On Error Exponents of Modulo Lattice Additive Noise Channels

Tie Liu, Pierre Moulin, and Ralf Koetter *

May 28, 2004, Revised July 1 and October 21, 2005

Abstract

Modulo lattice additive noise (MLAN) channels appear in the analysis of structured binning codes for Costa's dirty-paper channel and of nested lattice codes for the additive white Gaussian noise (AWGN) channel. In this paper, we derive a new lower bound on the error exponents of the MLAN channel. With a proper choice of the shaping lattice and the scaling parameter, the new lower bound coincides with the random-coding lower bound on the error exponents of the AWGN channel at the same signal-to-noise ratio (SNR) in the sphere-packing and straight-line regions. This result implies that, at least for rates close to channel capacity, (1) writing on dirty paper is as reliable as writing on clean paper; and (2) lattice encoding and decoding suffer no loss of error exponents relative to the optimal codes (with maximum-likelihood decoding) for the AWGN channel.

Keywords: Additive white Gaussian noise channel, Costa's dirty-paper channel, error exponents, lattice decoding, modulo lattice additive noise channel, nested lattice codes

*This work was supported by NSF under ITR grants CCR 00-81268 and CCR 03-25924. The material in this paper was presented in part at the IEEE Information Theory Workshop, San Antonio, TX, October 2004. The authors are with the Department of Electrical and Computer Engineering at the University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA (e-mail: {tieliu, moulin}@ifp.uiuc.edu; koetter@uiuc.edu).

1 Introduction

Consider Costa's dirty-paper channel [1]

$$\mathbf{Y} = \mathbf{X} + \mathbf{S} + \mathbf{N} \quad (1)$$

where the channel input $\mathbf{X} = (X_1, \dots, X_n)$ satisfies the average power constraint

$$\mathbb{E} \left[\frac{1}{n} \sum_{i=1}^n X_i^2 \right] \leq P_X, \quad (2)$$

the interference $\mathbf{S} \sim \mathcal{N}(\mathbf{0}, P_S \mathbf{I}_n)$ and the noise $\mathbf{N} \sim \mathcal{N}(\mathbf{0}, P_N \mathbf{I}_n)$ are length- n white Gaussian vectors with zero mean and power P_S and P_N respectively ($\mathbf{I}_n$ is the $n \times n$ identity matrix), and $\mathbf{S}$ and $\mathbf{N}$ are statistically independent. The interference $\mathbf{S}$ is *noncausally* known to the transmitter, and this knowledge can be used to encode the message over the entire block. On the other hand, only the statistics of $\mathbf{S}$ are known to the receiver.

The capacity of Costa's dirty-paper channel cannot exceed that obtained when $\mathbf{S}$ is also known to the receiver. In the latter case, (1) reduces to the zero-interference AWGN channel

$$\mathbf{Y} = \mathbf{X} + \mathbf{N} \quad (3)$$

under the same average power constraint (2). Costa [1] proved the amazing result that the capacity of the dirty-paper channel (1) actually coincides with that of the zero-interference AWGN channel (3). That is, the lack of knowledge of the interference at the receiver does not cause any loss in channel capacity. A natural question to ask next is whether this lack of knowledge causes any degradation in channel reliability.

Costa's result [1] followed from a direct evaluation of a capacity formula for generic Gel'fand-Pinsker channels [2] and was based on a *random binning* argument. Zamir, Shamai and Erez [3] proposed a *structured binning* scheme which they showed is also capacity-achieving for Costa's dirty-paper channel. The main idea there is to use a *scaled-lattice* strategy to transform Costa's dirty-paper channel (1) into a modulo lattice additive noise (MLAN) channel and show that the capacity of the induced MLAN channel is asymptotically the same as that of the zero-interference AWGN channel (3). Interestingly, using essentially the same idea, Erez and Zamir [3, 4, 5, 6] cracked the long-standing problem of achieving capacity of AWGN channels with lattice encoding and decoding.

Furthermore, Erez and Zamir [5, 6] studied the error exponents of the MLAN channel and showed that they are lower-bounded by the Poltyrev exponents which were previously derived in the context of coding for the *unconstrained* AWGN channel [7]. However, the Poltyrev

exponent is strictly inferior to the random-coding lower bound on the error exponents of the AWGN channel at the same SNR for all rates below channel capacity. On the other hand, the random-coding lower bound on the error exponents of the AWGN channel is known to be tight in the sphere-packing region [8], [9, p. 338]. Therefore, determining the reliability function of the MLAN channel (and hence of Costa's dirty-paper channel) remained an open problem.

In this paper, we derive a new lower bound on the error exponents of the MLAN channel. By a data-processing argument, it is also a lower bound on the error exponents of Costa's dirty-paper channel. With a proper choice of the shaping lattice and the scaling parameter, the new lower bound *coincides* with the random-coding lower bound on the error exponents of the AWGN channel at the same SNR in the sphere-packing and straight-line regions. Therefore, at least for rates close to channel capacity, writing on dirty paper is as reliable as writing on clean paper.

Before aliasing, the effective noise in a MLAN channel is not strictly Gaussian but rather approaches a Gaussian distribution as the dimension of the shaping lattice tends to infinity (when the shaping lattice is appropriately chosen). As illustrated by Forney [10], this vanishing "non-Gaussianity" does not affect the channel capacity. It does, however, impact the error exponents because channel reliability is known to be determined by the large deviation of the channel law rather than by its limiting behavior. It turns out that this fact has important consequences for the optimal choice of the lattice-scaling parameter α . Selecting α according to the minimum mean-square error (MMSE) principle (this choice of α is then denoted by α_{MMSE}) is asymptotically optimal for reliable communication at the capacity limit. However, α_{MMSE} is strictly suboptimal in maximizing the new lower bound (on the error exponents) for all rates below channel capacity. The best error exponents are achieved by using lattice-scaling parameters determined by large-deviation analysis.

The rest of the paper is organized as follows. In Section 2, we formalize the transformation from Costa's dirty-paper channel to the MLAN channel and summarize the known results on the capacity and error exponents of the MLAN channel. In Section 3, we derive a new lower bound on the error exponents of the MLAN channel. In Section 4, we give some numerical examples to illustrate the new results. In Section 5, we extend our results to the AWGN channel with lattice encoding and decoding. Finally, we give some concluding remarks in Section 6.

2 The (Λ, α) -MLAN Channel

We first recall some notation and results from lattice theory. An n -dimensional real lattice Λ is a discrete additive subgroup of $\mathbb{R}^n$ defined as $\Lambda = \{\mathbf{u}\mathbf{G} : \mathbf{u} \in \mathbb{Z}^n\}$ where $\mathbf{G}$ is an $n \times n$ full-rank generator matrix and $\mathbf{u}$ is a vector with integer components. (All vectors in this paper are row vectors.) A basic Voronoi region $\mathcal{V}_0(\Lambda)$ is the set of points $\mathbf{x} \in \mathbb{R}^n$ closer to $\mathbf{0}$ than to any other point in Λ , i.e.,

$$\mathcal{V}_0(\Lambda) \stackrel{\text{def}}{=} \{\mathbf{x} : \|\mathbf{x}\| \leq \|\mathbf{x} - \boldsymbol{\lambda}\|, \forall \boldsymbol{\lambda} \in \Lambda\} \quad (4)$$

where ties can be broken in any systematic fashion such that $\mathcal{V}_0(\Lambda)$ includes one and only one representative from each coset of Λ in $\mathbb{R}^n$. The second moment per dimension associated with Λ is defined as

$$\sigma^2(\Lambda) \stackrel{\text{def}}{=} \frac{1}{nV(\Lambda)} \int_{\mathcal{V}_0(\Lambda)} \|\mathbf{x}\|^2 d\mathbf{x}. \quad (5)$$

Here, $V(\Lambda) \stackrel{\text{def}}{=} \text{Vol}(\mathcal{V}_0(\Lambda))$ is the volume of $\mathcal{V}_0(\Lambda)$, and $\|\mathbf{x}\|$ denotes the Euclidean norm of $\mathbf{x}$. The normalized second moment is defined as

$$G(\Lambda) \stackrel{\text{def}}{=} V(\Lambda)^{-2/n} \sigma^2(\Lambda) \geq (2\pi e)^{-1} \quad (6)$$

where $(2\pi e)^{-1}$ is the normalized second moment of an n -dimensional ball as $n \rightarrow \infty$.

The covering radius $r_{\text{cov}}(\Lambda)$ is the radius of the smallest n -dimensional ball centered at the origin that contains $\mathcal{V}_0(\Lambda)$. The effective radius $r_{\text{eff}}(\Lambda)$ is the radius of an n -dimensional ball whose volume is equal to $V(\Lambda)$. A lattice Λ is *good for covering* if the covering efficiency

$$\rho_{\text{cov}}(\Lambda) \stackrel{\text{def}}{=} \frac{r_{\text{cov}}(\Lambda)}{r_{\text{eff}}(\Lambda)} \rightarrow 1 \quad (7)$$

and is *good for mean-square error quantization* if $G(\Lambda) \rightarrow (2\pi e)^{-1}$ for sufficiently-large dimension of Λ . Following a result of Rogers, Zamir and Feder [11] showed that there exist lattices which are simultaneously good for covering and mean-square error quantization. They referred to such lattices as *Rogers-good*.

The direct product of two lattices Λ_1 and Λ_2 is defined as

$$\Lambda_1 \times \Lambda_2 = \{(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2) : \boldsymbol{\lambda}_1 \in \Lambda_1, \boldsymbol{\lambda}_2 \in \Lambda_2\}, \quad (8)$$

which results in a new lattice with basic Voronoi region $\mathcal{V}_0(\Lambda_1) \times \mathcal{V}_0(\Lambda_2)$ and covering radius

$$\sqrt{r_{\text{cov}}^2(\Lambda_1) + r_{\text{cov}}^2(\Lambda_2)}. \quad (9)$$

It can be shown that the covering efficiency of the k -fold direct product lattice Λ^k satisfies

$$\lim_{k \rightarrow \infty} \ln \rho_{\text{cov}}(\Lambda^k) = \ln \rho_{\text{cov}}(\Lambda) + \frac{1}{2} \ln(2\pi e G_n^*) + \frac{1}{2} \ln \left(1 + \frac{2}{n}\right) \quad (10)$$

where G_n^* denotes the normalized second moment of an n -dimensional ball.

Figure 1: MLAN channel transformation of Costa's dirty-paper channel $\mathbf{Y} = \mathbf{X} + \mathbf{S} + \mathbf{N}$.

2.1 MLAN Channel Transformation

Let Λ be an n -dimensional lattice with second moment per dimension P_X and let $Q_\Lambda(\cdot)$ be the corresponding (Euclidean) lattice quantizer. Let $\mathbf{U} \sim \text{Unif}(\mathcal{V}_0(\Lambda))$ be a dither vector uniformly distributed over $\mathcal{V}_0(\Lambda)$. Referring to Figure 1, consider the following modulo- Λ transmission scheme for Costa's dirty-paper channel (1).

- *Transmitter:* For any $\mathbf{v} \in \mathcal{V}_0(\Lambda)$, the transmitter sends

$$\mathbf{x} = [\mathbf{v} - \alpha \mathbf{s} - \mathbf{u}] \mod \Lambda \quad (11)$$

where $\alpha \in [0, 1]$ is a scaling parameter, and $\mathbf{x} \mod \Lambda \stackrel{\text{def}}{=} \mathbf{x} - Q_\Lambda(\mathbf{x})$.

- *Receiver:* The receiver computes

$$\mathbf{y}' = [\alpha \mathbf{y} + \mathbf{u}] \mod \Lambda. \quad (12)$$

Due to the uniform distribution of the dither vector $\mathbf{U}$, for any $\mathbf{v}$, the channel input $\mathbf{X}$ is also uniformly distributed over $\mathcal{V}_0(\Lambda)$ [6, Lemma 1], [10, Lemma 2]. Thus, the average transmitted power is P_X and the input constraint (2) is satisfied. The resulting channel is the (Λ, α) -MLAN channel defined below.

Lemma 1 ([4, 5, 6]) *The channel defined by (1), (11) and (12) is equivalent in distribution to the (Λ, α) -MLAN channel*

$$\mathbf{y}' = [\mathbf{v} + \mathbf{N}'_{\text{eff}}] \mod \Lambda \quad (13)$$

with

$$\mathbf{N}'_{\text{eff}} = [(1 - \alpha)\mathbf{U} + \alpha\mathbf{N}] \mod \Lambda. \quad (14)$$

2.2 Summary of Known Results

The capacity-achieving distribution for the (Λ, α) -MLAN channel (13) is $\mathbf{V} \sim \text{Unif}(\mathcal{V}_0(\Lambda))$. Define

$$C(\text{SNR}, \alpha) \stackrel{\text{def}}{=} \frac{1}{2} \ln \frac{\text{SNR}}{(1 - \alpha)^2 \text{SNR} + \alpha^2} \quad (15)$$

where $\text{SNR} \stackrel{\text{def}}{=} P_X/P_N$. The channel capacity (in nats per dimension) is given by

$$C_\Lambda(\text{SNR}, \alpha) = \frac{1}{n} I(\mathbf{Y}; \mathbf{Y}') \quad (16)$$

$$= \frac{1}{n} h(\mathbf{Y}') - \frac{1}{n} h(\mathbf{N}'_{\text{eff}}) \quad (17)$$

$$= \frac{1}{2} \ln \frac{P_X}{G(\Lambda)} - \frac{1}{n} h(\mathbf{N}'_{\text{eff}}) \quad (18)$$

$$\geq \frac{1}{2} \ln \frac{P_X}{G(\Lambda)} - \frac{1}{n} h(\mathbf{N}_{\text{eff}}) \quad (19)$$

$$\geq \frac{1}{2} \ln \frac{P_X}{\mathbb{E} \left[\frac{1}{n} \|\mathbf{N}_{\text{eff}}\|^2 \right]} - \frac{1}{2} \ln (2\pi e G(\Lambda)) \quad (20)$$

$$= C(\text{SNR}, \alpha) - \frac{1}{2} \ln (2\pi e G(\Lambda)) \quad (21)$$

where $\mathbf{N}_{\text{eff}} \stackrel{\text{def}}{=} (1 - \alpha)\mathbf{U} + \alpha\mathbf{N}$ is the effective noise of the (Λ, α) -MLAN channel (13) before the “mod Λ ” operation (the aliasing). Here, (18) follows from the uniform distribution of $\mathbf{Y}'$ over $\mathcal{V}_0(\Lambda)$; (19) follows from $h(\mathbf{N}'_{\text{eff}}) \leq h(\mathbf{N}_{\text{eff}})$ because the many-to-one mapping “mod Λ ” can only reduce the (differential) entropy; (20) follows from the fact that the entropy of $\mathbf{N}_{\text{eff}}$ is upper bounded by that of a white Gaussian vector with the same second moment [9, p. 372]; and (21) follows from the definition of $C(\text{SNR}, \alpha)$ in (15).

Next, α may be chosen to maximize $C(\text{SNR}, \alpha)$. The maximizing α is $\alpha_{\text{MMSE}} \stackrel{\text{def}}{=} \frac{\text{SNR}}{1 + \text{SNR}}$, and the corresponding value of $C(\text{SNR}, \alpha)$ is the zero-interference AWGN capacity

$$C_{\text{AWGN}}(\text{SNR}) \stackrel{\text{def}}{=} \frac{1}{2} \ln (1 + \text{SNR}). \quad (22)$$

Finally, choosing Λ to be Rogers-good so that $G(\Lambda) \downarrow (2\pi e)^{-1}$ as $n \rightarrow \infty$, we obtain $C_\Lambda(\text{SNR}, \alpha) \uparrow C_{\text{AWGN}}(\text{SNR})$. We conclude that the capacity of the MLAN channel asymptotically approaches that of the AWGN channel at the same SNR, in the limit as the lattice dimension n tends to infinity.

An estimation-theoretic explanation for the choice $\alpha = \alpha_{\text{MMSE}}$ was given by Forney [10]. Note that, with this choice of α , the effective noise $\mathbf{N}'_{\text{eff}}$ in the (Λ, α) -MLAN channel (13)

involves a uniform component $(1 - \alpha)\mathbf{U}$ and hence is not strictly Gaussian before aliasing. The lower bound on the right side of (20), on the other hand, is *asymptotically* tight because, when the shaping lattice Λ is chosen to be Rogers-good, the dither vector $\mathbf{U}$ uniformly distributed over $\mathcal{V}_0(\Lambda)$ approaches in entropy rate a white Gaussian vector with the same second moment [11].

Erez and Zamir [5, 6] also studied the error exponents of the MLAN channel. They showed that the error exponent $E_\Lambda(R; \text{SNR}, \alpha)$ of the (Λ, α) -MLAN channel (13) satisfies

$$E_\Lambda(R; \text{SNR}, \alpha) \geq E_P \left(e^{2(C(\text{SNR}, \alpha) - R - \zeta_2(\Lambda))} \right) - \zeta_1(\Lambda) \quad (23)$$

where $E_P(\cdot)$ is the Poltyrev exponent given by

$$2E_P(\mu) = \begin{cases} \mu - \ln(e\mu), & 1 \leq \mu \leq 2 \\ \ln \frac{e\mu}{4}, & 2 \leq \mu \leq 4 \\ \frac{\mu}{4}, & \mu \geq 4, \end{cases} \quad (24)$$

and $\zeta_1(\Lambda)$ and $\zeta_2(\Lambda)$ are defined as

$$\zeta_1(\Lambda) \stackrel{\text{def}}{=} \ln \rho_{\text{cov}}(\Lambda) + \frac{1}{2} \ln(2\pi e G_n^*) + \frac{1}{n} \quad (25)$$

and

$$\zeta_2(\Lambda) \stackrel{\text{def}}{=} \ln \rho_{\text{cov}}(\Lambda) + \frac{1}{2} \ln(2\pi e G(\Lambda)). \quad (26)$$

This succinct parametrization (24) of the Poltyrev exponent is due to Forney, Trott and Chung [12]¹. The parameter μ represents the “volume-to-noise ratio” of the channel. (The factor of 2 before $E_P(\mu)$ shows that everything should really be measured per two real dimensions.)

The error exponent $E_{\text{Costa}}(R; \text{SNR})$ of Costa’s dirty-paper channel (1) satisfies

$$E_{\text{Costa}}(R; \text{SNR}) \geq \sup_{\alpha} \sup_{\Lambda} E_\Lambda(R; \text{SNR}, \alpha) \quad (27)$$

$$\geq \sup_{\alpha} \sup_{\Lambda} \left\{ E_P \left(e^{2(C(\text{SNR}, \alpha) - R - \zeta_2(\Lambda))} \right) - \zeta_1(\Lambda) \right\} \quad (28)$$

$$= \sup_{\alpha} E_P \left(e^{2(C(\text{SNR}, \alpha) - R)} \right) \quad (29)$$

$$= E_P \left(e^{2(C_{\text{AWGN}}(\text{SNR}) - R)} \right) \quad (30)$$

where (27) follows from the data-processing argument; (28) follows from (23); (29) is attained by lattices which are Rogers-good; and (30) is uniquely attained by $\alpha = \alpha_{\text{MMSE}}$.

¹While the parametrization of (24) is indeed due to [12], Poltyrev’s parametrization [7] differs from that of [12] only by some multiplicative factors like 2 and 2π .

Recall from [9, pp. 338-343] that the random-coding lower bound on the error exponent $E_{\text{AWGN}}(R; \text{SNR})$ of the AWGN channel (3) is given by

$$E_{\text{AWGN}}(R; \text{SNR}) \geq E_{\text{AWGN}}^{\text{rc}}(e^{2(C_{\text{AWGN}}(\text{SNR})-R)}; \text{SNR}) \quad (31)$$

where

$$\begin{aligned} & 2E_{\text{AWGN}}^{\text{rc}}(\mu; \text{SNR}) \\ &= \frac{\text{SNR}}{2(\text{SNR}+1)} \left\{ (\text{SNR}+1+\mu) - (\text{SNR}+1-\mu) + \sqrt{1 + \frac{\text{SNR}+1}{\text{SNR}} \frac{4}{\text{SNR}+1-\mu}} \right\} \\ &+ \ln \left\{ \frac{\text{SNR}+1}{\mu} - \frac{\text{SNR}}{2\mu} (\text{SNR}+1-\mu) \left(\sqrt{1 + \frac{\text{SNR}+1}{\text{SNR}} \frac{4}{\text{SNR}+1-\mu}} - 1 \right) \right\} \end{aligned} \quad (32)$$

$$\text{for } 1 \leq \mu \leq \frac{2(\text{SNR}+1)}{\text{SNR}} \left(1 + \frac{\text{SNR}}{2} - \sqrt{1 + \frac{\text{SNR}^2}{4}} \right);$$

$$\begin{aligned} & 2E_{\text{AWGN}}^{\text{rc}}(\mu; \text{SNR}) \\ &= \left(1 + \frac{\text{SNR}}{2} - \sqrt{1 + \frac{\text{SNR}^2}{4}} \right) + \ln \left\{ \frac{\mu}{2(\text{SNR}+1)} \left(\sqrt{1 + \frac{\text{SNR}^2}{4}} + 1 \right) \right\} \end{aligned} \quad (33)$$

$$\text{for } \frac{2(\text{SNR}+1)}{\text{SNR}} \left(1 + \frac{\text{SNR}}{2} - \sqrt{1 + \frac{\text{SNR}^2}{4}} \right) \leq \mu \leq \frac{8(\text{SNR}+1)}{\text{SNR}^2} \left(\sqrt{1 + \frac{\text{SNR}^2}{4}} - 1 \right); \text{ and}$$

$$2E_{\text{AWGN}}^{\text{rc}}(\mu; \text{SNR}) = \frac{\text{SNR}}{2} \left(1 - \sqrt{1 - \frac{\mu}{1+\text{SNR}}} \right) \quad (34)$$

$$\text{for } \mu \geq \frac{8(\text{SNR}+1)}{\text{SNR}^2} \left(\sqrt{1 + \frac{\text{SNR}^2}{4}} - 1 \right).$$

Figure 2 compares the Poltyrev exponent $E_P(e^{2(C_{\text{AWGN}}(\text{SNR})-R)})$ with the random-coding lower bound $E_{\text{AWGN}}^{\text{rc}}(e^{2(C_{\text{AWGN}}(\text{SNR})-R)}; \text{SNR})$ on the error exponents of the AWGN channel at the same SNR. Clearly, the Poltyrev exponent is strictly inferior to the random-coding lower bound on the error exponents of the AWGN channel, and the gap is particularly large at low rates in the low-SNR regime.

Erez and Zamir [6] proved (23) by directly evaluating an error-exponent formula for generic modulo-additive noise channels with transmitter side information [13]. Note that, in their derivation, Erez and Zamir again used the Gaussian-type bounds, e.g., see [6, Appendix A]. However, these bounds might be loose because, while channel capacity is determined by the limiting distribution of the channel noise, what also matters for the error

Figure 2: Comparison of the Poltyrev exponent (solid lines) and the random-coding lower bound on the error exponents of the AWGN channel (dashed lines), both normalized by SNR. Diamonds (“ $\diamond$ ”) and circles (“ $\circ$ ”) separate the sphere-packing, straight-line and expurgation regions of the Poltyrev exponent and the random-coding lower bound on the error exponents of the AWGN channel, respectively.

exponents is how the noise distribution approaches that limit. The rationale behind using the Gaussian-type bounds offered in [6] seems to be only a computational concern: the error-exponent formula [13] is an adaptation of the Gallager bound [9] (to modulo-additive noise channels with transmitter side information), and the Gallager bound is known to be hard to evaluate for channels with memory, e.g., the MLAN channel, because it cannot be factored into single-letter expressions.

3 A New Lower Bound on the Error Exponents of the (Λ, α) -MLAN Channel

In this section, we derive a new lower bound on the error exponents of the MLAN channel. Unlike in [6], our derivation does not depend on any previous results on the error exponents of side-information channels. Instead, we shall start from first principles and proceed with an estimate of the distance distribution of the optimal codes. Such an approach makes it possible to analyze the error probability of the MLAN channel *geometrically* rather than by the Gallager bound. As we shall see, the geometry of the problem plays a central role in the analysis of the typical error events, therefore allowing for a large-deviation analysis directly in the high-dimensional space.

3.1 The Encoder and Decoder

A (k, R) -block code for the n -dimensional (Λ, α) -MLAN channel (13) is a set of $M = \lceil e^{knR} \rceil$ codewords

$$\{(\mathbf{v}_w^{(1)}, \dots, \mathbf{v}_w^{(k)}) : \mathbf{v}_w^{(i)} \in \mathcal{V}_0(\Lambda), i = 1, \dots, k; w = 1, \dots, M\} \quad (35)$$

where k is the block length and R is the transmission rate measured by nats *per dimension* (rather than by nats per channel use). When the sub-codeword $\mathbf{v}_w^{(i)}$ is input to the (Λ, α) -MLAN channel (13), the output is

$$\mathbf{y}'^{(i)} = \left[\mathbf{v}_w^{(i)} + \mathbf{N}_{\text{eff}}'^{(i)} \right] \mod \Lambda, \quad i = 1, \dots, k \quad (36)$$

where

$$\mathbf{N}_{\text{eff}}'^{(i)} \stackrel{\text{def}}{=} [(1 - \alpha)\mathbf{U}^{(i)} + \alpha\mathbf{N}^{(i)}] \mod \Lambda, \quad (37)$$

and $\mathbf{U}^{(i)}$ and $\mathbf{N}^{(i)}$ are independently and identically distributed as $\text{Unif}(\mathcal{V}_0(\Lambda))$ and $\mathcal{N}(\mathbf{0}, P_N \mathbf{I}_n)$, respectively.

Define $\mathbf{v}_w \stackrel{\text{def}}{=} (\mathbf{v}_w^{(1)}, \dots, \mathbf{v}_w^{(k)})$ and $\mathbf{y}' \stackrel{\text{def}}{=} (\mathbf{y}'^{(1)}, \dots, \mathbf{y}'^{(k)})$. The channel from $\mathbf{v}_w$ to $\mathbf{y}'$ can be equivalently viewed as a super (Λ^k, α) -MLAN channel where the shaping lattice Λ^k is the k -fold direct product of Λ . (Direct product of lattices are defined in (8).) The effective noise of the (Λ^k, α) -MLAN channel is

$$\mathbf{N}'_{\text{eff}} = [(1 - \alpha)\mathbf{U} + \alpha\mathbf{N}] \mod \Lambda^k \quad (38)$$

where $\mathbf{U} \stackrel{\text{def}}{=} (\mathbf{U}^{(1)}, \dots, \mathbf{U}^{(k)}) \sim \text{Unif}(\mathcal{V}_0(\Lambda^k))$ and $\mathbf{N} \stackrel{\text{def}}{=} (\mathbf{N}^{(1)}, \dots, \mathbf{N}^{(k)}) \sim \mathcal{N}(\mathbf{0}, P_N \mathbf{I}_{kn})$.

Our encoding and decoding scheme is as follows.

- *Encoder*: Map each message $w \in \{1, \dots, M\}$ to a codeword $\mathbf{v}_w$.
- *Decoder*: The estimated message is given by

$$\hat{w} = \arg \min_{w \in \{1, \dots, M\}} \left(\min_{\boldsymbol{\lambda} \in \Lambda^k} \|\mathbf{y}' - (\mathbf{v}_w + \boldsymbol{\lambda})\| \right). \quad (39)$$

In words, the decoder first decodes the received vector $\mathbf{y}'$ to the nearest (in Euclidean metric) codeword in an extended codebook

$$\tilde{\mathcal{C}} \stackrel{\text{def}}{=} \bigcup_{w=1}^M \{\mathbf{v}_w + \Lambda^k\}. \quad (40)$$

It then decides that the transmitted message is the one corresponding to the coset representative of the estimated codeword in $\mathcal{V}_0(\Lambda^k)$. Note that the decision rule (39) might be slightly suboptimal, but this will only strengthen our achievability results. An illustration of the above decoding procedure when $k = 1$, $n = 2$ and $M = 9$ is shown in Figure 3.

3.2 The Error Probability

The probability of decoding error $P_{e,w}$ given the transmitted message w is given by

$$P_{e,w} = \Pr[(\mathbf{v}_w + \mathbf{N}'_{\text{eff}}) \notin \mathcal{V}_{\tilde{\mathcal{C}}}(\mathbf{v}_w)] \quad (41)$$

where $\mathcal{V}_{\tilde{\mathcal{C}}}(\mathbf{v}_w)$ is a nearest neighbor decoding region of $\mathbf{v}_w$ in $\tilde{\mathcal{C}}$. The extended codebook $\tilde{\mathcal{C}}$ has an infinite number of codewords and hence an infinite rate. The right side of (41) thus reminds us of Poltyrev's notion of coding for the *unconstrained* AWGN channel [7] for which the decoding error probability is measured against the density of any infinite input

Figure 3: An illustration of the decoding procedure. Filled circles (“●”) denote the codewords within the basic Voronoi region $\mathcal{V}_0(\Lambda)$, and circles (“○”) denote their translations in the extended codebook $\tilde{\mathcal{C}}$. The number next to a codeword is the index of the associated message. Dotted lines separate the decoding regions of codewords in $\tilde{\mathcal{C}}$. The received vector $\mathbf{y}' = [\mathbf{v}_1 + \mathbf{N}'_{\text{eff}}] \bmod \Lambda$ correctly decodes to message 1.

constellation. The main difference is that in a MLAN channel the noise is additive (i.e., independent of the transmitted codeword) but generally non-Gaussian (due to the uniform component and the aliasing).

In fact, Poltyrev's technique [7] for estimating the decoding error probability is based on the distance distribution of the input code and can be used to upper bound the nearest neighbor decoding error probability of any unconstrained additive noise channel in which the noise is *spherically distributed*. To apply Poltyrev's technique [7] to analyze the decoding error probability (41), we first need to "sphericalize" the effective noise $\mathbf{N}'_{\text{eff}}$.

Lemma 2 *The decoding error probability (41) is bounded from above as*

$$P_{e,w} \leq \Pr[(\mathbf{v}_w + \mathbf{N}_{\text{eff}}) \notin \mathcal{V}_{\tilde{\mathcal{C}}}(\mathbf{v}_w)] \quad (42)$$

where $\mathbf{N}_{\text{eff}} = (1 - \alpha)\mathbf{U} + \alpha\mathbf{N}$ is the effective noise of the (Λ^k, α) -MLAN channel before aliasing.

Proof: We first note that

$$\{(\mathbf{v}_w + \mathbf{N}'_{\text{eff}}) \in \mathcal{V}_{\tilde{\mathcal{C}}}(\mathbf{v}_w)\} = \bigcup_{\boldsymbol{\lambda} \in \Lambda^k} \{(\mathbf{v}_w + \mathbf{N}_{\text{eff}}) \in \mathcal{V}_{\tilde{\mathcal{C}}}(\mathbf{v}_w + \boldsymbol{\lambda})\}. \quad (43)$$

Since $\mathbf{0}$ is always a lattice point, the statement of the lemma follows from the fact that $\{(\mathbf{v}_w + \mathbf{N}'_{\text{eff}}) \notin \mathcal{V}_{\tilde{\mathcal{C}}}(\mathbf{v}_w)\}$ is a subset of $\{(\mathbf{v}_w + \mathbf{N}_{\text{eff}}) \notin \mathcal{V}_{\tilde{\mathcal{C}}}(\mathbf{v}_w)\}$. $\square$

Lemma 3 ([6]) *Let $m = kn$ be the dimension of the code and let $\mathbf{B} \sim \text{Unif}(\mathcal{B}_m(\mathbf{0}, \sqrt{mP'_X}))$ be a random vector uniformly distributed over the m -dimensional ball of center $\mathbf{0}$ and radius $\sqrt{mP'_X}$ with $P'_X \stackrel{\text{def}}{=} r_{\text{cov}}^2(\Lambda^k)/m$. Assume $\mathbf{B}$ and $\mathbf{N}$ are statistically independent. Then*

$$\Pr[(\mathbf{v}_w + \mathbf{N}_{\text{eff}}) \notin \mathcal{V}_{\tilde{\mathcal{C}}}(\mathbf{v}_w)] \leq e^{m \ln \rho_{\text{cov}}(\Lambda^k)} \Pr[(\mathbf{v}_w + \mathbf{Z}) \notin \mathcal{V}_{\tilde{\mathcal{C}}}(\mathbf{v}_w)] \quad (44)$$

with

$$\mathbf{Z} \stackrel{\text{def}}{=} (1 - \alpha)\mathbf{B} + \alpha\mathbf{N}. \quad (45)$$

Note that the random vector $\mathbf{Z}$ defined in (45) is spherically distributed. Furthermore, if we suppose the shaping lattice Λ is Rogers-good, by (9) and (10) we have $\lim_{k \rightarrow \infty} \ln \rho_{\text{cov}}(\Lambda^k) = \varepsilon_1(\Lambda)$ and $P'_X = (1 + \varepsilon_2(\Lambda)) P_X$ where

$$\varepsilon_1(\Lambda) \stackrel{\text{def}}{=} \ln \rho_{\text{cov}}(\Lambda) + \frac{1}{2} \ln(2\pi e G_n^*) + \frac{1}{2} \ln \left(1 + \frac{2}{n}\right) \quad (46)$$

and

$$\varepsilon_2(\Lambda) \stackrel{\text{def}}{=} \frac{\rho_{\text{cov}}^2(\Lambda)G_n^*}{G(\Lambda)} \left(1 + \frac{2}{n}\right) - 1, \quad (47)$$

both of which approach zero as the lattice dimension $n \rightarrow \infty$. In this case, the “spherical” upper bound on the right side of (44) only incurs an asymptotically small increase in the noise power and an exponentially small boost in the decoding error probability. A rigorous proof of Lemma 3 can be found in [6, Appendix A].

3.3 Random-Coding Analysis

We now turn to the right side of (44) and derive a random-coding upper bound on

$$\lambda(\mathbf{v}_w) \stackrel{\text{def}}{=} \Pr[(\mathbf{v}_w + \mathbf{Z}) \notin \mathcal{V}_{\tilde{\mathcal{C}}}(\mathbf{v}_w)] \quad (48)$$

assuming that the codewords $\mathbf{v}_w$, $w = 1, \dots, M$, are independently and identically chosen according to $\text{Unif}(\mathcal{V}_o(\Lambda^k))$. Following the footsteps of Poltyrev [7], we have

$$\lambda(\mathbf{v}_w) \leq \Pr[(\mathbf{v}_w + \mathbf{Z}) \notin \mathcal{V}_{\tilde{\mathcal{C}}}(\mathbf{v}_w) \text{ and } \|\mathbf{Z}\| \leq d] + \Pr[\|\mathbf{Z}\| \geq d] \quad (49)$$

$$\leq \sum_{\{\mathbf{v} \in \tilde{\mathcal{C}} \setminus \{\mathbf{v}_w\} : \|\mathbf{v} - \mathbf{v}_w\| \leq 2d\}} \Pr[\mathbf{Z} \in \mathcal{D}_m(d, \|\mathbf{v} - \mathbf{v}_w\|)] + \Pr[\|\mathbf{Z}\| \geq d], \quad \forall d > 0 \quad (50)$$

where $\mathcal{D}_m(d, \rho)$ is the section of the m -dimensional ball $\mathcal{B}_m(\mathbf{0}, d)$ cut off by the hyperplane that slices $\mathcal{B}_m(\mathbf{0}, d)$ at distance $\rho/2$ from the center, and (50) follows from the union bound. Note that, in the right side of (50), we used the fact that $\mathbf{Z}$ is spherically distributed so the pairwise error probability $\Pr[\mathbf{Z} \in \mathcal{D}_m(d, \|\mathbf{v} - \mathbf{v}_w\|)]$ is only a function of the Euclidean distance between $\mathbf{v}$ and $\mathbf{v}_w$. We may then rewrite (50) using the distance distribution of $\tilde{\mathcal{C}}$ with respect to $\mathbf{v}_w$. For $\Delta > 0$, let $M_i(\mathbf{v}_w)$ be the number of codewords $\mathbf{v} \in \tilde{\mathcal{C}}$ such that $(i-1)\Delta < \|\mathbf{v} - \mathbf{v}_w\| \leq i\Delta$. We have

$$\lambda(\mathbf{v}_w) \leq \sum_{i=1}^{\lceil 2d/\Delta \rceil} M_i(\mathbf{v}_w) \Pr[\mathbf{Z} \in \mathcal{D}_m(d, (i-1)\Delta)] + \Pr[\|\mathbf{Z}\| \geq d]. \quad (51)$$

Since $\mathbf{v}_w$, $w = 1, \dots, M$, are independently and identically chosen according to $\text{Unif}(\mathcal{V}_0(\Lambda^k))$ and $\tilde{\mathcal{C}}$ is generated by tiling $\mathbb{R}^n$ with translations of $\mathcal{V}_0(\Lambda^k)$ relative to Λ^k , the ensemble average of $M_i(\mathbf{v}_w)$ is proportional to the volume of the spherical shell $\mathcal{B}_m(\mathbf{v}_w, i\Delta) \setminus \mathcal{B}_m(\mathbf{v}_w, (i-1)\Delta)$.

1) Δ). Thus, we have

$$\mathbb{E}[M_i(\mathbf{v}_w)] = \frac{M}{V(\Lambda^k)} \cdot \text{Vol}(\mathcal{B}_m(\mathbf{v}_w, i\Delta) \setminus \mathcal{B}_m(\mathbf{v}_w, (i-1)\Delta)) \quad (52)$$

$$\leq \frac{M}{V(\Lambda^k)} \cdot \frac{m\pi^{m/2}}{\Gamma(m/2+1)} (i\Delta)^{m-1} \Delta \quad (53)$$

$$= \frac{e^{m\delta} m\pi^{m/2}}{\Gamma(m/2+1)} (i\Delta)^{m-1} \Delta \quad (54)$$

with

$$\delta \stackrel{\text{def}}{=} \frac{1}{m} \ln \frac{M}{V(\Lambda^k)} \quad (55)$$

and the geometric expression $\frac{m\pi^{m/2}}{\Gamma(m/2+1)}$ being the surface area of a unit m -dimensional ball. Averaging (51) over the ensemble and letting $\Delta \rightarrow 0$, we obtain

$$\mathbb{E}[\lambda(\mathbf{v}_w)] \leq \frac{e^{m\delta} m\pi^{m/2}}{\Gamma(m/2+1)} \int_0^{2d} \rho^{m-1} \Pr[\mathbf{Z} \in \mathcal{D}_m(d, \rho)] d\rho + \Pr[\|\mathbf{Z}\| \geq d]. \quad (56)$$

We note that the right side of (56) is independent of the choice of message w , so it may also serve as an upper bound on the ensemble average of

$$\bar{\lambda} \stackrel{\text{def}}{=} \frac{1}{M} \sum_{w=1}^M \lambda(\mathbf{v}_w). \quad (57)$$

We thus have proved the following result.

Lemma 4 *There exists a (k, R) -block code for the (Λ, α) -MLAN channel (13) such that*

$$\bar{\lambda} \leq \frac{e^{m\delta} m\pi^{m/2}}{\Gamma(m/2+1)} \int_0^{2d} \rho^{m-1} \Pr[\mathbf{Z} \in \mathcal{D}_m(d, \rho)] d\rho + \Pr[\|\mathbf{Z}\| \geq d], \quad \forall d > 0. \quad (58)$$

The upper bound on the right side of (58) can be improved for small values of R by means of an expurgation procedure. The result is summarized in the following lemma.

Lemma 5 *There exists a (k, R) -block code for the (Λ, α) -MLAN channel (13) such that*

$$\bar{\lambda} \leq 16 \left(\frac{e^{m\delta'} m\pi^{m/2}}{\Gamma(m/2+1)} \int_{\sqrt{m\rho_0}}^{2d} \rho^{m-1} \Pr[\mathbf{Z} \in \mathcal{D}_m(d, \rho)] d\rho + \Pr[\|\mathbf{Z}\| \geq d] \right), \quad \forall d > 0. \quad (59)$$

Proof: See Appendix A. □

Next, we provide some results regarding the tail of the spherical noise vector $\mathbf{Z}$ defined in (45). As we shall see, the distribution of $\mathbf{Z}$ possesses rather different decay behavior than the exponentially-quadratic decay of Gaussian tails.

Lemma 6 *Let $f_Z(\mathbf{z})$ be the probability density function of $\mathbf{Z}$ defined in (45). We have*

$$f_Z(\sqrt{m}\mathbf{z}) \leq \frac{(m-1)\Gamma(m/2+1)}{\sqrt{4\pi\alpha^2 P_N}\Gamma((m+1)/2)} \exp\left[-(m-2)E_Z\left(\frac{\|\mathbf{z}\|^2}{\sigma'^2}; \text{SNR}', \alpha\right)\right], \quad \forall \mathbf{z} \in \mathbb{R}^m \quad (60)$$

with $\sigma'^2 \stackrel{\text{def}}{=} (1-\alpha)^2 P'_X + \alpha^2 P_N$ and $\text{SNR}' \stackrel{\text{def}}{=} P'_X/P_N$. The exponent $E_Z(\cdot)$ satisfies

$$2E_Z(\mu; \text{SNR}, \alpha) = \begin{cases} E_0(\mu; \text{SNR}, \alpha) + \ln(2\pi e\sigma'^2\mu), & 0 \leq \mu \leq \mu_0(\text{SNR}, \alpha) \\ E_{\text{sp}}(\mu; \text{SNR}, \alpha) + \ln(2\pi e\sigma'^2\mu), & \mu \geq \mu_0(\text{SNR}', \alpha) \end{cases} \quad (61)$$

where

$$E_0(\mu; \text{SNR}, \alpha) \stackrel{\text{def}}{=} \ln \frac{(1-\alpha)^2 \text{SNR}}{((1-\alpha)^2 \text{SNR} + \alpha^2)\mu}, \quad (62)$$

$$\mu_0(\text{SNR}, \alpha) \stackrel{\text{def}}{=} \max \left\{ \frac{(1-\alpha)^2 \text{SNR} - \alpha^2}{(1-\alpha)^2 \text{SNR} + \alpha^2}, 0 \right\}, \quad (63)$$

$$E_{\text{sp}}(\mu; \text{SNR}, \alpha) \stackrel{\text{def}}{=} \frac{(1-\alpha)^2 \text{SNR} + \alpha^2}{\alpha^2} \left(1 + \mu - \frac{2\mu}{g_1(\mu, \text{SNR}, \alpha)} \right) - \ln g_1(\mu, \text{SNR}, \alpha), \quad (64)$$

and

$$g_1(\mu; \text{SNR}, \alpha) \stackrel{\text{def}}{=} \frac{\sqrt{(4(1-\alpha)^4 \text{SNR}^2 + 4\alpha^2(1-\alpha)^2 \text{SNR})\mu + \alpha^4} - \alpha^2}{2(1-\alpha)^2 \text{SNR}}. \quad (65)$$

Proof: See Appendix B. □

Figure 4 displays $E_Z(\mu; \text{SNR}, \alpha)$, normalized by SNR, as a function of μ for different values of α . Note that, whereas a Gaussian vector ($\alpha = 1$) has a normal tail, a uniform vector over a ball ($\alpha = 0$) has no tail. For α between 0 and 1, the distribution of the spherical noise vector $\mathbf{Z}$ “improves” from a normal tail to “no-tail” as α decreases.

Lemma 7 *Let $\lambda_Z(\xi) \stackrel{\text{def}}{=} \mathbb{E} \left[e^{\xi \|\mathbf{Z}\|^2} \right]$ be the moment-generating function of $\|\mathbf{Z}\|^2$. Then*

$$\lambda_Z(\xi) \leq \exp \left[-m \left(\frac{1}{2} \ln(1 - 2\alpha^2 P_N \xi) - \frac{(1-\alpha)^2 P'_X \xi}{1 - 2\alpha^2 P_N \xi} \right) \right] \quad (66)$$

for any $0 \leq \xi < (2\alpha^2 P_N)^{-1}$.

Figure 4: Exponent $E_Z(\mu; \text{SNR}, \alpha)$ of (61), normalized by SNR, as a function of μ for different values of α .

Proof: See Appendix C. □

Combine Lemmas 2, 3, 4 and 5. By investigating the asymptotic behavior of the exponents of the right side of (56) and (59) (with the help of Lemmas 6 and 7), we obtain a new lower bound on the error exponents of the (Λ, α) -MLAN channel (13) stated in the following theorem.

Theorem 8 *The error exponent $E_\Lambda(R; \alpha)$ of the (Λ, α) -MLAN channel (13) satisfies*

$$E_\Lambda(R; \alpha) \geq E\left(e^{2(C(\text{SNR}', \alpha) - R - \varepsilon_1(\Lambda))}; \text{SNR}', \alpha\right) - \varepsilon_1(\Lambda) \quad (67)$$

where $C(\cdot)$ is defined in (15), $\text{SNR}' \stackrel{\text{def}}{=} (1 + \varepsilon_2(\Lambda))\text{SNR}$, and $\varepsilon_1(\cdot)$ and $\varepsilon_2(\cdot)$ are defined in (46) and (47), respectively. The exponent $E(\cdot)$ satisfies

$$2E(\mu; \text{SNR}, \alpha) = \begin{cases} E_{\text{sp}}(\mu; \text{SNR}, \alpha), & 1 \leq \mu \leq \mu_{\text{cr}}(\text{SNR}, \alpha) \\ E_{\text{sl}}(\mu; \text{SNR}, \alpha), & \mu_{\text{cr}}(\text{SNR}, \alpha) \leq \mu \leq 2\mu_{\text{cr}}(\text{SNR}, \alpha) \\ E_{\text{ex}}(\mu; \text{SNR}, \alpha), & \mu \geq 2\mu_{\text{cr}}(\text{SNR}, \alpha) \end{cases} \quad (68)$$

where $E_{\text{sp}}(\cdot)$ is defined in (64),

$$E_{\text{sl}}(\mu; \text{SNR}, \alpha) \stackrel{\text{def}}{=} E_{\text{sp}}(\mu_{\text{cr}}(\text{SNR}, \alpha); \text{SNR}, \alpha) - \ln \frac{\mu_{\text{cr}}(\text{SNR}, \alpha)}{\mu}, \quad (69)$$

$$E_{\text{ex}}(\mu; \text{SNR}, \alpha) \stackrel{\text{def}}{=} E_{\text{sp}}(g_2(\mu; \text{SNR}, \alpha); \text{SNR}, \alpha) - \ln \left(1 - \frac{\mu}{4g_2(\mu; \text{SNR}, \alpha)}\right), \quad (70)$$

$$\mu_{\text{cr}}(\text{SNR}, \alpha) \stackrel{\text{def}}{=} \frac{(1 - \alpha)^2 \text{SNR} + 3\alpha^2 + \sqrt{((1 - \alpha)^2 \text{SNR} + 3\alpha^2)^2 - 8\alpha^4}}{2((1 - \alpha)^2 \text{SNR} + \alpha^2)}, \quad (71)$$

and $g_2(\mu; \text{SNR}, \alpha)$ is the unique zero of the function

$$f(x) = \sqrt{(4(1 - \alpha)^4 \text{SNR}^2 + 4\alpha^2(1 - \alpha)^2 \text{SNR})x + \alpha^4 + \alpha^2} - 2((1 - \alpha)^2 \text{SNR} + \alpha^2)x + \frac{2\alpha^2 \mu}{4x - \mu}, \quad \frac{\mu}{4} \leq x \leq \frac{\mu}{2}, \quad \mu \geq 2\mu_{\text{cr}}. \quad (72)$$

Proof: See Appendix D. □

Proposition 9 *The error exponent $E_{\text{Costa}}(R; \text{SNR})$ of the Costa's dirty-paper channel is lower-bounded by the random-coding lower bound $E_{\text{AWGN}}^{\text{rc}}(R; \text{SNR})$ on the error exponents of the AWGN channel at the same SNR in the sphere-packing and straight-line regions, i.e.,*

$$E_{\text{Costa}}(R; \text{SNR}) \geq E_{\text{AWGN}}^{\text{rc}}(R; \text{SNR}), \quad R_{\text{AWGN}}^{\text{ex}}(\text{SNR}) \leq R \leq C_{\text{AWGN}}(\text{SNR}) \quad (73)$$

where

$$R_{\text{AWGN}}^{\text{ex}}(\text{SNR}) \stackrel{\text{def}}{=} \frac{1}{2} \ln \left(\frac{1}{2} + \frac{1}{2} \sqrt{1 + \frac{\text{SNR}^2}{4}} \right). \quad (74)$$

Proof: We have

$$E_{\text{Costa}}(R; \text{SNR}) \geq \sup_{\alpha} \sup_{\Lambda} E_{\Lambda}(R; \text{SNR}, \alpha) \quad (75)$$

$$\geq \sup_{\alpha} \sup_{\Lambda} \left\{ E \left(e^{2(C(\text{SNR}', \alpha) - R - \varepsilon_1(\Lambda))}; \text{SNR}', \alpha \right) - \varepsilon_1(\Lambda) \right\} \quad (76)$$

$$= \sup_{\alpha} E \left(e^{2(C(\text{SNR}, \alpha) - R)}; \text{SNR}, \alpha \right) \quad (77)$$

where (76) follows from Theorem 8, and (77) follows from the facts that $\varepsilon_1(\Lambda) \rightarrow 0$ and $\text{SNR}' \rightarrow \text{SNR}$ with Λ chosen to be Rogers-good. The desired result (73) thus follows from the explicit solution to the optimization problem on the right side of (77). The optimal lattice-scaling parameter α_{LD} (the subscript “LD” stands for “Large Deviation”) is given by

$$\alpha_{\text{LD}} = \begin{cases} 1 + \frac{\text{SNR}}{2} - \sqrt{1 + \frac{\text{SNR}^2}{4}}, & R_{\text{AWGN}}^{\text{ex}}(\text{SNR}) \leq R \leq R_{\text{AWGN}}^{\text{cr}}(\text{SNR}) \\ \sqrt{\frac{\beta^2}{4} + \beta} - \frac{\beta}{2}, & R_{\text{AWGN}}^{\text{cr}}(\text{SNR}) \leq R \leq C_{\text{AWGN}}(\text{SNR}) \end{cases} \quad (78)$$

where

$$R_{\text{AWGN}}^{\text{cr}}(\text{SNR}) \stackrel{\text{def}}{=} \frac{1}{2} \ln \left(\frac{1}{2} + \frac{\text{SNR}}{4} + \frac{1}{2} \sqrt{1 + \frac{\text{SNR}^2}{4}} \right) \quad (79)$$

and $\beta \stackrel{\text{def}}{=} \text{SNR}(1 - e^{-2R})$. □

3.4 On the Achievability of the Poltyrev Exponents

Here, we comment on the achievability of the Poltyrev exponents for the MLAN channel. In our derivation of the new lower bound (67), Lemmas 2 and 3 are used to connect the error probability of the (Λ^k, α) -MLAN channel to that of an unconstrained additive noise channel in which the noise is a weighted sum of a white Gaussian and a spherically-uniform vector. It is very tempting to go further down that road and connect the error probability of the (Λ^k, α) -MLAN channel to that of an unconstrained AWGN channel. The following lemma establishes the connection.

Lemma 10 Let $\mathbf{G} \sim \mathcal{N}(\mathbf{0}, P_X \mathbf{I}_m)$ be statistically independent of $\mathbf{N} \sim \mathcal{N}(\mathbf{0}, P_N \mathbf{I}_m)$. For $\mathbf{Z}$ defined in (45), we have

$$\Pr[(\mathbf{v}_w + \mathbf{Z}) \notin \mathcal{V}_{\tilde{\mathcal{C}}}(\mathbf{v}_w)] \leq e^{m\varepsilon_3(\Lambda^k)} \Pr[(\mathbf{v}_w + \mathbf{N}_1) \notin \mathcal{V}_{\tilde{\mathcal{C}}}(\mathbf{v}_w)] \quad (80)$$

where

$$\mathbf{N}_1 \stackrel{\text{def}}{=} (1 - \alpha)\mathbf{G} + \alpha\mathbf{N} \quad (81)$$

and

$$\lim_{k \rightarrow \infty} \varepsilon_3(\Lambda^k) = \frac{1}{2} \left(\frac{P'_X}{P_X} - \ln \frac{P'_X}{P_X} - 1 \right). \quad (82)$$

Now (80) is a Gaussian-type bound in that, in contrast to $\mathbf{Z}$, $\mathbf{N}_1$ is strictly Gaussian. The achievability of the Poltyrev exponents for the MLAN channel thus follows from Lemmas 2, 3, 10 and Poltyrev's results on the error exponents of the unconstrained AWGN channel [7]. This gives an alternative proof of Erez and Zamir's result on the achievability of the Poltyrev exponents for the MLAN channel. However, the use of Gaussian-type bounds is in the same spirit.

4 Numerical Examples and Discussion

In this section, we provide some numerical examples to illustrate the results of Section 3. In Figures 5-7, we plot the exponents $E_P(e^{2(C(\text{SNR}, \alpha_{\text{MMSE}}) - R)})$, $E_{\text{AWGN}}^{\text{rc}}(e^{2(C_{\text{AWGN}}(\text{SNR}) - R)}; \text{SNR})$, $E(e^{2(C(\text{SNR}, \alpha_{\text{MMSE}}) - R)}; \text{SNR}, \alpha_{\text{MMSE}})$ and $E(e^{2(C(\text{SNR}, \alpha_{\text{LD}}) - R)}; \text{SNR}, \alpha_{\text{LD}})$, all normalized by SNR, as a function of R for SNR = -10, 0, 10 dB, respectively. We have also plotted α_{LD} , normalized by α_{MMSE} , as a function of R for the same SNRs. A few observations and remarks are now in order.

1. Fix $\alpha = \alpha_{\text{MMSE}}$. We observe from these examples that $E_P(e^{2(C(\text{SNR}, \alpha_{\text{MMSE}}) - R)})$ is strictly smaller than $E(e^{2(C(\text{SNR}, \alpha_{\text{MMSE}}) - R)}; \text{SNR}, \alpha_{\text{MMSE}})$ for all rates below channel capacity. Therefore, Erez and Zamir's conjecture (in a preliminary version of [6]) on the asymptotic optimality of the Poltyrev exponents for the MLAN channel is *not* true.
2. In the high-SNR regime, $E_{\Lambda}(e^{2(C(\text{SNR}, \alpha_{\text{LD}}) - R)}; \text{SNR}, \alpha_{\text{LD}}) \approx E_P(e^{2(C(\text{SNR}, \alpha_{\text{MMSE}}) - R)})$ for all rates below channel capacity (e.g., see Figure 7). This suggests that the MLAN channel (with a proper choice of the shaping lattice and the scaling parameter) is asymptotically equivalent to Poltyrev's unconstrained AWGN channel at the same volume-to-noise ratio in the limit as SNR tends to infinity. The reason is that, in the high-SNR regime, the optimal scaling parameter $\alpha_{\text{LD}} \approx 1$ and the effective noise becomes Gaussian before aliasing.

Figure 5: (a) Error exponents, normalized by SNR, as a function of R for $\text{SNR} = -10$ dB; (b) The optimal lattice-scaling parameter α_{LD} , normalized by α_{MMSE} , as a function of R for $\text{SNR} = -10$ dB.

Figure 6: (a) Error exponents, normalized by SNR, as a function of R for $\text{SNR} = 0$ dB; (b) The optimal lattice-scaling parameter α_{LD} , normalized by α_{MMSE} , as a function of R for $\text{SNR} = 0$ dB.

Figure 7: (a) Error exponents, normalized by SNR, as a function of R for $\text{SNR} = 10$ dB; (b) The optimal lattice-scaling parameter α_{LD} , normalized by α_{MMSE} , as a function of R for $\text{SNR} = 10$ dB.

3. α_{MMSE} is strictly suboptimal in maximizing $E(e^{2(C(\text{SNR}, \alpha) - R)}; \text{SNR}, \alpha)$ except for the rate equal to channel capacity. This is because, when the Gaussian-type bounds are used, α affects the Poltyrev exponent $E_P(e^{2(C(\text{SNR}, \alpha) - R)})$ only through the variance of the Gaussian noise (or, equivalently, the volume-to-noise ratio $e^{2(C(\text{SNR}, \alpha) - R)}$) for which α_{MMSE} is the unique minimizer. The new lower bound, on the other hand, takes into account the tail heaviness of the effective noise which is also controlled by the scaling parameter α (e.g, see Figure 4). The deviation of α_{LD} from α_{MMSE} indicates that there is a tradeoff between tail heaviness and variance of the noise in optimizing the error exponent of the MLAN channel. Whereas α_{MMSE} is second-moment optimal, α_{LD} is *large-deviation* optimal. Note that, while a Gaussian distribution has a normal tail, a uniform distribution over a ball has no tail. One would thus expect α to be smaller to favor the latter as we balance the large-deviation exponents. When the transmission rate is approaching the channel capacity, we begin to exit the large-deviation regime and enter the central-limit-theorem regime. “Large-deviation optimal” is replaced by “second-moment optimal” (which implies “mutual-information optimal” in this case), and α_{LD} approaches α_{MMSE} .
4. This tail-heaviness/variance tradeoff reminds us of the Gaussian arbitrarily-varying channel in which the worst-case noise is equivalent in distribution (induced by the stochastic encoder/decoder) to the sum of a white Gaussian and a uniform (over the surface of a ball) vector. A surprising result of [14] is that the error exponents of a Gaussian arbitrarily-varying channel are actually better than those of the AWGN channel at the same SNR.
5. Even though $E(e^{2(C(\text{SNR}, \alpha_{\text{LD}}) - R)}; \text{SNR}, \alpha_{\text{LD}}) = E_{\text{AWGN}}^{\text{rc}}(e^{2(C_{\text{AWGN}}(\text{SNR}) - R)}; \text{SNR})$ in the sphere-packing and straight-line regions, a gap remains in the expurgation region between the new lower bound and the random-coding lower bound on the error exponents of the AWGN channel at the same SNR. We suspect that this gap is inherent to the “inflation-receiving” scheme and can only be bridged by exploring more complicated receiving schemes, possibly involving the local geometry of the input code.

5 Lattice Encoding and Decoding for AWGN Channels

Motivated by their structured binning scheme for Costa’s dirty-paper channel, Erez and Zamir [4, 5, 6] showed that nested lattice codes in conjunction with lattice decoding can achieve capacity of the AWGN channel, thus cracking the long-standing open problem of achieving capacity of AWGN channels with lattice encoding and decoding. In this section, we extend our results to AWGN channels and show that Erez and Zamir’s lattice encoding and decoding scheme not only is capacity-achieving, but also is error-exponent-lossless relative to

Figure 8: MLAN channel transformation of the AWGN channel $\mathbf{Y} = \mathbf{X} + \mathbf{N}$.

the optimal codes (with maximum-likelihood decoding) for rates sufficiently close to channel capacity.

The key idea of [4, 5, 6] is to use an inflated-lattice strategy to transform the AWGN channel (3) into a MLAN channel. A diagram of the transformation scheme is shown in Figure 8. The similarity between Figures 8 and 1 is obvious. The resulting (Λ, α) -MLAN channel is again given by (13). Recall that our analysis of the error exponents of the (Λ, α) -MLAN channel (13) takes on a random-code ensemble in which the codewords are independently identically chosen according to a uniform distribution over the basic Voronoi region $\mathcal{V}_0(\Lambda)$ of the shaping lattice Λ . What makes the MLAN transformation interesting is that the same random-coding performance can be attained by the more structured *nested lattice codes*.

A lattice Λ (the coarse lattice) is nested in Λ_1 (the fine lattice) if $\Lambda \subseteq \Lambda_1$, i.e., if Λ is a sublattice of Λ_1 . The set of coset leaders of Λ relative to Λ_1 ,

$$\mathcal{C} \stackrel{\text{def}}{=} \{\Lambda_1 \cap \mathcal{V}_0(\Lambda)\}, \quad (83)$$

is called a nested lattice code. The rate of the nested lattice code is

$$R = \frac{1}{n} \ln |\mathcal{C}| = \frac{1}{n} \ln \frac{V(\Lambda)}{V(\Lambda_1)}. \quad (84)$$

When the nested lattice code $\mathcal{C}$ is used, the extended codebook $\tilde{\mathcal{C}}$ becomes the fine lattice Λ_1 . The decoding rule (39) is equivalent to producing an estimate

$$\hat{\mathbf{c}} = Q_{\Lambda_1}(\mathbf{y}') \bmod \Lambda \quad (85)$$

for the transmitted codeword $\mathbf{c}$. Note that (85) describes a (minimum-Euclidean-distance) lattice decoder which finds the closest lattice point, ignoring the boundary of the code. Such an unconstrained search preserves the lattice symmetry in the decoding process and reduces complexity. An ensemble of “good” nested lattice codes can be constructed using the following steps [15, 16]:

1. Let p be a prime number. Draw a generating vector $\mathbf{g} = (g_1, \dots, g_n)$ in which g_i , $i = 1, \dots, n$, are independently identically chosen according to $\text{Unif}(\{0, \dots, p-1\})$.
2. Define the codebook

$$\mathcal{C} \stackrel{\text{def}}{=} \{\mathbf{x} \in \mathbb{Z}_p^n : \mathbf{x} = q\mathbf{g} \pmod{p}, q = 0, \dots, p-1\}. \quad (86)$$

3. Apply Construction A [16] to lift $\mathcal{C}$ to $\mathbb{R}^n$ and form the lattice

$$\Lambda'_1 = p^{-1}\mathcal{C} + \mathbb{Z}^n. \quad (87)$$

4. Note that the cubic lattice $\mathbb{Z}^n$ may be viewed as nested in Λ'_1 . Let $\mathbf{G}$ be the generator matrix of a Rogers-good lattice. Apply the linear transformation $\mathbf{G}$ to Λ'_1 . It follows that $\Lambda = \mathbb{Z}^n \mathbf{G}$ is a sublattice of $\Lambda_1 = \Lambda'_1 \mathbf{G}$.

Since linear transformations preserve the nesting ratio, the coding rate (84) is

$$R = \frac{1}{n} \ln \frac{V(\mathbb{Z}^n)}{V(\Lambda'_1)} = \frac{1}{n} \ln p. \quad (88)$$

For a given rate R , we therefore must choose $p = \lceil e^{nR} \rceil_p$ where $\lceil \cdot \rceil_p$ denotes the operation of ceiling to the smallest prime number.

Note that p is exponentially increasing with the dimension n . For large p , the resulting ensemble is “matched” to the (Λ, α) -MLAN channel in that the codewords of the nested code become uniform over the basic Voronoi region $\mathcal{V}_0(\Lambda)$ [15]. Hence, a typical member of the ensemble approaches the optimal random-coding exponents of this channel [6, Appendix C]. In light of Theorem 8 and Proposition 9, we conclude that lattice encoding and decoding suffers no loss of error exponents relative to the optimal codes (with maximum-likelihood decoding) for rates sufficiently close to channel capacity.

6 Concluding Remarks

We derived a new lower bound on the error exponents of the MLAN channel. Whereas Erez and Zamir derived the Poltyrev exponents as a lower bound on the error exponents of their scheme, we established the new lower bound using Poltyrev’s random-coding bounds as a starting point. (Our development thus gives a concise rationale for why α_{MMSE} is the unique value of α that maximizes the Erez-Zamir-Poltyrev exponents.) The new lower bound is obtained by seeking the tradeoff between tail heaviness and variance of the effective noise that

maximizes the error exponents of the MLAN channel. As a consequence, the optimal lattice-scaling parameter α_{LD} becomes rate-adaptive and is chosen according to the large-deviation principle. The fact that α_{LD} differs from α_{MMSE} is barely surprising considering that the MMSE estimator is optimum when the quantity to be optimized is mutual information (in a linear Gaussian channel), but not necessarily for other optimization problems.

With a proper choice of the shaping lattice and the scaling parameter, the new lower bound coincides with the random-coding lower bound on the error exponents of the AWGN channel at the same SNR in the sphere-packing and straight-line regions. Therefore, at least for rates close to channel capacity, (1) writing on dirty paper is as reliable as writing on clean paper; and (2) lattice encoding and decoding suffer no loss of exponents relative to the optimal codes (with maximum-likelihood decoding) for the AWGN channel.

Finally, we would like to point out that the main thing that is currently missing in this paper is an explanation for the surprising *zero loss* of error exponents (at least for rates close to channel capacity) of the MLAN channel transformation (with the optimal choice of the shaping lattice and the scaling parameter α). Our large-deviation analysis discovers this fact “by surprise”, i.e., by comparing the obtained expression with the optimal one. We believe that such a coincidence should have a more fundamental explanation! The pursuit of such an explanation is worthy of future research.

A Proof of Lemma 5

We start from a code ensemble of $M' = 2M$ codewords $\mathbf{v}_w$, $w = 1, \dots, M'$, independently and identically drawn according to $\text{Unif}(\mathcal{V}_o(\Lambda^k))$. Define

$$\tilde{\mathcal{C}}' \stackrel{\text{def}}{=} \bigcup_{w=1}^{M'} \{\mathbf{v}_w + \Lambda^k\} \quad (89)$$

and denote by $M_\rho(\mathbf{v}_w)$ the number of codewords $\mathbf{v} \in \tilde{\mathcal{C}}' \setminus \{\mathbf{v}_w\}$ such that $\|\mathbf{v} - \mathbf{v}_w\| < \sqrt{m\rho}$. Let ρ_0 satisfy the equation

$$\mathbb{E} \left[\frac{1}{M'} \sum_{w=1}^{M'} M_{\rho_0}(\mathbf{v}_w) \right] = 0.05. \quad (90)$$

Thus,

$$\rho_0 = \frac{1}{m} \left(\frac{0.05 \cdot V(\Lambda^k)}{M' \cdot \text{Vol}(\mathcal{B}_m(\mathbf{0}, 1))} \right)^{2/m}. \quad (91)$$

Given ρ_0 , define

$$\lambda_1(\mathbf{v}_w) \stackrel{\text{def}}{=} \Pr[(\mathbf{v}_w + \mathbf{Z}) \notin \mathcal{V}_{\tilde{\mathcal{C}}'}(\mathbf{v}_w) \text{ and } \|\mathbf{Z}\| < \sqrt{m\rho_0}] \quad (92)$$

and

$$\lambda_2(\mathbf{v}_w) \stackrel{\text{def}}{=} \Pr[(\mathbf{v}_w + \mathbf{Z}) \notin \mathcal{V}_{\tilde{C}'}(\mathbf{v}_w) \text{ and } \|\mathbf{Z}\| \geq \sqrt{m\rho_0}]. \quad (93)$$

Following the derivation of Lemma 4, we obtain

$$\mathbb{E} \left[\frac{1}{M'} \sum_{w=1}^{M'} \lambda_2(\mathbf{v}_w) \right] \leq \frac{e^{m\delta'} m \pi^{m/2}}{\Gamma(m/2 + 1)} \int_{\sqrt{m\rho_0}}^{2d} \rho^{m-1} \Pr[\mathbf{Z} \in \mathcal{D}_m(d, \rho)] d\rho + \Pr[\|\mathbf{Z}\| \geq d], \quad \forall d > 0 \quad (94)$$

where

$$\delta' = \delta + (1/m) \ln 2. \quad (95)$$

Define the events

$$A \stackrel{\text{def}}{=} \left\{ \frac{1}{M'} \sum_{w=1}^{M'} \lambda_2(\mathbf{v}_w) < 4 \left(\frac{e^{m\delta'} m \pi^{m/2}}{\Gamma(m/2 + 1)} \int_{\sqrt{m\rho_0}}^{2d} \rho^{m-1} \Pr[\mathbf{Z} \in \mathcal{D}_m(d, \rho)] d\rho + \Pr[\|\mathbf{Z}\| \geq d] \right) \right\} \quad (96)$$

and

$$B \stackrel{\text{def}}{=} \left\{ \frac{1}{M'} \sum_{w=1}^{M'} M_{\rho_0}(\mathbf{v}_w) < 0.2 \right\}. \quad (97)$$

By (90), (94) and the Chebyshev inequality, we have $\Pr[A^c] \leq 0.25$ and $\Pr[B^c] \leq 0.25$. It follows that

$$\Pr[A \cap B] = 1 - \Pr[A^c \cup B^c] \quad (98)$$

$$\geq 1 - \Pr[A^c] - \Pr[B^c] \quad (99)$$

$$\geq 0.5, \quad (100)$$

i.e., there exists a set of M' codewords $\mathbf{v}_w$, $w = 1, \dots, M'$, such that

$$\frac{1}{M'} \sum_{w=1}^{M'} \lambda_2(\mathbf{v}_w) < 4 \left(\frac{e^{m\delta'} m \pi^{m/2}}{\Gamma(m/2 + 1)} \int_{\sqrt{m\rho_0}}^{2d} \rho^{m-1} \Pr[\mathbf{Z} \in \mathcal{D}_m(d, \rho)] d\rho + \Pr[\|\mathbf{Z}\| \geq d] \right) \quad (101)$$

and

$$\frac{1}{M'} \sum_{w=1}^{M'} M_{\rho_0}(\mathbf{v}_w) < 0.2. \quad (102)$$

Similarly, applying the Chebyshev inequality to (101) and (102), we conclude that there exists at least a subset of $0.5M' = M$ codewords from $\mathbf{v}_w$, $w = 1, \dots, M'$, such that

$$\lambda_2(\mathbf{v}_w) < 16 \left(\frac{e^{m\delta'} m \pi^{m/2}}{\Gamma(m/2 + 1)} \int_{\sqrt{m\rho_0}}^{2d} \rho^{m-1} \Pr[\mathbf{Z} \in \mathcal{D}_m(d, \rho)] d\rho + \Pr[\|\mathbf{Z}\| \geq d] \right) \quad (103)$$

and

$$M_{\rho_0}(\mathbf{v}_w) < 0.8 \quad (104)$$

for every $\mathbf{v}_w$ in the chosen subset. Note that (104) implies that $\lambda_1(\mathbf{v}_w) = 0$ and hence $\lambda(\mathbf{v}_w) = \lambda_1(\mathbf{v}_w) + \lambda_2(\mathbf{v}_w) = \lambda_2(\mathbf{v}_w)$. This completes the proof of Lemma 5.

B Proof of Lemma 6

Denote by f_N and f_B the probability density function of random vectors $\mathbf{N} \sim \mathcal{N}(\mathbf{0}, P_N \mathbf{I})$ and $\mathbf{B} \sim \text{Unif}(\mathcal{B}_m(\mathbf{0}, \sqrt{mP'_X}))$, respectively. Since $\mathbf{Z} = (1 - \alpha)\mathbf{B} + \alpha\mathbf{N}$ in which $\mathbf{B}$ and $\mathbf{N}$ are statistically independent, we have

$$f_Z(\sqrt{m}\mathbf{z}) = \frac{1}{\alpha^m(1 - \alpha)^m} \int_{\mathbb{R}^m} f_N\left(\frac{\sqrt{m}\mathbf{z} - \mathbf{x}}{\alpha}\right) f_B\left(\frac{\mathbf{x}}{1 - \alpha}\right) d\mathbf{x} \quad (105)$$

$$= \frac{m^{m/2}}{\alpha^m(1 - \alpha)^m} \int_{\mathbb{R}^m} f_N\left(\frac{\sqrt{m}\mathbf{z} - \sqrt{m}\mathbf{x}}{\alpha}\right) f_B\left(\frac{\sqrt{m}\mathbf{x}}{1 - \alpha}\right) d\mathbf{x} \quad (106)$$

$$= \frac{\Gamma(m/2 + 1)}{(2\pi^2\alpha^2P_N(1 - \alpha)^2P'_X)^{m/2}} \int_{\mathcal{B}_m(0, \sqrt{(1 - \alpha)^2P'_X})} \exp\left[-\frac{m\|\mathbf{z} - \mathbf{x}\|^2}{2\alpha^2P_N}\right] d\mathbf{x} \quad (107)$$

where (106) follows from a linear change of variable from $\mathbf{x}$ to $\sqrt{m}\mathbf{x}$. Expanding the integral on the right side of (107) in spherical coordinates (refer to Figure 9), we obtain

$$\begin{aligned} & \int_{\mathcal{B}_m(0, \sqrt{(1 - \alpha)^2P'_X})} \exp\left[-\frac{m\|\mathbf{z} - \mathbf{x}\|^2}{2\alpha^2P_N}\right] d\mathbf{x} \\ &= \frac{(m - 1)\pi^{(m-1)/2}}{\Gamma((m + 1)/2)} \int_0^{\sqrt{(1 - \alpha)^2P'_X}} \int_0^\pi \exp\left[-m\frac{r^2 + \|\mathbf{z}\|^2 - 2r\|\mathbf{z}\|\cos\theta}{2\alpha^2P_N}\right] r^{m-1}(\sin\theta)^{m-2} dr d\theta. \end{aligned} \quad (108)$$

Now, define

$$A(r, \theta) \stackrel{\text{def}}{=} \frac{r^2 + \|\mathbf{z}\|^2 - 2r\|\mathbf{z}\|\cos\theta}{2\alpha^2P_N} - \ln(r \sin\theta). \quad (109)$$

The double integral on the right side of (108) can be bounded from above as

$$\begin{aligned} & \int_0^{\sqrt{(1 - \alpha)^2P'_X}} \int_0^\pi \exp\left[-m\frac{r^2 + \|\mathbf{z}\|^2 - 2r\|\mathbf{z}\|\cos\theta}{2\alpha^2P_N}\right] r^{m-1}(\sin\theta)^{m-2} dr d\theta \\ &\leq \sqrt{(1 - \alpha)^2P'_X} \int_0^{\sqrt{(1 - \alpha)^2P'_X}} \int_0^\pi \exp[-(m - 2)A(r, \theta)] dr d\theta \end{aligned} \quad (110)$$

$$\leq \pi(1 - \alpha)^2P'_X \exp\left[-(m - 2)\min_{(r, \theta)} A(r, \theta)\right] \quad (111)$$

where the minimization is over the set $\{(r, \theta) : 0 \leq r \leq \sqrt{(1 - \alpha)^2P'_X}, 0 \leq \theta \leq \pi\}$. Simple calculations yield

$$\min_{(r, \theta)} A(r, \theta) = \begin{cases} \frac{1}{2} \ln \frac{e}{\alpha^2P_N}, & 0 \leq \frac{\|\mathbf{z}\|^2}{\sigma'^2} \leq \mu_0(\text{SNR}', \alpha) \\ E_{\text{sp}}\left(\frac{\|\mathbf{z}\|^2}{\sigma'^2}; \text{SNR}', \alpha\right) + \frac{1}{2} \ln \frac{e\|\mathbf{z}\|^2}{\alpha^2P_N(1 - \alpha)^2P'_X}, & \frac{\|\mathbf{z}\|^2}{\sigma'^2} \geq \mu_0(\text{SNR}', \alpha) \end{cases} \quad (112)$$

where $\mu_0(\cdot)$ is defined in (63). Substituting (108), (111) and (112) into (107), we obtain the desired result (60). This completes the proof of Lemma 6.

Figure 9: Plane of cone of half-angle θ .

C Proof of Lemma 7

The moment-generating function $\lambda_Z(\cdot)$ for $\|\mathbf{Z}\|^2$ is given by

$$\lambda_Z(\xi) = \mathbb{E} \left[e^{\xi \|\mathbf{Z}\|^2} \right] \quad (113)$$

$$= \mathbb{E} \left[\mathbb{E} \left[e^{\xi \|\mathbf{Z}\|^2} \mid \mathbf{B} \right] \right] \quad (114)$$

$$= \mathbb{E} \left[\exp \left[-\frac{m}{2} \log_e (1 - 2\alpha^2 P_N \xi) + \frac{(1 - \alpha)^2 \xi}{1 - 2\alpha^2 P_N \xi} \|\mathbf{B}\|^2 \right] \right] \quad (115)$$

for any $\xi < (2\alpha^2 P_N)^{-1}$. Here, (115) follows from the fact that, conditioned on $\mathbf{B}$, $(\alpha^2 P_N)^{-1} \|\mathbf{Z}\|^2$ follows the noncentral Chi-square distribution

$$\chi_m^2 \left(\frac{(1 - \alpha)^2 \|\mathbf{B}\|^2}{\alpha^2 P_N} \right). \quad (116)$$

The desired result (66) thus follows from (115) and the fact that $\|\mathbf{B}\|^2 \leq mP'_X$ with probability one. This completes the proof of Lemma 7.

D Proof of Theorem 8

We now use Lemmas 6 and 7 (combined with some geometric analysis) to investigate the exponent of the right side of (56). Let $d = \sqrt{m\rho_d}$ for some $\rho_d > 0$ and define

$$I_1(d) \stackrel{\text{def}}{=} \Pr[\|\mathbf{Z}\| \geq d] \quad \text{and} \quad I_2(d) \stackrel{\text{def}}{=} \frac{e^{m\delta} m \pi^{m/2}}{\Gamma(m/2 + 1)} \int_0^{2d} \rho^{m-1} \Pr[\mathbf{Z} \in \mathcal{D}_m(d, \rho)] d\rho. \quad (117)$$

First, $I_1(d)$ can be bounded from above as

$$I_1(d) = \Pr[\|\mathbf{Z}\|^2 \geq m\rho_d] \quad (118)$$

$$\leq \exp[-\xi m\rho_d + \ln \lambda_Z(\xi)] \quad (119)$$

$$\leq \exp \left[-m \left(\xi\rho_d + \frac{1}{2} \ln(1 - 2\alpha^2 P_N \xi) - \frac{(1 - \alpha)^2 P'_X \xi}{1 - 2\alpha^2 P_N \xi} \right) \right] \quad (120)$$

for any $0 \leq \xi < (2\alpha^2 P_N)^{-1}$. Here, (119) follows from the Chernoff bound on $\|\mathbf{Z}\|^2$, and (120) follows from Lemma 7. Choosing ξ to minimize the right side of (120), we obtain

$$-\frac{2}{m} \ln I_1(d) \geq \begin{cases} E_{\text{sp}}\left(\frac{\rho_d}{\sigma'^2}; \text{SNR}', \alpha\right), & \frac{\rho_d}{\sigma'^2} \geq 1 \\ 0, & 0 < \frac{\rho_d}{\sigma'^2} \leq 1. \end{cases} \quad (121)$$

Second, following a change of variable from ρ to $\sqrt{m\rho}$, $I_2(d)$ can be written as

$$I_2(d) = \frac{e^{m\delta} m^{m/2} \pi^{m/2}}{\Gamma(m/2)} \int_0^{4\rho_d} \rho^{m/2-1} \Pr[\mathbf{Z} \in \mathcal{D}_m(\sqrt{m\rho_d}, \sqrt{m\rho})] d\rho. \quad (122)$$

Expanding $\Pr[\mathbf{Z} \in \mathcal{D}_m(\sqrt{m\rho_d}, \sqrt{m\rho})]$ in spherical coordinates yields

$$\begin{aligned} & \Pr[\mathbf{Z} \in \mathcal{D}_m(\sqrt{m\rho_d}, \sqrt{m\rho})] \\ &= \frac{(m-1)\pi^{(m-1)/2}}{\Gamma((m+1)/2)} \int_{\sqrt{m\rho}/4}^{\sqrt{m\rho_d}} \int_0^{\arccos \sqrt{m\rho/(4r^2)}} \varphi_Z(r) r^{m-1} (\sin \theta)^{m-2} dr d\theta \end{aligned} \quad (123)$$

$$= \frac{(m-1)m^{m/2}\pi^{(m-1)/2}}{\Gamma((m+1)/2)} \int_{\rho/4}^{\rho_d} \int_0^{\arccos \sqrt{\rho/(4r)}} \varphi_Z(\sqrt{mr}) r^{m/2-1} (\sin \theta)^{m-2} dr d\theta \quad (124)$$

$$\leq \frac{(m-1)m^{m/2}\pi^{(m-1)/2}}{\Gamma((m+1)/2)} \int_{\rho/4}^{\rho_d} \varphi_Z(\sqrt{mr}) (r - \rho/4)^{m/2-1} dr \quad (125)$$

where $\varphi_Z(r) \stackrel{\text{def}}{=} f_Z(\mathbf{x})$ for $r = \|\mathbf{x}\|$. Here, (124) follows from the change of variable from r to $\sqrt{mr}$, and (125) follows from the monotonicity of the $\sin \theta$ for $0 \leq \theta \leq \arccos \sqrt{\rho/(4r)}$.

Substituting (60) of Lemma 6 into (125), we obtain

$$\begin{aligned} & \Pr[\mathbf{Z} \in \mathcal{D}_m(\sqrt{m\rho_d}, \sqrt{m\rho})] \\ & \leq A_m \int_{\rho/4}^{\rho_d} \exp \left[-(m-2) \left(E_Z \left(\frac{r}{\sigma'^2} \right) - \frac{1}{2} \ln \left(r - \frac{\rho}{4} \right) \right) \right] dr \end{aligned} \quad (126)$$

$$\leq \rho_d A_m \exp \left[-(m-2) \min_{\rho/4 \leq r \leq \rho_d} \left\{ E_Z \left(\frac{r}{\sigma'^2} \right) - \frac{1}{2} \ln \left(r - \frac{\rho}{4} \right) \right\} \right] \quad (127)$$

where

$$A_m \stackrel{\text{def}}{=} \frac{(m-1)^2 m^{m/2} \pi^{m/2} \Gamma(m/2 + 1)}{2\alpha^2 P_N \Gamma^2((m+1)/2)}. \quad (128)$$

Substituting (127) into (122) gives

$$I_2(d) \leq B_m \int_0^{4\rho_d} \exp \left[-(m-2) \min_{0 \leq r \leq \rho_d} \left\{ E_Z \left(\frac{r}{\sigma'^2} \right) - \frac{1}{2} \ln \left(r\rho - \frac{\rho^2}{4} \right) \right\} \right] d\rho \quad (129)$$

$$\leq 4\rho_d B_m \exp \left[-(m-2) \min_{0 \leq \rho \leq 4\rho_d} \min_{\rho/4 \leq r \leq \rho_d} \left\{ E_Z \left(\frac{r}{\sigma'^2} \right) - \frac{1}{2} \ln \left(r\rho - \frac{\rho^2}{4} \right) \right\} \right] \quad (130)$$

$$= 4\rho_d B_m \exp \left[-(m-2) \min_{0 \leq r \leq \rho_d} \min_{0 \leq \rho \leq 4r} \left\{ E_Z \left(\frac{r}{\sigma'^2} \right) - \frac{1}{2} \ln \left(r\rho - \frac{\rho^2}{4} \right) \right\} \right] \quad (131)$$

$$= 4\rho_d B_m \exp \left[-(m-2) \min_{0 \leq r \leq \rho_d} \left\{ E_Z \left(\frac{r}{\sigma'^2} \right) - \frac{1}{2} \ln \max_{0 \leq \rho \leq 4r} \left\{ r\rho - \frac{\rho^2}{4} \right\} \right\} \right] \quad (132)$$

$$= 4\rho_d B_m \exp \left[-(m-2) \min_{0 \leq r \leq \rho_d} \left\{ E_Z \left(\frac{r}{\sigma'^2} \right) - \ln r \right\} \right] \quad (133)$$

where

$$B_m \stackrel{\text{def}}{=} \frac{\rho_d (m-1)^2 e^{m\delta} m^{m+1} \pi^m}{4\alpha^2 P_N \Gamma^2((m+1)/2)}. \quad (134)$$

Simple calculations yield

$$\begin{aligned} & 2 \min_{0 \leq r \leq \rho_d} \left\{ E_Z \left(\frac{r}{\sigma'^2} \right) - \ln r \right\} \\ & = \begin{cases} E_{\text{sl}} \left(\frac{\rho_d}{\sigma'^2}; \text{SNR}', \alpha \right) - \ln \frac{\rho_d}{2\pi e}, & \frac{\rho_d}{\sigma'^2} \geq \mu_{\text{cr}}(\text{SNR}', \alpha) \\ E_{\text{sp}} \left(\frac{\rho_d}{\sigma'^2}; \text{SNR}', \alpha \right) - \ln \frac{\rho_d}{2\pi e}, & \mu_0(\text{SNR}', \alpha) \leq \frac{\rho_d}{\sigma'^2} \leq \mu_{\text{cr}}(\text{SNR}', \alpha) \\ E_0 \left(\frac{\rho_d}{\sigma'^2}; \text{SNR}', \alpha \right) - \ln \frac{\rho_d}{2\pi e}, & 0 < \frac{\rho_d}{\sigma'^2} \leq \mu_0(\text{SNR}', \alpha). \end{cases} \end{aligned} \quad (135)$$

Note from (55) that $\delta \rightarrow \frac{1}{2} \ln (e^{2R} G(\Lambda)/P_X)$ as the block length $k \rightarrow \infty$. Applying the Stirling approximation

$$\Gamma(z) = \sqrt{2\pi} e^{-z} z^{z-1/2} \left(1 + \frac{1}{12z} + \frac{1}{288z^2} + \cdots \right), \quad (136)$$

we obtain from (134) that

$$\lim_{k \rightarrow \infty} \left\{ -\frac{1}{m} \ln(4\rho_d B_m) \right\} = \frac{1}{2} \ln \frac{P_X}{4\pi^2 e^{2(R+1)} G(\Lambda)}. \quad (137)$$

Substituting (135) and (137) into (133), we conclude that

$$\begin{aligned} & \limsup_{k \rightarrow \infty} \left\{ -\frac{2}{m} \ln I_2(d) \right\} \\ & \geq \begin{cases} E_{\text{sl}}\left(\frac{\rho_d}{\sigma'^2}; \text{SNR}', \alpha\right) - \ln \frac{2\pi e^{2R+1} G(\Lambda) \rho_d}{P_X}, & \frac{\rho_d}{\sigma'^2} \geq \mu_{\text{cr}}(\text{SNR}', \alpha) \\ E_{\text{sp}}\left(\frac{\rho_d}{\sigma'^2}; \text{SNR}', \alpha\right) - \ln \frac{2\pi e^{2R+1} G(\Lambda) \rho_d}{P_X}, & \mu_0(\text{SNR}', \alpha) \leq \frac{\rho_d}{\sigma'^2} \leq \mu_{\text{cr}}(\text{SNR}', \alpha) \\ E_0\left(\frac{\rho_d}{\sigma'^2}; \text{SNR}', \alpha\right) - \ln \frac{2\pi e^{2R+1} G(\Lambda) \rho_d}{P_X}, & 0 < \frac{\rho_d}{\sigma'^2} \leq \mu_0(\text{SNR}', \alpha). \end{cases} \end{aligned} \quad (138)$$

Finally, choosing

$$\rho_d = \frac{P_X}{2\pi e^{2R+1} G(\Lambda)}, \quad (139)$$

we obtain from (58), (121) and (138) that

$$\limsup_{k \rightarrow \infty} \left\{ -\frac{2}{m} \ln \bar{\lambda} \right\} \geq \begin{cases} E_{\text{sl}}(\mu; \text{SNR}', \alpha), & \mu \geq \mu_{\text{cr}}(\text{SNR}', \alpha) \\ E_{\text{sp}}(\mu; \text{SNR}', \alpha), & 1 \leq \mu \leq \mu_{\text{cr}}(\text{SNR}', \alpha) \\ 0, & 0 < \mu \leq 1 \end{cases} \quad (140)$$

where $\mu = \frac{\rho_d}{\sigma'^2} = e^{2(C(\text{SNR}', \alpha) - R - \varepsilon_1(\Lambda))}$.

To investigate the exponent of the right side of (59), let us define

$$\tilde{I}_2(d) \stackrel{\text{def}}{=} \frac{e^{m\delta} m \pi^{m/2}}{\Gamma(m/2 + 1)} \int_{\sqrt{m\rho_0}}^{2d} \rho^{m-1} \Pr[\mathbf{Z} \in \mathcal{D}_m(d, \rho)] d\rho. \quad (141)$$

Following similar steps as in the derivation of the upper bound for $I_2(d)$, we have

$$\tilde{I}_2(d) \leq 4\rho_d B_m \exp \left[-(m-2) \min_{\rho_0 \leq \rho \leq 4\rho_d} \min_{\rho/4 \leq r \leq \rho_d} \left\{ E_Z\left(\frac{r}{\sigma'^2}\right) - \frac{1}{2} \ln \left(r\rho - \frac{\rho^2}{4} \right) \right\} \right] \quad (142)$$

$$= 4\rho_d B_m \exp \left[-(m-2) \min_{\rho_0/4 \leq r \leq \rho_d} \min_{\rho_0 \leq \rho \leq 4r} \left\{ E_Z\left(\frac{r}{\sigma'^2}\right) - \frac{1}{2} \ln \left(r\rho - \frac{\rho^2}{4} \right) \right\} \right] \quad (143)$$

$$= 4\rho_d B_m \exp \left[-(m-2) \min_{\rho_0/4 \leq r \leq \rho_d} \left\{ E_Z\left(\frac{r}{\sigma'^2}\right) - \frac{1}{2} \ln \max_{\rho_0 \leq \rho \leq 4r} \left\{ r\rho - \frac{\rho^2}{4} \right\} \right\} \right] \quad (144)$$

where B_m is defined in (134). Recall from (91) and (139) that $\rho_0 \rightarrow \rho_d$ as the block length $k \rightarrow \infty$ and note that

$$\max_{\rho_0 \leq \rho \leq 4r} \left\{ r\rho - \frac{\rho^2}{4} \right\} = \begin{cases} \rho_0 r - \frac{\rho_0^2}{4}, & \frac{\rho_0}{4} \leq r \leq \frac{\rho_0}{2} \\ r^2, & r \geq \frac{\rho_0}{2}. \end{cases} \quad (145)$$

After some (rather complicated) calculations, we arrive at

$$\limsup_{k \rightarrow \infty} \left\{ -\frac{2}{m} \ln \tilde{I}_2(d) \right\} \geq \begin{cases} E_{\text{ex}} \left(\frac{\rho_d}{\sigma'^2}; \text{SNR}', \alpha \right), & \frac{\rho_d}{\sigma'^2} \geq 2\mu_{\text{cr}}(\text{SNR}', \alpha) \\ E_{\text{sl}} \left(\frac{\rho_d}{\sigma'^2}; \text{SNR}', \alpha \right), & \mu_{\text{cr}}(\text{SNR}', \alpha) \leq \frac{\rho_d}{\sigma'^2} \leq 2\mu_{\text{cr}}(\text{SNR}', \alpha) \\ E_{\text{sp}} \left(\frac{\rho_d}{\sigma'^2}; \text{SNR}', \alpha \right), & \mu_0(\text{SNR}', \alpha) \leq \frac{\rho_d}{\sigma'^2} \leq \mu_{\text{cr}}(\text{SNR}', \alpha) \\ E_0 \left(\frac{\rho_d}{\sigma'^2}; \text{SNR}', \alpha \right), & 0 < \frac{\rho_d}{\sigma'^2} \leq \mu_0(\text{SNR}', \alpha). \end{cases} \quad (146)$$

Putting together (59), (121) and (146), we have

$$\limsup_{k \rightarrow \infty} \left\{ -\frac{2}{m} \ln \bar{\lambda} \right\} \geq \begin{cases} E_{\text{ex}}(\mu; \text{SNR}', \alpha), & \mu \geq 2\mu_{\text{cr}}(\text{SNR}', \alpha) \\ E_{\text{sl}}(\mu; \text{SNR}', \alpha), & \mu_{\text{cr}}(\text{SNR}', \alpha) \leq \mu \leq 2\mu_{\text{cr}}(\text{SNR}', \alpha) \\ E_{\text{sp}}(\mu; \text{SNR}', \alpha), & 1 \leq \mu \leq \mu_{\text{cr}}(\text{SNR}', \alpha) \\ 0, & 0 < \mu \leq 1 \end{cases} \quad (147)$$

where $\mu = \frac{\rho_d}{\sigma'^2} = e^{2(C(\text{SNR}', \alpha) - R - \varepsilon_1(\Lambda))}$.

Apply Lemmas 2, 3, 4 and 5 successively. The desired result (67) then follows from a comparison of (140) and (147). This completes the proof of Theorem 8.

Acknowledgment

The authors wish to thank Prof. G. David Forney, Jr. and two anonymous referees for their very helpful comments and suggestions.

References

- [1] M. H. M. Costa, "Writing on dirty paper," *IEEE Trans. Inform. Theory*, vol. 29, pp. 439-441, May 1983.
- [2] S. I. Gel'fand and M. S. Pinsker, "Coding for channel with random parameters," *Prob. Contr. Inform. Theory*, vol. 9(1), pp. 19-31, 1980.
- [3] R. Zamir, S. Shamai (Shitz), and U. Erez, "Nested linear/lattice codes for structured multiterminal binning," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1250-1276, June 2002.
- [4] U. Erez and R. Zamir, "Lattice decoding can achieve $\frac{1}{2} \log(1 + \text{SNR})$ on the AWGN channel using nested codes," in *Proc. IEEE Int. Symp. Inform. Theory*, Washington, DC, June 2001, p. 125.

- [5] U. Erez and R. Zamir, "Lattice decoded nested codes achieve the Poltyrev exponent," in *Proc. IEEE Int. Symp. Inform. Theory*, Lausanne, Switzerland, June-July 2002, p. 395.
- [6] U. Erez and R. Zamir, "Achieving $\frac{1}{2} \log(1 + \text{SNR})$ on the AWGN channel with lattice encoding and decoding," *IEEE Trans. Inform. Theory*, vol. 50, pp. 2293-2314, October 2004.
- [7] G. Poltyrev, "On coding without restrictions for the AWGN channel," *IEEE Trans. Inform. Theory*, vol. 40, pp. 409-417, March 1994.
- [8] C. E. Shannon, "Probability of error for optimal codes in a Gaussian channel," *Bell Syst. Tech. J.*, vol. 38, pp. 611-656, May 1959.
- [9] R. G. Gallager, *Information Theory and Reliable Communication*, New York: Wiley, 1968.
- [10] G. D. Forney Jr., "On the role of MMSE estimation in approaching the information-theoretic limits of linear Gaussian channels: Shannon meets Wiener," in *Proc. 41th Annal Allerton Conf. Comm., Contr., and Computing*, Monticello, IL, October 2003, pp. 430-439.
- [11] R. Zamir and M. Feder, "On lattice quantization noise," *IEEE Trans. Inform. Theory*, vol. 42, pp. 1152-1159, July 1996.
- [12] G. D. Forney Jr., M. D. Trott, and S. Y. Chung, "Sphere-bound-achieving coset codes and multilevel coset codes," *IEEE Trans. Inform. Theory*, vol. 46, pp. 820-850, May 2000.
- [13] U. Erez and R. Zamir, "Error exponents of modulo-additive noise channels with side information at the transmitter," *IEEE Trans. Inform. Theory*, vol. 47, pp. 210-218, January 2001.
- [14] T. G. Thomas and B. Hughes, "Exponential error bounds for random codes on Gaussian arbitrarily varying channels," *IEEE Trans. Inform. Theory*, vol. 37, pp. 643-649, May 1991.
- [15] U. Erez, S. Litsyn, and R. Zamir, "Lattices which are good for (almost) everything," *IEEE Trans. Inform. Theory*, vol. 51, pp. 3401-3416, October 2005.
- [16] H. A. Loeliger, "Averaging bounds for lattice and linear codes," *IEEE Trans. Inform. Theory*, vol. 43, pp. 1767-1773, November 1997.

Biography

Tie Liu received his B.S. (with honors) and M.S. degrees in 1998 and 2000 respectively, both from the Electrical Engineering department of Tsinghua University, Beijing, China. From 2000 to 2001 he was a graduate student of Massachusetts Institute of Technology, Cambridge, MA. Since 2001 he has been with University of Illinois at Urbana-Champaign, Urbana, IL where he obtained an M.S. degree in Mathematics in December 2004 and is currently pursuing a Ph.D. degree in Electrical Engineering. His research interests are in the areas of information theory, communication theory and systems, and signal processing.

Pierre Moulin received his doctoral degree in 1990, after which he worked for Bell Communications Research for 5 years. He joined the University of Illinois at Urbana-Champaign in 1996. He is currently Professor in the Department of Electrical and Computer Engineering and Affiliate Professor in the Department of Statistics. His fields of professional interest are information theory, image and video processing, statistical signal processing, compression, and information hiding. He has served on the editorial boards of the IEEE Transactions on Information Theory and the Transactions on Image Processing and is the editor in chief of the upcoming IEEE Transactions on Information Forensics and Security. He is an IEEE fellow, recipient of 1997 and 2002 best paper awards from the IEEE Signal Processing Society, 2003 Associate of the UIUC Center for Advanced Study, and 2005 Sony Faculty Scholar.

Ralf Koetter (S '91, M '96) received a Diploma in Electrical Engineering from the Technical University Darmstadt, Germany in 1990 and a Ph.D. degree from the Department of Electrical Engineering at Linköping University, Sweden. From 1996 to 1998, he was a visiting scientist at the IBM Almaden Research Lab., San Jose, California. Dr. Koetter was a Visiting Assistant Professor at the University of Illinois at Urbana-Champaign and Visiting Scientist at CNRS in Sophia Antipolis, France. He joined the faculty of the University of Illinois at Urbana-Champaign in 1999 and is currently an Associate Professor at the Coordinated Science Laboratory at the University. Dr. Koetter's research interests include coding and information theory and their applications to communication systems.

In the years 1999-2001, he served as Associate Editor of coding theory & techniques for the IEEE TRANSACTIONS ON COMMUNICATIONS. In 2000, he started a term as Associate Editor of coding theory for the IEEE TRANSACTIONS ON INFORMATION THEORY. He received an IBM Invention Achievement Award in 1997, an NSF CAREER Award in 2000, and an IBM Partnership Award in 2001. He is a member of the Board of Governors of the IEEE Information Theory Society. In 2004, he received the IEEE Information Theory Society Paper Award.