A Simple Proof for the Existence of "Good" Pairs of Nested Lattices

Or Ordentlich
Department of EE-Systems
Tel Aviv University
Ramat Aviv, 69978
Israel
ordent@eng.tau.ac.il

Uri Erez
Department of EE-Systems
Tel Aviv University
Ramat Aviv, 69978
Israel
uri@eng.tau.ac.il

***Abstract*—This paper gives a simplified proof for the existence of nested lattice codebooks that allow to achieve the capacity of the additive white Gaussian noise channel. The proof is self-contained and relies only on basic probabilistic and geometrical arguments. An ensemble of nested lattices which is different than the one used in previous proofs is introduced. This ensemble, in addition to giving rise to a simple proof, can be easily generalized to an ensemble of nested lattices chains. As a result, the proof technique given here easily extends to showing the existence of "good" chains of nested lattices.**

I. Introduction

Linear codes codes have attracted much attention from researchers ever since the first days of Information Theory, as a mean of approaching the capacity of additive discrete memoryless channels (DMC) with a reasonable encoding and decoding complexity. Interestingly, lattice codes, which are their Euclidean counterparts, were first considered for communication over additive white Gaussian noise (AWGN) channels only in the mid 1970s by de Buda [1]. The relatively late interest in lattice codes may be in part due to the power constrained nature of the AWGN channel, which requires an additional shaping procedure, that is not needed for linear codes over DMCs.

Lattice codes are attractive candidates for coding over the AWGN channel due to their structure, which can be exploited by encoding and decoding procedures with reduced complexity compared to non-structured codes. The question as to whether such codes can achieve the capacity of the AWGN channel with lattice encoding and decoding was fully resolved in [2] where a coding scheme based on nested lattice codes in conjunction with Wiener estimation was proposed. Using this scheme, which we describe in detail in Section II-A, it was shown that for the AWGN channel, lattice codes with lattice encoding and decoding can achieve good error exponents. As a corollary, it followed that the AWGN channel's capacity is achievable with lattice encoding and decoding. In [3] it was shown that for rates sufficiently close to capacity (i.e., above the critical rate of the channel) nested lattice codes actually achieve the optimal error exponent.

In the last decade lattice codes were found to play a new role in Information Theory. Following the emergence of wireless communication as a leading technology, the interest in understanding the fundamental limits of Gaussian networks had grown. In such networks, there are multiple transmitters and receivers, where each receiver sees a linear combination of the signals sent by the different transmitters, corrupted by AWGN. Surprisingly, in networks of this type, coding schemes that utilize lattice codes often outperform the best known random coding schemes. Thus, lattice codes are used in order to obtain new achievable rate regions, rather than reducing complexity.

As a consequence, the nested lattices coding scheme of [2] has been extensively used for proving new coding theorems for Gaussian networks. Since for the majority of these networks the capacity is not known, the question of finding the best error exponent is far out of scope. Therefore, when [2] is used in the context of Gaussian networks, it is usually the capacity result that is used and not the error exponent results.

The aim of this paper is to make the capacity result from [2] more accessible. To this end, we derive this result directly, without going through error exponent analysis. This leads to a simplified proof, that is completely self contained and uses only elementary probabilistic and geometrical arguments. A major advantage of the proof technique is that it can be easily extended to more complicated structures of nested lattices.

There are two main differences between the approach taken in this work and previous works. In previous work, a two-step construction was considered. In the first step, assuming a coarse lattice that is good for covering (and hence also for quantization) is given, a dense self-similar nested lattice pair is generated via scaling. In the second step, the fine lattice is diluted. In this paper we simultaneously construct the nested lattices by taking a chain of subcodes of one linear code. This is a direct extension of the construction of nested linear binary codes proposed by Zamir and Shamai in [4]. Another difference from previous works is that rather than requiring that the coarse lattice will be good for covering, we only require that it will be good for quantization

This work was supported in part by the Israel Science Foundation under grant 1557/10 and the Binational Science Foundation under grant 2008455. The work of O. Ordentlich was supported by the Adams Fellowship Program of the Israel Academy of Sciences and Humanities.

II. Preliminaries on Lattice Codes

A lattice Λ is a discrete subgroup of $\mathbb{R}^n$ which is closed under reflection and real addition. Any lattice Λ in $\mathbb{R}^n$ is spanned by some $n \times n$ matrix $\mathbf{F}$ such that

$$\Lambda = \{\mathbf{t} = \mathbf{F}\mathbf{a} : \mathbf{a} \in \mathbb{Z}^n\}.$$

We denote the nearest neighbor quantizer associated with the lattice Λ by

$$Q_\Lambda(\mathbf{x}) \triangleq \arg\min_{\mathbf{t}\in\Lambda} \|\mathbf{x} - \mathbf{t}\|. \tag{1}$$

The basic Voronoi region of Λ, denoted by $\mathcal{V}$, is the set of all points in $\mathbb{R}^n$ which are quantized to the zero vector, where ties in (1) are broken in a systematic manner. The modulo operation returns the quantization error w.r.t. the lattice,

$$[\mathbf{x}] \bmod \Lambda \triangleq \mathbf{x} - Q_\Lambda(\mathbf{x}),$$

and satisfies the distributive law,

$$\big[[\mathbf{x}] \bmod \Lambda + \mathbf{y}\big] \bmod \Lambda = [\mathbf{x} + \mathbf{y}] \bmod \Lambda.$$

Let $V(\Lambda)$ be the volume of a fundamental cell of Λ, e.g., the volume of $\mathcal{V}$, and let $\mathbf{U}$ be a random variable uniformly distributed over $\mathcal{V}$. We define the second moment per dimension associated with Λ as

$$\sigma^2(\Lambda) \triangleq \frac{1}{n}\mathbb{E}\|\mathbf{U}\|^2 = \frac{1}{n}\frac{\int_{\mathcal{V}} \|\mathbf{x}\|^2 d\mathbf{x}}{V(\Lambda)}.$$

The normalized second moment (NSM) of a lattice Λ is defined by

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

Note that this quantity is invariant to scaling of the lattice Λ.

It is often useful to compare the properties of the Voronoi region $\mathcal{V}$ with those of a ball.

Definition 1: Let

$$\mathcal{B}(\mathbf{s}, r) \triangleq \{\mathbf{x} \in \mathbb{R}^n \ : \ \|\mathbf{x} - \mathbf{s}\| \le r\},$$

denote the closed n-dimensional ball with radius r centered at $\mathbf{s}$. We refer to the volume of an n-dimensional ball with unit radius as V_n. In general $V(\mathcal{B}(\mathbf{s}, r)) = V_n r^n$. Note that $nV_n^{2/n} < 2\pi e$ for all n [5], and

$$\lim_{n\to\infty} nV_n^{\frac{2}{n}} = 2\pi e. \tag{2}$$

The ball $\mathcal{B}(\mathbf{0}, r)$ has the smallest second moment per dimension out of all sets in $\mathbb{R}^n$ with volume $V_n r^n$, and it is given by

$$\begin{aligned}\sigma^2(\mathcal{B}(\mathbf{0}, r)) &= \frac{1}{n}\frac{1}{V_n r^n}\int_{\mathbf{x}\in\mathcal{B}(\mathbf{0},r)} \|\mathbf{x}\|^2 d\mathbf{x} \\ &= \frac{1}{n}\frac{1}{V_n r^n}\int_0^r r'^2 d(V_n r'^n) \\ &= \frac{1}{n}\frac{1}{V_n r^n}\frac{nV_n r^{n+2}}{n+2} \\ &= \frac{r^2}{n+2}.\end{aligned} \tag{3}$$

It follows that $\mathcal{B}(\mathbf{0}, r)$ has the smallest possible NSM

$$G(\mathcal{B}(\mathbf{0}, r)) = \frac{\sigma^2(\mathcal{B}(\mathbf{0}, r))}{V(\mathcal{B}(\mathbf{0}, r))} = \frac{1}{n+2}V_n^{-\frac{2}{n}}, \tag{4}$$

which approaches $1/(2\pi e)$ from above as $n \to \infty$.

Thus, the NSM of any lattice in any dimension satisfies $G(\Lambda) \ge 1/(2\pi e)$.

A sequence of lattices $\Lambda^{(n)}$ with growing dimension is called good for mean squared error (MSE) quantization if

$$\lim_{n\to\infty} G\left(\Lambda^{(n)}\right) = \frac{1}{2\pi e}.$$

We define the effective radius $r_{\text{eff}}(\Lambda)$ as the radius of a ball which has the same volume as Λ, i.e.,

$$r_{\text{eff}}^2(\Lambda) \triangleq \frac{V(\Lambda)^{2/n}}{V_n^{2/n}}. \tag{5}$$

Since $\mathcal{B}(\mathbf{0}, r_{\text{eff}}(\Lambda))$ has the smallest second moment of all sets in $\mathbb{R}^n$ with volume $V(\Lambda)$, we have

$$\sigma^2(\mathcal{B}(\mathbf{0}, r_{\text{eff}}(\Lambda))) = \frac{r_{\text{eff}}^2(\Lambda)}{n+2} \le \sigma^2(\Lambda).$$

Thus,

$$r_{\text{eff}}(\Lambda) \le \sqrt{(n+2)\sigma^2(\Lambda)}. \tag{6}$$

Note that for large n we have

$$\frac{r_{\text{eff}}^2(\Lambda)}{n} \approx \frac{V(\Lambda)^{2/n}}{2\pi e}.$$

A lattice Λ_c is said to be nested in Λ_f if $\Lambda_c \subset \Lambda_f$. The lattice Λ_c is referred to as the coarse lattice and Λ_f as the fine lattice. The *nesting ratio* is defined as $(V(\Lambda_c)/V(\Lambda_f))^{1/n}$.

Definition 2: The operation of coset nearest neighbor decoding with respect to the pair of nested lattices $\Lambda_c \subset \Lambda_f$ is defined by

$$g(\mathbf{Y}) = \big[Q_{\Lambda_f}(\mathbf{Y})\big] \bmod \Lambda_c.$$

In words, coset nearest neighbor decoding refers to finding the closest lattice point in Λ_f and identifying its coset leader by reducing modulo Λ_c.

Definition 3: We say that a sequence in n of random noise vectors $\mathbf{Z}^{(n)}$ of length n with (finite) effective variance $\sigma_{\mathbf{Z}}^2 \triangleq \frac{1}{n}\mathbb{E}\|\mathbf{Z}^{(n)}\|^2$, is *semi norm-ergodic* if for any $\epsilon > 0$, $\delta > 0$ and n large enough

$$\Pr\left(\mathbf{Z}^{(n)} \notin \mathcal{B}(\mathbf{0}, \sqrt{(1+\delta)n\sigma_{\mathbf{Z}}^2}\right) \le \epsilon. \tag{7}$$

Note that by the law of large numbers, any i.i.d. noise is semi norm-ergodic. However, even for non i.i.d. noise, the requirement (7) is not very restrictive. In the sequel we omit the dimension index, and denote the sequence $\mathbf{Z}^{(n)}$ simply by $\mathbf{Z}$.

Next, we define "good" pairs of nested lattices. Our definition for the "goodness" of nested lattices pairs is different from the one used in [2].

Definition 4: A sequence of pairs of nested lattices $\Lambda_c^{(n)} \subset \Lambda_f^{(n)}$ is called "good" if:

1) The sequence of coarse lattices $\Lambda_c^{(n)}$ is good for MSE quantization.
2) For any lattice point $\mathbf{t} \in \Lambda_f^{(n)}$, and additive semi norm-ergodic noise $\mathbf{Z}$ with effective variance $\sigma_{\mathbf{Z}}^2 = \frac{1}{n}\mathbb{E}\|\mathbf{Z}\|^2$

$$\lim_{n\to\infty} \Pr\left(\left[Q_{\Lambda_f^{(n)}}(\mathbf{t}+\mathbf{Z})\right] \bmod \Lambda_c^{(n)} \neq \mathbf{t} \bmod \Lambda_c^{(n)}\right) = 0,$$

provided that[1]

$$\lim_{n\to\infty} V(\Lambda_f^{(n)})^{\frac{2}{n}} > 2\pi e\sigma_{\mathbf{Z}}^2.$$

That is, the error probability under coset nearest neighbor decoding in the presence of semi norm-ergodic additive noise $\mathbf{Z}$ vanishes with n if $r_{\text{eff}}(\Lambda_f^{(n)}) > \sqrt{n\sigma_{\mathbf{Z}}^2}$.

A. The mod-Λ Transmission Scheme

Consider the AWGN channel

$$Y = X + N,$$

where $N \sim \mathcal{N}(0,1)$ and X is subject to the average power constraint $\mathbb{E}(X^2) \leq \mathsf{SNR}$. We now briefly recall the nested lattices coding scheme proposed in [2] for communication over this channel.

A pair of nested lattices $\Lambda_c \subset \Lambda_f$ is used in order to construct the codebook $\mathcal{C} = \Lambda_f \cap \mathcal{V}_c$ with rate[2]

$$R = \frac{1}{n}\log\left(\frac{V(\Lambda_c)}{V(\Lambda_f)}\right) = \log(\text{nesting ratio}).$$

Each of the 2^{nR} messages is mapped to a codeword in $\mathcal{C}$. Assume the transmitter wants to send the message w which corresponds to the codeword $\mathbf{t} \in \mathcal{C}$. It transmits

$$\mathbf{X} = [\mathbf{t} - \mathbf{U}] \bmod \Lambda_C,$$

where $\mathbf{U}$ is a random dither that is uniformly distributed over $\mathcal{V}_c$ and is statistically independent of $\mathbf{t}$, and it is common randomness known to both the transmitter and the receiver. Due to the Crypto Lemma [2, Lemma 1], $\mathbf{X}$ is also uniformly distributed over $\mathcal{V}$ and is statistically independent of $\mathbf{t}$. Thus, the average transmission power is $\frac{1}{n}\mathbb{E}\|\mathbf{X}\|^2 = \sigma^2(\Lambda_c)$. In the sequel we assume that Λ_c is scaled such that $\sigma^2(\Lambda_c) = \mathsf{SNR}$, which satisfies the power constraint.

The receiver scales its observation by a factor $\alpha > 0$ to be specified later, adds back the dither $\mathbf{U}$ and reduces the result modulo the coarse lattice

$$\begin{aligned}\mathbf{Y}_{\text{eff}} &= [\alpha\mathbf{Y} + \mathbf{U}] \bmod \Lambda_c \\ &= [\mathbf{X} + \mathbf{U} + (\alpha-1)\mathbf{X} + \alpha\mathbf{N}] \bmod \Lambda_c \\ &= [\mathbf{t} + (\alpha-1)\mathbf{X} + \alpha\mathbf{N}] \bmod \Lambda_c \\ &= [\mathbf{t} + \mathbf{Z}_{\text{eff}}] \bmod \Lambda_c, \end{aligned} \tag{8}$$

where

$$\mathbf{Z}_{\text{eff}} = (\alpha-1)\mathbf{X} + \alpha\mathbf{N} \tag{9}$$

is effective noise statistically independent of $\mathbf{t}$ with effective variance

$$\sigma_{\text{eff}}^2 \triangleq \frac{1}{n}\mathbb{E}\|\mathbf{Z}_{\text{eff}}\|^2 = \alpha^2 + (1-\alpha)^2\mathsf{SNR}. \tag{10}$$

Thus, using nested lattice codes, the original AWGN channel is transformed to an effective modulo-additive channel whose output is a lattice point from Λ_f corrupted by effective noise statistically independent of it.

It was shown in [2], [6] that there exist sequences of pairs of nested lattices $\Lambda_c^{(n)} \subset \Lambda_f^{(n)}$ that achieve a good error exponent with coset nearest neighbor decoding using this transformation. As a corollary, it followed that such pairs of nested lattices achieve any rate below the capacity of the AWGN channel $C = \frac{1}{2}\log(1+\mathsf{SNR})$.

Here, we prove a weaker result: we only prove the existence of sequences of pairs $\Lambda_c^{(n)} \subset \Lambda_f^{(n)}$ that allow to achieve vanishing error probability for any rate below the capacity of the AWGN channel with coset nearest neighbor decoding. Namely, we show that there exist sequences of pairs that are "good" according to Definition 4, and that the "goodness" conditions from Definition 4 suffice to achieve any rate below capacity. Although these "goodness" conditions do not suffice for achieving good error exponents, using them instead of those consider in [2], gives rise to simpler proofs, which is the aim of this paper.

Note that since the coarse lattice is normalized such that $\sigma^2(\Lambda_c^{(n)}) = \mathsf{SNR}$, if $\Lambda_c^{(n)}$ is good for MSE quantization than $\frac{1}{n}\log V\left(\Lambda_c^{(n)}\right) \to \frac{1}{2}\log(2\pi e\mathsf{SNR})$ as $n \to \infty$. In the next section we show that for any lattice density in the range

$$\frac{1}{2}\log\left(2\pi e\sigma_{\text{eff}}^2\right) < \frac{1}{n}\log V(\Lambda_f^{(n)}) < \frac{1}{2}\log(2\pi e\mathsf{SNR})$$

there exists a sequence of "good" pairs of nested lattices. We also show that for such sequences of pairs, $\mathbf{Z}_{\text{eff}}$ is semi norm-ergodic. Therefore, any rate satisfying

$$\begin{aligned} R &< \lim_{n\to\infty}\left[\frac{1}{n}\log V(\Lambda_c^{(n)}) - \frac{1}{n}\log V(\Lambda_f^{(n)})\right] \\ &< \frac{1}{2}\log\left(2\pi e\mathsf{SNR}\right) - \frac{1}{2}\log\left(2\pi e\sigma_{\text{eff}}^2\right) \\ &= \frac{1}{2}\log\left(\frac{\mathsf{SNR}}{\sigma_{\text{eff}}^2}\right). \end{aligned} \tag{11}$$

is achievable using nested lattice codes. Setting $\alpha = \mathsf{SNR}/(1+\mathsf{SNR})$ in (10), which is the linear minimum MSE coefficient for estimating $\mathbf{X}$ from $\mathbf{Y}$, we see that any rate below $C = \frac{1}{2}\log(1+\mathsf{SNR})$ is achievable with nested lattice codes and coset nearest neighbor decoding.

III. Existence Proof

Previous proofs for the existence of capacity achieving pairs of nested lattices used random Construction A, introduced by Loeliger [7], for creating a fine lattice, and then rotated it using

[1] In [2] the volume-to-noise ratio (VNR) was defined as

$$\mu = V(\Lambda_f^{(n)})^{\frac{2}{n}}/2\pi e\sigma_{\mathbf{Z}}^2.$$

Thus, the condition $V(\Lambda_f^{(n)})^{\frac{2}{n}} > 2\pi e\sigma_{\mathbf{Z}}^2$ is equivalent to VNR > 1.

[2] All logarithms in this paper are to the base 2, and therefore all rates are expressed in bits per (real) channel use.

a lattice that is good for MSE quantization.[3] Here, we take a slightly different approach that is a direct extension of the original approach of [4] to creating nested binary linear codes. We use random Construction A to simultaneously create both the fine and the coarse lattice. Namely, we randomly draw a linear code and lift it to the Euclidean space in order to obtain the fine lattice. The coarse lattice is obtained by lifting a subcode from the same linear code to the Euclidean space. The ensemble of nested lattice pairs is defined as follows.

Let $\gamma = 2\sqrt{n\mathsf{SNR}}$, and $p = \xi n^{\frac{3}{2}}$, where ξ is chosen as the largest number in the interval $[1/2, 1)$ such that p is prime. For some $\epsilon_1 > 0$, to be specified later, choose k_1 such that[4]

$$\frac{k_1}{n}\log p = \frac{1}{2}\log\left(\frac{4}{V_n^{2/n}}\right) + \epsilon_1. \tag{12}$$

1) Draw a generating $k \times n$ ($k > k_1$), matrix $\mathbf{G}_f$ whose entries are i.i.d. and uniform over the elements of the prime field $\mathbb{Z}_p = \{0, 1, \cdots, p-1\}$. Let $\mathbf{G}_c$ be a $k_1 \times n$ matrix whose rows are the first k_1 rows of $\mathbf{G}_f$.
2) Define the discrete codebooks

$$\mathcal{C}_c = \left\{\mathbf{x} \in \mathbb{Z}_p^n \ : \ \mathbf{x} = [\mathbf{w}^T\mathbf{G}_c] \bmod p \quad \mathbf{w} \in \mathbb{Z}_p^{k_1}\right\}$$

and

$$\mathcal{C}_f = \left\{\mathbf{x} \in \mathbb{Z}_p^n \ : \ \mathbf{x} = [\mathbf{w}^T\mathbf{G}_f] \bmod p \quad \mathbf{w} \in \mathbb{Z}_p^{k}\right\}.$$

3) Apply Construction A to lift $\mathcal{C}_c$ and $\mathcal{C}_f$ to $\mathbb{R}^n$ and form the lattices

$$\Lambda_c = \gamma p^{-1}\mathcal{C}_c + \gamma\mathbb{Z}^n, \quad \Lambda_f = \gamma p^{-1}\mathcal{C}_f + \gamma\mathbb{Z}^n.$$

Remark 1: We have chosen to specify our ensemble in terms of the linear codes' generating matrices

$$\mathbf{G}_f = \begin{bmatrix} \mathbf{G}_c \\ - - - \\ \mathbf{G}' \end{bmatrix}.$$

We could have equally defined the ensemble using the linear codes' parity check matrices

$$\mathbf{H}_c = \begin{bmatrix} \mathbf{H}_f \\ - - - \\ \mathbf{H}' \end{bmatrix},$$

as done in [4], [8] for ensembles of nested binary linear codes.

Clearly, in this construction $\Lambda_c \subset \Lambda_f$. Moreover, if the linear code's generating matrix $\mathbf{G}_f$ is full-rank over $\mathbb{Z}_p$, which by construction implies that $\mathbf{G}_c$ is also full-rank, we have $V(\Lambda_c) = \gamma^n p^{-k_1}$ and $V(\Lambda_f) = \gamma^n p^{-k}$. Thus, the nesting ratio is $p^{\frac{k-k_1}{n}}$ and the rate of the induced codebook $\mathcal{C} = \Lambda_f \cap \mathcal{V}_c$ is therefore $R = \frac{k-k_1}{n}\log p$. The probability that $\mathbf{G}_f$ is not full-rank can be bounded [6] by

$$\Pr\left(\mathrm{rank}(\mathbf{G}_f) < k\right) < p^{k-n},$$

which vanishes as long as $k < \beta n$ for some $0 < \beta < 1$. We take k to satisfy this restriction, which ensures that $\mathbf{G}_c$ is full-rank with high probability. Thus, in the sequel we will assume that $\mathbf{G}_c$ is indeed full-rank. Otherwise, the obtained nested lattices pair is regarded as "bad".

Next, we show that almost all members of the ensemble satisfy the first "goodness" condition from Definition 4, and that almost all members of the ensemble satisfy the second condition of Definition 4. It follows that almost all pairs of nested lattices in the ensemble satisfy both conditions simultaneously, and are therefore "good". We then show that for coarse lattices Λ_c which are good for MSE quantization, a random dither uniformly distributed over the Voronoi region is semi norm-ergodic. This in turn implies that for such lattices, the effective noise $\mathbf{Z}_{\text{eff}}$ from the mod-Λ transmission scheme is also semi-spherical. Finally, we conclude that with "good" pairs of nested lattices, the capacity of the AWGN channel can be achieved using the mod-Λ transmission scheme.

Before going into the proofs we need to introduce some more notation. Let $\mathrm{CUBE} \triangleq [-1/2, 1/2)^n$ denote the unit cube centered at zero. Also, denote the operation of reducing each component of $\mathbf{x} \in \mathbb{R}^n$ modulo γ by $\mathbf{x}^* \triangleq [\mathbf{x}] \bmod \gamma\mathbb{Z}^n$. If $\mathcal{S}$ is a set of points in $\mathbb{R}^n$, $\mathcal{S}^*$ is the set obtained by reducing all points in $\mathcal{S}$ modulo $\gamma\mathbb{Z}^n$. If $\mathcal{S}$ and $\mathcal{T}$ are sets, $\mathcal{S} + \mathcal{T}$ is their Minkowski sum. In the sequel, we use the following lemma, which follows from simple geometric arguments and is illustrated in Figure 1.

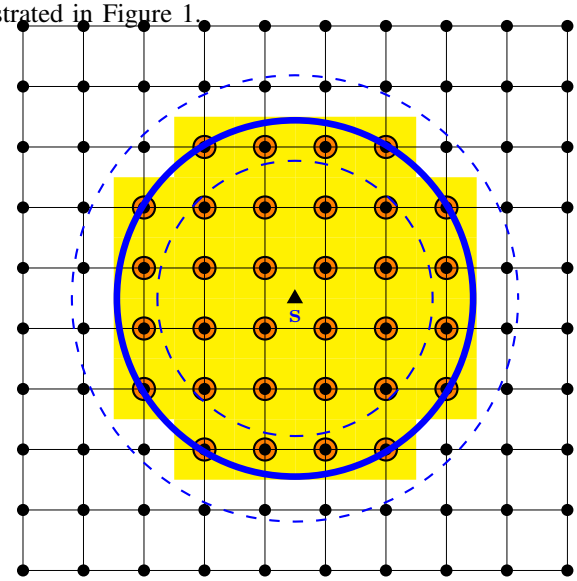

Fig. 1. An illustration of Lemma 1. The solid circle is the boundary of $\mathcal{B}(\mathbf{s}, r)$, and the points inside the small bright circles are the members of the set $\mathbb{Z}^n \cap \mathcal{B}(\mathbf{s}, r)$. The set $= \mathbb{Z}^n \cap \mathcal{B}(\mathbf{s}, r) + \mathrm{CUBE}$ is the shaded area, and as the lemma indicates, it contains $\mathcal{B}(\mathbf{s}, r - \frac{\sqrt{n}}{2})$ and is contained in $\mathcal{B}(\mathbf{s}, r + \frac{\sqrt{n}}{2})$, whose boundaries are plotted in dashed circles.

Lemma 1: For any $\mathbf{s} \in \mathbb{R}^n$ and $r > 0$ the number of points of $\mathbb{Z}^n$ inside $\mathcal{B}(\mathbf{s}, r)$ can be bounded as

$$\left(\max\left\{r - \frac{\sqrt{n}}{2}, 0\right\}\right)^n V_n \le |\mathbb{Z}^n \cap \mathcal{B}(\mathbf{s}, r)| \le \left(r + \frac{\sqrt{n}}{2}\right)^n V_n$$

[3] More precisely, the coarse lattice in the previous proofs was Rogers good, which also implies goodness for MSE quantization.

[4] For ease of exposition we disregard the fact that by this definition k_1 may not take an integer value. We can always round k_1 to the nearest larger integer with no effect on the derivations in the sequel.

Proof: Let $\mathcal{S} \triangleq (\mathbb{Z}^n \cap \mathcal{B}(\mathbf{s}, r)) + \text{CUBE}$, and note that $|\mathbb{Z}^n \cap \mathcal{B}(\mathbf{s}, r)| = \text{Vol}(\mathcal{S})$. We have

$$\mathcal{B}\left(\mathbf{s}, r - \frac{\sqrt{n}}{2}\right) \subseteq \mathcal{S}. \tag{13}$$

To see this, note that any $\mathbf{x} \in \mathcal{B}\left(\mathbf{s}, r - \frac{\sqrt{n}}{2}\right)$ lies inside $\mathbf{a} + \text{CUBE}$ for some $\mathbf{a} \in \mathbb{Z}^n$, and for this $\mathbf{a}$ the inequality $\|\mathbf{a} - \mathbf{x}\| \leq \sqrt{n}/2$ holds. Applying the triangle inequality gives

$$\|\mathbf{a} - \mathbf{s}\| = \|(\mathbf{a} - \mathbf{x}) + (\mathbf{x} - \mathbf{s})\| \leq \|(\mathbf{a} - \mathbf{x})\| + \|(\mathbf{x} - \mathbf{s})\| \leq r,$$

which implies (13). On the other hand,

$$\mathcal{S} \subseteq \mathcal{B}(\mathbf{s}, r) + \text{CUBE} \subseteq \mathcal{B}(\mathbf{s}, r) + \mathcal{B}(0, \frac{\sqrt{n}}{2}) = \mathcal{B}(\mathbf{s}, r + \frac{\sqrt{n}}{2}).$$

Thus,

$$\text{Vol}\left(\mathcal{B}\left(\mathbf{s}, r - \frac{\sqrt{n}}{2}\right)\right) \leq \text{Vol}\left(\mathcal{S}\right) \leq \text{Vol}\left(\mathcal{B}\left(\mathbf{s}, r + \frac{\sqrt{n}}{2}\right)\right).$$

■

A. Goodness for MSE Quantization

In this subsection we prove the following lemma, which follows from rate-distortion considerations.

Lemma 2: For any $\delta_1 > 0$ and n large enough, almost all lattices Λ_c in the defined Construction A ensemble satisfy

$$G(\Lambda_c) < (1 + \delta_1)\frac{1}{2\pi e}.$$

Proof: We begin by bounding the average MSE distortion the ensemble achieves, and then we connect it to the normalized second moment. For any $\mathbf{x} \in \mathbb{R}^n$, define

$$\begin{aligned} d(\mathbf{x}, \Lambda_c) &\triangleq \frac{1}{n} \min_{\lambda \in \Lambda_c} \|\mathbf{x} - \lambda\|^2 \\ &= \frac{1}{n} \min_{\mathbf{a} \in \mathbb{Z}^n, \mathbf{c} \in \mathcal{C}_c} \|\mathbf{x} - \gamma p^{-1}\mathbf{c} - \gamma \mathbf{a}\|^2 \\ &= \frac{1}{n} \min_{\mathbf{c} \in \mathcal{C}_c} \|(\mathbf{x} - \gamma p^{-1}\mathbf{c})^*\|^2. \end{aligned}$$

Note that $d(\mathbf{x}, \Lambda_c) \leq \gamma^2/4$ for any $\mathbf{x} \in \mathbb{R}^n$, regardless of $\mathcal{C}_c$.

For any $\mathbf{w} \in \mathbb{Z}_p^{k_1} \setminus \mathbf{0}$, define the random vector $\mathbf{C}(\mathbf{w}) = \left[\mathbf{w}^T \mathbf{G}_c\right] \bmod p$, and note that $\mathbf{C}(\mathbf{w})$ is uniformly distributed over $\mathbb{Z}_p^n$. For all $\mathbf{w} \in \mathbb{Z}_p^{k_1} \setminus \mathbf{0}$ and $\mathbf{x} \in \mathbb{R}^n$, we have

$$\begin{aligned} \varepsilon &\triangleq \Pr\left(\frac{1}{n}\| \left(\mathbf{x} - \gamma p^{-1}\mathbf{C}(\mathbf{w})\right)^* \|^2 \leq \mathsf{SNR}\right) \\ &= p^{-n} \left|(\gamma p^{-1}\mathbb{Z}_p^n) \bigcap \mathcal{B}^*(\mathbf{x}, \sqrt{n\mathsf{SNR}})\right| \\ &= p^{-n} \left|(\gamma p^{-1}\mathbb{Z}^n) \bigcap \mathcal{B}(\mathbf{x}, \sqrt{n\mathsf{SNR}})\right| \qquad (14) \\ &\geq p^{-n} V_n \left(p\gamma^{-1}\sqrt{n\mathsf{SNR}} - \frac{\sqrt{n}}{2}\right)^n \qquad (15) \\ &= V_n(\gamma^{-2} n\mathsf{SNR})^{\frac{n}{2}} \left(1 - \frac{\gamma}{2p\sqrt{\mathsf{SNR}}}\right)^n \\ &= V_n \left(\frac{1}{4}\right)^{\frac{n}{2}} \left(1 - \frac{\sqrt{n}}{p}\right)^n, \qquad (16) \end{aligned}$$

where (14) follows since $\gamma = 2\sqrt{n\mathsf{SNR}}$, and hence, for any two distinct points $\mathbf{b}_1, \mathbf{b}_2 \in \mathcal{B}(\mathbf{x}, \sqrt{n\mathsf{SNR}})$ we have $\mathbf{b}_1^* \neq \mathbf{b}_2^*$ (that is, the ball $\mathcal{B}(\mathbf{x}, \sqrt{n\mathsf{SNR}})$ is contained in a cube with side γ), and (15) follows from Lemma 1. Substituting $p = \xi n^{\frac{3}{2}}$ gives

$$\begin{aligned} \varepsilon &> V_n 2^{-n} \left(1 - \frac{1}{\xi n}\right)^n \\ &> V_n 2^{-n} \left(1 - \frac{2}{n}\right)^n \\ &> \frac{1}{n^2} V_n 2^{-n}, \end{aligned} \tag{17}$$

where (17) holds for all $n > 4$. Let $M \triangleq p^{k_1} - 1$. Label each of the vectors $\mathbf{w} \in \mathbb{Z}_p^{k_1} \setminus \mathbf{0}$ by an index $i = 1, \ldots, M$, and refer to its corresponding codeword as $\mathbf{C}_i$. Define the indicator random variable related to the point $\mathbf{x} \in \mathbb{R}^n$

$$\chi_i = \begin{cases} 1 & \text{if } \frac{1}{n}\| \left(\mathbf{x} - \gamma p^{-1}\mathbf{C}_i\right)^* \|^2 \leq \mathsf{SNR} \\ 0 & \text{otherwise} \end{cases}.$$

Since each χ_i occurs with probability ε, we have

$$\begin{aligned} \Pr\left(\sum_{i=1}^M \chi_i = 0\right) &= \Pr\left(\frac{1}{M}\sum_{i=1}^M \chi_i - \varepsilon = -\varepsilon\right) \\ &\leq \Pr\left(\left|\frac{1}{M}\sum_{i=1}^M \chi_i - \varepsilon\right| \geq \varepsilon\right) \\ &\leq \frac{\text{Var}\left(\frac{1}{M}\sum_{i=1}^M \chi_i\right)}{\varepsilon^2}, \end{aligned} \tag{18}$$

where the last inequality follows from Chebyshev's inequality. In order to further bound the variance term from (18), we note that $\mathbf{C}(\mathbf{w}_1)$ and $\mathbf{C}(\mathbf{w}_2)$ are statistically independent unless $\mathbf{w}_1 = [a\mathbf{w}_2] \bmod p$ for some $a \in \mathbb{Z}_p$. Therefore, each χ_i is statistically independent of all but p different χ_j's. Thus,

$$\begin{aligned} \text{Var}\left(\frac{1}{M}\sum_{i=1}^M \chi_i\right) &= \frac{1}{M^2}\sum_{i=1}^M \sum_{j=1}^M \text{Cov}(\chi_i, \chi_j) \\ &\leq \frac{Mp\varepsilon}{M^2}. \end{aligned}$$

Substituting into (18) and using (17), we see that for any $\mathbf{x} \in \mathbb{R}^n$

$$\begin{aligned} \Pr\left((d(\mathbf{x}, \Lambda_c) > \mathsf{SNR})\right) &= \Pr\left(\sum_{i=1}^M \chi_i = 0\right) \\ &< \frac{p}{M\varepsilon} \\ &< n^{\frac{3}{2}} \frac{1}{p^{k_1} - 1} n^2 2^n V_n^{-1} \\ &< 2n^{\frac{7}{2}} p^{-k_1} 2^n V_n^{-1}. \end{aligned} \tag{19}$$

It follows that for any distribution on $\mathbf{X}$ we have

$$\begin{aligned}
&\mathbb{E}_{\mathbf{X},\Lambda_c}\left(d(\mathbf{X},\Lambda_c)\right)\\
&\quad\le \mathsf{SNR}\Pr\left(d(\mathbf{X},\Lambda_c)\le \mathsf{SNR}\right)+\frac{\gamma^2}{4}\Pr\left(d(\mathbf{X},\Lambda_c)>\mathsf{SNR}\right)\\
&\quad\le \mathsf{SNR}\left(1+2n^{\frac{9}{2}}p^{-k_1}V_n^{-1}2^n\right)\\
&\quad= \mathsf{SNR}\left(1+2n^{\frac{9}{2}}2^{-n\left[\frac{k_1}{n}\log p-\frac{1}{2}\log\left(\frac{4}{V_n^{2/n}}\right)\right]}\right).
\end{aligned}$$

Our choice of k_1, as given in (12), ensures that the upper bound on the distortion averaged over $\mathbf{X}$ and over the ensemble of coarse lattices Λ_c becomes arbitrary close to SNR as n increases. Since this is true for all distributions on $\mathbf{X}$, we may take $\mathbf{X}\sim \mathrm{Unif}\left(\gamma[0,1)^n\right)$. Let $\mathbf{U}$ be a random variable uniformly distributed over the Voronoi region $\mathcal{V}_c$ of a lattice Λ_c randomly drawn from the ensemble. By construction, for any lattice Λ_c in the defined ensemble $\left[\gamma p^{-1}\mathcal{C}_c+\mathcal{V}_c\right]^*=\gamma[0,1)^n$. Moreover, reducing the set $\gamma p^{-1}\mathcal{C}_c+\mathcal{V}_c$ modulo $\gamma\mathbb{Z}^n$ does not change its volume. Therefore

$$\mathbb{E}_{\Lambda_c}\left(\sigma^2(\Lambda_c)\right)=\mathbb{E}_{\mathbf{U},\Lambda_c}\left(\frac{1}{n}\|\mathbf{U}\|^2\right)=\mathbb{E}_{\mathbf{X},\Lambda_c}\left(d(\mathbf{X},\Lambda_c)\right).$$

It follows that

$$\lim_{n\to\infty}\mathbb{E}_{\Lambda_c}\left(\sigma^2(\Lambda_c)\right)\le \mathsf{SNR}.$$

The normalized volume of a fundamental cell of Λ_c is lower bounded by

$$\begin{aligned}
V(\Lambda_c)^{\frac{2}{n}}&\ge\left(\gamma^n p^{-k_1}\right)^{\frac{2}{n}}\\
&=4n\mathsf{SNR}p^{-2\frac{k_1}{n}}\\
&=2^{-\epsilon_1}nV_n^{\frac{2}{n}}\mathsf{SNR},
\end{aligned}\tag{20}$$

with equality if and only if $\mathbf{G}_c$ is full-rank. This implies that

$$\begin{aligned}
\lim_{n\to\infty}\mathbb{E}_{\Lambda_c}\left(G(\Lambda_c^{(n)})\right)&=\lim_{n\to\infty}\mathbb{E}_{\Lambda_c}\left(\frac{\sigma^2(\Lambda_c)}{V(\Lambda_c)}\right)\\
&\le\lim_{n\to\infty}\frac{\mathbb{E}_{\Lambda_c}\left(\sigma^2(\Lambda_c)\right)}{2^{-\epsilon_1}nV_n^{\frac{2}{n}}\mathsf{SNR}}\\
&=2^{\epsilon_1}\lim_{n\to\infty}\frac{1}{n}V_n^{-\frac{2}{n}}\\
&=2^{\epsilon_1}\frac{1}{2\pi e}.
\end{aligned}\tag{21}$$

The normalized second moment of any set of points in $\mathbb{R}^n$ cannot be smaller than that of an n-dimensional ball, which approaches $1/(2\pi e)$ as n increases. Therefore, for all lattices in the ensemble $G(\Lambda_c)\ge 1/2\pi e$. Applying Markov's inequality for the non-negative random variable $T(\Lambda_c)=G(\Lambda_c)-1/(2\pi e)$ shows that for any $m>1$ at least a fraction of $(m-1)/m$ of the sequences of lattices in the ensemble satisfy

$$\lim_{n\to\infty}G(\Lambda_c^{(n)})\le\left(1+m(2^{\epsilon_1}-1)\right)\frac{1}{2\pi e}.$$

Choosing $\epsilon_1=\log(1+\delta_1/(2m))$ gives

$$\lim_{n\to\infty}G(\Lambda_c^{(n)})\le(1+\delta_1/2)\frac{1}{2\pi e},$$

for at least a fraction of $(m-1)/m$ of the sequences of lattices in the ensemble. Taking m as large as desired establishes the lemma.

Note that for $m>2\log(e)$ we have $\epsilon_1<\delta_1$. We will use this fact in Subsection III-D. ■

B. Goodness for Coding Under Nearest Neighbor Coset Decoding

The next lemma shows that most nested lattice pairs in the ensemble achieve arbitrary low error probabilities in the presence of additive semi norm-ergodic noise as long as their point density is not too high.

Lemma 3: Let $\mathbf{Z}$ be an additive semi norm-ergodic noise with effective variance $\sigma_{\mathbf{Z}}^2=\frac{1}{n}\mathbb{E}\|\mathbf{Z}\|^2$. For any $\mathbf{t}\in\Lambda_f$, $\epsilon>0$, $\delta>0$ and n large enough, for almost all lattices in the ensemble, the error probability under coset nearest neighbor decoding is bounded by

$$P_e\triangleq\Pr\left(\left[Q_{\Lambda_f}(\mathbf{t}+\mathbf{Z})\right]\bmod\Lambda_c\ne\mathbf{t}\bmod\Lambda_c\right)\le\varepsilon,$$

provided that

$$\frac{V(\Lambda_f)^{2/n}}{2\pi e\sigma_{\mathbf{Z}}^2}>1+\delta.$$

An error event under coset nearest neighbor decoding has two possible sources. Either the additive noise $\mathbf{Z}$ is too large, which rarely happens since it is semi norm-ergodic, or there are competing codewords which are "too close" to the transmitted codeword $\mathbf{t}$. Without loss of generality, we may assume that the zero codeword was transmitted. Before proving Lemma 3, we would first like to upper bound the probability that a competing codeword (i.e., a codeword that does not belong to the coset of the zero codeword Λ_c) falls inside a ball centered at the origin. Rather than computing this probability directly, we upper bound the probability of a larger event, namely the event that a lattice point from $\Lambda_f\setminus\gamma\mathbb{Z}^n$ (rather than $\Lambda_f\setminus\Lambda_c$) falls inside a ball centered at zero. This bound is given in the following proposition, whose proof follows from straightforward volume considerations.

Proposition 1: Let $\mathbf{Z}$ be some n-dimensional random vector with effective variance $\sigma_{\mathbf{Z}}^2=\frac{1}{n}\mathbb{E}\|\mathbf{Z}\|^2$ and let $r_{\mathbf{Z}}=\sqrt{(1+\rho)n\sigma_{\mathbf{Z}}^2}$ for some $\rho>0$. Then for any $\epsilon>0$, $\rho>0$ and n large enough, for almost all lattices in the ensemble

$$\Pr\left((\Lambda_f\setminus\gamma\mathbb{Z}^n)\bigcap\mathcal{B}(\mathbf{Z},r_{\mathbf{Z}})\ne\emptyset\right)<\epsilon,$$

provided that

$$\frac{V(\Lambda_f)^{2/n}}{2\pi e(1+\rho)\sigma_{\mathbf{Z}}^2}>1+\rho.$$

Proof of Proposition 1: We show that the average probability over the ensemble vanishes with n, and therefore

for almost all members of the ensemble this probability is small. Let $\mathbb{1}(\mathcal{A})$ be the indicator function of the event $\mathcal{A}$.

$$\begin{aligned}
&\mathbb{E}_{\Lambda_f}\left(\Pr\left((\Lambda_f \setminus \gamma\mathbb{Z}^n)\bigcap \mathcal{B}(\mathbf{Z}, r_{\mathbf{Z}}) \neq \emptyset\right)\right) \\
&= \mathbb{E}_{\Lambda_f}\mathbb{E}_{\mathbf{Z}}\left(\mathbb{1}\left((\gamma p^{-1}\mathcal{C}_f \setminus \mathbf{0})\bigcap \mathcal{B}^*(\mathbf{Z}, r_{\mathbf{Z}}) \neq \emptyset\right) \;\middle|\; \Lambda_f\right) \\
&= \mathbb{E}_{\mathbf{Z}}\mathbb{E}_{\Lambda_f}\left(\mathbb{1}\left((\gamma p^{-1}\mathcal{C}_f \setminus \mathbf{0})\bigcap \mathcal{B}^*(\mathbf{Z}, r_{\mathbf{Z}}) \neq \emptyset\right) \;\middle|\; \mathbf{Z}\right) \\
&= \mathbb{E}_{\mathbf{Z}}\Pr\left((\gamma p^{-1}\mathcal{C}_f \setminus \mathbf{0})\bigcap \mathcal{B}^*(\mathbf{Z}, r_{\mathbf{Z}}) \neq \emptyset \;\middle|\; \mathbf{Z}\right). \qquad (22)
\end{aligned}$$

Since each codeword in $\mathcal{C}_f \setminus \mathbf{0}$ is uniformly distributed over $\mathbb{Z}_p^n$, and there are less than p^k such codewords (i.e., $p^k - 1$), applying the union bound gives

$$\begin{aligned}
&\mathbb{E}_{\Lambda_f}\left(\Pr\left((\Lambda_f \setminus \gamma\mathbb{Z}^n)\bigcap \mathcal{B}(\mathbf{Z}, r_{\mathbf{Z}}) \neq \emptyset\right)\right) \\
&\leq \mathbb{E}_{\mathbf{Z}}\left(p^{k-n}\cdot\left|\gamma p^{-1}\mathbb{Z}_p^n \bigcap \mathcal{B}^*(\mathbf{Z}, r_{\mathbf{Z}})\right| \;\middle|\; \mathbf{Z}\right) \\
&\leq \mathbb{E}_{\mathbf{Z}}\left(p^{k-n}\cdot\left|\gamma p^{-1}\mathbb{Z}^n \bigcap \mathcal{B}(\mathbf{Z}, r_{\mathbf{Z}})\right| \;\middle|\; \mathbf{Z}\right) \\
&\leq p^{k-n}V_n\left(\frac{p}{\gamma}r_{\mathbf{Z}} + \frac{\sqrt{n}}{2}\right)^n \qquad (23) \\
&= p^k\gamma^{-n}V_n r_{\mathbf{Z}}^n\left(1 + \frac{\gamma\sqrt{n}}{2p\; r_{\mathbf{Z}}}\right)^n \\
&\leq p^k\gamma^{-n}V_n r_{\mathbf{Z}}^n\left(1 + \frac{1}{n}\sqrt{\frac{4\mathsf{SNR}}{\sigma_{\mathbf{Z}}^2}}\right)^n \\
&< \frac{V_n((1+\rho)n\sigma_{\mathbf{Z}}^2)^{n/2}}{\gamma^n p^{-k}} e^{\sqrt{\frac{4\mathsf{SNR}}{\sigma_{\mathbf{Z}}^2}}} \\
&= \left(\frac{nV_n^{2/n}(1+\rho)\sigma_{\mathbf{Z}}^2}{(\gamma^n p^{-k})^{2/n}}\right)^{n/2} e^{\sqrt{\frac{4\mathsf{SNR}}{\sigma_{\mathbf{Z}}^2}}} \\
&< \left(\frac{2\pi e(1+\rho)\sigma_{\mathbf{Z}}^2}{(\gamma^n p^{-k})^{2/n}}\right)^{n/2} e^{\sqrt{\frac{4\mathsf{SNR}}{\sigma_{\mathbf{Z}}^2}}} \qquad (24)
\end{aligned}$$

where (23) follows from Lemma 1 and (24) from the fact that $nV_n^{2/n} < 2\pi e$. If $\mathbf{G}_f$ is full-rank, as is the case for almost all lattices in the ensemble, we have

$$\begin{aligned}
&\mathbb{E}_{\Lambda_f}\left(\Pr\left((\Lambda_f \setminus \gamma\mathbb{Z}^n)\bigcap \mathcal{B}(\mathbf{Z}, r_{\mathbf{Z}}) \neq \emptyset\right)\right) \\
&< \left(\frac{2\pi e(1+\rho)\sigma_{\mathbf{Z}}^2}{V(\Lambda_f)^{2/n}}\right)^{n/2} e^{\sqrt{\frac{4\mathsf{SNR}}{\sigma_{\mathbf{Z}}^2}}} \\
&< (1+\rho)^{-\frac{n}{2}} e^{\sqrt{\frac{4\mathsf{SNR}}{\sigma_{\mathbf{Z}}^2}}},
\end{aligned}$$

which vanishes for large n. ∎

Proof of Lemma 3: We upper bound the error probability of the nearest neighbor coset decoder using the *bounded distance* coset decoder, which is inferior. More precisely, we analyze the performance of a decoder that finds all lattice points of Λ_f within an Euclidean distance r from $\mathbf{t} + \mathbf{Z}$, and reduces them modulo the coarse lattice Λ_c. Namely, the decoder's output is the set $[\Lambda_f \bigcap \mathcal{B}(\mathbf{t} + \mathbf{Z}, r)] \bmod \Lambda_c$. Note that all points in this set are codewords in $\mathcal{C} = \Lambda_f \cap \mathcal{V}_c$. If there is a unique codeword in this set, this is the decoded codeword. Otherwise, the decoder declares an error. It is easy to see that regardless of the choice of r, the nearest neighbor coset decoder makes the correct decision whenever the bounded distance coset decoder does. Therefore, the error probability of the nearest neighbor coset decoder is upper bounded by that of the bounded distance coset decoder.

The bounded distance coset decoder makes an error only if at least one of the events

$$\mathbf{t} \notin \mathcal{B}(\mathbf{t} + \mathbf{Z}, r), \qquad (25)$$

or

$$(\Lambda_f \setminus (\mathbf{t} + \Lambda_c))\bigcap \mathcal{B}(\mathbf{t} + \mathbf{Z}, r) \neq \emptyset \qquad (26)$$

occurs. The event in (25) is equivalent to $\mathbf{Z} \notin \mathcal{B}(0, r)$, whereas the event in (26) is equivalent to $(\Lambda_f \setminus \Lambda_c)\bigcap \mathcal{B}(\mathbf{Z}, r) \neq \emptyset$ which is included in the event $(\Lambda_f \setminus \gamma\mathbb{Z}^n)\bigcap \mathcal{B}(\mathbf{Z}, r) \neq \emptyset$.

Therefore, for any $r > 0$, we may apply the union bound, which gives

$$P_e \leq \Pr\left(\mathbf{Z} \notin \mathcal{B}(0, r)\right) + \Pr\left((\Lambda_f \setminus \gamma\mathbb{Z}^n)\bigcap \mathcal{B}(\mathbf{Z}, r) \neq \emptyset\right).$$

Let $\rho = \sqrt{1+\delta} - 1 > 0$ and note that

$$\frac{V(\Lambda_f)^{2/n}}{2\pi e\sigma_Z^2} > 1 + \delta$$

implies

$$\frac{V(\Lambda_f)^{2/n}}{2\pi e(1+\rho)\sigma_Z^2} > 1 + \rho.$$

Thus, setting $r = r_{\mathbf{Z}} = \sqrt{(1+\rho)n\sigma_{\mathbf{Z}}^2}$, we have

$$\begin{aligned}
P_e &\leq \Pr\left(\mathbf{Z} \notin \mathcal{B}(0, r_{\mathbf{Z}})\right) \\
&\quad + \Pr\left((\Lambda_f \setminus \gamma\mathbb{Z}^n)\bigcap \mathcal{B}(\mathbf{Z}, r_{\mathbf{Z}}) \neq \emptyset\right). \qquad (27)
\end{aligned}$$

The first summand in (27) can be made smaller than $\epsilon/2$ for n large enough, since $\mathbf{Z}$ is semi norm-ergodic. The second summand is smaller than $\epsilon/2$ for almost all members in the ensemble, by Proposition 1. It follows that for almost all members of the ensemble $P_e < \epsilon$. ∎

C. Mixture Noise Is Semi Norm-Ergodic

Our aim is to show that a mixture noise composed of AWGN and a dither from a lattice that is good for MSE quantization, as appears in the mod-Λ channel (8), is semi norm-ergodic. First, we show that if Λ_c is good for MSE quantization, i.e., if its normalized second moment is close to $1/2\pi e$, then a random dither uniformly distributed over $\mathcal{V}_c$ is semi norm-ergodic. To that end, we first prove the following lemma, which is a simple extension of [9].

Lemma 4: Let $\mathcal{S} \in \mathbb{R}^n$ be a set of points with volume $V(\mathcal{S})$ and normalized second moment

$$G(\mathcal{S}) = \frac{1}{nV(\mathcal{S})}\frac{\int_{\mathcal{S}}\|\mathbf{x}\|^2 d\mathbf{x}}{V(\mathcal{S})^{\frac{2}{n}}}.$$

Let r_{eff} be the radius of an n-dimensional ball with the same volume as $V(\mathcal{S})$, i.e., $V(\mathcal{S}) = V_n r_{\text{eff}}^n$. For any $0 < \epsilon < 1$ define

$$r_\epsilon \triangleq \sqrt{\frac{2\pi e G(\mathcal{S}) - \frac{n}{n+2}(1-\epsilon)^{1+\frac{2}{n}}}{\epsilon}} r_{\text{eff}}.$$

The probability that a random variable $\mathbf{U} \sim \text{Unif}(\mathcal{S})$ leaves a ball with radius r_ϵ is upper bounded by

$$\Pr\left(\mathbf{U} \notin \mathcal{B}(\mathbf{0}, r_\epsilon)\right) \leq \epsilon.$$

Proof: Let $\tilde{r}_\epsilon$ be the radius of a ball that contains exactly a fraction of $1-\epsilon$ of the volume of $\mathcal{S}$, i.e.,

$$\text{Vol}\left(\mathcal{S} \bigcap \mathcal{B}(\mathbf{0}, \tilde{r}_\epsilon)\right) = (1-\epsilon)V(\mathcal{S}).$$

Clearly, $\Pr\left(\mathbf{U} \notin \mathcal{B}(\mathbf{0}, \tilde{r}_\epsilon)\right) = \epsilon$. In order to establish the lemma we have to show that $\tilde{r}_\epsilon < r_\epsilon$. To that end, we write

$$\begin{aligned} nG(\mathcal{S})V(\mathcal{S})^{\frac{2}{n}} &= \frac{1}{V(\mathcal{S})}\int_{\mathbf{x}\in\mathcal{S}} \|\mathbf{x}\|^2 d\mathbf{x} \\ &= \frac{1}{V(\mathcal{S})}\Bigg(\int_{\mathbf{x}\in(S\cap\mathcal{B}(\mathbf{0},\tilde{r}_\epsilon))} \|\mathbf{x}\|^2 d\mathbf{x} \\ &\qquad + \int_{\mathbf{x}\in(S\cap(\mathbb{R}^n\setminus\mathcal{B}(\mathbf{0},\tilde{r}_\epsilon)} \|\mathbf{x}\|^2 d\mathbf{x}\Bigg). \quad (28) \end{aligned}$$

The first integral in (28) may be lower bounded by replacing its integration boundaries with an n-dimensional ball $\mathcal{B}(\mathbf{0}, \rho_\epsilon)$, where

$$\rho_\epsilon^2 = V_n^{-\frac{2}{n}}(1-\epsilon)^{\frac{2}{n}} V(\mathcal{S})^{\frac{2}{n}} \quad (29)$$

is chosen such that $V_n\rho_\epsilon^n = (1-\epsilon)V(\mathcal{S})$. Thus

$$\begin{aligned} \int_{\mathbf{x}\in(S\cap\mathcal{B}(\mathbf{0},\tilde{r}_\epsilon)} \|\mathbf{x}\|^2 d\mathbf{x} &\geq \int_{\mathbf{x}\in\mathcal{B}(\mathbf{0},\rho_\epsilon)} \|\mathbf{x}\|^2 d\mathbf{x} \\ &= nV_n\rho_\epsilon^n\sigma^2\left(\mathcal{B}(\mathbf{0},\rho_\epsilon)\right) \\ &= \frac{n}{n+2}V_n\rho_\epsilon^n\rho_\epsilon^2 \quad (30) \\ &= \frac{n}{n+2}\frac{V(\mathcal{S})^{1+\frac{2}{n}}(1-\epsilon)^{1+\frac{2}{n}}}{V_n^{\frac{2}{n}}} \\ &= \frac{n}{n+2}V(\mathcal{S})(1-\epsilon)^{1+\frac{2}{n}} r_{\text{eff}}^2, \quad (31) \end{aligned}$$

where we have used (3) to get (30). The second integral in (28) is over a set of points with volume $\epsilon V(\mathcal{S})$ which are all at distance greater than $\tilde{r}_\epsilon$ from the origin. Therefore, it can be bounded as

$$\int_{\mathbf{x}\in(S\cap(\mathbb{R}^n\setminus\mathcal{B}(\mathbf{0},\tilde{r}_\epsilon)))} \|\mathbf{x}\|^2 d\mathbf{x} \geq \epsilon V(\mathcal{S})\tilde{r}_\epsilon^2. \quad (32)$$

Substituting (31) and (32) into (28) gives

$$nG(\mathcal{S})V(\mathcal{S})^{\frac{2}{n}} \geq \left(\frac{n}{n+2}(1-\epsilon)^{1+\frac{2}{n}} r_{\text{eff}}^2 + \epsilon\tilde{r}_\epsilon^2\right). \quad (33)$$

Using the fact that $V(\mathcal{S})^{\frac{2}{n}} = V_n^{\frac{2}{n}} r_{\text{eff}}^2$, (33) reduces to

$$\begin{aligned} \tilde{r}_\epsilon^2 &\leq \frac{nV_n^{\frac{2}{n}}G(\mathcal{S}) - \frac{n}{n+2}(1-\epsilon)^{1+\frac{2}{n}}}{\epsilon} r_{\text{eff}}^2 \\ &\leq \frac{2\pi e G(\mathcal{S}) - \frac{n}{n+2}(1-\epsilon)^{1+\frac{2}{n}}}{\epsilon} r_{\text{eff}}^2, \end{aligned}$$

as desired. ■

We would like to apply Lemma 4 to a sequence of random variables uniformly distributed over the Voronoi regions $\mathcal{V}_c^{(n)}$ of a sequence of lattices $\Lambda_c^{(n)}$. From (6), we have

$$r_{\text{eff}}\left(\Lambda_c^{(n)}\right) \leq \sqrt{(n+2)\sigma^2\left(\Lambda_c^{(n)}\right)}. \quad (34)$$

If the sequence of lattices $\Lambda_c^{(n)}$ is good for MSE quantization, then for any $\delta_1 > 0$ and n large enough

$$G(\Lambda_c^{(n)}) < (1+\delta_1)\frac{1}{2\pi e}. \quad (35)$$

Combining (34) and (35), it follows that if $\Lambda_c^{(n)}$ is good for MSE quantization, for any $\delta_1 > 0$ and n large enough r_ϵ of Lemma 4 satisfies

$$r_\epsilon \leq \sqrt{\left(1+\frac{\delta_1}{\epsilon}\right)(n+2)\sigma^2(\Lambda_c^{(n)})}. \quad (36)$$

For any $\delta > 0$ and $\epsilon > 0$ we may choose $\delta_1 = \delta\epsilon/2$, which gives the following proposition.

Proposition 2: If the sequence of lattices $\Lambda_c^{(n)}$ is good for MSE quantization, then the sequence of random vectors $\mathbf{U} \sim \text{Unif}(\mathcal{V}_c^{(n)})$ is semi norm-ergodic.

The next lemma states that any linear combination of semi norm-ergodic noise and a dither from a lattice which is good for MSE quantization is itself semi norm-ergodic.

Lemma 5: Let $\mathbf{Z} = \alpha\mathbf{N} + \beta\mathbf{U}$, where $\alpha, \beta \in \mathbb{R}$, $\mathbf{N}$ is semi norm-ergodic noise with effective variance $\sigma_{\mathbf{N}}^2 = \frac{1}{n}\mathbb{E}\|\mathbf{N}\|^2$, and $\mathbf{U}$ is a dither statistically independent of $\mathbf{N}$ with effective variance $\sigma_U^2 = \frac{1}{n}\mathbb{E}\|\mathbf{U}\|^2$, uniformly distributed over the Voronoi region $\mathcal{V}$ of a lattice Λ which is good for MSE quantization. Then, the sequence of random vectors $\mathbf{Z}$ is semi norm-ergodic.

Proof: Since $\mathbf{N}$ and $\mathbf{U}$ are statistically independent, the effective variance of $\mathbf{Z}$ is

$$\sigma_{\mathbf{Z}}^2 = \frac{1}{n}\mathbb{E}\|\mathbf{Z}\|^2 = \alpha^2\sigma_{\mathbf{N}}^2 + \beta^2\sigma_{\mathbf{U}}^2.$$

We have to prove that for any $\epsilon > 0$, $\delta > 0$ and n large enough

$$\Pr\left(\mathbf{Z} \notin \mathcal{B}(\mathbf{0}, \sqrt{(1+\delta)n\sigma_{\mathbf{Z}}^2})\right) < \epsilon.$$

For any $\epsilon > 0$, $\delta > 0$ and n large enough we have

$$\begin{aligned}
&\Pr\left(\mathbf{Z} \notin \mathcal{B}(\mathbf{0}, \sqrt{(1+\delta)n\sigma_{\mathbf{Z}}^2})\right) \\
&= \Pr\left(\|\mathbf{Z}\|^2 > (1+\delta)n\sigma_{\mathbf{Z}}^2\right) \\
&= \Pr\left(\|\mathbf{N}\|^2 > (1+\delta)n\sigma_{\mathbf{N}}^2\right) \\
&\quad \cdot \Pr\left(\|\mathbf{Z}\|^2 > (1+\delta)n\sigma_{\mathbf{Z}}^2 \;\middle|\; \|\mathbf{N}\|^2 > (1+\delta)n\sigma_{\mathbf{N}}^2\right) \\
&+ \Pr\left(\|\mathbf{N}\|^2 \le (1+\delta)n\sigma_{\mathbf{N}}^2\right) \\
&\quad \cdot \Pr\left(\|\mathbf{Z}\|^2 > (1+\delta)n\sigma_{\mathbf{Z}}^2 \;\middle|\; \|\mathbf{N}\|^2 \le (1+\delta)n\sigma_{\mathbf{N}}^2\right) \\
&\le \frac{\epsilon}{3} + \Pr\Bigg(\beta^2\|\mathbf{U}\|^2 + 2\alpha\beta\mathbf{N}^T\mathbf{U} \\
&\qquad\qquad > (1+\delta)n\beta^2\sigma_{\mathbf{U}}^2 \;\Big|\; \|\mathbf{N}\|^2 \le (1+\delta)n\sigma_{\mathbf{N}}^2\Bigg) \qquad (37) \\
&\le \frac{\epsilon}{3} + \Pr\left(\beta^2\|\mathbf{U}\|^2 > n\beta^2\sigma_{\mathbf{U}}^2(1+\delta/2)\right) \\
&\quad + \Pr\left(2\alpha\beta\mathbf{N}^T\mathbf{U} > n\beta^2\sigma_{\mathbf{U}}^2\delta/2 \;\middle|\; \|\mathbf{N}\|^2 \le (1+\delta)n\sigma_{\mathbf{N}}^2\right) \qquad (38) \\
&\le \frac{2\epsilon}{3} + \Pr\left(2\alpha\beta\mathbf{N}^T\mathbf{U} > n\beta^2\sigma_{\mathbf{U}}^2\delta/2 \;\middle|\; \|\mathbf{N}\|^2 \le (1+\delta)n\sigma_{\mathbf{N}}^2\right), \qquad (39)
\end{aligned}$$

where (37) follows from the fact that $\mathbf{N}$ is semi norm-ergodic, (38) from the union bound and (39) from the fact that $\mathbf{U}$ is semi norm-ergodic due to Proposition 2. We are left with the task of showing that the last probability in (39) can be made smaller than $\epsilon/3$ for n large enough. This requires some more work.

Since $\mathbf{U}$ is semi norm-ergodic noise, than for any $\epsilon_2 > 0$, $\delta_2 > 0$ and n large enough

$$\Pr\left(\|\mathbf{U}\| > \sqrt{(1+\delta_2)n\sigma_{\mathbf{U}}^2}\right) < \epsilon_2.$$

Let $r_{\mathbf{U}} = \sqrt{(1+\delta_2)n\sigma_{\mathbf{U}}^2}$, and $f_{\mathbf{U}}(\mathbf{u})$ be the probability density function (pdf) of $\mathbf{U}$. For any $r > 0$ we have

$$\begin{aligned}
\Pr\left(\mathbf{N}^T\mathbf{U} > r \;\middle|\; \mathbf{N} = \mathbf{n}\right) &= \int_{|\mathbf{u}| \le r_{\mathbf{u}}} f_{\mathbf{U}}(\mathbf{u})\mathbb{1}(\mathbf{n}^T\mathbf{u} > r)d\mathbf{u} \\
&\quad + \int_{|\mathbf{u}| > r_{\mathbf{u}}} f_{\mathbf{U}}(\mathbf{u})\mathbb{1}(\mathbf{n}^T\mathbf{u} > r)d\mathbf{u} \\
&\le \int_{|\mathbf{u}| \le r_{\mathbf{u}}} \frac{1}{V(\Lambda)}\mathbb{1}(\mathbf{n}^T\mathbf{u} > r)d\mathbf{u} + \epsilon_2 \\
&= \frac{V(\mathcal{B}(\mathbf{0}, r_{\mathbf{U}}))}{V(\Lambda)} \int_{|\mathbf{u}| \le r_{\mathbf{u}}} \frac{1}{V(\mathcal{B}(\mathbf{0}, r_{\mathbf{U}}))}\mathbb{1}(\mathbf{n}^T\mathbf{u} > r)d\mathbf{u} + \epsilon_2.
\end{aligned}$$

Using the fact that Λ is good for MSE quantization we have $V(\Lambda)^{2/n} \to 2\pi e\sigma_{\mathbf{U}}^2$, and hence, for n large enough,

$$\left(\frac{V(\mathcal{B}(\mathbf{0}, r_{\mathbf{U}}))}{V(\Lambda)}\right)^{\frac{2}{n}} < (1+2\delta_2).$$

Let $\tilde{\mathbf{U}}$ be a random vector uniformly distributed over $\mathcal{B}(\mathbf{0}, r_{\mathbf{U}})$. We have

$$\Pr\left(\mathbf{N}^T\mathbf{U} > r \;\middle|\; \mathbf{N} = \mathbf{n}\right) < \epsilon_2 + (1+2\delta_2)^{\frac{n}{2}} \Pr(\mathbf{n}^T\tilde{\mathbf{U}} > r). \qquad (40)$$

Let $\tilde{\mathbf{Z}}$ be AWGN with zero mean and variance $r_{\mathbf{U}}^2/n$. Using a similar approach to that taken in [2, Lemma 11], we would now like to upper bound the pdf of $\tilde{\mathbf{U}}$ using that of $\tilde{\mathbf{Z}}$. For any $\mathbf{x} \in \mathbb{R}^n$ we have

$$\frac{f_{\tilde{\mathbf{U}}}(\mathbf{x})}{f_{\tilde{\mathbf{Z}}}(\mathbf{x})} = \frac{f_{\tilde{\mathbf{U}}}(\|\mathbf{x}\|)}{f_{\tilde{\mathbf{Z}}}(\|\mathbf{x}\|)} \le \frac{f_{\tilde{\mathbf{U}}}(r_{\mathbf{U}})}{f_{\tilde{\mathbf{Z}}}(r_{\mathbf{U}})} = \left(\frac{2\pi e}{nV_n^{2/n}}\right)^{\frac{n}{2}}.$$

Thus, for any $\mathbf{x} \in \mathbb{R}^n$

$$f_{\tilde{\mathbf{U}}}(\mathbf{x}) \le 2^{\frac{n}{2}\log\left(\frac{2\pi e}{n}V_n^{-2/n}\right)} f_{\tilde{\mathbf{Z}}}(\mathbf{x}).$$

We can further bound (40) for large enough n as

$$\begin{aligned}
&\Pr\left(\mathbf{N}^T\mathbf{U} > r \;\middle|\; \mathbf{N} = \mathbf{n}\right) \\
&\quad \le \epsilon_2 + 2^{\frac{n}{2}\log\left((1+2\delta_2)\frac{2\pi e}{n}V_n^{-2/n}\right)} \Pr(\mathbf{n}^T\tilde{\mathbf{Z}} > r) \\
&\quad = \epsilon_2 + 2^{\frac{n}{2}\log\left((1+2\delta_2)\frac{2\pi e}{n}V_n^{-2/n}\right)} Q\left(\frac{\sqrt{n}r}{\|\mathbf{n}\|r_{\mathbf{U}}}\right),
\end{aligned}$$

where $Q(\cdot)$ is the standard Q-function, which satisfies $Q(x) < e^{-x^2/2}$. It follows that

$$\begin{aligned}
&\Pr\left(2\alpha\beta\mathbf{N}^T\mathbf{U} > n\beta^2\sigma_{\mathbf{U}}^2\delta/2 \;\middle|\; \|\mathbf{N}\|^2 \le (1+\delta)n\sigma_{\mathbf{N}}^2\right) \\
&\le \epsilon_2 \\
&+ 2^{\frac{n}{2}\log\left((1+2\delta_2)\frac{2\pi e}{n}V_n^{-2/n}\right)} Q\left(\frac{\sqrt{n}\beta\sigma_{\mathbf{U}}\delta/2}{2\alpha\sigma_{\mathbf{N}}\sqrt{(1+\delta)(1+2\delta_2)}}\right).
\end{aligned}$$

Taking δ_2 sufficiently smaller than δ and $\epsilon_2 < \epsilon/6$, for n large enough we have

$$\Pr\left(2\alpha\beta\mathbf{N}^T\mathbf{U} > n\beta^2\sigma_{\mathbf{U}}^2\delta/2 \;\middle|\; \|\mathbf{N}\|^2 \le (1+\delta)n\sigma_{\mathbf{N}}^2\right) < \frac{\epsilon}{3}.$$

■

The next two corollaries are simple consequences of Lemma 5. The first follows since any i.i.d. noise is semi norm-ergodic, and the second follows by iterating over Lemma 5.

Corollary 1: Let $\mathbf{N}$ be i.i.d. random vector and $\mathbf{U}$ be a dither random vector that is statistically independent of $\mathbf{N}$ and is uniformly distributed over the Voronoi region $\mathcal{V}$ of a lattice Λ that is good for MSE quantization. Then, any linear combination of the form $\mathbf{Z} = \alpha N + \beta\mathbf{U}$, where $\alpha, \beta \in \mathbb{R}$, is semi norm-ergodic.

Corollary 2: Let $\mathbf{U}_1, \cdots, \mathbf{U}_K$ be statistically independent dither random vectors, each uniformly distributed over the Voronoi region $\mathcal{V}_k$ of Λ_k, $k = 1, \ldots, K$, that are all good for MSE quantization. Let $\mathbf{N}$ be an i.i.d. random vector statistically independent of $\{\mathbf{U}_1, \cdots, \mathbf{U}_K\}$. For any $\alpha, \beta_1, \cdots, \beta_K \in \mathbb{R}$ the random vector $\mathbf{Z} = \alpha\mathbf{N} + \sum_{k=1}^K \beta_k\mathbf{U}_k$ is semi norm-ergodic.

D. Tying It All Together

First, we combine Lemma 2 and Lemma 3 to obtain an existence theorem for nested lattice pairs which are "good" in the sense of Definition 4. Note that in the considered lattice construction, the normalized volume of the coarse lattice is

given by

$$\frac{1}{n}\log V(\Lambda_c) = \frac{1}{n}\log(\gamma^n p^{-k_1}) = \log(\gamma) - \frac{k_1}{n}\log p$$
$$= \log\left(\frac{\gamma\sqrt{V_n^{2/n}}}{2}\right) - \epsilon_1$$
$$\stackrel{n\to\infty}{\longrightarrow} \frac{1}{2}\log(2\pi e\mathsf{SNR}) - \epsilon_1. \quad (41)$$

Using the parameter $k > k_1$ in the definition of the nested lattices ensemble, we can set the normalized volume of the fine lattice to any desired value of $\frac{1}{n}\log V(\Lambda_f) < \frac{1}{2}\log(2\pi e\mathsf{SNR}) - \epsilon_1$. Since δ_1 from Lemma 2 is larger than ϵ_1, this means that we can set the normalized volume of the fine lattice to any desired value of $\frac{1}{n}\log V(\Lambda_f) < \frac{1}{2}\log(2\pi e\mathsf{SNR}) - \delta_1$. Thus, we have the following theorem.

Theorem 1: For any $\sigma_{\mathbf{Z}}^2 > 0$, $\epsilon > 0$, $\delta > 0$ and

$$\frac{1}{2}\log\left(2\pi e\sigma_{\mathbf{Z}}^2\right) + \delta \quad < \frac{1}{n}\log V(\Lambda_f) < \frac{1}{2}\log(2\pi e\mathsf{SNR}) - \delta,$$

for n large enough, there exists a pair of nested lattices $\Lambda_c \subset \Lambda_f$ such that

1) $\sigma^2(\Lambda_c) = \mathsf{SNR}$ and $G(\Lambda_c) < \frac{1+\delta}{2\pi e}$.
2) For any lattice point $\mathbf{t} \in \Lambda_f$, and additive semi norm-ergodic noise $\mathbf{Z}$ with effective variance $\sigma_{\mathbf{Z}}^2 = \frac{1}{n}\mathbb{E}\|\mathbf{Z}\|^2$

$$\Pr\left(\left[Q_{\Lambda_f}(\mathbf{t}+\mathbf{Z})\right] \bmod \Lambda_c \neq \mathbf{t} \bmod \Lambda_c\right) < \epsilon.$$

We now apply Theorem 1 to show that nested lattice codes achieve the capacity of the AWGN channel using the mod-Λ transmission scheme with coset nearest neighbor decoding. Theorem 1 states that nested lattice codes achieve arbitrary low error probabilities over channels of the form $\mathbf{Y} = \mathbf{t} + \mathbf{Z}$, where $\mathbf{Z}$ is semi norm-ergodic additive noise. However, the output of the equivalent channel when applying the mod-Λ transmission scheme is not exactly of this form. Rather, it is given by

$$\mathbf{Y}_{\text{eff}} = [\mathbf{t} + \mathbf{Z}_{\text{eff}}] \bmod \Lambda_c$$
$$= \mathbf{t} - Q_c(\mathbf{t} + \mathbf{Z}_{\text{eff}}) + \mathbf{Z}_{\text{eff}}.$$

Thus, it is equivalent to an additive channel with input $\mathbf{t} - Q_c(\mathbf{t} + \mathbf{Z}_{\text{eff}})$. Nevertheless, as

$$[\mathbf{t} - Q_c(\mathbf{t} + \mathbf{Z}_{\text{eff}})] \bmod \Lambda_c = \mathbf{t},$$

the coset nearest neighbor decoder cannot distinguish between the inputs $\mathbf{t}$ and $\mathbf{t} - Q_c(\mathbf{t} + \mathbf{Z}_{\text{eff}})$. Therefore, when the coset nearest neighbor decoder is used, the output of the induced channel is equivalent to $\mathbf{t} + \mathbf{Z}_{\text{eff}}$. See Figure 2 for an illustration of the nearest neighbor coset decoder's operation.

The effective noise $\mathbf{Z}_{\text{eff}}$ is a linear combination of an AWGN and a dither uniformly distributed over $\mathcal{V}_c$. When the coarse lattice Λ_c is good for MSE quantization, $\mathbf{Z}_{\text{eff}}$ is semi norm-ergodic by Corollary 1.

It follows that for "good" pairs of nested lattices, the error probability in coset nearest neighbor decoding of $\mathbf{Y}_{\text{eff}}$, the output of the equivalent channel induced by the mod-Λ scheme, can be made arbitrary small for all

$$R < \frac{1}{2}\log\left(\frac{\mathsf{SNR}}{\sigma_{\text{eff}}^2}\right),$$

where σ_{eff}^2 is the effective variance of $\mathbf{Z}_{\text{eff}}$.

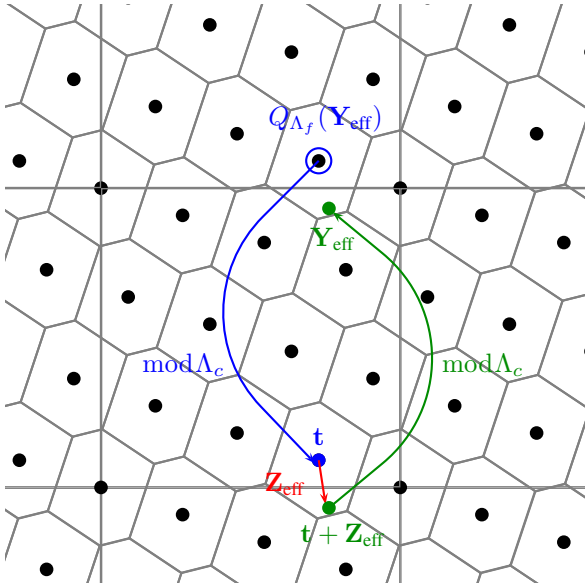

Fig. 2. An illustration of the coset nearest neighbor decoding process. The lattice point $\mathbf{t}$ was transmitted. The output of the induced channel when the mod-Λ transmission scheme is applied is $\mathbf{Y}_{\text{eff}} = [\mathbf{t} + \mathbf{Z}_{\text{eff}}] \bmod \Lambda_c$. The decoder quantizes $\mathbf{Y}_{\text{eff}}$ to the nearest lattice point in Λ_f and reduces the quantized output modulo Λ_c. In the figure the coarse lattice Λ_c is the integer lattice.

Remark 2: Note that we have not used the Gaussianity of $\mathbf{N}$ in the proof. Therefore, any rate below $\frac{1}{2}\log(\mathsf{SNR}/\sigma_{\text{eff}}^2)$ is in fact achievable using nested lattice codes and coset nearest neighbor decoding for any additive i.i.d. noise $\mathbf{N}$, or more generally, any additive semi norm-ergodic noise. This is analogous to the results of [10] where it is shown that $\frac{1}{2}\log(1+\mathsf{SNR})$ is the highest rate a coding scheme that utilizes Gaussian codebooks and nearest neighbor decoding can achieve over an additive ergodic noise channel.

Remark 3: Note that for unshaped transmission, i.e., where the coarse lattice is the cubic lattice, a random dither $\mathbf{U}$ uniformly distributed over its Voronoi region is an i.i.d. random vector. Hence, a linear combination of such a dither and AWGN is semi norm-ergodic. It follows that the effective noise in the mod-Λ transmission scheme is semi norm-ergodic even when no shaping is used. If the fine lattice is good for coding in the presence of additive semi norm-ergodic noise, which as Lemma 3 indicates, is the case for almost all Construction A lattices, then any rate satisfying

$$R < \frac{1}{2}\log\left(\frac{\mathsf{SNR}}{\sigma_{\text{eff}}^2}\right) - \frac{1}{2}\log\left(\frac{2\pi e}{12}\right) \quad (42)$$

is achievable with a one-dimensional coarse lattice.

IV. Extensions

During the last decade, nested lattice codes have been extensively used in the literature in the context of Gaussian network problems, see e.g. [11]–[15]. For such problems, a pair of nested lattices is often not sufficient, and more complicated *chains* of nested lattices are required. A major advantage of the nested lattice ensemble we have used in Section III for proving the existence of "good" pairs, is that it can be naturally extended to construct any (finite) chain of nested lattices. Specifically, we can draw a linear code with a generating matrix $\mathbf{G}$ of dimensions $k \times n$ over $\mathbb{Z}_p$, and then for $k \geq k_L > \cdots > k_1$ construct a sequence of nested linear codebooks $\mathcal{C}_1 \subset \cdots \subset \mathcal{C}_L$, by taking the generating matrix of the ℓth codebook to be the first k_ℓ rows of $\mathbf{G}$. These codewords can be lifted to the Euclidean space using Construction A in order to form a chain of nested lattices.

In Lemma 2 we have shown that almost all random Construction A lattices are good for MSE quantization. In Lemma 3 we have shown that almost all random Construction A pairs of nested lattices achieve low error probabilities in the presence of additive semi norm-ergodic noise, under coset nearest neighbor decoding. It follows that for any finite number L and any $\frac{1}{n}\log V(\Lambda_L) < \cdots < \frac{1}{n}\log V(\Lambda_1)$, there exists a chain of nested lattices $\Lambda_1 \subset \cdots \subset \Lambda_L$ such that all lattices are good for MSE quantization and all pairs within the chain achieve low error probabilities under coset nearest neighbor decoding in the presence of additive semi norm-ergodic noise.

In certain applications, chains of nested lattice codes are used in order to convert a Gaussian multiple access channel (MAC) into an effective modulo-lattice channel whose output is a fine lattice point plus effective noise reduced modulo a coarse lattice. Such a situation arises for example in the compute-and-forward framework [12], where a receiver is interested in decoding linear combinations with integer valued coefficients of the codewords transmitted by the different users of the MAC. In such applications, the effective noise is often a linear combination of AWGN and *multiple* statistically independent dithers uniformly distributed over the Voronoi region of the coarse lattice. By Corollary 2, such an effective noise is semi norm-ergodic regardless of the number of dithers contributing to it, as long as they are all independent and are induced by lattices which are good for MSE quantization.

In our proofs we took the cardinality p of the finite field over which we construct the linear codes to grow polynomially with n. While this facilitates the derivations, it does not seem to be necessary in practice when targeting a fixed (small) gap to capacity. Indeed, in [16] it was shown that construction A lattices with small values of p, such as 2, 3 and 5, achieve an NSM very close to $1/(2\pi e)$. For additive noise channels with a PAM input constellation with prime cardinality, linear codes perform just as well as random codes. Therefore, it seems that for low /moderate values of SNR the ensemble of nested lattice codes we have used can achieve rates close to the capacity of the AWGN channel even with small values of p.

Acknowledgment

The authors thank Bobak Nazer, Yair Yona and Ram Zamir for discussions that helped prompt this work.

References

[1] R. de Buda, "The upper error bound of a new near-optimal code," *IEEE Transactions on Information Theory*, vol. 21, no. 4, pp. 441–445, Jul. 1975.

[2] U. Erez and R. Zamir, "Achieving $\frac{1}{2}\log(1 + \text{SNR})$ on the AWGN channel with lattice encoding and decoding," *IEEE Trans. Information Theory*, vol. IT-50, pp. 2293–2314, Oct. 2004.

[3] T. Liu, P. Moulin, and R. Koetter, "On error exponents of modulo lattice additive noise channels," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 454–471, Feb. 2006.

[4] R. Zamir and S. Shamai, "Nested linear/lattice codes for Wyner-Ziv encoding," in *Information Theory Workshop*, Jun. 1998, pp. 92 –93.

[5] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups*. New York: Springer-Verlag, 1988.

[6] U. Erez, S. Litsyn, and R. Zamir, "Lattices which are good for (almost) everything," *IEEE Transactions on Information Theory*, vol. 51, no. 10, pp. 3401–3416, Oct. 2005.

[7] H.-A. Loeliger, "Averaging bounds for lattices and linear codes," *IEEE Transactions on Information Theory*, vol. 43, no. 6, pp. 1767–1773, Nov. 1997.

[8] R. Zamir, S. Shamai (Shitz), and U. Erez, "Nested linear/lattice codes for structured multiterminal binning," *IEEE Trans. Information Theory*, vol. 48, no. 1, pp. 1250–1276, Jun. 2002.

[9] U. Erez and R. Zamir, "Bounds on the ϵ-covering radius of linear codes with applications to self noise in nested Wyner-Ziv coding," *dept. Elec. Eng-Syst., Tel-Aviv Univ., Technical Report*, 2002, available online: http://www.eng.tau.ac.il/ zamir/techreport/selfnoiseTR.pdf.

[10] A. Lapidoth, "Nearest neighbor decoding for additive non-Gaussian noise channels," *IEEE Trans. Information Theory*, vol. 42, no. 5, pp. 1520–1529, Sep 1996.

[11] T. Philosof, R. Zamir, U. Erez, and A. J. Khisti, "Lattice strategies for the dirty multiple access channel," *IEEE Transactions on Information Theory*, vol. 57, no. 8, pp. 5006–5035, Aug. 2011.

[12] B. Nazer and M. Gastpar, "Compute-and-forward: Harnessing interference through structured codes," *IEEE Transactions on Information Theory*, vol. 57, no. 10, pp. 6463–6486, Oct. 2011.

[13] W. Nam, S.-Y. Chung, and Y. H. Lee, "Capacity of the Gaussian two-way relay channel to within $1/2$ bit," *IEEE Transactions on Information Theory*, vol. 56, no. 11, pp. 5488–5494, Nov. 2010.

[14] G. Bresler, A. Parekh, and D. N. C. Tse, "The approximate capacity of the many-to-one and one-to-many Gaussian interference channels," *IEEE Trans. Information Theory*, vol. 56, pp. 4566–4592, Sep. 2010.

[15] O. Ordentlich, U. Erez, and B. Nazer, "The approximate sum capacity of the symmetric Gaussian K-user interference channel," *IEEE Trans. Information Theory*, Submitted May 2012, see http://arxiv.org/abs/1206.0197.

[16] K. V. Yurkov and B. D. Kudryashov, "Random quantization bounds for lattices over q-ary linear codes," in *Information Theory, 2007. ISIT 2007. IEEE International Symposium on*, June 2007, pp. 236 –240.