§ 18. Lattice codes (by O. Ordentlich)

Consider the n-dimensional additive white Gaussian noise (AWGN) channel

$$\mathbf{Y} = \mathbf{X} + \mathbf{Z}$$

where $\mathbf{Z} \sim \mathcal{N}(0, \mathbf{I}_{n\times n})$ is statistically independent of the input $\mathbf{X}$. Our goal is to communicate reliably over this channel, under the power constraint

$$\frac{1}{n}\|\mathbf{X}\|^2 \le \mathsf{SNR}$$

where SNR is the *signal-to-noise-ratio.* The capacity of the AWGN channel is

$$C = \tfrac{1}{2}\log(1 + \mathsf{SNR}) \text{ bits/channel use,}$$

and is achieved with high probability by a codebook drawn at random from the Gaussian i.i.d. ensemble. However, a typical codebook from this ensemble has very little structure, and is therefore not applicable for practical systems. A similar problem occurs in discrete additive memoryless stationary channels, e.g., BSC, where most members of the capacity achieving i.i.d. uniform codebook ensemble have no structure. In the discrete case, engineers resort to linear codes to circumvent the lack of structure. Lattice codes are the Euclidean space counterpart of linear codes, and as we shall see, enable to achieve the capacity of the AWGN channel with much more structure than random codes. In fact, we will construct a lattice code with rate that approaches $\frac{1}{2}\log(1+\mathsf{SNR})$ that is guaranteed to achieve small error probability for essentially all additive noise channels with the same noise second moment. More precisely, our scheme will work if the noise vector $\mathbf{Z}$ is *semi norm-ergodic.*

Definition 18.1. We say that a sequence in n of random noise vectors $\mathbf{Z}^{(n)}$ of length n with (finite) effective variance $\sigma^2_{\mathbf{Z}} \triangleq \frac{1}{n}\mathbb{E}\|\mathbf{Z}^{(n)}\|^2$, is *semi norm-ergodic* if for any $\epsilon, \delta > 0$ and n large enough

$$\Pr\left(\mathbf{Z}^{(n)} \notin \mathcal{B}(\sqrt{(1+\delta)n\sigma^2_{\mathbf{Z}}}\right) \le \epsilon, \tag{18.1}$$

where $\mathcal{B}(r)$ is an n-dimensional ball of radius r.

18.1 Lattice Definitions

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 $\mathbf{G}$ such that

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

We will assume $\mathbf{G}$ is full-rank. Denote the nearest neighbor quantizer associated with the lattice Λ by

$$Q_\Lambda(\mathbf{x}) \triangleq \arg\min_{\mathbf{t}\in\Lambda} \|\mathbf{x}-\mathbf{t}\|, \tag{18.2}$$

where ties are broken in a systematic manner. We define the modulo operation w.r.t. a lattice Λ as

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

and note that it satisfies the distributive law,

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

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. The systematic tie-breaking in (18.2) ensures that

$$\biguplus_{\mathbf{t}\in\Lambda} (\mathcal{V}+\mathbf{t}) = \mathbb{R}^n,$$

where $\biguplus$ denotes disjoint union. Thus, $\mathcal{V}$ is a *fundamental cell* of Λ.

Definition 18.2. A measurable set $S \in \mathbb{R}^n$ is called a *fundamental cell* of Λ if

$$\biguplus_{\mathbf{t}\in\Lambda} (S+\mathbf{t}) = \mathbb{R}^n.$$

We denote the volume of a set $S \in \mathbb{R}^n$ by $\mathrm{Vol}(S)$.

Proposition 18.1. *If S is a fundamental cell of Λ, then* $\mathrm{Vol}(S) = \mathrm{Vol}(\mathcal{V})$. *Furthermore*

$$S \bmod \Lambda = \{[\mathbf{s}] \bmod \Lambda \ : \ \mathbf{s}\in S\} = \mathcal{V}.$$

Proof ([Zam14]). For any $\mathbf{t}\in\Lambda$ define

$$\mathcal{A}_\mathbf{t} \triangleq S \cap (\mathbf{t}+\mathcal{V}); \quad \mathcal{D}_\mathbf{t} \triangleq \mathcal{V} \cap (\mathbf{t}+S).$$

Note that

$$\begin{aligned}\mathcal{D}_\mathbf{t} &= [(-\mathbf{t}+\mathcal{V})\cap S] + \mathbf{t} \\ &= \mathcal{A}_{-\mathbf{t}} + \mathbf{t}.\end{aligned}$$

Thus

$$\mathrm{Vol}(S) = \sum_{\mathbf{t}\in\Lambda} \mathrm{Vol}(\mathcal{A}_\mathbf{t}) = \sum_{\mathbf{t}\in\Lambda} \mathrm{Vol}(\mathcal{A}_{-\mathbf{t}}+\mathbf{t}) = \sum_{\mathbf{t}\in\Lambda} \mathrm{Vol}(\mathcal{D}_\mathbf{t}) = \mathrm{Vol}(\mathcal{V}).$$

Moreover

$$S = \biguplus_{\mathbf{t}\in\Lambda} \mathcal{A}_\mathbf{t} = \biguplus_{\mathbf{t}\in\Lambda} \mathcal{A}_{-\mathbf{t}} = \biguplus_{\mathbf{t}\in\Lambda} \mathcal{D}_\mathbf{t} - \mathbf{t},$$

and therefore

$$[S] \bmod \Lambda = \biguplus_{\mathbf{t}\in\Lambda} \mathcal{D}_\mathbf{t} = \mathcal{V}.$$

□

Corollary 18.1. *If S is a fundamental cell of a lattice Λ with generating matrix $\mathbf{G}$, then* $\mathrm{Vol}(S) = |\det(\mathbf{G})|$. *In Particular,* $\mathrm{Vol}(\mathcal{V}) = |\det(\mathbf{G})|$.

Proof. Let $\mathcal{P} = \mathbf{G} \cdot [0,1)^n$ and note that it is a fundamental cell of Λ as $\mathbb{R}^n = \mathbb{Z}^n + [0,1)^n$. The claim now follows from Proposition 18.1 since $\mathrm{Vol}(\mathcal{P}) = |\det(\mathbf{G})| \cdot \mathrm{Vol}([0,1)^n) = |\det(\mathbf{G})|$. □

Definition 18.3 (Lattice decoder)**.** A lattice decoder w.r.t. the lattice Λ returns for every $\mathbf{y} \in \mathbb{R}^n$ the point $Q_\Lambda(\mathbf{y})$.

Remark 18.1. Recall that for linear codes, the ML decoder merely consisted of mapping syndromes to shifts. Similarly, it can be shown that a lattice decoder can be expressed as

$$Q_\Lambda(\mathbf{y}) = \mathbf{y} - g_{\text{synd}}\left([\mathbf{G}^{-1}\mathbf{y}] \bmod 1\right), \tag{18.3}$$

for some $g_{\text{synd}} : [0,1)^n \mapsto \mathbb{R}^n$, where the $\bmod 1$ operation above is to be understood as componentwise modulo reduction. Thus, a lattice decoder is indeed much more "structured" than ML decoder for a random code.

Note that for an additive channel $\mathbf{Y} = \mathbf{X} + \mathbf{Z}$, if $\mathbf{X} \in \Lambda$ we have that

$$P_e = \Pr\left(Q_\Lambda(\mathbf{Y}) \neq \mathbf{X}\right) = \Pr(\mathbf{Z} \notin \mathcal{V}). \tag{18.4}$$

We therefore see that the resilience of a lattice to additive noise is dictated by its Voronoi region. Since we know that $\mathbf{Z}$ will be inside a ball of radius $\sqrt{n(1+\delta)}$ with high probability, we would like the Voronoi region to be as close as possible to a ball. We define the effective radius of a lattice, denoted $r_{\text{eff}}(\Lambda)$ as the radius of a ball with the same volume as $\mathcal{V}$, namely $\mathrm{Vol}\left(\mathcal{B}\left(r_{\text{eff}}(\Lambda)\right)\right) = \mathrm{Vol}(\mathcal{V})$.

Definition 18.4 (Goodness for coding)**.** A sequence of lattices $\Lambda^{(n)}$ with growing dimension, satisfying

$$\lim_{n\to\infty} \frac{r_{\text{eff}}^2(\Lambda^{(n)})}{n} = \Phi$$

for some $\Phi > 0$, is called *good for channel coding* if for any additive semi norm-ergodic noise sequence $\mathbf{Z}^{(n)}$ with effective variance $\sigma_{\mathbf{Z}}^2 = \frac{1}{n}\mathbb{E}\|\mathbf{Z}\|^2 < \Phi$

$$\lim_{n\to\infty} \Pr\left(\mathbf{Z}^{(n)} \notin \mathcal{V}^{(n)}\right) = 0.$$

An alternative interpretation of this property, is that for a sequence $\Lambda^{(n)}$ that is good for coding, for any $0 < \delta < 1$ holds

$$\lim_{n\to\infty} \frac{\mathrm{Vol}\left(\mathcal{B}\left((1-\delta)r_{\text{eff}}(\Lambda^{(n)})\right) \cap \mathcal{V}^{(n)}\right)}{\mathrm{Vol}\left(\mathcal{B}\left((1-\delta)r_{\text{eff}}(\Lambda^{(n)})\right)\right)} = 1.$$

Roughly speaking, the Voronoi region of a lattice that is good for coding is as resilient to semi norm-ergodic noise as a ball with the same volume.

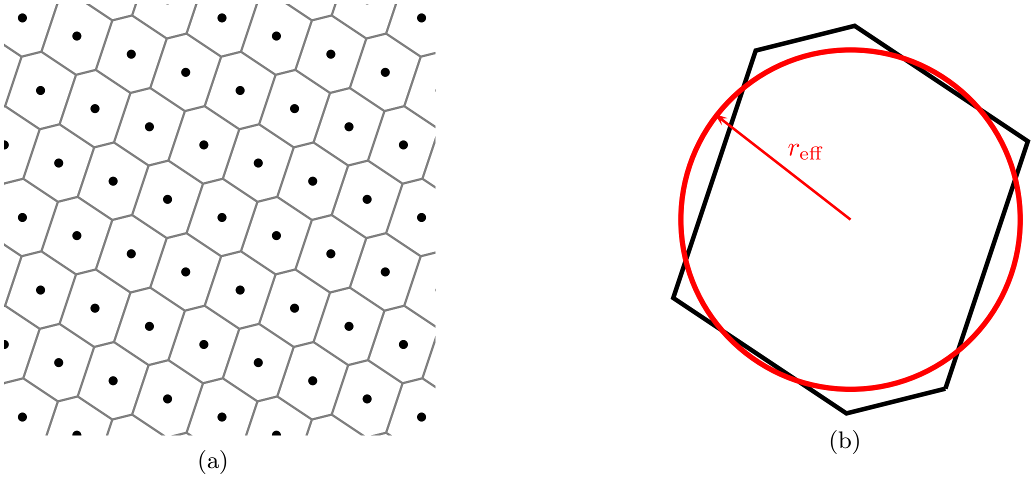

Figure 18.1: (a) shows a lattice in $\mathbb{R}^2$, and (b) shows its Voronoi region and the corresponding effective ball.

18.2 First Attempt at AWGN Capacity

Assume we have a lattice $\Lambda \subset \mathbb{R}^n$ with $r_{\text{eff}}(\Lambda) = \sqrt{n(1+\delta)}$ that is good for coding, and we would like to use it for communicating over an additive noise channel. In order to meet the power constraint, we must first intersect Λ, or a shifted version of Λ, with some compact set S that enforces the power constraint. The most obvious choice is taking S to be a ball with radius $\sqrt{n\mathsf{SNR}}$, and take some shift $\mathbf{v} \in \mathbb{R}^n$, such that the codebook

$$\mathcal{C} = (\mathbf{v} + \Lambda) \bigcap \mathcal{B}(\sqrt{n\mathsf{SNR}}) \tag{18.5}$$

satisfies the power constraint. Moreover [Loe97], there exist a shift $\mathbf{v}$ such that

$$\begin{aligned} |\mathcal{C}| &\geq \frac{\text{Vol}\,(S)}{\text{Vol}(\mathcal{V})} \\ &= \left(\frac{\sqrt{n\mathsf{SNR}}}{r_{\text{eff}}(\Lambda)}\right)^n \\ &= 2^{\frac{n}{2}(\log(\mathsf{SNR}) - \log(1+\delta))}. \end{aligned}$$

To see this, let $\mathbf{V} \sim \text{Uniform}(\mathcal{V})$, and write the expected size of $|\mathcal{C}|$ as

$$\begin{aligned} \mathbb{E}|\mathcal{C}| &= \mathbb{E} \sum_{\mathbf{t}\in\Lambda} \mathbb{1}((\mathbf{t}+\mathbf{V}) \in S) \\ &= \frac{1}{\text{Vol}(\mathcal{V})} \int_{\mathbf{v}\in\mathcal{V}} \sum_{\mathbf{t}\in\Lambda} \mathbb{1}((\mathbf{t}+\mathbf{v}) \in S) d\mathbf{v} \\ &= \frac{1}{\text{Vol}(\mathcal{V})} \int_{\mathbf{x}\in\mathbb{R}^n} \mathbb{1}(\mathbf{x} \in S) d\mathbf{x} \\ &= \frac{\text{Vol}(S)}{\text{Vol}(\mathcal{V})}. \end{aligned} \tag{18.6}$$

For decoding, we will simply apply the lattice decoder $Q_\Lambda(\mathbf{Y} - \mathbf{v})$ on the shifted output. Since $\mathbf{Y} - \mathbf{v} = \mathbf{t} + \mathbf{Z}$ for some $\mathbf{t} \in \Lambda$, the error probability is

$$P_e = \Pr\left(Q_\Lambda(\mathbf{Y} - \mathbf{v}) \neq \mathbf{t}\right) = \Pr(\mathbf{Z} \notin \mathcal{V}).$$

Since Λ is good for coding and $\frac{r_{\text{eff}}^2(\Lambda)}{n} = (1+\delta) > \frac{1}{n}\mathbb{E}\|\mathbf{Z}\|^2$, the error probability of this scheme over an additive semi norm-ergodic noise channel will vanish with n. Taking $\delta \to 0$ we see that any rate $R < \frac{1}{2}\log(\mathsf{SNR})$ can be achieved reliably. Note that for this coding scheme (encoder+decoder) the average error probability and the maximal error probability are the same.

The construction above gets us close to the AWGN channel capacity. We note that a possible reason for the loss of +1 in the achievable rate is the suboptimality of the lattice decoder for the codebook $\mathcal{C}$. The lattice decoder assumes all points of Λ were equally likely to be transmitted. However, in $\mathcal{C}$ only lattice points inside the ball can be transmitted. Indeed, it was shown [UR98] that if one replaces the lattice decoder with a decoder that takes the shaping region into account, there exist lattices and shifts for which the codebook $(\mathbf{v}+\Lambda)\bigcap\mathcal{B}(\sqrt{n\mathsf{SNR}})$ is capacity achieving. The main drawback of this approach is that the decoder no longer exploits the full structure of the lattice, so the advantages of using a lattice code w.r.t. some typical member of the Gaussian i.i.d. ensemble are not so clear anymore.

18.3 Nested Lattice Codes/Voronoi Constellations

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

$$\Gamma(\Lambda_f,\Lambda_c) \triangleq \left(\frac{\text{Vol}(\mathcal{V}_c)}{\text{Vol}(\mathcal{V}_f)}\right)^{1/n} \tag{18.7}$$

A *nested lattice code* (sometimes also called "Voronoi constellation") based on the nested lattice pair $\Lambda_c \subset \Lambda_f$ is defined as [CS83, For89, EZ04]

$$\mathcal{L} \triangleq \Lambda_f \cap \mathcal{V}_c. \tag{18.8}$$

Proposition 18.2.

$$|\mathcal{L}| = \frac{\text{Vol}(\mathcal{V}_c)}{\text{Vol}(\mathcal{V}_f)}.$$

Thus, the codebook $\mathcal{L}$ has rate $R = \frac{1}{n}\log|\mathcal{L}| = \log\Gamma(\Lambda_f,\Lambda_c)$.

Proof. First note that

$$\Lambda_f \triangleq \biguplus_{\mathbf{t}\in\mathcal{L}}(\mathbf{t}+\Lambda_c).$$

Let

$$S \triangleq \biguplus_{\mathbf{t}\in\mathcal{L}}(\mathbf{t}+\mathcal{V}_f),$$

and note that

$$\begin{aligned}
\mathbb{R}^n &= \biguplus_{\mathbf{b}\in\Lambda_f}(\mathbf{b}+\mathcal{V}_f) \\
&= \biguplus_{\mathbf{a}\in\Lambda_c}\biguplus_{\mathbf{t}\in\mathcal{L}}(\mathbf{a}+\mathbf{t}+\mathcal{V}_f) \\
&= \biguplus_{\mathbf{a}\in\Lambda_c}\left(\mathbf{a}+\left(\biguplus_{\mathbf{t}\in\mathcal{L}}(\mathbf{t}+\mathcal{V}_f)\right)\right) \\
&= \biguplus_{\mathbf{a}\in\Lambda_c}(\mathbf{a}+S).
\end{aligned}$$

Thus, S is a fundamental cell of Λ_c, and we have

$$\mathrm{Vol}(\mathcal{V}_c) = \mathrm{Vol}(S) = |\mathcal{L}| \cdot \mathrm{Vol}(\mathcal{V}_f).$$

□

We will use the codebook $\mathcal{L}$ with a standard lattice decoder, ignoring the fact that only points in $\mathcal{V}_c$ were transmitted. Therefore, the resilience to noise will be dictated mainly by Λ_f. The role of the coarse lattice Λ_c is to perform *shaping*. In order to maximize the rate of the codebook $\mathcal{L}$ without violating the power constraint, we would like $\mathcal{V}_c$ to have the maximal possible volume, under the constraint that the average power of a transmitted codeword is no more than $n\mathsf{SNR}$.

The average transmission power of the codebook $\mathcal{L}$ is related to a quantity called the *second moment of a lattice*. Let $\mathbf{U} \sim \mathrm{Uniform}(\mathcal{V})$. The second moment of Λ is defined as $\sigma^2(\Lambda) \triangleq \frac{1}{n}\mathbb{E}\|\mathbf{U}\|^2$. Let $\mathbf{W} \sim \mathrm{Uniform}(\mathcal{B}(r_{\mathrm{eff}}(\Lambda))$. By the isoperimetric inequality [Zam14]

$$\sigma^2(\Lambda) \geq \frac{1}{n}\mathbb{E}\|\mathbf{W}\|^2 = \frac{r_{\mathrm{eff}}^2(\Lambda)}{n+2}.$$

A lattice Λ exhibits a good tradeoff between average power and volume if its second moment is close to that of $\mathcal{B}(r_{\mathrm{eff}}(\Lambda)$.

Definition 18.5 (Goodness for MSE quantization)**.** A sequence of lattices $\Lambda^{(n)}$ with growing dimension, is called *good for MSE quantization* if

$$\lim_{n\to\infty} \frac{n\sigma^2\left(\Lambda^{(n)}\right)}{r_{\mathrm{eff}}^2\left(\Lambda^{(n)}\right)} = 1.$$

Remark 18.2. Note that both "goodness for coding" and "goodness for quantization" are scale invariant properties: if Λ satisfy them, so does $\alpha\Lambda$ for any $\alpha \in \mathbb{R}$.

Theorem 18.1 ([OE15])**.** *If Λ is good for MSE quantization and $\mathbf{U} \sim \mathrm{Uniform}(\mathcal{V})$, then $\mathbf{U}$ is semi norm-ergodic. Furthermore, if $\mathbf{Z}$ is semi norm-ergodic and statistically independent of $\mathbf{U}$, then for any $\alpha, \beta \in \mathbb{R}$ the random vector $\alpha\mathbf{U} + \beta\mathbf{Z}$ is semi norm-ergodic.*

Theorem 18.2 ([ELZ05, OE15])**.** *For any finite nesting ratio $\Gamma(\Lambda_f, \Lambda_c)$, there exist a nested lattice pair $\Lambda_c \subset \Lambda_f$ where the coarse lattice Λ_c is good for MSE quantization and the fine lattice Λ_f is good for coding.*

We now describe the Mod-Λ coding scheme introduced by Erez and Zamir [EZ04]. Let $\Lambda_c \subset \Lambda_f$ be a nested lattice pair, where the coarse lattice is good for MSE quantization and has $\sigma^2(\Lambda_c) = \mathsf{SNR}(1-\epsilon)$, whereas the fine lattice is good for coding and has $r_{\mathrm{eff}}^2(\Lambda_f) = n\frac{\mathsf{SNR}}{1+\mathsf{SNR}}(1+\epsilon)$. The rate is therefore

$$\begin{aligned} R &= \frac{1}{n}\log\left(\frac{\mathrm{Vol}(\mathcal{V}_c)}{\mathrm{Vol}(\mathcal{V}_f)}\right) \\ &= \frac{1}{2}\log\left(\frac{r_{\mathrm{eff}}^2(\Lambda_c)}{r_{\mathrm{eff}}^2(\Lambda_f)}\right) \\ &\to \frac{1}{2}\log\left(\frac{\mathsf{SNR}(1-\epsilon)}{\frac{\mathsf{SNR}}{1+\mathsf{SNR}}(1+\epsilon)}\right) \\ &\to \frac{1}{2}\log\left(1+\mathsf{SNR}\right), \end{aligned} \tag{18.9}$$

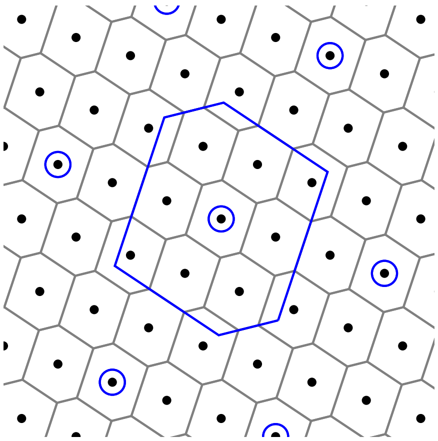

Figure 18.2: An example of a nested lattice code. The points and Voronoi region of Λ_c are plotted in blue, and the points of the fine lattice in black.

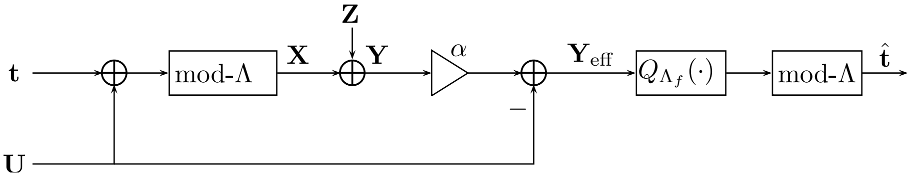

Figure 18.3: Schematic illustration of the Mod-Λ scheme.

where in (18.9) we have used the goodness of Λ_c for MSE quantization, that implies $\frac{r^2_{\text{eff}}(\Lambda_c)}{n} \to \sigma^2(\Lambda_c)$. The scheme also uses common randomness, namely a dither vector $\mathbf{U} \sim \text{Uniform}(\mathcal{V}_c)$ statistically independent of everything, known to both the transmitter and the receiver. In order to transmit a message $w \in [1, \dots, 2^{nR}]$ the encoder maps it to the corresponding point $\mathbf{t} = \mathbf{t}(w) \in \mathcal{L}$ and transmits

$$\mathbf{X} = [\mathbf{t} + \mathbf{U}] \bmod \Lambda. \tag{18.10}$$

Lemma 18.1 (Crypto Lemma)**.** *Let Λ be a lattice in $\mathbb{R}^n$, let $\mathbf{U} \sim \text{Uniform}(\mathcal{V})$ and let $\mathbf{V}$ be a random vector in $\mathbb{R}^n$, statistically independent of $\mathbf{U}$. The random vector $\mathbf{X} = [\mathbf{V} + \mathbf{U}] \bmod \Lambda$ is uniformly distributed over $\mathcal{V}$ and statistically independent of $\mathbf{V}$.*

Proof. For any $\mathbf{v} \in \mathbb{R}^n$ the set $\mathbf{v} + \mathcal{V}$ is a fundamental cell of Λ. Thus, by Proposition 18.1 we have that $[\mathbf{v} + \mathcal{V}] \bmod \Lambda = \mathcal{V}$ and $\text{Vol}(\mathbf{v} + \mathcal{V}) = \text{Vol}(\mathcal{V})$. Thus, for any $\mathbf{v} \in \mathbb{R}^n$

$$\mathbf{X}|\mathbf{V} = \mathbf{v} \sim [\mathbf{v} + \mathbf{U}] \bmod \Lambda \sim \text{Uniform}(\mathcal{V}).$$

□

The Crypto Lemma ensures that $\frac{1}{n}\mathbb{E}\|\mathbf{X}\|^2 = (1-\epsilon)\mathsf{SNR}$, but our power constraint was $\|\mathbf{X}\|^2 \le n\mathsf{SNR}$. Since $\mathbf{X}$ is uniformly distributed over $\mathcal{V}_c$ and Λ_c is good for MSE quantization, Theorem 18.1 implies that $\|\mathbf{X}\|^2 \le n\mathsf{SNR}$ with high probability. Thus, whenever the power constraint is violated we can just transmit $\mathbf{0}$ instead of $\mathbf{X}$, and this will have a negligible effect on the error probability of the scheme.

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

$$\begin{aligned}\mathbf{Y}_{\text{eff}} &= [\alpha\mathbf{Y} - \mathbf{U}] \bmod \Lambda_c \\ &= [\mathbf{X} - \mathbf{U} + (\alpha - 1)\mathbf{X} + \alpha\mathbf{Z}] \bmod \Lambda_c \\ &= [\mathbf{t} + (\alpha - 1)\mathbf{X} + \alpha\mathbf{Z}] \bmod \Lambda_c \qquad (18.11) \\ &= [\mathbf{t} + \mathbf{Z}_{\text{eff}}] \bmod \Lambda_c, \qquad (18.12)\end{aligned}$$

where we have used the modulo distributive law in (18.11), and

$$\mathbf{Z}_{\text{eff}} = (\alpha - 1)\mathbf{X} + \alpha\mathbf{Z} \qquad (18.13)$$

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

$$\sigma^2_{\text{eff}}(\alpha) \triangleq \frac{1}{n}\mathbb{E}\|\mathbf{Z}_{\text{eff}}\|^2 < \alpha^2 + (1 - \alpha)^2\mathsf{SNR}. \qquad (18.14)$$

Since $\mathbf{Z}$ is semi norm-ergodic, and $\mathbf{X}$ is uniformly distributed over the Voronoi region of a lattice that is good for MSE quantization, Theorem 18.1 implies that $\mathbf{Z}_{\text{eff}}$ is semi norm-ergodic with effective variance $\sigma^2_{\text{eff}}(\alpha)$. Setting $\alpha = \mathsf{SNR}/(1+\mathsf{SNR})$, such as to minimize the upper bound on $\sigma^2_{\text{eff}}(\alpha)$ results in effective variance $\sigma^2_{\text{eff}} < \mathsf{SNR}/(1+\mathsf{SNR})$.

The receiver next computes

$$\begin{aligned}\hat{\mathbf{t}} &= [Q_{\Lambda_f}(\mathbf{Y}_{\text{eff}})] \bmod \Lambda_c \\ &= \left[Q_{\Lambda_f}(\mathbf{t} + \mathbf{Z}_{\text{eff}})\right] \bmod \Lambda_c, \qquad (18.15)\end{aligned}$$

and outputs the message corresponding to $\hat{\mathbf{t}}$ as its estimate. Since Λ_f is good for coding, $\mathbf{Z}_{\text{eff}}$ is semi norm-ergodic, and

$$\frac{r^2_{\text{eff}}(\Lambda_f)}{n} = (1 + \epsilon)\frac{\mathsf{SNR}}{1 + \mathsf{SNR}} > \sigma^2_{\text{eff}},$$

we have that $\Pr(\hat{\mathbf{t}} \neq \mathbf{t}) \to 0$ as the lattice dimension tends to infinity. Thus, we have proved the following.

Theorem 18.3. *There exist a coding scheme based on a nested lattice pair, that reliably achieves any rate below* $\frac{1}{2}\log(1 + \mathsf{SNR})$ *with lattice decoding for all additive semi norm-ergodic channels. In particular, if the additive noise is AWGN, this scheme is capacity achieving.*

Remark 18.3. In the Mod-Λ scheme the error probability does not depend on the chosen message, such that $P_{e,\max} = P_{e,\text{avg}}$. However, this required common randomness in the form of the dither $\mathbf{U}$. By a standard averaging argument it follows that there exist some fixed shift $\mathbf{u}$ that achieves the same, or better, $P_{e,\text{avg}}$. However, for a fixed shift the error probability is no longer independent of the chosen message.

18.4 Dirty Paper Coding

Assume now that the channel is

$$\mathbf{Y} = \mathbf{X} + \mathbf{S} + \mathbf{Z},$$

where $\mathbf{Z}$ is a unit variance semi norm-ergodic noise, $\mathbf{X}$ is subject to the same power constraint $\|\mathbf{X}\|^2 \leq n\mathsf{SNR}$ as before, and $\mathbf{S}$ is some arbitrary interference vector, known to the transmitter but *not* to the receiver.

Naively, one can think that the encoder can handle the interference $\mathbf{S}$ just by subtracting it from the transmitted codeword. However, if the codebook is designed to exactly meet the power constraint, after subtracting $\mathbf{S}$ the power constraint will be violated. Moreover, if $\|\mathbf{S}\|^2 > n\mathsf{SNR}$, this approach is just not feasible.

Using the Mod-Λ scheme, $\mathbf{S}$ can be cancelled out with no cost in performance. Specifically, instead of transmitting $\mathbf{X} = [\mathbf{t} + \mathbf{U}] \bmod \Lambda_c$, the transmitted signal in the presence of known interference will be

$$\mathbf{X} = [\mathbf{t} + \mathbf{U} - \alpha\mathbf{S}] \bmod \Lambda_c.$$

Clearly, the power constraint is not violated as $\mathbf{X} \sim \text{Uniform}(\mathcal{V}_c)$ due to the Crypto Lemma (now, $\mathbf{U}$ should also be independent of $\mathbf{S}$). The decoder is exactly the same as in the Mod-Λ scheme with no interference. It is easy to verify that the interference is completely cancelled out, and any rate below $\frac{1}{2}\log(1 + \mathsf{SNR})$ can still be achieved.

Remark 18.4. When $\mathbf{Z}$ is Gaussian and $\mathbf{S}$ is Gaussian there is a scheme based on random codes that can reliably achieve $\frac{1}{2}\log(1 + \mathsf{SNR})$. For arbitrary $\mathbf{S}$, to date, only lattice based coding schemes are known to achieve the interference free capacity. There are many more scenarios where lattice codes can reliably achieve better rates than the best known random coding schemes.

18.5 Construction of Good Nested Lattice Pairs

We now briefly describe a method for constructing nested lattice pairs. Our construction is based on starting with a linear code over a prime finite field, and embedding it periodically in $\mathbb{R}^n$ to form a lattice.

Definition 18.6 (p-ary Construction A)**.** Let p be a prime number, and let $\mathbf{F} \in \mathbb{Z}_p^{k\times n}$ be a $k \times n$ matrix whose entries are all members of the finite field $\mathbb{Z}_p$. The matrix $\mathbf{F}$ generates a linear p-ary code

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

The p-ary Construction A lattice induced by the matrix $\mathbf{F}$ is defined as

$$\Lambda(\mathbf{F}) \triangleq p^{-1}\mathcal{C}(\mathbf{F}) + \mathbb{Z}^n.$$

Note that any point in $\Lambda(\mathbf{F})$ can be decomposed as $\mathbf{x} = p^{-1}\mathbf{c} + \mathbf{a}$ for some $\mathbf{c} \in \mathcal{C}(\mathbf{F})$ (where we identify the elements of $\mathbb{Z}_p$ with the integers $[0, 1, \ldots, p-1]$) and $\mathbf{a} \in \mathbb{Z}^n$. Thus, for any $\mathbf{x}_1, \mathbf{x}_2 \in \Lambda(\mathbf{F})$ we have

$$\begin{aligned}
\mathbf{x}_1 + \mathbf{x}_2 &= p^{-1}(\mathbf{c}_1 + \mathbf{c}_2) + \mathbf{a}_1 + \mathbf{a}_2 \\
&= p^{-1}([\mathbf{c}_1 + \mathbf{c}_2] \bmod p + p\mathbf{a}) + \mathbf{a}_1 + \mathbf{a}_2 \\
&= p^{-1}\tilde{\mathbf{c}} + \tilde{\mathbf{a}} \\
&\in \Lambda(\mathbf{F})
\end{aligned}$$

where $\tilde{\mathbf{c}} = [\mathbf{c}_1 + \mathbf{c}_2] \bmod p \in \mathcal{C}(\mathbf{F})$ due to the linearity of $\mathcal{C}(\mathbf{F})$, and $\mathbf{a}$ and $\tilde{\mathbf{a}}$ are some vectors in $\mathbb{Z}^n$. It can be verified similarly that for any $\mathbf{x} \in \Lambda(\mathbf{F})$ it holds that $-\mathbf{x} \in \Lambda(\mathbf{F})$, and that if all codewords in $\mathcal{C}(\mathbf{F})$ are distinct, then $\Lambda(\mathbf{F})$ has a finite minimum distance. Thus, $\Lambda(\mathbf{F})$ is indeed a lattice. Moreover, if $\mathbf{F}$ is full-rank over $\mathbb{Z}_p$, then the number of distinct codewords in $\mathcal{C}(\mathbf{F})$ is p^k. Consequently, the number of lattice points in every integer shift of the unit cube is p^k, so the corresponding Voronoi region must satisfy $\mathrm{Vol}(\mathcal{V}) = p^{-k}$.

Similarly, we can construct a nested lattice pair from a linear code. Let $0 \le k' < k$ and let $\mathbf{F}'$ be the sub-matrix obtained by taking only the first k' rows of $\mathbf{F}$. The matrix $\mathbf{F}'$ generates a linear code $\mathcal{C}'(\mathbf{F}')$ that is nested in $\mathcal{C}(\mathbf{F})$, i.e., $\mathcal{C}'(\mathbf{F}') \subset \mathcal{C}(\mathbf{F})$. Consequently we have that $\Lambda(\mathbf{F}') \subset \Lambda(\mathbf{F})$, and the nesting ratio is

$$\Gamma(\Lambda(\mathbf{F}), \Lambda(\mathbf{F}')) = p^{\frac{k-k'}{n}}.$$

An advantage of this nested lattice construction for Voronoi constellations is that there is a very simple mapping between messages and codewords in $\mathcal{L} = \Lambda_f \cap \mathcal{V}_c$. Namely, we can index our set of $2^{nR} = p^{k-k'}$ messages by all vectors in $\mathbb{Z}_p^{k-k'}$. Then, for each message vector $\mathbf{w} \in \mathbb{Z}_p^{k-k'}$, the corresponding codeword in $\mathcal{L} = \Lambda(\mathbf{F}) \cap \mathcal{V}(\Lambda(\mathbf{F}'))$ is obtained by constructing the vector

$$\tilde{\mathbf{w}}^T = [\underbrace{0 \cdots 0}_{k' \text{ zeros}} \mathbf{w}^T] \in \mathbb{Z}_p^k, \tag{18.16}$$

and taking $\mathbf{t} = \mathbf{t}(\mathbf{w}) = \left[[\tilde{\mathbf{w}}^T \mathbf{F}] \bmod p\right] \bmod \Lambda(\mathbf{F}')$. Also, in order to specify the codebook $\mathcal{L}$, only the (finite field) generating matrix $\mathbf{F}$ is needed.

If we take the elements of $\mathbf{F}$ to be i.i.d. and uniform over $\mathbb{Z}_p$, we get a random ensemble of nested lattice codes. It can be shown that if p grows fast enough with the dimension n (taking $p = O(n^{(1+\epsilon)/2})$ suffices) almost all pairs in the ensemble have the property that both the fine and coarse lattice are good for both coding and for MSE quantization [OE15].

Disclaimer: *This text is a very brief and non-exhaustive survey of the applications of lattices in information theory. For a comprehensive treatment, see [Zam14].*