An Achievable Error Exponent for the Mismatched Multiple-Access Channel

Jonathan Scarlett
University of Cambridge
jms265@cam.ac.uk

Albert Guillén i Fàbregas
ICREA & Universitat Pompeu Fabra
University of Cambridge
guillen@ieee.org

***Abstract*—This paper considers channel coding for the discrete memoryless multiple-access channel with a given (possibly suboptimal) decoding rule. Using constant-composition random coding, an achievable error exponent is obtained which is tight with respect to the ensemble average, and positive for all rate pairs in the interior of Lapidoth's achievable rate region.**

I. Introduction

The problem of channel coding with a mismatched decoding rule arises in numerous settings [1]–[5]. For example, in practical systems the decoder may have imperfect knowledge of the channel, or implementation constraints may prohibit the use of an optimal decoder. The problem of finding the mismatched capacity is an open problem in general, and most existing results are on achievable rates via random coding. Of particular note is the *LM rate*, which can be obtained via constant-composition random coding [3], [4] or i.i.d. random coding with a cost constraint [5].

The mismatched *multiple-access channel* (MAC) was first considered by Lapidoth [1], who obtained an achievable rate region and showed the surprising fact that the single-user LM rate can be improved by treating the single-user channel as a MAC. As an example, Lapidoth considered the channel in Figure 1 consisting of two parallel binary symmetric channels (BSCs) with crossover probabilities $\delta_1 < 0.5$ and $\delta_2 < 0.5$. The mismatched decoder assumes that both crossover probabilities are equal to $\delta < 0.5$. By treating the channel as a mismatched single-user channel from (x_1, x_2) to (y_1, y_2) and using random coding with a uniform distribution on the quaternary input alphabet, one can only achieve rates R satisfying

$$R < 2\left(1 - H_2\left(\frac{\delta_1 + \delta_2}{2}\right)\right) \tag{1}$$

where H_2 is the binary entropy function in bits. On the other hand, by treating the channel as a mismatched MAC from x_1 and x_2 to (y_1, y_2) and using random coding with equiprobable input distributions on each binary input alphabet, one can achieve any sum-rate R satisfying

$$R < \big(1 - H_2(\delta_1)\big) + \big(1 - H_2(\delta_2)\big). \tag{2}$$

This is the best rate possible even under maximum-likelihood (ML) decoding.

This work has been supported in part by the European Research Council under ERC grant agreement 259663.

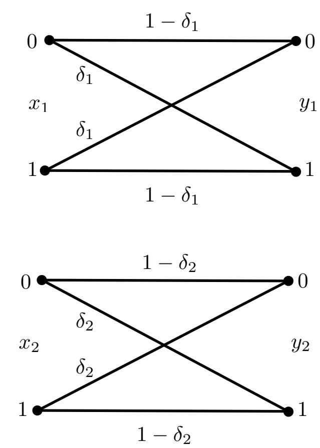

Figure 1. Lapidoth's Parallel BSC Example

A *random coding converse* was also given in [1], showing that the the random-coding error probability tends to one for rate pairs outside the given achievable rate region. In this paper, we strengthen the results of [1] by obtaining random-coding error exponents for the constant-composition ensemble which are tight with respect to the ensemble average, and positive within the interior of the rate region given in [1].

The problem of finding the best achievable error exponents for the MAC is unsolved even in the matched regime; see [6] and references therein. To our knowledge, no complete results on the ensemble tightness of error exponents for the MAC have been given previously, even in the matched regime.

While error exponents for a given decoding rule are presented in [6], they are not tight enough to prove the achievability of Lapidoth's rate region. For example, for the parallel BSC example in Figure 1 with uniform input distributions, the exponents of [6] are only positive when the sum-rate R satisfies (1), whereas the ensemble-tight exponent can be positive for all sum-rates satisfying (2). We will see that the key difference in our analysis is a refined application of the union bound (see Section III-B). Improving the standard use of the union bound was also a key idea in [1], but it was done differently to the present paper.

A. System Setup

We consider a 2-user discrete memoryless MAC (DM-MAC) $W(y|x_1,x_2)$ with input alphabets $\mathcal{X}_1$ and $\mathcal{X}_2$ and output alphabet $\mathcal{Y}$. The decoding metric is denoted by $q(x_1,x_2,y)$. We write $W(\boldsymbol{y}|\boldsymbol{x}_1,\boldsymbol{x}_2)$ and $q(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y})$ as a shorthand for $\prod_{i=1}^n W(y_i|x_{1,i},x_{2,i})$ and $\prod_{i=1}^n q(x_{1,i},x_{2,i},y_i)$ respectively, where y_i is the i-th entry of $\boldsymbol{y}$ and similarly for $x_{1,i}$ and $x_{2,i}$.

The encoders and decoder operate as follows. Encoder ν ($\nu \in \{1,2\}$) selects a message m_ν equiprobably from the set $\{1,...,M_\nu\}$, and transmits the corresponding codeword $\boldsymbol{x}_\nu^{(m_\nu)}$ from the codebook $\mathcal{C}_\nu = \{\boldsymbol{x}_\nu^{(1)},...,\boldsymbol{x}_\nu^{(M_\nu)}\}$. Upon receiving the signal $\boldsymbol{y}$ at the output of the channel, the decoder forms an estimate $(\widehat{m}_1,\widehat{m}_2)$ of the messages, given by

$$(\widehat{m}_1,\widehat{m}_2) = \operatorname*{arg\,max}_{i\in\{1,...,M_1\},j\in\{1,...,M_2\}} q(\boldsymbol{x}_1^{(i)},\boldsymbol{x}_2^{(j)},\boldsymbol{y}). \quad (3)$$

We assume that ties are broken at random. An error is said to have occurred if the estimate $(\widehat{m}_1,\widehat{m}_2)$ differs from (m_1,m_2). We distinguish between the following three types of error:

(Type 1) $\widehat{m}_1 \neq m_1$ and $\widehat{m}_2 = m_2$
(Type 2) $\widehat{m}_1 = m_1$ and $\widehat{m}_2 \neq m_2$
(Type 12) $\widehat{m}_1 \neq m_1$ and $\widehat{m}_2 \neq m_2$

The probabilities of these events are denoted by $p_{e,1}$, $p_{e,2}$ and $p_{e,12}$ respectively, and the overall error probability is denoted by p_e. The ensemble-average error probabilities for a given random-coding ensemble are denoted by $\overline{p}_{e,1}$, $\overline{p}_{e,2}$, $\overline{p}_{e,12}$ and $\overline{p}_e$ respectively. Clearly we have

$$\max\{p_{e,1},p_{e,2},p_{e,12}\} \le p_e \le p_{e,1}+p_{e,2}+p_{e,12} \quad (4)$$

and similarly for $\overline{p}_e$.

A rate pair (R_1,R_2) is said to be achievable if there exist sequences of codebooks with $M_1 = \exp(nR_1)$ and $M_2 = \exp(nR_2)$ codewords of length n for users 1 and 2 respectively such that $p_e \to 0$. We say that $E(R_1,R_2)$ is an *achievable error exponent* if there exist sequences of codebooks with $M_1 = \exp(nR_1)$ and $M_2 = \exp(nR_2)$ codewords of length n such that

$$\lim_{n\to\infty} -\frac{1}{n}\log p_e \ge E(R_1,R_2). \quad (5)$$

For a given random-coding ensemble, we say that the random-coding error exponent $E_r(R_1,R_2)$ exhibits *ensemble tightness* if

$$\lim_{n\to\infty} -\frac{1}{n}\log \overline{p}_e = E_r(R_1,R_2). \quad (6)$$

B. Notation

The set of all probability distributions on an alphabet $\mathcal{A}$ is denoted by $\mathcal{P}(\mathcal{A})$. The set of all sequences with a given type P_X is denoted by $T(P_X)$, and similarly for joint types. We refer the reader to [7], [8] for an introduction to the method of types. The main properties of types used in this paper are outlined in the Appendix.

The probability of an event is denoted by $\mathbb{P}[\cdot]$. The symbol $\sim$ means "distributed as". The marginals of a joint distribution $P_{XY}(x,y)$ are denoted by $P_X(x)$ and $P_Y(y)$. Similarly, $P_{Y|X}(y|x)$ denotes the conditional distribution induced by $P_{XY}(x,y)$. We write $P_X = \widetilde{P}_X$ to denote element-wise equality between two probability distributions on the same alphabet. For a joint distribution $P_{XY}(x,y)$, expectations are denoted by $\mathbb{E}_P[\cdot]$. When the probability distribution is understood from the context, we simply write $\mathbb{E}[\cdot]$.

Given a distribution $Q(x)$ and a conditional distribution $W(y|x)$, we write $Q \times W$ to denote the joint distribution $Q(x)W(y|x)$, and similarly when there are more than two distributions. For example, given $Q_1(x_1)$, $Q_2(x_2)$ and $W(y|x_1,x_2)$ we have

$$Q_1 \times Q_2 \times W \sim Q_1(x_1)Q_2(x_2)W(y|x_1,x_2). \quad (7)$$

Mutual information with respect to a joint distribution $P_{XY}(x,y)$ is written with a subscript, e.g.

$$I_P(X;Y) \triangleq \sum_{x,y} P_{XY}(x,y)\log\frac{P_{XY}(x,y)}{P_X(x)P_Y(y)}. \quad (8)$$

For two sequences $f(n)$ and $g(n)$, we write $f(n) \doteq g(n)$ if $\lim_{n\to\infty}\frac{1}{n}\log\frac{f(n)}{g(n)} = 0$, and similarly for $\dot{\le}$ and $\dot{\ge}$. All logarithms have base e, and all rates are in units of nats except in the examples, where bits are used. We define $[c]^+ = \max\{0,c\}$, and denote the indicator function by $\mathbb{1}\{\cdot\}$.

II. Error Exponent for Constant-Composition Random Coding

In this section, we present the ensemble-tight random-coding error exponent for the constant-composition ensemble, in which each codeword of a given user has the same empirical distribution. Using this exponent, we prove the achievability of Lapidoth's achievable rate region [1]. The derivation of the error exponent and the discussion of the analysis are postponed until Section III.

We let $\boldsymbol{X}_\nu^{(i)}$ be the random variable corresponding to the i-th codeword of user ν, and let $\boldsymbol{Y}$ denote the random sequence at the output of the channel. The codewords are distributed according to

$$\Big(\{\boldsymbol{X}_1^{(i)}\}_{i=1}^{M_1},\{\boldsymbol{X}_2^{(j)}\}_{j=1}^{M_2}\Big) \sim \prod_{i=1}^{M_1} Q_{\boldsymbol{X}_1}(\boldsymbol{x}_1^{(i)})\prod_{j=1}^{M_2} Q_{\boldsymbol{X}_2}(\boldsymbol{x}_2^{(j)}) \quad (9)$$

where $Q_{\boldsymbol{X}_\nu}$ is the codeword distribution for user ν. We assume without loss of generality that message $(1,1)$ is transmitted, and write $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ in place of $\boldsymbol{X}_1^{(1)}$ and $\boldsymbol{X}_2^{(1)}$. We write $\overline{\boldsymbol{X}}_1$ and $\overline{\boldsymbol{X}}_2$ to denote arbitrary codewords which are generated independently of $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$.

We fix $Q_1(x_1) \in \mathcal{P}(\mathcal{X}_1)$ and take $Q_{\boldsymbol{X}_1}$ to be the uniform distribution over the type class $T(Q_{1,n})$, where $Q_{1,n} \in \mathcal{P}_n(\mathcal{X}_1)$ is the most probable type under Q_1; similarly for $Q_2(x_2)$ and $Q_{2,n}$. That is,

$$Q_{\boldsymbol{X}_1}(\boldsymbol{x}_1) = \frac{1}{|T(Q_{1,n})|}\mathbb{1}\{\boldsymbol{x}_1 \in T(Q_{1,n})\} \quad (10)$$

$$Q_{\boldsymbol{X}_2}(\boldsymbol{x}_2) = \frac{1}{|T(Q_{2,n})|}\mathbb{1}\big\{\boldsymbol{x}_2 \in T(Q_{2,n})\big\}. \tag{11}$$

To ease notation, we write $f(\boldsymbol{Q})$ to denote a function f which depends on Q_1 and Q_2, and similarly for $\boldsymbol{Q}_n$.

Remark: All of the results in this paper can easily be extended to the ensemble in which the codewords are generated conditionally on a *time-sharing sequence* $\boldsymbol{u}$, such that the joint type of $(\boldsymbol{u}, \boldsymbol{x}_1)$ is fixed for every user-1 codeword $\boldsymbol{x}_1$, and similarly for user 2 (e.g. see [6]). However, in the mismatched setting there are some subtle differences between the performance of this ensemble and that of explicit time-sharing, and their study is beyond the scope of this paper.

The error exponents and achievable rates will be expressed in terms of the sets

$$\mathcal{S}(\boldsymbol{Q}) \triangleq \Big\{P_{X_1X_2Y} \in \mathcal{P}(\mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y}) \,:\, P_{X_1} = Q_1,\, P_{X_2} = Q_2\Big\} \tag{12}$$

$$\mathcal{T}_1(P_{X_1X_2Y}) \triangleq \Big\{\widetilde{P}_{X_1X_2Y} \in \mathcal{P}(\mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y}) \,:\, \widetilde{P}_{X_1} = P_{X_1}, \widetilde{P}_{X_2Y} = P_{X_2Y}, \mathbb{E}_{\widetilde{P}}[\log q(X_1,X_2,Y)] \ge \mathbb{E}_P[\log q(X_1,X_2,Y)]\Big\} \tag{13}$$

$$\mathcal{T}_2(P_{X_1X_2Y}) \triangleq \Big\{\widetilde{P}_{X_1X_2Y} \in \mathcal{P}(\mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y}) \,:\, \widetilde{P}_{X_2} = P_{X_2}, \widetilde{P}_{X_1Y} = P_{X_1Y}, \mathbb{E}_{\widetilde{P}}[\log q(X_1,X_2,Y)] \ge \mathbb{E}_P[\log q(X_1,X_2,Y)]\Big\} \tag{14}$$

$$\mathcal{T}_{12}(P_{X_1X_2Y}) \triangleq \Big\{\widetilde{P}_{X_1X_2Y} \in \mathcal{P}(\mathcal{X}_1 \times \mathcal{X}_2, \mathcal{Y}) \,:\, \widetilde{P}_{X_1} = P_{X_1}, \widetilde{P}_{X_2} = P_{X_2}, \widetilde{P}_Y = P_Y, \mathbb{E}_{\widetilde{P}}[\log q(X_1,X_2,Y)] \ge \mathbb{E}_P[\log q(X_1,X_2,Y)]\Big\}. \tag{15}$$

The following theorem gives the random-coding error exponent for each error type.

Theorem 1. *The random-coding error probabilities for the constant-composition ensemble in* (9)–(11) *satisfy*

$$\overline{p}_{e,1} \doteq \exp\big(-nE_{r,1}(\boldsymbol{Q}, R_1)\big) \tag{16}$$

$$\overline{p}_{e,2} \doteq \exp\big(-nE_{r,2}(\boldsymbol{Q}, R_2)\big) \tag{17}$$

$$\overline{p}_{e,12} \doteq \exp\big(-nE_{r,12}(\boldsymbol{Q}, R_1, R_2)\big) \tag{18}$$

where

$$E_{r,1}(\boldsymbol{Q}, R_1) \triangleq \min_{P_{X_1X_2Y} \in \mathcal{S}(\boldsymbol{Q})} \min_{\widetilde{P}_{X_1X_2Y} \in \mathcal{T}_1(P_{X_1X_2Y})} D(P_{X_1X_2Y} \| Q_1 \times Q_2 \times W) + \big[I_{\widetilde{P}}(X_1; X_2, Y) - R_1\big]^+ \tag{19}$$

$$E_{r,2}(\boldsymbol{Q}, R_2) \triangleq \min_{P_{X_1X_2Y} \in \mathcal{S}(\boldsymbol{Q})} \min_{\widetilde{P}_{X_1X_2Y} \in \mathcal{T}_2(P_{X_1X_2Y})} D(P_{X_1X_2Y} \| Q_1 \times Q_2 \times W) + \big[I_{\widetilde{P}}(X_2; X_1, Y) - R_2\big]^+ \tag{20}$$

$$E_{r,12}(\boldsymbol{Q}, R_1, R_2) \triangleq \min_{P_{X_1X_2Y} \in \mathcal{S}(\boldsymbol{Q})} \min_{\widetilde{P}_{X_1X_2Y} \in \mathcal{T}_{12}(P_{X_1X_2Y})} D(P_{X_1X_2Y} \| Q_1 \times Q_2 \times W) + \max_{\nu \in \{1,2\}} \Psi_\nu(\widetilde{P}_{X_1X_2Y}, R_1, R_2) \tag{21}$$

and

$$\Psi_1(\widetilde{P}_{X_1X_2Y}, R_1, R_2) \triangleq \Big[I_{\widetilde{P}}(X_2; Y) + \big[I_{\widetilde{P}}(X_1; X_2, Y) - R_1\big]^+ - R_2\Big]^+ \tag{22}$$

$$\Psi_2(\widetilde{P}_{X_1X_2Y}, R_1, R_2) \triangleq \Big[I_{\widetilde{P}}(X_1; Y) + \big[I_{\widetilde{P}}(X_2; X_1, Y) - R_2\big]^+ - R_1\Big]^+. \tag{23}$$

Proof: See Section III. ■

Due to the lack of converse results in mismatched decoding, it is important to determine whether the weakness in the achievability results is due to the ensemble itself, or the bounding techniques used in the analysis. Theorem 1 states that the overall error exponent

$$E_r(\boldsymbol{Q}, R_1, R_2) \triangleq \min\Big\{E_{r,1}(\boldsymbol{Q}, R_1), E_{r,2}(\boldsymbol{Q}, R_2), E_{r,12}(\boldsymbol{Q}, R_1, R_2)\Big\} \tag{24}$$

is not only achievable, but it is also tight with respect to the ensemble average.

The following achievable rate region follows from Theorem 1 in a straightforward fashion, and coincides with the ensemble-tight achievable rate region of [1].

Theorem 2. *The overall error exponent $E_r(\boldsymbol{Q}, R_1, R_2)$ is positive for all rate pairs (R_1, R_2) in the interior of $\mathcal{R}^{\mathrm{LM}}(\boldsymbol{Q})$, where $\mathcal{R}^{\mathrm{LM}}(\boldsymbol{Q})$ is the set of all rate pairs (R_1, R_2) satisfying*

$$R_1 \le \min_{\widetilde{P}_{X_1X_2Y} \in \mathcal{T}_1(Q_1 \times Q_2 \times W)} I_{\widetilde{P}}(X_1; X_2, Y) \tag{25}$$

$$R_2 \le \min_{\widetilde{P}_{X_1X_2Y} \in \mathcal{T}_2(Q_1 \times Q_2 \times W)} I_{\widetilde{P}}(X_2; X_1, Y) \tag{26}$$

$$R_1 + R_2 \le \min_{\substack{\widetilde{P}_{X_1X_2Y} \in \mathcal{T}_{12}(Q_1 \times Q_2 \times W) \\ I_{\widetilde{P}}(X_1;Y) \le R_1,\, I_{\widetilde{P}}(X_2;Y) \le R_2}} D(\widetilde{P}_{X_1X_2Y} \| Q_1 \times Q_2 \times \widetilde{P}_Y). \tag{27}$$

Proof: The conditions in (25)–(27) are obtained from the error exponents in (19)–(21) respectively. Focusing on (27), we note that the objective in (21) is always positive when $D(P_{X_1X_2Y}\|Q_1\times Q_2\times W)>0$. In the case that the minimizing $P_{X_1X_2Y}$ satisfies $D(P_{X_1X_2Y}\|Q_1\times Q_2\times W)=0$, we obtain that $P_{X_1X_2Y}=Q_1\times Q_2\times W$, and hence $E_{r,12}$ is positive provided that either

$$R_2 \le I_{\widetilde{P}}(X_2;Y)+[I_{\widetilde{P}}(X_1;X_2,Y)-R_1]^+ \quad (28)$$

or

$$R_1 \le I_{\widetilde{P}}(X_1;Y)+[I_{\widetilde{P}}(X_2;X_1,Y)-R_2]^+ \quad (29)$$

under the minimizing $\widetilde{P}_{X_1X_2Y}$ in (21). The condition in (28) corresponds to the the case that Ψ_1 achieves the maximum in (21), and (29) corresponds to the case that Ψ_2 achieves the maximum. Finally, (28) and (29) can be combined to obtain (27) by noting that (28) (respectively, (29)) *always* holds when $I_{\widetilde{P}}(X_2;Y)>R_2$ (respectively, $I_{\widetilde{P}}(X_1;Y)>R_1$), and using

$$I_{\widetilde{P}}(X_2;Y)+I_{\widetilde{P}}(X_1;X_2,Y)=D(\widetilde{P}_{X_1X_2Y}\|Q_1\times Q_2\times \widetilde{P}_Y) \quad (30)$$

$$I_{\widetilde{P}}(X_1;Y)+I_{\widetilde{P}}(X_2;X_1,Y)=D(\widetilde{P}_{X_1X_2Y}\|Q_1\times Q_2\times \widetilde{P}_Y). \quad (31)$$

■

Using the usual time-sharing argument [1], [9], it follows from Theorem 2 that we can achieve any rate pair in the convex hull of

$$\bigcup_{\boldsymbol{Q}}\mathcal{R}^{\mathrm{LM}}(\boldsymbol{Q})$$

where the union is over all input distributions Q_1 and Q_2 on $\mathcal{X}_1$ and $\mathcal{X}_2$ respectively.

In the proof of Theorem 1, it will be shown that a weaker analysis yields the achievable type-12 error exponent

$$E'_{r,12}(\boldsymbol{Q},R_1,R_2)\triangleq \min_{P_{X_1X_2Y}\in\mathcal{S}(\boldsymbol{Q})}\ \min_{\widetilde{P}_{X_1X_2Y}\in\mathcal{T}_{12}(P_{X_1X_2Y})} D(P_{X_1X_2Y}\|Q_1\times Q_2\times W) + \big[D(\widetilde{P}_{X_1X_2Y}\|Q_1\times Q_2\times\widetilde{P}_Y)-(R_1+R_2)\big]^+ \quad (32)$$

which coincides with an achievable exponent given in [6]. Using a similar argument to the proof of Theorem 2, we see that (32) yields a similar achievable rate region to (25)–(25), but with (27) replaced by

$$R_1+R_2 \le \min_{\widetilde{P}_{X_1X_2Y}\in\mathcal{T}_{12}(Q_1\times Q_2\times W)} D(\widetilde{P}_{X_1X_2Y}\|Q_1\times Q_2\times\widetilde{P}_Y). \quad (33)$$

In the following subsection, we compare the ensemble-tight type-12 exponent and the corresponding achievable rate region with that of (32)–(33).

A. Numerical Example

We now return to the parallel BSC example given in Figure 1, where the output is given by $y=(y_1,y_2)$. As mentioned in the introduction, the decoder assumes that both crossover probabilities are equal. It is straightforward to show that the corresponding decoding rule is equivalent to minimizing the sum of t_1 and t_2, where t_ν is the number of bit flips from the input sequence $\boldsymbol{x}_\nu$ to the output sequence $\boldsymbol{y}_\nu$. As noted in [1], this decision rule is in fact equivalent to ML. This channel could easily be analyzed by treating the two subchannels separately, but we treat it as a MAC because it serves as a good example for comparing the ensemble-tight achievability results with (32)–(33).

We let both Q_1 and Q_2 be the uniform distribution on $\{0,1\}$. With this choice of input distributions, it was shown in [1] that the right-hand side of (33) is no greater than

$$2\left(1-H_2\left(\frac{\delta_1+\delta_2}{2}\right)\right). \quad (34)$$

On the other hand, the refined condition in (27) can be used to prove the achievability of *any* (R_1,R_2) within the rectangle defined by the corners $(0,0)$ and (C_1,C_2), where $C_\nu \triangleq 1-H_2(\delta_\nu)$ [1]. This observation is analogous to the comparison between (1) and (2) in the introduction. The main difference is that the weakness in (1) is in the random-coding ensemble itself, whereas the weakness in (34) is in the bounding techniques used in the analysis.

We evaluate the error exponents using the optimization software YALMIP [10]. Figure 2 plots each of the exponents as a function of α, where the rate pairs are given by $(R_1,R_2)=(\alpha C_1,\alpha C_2)$. While the overall error exponent $E_r(\boldsymbol{Q},R_1,R_2)$ in (24) is unchanged at low to moderate values of α when $E'_{r,12}$ is used in place of $E_{r,12}$, this is not true for high values of α. Furthermore, consistent with the preceding discussion, $E'_{r,12}$ is non-zero only for $\alpha<0.865$, whereas $E_{r,12}$ is positive for all $\alpha<1$.

It is interesting to note that the curves $E_{r,12}$ and $E'_{r,12}$ coincide at low values of α. Roughly speaking, the reason for this is that the the arguments to the $[\cdot]^+$ functions in (22)–(23) are positive when the rates are sufficiently small. This is consistent with [6, Corollary 5], which states that (32) is ensemble-tight at low rates.

III. Proof of Theorem 1

While the random-coding error probabilities $\overline{p}_{e,1}$ and $\overline{p}_{e,2}$ can be handled very similarly to the single-user setting [8], $\overline{p}_{e,12}$ requires a more refined analysis. Furthermore, equivalent error exponents to (19)–(20) are given in [6]; we therefore focus exclusively on $\overline{p}_{e,12}$. We first write

$$\overline{p}_{e,12} = c_{12}\mathbb{E}\left[\mathbb{P}\left[\bigcup_{i\neq 1,j\neq 1}\left\{\frac{q(\boldsymbol{X}_1^{(i)},\boldsymbol{X}_2^{(j)},\boldsymbol{Y})}{q(\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{Y})}\ge 1\right\}\middle|\boldsymbol{X}_1,\boldsymbol{X}_2,\boldsymbol{Y}\right]\right] \quad (35)$$

for some $c_{12}\in[\frac{1}{2},1]$. Setting $c_{12}=1$ yields the average probability of error when ties are decoded as errors, and the condition $c_{12}\in[\frac{1}{2},1]$ arises since decoding ties at random reduces the error probability by at most a factor of two.

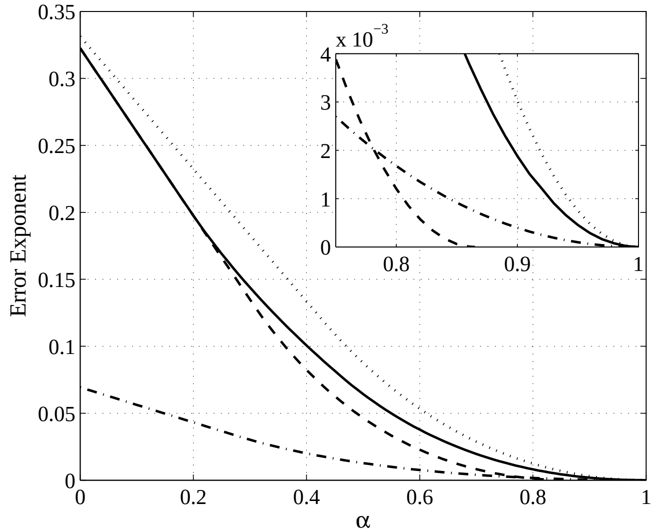

Figure 2. Error exponents $E_{r,1}$ (dotted), $E_{r,2}$ (dash-dot), $E_{r,12}$ (solid) and $E'_{r,12}$ (dashed) for the parallel channel shown in Figure 1 using $\delta_1 = 0.05$, $\delta_2 = 0.25$ and equiprobable input distributions.

We will rewrite (35) in terms of the possible joint types of $(\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y})$ and $(\boldsymbol{X}_1^{(i)}, \boldsymbol{X}_2^{(j)}, \boldsymbol{Y})$. To this end, we define

$$\mathcal{S}_n(\boldsymbol{Q}_n) \triangleq \Big\{ P_{X_1X_2Y} \in \mathcal{P}_n(\mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y}) : P_{X_1} = Q_{1,n},\, P_{X_2} = Q_{2,n} \Big\} \quad (36)$$

$$\mathcal{T}_{12,n}(P_{X_1X_2Y}) \triangleq \Big\{ \widetilde{P}_{X_1X_2Y} \in \mathcal{P}_n(\mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{Y}) : \widetilde{P}_{X_1} = P_{X_1}, \widetilde{P}_{X_2} = P_{X_2}, \widetilde{P}_Y = P_Y, \mathbb{E}_{\widetilde{P}}[\log q(X_1,X_2,Y)] \geq \mathbb{E}_P[\log q(X_1,X_2,Y)] \Big\}. \quad (37)$$

Roughly speaking, $\mathcal{S}_n$ is the set of possible joint types of $(\boldsymbol{X}_1^{(1)}, \boldsymbol{X}_2^{(1)}, \boldsymbol{Y})$, and $\mathcal{T}_{12,n}$ is the set of types of $(\boldsymbol{X}_1^{(i)}, \boldsymbol{X}_2^{(j)}, \boldsymbol{Y})$ which lead to decoding errors when $(\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y}) \in T(P_{X_1X_2Y})$. The constraints on P_{X_ν} and $\widetilde{P}_{X_\nu}$ arise from the fact that we are using constant-composition random coding, and the constraint $\mathbb{E}_{\widetilde{P}}[\log q(X_1,X_2,Y)] \geq \mathbb{E}_P[\log q(X_1,X_2,Y)]$ holds if and only if $\frac{q(\overline{\boldsymbol{x}}_1,\overline{\boldsymbol{x}}_2,\boldsymbol{y})}{q(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y})} \geq 1$ for $(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y}) \in T(P_{X_1X_2Y})$ and $(\overline{\boldsymbol{x}}_1,\overline{\boldsymbol{x}}_2,\boldsymbol{y}) \in T(\widetilde{P}_{X_1X_2Y})$. Fixing $P_{X_1X_2Y} \in \mathcal{S}_n(\boldsymbol{Q}_n)$ and letting $(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y})$ be a triplet of sequences such that $(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y}) \in T(P_{X_1X_2Y})$, it follows that the event

$$\bigcup_{i\neq1,j\neq1} \left\{ \frac{q(\boldsymbol{X}_1^{(i)}, \boldsymbol{X}_2^{(j)}, \boldsymbol{y})}{q(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y})} \geq 1 \right\} \quad (38)$$

can be written as

$$\bigcup_{i\neq1,j\neq1} \bigcup_{\widetilde{P}_{X_1X_2Y} \in \mathcal{T}_{12,n}(P_{X_1X_2Y})} \Big\{ (\boldsymbol{X}_1^{(i)}, \boldsymbol{X}_2^{(j)}, \boldsymbol{y}) \in T(\widetilde{P}_{X_1X_2Y}) \Big\}. \quad (39)$$

Expanding the probability and expectation in (35), substituting (39), and interchanging the order of the unions, we obtain

$$\overline{p}_{e,12} = c_{12} \times \sum_{P_{X_1X_2Y} \in \mathcal{S}_n(\boldsymbol{Q}_n)} \mathbb{P}\big[(\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y}) \in T(P_{X_1X_2Y})\big] \times \mathbb{P}\left[\bigcup_{\widetilde{P}_{X_1X_2Y} \in \mathcal{T}_{12,n}(P_{X_1X_2Y})} \mathcal{E}(\widetilde{P}_{X_1X_2Y}) \,\middle|\, \boldsymbol{Y} \in T(P_Y) \right] \quad (40)$$

where

$$\mathcal{E}(\widetilde{P}_{X_1X_2Y}) \triangleq \bigcup_{i\neq1,j\neq1} \Big\{ (\boldsymbol{X}_1^{(i)}, \boldsymbol{X}_2^{(j)}, \boldsymbol{Y}) \in T(\widetilde{P}_{X_1X_2Y}) \Big\}. \quad (41)$$

We define the conditional probability

$$p_{\mathcal{E}}(\widetilde{P}_{X_1X_2Y}) \triangleq \mathbb{P}\Big[\mathcal{E}(\widetilde{P}_{X_1X_2Y}) \,\big|\, \boldsymbol{Y} \in T(\widetilde{P}_Y) \Big]. \quad (42)$$

We have replaced the condition $\boldsymbol{Y} \in T(P_Y)$ by $\boldsymbol{Y} \in T(\widetilde{P}_Y)$ since we have that $\widetilde{P}_Y = P_Y$ for all $\widetilde{P}_{X_1X_2Y} \in \mathcal{T}_{12,n}(P_{X_1X_2Y})$.

Applying the union bound to (40) and using the fact that the number of joint types is polynomial in n, we obtain

$$\overline{p}_{e,12} \doteq \max_{P_{X_1X_2Y} \in \mathcal{S}_n(\boldsymbol{Q}_n)} \mathbb{P}\big[(\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y}) \in T(P_{X_1X_2Y})\big] \times \max_{\widetilde{P}_{X_1X_2Y} \in \mathcal{T}_{12,n}(P_{X_1X_2Y})} p_{\mathcal{E}}(\widetilde{P}_{X_1X_2Y}) \quad (43)$$

$$\doteq \max_{P_{X_1X_2Y} \in \mathcal{S}_n(\boldsymbol{Q}_n)} \exp\big(-nD(P_{X_1X_2Y} \| Q_1 \times Q_2 \times W) \big) \times \max_{\widetilde{P}_{X_1X_2Y} \in \mathcal{T}_{12,n}(P_{X_1X_2Y})} p_{\mathcal{E}}(\widetilde{P}_{X_1X_2Y}) \quad (44)$$

where (44) follows from the property of types in (70). It remains to determine the exponential behavior of $p_{\mathcal{E}}$.

Lemma 3. *The probability $p_{\mathcal{E}}(\widetilde{P}_{X_1X_2Y})$ satisfies*

$$p_{\mathcal{E}}(\widetilde{P}_{X_1X_2Y}) \doteq \exp\left(-n \max_{\nu\in\{1,2\}} \Psi_\nu(\widetilde{P}_{X_1X_2Y}, R_1, R_2) \right) \quad (45)$$

for any $\widetilde{P}_{X_1X_2Y}$ such that $\widetilde{P}_{X_1X_2Y} \in \mathcal{T}_{12,n}(P_{X_1X_2Y})$ for some $P_{X_1X_2Y} \in \mathcal{S}_n(\boldsymbol{Q})$, where Ψ_1 and Ψ_2 are defined in (22) *and* (23) *respectively.*

Proof: See Section III-A for the upper bound, and Section III-C for the matching lower bound. ∎

Substituting (45) into (44) and noting that the sets $\mathcal{S}_n$ and $\mathcal{T}_{12,n}$ can be replaced by $\mathcal{S}$ and $\mathcal{T}_{12}$ respectively, we recover the exponent in (21), and the proof is complete.

A. Upper Bound on $p_{\mathcal{E}}(\widetilde{P}_{X_1X_2Y})$

In this subsection, it will be convenient to write $p_{\mathcal{E}}$ as

$$p_{\mathcal{E}}(\widetilde{P}_{X_1X_2Y}) = \mathbb{P}\bigg[\bigcup_{i\neq 1}\bigcup_{j\neq 1}(\boldsymbol{X}_1^{(i)},\boldsymbol{X}_2^{(j)},\boldsymbol{y}) \in T(\widetilde{P}_{X_1X_2Y})\bigg] \quad (46)$$

where $\boldsymbol{y}$ is an arbitrary sequence such that $\boldsymbol{y} \in T(\widetilde{P}_Y)$. We upper bound the probability in (46) by applying the truncated union bound to *one union at a time*. Since there are $M_2 - 1$ identically distributed codewords for user 2, we have

$$p_{\mathcal{E}}(\widetilde{P}_{X_1X_2Y}) \le \min\bigg\{1, (M_2-1) \times \mathbb{P}\bigg[\bigcup_{i\neq 1}(\boldsymbol{X}_1^{(i)},\overline{\boldsymbol{X}}_2,\boldsymbol{y}) \in T(\widetilde{P}_{X_1X_2Y})\bigg]\bigg\} \quad (47)$$

$$= \min\bigg\{1, (M_2-1) \times \mathbb{E}\bigg[\mathbb{P}\bigg[\bigcup_{i\neq 1}(\boldsymbol{X}_1^{(i)},\overline{\boldsymbol{X}}_2,\boldsymbol{y}) \in T(\widetilde{P}_{X_1X_2Y}) \,\bigg|\, \overline{\boldsymbol{X}}_2\bigg]\bigg\}\bigg]. \quad (48)$$

Similarly, since there are $M_1 - 1$ identically distributed codewords for user 1, we obtain

$$p_{\mathcal{E}}(\widetilde{P}_{X_1X_2Y}) \le \min\bigg\{1, (M_2-1)\mathbb{E}\bigg[\min\bigg\{1, (M_1-1)\mathbb{P}\big[(\overline{\boldsymbol{X}}_1,\overline{\boldsymbol{X}}_2,\boldsymbol{y}) \in T(\widetilde{P}_{X_1X_2Y}) \,\big|\, \overline{\boldsymbol{X}}_2\big]\bigg\}\bigg]\bigg\}. \quad (49)$$

The inner probability in (49) is zero unless $(\overline{\boldsymbol{X}}_2,\boldsymbol{y}) \in T(\widetilde{P}_{X_2Y})$, since any other joint marginal must give a joint type of $(\overline{\boldsymbol{X}}_1,\overline{\boldsymbol{X}}_2,\boldsymbol{y})$ which differs from $\widetilde{P}_{X_1X_2Y}$. Hence, instead of writing the expectation in (49) as a summation over joint types of $(\overline{\boldsymbol{X}}_2,\boldsymbol{y})$, we can limit attention to the case that $(\overline{\boldsymbol{X}}_2,\boldsymbol{y}) \in T(\widetilde{P}_{X_2Y})$, yielding

$$p_{\mathcal{E}}(\widetilde{P}_{X_1X_2Y}) \le \min\bigg\{1, (M_2-1)\mathbb{P}\big[(\overline{\boldsymbol{X}}_2,\boldsymbol{y}) \in T(\widetilde{P}_{X_2Y})\big] \times \min\bigg\{(M_1-1)\mathbb{P}\big[(\overline{\boldsymbol{X}}_1,\overline{\boldsymbol{x}}_2,\boldsymbol{y}) \in T(\widetilde{P}_{X_1X_2Y})\bigg\}\bigg\} \quad (50)$$

where $\overline{\boldsymbol{x}}_2$ is an arbitrary sequence such that $(\overline{\boldsymbol{x}}_2,\boldsymbol{y}) \in T(\widetilde{P}_{X_2Y})$. Substituting the properties of types in (68) and (69) into (50) yields

$$p_{\mathcal{E}}(\widetilde{P}_{X_1X_2Y}) \,\dot{\le}\, \exp\big(-n\Psi_1(\widetilde{P}_{X_1X_2Y},R_1,R_2)\big) \quad (51)$$

where Ψ_1 is defined in (22). By following the steps from (47)–(51) with the union bounds applied in the opposite order, it can similarly be shown that

$$p_{\mathcal{E}}(\widetilde{P}_{X_1X_2Y}) \,\dot{\le}\, \exp\big(-n\Psi_2(\widetilde{P}_{X_1X_2Y},R_1,R_2)\big) \quad (52)$$

where Ψ_1 is defined in (23). We therefore obtain the right-hand side of (45).

B. Discussion

The key idea used in Section III-A is to apply the union bound to (46) one union at a time. If the union bound was instead applied to all $(M_1-1)(M_2-1)$ events at once, then the inner $[\cdot]^+$ functions of (22) and (23) would have been replaced by their argument, yielding the exponent $E'_{r,12}$ in (32) and the corresponding achievable rate condition in (33). Hence, only the refined analysis is powerful enough to yield the ensemble-tight error exponent.

We state without proof that under ML decoding (i.e. $q(x_1,x_2,y) = W(y|x_1,x_2)$), the overall error exponent $E_r(\boldsymbol{Q},R_1,R_2)$ given in (24) is unchanged when $E'_{r,12}$ in (32) is used in place of $E_{r,12}$.[1] That is, while the refined analysis of Section III-A is necessary to obtain the ensemble-tight error exponent $E_r(\boldsymbol{Q},R_1,R_2)$ under mismatched decoding, the analysis of [6] suffices under ML decoding.

C. Lower Bound on $p_{\mathcal{E}}(\widetilde{P}_{X_1X_2Y})$

In order to lower bound $p_{\mathcal{E}}$, we will make use of the following result due to de Caen [11].

Proposition 4. [11] *Let $A_1,...,A_k$ be a sequence of probabilistic events. Then*

$$\mathbb{P}\left[\bigcup_{i=1}^k A_i\right] \ge \sum_{i=1}^k \frac{\mathbb{P}[A_i]^2}{\sum_{j=1}^k \mathbb{P}[A_i \cap A_j]}. \quad (53)$$

We begin by rewriting (42) as

$$p_{\mathcal{E}}(\widetilde{P}_{X_1X_2Y}) = \mathbb{P}\left[\bigcup_{i\neq 1, j\neq 1}\big\{\mathcal{E}_{ij}(\widetilde{P}_{X_1X_2Y})\big\} \,\middle|\, \boldsymbol{Y} \in T(P_Y)\right] \quad (54)$$

where $\mathcal{E}_{ij}(\widetilde{P}_{X_1X_2Y})$ is the event that $(\boldsymbol{X}_1^{(i)},\boldsymbol{X}_2^{(j)},\boldsymbol{Y}) \in T(\widetilde{P}_{X_1X_2Y})$. Using (53), we obtain from (54) that

$$p_{\mathcal{E}}(\widetilde{P}_{X_1X_2Y}) \ge \sum_{i\neq 1, j\neq 1} \frac{\mathbb{P}[\mathcal{E}_{ij}]^2}{\sum_{i'\neq 1, j'\neq 1}\mathbb{P}[\mathcal{E}_{ij} \cap \mathcal{E}_{i'j'}]} \quad (55)$$

where the argument to $\mathcal{E}_{ij}$ and the conditioning of the probabilities on the event $\boldsymbol{Y} \in T(\widetilde{P}_Y)$ is kept implicit.

[1] While it is possible that $E_{r,12} > E'_{r,12}$ under ML decoding, it can be shown that this never occurs in the region where $E_{r,12}$ is the dominant exponent (i.e. achieves the minimum in (24))).

We claim that the pairwise probabilities of the events $\mathcal{E}_{ij}(\widetilde{P}_{X_1X_2Y})$ satisfy

$$\mathbb{P}[\mathcal{E}_{ij} \cap \mathcal{E}_{ij}] = e^{-nD(\widetilde{P}_{X_1X_2Y}\|Q_1\times Q_2\times \widetilde{P}_Y)} \tag{56}$$

$$\mathbb{P}[\mathcal{E}_{ij} \cap \mathcal{E}_{i'j}] = e^{-n\left(D(\widetilde{P}_{X_1X_2Y}\|Q_1\times Q_2\times \widetilde{P}_Y)+I_{\widetilde{P}}(X_1;X_2,Y)\right)} \tag{57}$$

$$\mathbb{P}[\mathcal{E}_{ij} \cap \mathcal{E}_{ij'}] = e^{-n\left(D(\widetilde{P}_{X_1X_2Y}\|Q_1\times Q_2\times \widetilde{P}_Y)+I_{\widetilde{P}}(X_2;X_1,Y)\right)} \tag{58}$$

$$\mathbb{P}[\mathcal{E}_{ij} \cap \mathcal{E}_{i'j'}] = e^{-n\,2D(\widetilde{P}_{X_1X_2Y}\|Q_1\times Q_2\times \widetilde{P}_Y)} \tag{59}$$

where $i \neq i'$ and $j \neq j'$. The first case follows from the property of types in (71), and the final case follows from the pairwise conditional independence of $\mathcal{E}_{ij}$ and $\mathcal{E}_{i'j'}$ when $i \neq i'$ and $j \neq j'$. The second case follows from the property of types in (72), whose proof is outlined in the Appendix. The third case is analogous to the second with the roles of users 1 and 2 reversed.

An inspection of the denominator in (55) reveals that there are 1, M_1-2, M_2-2 and $(M_1-2)(M_2-2)$ terms in the sum corresponding to the four cases in (56)–(59) respectively. Furthermore, by symmetry, each term in the outer summation of (55) is equal. Hence, substituting (56)–(59) into (55) and canceling a common term of $e^{-nD(\widetilde{P}_{X_1X_2Y}\|Q_1\times Q_2\times \widetilde{P}_Y)}$ from the numerator and denominator, we obtain

$$p_{\mathcal{E}}(\widetilde{P}_{X_1X_2Y}) \dot{\geq} (M_1-1)(M_2-1)e^{-nD(\widetilde{P}_{X_1X_2Y}\|Q_1\times Q_2\times \widetilde{P}_Y)} \times\Bigg(1+(M_1-2)e^{-nI_{\widetilde{P}}(X_1;X_2,Y)}+(M_2-2)e^{-nI_{\widetilde{P}}(X_2;X_1,Y)} + (M_1-2)(M_2-2)e^{-nD(\widetilde{P}_{X_1X_2Y}\|Q_1\times Q_2\times \widetilde{P}_Y)}\Bigg)^{-1} \tag{60}$$

$$\doteq M_1M_2e^{-nD(\widetilde{P}_{X_1X_2Y}\|Q_1\times Q_2\times \widetilde{P}_Y)} \times\Bigg(1+M_1M_2e^{-nD(\widetilde{P}_{X_1X_2Y}\|Q_1\times Q_2\times \widetilde{P}_Y)} + \max\Big\{M_1e^{-nI_{\widetilde{P}}(X_1;X_2,Y)}, M_2e^{-nI_{\widetilde{P}}(X_2;X_1,Y)}\Big\}\Bigg)^{-1} \tag{61}$$

where (61) follows since the sum of two terms has the same exponential behavior as their maximum. Let us first consider the case that the maximum in (61) is achieved by $M_1e^{-nI_{\widetilde{P}}(X_1;X_2,Y)}$. In this case, we have

$$p_{\mathcal{E}}(\widetilde{P}_{X_1X_2Y}) \dot{\geq} M_1M_2e^{-nD(\widetilde{P}_{X_1X_2Y}\|Q_1\times Q_2\times \widetilde{P}_Y)}\Bigg(1 +M_1e^{-nI_{\widetilde{P}}(X_1;X_2,Y)}+M_1M_2e^{-nD(\widetilde{P}_{X_1X_2Y}\|Q_1\times Q_2\times \widetilde{P}_Y)}\Bigg)^{-1} \tag{62}$$

$$= M_2e^{-nI_{\widetilde{P}}(X_2;Y)}\frac{M_1e^{-nI_{\widetilde{P}}(X_1;X_2,Y)}}{1+M_1e^{-nI_{\widetilde{P}}(X_1;X_2,Y)}} \times\left(1+M_2e^{-nI_{\widetilde{P}}(X_2;Y)}\frac{M_1e^{-nI_{\widetilde{P}}(X_1;X_2,Y)}}{1+M_1e^{-nI_{\widetilde{P}}(X_1;X_2,Y)}}\right)^{-1} \tag{63}$$

where (63) follows by dividing the numerator and denominator by $1+M_1e^{-nI_{\widetilde{P}}(X_1;X_2,Y)}$ and making use of (30). Applying the inequality $\frac{a}{1+a} \geq \frac{1}{2}\min\{1,a\}$ twice, we obtain

$$p_{\mathcal{E}}(\widetilde{P}_{X_1X_2Y}) \dot{\geq} \min\Big\{1, M_2e^{-nI_{\widetilde{P}}(X_2;Y)} \times\min\big\{1, M_1e^{-nI_{\widetilde{P}}(X_1;X_2,Y)}\big\}\Big\} \tag{64}$$

$$= \exp\big(-n\Psi_1(\widetilde{P}_{X_1X_2Y}, R_1, R_2)\big) \tag{65}$$

where Ψ_1 is defined in (22). Similarly, in the case that the maximum in (61) is achieved by $M_2e^{-nI_{\widetilde{P}}(X_2;X_1,Y)}$, we obtain

$$p_{\mathcal{E}}(\widetilde{P}_{X_1X_2Y}) \dot{\geq} \exp\big(-n\Psi_2(\widetilde{P}_{X_1X_2Y}, R_1, R_2)\big) \tag{66}$$

where Ψ_2 is defined in (23). Combining (65) and (66), we obtain the right-hand side of (45).

APPENDIX

Here we state the properties of types used in this paper. We use the notation and definitions given at the beginning of Section II. The random variables $(\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y}, \overline{\boldsymbol{X}}_1, \overline{\boldsymbol{X}}_2, \overline{\overline{\boldsymbol{X}}}_1, \overline{\overline{\boldsymbol{X}}}_2)$ are distributed according to

$$Q_{\boldsymbol{X}_1}(\boldsymbol{x}_1)Q_{\boldsymbol{X}_2}(\boldsymbol{x}_2)W(\boldsymbol{y}|\boldsymbol{x}_1,\boldsymbol{x}_2) \times Q_{\boldsymbol{X}_1}(\overline{\boldsymbol{x}}_1)Q_{\boldsymbol{X}_2}(\overline{\boldsymbol{x}}_2)Q_{\boldsymbol{X}_1}(\overline{\overline{\boldsymbol{x}}}_1)Q_{\boldsymbol{X}_2}(\overline{\overline{\boldsymbol{x}}}_2). \tag{67}$$

We then have the following.

1) For $\nu \in \{1,2\}$, if $\boldsymbol{y} \in T(P_Y)$ then

$$\mathbb{P}\Big[(\overline{\boldsymbol{X}}_\nu, \boldsymbol{y}) \in T(\widetilde{P}_{X_\nu Y})\Big] \doteq \exp\big(-nI_{\widetilde{P}}(X_\nu;Y)\big) \tag{68}$$

for any $\widetilde{P}_{X_\nu Y}$ such that $\widetilde{P}_{X_\nu} = Q_{\nu,n}$ and $\widetilde{P}_Y = P_Y$.

2) If $(\boldsymbol{x}_2, \boldsymbol{y}) \in T(P_{X_2Y})$ then

$$\mathbb{P}\Big[(\overline{\boldsymbol{X}}_1, \boldsymbol{x}_2, \boldsymbol{y}) \in T(\widetilde{P}_{X_1X_2Y})\Big] \doteq \exp\big(-nI_{\widetilde{P}}(X_1;X_2,Y)\big) \tag{69}$$

for any $\widetilde{P}_{X_1X_2Y}$ such that $\widetilde{P}_{X_1} = Q_{1,n}$ and $\widetilde{P}_{X_2Y} = P_{X_2Y}$.

3) The probability of $(\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y})$ having a given type satisfies

$$\mathbb{P}\left[(\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y}) \in T(P_{X_1X_2Y})\right] \doteq \exp\big(-nD(P_{X_1X_2Y}\|Q_1\times Q_2\times W)\big) \tag{70}$$

for any $P_{X_1X_2Y}$ with marginals $P_{X_1} = Q_{1,n}$ and $P_{X_2} = Q_{2,n}$.

4) If $\boldsymbol{y} \in T(P_Y)$, then

$$\mathbb{P}\Big[(\overline{\boldsymbol{X}}_1, \overline{\boldsymbol{X}}_2, \boldsymbol{y}) \in T(\widetilde{P}_{X_1X_2Y})\Big] \doteq \exp\Big(-nD\big(\widetilde{P}_{X_1X_2Y}\|Q_1\times Q_2\times \widetilde{P}_Y\big)\Big) \tag{71}$$

$$\mathbb{P}\Big[(\overline{\boldsymbol{X}}_1, \overline{\boldsymbol{X}}_2, \boldsymbol{y}) \in T(\widetilde{P}_{X_1X_2Y}) \cap (\overline{\overline{\boldsymbol{X}}}_1, \overline{\boldsymbol{X}}_2, \boldsymbol{y}) \in T(\widetilde{P}_{X_1X_2Y})\Big] \doteq \exp\bigg(-n\Big(D\big(\widetilde{P}_{X_1X_2Y} \| Q_1 \times Q_2 \times \widetilde{P}_Y\big) + I_{\widetilde{P}}(X_1; X_2, Y)\Big)\bigg) \quad (72)$$

where each equation holds for any $\widetilde{P}_{X_1X_2Y}$ such that $\widetilde{P}_{X_1} = Q_{1,n}$, $\widetilde{P}_{X_2} = Q_{2,n}$ and $\widetilde{P}_Y = P_Y$.

The proofs of (68)–(71) are omitted, since each is either a known property or a straightforward extension thereof, e.g. see [7], [8]. To prove (72), we write the left-hand side as

$$\mathbb{P}\left[(\overline{\boldsymbol{X}}_2, \boldsymbol{y}) \in T(\widetilde{P}_{X_2Y})\right] \mathbb{P}\left[(\overline{\boldsymbol{X}}_1, \overline{\boldsymbol{x}}_2, \boldsymbol{y}) \in T(\widetilde{P}_{X_1X_2Y})\right]^2 \quad (73)$$

where $\overline{\boldsymbol{x}}_2$ is an arbitrary sequence such that $(\overline{\boldsymbol{x}}_2, \boldsymbol{y}) \in T(\widetilde{P}_{X_2Y})$. Substituting (68) and (69) into (73) and using the identity in (30), we obtain (72).

Acknowledgement

The authors would like to thank Alfonso Martinez for helpful discussions.

References

[1] A. Lapidoth, "Mismatched decoding and the multiple-access channel," *IEEE Trans. Inf. Theory*, vol. 42, no. 5, pp. 1439–1452, Sep. 1996.

[2] N. Merhav, G. Kaplan, A. Lapidoth, and S. Shamai, "On information rates for mismatched decoders," *IEEE Trans. Inf. Theory*, vol. 40, no. 6, pp. 1953–1967, Nov. 1994.

[3] J. Hui, "Fundamental issues of multiple accessing," Ph.D. dissertation, MIT, 1983.

[4] I. Csiszár and P. Narayan, "Channel capacity for a given decoding metric," *IEEE Trans. Inf. Theory*, vol. 45, no. 1, pp. 35–43, Jan. 1995.

[5] A. Ganti, A. Lapidoth, and E. Telatar, "Mismatched decoding revisited: general alphabets, channels with memory, and the wide-band limit," *IEEE Trans. Inf. Theory*, vol. 46, no. 7, pp. 2315–2328, Nov. 2000.

[6] A. Nazari, A. Anastasopoulos, and S. S. Pradhan, "Error exponent for multiple-access channels: Lower bounds," arXiv:1010.1303v1 [cs.IT].

[7] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, 2nd ed. Cambridge University Press, 2011.

[8] R. Gallager, "Fixed composition arguments and lower bounds to error probability," http://web.mit.edu/gallager/www/notes/notes5.pdf.

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

[10] J. Löfberg, "YALMIP : A toolbox for modeling and optimization in MATLAB," in *Proc. CACSD Conf.*, Taipei, Taiwan, 2004.

[11] D. de Caen, "A lower bound on the probability of a union," *Discrete Mathematics*, vol. 169, pp. 217–220, 1997.