The Modulo Loss in Lattice Dirty-Paper Coding

Ram Zamir
Department of Electrical Engineering-Systems
Tel Aviv University, Israel
zamir@eng.tau.ac.il

***Abstract*—Lattice decoding of a lattice-shaped codebook is a simple alternative for ML decoding, and it is equivalent to ML decoding after modulo-lattice reduction of the channel output. For good (high-dimensional) lattices, this modulo operation is information lossless in the presence of AWGN. At a finite shaping dimension, however, the lattice decoder is inferior to direct ML decoding from the channel output. The "modulo loss" is particularly large at low SNR, and it gets up to 4dB for scalar shaping. We consider the effect of a known interference (i.e., a dirty-paper channel) on the gap between the two decoders. We show that in the limit of a strong interference, the modulo output becomes a *sufficient statistic* for decoding the input. Thus, in the strong-interference regime, ML decoding suffers the same "modulo loss" as lattice decoding.**

Key words: lattice encoding and decoding, ML decoding, sufficient statistics, modulo-lattice, structured binning, dirty-paper coding, known interference.

I. The dirty-paper channel

The "dirty-paper channel" (DPC) consists of an additive white-Gaussian noise (AWGN) component $Z \sim N(0, N)$, and an interference component S, which is known at the encoder but not at the decoder, [1]:

$$Y = X + S + Z, \tag{1}$$

where X and Y are the channel input and output, respectively. The name comes from viewing transmission of information as writing of X over a dirty paper S, in order to convey a message to a reader of the dirty-and-noisy manuscript Y.

A rate-R / block-length-n coding scheme consists of an encoding function f and decoding function g. Let us use boldface letters for n-vectors, e.g., $\mathbf{X} = (X_1, \ldots, X_n)$. The encoder f generates the channel-input vector $\mathbf{X} = f(m, \mathbf{S})$, depending on the message $m \in \{1, \ldots, M\}$ to be transmitted over the channel, and the forth coming interference vector $\mathbf{S}$, where the number of possible messages is given by $M = 2^{nR}$. The encoder must guarantee that the channel input satisfies a power constraint

$$\frac{1}{n} E\|\mathbf{X}\|^2 \leq P. \tag{2}$$

The decoder g decodes the message m from the received channel output vector; i.e., $\hat{m} = g(\mathbf{Y})$.

Costa [1] derived the capacity C of this channel (i.e., the largest rate R that can be transmitted reliably over the channel), based on the Gelfand-Pinsker formula for general channels with side-information at the encoder [2]. Costa assumed that the interference S is white-Gaussian, $S \sim N(0, Q)$, and showed that C is equal to the capacity of a pure-AWGN channel $Y = X + Z$; i.e.,

$$C = C_{\text{AWGN}} = \frac{1}{2} \log \left(1 + \frac{P}{N} \right), \tag{3}$$

independent of the interference power Q. This result is somewhat surprising, because, unlike the decoder, the encoder cannot simply subtract the interference S without a power penalty. Later work showed that Costa's result holds regardless of the distribution of the interference S; and in fact, with the aid of a common randomness, the result holds for an arbitrarily varying interference [3], [4].

The Gelfand-Pinsker capacity for a general channel with encoder side information S, is given by

$$C = \max_{U,X} \{ I(U;Y) - I(U;S), \} \tag{4}$$

where the maximization is over all auxiliary random variables U, and (possibly constrained) channel input variables X, such that $U \leftrightarrow (X, S) \leftrightarrow Y$ form a Markov chain. In the Costa solution of the dirty-paper channel, the maximization above is achieved by $U = X + \alpha^* S$, where $X \sim N(0, P)$ is independent of S, and

$$\alpha^* = \frac{P}{P + N} \tag{5}$$

is the Wiener-estimation coefficient. The achievability proof is based on the Gelfand-Pinsker *random-binning* scheme, specialized to this (U, X) pair.

II. Lattice dirty-paper coding

Lattice dirty-paper coding is a structured (non-random) technique to approach capacity [4], [5]. It can be viewed as a multi-dimensional extension of Tomlinson-Harashima precoding for inter-symbol interference. Unlike random coding, the performance of lattice dirty-paper coding can be characterized for *any* block length (even for $n = 1$).

The lattice dirty-paper coding system is based on a pair of nested lattices $\Lambda_2 \subset \Lambda_1$ in $\mathbb{R}^n$. We refer to Λ_1 as the *coding lattice* and to Λ_2 as the *shaping lattice*. See [6], [7] for common terminology associated with lattice codes. Each coset in Λ_1/Λ_2 represents a message, hence the coding rate is given by:

$$R = \frac{1}{n} \log |\Lambda_1/\Lambda_2|. \tag{6}$$

The coarse lattice Λ_2 is scaled such that its second moment is equal to the power constraint; i.e.,

$$\sigma^2(\Lambda_2) = P. \tag{7}$$

The fine lattice Λ_1 is randomized by a dither vector $\mathbf{U}$, known to both encoder and decoder, which is uniform over a fundamental cell of the coarse lattice. The encoder and decoder functions are given by

$$\textbf{Encoding:}\quad \mathbf{x} = (\mathbf{v}_m + \mathbf{u} - \alpha\mathbf{s}) \bmod \Lambda_2 \tag{8a}$$

$$\textbf{Decoding:}\quad \hat{\mathbf{v}} = \left(\arg\min_{\lambda\in\Lambda_1} \|\alpha\mathbf{Y} - \mathbf{u} - \lambda\|\right) \bmod \Lambda_2, \tag{8b}$$

where the vector $\mathbf{v}_m$ is a message (coset) representative, for $m = 1, \ldots, M$, and $\hat{\mathbf{v}}$ is the decoded message representative. The modulo-lattice operation $\mathbf{x} \bmod \Lambda_2$ means reducing $\mathbf{x}$ to the fundamental Voronoi cell $\mathcal{V}_0$ of Λ_2 by adding/subtracting a point in Λ.[1] By the basic property of dithered quantization [9], the average power (for each message $\mathbf{v}_m$) is given by $\sigma^2(\Lambda_2) = P$.

The minimization in (8b) is called *lattice decoding*, because it depends only on Λ_1, while ignoring the shaping lattice Λ_2. Lattice decoding has some practical advantages over maximum likelihood (ML) decoding, which decodes $\mathbf{Y}$ to the most probable $\mathbf{v}_m$ (e.g., for $\mathbf{s} = 0$, ML amounts to the nearest point in $\{(\mathbf{v}_m + \mathbf{u}) \bmod \Lambda_2\}_{m=1}^{M}$); see [10], [5]. The nesting relation between the two lattices implies that lattice decoding is, in fact, equivalent to ML decoding from the *modulo output*,

$$\tilde{\mathbf{Y}} = [\alpha\mathbf{Y} - \mathbf{U}] \bmod \Lambda_2. \tag{9}$$

To asses the potential capacity of the scheme, it was shown in [4] that the mapping from the input $\mathbf{v}$ to the modulo output $\tilde{\mathbf{Y}}$, assuming a uniform dither $\mathbf{U}$ and any estimation coefficient $0 < \alpha \leq 1$, is equivalent to a modulo-lattice additive noise channel:

$$\tilde{\mathbf{Y}} = \mathbf{v} + \underbrace{[\alpha\mathbf{Z} + (1-\alpha)\mathbf{U}]}_{\mathbf{Z}_{\text{eq}}} \bmod \Lambda_2, \tag{10}$$

where $\mathbf{Z}_{\text{eq}}$ is the equivalent noise. Note that this channel is *independent of the interference*. It was also shown that for the Wiener coefficient α^*, the mutual information between the input and output of this equivalent channel is greater than or equal to $C_{\text{AWGN}} - 1/2\log(2\pi e G(\Lambda_2))$, where $G(\Lambda_2)$ is the *normalized second moment* of Λ_2 [6]. Furthermore, for a given decoding-error probability P_e, the coding rate $R = 1/n \log|\Lambda_1/\Lambda_2|$ is given by

$$R = C_{\text{AWGN}} - \frac{1}{2}\log\Big(G(\Lambda_2)\cdot\mu(\Lambda_1, \mathbf{Z}_{\text{eq}}, P_e)\Big), \tag{11}$$

where $\mu(\Lambda_1, \mathbf{Z}_{\text{eq}}, P_e)$ is the *normalized volume-to-noise ratio* of Λ_1 with respect to the equivalent noise [11], [8]. For a sequence of "good" (nested) lattices, $G(\Lambda_2) \to 1/2\pi e$ and $\mu(\Lambda_1, \mathbf{Z}_{\text{eq}}, P_e) \to 2\pi e$, for all $P_e > 0$, in the limit as the lattice dimension n goes to infinity [12], [8]. Hence, the capacity loss term in (11) vanishes.

[1] To reduce complexity, the modulo operation at the decoder can be taken with respect to *any* complete set of coset representatives; e.g., a parallelepiped (instead of Voronoi) fundamental cell of Λ_2; see [8].

III. The modulo loss in the clean-paper case

So far so good. But in practice, the potential shaping gain $G(\mathbb{Z})/G(\Lambda_2)$ of a good Λ_2 over a simple scalar lattice $\mathbb{Z}$ does not justify the extra cost in complexity. Hence, the lattice dirty-paper coding scheme employs a simple (even one-dimensional) shaping lattice Λ_2. Beyond the possible loss in capacity, this non-ideal choice causes an additional loss in performance of the lattice decoder compared to the ML decoder. We shall call that a *modulo loss*.

Let us start with the zero-interference case ($S = 0$). For an ideal (good high-dimensional) coding lattice Λ_1, the modulo loss can be measured by the reduction in mutual information when replacing the channel output and (known) dither $(\mathbf{Y}, \mathbf{U})$ by the modulo output $\tilde{\mathbf{Y}}$ in (9). That is,

$$\text{Loss} \triangleq \frac{1}{n}\Big[I(\mathbf{V};\mathbf{Y},\mathbf{U}) - I(\mathbf{V};\tilde{\mathbf{Y}})\Big], \tag{12}$$

where both mutual-information terms are computed with respect to an input vector $\mathbf{V}$ which is uniform over a fundamental cell of Λ_2, and where the channel input is given by $\mathbf{X} = (\mathbf{V} + \mathbf{U} - \alpha\mathbf{S}) \bmod \Lambda_2$ as in (8a). Note that this difference is always nonnegative due to the Markov chain relation $\mathbf{V} \leftrightarrow (\mathbf{Y}, \mathbf{U}) \leftrightarrow \tilde{\mathbf{Y}}$. Note also that for $S = 0$, the first (ML) term in (12) can be replaced by

$$I(\mathbf{V};\mathbf{Y},\mathbf{U}) = I(\mathbf{V};\mathbf{V}+\mathbf{Z}). \tag{13}$$

We first observe that the modulo loss (12) is zero in the limit of high signal-to-noise ratio (SNR) $P/N \to \infty$.

Lemma 1. (Modulo loss at high SNR) *For a zero-interference channel ($S = 0$), and a lattice coding scheme with $\alpha = 1$ in (8),*

$$\lim_{N\to 0} I(\mathbf{V};\mathbf{V}+\mathbf{Z}) - I(\mathbf{V};\tilde{\mathbf{Y}}) = 0. \tag{14}$$

Proof: It is easy to verify that in the high SNR limit, both mutual-information terms in (12) (see also (13)) are approximately equal to $1/n[\log V(\Lambda_2) - h(\mathbf{Z})]$, where $V(\Lambda_2)$ is the cell volume of Λ_2. $\Box$

Note that $1/n \log V(\Lambda_2) = 1/2\log(P/G(\Lambda_2))$; see [4]. Thus, the gap to the zero-interference capacity C_{AWGN} in the high-SNR regime is the *shaping gain*,

$$\frac{1}{2}\log(2\pi e G(\Lambda_2)) \ \ \text{bit} \tag{15}$$

(or 1.5 dB) for both ML and lattice decoders.

Note also that – since the existence of the interference S cannot improve the performance under ML decoding (e.g., it could be replaced by a common randomness), while the performance of lattice decoding is invariant of the interference – Lemma 1 holds also in the presence of interference (i.e., with $S \neq 0$). Thus, the modulo loss at high SNR is zero also for the dirty-paper channel.

In the general SNR case, however, the modulo loss is strictly positive. In particular, for low SNR, the *non*-modulo AWGN channel $\mathbf{V} \to \mathbf{V} + \mathbf{Z}$ enjoys a "natural shaping": $\mathbf{V} + \mathbf{Z}$ is approximately white-Gaussian, hence $1/nI(\mathbf{V};\mathbf{V}+\mathbf{Z}) \approx C_{\text{AWGN}}$, *regardless of the distribution of* $\mathbf{V}$. This low-SNR phenomena is due in reality to the majority of codewords at the border of the shaping region (i.e., codewords whose decision regions are much larger). In contrast, as can be seen in Figure 1, $1/nI(\mathbf{V};\tilde{\mathbf{Y}})$ is strictly below C_{AWGN} for a scalar ($\mathbb{Z}$) shaping lattice. In particular, while the maximum derivative of $C_{\text{AWGN}}(\text{SNR})$ is ≈ 0.7 bit per unit cost (achieved at SNR=0), the maximum value of $1/nI(\mathbf{V};\tilde{\mathbf{Y}})/\text{SNR}$ is only ≈ 0.3 (achieved at SNR ≈ 1). This amounts to a loss of 4dB in performance in the low SNR regime.

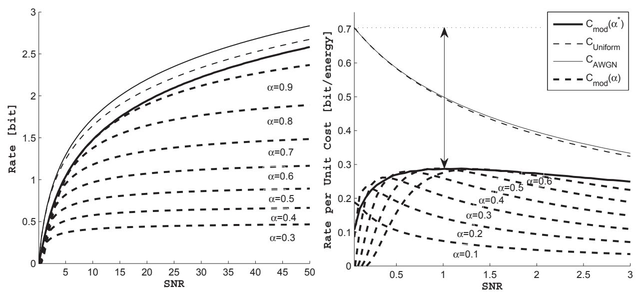

Fig. 1. The left figure shows the capacity $C_{\text{mod}-\mathbb{Z}}(\text{SNR},\alpha)$ of the mod-$\mathbb{Z}$ channel as a function of the SNR for several values of α, compared to the Shannon capacity $C_{AWGN} = 1/2\log(1+SNR)$, and the uniform-input (non-modulo) capacity $I(U;U+Z)$. The right figure zooms into the low SNR regime, and shows the corresponding graphs of capacity-per-unit cost, $C_{\text{mod}-\mathbb{Z}}(\text{SNR},\alpha)/\text{SNR}$. For very low SNR (SNR <0.5), the optimum linear coefficient α^{**} is, somewhat surprisingly, slightly larger than the Wiener coefficient α^{*}. Nevertheless, the best operation point in terms of energy per bit is around SNR=1, where $\alpha^{**}\approx\alpha^{*}=1/2$.

IV. Sufficient statistics for a very dirty paper

The negative conclusion above (for the general SNR case) assumed no interference.[2] In the presence of interference, the codebook expands over a larger region, and the significance of the border codewords (which improves ML decoding performance in the zero-interference case) decreases. Hence, it is natural to ask how interference affects the modulo loss (12).

We show next that in the limit of a *strong* interference, ML decoding cannot improve upon lattice decoding. In other words, the modulo output $\tilde{\mathbf{Y}}$ becomes a sufficient statistic, and the modulo loss (12) vanishes.

Theorem 1. (Modulo loss for a strong interference) *For a Gaussian interference $S\sim N(0,Q)$, and a lattice dirty-paper coding scheme (8) with any parameter $0<\alpha\leq 1$, the modulo loss (12) satisfies:*

$$\lim_{Q\to\infty} I(\mathbf{V};\mathbf{Y},\mathbf{U}) - I(\mathbf{V};\tilde{\mathbf{Y}}) = 0. \tag{16}$$

where $\tilde{\mathbf{Y}}$ is the modulo output (9).

The proof is given below after some discussion and preliminaries.

[2]For ML decoding we used the simplified expression (13), while lattice decoding is invariant of the interference; see the equivalent mod-Λ channel (10).

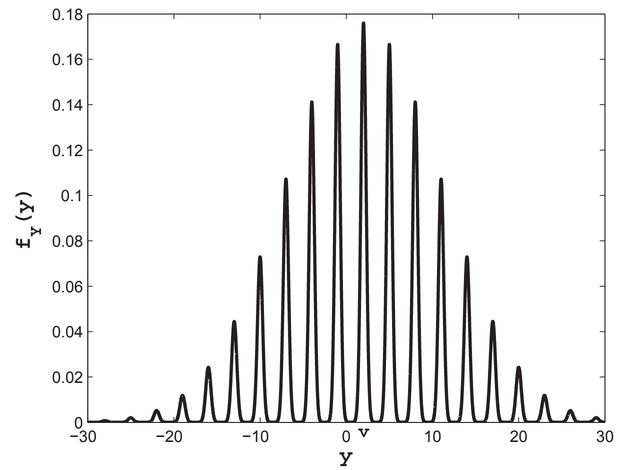

Fig. 2. Interference concentration. Channel output distribution for a specific message v in the presence of a strong Gaussian interference ($\sigma_s \gg \Delta \gg \sigma_z$). The figure assumes $u=0$, so the coset shift is equal to v.

A. Lattice-periodic channel

To get some intuition, consider a hypothetical situation, where the addition operation in the dirty-paper channel (1) is *modulo* some multiple of the coarse lattice cell, and the interference S is *uniform* over that extended cell; say, for simplicity, the coarse lattice is a step-Δ scalar lattice, the addition in (1) is modulo $N\Delta$ for some integer $N>1$, and $S\sim\text{Unif}(0,N\Delta)$. In this case, the conditional channel-output distribution $p(y|\text{message}=m,\text{dither}=u)$ is a cyclic-periodic function of y with a period Δ; i.e., $p(y\oplus\Delta|m,u) = p(y|m,u)$, for all m and u, where $\oplus$ denotes addition modulo $N\Delta$. It follows that $Y_q = Q_\Delta(Y)$ is uniformly distributed over $\{0,\ldots,N-1\}$, independent of (Y_e,m,u), where $Y_e = Y \bmod \Delta = Y - Y_q$:

$$p(y=y_q+y_e|m,u) = \frac{1}{N}p(y_e|m,u),$$

for all y_q in $\{0,\ldots,N-1\}$. Thus, Y_e (which is equivalent to the decision vector $\tilde{Y} = [Y_e - u] \bmod \Delta$) is a *sufficient statistic* for decoding m from (Y,u), so the modulo loss is *zero*.

For a general interference S and regular addition in (1), however, the output distribution in the example above is not quite periodic in Δ; hence, Y_q depends on (m,u), and ignoring it may result in loss of performance; see Figure 2. Yet, intuitively, the additional ambiguity induced by the interference S weakens the dependence of Y_q on the message m, and eliminates it completely as S becomes "strong".

B. Semi lattice-periodic channel

The following lemma provides a formal statement for this intuition. It implies that corrupting the output of a real-valued channel by a discrete noise which is nearly uniform over a lattice, is information-wise equivalent to reducing the output modulo the lattice.

Lemma 2. (Strong lattice-distributed noise) *If* $\mathbf{Y}$ *and* $\mathbf{Y}'$ *have finite second moments, then for any* $\mathbf{S}$ *which is independent of* $(\mathbf{X}, \mathbf{Y}, \mathbf{Y}')$*, has a p.d.f. and a finite second moment,*

$$\lim_{a\to\infty} I(\mathbf{X}; \mathbf{Y} + Q_\Lambda[a\mathbf{S} + \mathbf{Y}']) = I(\mathbf{X}; \mathbf{Y} \bmod \Lambda), \quad (17)$$

where $Q_\Lambda[\cdot]$ *denotes quantization to the lattice* Λ*; i.e.,*

$$\mathbf{x} = Q_\Lambda[\mathbf{x}] + [\mathbf{x} \bmod \Lambda]. \quad (18)$$

Note that the variable $Q_\Lambda[a\mathbf{S} + \mathbf{Y}']$ is asymptotically independent of $\mathbf{Y}$, and is nearly uniform over the points of Λ inside a region that increases with a.

The proof is due to Tamas Linder [13], and can be found in [8].

C. Proof of Theorem

For simplicity of the exposition, assume $\alpha = 1$; so $\mathbf{X} = [\mathbf{V} + \mathbf{u} - \mathbf{S}] \bmod \Lambda_2$ and $\tilde{\mathbf{Y}} = [\mathbf{Y} - \mathbf{u}] \bmod \Lambda_2$, for a given dither value $\mathbf{U} = \mathbf{u}$. Note that for any vector $\mathbf{a}$, we have from (18)

$$\begin{aligned}[] [\mathbf{a} - \mathbf{S}] \bmod \Lambda_2 + \mathbf{S} &= [\mathbf{a} - \mathbf{S}] - Q_{\Lambda_2}[\mathbf{a} - \mathbf{S}] + \mathbf{S} && (19)\\ &= \mathbf{a} + Q_{\Lambda_2}[\mathbf{S} - \mathbf{a}], && (20)\end{aligned}$$

Hence, setting $\mathbf{a} = \mathbf{V} + \mathbf{u}$, we have for each dither value $\mathbf{U} = \mathbf{u}$,

$$\begin{aligned} I(\mathbf{V}; \mathbf{Y}) &= I\Big(\mathbf{V}; \mathbf{V} + \mathbf{u} + \mathbf{Z} + Q_{\Lambda_2}[\mathbf{S} - \mathbf{V} - \mathbf{u}]\Big) && (21)\\ &\to I\Big(\mathbf{V}; [\mathbf{V} + \mathbf{u} + \mathbf{Z}] \bmod \Lambda_2\Big) && (22)\\ &= I(\mathbf{V}; \mathbf{Y} \bmod \Lambda_2) && (23)\\ &= I(\mathbf{V}; \tilde{\mathbf{Y}}), && (24)\end{aligned}$$

where the limit as $Q \to \infty$ in (22) follows from Lemma 2 above, and (23) follows from the *distributive law* of the modulo operation [5]. The equivalent statement $I(\mathbf{V}; \mathbf{Y}|\tilde{\mathbf{Y}}) \to 0$ follows from the chain rule for mutual information.

The extension to a general estimation coefficient $0 < \alpha \le 1$ is similar [8].[3] $\square$

V. Conclusion

The main messages of the paper are summarized as follows:

1) Lattice decoding amounts to ML decoding of the modulo-lattice output of the channel.

2) At high SNR, direct ML decoding and lattice decoding are equivalent, and the shaping loss of scalar shaping (1.5dB) is unavoidable for both decoders.

3) At low SNR and zero interference (i.e., a pure AWGN channel), ML decoding enjoys a "natural shaping"; i.e., even with scalar shaping it achieves the full AWGN channel capacity.

4) In contrast, for the same channel, lattice decoding suffers a "modulo loss", which can get up to 3 dB.

5) In the limit of a strong Gaussian interference, the modulo output becomes a sufficient statistic, and the modulo loss vanishes.[4]

The negative implication of this result is that there is no free lunch, i.e., no "natural shaping" in the strong interference regime. To approach the full dirty-paper channel capacity, a good shaping is necessary even at low SNR.[5]

We further note that as of today, the best known dirty-paper coding schemes in the presence of a strong interference are based on lattices. Thus, although the results presented in this paper are limited to lattice-based transmission, they may indicate on a more general behavior.

Acknowledgment

I thank Ilai Bistritz for computing the capacity curves in Figure 1, and the semi-periodic output distribution in Figure 2. I thank Uri Erez and Tamas Linder for helpful discussions about Lemma 2. I thank Chen Feng for detecting an error in the proof in an early version of the paper.

References

[1] M. Costa, "Writing on dirty paper," *IEEE Trans. Information Theory*, vol. IT-29, pp. 439–441, May 1983.

[2] S. Gelfand and M. S. Pinsker, "Coding for channels with random parameters," *Problemy Pered. Inform. (Problems of Inform. Trans.)*, vol. 9, No. 1, pp. 19–31, 1980.

[3] A. S. Cohen and A. Lapidoth, "The Gaussian watermarking game," *IEEE Trans. Information Theory*, June, 2002, 1639–1667. See also, *On the Gaussian Watermarking Game*, Proc. Int. Symp. Inform. Theory (ISIT), page 48, Sorrento, Italy, June 2000.

[4] U. Erez, S. S. (Shitz), and R. Zamir, "Capacity and lattice strategies for cancelling known interference," *IEEE Trans. Information Theory*, pp. 3820–3833, Nov. 2005.

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

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

[7] G. Forney, "On the role of MMSE estimation in approaching the information-theoretic limits of linear gaussian channels: Shannon meets Wienner," in *Allerton Conference*, Allerton House, Urbana, Illinois, Oct. 2003.

[8] R. Zamir, *Lattice Coding for Signals and Networks*. Cambridge University Press, to appear June 2014.

[9] R. Zamir and M. Feder, "On lattice quantization noise," *IEEE Trans. Information Theory*, pp. 1152–1159, July 1996.

[10] R. Urbanke and B. Rimoldi, "Lattice codes can achieve capacity on the AWGN channel," *IEEE Trans. Information Theory*, vol. IT-44, pp. 273–278, Jan. 1998.

[11] R. Zamir, "Lattices are everywhere," in *Inf. Theory Appl. Wkshp.* UCSD, (La Jolla, CA), February 2009, p. 392421.

[12] U. Erez, S. Litsyn, and R. Zamir, "Lattices which are good for (almost) everything," *IEEE Trans. Information Theory*, pp. pp. 3401–3416, Oct. 2005.

[13] T. Linder, "Private communication."

[14] S. Borade and L. Zheng, "Writing on fading paper, dirty tape with little ink: Wideband limits for causal transmitter csi," *IEEE Trans. Information Theory*, vol. IT-58, pp. 5388–5397, Aug. 2012.

[3] Follow the derivation of the equivalent modulo-lattice channel (10) (see [4]), and use the distributive law, to show that $\alpha\mathbf{Y} = [\mathbf{V} + \mathbf{u}] \bmod \Lambda_2 + \mathbf{Z}_{\rm eq} + Q_{\Lambda_2}(\alpha\mathbf{S})$. Then use the "high-resolution quantization property", to conclude that $\mathbf{Z}_{\rm eq}$ and $Q_{\Lambda_2}(\alpha\mathbf{S})$ are asymptotically independent as $Q \to \infty$, so Lemma 2 applies.

[4] This result can be easily extended to any interference with a smooth density ($S = \beta S_0$, as β goes to infinity), or to an arbitrary interference (in the minimax sense); see [8].

[5] Though the order of limits (interference / SNR) matters; see [14].