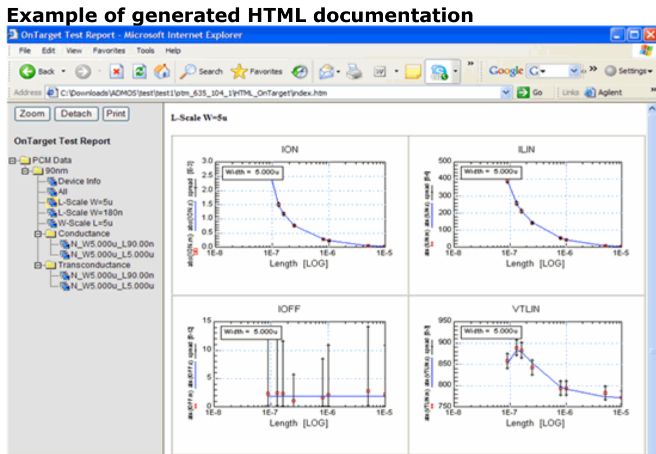
1. Once a Plot is open, you can apply the IC-CAP PlotOptimizer
2. Select the parameters to be tuned/optimized
3. Run the Tuner or the Optimizer
4. Note that the diagrams show Min/Max spreads for each PCM data

Understanding the Example:Report Generation



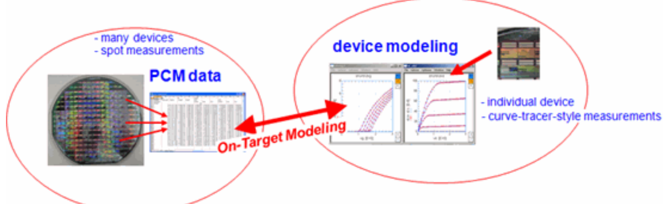
1. Select the "Report" Tab
2. Single and Multiplots diagrams are available on the left side tree.
3. The report structure is shown on the right side.
4. User may add/delete/move diagrams in the report structure
5. Select "Reports/Configure" and hit "Generate Report" to automatically generate HTML report.

Example of generated HTML documentation



Conclusions
On Target Modeling enables IC-CAP users to:

- Adjust a simulation model which was extracted from one set of golden diesto data representing the overall trend of the process.
- Perform comparisons between different simulated models and silicon data.
- Provide documentation of the improved model.



Direct Measurement Data Reduction for Scalable Parameter Extraction

This is an updated slideset of the one presented at the Europ. IC-CAP User Meeting in Prague, 2003, by Pietro Brenner, Infineon Munich, and Franz Sischka, Agilent-EEsof.

It now features the LSYNC sweep of IC-CAP, available since IC-CAP 2006B, Update3 (Sept.2007)

Also, a new IC-CAP demo has been developed. See demo_features/1_BASIC_MDLG_EXAMPLES/24_Scaled_Parameters_Mdlg/scaled_parameterized_modeling.mdl

Introduction

What is the fastest way to get scalable model parameters for extensive transistor libraries without performing a lot of single transistor measurements and a lot of parameter extractions ?
Perform direct extraction of *normalized* parameters on *normalized* measurement data.
Provided you know:

- transistor layout measures,
- vertical transistor geometry and transistor configurations,
- absolute, geometry independent parameters (e.g. NF, VAF, VAR)and geometrically normalized parameters: e.g. is_a, is_p, cj_a, cj_p, re_a, rb_sh, rb_l
  as well as the correct parameter scaling rules :

```
e.g.  IS = is_a * AREA_E  +  is_p * PERIM_E,
CJ = cj_a * AREA_E  +  cj_p * PERIM_E,
tf0 = tf0_a * (1+ a1*PERIM_E/AREA_E)
```

You can calculate parameter sets for arbitrary device geometries applying parameter calculation scripts (see e.g. Tradica /1/).

New Approach for Direct Extraction of Geometrically Normalized Device Parameters

Basic Idea

The basic idea for Direct extraction of Geometrically Normalized Device Parameters is as follows:

- Measure DC- and C(V)-characteristics from a set of deviceswith different geometries, e.g.
  - 3-5 different emitter lengths LE and
  - 4-5 different emitter widths WE.
- Transform measurements of device X Cx(v) into normalized per-area CAREAx(v) and per-perimeter CPERIMx(v) pseudo-measurements, e.g.

$$Cx(v) = C_{AREAx}(v) \cdot AREAx + C_{PERIMx}(v) \cdot PERIMx \quad (1)$$

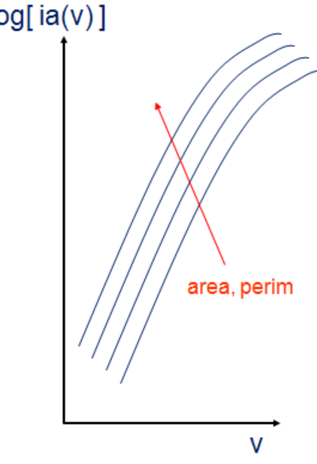
$$\frac{Cx(v)}{AREAx} = C_{AREAx}(v) + C_{PERIMx}(v) \cdot \frac{PERIMx}{AREAx} \quad (2)$$

= known data

- Perform multiple linear regression on equation (2) for each bias point vs. slope leads to perimeter-normalized meas.data CPERIM(v) and the intercept points lead to area-normalized meas.data CAREA(v).
- Perform parameter extraction only on the normalized data plots CPERIM(v) and CAREA(v).
  Only the parameters of the per-area and per-perimeter normalized capacitance pseudo-measurements have to be extracted.

Example: Diode DC Characteristics

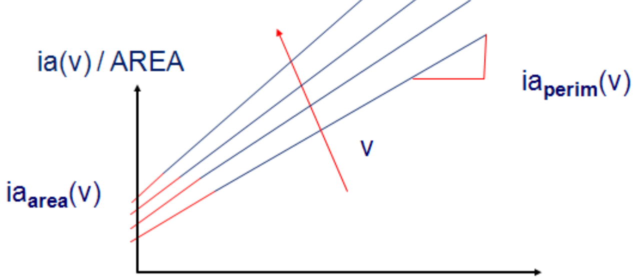
1. Measure data of devices with different geometry as a function of area and perimeter.



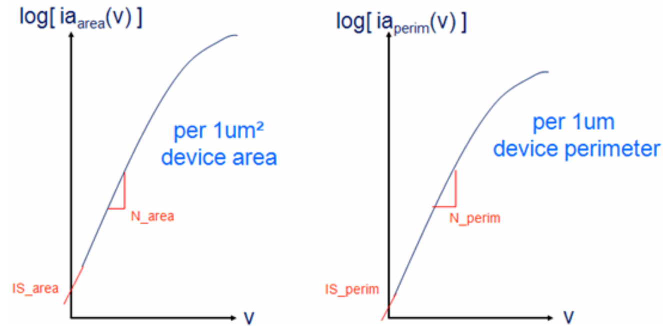
2. In a second step, display the data as a function of area and perimeter, after a known scaling function like e.g.

$$ia(v) = ia_{area}(v) * AREA + ia_{perim}(v) * PERIM$$

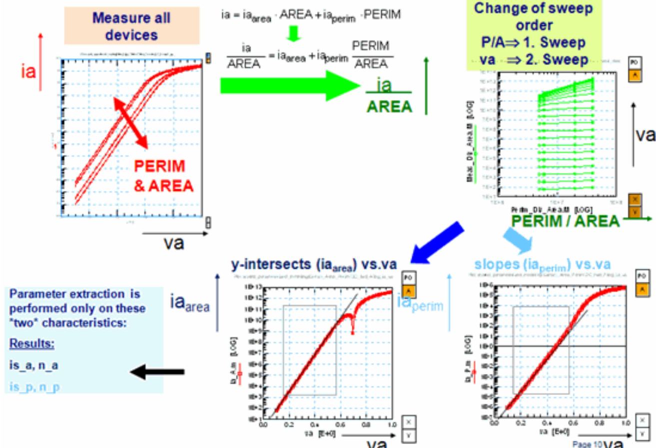
$$\frac{ia(v)}{AREA} = ia_{area}(v) + ia_{perim}(v) * \frac{PERIM}{AREA}$$



3. From the y-axis intersect, calculate iaarea(v) and from the slope, calculate iaperimeter(v), for every individual 'v'.
4. Display the normalized per-area and per-perimeter measurements against the original stimulus.



5. Extract the per-area (IS_area, N_area) and per-perimeter parameters (IS_perim, N_perim)

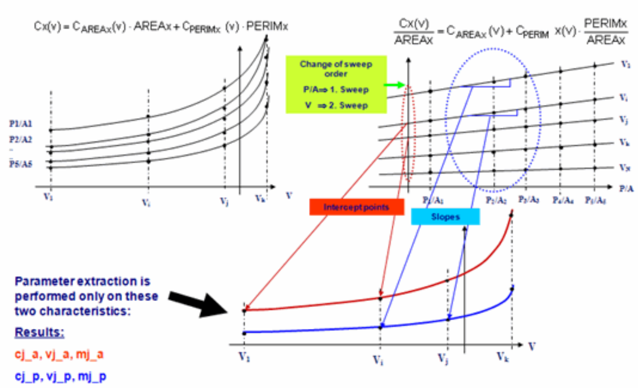


NOTE: In the Plot **'Change of sweep order'**, the curves represent lines, provided that the prerequisite equation is fulfilled. As can be seen, for low va, i.e. low ia values, the noise overlies the prerequisite, and for high va values, i.e. the ohmic range of the diode current, the prerequisite is not met too. Nevertheless, we can apply linfits to all traces of va, and later exclude those data in the parameter extractions behind the Plots 'y_intersects' and 'slopes'. See the specified parameter-extraction-boxes in these Plots.

NOTE:

- If the traces represent lines with a positive slope, the meas. data include dependencies of both, area and perimeter.
- If the traces in Plot 'Change of sweep order' represent flat lines (slope=0), then there is no dependency of perimeter for the meas. data.
- If the traces represent lines with a negative slope, then the assumed (drawn) area and parameter values are not valid for the produced device on the wafer.
  The area and perimeter are typically both smaller on the real device than in the drawn masks.
- If the per-area parameters (Plot y-intersects) are negative, then the parimeter effect dominates the area effect, and thus, the prerequisite of splitting into area and perimeter is no longer valid.
  For details on the area-perimeter scaling, see J.Berkner, Kompaktmodelle fuer Bipolartransistoren, Expert-Verlag, Renningen, Germany, ISBN 3-8169-2085-3, Chapter 7.1.2

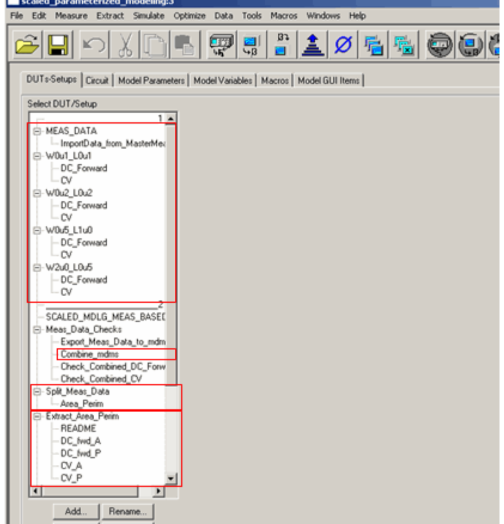
Example: Diode CV Characteristics

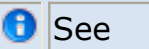


IC-CAP Implementation

An IC-CAP Model file example has been developed for the proposed scaled modeling method,

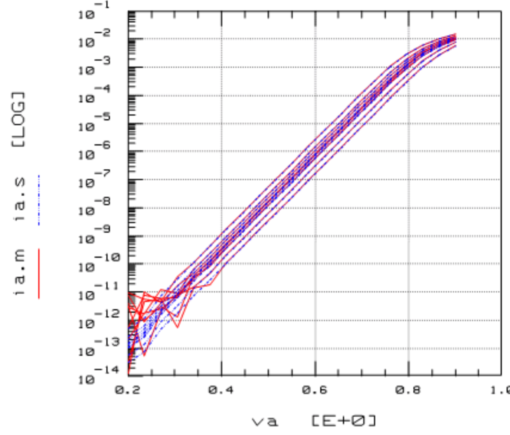
- for area and perimeter scaling,
- DC and CV.
  Other scaling functions can be implemented in a very similar way.



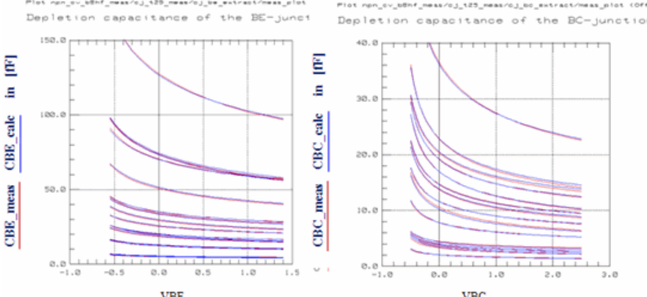

See demo_features/1_BASIC_MDLG_EXAMPLES/24_Scaled_Parameters_Mdlg/scaled_parameterized_modeling.mdl

Results

Diode DC Modeling Result, all devices
Measured (red) and simulated (blue) diode DC characteristics for all different geometries



NPN Junction Capacitance Modeling Result, all devices
Measured (red) BE and BC junction capacitance's and simulated (blue) ´BE and BC junction capacitances of all different geometries



Advantages of Direct Measurement Data Reduction

Advantages:

- Best fit to measured data is implicitly given
- Required measurement data processing is easy to implement
- Single transistor parameter extraction on a few edge devices
- Much lower effort and much faster parameter extraction

Requirements:

- Scaling rules for device characteristics
- Device characteristics must be measured for all devices at the same stimulus conditions
- Availability of enough device geometries and geometry variations

Summary

- Transforming measured data from a lot of different device geometries directly into per-area, per-perimeter and per-sheet normalized characteristics
  - Saves plenty of parameter extraction work
  - Gives automatically best fit to measurement data base
  - Provides inherent information on data distribution Statistics.
- All of the required normalized parameters can be extracted with few efforts using intelligent test structures.
- Using the presented simple test structures in production measurement setups (PCM) will lead to higher accuracy in determining important device parameters and will give excellent parameter statistics.

References

1. M. Schroeter, "TRADICA An Integrated Modeling Tool Linking Process And Circuit Design"Model Parameter Generation and Sizing of Integrated Bipolar Transistors. Manual Version 5. Nov. 2002.
2. J.Berkner, Kompaktmodelle fuer Bipolartransistoren, Expert-Verlag, Renningen, Germany, ISBN 3-8169-2085-3, Chapter 7.1.2

Statistics

Contents

- *Parametric Statistics* (iccapmhb)
- *Non-Parametric Statistics* (iccapmhb)

On the Web

- Articles
- Knowledge Center
- Technical Support
- Training
- Videos

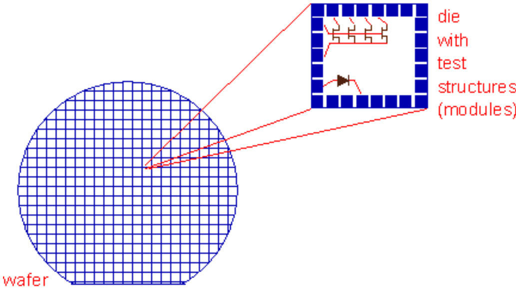
Parametric Statistics

Contents

- *Measurement Based Statistics* (iccapmhb)
- *Model Parameter Based Statistics* (iccapmhb)
- *Conclusions* (iccapmhb)

Measurement Related Statistics

A wafer, containing different dies with test structures on them, is tested. Each test die contains different MOS transistor geometries (large, narrow(s), short(s), small(s)), but also other test structures like sheet resistance or diodes. The wafer is loaded onto the chuck of the prober, and the prober then steps from test die to test die. All pads of the die are contacted together using a switching matrix. Selecting the specific connections, all different test structures (modules) on the die can then be measured independently. The following figure schematically represents the type of measurements used in this section:

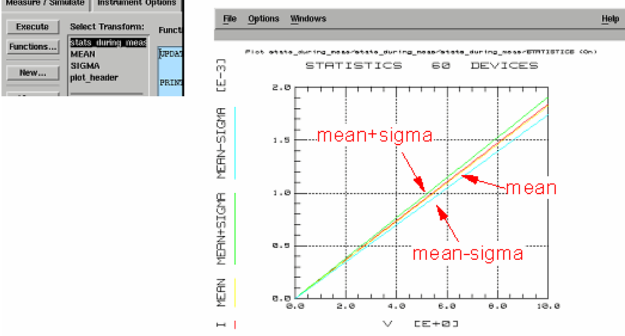


Wafer Scheme of a die containing MOS test transistors, sheet resistors, and diodes modules

Measurement related statistics can be performed by two different methods: either building the mean and sigma value of measurement data during the measurements itself (on the fly), or to firstly perform the measurements, store the results in an IC-CAP .mdm file and do the mean and sigma analysis based on the whole data of the .mdm file. Examples can be found at the following location:

- demo_features/9iccap4experts/7statistics/1meas_data_stats/stats_during_meas.mdl
- demo_features/9iccap4experts/7statistics/1meas_data_stats/stats_on_mdm.mdl

Referring to the IC-CAP file stats_during_meas.mdl, the measurements are performed, and, provided they were valid, the data is added to the existing mean and sigma data array. The statistics plot then shows the actual measurement and the updated mean and mean+-sigma plots. The following figure shows such a result, based on the measurement of resistor modules on 60 dies.



Measurement-based 'on the fly' mean and sigma statistics

Since after every measurement, the user is asked if the measurement was valid (can of course also be done by some macro), there are no outliers (bad measurement data) included in this type of measurement statistics.

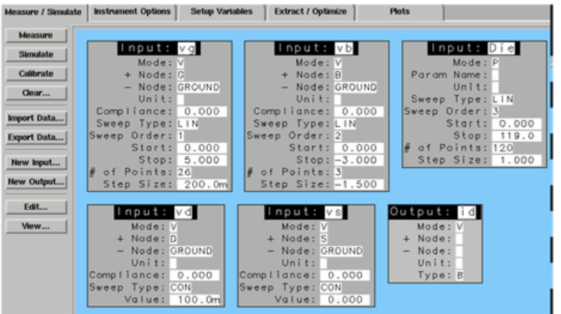
When all the test structures have been measured, a set of model parameters can be extracted from the mean data trace, as well as the mean+-sigma measurement data traces.

The other method of measurement based statistics, which is covered in the second IC-CAP file mentioned above, is related to already performed and stored measurements. In this case, IC-CAP has controlled the wafer prober, a switching matrix if necessary, and performed the measurements of different transistors. During that procedure, all the data have been stored in an IC-CAP .mdm file.

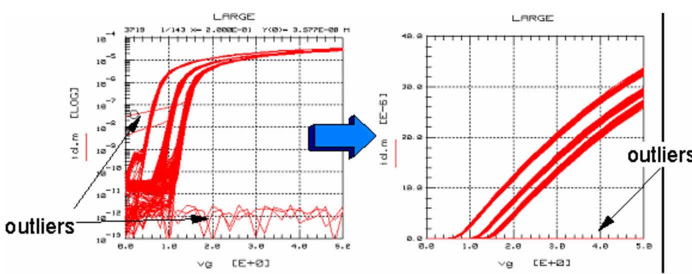
For an example about such an automated measurement procedure, see file demo_features/9iccap4experts/6wafertest/prober_matrix_mdm.mdl

In order to visualize the acquired measurement data, the complete .mdm file is re-imported into the setups of DUT mdm_data of our model file. This is depicted in fig.3 (top). For this example, module containinga a large MOS transistor was selected: Both, vG and vB are swept, while vD is small (100mV). The 3rd order sweep is then the die number.

Such type of MOS measurements is usually performed for the modeling of the threshold voltage and the mobility parameters of the MOS model.

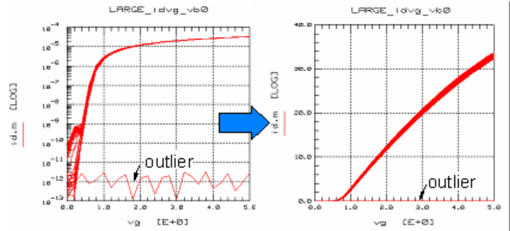


Setup to re-import all the measurement data for the large transistor of all dies



Visualization of the data including the outliers

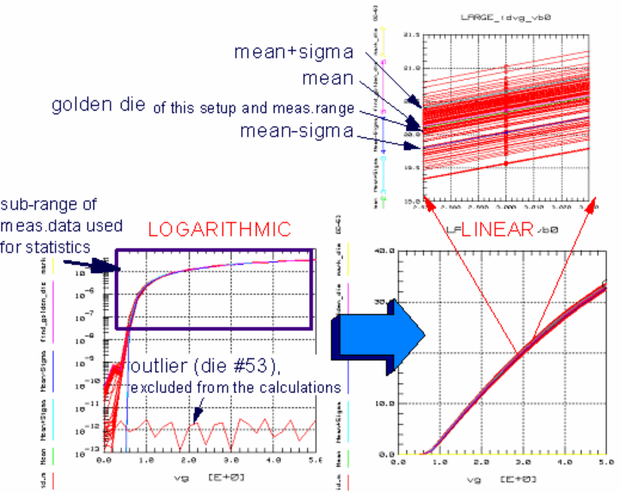
In order to simplify the statistics, we will reduce the bias sweep to only 1st order. I.e. we will re-import only measurement data belonging to a constant vB. In this way, we have only one single bias sweep (vG), and the 2nd order sweep is simply the die number. Fig 3 shows the resulting plots.



Selecting a single Bulk voltage out of the measurement data

In the next step, the measurement outliers have to be marked, such that they are not considered when calculating the mean and sigma curves. Back to the IC-CAP file, this is done by the transform 'mark_outliers'. It asks the user to specify the die (transistor) number of the outlier(s), and then enters a flag into the setup variable outliersi . In order to identify the outlier die numbers, click with the middle mouse button on the specific trace of fig.4.
IC-CAP returns a data index, consisting of firstly the total point index, followed by a number, like e.g. 19/25. This number means: 19th first-order bias sweep (vG) of die no.25. This no.25 is what has to be entered when running transform 'mark_outliers' in order to set the flag.

After the transform 'mark_outliers' was run, we are now ready to calculate the mean and sigma values of our measurement data array. This is done by the transform 'box_data_statistics', and finally 'Mean' and 'Sigma'. After these programs have been executed, program 'find_golden_die' identifies the very die which is closest to the calculated mean data. This is depicted with the plots of fig.5.



Calculated mean, mean+-sigma and identified 'golden die', related to measurements within the marked sub-range, and with eliminated outlier die #53, based on the 'large' transistor measurement data

About The Difficulties To Identify The Golden Die Using Measurement-Related Statistics

So far, the basically two different methods of measurement-related statistics have been introduced. Let's now consider the task to identify the real 'golden die', containing those transistor modules which should be used for model parameter extraction. Back to fig.5, it can be shown that the identified large 'golden die' transistor number is heavily dependent on the user-selected subrange of measured data. Referring to the data in the actual model file setup 'large_idvg_vb0', we obtain the following identified die number depending on the box limits defined by X_LOW and X_HIGH (defining the two Setup Variables and running program RUN_ALL):

X_LOW	X_HIGH	die #
0.5	6	74
0.5	4	74
1	3	5
2	3	5
4	5	9

In other words: based on a single Bulk voltage bias of the transconductance curve and based on different sub-ranges of measured data, we can either identify die #74, #5 or #9 to be the representative 'golden die' for later parameter extractions. If we would apply this procedure also to the other Bulk bias voltages, we would get even more die numbers which could be candidates for the mean 'large', i.e. the 'golden' large MOS transistor for later model parameter extraction.

This confusing picture becomes even more weird, when taking also the other measurement ranges of a specific transistor module into account. For this experiment, we will consider the data of the short transistor, because it is with this device where the demo file also has output characteristics data measured in addition to the so far discussed transfer curves.

Before we start, we check all available plots of this short transistor for outliers. All available plots, linear and logarithmic ones, of both measurement domains and for all bias conditions (setups meas_data_stats_short/short_idvg_vbxx and short_idvd;vgxx) , are inspected. This gives the following outlier die numbers:

25, 26, 27, 28, 29, 61, 71

Neglecting these outlier dies, we get the following table of results, when applying measurement-related statistics to the data:

Setup	Note	X_LOW /*	X_HIGH /*	die #	% error
short_idvg_vb0 15 3	vB=0 vB=-1.5 vB=-3	all data >0.5 /1 all data >0.5 /1 all data >0.5 /1	all data >0.5 /1	89 89 95	1.77 6.54 7.03
short_idvd_vg1 2 3 4	vG=1 vG=2 vG=3 vG=4	all data all data all data all data	all data all data all data all data	95 95 96 96	1.28 0.52 0.68 1.07
short_idvg_vb0 15 3	vB=0 vB=-1.5 vB=-3	1 1 1	3 3 3	119 118 119	0.15 0.34 0.32
short_idvd_vg1 2 3 4	vG=1 vG=2 vG=3 vG=4	1 1 1 1	3 3 3 3	95 95 96 89	0.17 0.05 0.01 0.02
short_idvg_vb0 15 3	vB=0 1.5 3	2 2 2	4 4 4	16 119 119	0.02 0.05 0.12
short_idvd_vg1 2 3 4	vG=1 vG=2 vG=3 vG=4	2 2 2 2	4 4 4 4	95 95 96 89	0.08 0.10 0.01 0.01
short_idvg_vb0 15 3	vB=0 vB=-1.5 vB=-3	3 3 3	5 5 5	100 89 88	0.02 0.04 0.06
short_idvd_vg1 2 3 4	vG=1 vG=2 vG=3 vG=4	3 3 3 3	5 5 5 5	95 119 96 89	0.14 0.16 0.05 0.02

Note:
/* to be consistent, instead of Setup Variables, DUT Variables are used
/1 to exclude measurement limitations (pA resolution), only vG >=0.5 is used

From this table, we have now the choice to select either die no. 16, 88, 89,95, 96 100, 118 or 119 which might be considered as the 'golden' short transistor module. If we would then also take the large transistor modules, the narrows and the smalls into account, the spread of choosable 'golden dies' would even be enlarged!
Remember: candidates for the 'golden' large transistor die were die numbers 5, 9, 74, and these numbers do not show up with the candidate numbers of the short 'golden' die.

Therefore, we are now really hit by the question:
How can we solve this dilemma to identify the die which contains the typical, i.s. the 'golden' transistor modules LARGE_golden, SHORT_golden, NARROW_golden etc.?

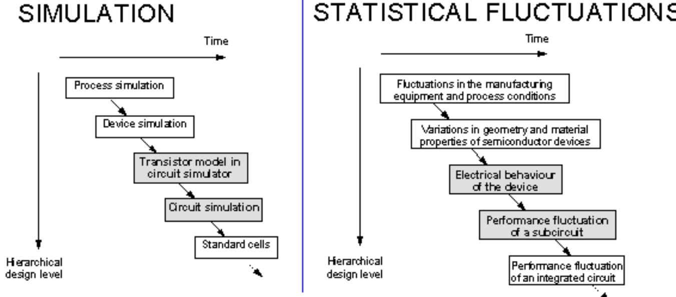
The only way out of this 'golden die' selection dilemma is non-parametric statistics, covered in the next topic.

Model Parameter Related Statistics

As we have seen in the preceding chapters, modeling of electronic components often means to determine quite a number of model parameters. This is especially true for MOS transistors and their spread of different geometries. As a consequence, the modeling department is under a permanent time pressure related to measurements, to model parameter extraction and, thus, to cost. Therefore, other approaches should be taken into account to respond to these challenges.

This is possible, provided we have enough statistical data in order to relate the big number of model parameters to a smaller number of dominant parameters. In case adding even production process data to this type of analysis, these process data then can become the dominant parameters. This means that, provided the process has not changed, and we get new process data, we are ready to calculate the modeling parameters with a given certainty of e.g. 90% out of these process data without a need for further modeling measurements or parameter extractions.

This procedure is called 'model parametric statistics'. Additionally, this type of statistical analysis can, if combined with process and device simulation results, also predict the performance of transistors which are based on new designs, without manufacturing them first.



Semiconductor modeling hierarchy from process simulation to model libraries

The hierarchy of simulation tools or statistical influences is illustrated in the above figure. The lowest simulation (statistical analysis) hierarchy is the process simulation, which models the different manufacturing steps in a semiconductor production. Results of these process simulations are given in form of doping profiles and device geometries of a single transistor. With this information, the device simulator calculates the electrical behavior of this single transistor, when external voltages or currents are applied to it. The results are typically given as a current vs. voltage. From these curves, the model parameters can be extracted for simulation of the transistor in a circuit simulation program like SPICE. These single transistors are the basis for further more complex circuits e.g. digital standard cells.

Let's come back to the model parameter related statistics. In order to find the dominant parameters, be it process or model parameters, out of the big total number of parameters, a big amount of extractions has to be performed first. As a general rule, for N model parameters, there should be at least 10*N parameter sets been extracted, better are N*N parameter sets. For a MOS transistor, this may be up to thousands of parameter sets.

The most commonly used tools for model parameter statistics are factor or principal components analysis. While basically describing the same effect, their difference lies in the definition of the fluctuation contexts. Interesting enough, both methods have not been developed for technical applications originally, but instead for social science. Their main objective was to 'measure' non-measurable properties like intelligence etc. In such statistical experiments, test persons are asked to fulfill some tests like solving mathematical problems, puzzles or other tests. From the test results, it is attempted to measure the non-measurable property 'intelligence'.

Principal Component Analysis

In order to discuss briefly the properties of the two main statistical methods, we will start with some general infos about the principal components analysis (PCA). Beginning with N correlated statistical variables (model parameters), this method calculates a user-defined amount of m<n uncorrelated variables: the principal components. These new principal components describe the statistical spread of the original n correlated variables. The important fact is, however, that in practice, less than n principal components are required to describe the original n correlated variables.

> **Note**
> These principal components should be considered as virtual variables, which do not necessarily represent physical parameters, and which, therefore, cannot be measured physically.

Definition wise, the principal components can be considered as a linear function of the original correlated variables X:

$$PC_m = w_{m,1}X_1 + w_{m,2}X_2 + \dots + w_{m,n}X_n$$

with:

- PC non-measurable principal component, m<n
- w weighting factors
- X normalized measurable, possibly correlated statistical variable
  (here: the n model parameters)

If the variables X have a wide range of values, they should be first normalized as follows:

$$X_n = \frac{x - \mu_x}{\sigma_x}$$

with:

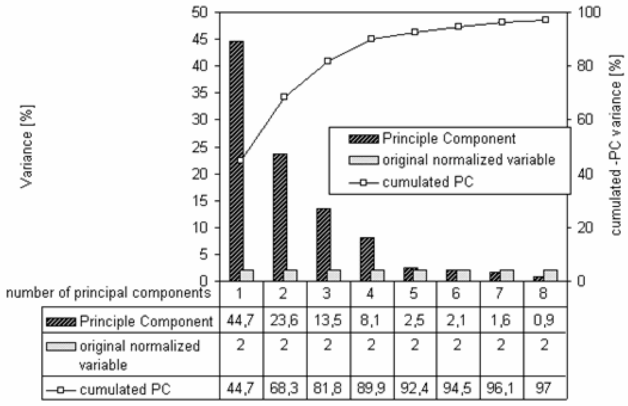
- !mrps_equation2.1.gif!mean of the statistical variable
- !mrps_equation2.2.gif!standard deviation of the statistical variable
- x statistical variable
- !mrps_equation2.3.gif!normalized statistical variable

After this normalization, all the variables have the mean value 0 and a standard deviation of 1. This normalization is especially important for device model parameters, because their numerical values can range from very low values (e.g. saturation current IS in the 10-20 range) to very big ones (doping concentration NCH in the 10 ^{16} range).

In order to calculate the weighting factors w in (1), the correlation matrix of the normalized statistical variables (model parameters) has to be calculated first. Out of this correlation matrix, the eigenvector U and the eigenvalues Λ can be derived. After that, the principal components can be calculated by:

$$PC = \Lambda^{-\frac{1}{2}} U^{-1} X_n$$

The following figure visualizes the effect of the principal components, related to modeling parameters. Starting point were 50 MOS model parameters. Each of them, because normalized, would explain 1/50 = 2% of the total variance of the MOS transistors. After the principal component analysis has been performed, the first principal component can explain 44.7 % of the variance of the original set of model parameters! The second one explains another 23.6 %. Both together can explain already 68.3 %. And if we take the first 5 principal components into account, over 92% of the MOS transistor variance can be explained.



number of principal components	1	2	3	4	5	6	7	8
Principle Component	44,7	23,6	13,5	8,1	2,5	2,1	1,6	0,9
original normalized variable	2	2	2	2	2	2	2	2
cumulated PC	44,7	68,3	81,8	89,9	92,4	94,5	96,1	97

Comparing the variance of the original normalized 50 model parameters and the 8 obtained principal components>

This analysis result can also be interpreted that among the originally 50 model parameters, there are 5 dominant model parameters. Therefore, the remaining last step is to associate the dominant model parameters with the determined principal components.

Factor Analysis

The other method to reduce the number of model parameter is the factor analysis. While the so far discussed principal components method assumes that the total variance of a statistical variable contributes to the principal component, the factor analysis is based on the assumption that the variance of a statistical variable is composed of two parts: one part is correlated with the variances of the other variables, while the second part is related only to the variable itself.

$$X_m = v_{m,1}F_1 + v_{m,2}F_2 + \ldots + v_{m,n}F_n + e_m$$

with:

- F non-measurable common factor
- e variable specific factor
- X normalized observable statistical variable
- v factor weighting

A little example should explain this idea:
In production, the weight and the diameter of a metal cylinder are measured. It is obvious that these two parameters are correlated, because a bigger diameter means immediately a bigger weight. However, each property is measured by another measurement procedure, which, on the other hand, is not ideal and adds by itself some error to the measurement result. This measurement specific error is not correlated with the other measurement parameter.

Like with the principal component analysis, there are less factors F than observable statistical variables X. And, again, some few dominant factors can cover the major part (usually >90%) of the original model parameter fluctuation.

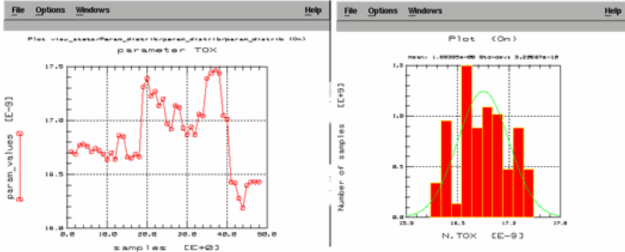
Practical Example

Referring to the modeling of MOS transistors from a single wafer or also from different wafers from different production lots, we need to create a spreadsheet of extracted model parameters first. It is very important to note that the spreadsheet should be based on directly extracted model parameters and not more than required on parameter optimization. As already stated in the previous chapter about measurement related statistics, finding model parameters by optimization may add so-called optimizer noise to the model parameter values. This is due to the fact that, depending on the starting conditions before the optimization, the mathematical curve fitting between measured and simulated data can be the same for different sets of model parameter values.

After the parameter spreadsheet is filled out, we have to check the validity of the data first.

> **Note**
> best statistical results are obtained if the spread of all the model parameter values is Gaussian. Of course, in reality, this is difficult to achieve, and therefore a pre-inspection of the data is very important.

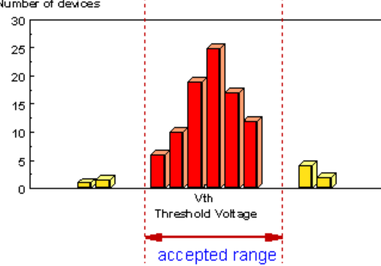
First, we should try to get an overview about the trend of the extracted model parameters versus the die number. Usually, this die number is related to certain manufacturing dates, lot numbers etc. Although the total spread of the model parameter for all dies on all wafers may be quite Gaussian, this graphical analysis can give us some insight about the process fluctuations included in our statistical analysis. In the example on the left of fig.3, it is obvious that our data is affected by changes in the production process. The parameter TOX, given in that plot, indicates that the process (oxide thickness) has been adjusted at least four times within the time slot related to our extracted model parameters. If, however, we only had plotted the histogram of all the TOX data, this property would not have become obvious.



Analyzing the distribution of the spreadsheet's model parameters versus production fluctuations

> **Note**
> an example for such a graphical analysis can be found under *demo_features/9iccap4experts/tstatistics/2param_stats/2visu_statistics/visu_stats.mdl* and running macro PARAM_DISTRIBUTION

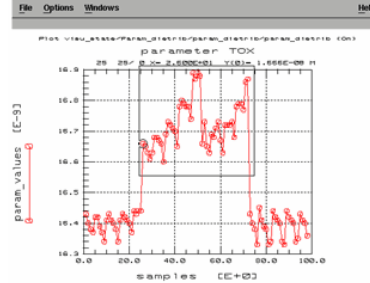
In a second step, when applying histogram plots to the parameter data directly, we can check the model parameter spread itself, independently of the manufacturing date. This is done in the following figure:



Distribution of a model parameter (VTH) to check for outliers>

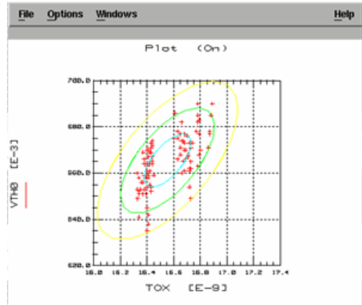
In the above figure, the outliers represent so-called spot-defects on the wafer: shorts due to dust particles, too big leakage currents etc. The inner part of the distribution, marked with 'accepted range', is what we want to include in our statistical analysis. Because this range represents the effects, which we are interested in for our statistical analysis: tolerances like mask positioning, implant effects, temperature effects during production, sputtering energy etc. Therefore, the spreadsheet rows, containing outliers, have to be deactivated before we apply statistical analysis.
This can be done by activating the actual parameter column and clicking Date/Date Filter.

An interesting aspect of model parameter analysis is to check for multi-modal parameter distributions and how to avoid them. Fig. 5 gives an example, based on the shipped example file examples/icstat/bsim3.sdf . Here, it becomes obvious that the data stems from different test dies (5 dies on different wafer locations, the small stepping) and that the wafers stem from different production lots (the big steps)



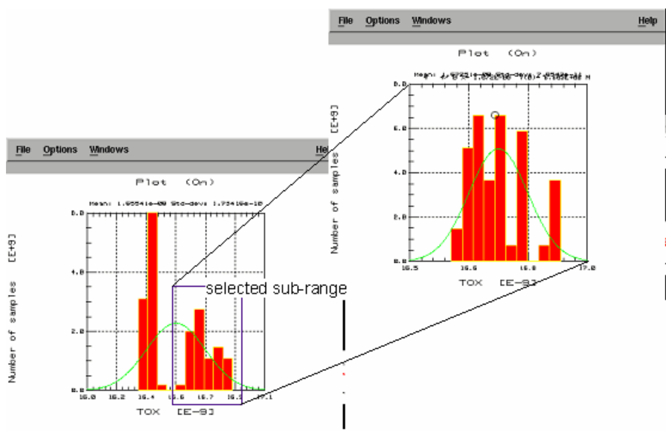
Bi-modal distribution of the model parameter TOX

Since TOX is always related to the MOS transistor's VTH, the bi-modal parameter distribution suspected in fig.5 becomes also visible with the scatter plot of these two parameters. This is depicted in fig.6.



Scatter plot of the bi-modal distribution of the MOS parameters TOX and VTH

Now, the suspicion of bi-modal parameter distribution has been proofed. Therefore, we display the histogram plot of parameter TOX and select the upper sub-range for further statistical analysis. The new histogram of this sub-range is given in the upper right side of fig.7. It represents a much nicer Gaussian distribution than the original data set.

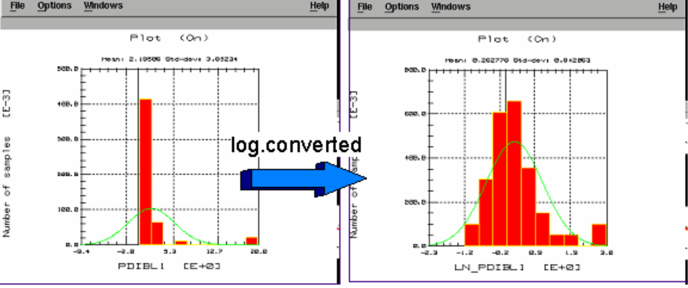


Eliminating bi-modal parameter distributions using the histogram plot

In a next step, we have to check the distributions of the other model parameters of the selected sub-range. In this step, the main objective is to non-linearly transform the model parameter values in order to obtain histogram plots with as similar as possible Gaussian distributions before we can start with the very statistical analysis.

Note
for the non-parametric analysis of the previous chapter (measurement related statistics), the Gaussian distribution of the model parameters was not required.

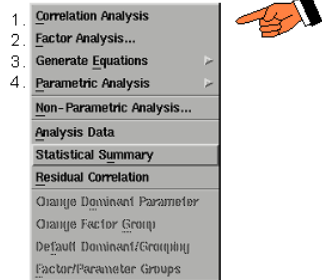
MOS parameter PDIBL is a candidate for such a check. As can be seen in fig.8, the distribution of this parameter looks still quite confusing (although we have already deactivated the bi-modal TOX-VTH distribution). After a logarithmic conversion of the parameter values, however, we obtain a very nice Gaussian distribution.



Converting the model parameter non-linearly in order to achieve a distribution as close as possible to Gaussian

After all, the model parameters have been checked and possibly transformed for Gaussian distribution, so that a final outlier elimination of the model parameters can be performed. From the 'Statistical Summary' window of the statistics package, we obtain the mean and standard deviation of the (transformed) model parameters, and, using a small macro in IC-CAP, which can access the rows and columns of the statistics parameter spread sheet, we can deactivate all those dies whose parameter values are outside the mean?sigma (or similar) limits.

Then, we are finally ready to go for the model parameter related statistics, and, after that, for the parametric statistical analysis. We click the menu pick 'Analysis', and, as can be seen in fig. 9, we simply follow the procedures top-down:

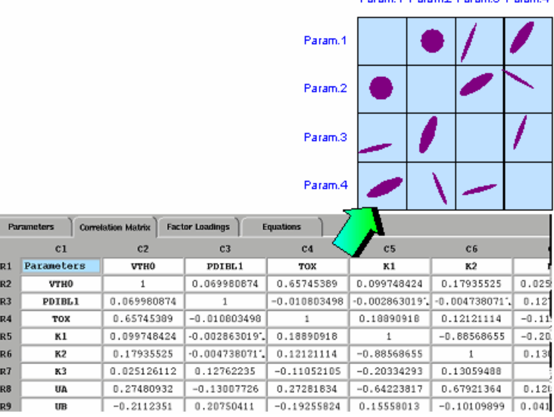


Steps for the model-parameter related statistics

Description of the Steps

Correlation Analysis:

This step returns the model parameter correlation matrix. The correlation analysis has to be performed first, giving the relationship among the model parameters. Fig.10 shows the table and also a schematic visualization for easier interpretation. All the correlation factors are printed and all the parameter relationships are plotted against each other. If there is a circle-like 'cloud', then there is no relationship among the two parameters. If the 'cloud' goes upwards, then the dependency is a linear one with positive slope, and if it tends downwards, the dependency goes with a negative slope. If we have several 'clouds' within a specific sub-plot of fig.10, this would hint to another bi-modal parameter distribution which should be eliminated before further proceeding with the statistical analysis.



	C1	C2	C3	C4	C5	C6	
R1	Parameters	VTH0	PDIBL1	TOX	K1	K2	
R2	VTH0	1	0.069980874	0.65745389	0.099748424	0.17935525	0.025
R3	PDIBL1	0.069980874	1	-0.010803498	-0.002863019'.	-0.004738071'.	0.12
R4	TOX	0.65745389	-0.010803498	1	0.18890918	0.12121114	-0.11
R5	K1	0.099748424	-0.002863019'.	0.18890918	1	-0.88568655	-0.20
R6	K2	0.17935525	-0.004738071'.	0.12121114	-0.88568655	1	0.13
R7	K3	0.025126112	0.12762235	-0.11052105	-0.20334293	0.13059488	
R8	UA	0.27480932	-0.13007726	0.27281834	-0.64223817	0.67921364	0.12
R9	UB	-0.2112351	0.20750411	-0.19255824	0.15558013	-0.10109899	0.041

Fig.10: correlation matrix and a schematized scatter plot matrix of the parameters.

Factor Analysis:

This step allows you to:

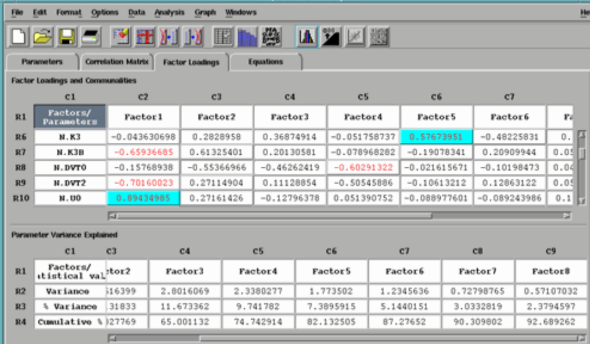
- Select the number of factors and the type of statistical analysis method (e.g. factor analysis, principal component analysis etc.).
- Returns the Factor Loadings spreadsheet containing the factors, their relationship to the parameters and another spreadsheet containing the factor variances and, last not least, the cumulative factor variances.

Then, applying factor analysis (step 2), we select the number of factors (here 8) and get the factor loading matrix, see fig.11. First of all, we should inspect the distribution of the factors. It is important to have, based on our 50 model parameters, not more than ~10 dominant factors which can explain >90% of the original model parameter variations, and also that the first factor is big compared to the second and so on. Topic 1 is ok, because, at the lower right in fig. 11, the cumulative percentage of the selected 8 factors is 92.7%. A further inspection of the distribution of the factors, depicted in fig.12, also indicates that the basis of our statistical analysis, the quality of the model parameters, was quite good. This can be seen from the fact, that the first factor is represented by a big variance value, followed by rapidly decreasing values of the next factors. This is an indication that there are real dominant model parameters present in our parameter spreadsheet.

(see IC-CAP demo file demo_features/9iccap4experts/tstatistics/2param_stats/2visu_statistics/visu_stats.mdl and running macro CUMUL_PARAM_VARIANCE)

Another important check at this point is the data in the upper table of fig.11. Here, we can identify the association of the factors with the original model parameters. We should check, if the factors, selected by the mathematical algorithm, are the wanted dominant model or process parameters or not. The user can change these associations in order to help with achieving this goal.

Hint: it has been shown that, when applying optimization for model parameter extraction rather than the recommended direct model parameter extraction, the automatically selected factor associations would link the factors to 2nd order model parameters and not to the desired 1st order parameters!



File Edit Format Options Data Analysis Graph Windows Help

Parameters | Correlation Matrix | Factor Loadings | Equations

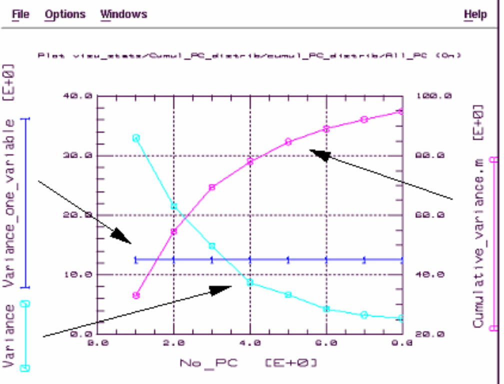
Factor Loadings and Communalities

	C1	C2	C3	C4	C5	C6	C7	
R1	Factors/ Parameters	Factor1	Factor2	Factor3	Factor4	Factor5	Factor6	Fa
R6	N.K3	-0.043630698	0.2828958	0.36874914	-0.051758737	0.57673951	-0.48225831	0.
R7	N.K3B	-0.65936685	0.61325401	0.20130581	-0.078968282	-0.19078341	0.20909944	0.0
R8	N.DVT0	-0.15768938	-0.55366966	-0.46262419	-0.60291322	-0.021615671	-0.10198473	0.0
R9	N.DVT2	-0.70160023	0.27114904	0.11128854	-0.50545886	-0.10613212	0.12863122	0.0
R10	N.U0	0.89434985	0.27161426	-0.12796378	0.051390752	-0.088977601	-0.089243986	0.1

Parameter Variance Explained

	C1	C3	C4	C5	C6	C7	C8	C9
R1	Factors/ itistical val	tor2	Factor3	Factor4	Factor5	Factor6	Factor7	Factor8
R2	Variance	16399	2.8016069	2.3380277	1.773502	1.2345636	0.72798765	0.57107032
R3	% Variance	31833	11.673362	9.741782	7.3895915	5.1440151	3.0332819	2.3794597
R4	Cumulative %	27769	65.001132	74.742914	82.132505	87.27652	90.309802	92.689262

Factor loading matrix after having applied principal components analysis with 10 factors



Cumulative variance of the calculated factors: the first factor alone can explain

already ~35% of the model parameter variance, while the first three factors together can explain more than 65%, and the 8 factors altogether explain more than 92% of the variance.

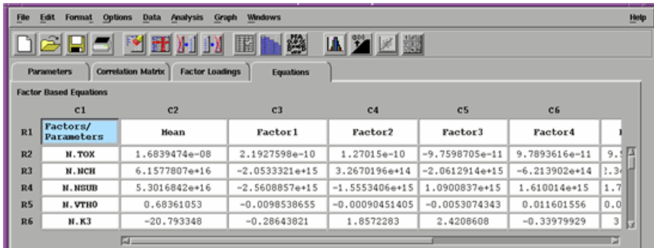
Generate Equations

submenu /Factors: returns the Equation spreadsheet, containing the context between the

parameters and the normalized factors

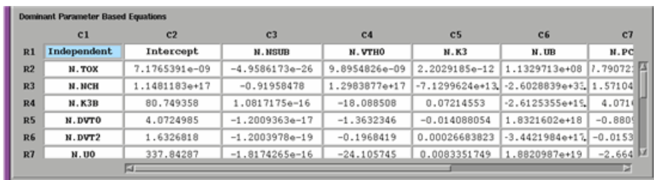
submenu /Dominant Parameters: calculates the set of the new, dependent model parameters based on the independent model parameters.

With this step, we have nearly reached our goal: to calculate the new model parameters out of the factors, so that they have values and a spread as similar as possible to the original, really extracted ones. First, after selecting the submenu /Factors, we obtain a list of the new model parameters as a function of the factors. Fig.13 shows that list.



	C1	C2	C3	C4	C5	C6
R1	Factors/ Parameters	Mean	Factor1	Factor2	Factor3	Factor4
R2	N.TOX	1.6839474e-08	2.1927598e-10	1.27015e-10	-9.7598705e-11	9.7893616e-11
R3	N.NCH	6.1577807e+16	-2.0533321e+15	3.2670196e+14	-2.0612914e+15	-6.213902e+14
R4	N.NSUB	5.3016842e+16	-2.5608857e+15	-1.5553406e+15	1.0900837e+15	1.610014e+15
R5	N.VTH0	0.68361053	-0.0098538655	-0.00090451405	-0.0053074343	0.011601556
R6	N.K3	-20.793348	-0.28643821	1.8572283	2.4208608	-0.33979929

Finding the context between the determined factors and the model parameters.



	C1	C2	C3	C4	C5	C6	C7
R1	Independent	Intercept	N.NSUB	N.VTH0	N.K3	N.UB	N.PC
R2	N.TOX	7.1765391e-09	-4.9586173e-26	9.8954826e-09	2.2029185e-12	1.1329713e+08	[illegible]
R3	N.NCH	1.1481183e+17	-0.91958478	1.2983877e+17	-7.1299624e+13	-2.6028839e+3	1.57104
R4	N.K3B	80.749358	1.0817175e-16	-18.088508	0.07214553	-2.6125355e+1	4.071
R5	N.DVT0	4.0724985	-1.2009363e-17	-1.3632346	-0.014088054	1.8321602e+18	-0.880
R6	N.DVT2	1.6326818	-1.2003978e-19	-0.1968419	0.00026683823	-3.4421984e+1	-0.0153
R7	N.U0	337.84287	-1.8174265e-16	-24.105745	0.0083351749	1.8820987e+19	-2.664

Calculating the dependent model parameter based on the independent parameters

identified by the analysis depicted in the upper part of fig.11.
Then, after selecting the submenu Dominant Parameters, we get the final goal of our attempts: the list of the dependent model parameters and their linear context with the independent ones. The independent ones were identified with the factor loadings matrix in fig.11.

In other words, this last spreadsheet of our statistical analysis closes the loop between the model parameter <~~> factor context of fig. 13, and the factor <~~> dominant model parameter context of fig.11.

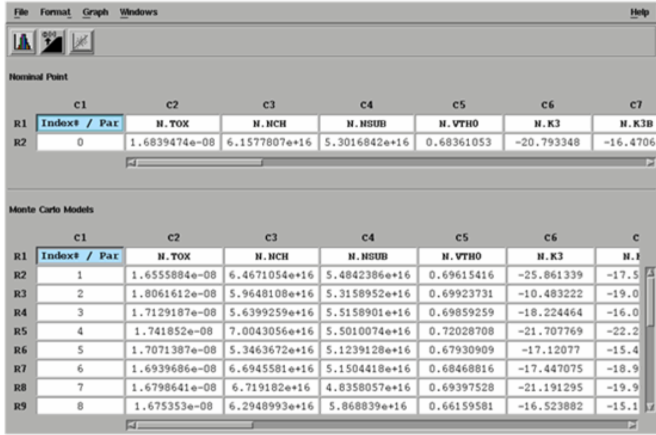
This ends the description of how to identify the dependent model parameters and of how to relate them to the independent ones.

But, we can also continue with our statistical experiment and apply methods, which can produce artificial complete model parameter sets with basically the same properties like the underlying old ones. Such a method could be used for design centering analysis.

Based on the previous statistical results, we can now also perform a parametric analysis (step 4 in fig.9). Parametric means, that we can use the results of our statistical analysis, and apply parametric techniques to predict sets of model parameters, which behave like the underlying original model parameter sets.

Parametric Analysis

submenu /Factor Equations: Asks for selecting the parametric analysis method, e.g. Monte Carlo Analysis, and for the
number of outcomes (model parameter sets).
Returns a new spreadsheet containing the calculated parameters, based on the statistical factors from above.



Nominal Point

	C1	C2	C3	C4	C5	C6	C7
R1	Index# / Par	N.TOX	N.NCH	N.NSUB	N.VTH0	N.K3	N.K3B
R2	0	1.6839474e-08	6.1577807e+16	5.3016842e+16	0.68361053	-20.793348	-16.4706

Monte Carlo Models

	C1	C2	C3	C4	C5	C6	C
R1	Index# / Par	N.TOX	N.NCH	N.NSUB	N.VTH0	N.K3	N.I
R2	1	1.6555884e-08	6.4671054e+16	5.4842386e+16	0.69615416	-25.861339	-17.5
R3	2	1.8061612e-08	5.9648108e+16	5.3158952e+16	0.69923731	-10.483222	-19.0
R4	3	1.7129187e-08	5.6399259e+16	5.5158901e+16	0.69859259	-18.224464	-16.0
R5	4	1.741852e-08	7.0043056e+16	5.5010074e+16	0.72028708	-21.707769	-22.2
R6	5	1.7071387e-08	5.3463672e+16	5.1239128e+16	0.67930909	-17.12077	-15.4
R7	6	1.6939686e-08	6.6945581e+16	5.1504418e+16	0.68468816	-17.447075	-18.9
R8	7	1.6798641e-08	6.719182e+16	4.8358057e+16	0.69397528	-21.191295	-19.9
R9	8	1.675353e-08	6.2948993e+16	5.868839e+16	0.66159581	-16.523882	-15.1

New set of model parameters generated by the factors of the principal component analysis

First, we specify which statistical method we want to apply to generate a spreadsheet of new model parameters (which are based on the determined dominant factors). Choosing Monte Carlo analysis for example, (which alters the dominant factors F1..Fn between -1 and +1 in fig.13), we get the new model parameter sets as shown in fig.15. These sets of parameters have basically the same statistical properties like the original sets of model parameter, i.e. the same mean and standard deviation, however, is has been calculated out of the dominant factors/parameters.

Parameter Based Statistics: Wrap-Up

This means that applying parameter related statistical analysis methods (factor analysis, principal component analysis) to device modeling, we are able to reduce the number of model parameters to the dominant ones. This works well, provided the spread of the underlying fabrication process has been described properly by eliminating outliers, eliminating multi-modal data fluctuations, and transforming model parameters to Gaussian distribution. Furthermore, we get a deeper insight into the process fluctuations and its stability.

Also, by applying parametric statistical methods like Monte Carlo Analysis to these techniques, we can also predict the performance of new, not yet produced components, based on the knowledge of the dominant factors of our production process and the expected new values of the dominant process parameters.

Literature

/1/: T. Gneiting, Handout of the Workshop on Statistical Modeling, Nov.4 1997,
HP Office Taufkirchen/Munich, Germany (in German)

/2/: IC-CAP 5.0, Manual Statistical Analysis, June 1997, HP Part no. 85190-90050

/3/: T. Gneiting, I. P. Jalowiecki, "The Prediction of Circuit Performance Variations for Deep Submicron CMOS Processes", IEEE International Symposium on Defect and Fault Tolerance in VLSI Systems, Boston/MA, November 1996

/4/: T. Gneiting, I. P. Jalowiecki, "The Influence of Process Parameter Variations on the Signal Distribution Behaviour of WSI Devices", IEEE Transactions on Components, Packaging and Manufacturing Technology, Part B, Vol. 18, No. 3, August 1995

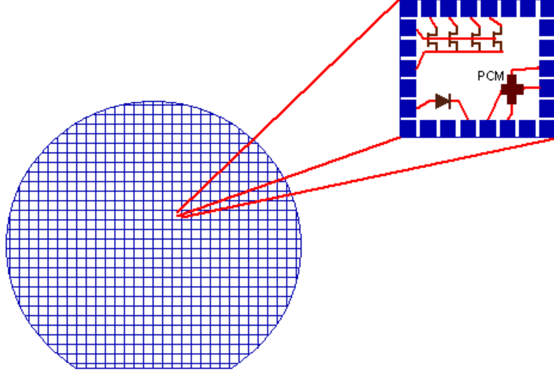
/5/: Chow, "Statistical Circuit Simulation of a Wideband Amplifier: A Case Study in Design for Manufacturability", Hewlett-Packard Journal, October 1990

/6/: J. Power, D. Barry, A. Mathewsen, W. Lane; "Worst Case Simulation Using Principal Component Analysis Techniques: An Investigation", Microelectronic Engineering 15-1991, pp 213-216, Elsevier

Conclusions

After the two basic statistical principles, measurement related and model parameter related statistics have been introduced, the question still remains, how to combine them in order to save time without sacrificing accuracy.

A possibility might be to combine the non-parametric statistics, however applied to process data (Process Control Measurement data) instead of measurement data, with the model parameter related, factor analysis based statistics. In this case, the PCM data from production, where the required big amount of data is acquired anyhow, are written to an ASCII file and then loaded into the statistics spreadsheet. Then, applying non-parametric statistics to these PCM data, we can identify the 'golden die' and the boundary dies (from the dies containing the PCM modules). Provided these dies contain, besides the PCM test modules, also the required transistor modules (see figure below), it is exactly those transistors which are then measured and modeled conventionally. The obtained model parameter sets are then exported one after the other to a new statistics spreadsheet, and a factor analysis based statistics can be applied. This method would automatically exclude outliers, and generate reliable model parameter sets, so that model parameter filtering should be nearly not necessary.



Measurements of the test modules for PCM and Modeling

Applying the non-parametric statistics to PCM data can also be used to identify those few dies which should be used for model parameter extraction, if the absolute minimal number of measurements is requested in order to get just the 'golden model parameter set' and a few boundary model parameter sets.

Publications

David McFeely, Dung Pham: Generating Statistical Models in IC-CAP, Hewlett-Packard Circuit Modeling Seminar, 1992.

Farshid Iravani, Masa Habu, Ebrahim Khalily: Statistical Modeling Tools, Methods and Applications for Integrated Circuit Manufacturability, Hewlett-Packard Circuit Modeling Seminar, 1994.

Dan Stoneking: Statistical Circuit Design and IC-CAP's Non-Parametric Boundary Analysis, Proceedings of the 4th European IC-CAP User Meeting, Berlin/Germany, Oct.1.-2., 1997

J.A. Power et al: Relating Statistical MOSFET Model Parameter Variabilities to IC Manufacturing Process Fluctuations Enabling Realistic Worst Case Design, IEEE Transactions on Semiconductor Manufacturing, August 1994.

J. Edward Jackson: A Users Guide to Principal Components, Wiley, 1991.

J.C. Zhang and M.A. Styblinski: Yield and Variability Optimization of Integrated Circuits, Kluwer Academic Publishers.

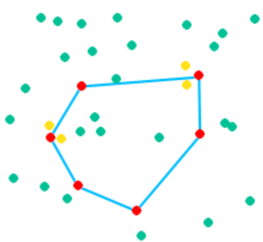
Sam Kachigan: Multivariate Statistical Analysis: A Conceptual Introduction, Radius Press

A very detailed paper on principal components analysis and correlation rotation is:
K.McKarthy, A.Mathewson, Extraction of Statistically Valid Parameters for IC Design, IEEE International Conference on Microelectronic Test Structures ICMTS 1999 Tutorial Short Course, March 15-18,1999, Gothenburg, Sweden.

Non-Parametric Statistics

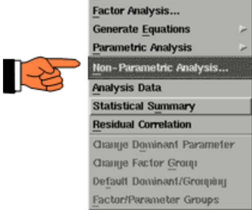
Contents

- Introduction to Non-Parametric Statistics
- Identify the 'golden device' and the 'limits devices'
- Non-Parametric Boundary Analysis
- Improving The Manufacturability Of Electronic Designs



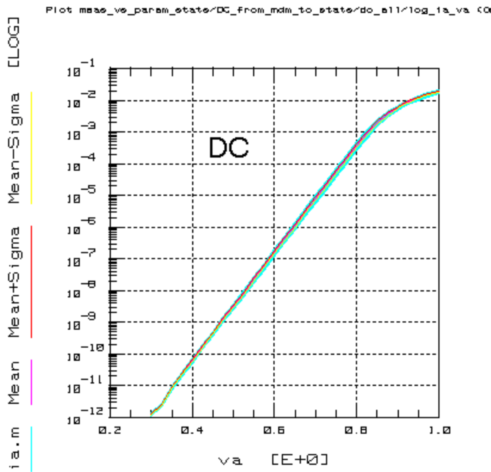
Introduction to Non-Parametric Statistics

Non-parametric statistics is a method which identifies in the multi-dimensional model parameter room that die which is closest to the center of that room: the 'golden die'. Therefore, this method takes into account all measurement setups and overcomes the 'golden die' identification problems described in the above chapter. Of course, with this method, it is required that the model parameters for all the measured setups have been determined accurately.

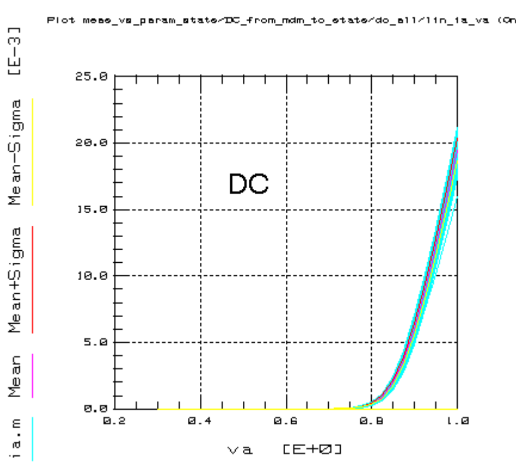


Let us refer to IC-CAP file
demo_features/9iccap4experts/7statistics/meas_vs_param_stats.mdl

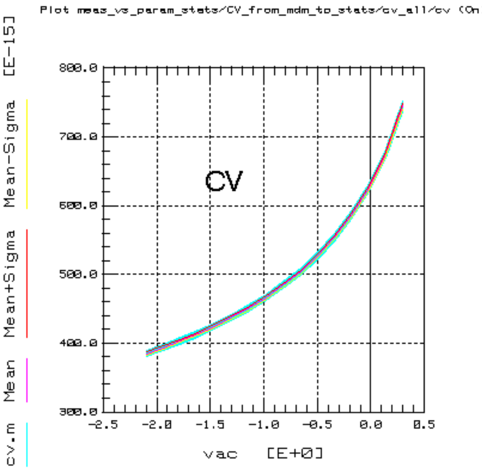
Again, we try to simplify the things as much as possible. Therefore, for this experiment, we choose a diode for the statistical analysis. Only two types of measurements, DC and CV, are used, over a total of 98 dies. Outliers have been eliminated from the measured data, before they were stored to an IC-CAP .mdm file. Fig. 6 depicts the results:



DC characteristics of 98 diode modules



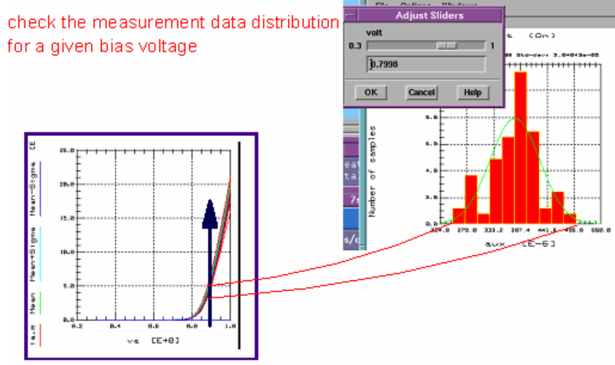
CV characteristics of 98 diode modules



Note
Back to the IC-CAP file, setup 'dc_all' and 'cv_all' contain the re-imported .mdm data, and in both setups, transforms 'meas_data_statistics', 'Mean' and 'Sigma', are used to perform the measurement-related statistics. Transform 'find_golden_die' again identifies the die which is closest to the calculated mean data.

Like with the example about MOS transistors, we have the same dilemma: for DC, it is the diode module on die 59 which is closest to the mean of all DC data (0.45% error), while for CV, it is rather the diode module on die 5 with an error of 0.03%.

As stated above, the goal of this section is to export the model parameters to the statistics spreadsheet. Before we do that, the measurement data distribution can be analyzed. This can be done with transform 'check_distrib'. This transform cuts through the die distribution for a given diode bias voltage and displays the result. Fig. 7 gives an example for V~0.8V.



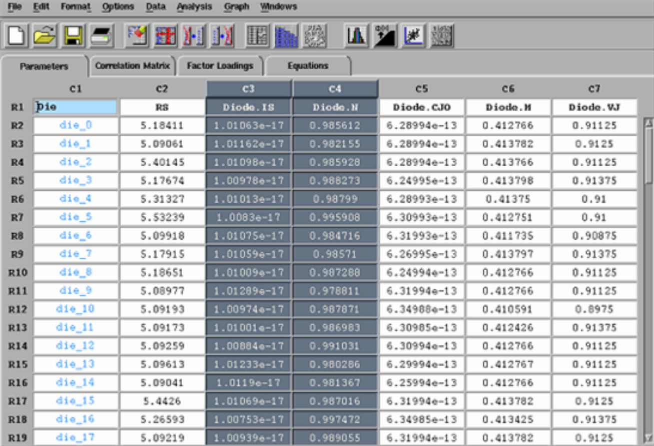
Checking the measurement data distribution for a given DC bias voltage

After this final check has been performed, we are now ready to extract the model parameter for both, DC and CV, and to export them into the statistics spreadsheet. This is performed by macro 'mdm2stats'.

It should be mentioned that it is especially important to use as much as possible 'direct extraction' instead of parameter optimization. 'Direct extraction', where the model parameters are calculated very accurately and directly from the measured data, produces always a clear and direct relationship between the measured data and the model parameters. Optimization, however, can add so-called 'optimization noise' to the model parameters. This effect comes from the fact that different starting conditions for the parameter values will lead to possibly the same fitting quality, however with slightly different final parameters. And this slight variations in the final model parameters can affect statistical analysis very badly. Modeling people call this effect 'optimizer noise'. Therefore, again, it is important to use rather the 'direct parameter extraction' method. (This is done with the extractions of the actual model file).

After the macro has been executed, the model parameter spreadsheet is filled up, see fig.8.

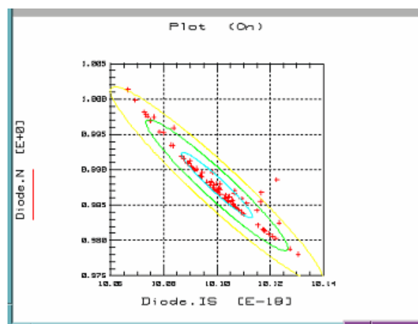
Note
a complete IC-CAP .sdf file of the extracted DC and CV diode parameters can be loaded into the statistics package from
demo_features/9iccap4experts/7statistics/meas_vs_param_stats_full.sdf



	C1	C2	C3	C4	C5	C6	C7
R1	Die	RS	Diode.IS	Diode.N	Diode.CJO	Diode.M	Diode.VJ
R2	die_0	5.18411	1.01063e-17	0.985612	6.28994e-13	0.412766	0.91125
R3	die_1	5.09061	1.01162e-17	0.982155	6.28994e-13	0.413782	0.9125
R4	die_2	5.40145	1.01098e-17	0.985928	6.28994e-13	0.413766	0.91125
R5	die_3	5.17674	1.00978e-17	0.988273	6.24995e-13	0.413798	0.91375
R6	die_4	5.31327	1.01013e-17	0.98799	6.28993e-13	0.41375	0.91
R7	die_5	5.53239	1.0083e-17	0.995908	6.30993e-13	0.412751	0.91
R8	die_6	5.09918	1.01075e-17	0.984716	6.31993e-13	0.411735	0.90875
R9	die_7	5.17915	1.01059e-17	0.98571	6.26995e-13	0.413797	0.91375
R10	die_8	5.18651	1.01009e-17	0.987288	6.24994e-13	0.412766	0.91125
R11	die_9	5.08977	1.01289e-17	0.978811	6.31994e-13	0.412766	0.91125
R12	die_10	5.09193	1.00974e-17	0.987871	6.34988e-13	0.410591	0.8975
R13	die_11	5.09173	1.01001e-17	0.986983	6.30985e-13	0.412426	0.91375
R14	die_12	5.09259	1.00884e-17	0.991031	6.30994e-13	0.412766	0.91125
R15	die_13	5.09613	1.01233e-17	0.980286	6.29994e-13	0.412767	0.91125
R16	die_14	5.09041	1.0119e-17	0.981367	6.25994e-13	0.412766	0.91125
R17	die_15	5.4426	1.01069e-17	0.987016	6.31994e-13	0.413782	0.9125
R18	die_16	5.26593	1.00753e-17	0.997472	6.34985e-13	0.413425	0.91375
R19	die_17	5.09219	1.00939e-17	0.989055	6.31994e-13	0.413782	0.9125

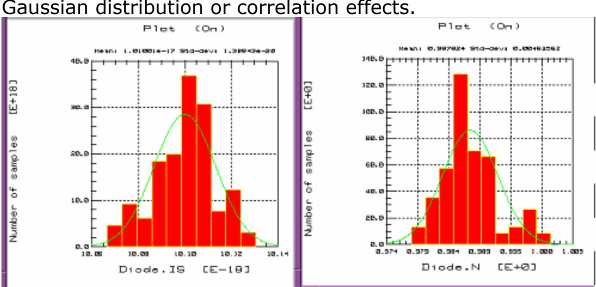
Filled statistics parameter spreadsheet

For the scatter plot analysis depicted in fig.9, the parameter columns Diode.IS and Diode.N are already marked.



Scatter plot of the parameters Diode.IS and Diode.N

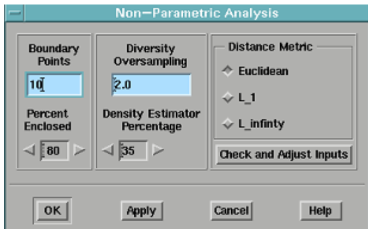
As shown in figures 9 and 10, as well as in table 1, different data analysis methods can be applied to the extracted model parameter sets of the 98 dies, in order to check for Gaussian distribution or correlation effects.



histogram plots of the distribution of model parameters Diode.IS and Diode.N

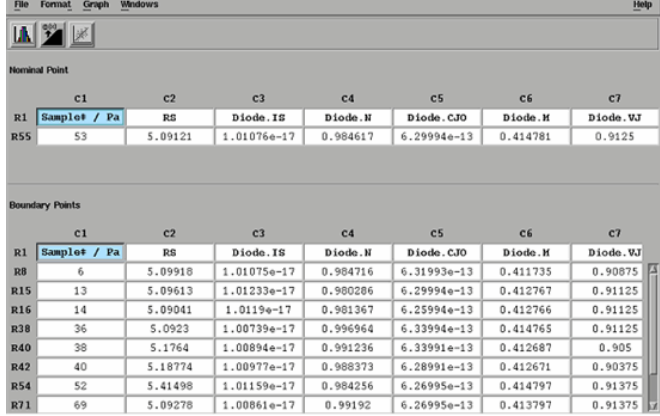
	RS	Diode.IS	Diode.N	Diode.CJO	Diode.M	Diode.VJ
Mean	5.17326	1.01001e-17	0.98782	6.31167e-13	0.414133	0.912216
Variance	0.04052	1.95562e-40	2.14889e-05	1.10783e-29	1.21248e-06	.60971e-05
Std.Deviation	0.20129	1.39843e-20	0.00464	3.32840e-15	0.001101	0.004012
Skewness	4.47606	-0.28516	0.69743	-0.016377	0.238159	0.238159
Kurtosis	24.9556	-0.10185	0.47870	-0.801842	2.291947	13.49792
Minimum	5.08958	1.0065e-17	0.98782	6.24985e-13	0.410591	0.8975
Maximum	6.58694	1.0132e-17	1.00132	6.37994e-13	0.418707	0.93625
Median	5.09256	1.0101e-17	0.98782	6.30994e-13	0.413798	0.9125
Med.abs.dev.	0.00213	8.7e-21	0.00248	2.003e-15	0.000983	0.00125
Rob.Est.Scale	0.00316	1.2898e-20	0.00367	2.9696e-15	0.001457	0.001853

With the use of 'non-parametric statistical analysis', we are now ready to identify the real existing die which contains the 'golden module' of our diodes. As shown in fig.11, we specify 10 boundary points (diode modules) which can be considered to enclose 80% of the parameter samples.



Non-parametric statistical analysis to identify the golden model parameter set nd, thus, the golden diode module, as well as 10 boundary parameter sets, which will include 80% of the 98 parameter sets.

From this analysis, see fig. 12, diode module no.53 is identified to represent best the 'golden device' for both, measured DC and CV data. Modules #6, 13, 14, 36 etc. represent the boundary chips on the wafer.



Nominal Point

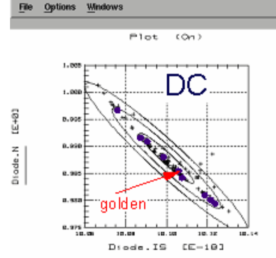
	C1	C2	C3	C4	C5	C6	C7
R1	Sample# / Pa	RS	Diode.IS	Diode.N	Diode.CJO	Diode.M	Diode.VJ
R55	53	5.09121	1.01076e-17	0.984617	6.29994e-13	0.414781	0.9125

Boundary Points

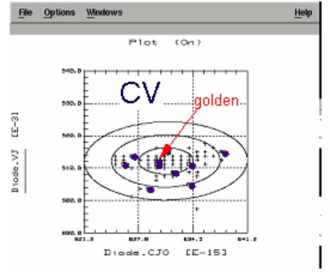
	C1	C2	C3	C4	C5	C6	C7
R1	Sample# / Pa	RS	Diode.IS	Diode.N	Diode.CJO	Diode.M	Diode.VJ
R8	6	5.09918	1.01075e-17	0.984716	6.31993e-13	0.411735	0.90875
R15	13	5.09613	1.01233e-17	0.980286	6.29994e-13	0.412767	0.91125
R16	14	5.09041	1.0119e-17	0.981367	6.25994e-13	0.412766	0.91125
R38	36	5.0923	1.00739e-17	0.996964	6.33994e-13	0.414765	0.91125
R40	38	5.1764	1.00894e-17	0.991236	6.33991e-13	0.412687	0.905
R42	40	5.18774	1.00977e-17	0.988373	6.28991e-13	0.412671	0.90375
R54	52	5.41498	1.01159e-17	0.984256	6.26995e-13	0.414797	0.91375
R71	69	5.09278	1.00861e-17	0.99192	6.26995e-13	0.413797	0.91375

Result of the non-parametric statistical analysis

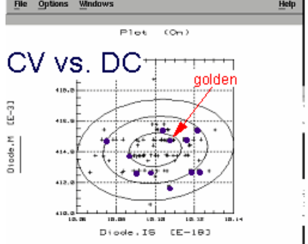
After having executed the non-parametric statistical analysis, we can re-execute a scatter plot for the most important DC and CV model parameter in order to check the validity of the identified 'golden' and 'boundary' modules. The two plots of fig. 13 show the result of such a check. It can be seen clearly, that the identified die #53 represents that diode module whose DC and CV data represent best the 'golden' module in question. Also, the boundary modules look very reasonable, again covering both domains, the DC and CV data.



DC parameters IS vs. N

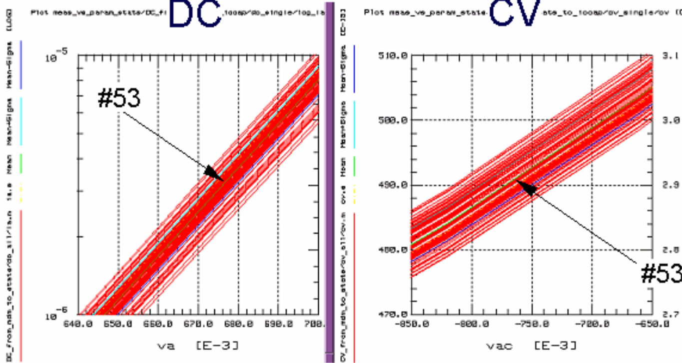


CV parameters VJ vs. CJO



CV parameter M vs. CJO

Using macro 'stats2mdm' of our IC-CAP model file, we can export the model parameters of our identified diode modules back into IC-CAP, run a simulation and display this simulation result with the measurement distribution. As can be seen in fig. 14, the golden no. 53 fits nearly ideally the mean CV curve, but is a little more aside the mean DC trace (see the
IC-CAP file plots for details). However, it is the module which represents best both, the DC and CV domain.



Identified 'golden diode module' no.53 and its location in the measurement data.

Note
if we had calculated simply the mean data of the DC curves and the mean data of the CV curves, and if we had then extracted, based on these two mean data traces, the 'golden model parameters', this set of parameters would correspond to no real existing diode. (Because the mean trace of the DC and the mean trace of the CV would not correspond to a specific existing module).
And if we had done so for our MOS transistor example, with the many different transistor geometries and measurement ranges, the result would have been even more non-physical and, therefore, even more doubtful.

Back to our starting example with the MOS transistors, a similar analysis can be performed. The result of a non-parametric statistical analysis (with the mentioned outliers taken into account) is die no.19 with its different transistor geometry modules, which is closest to the mean data of all dies, all modules and measurement ranges.

Besides this non-parametric type of statistical analysis, the IC-CAP statistics package covers also the 'classical' factor analysis and principal component analysis method, as well as the ability to generate Monte Carlo based sets of model parameters, with a distribution corresponding to the original model parameter distribution. More details about these methods can be found in the following chapter about parameter-related statistics.

Applying Non-Parametric Statistics Directly To Measured Data

Besides referring to extracted model parameters, non-parametric statistics can also be applied to measured data directly. In this case, a representative selection of measurement data points is gathered during the measurement of the whole wafer in an ASCII file and this file is imported into the statistics tool. Representative means that some characteristic measurement points, e.g. 10-20 points of an output characteristics, are exported to that ASCII file and not all the data. Especially noisy data have to be ignored for this type of analysis.

After this ASCII file is imported into IC-CAP's statistics, it may look like the following table:

chip#	gummel1	...	gummelN	output1	...	outputM	cv1	...
1	1.3u	24u	144u	1.3m	1.5m	1.6m	1.2p	1.5p
2	1.4u	25u	146u	1.4m	1.6m	1.7m	1.1p	1.4p
3	1.2u	24u	145u	1.3m	1.5m	1.7m	1.2p	1.6p
4	1.3u	23u	144u	1.5m	1.7m	1.8m	1.3p	1.5p

Applying a non-parametric analysis to this data will again identify the chip which is closest to the mean of all data of all measurement setups. Compared to the measurement-based

type of statistics of the first chapter, this method includes all kinds of measurement setups and is, therefore, identifying the real 'golden die' for all types of measurements and not simply that one related to a specific measurement setup.

Literature about non-parametric statistics

IC-CAP 5.0, Statistical Analysis Manual, June 1997, HP Part no. 85190-90050

Identify the 'golden device' and the 'limits devices'

In this section you will learn:

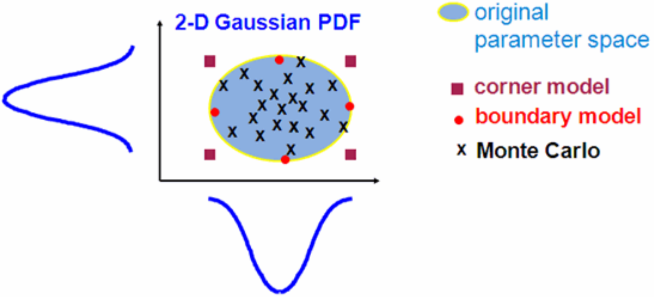
- Conventional Statistics Methods (e.g. Factor Analysis)
- Introducing The 'Non-Parametric Statistics' Concept
- How does 'Non-Parametric Statistics' work?
- An Example
- Conclusions

Conventional Statistics Methods

Example: Factor Analysis

- measure many devices
- extract model parameters
- add process parameters (PCM) to statistics spreadsheet
- transform each parameter distribution to Gaussian
- apply statistics
- find dominant parameters (= parametric statistics)
- apply e.g. Monte Carlo Analysis to generate new parameters which represent the original distribution

What are Corners, Boundaries, and Monte Carlo statistical models, obtained with conventional statistics?



This is a simple geometric representation of what corners, boundaries, and Monte Carlo are. In 2 dimensions, we select 4 evenly spaced models which lie on an N sigma boundary of the joint PDF (probability density function). Corners are outside the joint PDF (in purple), which produces too pessimistic results. Parametric boundary models (in red) are right on the boundary of joint PDF so it provides a more realistic worst-case model. Worst-case modeling includes both corner and boundary methods.
Monte Carlo (in black) is a sampling from the complete distribution that leads to a yield estimate when each sample is simulated in the final circuit and the circuits performance is checked against the pertinent performance specifications,

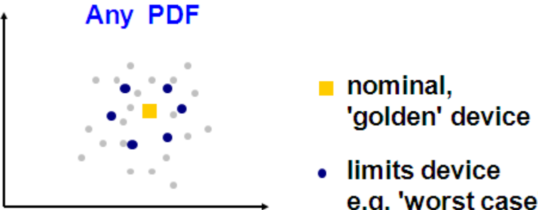
Introducing The Non-Parametric Statistics Concept

This type of statistics analyzes the hyper-cube of parameters:

- Identifies the chip closest to the parameter center
- Chips closest to an e.g. 90% enclosure hyper space

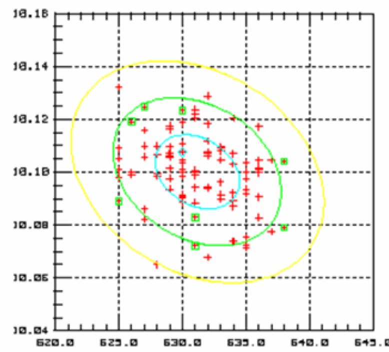
Therefore, it can easily be applied to model parameters and also to measurement data

What means typical component and limits components identification ?

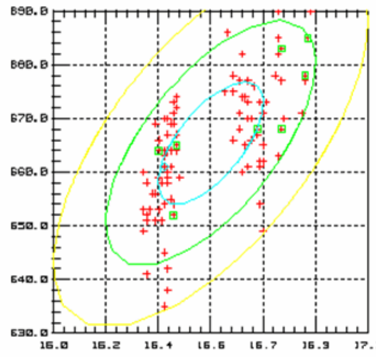


limits device: the user specifies the enclosure percentage for the data of the hyper space.
This kind of statistics is called 'Non-Parametric'

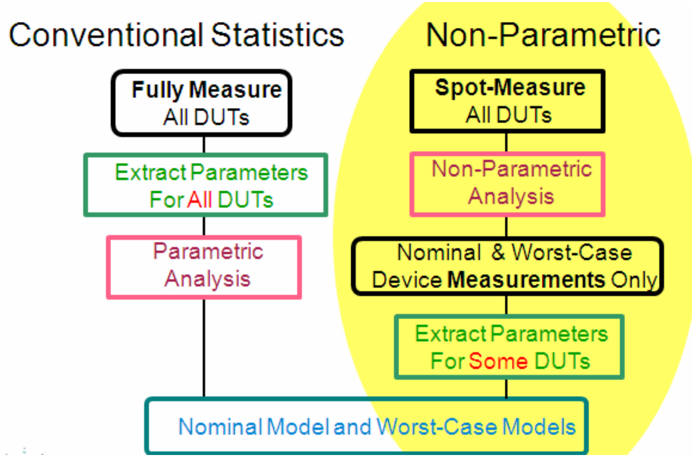
Non-Parametric Statistics Works on Gaussian Distribution



Non-Parametric Statistics Works on Non-Gaussian Distributions



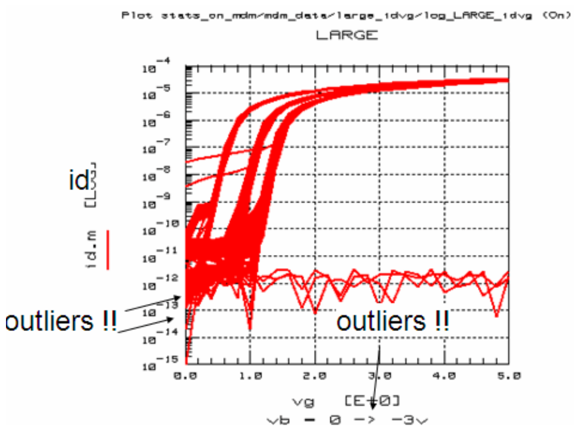
Conventional Statistics vs. Non-Parametric Statisticswrap-up



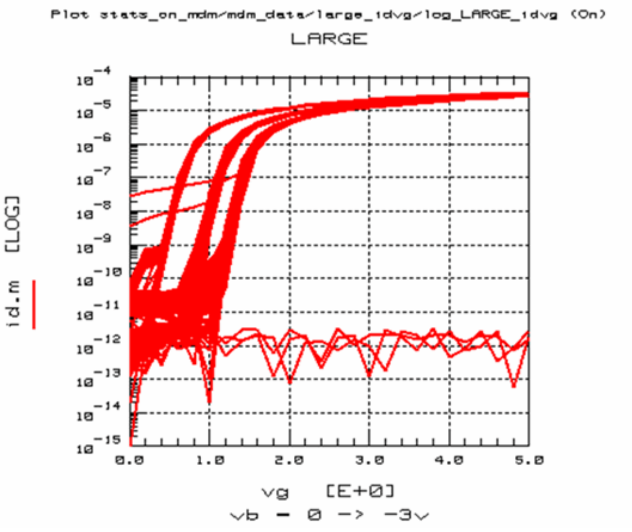
Mode 2 may lead to better congruence between boundary model and the boundary devices in terms of final electrical behavior
Mode 2 is better because we do not expend considerable effort extracting many models that we ultimately may throw away. If we are going to do direct or generated Monte Carlo yield estimation then the effort is not in fact wasted.
Mode 2 is also better because one get two outputs — nominal and worst-case vectors for both the measurements and the models.

How does 'Non-Parametric' Statistics work?

First of all: Identify outliers and eliminate them



Then:Select some spot data from DC, CV and Spar and export them to the statistics spreadsheet



Then, the Data Sets are Normalized

e.g., for a diode, id(at vd): is in the mA range C(at vd): is in the fF range Spar at freq and vd: are -1 ... 1 in MAG, 0 ... 360' in PHASE

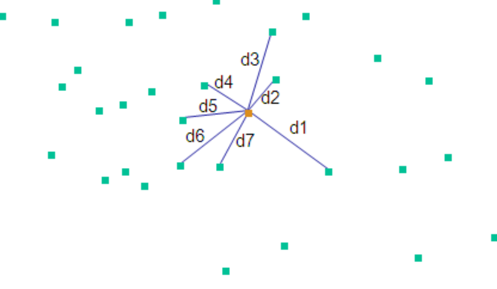
-> Normalize each data set to A Unit Hypercube

Estimating Density Using Nearest Neighbors Method

Number of NN = 7

$$\text{Average Distance} = \frac{d1 + d2 + \ldots + d7}{7}$$

$$\text{Density Metric} = \frac{1}{\text{AverageDistance}}$$

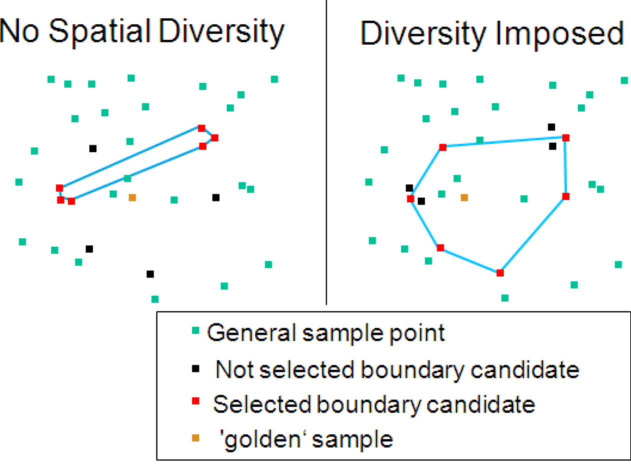


The user needs to understand the trade off between setting Density Estimator Percentage, which directly controls the number of nearest neighbors, either too high or too low. If set

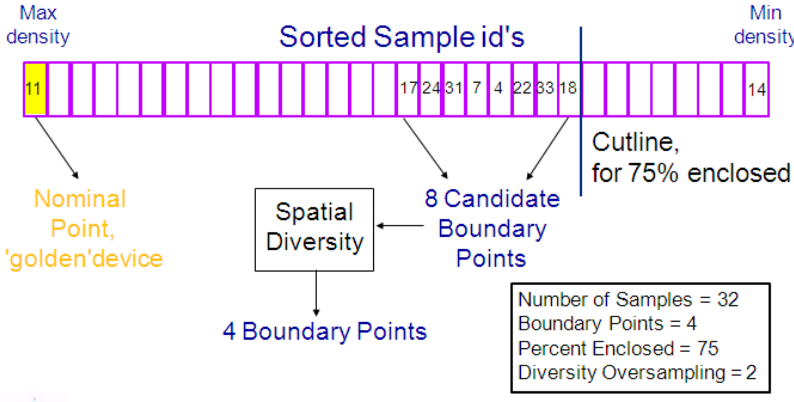
too high, density estimate is overly smooth and the algorithm can fail to resolve peaks in the PDF that are close together. Setting it too low leads to a different problem. The problem is that the distance averaging is a Monte Carlo process and consequently introduces Monte Carlo noise into the density estimate. This noise can lead to poorly chosen nominal and boundary models. The problem is more acute for the nominal model, as the nominal model is in a pronounced region of low slope in the density estimate.

Spatial Diversity

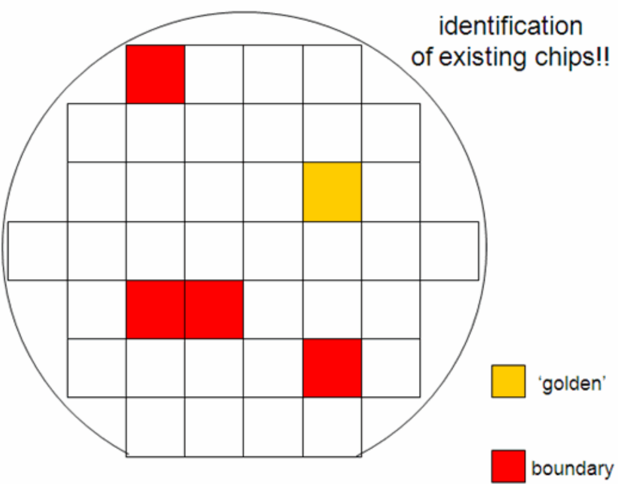
Spatial diversity is done by oversampling and then iteratively forming all possible pairs and discarding a point from the pair that is closest together. The procedure terminates when only the desired number of boundary points remain.
Do not know best limits for Spatial Diversity Oversampling factor. At the present 1.5 to 3.0 seems like a good range. Clearly, it cannot be 1 or lower since 1 implies no oversampling. Runs with it set above 3 show no clear advantage. High settings do have a significant performance penalty for large numbers of requested boundary points.



Nominal and Boundary Point Selection



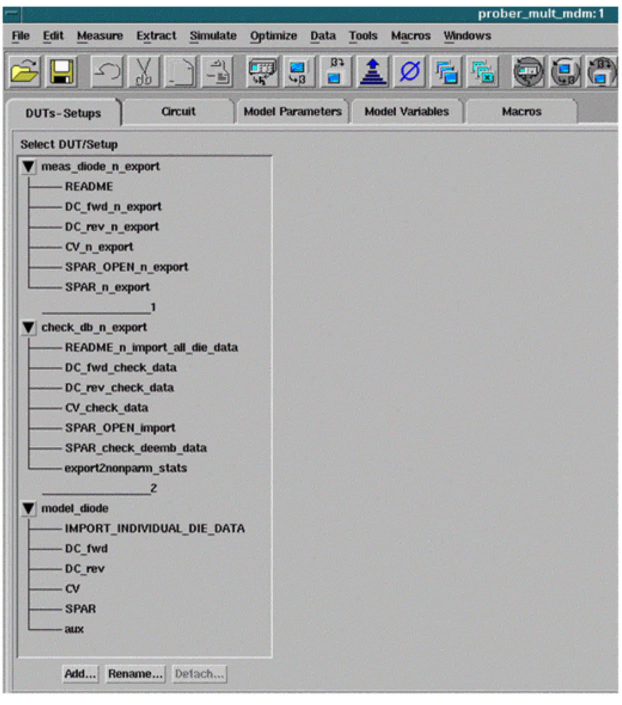
The Result



Note
The 'golden' device is NOT the mean(DC), mean(CV), mean(Spar) and the 'boundary' devices are NOT the ±sigma(DC), ±sigma(CV), ±sigma(Spar). The 'golden' device is rather the device which is the most centered to the 'cloud' of dataof all types of measurements and the 'boundary' devices are thosewhich include a given data variation for all types of measurements.

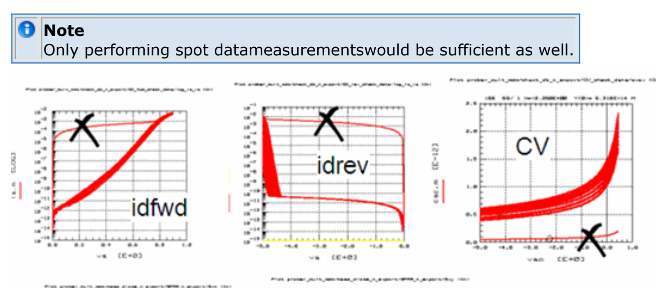
An Example: Diode with measurements from DC to S-parameters

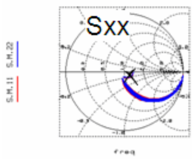
- measure and save data
- identify outliers, export to stats
- model individual die

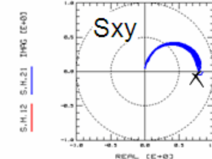


measurement results (here full curves, but spot data wouldbe sufficient), outlier identification, export of spot data to statistics spreadsheet

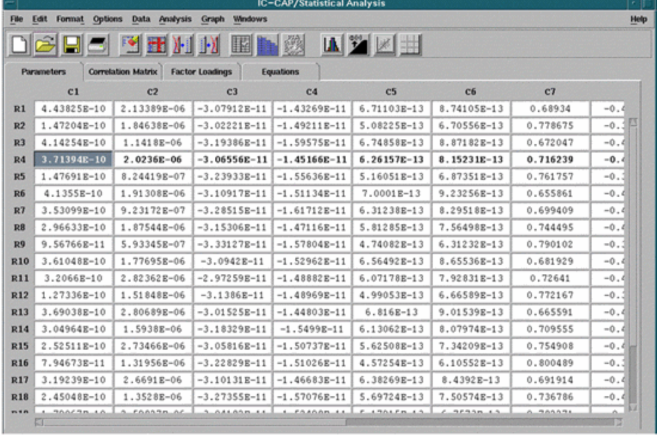
Note
Only performing spot datameasurementswould be sufficient as well.





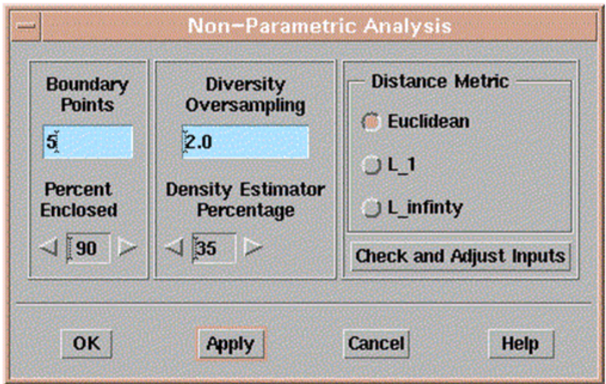


Quick-extraction parameters, PCM data and/or spot data fromoperating points are imported into the statistics spreadsheet

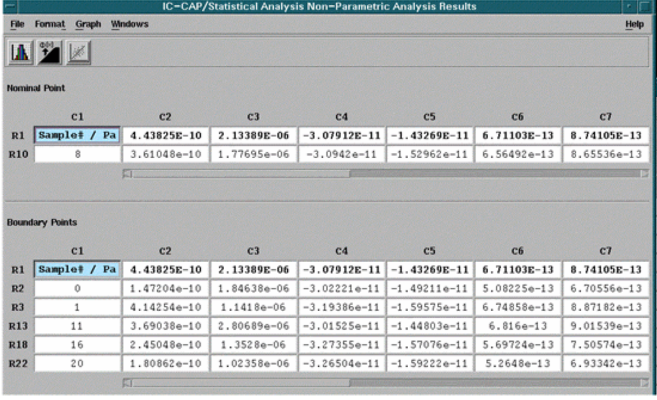


	C1	C2	C3	C4	C5	C6	C7
R1	4.43825E-10	2.13389E-06	-3.07912E-11	-1.43269E-11	6.71103E-13	8.74105E-13	0.68934
R2	1.47204E-10	1.84638E-06	-3.02221E-11	-1.49211E-11	5.08225E-13	6.70556E-13	0.778675
R3	4.14254E-10	1.1418E-06	-3.19386E-11	-1.59575E-11	6.74858E-13	8.87182E-13	0.672047
R4	3.71394E-10	2.0236E-06	-3.06556E-11	-1.45166E-11	6.26157E-13	8.15231E-13	0.716239
R5	1.47691E-10	8.24419E-07	-3.23933E-11	-1.55636E-11	5.16051E-13	6.87351E-13	0.761757
R6	4.1355E-10	1.91308E-06	-3.10917E-11	-1.51134E-11	7.0001E-13	9.23256E-13	0.655861
R7	3.53099E-10	9.23172E-07	-3.28515E-11	-1.61712E-11	6.31238E-13	8.29518E-13	0.699409
R8	2.96633E-10	1.87544E-06	-3.15306E-11	-1.47116E-11	5.81285E-13	7.56498E-13	0.744495
R9	9.56766E-11	5.93345E-07	-3.33127E-11	-1.57804E-11	4.74082E-13	6.31232E-13	0.790102
R10	3.61048E-10	1.77695E-06	-3.0942E-11	-1.52962E-11	6.56492E-13	8.65536E-13	0.681929
R11	3.2066E-10	2.82362E-06	-2.97259E-11	-1.48882E-11	6.07178E-13	7.92831E-13	0.72641
R12	1.27336E-10	1.51848E-06	-3.1386E-11	-1.48969E-11	4.99053E-13	6.66589E-13	0.772167
R13	3.69038E-10	2.80689E-06	-3.01525E-11	-1.44803E-11	6.816E-13	9.01539E-13	0.665591
R14	3.04964E-10	1.5938E-06	-3.18329E-11	-1.5499E-11	6.13062E-13	8.07974E-13	0.709555
R15	2.52511E-10	2.73466E-06	-3.05816E-11	-1.50737E-11	5.62508E-13	7.34209E-13	0.754908
R16	7.94673E-11	1.31956E-06	-3.22829E-11	-1.51026E-11	4.57254E-13	6.10552E-13	0.800489
R17	3.19239E-10	2.6691E-06	-3.10131E-11	-1.46683E-11	6.38269E-13	8.4392E-13	0.691914
R18	2.45048E-10	1.3528E-06	-3.27355E-11	-1.57076E-11	5.69724E-13	7.50574E-13	0.736786

The non-parametric conditions are set and the statistics is applied



The result in table form:



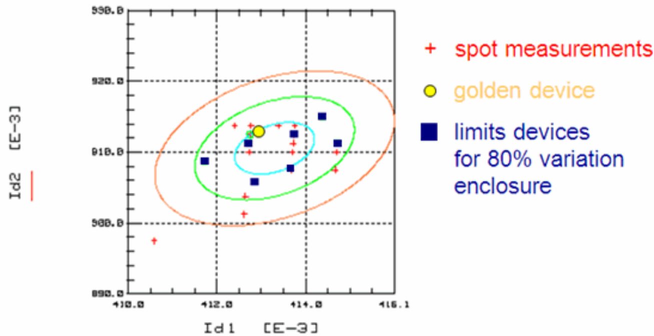
Nominal Point

	C1	C2	C3	C4	C5	C6	C7
R1	Sample# / Pa	4.43825E-10	2.13389E-06	-3.07912E-11	-1.43269E-11	6.71103E-13	8.74105E-13
R10	8	3.61048e-10	1.77695e-06	-3.0942e-11	-1.52962e-11	6.56492e-13	8.65536e-13

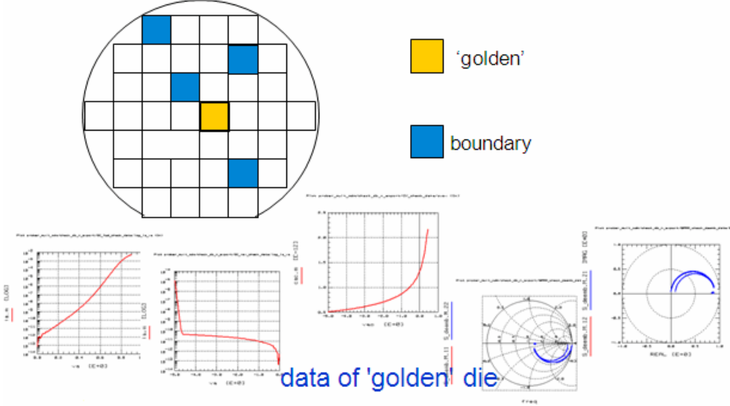
Boundary Points

	C1	C2	C3	C4	C5	C6	C7
R1	Sample# / Pa	4.43825E-10	2.13389E-06	-3.07912E-11	-1.43269E-11	6.71103E-13	8.74105E-13
R2	0	1.47204e-10	1.84638e-06	-3.02221e-11	-1.49211e-11	5.08225e-13	6.70556e-13
R3	1	4.14254e-10	1.1418e-06	-3.19386e-11	-1.59575e-11	6.74858e-13	8.87182e-13
R13	11	3.69038e-10	2.80689e-06	-3.01525e-11	-1.44803e-11	6.816e-13	9.01539e-13
R18	16	2.45048e-10	1.3528e-06	-3.27355e-11	-1.57076e-11	5.69724e-13	7.50574e-13
R22	20	1.80862e-10	1.02358e-06	-3.26504e-11	-1.59222e-11	5.2648e-13	6.93342e-13

The result graphically:



Now, that the 'golden' and boundary devices are identified on the wafer, we measure them completely, and model them (and only them)



Modeling Results

The model parameters obtained will represent the centereddevice and an e.g. 80% (boundary) yield parameter setrelated to the original spot measurement data distribution, depending on the selected enclosure rate.

A simple method to identifythe 'golden device' and the 'limits devices'has been proposed.

It has been implemented in IC-CAP.
This method reduces the amountof measurements and time considerably.

Non-Parametric Boundary Analysis

In this section, you will learn about:

- Motivation
- Circuit and System Design Methods
- The Non-Parametric Boundary Analysis Algorithm
- Validating the Algorithm
- Applying the Tool
- Conclusion

Motivation

- Fabrication processes inherently stochastic
- Boost yields and reduce design cycle times
- Control process tolerances
- Accurate nominal models
- Assess manufacturability before production

Note
Assess manufacturability via yield estimation and worst-case model analysis.

Circuit and System Design Methods

Non-statistical design:

- Traditional methods: Cut and try. Use hopefully nominal models. No assessment of manufacturability prior to production.
- Sensitivity based analysis: Add sensitivity analysis to traditional flow. Better but no direct assessment of manufacturability.

Statistical design

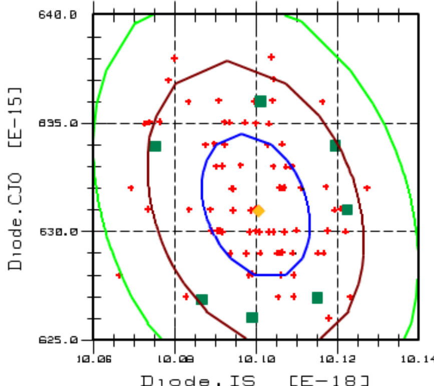
- Yield optimization: Must be able to estimate yield and need a global optimization method. Best of the statistical design techniques.
- Response surface methods: Probe circuit behavior with design of experiments so as to build a low order polynomial model of the circuit's performance curves. No direct information about yield.
- Yield estimation: Direct or generated Monte Carlo. Ensemble directly or (PDF fit and sample)
- Worst-case analyses
  - Corner
  - Parametric boundary
  - Non-parametric boundary

What are Corners, Boundaries, and Monte Carlo?

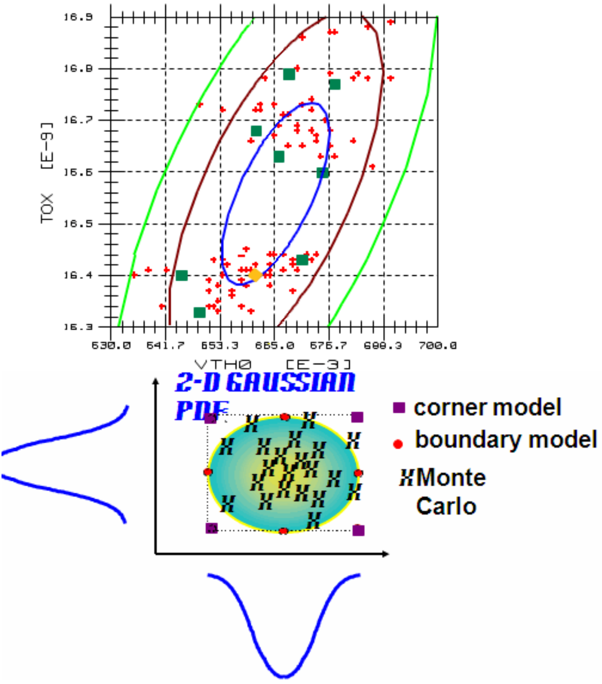
This is a simple geometric representation of what corners, boundaries, and Monte Carlo are. In 2 dimensions, we select 4 evenly spaced models which lie on an N sigma boundary of the joint PDF (probability density function). Corners are outside the joint PDF (in purple), which produces too pessimistic results. Parametric boundary models (in red) are right on the boundary of joint PDF so it provides a more realistic worst-case model. Worst-case modeling includes both corner and boundary methods.
Monte Carlo (in black) is a sampling from the complete distribution that leads to a yield estimate when each sample is simulated in the final circuit and the circuits performance is checked against the pertinent performance specifications,

Non-Parametric Boundary Analysis

A statistically robust method for obtaining nominal and worst-case vectors from a collection of vectors drawn from an arbitrary multidimensional stochastic process
Works on Gaussian PDFs



And on Non-Gaussian PDFs !



Primary Inputs to Non-Parametric Boundary Analysis

- Real valued dataset of S sample vectors with each vector having P parameters
- Number or Boundary Points
  - desired number of worst-case models
- Percent Enclosed
  - Percentage of data enclosed by boundary points
    We conjecture that Percent Enclosed may be a lower limit on the yield under some circumstances. First, each boundary model when simulated must indicate that the circuit in which it is embedded passes all of its performance specifications. Second, the mapping functions from the model parameter space to the specific circuit performance space are monotonic. We have neither proven this mathematically or demonstrated it empirically at this time.

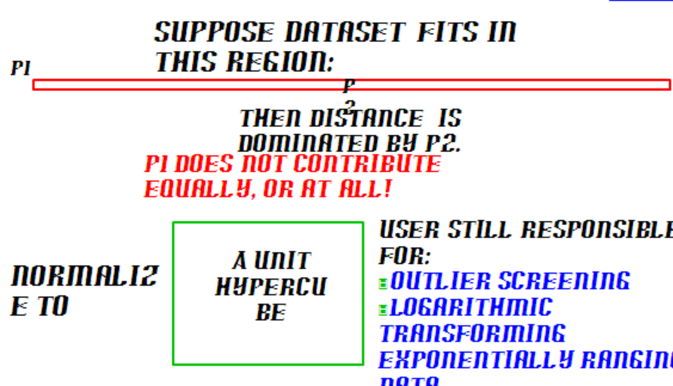
Secondary Inputs

- Density Estimator Percentage
  - Number of nearest neighbors for density estimation
- Diversity Oversampling
  - Oversampling factor for ensuring spatial diversity
- Distance Metric
  - Formula for distance calculation

Outputs

- Nominal vector
- Boundary vectors
- Both are selected, not constructed

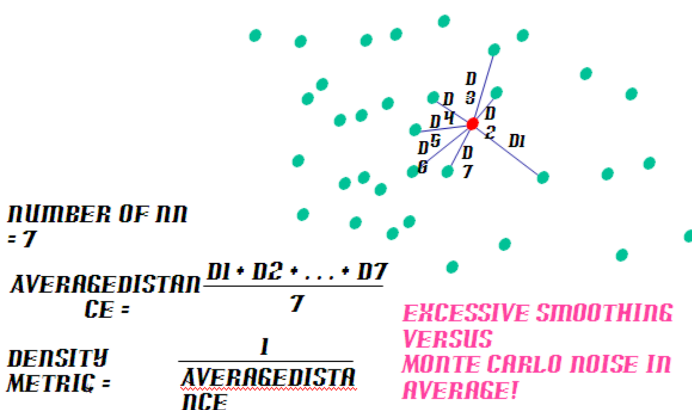
Normalize the Dataset



Outlier screening is necessary because outliers cause the true data to lie along some short

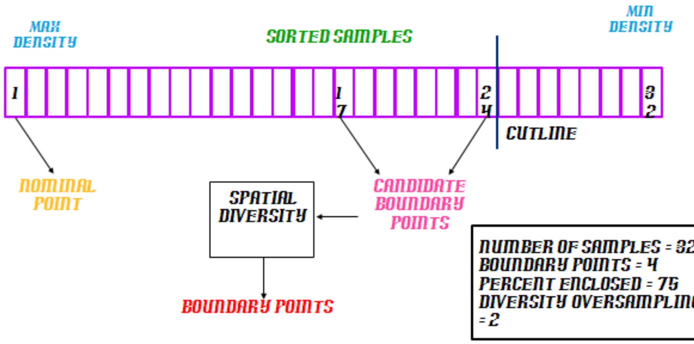
interval of one or more of the unit parameter axes. Statistical variation in that parameter is lost in distance calculation.
Logarithmic transforming is necessary for similar reason because the higher magnitude data skews the results. Variance is more evenly reported if the data is logarithmically transformed to make it linearly varying.

Estimating Density Using Nearest Neighbors Method

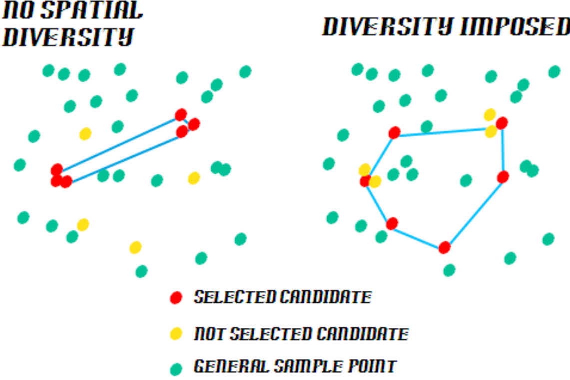


User needs to understand the trade off between setting Density Estimator Percentage, which directly controls the number of nearest neighbors, either too high or too low. If set too high, density estimate is overly smooth and the algorithm can fail to resolve peaks in the PDF that are close together. Setting it too low leads to a different problem. The problem is that the distance averaging is a Monte Carlo process and consequently introduces Monte Carlo noise into the density estimate. This noise can lead to poorly chosen nominal and boundary models. The problem is more acute for the nominal model, as the nominal model is in a pronounced region of low slope in the density estimate.

Nominal and Boundary Point Selection



Spatial Diversity---An Illustration

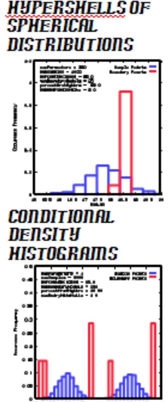


Spatial diversity is done by oversampling and then iteratively forming all possible pairs and discarding a point from the pair that is closest together. The procedure terminates when only the desired number of boundary points remain.
Do not know best limits for Spatial Diversity Oversampling factor. At the present 1.5 to

3.0 seems like a good range. Clearly, it cannot be 1 or lower since 1 implies no oversampling. Runs with it set above 3 show no clear advantage. High settings do have a significant performance penalty for large numbers of requested boundary points.

Validating the Algorithm

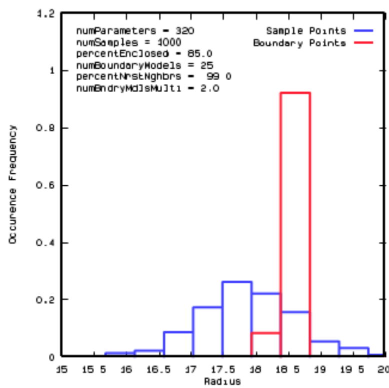
Constant density curves for 2-D distributions



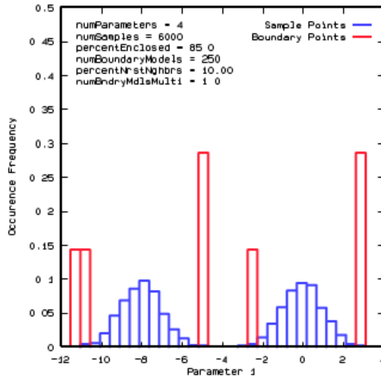
2-D Joint Gaussian Distribution with Nominal and Boundary Points

Handles this bimodal distribution nicely. Ellipses are distorted due to unequal scales on axes.
Discuss what happen when the distributions are moved close together. NEAREST NEIGHBORS TOO HIGH.
Point out how the nominal can wander around for NEAREST NEIGHBORS TOO LOW.

Hypershell in 320-D Spherical Gaussian Distribution



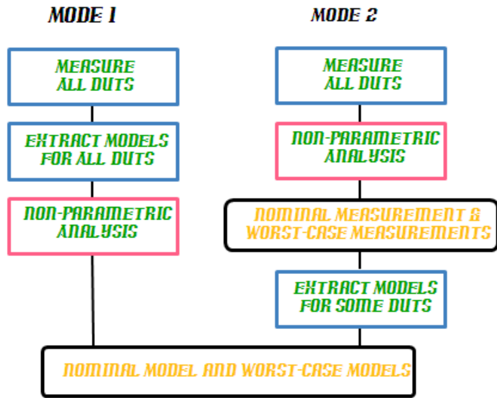
Conditional Density Histogram on 4-D Bimodal Gaussian Distribution



Conditional density histogram are done by selecting point from the distribution such that all but one of the parameters are constrained to be in some reasonably narrow band of values for each of the parameters. The remaining parameter is allow to range over its

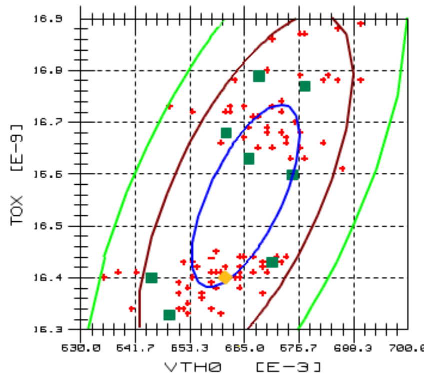
natural region of support. This is equivalent to saying that we wish to repeatedly slice up the space, except for the last parameter. We then form a histogram on the sample points that are in the constrained region and a histogram over the boundary points that are also in that same region. This procedure is somewhat related to the computing of margin density histograms.
In out case, the expected outcome is to see two approximately Gaussian sample histograms with the boundary points be in the tails of the two distribution. This is the result that we observe.

Applying the Tool

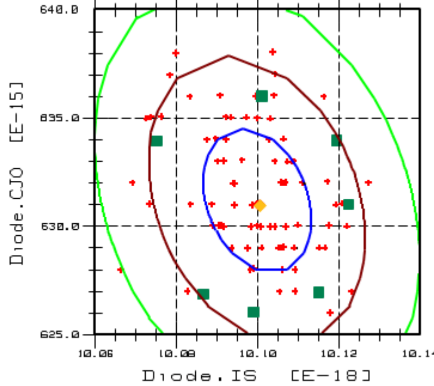


Mode 2 may lead to better congruence between boundary model and the boundary devices in terms of final electrical behavior
Mode 2 is better because we do not expend considerable effort extracting many models that we ultimately may throw away. If we are going to do direct or generated Monte Carlo yield estimation then the effort is not in fact wasted.
Mode 2 is also better because one get two outputs — nominal and worst-case vectors for both the measurements and the models.

2-D Analysis of BSIM3 Data



Analysis of Diode Data



Conclusion

- Statistical circuit design still beckons

  - More upfront infrastructure and effort
  - Fewer problems and higher yields during production
- Non-parametric boundary analysis is a powerful new tool for statistical circuit design
  - Easy to use
  - Accommodates wide range of stochastic processes
  - Provides nominal and worst-case models via a statistical robust technique

Improving The Manufacturability Of Electronic Designs

Dan Stoneking, Hewlett-Packard,
IEEE Spectrum, June 1999, Volume 36, Number 6

This material appears here with permission of the IEEE. Such permission of the IEEE does not in any way imply IEEE endorsement of any of Hewlett-Packard's products or services. Internal or personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution must be obtained from the IEEE by sending a blank email message to pubs-permissions@ieee.org. By choosing to view this document, you agree to all provisions of the copyright laws protecting it.

After many weeks or months of laboring over your design with the best available models and a favorite simulator, you and your management commit it to production. You breathe a sigh of relief as the first manufactured units work their way to final test. Then, just as you are transitioning to the next project, the bad news comes: yields are unacceptably low.

A likely cause is that you overlooked the effects of the stochastic nature of the manufacturing process. Or perhaps you assumed that random fluctuations in the myriad of process steps would more or less cancel each other out--and this time they seem to be canceling rather less than more. Whatever the reason, too many of the units failed to measure up and became scrap.

Now management is on edge. Customers are unsettled and looking around for options. The part can be produced, but the low yield means that manufacturing costs will be too high. Yields can be improved, but that will require more engineering time. In retrospect, it seems likely that the application of statistical circuit design techniques could have largely prevented these problems.

Nonparametric boundary analysis

Statistical circuit design techniques analyze the yields of circuit designs whose underlying components exhibit random fluctuations. These techniques can help produce more robust designs by calling attention to areas where statistical variations are likely to combine in such a way as to cause circuit failure. Nonparametric boundary analysis (NBA), see "How nonparametric bounary analysis works" a technique introduced in Hewlett-Packard EEsof's IC-CAP 5.0, permits yield analysis when the random fluctuations result from an arbitrary stochastic process, in addition to well-studied processes such as the Gaussian. Its ability to analyze data from all stochastic processes is why NBA takes the label nonparametric, which indicates that it does not attempt to fit a parametric, closed-form probability density function to the input data.
NBA captures the statistics of the process from data vectors-collections of parameter values for all of a design's statistically variant components. These parameters can be just about anything so long as a given set of values uniquely represents each component sample. Herein lies NBA's flexibility. It allows the input data vectors to be formed from measurements or from model parameters that represent each of the design's components. Each component's model can have any number of parameters, and the number of components in the design can be quite large.

The technique is a straightforward means of estimating a lower bound on yield, and requires much less computational effort than the well-known technique of Monte Carlo yield estimation. The Monte Carlo approach, on the other hand, provides an actual yield estimate, not just a lower bound. It randomly samples the distribution of combinations, performs a complete circuit simulation, and counts how many give an acceptable result. For typical distributions, most of the computation effort is expended on simulations that have near-nominal responses.

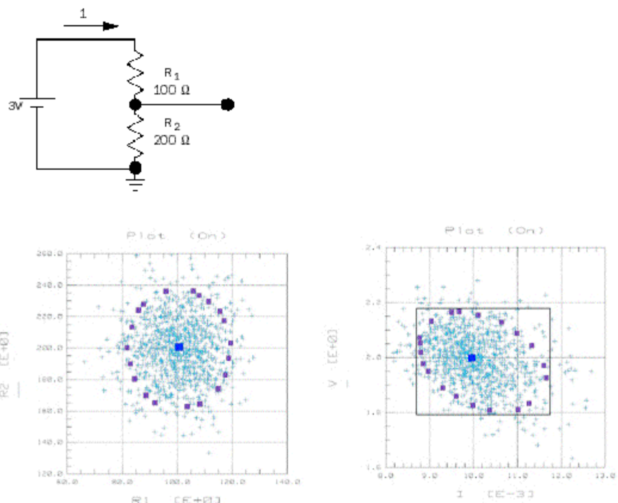
NBA requires less computational effort because it first examines the overall distribution of model parameter values or measurements of components and then selects only those combinations in the tails of the distribution-the combinations most likely to cause yield problems-for a complete circuit simulation. The combinations selected enclose a user-specified fraction of the distribution. NBA's lower bound on yield is obtained by simulating all the selected combinations and determining whether they all result in circuits with acceptable responses. If they do, then the yield will be at least as large as the user-specified enclosure fraction. If not, then the enclosure fraction is reduced and the process is repeated until a fraction is found for which all of the circuits work properly.

In addition to analyzing yield, the technique identifies a nominal model for each of the statistically variant components in a design. Nominal models are useful for initiating the design process because they emulate the various components' ideal behaviors.

Applying NBA

To see how NBA works, consider the voltage divider of Fig. 1. The voltage supply is ideal with an output of 3 V. The resistors, however, are statistically variant. The NBA inputs are pairs of R1 and R2 sample values. The responses of interest are V and I. By construction, the R1-R2 distribution is Gaussian. The enclosure percentage for the R1-R2 distribution is 85 percent. Since the dataset is two-dimensional, the enclosure boundary is clearly evident for the R1-R2 and I-V spaces. Note that the points in the I-V space that correspond to the boundary points in the R1-R2 space also create a boundary. This

boundary encloses the same percentage-85 percent-of the I-V distribution. Note also the position of the nominal points, shown in blue, in the input and response distributions. They are both close to the known means.



The voltage-divider circuit example (top left)

The above figure has statistically variant resistors that illustrate equal-enclosure percentages in the input (R1-R2) and response(I-V) spaces below . In the I-V plot, the corner areas within the white rectangle but outside the area defined by the green boundary points illustrate why nonparametric boundary analysis provides only a lower bound on yield.

We can also see why NBA provides only a lower bound on the yield and why this bound can be pessimistic. Compare the 85 percent enclosure in the I-V space with the hypothetical acceptability region outlined in black. NBA provides only a lower bound because the 85 percent enclosure is round and the acceptability region has corners that are not enclosed. Grossly pessimistic yield estimates can be obtained by applying NBA only once and noting that all the outcomes are acceptable. This situation would pertain if, for example, the hypothetical acceptability region were expanded. This difficulty can be corrected simply by iterating the NBA enclosure percentage upward until a failure is found.

NBA has a number of other constraints and limitations on its practical application. An important constraint is that the functions that map the input space to the response space must be approximately monotonic if the property of equal enclosures is to hold. Without equal enclosures in the two spaces, the algorithm does not provide a lower bound on yield. Approximately monotonic means that the mapping functions can have some ripple. However, the ripple must be small relative to the overall change in the response functions across the acceptable response region.

Although approximate monotonicity is difficult to guarantee in all cases, the author conjectures that the vast majority of practical circuit designs are approximately monotonic. The monotonicity of a design can be explored by artificially adjusting the component parameters to trace paths along rays emanating from the nominal design while verifying that the circuit responses are monotonically increasing or decreasing. Another method consists of iteratively increasing the NBA enclosure percentage while verifying for all pair-wise combinations of the response variables that the area (or volume) of enclosure in the response space is constantly increasing.

Ultimately, even if the circuit is not approximately monotonic with respect to its parameters, NBA can be used to explore for worst-case designs in the performance space.

Dimensionality considerations
Another concern when applying NBA is to limit the dimensionality (the number of real numbers per input vector) of each design's parameter space. This is important for several reasons. As the dimensionality goes up, the user must populate that high-dimensional space with data. The amount of data needed grows rapidly. This is mostly a concern for components that require many parameters for a unique representation. Getting enough samples of the overall design is not normally a problem because of the large numbers of combinations that can be formed from modest numbers of samples for each component. Another point to watch regarding dimensionality is that the method's density estimate degrades as dimensionality increases. The author has shown proper functionality with dimensionalities as high as 320. Proper performance into low- to mid-thousands is viable.

Beyond the logistics of getting the data and algorithm performance, the user should be aware of an interesting phenomena of very high-dimensional spaces: as the number of dimensions goes to infinity, the volume contained in a very thin shell on the surface of a hypersphere containing one unit of volume also goes to unity. That is, all the volume of a high-dimensional hypersphere is on its surface. This phenomena ultimately limits NBA's ability to return a meaningful set of boundary points for a given enclosure percentage.

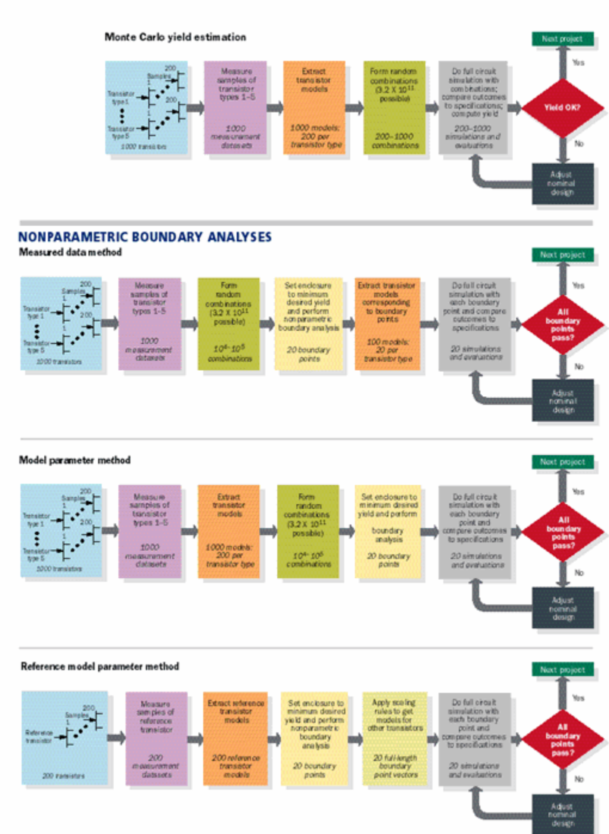
Several tools are available to help the user of NBA limit dimensionality. Direct knowledge and engineering judgment assist in selecting appropriate model parameters or measurements to represent each component in the overall design. Obviously, parameters to which the component's responses are not sensitive should be avoided. If using measurements, do not choose data that shows little variation across the samples. Avoid using parameters or measurements that are highly correlated with one another. Beyond engineering judgment, employ correlation and sensitivity analyses to determine the important parameters or measurements.

NBA has different data requirements from other statistical circuit design methodologies. This is both good and bad. To apply other methods, it is normal to use a compact set of statistics derived from the raw data-mean vectors and covariance matrices for Gaussian distributions or more exotic distributions along with their pertinent statistics. NBA, however, needs the raw data. This is not as onerous a demand as it may at first appear, however. To get the compact statistics in the first place, the raw data must be gathered-no extra work here. Although the volume of NBA's data is much larger, it is not impractically so, and today's Internet makes user access simple and quick.

A big advantage of NBA is that it never requires that a probability density function be fitted to the data. It thereby avoids a host of problems, not the least being the often unjustified assumption that the data fits a Gaussian model. Even if the probability density model is appropriate, fitting any model to any data inherently leads to an imperfect emulation.

The last major concern for NBA is how large a circuit can be analyzed. To answer that question, it is helpful first to consider the various statistical circuit design flows that use NBA. As is turns out, there are several answers depending on the flow that one employs.

Statistical circuit design flows using NBA
Different statistical circuit design flows are possible for more complicated circuits Fig. 2 . The first flow in the figure shows the application of Monte Carlo yield estimation (MCYE). The last three flows are variations of NBA-the measured data (MD), model parameter (MP), and reference model parameter (RMP) methods. All focus on designs in which transistors are the prominent cause of circuit variability. Designs with other statistically variant components would use identical flows. As a visual cue to the reader, similar functionality steps carry the same color coding in all the flows.
In common with all the others, the MCYE flow assumes that the design under analysis has a total of five transistors of different sizes light blue block . Two hundred samples of each of the transistors are measured to produce 1000 measurement datasets light purple block . One thousand models are then extracted orange block . Next, form as many combinations of five transistors as it is practical to simulate (200-1000) by randomly selecting a model from each pool of 200 transistor models green block . Then run full-circuit simulations with the combinations and get a yield estimate gray block . If the yield estimate is low, repeat as shown until the yield becomes satisfactory black block .



Statistical circuit design flows are diagrammed for comparison .

From top to bottom, they are:traditional Monte Carlo yield estimation; nonparametric boundary analysis operating on measured data from statistically variant components; nonparametric boundary analysis applied to model parameters of the components; and use of a reference model and scaling rules to decrease effort needed for a nonparametric boundary analysis.

The NBA-MD flow introduces an extra step into the MCYE flow-perform NBA after choosing a minimum desired yield tan block . In addition, it transposes the steps for forming the random combinations and extracting models. This transposition matters both from the standpoint of how many combinations are considered (coverage of a high-dimensional space) and because of the amount of effort expended on extracting the models and running the full-circuit simulations.

The NBA-MP flow is essentially that of the MCYE flow. The NBA-MP flow has a "perform NBA" block interjected between the "form random combinations" and "full-circuit simulation" blocks of the MYCE analysis. This flow diminishes the effort in the iteration loop but still requires extraction of a model for all the measurement datasets.

The NBA-RMP flow begins like the NBA-MP flow but relies on a priori knowledge of how to translate the model parameter values of the measured reference transistor into the parameter values for all the other transistors. If this translation can be done effectively, less effort by far is needed in gathering the measured datasets.

The main difference between MCYE and the NBA methods is the number of full-circuit simulations required to interrogate the nominal circuit design. For the present example, which assumes typical counts, MCYE requires a factor of 10-50 more full-circuit simulations per iteration. Given that simulation time can be large and that the full-circuit simulations are in an iteration loop, significant savings in engineering time will result from using NBA.

Another key point in comparing MCYE to NBA-MD and NBA-MP is the extent to which the methods explore the input data space. Note the number of combinations formed for the MCYE, NBA-MD and NBA-MP methods. The NBA methods accommodate far greater numbers of combinations because they select only a tiny subset of the input points, which then are run through full-circuit simulation. MCYE's weakness is that it runs all the combinations through full-circuit simulation.

Comparing the NBA methods indicates that NBA-RMP requires the least effort and NBA-MP the most. NBA-RMP enjoys its position because fairly few transistors are measured and extracted. The weak link in NBA-RMP is the transistor scaling operation light green block . For example, it implicitly assumes that all of a design's transistors are deterministically

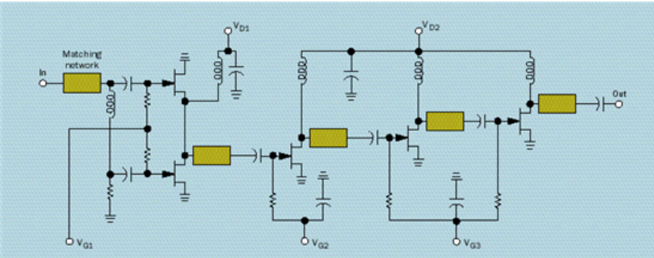
related-that there are no intra-die variations.

NBA-MP takes the most effort, and may not offer any advantage in accuracy over NBA-MD, which requires the mid-range effort. Low data dimensionality and data fidelity drive accuracy in any NBA method. If the underlying models are sufficiently compact, then NBA-MP offers lower dimensionality. All the same, this lower dimensionality from models as compared to measurements is not a given. NBA-MD always offers higher fidelity than NBA-MP because extracting a model from measurements always introduces noise into the NBA input data.

How big is big?

Returning to the question of the size of circuit amenable to an NBA analysis, assume that the maximum dimensionality of the input data to the NBA algorithm is 2000. Assume also that the average number of measurement data or model parameters per statistically variant device is 20. Then using the NBA-MD or NBA-MP methods would permit analysis of circuits with 100 statistically variant devices-small to maybe modest-sized analog circuits. Keep in mind that if identical devices are used more than once, then the maximum circuit size is leveraged by the number of times each unique device is used.
Far larger circuits can be accommodated if the NBA-RMP method is employed. With this method, each of 100 statistically variant reference devices can represent tens to even thousands of individual devices in the circuit designs. The number of devices in an amenable circuit then grows into the 1 000 to 100 000 range.

NBA in action
To see how NBA works, consider how it was actually used in the design of a microwave amplifier. NBA-RMP was applied to a high-gain, wide-bandwidth HMMC5040 amplifier manufactured by Hewlett-Packard's Microwave Technology Division. The amplifier's overall design is pictorially rendered in the schematic of Fig. 3. The procedure begins with 108 samples of input data concatenated with response data. The inputs are data derived from reference models of transistors in the process control monitor die. The response data are amplifier performances of interest. Table 1 summarizes the input and response data.



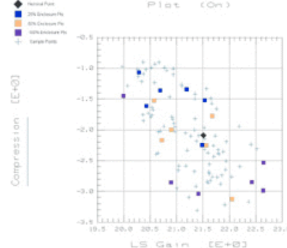
The five transistors are the principal sources of circuit variability in this representation of the Hewlett-Packard HMMC5040 RF amplifier above . The amplifier stages are connected by matching networks of proprietary design.

1. The input and response variables for example analysis on HMMC5040

Input data			Response data	
Symbol	**Description**	**V_{ds}**	**Symbol**	**Description**
ft_{x}, ft_{2x}	Transition frequency	5 V, 2 V	$\|P_{out}\|$	Output power magnitude at 0 dBm input power
I_{max}	Maximum forward-bias gate current	—	$Re(P_{out})$	Real output power at 0 dBm input power
I_{dss}, I_{dss2}	Drain current in	5 V, 2 V	Compression	Gain compression
g_{mx}, g_{m2}	Transconductance	5 V, 2 V	P_{sat}	Saturated output power
V_p, V_{p2}	Pinch-off voltage	5 V, 2 V	LS_gain	Large-signal gain
C_{ds}	Equivalent circuit model drain-source capacitance	—	PAE	Power-added efficiency
C_{gs}, C_{gs2}	Equivalent circuit model gate-source capacitance	5 V, 2 V	RL_{out}	Output return loss
C_{gd}	Equivalent circuit model gate-drain capacitance	—	$<S_{21}$	Phase of S_{21}

The result of most interest is how well nonparametric boundary analysis helps determine a lower bound for the yield. Since acceptable circuit responses are usually specified in terms of independent upper and lower limits for each response, acceptability regions are hyperboxes. (A hyperbox is the generalization of a rectangle in two dimensions and a box in three dimensions to higher-dimensional spaces.) Three enclosure percentages are compared with Monte Carlo yield estimates. The Monte Carlo estimates are then 34, 89, and 97 percent for the hyperboxes enclosing the 25-, 50-, and 100-percent boundary point sets. The expected result is that the analysis enclosure percentages represent lower bounds on the actual yields.

Calculating a lower bound on yield is also possible with parametric worst-case methods so long as the underlying data distributions are Gaussian. Inspection of Fig. 4 illustrates that the compression versus large-signal (LS) gain response data are clearly non-Gaussian. Also, we see overlays of the nominal point and the boundary points for 25, 50, and 100 percent enclosures. Visually, the nominal and boundary points appear consistent with the response scatter although the analysis was done on the input variables.



The response data in this plot of compression versus large-signal gain right is clearly not Gaussian, so parametric methods cannot be used for yield estimation. The nonparametric method, on the other hand, does yield useful results.

The data presented is a projection of the responses, which are eight-dimensional, onto the compression versus LS-gain plane. Interesting problems arise when a human being tries to examine multidimensional boundaries on a two-dimensional sheet of paper-to wit, they can't be seen. This does not mean that the boundary is absent, but rather that the boundary points do not form an enclosed region in this two-dimensional projection.

Acknowledgments
The analysis of the HMMC5040 would not have been possible without the efforts of John Wood and David Root. David Root and Ryan Zurakowski contributed a a number of important ideas over the course of nonparametric boundary analysis's development. Many thanks go to Brian Hughes for his insightful suggestions regarding the flow and presentation of this article.

To probe further
In Handbook of Statistical Methods for Engineers and Scientists, editor Harrison M. Wadsworth supplies an accessible overview of statistical methods for the practicing engineer (McGraw-Hill, New York, 1990). His handbook has considerable breadth, but is geared for those who just want to know how to get the job done.
Michael D. Meehan and John Purviance present a lively review of statistical circuit design in their book Yield and Reliability in Microwave Circuit and System Design (Artech House, Boston, 1993). The authors discuss not only the many algorithms in the field, but also when and how to apply these methods.

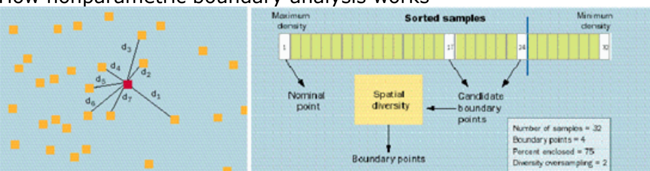
David W. Scott's book, Multivariate Density Estimation Theory, Practice, and Visualization (John Wiley & Sons, New York, 1992), provides the reader with an introduction to density estimation in multivariate (multidimensional) spaces.

Further information about the theory and application of nonparametric boundary analysis is available in the form of U.S. Patent No. 5,835,891, "Device Modeling Using Non-Parametric Statistical Determination of Boundary Data Vectors," which was issued to the author, Dan Stoneking, on 10 November 1998. Dan also wrote "Statistical Circuit Design and IC-CAP's Non-Parametric Boundary Analysis," which is in HP EEsof's Characterization Solutions (Order No. 5965-8931EUS).

About the author
Dan Stoneking joined Hewlett-Packard's EEsof Division in Santa Rosa, Calif., as a software design engineer in 1994. His current responsibilities and interests include behavioral modeling of nonlinear analog circuits and the advancement of statistical circuit design techniques. Before joining HP, he was a senior engineer with M/A-COM in Lowell, Mass. His work there focused on applying statistical techniques to improve both manufacturing and design processes for electronic circuits.

Spectrum editor: Michael J. Riezenman
IEEE Spectrum June 1999 Volume 36 Number 6
(c) Copyright 1999, The Institute of Electrical and Electronics Engineers, Inc.

How nonparametric boundary analysis works



Nonparametric boundary analysis (NBA) finds points lying near the surfaces of equi-density that occur in collections of multidimensional data. The collections are either measurement data or else model parameters that represent specific parts, transistors, or components. The algorithm uses equi-density surfaces because the region interior to the surface contains the most frequently occurring parts. The probability of manufacturing a part that falls in the region is directly related to the manufacturing process yield. Further, it is straightforward to find points near such a surface.

To perform these operations, the NBA algorithm must estimate density for all points in the

data collection and have a technique for finding the equi-density surface appropriate for a given yield. The density estimator uses a nearest neighbor method figure at right . The nearest neighbor method sorts the data, repetitively locates nearest neighbors for each point, calculates distances, and finally computes density estimates.

A specific feature of NBA's method is the density metric formulation. This formulation is advantageous because it does not aggravate the inherent Monte Carlo noise present in the computed average distance to the nearest neighbors. A dimensionally correct density estimate suffers from ever poorer signal-to-noise ratio as the dimensionality of the problem increases. To some extent, the noise can be mitigated by increasing the number of nearest neighbors. But although this smoothes the density estimate, excessive smoothing can obscure important features in the density function. That, in turn, can lead to erroneous results. A good rule of thumb in this regard is to keep the number of nearest neighbors below the number of points in the smallest data cluster.

Once the density estimates are available for all the points, NBA sorts them into order per their density estimate above . NBA then finds the points near a surface of equi-density. The surface of interest is the one that encloses the desired percentage of the distribution. NBA proceeds by placing a cutline in the sorted list that encloses the desired percentage of points as shown. The boundary points are then those to the left of the cutline.

Since clustering of the returned boundary points is not desirable, NBA oversamples to the left of the cutline to produce the candidate boundary points. Obviously, high oversampling degrades the approximation that the candidate boundary points are close to the desired equi-density surface. The algorithm has a heuristic constraint to prevent excessive oversampling.

The candidate boundary points then pass into the spatial diversity algorithm, which outputs the final boundary points to the user. The spatial diversity algorithm works by iteratively forming all candidate boundary points pairs, identifying the pair closest together, and then discarding one point from the closest pair.

The net result of this iterative process is that the final boundary points are spatially distributed along the desired surface of equi-density