Approximately achieving Gaussian relay network capacity with lattice codes

Ayfer Özgür and Suhas Diggavi

Abstract

Recently, it has been shown that a quantize-map-and-forward scheme approximately achieves (within a constant number of bits) the Gaussian relay network capacity for arbitrary topologies [1]. This was established using Gaussian codebooks for transmission and random mappings at the relays. In this paper, we show that the same approximation result can be established by using lattices for transmission and quantization along with structured mappings at the relays.

I. Introduction

Characterizing the capacity of relay networks has been a long-standing open question in network information theory. The seminal work of Cover and El-Gamal [2] has established the basic achievability schemes for relay channels. More recently there has been extension of these techniques to larger networks (see [3] and references therein). In [1], motivated by a deterministic model of wireless communication, it was shown that the quantize-map-and-forward scheme achieves within a constant number of bits from the information-theoretic cutset upper bound. This constant is universal in the sense that it is independent of the channel gains and the operating SNR, though it could depend on the network topology (like the number of nodes).

In the quantize-map-and-forward scheme analyzed in [1], each relay node first quantizes its received signal at the noise level, then randomly maps it to a Gaussian codeword and transmits it. A natural question that we address in this paper is whether lattice codes retain the approximate optimality of the above scheme. This is motivated in part since lattice codes along with lattice decoding could enable computationally tractable encoding and decoding methods. For example lattice codes were used for linear function computation over multiple-access networks [4] and for communication over multiple-access relay networks (with orthogonal broadcast) in [5]. The main result of this paper is to show that the quantize-map-and-forward scheme using nested lattice codes for transmission and quantization, still achieves the Gaussian relay network capacity within a constant. This result is summarized in Theorem 2.1. The use of structured codes allows to specify a structured mapping between the quantization and transmission codebooks at each relay. The nested lattice codebooks considered in this paper are based on the random construction in [6], where they are shown to achieve the capacity of the AWGN channel.

A. Özgür is with EPFL, Lausanne, Switzerland e-mail:ayfer.ozgur@epfl.ch. S. Diggavi is with UCLA, USA and EPFL, Switzerland e-mail: suhas@ee.ucla.edu.
This paper is organized as follows: In Section II, we state the network model and our main result. In Section III, we summarize the construction of the nested lattice ensemble. In Section IV, we describe the network operation. In particular, we specify how we use the nested lattice codes of Section III for encoding at the source, quantization, mapping and transmission at the relay nodes, and decoding at the destination node. In Section V, we analyse the performance achieved by the scheme.

II. MAIN RESULT

We consider a Gaussian relay network where a source node $s$ wants to communicate to a destination node $d$, with the help of $N$ relay nodes, denoted $\mathcal{N}$. The signal received by node $i \in \{s, d, \mathcal{N}\}$ is given by

$$y_i = \sum_{j \neq i} H_{ij} x_j + z_i$$

where $H_{ij}$ is the $N_i \times M_j$ channel matrix from node $j$ comprising $M_j$ transmit antennas to node $i$ comprising $N_i$ receive antennas. Each element of $H_{ij}$ represents the complex channel gain from a transmitting antenna of node $j$ to a receiving antenna of node $i$. The noise $z_i$ is complex circularly-symmetric Gaussian vector $CN(0, \sigma^2 I)$ and is i.i.d. for different nodes. The transmitted signals $x_j$ are subject to an average power constraint $P$.

The following theorem is the main result of this paper.

**Theorem 2.1:** Using nested lattice codes for transmission and quantization along with structured mappings at the relays, we can achieve all rates $R \leq \min_{\Omega} I(X_\Omega; Y_{\Omega^c} | X_{\Omega^c}) - \sum_{i \in \mathcal{N}} N_i$ between $s$ and $d$, where $\Omega$ is a source-destination cut of the network and $X_\Omega = \{X_i, i \in \Omega\}$ are i.i.d. $CN(0, (P/M_i)I)$. It has been shown in [1] that the restriction to i.i.d. Gaussian input distributions is within $\sum_{i \in \mathcal{N},d} N_i$ bits/s/Hz of the cut-set upper bound. Therefore the rate achieved using lattice codes in the above theorem is within $2\sum_{i \in \mathcal{N},d} N_i$ bits/s/Hz to the cut-set upper bound of the network.

For simplicity of presentation, in the rest of the paper we concentrate on a layered network where every node has a single transmit and receive antenna. More precisely, the signal received by node $i$ in layer $l$, $0 \leq l \leq l_d$, denoted $i \in \mathcal{N}_l$, is given by

$$y_i = \sum_{j \in \mathcal{N}_{l-1}} h_{ij} x_j + z_i$$

where $h_{ij}$ is the real scalar channel coefficient from node $j$ to node $i$ and $s \in \mathcal{N}_0$, $d \in \mathcal{N}_{l_d}$. The analysis can be extended to non-layered networks by following the time-expansion argument of [1] and to multicast traffic with multiple destination nodes as well as to multiple multicast where multiple source nodes multicast to a group of destination nodes. The complex case follows by representing each complex number as a two-dimensional real vector. The extension to multiple antennas is discussed inside the text.
III. CONSTRUCTION OF THE NESTED LATTICE ENSEMBLE

Consider a lattice Λ (or more precisely, a sequence of lattices Λ^{(n)} indexed by the lattice dimension n) with $V$ denoting the Voronoi region of Λ. The second moment per dimension of Λ is defined as

$$\sigma^2(\Lambda) = \frac{1}{n} \frac{1}{V} \int_V \|x\|^2 \, dx$$

where $V$ denotes the volume of $V$. We also define the normalized second moment of Λ,

$$G(\Lambda) = \frac{\sigma^2(\Lambda)}{V^{2/n}}.$$

We assume that Λ (or more precisely, the sequence of lattices Λ^{(n)}) is both Roger’s and Poltyrev good. The existence of such lattices has been shown in [7]. Formally, Λ satisfies,

- (Roger’s good) Let $R_u$ and $R_l$ be the covering and effective radius of the lattice Λ. Λ (more precisely the sequence of lattices Λ^{(n)}) is called Roger’s good if its covering efficiency approaches 1 as the dimension $n$ grows,

$$\rho_{cov}(\Lambda) = \frac{R_u}{R_l} \to 1. \quad (1)$$

It is known that a lattice that is good for covering is necessarily good for quantization. A lattice is called good for quantization if

$$G(\Lambda) \to G^*_n \quad (2)$$

where $G^*_n$ is the normalized second moment of an $n$-dimensional sphere and $G^*_n \to \frac{1}{2\pi e}$ when the dimension $n$ becomes large. (2) follows from (1) and the relation

$$G(\Lambda) \leq \frac{n+2}{n} G^*_n (\rho_{cov})^2.$$

- (Poltyrev good) Let $Z$ be a Gaussian random vector whose components are i.i.d. $\mathcal{N}(0, \sigma^2)$, such that $\sigma^2 \leq \sigma^2(\Lambda)$. The volume to noise ratio of the lattice Λ relative to $\mathcal{N}(0, \sigma^2)$ is defined as $\mu = \sigma^2(\Lambda)/\sigma^2$. Then, Λ (more precisely the sequence of such lattices Λ^{(n)}) is called Poltyrev-good if

$$\mathbb{P}(Z \notin V) < e^{-n[E_\mu(\mu) - o_n(1)]}$$

where $E_\mu(\mu)$ is the Poltyrev exponent.

Let the $n \times n$ full-rank generator matrix of Λ be denoted by $G_\Lambda$, i.e., $\Lambda = G_\Lambda \mathbb{Z}^n$. This fixed lattice Λ will serve as the coarse lattice for all our nested lattice constructions below. The fine lattice $\Lambda_1$ is constructed using Loeliger’s type-A construction [8]. Let $k, n, p$ be integers such that $k \leq n$ and $p$ is prime. The fine lattice is constructed using the following steps:

- Draw an $n \times k$ matrix $G$ such that each of its entries is i.i.d according to the uniform distribution over $\mathbb{Z}_p = \{0, 1, \ldots, p - 1\}$. 

---

1For any operation $f : \mathbb{R}^n \to \mathbb{R}^n$ and a set $A \subseteq \mathbb{R}^n$, $f(A) \subseteq \mathbb{R}^n$ denotes $f(A) = \{f(a) : a \in A\}$. 

May 10, 2010 DRAFT
• Form the linear code
\[ C = \{ c : c = G \cdot w, w \in \mathbb{Z}_p^k \}, \] (3)

where "." denotes modulo-p multiplication.

• Lift \( C \) to \( \mathbb{R}^n \) to form
\[ \Lambda'_1 = p^{-1}C + \mathbb{Z}^n. \]

where for two sets \( A \subset \mathbb{R}^n \) and \( B \subset \mathbb{R}^n \), the sum set \( A + B \subset \mathbb{R}^n \) denotes \( A + B = \{ a + b : a \in A, b \in B \} \).

• \( \Lambda_1 = G_{\Lambda} \Lambda'_1 \) is the desired fine lattice. Note that since \( \mathbb{Z}^n \subset \Lambda'_1 \), we have \( \Lambda \subset \Lambda_1 \).

• Draw \( v \) uniformly over \( p^{-1}\Lambda \cap \mathcal{V} \) and translate the lattice \( \Lambda_1 \) by \( v \). The nested lattice codebook consists of all points of the translated fine lattice inside the Voronoi region of the coarse lattice,
\[ \Lambda^* = (v + \Lambda_1) \mod \Lambda = (v + \Lambda_1) \cap \mathcal{V}. \] (4)

In the above equation, we define \( x \mod \Lambda \) as the quantization error of \( x \in \mathbb{R}^n \) with respect to the lattice \( \Lambda \), i.e.,
\[ x \mod \Lambda = x - Q_{\Lambda}(x), \] (5)

where \( Q_{\Lambda}(x) : \mathbb{R}^n \rightarrow \Lambda \) is the nearest-neighbor lattice quantizer defined as,
\[ Q_{\Lambda}(x) = \arg \min_{\lambda \in \Lambda} \| x - \lambda \|. \]

Note that the quantization and mod operations with respect to a lattice can be defined in different ways. The mod operation in (5) maps \( x \in \mathbb{R}^n \) to the Voronoi region \( \mathcal{V} \) of the lattice. More generally, it is possible to define a mod or quantization operation with respect to any fundamental region of the lattice. In particular, when we consider the integer lattice \( \mathbb{Z}^n \) in the sequel, or more generally its multiples \( p \mathbb{Z}^n \) where \( p \) is a positive integer, we will assume that
\[ x \mod p \mathbb{Z}^n = x - \lfloor x \rfloor_p \]
where \( \lfloor x \rfloor_p \) denotes component-wise rounding to the nearest smaller integer multiple of \( p \). In other words, the mod operation with respect to \( p \mathbb{Z}^n \) will map the point \( x \in \mathbb{R}^n \) to the region \( p[0,1)^n \).

The above construction yields a random ensemble of nested lattice codes that has the following desired properties:

• There is a bijection between
\[ \mathbb{Z}_p^n \leftrightarrow p^{-1} \mathbb{Z}_p^n = p^{-1} \mathbb{Z}^n \cap [0,1)^n \leftrightarrow p^{-1} \Lambda \cap G_{\Lambda} [0,1)^n \leftrightarrow p^{-1} \Lambda \cap \mathcal{V}. \]

The last bijection follows from the fact that both \( G_{\Lambda} [0,1)^n \) and \( \mathcal{V} \) are fundamental regions of the lattice \( \Lambda \), i.e., they both tile \( \mathbb{R}^n \). Since \( C \subset \mathbb{Z}_p^n \), the above bijection restricted to \( C \) yields,
\[ C \leftrightarrow p^{-1}C = \Lambda'_1 \cap [0,1)^n \leftrightarrow \Lambda_1 \cap G_{\Lambda} [0,1)^n \leftrightarrow \Lambda_1 \cap \mathcal{V} \leftrightarrow \Lambda^*. \] (6)

In the sequel, we slightly abuse notation by using \( C \) to denote both the code over the finite field and its projection to the reals. Hence, the codewords \( c \) are either considered as vectors in \( \mathbb{Z}_p^n \), in which case they are subject to finite field operations, or they are considered as vectors in \( \mathbb{R}^n \) subject to real field operations. It is to be deduced from the context to which of these two cases the notation refers to.
Note, therefore, that $\Lambda^* \subseteq p^{-1}\Lambda \cap \mathcal{V}$. The bijections above can be explicitly specified in both directions and we will make use of this fact in the next section. Note that $w$ in (3) runs through all the $p^k$ vectors in $\mathbb{Z}_p^k$. Let us index these vectors as $w(i)$, $i = 0, \ldots, p^k - 1$. Let us index the corresponding codewords in $C$ as $C(i) = G \cdot w(i)$, $i = 0, \ldots, p^k - 1$. The $p^k$ codewords in $C$ need not be distinct. By the bijection in (6), each codeword in $C$ corresponds to one fine lattice point in $\Lambda_1 \cap \mathcal{V}$ and one codeword of $\Lambda^*$. Let us similarly index the points in $\Lambda_1 \cap \mathcal{V}$ as $\Lambda_1(i)$ and the corresponding codewords of $\Lambda^*$ as $\Lambda^*(i)$, for $i = 0, \ldots, p^k - 1$. We have,

$$\Lambda_1(i) = G_A p^{-1} C(i) \mod \Lambda, \quad \Lambda^*(i) = (v + \Lambda_1(i)) \mod \Lambda.$$  

The random codebook $\Lambda^*$ has the following statistical properties:

- Let $\lambda \in p^{-1}\Lambda \cap \mathcal{V}$,
  $$\mathbb{P}(\Lambda^*(i) = \lambda) = \frac{1}{|p^{-1}\Lambda \cap \mathcal{V}|} = \frac{1}{p^n},$$  

- Let $\lambda_1, \lambda_2 \in p^{-1}\Lambda \cap \mathcal{V}, \quad \forall i \neq j$,
  $$\mathbb{P}(\Lambda^*(i) = \lambda_1, \Lambda^*(j) = \lambda_2) = \frac{1}{|p^{-1}\Lambda \cap \mathcal{V}|^2} = \frac{1}{p^{2n}}.$$  

In other words, the construction in this section yields an ensemble of nested lattice codes such that each codeword of the random codebook $\Lambda^*$ is uniformly distributed over $p^{-1}\Lambda \cap \mathcal{V}$ and the codewords of $\Lambda^*$ are pairwise independent. These two properties suffice to prove the random coding result of this paper.

**Proof of properties (7) and (8):** The first property (7) simply follows from the fact that $v$ is uniformly distributed on $p^{-1}\Lambda \cap \mathcal{V}$. For the second probability, we have

$$\mathbb{P}(\Lambda^*(i) = \lambda_1, \Lambda^*(j) = \lambda_2)
\begin{align*}
&= \mathbb{P}((v + \Lambda_1(i)) \mod \Lambda = \lambda_1, \quad (v + \Lambda_1(j)) \mod \Lambda = \lambda_2) \\
&= \mathbb{P}((v + \Lambda_1(i)) \mod \Lambda = \lambda_1, \quad (v + \Lambda_1(j)) \mod \Lambda - (v + \Lambda_1(i)) \mod \Lambda = \lambda_2 - \lambda_1) \\
&= \mathbb{P}(\Lambda_1(i) = (\lambda_1 - v) \mod \Lambda, \quad (\Lambda_1(j) - \Lambda_1(i)) \mod \Lambda = (\lambda_2 - \lambda_1) \mod \Lambda) \\
&= \mathbb{P}((\Lambda_1(j) - \Lambda_1(i)) \mod \Lambda = (\lambda_2 - \lambda_1) \mod \Lambda) \\
&\quad \times \mathbb{P}(\Lambda_1(i) = (\lambda_1 - v) \mod \Lambda, (\Lambda_1(j) - \Lambda_1(i)) \mod \Lambda = (\lambda_2 - \lambda_1) \mod \Lambda). 
\end{align*}$$

Note that the first probability in (9) is independent of $v$. Let us denote $\lambda = (\lambda_2 - \lambda_1) \mod \Lambda \in p^{-1}\Lambda \cap \mathcal{V}$, we have

$$\begin{align*}
(\Lambda_1(j) - \Lambda_1(i)) \mod \Lambda &= \lambda \\
\Leftrightarrow & (G_A p^{-1} C(j) - G_A p^{-1} C(i)) \mod \Lambda = \lambda \\
\Leftrightarrow & (G_A p^{-1} C(j) - G_A p^{-1} C(i)) = \lambda + x, \quad x \in \Lambda \\
\Leftrightarrow & (C(j) - C(i)) = p G_A^{-1} \lambda + p G_A^{-1} x, \quad p G_A^{-1} x \in p \mathbb{Z}^n \\
\Leftrightarrow & (C(j) - C(i)) \mod p \mathbb{Z}^n = p G_A^{-1} \lambda \mod p \mathbb{Z}^n \\
\Leftrightarrow & G \cdot (w(j) - w(i)) = c, \\
\end{align*}$$

May 10, 2010  DRAFT
where all equations except the last one are over the reals. The last equation (11) is a restatement of (10) in terms of finite field operations with $c = p \cdot \mathbf{G}^{-1} \cdot \lambda \mod p \mathbb{Z}^n$. Since $j \neq i$, the vector $\mathbf{w}(j) - \mathbf{w}(i)$ has at least one nonzero entry. Since the corresponding column of $G$ is uniformly distributed over $\mathbb{Z}^n_p$, we have

$$P(G \cdot (\mathbf{w}(j) - \mathbf{w}(i)) = c) = P((\Lambda_1(j) - \Lambda_1(i)) \mod \Lambda = (\lambda_2 - \lambda_1) \mod \Lambda) = \frac{1}{p^n}. $$

For the second probability in (9), it is easy to observe that for any realization of $G$, hence $\Lambda_1(i)$, there is exactly one choice of $v$ out of $p^n$ possible choices that satisfies the equality $\Lambda_1(i) = (\lambda_1 - v) \mod \Lambda$. Combining these observations yields the conclusion in (8).

The above construction yields a random ensemble of nested lattice pairs $\Lambda \subseteq \Lambda_1$ with coding rate,

$$R = \frac{1}{n} \log |\Lambda^*|$$

which can be tuned by choosing the precise magnitudes of $k$ and $p$. Note that $|\Lambda^*| = p^k$ if the random matrix $G$ in (3) is full rank. The probability that $G$ is not full rank can be upper bounded by

$$P(\text{rank}(G) < k) = \sum_{\mathbf{w} \in \mathbb{Z}^n_p, \mathbf{w} \neq 0} P(G \cdot \mathbf{w} = 0) = (p^k - 1) p^{-n}. $$

Therefore if $k \leq \beta n$ for $\beta < 1$, the above probability decreases to zero at least exponentially as $n$ increases ($p$ may also grow with $n$). This will be the case for our nested lattice codes in the next section.

**IV. ENCODING, MAPPING AND DECODING**

In the previous section, we have constructed an ensemble of nested lattices where the coarse lattice $\Lambda$ is fixed and the fine lattice $\Lambda_1$ is randomized. It has been shown in [9] that with high probability, a nested lattice $(\Lambda_1, \Lambda)$ in this ensemble is such that both $\Lambda_1$ and $\Lambda$ are Roger’s and Poltyrev-good. (The fixed lattice $\Lambda$ is Roger’s and Poltyrev-good by construction.) For quantization, we fix one such good member of the ensemble and use it at all the relay nodes. For transmission, we draw a random nested lattice codebook from this ensemble, independently at each relay. The mapping between the quantization and transmission codebooks at each relay is specified below.

**Source:** The source has $p^k$ messages, where $p$ is prime and $k \leq n$. The messages are represented as length-$k$ vectors over the finite field $\mathbb{Z}_p$ and mapped to a random nested lattice codebook $\Lambda^*$ following the construction in Section III. In the construction, the coarse lattice $\Lambda$ is scaled such that its second moment,

$$\sigma^2(\Lambda^T) = \frac{n}{n+2} \cdot \frac{1}{\rho_{\text{conv}}(\Lambda^T)^2} P,$$

where $\Lambda^T$ now denotes the scaled version of the lattice $\Lambda$ to satisfy the power constraint. Note that $\sigma^2(\Lambda^T) \to P$ as $n$ increases since $\Lambda^T$ is Roger’s good. This choice ensures that every codeword of $\Lambda^*$ satisfies the power constraint $P$. This result is stated in the Proposition 4.1 below. The information rate of the code is given by

$$R = \frac{1}{n} \log p^k.$$
Let us denote by $x_s^{(w)}$, $w \in \{1, \ldots, e^{nR}\}$ the random transmit codewords corresponding to each message $w$ of the source node.

**Proposition 4.1:** Each transmitted codeword $x_s^{(w)}$ satisfies the transmit power constraint $P$.

**Proof of Proposition 4.1:** Since every transmitted codeword $x_s^{(w)} \in V^T$, we have

$$\frac{1}{n} \|x_s^{(w)}\|^2 \leq \frac{1}{n} (R_u^T)^2,$$

where $R_u^T$ is the covering radius of $\Lambda^T$. We now relate the covering radius $R_u^T$ of $\Lambda^T$ to its second moment $\sigma^2(\Lambda^T)$. Let $G_n^*$ be the normalized second moment of the $n$-dimensional sphere $B(R_l^T)$ of radius $R_l^T$. We have the identity

$$G_n^* \left| B(R_l^T) \right|^{2/n} = \left( \frac{(R_l^T)^2}{n+2} \right).$$

Since $|V^T| = |B(R_l^T)|$ when $R_l^T$ is the effective radius of $\Lambda^T$, we have

$$R_l^T = \sqrt{\frac{n+2}{n} \frac{G_n^*}{G(\Lambda^T)} \sqrt{n} \sigma^2(\Lambda^T)}.$$

Thus, the covering radius of the lattice $\Lambda^T$ is given by

$$R_u^T = \rho_{cov}(\Lambda^T) \sqrt{\frac{n+2}{n} \frac{G_n^*}{G(\Lambda^T)} \sqrt{n} \sigma^2(\Lambda^T)} \quad (13)$$

This expression together with our choice in (12), yields

$$\frac{1}{n} \|x_s^{(w)}\|^2 \leq P.$$

**Relays:** The relay node $i$ receives the signal $y_i$.

**Quantize:** The signal $y_i$ is first quantized by using a nested lattice codebook that has been generated by the construction in Section III. It is shown in [9] that this construction yields nested lattices where the fine lattice is Roger’s good with high probability if $k \geq (\log n)^2$. (The coarse lattice is both Roger’s and Poltyrev good by construction.) We fix one such good nested lattice $(\Lambda_Q^1, \Lambda_Q^2)$ and use the corresponding codebook $\Lambda_Q^* = \Lambda_Q^1 \mod \Lambda_Q^2$ at all the relay nodes for quantization. Therefore our quantization codebook is not random but fixed and moreover same for all relay nodes. We assume that the nested lattice $(\Lambda_Q^1, \Lambda_Q^2)$ has been generated by using the following parameters: Let

$$D_s = \max_i \sum_{j \in N_i-1} |h_{ij}|^2 P. \quad (14)$$

The coarse lattice $\Lambda_Q^1$ is a scaled version of the lattice $\Lambda$ such that

$$\sigma^2(\Lambda_Q^1) = 2\eta(D_s + \sigma^2) \quad (15)$$

for a constant $\eta > 0$. Recall that $\sigma^2$ is the noise variance. We denote the generator matrix of the scaled coarse lattice $\Lambda_Q^1$ by $G_{\Lambda_Q^1}$. The parameters $k_r$ and $p_r$ are chosen such that $k_r = (\log n)^2$ and $p_r$ is the prime number.
such that
\[ p_r^{k_r} = e^{nR_r}, \quad \text{where} \quad R_r = \frac{1}{2} \log \frac{\sigma^2(\Lambda^Q)}{\sigma^2}. \] (16)

Note that since \( R_r \) is independent of \( n \), \( p_r = e^{\frac{nR_r}{\log n^2}} \), i.e., \( p_r \to \infty \) as \( n \to \infty \). With the choice in (16) for \( R_r \), the second moment of \( \Lambda^Q \) is given by
\[ \sigma^2(\Lambda^Q) = \frac{G(\Lambda^Q)}{G(\Lambda)} \sigma^2. \] (17)

Since both \( \Lambda^Q \) and \( \Lambda \) are Roger’s good, \( \sigma^2(\Lambda^Q) \to \sigma^2 \) when \( n \) increases. Therefore, we are effectively quantizing at the noise level.

The quantized signal is given by
\[ \hat{y}_i = Q_i(y_i + u_i) \mod \Lambda^Q \]
where \( u_i \) is a random dither known at the destination node and uniformly distributed over the Voronoi region \( V^Q_i \) of the fine lattice \( \Lambda^Q_i \). The dithers \( u_i \) are independent for different nodes.

**Map and Forward:** Let us scale the coarse lattice \( \Lambda \) such that its second moment \( \sigma^2(\Lambda^T) \) is given by (12). Let \( G_{\Lambda^T} \) denote the generator matrix of the scaled coarse lattice. The quantized signal \( \hat{y}_i \) at relay \( i \) is mapped to the transmitted signal \( x_i \) by the following mapping,
\[ x_i = G_{\Lambda^T} p_r^{-1}(G_i p_r(G_{\Lambda^Q_i} \hat{y}_i \mod \mathbb{Z}^n) \mod p_r \mathbb{Z}^n) + v_i \mod \Lambda^T, \] (18)
where \( G_i \) is an \( n \times n \) random matrix with its entries uniformly and independently distributed in \( 0, 1, \ldots, p_r - 1 \) and \( v_i \) is a random vector uniformly distributed over \( p_r^{-1} \Lambda^T \cap V^T \), where \( V^T \) is the Voronoi region of \( \Lambda^T \). \( G_i \) and \( v_i \) are independent for different relay nodes. We index the \( e^{nR_r} \) codewords of \( \Lambda^Q \) as \( \hat{y}_i^{(k_i)} \), \( k_i \in \{1, \ldots, e^{nR_r}\} \). The corresponding sequence that the codeword \( \hat{y}_i^{(k_i)} \) is mapped to in (18), is denoted by \( x_i^{(k_i)} \).

**Proposition 4.2:** The above mapping has the following properties:

- At each relay \( i \), the transmitted sequences \( x_i \in \Lambda^*_i \), where \( \Lambda^*_i \) is a random nested lattice codebook.

- The mapping induces a pairwise independent and uniform distribution over \( p_r^{-1} \Lambda^T \cap V^T \). Formally, each quantization codeword \( \hat{y}_i^{(k_i)} \in \Lambda^*_Q \) is mapped uniformly at random to the set \( p_r^{-1} \Lambda^T \cap V^T \). Two codewords \( \hat{y}_i^{(k_i)}, \hat{y}_i^{(k'_i)} \in \Lambda^*_Q \) such that \( k_i \neq k'_i \) are mapped independently.

- The mapping induces an independent distribution across the relays.

**Proof of Proposition 4.2:** The proposition says that the quantization codebooks at each relay are independently mapped to a random nested lattice codebook from the ensemble constructed in the earlier section. The proof is based on the bijection given in (6): There is one-to-one correspondence between the codebook \( \Lambda^*_Q \) and its underlying finite field codebook \( C^Q \). The mapping in (18) first takes the codeword \( \hat{y}_i \in \Lambda^*_Q \) to its corresponding codeword in

---

\(^3\)To be more precise, one can take \( p_r \) to be the largest prime number such that \( p_r \leq e^{nR_r/k} \) in which case the rate of the code is \( \frac{1}{n} \log p_r^{k_r} \leq R_r \). When \( n \) is large, the difference becomes negligible and is therefore ignored.
\( C_1 \). Note that
\[
\hat{y}_i \in \Lambda_1^Q \quad \Rightarrow \quad G_A^{-1} \hat{y}_i \in p_r^{-1} \mathbb{Z}^n
\]
\[
\Rightarrow \quad G_A^{-1} \hat{y}_i \mod \mathbb{Z}^n \in p_r^{-1} \mathbb{Z}^n \cap [0,1)^n
\]
\[
\Rightarrow \quad p_r \left( G_A^{-1} \hat{y}_i \mod \mathbb{Z}^n \right) \in \mathbb{Z}_p^n.
\]
Therefore, \( p_r \left( G_A^{-1} \hat{y}_i \mod \mathbb{Z}^n \right) \in C_1 \). This codeword in \( C_1 \) is then mapped to a random finite-field codebook \( C_i = \{ c' : c' = G_i \cdot c, \ c \in C_1 \} \). We finally form the nested lattice codebook \( \Lambda_i^* \) corresponding to \( C_i \) following again the construction of the earlier section. Note that, for \( c' \in C_i \),
\[
G_{\Lambda_r} p_r^{-1} c' + v_i \mod \Lambda_i^* \in \Lambda_i^*,
\]
where \( \Lambda_i^* = (v_i + \Lambda_1^T_{1,i}) \mod \Lambda^T \) and \( \Lambda_1^T_{1,i} \) is the fine lattice generated by \( C_i \). The second property follows by similar observations as in Section III: The random matrix \( G_i \) maps every nonzero vector \( c \in C_1 \) uniformly at random to another finite field vector in \( \mathbb{Z}_p^n \). The third property follows from the independence of the \( G_i \)'s and \( v_i \)'s for different nodes \( i \).

The mapping in (18) can be simplified to the form,
\[
x_i = G_{\Lambda_r} G_i G_A^{-1} \hat{y}_i + v_i \mod \Lambda_i^*.
\]
Effectively, it takes the quantization codebook \( \Lambda_1^* \), expands it by multiplying with a random matrix with large entries (of the order of \( p_r \)) and then folds it to the Voronoi region of \( \Lambda_i^* \). Since the entries of \( G_i \) are potentially very large, even if two codewords are close in \( \Lambda_i^* \), they are mapped independently to the codewords of the transmit codebook. Note that the complexity of the mapping is polynomial in \( n \), while random mapping of the form in [1] has exponential complexity in \( n \).

**Destination:** Given its received signal \( y_d \), together with the knowledge of all codebooks, mappings, dithers and channel gains, the decoder performs a consistency check to recover the transmitted message. For each relay \( i \) and quantization codeword \( \hat{y}_i^{(k_i)} \), it first forms the signals
\[
\tilde{y}_i^{(k_i)} = \hat{y}_i^{(k_i)} - u_i \mod \Lambda_i^*.
\]
Note that for \( i \in \mathcal{N}_i \)
\[
\tilde{y}_i = \hat{y}_i - u_i \mod \Lambda_i^* = Q_{\Lambda_2^Q} (y_i + u_i) - u_i \mod \Lambda_i^* = (y_i - (y_i + u_i) \mod \Lambda_i^*) \mod \Lambda_i^* = \sum_{j \in \mathcal{N}_i - i} h_{ij} x_j + z_i - u_i' \mod \Lambda_i^*,
\]
where \( u_i' = (y_i + u_i) \mod \Lambda_i^* \). \( u_i' \) is independent of \( y_i \) and is uniform over the Voronoi region of \( \Lambda_i^* \) (Crypto Lemma, see [6]).
The decoder then checks the set \( \mathcal{W} \) of messages \( \hat{w} \) for which there exists some indices \( k_i \) such that

\[
(x_s^{(\hat{w})}, y_d, \{ y_i^{(k_i)}, x_i^{(k_i)} \}_{i \in N}) \in \mathcal{A}_c
\]

where \( \mathcal{A}_c \) denotes consistency. We define consistency as follows: For a given set of indices \( \{k_i\}_{i \in N} \), we say

\[
(x_s^{(\hat{w})}, y_d, \{ y_i^{(k_i)}, x_i^{(k_i)} \}_{i \in N}) \in \mathcal{A}_c \text{ if }

\| (y_i^{(k_i)} - \sum_{j \in N_{i-1}} h_{ij} x_j^{(k_j)}) \mod \Lambda^Q \|^2 \leq n \sigma_c^2,
\]

(22)

for all \( i \in N_i, 1 \leq l \leq l_d \) where for convenience of notation we have denoted \( x_s^{(\hat{w})} = x_j^{(k_j)}, j \in N_0 \), and \( y_d = y_i^{(k_i)}, i \in N_d \). We choose

\[
\sigma_c^2 = (1 + \epsilon) 2\sigma^2
\]

(23)

for a constant \( \epsilon > 0 \) that can be taken arbitrarily small. We can interpret the consistency check as follows: For each layer \( l = 1, \ldots, l_d - 1 \) the decoder picks a set of potential (quantized) received sequences \( \{ y_i^{(k_i)} \}_{i \in N_i} \) and the transmit sequences corresponding to them \( \{ x_i^{(k_i)} \}_{i \in N_i} \). It checks for each layer \( l \), whether the inputs and outputs are consistent, i.e., whether the examined inputs \( \{ x_i^{(k_i)} \}_{i \in N_{i-1}} \) of the layer \( l \) could have generated the examined outputs \( \{ y_i^{(k_i)} \}_{i \in N_i} \) based on the second order statistics of the thermal and quantization noise. Note that the definition of consistency in (22) is closely related to weak typicality. Indeed, it is a variant of the weak typicality condition for Gaussian vectors. Therefore, effectively our decoder is a typicality decoder.

### A. Multiple Antennas

A slightly modified version of the above scheme applies to the case of multiple transmit and receive antennas at each node. Let \( M_i \) be the number of transmit and \( N_i \) be the number of receive antennas at each node.

**Source:** The source node \( s \) maps its message to \( M_s \) independent nested lattice codebooks \( \Lambda_{1,s}^*, \ldots, \Lambda_{M_s}^* \) and transmits its codeword from its corresponding transmit antenna.

**Relays:** The relay node \( i \) receives \( N_i \) signals denoted \( y_{i,1}, \ldots, y_{i,N_i} \). It individually quantizes each signal by adding an independent random dither,

\[
\tilde{y}_{i,l} = Q\Lambda^Q (y_{i,l} + u_{i,l}) \mod \Lambda^Q, \quad l = 1, \ldots, N_i.
\]

The transmitted codeword from the \( m \)'th transmit antenna of node \( i \) is given by

\[
x_{i,m} = G_{\Lambda^T} \sum_{l=1}^{N_i} G_{i,m,l}^{-1} \tilde{y}_{i,l} + v_{i,m} \mod \Lambda^T.
\]

(24)

where \( G_{i,m,l} \) is \( n \times n \) random matrix independent across \( i, m \) and \( l \). The mapping is modified from (19) so that at each relay, the set of quantization codewords \( \tilde{y}_{i,1}^{(k_i)}, \ldots, \tilde{y}_{i,N_i}^{(k_i)} \) is mapped independently to \( M_i \) random nested lattice codebooks. For each of the \( M_i \) random codebooks, two different sets of quantization codewords
The source node is naturally in the distinguishability set $\Omega$ as introducing a notion of distinguishability. The relay nodes in $k_i$ indices we can carry out the entire calculation conditioned on this and then average over it. The summation over the $k_i$ indices

$$N$$

is a source-destination cut of the network, i.e., $\Omega = \{s, N_\Omega \}$ where $N_\Omega$ is a subset of the relaying nodes $N$. The first summation runs over all possible source-destination cuts $\Omega$ of the network, or equivalently over all subsets $N_\Omega$ of the relaying nodes $N$. Following [1], the rearrangement of the summation above can be interpreted as introducing a notion of distinguishability. The relay nodes in $\Omega$ are the ones that can distinguish between $w$ and $w'$ because $y_i^{(k_i)} \neq y_i^{(k_i)}$, when the relay nodes in $\Omega^c$ cannot distinguish between $w$ and $w'$ because $y_i^{(k_i)} = y_i^{(k_i)}$.

The source node is naturally in the distinguishability set $\Omega$ and the destination node is in $\Omega^c$. Thus in (25), we sum over all possible cases for the distinguishability set $\Omega$.

Now, let us examine the probability denoted by $P_e$. For a given set of $\{k_i^j\}_{i \in N}$ such that $k_i^j = k_i$, $i \in N_\Omega^c$ and $k_i^j \neq k_i$, $i \in N_\Omega$, the consistency condition is given by (22) as

$$\|y_i^{(k_i)} - \sum_{j \in N_{i-1}} h_{ij}x_j^{(k_i)}\| \mod Q^i \leq n_\sigma^2, \quad \forall i \in \{N, d\},$$

V. Error Analysis

An error occurs if the transmitted message $w$ is not in the list, i.e., $w \notin \hat{W}$ or when $w' \neq w$ is also in the list $\hat{W}$. The correct message $w$ is in the list with high probability.

We concentrate on the probability that there exist an error because $w$ is not the unique message in $\hat{W}$. This probability can be upper bounded by concentrating on the pair-wise error probabilities, i.e.,

$$P_e \leq e^{nR} P(w \rightarrow w')$$

where $P(w \rightarrow w')$ denotes the probability that $w' \in \hat{W}$ given $w$ was the transmitted message. This probability is given by

$$P\left(\exists \{k_i^j\}_{i \in N^\Omega} \text{ s.t. } (x_s^{(w')}, y_d, \{y_i^{(k_i^j)}\}_{i \in N^\Omega}) \in A_e \right) \leq \sum_{k_i^j, k_i^{j'}} P\left((x_s^{(w')}, y_d, \{y_i^{(k_i^j)}, x_i^{(k_i^{j'})}\}_{i \in N^\Omega}) \in A_e \right)$$

We can condition on the event that the correct message produces indices $\{k_i\}$, and since this is a generic index, we can carry out the entire calculation conditioned on this and then average over it. The summation over the $N$ indices $k_i^1, \ldots, k_i^{N_i}$ above can be rearranged to yield

$$\sum_{\Omega} \sum_{k_i^j, k_i^{j'}, \Omega} P\left((x_s^{(w')}, y_d, \{y_i^{(k_i^j)}, x_i^{(k_i^{j'})}\}) \in A_e \text{ s.t. } k_i^j = k_i, \ i \in N_\Omega^c, \ k_i^{j'} \neq k_i \right),$$

where $\Omega$ is a source-destination cut of the network, i.e., $\Omega = \{s, N_\Omega \}$ where $N_\Omega$ is a subset of the relaying nodes $N$. The first summation runs over all possible source-destination cuts $\Omega$ of the network, or equivalently over all subsets $N_\Omega$ of the relaying nodes $N$. Following [1], the rearrangement of the summation above can be interpreted as introducing a notion of distinguishability. The relay nodes in $\Omega$ are the ones that can distinguish between $w$ and $w'$ because $y_i^{(k_i)} = y_i^{(k_i)}$, when the relay nodes in $\Omega^c$ cannot distinguish between $w$ and $w'$ because $y_i^{(k_i)} = y_i^{(k_i)}$.

The source node is naturally in the distinguishability set $\Omega$ and the destination node is in $\Omega^c$. Thus in (25), we sum over all possible cases for the distinguishability set $\Omega$.
where for convenience of notation we denote $y_d = y^{(k_i)}_i$, $i \in \mathcal{N}_d$, and $x^{(w')}_j = x^{(k_j')}_j$, $j \in \mathcal{N}_0$. The condition in (26) takes two different forms depending on whether $i \in \mathcal{N}_1$ or $i \in \Omega^c$. For nodes $i \in \Omega^c$, the condition is equivalent to
\[
\mathcal{A}_i = \{\|\left( \sum_{j \in \Omega_{i-1}} h_{ij}(x^{(k_i)}_j - x^{(k'_j)}_j) + z_i - u'_i \right) \mod \Lambda^Q\|^2 \leq n \sigma^2 \},
\]
where $\Omega_{i-1} = \Omega \cap \mathcal{N}_{i-1}$ and we denote this event by $\mathcal{A}_i$.\(^4\) For nodes $i \in \mathcal{N}_1$, the condition yields
\[
\mathcal{B}_i = \{\|\left( \tilde{y}_i^{(k'_i)} - \sum_{j \in \Omega_{i-1}} h_{ij}x^{(k_j)}_j - \sum_{j \in \Omega_{i-1}} h_{ij}x^{(k'_j)}_j \right) \mod \Lambda^Q\|^2 \leq n \sigma^2 \},
\]
where $\Omega_{i-1}^c = \Omega^c \cap \mathcal{N}_{i-1}$ and we denote this event by $\mathcal{B}_i$. We have,
\[
\mathcal{P} = \mathbb{P}(\{\mathcal{A}_i, i \in \Omega^c\}, \{\mathcal{B}_i, i \in \mathcal{N}_1\}) = \mathbb{P}(\mathcal{A}_i, i \in \Omega^c) \mathbb{P}(\mathcal{B}_i, i \in \mathcal{N}_1 | \mathcal{A}_i, i \in \Omega^c).
\]
Note that due to Proposition 4.2, $x^{(k_j)}, x^{(k'_j)}_j, j \in \{s, \mathcal{N}\}$ in expressions (27) and (28) are a set of independent random variables, uniformly distributed over $p^{-1} \Lambda^T \cap \mathcal{V}^T$.\(^5\) Due to the dithering in (20), $y_i^{(k_i)}$ in (28) is uniformly distributed over the Voronoi region $\mathcal{V}_i^Q$ of the quantization lattice point $\hat{y}_i^{(k_i)}$.

We will first bound the probability $\mathbb{P}(\mathcal{A}_i, i \in \Omega^c)$ by conditioning on the event defined in the following lemma.

**Lemma 5.1:** Let us define $\mathcal{E}_1$ to be the following event,
\[
\{\exists i \in \{\mathcal{N}, d\}, \exists \{k_j, k'_j\} s.t. \sum_j h_{ij}(x^{(k_j)}_j - x^{(k'_j)}_j) + z_i - u'_i \notin \mathcal{V}^Q \}.
\]
We have $\mathbb{P}(\mathcal{E}_1) \to 0$ as $n \to \infty$.

When $\mathcal{E}_1$ is true, we declare this as an error. This adds a vanishing term to the decoding error probability by the above lemma. Conditioning on the complement of $\mathcal{E}_1$ allows us to get rid of the mod operation w.r.t $\Lambda^Q$ in (27). Given $\mathcal{E}_1^c$, the condition $\mathcal{A}_i$ is equivalent to
\[
\mathcal{A}_i' = \{\|\left( \sum_{j \in \Omega_{i-1}} h_{ij}(x^{(k_j)}_j - x^{(k'_j)}_j) + z_i - u'_i \right)\|^2 \leq n \sigma^2 \}.
\]
Therefore, we have
\[
\mathbb{P}(\mathcal{A}_i, i \in \Omega^c | \mathcal{E}_1^c) = \mathbb{P}(\mathcal{A}_i', i \in \Omega^c | \mathcal{E}_1^c) \leq \frac{\mathbb{P}(\mathcal{A}_i', i \in \Omega^c)}{\mathbb{P}(\mathcal{E}_1^c)}
\]
We upperbound the last probability in the following lemma.

\(^4\)The condition is slightly different for the destination node $d$, in particular it does not contain the term $u'_i$ in (27), since we operate directly on the observation $y_d$ and not it's quantized version. This fact is ignored since it does not create any significant difference in the below analysis. Alternatively, it can be assumed that the destination node first quantizes its received signal and then performs the consistency check.

\(^5\)For the source node, $x^{(k_j)}_j$ and $x^{(k'_j)}_j$ or equivalently $x^{(w)}_j$ and $x^{(w')}_j$ are uniformly distributed over $p^{-1} \Lambda^T \cap \mathcal{V}^T$ where $p$ is different than $p_T$. However, this fact does not create any difference in the following analysis and is therefore ignored.
Lemma 5.2:
\[
P\left( \left\| \sum_{j \in \Omega_{t-1}} h_{ij}(x_j^{(k_i)} - x_j^{(k_j')}) + z_i - u_i' \right\|^2 \leq n \sigma_c^2, \forall i \in \Omega^c \right) \\
\leq e^{-n\left(I(X_{it}; HX_{it} + Z_{it^c}) - \frac{1}{2} |\Omega^c| (1+\log(1+\epsilon)) - o_n(1)\right)},
\]

where \(X_{it}, i \in \Omega\) are i.i.d Gaussian random variables \(\mathcal{N}(0, P)\), \(Z_{it^c}\) are i.i.d Gaussian random variables \(\mathcal{N}(0, \sigma^2)\) and \(H\) is the channel transfer matrix from nodes in \(\Omega\) to nodes in \(\Omega^c\).

The proof of the lemma involves two main steps. Recall that \(x_j^{(k_i)}, x_j^{(k_j')}, j \in \Omega\) are discrete random variables, independently and uniformly distributed over \(p_r^{-1} \Lambda^n \cap \nu^n\). We first show that the probability in the lemma is upper bounded by
\[
e^{-n(\epsilon)} P\left( \left\| \sum_{j \in \Omega_{t-1}} h_{ij}(x_j - x_j') + z_i - z_i' \right\|^2 \leq n \sigma_c^2, \forall i \in \Omega^c \right) \tag{29}
\]
where \(x_j, x_j', j \in \Omega\) and \(z_i', i \in \Omega^c\) are all independent Gaussian random variables such that \(x_j, x_j' \sim \mathcal{N}(0, \sigma^2 I_n)\), \(z_i' \sim \mathcal{N}(0, \sigma^2 I_n)\) and \(\sigma^2 \to \sigma^2(\Lambda^n) \to P\) as \(n \to \infty\) when \(\Lambda^n \) is Roger's good, \(\sigma^2 \to \sigma^2(\Lambda_1^n) \to \sigma^2\) as \(n \to \infty\) when \(\Lambda_1^n \) is Roger's good which is our case here. \(\epsilon \to 0\) when \(n\) increases, again if \(\Lambda^n \) and \(\Lambda_1^n \) are Roger's good. Given this translation to Gaussian distributions the problem becomes very similar to the one for Gaussian codebooks in [1]. The second step is to bound the probability in (29) by following a similar approach to [1]. The proof is given in the appendix.

We will now upper bound the term
\[
\sum_{k'_{i}, i \in \mathcal{N}_{\Omega} \mid A_{i} , i \in \Omega^c} P( B_{i}, i \in \mathcal{N}_{\Omega} \mid A_{i} , i \in \Omega^c) \tag{30}
\]
by first removing the condition \(k'_{i} \neq k_{i}\) in the summation above and then noting that this term is equal to \(e^{n|\mathcal{N}_{\Omega}| n R_r}\) times the probability \(P( B_{i}, i \in \mathcal{N}_{\Omega} \mid A_{i} , i \in \Omega^c)\) evaluated for a randomly and independently chosen set of indices \(\{k'_{i}\}_{i \in \mathcal{N}_{\Omega}}\). When each of the indices \(\{k'_{i}\}_{i \in \mathcal{N}_{\Omega}}\) is chosen uniformly at random, \(\nu_{i}(k_{i})\) in (28) is a random variable uniformly distributed over \(\nu^Q\). This is due to the dithering over the Voronoi region \(\nu^Q\) of the fine lattice and the mod operation with respect to the coarse lattice \(\Lambda^Q\) in (20). Moreover, by the Crypto Lemma,
\[
\nu_{i} = \nu_{i}(k_{i}) - \sum_{j \in \Omega_{t-1}} h_{ij}(x_j^{(k_{i})}) - \sum_{j \in \Omega_{t-1}} h_{ij}(x_j^{(k_{i})}) \mod \Lambda^Q
\]
is also uniformly distributed over \(\nu^Q\) and is independent of
\[
\sum_{j \in \Omega_{t-1}} h_{ij}(x_j^{(k_{i})}) + \sum_{j \in \Omega_{t-1}} h_{ij}(x_j^{(k_{i})}).
\]
This is due to the fact that \(\nu_{i}(k_{i})\) is independent of this term, which is due to the fact the index \(k'_{i}\) and the dither \(u_{i}\) are chosen independently of everything else. Therefore (30) is upper bounded by
\[
\sum_{k'_{i}, i \in \mathcal{N}_{\Omega} \mid A_{i} , i \in \Omega^c} P( B_{i}, i \in \mathcal{N}_{\Omega} \mid A_{i} , i \in \Omega^c) = e^{n|\mathcal{N}_{\Omega}| n R_r} \prod_{i \in \mathcal{N}_{\Omega}} P\left( \|u_i\|^2 \leq n \sigma_c^2 \right) \leq e^{n|\mathcal{N}_{\Omega}| n \frac{1}{2}(\log(2(1+\epsilon)) + 1 + o_n(1))} \tag{31}
\]
where the last inequality follows from the below lemma.\footnote{An alternative way to upper bound (30) is to randomly choose the quantization lattices at each relay instead of using a fixed lattice.}

**Lemma 5.3:** Let $\nu$ be uniformly distributed over $\mathcal{V}^Q$. We have,

$$
P(\|\nu\|_2^2 \leq n \sigma_c^2) \leq e^{-\frac{n}{2} \left( \log \left( 1 + \frac{\sigma_c^2}{\sigma^2} \right) - 1 - o_n(1) \right)}.
$$

Combining the results of Lemma 5.2 and (31), together with the summation over all possible source-destination cuts in (25), proves the main result of this paper which is stated in Theorem 2.1.

**REFERENCES**


VI. APPENDIX

We first introduce the following two technical lemmas that we use repeatedly in this appendix.

**Lemma 6.1:** (Lemma 11 of [6])

(a) Let $u \sim \text{unif}(B(R))$. Let us denote $\frac{1}{n} \mathbb{E}[\|u\|^2] = \frac{R^2}{n+2} := \sigma^2$. Let $z \sim \mathcal{N}(0, \sigma^2 I_n)$. Then,

$$
f_u(x) \leq f_z(x) e^{n\epsilon_2},
$$

where $\epsilon_2 = \frac{1}{2} \log(2\pi e G_n^* G_n^*) + \frac{1}{n}$.

(b) Let $u \sim \text{unif}(\mathcal{V})$ where $\mathcal{V}$ is the Voronoi region of a lattice $\Lambda$. Note that $\frac{1}{n} \mathbb{E}[\|u\|^2] = \sigma^2(\Lambda)$. Let $z \sim \mathcal{N}(0, \sigma^2 I_n)$ such that

$$
\sigma^2 = \frac{G_n^*}{G(\Lambda)} (\rho_{\text{cov}}(\Lambda))^2 \sigma^2(\Lambda).
$$

Then,

$$
f_u(x) \leq f_z(x) e^{n\epsilon_2(\Lambda)},
$$

where $\epsilon_2(\Lambda) = \log(\rho_{\text{cov}}(\Lambda)) + \frac{1}{2} \log(2\pi e G_n^* G_n^*) + \frac{1}{n}$.
The significance of the above lemma is that it allows to upper bound the probability distribution of a random variable \( u \), either uniformly distributed on an n-dimensional sphere or over the Voronoi region of a Roger’s good lattice, with the probability distribution of a Gaussian vector of identity covariance matrix and of the same variance with \( u \). Note that \( \epsilon_2 \) in part (a) of the lemma goes to zero with increasing dimension \( n \). Similarly in part (b), \( \epsilon_2(\Lambda) \to 0 \) and \( \sigma^2 \to \sigma^2(\Lambda) \) as \( n \) increases if \( \Lambda \) is Roger’s good.

Lemma 6.2: Let \( z_i, i = 1, \ldots, n \) be i.i.d random variables with distribution \( \mathcal{N}(0, \sigma_i^2) \).

\[
P\left( \sum_{i=1}^{n} z_i^2 \leq nc \right) \leq e^{-\left(\sum_{i=1}^{n} \frac{1}{2} \log \left(1 + \frac{c^2}{\sigma_i^2} \right) - \frac{ct}{2}\right)}.
\]

Proof of Lemma 6.2: The proof of the lemma follows by a simple application of the exponential Chebyshev’s inequality. For any \( t > 0 \), we have

\[
P\left( \sum_{i=1}^{n} z_i^2 \leq nc \right) = P\left( e^{-t \sum_{i=1}^{n} z_i^2} \geq e^{-ntc} \right) \leq E[e^{-t \sum_{i=1}^{n} z_i^2}] e^{ntc} = \prod_{i=1}^{n} E[e^{-t z_i^2}] e^{ntc} = \prod_{i=1}^{n} \left( \frac{1}{\sqrt{1 + 2\sigma_i^2 t}} \right) e^{ntc} = e^{-\left(\frac{t}{2} \sum_{i=1}^{n} \log(1 + 2\sigma_i^2 t) - ntc\right)}.
\]

Therefore by choosing \( t = 1/2c \), we have

\[
P\left( \sum_{i=1}^{n} z_i^2 \leq nc \right) \leq e^{-\sup_{t \geq 0} \left(\frac{t}{2} \sum_{i=1}^{n} \log(1 + 2\sigma_i^2 t) - ntc\right)} \leq e^{\left(\frac{t}{2} \sum_{i=1}^{n} \log \left(1 + \frac{c^2}{\sigma_i^2} \right) - \frac{ct}{2}\right)}.
\]

The proof of Lemma 5.3 follows by a straightforward application of the above two lemmas.

Proof of Lemma 5.3: If \( \nu \) is uniformly distributed over \( \mathcal{V}^Q \), by part-(b) of Lemma 6.1 we have

\[
P\left( \|\nu\|^2 \leq n \sigma^2 \right) \leq e^{n \epsilon_2(\Lambda^Q)} P\left( \|\nu\|^2 \leq n \sigma^2 \right)
\]

where \( \nu' \sim \mathcal{N}(0, \sigma^2 \nu I_n) \) with

\[
\sigma^2 = \frac{G_n^*}{G(\Lambda^Q)} \left( \rho_{\text{conv}}(\Lambda^Q) \right)^2 \sigma^2(\Lambda^Q) = (1 + o_n(1))\sigma^2(\Lambda^Q).
\]

Applying Lemma 6.2 for the case of equal variances, yields the result

\[
P\left( \|\nu\|^2 \leq n \sigma^2 \right) \leq e^{-\frac{t}{2} \left( \log \left(1 + \frac{(1 + o_n(1))\sigma^2(\Lambda^Q)}{\sigma^2} \right) - 1\right)},
\]

and therefore

\[
P\left( \|\nu\|^2 \leq n \sigma^2 \right) \leq e^{n \epsilon_2(\Lambda^Q)} e^{-\frac{t}{2} \left( \log \left(1 + \frac{(1 + o_n(1))\sigma^2(\Lambda^Q)}{\sigma^2} \right) - 1\right)}.
\]

To prove Lemmas 5.1 and 5.2, we introduce the following lemma as an intermediate step:

Lemma 6.3: Let \( x_j, x_j', j = 1, \ldots, N_1 \) be independent discrete random variables uniformly distributed over the \( p_r^n \) lattice points \( p_r^{-1} \Lambda T \cap \mathcal{V}^T \). Let \( z_i \) and \( u_i' \), \( i = 1, \ldots, N_2 \) be independent random variables with distributions

May 10, 2010 DRAFT
\( z_i \sim \mathcal{N}(0, \sigma_i^2 I_n) \) and \( u'_i \sim \text{unif}(\mathcal{V}_1^Q) \) where \( \mathcal{V}_1^Q \) denotes the Voronoi region of the lattice \( \Lambda_1^Q \). Let \( S_1, \ldots, S_{N_2} \subseteq \mathbb{R}^n \). Then,

\[
\mathbb{P} \left( \sum_{j=1}^{N_1} h_{ij}(x_j - x'_j) + z_i - u'_i \in S_i, \ \forall i = 1, \ldots, N_2 \right) \\
\leq \left( 1 + \epsilon_4(\Lambda^T) \right)^{N_2} e^{\epsilon_2(\Lambda^T) + \epsilon_2} \left( e^{\epsilon_2(\Lambda_1^Q)} \right)^{N_2} \mathbb{P} \left( \sum_{j=1}^{N_1} h_{ij}(\tilde{x}_j - \tilde{x}'_j) + \tilde{z}_i \in S_i, \ \forall i = 1, \ldots, N_2 \right)
\]

where \( \tilde{x}_j, \tilde{x}'_j, j = 1, \ldots, N_1, \tilde{z}_i, i = 1, \ldots, N_2 \) are all independent Gaussian random variables such that \( \tilde{x}_j, \tilde{x}'_j \sim \mathcal{N}(0, \sigma_i^2 I_n) \) with

\[
\sigma_i^2 = (1 + \epsilon_5)^{2N_1} \left( \sigma^2 + \frac{G^*_n}{G(\Lambda^T)} \right)^{\epsilon_2(\Lambda_1^Q)}
\]

and \( \tilde{z}_i \sim \mathcal{N}(0, \sigma_i^2 I_n) \),

where \( \epsilon_1(\Lambda^T), \epsilon_2, \epsilon_3, \epsilon_4(\Lambda_1^Q), \epsilon_5 \to 0 \) as \( n \to \infty \). Furthermore \( \sigma_i^2 \to \sigma^2(\Lambda^T) \) and \( \sigma_i^2 \to \sigma^2 + \sigma^2(\Lambda_1^Q) \) since both \( \Lambda^T \) and \( \Lambda_1^Q \) are Roger’s good.

**Proof of Lemma 6.3:** First, by using Part-(b) of Lemma 6.1, we can upper bound the probability

\[
\mathbb{P} \left( \sum_{j=1}^{N_1} h_{ij}(x_j - x'_j) + z_i - u'_i \in S_i, \ \forall i = 1, \ldots, N_2 \right)
\]

by

\[
\left( e^{\epsilon_2(\Lambda_1^Q)} \right)^{N_2} \mathbb{P} \left( \sum_{j=1}^{N_1} h_{ij}(x_j - x'_j) + z_{eq,i} \in S_i, \ \forall i = 1, \ldots, N_2 \right),
\]

where \( z_{eq,i} \) are i.i.d with distribution \( \mathcal{N}(0, \sigma_{eq,i}^2 I_n) \),

\[
\sigma_{eq,i}^2 = \sigma^2 + \frac{G^*_n}{G(\Lambda_1^Q)} \left( \rho_{cov}(\Lambda_1^Q) \right)^2 \sigma^2(\Lambda_1^Q).
\]

Since \( \Lambda_1^Q \) is Roger’s good, \( \epsilon_2(\Lambda_1^Q) \) given in the lemma vanishes with increasing \( n \). The probability in (32) can be expressed as,

\[
\mathbb{P} \left( \sum_{j=1}^{N_1} h_{ij}(x_j - x'_j) + z_{eq,i} \in S_i, \ \forall i = 1, \ldots, N_2 \right) = \left( p_r^{-n} \right)^{2N_1} \sum_{\mathbf{x}_1, \ldots, \mathbf{x}_N \in \mathcal{V}} \mathbb{P} \left( \sum_{j=1}^{N_1} h_{ij}(x_j - x'_j) + z_{eq,i} \in S_i, \ \forall i = 1, \ldots, N_2 \right).
\]

The last probability is only over \( z_{eq,i} \)'s and note that the \( x_j \) and \( x'_j \)’s now denote the dummy variables of the summation. Consider one of the summations above of the form,

\[
p_r^{-n} |\mathcal{V}| \sum_{\mathbf{x}_i \in p_r^{-1} \mathcal{A} \cap \mathcal{V}} \mathbb{P} \left( \sum_{j=1}^{N_1} h_{ij}(x_j - x'_j) + z_{eq,i} \in S_i, \ \forall i = 1, \ldots, N_2 \right).
\]
where \( x_1 \) denotes the dummy variable of the summation and \( x_2, \ldots, x_{N_1}, x_1', \ldots, x_{N_1} \) are fixed vectors. We show below that this summation is upper bounded by

\[
(1 + \epsilon_4(A^T))^{N_2} \int_{V^T + p_r^{-1}V^T} dx_1 \mathbb{P} \left( \sum_{j=1}^{N_1} h_{ij}(x_j - x_j') + z_{eq,i} \in S_i, \forall i = 1, \ldots, N_2 \right) \tag{34}
\]

where \( z'_{eq,i} \sim \mathcal{N}(0, (1 + \epsilon_5)\sigma_{eq}^2(I_n)) \) and both \( \epsilon_4(A) \) and \( \epsilon_5 \to 0 \) as \( n \to \infty \). For two sets \( A \subset \mathbb{R}^n \) and \( B \subset \mathbb{R}^n \), the sum set \( A + B \subset \mathbb{R}^n \) denotes \( A + B = \{ a + b : a \in A, b \in B \} \). Applying this upper bound recursively to all the summations in (33) yields

\[
\mathbb{P} \left( \sum_{j=1}^{N_1} h_{ij}(x_j - x_j') + z_{eq,i} \in S_i, \forall i = 1, \ldots, N_2 \right) \leq (1 + \epsilon_4(A^T))^{2N_2N_1} \frac{1}{|V^T|} \int_{V^T + p_r^{-1}V^T} \cdots \int_{V^T + p_r^{-1}V^T} dx_1 \cdots dx_{N_1} dx_1' \cdots dx_{N_1}' \mathbb{P} \left( \sum_{j=1}^{N_1} h_{ij}(x_j - x_j') + \tilde{z}_i \in S_i, \forall i = 1, \ldots, N_2 \right) \tag{35}
\]

where \( \tilde{z}_i \sim \mathcal{N}(0, (1 + \epsilon_5)\sigma_{eq}^2(I_n)) \). \( R_u^T \) in the last inequality denotes the covering radius of \( V^T \) and \( B \left( (1 + p_r^{-1})R_u^T \right) \) denotes an \( n \)-dimensional sphere in \( \mathbb{R}^n \) of radius \( (1 + p_r^{-1})R_u^T \). The last inequality follow follows the fact that \( V^T + p_r^{-1}V^T \subseteq B \left( (1 + p_r^{-1})R_u^T \right) \) which in turn follows from the definition of \( R_u^T \). We can rewrite (35) as

\[
(1 + \epsilon_4(A^T))^{2N_2N_1} \left( e^{\epsilon_1(A^T)} \right)^{N_1} \frac{1}{|V^T|} \int_{B \left( (1 + p_r^{-1})R_u^T \right)} \cdots \int_{B \left( (1 + p_r^{-1})R_u^T \right)} dx_1 \cdots dx_{N_1} dx_1' \cdots dx_{N_1}' \mathbb{P} \left( \sum_{j=1}^{N_1} h_{ij}(x_j - x_j') + \tilde{z}_i \in S_i, \forall i = 1, \ldots, N_2 \right) \tag{36}
\]

where,

\[
\frac{|B \left( (1 + p_r^{-1})R_u^T \right)|}{|V^T|} = \left( \frac{|B \left( (1 + p_r^{-1})R_u^T \right)|}{|B \left( R_u^T \right)|} \right)^n = \left( \frac{(1 + p_r^{-1})R_u^T}{R_u^T} \right)^n = e^{\epsilon_1(A^T)}
\]

and \( \epsilon_1(A^T) = \log(1 + p_r^{-1}) + \log \rho_{cov}(A^T) \). Recall that the effective radius \( R_u^T \) of the lattice \( A^T \) is defined as the radius of a sphere having the same volume as the Voronoi region of \( A^T \). Since \( A^T \) is Roger’s good and \( p_r \to \infty \) as \( n \to \infty \), we have \( \epsilon_1(A^T) \to 0 \). We can upper bound (36) by applying Part-(a) of Lemma 6.1 which gives

\[
\mathbb{P} \left( \sum_{j=1}^{N_1} h_{ij}(x_j - x_j') + z_{eq,i} \in S_i, \forall i = 1, \ldots, N_2 \right) \leq \left( (1 + \epsilon_4(A^T))^{N_2} e^{\epsilon_1(A^T) + \epsilon_2} \right)^{2N_1} \mathbb{P} \left( \sum_{j=1}^{N_1} h_{ij}(\tilde{x}_j - \tilde{x}_j') + \tilde{z}_i \in S_i, \forall i = 1, \ldots, N_2 \right), \tag{37}
\]
where $\mathbf{x}_j, \mathbf{x}'_j, j = 1, \ldots, N_1$ are independent $\mathcal{N}(0, \sigma^2_x I_n)$ with
\[
\sigma^2_x = \frac{(1 + p^{-1}_e R^2_x)}{n + 2}.
\] (38)

Plugging the expression in (13) to (38), yields
\[
\sigma^2_x = \frac{(1 + p^{-1}_e R^2_x)^2}{n + 2} = (1 + p^{-1}_e)^2 \frac{G^2_x}{G(A^T)} \sigma^2(A^T).
\]

The upper bounds (32) and (37) together yield the result stated in the lemma.

It remains to prove (34). We will first show that
\[
p^{-n}_x |Y^T| \mathbb{P} \left( \sum_{j=1}^{N_1} h_{ij}(\mathbf{x}_j - \mathbf{x}'_j) + \mathbf{z}_{eq,i} \in S_i, \forall i = 1, \ldots, N_2 \right)
\leq (1 + \epsilon_4(A^T))^N \int_{p^{-1}_x |Y^T|} ds \mathbb{P} \left( h_{i1}s + \sum_{j=1}^{N_1} h_{ij}(\mathbf{x}_j - \mathbf{x}'_j) + \mathbf{z}'_{eq,i} \in S_i, \forall i = 1, \ldots, N_2 \right)
\] (39)

where $\mathbf{x}_j, \mathbf{x}'_j$‘s are fixed vectors and $\mathbf{z}_{eq,i} \sim \mathcal{N}(0, \sigma^2_x I_n)$, $\mathbf{z}'_{eq,i} \sim \mathcal{N}(0, (1 + \epsilon_5)\sigma^2_x I_n)$ and both $\epsilon_4(A)$ and $\epsilon_5 \to 0$ as $n \to 0$.

First, note that for $\mathbf{z}'_{eq,i} \sim \mathcal{N}(0, \delta^2 I_n)$, $i = 1, \ldots, N_2$,
\[
\mathbb{P} \left( h_{i1}s + \sum_{j=1}^{N_1} h_{ij}(\mathbf{x}_j - \mathbf{x}'_j) + \mathbf{z}'_{eq,i} \in S_i, \forall i = 1, \ldots, N_2 \right)
= \prod_{i=1}^{N_2} \mathbb{P} \left( h_{i1}s + \sum_{j=1}^{N_1} h_{ij}(\mathbf{x}_j - \mathbf{x}'_j) + \mathbf{z}'_{eq,i} \in S_i, \right)
= \prod_{i=1}^{N_2} \int_{S_i} f_{z'_{eq,i}} \left( \mathbf{z}_i - h_{i1}s - \sum_{j=1}^{N_1} h_{ij}(\mathbf{x}_j - \mathbf{x}'_j) \right) d\mathbf{z}_i.
\] (40)

The probability density function $f_{z'_{eq,i}}(\mathbf{c})$ of $\mathbf{z}'_{eq,i}$ depends only on $||\mathbf{c}||$. By the triangle inequality, for any two vectors $\mathbf{a}$ and $\mathbf{b}$, we have
\[
||\mathbf{a} + \mathbf{b}||^2 \leq ||\mathbf{a}||^2 + 2||\mathbf{a}|| ||\mathbf{b}|| + ||\mathbf{b}||^2.
\]

Also for any $t > 0$,
\[
||\mathbf{a}|| ||\mathbf{b}|| \leq \frac{||\mathbf{a}||^2}{t} + t ||\mathbf{b}||^2.
\]

Therefore, for any $t > 0$,
\[
||\mathbf{a} + \mathbf{b}||^2 \leq \left( 1 + \frac{2}{t} \right) ||\mathbf{a}||^2 + (1 + 2t)||\mathbf{b}||^2.
\]

Using this inequality, we obtain
\[
f_{z'_{eq,i}}(\mathbf{a} + \mathbf{b}) \propto e^{-\frac{||\mathbf{a} + \mathbf{b}||^2}{2\sigma^2_x}} \geq e^{-\frac{(1 + \frac{2}{t})||\mathbf{a}||^2}{2\sigma^2_x}} e^{-\frac{(1 + 2t)||\mathbf{b}||^2}{2\sigma^2_x}} \propto f_{z_{eq,i}}(\mathbf{a}) e^{-\frac{(1 + 2t)||\mathbf{b}||^2}{2\sigma^2_x}}
\]
where \( z_{eq,i} \sim \mathcal{N}(0, \sigma^2_{eq} I_n) \) with \( \sigma^2_{eq} = (1 + \frac{1}{\epsilon})^{-1} \delta^2 \). Applying this inequality to (40) with \( a_i = z_i - \sum_{j=1}^{N_i} h_{ij}(x_j - x'_j) \) and \( b_i = h_{i1}s \) yields

\[
\mathbb{P} \left( h_{i1}s + \sum_{j=1}^{N_i} h_{ij}(x_j - x'_j) + z'_{eq,i} \in S_i, \forall i = 1, \ldots, N_2 \right) 
\geq \prod_{i=1}^{N_2} \int_{S_i} e^{-\frac{(x_1^2 + 2x_1s + 2)}{2\sigma^2_{eq}}} f_{z_{eq,i}} \left( z_i - \sum_{j=1}^{N_i} h_{ij}(x_j - x'_j) \right) dz_i,
\]

\[
\geq e^{-\frac{(1+2\epsilon)N\sigma^2_{eq}}} \prod_{i=1}^{N_2} \int_{S_i} e^{-\frac{(x_1^2 + 2x_1s + 2)}{2\sigma^2_{eq}}} f_{z_{eq,i}} \left( z_i - \sum_{j=1}^{N_i} h_{ij}(x_j - x'_j) \right) dz_i,
\]

\[
(41)
\]

where we make use of the inequality

\[
\|h_{i1}s\| = |h_{i1}||s| \leq |h_{i1}| p_{\epsilon}^{-1} R_u^T.
\]

(42)

From (13) for \( R_u^T \) and the choice for \( p_{\epsilon} \) in (16), we know that \( p_{\epsilon}^{-1} R_u^T = O(\sqrt{n}e^{-\frac{nR^2_u}{2\delta^2}}) \rightarrow 0 \) as \( n \rightarrow 0 \). We choose \( t \) such that \( t^{-1} \rightarrow 0 \), while \( t p_{\epsilon}^{-1} R_u^T \rightarrow 0 \). For example, choose \( t = n \). Integrating both sides of the inequality (41) with respect to \( s \) over the region \( p_{\epsilon}^{-1}V^T \), this yields the desired result in (39) where we denote \( 1 + \epsilon_4(\Lambda^T) = e^{\frac{(1+2\epsilon)N\sigma^2_{eq}}{2\sigma^2_{eq}}} p_{\epsilon}^{-1} R_u^T \) and \( 1 + \epsilon_5 = \left(1 + \frac{1}{\epsilon}\right)^2 \).

The conclusion in (34) follows by combining (39) with the following observation,

\[
\sum_{x_i \in p_{\epsilon}^{-1}\Lambda^T \cap V^T} ds \int_{p_{\epsilon}^{-1}V^T} d\mathbf{x} \mathbb{P} \left( h_{i1}(x_1 + s) - h_{i1}x'_1 + \sum_{j=2}^{N_i} h_{ij}(x_j - x'_j) + z_{eq,i} \in S_i, \forall i = 1, \ldots, N_2 \right) 
\leq \int_{V^T + p_{\epsilon}^{-1}V^T} d\mathbf{x} \mathbb{P} \left( h_{i1}x_1 - h_{i1}x'_1 + \sum_{j=2}^{N_i} h_{ij}(x_j - x'_j) + z_{eq,i} \in S_i, \forall i = 1, \ldots, N_2 \right).
\]

This observation simply follows from the fact that the summation and the integration in the first case, together correspond to integrating the function

\[
\mathbb{P} \left( h_{i1}x_1 - h_{i1}x'_1 + \sum_{j=2}^{N_i} h_{ij}(x_j - x'_j) + z_{eq,i} \in S_i, \forall i = 1, \ldots, N_2 \right)
\]

over the sum region \( p_{\epsilon}^{-1}\Lambda^T \cap V^T + p_{\epsilon}^{-1}V^T \) which lies inside the second region \( V^T + p_{\epsilon}^{-1}V^T \). \( \square \)

**Proof of Lemma 5.1:** For a given \( i \in \{N, d\} \) and a set of indices \( k_j, k'_j, j = 1, \ldots, N_i \) we first consider the probability

\[
\mathbb{P} \left( \sum_{j=1}^{N_i} h_{ij}(x^{(k_j)}_j - x^{(k'_j)}_j) + z_i - u'_i \notin V^Q \right),
\]

(43)

where \( N_i \) denotes the number of nodes \( j \) that have non-zero channel coefficients to node \( i \). \( x^{(k_j)}_j \) and \( x^{(k'_j)}_j \) are independent and uniformly distributed over the \( p_{\epsilon}^n \) lattice points \( p_{\epsilon}^{-1}\Lambda^T \cap V^T \), \( z_i \sim \mathcal{N}(0, \sigma^2) \), and \( u'_i \sim \text{unif}(V^Q) \).

Note that we can immediately apply Lemma 6.3 by identifying \( S_1 \) in the lemma as the complement of \( V^Q \), and switch from the discrete distribution over the lattice points \( p_{\epsilon}^{-1}\Lambda^T \cap V^T \) for \( x^{(k_j)}_j \) and \( x^{(k'_j)}_j \) to a Gaussian distribution.
More precisely the above probability is upper bounded by
\[
\left( (1 + \epsilon_4(\Lambda^T)) e^{n\epsilon_2(\Lambda^T) + n\epsilon_2} \right)^{2N_i} e^{n\epsilon_2(\Lambda^T)} \mathbb{P} \left( \sum_{j=1}^{N_i} h_{ij}(\tilde{x}_j - \tilde{x}_j') + \tilde{z}_i \notin \mathbb{V}^Q \right),
\]
where \( \tilde{x}_j, \tilde{x}_j', j = 1, \ldots, N_i \) are independent \( \sim \mathcal{N}(0, \sigma_z^2 I_n) \) with
\[
\sigma_z^2 = (1 + p_r^{-1})^2 (\rho_{\text{cov}}(\Lambda^T))^2 \frac{G_n^*}{G(\Lambda^T)} \sigma^2(\Lambda^T),
\]
and \( \tilde{z}_i \sim \mathcal{N}(0, \sigma_z^2 I_n) \),
\[
\sigma_z^2 = (1 + \epsilon_5)^2 2N_i \left( \sigma^2 + \frac{G_n^*}{G(\Lambda^T)} (\rho_{\text{cov}}(\Lambda^Q))^2 \sigma^2(\Lambda^Q) \right)
\]
where all \( \epsilon_1(\Lambda^T), \epsilon_2, \epsilon_3(\Lambda^Q) \epsilon_4(\Lambda^T), \epsilon_5 \to 0 \) as \( n \to \infty \).

Note that \( \sum_{j=1}^{N_i} h_{ij}(\tilde{x}_j - \tilde{x}_j') + \tilde{z}_i \) has distribution \( \mathcal{N}(0, \sigma_z^2 I_n) \), where
\[
\sigma_z^2 = 2\sigma^2 + 2 \sum_{j=1}^{N_i} |h_{ij}|^2 P + o_n(1),
\]
which follows from our choices for \( \sigma^2(\Lambda^Q) \) and \( \sigma_z^2(\Lambda^T) \) in (17) and (12) respectively. Note that both \( \Lambda^T \) and \( \Lambda^Q \) are Roger’s good and from (16), \( p_r = e^{\frac{\alpha R}{\log n^2}} \) and hence \( p_r^{-1} \to 0 \) as \( n \to 0 \). Since \( \Lambda^Q \) is Poltyrev good, we have
\[
\mathbb{P} \left( \sum_{j=1}^{N_i} h_{ij}(\tilde{x}_j - \tilde{x}_j') + \tilde{z}_i \notin \mathbb{V}^Q \right) \leq e^{-n[E_P(\mu_i) - o_n(1)]}
\]
(44)
where \( E_P(\mu_i) \) is the Poltyrev exponent,
\[
E_P(\mu_i) = \begin{cases} 
\frac{1}{2}[(\mu_i - 1) - \log \mu_i] & 1 < \mu_i \leq 2 \\
\frac{1}{2} \log \frac{\mu_i}{4} & 2 \leq \mu_i \leq 4 \\
\frac{\mu_i}{8} & \mu_i \geq 4 
\end{cases}
\]
and \( \mu_i = \sigma^2(\Lambda^Q)/\sigma_z^2 \). By the union bound, for node \( i \in \{N, d\} \),
\[
\mathbb{P} \left( \exists \{k_j, k_j'\} \text{ s.t. } \sum_{j=1}^{N_i} h_{ij}(x_j^{(k_j)} - x_j^{(k_j')}) + z_i - u_i' \notin \mathbb{V}^Q \right)
\]
\[
\leq \left( (1 + \epsilon_4(\Lambda^T)) e^{n\epsilon_2(\Lambda^T) + n\epsilon_2} \right)^{2N_i} e^{n\epsilon_2(\Lambda^T)} \left( e^{2nR} \right)^{N_i} e^{-n[E_P(\mu_i) - o_n(1)]}
\]
since for every \( j = 1, \ldots, N_i \), \( k_j \) and \( k_j' \) run over the \( e^{nR} \) possible transmit codewords. Finally,
\[
\mathbb{P} \left( \exists i \in \{N, d\}, \{k_j, k_j'\} \text{ s.t. } \sum_j h_{ij}(x_j^{(k_j)} - x_j^{(k_j')}) + z_i - u_i' \notin \mathbb{V}^Q \right)
\]
\[
\leq \left( (1 + \epsilon_4(\Lambda^T)) e^{n\epsilon_2(\Lambda^T) + n\epsilon_2} \right)^{2N_i} e^{n\epsilon_2(\Lambda^T)} (N + 1) e^{-n[E_P(\mu) - 2R, N_s - o_n(1)]}
\]
(46)
where \( N_s = \max_{i \in \{N, d\}} N_i \) and \( \mu = \sigma^2(\Lambda^Q)/\sigma_z^2 \) with
\[
\sigma_z^2 = 2\sigma^2 + 2D_s + o_n(1).
\]
Recall from (14) that \( D_s = \max_{i \in \mathcal{N}, d} \sum_j |h_{ij}|^2 P \). We have chosen in (15) and (16)

\[
R_c = \frac{1}{2} \log \frac{\sigma^2(\Lambda^Q)}{\sigma^2} \quad \text{and} \quad \sigma^2(\Lambda^Q) = 2\eta(\sigma^2 + D_s)
\]

for some \( \eta > 0 \). Therefore \( R_c \) increases logarithmically in \( \eta \) while the Poltyrev exponent is linear in \( \mu \) (and hence in \( \eta \)) in the third regime in (45). By choosing the constant \( \eta \) large enough, we can ensure that the exponent in (46) is negative and hence the probability decreases to zero when \( n \) increases.

\[ \square \]

**Proof of Lemma 5.2:** Let us denote \( N_L = |\Omega| \) and \( N_R = |\Omega^c| \). We want to evaluate the probability

\[
P \left( \left\| \sum_{j \in \Omega} h_{ij}(x_{ij}^{(k_j)}) - x_{ij}^{(k_j)} \right\|^2 + \sum_{i} \left( z_i - u_i \right)^2 \leq n \sigma_e^2, \forall i \in \Omega^c \right).
\]

where \( x_{ij}^{(k_j)} \) and \( x_{ij}^{(k_j)} \), \( j \in \Omega \) are independent and uniformly distributed over the \( p^x \) lattice points \( p^{-1}_r \Lambda^T \cap \mathcal{V}^T \), \( z_i \sim \mathcal{N}(0, \sigma^2) \), and \( u_i \sim \text{unif}(\mathcal{V}^Q) \). We can rewrite the above expression in the form

\[
P \left( \left\| \sum_{j \in \Omega} h_{ij}(x_{ij}^{(k_j)}) - x_{ij}^{(k_j)} \right\|^2 + \sum_{i} \left( z_i - u_i \right)^2 \leq n \sigma_e^2, \forall i \in \Omega^c \right).
\]

with the understanding that \( h_{ij} \) is only non zero if \( i \in \mathcal{N}_l \) and \( j \in \mathcal{N}_{l-1} \) for some \( l = 1, \ldots, l_d \). Note that we can immediately apply Lemma 6.3 by identifying \( S_i \) in the lemma as \( B(\sqrt{n \sigma_e^2}) \), and switch from the discrete distribution over the lattice points \( p^{-1}_r \Lambda^T \cap \mathcal{V}^T \) for \( x_{ij}^{(k_j)} \) and \( x_{ij}^{(k_j)} \), \( j \in \Omega \) to a Gaussian distribution. More precisely, the above probability is upper bounded by

\[
\left((1 + \epsilon_1(\Lambda^T)^N) e^{n \epsilon_2(\Lambda^T)^N} \right)^{2N_L} \left((1 + \epsilon_2(\Lambda^Q)^N) e^{n \epsilon_3(\Lambda^Q)^N} \right)^{2N_R} P \left( \left\| \sum_{j \in \Omega} h_{ij}(\tilde{x}_j - \tilde{x}_j') + \tilde{z}_i \right\|^2 \leq n \sigma^2, \forall i \in \Omega^c \right), \quad (47)
\]

where \( \tilde{x}_j, \tilde{x}_j', j \in \Omega \) are independent \( \sim \mathcal{N}(0, \sigma^2 I_n) \) with

\[
\sigma^2 = (1 + p^{-1}_r)^2 (\rho_{\text{cov}}(\Lambda^T))^2 \frac{G_n}{G(\Lambda^T)} \sigma^2(\Lambda^T),
\]

and \( \tilde{z}_i, i \in \Omega \) are independent \( \sim \mathcal{N}(0, \sigma^2 I_n) \),

\[
\sigma^2 = (1 + \epsilon_5)^2 e^N \left( \sigma^2 + \frac{G_n}{G(\Lambda^Q)} (\rho_{\text{cov}}(\Lambda^Q))^2 \sigma^2(\Lambda^Q) \right)
\]

where all \( \epsilon_1(\Lambda^T), \epsilon_2, \epsilon_2(\Lambda^Q), \epsilon_4(\Lambda^T), \epsilon_5 \rightarrow 0 \) as \( n \rightarrow \infty \). Furthermore \( \sigma^2 \rightarrow P \) and \( \sigma^2 \rightarrow 2\sigma^2 \) as \( n \rightarrow \infty \) since all \( \Lambda^T, \Lambda^Q \) and \( \Lambda^Q \) are Roger’s good.

The probability in (47) can be upper bounded as follows:

\[
P \left( \left\| \sum_{j \in \Omega} h_{ij}(\tilde{x}_j - \tilde{x}_j') + \tilde{z}_i \right\|^2 \leq n \sigma_e^2, \forall i \in \Omega^c \right)
\]

\[
\leq P \left( \left\| H (\tilde{X} - \tilde{X}') + \tilde{Z} \right\|_2^2 \leq N_R n \sigma^2 \right) \quad (48)
\]

\[
= P \left( \left\| \Sigma (\tilde{X} - \tilde{X}') + \tilde{Z} \right\|_2^2 \leq N_R n \sigma^2 \right) \quad (49)
\]

\[
\leq P \left( \sum_{i=1}^{\min(N_R, N_L)} \left\| \sigma_i (\tilde{x}_i - \tilde{x}_i') + \tilde{z}_i \right\|^2 + \sum_{i=1}^{N_R-N_L} \left\| \tilde{z}_i \right\|^2 \leq N_R n \sigma^2 \right), \quad (50)
\]
where $H$ is the $N_R \times N_L$ matrix from the nodes in $\Omega$ to the nodes in $\Omega^c$ and $\Sigma$ is a diagonal matrix containing the singular values $\sigma_i$, $i = 1, \ldots, \min(N_R, N_L)$ of $H$. $\tilde{X}$ and $\tilde{X}'$ are $N_L \times n$ matrices, their $j$'th row containing the vectors $\tilde{x}_j$ and $\tilde{x}'_j$ respectively. $\tilde{Z}$ is $N_R \times n$ matrix, its $i$'th row containing the vector $\tilde{z}_i$. The entries of the matrix $\tilde{X} - \tilde{X}'$ are i.i.d. with distribution $\mathcal{N}(0, 2\sigma_z^2)$ and the entries of the matrix $\tilde{Z}$ are i.i.d. with distribution $\mathcal{N}(0, \sigma^2)$. Inequality (48) follows from the definition of the Frobenius norm for matrices. (49) is obtained by replacing $1$ in the last expression we identify $t$. Choosing $\sigma$ as $n$, we identify $t \to \infty$. Combining everything together yields, 

$$
\mathbb{P} \left( \min(N_R, N_L) \sum_{i=1}^{(N_R-N_L)^+} \|\sigma_i(\tilde{x}_i - \tilde{x}'_i) + \tilde{z}_i\|^2 + \sum_{i=1}^{(N_R-N_L)^+} \|\tilde{z}_i\|^2 \leq N_R n \sigma^2 \right) 
\leq e \cdot \frac{n}{2} \left( \sum_{i=1}^{(N_R-N_L)^+} \log \left( \frac{2\sigma_i^2}{\sigma_c^2} + \frac{\sigma_z^2}{\sigma_c^2} \right) + \sum_{i=1}^{(N_R-N_L)^+} \log \left( \frac{1 + \frac{\sigma_z^2}{\sigma_c^2}}{1 + \epsilon} \right) 
\right) 
$$

Choosing $t = 1/2\sigma^2_c$, yields an exponent

$$
\leq e \cdot \frac{n}{2} \left( \sum_{i=1}^{(N_R-N_L)^+} \log \left( \frac{2\sigma_i^2}{\sigma_c^2} + \frac{\sigma_z^2}{\sigma_c^2} \right) + \sum_{i=1}^{(N_R-N_L)^+} \log \left( \frac{1 + \frac{\sigma_z^2}{\sigma_c^2}}{1 + \epsilon} \right) 
\right) 
$$

in the above expression. We have

$$
\frac{2\sigma_i^2}{\sigma_c^2} + \frac{\sigma_z^2}{\sigma_c^2} \to \frac{\sigma_i^2 P + \sigma_z^2}{(1 + \epsilon)\sigma^2} \to \frac{1}{1 + \epsilon}, 
$$

as $n \to \infty$. Combining everything together yields,

$$
\mathbb{P} \left( \ \left\| \sum_{j \in \Omega \setminus \Omega} h_{ij}(x_j^{(k_j)} - x_j^{(k'_j)}) + z_i - u_i' \right\|^2 \leq n \sigma^2, \ \forall i \in \Omega^c \right) 
\leq e \cdot \left( \sum_{i=1}^{(N_R-N_L)^+} \log \left( \frac{1 + \frac{\sigma_z^2}{\sigma_c^2}}{1 + \epsilon} \right) 
\right) 
$$

In the last expression we identify $\frac{1}{2} \sum_{i=1}^{(N_R-N_L)^+} \log \left( \frac{1 + \frac{\sigma_z^2}{\sigma_c^2}}{1 + \epsilon} \right)$ as $I(X_\Omega; HX_\Omega + Z_{\Omega^c})$, where $X_\Omega$ is an $N_L \times 1$ Gaussian vector with i.i.d entries of variance $P$ and $Z_{\Omega^c}$ is an $N_R \times 1$ Gaussian vector with i.i.d entries of variance $\sigma^2$ and $H$ is the corresponding transfer matrix between nodes in $\Omega$ and $\Omega^c$. 

May 10, 2010 DRAFT