Expurgation for Discrete Multiple-Access Channels via Linear Codes

Eli Haim
Dept. of EE-Systems, TAU
Tel Aviv, Israel
Email: elih@eng.tau.ac.il

Yuval Kochman
School of CSE, HUJI
Jerusalem, Israel
Email: yuvalko@cs.huji.ac.il

Uri Erez*
Dept. of EE-Systems, TAU
Tel Aviv, Israel
Email: uri@eng.tau.ac.il

***Abstract*—We consider the error exponent of the memoryless multiple-access (MAC) channel. We show that if the MAC channel is modulo-additive, then any error probability, and hence any error exponent, achievable by a linear code for the corresponding single-user channel, is also achievable for the MAC channel. Specifically, for an alphabet of prime cardinality, where linear codes achieve the best known exponents in the single-user setting (and the optimal exponent above the critical rate), this performance carries over to the MAC setting. At least at low rates, where expurgation is needed, our approach strictly improves performance over previous results, where expurgation was used at most for one of the users. Even when the MAC channel is not additive, it may be transformed into such a channel. While the transformation is lossy, we show that the distributed structure gain in some "nearly additive" cases outweighs the loss, and thus we can improve upon the best known exponent for these cases as well. This approach is related to that previously proposed for the Gaussian MAC channel, and is based on "distributed structure".**

I. Introduction

The error exponent of the multiple access (MAC) channel is a long-standing open problem. While superposition and successive decoding methods lead to capacity, they may not be optimal in the sense of error probability: the decoding process may be improved by considering that the transmission of other users is a codeword, rather than noise. However, finding the optimal performance is a difficult task. Early results include the works of Slepian and Wolf [1], Gallager [2] and Pokorny and Wallmeier [3]. Applying the results of [2] to the important special case of a (modulo) additive MAC channel, e.g., the binary symmetric case, it follows that the random-coding exponent of the corresponding single-user channel is achievable for the MAC channel. This exponent is optimal above the critical rate [4] and gives the best known performance above the expurgation rate. However, for yet lower rates it is outperformed by the expurgated exponent (in the single-user case). The reason that the expurgated exponent is not achieved in [2] is that the sum of such two good (expurgated) single-user codebooks does not result in a good single-user one, and in particular, the sum of two codebooks with good minimum-distance properties may not be good in that respect. Liu and Hughes [5] and recently Nazari et al. [6] have proposed improvements over the previous results. Specifically, Nazari et al. suggest to use expurgation on one of the codebooks. While this certainly improves performance, it still does not allow to achieve the single-user expurgated exponent for additive MAC channels.

In [7] the exponent of a Gaussian MAC channel is considered. It is suggested to use "distributed structure": the users use lattice codebooks, where the codebook of one user is nested in that of the other. This scheme has the advantage that the sum of the codebooks, as seen by the decoder, is a single linear code; since linear codes are inherently expurgated, the exponent at low rates is improved. However, the exponent obtained is inferior to the single-user exponent. It can be explained by viewing the joint codebook as tiling of the codebook of the user with weaker power (finest lattice) in the shaping region of the associated single-user channel. There is a loss since it is not a perfect tiling, i.e. not filling this whole shaping region.

Favorably, when considering a discrete memoryless MAC channel without a cost constraint, the situation is more simple. For additive MAC channels we make the basic observation, that by "splitting" a linear codebook between the users, any error probability achievable in the corresponding single-user channel using linear codebooks is achievable for the MAC channel as well. This implies for prime (e.g. binary) alphabets, that the best currently known single-user error exponents for *any* code (not necessarily linear) are achievable for the MAC channel, including the random-coding and expurgated exponents. The improvement over previous results stems from the use of linear codes, which are inherently expurgated; thus using them provides "joint expurgation" even in a distributed setting. In comparison, [2] performs no expurgation, and in [6] expurgation is performed only for one user.

But what happens outside the special case of additive channels? We are inspired by the fact that in the context of first-order (capacity) analysis of networks, the advantage of linear codes has indeed been extended to some non-additive channels [8].

The application of linear codes to additive communication networks has a capacity advantage in many interesting scenarios, see e.g. [9], [10], [11]. In [8], a modulo-lattice transformation is derived, that allows to obtain a virtual additive channel from any original MAC channel, albeit with a loss of capacity. It is shown in [8] that in some situations, the gain offered by the ability to use linear codes outweighs the loss inflicted by

* This work was supported in part by the U.S. - Israel Binational Science Foundation under grant 2008/455.

978-1-4673-2579-0/12/$31.00 ©2012 IEEE

the transformation. In this work we adopt the same ideas to the MAC exponent problem: we show that for MAC channels that are "nearly additive", indeed the transformation improves over the best known exponents so far at low rates. We note that when one considers less symmetric channels, the results of [6] outperform those of the new scheme.

II. Channel Model and Exponents

A. Single-User Channel

Consider the single-user discrete memoryless channel (DMC) defined by $P_{Y|X}(\cdot|\cdot)$, where X and Y are the channel input and output, respectively, with discrete alphabets $\mathcal{X}$ and $\mathcal{Y}$. We recall some results regarding the error exponent of this channel, see [4].

The error exponent is defined as

$$E^{\text{SU}}(R) = \limsup_{n\to\infty} -\frac{1}{n}\log \epsilon_n, \tag{1}$$

where ϵ_n is the minimal possible error probability of codes (averaged over the codewords) with block length n and rate R.

The best known achievable error exponent for this channel, denoted by $\underline{E}^{\text{SU}}(R)$, is given by the maximum between the expurgated error exponent $E_{ex}^{\text{SU}}(R)$ and the random-coding error exponent $E_r^{\text{SU}}(R)$. The expurgated exponent is larger than the random-coding exponent below some rate R_{ex} (this range is thus called "the expurgation region"). Above the critical rate R_{cr}, the random-coding exponent is known to be optimal.

B. MAC Channel

Consider a two-user discrete memoryless MAC channel $P_{Y|X_1,X_2}$, where X_1, X_2 are the channel inputs and Y is its output. Denote the codebook of user i by $\mathcal{C}_i$, and its rate by $R_i = {}^1\!/_n \log|\mathcal{C}_i|$.

The error event is defined as the event that at least one of the messages from the message pair is decoded in error.[1] The error exponent of the MAC channel is defined as

$$E^{\text{MAC}}(R_1, R_2) = \limsup_{n\to\infty} -\frac{1}{n}\log \epsilon_n, \tag{2}$$

where ϵ_n is the minimal possible error probability for codes of length n, with the rate-pair (R_1, R_2).

Gallager [2] found an achievable error exponent that is the minimum of three random coding error exponents corresponding to different error events. The first two correspond to making a erroneous decision on one message, by a "genie" aided decoder, i.e., one that has knowledge of the message of the other user as side information. The third error event corresponds to making an erroneous decision regarding the two messages as a combined message. Each of these amounts to an error event over a single-user channel. Therefore, each exponent is equal to by Gallager's random coding error exponent [4] for the associated single-user channel.

[1]Other definitions, leading to an error exponent region, were considered in [12].

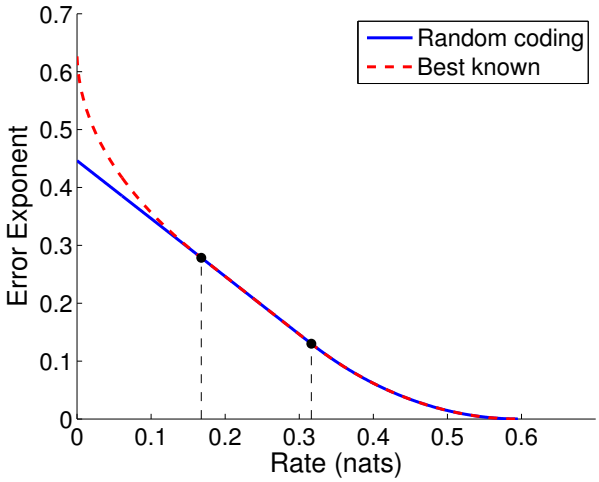

Fig. 1. Comparing the random coding error exponent of a additive-noise single-user channel, with the best known error exponent. The channel is additive binary MAC with noise $\sim$ Bernulli(0.02). The two dots show the expurgation rate and the critical rate of this channel respectively.

The best known error-exponent of this channel is given by Nazari et al. [6]. This bound however is not given in closed form and is hard to compute.

C. Additive-Noise Single-User Channel

Consider the following channel:

$$Y = X \oplus N, \tag{3}$$

where all variables are defined over the alphabet $\mathbb{Z}_m = \{0, 1, \ldots, m-1\}$ and $\oplus$ denotes addition over this alphabet, i.e., modulo an integer number m. The noise N is additive, i.e., statistically independent of the channel input X.

In the expurgation region, the best known error exponent of this channel is larger than the random coding error exponent, as can be seen in Figure 1.

In the context of additive channels, it is important to consider *linear* codes. We define a linear code $\mathcal{C}$ via a $k \times n$ generating matrix G,[2] by

$$\mathcal{C} = \{\mathbf{c} : \mathbf{c} = \mathbf{u}G,\ \mathbf{u} \in \mathbb{Z}_m^k\}, \tag{4}$$

The rate is equal to $R = {}^k\!/_n \cdot \log m$. Clearly, every rate is possible asymptotically as $n \to \infty$. We define $E_{\text{L}}^{\text{SU}}(R)$ to be the error exponent of linear codes, i.e., as (1), except that ϵ_n is the minimal possible error probability of *linear* codes only. We note that for single-user additive channels of the form (3), when the alphabet size m is a prime, the best known error exponent of linear codes, denoted by $\underline{E}_{\text{L}}^{\text{SU}}(R)$, is equal to the best known error exponent of the channel $\underline{E}^{\text{SU}}(R)$ (see [13], [4], [14]), and in particular is optimal above the critical rate. In addition, we note that for linear codes the average error probability (over the codewords) is equal to the maximal error probability, due to their structure.

[2]We assume a full-rank matrix G.

D. Additive-Noise MAC Channel

A channel which is of particular interest in this work is the additive MAC channel

$$Y = X_1 \oplus X_2 \oplus N, \tag{5}$$

where all variables are defined over the alphabet $\mathbb{Z}_m = \{0, 1, \ldots, m-1\}$ and $\oplus$ denotes addition over this alphabet, i.e., modulo an integer number m. The noise N is additive, i.e., statistically independent of the pair (X_1, X_2).

Viewing the joint codebook $X = X_1 \oplus X_2$ as a single-user codebook, we get the following channel (over the same alphabet)

$$Y = X \oplus N, \tag{6}$$

which we call the *associated single-user channel* of (5). Notice that when comparing the MAC channel to its associated single-user one,

$$E^{\text{MAC}}(R_1, R_2) \leq E^{\text{SU}}(R_1 + R_2),$$

since the last is equivalent to cooperation between the encoders. In [2] it is shown that[3]

$$E^{\text{MAC}}(R_1, R_2) \geq E_r^{\text{SU}}(R_1 + R_2).$$

Thus, it is equal to $\underline{E}^{\text{SU}}(R_1 + R_2)$ above the expurgation rate and optimal above the critical rate. However, the best known error exponent for the associated single-user though, is larger in the expurgation region, as was shown in Figure 1.

We note that simple time sharing, where every user uses an expurgated codebook, improves on Gallager's random-coding bound in some cases, particularly for small enough rates and as the channel noise become smaller. Since [2], there were several improvements [5], [6] to the achievable error exponent. However these do not close the gap to the best known error exponent of the associated single-user channel. In the next section, we close this gap (for modulo-additive MAC channels) by attaining expurgation for all users.

III. Coding for Modulo-Additive Discrete MAC Channels

In this section we first describe a coding scheme for additive-noise MAC channels with prime alphabet size, which achieves the best known error exponent of its associated single-user channel. This is equivalent to full cooperation of the encoders, and thus it is optimal (in terms of error exponent) whenever the optimum is known for the single-user channel (i.e., above its critical rate).

Consider the additive-noise MAC channel, as given in (5), with prime alphabet size m. We construct a codebook pair for the MAC channel using linear codes. We use a good linear code for the associated single-user channel (6), which we decompose into two sub linear codes, one for each user.

Let G be a $k \times n$ generating matrix of a linear code $\mathcal{C}$ (see (4)) with rate $R = {}^{k}\!/_{n} \cdot \log m$. For some integers $k_1 + k_2 = k$, define the rates

$$R_i = \frac{k_i}{n} \log m, \ i = 1, 2.$$

Decompose the codeword $\mathbf{c}$ into two codewords:

$$\mathbf{c} = \mathbf{c}_1 \oplus \mathbf{c}_2 = (\mathbf{u}_{1\times k_1} \mid \mathbf{u}_{1\times k_2}) \begin{pmatrix} G_{k_1 \times n} \\ G_{k_2 \times n} \end{pmatrix} \tag{7}$$

$$\stackrel{\triangle}{=} \mathbf{u}_1 G_1 \oplus \mathbf{u}_2 G_2. \tag{8}$$

Thus, we have a pair of codebooks:

$$\mathcal{C}_i = \{\mathbf{c}_i : \mathbf{c}_i = \mathbf{u}_i G_i, \ \mathbf{u}_i \in \mathbb{Z}_m^{k_i}\}; \quad i = 1, 2. \tag{9}$$

Therefore, the sum of codewords is indistinguishable from a codeword of the single-user code with $R = R_1 + R_2$. Clearly, for every rate pair such a construction is possible asymptotically as $n \to \infty$. A similar claim holds for a general number of users as well.

Proposition 1: The coding technique above achieves the best error probability of linear codes for the single-user channel. Thus, it achieves the exponent $\underline{E}_{\text{L}}^{\text{SU}}(R)$, and for prime m it achieves $\underline{E}^{\text{SU}}(R)$ as well.

The previously best known error exponent for general memoryless discrete MAC channels is given by Nazari et al. [6]. In their derivation, codewords are expurgated from only one of the codebooks. Since our bound achieves the best known error exponent of the associated single-user channel (for the special case of additive MAC channels with prime alphabet size), it must be at least as good the one found by Nazari et al. Moreover, for small enough rate-pairs we expect our bound to be strictly better, since full expurgation is required in order to achieve the error exponent of the associated single-user channel. When considering more than two users, the gap is expected to increase since expurgation of one user becomes less significant. In the sequel, we show how this advantage can be leveraged to non-additive MAC channels.

IV. Transforming a General Discrete MAC Channel into an Additive Channel

With the aim of applying a similar scheme to general (non-additive) discrete memoryless MAC channels $P_{Y|X_1,X_2}$, in this section we describe a method for transforming such channels into additive-noise MAC channels. We denote the obtained channel after the transformation as the resulting *virtual* channel. The transformation is a discrete and scalar modification of the *Modulo-Lattice Transformation* for continuous MAC channels [8].

The transformation is defined for any finite alphabet size m. Let $v_i \in \mathbb{Z}_m$ be the input of the ith user to the virtual channel, and $U_i \sim \text{Uniform}(\mathbb{Z}_m)$ be its dither (i.e., common randomness at the ith transmitter and at the receiver), which is statistically independent of the dither of the other user and of v_1, v_2. Each encoder computes $X_i' = v_i \oplus U_i$ and applies a scalar precoding function $f_i : \mathbb{Z}_m \to \mathcal{X}$ on it. The inputs to the channel are therefore given by

$$X_i = f_i(X_i'). \tag{10}$$

[3] Since it is shown that for the discrete additive MAC channel, out of the three error exponents discussed above, the third always dominates.

Note that due to the dither, X_i' is uniformly distributed over $\mathbb{Z}_m$ and is statistically independent of v_1, v_2. Let

$$S = k_1 X_1' + k_2 X_2',$$

where $k_i \in \mathbb{Z}_m$, and let $\hat{S} = g(Y)$ be some scalar estimator function of S from the channel output Y. Denote the estimation error by $N = \hat{S} - S$. We define the output of the virtual channel as

$$\begin{aligned} Y' &\triangleq [\hat{S} - (k_1 U_1 + k_2 U_2)] \bmod m \\ &= [\hat{S} - S + S - (k_1 U_1 + k_2 U_2)] \bmod m \\ &= [(k_1(v_1 + U_1) \bmod m) + (k_2(v_2 + U_2) \bmod m) \\ &\quad + N - (k_1 U_1 + k_2 U_2)] \bmod m \\ &= (k_1 v_1 + k_2 v_2 + N) \bmod m \\ &= k_1 v_1 \oplus k_2 v_2 \oplus (N \bmod m) \\ &= k_1 v_1 \oplus k_2 v_2 \oplus \tilde{N}, \end{aligned}$$

where $\tilde{N} = N \bmod m$.

Proposition 2 (The virtual MAC channel): Applying the transformation leads to the following virtual channel:

$$Y' = k_1 v_1 \oplus k_2 v_2 \oplus \tilde{N}, \tag{11}$$

where $N = \hat{S} - S$ is statistically independent of the channel inputs (v_1, v_2), and $\tilde{N} = N \mod m$.

Notice that the transformation is not unique, and one is free to choose the alphabet size m, the precoding functions $f_i(\cdot)$ and the estimator of S. We call any virtual MAC channel (11) that can be obtained by some choice of parameters, a feasible virtual MAC channel. Applying this transformation to any MAC channel, we have the following.

Proposition 3: Let ϵ_n be the best error probability achievable with a code of length n on a MAC channel. Then

$$\epsilon_n \leq \tilde{\epsilon}_n,$$

where $\tilde{\epsilon}_n$ is the best error probability achievable by a *linear* code of the same length on a feasible virtual MAC channel (11).

A feasible single-user channel:

$$Y = X \oplus \tilde{N}, \tag{12}$$

is the associated single-user channel of a feasible virtual MAC channel (11). Applying Proposition 3 to exponents, leads to our main result:

Theorem 1: For any MAC channel, and any feasible single-user channel (12) of that channel,

$$E^{\text{MAC}}(R_1, R_2) \geq \underline{E}_{\text{L}}^{\text{SU}}(R_1 + R_2). \tag{13}$$

This has great significance when linear codes are known to perform well:

Corollary 1: For any MAC channel, and any feasible single-user channel (12) of that channel with alphabet of prime cardinality,

$$E^{\text{MAC}}(R_1, R_2) \geq \underline{E}^{\text{SU}}(R_1 + R_2).$$

Remarks:

- Notice that this transformation is lossy in terms of capacity. However, since the resulting channel is an additive-noise channel, efficient coding techniques and known bounds can be easily applied. In particular, for MAC channels, expurgation in all the users can be applied by using linear codes as in Section III.
- We expect the benefit from this coding technique to outweigh the loss when the channel is "close" to additive. In the next section, after studying the binary case, we give a binary example which illustrates this property with a single parameter.
- We note that this transformation is applicable to various non-additive network problems, where structure can improve the best-known achievable *rate region* (see e.g. [9], [10], [11]). In such settings, the gain will appear also as a "capacity gain" rather than only in the error exponent.

V. Binary Case

In this section we confine the discussion to binary MAC channels, i.e. channels with binary inputs and a binary output. We denote this general channel $P_{Y|X_1,X_2}$ as:

$$Y = X_1 \oplus X_2 \oplus N, \tag{14}$$

where the additive noise $N = Y \oplus (X_1 \oplus X_2)$ *depends* on the channel input pair (X_1, X_2).

A. Analysis of the Virtual Channel

A natural choice for the parameter m of the transformation is clearly $m = 2$. We select $f(x) = x$, $k_1 = k_2 = 1$ and $\hat{S} = g(Y) = Y$. This selection leads to the following effective additive noise at the virtual channel is

$$\tilde{N} \sim \text{Bernulli}(p), \tag{15}$$

with

$$p = \frac{1}{4} \sum_{x_1, x_2 \in \{0,1\}} \Pr(N = 1 | X_1 = x_1, X_2 = x_2). \tag{16}$$

The virtual channel is then an additive MAC channel given by:

$$Y = V_1 \oplus V_2 \oplus \tilde{N}, \tag{17}$$

where $\tilde{N}$, given in (15)-(16), is statistically independent of (V_1, V_2).

B. Example: Almost Additive Binary MAC Channel

We now use the analysis of the previous subsection in order to study an example of an almost additive-noise binary MAC channel. Specifically, we consider the MAC channel characterized by Table I, which gives its transition probabilities.

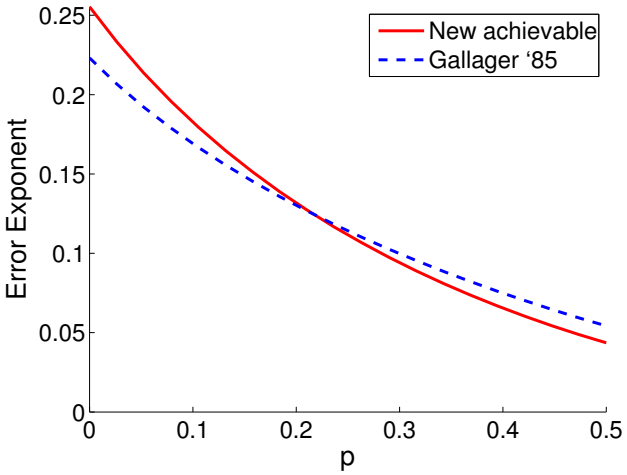

Fig. 2. Comparing error exponents for the almost additive binary MAC in Table I (at zero rate). The dashed line is Gallager's bound [2], while the solid line is the error exponent of the virtual channel, which is achieved according to Corollary 1. Here $q = 0.1$.

x_1	x_2	$P_{Y\|X_1,X_2}(1\|x_1,x_2)$	$P_{N\|X_1,X_2}(1\|x_1,x_2)$
0	0	q	q
0	1	$1-[p(1-q)+(1-p)q]$	$p(1-q)+(1-p)q$
1	0	$1-[p(1-q)+(1-p)q]$	$p(1-q)+(1-p)q$
1	1	q	q

TABLE I
ALMOST ADDITIVE BINARY MAC. THE VALUE OF p DETERMINES THE DEVIATION OF THE CHANNEL FROM ADDITIVITY.

The value of p determines the deviation of the channel from additivity. For small p the channel is nearly an additive MAC channel.

In Figure 2 we compare the resulting error exponent of the virtual channel with Gallager's [2] random coding exponent for symmetric rate pairs.[4] For the comparison we take the limit of zero rate-pair, where the gain due to expurgation is maximal. As p increases, the coding technique developed in this paper gains less since the channel transformation looses more as the channel is less additive.

For small enough p and for small enough rates we expect this bound to be strictly larger then the best known error exponent for this channel [6]. This is since [6] applies expurgation only to the user with larger rate, while the bound presented here achieves two-user expurgation. Nazari et al. [6] studied a non-symmetric example, where $\Pr(N = 1|X_1 = 1, X_2 = 1) = \frac{1}{2}$, and all the other conditional probabilities of N are equal to 0.01. In this case the coding scheme described here is inferior to the one of [6], as expected since the channel is far from being an additive.

[4]For the channel parameter of Fig. 2, the exponent of time sharing between expurgated codebooks is below Gallager's random coding exponent.

VI. CONCLUSION

By using linear codes, we have shown that for modulo-additive MAC channels with a prime alphabet size, the achievable error exponent is equal to the best known exponent of the associated single-user channel. In addition, we have demonstrated that linear codes offer improvement to the best known MAC error exponent region for "almost additive" channels.

While we chose to present the results using two-user MAC channels, the approach immediately extends to any number of users, allowing for full expurgation of the combined linear code. It is therefore reasonable to expect that at low rates and at least for modulo-additive channels, the gain over the best previously known achievable error exponent will increase with the number of users. The approach of transforming a MAC channel into an additive one is also applicable to a wide variety of non-additive network setups, where structure is beneficial in terms of capacity.

REFERENCES

[1] D. Slepian and J. K. Wolf, "A coding theorem for multiple access channels with correlated sources," *Bell System Tech. J.*, vol. vol. 52, pp. 1037–1076, Sept. 1973.
[2] R. G. Gallager, "A perspective on multiaccess channels," *IEEE Trans. Information Theory*, vol. 31, no. 2, pp. 124 – 142, Mar. 1985.
[3] J. Pokorny and H. Wallmeier, "Random coding bound and codes produced by permutations for the multiple-access channel," *IEEE Trans. Information Theory*, vol. 31, no. 6, pp. 741 – 750, nov 1985.
[4] R. G. Gallager, *Information Theory and Reliable Communication*. John Wiley & Sons, 1968.
[5] Y.-S. Liu and B. Hughes, "A new universal random coding bound for the multiple-access channel," *IEEE Trans. Information Theory*, vol. 42, no. 2, pp. 376 –386, Mar. 1996.
[6] A. Nazari, A. Anastasopoulos, and S. S. Pradhan, "Error exponent for multiple-access channels:lower bounds," *CoRR*, vol. abs/1010.1303, 2010.
[7] E. Haim, Y. Kochman, and U. Erez, "Improving the MAC error exponent using distributed structure," in *Proc. Int. Symp. Info. Theory*, Aug. 2011.
[8] U. Erez and R. Zamir, "A modulo-lattice transformation for multiple-access channels," in *Electrical and Electronics Engineers in Israel, 2008. IEEEI 2008. IEEE 25th Convention of*, Dec. 2008, pp. 836 –840.
[9] D. Krithivasan and S. Pradhan, "Lattices for distributed source coding: Jointly Gaussian sources and reconstruction of a linear function," *IEEE Trans. Information Theory*, vol. 55, no. 12, pp. 5628 –5651, 2009.
[10] M. P. Wilson, K. R. Narayanan, H. D. Pfister, and A. Sprintson, "Joint physical layer coding and network coding for bidirectional relaying," *IEEE Trans. Information Theory*, vol. IT-56, pp. 5641–5654, Nov. 2010.
[11] T. Philosof, R. Zamir, U. Erez, and A. Khisti, "Lattice strategies for the dirty multiple access channel," *IEEE Trans. Information Theory*, vol. IT-57, pp. 5006–5035, Aug. 2011.
[12] L. Weng, S. S. Pradhan, and A. Anastasopoulos, "Error exponent regions for Gaussian broadcast and multiple-access channels," *IEEE Trans. Information Theory*, vol. IT-54, pp. 2919–2942, Jul. 2008.
[13] R. L. Dobrushin, "Asymptotic optimality of group and systematic codes for some channels," *Theor. Probab. Appl.*, vol. 8, pp. 52–66, 1963.
[14] A. Barg and J. Forney, G.D., "Random codes: minimum distances and error exponents," *IEEE Trans. Information Theory*, vol. 48, no. 9, pp. 2568 – 2573, sep 2002.