

A Modulo-Lattice Transformation for Multiple-Access Channels

Uri Erez[†] and Ram Zamir^{*}

Dept. of Electrical Engineering - Systems
Tel-Aviv University
Tel-Aviv 69978, Israel
Email: {uri,zamir}@eng.tau.ac.il

Abstract—A simple lemma is derived that allows to transform a general single-user (non-Gaussian, non-additive) continuous-alphabet channel as well as a general multiple-access channel into a modulo-additive noise channel. While in general the transformation is information lossy, it allows to leverage linear coding techniques and capacity results derived for networks comprised of additive Gaussian nodes to more general networks.

Index Terms—Lattice codes, Network coding, Additive Noise.

I. INTRODUCTION

There has been considerable progress in recent years in deriving structured coding schemes and capacity results for single-user channels as well as networks with AWGN noise, e.g., [2], [7], [10], [4], [6], [1]. A key ingredient in these results is the use of linear/lattice codes that are particularly well suited for additive noise channels. In this work, we derive a lemma that allows to leverage these techniques to networks which have general multiple-access channels (MAC) nodes. Specifically, we provide a transformation that transforms a general (neither Gaussian nor additive) continuous-amplitude MAC, into an modulo-additive one, at the price of some information loss.

For some channels, the rate increase offered by lattice coding outweigh the loss in mutual information incurred by the transformation. More specifically, for any network of Gaussian channels for which lattice encoding schemes are exploited to derive an achievable rate region, using the derived lemma, one may obtain an analogous inner bound on the rate region for a network with the same topology but with general channels.

The transformation is a simple extension of the mod- Λ transformation derived for single-user channels as given in [2] and specifically the version given by Forney in [3].

II. MODULO LATTICE TRANSFORMATION

Consider a general multiple-access channel with M users. For simplicity, the derivation below will be written in terms of a two-user ($M = 2$) MAC. The extension to the general case is straightforward. We thus assume a two-user MAC with inputs X_1, X_2 and output Y , with a general transition

distribution $w(Y|X_1, X_2)$, where all random variables are real-valued and continuous. Unless stated otherwise we assume that all transmitters as well as the receiver have full channel knowledge. We further assume that the transmitters are subject to an input constraint. In the following the input constraint will be taken to be a power constraint so that the inputs must satisfy $E[X_i^2] \leq P_{X,i}$. We may transform the channel into a mod- Λ MAC as we describe next.

Let Λ be a lattice and let \mathcal{V} be some fundamental region of Λ . Let $\mathbf{v}_i \in \mathcal{V}$, $i = 1, 2$, be the information bearing signals of the users. Further, let $\mathbf{U}_i \sim \text{Unif}(\mathcal{V})$, $i = 1, 2$, be independent dithers uniformly distributed over \mathcal{V} . We assume that the dither \mathbf{U}_i is known to transmitter i as well as to the receiver¹. For any scalar function $f: \mathbb{R} \rightarrow \mathbb{R}$, when operating on an n -dimensional vector, we define its operation as the cartesian product extension. That is, for $\mathbf{x} \in \mathbb{R}^n$,

$$f(\mathbf{x}) = (f(x_1), \dots, f(x_n)).$$

With this notation, the transmitters operate as follows.

Transmitter i

Computes:

$$\mathbf{X}'_i = [\mathbf{v}_i + \mathbf{U}_i] \bmod \Lambda. \quad (1)$$

Sends the preprocessed signal:

$$\mathbf{X}_i = f_i(\mathbf{X}'_i). \quad (2)$$

where $f_i(\cdot)$ satisfies the power constraint,

$$E[f_i(\mathbf{U})^2] \leq P_{X,i}. \quad (3)$$

Note that due to the dither, \mathbf{X}'_i is uniformly distributed over \mathcal{V} (see [2], [3]), for any value of \mathbf{v}_i . Let $\mathbf{S} = k_1 \mathbf{X}'_1 + k_2 \mathbf{X}'_2$ where k_1, k_2 are integers. Let $\hat{\mathbf{S}} = g(\mathbf{Y})$ be an estimator of \mathbf{S} from the output \mathbf{Y} . Ideally, the criterion for estimation should be that of minimum noise entropy but any other estimator may be used. Denote the estimation error by $\mathbf{N} = \hat{\mathbf{S}} - \mathbf{S}$. The receiver processes the output as follows.

Receiver

Computes:

$$\begin{aligned} \mathbf{Y}' &= [\hat{\mathbf{S}} - k_1 \mathbf{U}_1 - k_2 \mathbf{U}_2] \bmod \Lambda \\ &= [g(\mathbf{Y}) - k_1 \mathbf{U}_1 - k_2 \mathbf{U}_2] \bmod \Lambda. \end{aligned} \quad (4)$$

^{*} This work was supported by the ISF under grant # 1259/08.

[†] This work was supported by the ISF under grant # 1234/08.

¹In practice, a pseudorandom sequence will be used at both transmission ends.

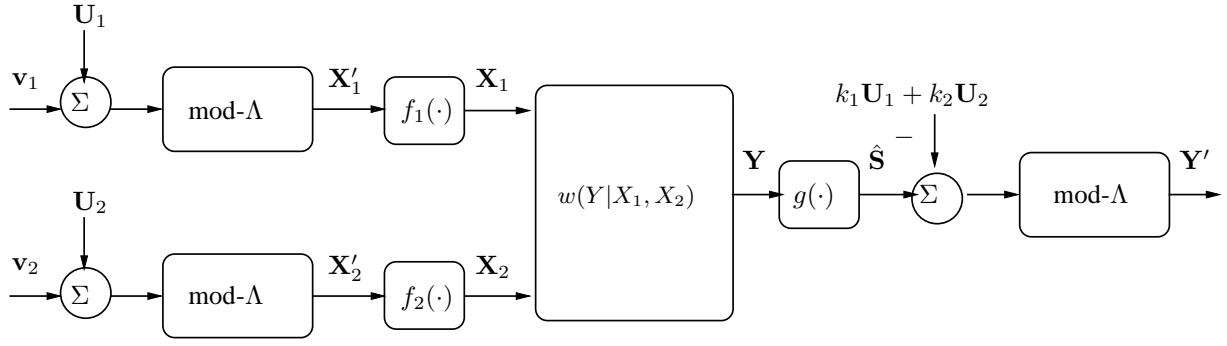


Fig. 1. The general mod Λ transformation.

We now observe that the induced channel from $\mathbf{v}_1, \mathbf{v}_2$ to \mathbf{Y}' is indeed additive modulo Λ . From (4) and using the definition of \mathbf{S} we obtain,

$$\begin{aligned}
 \mathbf{Y}' &= [\hat{\mathbf{S}} + (\mathbf{S} - \hat{\mathbf{S}}) - k_1 \mathbf{U}_1 - k_2 \mathbf{U}_2] \bmod \Lambda \\
 &= \left[\hat{\mathbf{S}} + k_1([\mathbf{v}_1 + \mathbf{U}_1] \bmod \Lambda) + k_2([\mathbf{v}_2 + \mathbf{U}_2] \bmod \Lambda) \right. \\
 &\quad \left. - \mathbf{S} - k_1 \mathbf{U}_1 - k_2 \mathbf{U}_2 \right] \bmod \Lambda \\
 &= [k_1 \mathbf{v}_1 + k_2 \mathbf{v}_2 + (\hat{\mathbf{S}} - \mathbf{S})] \bmod \Lambda \\
 &= [k_1 \mathbf{v}_1 + k_2 \mathbf{v}_2 + \mathbf{N}] \bmod \Lambda,
 \end{aligned} \tag{5}$$

where (5) follows since k_1 and k_2 are integers. Note that $(\mathbf{V}_1, \mathbf{V}_2, \mathbf{U}) \leftrightarrow (\mathbf{X}_1, \mathbf{X}_2) \leftrightarrow \mathbf{Y}$ forms a Markov chain. Since $\mathbf{X}_1, \mathbf{X}_2$ are independent of $\mathbf{V}_1, \mathbf{V}_2$, so is the pair $(\mathbf{S}, g(\mathbf{Y}))$ and hence \mathbf{N} . We thus obtain the following lemma,

Lemma 1: Applying the transmission scheme (2), (4), results in a mod- Λ additive noise channel:

$$\mathbf{Y}' = [k_1 \mathbf{v}_1 + k_2 \mathbf{v}_2 + \mathbf{N}] \bmod \Lambda \tag{6}$$

where $\mathbf{N} = \hat{\mathbf{S}} - \mathbf{S}$ is independent of the inputs \mathbf{v}_1 and \mathbf{v}_2 .

We note that the lemma generalizes the single-user version of [2] in several ways. First, it applies to any channel, which need not be Gaussian nor additive², and thus extends results even for a single-user channel. Second, it treats multiple-access channels as well. We further note that while in [3] the optimal estimator turned out to be linear since the channel considered was an AWGN channel, it is apparent that in a more general setting, non-linear estimation is beneficial.

We refer to the application of the functions $f_i(\cdot)$ at the transmitters (2) as “preprocessing” and to the application of the function $g(\cdot)$ at the receiver (4) as “postprocessing”. Allowing for preprocessing at the transmitters is of great value, as will be seen in the examples below. The ability of the transmitters to choose suitable preprocessing functions of course depends on their having channel knowledge. Furthermore, in general, determination of the optimal preprocessing and postprocessing functions calls for a joint optimization³. We note that pre-

²We note that as opposed to the single-user additive Gaussian case, in general there is information loss associated with having the receiver use only \mathbf{Y}' (i.e., it is not a sufficient statistic) rather than $(\mathbf{Y}', \mathbf{U})$.

³In a slowly varying environment, for instance, the pre/post processing functions could be performed at the receiver and the determined preprocessing functions would be fed back to the transmitters.

processing was not allowed for in previous works such as [8] where a time-varying channel model was assumed and no channel knowledge was available to the transmitters.

Consider now a single user mod Λ channel

$$\mathbf{Y}' = [k_i \mathbf{v}_i + \mathbf{N}] \bmod \Lambda. \tag{7}$$

Note first that since k_i is an integer, as long as it is non-zero, it has no effect on the capacity of the channel, which may be achieved by taking the input $\mathbf{V}_i \sim \text{Unif}(\mathcal{V})$, a choice that results also in $[k_i \mathbf{V}_i] \bmod \Lambda \sim \text{Unif}(\mathcal{V})$. We denote the resulting capacity by C_Λ , where we suppress the dependence on the particular choice of encoding functions f_i , the choice of lattice Λ and the choice of estimator $g(\cdot)$. We have,

$$\begin{aligned}
 C_\Lambda &= \frac{1}{n} \log |\mathcal{V}| - \frac{1}{n} h(\mathbf{N} \bmod \Lambda) \\
 &\geq \frac{1}{n} \log |\mathcal{V}| - \frac{1}{n} h(\mathbf{N}).
 \end{aligned}$$

Denote the set of (indices of) “active users”, for which $k_i \neq 0$, by \mathcal{A} . It is easy to see that the capacity region of the mod Λ MAC (6) is the set of all rate vectors satisfying,

$$\sum_{i \in \mathcal{A}} R_i \leq C_\Lambda, \quad \text{and} \quad R_i = 0, \text{ for } i \notin \mathcal{A}. \tag{8}$$

Note that in our formulation, all users transmit simultaneously⁴ and this assumption also affects the capacity for the other users. In particular, even for the case of an additive Gaussian MAC, the transformation is lossy as seen in Example IV-B below. Note also that the capacity region of the mod Λ MAC (8) is symmetric in the users, which means that in a sense “stronger” users suffer due to “weaker” ones.

One way to achieve a general rate vector \mathbf{R} is by time sharing of corner points of the form $(0, \dots, R_i, \dots, 0)$, corresponding to only one user, say user i sending a message at any given time, while for all other users $\mathbf{v}_j = 0, j \neq i$. Alternatively, one can have all users sending a message simultaneously, each user using a (different) codebook drawn uniformly over \mathcal{V} with rate close to R_i , and decode using a joint maximum likelihood (ML) decoder. In certain network

⁴This is true even for users for which $k_i = 0$, as dictated by (2) due to the dither. Nonetheless, choosing $k_i = 0$ for a user, still increases the capacity of the other users as it allows for better estimation.

settings however, further discussed below, the users will use neither of these coding approaches but will rather all send messages simultaneously while using the *same* codebook (which will be linear). In such a scenario, the decoder's task is to decode a linear combination of the messages (see, e.g., [6], [1]) rather than the individual ones (which will in general not be possible). In a network scenario, different MAC nodes may choose to focus on different subsets of users, i.e., they may set $k_i = 0$ for different users, thus working with different sets of "active users".

III. NETWORK CODING USING NESTED LATTICE CODES

As our motivation for the $\text{mod } \Lambda$ transformation is to allow to leverage the strengths of linear coding in a network coding scenario, we now briefly describe how this may be achieved using the framework of nested lattice codes.

As an example scenario, we may consider the Gaussian relay network proposed in [8], depicted in Figure 2, where M users wish to communicate with a destination (central decoder) through a layer of L relays. Each relay receives some weighted (by the fading coefficients) linear combination of the transmitted signals corrupted by AWGN, i.e., each relay is a Gaussian MAC.

We note that the approach we describe is different than the one taken in [8] which relies on treating the codewords as Gaussian sources. In particular the approach of [8] calls for decoding first the lattice Λ (necessitating Λ to be a good channel code), while the nested lattice approach does not. The nested lattice approach has been described in in the context of Gaussian networks in [6], [1], [9].

In general, for any lattice Λ used in the transformation, we may use a "fine" lattice Λ_1 such that $\Lambda \subset \Lambda_1$, i.e., (Λ, Λ_1) is a nested lattice pair [2]. The codebook for all users will then be $\mathcal{C} = \Lambda_1 \cap \mathcal{V}$. It is not hard to show that for any⁵ coarse lattice Λ , there exists a fine lattice Λ_1 such that the resulting code $\mathcal{C} = \Lambda_1 \cap \mathcal{V}$ will approach the capacity C_Λ as the dimension goes to infinity, when used over the channel (7). This may be shown by random coding as in [5] (see also [2]).

We now further restrict the choice of the nested lattice pair, for reasons that will become clear shortly. In the rest of this section, we will focus for simplicity on the case where $\Lambda = p \cdot \mathbb{Z}^n$ where p is a (large) prime number.⁶ That is, the lattice is a Cartesian product of a uniform one-dimensional lattice. We take $\mathcal{V} = [0, p)^n$ as the fundamental region. We further assume that the code \mathcal{C} is a linear code over \mathbb{Z}_p . That is, it is spanned by some $\kappa \times n$ ($1 \leq \kappa < n$) generating matrix whose elements are in the field \mathbb{Z}_p ,

$$\mathcal{C} = \{\mathbf{c} : \mathbf{c} = \mathbf{n} \cdot G \text{ mod } p \text{ for some } \mathbf{n} \in \mathbb{Z}_p^\kappa.\}$$

The fine lattice Λ_1 is obtained by "lifting" \mathcal{C} using $\Lambda_1 = \mathcal{C} + p \cdot \mathbb{Z}^n$, i.e., Λ_1 is generated from \mathcal{C} using Construction A.

⁵To be precise, the coarse lattice should be chosen to be a Cartesian product of Λ , see [2].

⁶The results may be generalized to obtain "shaping" along the lines of Section VII of [2].

The benefit of using a nested lattice approach as described above is that the coding problem is greatly simplified; it reduces to coding and decoding of a linear code over an additive-noise channel over the field \mathbb{Z}_p . We now examine the "network coding" aspect of the problem by considering the setup proposed in [8]. Recall that M denotes the number of users and suppose we have $L \geq M$ MACs (i.e., relays), all sharing the same inputs. Applying a $\text{mod } \Lambda$ transformation on each, we obtain L channels,

$$\mathbf{Y}'(l) = \left[\sum_{m=1}^M k_m(l) \mathbf{c}_m + \mathbf{N}(l) \right] \text{ mod } p, \quad l = 1, \dots, L.$$

Assuming correct decoding⁷ by some $L' \geq M$ relays, each of these relays will obtain knowledge of a linear combination $\bar{\mathbf{c}}(l) = [\sum_{m=1}^M k_{lm} \mathbf{c}_m] \text{ mod } p$. Assuming now that a central node obtains from each MAC its decoded output, as well as the coefficient vectors $\mathbf{k}(l) \text{ mod } p$, then it will be able to recover the messages if and only if the matrix

$$\begin{pmatrix} k_1(1) & k_1(2) & \dots & k_1(L') \\ k_2(1) & k_2(2) & \dots & k_2(L') \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ k_M(1) & k_M(2) & \dots & k_M(L') \end{pmatrix}$$

is full row rank over \mathbb{Z}_p .

IV. EXAMPLES AND APPLICATIONS

A. Multiplicative MAC with AWGN noise

Consider the channel,

$$Y = X_1 \cdot X_2 + Z,$$

where Z is zero-mean AWGN with variance P_Z and X_1 and X_2 are subject to power constraint P_X . Let $\Lambda = A\mathbb{Z}$ be the one-dimensional lattice of integers, so that the fundamental region is $\mathcal{V} = [0, A)$. The information bearing signals take values in $v_i \in \mathcal{V}$. We may take

$$\begin{aligned} X_i &= f_i(X'_i) \\ &= c \cdot e^{bX'_i}, \end{aligned}$$

where the constants b, c are such that the power constraint is satisfied. At the receiver we get,

$$Y = c^2 e^{b(X'_1 + X'_2)} + Z.$$

We define $S = X'_1 + X'_2$. Consider the following estimator,

$$\hat{S} = \frac{1}{b} \ln(|Y/c^2|).$$

Figure 3 depicts the probability density function (PDF) of the effective noise when $b = 1, c = 4$ and $P_Z = 1$.

⁷We implicitly assume that a relay that is not able to decode is aware of that and "declares an erasure".

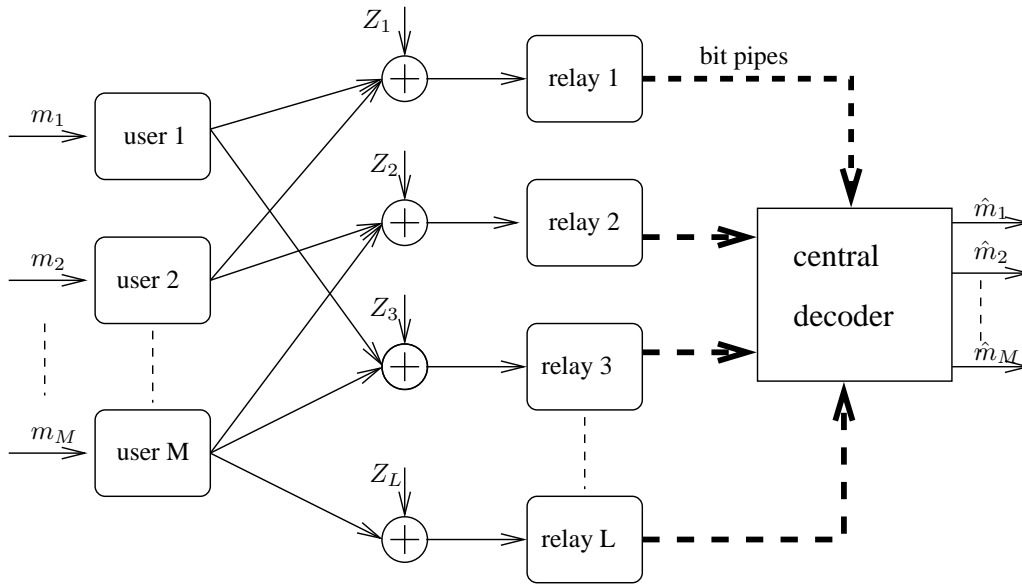


Fig. 2. A multi-relay multi-user network scenario.

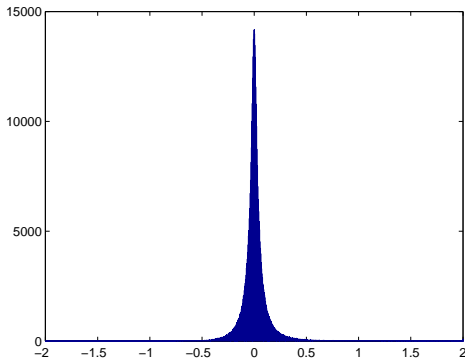


Fig. 3. PDF of effective noise in Example IV-A.

B. Gaussian MAC

Consider an M -user Gaussian MAC,

$$\mathbf{Y} = \sum_{i=1}^M h_i \mathbf{X}_i + \mathbf{Z}, \quad (9)$$

where \mathbf{Z} is zero-mean AWGN with unit variance and user i is subject to power constraint $P_{X,i}$. We also denote $\text{SNR}_i = P_{X,i} h_i^2$. We assume throughout the rest of the paper that all users are active, i.e., that $|\mathcal{A}| = M$.

1) *Symmetric case*: Assume that $\text{SNR}_i = \text{SNR}$ for all $i = 1, \dots, M$. Let \mathcal{V} be the fundamental Voronoi region of a lattice Λ . Assume further that Λ is normalized so that $\frac{1}{n} E[\|\mathbf{U}\|^2] = 1$ where $\mathbf{U} \sim \text{Unif}(\mathcal{V})$. The information bearing signals are $\mathbf{v}_i \in \mathcal{V}$ and the “dithered information bearing signals are given by (2). The output of the transmitters is thus,

$$\begin{aligned} \mathbf{X}_i &= \sqrt{P_{X,i}} \cdot \mathbf{X}'_i \\ &= \sqrt{P_{X,i}} \cdot ([\mathbf{v}_i + \mathbf{U}_i] \bmod \Lambda). \end{aligned}$$

Define $\mathbf{S} = \sum_{i=1}^M \mathbf{X}'_i$, i.e., we take $k_i = 1$ for all i . It seems reasonable to take a linear estimator in this example. The optimal linear MMSE estimator of \mathbf{S} from the output \mathbf{Y} is

$$g(\mathbf{Y}) = \frac{M \cdot \sqrt{\text{SNR}}}{M \cdot \text{SNR} + 1} \cdot \mathbf{Y}.$$

Taking a sequence of lattices with normalized second moment $G(\Lambda) \rightarrow \frac{1}{2\pi e}$ as $n \rightarrow \infty$, results in

$$C_\Lambda = \frac{1}{2} \log \left(\frac{1}{M} + \text{SNR} \right)^+,$$

where x^+ means $\max\{x, 0\}$. Comparing this capacity with the sum - capacity of the MAC, i.e., with $\frac{1}{2} \log(1 + M \cdot \text{SNR})$, we observe that we suffer from two losses⁸ (see also [10], [6], [8]). First, we lose a factor of M in power. This will be the dominant loss at high SNR. Second the “1” is replaced by $\frac{1}{M}$ which limits performance at low SNR⁹.

2) *Non-symmetric case*: Assume now that the users enjoy different SNRs. Without loss of generality, we label the users so that

$$\text{SNR}_1 \leq \text{SNR}_2 \leq \dots \leq \text{SNR}_M.$$

We also denote $\text{SNR}_{\min} = \text{SNR}_1$. It is now beneficial to allow the users to scale their inputs with a gain β_i that may be smaller than $\sqrt{P_{X,i}}$, i.e., we allow for some “power backoff”. This results in the effective channel,

$$\mathbf{Y} = \sum_{i=1}^M h_i \beta_i \cdot \mathbf{X}'_i + \mathbf{Z}. \quad (10)$$

⁸We note that these losses are not due to the modulo operation at the receiver and are encountered even when ML decoding is used at the receiver, see [6].

⁹Note also that at low SNR time sharing is beneficial.

Consider using a linear estimator $\hat{S} = \alpha \mathbf{Y}$ for $S = \sum_{i=1}^M k_i \mathbf{X}'_i$ where the k_i are integers to be chosen. Then the estimation error is given by

$$\mathbf{N} = \sum_{i=1}^M (k_i - \alpha \beta_i h_i) \cdot \mathbf{X}'_i + \alpha \mathbf{Z}.$$

Notice that we may always choose α and the gains β_i so as to ensure that $\alpha \beta_i h_i$ is a (non-zero) integer for all i . Indeed in the case of a single MAC (as opposed to a network setting), at high SNR, an optimal choice satisfies that $k_1 = 1$ (the choice of k_i , for $i > 1$ has no effect on C_Λ in this limit but does on the power consumed) and using

$$\alpha = 1/\sqrt{\text{SNR}_{\min}}$$

and

$$\beta_i = \sqrt{\text{SNR}_{\min}/\text{SNR}_i} \cdot \sqrt{P_{X_i}}.$$

With this choice we obtain a mod Λ channel with capacity $C_\Lambda = \frac{1}{2} \log\left(\frac{1}{M} + \text{SNR}_{\min}\right)$. For non-high SNRs, the choice of the vectors \mathbf{k} along with the vector of β_i and the value α involves a non trivial optimization similar to the one described in [8]. This optimization becomes even more involved when we consider a number of MACs as described in the previous section, as then we also have to require that the resulting equations be linearly independent.

We note that the “left over” power need not go unused. Using superposition coding along with “power backoff”, the “left over” power may be used successively to form layers of Gaussian MACs, each time with less users. That is, after the first layer, we will be left with a MAC with no more than $M - 1$ active users, then with $M - 2$ and so forth. The decoder will correspondingly use a successive decoding and stripping approach.

3) *Bi-Directional Gaussian relay*: Consider the bi-directional two-user Gaussian relay studied in [6], [1] where the “uplink” is a two-user MAC as described in (10) with $M = 2$, and where the “downlink” to user i is given by,

$$Y_i = h_i X + Z, \quad i = 1, 2$$

where Z_i is AWGN with unit power. The power of the relay is assumed to be at least $\max\{P_{X,1}, P_{X,2}\}$. Using the results of Section IV-B2 with the nested lattice approach described in Section III we recover the result of [6], [1] that a transfer rate of

$$2C_\Lambda = 2 \cdot \frac{1}{2} \log\left(\frac{1}{2} + \text{SNR}_{\min}\right)$$

is achievable, where we measure rate as total number of bits received by both users per channel use per *uplink* cycle.

C. On/off fading MAC

In this example we show how the recently proposed methods employing linear/lattice codes, when applied with the mod Λ transformation, may be extended to channel models which are not additive. If the channel is “not too far” from additive, methods based on linear coding may result in an achievable rate region which is larger than that obtained using “conventional”

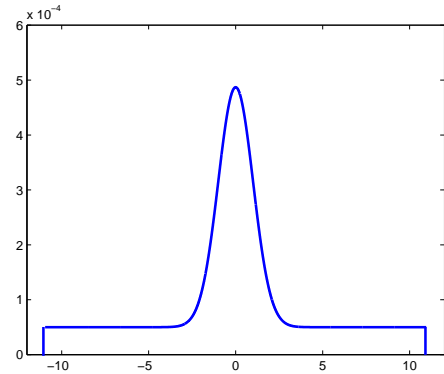


Fig. 4. PDF of effective noise in Example IV-C1.

approaches which do not exploit linearity. We demonstrate this by considering an on/off fading MAC.

1) *Single-user on/off fading channel*: Consider the channel,

$$Y = HX + Z, \quad (11)$$

where H is an i.i.d. Bernoulli random variable taking the value 0 with probability p_{out} and the value 1 with probability $1 - p_{\text{out}}$. The noise Z is AWGN with unit power and the input X is subject to power constraint P_X . We assume that H is unknown to both transmitter and receiver. A simple upper bound on the capacity of this channel is obtained by allowing the receiver knowledge of H , resulting in,

$$C \leq (1 - p_{\text{out}}) \cdot \frac{1}{2} \log(1 + P_X). \quad (12)$$

We now demonstrate how the channel may be transformed into an additive one. We use a very crude (far from optimal) choice of parameters. We take $\Lambda = A \cdot \mathbb{Z}$, with fundamental region is $\mathcal{V} = [-A/2, A/2)$ where $A = \sqrt{12P_X}$. We take the input to the channel to be,

$$X = f(X'_i) = X'_i.$$

Since the channel is a single-user channel, we define $S = X'$ and let the estimator be simply $\hat{S} = Y$. This results in effective noise,

$$\begin{aligned} N &= Y - X' \\ &= (H - 1)X + Z \end{aligned}$$

Note that the PDF of $N \bmod \Lambda$ is a mixture of a normal PDF (of unit variance) with weight $q = p_{\text{out}}$ and uniform PDF with weight $1 - q$. Figure 4 depicts the PDF of the effective noise for $p_{\text{out}} = 1/2$ and $P_X = 40$.

2) *Multiple access on/off fading channel*: We may generalize the channel (9) to an on/off fading MAC,

$$Y = H_1 X_1 + H_2 X_2 + Z. \quad (13)$$

For simplicity we consider only the symmetric case where H_1 and H_2 are both i.i.d. and independent of each other, both with outage probability p_{out} . We further assume that both users are

subject to the same power constraint P_X and that the noise Z is AWGN with unit variance. As for the single-user channel, a trivial upper bound on the capacity region is obtained by considering the case where the receiver has knowledge of both H_1 and H_2 . We only consider the sum rate in this example, for which this bound yields $C_{\text{sum}} \leq C_{\text{sum}}^U$, where,

$$C_{\text{sum}}^U = (1 - p_{\text{out}})^2 \cdot \frac{1}{2} \log(1 + 2P_X) + 2p_{\text{out}}(1 - p_{\text{out}}) \cdot \frac{1}{2} \log(1 + P_X). \quad (14)$$

Taking the exact same approach as for the single-user case, we let $X_i = X'_i$, $i = 1, 2$, and define $S = X'_1 + X'_2$ and $\hat{S} = Y$. This results in a mod Λ channel with effective noise,

$$N = (H_1 - 1)X_1 + (H_2 - 1)X_2 + Z.$$

Note that the PDF of $N \bmod \Lambda$ is as in the single-user case (depicted in Figure IV-C1) but now with $q = (1 - p_{\text{out}})^2$. The capacity of the resulting channel is

$$C_\Lambda = \frac{1}{2} \log(12P_X) - h(N). \quad (15)$$

It is not hard to show that (for M users), we have

$$C_\Lambda \geq (1 - p_{\text{out}})^M \cdot \frac{1}{2} \log(12P_X/2\pi e) - H_B(q).$$

Thus, as $p_{\text{out}} \rightarrow 0$, the gap to the upper bound behaves as that for the non-fading MAC. On the other hand, the gap becomes very large for large values of p_{out} , particularly so when M is large. In this sense, it is not surprising that the method proposed is of most value when the original channel is “not too far” from additive.

3) *Bi-directional relay channel with on/off fading*: Consider the bi-directional relay channel scenario of [6], [1] where we replace the Gaussian MAC with an “on/off fading” MAC. The “uplink” is described by (13) and the “downlink” is given by,

$$Y_i = H_i X + Z,$$

where both users as well as the relay are subject to the power constraint P_X . The capacity of the MAC (15) immediately provides $2C_\Lambda$ as lower bound on the capacity of this channel. Consider now a “traditional” approach where the relay decodes both messages and then relays the two messages in two (or one if conventional network coding at the bit/message level is used) communication rounds. A (loose) upper bound on the information exchange rate is then C_{sum}^U where C_{sum}^U is given in (14). A simple calculation shows that for example, for $p_{\text{out}} = 5\%$, the network coding scheme using the mod Λ transformation of Section IV-C2 outperforms the traditional approach (as evaluated by C_{sum}^U) for P_X greater than ≈ 7 dB.

V. ACKNOWLEDGEMENT

The authors are grateful to Yuval Kochman for helpful discussions.

REFERENCES

- [1] W. Nam, S.-Y. Chung, and Y. H. Lee, Capacity bounds for two-way relay channels, in *International Zurich Seminar on Communications*. In *IZS 2008, Zurich, Switzerland*, March 2008.
- [2] U. Erez and R. Zamir, Achieving $\frac{1}{2} \log(1 + \text{SNR})$ on the AWGN channel with lattice encoding and decoding. In *IEEE Trans. Information Theory*, IT-50, pp. 2293-2314, Oct. 2004.
- [3] G. D. Forney, Jr., On the role of MMSE estimation in approaching the information-theoretic limits of linear Gaussian channels: Shannon meets Wiener, In *Proceedings of 2003 Allerton Conf. (Monticello, IL)*, Oct. 2003.
- [4] D. Krithivasan and S. Pradhan, Lattices for distributed source coding: Jointly Gaussian sources and reconstruction of a linear function *IEEE Transactions on Information Theory*, submitted July 2007. See <http://arxiv.org/abs/0707.3461>.
- [5] H. A. Loeliger, Averaging bounds for lattices and linear codes. In *IEEE Trans. Information Theory*, IT-43, pp. 1767-1773, November 1997.
- [6] K. Narayanan, M. P. Wilson and A. Sprintson, Joint physical layer coding and network coding for bi-directional relaying. In *45th Annual Allerton Conference*, Monticello, IL, Sept., 2007.
- [7] B. Nazer and M. Gastpar, The case for structured random codes in network capacity theorems. In *Euro. Trans. Telecomm.* To appear.
- [8] B. Nazer and M. Gastpar, Compute-and-forward: Harnessing interference with structured codes. In *Proceedings of ISIT 2008*, July 6-11, Toronto, Canada.
- [9] B. Nazer and M. Gastpar, Compute-and-Forward: A Novel Strategy for Cooperative Networks. In *Proceedings of the 42nd Annual IEEE Asilomar Conference on Signals, Systems, and Computers*, Monterey, CA, October 2008.
- [10] T. Philosof, A. Khisti, U. Erez and R. Zamir, Lattice strategies for the dirty multiple access channel. In *Proceedings of ISIT 2007*, June 24-29, Nice, France.