Random coding error exponents for the two-user interference channel

Conference Paper

Author(s):
Huleihel, Wasim; Merhav, Neri

Publication date:
2016

Permanent link:
https://doi.org/10.3929/ethz-a-010645886

Random Coding Error Exponents for the Two-User Interference Channel

Wasim Huleihel and Neri Merhav
Technion - Israel Institute of Technology
Department of Electrical Engineering
Haifa 3200003, ISRAEL
E-mail: {wh@campus, merhav@ee}.technion.ac.il

***Abstract*—This paper is about deriving lower bounds on the error exponents for the two-user interference channel under the random coding regime for several ensembles. Specifically, we first analyze the standard random coding ensemble, where the codebooks are comprised of independently and identically distributed (i.i.d.) codewords. For this ensemble, we focus on optimum decoding, which is in contrast to other, suboptimal decoding rules that have been used in the literature (e.g., joint typicality decoding, treating interference as noise, etc.). The fact that the interfering signal is a codeword, rather than an i.i.d. noise process, complicates the application of conventional techniques of performance analysis of the optimum decoder. Also, unfortunately, these conventional techniques result in loose bounds. Using analytical tools rooted in statistical physics, as well as advanced union bounds, we derive single-letter formulas for the random coding error exponents. We compare our results with the best known lower bound on the error exponent, and show that our exponents can be strictly better. Then, in the second part of this paper, we consider more complicated coding ensembles, and find a lower bound on the error exponent associated with the celebrated Han-Kobayashi (HK) random coding ensemble, which is based on superposition coding.**

Keywords*—Random coding, error exponent, interference channels, superposition coding, Han-Kobayashi scheme, statistical physics, optimal decoding, multiuser communication.

I. Introduction

A. Previous Work

The two-user interference channel (IFC) models a general scenario of communication between two transmitters and two receivers (with no cooperation at either side), where each receiver decodes its intended message from an observed signal, which is interfered by the other user, and corrupted by channel noise. The information-theoretic analysis of this model has begun over more than four decades ago and has recently witnessed a resurgence of interest. Most of the previous work on multiuser communication, and specifically, on the IFC, has focused on obtaining inner and outer bounds to the capacity region (see, for example, [1, Ch. II.7]). In a nutshell, the study of this kind of channel has started in [2], and continued in [3], where simple inner and outer bounds to the capacity region were given. Then, in [4], by using the well-known superposition coding technique, the inner bound of [3] was strictly improved. In [5], various inner and outer bounds were obtained by transforming the IFC model into some multiple-access or broadcast channel. Unfortunately, the capacity region for the general interference channel is still unknown, although it has been solved for some very special cases [6, 7]. The best known inner bound is the Han-Kobayashi (HK) region, established in [8], and which will also be considered in this paper.

To our knowledge, [9, 10] are the only previous works which treat the error exponents for the IFC under optimal decoding. Specifically, [9] derives lower bounds on error exponents of random codebooks comprised of i.i.d. codewords uniformly distributed over a given type class, under maximum likelihood (ML) decoding at each user, that is, optimal decoding. Contrary to the error exponent analysis of other multiuser communication systems, such as the multiple access channel [11], the difficulty in analyzing the error probability of the optimal decoder for the IFC is due to statistical dependencies induced by the interfering signal. Indeed, for the IFC, the marginal channel determining each receiver's ML decoding rule is induced also by the codebook of the interfering user. This indeed extremely complicates the analysis, mostly because the interfering signal is a codeword and not an i.i.d. process. Another important observation, which was noticed in [9], is that the usual bounding techniques (e.g., Gallager's bounding technique) on the error probability fail to give tight results. To alleviate this problem, the authors of [9], combined some of the ideas from Gallager's bounding technique [12] to get an upper bound on the average probability of decoding error under ML decoding, the method of types [13], and used the method of distance enumerators, in the spirit of [14], which allows to avoid the use of Jensen's inequality in some steps. Finally, another relevant work is [15], where lower bounds on the error exponents of both standard and cognitive multiple-access channels (MACs), were derived assuming suboptimal successive decoding scheme.

B. Contributions

The main purpose of this paper is to extend the study of achievability schemes to the more refined analysis of error exponents achieved by the two users, similarly as in [9]. Specifically, we derive single-letter expressions for the error exponents associated with the average error probability, for the finite-alphabet two-user IFC, under several random coding ensembles. The main contributions of this paper are as follows:

• Similarly as in recent works (see, e.g., [11, 16-19] and references therein) on the analysis of error exponents, we derive single-letter lower bounds for the random coding error exponents. For the standard random coding ensemble, considered in Subsection II-B, we analyze the optimal decoder for each receiver, which is interested solely in its intended message. This is in contrast to usual decoding techniques analyzed for the IFC, in which each receiver decodes, in addition to its intended message, also part of (or all) the interfering codeword (that is, the other user's message), or other conventional achievability arguments [1, Ch. II.7], which are based on joint-typicality decoding, with restrictions on the decoder (such as, "treat interference as noise" or to "decode the interference"). This enables us to understand whether there is any significant degradation in performance due to the sub-optimality of the decoder. Also, since [9] analyzed the optimal decoder as well, we compare our formulas with those of [9], and show that our error exponent can be strictly better, which implies that the bounding technique in [9] is not tight. It is worthwhile to mention that the analytical formulas of our error exponents are simpler than the lower bound of [9].

• As was mentioned earlier, in [9] only random codebooks comprised of i.i.d. codewords (uniformly distributed over a type class) were considered. These ensembles are much simpler than the superposition codebooks of [8]. Unfortunately, it very tedious to analyze superposition codebooks using the methods of [9], and even if we do so, the tightness is questionable. In this paper, however, the new tools that we have derived enable us to analyze more involved random coding ensembles. Indeed, we can consider the coding ensemble used in the HK achievability scheme [8] and derive the respective error exponents. We also discuss an ensemble of hierarchical/tree codes [20].

• The analysis of the error exponents, carried out in this paper, turns out to be much more difficult than in previous works on point-to-point and multiuser communication problems, see, e.g., [11, 16-19]. Specifically, we encounter two main difficulties in our analysis: First, typically, when analyzing the probability of error, the first step is to apply the union bound. Usually, for point-to-point systems, under the random coding regime, the average error probability can be written as a union of pairwise independent error events. Accordingly, in this case, it is well known that the truncated union bound is exponentially tight [21, Lemma A.2]. This is no longer the case, however, when considering multiuser systems, and in particular, the IFC. For the IFC, the events comprising the union are strongly dependent, especially due to the fact that we are considering the optimal decoder. Indeed, recall that the optimal decoder for the first user, for example, declares that a certain message was transmitted if this message maximizes the likelihood pertaining to the marginal channel. This marginal channel[1] is the average of the actual channel over the messages of the interfering user, and thus depends on the whole codebook of the that user. Accordingly, the overall error event is the union of an exponential number of error events where each event depends on the marginal channel, and thus on the codebook of the interfering user. To alleviate this difficulty, following the ideas of [11], we derived new upper bounds on the probability of a union of events, which take into account the dependencies among the events. The second difficulty that we have encountered in our analysis is that in contrast to previous works, applying the type class enumerator method [14] is not simple, due to the reason mentioned above. Using some methods from large deviations theory, we were able to tackle this difficulty.

[1]The precise definition will be given in the sequel.

• Recently, in [15], the authors independently suggested lower bounds on the error exponents of both standard and cognitive multiple-access channels (MACs), assuming suboptimal successive decoding scheme, and using the standard random coding ensemble (considered in Subsection II-B). Although the motivation in [15] is different, the codebook construction and the decoding rule are the same as in the first part of this paper, and thus, essentially, their results apply also for the IFC. Now, despite the fact that the analysis in our paper is not the same as in [15], for the standard random coding ensemble, our lower bound coincides with that of [15]. More importantly, as was mentioned above, we consider also the more complicated ensemble pertaining to the HK scheme. Accordingly, the derivation of the lower bound on the error exponent of this ensemble is built upon the derivation of the lower bound on the error exponent of the standard random coding ensemble, and thus it makes useful and convenient to start with the analysis of the latter ensemble. We emphasize that the techniques used in [15] are not sufficient to analyze the ensemble pertaining to the HK scheme. Finally, we mention that the focus in [15] was on achievable rate region, rather than the error exponents, and thus no comparison to [9] was provided.

• We believe that by using the techniques and tools derived in this paper, other multiuser systems, such as the IFC with mismatched decoding, the MAC [11], the broadcast channel, the relay channel, etc., and accordingly, other coding schemes, such as binning [16], and hierarchical codes [20], can be analyzed.

C. Notation Conventions

Throughout this paper, scalar random variables (RVs) will be denoted by capital letters, their sample values will be denoted by the respective lower case letters, and their alphabets will be denoted by the respective calligraphic letters, e.g. X, x, and $\mathcal{X}$, respectively. A similar convention will apply to random vectors of dimension n and their sample values, which will be denoted with the same symbols in the boldface font. We also use the notation X_i^j $(j > i)$ to designate the sequence of RVs $(X_i, X_{i+1}, \ldots, X_j)$. The set of all n-vectors with components taking values in a certain finite alphabet, will be denoted as the same alphabet superscripted by n, e.g., $\mathcal{X}^n$. Generic channels will be usually denoted by the letters P, Q, or W. We shall mainly consider joint distributions of two RVs (X, Y) over the Cartesian product of two finite alphabets $\mathcal{X}$ and $\mathcal{Y}$. For brevity, we will denote any joint distribution, e.g. Q_{XY}, simply by Q, the marginals will be denoted by Q_X and Q_Y, and the conditional distributions will be denoted by $Q_{X|Y}$ and $Q_{Y|X}$. The joint distribution induced by Q_X and $Q_{Y|X}$ will be denoted by $Q_X \times Q_{Y|X}$, and a similar notation will be used when the roles of X and Y are switched.

The expectation operator will be denoted by $\mathbb{E}\{\cdot\}$, and when we wish to make the dependence on the underlying distribution Q clear, we denote it by $\mathbb{E}_Q\{\cdot\}$. Information measures induced by the generic joint distribution Q_{XY}, will be subscripted by Q, for example, $I_Q(X;Y)$ will denote the corresponding mutual information, etc. The divergence (or, Kullback-Liebler distance) between two probability measures Q and P will be denoted by $D(Q||P)$. The conditional information divergence between the conditional distributions

$Q_{Y|X}$ and $P_{Y|X}$, averaged over P_X, will be denoted by $D(Q_{Y|X}||P_{Y|X}|P_X)$. Logarithms are defined with respect to (w.r.t.) the natural basis, that is, $\log(\cdot) = \ln(\cdot)$, and finally, for a real number x, we denote $[x]_+ \triangleq \max\{0, x\}$.

II. Problem Formulation and Main Results

A. The IFC Model

Consider a two-user interference channel of two senders, two receivers, and a discrete memoryless channel (DMC), defined by a set of single-letter transition probabilities, $W_{Y_1Y_2|X_1X_2}(y_1y_2|x_1x_2)$, with finite input alphabets $\mathcal{X}_1, \mathcal{X}_2$ and finite output alphabets $\mathcal{Y}_1, \mathcal{Y}_2$. Here, each sender, $k \in \{1,2\}$, wishes to communicate an independent message $M_k \in \left\{1, 2, \ldots, 2^{nR_k}\right\}$ at rate R_k, and each receiver, $l \in \{1,2\}$, wishes to decode its respective message. Specifically, a $(2^{nR_1}, 2^{nR_2}, n)$ code $\mathcal{C}_n$ consists of:

- Two message sets $\mathcal{M}_1 \triangleq \left\{0, \ldots, 2^{nR_1} - 1\right\}$ and $\mathcal{M}_2 \triangleq \left\{0, \ldots, 2^{nR_2} - 1\right\}$ for the first and second users, respectively.
- Two encoders, where for each $k \in \{1,2\}$, the k-th encoder assigns a codeword $x_{k,i}^n \triangleq (x_{k,i,1}, x_{k,i,2}, \ldots, x_{k,i,n})$ to each message $i \in \mathcal{M}_k$.
- Two decoders, where each decoder $l \in \{1,2\}$ assigns an estimate $\hat{M}_l$ to M_l.

We assume that the message pair (M_1, M_2) is uniformly distributed over $\mathcal{M}_1 \times \mathcal{M}_2$. It is clear that the *optimal decoder* of the first user, for this problem, is given by

$$\hat{M}_1 = \arg\max_{i\in\mathcal{M}_1} P\left(y_1^n | x_{1,i}^n\right) \tag{1}$$

$$= \arg\max_{i\in\mathcal{M}_1} e^{-nR_2} \sum_{j=1}^{M_2-1} P\left(y_1^n | x_{1,i}^n, x_{2,j}^n\right) \tag{2}$$

where $P\left(y_1^n | x_{1,i}^n, x_{2,j}^n\right)$ is the marginal channel defined as

$$P\left(y_1^n | x_{1,i}^n, x_{2,j}^n\right) \triangleq \prod_{k=1}^{n} W_{Y_1|X_1X_2}(y_{1k}|x_{1,i,k}x_{2,j,k}), \tag{3}$$

and

$$\begin{aligned} &W_{Y_1|X_1X_2}(y_{1,k}|x_{1,i,k}x_{2,j,k}) \\ &\quad\triangleq \sum_{y_{2,k}\in\mathcal{Y}_2} W_{Y_1Y_2|X_1X_2}(y_{1,k}y_{2,k}|x_{1,i,k}x_{2,j,k}). \end{aligned} \tag{4}$$

The optimal decoder of the second user is defined similarly. Since there is no cooperation between the two receivers, the error probabilities for the code $\mathcal{C}_n$, are defined as:

$$\begin{aligned} &P_{e,i}(\mathcal{C}_n) \triangleq 2^{-n(R_1+R_2)} \cdot \\ &\sum_{m_1,m_2} \mathbb{P}\left\{\hat{M}_i(Y_i^n) \neq m_i | M_1 = m_1, M_2 = m_2\right\}, \quad i = 1,2. \end{aligned} \tag{5}$$

B. The Ordinary Random Coding Ensemble

In this subsection, we consider the ordinary random coding ensemble: For each $k \in \{1,2\}$, we select independently M_k codewords $x_{k,i}^n$, for $i \in \mathcal{M}_k$, under the uniform distribution across the type class $T(P_{X_k})$, for a given distribution P_{X_k} on $\mathcal{X}_k$. Our goal is to assess the exponential rate of $\bar{P}_{e,1}^{(n)} \triangleq \mathbb{E}\{P_{e,1}(\mathcal{C}_n)\}$, where the average is over the code ensemble, that is,

$$E_1^*(R_1, R_2) \triangleq \liminf_{n\to\infty} -\frac{1}{n} \log \bar{P}_{e,1}^{(n)}, \tag{6}$$

and similarly for the second user. Before stating the main result, we define some quantities. Given a joint distribution $Q_{X_1X_2Y_1}$ over $\mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y}_1$, consider the definitions in (7), shown at the top of the next page. Our main result is the following. Due to space limitation, the proofs of all the following results are omitted and can be found in [22].

Theorem 1 Let R_1 and R_2 be given, and let $E^*(R_1, R_2)$ be defined as in (6). Consider the ensemble of fixed composition codes of types P_{X_1} and P_{X_2}, for the first and second users, respectively. For a discrete memoryless two-user IFC, we have:

$$E_1^*(R_1, R_2) \geq \tilde{E}_1(R_1, R_2), \tag{8}$$

for any $R_1, R_2 \geq 0$.

Several remarks on Theorem 1 are in order.

- Due to symmetry, the error exponent for the second user, that is, $E_2^*(R_1, R_2)$ is simply obtained from Theorem 1 by swapping the roles of X_1, Y_1, and R_1, with X_2, Y_2, and R_2, respectively.
- An immediate byproduct of Theorem 1 is finding the set of rates $(R_1, R_2) \in \mathbb{R}_+^2$ for which $\tilde{E}_1(R_1, R_2) > 0$, namely, for which the probability of error vanishes exponentially as $n \to \infty$. It is not difficult to show that this set is given by:

$$\mathcal{R}_{\text{ordinary},1} = \hat{\mathcal{R}}_1 \cup \left\{(R_1, R_2) : \begin{array}{c} R_1 < I(X_1; Y_1|X_2) \\ R_1 + R_2 < I(X_1, X_2; Y_1) \end{array}\right\}, \tag{9}$$

evaluated with $P_{X_1X_2Y_1} = P_{X_1} \times P_{X_2} \times W_{Y_1|X_1X_2}$, where $\hat{\mathcal{R}}_1 \triangleq \{R_1 : R_1 < I(X_1; Y_1)\}$. Fig. 1 demonstrates a qualitative description of this region. The interpretation is as follows: The corner point $(I(X_1; Y_1|X_2), I(X_2; Y_1))$ is achieved by first decoding the interference (the second user), canceling it, and then decoding the first user. The sum-rate constraint can be achieved by joint decoding the two users (similarly to MAC), and thus, obviously, also by our optimal decoder. Finally, the region $R_1 < I(X_1; Y_1)$ and $R_2 \geq I(X_2; Y_1|X_1)$ means that we decode the first user while treating the interference as noise. Evidently, from the perspective of the first decoder, which is interested only in the message that is emitted from the first sender, the second sender can use any rate, and thus there is no bound on R_2 whenever $R_1 < I(X_1; Y_1)$. Note that this region was also obtained in [9], but from a lower bound on the error exponent. Accordingly, this means that according to [9], the achievable rate could be larger. Our results, however, show that one cannot do better when standard random coding is applied. Notice that $\mathcal{R}_{\text{ordinary},1}$ is well-known to be contained in the HK region [10, 23].
- *Existence of a single code:* our result holds true on the average, where the averaging is done over the random choice of codebooks. It can be shown (see, for example, [24, p. 2924]) that there exists deterministic sequence of fixed composition codebooks of increasing block length n for which the same asymptotic error performance can be achieved for *both* users simultaneously.

$$f\left(Q_{X_1X_2Y_1}\right) \triangleq \mathbb{E}_Q\left[\log W_{Y_1|X_1X_2}(Y_1|X_1X_2)\right], \tag{7a}$$

$$t_0(Q_{X_1Y_1}) \triangleq R_2 + \max_{\hat{Q}:\ \hat{Q}_{X_2}=P_{X_2},\ \hat{Q}_{X_1Y_1}=Q_{X_1Y_1},\ I_{\hat{Q}}(X_2;X_1,Y_1)\leq R_2}\left[f(\hat{Q}) - I_{\hat{Q}}(X_2;X_1,Y_1)\right], \tag{7b}$$

$$\mathcal{L}(\tilde{Q}_{X_1X_2Y_1}, Q_{X_1X_2Y_1}) \triangleq \left\{\hat{Q}:\ \max\left[t_0(Q_{X_1X_2Y_1}), f(Q_{X_1X_2Y_1})\right] \leq \max\left[f(\tilde{Q}_{X_1X_2Y_1}), f(\hat{Q}) + \left[R_2 - I_{\hat{Q}}(X_2;X_1,Y_1)\right]_+\right]\right\}, \tag{7c}$$

$$E_1(\tilde{Q}_{X_1X_2Y_1}, Q_{X_1X_2Y_1}) \triangleq \min_{\hat{Q}:\ \hat{Q}_{X_2}=P_{X_2},\ \hat{Q}_{X_1Y_1}=\tilde{Q}_{X_1Y_1},\ \hat{Q}\in\mathcal{L}(\tilde{Q}_{X_1X_2Y_1},Q_{X_1X_2Y_1})}\left[I_{\hat{Q}}(X_2;X_1,Y_1) - R_2\right]_+, \tag{7d}$$

$$\hat{E}_1(Q_{X_1X_2Y_1}, R_2) \triangleq \min_{\tilde{Q}:\ \tilde{Q}_{X_1}=P_{X_1},\ \tilde{Q}_{X_2Y_1}=Q_{X_2Y_1}}\left[I_{\tilde{Q}}(X_1;X_2,Y_1) + E_1(\tilde{Q}_{X_1X_2Y_1}, Q_{X_1X_2Y_1})\right], \tag{7e}$$

$$\tilde{E}_1(R_1,R_2) \triangleq \min_{Q_{Y_1|X_1X_2}:\ Q_{X_1}=P_{X_1}, Q_{X_2}=P_{X_2}}\left[D(Q_{Y_1|X_1X_2}||W_{Y_1|X_1X_2}|P_{X_1}\times P_{X_2}) + \left[\hat{E}_1(Q_{X_1X_2Y_1},R_2) - R_1\right]_+\right]. \tag{7f}$$

$$\bar{P}_{e,1}^{(n)} = \Pr\left[\bigcup_{i=1}^{M_1-1}\left\{\sum_{j=0}^{M_2-1} P\left(Y_1^n|X_{1,i}^n, X_{2,j}^n\right) \geq \sum_{j=0}^{M_2-1} P\left(Y_1^n|X_{1,0}^n, X_{2,j}^n\right)\right\}\right], \tag{10}$$

$$= \mathbb{E}\left\{\Pr\left[\bigcup_{i=1}^{M_1-1}\left\{\sum_{j=0}^{M_2-1} P\left(Y_1^n|X_{1,i}^n, X_{2,j}^n\right) \geq \sum_{j=0}^{M_2-1} P\left(Y_1^n|X_{1,0}^n, X_{2,j}^n\right)\right\}\middle|\mathcal{F}_0\right]\right\}, \tag{11}$$

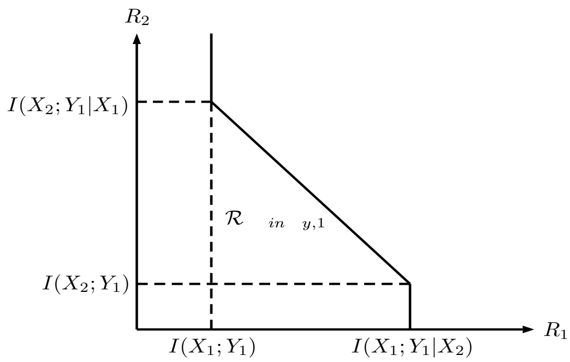

Fig. 1. Rate region $\mathcal{R}_{\text{ach},1}$ for which $E_1^*(R_1,R_2) > 0$.

• *On the proof:* it is instructive to discuss (in some more detail than earlier) one of the main difficulties in proving Theorem 1, which is customary to multiuser systems, such as the IFC. Without loss of generality, we assume throughout, that the transmitted codewords are $x_{1,0}^n$ and $x_{2,0}^n$. Accordingly, the average probability of error associated with the decoder (2) is given by (11), shown at the top of the next page, where $\mathcal{F}_0 \triangleq \left(X_{1,0}^n, X_{2,0}^n, Y_1^n\right)$. By the union bound and Shulman's inequality [21, Lemma A.2], we know that for a sequence of pairwise independent events, $\{\mathcal{A}_i\}_{i=1}^N$, the following holds:

$$\frac{1}{2}\min\left\{1, \sum_{i=1}^N \Pr\left\{\mathcal{A}_i\right\}\right\} \leq \Pr\left\{\bigcup_{i=1}^N \mathcal{A}_i\right\} \leq \min\left\{1, \sum_{i=1}^N \Pr\left\{\mathcal{A}_i\right\}\right\}, \tag{12}$$

which is a useful result when assessing the exponential behavior of such probabilities. Equation (12) is one of the building blocks of tight exponential analysis of previously considered point-to-point systems (see, e.g., [16-19], and many references therein). However, it is evident that in our case the various events are not pairwise independent, and therefore this result cannot be applied directly. Indeed, since we are interested in the optimal decoder, each event of the union in (11), depends on the whole codebook of the second user. One may speculate that this problem can be tackled by conditioning on the codebook of the second user, and then (12). However, the cost of this conditioning is a very complicated (if not intractable) large deviations analysis of some quantities. To alleviate this problem, we derived new upper bounds on the probability of union of events, which takes into account the dependencies among the events. This was done using the techniques of [11].

Another difficulty that arises in the error exponent analysis of the IFC model, is that in contrast to previous works, applying the distance enumerator[2] method [14], is not a simple task. Again, our optimal decoder compares two quantities (i.e., likelihoods) which are both depend on the whole codebook of the second user. The consequence of this situation, is that in order to analyze the probability of error, it is required to analyze the joint distribution of type class enumerators, and not just rely on their marginal distributions, as usually done, e.g., [16-19].

• *Comparison with [9]*: Similarly to [9], we present results

[2] For a given $y^n \in \mathcal{Y}^n$, and a given joint probability distribution Q_{XY} on $\mathcal{X}\times\mathcal{Y}$, the *distance enumerator* (or, *type class enumerator*), $N(Q_{XY})$, is the number of codewords $\{x_i^n\}$ in $\mathcal{C}_n$ whose conditional empirical joint distribution with y^n is Q_{XY}, namely, $N(Q_{XY}) = \left|x^n \in \mathcal{C}_n :\ \hat{Q}_{x^ny^n} = Q_{XY}\right|$, where $\hat{Q}_{x^ny^n}$ is the empirical joint distribution of x^n and y^n, and $|\mathcal{A}|$ designates the cardinality of a finite set $\mathcal{A}$.

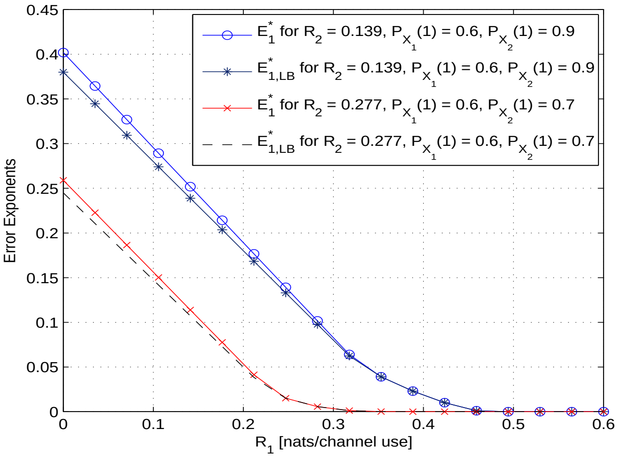

Fig. 2. Comparison between our error exponent $E_1^*(R_1, R_2)$ and the lower bound $E_{\rm LB}(R_1, R_2)$ of [9], as a function of R_1 for two different values of R_2 and fixed choices of P_{X_1} and P_{X_2}.

for the binary Z-channel model defined as follows: $Y_1 = X_1 \cdot X_2 \oplus Z$ and $Y_2 = X_2$, where $X_1, X_2, Y_1, Y_2 \in \{0, 1\}$, $Z \sim$ Bern(p), "$\cdot$" is multiplication, and "$\oplus$" is modulo-2 addition. In the numerical calculations, we fix $p = 0.01$. Fig. 2 presents the lower bound on the error exponent under optimal decoding, derived in this paper, compared to the lower bound $E_{\rm LB}(R_1, R_2)$ of [9], as a function of R_1, for different values of P_{X_1}, P_{X_2}, and R_2. It can be seen that our exponents can be strictly better than those of [9].

• *Generalization to other ensemble*: As was mentioned before, in [9] only random codebooks comprised of i.i.d. codewords were considered. These ensembles are much simpler than the superposition codebooks of [8]. Unfortunately, it very tedious to analyze superposition codebooks using the methods of [9], and even if we do so, the tightness is questionable. However, the new tools that we have derived enable us to analyze more involved random coding ensembles. Due to space limitations, we do not present the error exponents achieved by the following schemes. All the details can be found in [22, Subsection III.C]. For example, we can derive the error exponents for the HK scheme, which gives the best known inner bound. The idea in this scheme is to split the message M_1 into "private" and "common" messages, M_{11} and M_{12} at rates R_{11} and R_{12}, respectively, such that $R_1 = R_{11} + R_{12}$. Similarly M_2 is split into M_{21} and M_{22} at rates R_{21} and R_{22}, respectively, such that $R_2 = R_{21} + R_{22}$. Then, receiver $k = 1, 2$, recovers its intended message M_k, and the common message from the other sender (although it is not required to) each decoder. Also, using the same techniques, we can analyze the error exponents resulting from the *hierarchical code ensemble* [20], in which the case has a tree structure with two levels, where the first serves for "cloud centers", and the second for the "satellites".

Acknowledgment

This research was partially supported by The Israeli Science Foundation (ISF), grant no. 412/12.

References

[1] A. El Gamal and Y. H. Kim, *Network Information Theory*. Cambridge University Press, 2011.

[2] E. C. Shannon, "Two-way communication channels," in *Proc. 4th Berkeley Symp. on Mathematical Statistics and Probability*, vol. 1. Berkeley, CA: Univ. California Press, 1961, pp. 611–64.

[3] R. Ahlswede, "The capacity region of a channel with two senders and two receivers," *Annals Probabil.*, vol. 2, no. 5, pp. 805–814, 1974.

[4] A. B. Carleial, "Interference channels," *IEEE Trans. on Inf. Theory*, vol. IT-24, pp. 60–70, Jan. 1978.

[5] H. Sato, "Two-user communication channels," *IEEE Trans. on Inf. Theory*, vol. IT-23, pp. 295–304, May. 1977.

[6] A. B. Carleial, "A case where intereference does not reduce capacity," *IEEE Trans. on Inf. Theory*, vol. IT-21, pp. 569–570, Sep. 1975.

[7] R. Benzel, "The capacity region of a class of discrete additive degraded interference channels," *IEEE Trans. on Inf. Theory*, vol. IT-25, pp. 228–231, Mar. 1979.

[8] T. S. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. on Inf. Theory*, vol. IT, no. 27, pp. 49–60, Jan. 1981.

[9] R. Etkin, N. Merhav, and E. Ordentlich, "Error exponents of optimum decoding for the interference channel," *IEEE Trans. on Inf. Theory*, vol. 56, no. 1, pp. 40–56, Jan. 2010.

[10] C. Chang, R. Etkin, and E. Ordentlich, "Interference channel capacity region for randomized fixed-composition codes," *HP Labs Technical Report*, 2009. [Online]. Available: http://www.hpl.hp.com/techreports/2008/HPL-2008-194R1.html

[11] J. Scarlett, A. Martinez, and A. G. Fábregas, "Multiuser coding techniques for mismatched decoding," *submitted to IEEE Trans. on Inf. Theory*, Nov. 2013. [Online]. Available: arxiv.org/pdf/1311.6635

[12] R. G. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.

[13] I. Csiszár and J. Korner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Academic Press, 2011.

[14] N. Merhav, "Relations between random coding exponents and the statistical physics of random codes," *IEEE Trans. on Inf. Theory*, vol. 55, no. 1, pp. 83–92, Jan. 2009.

[15] J. Scarlett, A. Martinez, and A. G. Fábregas, "Mismatched multi-letter successive decoding for the multiple-access channel," in *Proc. ISIT 2014*, Jul. 2014, pp. 2539–2543.

[16] N. Merhav, "Exact random coding exponents of optimal bin index decoding," *IEEE Trans. on Inf. Theory*, vol. 60, no. 10, pp. 6024–6031, Oct. 2014.

[17] ——, "List decoding-random coding exponents and expurgated exponents," *IEEE Trans. on Inf. Theory*, vol. 60, no. 11, pp. 6749–6759, Nov. 2014.

[18] ——, "Exact correct-decoding exponent for the wiretap channel decoder," *IEEE Trans. on Inf. Theory*, vol. 60, no. 12, pp. 7606–7615, Dec. 2014.

[19] W. Huleihel, N. Weinberger, and N. Merhav, "Erasure/list random coding error exponents are not universally achievable," *submitted to IEEE Trans. on Inf. Theory*, Oct. 2014. [Online]. Available: http://arxiv.org/abs/1410.7005

[20] N. Merhav, "The generalized random energy model and its application to the statistical physics of ensembles of hierarchical codes," *IEEE Trans. on Inf. Theory*, vol. 55, no. 3, pp. 1250–1268, May 2009.

[21] N. Shulman, "*Communication over an unknown channel via common broadcasting*," Ph.D. dissertation, Tel-Aviv University, 2003, http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.407.7542.

[22] W. Huleihel and N. Merhav, "Exact random coding error exponents for the two-user interference channel," Mar. 2015. [Online]. Available: http://webee.technion.ac.il/people/merhav/papers/p179arxiv.pdf

[23] B. Bandemer, A. El Gamal, and Y. K. Kim, "Optimal achievable rates for interference networks with random codes," *submitted to IEEE Trans. on Inf. Theory*, 2012. [Online]. Available: http://arxiv.org/abs/1210.4596

[24] L. Weng, S. S. Pradhan, and A. Anastasopoulos, "Error exponent regions for Gaussian broadcast and multiple-access channels," *IEEE Trans. on Inf. Theory*, vol. 54, no. 7, pp. 2919–2942, July 2008.