

Error Exponent for Multiple-Access Channels: Lower Bounds

Ali Nazari, Achilleas Anastasopoulos and S. Sandeep Pradhan ,
Department of Electrical Engineering and Computer Science,
University of Michigan, Ann Arbor, MI 48109, USA

email: anazari@umich.edu, anastas@umich.edu, pradhanv@eecs.umich.edu

Abstract

A unified framework to obtain all known lower bounds (random coding, typical random coding and expurgated bound) on the reliability function of a point-to-point discrete memoryless channel (DMC) is presented. By using a similar idea for a two-user discrete memoryless (DM) multiple-access channel (MAC), three lower bounds on the reliability function are derived. The first one (random coding) is identical to the best known lower bound on the reliability function of DM-MAC. It is shown that the random coding bound is the performance of the average code in the constant composition code ensemble. The second bound (Typical random coding) is the typical performance of the constant composition code ensemble. To derive the third bound (expurgated), we eliminate some of the codewords from the codebook with larger rate. This is the first bound of this type that explicitly uses the method of expurgation for MACs. It is shown that the exponent of the typical random coding and the expurgated bounds are greater than or equal to the exponent of the known random coding bounds for all rate pairs. Moreover, an example is given where the exponent of the expurgated bound is strictly larger. All these bounds can be universally obtained for all discrete memoryless MACs with given input and output alphabets.

I. INTRODUCTION

In this paper, we consider the problem of communication over a multiple-access channel (MAC) without feedback in the discrete memoryless setting. In particular, we consider the error exponents for this channel model. In this model, two transmitters wish to communicate reliably two independent messages to a single decoder. A schematic is depicted in Figure 1.

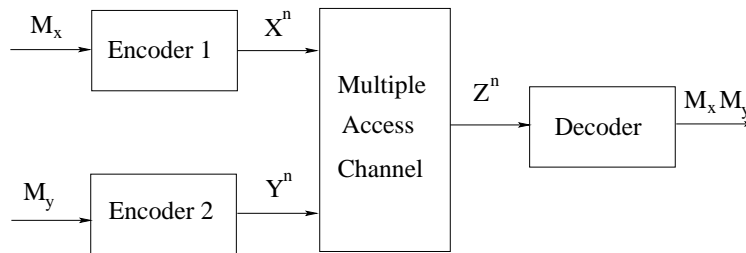


Fig. 1. A schematic of two-user multiple-access channel

Error exponents have been meticulously studied for point to point discrete memoryless channels (DMCs) in the literature [1]–[7]. The optimum error exponent $E(R)$ at some fixed transmission rate

This work was supported by NSF grants CCF 0427385 and CCF 0448115. The material in this paper was presented in part at the Information Theory and Applications conference, San Diego, and also at the Conference on Information Sciences and Systems, the John Hopkins University, Baltimore, 2009.

R (also known as the channel reliability function) gives the decoding error probability exponential rate of decay as a function of block-length for the best sequence of codes. Lower and upper bounds on the channel reliability function for the DMC are known. A lower bound, known as the random coding exponent, was developed by Fano [3] by upper-bounding the average error probability over an ensemble of codes. This bound is loose at low rates. Gallager [8] demonstrated that the random coding bound is the true average error exponent for the random code ensemble. This result illustrates that the weakness of the random coding bound, at low rates, is not due to upper-bounding the ensemble average. Rather, this weakness is due to the fact that the best codes perform much better than the average, especially at low rates. The random coding exponent is further improved at low rates by the process of “expurgation” [9]–[11]. The expurgated bound coincides with the upper bound on the reliability function at $R = 0$ [12, pg. 189]. Barg and Forney [13] investigated another lower bound for the binary symmetric channel (BSC), called the “typical” random coding bound. The authors showed that almost all codes in the standard random coding ensemble exhibit a performance that is as good as the one described by the typical random coding bound. In addition, they showed that the typical error exponent is larger than the random coding exponent and smaller than the expurgated exponent at low rates. Regarding discrete memoryless multiple-access channels (DM-MACs), stronger versions of Ahlswede and Liao’s coding theorem [14], [15], giving exponential upper and lower bounds for the error probability, were derived by several authors. Slepian and Wolf [16], Dyachkov [17], Gallager [18], Pokorny and Wallmeier [19], and Liu and Hughes [20] studied upper bounds on the error probability. Haroutunian [21] and Nazari [22]–[24] studied lower bounds on the error probability.

Comparing the state of the art in the study of error exponents for DMCs and DM-MACs, we observe that the latter is much less advanced. We believe the main difficulty in the study of error exponents for DM-MACs is the fact that error performance in a DM-MAC depends on the pair of codebooks (in the case of a two-user MAC) used by the two transmitters, while at the same time, each transmitter can only control its own codebook. This simple fact has important consequences. For instance, expurgation has not been studied in MAC, since by eliminating some of the “bad” codeword pairs, we may end up with a set of correlated input sequences, which is hard to analyze. In this paper, we develop two new lower bounds for the reliability function of DM-MACs. These bound outperform the bounds of [19], [20].

Toward this goal, we first revisit the point-to-point case and look at the techniques that are used for obtaining the lower bounds on the optimum error exponents. The techniques can be broadly classified into three categories. The first is the Gallager technique [8]. Although this yields expressions for the error exponents that are computationally easier to evaluate than others, the expressions themselves are harder to interpret. The second is the Csiszar-Korner technique [12]. This technique gives more intuitive expressions for the error exponents in terms of optimization of an objective function involving information quantities over probability distributions. This approach is more amenable to generalization to multi-user channels. The third is the graph decomposition technique using α -decoding [25]. α -decoding is a class of decoding procedures that includes maximum likelihood decoding and minimum entropy decoding. Although this technique gives a simpler derivation of the exponents, we believe that it is harder to generalize this to multi-user channels. All three classes of techniques give expressions for the random coding and expurgated exponents. The expressions obtained by the three techniques appear in different forms.

In developing our main result, we first develop a new simpler technique for deriving the random coding and expurgated exponents for the point-to-point channel using a constant composition code ensemble

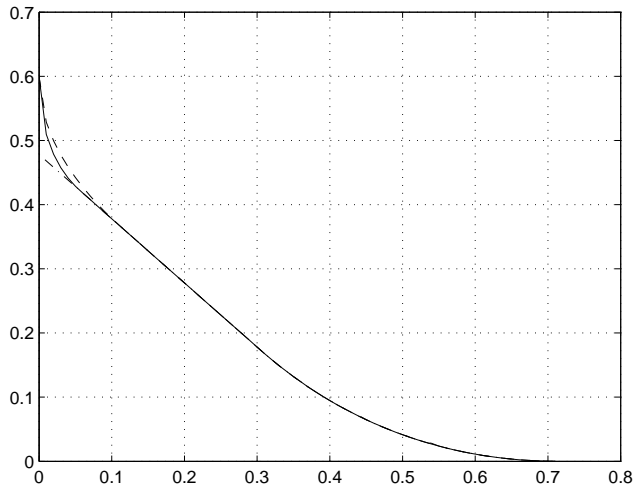


Fig. 2. Lower bounds on the reliability function for point-to-point channel (random coding $-$, typical random coding $-$, expurgated $-$)

with α -decoding. We present our results in the format given in [25]. This technique also gives upper bounds on the ensemble averages. As a bonus, we obtain the typical random coding exponent for this channel. This gives an exact characterization (lower and upper bounds that meet) of the error exponent of almost all codes in the ensemble. When specialized to the BSC, this reduces to the typical random coding bound of Barg and Forney [13]¹. Fig. 2 shows the random coding, the typical random coding, and the expurgated bounds for a BSC with crossover probability $p = 0.05$, which is representative of the general case. All the three lower bounds are expressed as minimizations of a single objective function under different constraint sets. The reasons for looking at typical performance are two-fold. The first is that the average error exponent is in general smaller than the typical error exponent at low rates, hence the latter gives a tighter characterization of the optimum error exponent of the channel. For example, for the BSC, although the average performance of the linear code ensemble is given by the random coding exponent of the Gallager ensemble, the typical performance is given by the expurgated exponent of the Gallager ensemble. In this direction, it was also noted recently in [26] that for the 8-PSK Gaussian channel, the typical performance of the ensemble of group codes over \mathbb{Z}_8 equals the expurgated exponent of the Gallager ensemble, whereas the typical performance of the ensemble of binary coset codes (under any mapping) is bounded away from the same. The second is that in some cases, expurgation may not be possible or may not be desirable. For example, (a) in the MAC, the standard expurgation is not possible, and (b) if one is looking at the performance of the best linear code for a channel, then expurgation destroys the linear structure which is not desirable. In the proposed technique we provide a unified way to derive all the three lower bounds on the optimum error exponents, and upper bounds on the ensemble average and the typical performance. We wish to note that the bounds derived in this paper are universal in nature. The proposed approach appears to be more amenable to generalization to multi-user channels.

A brief outline of the technique is given as follows. First, for a given constant composition code, we

¹Barg and Forney gave only a lower bound in [13].

define a pair of packing functions that are independent of the channel. For an arbitrary channel, we relate the probability of error of a code with α -decoding to its packing functions. Packing functions give pair-wise and triple-wise joint-type distributions of the code. This is similar in spirit to the concept of distance distribution of the code. Then we do random coding and obtain lower and upper bounds on the expected value of the packing functions of the ensemble without interfacing it with the channel. That is, these bounds do not depend on the channel. Finally, using the above relation between the packing function and the probability of error, we get single-letter expressions for the bounds on the optimum error exponents for an arbitrary channel.

Toward extending this technique to MACs, we follow a three-step approach. We start with a constant conditional composition ensemble identical to [20]. Then, we provide a new packing lemma in which the resulting code has better properties in comparison to the packing lemmas in [19] and [20]. This packing lemma is similar to Pokorny's packing lemma, in the sense that the channel conditional distribution does not appear in the inequalities. One of the advantages of our methodology is that it enables us to partially expurgate some of the codewords and end up with a new code with stronger properties. In particular, we do not eliminate pairs of codewords. Rather, we expurgate codewords from only one of the codebooks and analyze the performance of the expurgated code.

Contributions: In summary the key contributions of this work are

- An exact characterization of the typical error exponent for the constant composition code ensemble for the DMC.
- A new lower bound on the optimum error exponent for the MAC.
- An upper bound on the average error exponent of the constant composition code ensemble for the MAC.
- A characterization of the typical error exponent for the constant composition code ensemble for the MAC.

This paper is organized as follows: Section II introduces terminology, and Section III unifies the derivation of all lower bounds on the reliability function for a point-to-point DMC. Our main results for the DM-MAC are introduced in Section IV. Some numerical results are presented in Section V, and Section VI concludes the paper. The proofs of some of these results are given in the Appendix.

II. PRELIMINARIES

We will follow the notation of [12]. For any finite alphabet \mathcal{X} , let $\mathcal{P}(\mathcal{X})$ denote the set of all probability distributions on \mathcal{X} . For any sequence $\mathbf{x} \in \mathcal{X}^n$, let $P_{\mathbf{x}}$ denote its type. Let T_P denote the type class of type P . Let $\mathcal{P}_n(\mathcal{X})$ denote the set of all types on \mathcal{X} . Let T_V denote a V-shell, and $D(V||W|P)$ denote conditional I-divergence. In this paper, we consider channels without feedback.

Definition 1. A discrete memoryless channel (DMC) is defined by a stochastic matrix $W : \mathcal{X} \rightarrow \mathcal{Y}$, where \mathcal{X} , the input alphabet, and \mathcal{Y} , the output alphabet, are finite sets. The channel transition probability for n -sequences is given by

$$W^n(\mathbf{y}|\mathbf{x}) \triangleq \prod_{i=1}^n W(y_i|x_i),$$

where $\mathbf{x} \triangleq (x_1, \dots, x_n) \in \mathcal{X}^n$, $\mathbf{y} \triangleq (y_1, \dots, y_n) \in \mathcal{Y}^n$. An (n, M) code for a given DMC, W , is a set $C = \{(\mathbf{x}_i, D_i) : 1 \leq i \leq M\}$ with (a) $\mathbf{x}_i \in \mathcal{X}^n$, $D_i \subset \mathcal{Y}^n$ and (b) $D_i \cap D_{i'} = \emptyset$ for $i \neq i'$.

When message i is transmitted, the conditional probability of error of a code C is given by

$$e_i(C, W) \triangleq W^n(D_i^c | \mathbf{x}_i).$$

The average probability of error for this code is defined as

$$e(C, W) \triangleq \frac{1}{M} \sum_{i=1}^M e_i(C, W). \quad (1)$$

Definition 2. For the DMC, $W : \mathcal{X} \rightarrow \mathcal{Y}$, the average error exponent, at rate R , is defined as:

$$E_{av}^*(R) \triangleq \limsup_{n \rightarrow \infty} \max_{C \in \mathcal{C}} -\frac{1}{n} \log e(C, W), \quad (2)$$

where \mathcal{C} is the set of all codes of length n and rate R . The typical average error exponent of an ensemble \mathcal{C} , at rate R , is defined as:

$$E_{av}^T(R) \triangleq \liminf_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \max_{\tilde{\mathcal{C}}: \mathbb{P}(\tilde{\mathcal{C}}) > 1 - \delta} \min_{C \in \tilde{\mathcal{C}}} -\frac{1}{n} \log e(C, W). \quad (3)$$

where \mathbb{P} is the uniform distribution over \mathcal{C} .

The typical error exponent is basically the exponent of the average error probability of the worst code belonging to the best high probable collection of the ensemble.

Definition 3. A two-user DM-MAC is defined by a stochastic matrix $W : \mathcal{X} \times \mathcal{Y} \rightarrow \mathcal{Z}$, where \mathcal{X} , \mathcal{Y} , the input alphabets, and \mathcal{Z} , the output alphabet, are finite sets. The channel transition probability for n -sequences is given by

$$W^n(\mathbf{z} | \mathbf{x}, \mathbf{y}) \triangleq \prod_{i=1}^n W(z_i | x_i, y_i), \quad (4)$$

where $\mathbf{x} \triangleq (x_1, \dots, x_n) \in \mathcal{X}^n$, $\mathbf{y} \triangleq (y_1, \dots, y_n) \in \mathcal{Y}^n$, and $\mathbf{z} \triangleq (z_1, \dots, z_n) \in \mathcal{Z}^n$.

An (n, M, N) multi-user code for a given MAC, W , is a set $C = \{(\mathbf{x}_i, \mathbf{y}_j, D_{ij}) : 1 \leq i \leq M, 1 \leq j \leq N\}$ with

- $\mathbf{x}_i \in \mathcal{X}^n$, $\mathbf{y}_j \in \mathcal{Y}^n$, $D_{ij} \subset \mathcal{Z}^n$
- $D_{ij} \cap D_{i'j'} = \emptyset$ for $(i, j) \neq (i', j')$.

When message (i, j) is transmitted, the conditional probability of error of the two-user code C is given by

$$e_{ij}(C, W) \triangleq W^n(D_{ij}^c | \mathbf{x}_i, \mathbf{y}_j). \quad (5)$$

The average probability of error for the two-user code, C , is defined as

$$e(C, W) \triangleq \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N e_{ij}(C, W). \quad (6)$$

Definition 4. For the MAC, $W : \mathcal{X} \times \mathcal{Y} \rightarrow \mathcal{Z}$, the average error exponent at rate pair (R_X, R_Y) , is defined as:

$$E_{av}^*(R_X, R_Y) \triangleq \limsup_{n \rightarrow \infty} \max_{C \in \mathcal{C}_M} -\frac{1}{n} \log e(C, W), \quad (7)$$

where \mathcal{C}_M is the set of all codes of length n and rate pair (R_X, R_Y) . The typical average error exponent of an

ensemble \mathcal{C} , at rate pair (R_X, R_Y) , is defined as:

$$E_{av}^T(R_X, R_Y) \triangleq \liminf_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \max_{\tilde{\mathcal{C}} \subset \mathcal{C}: \mathbb{P}(\tilde{\mathcal{C}}) > 1 - \delta} \min_{C \in \tilde{\mathcal{C}}} -\frac{1}{n} \log e(C, W), \quad (8)$$

where \mathbb{P} is the uniform distribution over \mathcal{C} .

III. POINT TO POINT: LOWER BOUNDS ON RELIABILITY FUNCTION

A. Packing functions

Consider the class of DMCs with input alphabet \mathcal{X} and output alphabet \mathcal{Y} . In the following, we introduce a unified way to derive all known lower bounds on the reliability function of such a channel. We will follow the random coding approach. First, we choose a constant composition code ensemble. Then, we define a packing function, $\pi : \mathcal{C} \times \mathcal{P}(\mathcal{X} \times \mathcal{X}) \rightarrow \mathbb{R}$, on all codebooks in the ensemble. The packing function that we use is the average number of codeword pairs sharing a particular joint type, $V_{X\bar{X}}$. Specifically, for $P \in \mathcal{P}_n(\mathcal{X})$, $V_{X\bar{X}} \in \mathcal{P}_n(\mathcal{X} \times \mathcal{X})$, and any code $C = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M\} \subset T_P$, the packing function is defined as:

$$\pi(C, V_{X\bar{X}}) = \frac{1}{M} \sum_{i=1}^M \sum_{j \neq i} 1_{T_{V_{X\bar{X}}}}(\mathbf{x}_i, \mathbf{x}_j). \quad (9)$$

We call this the first order packing function. Using this packing function, we prove three different packing lemmas, each of which shows the existence of a code with some desired properties.

In the first packing lemma, tight upper and lower bounds on the expectation of the packing function over the ensemble are derived. By using this packing lemma, upper and lower bounds on the expectation of the average probability of error over the ensemble are derived. These bounds meet for all transmission rates below the critical rate². In the second packing lemma, by using the expectation and the variance of the packing function, we prove that for almost all codes in the constant composition code ensemble, the bounds in the first packing lemma are still valid. By using this tight bound on the performance of almost every code in the ensemble, we provide a tighter bound on the error exponent which we call the ‘‘typical’’ random coding bound. As we see later in the paper, the typical random coding bound is indeed the typical performance of the constant composition code ensemble. In the third packing lemma, we use one of the typical codes and eliminate some of its ‘‘bad’’ codewords. The resulting code satisfies some stronger constraints in addition to all the previous properties. By using this packing lemma and an efficient decoding rule, we re-derive the well-known expurgated bound.

To provide upper bounds on the average error exponents, such as those given below in Fact 1 and Theorem 1, for every $V_{X\bar{X}\hat{X}} \in \mathcal{P}_n(\mathcal{X} \times \mathcal{X} \times \mathcal{X})$, we define a second packing function $\lambda : \mathcal{C} \times \mathcal{P}(\mathcal{X} \times \mathcal{X} \times \mathcal{X}) \rightarrow \mathbb{R}$ on all codes in the constant composition code ensemble as follows:

$$\lambda(C, V_{X\bar{X}\hat{X}}) \triangleq \frac{1}{M} \sum_{i=1}^M \sum_{j \neq i} \sum_{k \neq i, j} 1_{T_{V_{X\bar{X}\hat{X}}}}(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k). \quad (10)$$

We call this the second order packing function. As it is clear from the definition, this quantity is the average number of codeword triplets sharing a common joint distribution in code C .

²This is essentially a re-derivation of the upper and lower bounds on the average probability of error obtained by Gallager in a different form. The present results are for constant composition codes.

B. Relation between packing function and probability of error

First, we consider the decoding rule at the receiver, and secondly we relate the average probability of error to the packing function.

Decoding Rule: In our derivation, error probability bounds using maximum-likelihood and minimum-entropy decoding rules will be obtained in a unified way. The reason is that both can be given in terms of a real-valued function on the set of distributions on $\mathcal{X} \times \mathcal{Y}$. This type of decoding rule was introduced in [25] as the α -decoding rule. For a given real-valued function α , a given code C , and for a received sequence $\mathbf{y} \in \mathcal{Y}^n$, the α -decoder accepts the codeword $\hat{\mathbf{x}} \in C$ for which the joint type of $\hat{\mathbf{x}}$ and \mathbf{y} minimizes the function α , i.e., the decoder accepts $\hat{\mathbf{x}}$ if

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x} \in C} \alpha(P \cdot V_{\mathbf{y}|\mathbf{x}}). \quad (11)$$

It was shown in [25] that for fixed composition codes, maximum-likelihood and minimum-entropy are special cases of this decoding rule. In particular, for maximum-likelihood decoding,

$$\alpha(P \cdot V) = D(V||W|P) + H(V|P), \quad (12)$$

and for minimum entropy decoding,

$$\alpha(P \cdot V) = H(V|P), \quad (13)$$

where P is the fixed composition of the codebook, and V is the conditional type of \mathbf{y} given \mathbf{x} .

Relation between probability of error and packing function: Next, for a given channel, we derive an upper bound and a lower bound on the average probability of error of an arbitrary constant composition code in terms of its first order and second order packing functions. The rest of the paper is built on this crucial derivation. Consider the following argument about the average probability of error of a code C used on a channel W .

$$\begin{aligned} e(C, W) &= \frac{1}{M} \sum_{i=1}^M W^n(D_i^c|\mathbf{x}_i) = \frac{1}{M} \sum_{i=1}^M W^n(\{\mathbf{y} : \alpha(P \cdot V_{\mathbf{y}|\mathbf{x}_i}) \geq \alpha(P \cdot V_{\mathbf{y}|\mathbf{x}_j}) \text{ for some } j \neq i\}|\mathbf{x}_i) \\ &= \sum_{V_{X\tilde{X}Y} \in \mathcal{P}_n^r} \left(2^{-n[D(V_{Y|X}||W|P) + H_V(Y|X)]} \left[\frac{1}{M} \sum_{i=1}^M A_i(V_{X\tilde{X}Y}, C) \right] \right), \end{aligned} \quad (14)$$

where \mathcal{P}_n^r and $A_i(V_{X\tilde{X}Y}, C)$ are defined as follows

$$\mathcal{P}_n^r \triangleq \left\{ V_{X\tilde{X}Y} \in \mathcal{P}_n(\mathcal{X} \times \mathcal{X} \times \mathcal{Y}) : V_X = V_{\tilde{X}} = P, \alpha(P \cdot V_{Y|\tilde{X}}) \leq \alpha(P, V_{Y|X}) \right\}, \quad (15)$$

$$A_i(V_{X\tilde{X}Y}, C) \triangleq |\{\mathbf{y} : (\mathbf{x}_i, \mathbf{x}_j, \mathbf{y}) \in T_{V_{X\tilde{X}Y}} \text{ for some } j \neq i\}|. \quad (16)$$

From the inclusion-exclusion principle, it follows that $A_i(V_{X\tilde{X}Y}, C)$ satisfies

$$B_i(V_{X\tilde{X}Y}, C) - C_i(V_{X\tilde{X}Y}, C) \leq A_i(V_{X\tilde{X}Y}, C) \leq B_i(V_{X\tilde{X}Y}, C), \quad (17)$$

where

$$B_i(V_{X\tilde{X}Y}, C) \triangleq \sum_{j \neq i} 1_{T_{V_{X\tilde{X}Y}}}(\mathbf{x}_i, \mathbf{x}_j) \left| \left\{ \mathbf{y} : \mathbf{y} \in T_{V_{Y|\mathbf{x}_i}}(\mathbf{x}_i, \mathbf{x}_j) \right\} \right|, \quad (18)$$

$$C_i(V_{X\tilde{X}Y}, C) \triangleq \sum_{j \neq i} \sum_{k \neq i, j} 1_{T_{V_{X\tilde{X}Y}}}(\mathbf{x}_i, \mathbf{x}_j) 1_{T_{V_{X\tilde{X}Y}}}(\mathbf{x}_i, \mathbf{x}_k) \left| \left\{ \mathbf{y} : \mathbf{y} \in T_{V_{Y|\mathbf{x}_i}}(\mathbf{x}_i, \mathbf{x}_j) \cap T_{V_{Y|\mathbf{x}_i}}(\mathbf{x}_i, \mathbf{x}_k) \right\} \right|. \quad (19)$$

Next, we provide an upper bound on the second term on the right hand side of (14) as follows.

$$\frac{1}{M} \sum_{i=1}^M A_i(V_{X\bar{X}Y}, C) \leq \frac{1}{M} \sum_{i=1}^M B_i(V_{X\bar{X}Y}, C) \quad (20a)$$

$$= \frac{1}{M} \sum_{i=1}^M \sum_{j \neq i} 1_{T_{V_{X\bar{X}}}(\mathbf{x}_i, \mathbf{x}_j)} \left| \left\{ \mathbf{y} : \mathbf{y} \in T_{V_{Y|X\bar{X}}}(\mathbf{x}_i, \mathbf{x}_j) \right\} \right| \quad (20b)$$

$$\leq \frac{1}{M} \sum_{i=1}^M \sum_{j \neq i} 1_{T_{V_{X\bar{X}}}(\mathbf{x}_i, \mathbf{x}_j)} 2^{nH(Y|X\bar{X})} \quad (20c)$$

$$= \pi(C, V_{X\bar{X}}) 2^{nH(Y|X\bar{X})} \quad (20d)$$

On the other hand

$$\left\{ \mathbf{y} : (\mathbf{x}_i, \mathbf{x}_j, \mathbf{y}) \in T_{V_{X\bar{X}Y}} \text{ for some } j \neq i \right\} \subset T_{V_{Y|X}}(\mathbf{x}_i), \quad (21)$$

so we can conclude that

$$\frac{1}{M} \sum_{i=1}^M A_i(V_{X\bar{X}Y}, C) \leq 2^{nH_V(Y|X)}. \quad (22)$$

Combining the above with (14), we have an upper bound on the probability of error in terms of the first order packing function as follows.

$$e(C, W) \leq \sum_{V_{X\bar{X}Y} \in \mathcal{P}_n^r} 2^{-n[D(V_{Y|X} || W|P)]} \min \left\{ 2^{-nI_V(\bar{X} \wedge Y|X)} \pi(C, V_{X\bar{X}}), 1 \right\} \quad (23)$$

Next, we consider the lower bound. For that, we provide a lower bound on B_i and upper bound on C_i as follows.

$$\begin{aligned} \frac{1}{M} \sum_{i=1}^M B_i(V_{X\bar{X}Y}, C) &= \frac{1}{M} \sum_{i=1}^M \sum_{j \neq i} 1_{T_{V_{X\bar{X}}}(\mathbf{x}_i, \mathbf{x}_j)} \left| \left\{ \mathbf{y} : \mathbf{y} \in T_{V_{Y|X\bar{X}}}(\mathbf{x}_i, \mathbf{x}_j) \right\} \right| \\ &\geq \pi(C, V_{X\bar{X}}) 2^{n[H(Y|X\bar{X}) - \delta]}, \end{aligned} \quad (24)$$

and

$$\begin{aligned} &\frac{1}{M} \sum_{i=1}^M C_i(V_{X\bar{X}Y}, C) \\ &= \frac{1}{M} \sum_{i=1}^M \sum_{j \neq i} \sum_{k \neq i, j} 1_{T_{V_{X\bar{X}}}(\mathbf{x}_i, \mathbf{x}_j)} 1_{T_{V_{X\bar{X}}}(\mathbf{x}_i, \mathbf{x}_k)} \left| \left\{ \mathbf{y} : \mathbf{y} \in T_{V_{Y|X\bar{X}}}(\mathbf{x}_i, \mathbf{x}_j) \cap T_{V_{Y|X\bar{X}}}(\mathbf{x}_i, \mathbf{x}_k) \right\} \right| \\ &= \sum_{\substack{V_{X\bar{X}\hat{X}Y}: \\ V_{X\bar{X}Y} = V_{X\bar{X}Y}}} \frac{1}{M} \sum_{i=1}^M \sum_{j \neq i} \sum_{k \neq i, j} 1_{T_{V_{X\bar{X}\hat{X}}}(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k)} \left| \left\{ \mathbf{y} : \mathbf{y} \in T_{V_{Y|X\bar{X}\hat{X}}}(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k) \right\} \right| \\ &\leq \sum_{\substack{V_{X\bar{X}\hat{X}Y}: \\ V_{X\bar{X}Y} = V_{X\bar{X}Y}}} 2^{nH(Y|X\bar{X}\hat{X})} \lambda(C, V_{X\bar{X}\hat{X}}) \end{aligned} \quad (25)$$

Combining (14), (24), and (25) we have the following lower bound on the average probability of error.

$$e(C, W) \geq \sum_{V_{X\tilde{X}Y} \in \mathcal{P}_n^+} 2^{-n[D(V_{Y|X}||W|P) + I_V(\tilde{X}\wedge Y|X) + \delta]} \left| \pi(C, V_{X\tilde{X}}) - \sum_{\substack{V_{X\tilde{X}\hat{X}Y}: \\ V_{X\tilde{X}Y} = V_{X\hat{X}Y}}} 2^{-n[I_V(\hat{X}\wedge Y|X\tilde{X})]} \lambda(C, V_{X\tilde{X}\hat{X}}) \right| \quad (26)$$

Observe that these upper and lower bounds apply for every code C . We have accomplished the task of relating the average probability of error to the two packing functions. The key results of this subsection are given by (23) and (26). Next we use the packing lemmas to derive the bounds on the error exponents.

C. Random Coding Packing Lemmas

Lemma 1. (Random Coding Packing Lemma) Fix $R > 0$, $\delta > 0$, a sufficient large n and any type P of sequences in \mathcal{X}^n satisfying $H(P) > R$. For any $V_{X\tilde{X}} \in \mathcal{P}_n(\mathcal{X} \times \mathcal{X})$, the expectation of the first order packing function over the constant composition code ensemble is bounded by

$$2^{n(R - I_V(X\wedge\tilde{X}) - \delta)} \leq \mathbb{E}(\pi(X^M, V_{X\tilde{X}})) \leq 2^{n(R - I_V(X\wedge\tilde{X}) + \delta)}, \quad (27)$$

where $X^M \triangleq (X_1, X_2, \dots, X_M) \subset T_P$ are independent and X_i s are uniformly distributed on T_P , and $2^{n(R-\delta)} \leq M \leq 2^{nR}$. Moreover, the following inequality holds for the second order packing function:

$$\mathbb{E}(\lambda(X^M, V_{X\tilde{X}\hat{X}})) \leq 2^{n[2R - I_V(X\wedge\tilde{X}) - I_V(\hat{X}\wedge X\tilde{X}) + 4\delta]} \quad \text{for all } V_{X\tilde{X}\hat{X}} \in \mathcal{P}_n(\mathcal{X} \times \mathcal{X} \times \mathcal{X}). \quad (28)$$

Proof: The proof follows directly from the fact that two words drawn independently from T_P have a joint type $V_{X\tilde{X}}$ with probability close to $2^{-nI(X\wedge\tilde{X})}$. The details are provided in the Appendix. ■

Lemma 2. (Typical Random Code Packing Lemma) Fix $R > 0$, $\delta > 0$, a sufficient large n and any type P of sequences in \mathcal{X}^n satisfying $H(P) > R$. Almost every code, C^t , with $2^{n(R-\delta)} \leq M \leq 2^{nR}$ codewords, in the constant composition code ensemble satisfies the following inequalities

$$2^{n[R - I_V(X\wedge\tilde{X}) - 2\delta]} \leq \pi(C^t, V_{X\tilde{X}}) \leq 2^{n[R - I_V(X\wedge\tilde{X}) + 2\delta]} \quad \text{for all } V_{X\tilde{X}} \in \mathcal{P}_n(\mathcal{X} \times \mathcal{X}), \quad (29)$$

and

$$\lambda(C^t, V_{X\tilde{X}\hat{X}}) \leq 2^{n[2R - I_V(X\wedge\tilde{X}) - I_V(\hat{X}\wedge X\tilde{X}) + 4\delta]} \quad \text{for all } V_{X\tilde{X}\hat{X}} \in \mathcal{P}_n(\mathcal{X} \times \mathcal{X} \times \mathcal{X}). \quad (30)$$

Proof: The proof is provided in the Appendix. In the proof, we evaluate the variance of the packing function and use Chebyshev's inequality to show that with high probability the packing function is close to its expected value. ■

Lemma 3. (Expurgated Packing Lemma) For every sufficiently large n , every $R > 0$, $\delta > 0$ and every type P of sequences in \mathcal{X}^n satisfying $H(P) > R$, there exists a set of codewords $C^{ex} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{M^*}\} \subset T_P$ with $M^* \geq \frac{2^{n(R-\delta)}}{2}$, such that for any $V_{X\tilde{X}} \in \mathcal{P}_n(\mathcal{X} \times \mathcal{X})$,

$$\pi(C^{ex}, V_{X\tilde{X}}) \leq 2^{n(R - I_V(X\wedge\tilde{X}) + 2\delta)}, \quad (31)$$

and for every sequence $\mathbf{x}_i \in C^{ex}$,

$$|T_{V_{\tilde{X}|X}}(\mathbf{x}_i) \cap C^{ex}| \leq 2^{n(R - I_V(X\wedge\tilde{X}) + 2\delta)}. \quad (32)$$

Proof: The Proof is provided in the Appendix. The basic idea of the proof is simple. From Lemma 1 we know that for every $V_{X\tilde{X}}$, there exists a code whose packing function is upper bounded by a number that is close to $2^{n(R-I_V(X\wedge\tilde{X}))}$. Since the packing function is an average over all codewords in the code, we infer that for at least half of the codewords, the corresponding property (32) is satisfied. In the Appendix, we show that there exists a single code that works for every joint type. ■

D. Error Exponent Bounds

Now, we obtain the bounds on the error exponents using the results from the previous three subsections. We present three lower bounds and two upper bounds. The lower bounds are the random coding exponent, typical random coding exponent and expurgated exponent. All the three lower bounds are expressed as minimization of the same objective function under different constraint sets. Similar structure is manifested in the case of upper bounds. For completeness, we first rederive the well-known result of random coding exponent.

Fact 1. (Random Coding Bound) For every type P of sequences in \mathcal{X}^n and $0 \leq R \leq H(P)$, $\delta > 0$, every DMC, $W : \mathcal{X} \rightarrow \mathcal{Y}$, and $2^{n(R-\delta)} \leq M \leq 2^{nR}$, the expectation of the average error probability over the constant composition code ensemble with M codewords of type P , can be bounded by

$$2^{-n[E_{rL}(R,P,W)+3\delta]} \leq \bar{P}_e \leq 2^{-n[E_r(R,P,W)-2\delta]}, \quad (33)$$

whenever $n \geq n_1(|\mathcal{X}|, |\mathcal{Y}|, \delta)$, where

$$E_r(R, P, W) \triangleq \min_{V_{X\tilde{X}Y} \in \mathcal{P}^r} D(V_{Y|X} \| W|P) + |I_V(\tilde{X} \wedge XY) - R|^+, \quad (34)$$

$$E_{rL}(R, P, W) \triangleq \min_{\substack{V_{X\tilde{X}Y} \in \mathcal{P}^r: \\ I_V(\tilde{X} \wedge XY) \geq R}} D(V_{Y|X} \| W|P) + I_V(\tilde{X} \wedge XY) - R, \quad (35)$$

and

$$\mathcal{P}^r \triangleq \{V_{X\tilde{X}Y} \in \mathcal{P}(\mathcal{X} \times \mathcal{X} \times \mathcal{Y}) : V_X = V_{\tilde{X}} = P, \alpha(P, V_{Y|\tilde{X}}) \leq \alpha(P, V_{Y|X})\}. \quad (36)$$

In particular, there exists a set of codewords $C^r = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M\} \subset T_P$, with $M \geq 2^{n(R-\delta)}$, such that for every DMC, $W : \mathcal{X} \rightarrow \mathcal{Y}$,

$$e(C^r, W) \leq 2^{-n[E_r(R,P,W)-3\delta]}. \quad (37)$$

Proof: The proof is straightforward and is outlined in the Appendix. ■

It is well known that for $R \geq R_{crit}$, the random coding error exponent is equal to the sphere packing error exponent, and as a result the random coding bound is a tight bound. In addition, the following is true.

Corollary 1. For any $R \leq R_{crit}$,

$$\max_{P \in \mathcal{P}(\mathcal{X})} E_{rL}(R, P, W) = \max_{P \in \mathcal{P}(\mathcal{X})} E_r(R, P, W). \quad (38)$$

Proof: The proof is provided in the Appendix. ■

Next we have an exact characterization of the typical performance of the constant composition code ensemble.

Theorem 1. (Typical random Coding Bound) For every type P of sequences in \mathcal{X}^n , $\delta > 0$, and every transmission rate satisfying $0 \leq R \leq H(P)$, almost all codes, $C^t = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M\}$ with $\mathbf{x}_i \in T_P$ for all i , $M \geq 2^{n(R-\delta)}$, satisfy

$$2^{-n[E_{TL}(R,P,W)+4\delta]} \leq e(C^t, W) \leq 2^{-n[E_T(R,P,W)-3\delta]}, \quad (39)$$

for every DMC, $W : \mathcal{X} \rightarrow \mathcal{Y}$, whenever $n \geq n_1(|\mathcal{X}|, |\mathcal{Y}|, \delta)$. Here,

$$E_T(R, P, W) \triangleq \min_{V_{X\tilde{X}Y} \in \mathcal{P}^t} D(V_{Y|X} \| W|P) + |I_V(\tilde{X} \wedge XY) - R|^+, \quad (40)$$

$$E_{TL}(R, P, W) \triangleq \min_{\substack{V_{X\tilde{X}Y} \in \mathcal{P}^t: \\ I_V(\tilde{X} \wedge XY) \geq R}} D(V_{Y|X} \| W|P) + I_V(\tilde{X} \wedge XY) - R, \quad (41)$$

where

$$\mathcal{P}^t \triangleq \{V_{X\tilde{X}Y} \in \mathcal{P}(\mathcal{X} \times \mathcal{X} \times \mathcal{Y}) : V_X = V_{\tilde{X}} = P, I_V(X \wedge \tilde{X}) \leq 2R, \alpha(P, V_{Y|\tilde{X}}) \leq \alpha(P, V_{Y|X})\}. \quad (42)$$

Proof: The proof is provided in the Appendix. ■

In Theorem 1, we proved the existence of a high probability (almost 1) collection of codes such that every code in this collection satisfies (39). This provides a lower bound on the typical average error exponent for the constant composition code ensemble as defined in Definition 2. In the following, we show that the typical performance of the best high-probability collection cannot be better than that given in Theorem 1.

Corollary 2. For every type P of sequences in \mathcal{X}^n , $\delta > 0$, and every transmission rate satisfying $0 \leq R \leq H(P)$,

$$E_T(R, P, W) \leq E_{av}^T(R) \leq E_{TL}(R, P, W), \quad (43)$$

for the constant composition code ensemble.

Proof: The proof is provided in the Appendix. ■

Clearly, since the random coding bound is tight for $R \geq R_{crit}$, the same is true for the typical random coding bound. For $R \leq R_{crit}$ we have the following result.

Corollary 3. For any $R \leq R_{crit}$,

$$\max_{P \in \mathcal{P}(\mathcal{X})} E_{TL}(R, P, W) = \max_{P \in \mathcal{P}(\mathcal{X})} E_T(R, P, W). \quad (44)$$

Proof: The proof is very similar to that of Corollary 1 and is omitted. ■

It can be seen that the typical random coding bound is the true error exponent for almost all codes, with M codewords, in the constant composition code ensemble. A similar lower bound on the typical random coding bound was derived by Barg and Forney [13] for the binary symmetric channel. Although the approach used here is completely different from the one in [13], in the following corollary we show that these two bounds coincide for binary symmetric channels.

Corollary 4. For a binary symmetric channel with crossover probability p , and for $0 \leq R \leq R_{crit}$

$$E_T(R, P, W) = E_{TRC}(R), \quad (45)$$

where E_{TRC} is the lower bound for the error exponent of a typical random code in [13].

Finally, we re-derive the well-known expurgated error exponent in a rather straightforward way.

Fact 2. (Expurgated Bound) For every type P of sequences in \mathcal{X}^n and $0 \leq R \leq H(P)$, $\delta > 0$, there exists a set of codewords $C^{ex} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{M^*}\} \subset T_P$ with $M^* \geq \frac{2^{n(R-\delta)}}{2}$, such that for every DMC, $W : \mathcal{X} \rightarrow \mathcal{Y}$,

$$e(C^{ex}, W) \leq 2^{-n[E_{ex}(R,P,W)-3\delta]} \quad (46)$$

whenever $n \geq n_1(|\mathcal{X}|, |\mathcal{Y}|, \delta)$, where

$$E_{ex}(R, P, W) \triangleq \min_{V_{X\tilde{X}Y} \in \mathcal{P}^{ex}} D(V_{Y|X} || W|P) + |I_V(\tilde{X} \wedge XY) - R|^+ \quad (47)$$

where

$$\mathcal{P}^{ex} \triangleq \{V_{X\tilde{X}Y} \in \mathcal{P}(\mathcal{X} \times \mathcal{X} \times \mathcal{Y}) : V_X = V_{\tilde{X}} = P, I_V(X \wedge \tilde{X}) \leq R, \alpha(P, V_{Y|\tilde{X}}) \leq \alpha(P, V_{Y|X})\} \quad (48)$$

Proof: The proof is provided in the Appendix. ■

Note that none of the mentioned three bounds have their “traditional format” as found in [12], [9], but rather the format introduced in [25] by Csiszar and Korner. It was shown in [25] that the new random coding bound is equivalent to the original one for maximum likelihood and minimum entropy decoding rule. Furthermore, the new format for the expurgated bound is equivalent to the traditional one for maximum likelihood-decoding and it results in a bound that is the maximum of the traditional expurgated and random coding bounds.

IV. MAC: LOWER BOUNDS ON RELIABILITY FUNCTION

Consider a DM-MAC, W , with input alphabets \mathcal{X} and \mathcal{Y} , and output alphabet \mathcal{Z} . In this section, we present three achievable lower bounds on the reliability function (upper bound on the average error probability) for this channel. The method we are using is very similar to the point-to-point case. Again, the goal is first proving the existence of a good code and then analyzing its performance. The first step is choosing the ensemble. The ensemble, \mathcal{C} , we are using is similar to the ensemble in [20]. For a fixed distribution, $P_U P_{X|U} P_{Y|U}$, the codewords of each code in the ensemble are chosen from $T_{P_{X|U}}(\mathbf{u})$ and $T_{P_{Y|U}}(\mathbf{u})$ for some sequence $\mathbf{u} \in T_{P_U}$. Intuitively, we expect that the codewords in a “good” code must be far from each other. In accordance with the ideas of Csiszar and Korner [12], we use conditional types to quantify this statement. We select a prescribed number of sequences in \mathcal{X}^n and \mathcal{Y}^n so that the shells around each pair have small intersections with the shells around other sequences. In general, two types of packing lemmas have been studied in the literature based on whether the shells are defined on the channel input space or channel output space. The packing lemma in [19] belongs to the first type, and the one in [20] belongs to the second type. All the inequalities in the first type depend only on the channel input sequences. However, in the second type, the lemma incorporates the channel output into the packing inequalities. In this work, we use the first type. In the following, we follow a four step procedure to arrive at the error exponent bounds. In step one, we define first-order and second-order packing functions. These functions are independent of the channel statistics. Next, in step two, for any constant composition code and any DM-MAC, we provide upper and lower bounds on the probability of decoding error in terms of these packing functions. In step three, by using a random coding argument on the constant composition code ensemble, we show the existence of codes whose packing functions satisfy certain conditions. Finally, in step four, by connecting the results in step two and three, we provide lower and upper bounds on the error exponents. Our results include a new tighter lower bound on the

error exponent for DM-MAC using a new partial expurgation method for multi-user codes. We also give a tight characterization of the typical performance of the constant composition code ensemble. Both the expurgated bound as well as the typical bound outperform the random coding bound of [20], which is derived as special case of our methodology.

A. Definition of Packing Functions

Let $C_X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{M_X}\}$ and $C_Y = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_{M_Y}\}$ be constant composition codebooks with $\mathbf{x}_i \in T_{P_{X|U}}(\mathbf{u})$ and $\mathbf{y}_j \in T_{P_{Y|U}}(\mathbf{u})$, for some $\mathbf{u} \in T_{P_U}$. In the following, for a two-user code $C = C_X \times C_Y$, we define the following quantities that we will use later in this section.

Definition 5. Fix a finite set \mathcal{U} , and a joint type $V_{U_{XY}\bar{X}\bar{Y}} \in \mathcal{P}_n(\mathcal{U} \times (\mathcal{X} \times \mathcal{Y})^2)$. For code C , the first-order packing functions are defined as follows:

$$N_U(C, V_{U_{XY}}) \triangleq \frac{1}{M_X M_Y} \sum_{i=1}^{M_X} \sum_{j=1}^{M_Y} 1_{T_{V_{U_{XY}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j)}, \quad (49a)$$

$$N_X(C, V_{U_{XY}\bar{X}}) \triangleq \frac{1}{M_X M_Y} \sum_{i=1}^{M_X} \sum_{j=1}^{M_Y} \sum_{k \neq i} 1_{T_{V_{U_{XY}\bar{X}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j, \mathbf{x}_k)}, \quad (49b)$$

$$N_Y(C, V_{U_{XY}\bar{Y}}) \triangleq \frac{1}{M_X M_Y} \sum_{i=1}^{M_X} \sum_{j=1}^{M_Y} \sum_{l \neq j} 1_{T_{V_{U_{XY}\bar{Y}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j, \mathbf{y}_l)}, \quad (49c)$$

$$N_{XY}(C, V_{U_{XY}\bar{X}\bar{Y}}) \triangleq \frac{1}{M_X M_Y} \sum_{i=1}^{M_X} \sum_{j=1}^{M_Y} \sum_{k \neq i} \sum_{l \neq j} 1_{T_{V_{U_{XY}\bar{X}\bar{Y}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j, \mathbf{x}_k, \mathbf{y}_l)}. \quad (49d)$$

Moreover, for any $V_{U_{XY}\bar{X}\bar{Y}\bar{X}'\bar{Y}'\bar{X}''\bar{Y}''} \in \mathcal{P}_n(\mathcal{U} \times (\mathcal{X} \times \mathcal{Y})^3)$, we define a set of second-order packing functions as follows:

$$\Lambda_X(C, V_{U_{XY}\bar{X}\bar{X}'}) \triangleq \frac{1}{M_X M_Y} \sum_{i,j} \sum_{k \neq i} \sum_{k' \neq i,k} 1_{T_{V_{U_{XY}\bar{X}\bar{X}'}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j, \mathbf{x}_k, \mathbf{x}_{k'})}, \quad (50a)$$

$$\Lambda_Y(C, V_{U_{XY}\bar{Y}\bar{Y}'}) \triangleq \frac{1}{M_X M_Y} \sum_{i,j} \sum_{l \neq j} \sum_{l' \neq j,l} 1_{T_{V_{U_{XY}\bar{Y}\bar{Y}'}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j, \mathbf{y}_l, \mathbf{y}_{l'})}, \quad (50b)$$

$$\Lambda_{XY}(C, V_{U_{XY}\bar{X}\bar{Y}\bar{X}'\bar{Y}'}) \triangleq \frac{1}{M_X M_Y} \sum_{i,j} \sum_{k \neq i} \sum_{\substack{k' \neq i,k \\ l \neq j \\ l' \neq j,l}} 1_{T_{V_{U_{XY}\bar{X}\bar{Y}\bar{X}'\bar{Y}'}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j, \mathbf{x}_k, \mathbf{y}_l, \mathbf{x}_{k'}, \mathbf{y}_{l'})}. \quad (50c)$$

The second-order packing functions are used to prove the tightness of the results of Theorem 2 and Theorem 3. Next we will obtain upper and lower bounds on the probability of decoding error for an arbitrary two-user code that depend on its packing functions defined above.

B. Relation between probability of error and packing functions

Consider the multiuser code C as defined above, and a function $\alpha : \mathcal{P}(\mathcal{U} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}) \rightarrow \mathbb{R}$. Taking into account the given \mathbf{u} , α -decoding yields the decoding sets

$$D_{ij} = \{\mathbf{z} : \alpha(P_{\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j, \mathbf{z}}) \leq \alpha(P_{\mathbf{u}, \mathbf{x}_k, \mathbf{y}_l, \mathbf{z}}) \text{ for all } (k, l) \neq (i, j)\}. \quad (51)$$

The average error probability of this multiuser code on DM-MAC W , can be written as

$$\begin{aligned}
e(C, W) &\triangleq \frac{1}{M_X M_Y} \sum_{i,j} W^n(D_{ij}^c | \mathbf{x}_i, \mathbf{y}_j) \\
&= \frac{1}{M_X M_Y} \sum_{i,j} W^n\left(\bigcup_{k \neq i} D_{kj} | \mathbf{x}_i, \mathbf{y}_j\right) + \frac{1}{M_X M_Y} \sum_{i,j} W^n\left(\bigcup_{l \neq j} D_{il} | \mathbf{x}_i, \mathbf{y}_j\right) + \frac{1}{M_X M_Y} \sum_{i,j} W^n\left(\bigcup_{\substack{k \neq i \\ l \neq j}} D_{kl} | \mathbf{x}_i, \mathbf{y}_j\right).
\end{aligned} \tag{52}$$

The first term on the right side of (52) can be written as

$$\begin{aligned}
&\frac{1}{M_X M_Y} \sum_{i,j} W^n\left(\bigcup_{k \neq i} D_{kj} | \mathbf{x}_i, \mathbf{y}_j\right) \\
&= \frac{1}{M_X M_Y} \sum_{i,j} W^n\left(\left\{\mathbf{z} : \alpha(P_{\mathbf{u}, \mathbf{x}_k, \mathbf{y}_j, \mathbf{z}}) \leq \alpha(P_{\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j, \mathbf{z}}), \text{ for some } k \neq i\right\} | \mathbf{u}, \mathbf{x}_i, \mathbf{y}_j\right) \\
&= \frac{1}{M_X M_Y} \sum_{i,j} \sum_{\substack{\mathbf{z}: \\ \alpha(P_{\mathbf{u}, \mathbf{x}_k, \mathbf{y}_j, \mathbf{z}}) \leq \alpha(P_{\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j, \mathbf{z}}) \\ \text{for some } k \neq i}} W^n(\mathbf{z} | \mathbf{u}, \mathbf{x}_i, \mathbf{y}_j) \\
&= \frac{1}{M_X M_Y} \sum_{i,j} \sum_{V_{UXY\bar{X}Z} \in \mathcal{V}_{X,n}^r} \sum_{\substack{\mathbf{z}: \\ \alpha(P_{\mathbf{u}, \mathbf{x}_k, \mathbf{y}_j, \mathbf{z}}) \leq \alpha(P_{\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j, \mathbf{z}}) \\ \text{for some } k \neq i}} 1_{T_{V_{UXY\bar{X}Z}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j, \mathbf{x}_k, \mathbf{z}) W^n(\mathbf{z} | \mathbf{u}, \mathbf{x}_i, \mathbf{y}_j) \\
&= \sum_{V_{UXY\bar{X}Z} \in \mathcal{V}_{X,n}^r} 2^{-n[D(V_Z | XYU) + H(V_Z | XYU)]} \cdot \left[\frac{1}{M_X M_Y} \sum_{i,j} 1_{T_{V_{UXY\bar{X}Z}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j) \cdot A_{i,j}^X(V_{UXY\bar{X}Z}, C) \right],
\end{aligned} \tag{53}$$

where

$$\begin{aligned}
A_{i,j}^X(V_{UXY\bar{X}Z}, C) &\triangleq |\{\mathbf{z} : (\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j, \mathbf{x}_k, \mathbf{z}) \in T_{V_{UXY\bar{X}Z}} \text{ for some } k \neq i\}| \\
\mathcal{V}_{X,n}^r &\triangleq \{V_{UXY\bar{X}Z} : \alpha(V_{UXYZ}) \geq \alpha(V_{U\bar{X}YZ}), V_{UX} = V_{U\bar{X}} = P_{UX}, V_{UY} = P_{UY}\}.
\end{aligned} \tag{54}$$

Note that $\mathcal{V}_{X,n}^r$ is a set of types of resolution n , therefore, we use a subscript n to define it. Similarly, the second and third term term on the right side of (52) can be written as follows:

$$\begin{aligned}
&\frac{1}{M_X M_Y} \sum_{i,j} W^n\left(\bigcup_{l \neq j} D_{il} | \mathbf{x}_i, \mathbf{y}_j\right) \\
&= \sum_{V_{UXY\bar{Y}Z} \in \mathcal{V}_{Y,n}^r} 2^{-n[D(V_Z | XYU) + H(V_Z | XYU)]} \cdot \left[\frac{1}{M_X M_Y} \sum_{i,j} 1_{T_{V_{UXY\bar{Y}Z}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j) \cdot A_{i,j}^Y(V_{UXY\bar{Y}Z}, C) \right],
\end{aligned} \tag{55}$$

where

$$\begin{aligned}
A_{i,j}^Y(V_{UXY\bar{Y}Z}, C) &\triangleq |\{\mathbf{z} : (\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j, \mathbf{y}_l, \mathbf{z}) \in T_{V_{UXY\bar{Y}Z}} \text{ for some } l \neq j\}| \\
\mathcal{V}_{Y,n}^r &\triangleq \{V_{UXY\bar{Y}Z} : \alpha(V_{UXYZ}) \geq \alpha(V_{UX\bar{Y}Z}), V_{UX} = P_{UX}, V_{UY} = V_{U\bar{Y}} = P_{UY}\},
\end{aligned} \tag{56}$$

and,

$$\begin{aligned} & \frac{1}{M_X M_Y} \sum_{i,j} W^n \left(\bigcup_{\substack{k \neq i \\ l \neq j}} D_{kl} | \mathbf{x}_i, \mathbf{y}_j \right) \\ &= \sum_{V_{UXY\tilde{X}\tilde{Y}Z} \in \mathcal{V}_{XY,n}^r} 2^{-n[D(V_{Z|XYU}) + W(V_{XYU}) + H_V(Z|XYU)]} \cdot \left[\frac{1}{M_X M_Y} \sum_{i,j} 1_{T_{V_{UXY}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j) \cdot A_{i,j}^{XY}(V_{UXY\tilde{X}\tilde{Y}Z}, C) \right], \end{aligned} \quad (57)$$

where

$$\begin{aligned} A_{i,j}^{XY}(V_{UXY\tilde{X}\tilde{Y}Z}, C) &\triangleq |\{\mathbf{z} : (\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j, \mathbf{x}_k, \mathbf{y}_l, \mathbf{z}) \in T_{V_{UXY\tilde{X}\tilde{Y}Z}} \text{ for some } k \neq i, l \neq j\}| \\ \mathcal{V}_{XY,n}^r &\triangleq \{V_{UXY\tilde{X}\tilde{Y}Z} : \alpha(V_{UXYZ}) \geq \alpha(V_{U\tilde{X}\tilde{Y}Z}), V_{UX} = V_{U\tilde{X}} = P_{UX}, V_{UY} = V_{U\tilde{Y}} = P_{UY}\}. \end{aligned} \quad (58)$$

Clearly, $A_{i,j}^X(V_{UXY\tilde{X}Z})$ satisfies

$$B_{i,j}^X(V_{UXY\tilde{X}Z}, C) - C_{i,j}^X(V_{UXY\tilde{X}Z}, C) \leq A_{i,j}^X(V_{UXY\tilde{X}Z}, C) \leq B_{i,j}^X(V_{UXY\tilde{X}Z}, C), \quad (59)$$

where

$$B_{i,j}^X(V_{UXY\tilde{X}Z}, C) \triangleq \sum_{k \neq i} 1_{T_{V_{UXY\tilde{X}}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j, \mathbf{x}_k) \cdot |\{\mathbf{z} : \mathbf{z} \in T_{V_{Z|UXY\tilde{X}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j, \mathbf{x}_k)\}|, \quad (60)$$

$$\begin{aligned} C_{i,j}^X(V_{UXY\tilde{X}Z}, C) &\triangleq \sum_{k \neq i} \sum_{k' \neq k, i} 1_{T_{V_{UXY\tilde{X}}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j, \mathbf{x}_k) 1_{T_{V_{UXY\tilde{X}}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j, \mathbf{x}_{k'}) \\ &\cdot |\{\mathbf{z} : \mathbf{z} \in T_{V_{Z|UXY\tilde{X}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j, \mathbf{x}_k) \cap T_{V_{Z|UXY\tilde{X}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j, \mathbf{x}_{k'})\}|. \end{aligned} \quad (61)$$

Having related the probability of error and the function $B_{i,j}^X$, $B_{i,j}^Y$ and $B_{i,j}^{XY}$, our next task is to provide a simple upper bound on these functions. This is done as follows.

$$\begin{aligned} & \frac{1}{M_X M_Y} \sum_{i,j} 1_{T_{V_{UXY}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j) B_{i,j}^X(V_{UXY\tilde{X}Z}, C) \\ &= \frac{1}{M_X M_Y} \sum_{i,j} \sum_{k \neq i} 1_{T_{V_{UXY\tilde{X}}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j, \mathbf{x}_k) \left| \left\{ \mathbf{z} : \mathbf{z} \in T_{V_{Z|UXY\tilde{X}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j, \mathbf{x}_k) \right\} \right| \\ &\leq 2^{nH(Z|UXY\tilde{X})} \frac{1}{M_X M_Y} \sum_{i,j} \sum_{k \neq i} 1_{T_{V_{UXY\tilde{X}}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j, \mathbf{x}_k) \\ &= 2^{nH(Z|UXY\tilde{X})} N_X(C, V_{UXY\tilde{X}}) \end{aligned} \quad (62)$$

Similarly, we can provide upper bounds for $B_{i,j}^Y$ and $B_{i,j}^{XY}$. Moreover, we can also provide trivial upper bounds on $A(\cdot)$ functions as was done in the point-to-point case.

$$A_{i,j}^X(V_{UXY\tilde{X}Z}, C) \leq 2^{nH_V(Z|XYU)}.$$

The same bound applies to A^Y and A^{XY} . Collecting all these results, we provide the following upper bound on the probability of error.

$$\begin{aligned}
e(C, W) &\leq \sum_{\substack{V_{UXY\tilde{X}Z} \\ \in \mathcal{V}_{X,n}^r}} 2^{-n[D(V_Z|XYU)||W|V_{XYU}]} \min \left\{ 2^{-nI_V(\tilde{X}\wedge Z|XYU)} N_X(C, V_{UXY\tilde{X}}, 1) \right\} \\
&+ \sum_{\substack{V_{UXY\tilde{Y}Z} \\ \in \mathcal{V}_{Y,n}^r}} 2^{-n[D(V_Z|XYU)||W|V_{XYU}]} \min \left\{ 2^{-nI_V(\tilde{Y}\wedge Z|XYU)} N_Y(C, V_{UXY\tilde{Y}}, 1) \right\} \\
&+ \sum_{\substack{V_{UXY\tilde{X}\tilde{Y}Z} \\ \in \mathcal{V}_{XY,n}^r}} 2^{-n[D(V_Z|XYU)||W|V_{XYU}]} \min \left\{ 2^{-nI_V(\tilde{X}\tilde{Y}\wedge Z|XYU)} N_{XY}(C, V_{UXY\tilde{X}\tilde{Y}}, 1) \right\} \quad (63)
\end{aligned}$$

Next, we consider lower bounds on $B(\cdot)$ functions and upper bounds on $C(\cdot)$ functions. One can use a similar argument to show the following

$$\frac{1}{M_X M_Y} \sum_{i,j} 1_{T_{V_{UXY}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j) B_{i,j}^X(V_{UXY\tilde{X}Z}, C) \geq 2^{n[H(Z|UXY\tilde{X})-\delta]} N_X(C, V_{UXY\tilde{X}}).$$

Similar lower bounds can be obtained for B^Y and B^{XY} . Moreover, we have the following arguments for bounding from above the function C^X .

$$\begin{aligned}
&\frac{1}{M_X M_Y} \sum_{i,j} 1_{T_{V_{UXY}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j) \cdot C_{i,j}^X(V_{UXY\tilde{X}Z}) \\
&= \frac{1}{M_X M_Y} \sum_{i,j} 1_{T_{V_{UXY}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j) \sum_{k \neq i} \sum_{k' \neq k, i} 1_{T_{V_{UXY\tilde{X}}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j, \mathbf{x}_k) 1_{T_{V_{UXY\tilde{X}}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j, \mathbf{x}_{k'}) \\
&\quad \cdot \left| \left\{ \mathbf{z} : \mathbf{z} \in T_{V_{Z|UXY\tilde{X}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j, \mathbf{x}_k) \cap T_{V_{Z|UXY\tilde{X}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j, \mathbf{x}_{k'}) \right\} \right| \\
&= \frac{1}{M_X M_Y} \sum_{i,j} \sum_{\substack{V_{UXY\tilde{X}\tilde{X}Z}: \\ V_{UXY\tilde{X}Z}=V_{UXY\tilde{X}\tilde{X}Z}}} \sum_{k \neq i} \sum_{k' \neq k, i} 1_{T_{V_{UXY\tilde{X}\tilde{X}}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j, \mathbf{x}_k, \mathbf{x}_{k'}) \left| \left\{ \mathbf{z} : \mathbf{z} \in T_{V_{Z|UXY\tilde{X}\tilde{X}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j, \mathbf{x}_k, \mathbf{x}_{k'}) \right\} \right| \\
&\leq \sum_{\substack{V_{UXY\tilde{X}\tilde{X}Z}: \\ V_{UXY\tilde{X}Z}=V_{UXY\tilde{X}\tilde{X}Z}}} 2^{nH(Z|UXY\tilde{X}\tilde{X})} \frac{1}{M_X M_Y} \sum_{i,j} \sum_{k \neq i} \sum_{k' \neq k, i} 1_{T_{V_{UXY\tilde{X}\tilde{X}}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j, \mathbf{x}_k, \mathbf{x}_{k'}) \\
&= \sum_{\substack{V_{UXY\tilde{X}\tilde{X}Z}: \\ V_{UXY\tilde{X}Z}=V_{UXY\tilde{X}\tilde{X}Z}}} 2^{nH(Z|UXY\tilde{X}\tilde{X})} \Lambda_X(C, V_{UXY\tilde{X}\tilde{X}}). \quad (64)
\end{aligned}$$

Similar relation can be obtained that relate C^Y and λ_Y , C^{XY} and λ_{XY} . Combining the lower bounds on $B(\cdot)$ -functions and upper bounds on $C(\cdot)$ -functions, we have the following lower bound on the probability of decoding error.

$e(C, W)$

$$\begin{aligned}
&\geq \sum_{\substack{V_{UXY\tilde{X}Z} \\ \in \mathcal{V}_{\tilde{X},n}^*}} 2^{-n[D(V_Z|_{XYU})\|W|V_{XYU})+I_V(\tilde{X}\wedge Z|XYU)+\delta]} \left| N_X - \sum_{\substack{V_{UXY\tilde{X}\tilde{X}Z}: \\ V_{UXY\tilde{X}Z}=V_{UXY\tilde{X}Z}}} 2^{nI(\tilde{X}\wedge Z|UXY\tilde{X})} \Lambda_X \right|^+ \\
&+ \sum_{\substack{V_{UXY\tilde{Y}Z} \\ \in \mathcal{V}_{\tilde{Y},n}^*}} 2^{-n[D(V_Z|_{XYU})\|W|V_{XYU})+I_V(\tilde{Y}\wedge Z|XYU)+\delta]} \left| N_Y - \sum_{\substack{V_{UXY\tilde{Y}\tilde{Y}Z}: \\ V_{UXY\tilde{Y}Z}=V_{UXY\tilde{Y}Z}}} 2^{nI(\tilde{Y}\wedge Z|UXY\tilde{Y})} \Lambda_Y \right|^+ \\
&+ \sum_{\substack{V_{UXY\tilde{X}\tilde{Y}Z} \\ \in \mathcal{V}_{\tilde{X}\tilde{Y},n}^*}} 2^{-n[D(V_Z|_{XYU})\|W|V_{XYU})+I_V(\tilde{X}\tilde{Y}\wedge Z|XYU)+\delta]} \left| N_{XY} - \sum_{\substack{V_{UXY\tilde{X}\tilde{Y}\tilde{Y}Z}: \\ V_{UXY\tilde{X}\tilde{Y}Z}=V_{UXY\tilde{X}\tilde{Y}Z}}} 2^{nI(\tilde{X}\tilde{Y}\wedge Z|UXY\tilde{X}\tilde{Y})} \Lambda_{XY} \right|^+. \tag{65}
\end{aligned}$$

This completes our task of relating the average probability of error of any code C in terms of the first and the second order packing functions. We next proceed toward obtaining lower bounds on the error exponents. The expressions for the error exponents that we derive are conceptually very similar to those derived for the point-to-point channels. However, since we have to deal with a bigger class of error events, the expressions for the error exponents become longer. To state our results concisely, in the next subsection, we define certain functions of information quantities and transmission rates. We will express our results in terms of these functions. The reader can skip this subsection, and move to the next subsection without losing the flow of the exposition. The reader can come back to it when we refer to it in the subsequent discussions.

C. Definition of Information Functions

In the following, we consider five definitions which are mainly used for conciseness.

Definition 6. For any fix rate pair $R_X, R_Y \geq 0$, and any distribution $V_{UXY\tilde{X}\tilde{Y}} \in \mathcal{P}(\mathcal{U} \times (\mathcal{X} \times \mathcal{Y})^2)$, we define

$$F_U(V_{UXY}) \triangleq I(X \wedge Y|U), \tag{66a}$$

$$F_X(V_{UXY\tilde{X}}) \triangleq I(X \wedge Y|U) + I_V(\tilde{X} \wedge XY|U) - R_X, \tag{66b}$$

$$F_Y(V_{UXY\tilde{Y}}) \triangleq I(X \wedge Y|U) + I(\tilde{Y} \wedge XY|U) - R_Y, \tag{66c}$$

$$F_{XY}(V_{UXY\tilde{X}\tilde{Y}}) \triangleq I(X \wedge Y|U) + I(\tilde{X} \wedge \tilde{Y}|U) + I(\tilde{X}\tilde{Y} \wedge XY|U) - R_X - R_Y. \tag{66d}$$

Moreover, for any $V_{UXY\tilde{X}\tilde{Y}\hat{X}\hat{Y}} \in \mathcal{P}(\mathcal{U} \times (\mathcal{X} \times \mathcal{Y})^3)$, we define

$$E_S^{\tilde{X}}(V_{UXY\tilde{X}\hat{X}}) \triangleq I(\hat{X} \wedge XY\tilde{X}|U) + I(\tilde{X} \wedge XY|U) + I(X \wedge Y|U) - 2R_X, \tag{67a}$$

$$E_S^{\tilde{Y}}(V_{UXY\tilde{Y}\hat{Y}}) \triangleq I(\hat{Y} \wedge XY\tilde{Y}|U) + I(\tilde{Y} \wedge XY|U) + I(X \wedge Y|U) - 2R_Y, \tag{67b}$$

$$E_S^{\tilde{X}\tilde{Y}}(V_{UXY\tilde{X}\tilde{Y}\hat{X}\hat{Y}}) \triangleq I(\hat{X}\hat{Y} \wedge XY\tilde{X}\tilde{Y}|U) + I(\tilde{X}\tilde{Y} \wedge XY|U) + I(X \wedge Y|U) + I(\tilde{X} \wedge \tilde{Y}|U) + I(\hat{X} \wedge \hat{Y}|U) - 2R_X - 2R_Y. \tag{67c}$$

Definition 7. For any given $R_X, R_Y \geq 0$, $P_{XYU} \in \mathcal{P}(\mathcal{X} \times \mathcal{Y} \times \mathcal{U})$, we define the sets of distributions \mathcal{V}_X^r , \mathcal{V}_Y^r and \mathcal{V}_{XY}^r as follows:

$$\mathcal{V}_X^r \triangleq \{V_{UXY\tilde{X}Z} : \alpha(V_{UXYZ}) \geq \alpha(V_{U\tilde{X}YZ}), V_{UX} = V_{U\tilde{X}} = P_{UX}, V_{UY} = P_{UY}\}, \quad (68a)$$

$$\mathcal{V}_Y^r \triangleq \{V_{UXY\tilde{Y}Z} : \alpha(V_{UXYZ}) \geq \alpha(V_{UX\tilde{Y}Z}), V_{UX} = P_{UX}, V_{UY} = V_{U\tilde{Y}} = P_{UY}\}, \quad (68b)$$

$$\mathcal{V}_{XY}^r \triangleq \{V_{UXY\tilde{X}\tilde{Y}Z} : \alpha(V_{UXYZ}) \geq \alpha(V_{U\tilde{X}\tilde{Y}Z}), V_{UX} = V_{U\tilde{X}} = P_{UX}, V_{UY} = V_{U\tilde{Y}} = P_{UY}\}. \quad (68c)$$

Moreover, $\mathcal{V}_X^{r,L}$, $\mathcal{V}_Y^{r,L}$ and $\mathcal{V}_{XY}^{r,L}$ are sets of distributions and defined as

$$\mathcal{V}_X^{r,L} \triangleq \{V_{UXY\tilde{X}Z} \in \mathcal{V}_X^r : I(\tilde{X} \wedge XYZ|U) \geq R_X\}, \quad (69a)$$

$$\mathcal{V}_Y^{r,L} \triangleq \{V_{UXY\tilde{Y}Z} \in \mathcal{V}_Y^r : I(\tilde{Y} \wedge XYZ|U) \geq R_Y\}, \quad (69b)$$

$$\mathcal{V}_{XY}^{r,L} \triangleq \{V_{UXY\tilde{X}\tilde{Y}Z} \in \mathcal{V}_{XY}^r : I(\tilde{X}\tilde{Y} \wedge XYZ|U) + I(\tilde{X} \wedge \tilde{Y}) \geq R_X + R_Y\}. \quad (69c)$$

Definition 8. For any given $R_X, R_Y \geq 0$, $P_{XYU} \in \mathcal{P}(\mathcal{X} \times \mathcal{Y} \times \mathcal{U})$, we define the sets of distributions \mathcal{V}_X^T , \mathcal{V}_Y^T and \mathcal{V}_{XY}^T as follows

$$\mathcal{V}_X^T \triangleq \left\{ \begin{array}{l} V_{UXY\tilde{X}} : \quad V_{XU} = V_{\tilde{X}U} = P_{XU}, V_{YU} = P_{YU} \\ \quad \quad \quad F_U(V_{UXY}), F_U(V_{U\tilde{X}Y}) \leq R_X + R_Y \\ \quad \quad \quad F_X(V_{UXY\tilde{X}}) \leq R_X + R_Y \\ \quad \quad \quad \alpha(V_{UXYZ}) \geq \alpha(V_{U\tilde{X}YZ}) \end{array} \right\} \quad (70a)$$

$$\mathcal{V}_Y^T \triangleq \left\{ \begin{array}{l} V_{UXY\tilde{Y}} : \quad V_{XU} = P_{XU}, V_{YU} = V_{\tilde{Y}U} = P_{YU} \\ \quad \quad \quad F_U(V_{UXY}), F_U(V_{UX\tilde{Y}}) \leq R_X + R_Y \\ \quad \quad \quad F_Y(V_{UXY\tilde{Y}}) \leq R_X + R_Y \\ \quad \quad \quad \alpha(V_{UXYZ}) \geq \alpha(V_{UX\tilde{Y}Z}) \end{array} \right\} \quad (70b)$$

$$\mathcal{V}_{XY}^T \triangleq \left\{ \begin{array}{l} V_{UXY\tilde{X}\tilde{Y}} : \quad V_{XU} = V_{\tilde{X}U} = P_{XU}, V_{YU} = V_{\tilde{Y}U} = P_{YU} \\ \quad \quad \quad F_U(V_{UXY}), F_U(V_{U\tilde{X}Y}), F_U(V_{UX\tilde{Y}}), F_U(V_{U\tilde{X}\tilde{Y}}) \leq R_X + R_Y \\ \quad \quad \quad F_X(V_{UXY\tilde{X}}), F_X(V_{UX\tilde{Y}\tilde{X}}) \leq R_X + R_Y \\ \quad \quad \quad F_Y(V_{UXY\tilde{Y}}), F_Y(V_{U\tilde{X}Y\tilde{Y}}) \leq R_X + R_Y \\ \quad \quad \quad F_{XY}(V_{UXY\tilde{X}\tilde{Y}}), F_{XY}(V_{U\tilde{X}Y\tilde{X}\tilde{Y}}) \leq R_X + R_Y \\ \quad \quad \quad \alpha(V_{UXYZ}) \geq \alpha(V_{U\tilde{X}\tilde{Y}Z}) \end{array} \right\} \quad (70c)$$

Moreover, $\mathcal{V}_X^{T,L}$, $\mathcal{V}_Y^{T,L}$ and $\mathcal{V}_{XY}^{T,L}$ are sets of distributions and defined as

$$\mathcal{V}_X^{T,L} \triangleq \{V_{UXY\tilde{X}Z} \in \mathcal{V}_X^T : I(\tilde{X} \wedge XYZ|U) \geq R_X\}, \quad (71a)$$

$$\mathcal{V}_Y^{T,L} \triangleq \{V_{UXY\tilde{Y}Z} \in \mathcal{V}_Y^T : I(\tilde{Y} \wedge XYZ|U) \geq R_Y\}, \quad (71b)$$

$$\mathcal{V}_{XY}^{T,L} \triangleq \{V_{UXY\tilde{X}\tilde{Y}Z} \in \mathcal{V}_{XY}^T : I(\tilde{X}\tilde{Y} \wedge XYZ|U) + I(\tilde{X} \wedge \tilde{Y}) \geq R_X + R_Y\}. \quad (71c)$$

Definition 9. For any given $R_X, R_Y \geq 0$, $P_{XYU} \in \mathcal{P}(\mathcal{X} \times \mathcal{Y} \times \mathcal{U})$, we define the sets of distributions \mathcal{V}_X^{ex} , \mathcal{V}_Y^{ex} ,

and \mathcal{V}_{XY}^{ex} as follows

$$\mathcal{V}_X^{ex} \triangleq \left\{ \begin{array}{l} V_{UXY\tilde{X}} : \quad V_{XU} = V_{\tilde{X}U} = P_{XU}, V_{YU} = P_{YU} \\ F_U(V_{UXY}), F_U(V_{U\tilde{X}Y}) \leq \min\{R_X, R_Y\} \\ F_X(V_{UXY\tilde{X}}) \leq \min\{R_X, R_Y\} \\ \alpha(V_{UXYZ}) \geq \alpha(V_{U\tilde{X}YZ}) \end{array} \right\} \quad (72a)$$

$$\mathcal{V}_Y^{ex} \triangleq \left\{ \begin{array}{l} V_{UXY\tilde{Y}} : \quad V_{XU} = P_{XU}, V_{YU} = V_{\tilde{Y}U} = P_{YU} \\ F_U(V_{UXY}), F_U(V_{U\tilde{Y}Y}) \leq \min\{R_X, R_Y\} \\ F_Y(V_{UXY\tilde{Y}}) \leq \min\{R_X, R_Y\} \\ \alpha(V_{UXYZ}) \geq \alpha(V_{U\tilde{Y}YZ}) \end{array} \right\} \quad (72b)$$

$$\mathcal{V}_{XY}^{ex} \triangleq \left\{ \begin{array}{l} V_{UXY\tilde{X}\tilde{Y}} : \quad V_{XU} = V_{\tilde{X}U} = P_{XU}, V_{YU} = V_{\tilde{Y}U} = P_{YU} \\ F_U(V_{UXY}), F_U(V_{U\tilde{X}Y}), F_U(V_{U\tilde{X}\tilde{Y}}), F_U(V_{U\tilde{X}\tilde{Y}}) \leq \min\{R_X, R_Y\} \\ F_X(V_{UXY\tilde{X}}), F_X(V_{U\tilde{X}\tilde{Y}\tilde{X}}) \leq \min\{R_X, R_Y\} \\ F_Y(V_{UXY\tilde{Y}}), F_Y(V_{U\tilde{X}\tilde{Y}\tilde{Y}}) \leq \min\{R_X, R_Y\} \\ F_{XY}(V_{UXY\tilde{X}\tilde{Y}}), F_{XY}(V_{U\tilde{X}\tilde{Y}\tilde{X}\tilde{Y}}) \leq \min\{R_X, R_Y\} \\ \alpha(V_{UXYZ}) \geq \alpha(V_{U\tilde{X}\tilde{Y}Z}) \end{array} \right\} \quad (72c)$$

Definition 10. For any given $R_X, R_Y \geq 0$, $P_{XYU} \in \mathcal{P}(\mathcal{X} \times \mathcal{Y} \times \mathcal{U})$, and $V_{UXY\tilde{X}\tilde{Y}} \in \mathcal{P}(\mathcal{U} \times (\mathcal{X} \times \mathcal{Y})^2)$, we define the following quantities

$$E_X(R_X, R_Y, W, P_{XYU}, V_{UXY\tilde{X}}) \triangleq D(V_{Z|XYU}||W|V_{XYU}) + I_V(X \wedge Y|U) + |I(\tilde{X} \wedge XYZ|U) - R_X|^+, \quad (73a)$$

$$E_Y(R_X, R_Y, W, P_{XYU}, V_{UXY\tilde{Y}}) \triangleq D(V_{Z|XYU}||W|V_{XYU}) + I_V(X \wedge Y|U) + |I(\tilde{Y} \wedge XYZ|U) - R_Y|^+, \quad (73b)$$

$$E_{XY}(R_X, R_Y, W, P_{XYU}, V_{UXY\tilde{X}\tilde{Y}}) \triangleq D(V_{Z|XYU}||W|V_{XYU}) + I_V(X \wedge Y|U) + |I(\tilde{X}\tilde{Y} \wedge XYZ|U) + I_V(\tilde{X} \wedge \tilde{Y}|U) - R_X - R_Y|^+. \quad (73c)$$

Moreover, we define

$$E_X^L(R_X, R_Y, W, P_{XYU}, V_{UXY\tilde{X}}) \triangleq D(V_{Z|XYU}||W|V_{XYU}) + I_V(X \wedge Y|U) + I(\tilde{X} \wedge XYZ|U) - R_X, \quad (74a)$$

$$E_Y^L(R_X, R_Y, W, P_{XYU}, V_{UXY\tilde{Y}}) \triangleq D(V_{Z|XYU}||W|V_{XYU}) + I_V(X \wedge Y|U) + I(\tilde{Y} \wedge XYZ|U) - R_Y, \quad (74b)$$

$$E_{XY}^L(R_X, R_Y, W, P_{XYU}, V_{UXY\tilde{X}\tilde{Y}}) \triangleq D(V_{Z|XYU}||W|V_{XYU}) + I_V(X \wedge Y|U) + I(\tilde{X}\tilde{Y} \wedge XYZ|U) + I_V(\tilde{X} \wedge \tilde{Y}|U) - R_X - R_Y, \quad (74c)$$

and,

$$E_\beta^\alpha(R_X, R_Y, W, P_{XYU}, \mathcal{V}_\beta^\alpha) \triangleq \min_{V_{UXY\tilde{\beta}Z} \in \mathcal{V}_\beta^\alpha} E_\beta(R_X, R_Y, W, P_{XYU}, V_{UXY\tilde{\beta}}), \quad (75a)$$

$$E_\beta^{\alpha,L}(R_X, R_Y, W, P_{XYU}, \mathcal{V}_\beta^\alpha) \triangleq \min_{V_{UXY\tilde{\beta}Z} \in \mathcal{V}_\beta^{\alpha,L}} E_\beta^L(R_X, R_Y, W, P_{XYU}, V_{UXY\tilde{\beta}}), \quad (75b)$$

for $\alpha \in \{r, T, ex\}$, and $\beta \in \{X, Y, XY\}$.

D. Packing Lemmas

As we did in the point-to-point case, here we perform random coding and derive bounds on the packing functions. The results will be stated as three lemmas, one for the average and one for the typical performance of the ensemble, and finally one for the expurgated ensemble. These results will be used in conjunction with the relation between the packing functions and the probability of error established in Section IV-B to obtain the bounds on the error exponents.

Lemma 4. Fix a finite set \mathcal{U} , $P_{XYU} \in \mathcal{P}_n(\mathcal{X} \times \mathcal{Y} \times \mathcal{U})$ such that $X - U - Y$, $R_X \geq 0$, $R_Y \geq 0$, $\delta > 0$, $2^{n(R_X - \delta)} \leq M_X \leq 2^{nR_X}$, $2^{n(R_Y - \delta)} \leq M_Y \leq 2^{nR_Y}$, and $\mathbf{u} \in T_{P_U}$. Let $X^{M_X} \triangleq \{X_1, X_2, \dots, X_{M_X}\}$ and $Y^{M_Y} \triangleq \{Y_1, Y_2, \dots, Y_{M_Y}\}$ are independent, and X_i s and Y_j s are uniformly distributed over $T_{P_{X|U}}(\mathbf{u})$ and $T_{P_{Y|U}}(\mathbf{u})$ respectively. For every joint type $V_{UXY\bar{X}\bar{Y}} \in \mathcal{P}_n(\mathcal{U} \times (\mathcal{X} \times \mathcal{Y})^2)$, the expectation of the packing functions over the random code $X^{M_X} \times Y^{M_Y}$ are bounded by

$$2^{-n[F_U(V_{UXY}) + \delta]} \leq \mathbb{E} \left[N_U(X^{M_X} \times Y^{M_Y}, V_{UXY}) \right] \leq 2^{-n[F_U(V_{UXY}) - 2\delta]}, \quad (76a)$$

$$2^{-n[F_X(V_{UXY\bar{X}}) + 3\delta]} \leq \mathbb{E} \left[N_X(X^{M_X} \times Y^{M_Y}, V_{UXY\bar{X}}) \right] \leq 2^{-n[F_X(V_{UXY\bar{X}}) - 4\delta]}, \quad (76b)$$

$$2^{-n[F_Y(V_{UXY\bar{Y}}) + 3\delta]} \leq \mathbb{E} \left[N_Y(X^{M_X} \times Y^{M_Y}, V_{UXY\bar{Y}}) \right] \leq 2^{-n[F_Y(V_{UXY\bar{Y}}) - 4\delta]}, \quad (76c)$$

$$2^{-n[F_{XY}(V_{UXY\bar{X}\bar{Y}}) + 4\delta]} \leq \mathbb{E} \left[N_{XY}(X^{M_X} \times Y^{M_Y}, V_{UXY\bar{X}\bar{Y}}) \right] \leq 2^{-n[F_{XY}(V_{UXY\bar{X}\bar{Y}}) - 4\delta]}, \quad (76d)$$

whenever $n \geq n_0(|\mathcal{U}|, |\mathcal{X}|, |\mathcal{Y}|, \delta)$. Moreover, for any $V_{UXY\bar{X}\bar{Y}\hat{X}\hat{Y}} \in \mathcal{P}_n(\mathcal{U} \times (\mathcal{X} \times \mathcal{Y})^3)$

$$\mathbb{E} \left[\Lambda_X(X^{M_X} \times Y^{M_Y}, V_{UXY\bar{X}\hat{X}}) \right] \leq 2^{-n(E_S^X(V_{UXY\bar{X}\hat{X}}) - 4\delta)}, \quad (77a)$$

$$\mathbb{E} \left[\Lambda_Y(X^{M_X} \times Y^{M_Y}, V_{UXY\bar{Y}\hat{Y}}) \right] \leq 2^{-n(E_S^Y(V_{UXY\bar{Y}\hat{Y}}) - 4\delta)}, \quad (77b)$$

$$\mathbb{E} \left[\Lambda_{XY}(X^{M_X} \times Y^{M_Y}, V_{UXY\bar{X}\bar{Y}\hat{X}\hat{Y}}) \right] \leq 2^{-n(E_S^{XY}(V_{UXY\bar{X}\bar{Y}\hat{X}\hat{Y}}) - 6\delta)}, \quad (77c)$$

whenever $n \geq n_0(|\mathcal{U}|, |\mathcal{X}|, |\mathcal{Y}|, \delta)$.

Proof: The proof is provided in the Appendix. ■

Lemma 5. Fix a finite set \mathcal{U} , $P_{XYU} \in \mathcal{P}_n(\mathcal{X} \times \mathcal{Y} \times \mathcal{U})$ such that $X - U - Y$, $R_X \geq 0$, $R_Y \geq 0$, $\delta > 0$, $2^{n(R_X - \delta)} \leq M_X \leq 2^{nR_X}$, $2^{n(R_Y - \delta)} \leq M_Y \leq 2^{nR_Y}$, and $\mathbf{u} \in T_{P_U}$. Almost every multi-user code $C = C_X \times C_Y$, $C_X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{M_X}\} \subset T_{P_{X|U}}(\mathbf{u})$ and $C_Y = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_{M_Y}\} \subset T_{P_{Y|U}}(\mathbf{u})$, in the constant composition code ensemble, \mathcal{C} , satisfies the following inequalities:

$$2^{-n[F_U(V_{UXY}) + 3\delta]} \leq N_U(C, V_{UXY}) \leq 2^{-n[F_U(V_{UXY}) - 3\delta]}, \quad (78a)$$

$$2^{-n[F_X(V_{UXY\bar{X}}) + 5\delta]} \leq N_X(C, V_{UXY\bar{X}}) \leq 2^{-n[F_X(V_{UXY\bar{X}}) - 5\delta]}, \quad (78b)$$

$$2^{-n[F_Y(V_{UXY\bar{Y}}) + 5\delta]} \leq N_Y(C, V_{UXY\bar{Y}}) \leq 2^{-n[F_Y(V_{UXY\bar{Y}}) - 5\delta]}, \quad (78c)$$

$$2^{-n[F_{XY}(V_{UXY\bar{X}\bar{Y}}) + 5\delta]} \leq N_{XY}(C, V_{UXY\bar{X}\bar{Y}}) \leq 2^{-n[F_{XY}(V_{UXY\bar{X}\bar{Y}}) - 5\delta]}, \quad (78d)$$

for all $V_{UXY\tilde{X}\tilde{Y}} \in \mathcal{P}_n(\mathcal{U} \times (\mathcal{X} \times \mathcal{Y})^2)$, and

$$\Lambda_X(C, V_{UXY\tilde{X}\tilde{X}}) \leq 2^{-n(E_S^X(V_{UXY\tilde{X}\tilde{X}}) - 5\delta)}, \quad (79a)$$

$$\Lambda_Y(C, V_{UXY\tilde{Y}\tilde{Y}}) \leq 2^{-n(E_S^Y(V_{UXY\tilde{Y}\tilde{Y}}) - 5\delta)}, \quad (79b)$$

$$\Lambda_{XY}(C, V_{UXY\tilde{X}\tilde{Y}\tilde{X}\tilde{Y}}) \leq 2^{-n(E_S^{XY}(V_{UXY\tilde{X}\tilde{Y}}) - 7\delta)}. \quad (79c)$$

for all $V_{UXY\tilde{X}\tilde{Y}\hat{X}\hat{Y}} \in \mathcal{P}_n(\mathcal{U} \times (\mathcal{X} \times \mathcal{Y})^3)$, whenever $n \geq n_0(|\mathcal{U}|, |\mathcal{X}|, |\mathcal{Y}|, \delta)$.

Proof: The proof is provided in the Appendix. ■

Lemma 6. For every finite set \mathcal{U} , $P_{XYU} \in \mathcal{P}_n(\mathcal{X} \times \mathcal{Y} \times \mathcal{U})$ such that $X - U - Y$, $R_X \geq 0$, $R_Y \geq 0$, $\delta > 0$, and $\mathbf{u} \in T_{P_U}$, there exist a multi-user code $C^* = C_X^* \times C_Y^*$, $C_X^* = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{M_X^*}\} \subset T_{P_{X|U}}(\mathbf{u})$ and $C_Y^* = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_{M_Y^*}\} \subset T_{P_{Y|U}}(\mathbf{u})$ with $M_X^* \geq \frac{2^{n(R_X - \delta)}}{2}$, $M_Y^* \geq \frac{2^{n(R_Y - \delta)}}{2}$, such that for every joint type $V_{UXY\tilde{X}\tilde{Y}} \in \mathcal{P}_n(\mathcal{U} \times (\mathcal{X} \times \mathcal{Y})^2)$,

$$N_U(C^*, V_{UXY}) \leq 2^{-n[F_U(V_{UXY}) - 6\delta]} \quad (80a)$$

$$N_X(C^*, V_{UXY\tilde{X}}) \leq 2^{-n[F_X(V_{UXY\tilde{X}}) - 6\delta]} \quad (80b)$$

$$N_Y(C^*, V_{UXY\tilde{Y}}) \leq 2^{-n[F_Y(V_{UXY\tilde{Y}}) - 6\delta]} \quad (80c)$$

$$N_{XY}(C^*, V_{UXY\tilde{X}\tilde{Y}}) \leq 2^{-n[F_{XY}(V_{UXY\tilde{X}\tilde{Y}}) - 6\delta]} \quad (80d)$$

and for any $1 \leq i \leq M_X^*$, and any $1 \leq j \leq M_Y^*$,

$$1_{T_{V_{UXY}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j) \leq 2^{-n[F_U(V_{UXY}) - \min\{R_X, R_Y\} - 6\delta]} \quad (81a)$$

$$\sum_{k \neq i} 1_{T_{V_{UXY\tilde{X}}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j, \mathbf{x}_k) \leq 2^{-n[F_X(V_{UXY\tilde{X}}) - \min\{R_X, R_Y\} - 6\delta]} \quad (81b)$$

$$\sum_{l \neq j} 1_{T_{V_{UXY\tilde{Y}}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j, \mathbf{y}_l) \leq 2^{-n[F_Y(V_{UXY\tilde{Y}}) - \min\{R_X, R_Y\} - 6\delta]} \quad (81c)$$

$$\sum_{k \neq i} \sum_{l \neq j} 1_{T_{V_{UXY\tilde{X}\tilde{Y}}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j, \mathbf{x}_k, \mathbf{y}_l) \leq 2^{-n[F_{XY}(V_{UXY\tilde{X}\tilde{Y}}) - \min\{R_X, R_Y\} - 6\delta]}, \quad (81d)$$

whenever

$$n \geq n_0(|\mathcal{U}|, |\mathcal{X}|, |\mathcal{Y}|, \delta).$$

Proof: The proof is provided in the Appendix. ■

As it is shown in the Appendix, the above property is derived by the method of expurgation. Unlike the point-to-point case, expurgation in the MAC is not a trivial procedure. To see that, observe that expurgating bad pairs of codewords results in a code with correlated messages, which is hard to analyze. Instead, what we do is a sort of ‘‘partial’’ expurgation. Roughly speaking, we start with a code whose existence is proved in Lemma 4 and eliminate some of the bad codewords from the code with the larger rate (as opposed to codeword pairs). By doing that, all messages in the new code are independent, and such a code is easier to analyze.

E. Error exponent bounds

We can now proceed in a fashion that is similar to the point-to-point case and derive a series of exponential bounds based on Lemmas 4, 5, and 6. In the following, we present three lower bounds, the

random coding, the typical random coding, and the expurgated bounds. As in the case of point-to-point channels, here too, all the lower bounds are expressed in terms of the optimization of a single objective function under different constraint sets.

Theorem 2. Fix a finite set \mathcal{U} , $P_{XYU} \in \mathcal{P}_n(\mathcal{X} \times \mathcal{Y} \times \mathcal{U})$ such that $X - U - Y$, $R_X \geq 0$, $R_Y \geq 0$, $\delta > 0$, $2^{n(R_X - \delta)} \leq M_X \leq 2^{nR_X}$, $2^{n(R_Y - \delta)} \leq M_Y \leq 2^{nR_Y}$, and $\mathbf{u} \in T_{P_U}$. Consider the ensemble, \mathcal{C} , of multi-user codes consisting of all pair of codebooks (C_X, C_Y) , where $C_X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{M_X}\} \subset T_{P_{X|U}}(\mathbf{u})$ and $C_Y = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_{M_Y}\} \subset T_{P_{Y|U}}(\mathbf{u})$. The expectation of the average probability of error over \mathcal{C} is bounded by

$$2^{-n[E_{rL}(R_X, R_Y, W, P_{XYU}) + 8\delta]} \leq \bar{P}_e \leq 2^{-n[E_r(R_X, R_Y, W, P_{XYU}) - 6\delta]} \quad (82)$$

whenever $n \geq n_1(|\mathcal{Z}|, |\mathcal{X}|, |\mathcal{Y}|, |\mathcal{U}|, \delta)$, where

$$E_r(R_X, R_Y, W, P_{XYU}) \triangleq \min_{\beta=X, Y, XY} E_{\beta}^r(R_X, R_Y, W, P_{UXY}, \mathcal{V}_{\beta}^r), \quad (83)$$

$$E_{rL}(R_X, R_Y, W, P_{XYU}) \triangleq \min_{\beta=X, Y, XY} E_{\beta}^{r,L}(R_X, R_Y, W, P_{UXY}, \mathcal{V}_{\beta}^{r,L}). \quad (84)$$

Proof: The proof is provided in the Appendix. ■

Corollary 5. In the low rate regime,

$$E_{rL}(R_X, R_Y, W, P_{XYU}) = E_r(R_X, R_Y, W, P_{XYU}). \quad (85)$$

We call this rate region as the critical region for W .

Proof: The proof is similar to the proof of corollary 1 and is omitted. ■

Theorem 3. Fix a finite set \mathcal{U} , $P_{XYU} \in \mathcal{P}_n(\mathcal{X} \times \mathcal{Y} \times \mathcal{U})$ such that $X - U - Y$, $R_X \geq 0$, $R_Y \geq 0$, $\delta > 0$, $2^{n(R_X - \delta)} \leq M_X \leq 2^{nR_X}$, $2^{n(R_Y - \delta)} \leq M_Y \leq 2^{nR_Y}$, and $\mathbf{u} \in T_{P_U}$. The average probability of error for almost all multi-user codes $C = C_X \times C_Y$, $C_X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{M_X}\} \subset T_{P_{X|U}}(\mathbf{u})$ and $C_Y = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_{M_Y}\} \subset T_{P_{Y|U}}(\mathbf{u})$, in ensemble \mathcal{C} , satisfies the following inequalities

$$2^{-n[E_{TL}(R_X, R_Y, W, P_{XYU}) + 7\delta]} \leq e(C, W) \leq 2^{-n[E_T(R_X, R_Y, W, P_{XYU}) - 6\delta]} \quad (86)$$

whenever $n \geq n_1(|\mathcal{Z}|, |\mathcal{X}|, |\mathcal{Y}|, |\mathcal{U}|, \delta)$, where

$$E_T(R_X, R_Y, W, P_{XYU}) \triangleq \min_{\beta=X, Y, XY} E_{\beta}^T(R_X, R_Y, W, P_{UXY}, \mathcal{V}_{\beta}^T) \quad (87)$$

$$E_{TL}(R_X, R_Y, W, P_{XYU}) \triangleq \min_{\beta=X, Y, XY} E_{\beta}^{T,L}(R_X, R_Y, W, P_{UXY}, \mathcal{V}_{\beta}^{T,L}). \quad (88)$$

Proof: The proof is provided in the Appendix. ■

Corollary 6. For every finite set \mathcal{U} , $\mathcal{P}_{XYU} \in \mathcal{P}_n(\mathcal{X} \times \mathcal{Y} \times \mathcal{U})$ such that $X - U - Y$, $R_X \geq 0$, $R_Y \geq 0$,

$$E_T(R_X, R_Y, P_{XYU}, W) \leq E_{av}^T(R_X, R_Y) \leq E_{TL}(R_X, R_Y, P_{XYU}, W). \quad (89)$$

Proof: The proof is very similar to the proof of Corollary 2. ■

Corollary 7. In the low rate regime,

$$E_{TL}(R_X, R_Y, P_{XYU}, W) = E_T(R_X, R_Y, P_{XYU}, W). \quad (90)$$

Proof: The proof is similar to the proof of Corollary 1 and is omitted. ■

Theorem 4. For every finite set \mathcal{U} , $\mathcal{P}_{XYU} \in \mathcal{P}_n(\mathcal{X} \times \mathcal{Y} \times \mathcal{U})$ such that $X - U - Y$, $R_X \geq 0$, $R_Y \geq 0$, $\delta > 0$, and $\mathbf{u} \in T_{P_U}$, there exists a multi-user code

$$C = \{(\mathbf{x}_i, \mathbf{y}_j, D_{ij}) : i = 1, \dots, M_X^*, j = 1, \dots, M_Y^*\} \quad (91)$$

with $\mathbf{x}_i \in T_{P_{X|U}}(\mathbf{u})$, $\mathbf{y}_j \in T_{P_{Y|U}}(\mathbf{u})$ for all i and j , $M_X^* \geq \frac{2^{n(R_X - \delta)}}{2}$, and $M_Y^* \geq \frac{2^{n(R_Y - \delta)}}{2}$, such that for every MAC $W : \mathcal{X} \times \mathcal{Y} \rightarrow \mathcal{Z}$

$$e(C, W) \leq 2^{-n[E_{ex}(R_X, R_Y, W, P_{XYU}) - 5\delta]} \quad (92)$$

whenever $n \geq n_1(|\mathcal{Z}|, |\mathcal{X}|, |\mathcal{Y}|, |\mathcal{U}|, \delta)$, where

$$E_{ex}(R_X, R_Y, W, P_{XYU}) \triangleq \min_{\beta=X, Y, XY} E_{\beta}^{ex}(R_X, R_Y, W, P_{U_{XY}}, \mathcal{V}_{\beta}^{ex}). \quad (93)$$

Proof: The proof is provided in the Appendix. ■

This exponential error bound can be universally obtained for all MAC's with given input and output alphabets, since the choice of the codewords does not depend on the channel.

In the following, we show that the bound in Theorem 2 is at least as good as the best known random coding bound, found in [20]. For this purpose, let us use the minimum equivocation decoding rule.

Definition 11. Given \mathbf{u} , for a multiuser code

$$C = \{(\mathbf{x}_i, \mathbf{y}_j, D_{ij}) : i = 1, \dots, M_X, j = 1, \dots, M_Y\}$$

we say that the D_{ij} are minimum equivocation decoding sets for \mathbf{u} , if $\mathbf{z} \in D_{ij}$ implies

$$H(\mathbf{x}_i \mathbf{y}_j | \mathbf{z} \mathbf{u}) = \min_{k,l} H(\mathbf{x}_k \mathbf{y}_l | \mathbf{z} \mathbf{u}).$$

It can be easily observed that these sets are equivalent to α -decoding sets, where $\alpha(\mathbf{u}, \mathbf{x}, \mathbf{y}, \mathbf{z})$ is defined as

$$\alpha(V_{U_{XYZ}}) \triangleq H_V(XY|ZU). \quad (94)$$

Here, $V_{U_{XYZ}}$ is the joint empirical distribution of $(\mathbf{u}, \mathbf{x}, \mathbf{y}, \mathbf{z})$.

Theorem 5. For every finite set \mathcal{U} , $\mathcal{P}_{XYU} \in \mathcal{P}(\mathcal{X} \times \mathcal{Y} \times \mathcal{U})$, $R_X \geq 0$, $R_Y \geq 0$, and $W : \mathcal{X} \times \mathcal{Y} \rightarrow \mathcal{Z}$, and an appropriate α -decoder (minimum equivocation),

$$E_{\beta}^r(R_X, R_Y, W, P_{XYU}) \geq E_{r\beta}^{Liu}(R_X, R_Y, W, P_{XYU}) \quad \beta = X, Y, XY, \quad (95a)$$

$$E_{\beta}^T(R_X, R_Y, W, P_{XYU}) \geq E_{r\beta}^{Liu}(R_X, R_Y, W, P_{XYU}) \quad \beta = X, Y, XY, \quad (95b)$$

$$E_{\beta}^{ex}(R_X, R_Y, W, P_{XYU}) \geq E_{r\beta}^{Liu}(R_X, R_Y, W, P_{XYU}) \quad \beta = X, Y, XY. \quad (95c)$$

Hence

$$E_r(R_X, R_Y, W, P_{XYU}) \geq E_r^{Liu}(R_X, R_Y, W, P_{XYU}), \quad (96a)$$

$$E_T(R_X, R_Y, W, P_{XYU}) \geq E_r^{Liu}(R_X, R_Y, W, P_{XYU}), \quad (96b)$$

$$E_{ex}(R_X, R_Y, W, P_{XYU}) \geq E_r^{Liu}(R_X, R_Y, W, P_{XYU}), \quad (96c)$$

for all $P_{XYU} \in \mathcal{P}(\mathcal{X} \times \mathcal{Y} \times \mathcal{U})$ satisfying $X - U - Y$. Here, E_r^{Liu} is the random coding exponent of [20]. $E_{r\beta}^{Liu}$ are also defined in [20] for $\beta = X, Y, XY$.

Proof: The proof is provided in the Appendix. ■

We expect our typical random coding and expurgated bound to be strictly better than the one in [20] at low rates. This is so, because all inequalities in (70a)-(70c) and (72a)-(72c) will be active at zero rates, and thus (due to continuity) at sufficiently low rates. Although we have not been able to prove this fact rigorously, in the next section, we show that this is true by numerically evaluating the expurgated bound for different rate pairs.

V. NUMERICAL RESULT

In this section, we calculate the exponent derived in Theorem 4 for a multiple-access channel very similar to the one used in [20]. This example shows that strict inequality can hold in (95c). Consider a discrete memoryless MAC with $\mathcal{X} = \mathcal{Y} = \mathcal{Z} = \{0, 1\}$ and the transition probability given in the following table.

x	y	z	$W(z xy)$
0	0	0	0.99
0	0	1	0.01
0	1	0	0.01
0	1	1	0.99
1	0	0	0.01
1	0	1	0.99
1	1	0	0.50
1	1	1	0.50

First, we choose some time-sharing alphabet \mathcal{U} of size $|\mathcal{U}| = 4$. Then some channel input distribution $P_U P_{X|U} P_{Y|U}$ is chosen randomly. The following table gives numerical values of the random coding exponent of [20], and the expurgated exponent we have obtained for selected rate pairs.

R_X	R_Y	$E_{ex}(R_X, R_Y, W, P_{UXY})$	$E_r^{Liu}(R_X, R_Y, W, P_{UXY})$
0.01	0.01	0.2672	0.2330
0.01	0.02	0.2671	0.2330
0.01	0.03	0.2671	0.2330
0.02	0.01	0.2458	0.2230
0.02	0.02	0.2379	0.2230
0.02	0.05	0.2379	0.2230
0.03	0.01	0.2279	0.2130
0.03	0.03	0.2183	0.2130
0.04	0.01	0.2123	0.2030
0.04	0.04	0.2040	0.2030
0.05	0.05	0.1930	0.1930
0.06	0.01	0.1856	0.1830
0.06	0.06	0.1830	0.1830
0.07	0.01	0.1740	0.1730
0.07	0.07	0.1730	0.1730

As we see in the table, in the low rate regime, we have strictly better results in comparison with the results of [20]. For larger rate pairs, the inequalities containing $\min\{R_X, R_Y\}$ will not be active anymore, thus, we will end up with result similar to [20].

VI. CONCLUSIONS

We studied a unified framework to obtain all known lower bounds (random coding, typical random coding and expurgated bound) on the reliability function of a point-to-point discrete memoryless channel. We showed that the typical random coding bound is the typical performance of the constant composition code ensemble. By using a similar idea with a two-user discrete memoryless multiple-access channel, we derived three lower bounds on the reliability function. The first one (random coding) is identical to the best known lower bound on the reliability function of DM-MAC. We also showed that the random coding bound is the average performance of the constant composition code ensemble. The second bound (typical random coding) is the typical performance of the constant composition code ensemble. To derive the third bound (expurgated), we eliminated some of the codewords from the codebook with a larger rate. This is the first bound of its type that explicitly uses the method of expurgation in a multi-user transmission system. We showed that the exponent of the typical random coding and expurgated bounds are greater than or equal to the exponent of the known random coding bounds for all rate pairs. By numerical evaluation of the random coding and the expurgated bounds for a simple symmetric MAC, we showed that, at low rates, the expurgated bound is strictly larger. All these bounds can be universally obtained for all discrete memoryless MACs with given input and output alphabets.

APPENDIX

1. Point to Point Proofs

This section contains the proof of all lemmas and theorems related to point to point result.

Proof: (Lemma 1) We use the method of random selection. Define M such that

$$2^{n(R-\delta)} \leq M \leq 2^{nR}.$$

In the following, we obtain the expectation of the packing function over the constant composition code ensemble. The expectation of $\pi(X^M, V_{X\tilde{X}})$ can be obtained as follows:

$$\begin{aligned} \mathbb{E}(\pi(X^M, V_{X\tilde{X}})) &= \frac{1}{M} \sum_{i=1}^M \sum_{j \neq i} \mathbb{E} \left(1_{T_{V_{X\tilde{X}}}}(X_i, X_j) \right) = \frac{1}{M} \sum_{i=1}^M \sum_{j \neq i} \mathbb{P} \left(X_j \in T_{V_{\tilde{X}|X}}(X_i) \right) \\ &= (M-1) \mathbb{P} \left(X_2 \in T_{V_{\tilde{X}|X}}(X_1) \right) \leq 2^{n(R - I_V(X \wedge \tilde{X}) + \delta)}. \end{aligned} \quad (97)$$

Similarly, it can be shown that for sufficiently large n ,

$$\mathbb{E}(\pi(X^M, V_{X\tilde{X}})) \geq 2^{n(R - I_V(X \wedge \tilde{X}) - \delta)}. \quad (98)$$

The expectation of λ over the ensemble can be written as

$$\mathbb{E}(\lambda(X^M, V_{X\tilde{X}\hat{X}})) = \frac{1}{M} \sum_{i=1}^M \sum_{j \neq i} \sum_{k \neq i, j} \mathbb{P}((X_i, X_j, X_k) \in T_{V_{X\tilde{X}\hat{X}}}). \quad (99)$$

Since

$$\frac{2^{n[H(\tilde{X}\tilde{X}|X)-\delta]}}{2^{nH(\tilde{X})}2^{nH(\tilde{X})}} \leq \mathbb{P}((X_i, X_j, X_k) \in T_{V_{X\tilde{X}\tilde{X}}}) \leq \frac{2^{nH(\tilde{X}\tilde{X}|X)}}{2^{n[H(\tilde{X})-\delta]}2^{n[H(\tilde{X})-\delta]}}, \quad (100)$$

it can be concluded that

$$2^{n[E_S(R, V_{X\tilde{X}\tilde{X}})-2\delta]} \leq \mathbb{E}(\lambda(X^M, V_{X\tilde{X}\tilde{X}})) \leq 2^{n[E_S(R, V_{X\tilde{X}\tilde{X}})+2\delta]}, \quad (101)$$

where

$$E_S(R, V_{X\tilde{X}\tilde{X}}) \triangleq 2R - I(X \wedge \tilde{X}) - I(\tilde{X} \wedge \tilde{X}X). \quad (102)$$

By using (97) and markov inequality, it can be concluded that

$$\mathbb{P}\left(\pi(X^M, V_{X\tilde{X}}) \geq 2^{n(R-I_V(X \wedge \tilde{X})+2\delta)} \text{ for some } V_{X\tilde{X}}\right) \leq \sum_{V_{X\tilde{X}}} \frac{\mathbb{E}(\pi(X^M, V_{X\tilde{X}}))}{2^{n(R-I_V(X \wedge \tilde{X})+2\delta)}} \leq 2^{-n\frac{\delta}{2}}, \quad (103)$$

therefore, there exists at least one code, C^r , with M codewords satisfying

$$\pi(C^r, V_{X\tilde{X}}) \leq 2^{n(R-I_V(X \wedge \tilde{X})+2\delta)}. \quad (104)$$

■

Proof: (Lemma 2) To prove that a specific property holds for almost all codes, with certain number of codewords, in the constant composition code ensemble, we use a second-order argument method. We already have obtained upper and lower bounds on the expectation of the desired function over the entire ensemble. In the following, we derive an upper bound on the variance of the packing function. Finally, by using the Chebychev's inequality, we prove that the desired property holds for almost all codes in the ensemble.

To find the variance of the packing function, let us define $U_{ij} \triangleq 1_{T_{V_{X\tilde{X}}}}(X_i, X_j)$, and $Y_{ij} \triangleq U_{ij} + U_{ji}$. We can rewrite $\pi(X^M, V_{X\tilde{X}})$ as

$$\pi(X^M, V_{X\tilde{X}}) = \frac{1}{M} \sum_{i=1}^M \sum_{j \neq i} U_{ij} = \frac{1}{M} \sum_{i=1}^M \sum_{j < i} (U_{ij} + U_{ji}) = \frac{1}{M} \sum_{i=1}^M \sum_{j < i} Y_{ij}. \quad (105)$$

It is easy to check that Y_{ij} 's are identically distributed pairwise independent random variables. Therefore, the variance of $\pi(X^M, V_{X\tilde{X}})$ can be written as

$$\text{Var}(\pi(X^M, V_{X\tilde{X}})) = \frac{1}{M^2} \sum_{i=1}^M \sum_{j < i} \text{Var}(Y_{ij}) = \frac{1}{M^2} \binom{M}{2} \text{Var}(Y_{21}). \quad (106)$$

To find the variance of Y_{21} , let us consider the following two cases for $V_{X\tilde{X}}$:

- $V_{X\tilde{X}}$ is a symmetric distribution. In this case $U_{12} = U_{21}$, therefore,

$$Y_{21} = \begin{cases} 2 & \text{with probability } p \leq 2^{-n[I(X \wedge \tilde{X})-\delta]} \\ 0 & \text{with probability } 1-p \end{cases},$$

and the variance is upper bounded by

$$\text{Var}(Y_{21}) \leq E(Y_{21}^2) = 4 \times 2^{-n[I(X \wedge \tilde{X})-\delta]} \quad (107)$$

- $V_{X\tilde{X}}$ is not a symmetric distribution. In this case, if $U_{ij} = 1 \Rightarrow U_{ji} = 0$. Therefore,

$$\mathbb{P}(Y_{12} = 1) = \mathbb{P}(U_{12} = 1 \text{ or } U_{21} = 1) = \mathbb{P}(U_{12} = 1) + \mathbb{P}(U_{21} = 1) \leq 2 \times 2^{-n[I(X \wedge \tilde{X}) - \delta]}, \quad (108)$$

therefore,

$$\text{Var}(Y_{21}) \leq E(Y_{21}^2) = 2 \times 2^{-n[I(X \wedge \tilde{X}) - \delta]}. \quad (109)$$

By using (107), and (109), we have

$$\text{Var}(\pi(X^M, V_{X\tilde{X}})) \leq \frac{1}{M^2} \binom{M}{2} 4 \times 2^{-n[I(X \wedge \tilde{X}) - \delta]} \leq 2 \times 2^{-n[I(X \wedge \tilde{X}) - \delta]}, \quad (110)$$

for any $V_{X\tilde{X}} \in \mathcal{P}(\mathcal{X} \times \mathcal{X})$. Now, by using Chebychev's inequality,

$$\begin{aligned} & \mathbb{P}(|\pi(X^M, V_{X\tilde{X}}) - \mathbb{E}(\pi(X^M, V_{X\tilde{X}}))| \geq 2^{n\delta} \text{ for some } V_{X\tilde{X}}) \\ & \leq \sum_{V_{X\tilde{X}}} \mathbb{P}(|\pi(X^M, V_{X\tilde{X}}) - \mathbb{E}(\pi(X^M, V_{X\tilde{X}}))| \geq 2^{n\delta}) \\ & \leq \sum_{V_{X\tilde{X}}} \frac{\text{Var}(\pi(X^M, V_{X\tilde{X}}))}{2^{2n\delta}} \leq \sum_{V_{X\tilde{X}}} \frac{2 \times 2^{-n[I(X \wedge \tilde{X}) - \delta]}}{2^{2n\delta}} \\ & = \sum_{V_{X\tilde{X}}} 2 \times 2^{-n(I(X \wedge \tilde{X}) + \delta)} \leq 2^{-n\frac{\delta}{2}}, \quad \text{for sufficiently large } n. \end{aligned} \quad (111)$$

Moreover, by using (101) and Markov's inequality, it can be concluded that

$$\mathbb{P}(\lambda(X^M, V_{X\tilde{X}\hat{X}}) \geq 2^{n[E_S(R, V_{X\tilde{X}\hat{X}}) + 4\delta]} \text{ for some } V_{X\tilde{X}\hat{X}}) \leq \sum_{V_{X\tilde{X}\hat{X}}} \frac{\mathbb{E}\lambda(X^M, V_{X\tilde{X}\hat{X}})}{2^{n[E_S(R, V_{X\tilde{X}\hat{X}}) + 4\delta]}} \leq 2^{-n\delta}. \quad (112)$$

Now, by combining (111) and (112) and using the bound on $\mathbb{E}(\pi(X^M, V_{X\tilde{X}}))$, we conclude that for any $V_{X\tilde{X}} \in \mathcal{P}(\mathcal{X} \times \mathcal{X})$, any $V_{X\tilde{X}\hat{X}} \in \mathcal{P}(\mathcal{X} \times \mathcal{X} \times \mathcal{X})$, for sufficiently large n

$$\begin{aligned} 2^{n(R - I(X \wedge \tilde{X}) - \delta)} & \leq \pi(X^M, V_{X\tilde{X}}) \leq 2^{n(R - I(X \wedge \tilde{X}) + \delta)}, \\ \lambda(X^M, V_{X\tilde{X}\hat{X}}) & \leq 2^{n[E_S(R, V_{X\tilde{X}\hat{X}}) + 4\delta]}, \end{aligned} \quad (113)$$

with probability $> 1 - 2 \times 2^{-n\frac{\delta}{2}}$. We put all the codebooks satisfying (113) in a set called \mathcal{C}^T . ■

Proof: (Lemma 3) Consider the code $C^r \triangleq \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M\}$ whose existence is asserted in random coding packing lemma. Let us define

$$\Pi(C^r) \triangleq \sum_{V_{X\tilde{X}}} 2^{-n(R - I_V(X \wedge \tilde{X}) + 3\delta)} \pi(C^r, V_{X\tilde{X}}). \quad (114)$$

Note that using Lemma 1 and using the fact that $\Pi(C^r) = \frac{1}{M} \sum_{i=1}^M \left\{ \sum_{V_{X\tilde{X}}} |T_{V_{\tilde{X}}|\mathbf{x}}(\mathbf{x}_i) \cap C^r| 2^{-n(R - I_V(X \wedge \tilde{X}) + 3\delta)} \right\}$, it can be concluded that

$$\Pi(C^r) \leq \sum_{V_{X\tilde{X}}} 2^{-n(R - I_V(X \wedge \tilde{X}) + 3\delta)} 2^{n(R - I_V(X \wedge \tilde{X}) + 2\delta)} < \frac{1}{2}. \quad (115)$$

As a result, it can be concluded that there exists $M^* \geq \frac{M}{2}$ codewords in C^r satisfying

$$\sum_{V_{X\tilde{X}}} |T_{V_{\tilde{X}|X}}(\mathbf{x}_i) \cap C^r| 2^{-n(R-I_V(X\wedge\tilde{X})+3\delta)} < 1. \quad (116)$$

Let us call this subset of the code as C^{ex} . Without loss of generality, we assume C^{ex} contains the first M^* sequences of C^r , i.e., $C^{ex} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{M^*}\}$. Since

$$|T_{V_{\tilde{X}|X}}(\mathbf{x}_i) \cap C^{ex}| \leq |T_{V_{\tilde{X}|X}}(\mathbf{x}_i) \cap C^r| \quad \forall \mathbf{x}_i \in C^{ex}, \quad (117)$$

it can be concluded that for all $\mathbf{x}_i \in C^{ex}$,

$$\sum_{V_{X\tilde{X}}} |T_{V_{\tilde{X}|X}}(\mathbf{x}_i) \cap C^{ex}| 2^{-n(R-I_V(X\wedge\tilde{X})+3\delta)} < 1. \quad (118)$$

Since all the terms in the summation are non-negative terms, we conclude that

$$|T_{V_{\tilde{X}|X}}(\mathbf{x}_i) \cap C^{ex}| < 2^{n(R-I_V(X\wedge\tilde{X})+3\delta)}, \quad (119)$$

for all $V_{X\tilde{X}} \in \mathcal{P}(\mathcal{X} \times \mathcal{X})$, and all $\mathbf{x}_i \in C^{ex}$. Also, by (119), it can be concluded that for all $V_{X\tilde{X}} \in \mathcal{P}(\mathcal{X} \times \mathcal{X})$

$$\pi(C^{ex}, V_{X\tilde{X}}) = \frac{1}{M^*} \sum_{i=1}^{M^*} |T_{V_{\tilde{X}|X}}(\mathbf{x}_i) \cap C^{ex}| \leq 2^{n(R-I_V(X\wedge\tilde{X})+3\delta)}. \quad (120)$$

■

Proof: (Fact 1) We will use the result of Lemma 1 and the relation between the probability of error and the packing functions. Let $X^M \triangleq (X_1, X_2, \dots, X_M)$ be independent sequences of independent random variable, where X_i s are uniformly distributed on T_P .

(Upper Bound): Taking expectation on both sides of (23), using Lemma 1 and using the continuity of information measures, it can be concluded that

$$\begin{aligned} \mathbb{E}(e(X^M, W)) &\leq \sum_{V_{X\tilde{X}Y} \in \mathcal{P}_n^r} 2^{-n[D(V_{Y|X}||W|P) + |I(\tilde{X}\wedge XY) - R|^+ - \delta]} \\ &\leq 2^{-n[E_r(R, P, W) - 2\delta]} \end{aligned} \quad (121)$$

whenever $n \geq n_1(|\mathcal{X}|, |\mathcal{Y}|, \delta)$, where

$$E_r(R, P, W) \triangleq \min_{V_{X\tilde{X}Y} \in \mathcal{P}^r} D(V_{Y|X}||W|P) + |I_V(XY \wedge \tilde{X}) - R|^+, \quad (122)$$

and \mathcal{P}^r is defined in (36).

(Lower Bound): Taking expectation on both sides of (26), and using Lemma 1 we have

$$\begin{aligned} \bar{P}_e = \mathbb{E}e(X^M, W) &\geq \sum_{V_{X\tilde{X}Y} \in \mathcal{P}_n^r} 2^{-n[D(V_{Y|X}||W|P) + I_V(\tilde{X}\wedge Y|X) + \delta]} \left| 2^{n(R - I(X\wedge\tilde{X}) - \delta)} - \right. \\ &\quad \left. \sum_{\substack{V_{X\tilde{X}\hat{X}Y}: \\ V_{X\tilde{X}Y} = V_{X\hat{X}Y}}} 2^{-n[I_V(\hat{X}\wedge Y|X\tilde{X})]} 2^{n(2R - I(X\wedge\tilde{X}) - I(\hat{X}\wedge X\tilde{X}) + 4\delta)} \right|^+ \end{aligned} \quad (123)$$

$$= \sum_{V_{X\tilde{X}Y} \in \mathcal{P}_n^r} 2^{-n[D(V_{Y|X}||W|P) + I_V(\tilde{X}\wedge XY) - R + 2\delta]} \left| 1 - \sum_{\substack{V_{X\tilde{X}\hat{X}Y}: \\ V_{X\tilde{X}Y} = V_{X\hat{X}Y}}} 2^{-n[I_V(\hat{X}\wedge X\tilde{X}Y) - R + 3\delta]} \right|^+ \quad (124)$$

Toward further simplification of this expression, we have the following lemma.

Lemma 7.

$$\min_{\substack{V_{X\tilde{X}\hat{X}Y}: \\ V_{X\tilde{X}Y} = V_{X\hat{X}Y}}} I(\hat{X} \wedge X\tilde{X}Y) = I(\tilde{X} \wedge XY). \quad (125)$$

Proof: Note that, for any $V_{X\tilde{X}\hat{X}Y}$,

$$I(\hat{X} \wedge X\tilde{X}Y) = I(\hat{X} \wedge XY) + I(\hat{X} \wedge \tilde{X}|XY) \geq I(\hat{X} \wedge XY), \quad (126)$$

therefore,

$$\min_{\substack{V_{X\tilde{X}\hat{X}Y}: \\ V_{X\tilde{X}Y} = V_{X\hat{X}Y}}} I(\hat{X} \wedge X\tilde{X}Y) \geq I(\hat{X} \wedge XY) = I(\tilde{X} \wedge XY). \quad (127)$$

Now, consider $V_{X\tilde{X}\hat{X}Y}^*$ defined as

$$V_{X\tilde{X}\hat{X}Y}^*(x, \tilde{x}, \hat{x}, y) = V_{\tilde{X}|XY}(\tilde{x}|x, y) V_{\hat{X}|XY}(\hat{x}|x, y) V_{XY}(x, y). \quad (128)$$

Note that $V_{X\tilde{X}Y}^* = V_{X\hat{X}Y}^*$, and $\tilde{X} - (X, Y) - \hat{X}$. Therefore,

$$I_{V^*}(\hat{X} \wedge X\tilde{X}Y) = I_V(\hat{X} \wedge XY) = I_V(\tilde{X} \wedge XY). \quad (129)$$

By combining (127) and (129), the proof is complete. \blacksquare

Therefore, using the above lemma, (124) can be rewritten as

$$\bar{P}_e \geq \sum_{\substack{V_{X\tilde{X}Y} \in \mathcal{P}_n^r \\ I(\tilde{X}\wedge XY) > R + 3\delta}} 2^{-n[D(V_{Y|X}||W|P) + I_V(\tilde{X}\wedge XY) - R + 3\delta]}. \quad (130)$$

By using the continuity of information measures, it can be concluded that

$$\mathbb{E}(e(X^M, W)) \geq 2^{-n[E_L(R, P, W) + 4\delta]}, \quad \text{for sufficient large } n \quad (131)$$

where

$$E_L(R, P, W) \triangleq \min_{\substack{V_{X\tilde{X}Y} \in \mathcal{P}_n^r \\ I(\tilde{X}\wedge XY) \geq R}} D(V_{Y|X}||W|P) + I_V(XY \wedge \tilde{X}) - R. \quad (132)$$

Now, by using Markov inequality and (121), we conclude that

$$\mathbb{P}\left(e(X^M, W) \geq 2^{-n[E_r(R,P,W)-3\delta]}\right) \leq \frac{\mathbb{E}(e(X^M, W))}{2^{-n[E_r(R,P,W)-3\delta]}} \leq 2^{-n\delta}. \quad (133)$$

Therefore, with probability greater than $1 - 2^{-n\delta}$, any selected code with M codewords form the constant composition code ensemble satisfies the desired property. Let us call one of these codebooks as C^r . ■

Proof: (Corollary 1) Consider the input distribution $P^* \in \mathcal{P}(\mathcal{X})$ maximizing the random coding bound, i.e.,

$$P^* \triangleq \arg \max_{P \in \mathcal{P}(\mathcal{X})} E_r(R, P, W). \quad (134)$$

Let us define

$$V_{X\tilde{X}Y}^* \triangleq \arg \min_{V_{X\tilde{X}Y}} E_r(R, P^*, W). \quad (135)$$

For any $R \leq R_{crit}$, the random coding bound is a straight line with slope -1 , and the term in $|\cdot|^+$ is active. Therefore,

$$E_r(R, P^*, W) = D(V_{Y|X}^* || W | P^*) + I_{V^*}(\tilde{X} \wedge XY) - R. \quad (136)$$

Here, $I_{V^*}(\tilde{X} \wedge XY) \geq R$. It is clear that $V_{X\tilde{X}Y}^*$ is the minimizing distribution in $E_{rL}(R, P^*, W)$, and as a result

$$E_{rL}(R, P^*, W) = E_r(R, P^*, W). \quad (137)$$

■

Proof: (Theorem 1) In the proof of Fact 1, we used the lower and upper bounds on the expected value of the first-order packing functions and an upper bound on the expected value of the second-order packing functions. In the following, we use similar techniques on the packing function of almost every codebook in the ensemble by using the bounds obtained in Lemma 2. Consider the code C whose existence is asserted in the typical random coding packing lemma. For all $V_{X\tilde{X}} \in \mathcal{P}(\mathcal{X} \times \mathcal{X})$, we have

$$\frac{1}{M} \sum_{i=1}^M |T_{V_{\tilde{X}|X}}(\mathbf{x}_i) \cap C| \leq 2^{n(R - I_V(X \wedge \tilde{X}) + 2\delta)}. \quad (138)$$

By multiplying both sides of inequality (138) by M , and using the proper upper bound on the number of sequences in C , we conclude that

$$|T_{V_{\tilde{X}|X}}(\mathbf{x}_i) \cap C| \leq 2^{n(2R - I_V(X \wedge \tilde{X}) + 2\delta)} \quad \forall i = 1, \dots, M, \quad (139)$$

for all $V_{X\tilde{X}} \in \mathcal{P}(\mathcal{X} \times \mathcal{X})$. We will obtain a higher error exponent for almost all codes by removing certain types from the constraint set \mathcal{P}_n^r . Consider any $V_{X\tilde{X}} \in \mathcal{P}(\mathcal{X} \times \mathcal{X})$ satisfying $I_V(X \wedge \tilde{X}) > 2(R + \delta)$. By (139),

$$|T_{V_{\tilde{X}|X}}(\mathbf{x}_i) \cap C| = 0 \text{ for all } i \Rightarrow \pi(C, V_{X\tilde{X}}) = 0. \quad (140)$$

Upper bound: Hence, by using (23) on C , and by using the result of Lemma 2, we have

$$e(C, W) \leq \sum_{V_{X\tilde{X}Y} \in \mathcal{P}_n^T(\delta)} 2^{-n[D(V_{Y|X} || W | P) + I_V(XY \wedge \tilde{X}) - R]^+ - 2\delta}$$

where

$$\mathcal{P}_n^T(\delta) \triangleq \{V_{X\tilde{X}Y} \in \mathcal{P}_n(\mathcal{X} \times \mathcal{X} \times \mathcal{Y}) : V_X = V_{\tilde{X}} = P, I_V(X \wedge \tilde{X}) \leq 2R + 2\delta, \alpha(P, V_{Y|\tilde{X}}) \leq \alpha(P, V_{Y|X})\} \quad (141)$$

Using the continuity of information measures, the upper bound as given by the theorem follows.

Lower bound: Using (26) on C and using Lemma 2, we have

$$\begin{aligned} e(C, W) &\geq \sum_{V_{X\tilde{X}Y} \in \mathcal{P}_n^T} 2^{-n[D(V_{Y|X}||W|P) + I_V(\tilde{X} \wedge Y|X) + \delta]} \left| \pi(C, V_{X\tilde{X}}) - \sum_{\substack{V_{X\tilde{X}\tilde{X}Y}: \\ V_{X\tilde{X}Y} = V_{X\tilde{X}Y}}} 2^{-n[I_V(\tilde{X} \wedge Y|X\tilde{X})]} \lambda(C, V_{X\tilde{X}\tilde{X}}) \right|^+ \\ &\geq \sum_{V_{X\tilde{X}Y} \in \mathcal{P}_n^T(\delta)} 2^{-n[D(V_{Y|X}||W|P) + I_V(\tilde{X} \wedge Y|X) + \delta]} \left| 2^{n(R - I(X \wedge \tilde{X}) - \delta)} - \sum_{\substack{V_{X\tilde{X}\tilde{X}Y}: \\ V_{X\tilde{X}Y} = V_{X\tilde{X}Y}}} 2^{-n[I_V(\tilde{X} \wedge Y|X\tilde{X})]} 2^{n(2R - I(X \wedge \tilde{X}) - I(\tilde{X} \wedge X\tilde{X}) + 2\delta)} \right|^+ \end{aligned} \quad (142)$$

$$= \sum_{V_{X\tilde{X}Y} \in \mathcal{P}_n^T(\delta)} 2^{-n[D(V_{Y|X}||W|P) + I_V(\tilde{X} \wedge XY) - R + 2\delta]} \left| 1 - \sum_{\substack{V_{X\tilde{X}\tilde{X}Y}: \\ V_{X\tilde{X}Y} = V_{X\tilde{X}Y}}} 2^{-n[I_V(\tilde{X} \wedge X\tilde{X}Y) - R - 3\delta]} \right|^+ \quad (143)$$

$$\geq \sum_{\substack{V_{X\tilde{X}Y} \in \mathcal{P}_n^T(\delta) \\ I(\tilde{X} \wedge XY) > R + 5\delta}} 2^{-n[D(V_{Y|X}||W|P) + I_V(\tilde{X} \wedge XY) - R + 3\delta]}, \quad (144)$$

Here, the last inequality follows from Lemma 7.

By using the continuity argument, and for sufficient large n ,

$$e(C, W) \geq 2^{-n[E_{LT}(R, P, W) + 4\delta]}, \quad (145)$$

where

$$E_{LT}(R, P, W) \triangleq \min_{\substack{V_{X\tilde{X}Y} \in \mathcal{P}^T \\ I(\tilde{X} \wedge XY) \geq R}} D(V_{Y|X}||W|P) + I_V(XY \wedge \tilde{X}) - R. \quad (146)$$

Proof: (Corollary 2) Fix $R \geq 0$, $\delta > 0$. By the result of Theorem 1 and for sufficiently large n , there exists a collection of codes, \mathcal{C}^* , with length n and rate R , such that

- $\mathbb{P}(\mathcal{C}^*) \geq 1 - \delta$,
- $2^{-n[E_{TL}(R, P, W) + 4\delta]} \leq e(C, W) \leq 2^{-n[E_T(R, P, W) - 3\delta]}$ for all $C \in \mathcal{C}^*$.

Note that

$$\max_{\hat{\mathcal{C}}: \mathbb{P}(\hat{\mathcal{C}}) > 1 - \delta} \min_{C \in \hat{\mathcal{C}}} -\frac{1}{n} \log e(C, W) \geq \min_{C \in \mathcal{C}^*} -\frac{1}{n} \log e(C, W) \geq E_T(R, P, W) - 3\delta. \quad (147)$$

Now, consider any high probability collection of codes with length n and rate R . Let us call this collection as $\hat{\mathcal{C}}$. Note that

$$\left. \begin{aligned} \mathbb{P}(\mathcal{C}^*) &\geq 1 - \delta \\ \mathbb{P}(\hat{\mathcal{C}}) &\geq 1 - \delta \end{aligned} \right\} \Rightarrow \mathbb{P}(\mathcal{C}^* \cap \hat{\mathcal{C}}) \geq 1 - 2\delta \Rightarrow \mathcal{C}^* \cap \hat{\mathcal{C}} \neq \emptyset. \quad (148)$$

Consider a code $C(\hat{\mathcal{C}}) \in \mathcal{C}^* \cap \hat{\mathcal{C}}$. It can be concluded that

$$\max_{\tilde{\mathcal{C}}: \mathbb{P}(\tilde{\mathcal{C}}) > 1-\delta} \min_{C \in \tilde{\mathcal{C}}} -\frac{1}{n} \log e(C, W) \leq \max_{\tilde{\mathcal{C}}: \mathbb{P}(\tilde{\mathcal{C}}) > 1-\delta} -\frac{1}{n} \log e(C(\tilde{\mathcal{C}}), W) \leq E_{LT}(R, P, W) + 4\delta. \quad (149)$$

The last inequality follows from the fact that $C(\hat{\mathcal{C}}) \in \mathcal{C}^*$. By combining (147) and (149), and by letting δ goes to zero and n goes to infinity, it can be concluded that

$$E_T(R, P, W) \leq E_{av}^T(R) \leq E_{TL}(R, P, W). \quad (150)$$

■

Proof: (Fact 2) First, we prove the following lemma.

Lemma 8. *Let C^{ex} be the collection of the codewords whose existence is asserted in Lemma 3. For any distribution $V_{X\tilde{X}} \in \mathcal{P}_n(\mathcal{X} \times \mathcal{X})$, satisfying $I_V(X \wedge \tilde{X}) > R + \delta$, the following holds:*

$$\pi(C^{ex}, V_{X\tilde{X}}) = 0. \quad (151)$$

Proof: By (32),

$$|T_{V_{\tilde{X}|X}}(\mathbf{x}_i) \cap C^{ex}| \leq 2^{n(R - I_V(X \wedge \tilde{X}) + 2\delta)}, \quad (152)$$

for every $\mathbf{x}_i \in C^{ex}$. Since $I_V(X \wedge \tilde{X}) > R + 2\delta$, it can be concluded that

$$|T_{V_{\tilde{X}|X}}(\mathbf{x}_i) \cap C^{ex}| = 0 \text{ for every } \mathbf{x}_i \in C^{ex} \Rightarrow \pi(C^{ex}, V_{X\tilde{X}}) = 0 \quad (153)$$

■

The rest of the proof is identical to the proof of random coding bound. ■

2. MAC Proofs

Proof: (Lemma 4) In this proof, we use a similar random coding argument that Pokorny and Wallmeier used in [19]. The main difference is that our lemma uses a different code ensemble which results in a tighter bound. Instead of choosing our sequences from T_{P_X} and T_{P_Y} , we choose our random sequences uniformly from $T_{P_{X|U}}(\mathbf{u})$, and $T_{P_{Y|U}}(\mathbf{u})$ for a given $\mathbf{u} \in T_{P_U}$. In [20], we see a similar random code ensemble, however, their packing lemma incorporates the channel output \mathbf{z} into the packing inequalities. One can easily show that, by using this packing lemma and considering the minimum equivocation decoding rule, we would end up with the random coding bound derived in [20].

Fix any \mathcal{U} , $P_{XYU} \in \mathcal{P}_n(\mathcal{U} \times \mathcal{X} \times \mathcal{Y})$ such that $X - U - Y$, $R_X \geq 0$, $R_Y \geq 0$, $\delta > 0$, and $\mathbf{u} \in T_{P_U}$. Define M_X, M_Y such that

$$2^{n(R_X - \delta)} \leq M_X \leq 2^{nR_X}, \quad 2^{n(R_Y - \delta)} \leq M_Y \leq 2^{nR_Y}.$$

First, we find upper bounds on the expectations of packing functions for a fixed α and $V_{U,XY\tilde{X}\tilde{Y}}$, with

respect to the random variables X_i and Y_j . Since X_i s and Y_j s are i.i.d random sequences, we have

$$\begin{aligned}
\mathbb{E}\left[N_U(X^{M_X} \times Y^{M_Y}, V_{UXY})\right] &\triangleq \mathbb{E}\left[\frac{1}{M_X M_Y} \sum_{i,j} 1_{T_{V_{UXY}}}(\mathbf{u}, X_i, Y_j)\right] \\
&= \mathbb{E}\left[1_{T_{V_{UXY}}}(\mathbf{u}, X_1, Y_1)\right] \\
&= \sum_{\mathbf{x}, \mathbf{y}} 1_{T_{V_{XY|U}}}(\mathbf{x}, \mathbf{y}|\mathbf{u}) \mathbb{P}(X_1 = \mathbf{x}|\mathbf{u}) \mathbb{P}(Y_1 = \mathbf{y}|\mathbf{u}) \\
&\leq \sum_{(\mathbf{x}, \mathbf{y}) \in T_{V_{XY|U}}(\mathbf{u})} 2^{-n[H_V(X|U) - \delta]} 2^{-n[H_V(Y|U) - \delta]} \\
&\leq 2^{nH_V(XY|U)} 2^{-n[H_V(X|U) - \delta]} 2^{-n[H_V(Y|U) - \delta]} \\
&= 2^{-n[I_V(X \wedge Y|U) - 2\delta]} = 2^{-n[F_U(V_{UXY}) - 2\delta]}. \tag{154}
\end{aligned}$$

On the other hand,

$$\begin{aligned}
\mathbb{E}\left[N_U(X^{M_X} \times Y^{M_Y}, V_{UXY})\right] &= \sum_{\mathbf{x}, \mathbf{y}} 1_{T_{V_{XY|U}}}(\mathbf{x}, \mathbf{y}|\mathbf{u}) \mathbb{P}(X_1 = \mathbf{x}|\mathbf{u}) \mathbb{P}(Y_1 = \mathbf{y}|\mathbf{u}) \\
&\geq \sum_{(\mathbf{x}, \mathbf{y}) \in T_{V_{XY|U}}(\mathbf{u})} 2^{-nH_V(X|U)} 2^{-nH_V(Y|U)} \\
&\geq 2^{n[H_V(XY|U) - \delta]} 2^{-nH_V(X|U)} 2^{-nH_V(Y|U)} \\
&= 2^{-n[I_V(X \wedge Y|U) + \delta]} = 2^{-n[F_U(V_{UXY}) + \delta]}. \tag{155}
\end{aligned}$$

Therefore, by (154) and (155),

$$2^{-n[F_U(V_{UXY}) + \delta]} \leq \mathbb{E}\left[N_U(X^{M_X} \times Y^{M_Y}, V_{UXY})\right] \leq 2^{-n[F_U(V_{UXY}) - 2\delta]}.$$
 \tag{156}

By using a similar argument,

$$\mathbb{E}\left[N_X(X^{M_X} \times Y^{M_Y}, V_{UXY\hat{X}})\right] \geq 2^{-n[F_X(V_{UXY\hat{X}}) - 4\delta]}.$$
 \tag{157}

On the other hand,

$$\begin{aligned}
\mathbb{E}\left[N_X(X^{M_X} \times Y^{M_Y}, V_{U_{XY}\tilde{X}})\right] &\geq (M_X - 1) \mathbb{E}\left[1_{T_{V_{U_{XY}}}}(\mathbf{u}, X_1, Y_1)1_{T_{V_{U_{XY}\tilde{X}}}}(\mathbf{u}, X_1, Y_1, X_2)\right] \\
&= (M_X - 1) \sum_{\mathbf{x}, \mathbf{y}} \mathbb{P}(X_1 = \mathbf{x}|\mathbf{u})\mathbb{P}(Y_1 = \mathbf{y}|\mathbf{u})1_{T_{V_{U_{XY}}}}(\mathbf{u}, \mathbf{x}, \mathbf{y}) \\
&\quad \cdot \sum_{\tilde{\mathbf{x}}} \mathbb{P}(X_2 = \tilde{\mathbf{x}}|\mathbf{u})1_{T_{V_{U_{XY}\tilde{X}}}}(\mathbf{u}, \mathbf{x}, \mathbf{y}, \tilde{\mathbf{x}}) \\
&\geq (M_X - 1) \sum_{\mathbf{x}, \mathbf{y} \in T_{V_{XY|U}}(\mathbf{u})} 2^{-nH_V(X|U)}2^{-nH_V(Y|U)} \\
&\quad \cdot \sum_{\tilde{\mathbf{x}} \in T_{V_{\tilde{X}|U_{XY}}(\mathbf{u}, \mathbf{x}, \mathbf{y})} 2^{-nH_V(\tilde{X}|U)} \\
&\geq (M_X - 1) 2^{n[H(XY|U)-\delta]}2^{-nH_V(X|U)}2^{-nH_V(Y|U)} \\
&\quad \cdot 2^{n[H_V(\tilde{X}|U_{XY})-\delta]}2^{-nH_V(\tilde{X}|U)} \\
&\geq 2^{-n[I_V(X \wedge Y|U)+I_V(\tilde{X} \wedge Y|U)+I_V(\tilde{X} \wedge X|U_{XY})-R_X+3\delta]} \\
&= 2^{-n[F_X(V_{U_{XY}\tilde{X}})+3\delta]}. \tag{158}
\end{aligned}$$

Therefore, by (157) and (158),

$$2^{-n[F_X(V_{U_{XY}\tilde{X}})+3\delta]} \leq \mathbb{E}\left[N_X(X^{M_X} \times Y^{M_Y}, V_{U_{XY}\tilde{X}})\right] \leq 2^{-n[F_X(V_{U_{XY}\tilde{X}})-4\delta]}. \tag{159}$$

By using a similar argument for $N_Y(X^{M_X} \times Y^{M_Y}, V_{U_{XY}\tilde{X}})$ and $N_{XY}(X^{M_X} \times Y^{M_Y}, V_{U_{XY}\tilde{X}\tilde{Y}})$, we can show that

$$2^{-n[F_Y(V_{U_{XY}\tilde{Y}})+3\delta]} \leq \mathbb{E}\left[N_Y(X^{M_X} \times Y^{M_Y}, V_{U_{XY}\tilde{Y}})\right] \leq 2^{-n[F_Y(V_{U_{XY}\tilde{Y}})-4\delta]}, \tag{160}$$

$$2^{-n[F_{XY}(V_{U_{XY}\tilde{X}\tilde{Y}})+4\delta]} \leq \mathbb{E}\left[N_{XY}(X^{M_X} \times Y^{M_Y}, V_{U_{XY}\tilde{X}\tilde{Y}})\right] \leq 2^{-n[F_{XY}(V_{U_{XY}\tilde{X}\tilde{Y}})-4\delta]}. \tag{161}$$

We can obtain an upper bound for $\mathbb{E} [\Lambda_{XY}(X^{M_X} \times Y^{M_Y}, V_{U_{XY}\tilde{X}\tilde{Y}\hat{X}\hat{Y}})]$ as follows

$$\begin{aligned}
& \mathbb{E} [\Lambda_{XY}(X^{M_X}, Y^{M_Y}, V_{U_{XY}\tilde{X}\tilde{Y}\hat{X}\hat{Y}})] \\
&= \mathbb{E} \left[\frac{1}{M_X M_Y} \sum_{i,j} \sum_{\substack{k \neq i \\ l \neq j}} \sum_{\substack{k' \neq i,k \\ l' \neq j,l}} 1_{T_{V_{U_{XY}\tilde{X}\tilde{Y}\hat{X}\hat{Y}}}(\mathbf{u}, X_i, Y_j, X_k, Y_l, X_{k'}, Y_{l'})} \right] \\
&\leq M_X^2 M_Y^2 \mathbb{E} [1_{T_{V_{U_{XY}}}(\mathbf{u}, X_1, Y_1)} 1_{T_{V_{U_{XY}\tilde{X}\tilde{Y}\hat{X}\hat{Y}}}(\mathbf{u}, X_1, Y_1, X_2, Y_2, X_3, Y_3)}] \\
&= M_X^2 M_Y^2 \sum_{\mathbf{x}, \mathbf{y}, \tilde{\mathbf{x}}, \tilde{\mathbf{y}}, \hat{\mathbf{x}}, \hat{\mathbf{y}}} \mathbb{P}(X_1 = \mathbf{x}, Y_1 = \mathbf{y}, X_2 = \tilde{\mathbf{x}}, Y_2 = \tilde{\mathbf{y}}, X_3 = \hat{\mathbf{x}}, Y_3 = \hat{\mathbf{y}} | \mathbf{u}) \\
&\quad \cdot 1_{T_{V_{U_{XY}}}(\mathbf{u}, \mathbf{x}, \mathbf{y})} \cdot 1_{T_{V_{U_{XY}\tilde{X}\tilde{Y}\hat{X}\hat{Y}}}(\mathbf{u}, \mathbf{x}, \mathbf{y}, \tilde{\mathbf{x}}, \tilde{\mathbf{y}}, \hat{\mathbf{x}}, \hat{\mathbf{y}})} \\
&= M_X^2 M_Y^2 \sum_{\mathbf{x}, \mathbf{y}} \mathbb{P}(X_1 = \mathbf{x} | \mathbf{u}) \Pr(Y_1 = \mathbf{y} | \mathbf{u}) \cdot 1_{T_{V_{U_{XY}}}(\mathbf{u}, \mathbf{x}, \mathbf{y})} \\
&\quad \cdot \sum_{\tilde{\mathbf{x}}} \mathbb{P}(X_2 = \tilde{\mathbf{x}} | \mathbf{u}) 1_{T_{V_{U_{XY}\tilde{X}}}(\mathbf{u}, \mathbf{x}, \mathbf{y}, \tilde{\mathbf{x}})} \sum_{\tilde{\mathbf{y}}} \mathbb{P}(Y_2 = \tilde{\mathbf{y}} | \mathbf{u}) 1_{T_{V_{U_{XY}\tilde{Y}}}(\mathbf{u}, \mathbf{x}, \mathbf{y}, \tilde{\mathbf{x}}, \tilde{\mathbf{y}})} \\
&\quad \cdot \sum_{\hat{\mathbf{x}}} \mathbb{P}(X_3 = \hat{\mathbf{x}} | \mathbf{u}) 1_{T_{V_{U_{XY}\tilde{X}\hat{X}}}(\mathbf{u}, \mathbf{x}, \mathbf{y}, \tilde{\mathbf{x}}, \tilde{\mathbf{y}}, \hat{\mathbf{x}})} \sum_{\hat{\mathbf{y}}} \mathbb{P}(Y_3 = \hat{\mathbf{y}} | \mathbf{u}) 1_{T_{V_{U_{XY}\tilde{Y}\hat{Y}}}(\mathbf{u}, \mathbf{x}, \mathbf{y}, \tilde{\mathbf{x}}, \tilde{\mathbf{y}}, \hat{\mathbf{x}}, \hat{\mathbf{y}})} \\
&\leq M_X^2 M_Y^2 \sum_{\mathbf{x}, \mathbf{y} \in T_{V_{XY|U}}(\mathbf{u})} 2^{-n[H_V(X|U) - \delta]} 2^{-n[H_V(Y|U) - \delta]} \sum_{\tilde{\mathbf{x}} \in T_{V_{\tilde{X}|U_{XY}}}(\mathbf{u}, \mathbf{x}, \mathbf{y})} 2^{-n[H_V(\tilde{X}|U) - \delta]} \\
&\quad \cdot \sum_{\tilde{\mathbf{y}} \in T_{V_{\tilde{Y}|U_{XY}\tilde{X}}}(\mathbf{u}, \mathbf{x}, \mathbf{y}, \tilde{\mathbf{x}})} 2^{-n[H_V(\tilde{Y}|U) - \delta]} \sum_{\hat{\mathbf{x}} \in T_{V_{\hat{X}|U_{XY}\tilde{X}\tilde{Y}}}(\mathbf{u}, \mathbf{x}, \mathbf{y}, \tilde{\mathbf{x}}, \tilde{\mathbf{y}})} 2^{-n[H_V(\hat{X}|U) - \delta]} \\
&\quad \cdot \sum_{\hat{\mathbf{y}} \in T_{V_{\hat{Y}|U_{XY}\tilde{X}\tilde{Y}\hat{X}}}(\mathbf{u}, \mathbf{x}, \mathbf{y}, \tilde{\mathbf{x}}, \tilde{\mathbf{y}}, \hat{\mathbf{x}})} 2^{-n[H_V(\hat{Y}|U) - \delta]} \\
&\leq M_X^2 M_Y^2 \cdot 2^{-nH(XY|U)} 2^{-n[H_V(X|U) - \delta]} 2^{-n[H_V(Y|U) - \delta]} 2^{-nH_V(\tilde{X}|U_{XY})} 2^{-n[H_V(\tilde{X}|U) - \delta]} 2^{-nH_V(\tilde{Y}|U_{XY}\tilde{X})} \\
&\quad \cdot 2^{-n[H_V(\tilde{Y}|U) - \delta]} 2^{-nH_V(\hat{X}|U_{XY}\tilde{X}\tilde{Y})} 2^{-n[H_V(\hat{X}|U) - \delta]} 2^{-nH_V(\hat{Y}|U_{XY}\tilde{X}\tilde{Y}\hat{X})} 2^{-n[H_V(\hat{Y}|U) - \delta]} \\
&\leq 2^{-n[I(\tilde{X}\tilde{Y} \wedge XY|U) + I(\tilde{X}\tilde{Y} \wedge XY\tilde{X}\tilde{Y}|U) + I(X \wedge Y|U) + I(\tilde{X} \wedge \tilde{Y}|U) + I(\hat{X} \wedge \hat{Y}|U) - 2R_X - 2R_Y - 6\delta]} \\
&= 2^{-n[E_S^{XY}(V_{U_{XY}\tilde{X}\tilde{Y}\hat{X}\hat{Y}}) - 6\delta]}. \tag{162}
\end{aligned}$$

By using a similar argument, we can obtain the following bounds

$$\mathbb{E} [\Lambda_X(X^{M_X} \times Y^{M_Y}, V_{U_{XY}\tilde{X}\hat{X}})] \leq 2^{-n[E_S^X(V_{U_{XY}\tilde{X}\hat{X}}) - 4\delta]} \tag{163}$$

$$\mathbb{E} [\Lambda_Y(X^{M_X} \times Y^{M_Y}, V_{U_{XY}\tilde{Y}\hat{Y}})] \leq 2^{-n[E_S^Y(V_{U_{XY}\tilde{Y}\hat{Y}}) - 4\delta]} \tag{164}$$

Here, E_S^X , E_S^Y and E_S^{XY} are defined in (67a)-(67c).

By using Markov inequality, it can be concluded that

$$\begin{aligned}
& \mathbb{P} \left(N_U(X^{M_X} \times Y^{M_Y}, V_{U_{XY}}) \geq 2^{-n[F_U(V_{U_{XY}}) - 3\delta]} \text{ for some } V_{U_{XY}} \right) \\
&\leq \sum_{\substack{V_{U_{XY}}: \\ V_{UX} = P_{UX} \\ V_{UY} = P_{UY}}} \frac{\mathbb{E} (N_U(X^{M_X} \times Y^{M_Y}, V_{U_{XY}}))}{2^{-n[F_U(V_{U_{XY}}) - 3\delta]}} \leq \sum_{\substack{V_{U_{XY}}: \\ V_{UX} = P_{UX} \\ V_{UY} = P_{UY}}} 2^{-n\delta} \leq 2^{-n\frac{\delta}{2}} \tag{165}
\end{aligned}$$

Similarly, it can be shown that

$$\mathbb{P}\left(N_X(X^{M_X} \times Y^{M_Y}, V_{U_{XY\tilde{X}}}) \geq 2^{-n[F_X(V_{U_{XY\tilde{X}}})-5\delta]} \text{ for some } V_{U_{XY\tilde{X}}}\right) \leq 2^{-n\frac{\delta}{2}}, \quad (166)$$

$$\mathbb{P}\left(N_Y(X^{M_X} \times Y^{M_Y}, V_{U_{XY\tilde{Y}}}) \geq 2^{-n[F_Y(V_{U_{XY\tilde{Y}}})-5\delta]} \text{ for some } V_{U_{XY\tilde{Y}}}\right) \leq 2^{-n\frac{\delta}{2}}, \quad (167)$$

$$\mathbb{P}\left(N_{XY}(X^{M_X} \times Y^{M_Y}, V_{U_{XY\tilde{X}\tilde{Y}}}) \geq 2^{-n[F_{XY}(V_{U_{XY\tilde{X}\tilde{Y}}})-5\delta]} \text{ for some } V_{U_{XY\tilde{X}\tilde{Y}}}\right) \leq 2^{-n\frac{\delta}{2}}. \quad (168)$$

Now, by combining (165)-(168), and using the union bound, it can be concluded that

$$\begin{aligned} &\mathbb{P}\left(N_U(X^{M_X} \times Y^{M_Y}, V_{U_{XY}}) \geq 2^{-n[F_U(V_{U_{XY}})-3\delta]} \text{ for some } V_{U_{XY}} \text{ or} \right. \\ &\quad N_X(X^{M_X} \times Y^{M_Y}, V_{U_{XY\tilde{X}}}) \geq 2^{-n[F_X(V_{U_{XY\tilde{X}}})-5\delta]} \text{ for some } V_{U_{XY\tilde{X}}} \text{ or} \\ &\quad N_Y(X^{M_X} \times Y^{M_Y}, V_{U_{XY\tilde{Y}}}) \geq 2^{-n[F_Y(V_{U_{XY\tilde{Y}}})-5\delta]} \text{ for some } V_{U_{XY\tilde{Y}}} \text{ or} \\ &\quad \left. N_{XY}(X^{M_X} \times Y^{M_Y}, V_{U_{XY\tilde{X}\tilde{Y}}}) \geq 2^{-n[F_{XY}(V_{U_{XY\tilde{X}\tilde{Y}}})-5\delta]} \text{ for some } V_{U_{XY\tilde{X}\tilde{Y}}}\right) \leq 4 \times 2^{-n\frac{\delta}{2}}, \end{aligned} \quad (169)$$

therefore, there exists at least a multi-user code with the desired properties mentioned in (76)-(77). \blacksquare

Proof: (Lemma 5) To prove that a specific property holds for almost all codes, with certain number of codewords, in the constant composition code ensemble, we use a second order argument method. We already have obtained upper and lower bounds on the expectation of the desired function over the entire ensemble. In the following, we derive an upper bound on the variance of the packing function. Finally, by using the Chebychev's inequality, we prove that the desired property holds for almost all codes in the ensemble. To find the variance of $N_U(X^{M_X} \times Y^{M_Y}, V_{U_{XY}})$, let us define $W_{ij} \triangleq 1_{T_{V_{U_{XY}}}(\mathbf{u}, X_i, Y_j)}$. Therefore, the variance of $N_U(X^{M_X} \times Y^{M_Y}, V_{U_{XY}})$ can be written as

$$\begin{aligned} \text{Var}\left(N_U(X^{M_X} \times Y^{M_Y}, V_{U_{XY}})\right) &= \text{Var}\left(\frac{1}{M_X M_Y} \sum_{i,j} 1_{T_{V_{U_{XY}}}(\mathbf{u}, X_i, Y_j)}\right) \\ &= \frac{1}{M_X^2 M_Y^2} \text{Var}\left(\sum_{i,j} W_{ij}\right). \end{aligned} \quad (170)$$

Since W_{ij} 's are pairwise independent random variables, (170) can be written as

$$\begin{aligned} \text{Var}\left(N_U(X^{M_X} \times Y^{M_Y}, V_{U_{XY}})\right) &= \frac{1}{M_X^2 M_Y^2} \sum_{i,j} \text{Var}(W_{ij}) \\ &\leq \frac{1}{M_X^2 M_Y^2} \sum_{i,j} \mathbb{E}(W_{ij}) \\ &\leq \frac{1}{M_X M_Y} \cdot 2^{-n[F_U(V_{U_{XY}})-2\delta]} \leq 2^{-n[F_U(V_{U_{XY}})+R_X+R_Y-2\delta]}. \end{aligned} \quad (171)$$

By defining $Q_{ik}^j \triangleq 1_{T_{V_{U_{XY\tilde{X}}}}(\mathbf{u}, X_i, Y_j, X_k)}$, the variance of $N_X(X^{M_X} \times Y^{M_Y}, V_{U_{XY\tilde{X}}})$ can be upper-

bounded as follows

$$\begin{aligned}
\text{Var} (N_X(X^{M_X} \times Y^{M_Y}, V_{U_{XY\bar{X}}})) &= \text{Var} \left(\frac{1}{M_X M_Y} \sum_{i,j} \sum_{k \neq i} 1_{T_{V_{U_{XY\bar{X}}}}(\mathbf{u}, X_i, Y_j, X_k)} \right) \\
&= \frac{1}{M_X^2 M_Y^2} \text{Var} \left(\sum_{i,j} \sum_{k \neq i} 1_{T_{V_{U_{XY\bar{X}}}}(\mathbf{u}, X_i, Y_j, X_k)} \right) \\
&= \frac{1}{M_X^2 M_Y^2} \text{Var} \left(\sum_j \sum_i \sum_{k \neq i} Q_{ik}^j \right) \\
&= \frac{1}{M_X^2 M_Y^2} \text{Var} \left(\sum_j \sum_i \sum_{k < i} Q_{ik}^j + Q_{ki}^j \right) = \frac{1}{M_X^2 M_Y^2} \text{Var} \left(\sum_j \sum_i \sum_{k < i} J_{i,k}^j \right), \tag{172}
\end{aligned}$$

where $J_{i,k}^j \triangleq Q_{ik}^j + Q_{ki}^j$, $k < i$. One can show that $J_{i,k}^j$'s are identically pairwise independent random variables. Therefore, the $\text{Var} (N_X(X^{M_X} \times Y^{M_Y}, V_{U_{XY\bar{X}}}))$ can be written as

$$\text{Var} (N_X(X^{M_X} \times Y^{M_Y}, V_{U_{XY\bar{X}}})) = \frac{1}{M_X^2 M_Y^2} \sum_j \sum_i \sum_{k < i} \text{Var} (J_{i,k}^j) \leq \frac{1}{2M_Y} \text{Var} (J_{2,1}^j). \tag{173}$$

To find the variance of $J_{2,1}^j$, let us consider the following two cases for $V_{U_{XY\bar{X}}}$:

- $V_{U_{XY\bar{X}}}$ is a symmetric distribution, i.e., $V_{U_{XY\bar{X}}} = V_{U_{\bar{X}YX}}$. In this case $Q_{12}^1 = Q_{21}^1$, therefore,

$$J_{2,1}^1 = \begin{cases} 2 & \text{with probability } p \approx 2^{-n[I_V(X \wedge Y|U) + I_V(\bar{X} \wedge XY|U)]} \\ 0 & \text{with probability } 1 - p \end{cases},$$

and the variance is upper bounded by

$$\text{Var}(J_{2,1}^1) \leq E(J_{2,1}^1)^2 = 4 \times 2^{-n[I_V(X \wedge Y|U) + I_V(\bar{X} \wedge XY|U)]}, \tag{174}$$

- $V_{U_{XY\bar{X}}}$ is not a symmetric distribution. In this case, if $Q_{ik}^j = 1 \Rightarrow Q_{ki}^j = 0$. Therefore,

$$\begin{aligned}
\mathbb{P}(J_{2,1}^1 = 1) &= \mathbb{P}(Q_{12}^1 = 1 \text{ or } Q_{21}^1 = 1) = \mathbb{P}(Q_{12}^1 = 1) + \mathbb{P}(Q_{21}^1 = 1) \\
&\leq 2 \times 2^{-n[I_V(X \wedge Y|U) + I_V(\bar{X} \wedge XY|U)]}, \tag{175}
\end{aligned}$$

therefore,

$$\text{Var}(J_{2,1}^1) \leq E(J_{2,1}^1)^2 = 2 \times 2^{-n[I_V(X \wedge Y|U) + I_V(\bar{X} \wedge XY|U)]}. \tag{176}$$

By combining the results in (173)-(175), it can be concluded that

$$\text{Var} (N_X(X^{M_X} \times Y^{M_Y}, V_{U_{XY\bar{X}}})) \leq 2^{-n[I_V(X \wedge Y|U) + I_V(\bar{X} \wedge XY|U) + R_Y - 3\delta]}. \tag{177}$$

Similarly, it can be shown that

$$\text{Var} (N_Y(X^{M_X} \times Y^{M_Y}, V_{U_{XY\bar{Y}}})) \leq 2^{-n[I_V(X \wedge Y|U) + I_V(\bar{Y} \wedge YX|U) + R_X - 3\delta]}. \tag{178}$$

By defining $R_{ik}^{jl} \triangleq 1_{T_{V_{U_{XY\bar{X}}}}(\mathbf{u}, X_i, Y_j, X_k, Y_l)}$, the variance of $N_{XY}(X^{M_X} \times Y^{M_Y}, V_{U_{XY\bar{X}\bar{Y}}})$ can be upper-

bounded as follows

$$\begin{aligned}
\text{Var} (N_{XY}(X^{M_X} \times Y^{M_Y}, V_{UXY\bar{X}\bar{Y}})) &= \text{Var} \left(\frac{1}{M_X M_Y} \sum_{i,j} \sum_{\substack{k \neq i \\ l \neq j}} 1_{T_{V_{UXY\bar{X}\bar{Y}}}(\mathbf{u}, X_i, Y_j, X_k, Y_l)} \right) \\
&= \frac{1}{M_X^2 M_Y^2} \text{Var} \left(\sum_{i,j} \sum_{\substack{k \neq i \\ l \neq j}} 1_{T_{V_{UXY\bar{X}\bar{Y}}}(\mathbf{u}, X_i, Y_j, X_k, Y_l)} \right) \\
&= \frac{1}{M_X^2 M_Y^2} \text{Var} \left(\sum_i \sum_j \sum_{k \neq i} \sum_{l \neq j} R_{ik}^{jl} \right) \\
&= \frac{1}{M_X^2 M_Y^2} \text{Var} \left(\sum_i \sum_j \sum_{k < i} \sum_{l < j} \left\{ R_{ik}^{jl} + R_{ki}^{jl} + R_{ik}^{lj} + R_{ki}^{lj} \right\} \right) \\
&= \frac{1}{M_X^2 M_Y^2} \text{Var} \left(\sum_i \sum_j \sum_{k < i} \sum_{j < l} S_{i,k}^{j,l} \right), \tag{179}
\end{aligned}$$

where $S_{i,k}^{j,l} \triangleq R_{ik}^{jl} + R_{ki}^{jl} + R_{ik}^{lj} + R_{ki}^{lj}$, $k < i$, $l < j$. It is easy to check that $S_{i,k}^{j,l}$'s are identically pairwise independent random variables. Therefore, the $\text{Var} (N_{XY}(X^{M_X} \times Y^{M_Y}, V_{UXY\bar{X}\bar{Y}}))$ can be written as

$$\text{Var} (N_{XY}(X^{M_X} \times Y^{M_Y}, V_{UXY\bar{X}\bar{Y}})) = \frac{1}{M_X^2 M_Y^2} \sum_i \sum_j \sum_{k < i} \sum_{l < j} \text{Var} (S_{i,k}^{j,l}) \leq \frac{1}{4} \text{Var} (S_{1,2}^{1,2}). \tag{180}$$

By using a similar argument to (174)-(175), the variance of To find the variance of $S_{1,2}^{1,2}$, can be upper bounded by

$$\text{Var} (S_{1,2}^{1,2}) \leq 16 \times 2^{-n[I_V(X \wedge Y|U) + I_V(\bar{X} \wedge \bar{Y}|U) + I_V(\bar{X}\bar{Y} \wedge XY|U) - 4\delta]}, \tag{181}$$

and therefore,

$$\text{Var} (N_{XY}(X^{M_X} \times Y^{M_Y}, V_{UXY\bar{X}\bar{Y}})) \leq 4 \times 2^{-n[I_V(X \wedge Y|U) + I_V(\bar{X} \wedge \bar{Y}|U) + I_V(\bar{X}\bar{Y} \wedge XY|U) - 4\delta]}. \tag{182}$$

Now, by using the Chebychev's inequality, we can obtain the following

$$\begin{aligned}
\mathbb{P} (|N_U(X^{M_X} \times Y^{M_Y}, V_{UXY}) - \mathbb{E} (N_U(X^{M_X} \times Y^{M_Y}, V_{UXY}))| \geq 2^{2n\delta} \text{ for some } V_{UXY}) \\
&\leq \sum_{V_{UXY}} \mathbb{P} (|N_U(X^{M_X} \times Y^{M_Y}, V_{UXY}) - \mathbb{E} (N_U(X^{M_X} \times Y^{M_Y}, V_{UXY}))| \geq 2^{2n\delta}) \\
&\leq \sum_{V_{UXY}} \frac{\text{Var} (N_U(X^{M_X} \times Y^{M_Y}, V_{UXY}))}{2^{4n\delta}} \\
&\leq \sum_V 2^{-n[F_U(V) + R_X + R_Y + 2\delta]} \leq 2^{-n\delta}. \tag{183}
\end{aligned}$$

Similarly, it can be shown that

$$\mathbb{P}(|N_X(X^{M_X} \times Y^{M_Y}, V_{UXY\tilde{X}}) - \mathbb{E}(N_X(X^{M_X} \times Y^{M_Y}, V_{UXY\tilde{X}}))| \geq 2^{2n\delta} \text{ for some } V_{UXY\tilde{X}}) \leq 2^{-n\delta} \quad (184)$$

$$\mathbb{P}(|N_Y(X^{M_X} \times Y^{M_Y}, V_{UXY\tilde{Y}}) - \mathbb{E}(N_Y(X^{M_X} \times Y^{M_Y}, V_{UXY\tilde{Y}}))| \geq 2^{2n\delta} \text{ for some } V_{UXY\tilde{Y}}) \leq 2^{-n\delta} \quad (185)$$

$$\mathbb{P}(|N_{XY}(X^{M_X} \times Y^{M_Y}, V_{UXY\tilde{X}\tilde{Y}}) - \mathbb{E}(N_{XY}(X^{M_X} \times Y^{M_Y}, V_{UXY\tilde{X}\tilde{Y}}))| \geq 2^{2n\delta} \text{ for some } V_{UXY\tilde{X}\tilde{Y}}) \leq 2^{-n\delta}. \quad (186)$$

Now, by using the result of Lemma 4 and Markov's inequality, it can be concluded that

$$\begin{aligned} & \mathbb{P}\left(\Lambda_X(X^{M_X} \times Y^{M_Y}, V_{UXY\tilde{X}\tilde{X}}) \geq 2^{-n(E_S^X(V_{UXY\tilde{X}\tilde{X}}) - 5\delta)} \text{ for some } V_{UXY\tilde{X}\tilde{X}}\right) \\ & \leq \sum_{V_{UXY\tilde{X}\tilde{X}}} \mathbb{P}\left(\Lambda_X(X^{M_X} \times Y^{M_Y}, V_{UXY\tilde{X}\tilde{X}}) \geq 2^{-n(E_S^X(V_{UXY\tilde{Y}\tilde{Y}}) - 5\delta)}\right) \\ & \leq \sum_{V_{UXY\tilde{X}\tilde{X}}} \frac{\mathbb{E}(\Lambda_X(X^{M_X} \times Y^{M_Y}, V_{UXY\tilde{X}\tilde{X}}))}{2^{-n(E_S^X(V_{UXY\tilde{X}\tilde{X}}) - 5\delta)}} \leq \sum_{V_{UXY\tilde{Y}\tilde{Y}}} 2^{-n\delta} \leq 2^{-n\frac{\delta}{2}}. \end{aligned} \quad (187)$$

Similarly,

$$\mathbb{P}\left(\Lambda_Y(X^{M_X} \times Y^{M_Y}, V_{UXY\tilde{Y}\tilde{Y}}) \geq 2^{-n(E_S^Y(V_{UXY\tilde{Y}\tilde{Y}}) - 5\delta)} \text{ for some } V_{UXY\tilde{Y}\tilde{Y}}\right) \leq 2^{-n\frac{\delta}{2}}, \quad (188)$$

and

$$\mathbb{P}\left(\Lambda_{XY}(X^{M_X} \times Y^{M_Y}, V_{UXY\tilde{X}\tilde{Y}\tilde{X}\tilde{Y}}) \geq 2^{-n(E_S^{XY}(V_{UXY\tilde{X}\tilde{Y}\tilde{X}\tilde{Y}}) - 7\delta)} \text{ for some } V_{UXY\tilde{X}\tilde{Y}\tilde{X}\tilde{Y}}\right) \leq 2^{-n\frac{\delta}{2}}. \quad (189)$$

Therefore, with probability $> 1 - 7 \times 2^{-n\frac{\delta}{2}}$, a code $C = C_X \times C_Y$ from random code ensemble satisfies the conditions given in the lemma. \blacksquare

Proof: (Lemma 6) Let $C_X^r = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{M_X}\}$ and $C_Y^r = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_{M_Y}\}$ be the collections of codewords whose existence is asserted in Lemma 4. Let us define

$$\begin{aligned} \Pi(C_X^r \times C_Y^r) & \triangleq \sum_{V_{UXY\tilde{X}\tilde{Y}}} \left\{ N_U(C_X^r \times C_Y^r, V_{UXY}) 2^{n[F_U(V_{UXY}) - 6\delta]} \right. \\ & \quad + N_X(C_X^r \times C_Y^r, V_{UXY\tilde{X}}) 2^{n[F_X(V_{UXY\tilde{X}}) - 6\delta]} \\ & \quad + N_Y(C_X^r \times C_Y^r, V_{UXY\tilde{Y}}) 2^{n[F_Y(V_{UXY\tilde{Y}}) - 6\delta]} \\ & \quad \left. + N_{XY}(C_X^r \times C_Y^r, V_{UXY\tilde{X}\tilde{Y}}) 2^{n[F_{XY}(V_{UXY\tilde{X}\tilde{Y}}) - 6\delta]} \right\} \end{aligned} \quad (190)$$

$$\leq \sum_{V_{UXY\tilde{X}\tilde{Y}}} 4 \times 2^{-n\delta} < \frac{1}{2} \quad (191)$$

For $C^r = C_X^r \times C_Y^r$, and the sequence \mathbf{u} defined in random coding packing lemma, we define

$$L_U(C^r, V_{UXY}, i, j) \triangleq 1_{T_{V_{UXY}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j), \quad (192)$$

$$L_X(C^r, V_{UXY\tilde{X}}, i, j) \triangleq \sum_{k \neq i} 1_{T_{V_{UXY\tilde{X}}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j, \mathbf{x}_k), \quad (193)$$

$$L_Y(C^r, V_{UXY\tilde{Y}}, i, j) \triangleq \sum_{l \neq j} 1_{T_{V_{UXY\tilde{Y}}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j, \mathbf{y}_l), \quad (194)$$

$$L_{XY}(C^r, V_{UXY\tilde{X}\tilde{Y}}, i, j) \triangleq \sum_{k \neq i} \sum_{l \neq j} 1_{T_{V_{UXY\tilde{X}\tilde{Y}}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j, \mathbf{x}_k, \mathbf{y}_l). \quad (195)$$

By definition of N_α , (190) can be written as

$$\Pi(C^r) = \frac{1}{M_X} \sum_{i=1}^{M_X} G(i), \quad \text{for } \alpha = U, X, Y, XY, \quad (196)$$

where $G(i)$ is defined as follows:

$$\begin{aligned} G(i) \triangleq & \frac{1}{M_Y} \sum_{j=1}^{M_Y} \sum_{V_{UXY\tilde{X}\tilde{Y}}} \left\{ L_U(C^r, V_{UXY}, i, j) 2^{n[F_U(V_{UXY})-6\delta]} \right. \\ & + L_X(C^r, V_{UXY\tilde{X}}, i, j) 2^{n[F_X(V_{UXY\tilde{X}})-6\delta]} \\ & + L_X(C^r, V_{UXY\tilde{Y}}, i, j) 2^{n[F_Y(V_{UXY\tilde{Y}})-6\delta]} \\ & \left. + L_{XY}(C^r, V_{UXY\tilde{X}\tilde{Y}}, i, j) 2^{n[F_{XY}(V_{UXY\tilde{X}\tilde{Y}})-6\delta]} \right\}. \end{aligned} \quad (197)$$

By using (191), we see that the average of $G(i)$ over C_X^r is upper bounded by $\frac{1}{2}$, therefore, there must exist $\hat{M}_X \geq \frac{M_X}{2}$ codewords, $\mathbf{x}_i \in C_X^r$, for which

$$G(i) < 1. \quad (198)$$

Let us call this set of codewords as C_X^{ex} . Without loss of generality, we assume C_X^{ex} contains the first \hat{M}_X sequences of C_X^r , i.e., $C_X^{ex} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{\hat{M}_X}\}$. Consider the multiuser code $C_1^{ex} \triangleq C_X^{ex} \times C_Y$. By definition of L_α , $\alpha = U, X, Y, XY$,

$$L_\alpha(C_1^{ex}, V, i, j) \leq L_\alpha(C^r, V, i, j) \quad \forall (\mathbf{x}_i, \mathbf{y}_j) \in C_1^{ex}. \quad (199)$$

By combining (198) and (199), we conclude that for all $i \in \{1, 2, \dots, \hat{M}_X\}$

$$\begin{aligned} \frac{1}{M_Y} \sum_{j=1}^{M_Y} \sum_{V_{UXY\tilde{X}\tilde{Y}}} \left\{ L_U(C_1^{ex}, V_{UXY}, i, j) 2^{n[F_U(V_{UXY})-6\delta]} \right. \\ + L_X(C_1^{ex}, V_{UXY\tilde{X}}, i, j) 2^{n[F_X(V_{UXY\tilde{X}})-6\delta]} \\ + L_X(C_1^{ex}, V_{UXY\tilde{Y}}, i, j) 2^{n[F_Y(V_{UXY\tilde{Y}})-6\delta]} \\ \left. + L_{XY}(C_1^{ex}, V_{UXY\tilde{X}\tilde{Y}}, i, j) 2^{n[F_{XY}(V_{UXY\tilde{X}\tilde{Y}})-6\delta]} \right\} < 1, \end{aligned} \quad (200)$$

which results in

$$\begin{aligned} \sum_{j=1}^{M_Y} \sum_{V_{UXY\tilde{X}\tilde{Y}}} \left\{ L_U(C_1^{ex}, V_{UXY}, i, j) 2^{n[F_U(V_{UXY})-R_Y-6\delta]} \right. \\ + L_X(C_1^{ex}, V_{UXY\tilde{X}}, i, j) 2^{n[F_X(V_{UXY\tilde{X}})-R_Y-6\delta]} \\ + L_X(C_1^{ex}, V_{UXY\tilde{Y}}, i, j) 2^{n[F_Y(V_{UXY\tilde{Y}})-R_Y-6\delta]} \\ \left. + L_{XY}(C_1^{ex}, V_{UXY\tilde{X}\tilde{Y}}, i, j) 2^{n[F_{XY}(V_{UXY\tilde{X}\tilde{Y}})-R_Y-6\delta]} \right\} < 1. \end{aligned} \quad (201)$$

Since all terms in the summation are non-negative, we conclude that

$$L_\alpha(C_1^{ex}, V, i, j) 2^{-n[F_\alpha(V)-R_Y-6\delta]} < 1 \quad (202)$$

for all $i \in \{1, 2, \dots, \hat{M}_X\}$, $j \in \{1, 2, \dots, M_Y\}$, all $V \in \mathcal{P}(\mathcal{U} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{X} \times \mathcal{Y})$, and all $\alpha = U, X, Y, XY$.

Therefore,

$$L_\alpha(C_1^{ex}, V, i, j) < 2^{-n[F_\alpha(V) - R_Y - 6\delta]}. \quad (203)$$

On the other hand, (191) can also be written as

$$\Pi(C^r) = \frac{1}{M_Y} \sum_{j=1}^{M_Y} H(j), \quad \text{for } \alpha = U, X, Y, XY, \quad (204)$$

where $H(j)$ is defined as

$$\begin{aligned} H(j) \triangleq & \frac{1}{M_X} \sum_{i=1}^{M_X} \sum_{V_{UXY\bar{X}\bar{Y}}} \left\{ L_U(C^r, V_{UXY}, i, j) 2^{n[F_U(V_{UXY}) - 6\delta]} \right. \\ & + L_X(C^r, V_{UXY\bar{X}}, i, j) 2^{n[F_X(V_{UXY\bar{X}}) - 6\delta]} \\ & + L_X(C^r, V_{UXY\bar{Y}}, i, j) 2^{n[F_X(V_{UXY\bar{Y}}) - 6\delta]} \\ & \left. + L_{XY}(C^r, V_{UXY\bar{X}\bar{Y}}, i, j) 2^{n[F_{XY}(V_{UXY\bar{X}\bar{Y}}) - 6\delta]} \right\}. \quad (205) \end{aligned}$$

By a similar argument as we did before, we can show that there exist $\hat{M}_Y \geq \frac{M_Y}{2}$ codewords, $\mathbf{y}_j \in C_Y^r$, for which

$$H(j) < 1. \quad (206)$$

Let us call this set of codewords as C_Y^{ex} . Without loss of generality, we assume C_Y^{ex} contains the first \hat{M}_Y sequences of C_Y^r , i.e., $C_Y^{ex} = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_{\hat{M}_Y}\}$. Consider the multiuser code $C_2^{ex} \triangleq C_X \times C_Y^{ex}$. By definition of L_α , $\alpha = U, X, Y, XY$, we have

$$L_\alpha(C_2^{ex}, V, i, j) \leq L_\alpha(C^r, V, i, j) \quad \forall (\mathbf{x}_i, \mathbf{y}_j) \in C_2^{ex}. \quad (207)$$

By a similar argument as we did before, we can show that

$$L_\alpha(C_2^{ex}, V, i, j) < 2^{-n[F_\alpha(V) - R_X - 6\delta]}. \quad (208)$$

for all $i \in \{1, 2, \dots, M_X\}$, $j \in \{1, 2, \dots, \hat{M}_Y\}$, all $V \in \mathcal{P}(\mathcal{U} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{X} \times \mathcal{Y})$, and all $\alpha = U, X, Y, XY$.

By combining (203) and (208), we conclude that, there exists a multiuser code $C^{ex} = C_X^* \times C_Y^*$ with $M_X^* \times M_Y^*$ messages

$$M_X^* \geq \frac{2^{n(R_X - \delta)}}{2}, \quad M_Y^* \geq \frac{2^{n(R_Y - \delta)}}{2}, \quad M_X^* \times M_Y^* \geq \frac{2^{n(R_X + R_Y - 2\delta)}}{2} \quad (209)$$

such that for any pair of messages $(\mathbf{x}_i, \mathbf{y}_j) \in C^{ex}$, all $V \in \mathcal{P}(\mathcal{U} \times \mathcal{X} \times \mathcal{Y} \times \mathcal{X} \times \mathcal{Y})$, and all $\alpha = U, X, Y, XY$,

$$L_\alpha(C^{ex}, V, i, j) < 2^{-n[F_\alpha(V) - \min\{R_X, R_Y\} - 6\delta]}. \quad (210)$$

It is easy to check that

$$\Pi(C^{ex}) \leq 2 \times \Pi(C^r) < 1, \quad (211)$$

therefore, C^{ex} , satisfies all the constraints in (80a)-(80d).

Here, by method of expurgation, we end up with a code with a similar average bound as we had for the original code. However, all pairs of codewords in the new code also satisfy (81a)-(81d). Therefore,

we did not lose anything in terms of average performance, however, as we will see in Theorem 2, we would end up with a tighter bound since we have more constraints on any particular pair of codewords in our codebook pair. \blacksquare

Proof: (Theorem 2) Let us do random coding. Fix any \mathcal{U} , $P_{XYU} \in \mathcal{P}_n(\mathcal{X} \times \mathcal{Y} \times \mathcal{U})$ such that $X - U - Y$, $R_X \geq 0$, $R_Y \geq 0$, $\delta > 0$, and $\mathbf{u} \in T_{P_U}$. Define M_X, M_Y such that

$$2^{n(R_X - \delta)} \leq M_X \leq 2^{nR_X} \quad 2^{n(R_Y - \delta)} \leq M_Y \leq 2^{nR_Y}$$

Let $X^{M_X} \triangleq (X_1, X_2, \dots, X_{M_X})$ and $Y^{M_Y} \triangleq (Y_1, Y_2, \dots, Y_{M_Y})$ be independent random variables, where X_i s are uniformly distributed on $T_{P_{X|U}}(\mathbf{u})$, and Y_j s are uniformly distributed on $T_{P_{Y|U}}(\mathbf{u})$.

Upper bound: By taking expectation over (63), applying Lemma 4, and using the continuity of information measures, we get the desired upper bound.

Lower bound: By taking expectation over (65), applying Lemma 4, we get

$$\begin{aligned} \mathbb{E}e(C, W) &\geq \sum_{\substack{V_{UXY\tilde{X}Z} \\ \in \mathcal{V}_{X,n}^r}} 2^{-n(E_{\tilde{X}}^L + 4\delta)} \left[1 - \sum_{\substack{V_{UXY\tilde{X}Z}: \\ V_{UXY\tilde{X}Z} = V_{UXY\tilde{X}Z}}} 2^{-n(I_V(\hat{X} \wedge XY\tilde{X}Z|U) - R_X - 7\delta)} \right] \\ &+ \sum_{\substack{V_{UXY\tilde{Y}Z} \\ \in \mathcal{V}_{Y,n}^r}} 2^{-n(E_{\tilde{Y}}^L + 4\delta)} \left[1 - \sum_{\substack{V_{UXY\tilde{Y}Z}: \\ V_{UXY\tilde{Y}Z} = V_{UXY\tilde{Y}Z}}} 2^{-n(I_V(\hat{Y} \wedge XY\tilde{Y}Z|U) - R_Y - 7\delta)} \right] \\ &+ \sum_{\substack{V_{UXY\tilde{X}\tilde{Y}Z} \\ \in \mathcal{V}_{XY,n}^r}} 2^{-n(E_{\tilde{X}\tilde{Y}}^L + 4\delta)} \left[1 - \sum_{\substack{V_{UXY\tilde{X}\tilde{Y}Z}: \\ V_{UXY\tilde{X}\tilde{Y}Z} = V_{UXY\tilde{X}\tilde{Y}Z}}} 2^{-n(I_V(\hat{X}\hat{Y} \wedge XY\tilde{X}\tilde{Y}Z|U) - R_X - R_Y - 7\delta)} \right] \end{aligned} \quad (212)$$

Toward further simplification of this expression, we use the following lemma.

Lemma 9.

$$\min_{\substack{V_{UXY\tilde{X}Z}: \\ V_{UXY\tilde{X}Z} = V_{UXY\tilde{X}Z}}} I_V(\hat{X} \wedge XY\tilde{X}Z|U) = I_V(\tilde{X} \wedge XYZ|U) \quad (213)$$

Proof: Note that, for any $V_{UXY\tilde{X}Z}$,

$$I_V(\hat{X} \wedge XY\tilde{X}Z|U) = I_V(\hat{X} \wedge XYZ|U) + I_V(\tilde{X} \wedge \hat{X}|UXYZ), \quad (214)$$

therefore,

$$\min_{\substack{V_{UXY\tilde{X}Z}: \\ V_{UXY\tilde{X}Z} = V_{UXY\tilde{X}Z}}} I_V(\hat{X} \wedge XY\tilde{X}Z|U) \geq I_V(\hat{X} \wedge XYZ|U) = I_V(\tilde{X} \wedge XYZ|U). \quad (215)$$

Now, consider $V_{UX\tilde{X}\hat{Y}Z}^*$ defined as

$$V_{UX\tilde{X}\hat{Y}Z}^*(u, x, y, \tilde{x}, \hat{y}, z) = V_{\tilde{X}|UXYZ}(\tilde{x}|u, x, y, z) V_{\hat{Y}|UXYZ}(\hat{y}|u, x, y, z) V_{XYZ}(u, x, y, z). \quad (216)$$

Note that $V_{UX\tilde{X}\hat{Y}Z}^* = V_{UX\tilde{X}\hat{Y}Z}$, and $\tilde{X} - (U, X, Y, Z) - \hat{X}$. Therefore,

$$I_{V^*}(\hat{X} \wedge XY\tilde{X}Z|U) = I_V(\hat{X} \wedge XYZ|U) = I_V(\tilde{X} \wedge XYZ|U). \quad (217)$$

By combining (215) and (217), the proof is complete. \blacksquare

Using the above lemma, the average probability of error can be bounded from below as

$$\bar{P}_e \geq \sum_{\substack{V_{UXY\tilde{X}Z} \in \mathcal{V}_{X,n}^r \\ I(\tilde{X} \wedge XYZ|U) > R_X + 12\delta}} 2^{-nE_{\tilde{X}}^L} + \sum_{\substack{V_{UXY\tilde{Y}Z} \in \mathcal{V}_{Y,n}^r \\ I(\tilde{Y} \wedge XYZ|U) > R_Y + 12\delta}} 2^{-nE_{\tilde{Y}}^L} + \sum_{\substack{V_{UXY\tilde{X}\tilde{Y}Z} \in \mathcal{V}_{XY,n}^r \\ I_V(\tilde{X}\tilde{Y} \wedge XY|U) + I_V(\tilde{X} \wedge \tilde{Y}|U) > R_X + R_Y + 14\delta}} 2^{-nE_{\tilde{X}\tilde{Y}}^L}$$

Using the continuity argument, the lower bound on the average error probability follows. \blacksquare

Proof: (Theorem 3) As was done in Theorem 1 for the point-to-point case, here, we will obtain higher error exponents for almost all codes by removing certain types from the constraint sets \mathcal{V}_X^r , \mathcal{V}_Y^r and \mathcal{V}_{XY}^r . Let us define the sets of n -types \mathcal{V}_X^t , \mathcal{V}_Y^t and \mathcal{V}_{XY}^t as follows:

$$\mathcal{V}_{X,n}^t \triangleq \left\{ \begin{array}{l} V_{UXY\tilde{X}} : V_{XU} = V_{\tilde{X}U} = P_{XU}, V_{YU} = P_{YU} \\ F_U(V_{UXY}), F_U(V_{U\tilde{X}Y}) \leq R_X + R_Y \\ F_X(V_{UXY\tilde{X}}) \leq R_X + R_Y \end{array} \right\} \quad (218)$$

$$\mathcal{V}_{Y,n}^t \triangleq \left\{ \begin{array}{l} V_{UXY\tilde{Y}} : V_{XU} = P_{XU}, V_{YU} = V_{\tilde{Y}U} = P_{YU} \\ F_U(V_{UXY}), F_U(V_{U\tilde{Y}Y}) \leq R_X + R_Y \\ F_Y(V_{UXY\tilde{Y}}) \leq R_X + R_Y \end{array} \right\} \quad (219)$$

$$\mathcal{V}_{XY,n}^t \triangleq \left\{ \begin{array}{l} V_{UXY\tilde{X}\tilde{Y}} : V_{UXY\tilde{X}}, V_{UX\tilde{Y}\tilde{X}} \in \mathcal{V}_X^t, V_{UXY\tilde{Y}}, V_{U\tilde{X}\tilde{Y}\tilde{Y}} \in \mathcal{V}_Y^t \\ F_{XY}(V_{UXY\tilde{X}\tilde{Y}}), F_{XY}(V_{U\tilde{X}\tilde{Y}\tilde{X}\tilde{Y}}) \leq R_X + R_Y \end{array} \right\} \quad (220)$$

Lemma 10. Let $C = C_X \times C_Y$ be one of the multiuser codes whose existence is asserted in the Typical random coding packing lemma. The following hold:

$$\text{If } V_{UXY\tilde{X}} \in (\mathcal{V}_{X,n}^t)^c \Rightarrow N_X(C, V_{UXY\tilde{X}}) = 0, \quad (221)$$

$$\text{If } V_{UXY\tilde{Y}} \in (\mathcal{V}_{Y,n}^t)^c \Rightarrow N_Y(C, V_{UXY\tilde{Y}}) = 0, \quad (222)$$

$$\text{If } V_{UXY\tilde{X}\tilde{Y}} \in (\mathcal{V}_{XY,n}^t)^c \Rightarrow N_{XY}(C, V_{UXY\tilde{X}\tilde{Y}}) = 0. \quad (223)$$

Proof: Consider $V_{UXY\tilde{X}} \in (\mathcal{V}_{X,n}^t)^c$. If $V_{XU} \neq P_{XU}$ or $V_{\tilde{X}U} \neq P_{XU}$ or $V_{YU} \neq P_{YU}$, it is clear that

$$N_X(C, V_{UXY\tilde{X}}) = 0. \quad (224)$$

Now, let us assume $F_U(V_{UXY}) > R_X + R_Y + 3\delta$. In this case, by using (78a), we conclude that

$$N_U(C, V_{UXY}) < 2^{-n(R_X + R_Y)} \Rightarrow \sum_{i=1}^{M_X} \sum_{j=1}^{M_Y} 1_{T_{V_{UXY}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j) < 1 \Rightarrow \sum_{i=1}^{M_X} \sum_{j=1}^{M_Y} 1_{T_{V_{UXY}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j) = 0, \quad (225)$$

and as a result, $N_U(C, V_{UXY}) = 0$. Now, note that

$$\begin{aligned} N_X(C, V_{UXY\tilde{X}}) &= \frac{1}{M_X M_Y} \sum_{i=1}^{M_X} \sum_{j=1}^{M_Y} \sum_{k \neq i} 1_{T_{V_{UXY\tilde{X}}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j, \mathbf{x}_k) \\ &\leq \frac{1}{M_X M_Y} \sum_{i=1}^{M_X} \sum_{j=1}^{M_Y} \sum_{k \neq i} 1_{T_{V_{UXY}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j) \\ &= 2^{nR_X} N_U(C, V_{UXY}) = 0, \end{aligned} \quad (226)$$

therefore, $N_X(C, V_{UXY\tilde{X}}) = 0$. Similarly, if $F_U(V_{U\tilde{X}Y}) > R_X + R_Y + 3\delta$,

$$N_U(C, V_{U\tilde{X}Y}) < 2^{-n(R_X+R_Y)} \Rightarrow \sum_{i=1}^{M_X} \sum_{j=1}^{M_Y} 1_{T_{V_{U\tilde{X}Y}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j) < 1 \Rightarrow \sum_{i=1}^{M_X} \sum_{j=1}^{M_Y} 1_{T_{V_{U\tilde{X}Y}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j) = 0, \quad (227)$$

and as a result, $N_U(C_X, C_Y, V_{U\tilde{X}Y}) = 0$. Also, note that

$$\begin{aligned} N_X(C, V_{UXY\tilde{X}}) &= \frac{1}{M_X M_Y} \sum_{i=1}^{M_X} \sum_{j=1}^{M_Y} \sum_{k \neq i} 1_{T_{V_{UXY\tilde{X}}}}(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j, \mathbf{x}_k) \\ &\leq \frac{1}{M_X M_Y} \sum_{i=1}^{M_X} \sum_{j=1}^{M_Y} \sum_{k \neq i} 1_{T_{V_{U\tilde{X}Y}}}(\mathbf{u}, \mathbf{x}_k, \mathbf{y}_j) = 0, \end{aligned} \quad (228)$$

therefore, $N_X(C, V_{UXY\tilde{X}}) = 0$. If $F_X(V_{UXY\tilde{X}}) > R_X + R_Y + 5\delta$, by the property of the code derived in Lemma 5, we observe that $N_X(C_X, C_Y, V_{UXY\tilde{X}}) = 0$. Similarly, by doing a similar argument, it can be concluded that

$$\text{If } V_{UXY\tilde{Y}} \in (\mathcal{V}_{Y,n}^t)^c \Rightarrow N_Y(C, V_{UXY\tilde{Y}}) = 0, \quad (229)$$

and

$$\text{If } V_{UXY\tilde{X}\tilde{Y}} \in (\mathcal{V}_{XY,n}^t)^c \Rightarrow N_{XY}(C, V_{UXY\tilde{X}\tilde{Y}}) = 0. \quad (230)$$

■

Upper bound: We will follow the techniques used in Theorem 2 to provide lower and upper bounds on the average probability of error of almost all codes in the random coding ensemble. For this, we will use the results of Lemma 6. Consider any typical two-user code $C = C_X \times C_Y$ whose existence was established in Lemma 5. Applying (63) on C , and using the continuity argument, we conclude that

$$\begin{aligned} e(C, W) &\leq \sum_{V_{UXY\tilde{X}Z} \in \mathcal{V}_{\tilde{X},n}^r \cap \mathcal{V}_{\tilde{X},n}^t} 2^{-n[D(V_{Z|XYU}||W|V_{XYU}) + I_V(X \wedge Y|U) + |I_V(\tilde{X} \wedge XYZ|U) - R_X|^+ - 5\delta]} \\ &+ \sum_{V_{UXY\tilde{Y}Z} \in \mathcal{V}_{\tilde{Y},n}^r \cap \mathcal{V}_{\tilde{Y},n}^t} 2^{-n[D(V_{Z|XYU}||W|V_{XYU}) + I_V(X \wedge Y|U) + |I_V(\tilde{Y} \wedge XYZ|U) - R_Y|^+ - 5\delta]} \\ &+ \sum_{\substack{V_{UXY\tilde{X}\tilde{Y}Z} \\ \in \mathcal{V}_{\tilde{X},n}^r \cap \mathcal{V}_{\tilde{X}\tilde{Y},n}^t}} 2^{-n[D(V_{Z|XYU}||W|V_{XYU}) + I_V(X \wedge Y|U) + |I_V(\tilde{X} \wedge \tilde{Y}|U) + I_V(\tilde{X}\tilde{Y} \wedge XYZ|U) - R_X - R_Y|^+ - 5\delta]} \\ &\leq 2^{-n[E_T(R_X, R_Y, W, P_{UXY}) - 6\delta]} \end{aligned} \quad (231)$$

whenever $n \geq n_1(|\mathcal{Z}|, |\mathcal{X}|, |\mathcal{Y}|, |\mathcal{U}|, \delta)$, where $E_T(R_X, R_Y, W, P_{XYU})$ is defined in the statement of the theorem.

Lower bound: In the following, we obtain a lower bound on the average error probability of code $C = C_X \times C_Y$. Applying (65) on C , then using (a) Lemma 5 and (b) the fact that for $V \notin \mathcal{V}_{\tilde{X},n}^t$, we have

$A_{i,j}^X \geq 0$, and similar such facts about A^Y and A^{XY} , we get

$$\begin{aligned}
e(C, W) \geq & \sum_{\substack{V_{UXY\tilde{X}Z} \\ \in \mathcal{V}_{X,n}^r \cap \mathcal{V}_{X,n}^t}} 2^{-n(E_X^L + 4\delta)} \left| 1 - \sum_{\substack{V_{UXY\tilde{X}Z}: \\ V_{UXY\tilde{X}Z} = V_{UXY\tilde{X}Z}}} 2^{-n(I_V(\tilde{X} \wedge XY \tilde{X} Z | U) - R_X - 7\delta)} \right|^+ \\
& + \sum_{\substack{V_{UXY\tilde{Y}Z} \\ \in \mathcal{V}_{Y,n}^r \cap \mathcal{V}_{Y,n}^t}} 2^{-n(E_Y^L + 4\delta)} \left| 1 - \sum_{\substack{V_{UXY\tilde{Y}Z}: \\ V_{UXY\tilde{Y}Z} = V_{UXY\tilde{Y}Z}}} 2^{-n(I_V(\tilde{Y} \wedge XY \tilde{Y} Z | U) - R_Y - 7\delta)} \right|^+ \\
& + \sum_{\substack{V_{UXY\tilde{X}\tilde{Y}Z} \\ \in \mathcal{V}_{XY,n}^r \cap \mathcal{V}_{XY,n}^t}} 2^{-n(E_{XY}^L + 4\delta)} \left| 1 - \sum_{\substack{V_{UXY\tilde{X}\tilde{Y}Z}: \\ V_{UXY\tilde{X}\tilde{Y}Z} = V_{UXY\tilde{X}\tilde{Y}Z}}} 2^{-n(I_V(\tilde{X}\tilde{Y} \wedge XY \tilde{X}\tilde{Y} Z | U) - R_X - R_Y - 7\delta)} \right|^+ \quad (232)
\end{aligned}$$

This expression can be simplified as follows.

$$\begin{aligned}
e(C, W) \geq & \sum_{\substack{V_{UXY\tilde{X}Z} \in \mathcal{V}_{X,n}^r \cap \mathcal{V}_{X,n}^t \\ I(\tilde{X} \wedge XY Z | U) > R_X + 12\delta}} 2^{-nE_X^L} + \sum_{\substack{V_{UXY\tilde{Y}Z} \in \mathcal{V}_{Y,n}^r \cap \mathcal{V}_{Y,n}^t \\ I(\tilde{Y} \wedge XY Z | U) > R_Y + 12\delta}} 2^{-nE_Y^L} + \sum_{\substack{V_{UXY\tilde{X}\tilde{Y}Z} \in \mathcal{V}_{XY,n}^r \cap \mathcal{V}_{XY,n}^t \\ I_V(\tilde{X}\tilde{Y} \wedge XY | U) + I_V(\tilde{X} \wedge \tilde{Y} | U) > R_X + R_Y + 14\delta}} 2^{-nE_{XY}^L}
\end{aligned}$$

Using the continuity argument, the lower bound on the average error probability follows. \blacksquare

Proof. (Theorem 4) Fix \mathcal{U} , $\mathcal{P}_{XYU} \in \mathcal{P}_n(\mathcal{X} \times \mathcal{Y} \times \mathcal{U})$ with $X - U - Y$, $R_X \geq 0$, $R_Y \geq 0$, $\delta > 0$, and $\mathbf{u} \in T_{P_U}$. Let $C^* = C_X^* \times C_Y^*$ be the multiuser code whose existence is asserted in Lemma 6. Taking into account the given \mathbf{u} , the α -decoding yields the decoding sets

$$D_{ij} = \{\mathbf{z} : \alpha(\mathbf{u}, \mathbf{x}_i, \mathbf{y}_j, \mathbf{z}) \leq \alpha(\mathbf{u}, \mathbf{x}_k, \mathbf{y}_l, \mathbf{z}) \text{ for all } (k, l) \neq (i, j)\}.$$

Let us define the collection of n -types $\mathcal{V}_{X,n}^x$, $\mathcal{V}_{Y,n}^x$ and $\mathcal{V}_{XY,n}^x$ as follows:

$$\mathcal{V}_{X,n}^x \triangleq \left\{ \begin{array}{l} V_{UXY\tilde{X}} : \quad V_{XU} = V_{\tilde{X}U} = P_{XU}, V_{YU} = P_{YU} \\ F_U(V_{UXY}), F_U(V_{U\tilde{X}Y}) \leq \min\{R_X, R_Y\} \\ F_X(V_{UXY\tilde{X}}) \leq \min\{R_X, R_Y\} \end{array} \right\} \quad (233)$$

$$\mathcal{V}_{Y,n}^x \triangleq \left\{ \begin{array}{l} V_{UXY\tilde{Y}} : \quad V_{XU} = P_{XU}, V_{YU} = V_{\tilde{Y}U} = P_{YU} \\ F_U(V_{UXY}), F_U(V_{U\tilde{Y}Y}) \leq \min\{R_X, R_Y\} \\ F_Y(V_{UXY\tilde{Y}}) \leq \min\{R_X, R_Y\} \end{array} \right\} \quad (234)$$

$$\mathcal{V}_{XY,n}^x \triangleq \left\{ \begin{array}{l} V_{UXY\tilde{X}\tilde{Y}} : \quad V_{UXY\tilde{X}}, V_{UX\tilde{Y}\tilde{X}} \in \mathcal{V}_X^x, \quad V_{UXY\tilde{Y}}, V_{U\tilde{X}\tilde{Y}\tilde{Y}} \in \mathcal{V}_Y^x \\ F_{XY}(V_{UXY\tilde{X}\tilde{Y}}), F_{XY}(V_{U\tilde{X}\tilde{Y}\tilde{Y}}) \leq \min\{R_X, R_Y\} \end{array} \right\} \quad (235)$$

Lemma 11. For the multiuser code $C^* = C_X^* \times C_Y^*$, the following holds:

$$\text{If } V_{UXY\tilde{X}} \in (\mathcal{V}_{X,n}^x)^c \Rightarrow N_X(C^*, V_{UXY\tilde{X}}) = 0, \quad (236)$$

$$\text{If } V_{UXY\tilde{Y}} \in (\mathcal{V}_{Y,n}^x)^c \Rightarrow N_Y(C^*, V_{UXY\tilde{Y}}) = 0, \quad (237)$$

$$\text{If } V_{UXY\tilde{X}\tilde{Y}} \in (\mathcal{V}_{XY,n}^x)^c \Rightarrow N_{XY}(C^*, V_{UXY\tilde{X}\tilde{Y}}) = 0. \quad (238)$$

Proof: The proof is very similar to the proof of lemma 10. \blacksquare

The average error probability of C^* can be obtained as follows in a similar way that used in the proof

of Theorem 2 and Theorem 3.

$$\begin{aligned}
e(C^*, W) \leq & \sum_{V_{UXY\tilde{X}Z} \in \mathcal{V}_{X,n}^r \cap \mathcal{V}_{\tilde{X},n}^x} 2^{-n[D(V_Z|_{XYU}||W|_{V_{XYU}})+I_V(X \wedge Y|U)-3\delta]} \\
& + \sum_{V_{UXY\tilde{Y}Z} \in \mathcal{V}_{Y,n}^r \cap \mathcal{V}_{\tilde{Y},n}^x} 2^{-n[D(V_Z|_{XYU}||W|_{V_{XYU}})+I_V(X \wedge Y|U)-3\delta]} \\
& + \sum_{\substack{V_{UXY\tilde{X}\tilde{Y}Z} \\ \in \mathcal{V}_{\tilde{X},n}^r \cap \mathcal{V}_{\tilde{Y},n}^x}} 2^{-n[D(V_Z|_{XYU}||W|_{V_{XYU}})+I_V(X \wedge Y|U)-3\delta]}. \tag{239}
\end{aligned}$$

Now using the continuity argument the statement of the theorem follows. \blacksquare

Proof: (Theorem 5) For any $V_{UXY\tilde{X}Z} \in \mathcal{V}_X^r$,

$$H_V(XY|ZU) \geq H_V(\tilde{X}Y|ZU), \tag{240}$$

therefore, by subtracting $H_V(Y|ZU)$ from both sides of (240), we can conclude that

$$H_V(X|U) - I_V(X \wedge YZ|U) \geq H_V(\tilde{X}|U) - I_V(\tilde{X} \wedge YZ|U), \tag{241}$$

Since $V_{XU} = V_{\tilde{X}U} = P_{XU}$, the last inequality is equivalent to

$$I_V(X \wedge YZ|U) \leq I_V(\tilde{X} \wedge YZ|U). \tag{242}$$

Since $I_V(\tilde{X} \wedge XYZ|U) \geq I_V(\tilde{X} \wedge YZ|U)$, it can be seen that for any $V_{UXY\tilde{X}Z} \in \mathcal{V}_X^r$

$$I_V(\tilde{X} \wedge XYZ|U) \geq I_V(X \wedge YZ|U). \tag{243}$$

Moreover, since

$$\mathcal{V}_X^r \subseteq \{V_{UXY\tilde{X}Z} : V_{UXYZ} \in \mathcal{V}(P_{UXY})\} \tag{244}$$

it can be easily concluded that

$$E_X^r(R_X, R_Y, W, P_{XYU}) \geq E_{r\tilde{X}}^{Liu}(R_X, R_Y, W, P_{XYU}).$$

Similarly, for any $V_{UXY\tilde{Y}Z} \in \mathcal{V}_Y^r$,

$$H_V(XY|ZU) \geq H_V(X\tilde{Y}|ZU). \tag{245}$$

By using the fact that, $V_{YU} = V_{\tilde{Y}U} = P_{YU}$, it can be concluded that

$$I_V(\tilde{Y} \wedge XYZ|U) \geq I_V(Y \wedge XZ|U). \tag{246}$$

Since

$$\mathcal{V}_Y^r \subseteq \{V_{UXY\tilde{Y}Z} : V_{UXYZ} \in \mathcal{V}(P_{UXY})\}, \tag{247}$$

we conclude that

$$E_Y^r(R_X, R_Y, W, P_{XYU}) \geq E_{r\tilde{Y}}^{Liu}(R_X, R_Y, W, P_{XYU}). \tag{248}$$

Similarly, we can conclude that, for any $V_{UXY\tilde{X}\tilde{Y}Z} \in \mathcal{V}_{XY}^r$,

$$I_V(\tilde{X}\tilde{Y} \wedge XYZ|U) + I(\tilde{X} \wedge \tilde{Y}|U) \geq I_V(XY \wedge Z|U) + I(X \wedge Y|U). \tag{249}$$

Since

$$\mathcal{V}_{XY}^r \subseteq \{V_{UXY\tilde{X}\tilde{Y}Z} : V_{UXYZ} \in \mathcal{V}(P_{UXY})\}, \quad (250)$$

it can be concluded that

$$E_{XY}^r(R_X, R_Y, W, P_{XYU}) \geq E_{rXY}^{Liu}(R_X, R_Y, W, P_{XYU}). \quad (251)$$

By combining (A), (248) and (251), we conclude that (96a) holds. Similarly, we can prove that (96b) and (96c) hold. ■

REFERENCES

- [1] P. Elias, "Coding for noisy channels," *IRE Convention Record*, pp. 37–46, 1955.
- [2] R. L. Dobrusin, "Asymptotic bounds of the probability of error for the transmission of messages over a discrete memoryless channel with a symmetric transition probability matrix," *Teor. Veroyatnost. i Primenen*, pp. 283–311, 1962.
- [3] R. M. Fano, *Transmission of Information: A Statistical Theory of Communication*. MIT Press, 1961.
- [4] C. E. Shannon, R. Gallager, and E. Berlekamp, "Lower bounds on error probability for coding on discrete memoryless channels (part i)," *Inform. and Control*, pp. 65–103, 1967.
- [5] C. E. Shannon, R. Gallager, and E. Berlekamp, "Lower bounds on error probability for coding on discrete memoryless channels (part ii)," *Inform. and Control*, pp. 522–552, 1967.
- [6] U. Augustin, "Gedachtnisfreie kannale for diskrete zeit," *Z. Wahrscheinlichkeits theory verw*, pp. 10–61, 1966.
- [7] A. Feinstein, "Error bounds in noisy channels without memory," *IEEE Trans. Information Theory*, vol. 1, pp. 13–14, Sept. 1955.
- [8] R. Gallager, "The random coding bound is tight for the average code," *IEEE Trans. Information Theory*, vol. 23, pp. 244–246, Mar. 1973.
- [9] R. G. Gallager, *Information theory and Reliable Communications*. New York: John Wiley & Sons, 1968.
- [10] R. G. Gallager, *Transmission of Information: A Statistical Theory of Communication*. MIT Press, 1961.
- [11] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: John Wiley & Sons, 1991.
- [12] I. Csiszar and J. Korner, *Information theory: Coding theorems for Discrete memoryless Systems*. 1981.
- [13] A. Barg and D. Forney, "Random codes: Minimum distances and error exponents," *IEEE Trans. Information Theory*, vol. 48, pp. 2568–2573, Sept. 2002.
- [14] R. Ahlswede, "Multi-way communication channels," in *Proc. International Symposium on Information Theory*, 1971.
- [15] H. Liao, "A coding theorem for multiple-access communications," in *Proc. International Symposium on Information Theory*.
- [16] D. Slepian and J. K. Wolf, "A coding theorem for multiple access channels with correlated sources," *bell Syst. tech. J.*, vol. 52, pp. 1037–1076, 1973.
- [17] A. G. Dyachkov, "Random constant composition codes for multiple-access channels," *Probl. of Control and Inform. Theory*, pp. 357–369, 1984.
- [18] R. Gallager, "A perspective on multi-access channels," *IEEE Trans. Information Theory*, vol. 31, pp. 124–142, Mar. 1985.
- [19] J. Pokorney and H. S. Wallmeier, "Random coding bounds and codes produced by permutations for the multiple-access channels," *IEEE Trans. Information Theory*, vol. 31, pp. 741–750, Nov. 1985.
- [20] Y. Liu and B. L. Hughes, "A new universal random coding bound for the multiple-access channels," *IEEE Trans. Information Theory*, vol. 42, pp. 376–386, Mar. 1996.
- [21] E. A. Haroutunian, "Lower bound for the error probability of multiple-access channels," *Problemy Peredachi Informatsii*, vol. 11, pp. 23–36, June 1975.
- [22] A. Nazari, S. S. Pradhan, and A. Anastasopoulos, "A new sphere-packing bound for maximal error exponent for multiple-access channels," in *Proc. International Symposium on Information Theory*, 2008. Online: <http://arxiv.org/abs/0803.3645>.
- [23] A. Nazari, S. S. Pradhan, and A. Anastasopoulos, "New bounds on the maximal error exponent for multiple-access channels," in *Proc. IEEE Int. Symp. Inf. Theory*, July 2009.
- [24] A. Nazari, S. S. Pradhan, and A. Anastasopoulos, "A new upper bound for the average error exponent for discrete memoryless multiple-access channels," in *Online: http://arxiv.org*, Dec 2009.
- [25] I. Csiszar and J. Korner, "Graph decomposition: A new key to coding theorems," *IEEE Trans. Information Theory*, vol. 1, pp. 5–12, Jan. 1981.
- [26] G. Como, "Group codes outperform binary coset codes on non-binary symmetric memoryless channels," *IEEE Trans. Information Theory (Submitted)*.