Error Exponents for Asymmetric Two-User Discrete Memoryless Source-Channel Coding Systems

Yangfan Zhong, *Student Member, IEEE*, Fady Alajaji, *Senior Member, IEEE*, and L. Lorne Campbell, *Life Fellow, IEEE*

***Abstract*—We study the transmission of two discrete memoryless correlated sources, consisting of a common and a private source, over a discrete memoryless multiterminal channel with two transmitters and two receivers. At the transmitter side, the common source is observed by both encoders but the private source can only be accessed by one encoder. At the receiver side, both decoders need to reconstruct the common source, but only one decoder needs to reconstruct the private source. We hence refer to this system by the asymmetric two-user source–channel coding system. We derive a universally achievable lossless joint source–channel coding (JSCC) error exponent pair for the two-user system by using a technique which generalizes Csiszár's type-packing lemma (1980) for the point-to-point (single-user) discrete memoryless source–channel system. We next investigate the largest convergence rate of asymptotic exponential decay of the system (overall) probability of erroneous transmission, i.e., the system JSCC error exponent. We obtain lower and upper bounds for the exponent. As a consequence, we establish a JSCC theorem with single-letter characterization and we show that the separation principle holds for the asymmetric two-user scenario. By introducing common randomization, we also provide a formula for the tandem (separate) source–channel coding error exponent. Numerical examples show that for a large class of systems consisting of two correlated sources and an asymmetric multiple-access channel with additive noise, the JSCC error exponent considerably outperforms the corresponding tandem coding error exponent.**

***Index Terms*—Asymmetric two-user source–channel system, broadcast channel, common and private message, common randomization, discrete memoryless correlated sources, error exponent, multiple-access channel, lossless joint source-channel coding (JSCC), separation principle, tandem coding, type packing lemma.**

I. Introduction

Recently, the study of the error exponent (reliability function) for point-to-point (single-user) source–channel coding systems (with or without memory) has illustrated substantial superiority of joint source–channel coding (JSCC) over the traditional tandem coding (i.e., separate source and channel coding) approach (e.g., [8], [29], [30]). It is of natural interest to study the JSCC error exponent for multiterminal source–channel systems.

Manuscript received March 01, 2008; revised November 28, 2008. Current version published March 18, 2009. This work was supported in part by the Natural Sciences and Engineering Research Council of Canada. The material in this paper was presented in part at the IEEE International Symposium on Information Theory, Nice, France, July 2007.

The authors are with the Department of Mathematics and Statistics, Queen's University, Kingston, ON K7L 3N6, Canada (e-mail: yangfan@mast.queensu.ca; fady@mast.queensu.ca; campbell@mast.queensu.ca).

Communicated by I. Kontoyiannis, Associate Editor for Shannon Theory.
Digital Object Identifier 10.1109/TIT.2009.2013018

In this work, we address the asymmetric two-user source–channel coding system depicted in Fig. 1. Two discrete memoryless correlated source messages $(\boldsymbol{s},\boldsymbol{l}) \in \mathcal{S}^{\tau n} \times \mathcal{L}^{\tau n}$ drawn from a joint distribution $Q_{SL} : \mathcal{S} \times \mathcal{L}$, consisting of a common source messages $\boldsymbol{s}$ and a private source message $\boldsymbol{l}$ of length τn, are transmitted over a discrete memoryless asymmetric communication channel described by $W_{YZ\,|\,UX} : \mathcal{U} \times \mathcal{X} \to \mathcal{Y} \times \mathcal{Z}$ with block codes of length n, where $\tau > 0$ (measured in source symbol/channel use) is the overall transmission rate. The common source can be accessed by both encoders, but the private source can only be observed by one encoder (say, Encoder 1). In this setup, the goal is to send the common information to both receivers, and send the private information to only one receiver (say, Decoder 1).

This asymmetric two-user system can be used to model [23] interference channels with cognitive radio, an emerging and promising wireless technology where wireless systems, equipped with flexible software, dynamically adapt to their environment (by, for example, adjusting the modulation format or the coding scheme) to harness unemployed spectral capabilities [25]–[27], [12], [13]. For example, it can model the practical situation where audio and video signals are modulated and transmitted to two receivers over a cognitive interference channel (without secrecy constraints) [23], with the cognitive receiver needing to decode both audio and video signals while the noncognitive receiver needing to only reconstruct the audio signal. Furthermore, it is worthy to point out that the asymmetric two-user system is a generalization of the following two classical asymmetric multiterminal scenarios which have been extensively studied in the literature.

i) The CS-AMAC system: If we remove Decoder 2 from Fig. 1, and let $|\mathcal{Z}| = 1$, then the channel reduces to a multiple-access channel $W_{Y\,|\,UX}$, and the coding problem reduces to transmitting two correlated sources (CS) over an asymmetric multiple-access channel (AMAC) with one receiver.
ii) The CS-ABC system: If we remove Encoder 2 from Fig. 1, and let $|\mathcal{U}| = 1$, then the channel reduces to a broadcast channel $W_{YZ\,|\,X}$, and the coding problem reduces to transmitting two CS over an asymmetric broadcast channel (ABC) with one transmitter.

The sufficient and necessary condition for the reliable transmission of CS over the AMAC—i.e., the lossless JSCC theorem for the CS-AMAC system—has been derived with single-letter characterization in [4]. The capacity region of the ABC has been determined in [21], and the JSCC theorem for CS-ABC system with arbitrary transmission rate can also be analogously carried

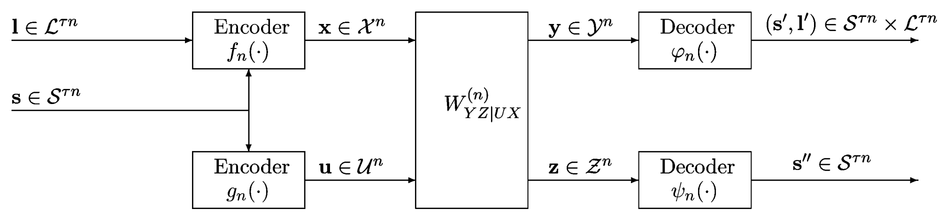

Fig. 1. Transmitting two CS over the asymmetric two-user communication channel.

out (e.g., [17]). In this work, we study a refined version of the JSCC theorem for the general asymmetric two-user system (depicted in Fig. 1), by investigating the achievable JSCC error exponent pair (for two receivers) as well as the system JSCC error exponent, i.e., the largest convergence rate of asymptotic exponential decay of the system (overall) probability of erroneous transmission. We also apply our results to the CS-AMAC and CS-ABC systems.

We outline our results as follows. We first extend Csiszár's type packing lemma [8] from a single-letter (one-dimension) type setting to a joint (two-dimensional) type setting. By employing the joint type packing lemma and generalized maximum mutual information (MMI) decoders, we establish achievable exponential upper bounds for the probabilities of erroneous transmission over an augmented two-user channel $W_{YZ\,|\,TUX}$ for a given triple of n-length sequences $(\boldsymbol{t}, \boldsymbol{u}, \boldsymbol{x})$; see Theorem 1. Here, the augmented channel $W_{YZ\,|\,TUX}$ is induced from the original two-user channel $W_{YZ\,|\,UX}$ by adding an auxiliary random variable (RV) T such that T, (UX), and (YZ), form a Markov chain in this order. We introduce the RV T because we will employ superposition encoding which maps a source message pair $(\boldsymbol{s}, \boldsymbol{l})$ to a codeword triplet $(\boldsymbol{t}, \boldsymbol{u}, \boldsymbol{x})$, where $\boldsymbol{t}$ is the auxiliary superposition codeword. For the asymmetric two-user system, since one of the encoders has full access to both sources, it knows the output of the other encoder. By properly designing the two (superposition) encoders, we apply Theorem 1 to establish a universally achievable error exponent pair for the two receivers (namely, the pair of exponents can be achieved by a sequence of source–channel codes independently of the statistics of the source and the channel); this generalizes Körner and Sgarro's exponent pair for ABC coding (with uniformly distributed message sets) [22]. We also employ Theorem 1 to establish a lower bound for the system JSCC error exponent; see Theorem 2. Note that one consequence of our results is a sufficient condition (forward part) for the JSCC theorem. In addition, we use Fano's inequality to prove a necessary condition (converse part) which coincides with the sufficient condition, and hence completes the JSCC theorem (Theorem 3). We next demonstrate that the separation principle holds for the two-user system, i.e., there exists a separate source and channel coding system which can achieve optimality from the point of view of reliable transmissibility.

Using an approach analogous to [8], we also obtain an upper bound for the system JSCC error exponent (Theorem 4). As applications, we then specialize these results to the CS-AMAC and CS-ABC systems. The computation of the lower and upper bounds for the system JSCC error exponent is partially studied for the CS-AMAC system when the channel admits a symmetric conditional distribution.

We next study the tandem coding error exponent for the asymmetric two-user system, which is the exponent resulting from separate and independent source and channel coding under common randomization. We derive a formula for the tandem coding error exponent in terms of the corresponding two-user source error exponent and the asymmetric two-user channel error exponent (Theorem 6). Finally, by numerically comparing the lower bound of the JSCC error exponent and the upper bound of the tandem coding error exponent, we illustrate that, as for the point-to-point systems ([29], [30]), JSCC can considerably outperform tandem coding in terms of error exponent for a large class of binary CS-AMAC systems with additive noise.

At this point, we pause to mention some related works in the literature on the multiterminal JSCC of CS. The JSCC theorem for transmitting two CS over a (symmetric) multiple-access channel (where each encoder can only access one source) has been studied in [1], [7], [14], [19], [20], [28], and the JSCC theorem for transmitting two CS over a (symmetric) broadcast channel (where each decoder needs to reconstruct one source) has been addressed in [5], [17]. These works focus on the case when the overall transmission rate τ is 1 and establish some sufficient and/or necessary conditions for which the sources can be reliably transmitted over the channel. However, for both (symmetric) systems, no matter whether the transmission rate τ is 1 or not, a tight sufficient and necessary condition (JSCC theorem) with single-letter characterization is still unknown.

The rest of the paper is organized as follows. In Section II, we introduce the notation and some basic facts regarding the method of types. A generalized joint type packing lemma is presented in Section III. In Section IV, we establish a universally achievable error exponent pair for the two-user system, as well as a lower and an upper bound for the system JSCC error exponent. A JSCC theorem with single-letter characterization is also given and we demonstrate that the reliable transmissibility condition can be achieved by separately performing source and channel coding. In Section V, we apply our results to the CS-AMAC and CS-ABC systems. We partially address the computation of the bounds for the system JSCC error exponent in Section VI. In Section VII, we provide an expression for the tandem coding error exponent for the two-user system and we then show that the JSCC error exponent can be strictly larger than the tandem coding error exponent for many CS-AMAC systems. Finally, we state our conclusions in Section VIII.

II. Preliminaries

The following notation and conventions are adopted from [8]–[10]. For any finite set (or alphabet) $\mathcal{X}$, the size of $\mathcal{X}$ is denoted by $|\mathcal{X}|$. The set of all probability distributions on $\mathcal{X}$

is denoted by $\mathcal{P}(\mathcal{X})$. The type of an n-length sequence $\boldsymbol{x} \triangleq (x_1, x_2, \ldots, x_n) \in \mathcal{X}^n$ is the empirical probability distribution $P_{\boldsymbol{x}} \in \mathcal{P}(\mathcal{X})$ defined by

$$P_{\boldsymbol{x}}(a) \triangleq \frac{1}{n} N(a|\boldsymbol{x}), \quad a \in \mathcal{X} \tag{1}$$

where $N(a\,|\,\boldsymbol{x})$ is the number of occurrences of a in $\boldsymbol{x}$. Let $\mathcal{P}_n(\mathcal{X}) \subseteq \mathcal{P}(\mathcal{X})$ be the collection of all types of sequences in $\mathcal{X}^n$. For any $P_X \in \mathcal{P}_n(\mathcal{X})$, the set of all $\boldsymbol{x} \in \mathcal{X}^n$ with type P_X is denoted by $\mathbb{T}_{P_X}$, or simply by $\mathbb{T}_X$ if P_X is understood. We also call $\mathbb{T}_{P_X}$ or $\mathbb{T}_X$ a type class.

Similarly, the joint type of n-length sequences $\boldsymbol{x} \in \mathcal{X}^n$ and $\boldsymbol{y} \triangleq (y_1, y_2, \ldots, y_n) \in \mathcal{Y}^n$ is the empirical joint probability distribution $P_{\boldsymbol{xy}} \in \mathcal{P}(\mathcal{X} \times \mathcal{Y})$ defined by

$$P_{\boldsymbol{xy}}(a, b) \triangleq \frac{1}{n} N(a, b|\boldsymbol{x}, \boldsymbol{y}), \quad (a, b) \in \mathcal{X} \times \mathcal{Y}.$$

Let $\mathcal{P}_n(\mathcal{X} \times \mathcal{Y}) \subseteq \mathcal{P}(\mathcal{X} \times \mathcal{Y})$ be the collection of all joint types of sequences in $\mathcal{X}^n \times \mathcal{Y}^n$. The set of all $\boldsymbol{x} \in \mathcal{X}^n$ and $\boldsymbol{y} \in \mathcal{Y}^n$ with joint type $P_{XY} \in \mathcal{P}_n(\mathcal{X} \times \mathcal{Y})$ is denoted by $\mathbb{T}_{P_{XY}}$, or simply by $\mathbb{T}_{XY}$.

For any finite sets $\mathcal{X}$ and $\mathcal{Y}$, the set of all conditional distributions $V_{Y\,|\,X} : \mathcal{X} \to \mathcal{Y}$ is denoted by $\mathcal{P}(\mathcal{Y}\,|\,\mathcal{X})$. The conditional type of $\boldsymbol{y} \in \mathcal{Y}^n$ given $\boldsymbol{x} \in \mathbb{T}_{P_X}$ is the empirical conditional probability distribution $P_{\boldsymbol{y}|\boldsymbol{x}} \in \mathcal{P}(\mathcal{Y}|\mathcal{X})$ defined by

$$P_{\boldsymbol{y}\,|\,\boldsymbol{x}}(b\,|\,a) = \frac{N(a, b\,|\,\boldsymbol{x}, \boldsymbol{y})}{N(a\,|\,\boldsymbol{x})}$$

whenever $N(a|\boldsymbol{x}) > 0$; otherwise (if $N(a\,|\,\boldsymbol{x}) = 0$), define $P_{\boldsymbol{y}\,|\,\boldsymbol{x}}(b\,|\,a) = 0$, $(a, b) \in \mathcal{X} \times \mathcal{Y}$.

Let $\mathcal{P}_n(\mathcal{Y}\,|\,P_X)$ be the collection of all conditional distributions $V_{Y\,|\,X}$ which are conditional types of $\boldsymbol{y} \in \mathcal{Y}^n$ given an $\boldsymbol{x} \in \mathbb{T}_{P_X}$. For any conditional type $V_{Y\,|\,X} \in \mathcal{P}_n(\mathcal{Y}\,|\,P_X)$, the set of all $\boldsymbol{y} \in \mathcal{Y}^n$ for a given $\boldsymbol{x} \in \mathbb{T}_{P_X}$ satisfying $P_{\boldsymbol{y}\,|\,\boldsymbol{x}} = V_{Y\,|\,X}$ is denoted by $\mathbb{T}_{V_{Y\,|\,X}}(\boldsymbol{x})$, or simply by $\mathbb{T}_{Y\,|\,X}(\boldsymbol{x})$, which is also called a conditional type class (V-shell) with respect to $\boldsymbol{x}$.

For finite sets $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Z}$ with joint distribution $P_{XYZ} \in \mathcal{P}(\mathcal{X} \times \mathcal{Y} \times \mathcal{Z})$, we use $P_X, P_{XY}, P_{YZ\,|\,X}$, etc., to denote the corresponding marginal and conditional probabilities induced by P_{XYZ}. Note that for a given joint type $P_{XY} \in \mathcal{P}_n(\mathcal{X} \times \mathcal{Y})$, $\mathbb{T}_{P_{Y\,|\,X}}(\boldsymbol{x}) = \{\boldsymbol{y} : (\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{T}_{P_{XY}}\}$. Note also that

$$\{P_X V_{Y\,|\,X} : P_X \in \mathcal{P}_n(\mathcal{X}), V_{Y\,|\,X} \in \mathcal{P}_n(\mathcal{Y}\,|\,P_X)\} = \mathcal{P}_n(\mathcal{X} \times \mathcal{Y}).$$

In addition, we denote

$$\mathcal{P}_n(\mathcal{Y}\,|\,\mathcal{X}) \triangleq \bigcup_{P_X \in \mathcal{P}_n(\mathcal{X})} \mathcal{P}_n(\mathcal{Y}\,|\,P_X) \subseteq \mathcal{P}(\mathcal{Y}\,|\,\mathcal{X}).$$

To distinguish between different distributions (or types) defined on the same alphabet, we use sub-subscripts, say, i, j, in P_{X_i}, $P_{X_iY_j}$, $\mathbb{T}_{X_iY_j}$, and so on. For example, $\mathbb{T}_{X_iY_j}$ is the type class of the joint type $P_{X_iY_j} \in \mathcal{P}_n(\mathcal{X} \times \mathcal{Y})$. For any distribution $P_{XYZ} \in \mathcal{P}(\mathcal{X} \times \mathcal{Y} \times \mathcal{Z})$, we use $H_{P_{XYZ}}(\cdot)$ and $I_{P_{XYZ}}(\cdot;\cdot)$ to denote the entropy and mutual information under P_{XYZ}, respectively, or simply by $H(\cdot)$ and $I(\cdot;\cdot)$ if P_{XYZ} is understood. $D(P_X \parallel Q_X)$ denotes the Kullback–Leibler divergence between distributions $P_X, Q_X \in \mathcal{P}(\mathcal{X})$. $D(V_{Y\,|\,X} \parallel W_{Y\,|\,X}\,|\,P_X)$ denotes the Kullback–Leibler divergence between stochastic matrices (conditional distributions) $V_{Y\,|\,X}, W_{Y\,|\,X} \in \mathcal{P}(\mathcal{Y}\,|\,\mathcal{X})$ conditional on distribution $P_X \in \mathcal{P}(\mathcal{X})$. For $\boldsymbol{x} \in \mathcal{X}^n$, $\boldsymbol{y} \in \mathcal{Y}^n$, and $\boldsymbol{z} \in \mathcal{Z}^n$, since the types $P_{\boldsymbol{x}}$, $P_{\boldsymbol{xy}}$, and $P_{\boldsymbol{xyz}}$ can also be represented as distributions of dummy RVs, we define the empirical entropy and mutual information by $H(\boldsymbol{x}) \triangleq H_{P_{\boldsymbol{x}}}(X)$, $I(\boldsymbol{x};\boldsymbol{y}) \triangleq I_{P_{\boldsymbol{xy}}}(X;Y)$, and $I(\boldsymbol{x};\boldsymbol{y}\,|\,\boldsymbol{z}) \triangleq I_{P_{\boldsymbol{xyz}}}(X;Y\,|\,Z)$. Given distributions $P_X \in \mathcal{P}(\mathcal{X})$ and $W_{Y\,|\,X} \in \mathcal{P}(\mathcal{Y}\,|\,\mathcal{X})$, let $P_X^{(n)}$ and $W_{Y\,|\,X}^{(n)}$ be their n-dimensional product distributions. All logarithms and exponentials throughout this paper are in base 2. The following facts will be frequently used throughout this paper.

Lemma 1 [10]:

i) $|\mathcal{P}_n(\mathcal{X})| \leq (n+1)^{|\mathcal{X}|}, |\mathcal{P}_n(\mathcal{Y}|\mathcal{X})| \leq (n+1)^{|\mathcal{Y}||\mathcal{X}|}$.

ii) For any $P_X, Q_X \in \mathcal{P}_n(\mathcal{X})$, we have

$$(n+1)^{-|\mathcal{X}|} 2^{nH_{P_X}(X)} \leq |\mathbb{T}_{P_X}| \leq 2^{nH_{P_X}(X)}$$

and

$$(n+1)^{-|\mathcal{X}|} 2^{-nD(P_X \| Q_X)} \leq Q_X^{(n)}(\mathbb{T}_{P_X}) \leq 2^{-nD(P_X \| Q_X)}.$$

iii) For any $\boldsymbol{x} \in \mathbb{T}_{P_X}$, $\boldsymbol{y} \in \mathbb{T}_{V_{Y|X}}(\boldsymbol{x})$ and $W_{Y|X}, V_{Y|X} \in \mathcal{P}_n(\mathcal{Y}|P_X)$, we have

$$\begin{aligned}&(n+1)^{-|\mathcal{X}||\mathcal{Y}|} 2^{nH_{P_X V_{Y|X}}(Y|X)} \\ &\quad \leq |\mathbb{T}_{V_{Y|X}}(\boldsymbol{x})| \leq 2^{nH_{P_X V_{Y|X}}(Y|X)} \\ &W_{Y|X}^{(n)}(\boldsymbol{y}|\boldsymbol{x}) \\ &\quad = 2^{-n[D(V_{Y|X} \| W_{Y|X} | P_X) + H_{P_X V_{Y|X}}(Y|X)]}\end{aligned}$$

and hence

$$\begin{aligned}&(n+1)^{-|\mathcal{X}||\mathcal{Y}|} 2^{-nD(V_{Y|X} \| W_{Y|X} | P_X)} \\ &\quad \leq W_{Y|X}^{(n)}(\mathbb{T}_{V_{Y|X}}(\boldsymbol{x})|\boldsymbol{x}) \leq 2^{-nD(V_{Y|X} \| W_{Y|X} | P_X)}.\end{aligned}$$

III. A Joint Type Packing Lemma

Let us first recall Csiszár's type packing lemma for JSCC [8], which is an essential tool to establish an exponentially achievable upper bound for the JSCC probability of error over a discrete memoryless channel.

Lemma 2 [6, Lemma 6]: Given finite set $\mathcal{A}$ and a sequence of positive integers $\{m_n\}$, for arbitrary (not necessarily distinct) types $P_{A_i} \in \mathcal{P}_n(\mathcal{A})$, and positive integers N_i, $i = 1, 2, \ldots, m_n$ with

$$\frac{1}{n} \log_2 N_i < H_{P_{A_i}}(A) - \delta \tag{1}$$

where

$$\delta \triangleq \frac{2}{n}[|\mathcal{A}|^2 \log_2(n+1) + \log_2 m_n + 1]$$

there exist m_n disjoint subsets

$$\Omega_i = \left\{\boldsymbol{a}_p^{(i)}\right\}_{p=1}^{N_i} \subseteq \mathbb{T}_{A_i} \triangleq \mathbb{T}_{P_{A_i}}$$

such that

$$\left|\mathbb{T}_{V_{A'\,|\,A}}\left(\boldsymbol{a}_p^{(i)}\right)\bigcap\Omega_k\right| \le N_k 2^{-n[I_{P_{A_i}V_{A'\,|\,A}}(A;A')-\delta]} \tag{2}$$

for every i,k,p and $V_{A'\,|\,A}\in\mathcal{P}_n(\mathcal{A}|\mathcal{A})$, with the exception of the case when both $i=k$ and $V_{A'\,|\,A}$ is the conditional distribution such that $V_{A'\,|\,A}(a'|a)$ is 1 if $a'=a$ and 0 otherwise.

Note that Lemma 2 is a generalization of the packing lemma in [10, p. 162, Lemma 5.1], where the later one is used for channel coding, while Lemma 2 is used for JSCC. Roughly and intuitively, if $\boldsymbol{a}$ is a transmitted codeword, then the possible sequences decoded as $\boldsymbol{a}$ can be seen as elements in the "sphere" $\mathbb{T}_{V_{A'\,|\,A}}(\boldsymbol{a})$ "centered" at $\boldsymbol{a}$ for some $V_{A'\,|\,A}$. Equation (2) in the packing lemma states that there exist disjoint sets Ω_k with bounded cardinalities such that the size of the intersection between the sphere $\mathbb{T}_{V_{A'\,|\,A}}(\boldsymbol{a})$ for every $\boldsymbol{a}\in\Omega_i$ and every set Ω_k is "exponentially small" compared with the size of each Ω_k. So the packing lemma can be used to prove the existence of good codes that have an exponentially small probability of error.

We herein extend Csiszár's above type packing lemma from the (one-dimensional) single-letter type setting to a (two-dimensional) joint type setting. This lemma will play a key role in establishing an exponentially achievable upper bound (in Theorem 1) for the probability of erroneous transmission for our asymmetric two-user source–channel system.

Lemma 3 (Joint Type Packing Lemma): Given finite sets $\mathcal{A}$ and $\mathcal{B}$, a sequence of positive integers $\{m_n\}$, and a sequence of positive integers $\{m'_{\text{in}}\}$ associated with every $i=1,2,\ldots,m_n$, for arbitrary (not necessarily distinct) types $P_{A_i}\in\mathcal{P}_n(\mathcal{A})$ and conditional types $P_{B_j|A_i}\in\mathcal{P}_n(\mathcal{B}\,|\,P_{A_i})$, and positive integers N_i and M_{ij}, $i=1,2,\ldots,m_n$, and $j=j(i)=1,2,\ldots,m'_{\text{in}}$ with

$$\frac{1}{n}\log_2 N_i < H_{P_{A_i}}(A)-\delta \tag{3}$$

and

$$\frac{1}{n}\log_2 M_{ij} < H_{P_{A_i}P_{B_j|A_i}}(B\,|\,A)-\delta \tag{4}$$

where

$$\delta \triangleq \frac{2}{n}[|\mathcal{A}|^2|\mathcal{B}|^2\log_2(n+1)+\log_2 m_n + \log_2(\max_i m'_{\text{in}})+\log_2 12]$$

there exist m_n disjoint subsets

$$\Omega_i=\left\{\boldsymbol{a}_p^{(i)}\right\}_{p=1}^{N_i}\subseteq\mathbb{T}_{A_i}$$

such that

$$\left|\mathbb{T}_{V_{A'\,|\,A}}\left(\boldsymbol{a}_p^{(i)}\right)\bigcap\Omega_k\right| \le N_k 2^{-n[I_{P_{A_i}V_{A'\,|\,A}}(A;A')-\delta]} \tag{5}$$

for every i,k,p, and $V_{A'\,|\,A}\in\mathcal{P}_n(\mathcal{A}\,|\,\mathcal{A})$, with the exception of the case when both $i=k$ and $V_{A'\,|\,A}$ is the conditional distribution such that $V_{A'\,|\,A}(a'\,|\,a)$ is 1 if $a'=a$ and 0 otherwise; furthermore, for every $\boldsymbol{a}_p^{(i)}\in\Omega_i$ and every i, there exist m'_{in} disjoint subsets

$$\Omega_{ij}(\boldsymbol{a}_p^{(i)})=\left\{\left(\boldsymbol{a}_p^{(i)},\boldsymbol{b}_{p,q}^{(j)}\right)\right\}_{q=1}^{M_{ij}}$$

such that $\boldsymbol{b}_{p,q}^{(j)}\in\mathbb{T}_{B_j|A_i}(\boldsymbol{a}_p^{(i)})\triangleq\mathbb{T}_{P_{B_j|A_i}}(\boldsymbol{a}_p^{(i)})$ and

$$\left|\mathbb{T}_{V_{A'B'\,|\,AB}}\left(\boldsymbol{a}_p^{(i)},\boldsymbol{b}_{p,q}^{(j)}\right)\bigcap\bigcup_{p'=1}^{N_k}\Omega_{kl}\left(\boldsymbol{a}_{p'}^{(k)}\right)\right| \le N_k M_{kl}2^{-n[I_{P_{A_i B_j}V_{A'B'\,|\,AB}}(A,B;A',B')-\delta]} \tag{6}$$

$$\left|\mathbb{T}_{V_{A'B'\,|\,AB}}\left(\boldsymbol{a}_p^{(i)},\boldsymbol{b}_{p,q}^{(j)}\right)\bigcap\bigcup_{p'=1}^{N_i}\Omega_{il}\left(\boldsymbol{a}_{p'}^{(i)}\right)\right| \le M_{il}2^{-n[I_{P_{A_i B_j}V_{A'B'\,|\,AB}}(B;B'|A)-\delta]} \tag{7}$$

for any i,j,k,l,p,q, and $V_{A'B'\,|\,AB}\in\mathcal{P}_n(\mathcal{A}\times\mathcal{B}|\mathcal{A}\times\mathcal{B})$, with the exception of the case when both $i=k$, $j=l$, and $V_{A'B'\,|\,AB}$ is the conditional distribution such that $V_{A'B'\,|\,AB}(a',b'\,|\,a,b)$ is 1 if $(a',b')=(a,b)$ and 0 otherwise.

The proof of the packing lemma is lengthy and is deferred to Appendix A. Compared with Lemma 2, it is seen that Csiszár's type packing lemma (5) is incorporated in our extended packing lemma, and we emphasize that here we need (6) and (7) to hold in addition to (5).

Similarly, for the two-user channel, if $(\boldsymbol{a},\boldsymbol{b})$ is a pair of transmitted codewords, then the possible sequences decoded as $(\boldsymbol{a},\boldsymbol{b})$ can be seen as elements in the "sphere" $\mathbb{T}_{V_{A'B'\,|\,AB}}(\boldsymbol{a},\boldsymbol{b})$ "centered" at $(\boldsymbol{a},\boldsymbol{b})$ for some $V_{A'B'\,|\,AB}$. As depicted in Fig. 2, (6) (similarly to (7)) states that there exist disjoint sets $\Omega_{kl}=\bigcup_{p'=1}^{N_k}\Omega_{kl}(\boldsymbol{a}_{p'}^{(k)})$ with bounded cardinalities such that the size of the intersection between the sphere $\mathbb{T}_{V_{A'B'\,|\,AB}}(\boldsymbol{a},\boldsymbol{b})$ for every $(\boldsymbol{a},\boldsymbol{b})\in\Omega_{ij}$ and every set Ω_{kl} is "exponentially small" compared with the size of each Ω_{kl}. Note also that the extended packing lemma is analogous to, but different from the one introduced by Körner and Sgarro in [22], which is used to prove a lower bound for the channel coding ABC exponent. Lemma 3 here is used for the asymmetric two-user JSCC problem.

IV. Transmitting CS Over the Asymmetric Two-User Channel

A. System

Let $\{W_{YZ\,|\,UX}:\mathcal{U}\times\mathcal{X}\to\mathcal{Y}\times\mathcal{Z}\}$ be a two-user discrete memoryless channel with finite input alphabet $\mathcal{U}\times\mathcal{X}$, finite output alphabet $\mathcal{Y}\times\mathcal{Z}$, and a transition distribution $W_{YZ\,|\,UX}(y,z|u,x)$ such that the n-tuple transition probability is

$$W_{YZ\,|\,UX}^{(n)}(\boldsymbol{y},\boldsymbol{z}|\boldsymbol{u},\boldsymbol{x})=\prod_{i=1}^{n}W_{YZ\,|\,UX}(y_i,z_i|u_i,x_i)$$

where $u\in\mathcal{U}$, $x\in\mathcal{X}$, $y\in\mathcal{Y}$, $z\in\mathcal{Z}$, $\boldsymbol{u}\triangleq(u_1,\ldots,u_n)\in\mathcal{U}^n$, $\boldsymbol{x}=(x_1,\ldots,,x_n)\in\mathcal{X}^n$, $\boldsymbol{y}=(y_1,\ldots,y_n)\in\mathcal{Y}^n$, and $\boldsymbol{z}\triangleq(z_1,\ldots,z_n)\in\mathcal{Z}^n$. Denote the marginal transition distributions of $W_{YZ\,|\,UX}$ at its Y-output (respectively, Z-output) by $W_{Y\,|\,UX}\triangleq\sum_Z W_{YZ\,|\,UX}$ (respectively, $W_{Z|UX}\triangleq\sum_Y W_{YZ\,|\,UX}$). The marginal distributions of $W_{YZ\,|\,UX}^{(n)}$ are denoted by $W_{Y\,|\,UX}^{(n)}$ and $W_{Z|UX}^{(n)}$, respectively.

Consider two discrete memoryless CS with a generic joint distribution $Q_{SL}(s,l)$ defined on the finite al-

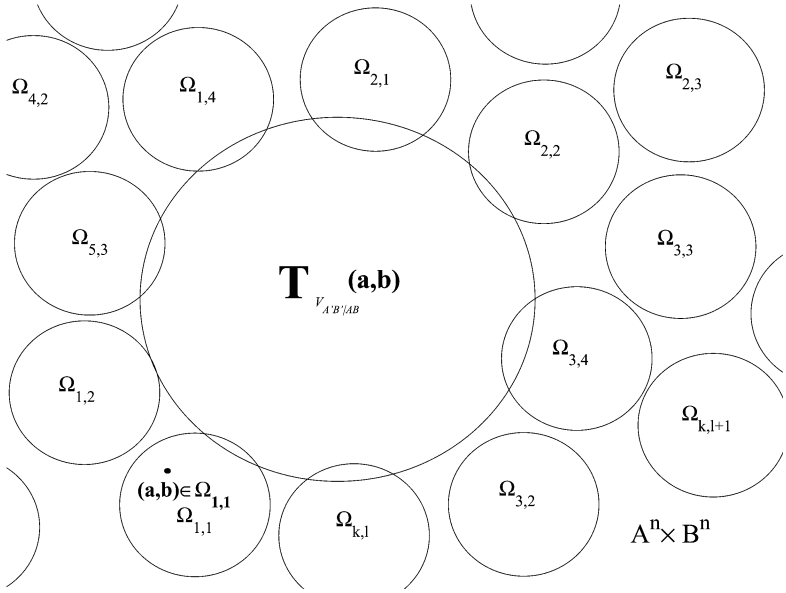

Fig. 2. A graphical illustration of the (two-dimensional) joint type packing lemma (Lemma 3): there exist disjoint subsets Ω_{ij}'s with bounded cardinalities in the "two-dimensional" space $\mathcal{A}^n \times \mathcal{B}^n$ such that for any $(\boldsymbol{a},\boldsymbol{b}) \in \Omega_{ij}$ (say, $(\boldsymbol{a},\boldsymbol{b}) \in \Omega_{1,1}$), the size of the intersection between the sphere $\mathbb{T}_{V_{A'B'\mid AB}}(\boldsymbol{a},\boldsymbol{b})$ and every set Ω_{kl} is "exponentially small" compared with the size of each Ω_{kl}.

phabet $\mathcal{S} \times \mathcal{L}$ such that the k-tuple joint distribution is $Q_{SL}^{(k)}(\boldsymbol{s},\boldsymbol{l}) = \prod_{i=1}^{k} Q_{SL}(s_i,l_i)$, where $(s,l) \in \mathcal{S}\times\mathcal{L}$, and $(\boldsymbol{s},\boldsymbol{l}) \triangleq ((s_1,l_1),\ldots,(s_k,l_k)) \in \mathcal{S}^k \times \mathcal{L}^k$. For each pair of source messages $(\boldsymbol{s},\boldsymbol{l})$ drawn from the above joint distribution, we need to transmit the *common message* $\boldsymbol{s}$ over the channel $W_{YZ\mid UX}$ to Receivers Y and Z and transmit the *private message* $\boldsymbol{l}$ only to Receiver Y. A joint source–channel (JSC) code with block length n and positive transmission rate τ (source symbol/channel use) for transmitting Q_{SL} through $W_{YZ\mid UX}$ is a quadruple of mappings, $(f_n, g_n, \varphi_n, \psi_n)$, where $f_n : \mathcal{S}^{\tau n} \times \mathcal{L}^{\tau n} \to \mathcal{X}^n$ and $g_n : \mathcal{S}^{\tau n} \to \mathcal{U}^n$ are called encoders, and $\varphi_n : \mathcal{Y}^n \to \mathcal{S}^{\tau n} \times \mathcal{L}^{\tau n}$ and $\psi_n : \mathcal{Z}^n \to \mathcal{S}^{\tau n}$ are referred to as Y-decoder and Z-decoder, respectively; see Fig. 1.

The probabilities of Y- and Z-error are given by

$$
\begin{aligned}
&P_{Ye}^{(n)}(Q_{SL}, W_{YZ\mid UX}, \tau)\\
&\quad\triangleq \Pr(\{\varphi_n(Y^n) \neq (S^{\tau n}, L^{\tau n})\})\\
&\quad= \sum_{\mathcal{S}^{\tau n}\times\mathcal{L}^{\tau n}} Q_{SL}^{(\tau n)}(\boldsymbol{s},\boldsymbol{l}) \sum_{\boldsymbol{y}:\varphi_n(\boldsymbol{y})\neq(\boldsymbol{s},\boldsymbol{l})} W_{Y\mid UX}^{(n)}(\boldsymbol{y}|\boldsymbol{u},\boldsymbol{x})
\end{aligned}
\tag{8}
$$

and

$$
\begin{aligned}
&P_{Ze}^{(n)}(Q_{SL}, W_{YZ\mid UX}, \tau)\\
&\quad\triangleq \Pr(\{\psi_n(Z^n) \neq S^{\tau n}\})\\
&\quad= \sum_{\mathcal{S}^{\tau n}} Q_{S}^{(\tau n)}(\boldsymbol{s}) \sum_{\boldsymbol{z}:\psi_n(\boldsymbol{z})\neq\boldsymbol{s}} W_{Z|UX}^{(n)}(\boldsymbol{z}|\boldsymbol{u},\boldsymbol{x})
\end{aligned}
\tag{9}
$$

where $\boldsymbol{x} \triangleq f_n(\boldsymbol{s},\boldsymbol{l})$ and $\boldsymbol{u} \triangleq g_n(\boldsymbol{s})$ are the corresponding codewords of the source message pair $(\boldsymbol{s},\boldsymbol{l})$ and the source message $\boldsymbol{s}$, and $\boldsymbol{y}$ and $\boldsymbol{z}$ are the received codewords at the Receivers Y and Z, respectively. Since we will study the exponential behavior of these probabilities using the method of types, it might be a better way to rewrite the probabilities of Y- and Z- error as a sum of probabilities of types

$$
\begin{aligned}
&P_{ie}^{(n)}(Q_{SL}, W_{YZ\mid UX}, \tau)\\
&\quad= \sum_{P_{SL}\in\mathcal{P}_{\tau n}(\mathcal{S}\times\mathcal{L})} Q_{SL}^{(\tau n)}(\mathbb{T}_{SL}) P_{ie}(\mathbb{T}_{SL}), \quad i = Y, Z
\end{aligned}
\tag{10}
$$

where $\mathbb{T}_{SL} \triangleq \mathbb{T}_{P_{SL}}$, and

$$
\begin{aligned}
&P_{Ye}(\mathbb{T}_{SL})\\
&\quad= \frac{1}{|\mathbb{T}_{SL}|} \sum_{(\boldsymbol{s},\boldsymbol{l})\in\mathbb{T}_{SL}} \sum_{\boldsymbol{y}:\varphi_n(\boldsymbol{y})\neq(\boldsymbol{s},\boldsymbol{l})} W_{Y\mid UX}^{(n)}(\boldsymbol{y}|\boldsymbol{u},\boldsymbol{x})
\end{aligned}
\tag{11}
$$

and

$$
\begin{aligned}
&P_{Ze}(\mathbb{T}_{SL})\\
&\quad= \frac{1}{|\mathbb{T}_{SL}|} \sum_{(\boldsymbol{s},\boldsymbol{l})\in\mathbb{T}_{SL}} \sum_{\boldsymbol{z}:\psi_n(\boldsymbol{z})\neq\boldsymbol{s}} W_{Z|UX}^{(n)}(\boldsymbol{z}|\boldsymbol{u},\boldsymbol{x}).
\end{aligned}
\tag{12}
$$

We say that the JSCC error exponent pair (E_{AY}, E_{AZ}) is achievable with respect to $\tau > 0$ if there exists a sequence of JSC codes $(f_n, g_n, \varphi_n, \psi_n)$ with transmission rate τ such that the probabilities of Y-error and Z-error are simultaneously bounded by

$$
P_{ie}^{(n)}(Q_{SL}, W_{YZ\mid UX}, \tau) \leq 2^{-n[E_{Ai}-\delta]}, \quad i = Y, Z \tag{13}
$$

for n sufficiently large and any $\delta > 0$. As the point-to-point system, we denote the system (overall) probability of error by

$$
\begin{aligned}
&P_{e}^{(n)}(Q_{SL}, W_{YZ\mid UX}, \tau)\\
&\quad\triangleq \Pr\left(\{\varphi_n(Y^n) \neq (S^{\tau n}, L^{\tau n})\} \bigcup \{\psi_n(Z^n) \neq S^{\tau n}\}\right)
\end{aligned}
\tag{14}
$$

where $(S^{\tau n}, L^{\tau n})$ are drawn according to $Q_{SL}^{(\tau n)}$.

Definition 1: Given Q_{SL}, $W_{YZ\mid UX}$, and $\tau > 0$, the system JSCC error exponent $E_J(Q_{SL}, W_{YZ\mid UX}, \tau)$ is defined as the supremum of the set of all numbers E for which there exists a sequence of JSC codes $(f_n, g_n, \varphi_n, \psi_n)$ with block length n and transmission rate τ such that

$$E \le \liminf_{n\to\infty} -\frac{1}{n}\log_2 P_e^{(n)}(Q_{SL}, W_{YZ\,|\,UX}, \tau). \tag{15}$$

Since the system probability of error must be larger than $P_{Ye}^{(n)}$ and $P_{Ze}^{(n)}$ defined by (8) and (9), and is also upper-bounded by the sum of the two, it follows that for any sequence of JSC codes $(f_n, g_n, \varphi_n, \psi_n)$

$$\liminf_{n\to\infty} -\frac{1}{n}\log_2 P_e^{(n)}(Q_{SL}, W_{YZ\,|\,UX}, \tau) = \liminf_{n\to\infty} -\frac{1}{n}\log_2 \max\left(P_{Ye}^{(n)}, P_{Ze}^{(n)}\right). \tag{16}$$

B. Superposition Encoding for Asymmetric Two-User Channels

Given an asymmetric two-user channel $W_{YZ\,|\,UX}$, at the encoder side, we can artificially augment the channel input alphabet by introducing an auxiliary (arbitrary and finite) alphabet $\mathcal{T}$, and then look at the channel as a discrete memoryless channel $W_{YZ\,|\,TUX} = W_{YZ\,|\,UX}$ with marginal distributions $W_{Y\,|\,TUX}$ and $W_{Z\,|\,TUX}$ such that $W_{YZ\,|\,TUX}(y,z\,|\,t,u,x) = W_{YZ\,|\,UX}(y,z\,|\,u,x)$ for any $t\in\mathcal{T}$, $u\in\mathcal{U}$, $x\in\mathcal{X}$, $y\in\mathcal{Y}$, and $z\in\mathcal{Z}$. In other words, we introduce a dummy RV $T\in\mathcal{T}$ such that T, (U,X),, and (Y,Z) form a Markov chain in this order, i.e., $T\to(U,X)\to(Y,Z)$.

The idea of superposition coding is described as follows. The encoder g_n first maps the source message $\boldsymbol{s}$ to a pair of n-length sequences $(\boldsymbol{t},\boldsymbol{u})\in\mathcal{T}^n\times\mathcal{U}^n$ with a fixed type, say P_{TU}, and then sends the codeword $\boldsymbol{u}$ over the channel, i.e., $g_n(\boldsymbol{s}) = \boldsymbol{u}$. The encoder f_n first maps each pair $(\boldsymbol{s},\boldsymbol{l})$ to a triple of sequences $(\boldsymbol{t},\boldsymbol{u},\boldsymbol{x})\in\mathcal{T}^n\times\mathcal{U}^n\times\mathcal{X}^n$ such that $\boldsymbol{x}\in\mathbb{T}_{P_{X\,|\,TU}}(\boldsymbol{t},\boldsymbol{u})$, then f_n sends the codeword $\boldsymbol{x}$ over the channel, i.e., $f_n(\boldsymbol{s},\boldsymbol{l}) = \boldsymbol{x}$. In other words, g_n and f_n map $(\boldsymbol{s},\boldsymbol{l})$ to a tuple of sequences $(\boldsymbol{t},\boldsymbol{u},\boldsymbol{x})$ with a joint type $P_{TU}P_{X\,|\,TU}$, although only $\boldsymbol{u}$ and $\boldsymbol{x}$ are sent to the channel, where $\boldsymbol{t}$ plays the role of a dummy codeword.

Since $W_{YZ\,|\,TUX}^{(n)}(\boldsymbol{y},\boldsymbol{z}|\boldsymbol{t},\boldsymbol{u},\boldsymbol{x})$ is equal to $W_{YZ\,|\,UX}^{(n)}(\boldsymbol{y},\boldsymbol{z}\,|\,\boldsymbol{u},\boldsymbol{x})$ and is independent of $\boldsymbol{t}$, transmitting the codewords $(\boldsymbol{u},\boldsymbol{x})$ through the channel $W_{YZ\,|\,UX}$ can be viewed as transmitting the codewords $(\boldsymbol{t},\boldsymbol{u},\boldsymbol{x})$ over the augmented channel $W_{YZ\,|\,TUX}$. Here, the common outputs of g_n and f_n, $(\boldsymbol{t},\boldsymbol{u})$'s, are called auxiliary *cloud centers* according to the traditional superposition coding notion [3], which convey the information of the common message $\boldsymbol{s}$, and the codewords $\boldsymbol{x}$'s corresponding to the same $(\boldsymbol{t},\boldsymbol{u})$ are called satellite codewords of $(\boldsymbol{t},\boldsymbol{u})$, which contain both the common and private information, see Fig. 3. At the decoding stage, Receiver Z only needs to figure out which cloud $(\boldsymbol{t},\boldsymbol{u})$ was transmitted, and Receiver Y needs to estimate not only the cloud but also the satellite codeword $\boldsymbol{x}$. The introduction of the auxiliary RV T is made to enlarge the channel input alphabet from $\mathcal{U}\times\mathcal{X}$ to $\mathcal{T}\times\mathcal{U}\times\mathcal{X}$, and the use of the superposition codeword $\boldsymbol{t}$ renders the cloud centers $(\boldsymbol{t},\boldsymbol{u})$ more distinguishable by both receivers. We next employ superposition encoding to derive the achievable error exponent pair and the lower bound of system JSCC error exponent.

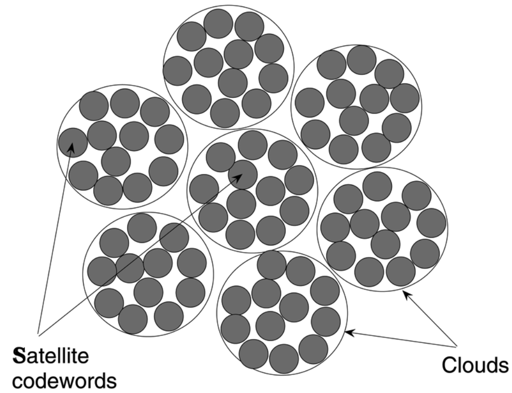

Fig. 3. Relation between clouds and satellite codewords in superposition coding.

C. Achievable Exponents and a Lower Bound for e_j

Given arbitrary and finite alphabet $\mathcal{T}$, for any joint distribution $P_{TUX}\in\mathcal{P}(\mathcal{T}\times\mathcal{U}\times\mathcal{X})$ and every $R_1>0$, $R_2>0$, define

$$\begin{aligned}&E_Y(R_1,R_2,W_{Y\,|\,TUX},P_{TUX})\\ &\quad\triangleq \min_{V_{Y\,|\,TUX}}\Big[D(V_{Y\,|\,TUX}\,\|\,W_{Y\,|\,TUX}|P_{TUX})\\ &\qquad+\min(|I_{P_{TUX}V_{Y\,|\,TUX}}(T,U,X;Y)-(R_1+R_2)|^+,\\ &\qquad|I_{P_{TUX}V_{Y\,|\,TUX}}(X;Y\,|\,T,U)-R_2|^+)\Big]\end{aligned} \tag{17}$$

and

$$\begin{aligned}&E_Z(R_1,R_2,W_{Z\,|\,TUX},P_{TUX})\\ &\quad\triangleq \min_{V_{Z\,|\,TUX}}\Big[D(V_{Z\,|\,TUX}\,\|\,W_{Z\,|\,TUX}|P_{TUX})\\ &\qquad+|I_{P_{TUX}V_{Z\,|\,TUX}}(T,U;Z)-R_1|^+\Big]\end{aligned} \tag{18}$$

where $|x|^+ = \max(0,x)$, and the outer minimum in (17) (respectively, (18)) is taken over all conditional distributions on $\mathcal{P}(\mathcal{Y}|\mathcal{T}\times\mathcal{U}\times\mathcal{X})$ (respectively, $\mathcal{P}(\mathcal{Z}|\mathcal{T}\times\mathcal{U}\times\mathcal{X})$). It immediately follows by definition that $E_Y(R_1,R_2,W_{Y\,|\,TUX},P_{TUX})$ is zero if and only if at least one of the following is satisfied:

$$R_1+R_2 \ge I_{P_{TUX}W_{Y\,|\,TUX}}(T,U,X;Y) = I(U,X;Y) \tag{19}$$

$$R_2 \ge I_{P_{TUX}W_{Y\,|\,TUX}}(X;Y\,|\,T,U) \tag{20}$$

and $E_Z(R_1,R_2,W_{Z\,|\,TUX},P_{TUX})$ is zero if and only if

$$R_1 \ge I_{P_{TUX}W_{Z\,|\,TUX}}(T,U;Z). \tag{21}$$

Using Lemma 3 and employing generalized MMI decoders at the two receivers, we can prove the following auxiliary bounds.

Theorem 1: Given finite sets $\mathcal{T}$, $\mathcal{U}$, $\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Z}$, a sequence of positive integers $\{m_n\}$, and a sequence of positive integers $\{m'_{\text{in}}\}$ associated with every $i=1,2,\ldots,m_n$ with

$$\frac{1}{n}\log_2 m_n \to 0 \quad\text{and}\quad \frac{1}{n}\log_2\max_i m'_{\text{in}} \to 0$$

for any $\delta>0$, n sufficiently large, arbitrary (not necessarily distinct) types $P_{(TU)_i}\in\mathcal{P}_n(\mathcal{T}\times\mathcal{U})$, and conditional types $P_{X_j\,|\,(TU)_i}\in\mathcal{P}_n(\mathcal{X}|P_{(TU)_i})$, and positive integers N_i and M_{ij}, $i=1,2,\ldots,m_n$ and $j=j(i)=1,2,\ldots,m'_{\text{in}}$ with $R_i < H_{P_{(TU)_i}}(T,U)-\delta$, and $R_{ij} < H_{P_{(TU)_i}P_{X_j\,|\,(TU)_i}}(X\,|\,T,U)-\delta$, where $R_i \triangleq \frac{1}{n}\log_2 N_i$ and $R_{ij}\triangleq\frac{1}{n}\log_2 M_{ij}$, there exist m_n

disjoint subsets $\Omega_i = \{(\boldsymbol{t},\boldsymbol{u})_p^{(i)}\}_{p=1}^{N_i} \subseteq \mathbb{T}_{(TU)_i}$, $m'_{\rm in}$ disjoint subsets

$$\Omega_{ij}((\boldsymbol{t},\boldsymbol{u})_p^{(i)}) = \left\{ \left((\boldsymbol{t},\boldsymbol{u})_p^{(i)}, \boldsymbol{x}_{p,q}^{(j)} \right) \right\}_{q=1}^{M_{ij}}$$

with $\boldsymbol{x}_{p,q}^{(j)} \in \mathbb{T}_{X_j\,|\,(TU)_i}((\boldsymbol{t},\boldsymbol{u})_p^{(i)})$ for every $(\boldsymbol{t},\boldsymbol{u})_p^{(i)} \in \Omega_i$ and every i, and a pair of mappings (decoding functions) $\varphi_n^{(0)}: \mathcal{Y}^n \to \Omega$ and $\psi_n^{(0)}: \mathcal{Z}^n \to \Omega$, where $\Omega \triangleq \bigcup_{ij} \Omega_{ij}$, where

$$\Omega_{ij} = \bigcup_{p=1}^{N_i} \Omega_{ij}((\boldsymbol{t},\boldsymbol{u})_p^{(i)})$$

such that the probabilities of erroneous transmission of a triplet $(\boldsymbol{t},\boldsymbol{u},\boldsymbol{x}) \in \Omega$ over the augmented channel $W_{YZ\,|\,TUX}$ using decoders $(\varphi_n^{(0)}, \psi_n^{(0)})$ are *simultaneously* bounded by

$$\begin{aligned} P_{Ye}^{(n)}(\boldsymbol{t},\boldsymbol{u},\boldsymbol{x}) &\triangleq \sum_{\boldsymbol{y}:\varphi_n^{(0)}(\boldsymbol{y}) \neq ((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x})} W_{Y\,|\,TUX}^{(n)}(\boldsymbol{y}|\boldsymbol{t},\boldsymbol{u},\boldsymbol{x}) \\ &\leq 2^{-n[E_Y(R_i,R_{ij},W_{Y\,|\,TUX},P_{(TU)_i}P_{X_j\,|\,(TU)_i})-\delta]} \end{aligned} \tag{22}$$

and

$$\begin{aligned} P_{Ze}^{(n)}(\boldsymbol{t},\boldsymbol{u},\boldsymbol{x}) &\triangleq \sum_{\boldsymbol{z}:\psi_n^{(0)}(\boldsymbol{z}) = ((\boldsymbol{t},\boldsymbol{u})',\boldsymbol{x}') \text{ such that } (\boldsymbol{t},\boldsymbol{u})' \neq (\boldsymbol{t},\boldsymbol{u})} W_{Z\,|\,TUX}^{(n)}(\boldsymbol{z}|\boldsymbol{t},\boldsymbol{u},\boldsymbol{x}) \\ &\leq 2^{-n[E_Z(R_i,R_{ij},W_{Z\,|\,TUX},P_{(TU)_i}P_{X_j\,|\,(TU)_i})-\delta]} \end{aligned} \tag{23}$$

if $((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x}) \in \Omega_{ij}$ for every i, j.

Proof: We apply the packing lemma (Lemma 3) and a generalized MMI decoding rule.[1] In the sequel of the proof, we look at the superletter (T,U) (respectively, X) as the RV A (respectively, B) in Lemma 3. For the $\{m_n\}$, $\{m'_{\rm in}\}$, $P_{(TU)_i}$, $P_{X_j\,|\,(TU)_i}$ given in Theorem 1, according to Lemma 3, there exist pairwise disjoint subsets Ω_i and $\Omega_{ij}((\boldsymbol{t},\boldsymbol{u})_p^{(i)})$ satisfying (5), (6), and (7) for every $1 \leq i \leq m_n$, $1 \leq j \leq m'_{\rm in}$, $1 \leq p \leq N_i$, $V_{(TU)'|TU} \in \mathcal{P}_n(\mathcal{T}\times\mathcal{U}|\mathcal{T}\times\mathcal{U})$, and $V_{(TU)'X'\,|\,TUX} \in \mathcal{P}_n(\mathcal{T}\times\mathcal{U}\times\mathcal{X}|\mathcal{T}\times\mathcal{U}\times\mathcal{X})$, with the exception of the two cases that $i = k$ and $V_{(TU)'\,|\,TU}$ is the conditional distribution such that $V_{(TU)'\,|\,TU}((t,u)'|(t,u))$ is 1 if $(t,u)' = (t,u)$ and 0 otherwise, and that $i = k$, $j = l$, and $V_{(TU)'X'\,|\,TUX}$ is the conditional distribution such that $V_{(TU)'X'\,|\,TUX}((t,u)', x'|t,u,x)$ is 1 if $(t,u)' = (t,u)$, $x' = x$ and 0 otherwise. Let

$$\Omega_{ij} = \bigcup_{p=1}^{N_i} \Omega_{ij}((\boldsymbol{t},\boldsymbol{u})_p^{(i)}) \quad \text{and} \quad \Omega = \bigcup_{ij} \Omega_{ij}.$$

We shall show that for such Ω_{ij}, there exists a pair of mappings $(\varphi_n^{(0)}, \psi_n^{(0)})$ such that (22) and (23) are satisfied.

[1]Note that for the symmetric multiple-access channel, it has been shown in [24] that the minimum conditional entropy (MCE) decoder leads to a larger channel error exponent than the MMI decoder; however, for the asymmetric two-user channel with superposition coding, MMI decoding is equivalent to MCE decoding.

We first show that there exists a Y-decoder $\varphi_n^{(0)}$ such that (22) holds. For any $((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x}) \in \Omega$ and $\boldsymbol{y} \in \mathcal{Y}^n$, let

$$\alpha((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x};\boldsymbol{y}) \triangleq I((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x};\boldsymbol{y}) - (R_i + R_{ij})$$

where $R_i = \frac{1}{n}\log_2 N_i$ and $R_{ij} = \frac{1}{n}\log_2 M_{ij}$ if $((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x}) \in \Omega_{ij}$. Define Y-decoder $\varphi_n^{(0)} : \mathcal{Y}^n \to \Omega$ by

$$\varphi_n^{(0)}(\boldsymbol{y}) \triangleq \arg \max_{((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x})\in\Omega} \alpha((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x};\boldsymbol{y}).$$

Using the decoder $\varphi_n^{(0)}$, we can upper-bound the probability of error (assuming that $((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x}) \in \Omega_{ij}$ is sent through the channel) as follows:

$$\begin{aligned} &P_{Ye}^{(n)}((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x}) \\ &\quad = W_{Y\,|\,TUX}^{(n)}\left(\left\{ \boldsymbol{y} : \varphi_n^{(0)}(\boldsymbol{y}) \neq ((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x}) \right\} \Big| (\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x} \right) \\ &\quad \leq \sum_{\hat{V}_{Y\,|\,TUX} \in \mathcal{P}_n(\mathcal{Y}\,|\,P_{(TU)_i X_j})} W_{Y\,|\,TUX}^{(n)} \\ &\qquad \left(\mathbb{T}_{\hat{V}_{Y\,|\,TUX}}((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x}) \right. \\ &\qquad \left. \bigcap \left\{ \boldsymbol{y} : \varphi_n^{(0)}(\boldsymbol{y}) \neq ((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x}) \right\} \Big| \boldsymbol{t},\boldsymbol{u},\boldsymbol{x} \right). \end{aligned} \tag{24}$$

For any particular $\hat{V}_{Y\,|\,TUX}$, since

$$\begin{aligned} &\left\{ \boldsymbol{y} : \varphi_n^{(0)}(\boldsymbol{y}) \neq ((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x}) \right\} \\ &\quad = \underbrace{\left\{ \boldsymbol{y} : \varphi_n^{(0)}(\boldsymbol{y}) = ((\boldsymbol{t},\boldsymbol{u})',\boldsymbol{x}'), (\boldsymbol{t},\boldsymbol{u})' \neq (\boldsymbol{t},\boldsymbol{u}) \right\}}_{\triangleq \mathcal{E}_1} \\ &\qquad \bigcup \underbrace{\left\{ \boldsymbol{y} : \varphi_n^{(0)}(\boldsymbol{y}) = ((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x}'), \boldsymbol{x}' \neq \boldsymbol{x} \right\}}_{\triangleq \mathcal{E}_2} \end{aligned}$$

we can upper-bound

$$\begin{aligned} &W_{Y\,|\,TUX}^{(n)}\left(\mathbb{T}_{\hat{V}_{Y\,|\,TUX}}((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x}) \right. \\ &\qquad \left. \bigcap \left\{ \boldsymbol{y} : \varphi_n^{(0)}(\boldsymbol{y}) \neq ((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x}) \right\} \Big| \boldsymbol{t},\boldsymbol{u},\boldsymbol{x} \right) \\ &\quad \leq \sum_{\boldsymbol{y} \in \mathbb{T}_{\hat{V}_{Y\,|\,TUX}}((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x}) \bigcap \mathcal{E}_1} W_{Y\,|\,TUX}^{(n)}(\boldsymbol{y}|\boldsymbol{t},\boldsymbol{u},\boldsymbol{x}) \\ &\qquad + \sum_{\boldsymbol{y} \in \mathbb{T}_{\hat{V}_{Y\,|\,TUX}}((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x}) \bigcap \mathcal{E}_2} W_{Y\,|\,TUX}^{(n)}(\boldsymbol{y}|\boldsymbol{t},\boldsymbol{u},\boldsymbol{x}). \end{aligned} \tag{25}$$

It can be shown by the type packing lemma (Lemma 3) and a standard counting argument (see Appendix B) what is displayed in (26) and (27) at the top of the following page. Using the identity (cf. Lemma 1) when $((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x}) \in \Omega_{ij} \subseteq \mathbb{T}_{(TU)_i X_j}$ and $\boldsymbol{y} \in \mathbb{T}_{\hat{V}_{Y\,|\,TUX}}((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x})$, and as shown by the second expression at the top of the following page, we obtain (28) and (29) at the bottom of the following page.Substituting (28) and (29) back into (25) and (24) successively, and noting that $|\mathcal{P}_n(\mathcal{Y}|P_{(TU)_i X_j})|$ is polynomial in n by Lemma 1, we obtain that, for any $\delta > 0$, there exists a Y-decoder $\varphi_n^{(0)}$ such that,

$$\left|\mathbb{T}_{\hat{V}_{Y\,|\,TUX}}((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x})\bigcap\mathcal{E}_1\right| \le m_n\left(\max_i m'_{\text{in}}\right)(n+1)^{|\mathcal{T}\times\mathcal{U}|^2|\mathcal{X}|^2|\mathcal{Y}|} \times 2^{n[H_{P_{(TU)_i X_j}\hat{V}_{Y\,|\,TUX}}(Y\,|\,T,U,X)-|I_{P_{(TU)_i X_j}\hat{V}_{Y\,|\,TUX}}(T,U,X;Y)-(R_i+R_{ij})|^+]} \tag{26}$$

and

$$\left|\mathbb{T}_{\hat{V}_{Y\,|\,TUX}}((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x})\bigcap\mathcal{E}_2\right| \le \left(\max_i m'_{\text{in}}\right)(n+1)^{|\mathcal{T}\times\mathcal{U}||\mathcal{X}|^2|\mathcal{Y}|} \times 2^{n[H_{P_{(TU)_i X_j}\hat{V}_{Y\,|\,TUX}}(Y\,|\,T,U,X)-|I_{P_{(TU)_i X_j}\hat{V}_{Y\,|\,TUX}}(X;Y\,|\,T,U)-R_{ij}|^+]}. \tag{27}$$

$$W^{(n)}_{Y\,|\,TUX}(\boldsymbol{y}\,|\,(\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x}) = 2^{-n[D(\hat{V}_{Y\,|\,TUX}\|W_{Y\,|\,TUX}|P_{(TU)_i X_j})+H_{P_{(TU)_i X_j}\hat{V}_{Y\,|\,TUX}}(Y\,|\,T,U,X)]}$$

---

given $((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x}) \in \Omega_{ij}$, the probability of Y-error is bounded by

$$P^{(n)}_{Ye}((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x}) \le 2^{-n[E_Y(R_i,R_{ij},W_{Y\,|\,TUX},P_{(TU)_i}P_{X_j\,|\,(TU)_i})-\delta]} \tag{30}$$

for sufficiently large n.

Similarly, we can design a decoder for Receiver Z as follows. For any $((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x}) \in \Omega$ and $\boldsymbol{z} \in \mathcal{Z}^n$, let

$$\beta((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x};\boldsymbol{z}) = \beta((\boldsymbol{t},\boldsymbol{u});\boldsymbol{z}) \triangleq I((\boldsymbol{t},\boldsymbol{u});\boldsymbol{z}) - R_i$$

where $R_i = \frac{1}{n}\log_2 N_i$ if $(\boldsymbol{t},\boldsymbol{u}) \in \Omega_i$. Note that $\beta((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x};\boldsymbol{z})$ is independent of $\boldsymbol{x}$. Let $\widetilde{\Omega} = \sum_{i=1}^{m_n}\Omega_i$. The Z-decoder $\psi^{(0)}_n : \mathcal{Z}^n \to \Omega$ is defined by the last equation at the bottom of the page. It can be shown in a similar manner by using (5) in Lemma 3 that, under the decoder $\psi^{(0)}_n$, the probability of the Z-error is bounded by

$$P^{(n)}_{Ze}((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x}) \le 2^{-n[E_Z(R_i,R_{ij},W_{Z\,|\,TUX},P_{(TU)_i}P_{X_j\,|\,(TU)_i})-\delta]} \tag{31}$$

for sufficiently large n. Finally, we remark that Lemma 3 ensures that there exist mappings $(\varphi^{(0)}_n, \psi^{(0)}_n)$ such that (31) holds simultaneously with (30). □

Theorem 1 is an auxiliary result for the channel coding problem for the two-user asymmetric channel. To apply it to our two-user source–channel system, we need to design encoders which can map a pair of correlated source messages to a particular $(\boldsymbol{t},\boldsymbol{u},\boldsymbol{x})$ with a joint type, so that the total probabilities of error still vanish exponentially. We hence can establish the following bounds.

Theorem 2: Given an arbitrary and finite alphabet $\mathcal{T}$, for any $\widetilde{P}_{TUX} \in \mathcal{P}(\mathcal{T}\times\mathcal{U}\times\mathcal{X})$, the following exponent pair is universally achievable:

$$E_{JY}(Q_{SL},W_{YZ\,|\,TUX},\tilde{P}_{TUX},\tau) \triangleq \min_{P_{SL}}[\tau D(P_{SL}\,\|\,Q_{SL}) + E_Y(\tau H_P(S), \tau H_P(L\,|\,S), W_{Y\,|\,TUX}, \tilde{P}_{TUX})] \tag{32}$$

and

$$E_{JZ}(Q_{SL},W_{YZ\,|\,TUX},\tilde{P}_{TUX},\tau) \triangleq \min_{P_{SL}}[\tau D(P_{SL}\,\|\,Q_{SL}) + E_Z(\tau H_P(S), \tau H_P(L\,|\,S), W_{Z\,|\,TUX}, \tilde{P}_{TUX})] \tag{33}$$

where $W_{Y\,|\,TUX}$ and $W_{Z\,|\,TUX}$ are marginal distributions of $W_{YZ\,|\,TUX}$, which is the augmented conditional distribution from $W_{YZ\,|\,UX}$. Furthermore, given Q_{SL}, $W_{YZ\,|\,UX}$, and τ, the system JSCC error exponent satisfies

$$E_J(Q_{SL},W_{YZ\,|\,UX},\tau) \ge \min_{P_{SL}}[\tau D(P_{SL}\,\|\,Q_{SL}) + E_r(\tau H_P(S), \tau H_P(L\,|\,S), W_{YZ\,|\,UX})] \tag{34}$$

$$\sum_{\boldsymbol{y}\in\mathbb{T}_{\hat{V}_{Y\,|\,TUX}}((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x})\bigcap\mathcal{E}_1} W^{(n)}_{Y\,|\,TUX}(\boldsymbol{y}|((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x})\in\Omega_{ij}) \le m_n\left(\max_i m'_{\text{in}}\right)(n+1)^{|\mathcal{T}\times\mathcal{U}|^2|\mathcal{X}|^2|\mathcal{Y}|} \times 2^{-n[D(\hat{V}_{Y\,|\,TUX}\|W_{Y\,|\,TUX}|P_{(TU)_i X_j})+|I_{P_{(TU)_i X_j}\hat{V}_{Y\,|\,TUX}}(T,U,X;Y)-(R_i+R_{ij})|^+]} \tag{28}$$

and

$$\sum_{\boldsymbol{y}\in\mathbb{T}_{\hat{V}_{Y\,|\,TUX}}((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x})\bigcap\mathcal{E}_2} W^{(n)}_{Y\,|\,TUX}(\boldsymbol{y}|((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x})\in\Omega_{ij}) \le \left(\max_i m'_{\text{in}}\right)(n+1)^{|\mathcal{T}\times\mathcal{U}||\mathcal{X}|^2|\mathcal{Y}|} \times 2^{-n[D(\hat{V}_{Y\,|\,TUX}\|W_{Y\,|\,TUX}|P_{(TU)_i X_j})+|I_{P_{(TU)_i X_j}\hat{V}_{Y\,|\,TUX}}(X;Y\,|\,T,U)-R_{ij}|^+]}. \tag{29}$$

$$\begin{aligned}\varphi^{(0)}_n(\boldsymbol{z}) &= \arg\max_{((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x})\in\Omega}\beta((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x};\boldsymbol{z}) \\ &= ((\boldsymbol{t},\boldsymbol{u})',\boldsymbol{x}') \quad \text{such that} \begin{cases}(\boldsymbol{t},\boldsymbol{u})' = \arg\max_{(\boldsymbol{t},\boldsymbol{u})\in\tilde{\Omega}}\beta((\boldsymbol{t},\boldsymbol{u});\boldsymbol{z}), \\ \boldsymbol{x}' \text{ is arbitrary.}\end{cases}\end{aligned}$$

where

$$E_r(R_1, R_2, W_{YZ\,|\,UX}) \triangleq \sup_{\mathcal{T}} \max_{P_{TUX}} E_r(R_1, R_2, W_{YZ\,|\,TUX}, P_{TUX}) \quad (35)$$

where the supremum is taken over all finite alphabets $\mathcal{T}$, and the maximum is taken over all the joint distributions on $\mathcal{P}(\mathcal{T} \times \mathcal{U} \times \mathcal{X})$ and $E_r(R_1, R_2, W_{YZ\,|\,TUX}, P_{TUX})$ is given by

$$\min\{E_Y(R_1, R_2, W_{Y\,|\,TUX}, P_{TUX}), E_Z(R_1, R_2, W_{Z\,|\,TUX}, P_{TUX})\}$$

where E_Y and E_Z are given by (17) and (18), respectively.

We remark that (32) and (33) can be achieved by a sequence of codes without the knowledge of Q_{SL} and $W_{YZ\,|\,UX}$, but the lower bound (34) is achieved by a sequence of codes that needs to know the statistics of the channel.

Proof of Theorem 2: We first prove the achievable error exponent pair (32) and (33). We need to show that, for any given $\tilde{P}_{TUX} \in \mathcal{P}(\mathcal{T} \times \mathcal{U} \times \mathcal{X})$ and $\delta > 0$, there exists a sequence of JSC codes such that both the probabilities of decoding error are upper-bounded by

$$P_{ke}^{(n)}(Q_{SL}, W_{YZ\,|\,UX}, \tau) \le 2^{-n[E_{Jk}(Q_{SL}, W_{YZ\,|\,TUX}, \tilde{P}_{TUX}, \tau) - \delta]}, \quad k = Y, Z$$

where E_{JY} and E_{JZ} are given by (32) and (33).

To apply Theorem 1, set $m_n \triangleq |\mathcal{P}_{\tau n}(\mathcal{S})|$. For each type $P_{S_i} \in \mathcal{P}_{\tau n}(\mathcal{S})$, $i = 1, 2, \ldots, m_n$, denote N_i to be the cardinalities of these type classes, $N_i \triangleq |\mathbb{T}_{S_i}|$, and set $m'_{\text{in}} \triangleq |\mathcal{P}_{\tau n}(\mathcal{L}|P_{S_i})|$. For each conditional type $P_{L_j|S_i} \in \mathcal{P}_{\tau n}(\mathcal{L}|P_{S_i})$, $j = 1, 2, \ldots, m'_{\text{in}}$, denote M_{ij} be the cardinalities of these type classes, $M_{ij} \triangleq |\mathbb{T}_{L_j|S_i}(\boldsymbol{s})|$ where $\boldsymbol{s}$ is an arbitrary sequence in $\mathbb{T}_{S_i}$. Note that $|\mathbb{T}_{L_j|S_i}(\boldsymbol{s})|$ is constant for all $\boldsymbol{s} \in \mathbb{T}_{S_i}$. R_i and R_{ij} are, respectively, given by $\frac{1}{n}\log_2 N_i$ and $\frac{1}{n}\log_2 M_{ij}$.

Now no matter whether the given $\tilde{P}_{TUX}$ belongs to $\mathcal{P}_n(\mathcal{T} \times \mathcal{U} \times \mathcal{X})$ or not, we always can find a sequence of joint types $\{P_{TUX} \in \mathcal{P}_n(\mathcal{T} \times \mathcal{U} \times \mathcal{X})\}_{n=1}^{\infty}$ such that $P_{TUX} \to \tilde{P}_{TUX}$ uniformly[2] as $n \to \infty$. Thus, we can choose, by the continuity of $E_k(R_i, R_{ij}, W_{k\,|\,TUX}, \tilde{P}_{TUX})$ with respect to $\tilde{P}_{TUX}$, for each $i = 1, 2, \ldots, m_n$, and $j = j(i) = 1, 2, \ldots, m'_{\text{in}}$, the joint type $P_{(TU)_i X_j} = P_{TUX}$ such that the following are satisfied:

$$\Big|E_k(R_i, R_{ij}, W_{k\,|\,TUX}, P_{TUX}) - E_k(R_i, R_{ij}, W_{k\,|\,TUX}, \tilde{P}_{TUX})\Big| < \frac{\delta}{4}, \quad k = Y, Z$$

for n sufficiently large. Since the type P_{TUX} can also be regarded as a joint distribution, let $P_{(TU)_i} = P_{TU} \in \mathcal{P}_n(\mathcal{T} \times \mathcal{U})$ be the marginal distribution on $\mathcal{T} \times \mathcal{U}$ induced by P_{TUX} for all $i = 1, 2, \ldots, m_n$ and let $P_{X_j\,|\,(TU)_i} = P_{X\,|\,TU} \in \mathcal{P}_n(\mathcal{X}|P_{TU})$ be the corresponding conditional distribution for all $i = 1, 2, \ldots, m_n$ and $j = 1, 2, \ldots, m'_{\text{in}}$, i.e., $P_{X\,|\,TU}(\boldsymbol{x}|\boldsymbol{t}, \boldsymbol{u}) = P_{TUX}(\boldsymbol{t}, \boldsymbol{u}, \boldsymbol{x})/P_{TU}(\boldsymbol{t}, \boldsymbol{u})$ for any $(\boldsymbol{t}, \boldsymbol{u}, \boldsymbol{x}) \in \mathbb{T}_{TUX}$.

[2] We say that a sequence of distributions $\{P_{X_i} \in \mathcal{P}(\mathcal{X})\}_{i=1}^{\infty}$ uniformly converges to $P_X^* \in \mathcal{P}(\mathcal{X})$ if the variational distance [10] between P_{X_i} and P_X^* converges to zero as $n \to \infty$.

Without loss of generality, we assume, for the choice of N_i, M_{ij}, $P_{(TU)_i}$, and $P_{X_j\,|\,(TU)_i}$, the following conditions are satisfied for $i = 1, 2, \ldots, \hat{m}_n$, $j = 1, 2, \ldots, \hat{m}'_{\text{in}}$:

$$R_i < H_{P_{(TU)_i}}(T, U) - \frac{\delta}{4}, \quad i = 1, 2, \ldots, \hat{m}_n \quad (36)$$

and

$$R_{ij} < H_{P_{(TU)_i X_j}}(X\,|\,T, U) - \frac{\delta}{4}, \quad i = 1, 2, \ldots, \hat{m}_n, \ j = j(i) = 1, 2, \ldots, \hat{m}'_{\text{in}} \quad (37)$$

where $\hat{m}_n \le m_n$ and $\hat{m}'_{\text{in}} \le m'_n$. Then according to Theorem 1, there exist pairwise disjoint subsets $\Omega_{ij} \subseteq \mathbb{T}_{(TU)_i X_j}$ with $|\Omega_{ij}| = N_i M_{ij}$, $i = 1, 2, \ldots, \hat{m}_n$, $j = 1, 2, \ldots, \hat{m}'_{\text{in}}$, and a pair of mappings $(\varphi_n^{(0)}, \psi_n^{(0)})$, such that the probabilities of erroneous transmission of a $((\boldsymbol{t}, \boldsymbol{u}), \boldsymbol{x}) \in \Omega_{ij}$ are *simultaneously* bounded for the channel $W_{YZ\,|\,TUX}$ as

$$P_{Ye}^{(n)}(\boldsymbol{t}, \boldsymbol{u}, \boldsymbol{x}) \le 2^{-n[E_Y(R_i, R_{ij}, W_{Y\,|\,TUX}, P_{(TU)_i X_j}) - \delta/4]} \le 2^{-n[E_Y(R_i, R_{ij}, W_{Y\,|\,TUX}, \tilde{P}_{TUX}) - \delta/2]} \quad (38)$$

and

$$P_{Ze}^{(n)}(\boldsymbol{t}, \boldsymbol{u}, \boldsymbol{x}) \le 2^{-n[E_Z(R_i, R_{ij}, W_{Z\,|\,TUX}, P_{(TU)_i X_j}) - \delta/4]} \le 2^{-n[E_Z(R_i, R_{ij}, W_{Z\,|\,TUX}, \tilde{P}_{TUX}) - \delta/2]}. \quad (39)$$

For the N_i, M_{ij}, $P_{(TU)_i}$, and $P_{X_j\,|\,(TU)_i}$ violating (36) or (37) (i.e., for $i > \hat{m}_n$ or $j > \hat{m}'_{\text{in}}$), (38) and (39) trivially hold for arbitrary choice of disjoint subsets Ω_{ij} since $E_Y\left(R_i, R_{ij}, W_{Y\,|\,TUX}, P_{(TU)_i X_j}\right)$ or $E_Z\left(R_i, R_{ij}, W_{Z\,|\,TUX}, P_{(TU)_i X_j}\right)$ would be less than $\delta/4$. In fact, the functions E_Y and E_Z are trivially bounded by the following linear functions of R_i and R_{ij} with slope -1 by definition:

$$E_Y(R_i, R_{ij}, W_{Y\,|\,TUX}, P_{(TU)_i X_j}) \le \min\Big\{I_{P_{(TU)_i X_j} W_{Y\,|\,TUX}}(T, U, X; Y) - R_i - R_{ij}, I_{P_{(TU)_i X_j} W_{Y\,|\,TUX}}(X; Y\,|\,T, U) - R_{ij}\Big\} \quad (40)$$

and

$$E_Z\left(R_i, R_{ij}, W_{Z\,|\,TUX}, P_{(TU)_i X_j}\right) \le I_{P_{(TU)_i X_j} W_{Z\,|\,TUX}}(T, U; Z) - R_i. \quad (41)$$

If $R_i \ge H_{P_{(TU)_i}}(T, U) - \frac{\delta}{4} \ge I_{P_{(TU)_i X_j} W_{Z\,|\,TUX}}(T, U; Z) - \frac{\delta}{4}$, then by (41) $E_Z\left(R_i, R_{ij}, W_{Z\,|\,TUX}, P_{(TU)_i X_j}\right) \le \frac{\delta}{4}$. Similarly, if $R_{ij} \ge H_{P_{(TU)_i X_j}}(X\,|\,T, U) - \frac{\delta}{4}$, then by (40) $E_Y\left(R_i, R_{ij}, W_{Y\,|\,TUX}, P_{(TU)_i X_j}\right) \le \frac{\delta}{4}$.

Therefore, we may construct the JSC code $(f_n, g_n, \varphi_n, \psi_n)$ for CS Q_{SL} and the two-user channel $W_{YZ\,|\,UX}$ as follows. Without the loss of generality, we assume that the alphabets $\mathcal{U}$ and $\mathcal{X}$ contain the element 0.

<u>*Encoder* g_n</u>: For the message $\boldsymbol{s} \in \mathbb{T}_{S_i}$ such that $i > \hat{m}_n$, let $g_n(\boldsymbol{s}) = \boldsymbol{0} \in \mathcal{U}^n$. Denote $\tilde{\Omega} = \bigcup_i \Omega_i$. For the $\boldsymbol{s} \in \mathbb{T}_{S_i}$ such that $i \le \hat{m}_n$, let $g_n^{(1)} : \mathcal{S}^{\tau n} \to \tilde{\Omega}$ be a bijection that maps each $\boldsymbol{s} \in \mathbb{T}_{S_i}$ to the corresponding $(\boldsymbol{t}, \boldsymbol{u}) \in \Omega_i$, by noting that $|\Omega_i| = |\mathbb{T}_{S_i}| = N_i$. Finally, let $g_n(\boldsymbol{s})$ be the second component $\boldsymbol{u}$ of $g_n^{(1)}(\boldsymbol{s})$.

<u>*Encoder* f_n</u>: For the message pair $(\boldsymbol{s}, \boldsymbol{l}) \in \mathbb{T}_{S_i L_j}$ such that $i > \hat{m}_n$ or $j > \hat{m}'_{\text{in}}$, let $f_n(\boldsymbol{s}, \boldsymbol{l}) = \boldsymbol{0} \in \mathcal{X}^n$. For the $(\boldsymbol{s}, \boldsymbol{l}) \in$

$\mathbb{T}_{S_iL_j}$ such that $i \le \hat{m}_n$ and $j \le \hat{m}'_{\text{in}}$, noting that $|\mathbb{T}_{L_j|S_i}(\boldsymbol{s})| = |\Omega_{ij}(\varphi_n(\boldsymbol{s}))| = M_{ij}$, if $\boldsymbol{s} \in \mathbb{T}_{S_i}$, let $f_n^{(1)}(\boldsymbol{s},\cdot) : \mathbb{T}_{L_j|S_i}(\boldsymbol{s}) \to \Omega_{ij}(g_n(\boldsymbol{s}))$ be a bijection such that $f_n^{(1)}(\boldsymbol{s},\boldsymbol{l}) = (g_n^{(1)}(\boldsymbol{s}),\boldsymbol{x}) \in \Omega_{ij}$. Let $f_n(\boldsymbol{s},\boldsymbol{l})$ be the third component $\boldsymbol{x}$ of $f_n^{(1)}(\boldsymbol{s},\boldsymbol{l})$.

Clearly, the JSC encoders (f_n, g_n), although working independently, they map each $(\boldsymbol{s},\boldsymbol{l}) \in \mathbb{T}_{S_iL_j}$ to a unique pair $(\boldsymbol{u},\boldsymbol{x})$ when $i \le \hat{m}_n$ and $j \le \hat{m}'_{\text{in}}$, and to $(\cdot,\mathbf{0})$, otherwise (in this case an error is declared).

Y-Decoder φ_n: The Y-decoder is defined by the first expression at the bottom of the page.

Z-Decoder ψ_n: The Z-decoder is defined by the second expression at the bottom of the page.

For such JSC code $(f_n,\, g_n,\, \varphi_n,\, \psi_n)$, the probabilities of Y-error and Z-error are bounded by

$$P_{Ye}^{(n)}(\boldsymbol{s},\boldsymbol{l}) \le 2^{-n[E_Y(R_i,R_{ij},W_{Y\,|\,TUX},\tilde{P}_{TUX})-\delta/2]}, \quad \text{if } (\boldsymbol{s},\boldsymbol{l}) \in \mathbb{T}_{S_iL_j} \qquad (42)$$

and

$$P_{Ze}^{(n)}(\boldsymbol{s},\boldsymbol{l}) \le 2^{-nE_Z(R_i,R_{ij},W_{Z\,|\,TUX},\tilde{P}_{TUX})-\delta/2]}, \quad \text{if } (\boldsymbol{s},\boldsymbol{l}) \in \mathbb{T}_{S_iL_j}. \qquad (43)$$

Substituting (42) and (43) into (10) and using the fact (Lemma 1) $Q_{SL}^{(\tau n)}(\mathbb{T}_{SL}) \le 2^{-n\tau D(P_{SL}\,\|\,Q_{SL})}$, we obtain, for n sufficiently large, $P_{Ye}^{(n)}(Q_{SL}, W_{YZ\,|\,UX}, \tau)$ and $P_{Ze}^{(n)}(Q_{SL}, W_{YZ\,|\,UX}, \tau)$, shown in (44) and (45) at the bottom of the page, where

$$o_1(n) = \frac{|\mathcal{S}|\log_2(\tau n+1)}{n} \text{ and } o_2(n) = \frac{|\mathcal{S}||\mathcal{L}|\log_2(\tau n+1)}{n}.$$

Finally, the bounds (32) and (33) follow from (44) and (45), and the fact that the cardinality of set of joint types $\mathcal{P}_{\tau n}(\mathcal{S}\times\mathcal{L})$ is upper-bounded by $(\tau n+1)^{|\mathcal{S}||\mathcal{L}|}$.

To prove the lower bound (34), we slightly modify the above approach by choosing $P_{(TU)_iX_j} = \tilde{P}^*_{(TU)_iX_j}$ which achieves the maximum and the supremum of $E_r(R_i, R_{ij}, W_{YZ\,|\,UX})$ in (35) for every R_i and R_{ij}, $i = 1, 2, \ldots, m_n$, $j = 1, 2, \ldots, m'_{\text{in}}$. Then the probabilities of Y-error and Z-error in (42) and (43) are bounded by

$$\begin{aligned} P_{Ye}^{(n)}(\boldsymbol{s},\boldsymbol{l}) &\le 2^{-n[E_Y(R_i,R_{ij},W_{Y\,|\,TUX},\tilde{P}^*_{(TU)_iX_j})-\delta/2]} \\ &\le 2^{-n[E_r(R_i,R_{ij},W_{YZ\,|\,UX})-\delta/2]}, \quad \text{if } (\boldsymbol{s},\boldsymbol{l}) \in \mathbb{T}_{S_iL_j} \end{aligned} \qquad (46)$$

and

$$\begin{aligned} P_{Ze}^{(n)}(\boldsymbol{s},\boldsymbol{l}) &\le 2^{-nE_Z(R_i,R_{ij},W_{Z\,|\,TUX},\tilde{P}^*_{TU_iX_j})-\delta/2]} \\ &\le 2^{-n[E_r(R_i,R_{ij},W_{YZ\,|\,UX})-\delta/2]}, \quad \text{if } (\boldsymbol{s},\boldsymbol{l}) \in \mathbb{T}_{S_iL_j} \end{aligned} \qquad (47)$$

for n sufficiently large. The rest of the proof is similar to the proofs of (32) and (33). $\square$

By examining the positivity of the lower bound to E_J, we obtain a sufficient condition for reliable transmissibility for the asymmetric two-user system. For the sake of completeness, we also prove a converse by using Fano's inequality, and hence establish the JSCC theorem for this system. Given $W_{YZ\,|\,UX}$, define

$$\mathcal{R}(W_{YZ\,|\,UX}) \triangleq \bigcup_{\mathcal{T}:|\mathcal{T}|\le|\mathcal{U}||\mathcal{X}|+1}\ \bigcup_{P_{TUX}\in\mathcal{P}(\mathcal{T}\times\mathcal{U}\times\mathcal{X})} \mathcal{R}(W_{YZ\,|\,TUX}, P_{TUX}) \qquad (48)$$

$$\varphi_n(\boldsymbol{y}) \triangleq \begin{cases} (\boldsymbol{s}',\boldsymbol{l}'), & \text{if } \exists(\boldsymbol{s}',\boldsymbol{l}') \in \mathcal{S}^n\times\mathcal{L}^n \text{ such that } f_n^{(1)}(\boldsymbol{s}',\boldsymbol{l}') = \varphi_n^{(0)}(\boldsymbol{y}) \\ (\mathbf{0},\mathbf{0}), & \text{otherwise.} \end{cases}$$

$$\psi_n(\boldsymbol{z}) \triangleq \begin{cases} \boldsymbol{s}', & \text{if } \exists\ \boldsymbol{s}' \in \mathcal{S}^n \text{ such that } g_n^{(1)}(\boldsymbol{s}') \text{ is equal to the first two components of } \psi_n^{(0)}(\boldsymbol{z}) \\ \mathbf{0}, & \text{otherwise.} \end{cases}$$

$$\begin{aligned} P_{Ye}^{(n)}(Q_{SL}, W_{YZ\,|\,UX}, \tau) &\le \sum_{i,j} 2^{-n[\tau D(P_{S_iL_j}\|Q_{SL})+E_Y(R_i,R_{ij},W_{Y\,|\,TUX},\tilde{P}_{TUX})-\delta/2]} \\ &\le \sum_{P_{SL}} 2^{-n[\tau D(P_{SL}\|Q_{SL})+E_Y(\tau H_P(S)-o_1(n),\tau H_P(L\,|\,S)-o_2(n),W_{Y\,|\,TUX},\tilde{P}_{TUX})-\delta/2]} \\ &\le \sum_{P_{SL}} 2^{-n[\tau D(P_{SL}\|Q_{SL})+E_Y(\tau H_P(S),\tau H_P(L\,|\,S),W_{Y\,|\,TUX},\tilde{P}_{TUX})-\delta]} \end{aligned} \qquad (44)$$

and

$$\begin{aligned} P_{Ze}^{(n)}(Q_{SL}, W_{YZ\,|\,UX}, \tau) &\le \sum_{i,j} 2^{-n[\tau D(P_{S_iL_j}\|Q_{SL})+E_Z(R_i,R_{ij},W_{Z\,|\,TUX},\tilde{P}_{TUX})-\delta/2]} \\ &\le \sum_{P_{SL}} 2^{-n[\tau D(P_{SL}\|Q_{SL})+E_Z(\tau H_P(S)-o_1(n),\tau H_P(L\,|\,S)-o_2(n),W_{Z\,|\,TUX},\tilde{P}_{TUX})-\delta/2]} \\ &\le \sum_{P_{SL}} 2^{-n[\tau D(P_{SL}\|Q_{SL})+E_Z(\tau H_P(S),\tau H_P(L\,|\,S),W_{Z\,|\,TUX},\tilde{P}_{TUX})-\delta]} \end{aligned} \qquad (45)$$

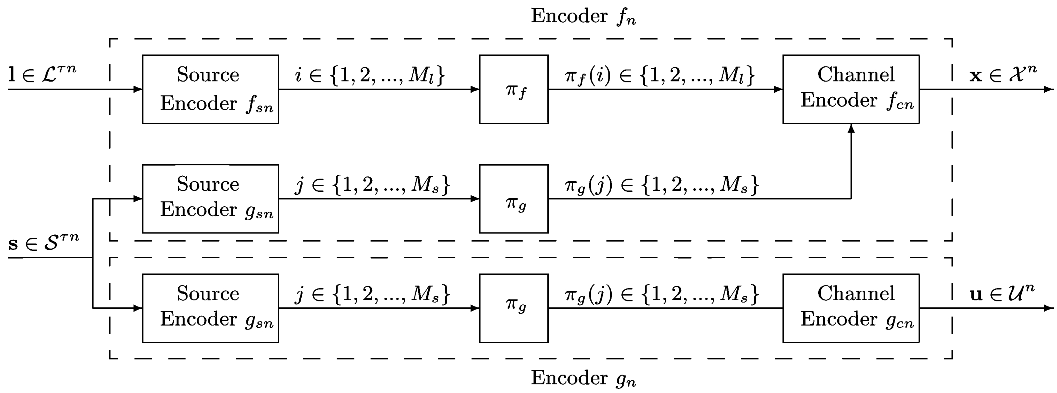

Fig. 4. Tandem source–channel coding system—encoders.

where

$$\mathcal{R}(W_{YZ\,|\,TUX}, P_{TUX}) \triangleq \left\{ (R_1, R_2) : \begin{array}{l} R_1 + R_2 < I(T,U,X;Y) = I(U,X;Y) \\ R_1 < I(T,U;Z) \\ R_2 < I(X;Y\,|\,T,U) \end{array} \right\}$$

where the mutual informations are taken under the joint distribution $P_{TUXYZ} = P_{TUX} W_{YZ\,|\,UX}$. Note that $\mathcal{R}(W_{YZ\,|\,UX})$ is convex and we denote $\bar{\mathcal{R}}(W_{YZ\,|\,UX})$ be the closure of $\mathcal{R}(W_{YZ\,|\,UX})$.

Theorem 3: (JSCC Theorem) Given Q_{SL}, $W_{YZ\,|\,UX}$, and $\tau > 0$, the following statements hold.

1) The sources Q_{SL} can be transmitted over the channel $W_{YZ\,|\,UX}$ with probability of error $P_e^{(n)} \to 0$ as $n \to \infty$ if $(\tau H_Q(S), \tau H_Q(L\,|\,S)) \in \mathcal{R}(W_{YZ\,|\,UX})$.
2) Conversely, if the sources Q_{SL} can be transmitted over the channel $W_{YZ\,|\,UX}$ with an arbitrarily small probability of error $P_e^{(n)}$ as $n \to \infty$, then $(\tau H_Q(S), \tau H_Q(L\,|\,S)) \in \bar{\mathcal{R}}(W_{YZ\,|\,UX})$.

Proof: See Appendix C. □

Observation 1: Theorem 3 implies that $\bar{\mathcal{R}}(W_{YZ\,|\,UX})$ is actually the capacity region for the asymmetric two-user channel $W_{YZ\,|\,UX}$, as the JSCC reduces to the asymmetric two-user channel coding if the sources come with a uniform joint distribution. It is shown in Appendix C that $\bar{\mathcal{R}}(W_{YZ\,|\,UX})$ can be equivalently written as

$$\mathcal{R}'(W_{YZ\,|\,UX}) \triangleq \bigcup_{\mathcal{T}:|\mathcal{T}|\le|\mathcal{U}||\mathcal{X}|+1} \; \bigcup_{P_{TUX}\in\mathcal{P}(\mathcal{T}\times\mathcal{U}\times\mathcal{X})} \mathcal{R}'(W_{YZ\,|\,TUX}, P_{TUX}) \quad (49)$$

where $\mathcal{R}'(W_{YZ\,|\,TUX}, P_{TUX})$ is given in the first expression at the bottom of the page. Recently, Liang *et al.* [23] also showed (using a different approach) that the capacity region for the asymmetric two-user channel is given by

$$\mathcal{R}''(W_{YZ\,|\,UX}) \triangleq \bigcup_{\mathcal{T}:|\mathcal{T}|\le|\mathcal{U}||\mathcal{X}|+1} \; \bigcup_{P_{TUX}\in\mathcal{P}(\mathcal{T}\times\mathcal{U}\times\mathcal{X})} \mathcal{R}''(W_{YZ\,|\,TUX}, P_{TUX}) \quad (50)$$

where $\mathcal{R}''(W_{YZ\,|\,TUX}, P_{TUX})$ is given in the second expression at the bottom of the page, where the mutual informations are taken under the joint distribution $P_{TUXYZ} = P_{TUX} W_{YZ\,|\,UX}$. They state that our capacity region, $\bar{\mathcal{R}}(W_{YZ\,|\,UX})$, is a subset of their region $\mathcal{R}''(W_{YZ\,|\,UX})$ described above by (50); this holds since $I(X;Y\,|\,T,U) \le I(X;Y\,|\,U)$. However, this is only partially correct, since noting that $\mathcal{R}''(W_{YZ\,|\,UX}) \subseteq \mathcal{R}'(W_{YZ\,|\,UX})$ and that $\mathcal{R}'(W_{YZ\,|\,UX}) = \bar{\mathcal{R}}(W_{YZ\,|\,UX})$ (as shown in Appendix C), one directly obtains that $\mathcal{R}''(W_{YZ\,|\,UX}) \subseteq \bar{\mathcal{R}}(W_{YZ\,|\,UX})$. Thus, the regions are all identical: $\bar{\mathcal{R}}(W_{YZ\,|\,UX}) = \mathcal{R}'(W_{YZ\,|\,UX}) = \mathcal{R}''(W_{YZ\,|\,UX})$.

D. Separation Principle for the Asymmetric Two-User System

It can be verified that the condition $(\tau H_Q(S), \tau H_Q(L\,|\,S)) \in \mathcal{R}(W_{YZ\,|\,UX})$ of Theorem 3 can be achieved by separate source and channel coding. The separate coding system of rate τ (source symbol/channel symbol) (we refer to it by the *tandem* coding system) is depicted in Figs. 4 and 5 (with π_f and π_g being identity mappings).

The encoder f_n is composed of two source encoders $f_{sn} : \mathcal{L}^{\tau n} \to \{1, 2, \ldots, M_l\}$ and $g_{sn} : \mathcal{S}^{\tau n} \to \{1, 2, \ldots, M_s\}$ with private source coding rate $\hat{R}_l \triangleq \frac{1}{\tau n} \log_2 M_l$ and common source coding rate $\hat{R}_s \triangleq \frac{1}{\tau n} \log_2 M_s$ and a channel encoder $\{1, 2, \ldots, M_l\} \times \{1, 2, \ldots, M_s\} \to \mathcal{X}^n$. Similarly, the encoder g_n is composed of a source encoder

$$\mathcal{R}'(W_{YZ\,|\,TUX}, P_{TUX}) \triangleq \left\{ (R_1, R_2) : \begin{array}{l} R_1 + R_2 \le \min\{I(U,X;Y), I(T,U;Z) + I(X;Y\,|\,T,U)\} \\ R_1 \le I(T,U;Z) \end{array} \right\}.$$

$$\mathcal{R}''(W_{YZ\,|\,TUX}, P_{TUX}) \triangleq \left\{ (R_1, R_2) : \begin{array}{l} R_1 + R_2 \le \min\{I(U,X;Y), I(T,U;Z) + I(X;Y\,|\,T,U)\} \\ R_1 \le I(T,U;Z) \\ R_2 \le I(X;Y\,|\,U) \end{array} \right\}$$

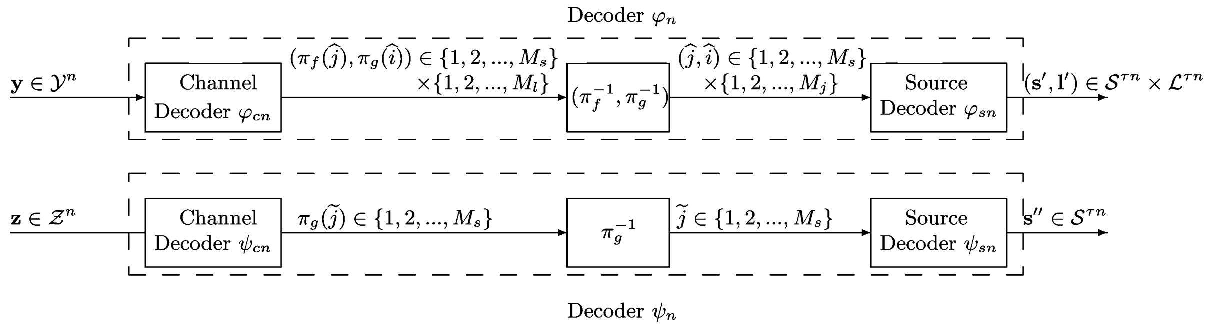

Fig. 5. Tandem source–channel coding system—decoders.

$g_{sn} : \mathcal{S}^{\tau n} \to \{1, 2, \ldots, M_s\}$ with common coding rate $\hat{R}_s$ and a channel encoder $g_{cn} : \{1, 2, \ldots, M_s\} \to \mathcal{U}^n$.

At the receiver side, the decoder φ_n is composed of a channel decoder $\varphi_{cn} : \mathcal{Y}^n \to \{1, 2, \ldots, M_l\} \times \{1, 2, \ldots, M_s\}$, and a source decoder $\varphi_{sn} : \{1, 2, \ldots, M_l\} \times \{1, 2, \ldots, M_s\} \to \mathcal{S}^{\tau n} \times \mathcal{L}^{\tau n}$ which outputs the approximation of the source messages $\boldsymbol{s}'$ and $\boldsymbol{l}'$. Similarly, the decoder ψ_n is composed of a channel decoder $\psi_{cn} : \mathcal{Z}^n \to \{1, 2, \ldots, M_s\}$, and a source decoder $\psi_{sn} : \{1, 2, \ldots, M_s\} \to \mathcal{S}^{\tau n}$.

To show that the condition

$$(\tau H_Q(S), \tau H_Q(L \,|\, S)) \in \mathcal{R}(W_{YZ\,|\,UX})$$

can be achieved by the above tandem system, we need to apply the following two-user source and channel coding theorems (we only state the forward parts of the theorems). Note that both of these theorems are special case of Theorem 3.

Let $(f_{sn}, g_{sn}, \varphi_{sn}, \psi_{sn})$ be a sequence of source codes for CS Q_{SL} with common source rate $\hat{R}_s$ and private source rate $\hat{R}_l$ as defined above. The probability of the overall two-user source coding error is given by

$$\begin{aligned} P_{es}^{(n)}&(\hat{R}_s, \hat{R}_l, Q_{SL}) \\ &\triangleq \Pr\Big(\{\varphi_{sn}(g_{sn}(S^{\tau n}), f_{sn}(L^{\tau n})) \neq (S^{\tau n}, L^{\tau n})\} \\ &\qquad \bigcup\{\psi_{sn}(g_{sn}(S^{\tau n})) \neq S^{\tau n}\}\Big). \end{aligned} \tag{51}$$

Then by the two-user source coding theorem, there exists a sequence of source codes $(f_{sn}, g_{sn}, \varphi_{sn}, \psi_{sn})$ with rates $\hat{R}_s$ and $\hat{R}_l$ such that $P_{es}^{(n)}(\hat{R}_s, \hat{R}_l, Q_{SL}) \to 0$ as $n \to \infty$ if the rates satisfy $\hat{R}_s > H_Q(S)$ and $\hat{R}_l > H_Q(L \,|\, S)$, i.e., $(\hat{R}_s, \hat{R}_l)$ lies in the upper-right infinite rectangle with vertex given by the point $(H_Q(S), H_Q(L \,|\, S))$.

We next state the forward part of channel coding theorem for the asymmetric two-user channel. Let the (common and private) message pair (j, i) be uniformly drawn from the finite set $\mathcal{M}_s \times \mathcal{M}_l$, where $\mathcal{M}_s \triangleq \{1, 2, \ldots, M_s\}$ and $\mathcal{M}_l \triangleq \{1, 2, \ldots, M_l\}$, and let $(f_{cn}, g_{cn}, \varphi_{cn}, \psi_{cn})$ be an asymmetric two-user channel code with block length n and common and private message sets $\mathcal{M}_s$ and $\mathcal{M}_l$. Let $R_s \triangleq \frac{1}{n}\log_2 M_s$ and $R_l \triangleq \frac{1}{n}\log_2 M_l$ be the common and private rates of the channel code, respectively. The average probability of error for asymmetric two-user channel coding is given by

$$\begin{aligned} P_{ec}^{(n)}&(R_s, R_l, W_{YZ\,|\,UX}) \\ &\triangleq \Pr(\{\varphi_{cn}(Y^n) \neq (J, I)\} \bigcup\{\psi_{cn}(Z^n) \neq J\}) \end{aligned} \tag{52}$$

where (J, I) are uniformly drawn from $\mathcal{M}_s \times \mathcal{M}_l$. The maximum probability for error of asymmetric two-user channel coding is given by

$$\begin{aligned} P_{ec,\max}^{(n)}&(R_s, R_l, W_{YZ\,|\,UX}) \\ &\triangleq \max_{(j,i)\in\mathcal{M}_s\times\mathcal{M}_l} \Pr\Big(\{\varphi_{cn}(Y^n) \neq (J, I)\} \\ &\qquad \bigcup\{\psi_{cn}(Z^n) \neq J\} \,|\, J = j, I = i\Big). \end{aligned} \tag{53}$$

Then there exists a sequence of channel codes $(f_{cn}, g_{cn}, \varphi_{cn}, \psi_{cn})$ such that $P_{ec}^{(n)}(R_s, R_l, W_{YZ\,|\,UX}) \to 0$ as $n \to \infty$ if $(R_s, R_l) \in \mathcal{R}(W_{YZ\,|\,UX})$. Furthermore, it can be readily shown by a standard expurgation argument [6, p. 204] that $P_{ec,\max}^{(n)}(R_s, R_l, W_{YZ\,|\,UX}) \to 0$ as $n \to \infty$ if $(R_s, R_l) \in \mathcal{R}(W_{YZ\,|\,UX})$.

Now by (14), the overall probability of error for the tandem system is given by

$$\begin{aligned} P_e^{(n)} \triangleq \Pr\Big(&\{\varphi_{sn}[\varphi_{cn}(Y^n)] \neq (S^{\tau n}, L^{\tau n})\} \\ &\bigcup\{\psi_{sn}[\psi_{cn}(Z^n)] \neq S^{\tau n}\}\Big). \end{aligned}$$

By the union bound, it is easy to see that $P_e^{(n)}$ is upper-bounded by

$$\begin{aligned} P_e^{(n)} &\leq \Pr\Big(\{\varphi_{sn}(g_{sn}(S^{\tau n}), f_{sn}(L^{\tau n})) \neq (S^{\tau n}, L^{\tau n})\} \\ &\qquad \bigcup\{\psi_{sn}(g_{sn}(S^{\tau n})) \neq S^{\tau n}\}\Big) \\ &\quad + \Pr\Big(\{\varphi_{cn}(Y^n) \neq (g_{sn}(S^{\tau n}), f_{sn}(L^{\tau n}))\} \\ &\qquad \bigcup\{\psi_{cn}(Z^n) \neq g_{sn}(S^{\tau n})\}\Big) \\ &= P_{es}^{(n)}(\hat{R}_s, \hat{R}_l, Q_{SL}) \\ &\quad + \sum_{(j,i)\in\mathcal{M}_s\times\mathcal{M}_l} \Pr\left(g_{sn}(S^{\tau n}) = j, f_{sn}(L^{\tau n}) = i\right) \\ &\quad \times \Pr\Big(\{\varphi_{cn}(Y^n) \neq (J, I)\} \\ &\qquad \bigcup\{\psi_{cn}(Z^n) \neq J\} \,|\, J = j, I = i\Big) \\ &\leq P_{es}^{(n)}(\hat{R}_s, \hat{R}_l, Q_{SL}) + P_{ec,\max}^{(n)}(\tau\hat{R}_s, \tau\hat{R}_l, W_{YZ\,|\,UX}) \end{aligned}$$

where $P_{ec,\max}^{(n)}(\tau\hat{R}_s, \tau\hat{R}_l, W_{YZ\,|\,UX})$ is the maximum channel coding probability of error with common rate $\tau\hat{R}_s$ and private rate $\tau\hat{R}_l$. Clearly, by combining the two-user source coding theorem and the asymmetric two-user channel coding theorem, if

$(\tau H_Q(S), \tau H_Q(L\,|\,S)) \in \mathcal{R}(W_{YZ\,|\,UX})$, then there exist a sequence of source codes $(f_{sn}, g_{sn}, \varphi_{sn}, \psi_{sn})$ and a sequence of channel codes $(f_{cn}, g_{cn}, \varphi_{cn}, \psi_{cn})$ such that the overall tandem system probability of error $P_e^{(n)} \to 0$ as $n \to \infty$. Therefore, separation of source and channel coding is optimal from the point of view of reliable transmissibility.

E. The Upper Bound to E_J

In [8], Csiszár also established an upper bound for the JSCC error exponent for the point-to-point discrete memoryless source–channel system in terms of the source and channel error exponents by a simple type counting argument. He shows that the JSCC error exponent is always less than the infimum of the sum of the source and channel error exponent, even though the channel error exponent is only partially known for high rates. This conceptual bound cannot currently be computed as the channel error exponent is not yet fully known for all achievable coding rates, but it directly implies that any upper bound for the channel error exponent yields a corresponding upper bound for the JSCC error exponent. For the asymmetric two-user channel, a similar bound can be shown.

Definition 2: The asymmetric two-user channel coding error exponent $E(R_1, R_2, W_{YZ\,|\,UX})$, for any $R_1 > 0$ and $R_2 > 0$, is defined by the supremum of the set of all numbers E_c for which there exists a sequence of asymmetric channel codes $(f_{cn}, g_{cn}, \varphi_{cn}, \psi_{cn})$ with block length n, the common rate no less than R_1, and the private rate no less than R_2, such that

$$E_c \le \liminf_{n\to\infty} -\frac{1}{n}\log_2 P_{ec}^{(n)}(R_1, R_2, W_{YZ\,|\,UX}). \tag{54}$$

Denote the probabilities of Y- and Z-error of the channel coding by

$$\begin{aligned} &P_{Yec}^{(n)}(R_s, R_l, W_{YY|UX}) \\ &\quad \triangleq \Pr(\{\varphi_{cn}(Y^n) \ne (J,I)\}) \\ &\quad = \frac{1}{2^{R_1+R_2}} \sum_{\mathcal{M}_s\times\mathcal{M}_l} \sum_{\boldsymbol{y}:\varphi_{cn}(\boldsymbol{y})\ne(j,i)} W_{Y|X}^{(n)}(\boldsymbol{y}|\boldsymbol{u},\boldsymbol{x}) \end{aligned} \tag{55}$$

and

$$\begin{aligned} &P_{Zec}^{(n)}(R_s, R_l, W_{YZ\,|\,UX}) \\ &\quad \triangleq \Pr(\{\psi_{cn}(Z^n) \ne J\}) \\ &\quad = \frac{1}{2^{R_1+R_2}} \sum_{\mathcal{M}_s\times\mathcal{M}_l} \sum_{\boldsymbol{z}:\psi_{cn}(\boldsymbol{z})\ne j} W_{Z|X}^{(n)}(\boldsymbol{z}\,|\,\boldsymbol{u},\boldsymbol{x}) \end{aligned} \tag{56}$$

where $\boldsymbol{x} \triangleq f_{cn}(j,i)$ and $\boldsymbol{u} \triangleq g_{cn}(j)$. Clearly, for any sequence of channel codes $(f_{cn}, g_{cn}, \varphi_{cn}, \psi_{cn})$, $P_e^{(n)}(R_1, R_2, W_{YZ\,|\,UX})$ must be larger than $P_{Yec}^{(n)}(R_1, R_2, W_{Y\,|\,UX})$ and $P_{Zec}^{(n)}(R_1, R_2, W_{Z|UX}))$ but less than the sum of the two, so we have

$$\begin{aligned} &\liminf_{n\to\infty} -\frac{1}{n}\log_2 P_{ec}^{(n)}(R_1, R_2, W_{YZ\,|\,UX}) \\ &\quad = \liminf_{n\to\infty} -\frac{1}{n}\log_2 \max\Big(P_{Yec}^{(n)}(R_1, R_2, W_{Y\,|\,UX}), \\ &\qquad\qquad P_{Zec}^{(n)}(R_1, R_2, W_{Z|UX})\Big). \end{aligned} \tag{57}$$

Our upper bound for the system JSCC error exponent E_J (defined in Definition 1) is stated as follows.

Theorem 4: Given Q_{SL}, $W_{YZ\,|\,UX}$, and τ, the system JSCC error exponent satisfies

$$\begin{aligned} &E_J(Q_{SL}, W_{YZ\,|\,UX}, \tau) \\ &\quad \le \inf_{P_{SL}} [\tau D(P_{SL}\,\|\,Q_{SL}) + E(\tau H_P(S), \\ &\qquad\qquad \tau H_P(L\,|\,S), W_{YZ\,|\,UX})] \end{aligned} \tag{58}$$

where $E(\cdot,\cdot,W_{YZ\,|\,UX})$ is the corresponding channel coding error exponent for the asymmetric two-user channel as defined above in Definition 2.

Proof: First, from (10) we can write

$$\begin{aligned} &P_{ie}^{(n)}(Q_{SL}, W_{YZ\,|\,UX}, \tau) \\ &\quad \ge \max_{P_{SL}\in\mathcal{P}_{\tau n}(\mathcal{S}\times\mathcal{L})} Q_{SL}^{(\tau n)}(\mathbb{T}_{SL}) P_{ie}(\mathbb{T}_{SL}), \quad i = Y, Z \end{aligned} \tag{59}$$

where $P_{Ye}(\mathbb{T}_{SL})$ and $P_{Ze}(\mathbb{T}_{SL})$ are given by (11) and (12), respectively. Comparing (11) with (55), and comparing (12) with (56), we note that $P_{Ye}(\mathbb{T}_{SL})$ and $P_{Ze}(\mathbb{T}_{SL})$ can be interpreted as the probabilities of Y-error and Z-error of the asymmetric two-user channel coding with (common and private) message sets $\mathbb{T}_{SL}$, since $(\boldsymbol{s},\boldsymbol{l})$ are uniformly distributed on $\mathbb{T}_{SL}$. For any $P_{SL} \in \mathcal{P}_{\tau n}(\mathcal{S}\times\mathcal{L})$, let P_S and $P_{L|S}$ be the marginal and conditional distributions induced by P_{SL}. Recall that for each $\boldsymbol{s} \in \mathbb{T}_S = \mathbb{T}_{P_S}$

$$\mathbb{T}_{L|S}(\boldsymbol{s}) \triangleq \mathbb{T}_{P_{L|S}}(\boldsymbol{s}) = \{\boldsymbol{l} : (\boldsymbol{s},\boldsymbol{l}) \in \mathbb{T}_{SL}\}$$

and that $\mathbb{T}_{L|S}(\boldsymbol{s})$ is the same set for all $\boldsymbol{s} \in \mathbb{T}_S$. Hence, we can write $\mathbb{T}_{SL}$ by the product of two sets $\mathbb{T}_{SL} = \mathbb{T}_S \times \mathbb{T}_{L|S}(\boldsymbol{s})$. Setting $\tilde{R}_1 = \frac{1}{n}\log_2|\mathbb{T}_S|$ and $\tilde{R}_2 = \frac{1}{n}\log_2|\mathbb{T}_{L|S}(\boldsymbol{s})|$, it follows that, by the definition of asymmetric two-user channel coding error exponent and (57)

$$\begin{aligned} &\liminf_{n\to\infty} -\frac{1}{n}\log_2 \max_{i=Y,Z} P_{ie}(\mathbb{T}_{SL}) \\ &\quad \le E(\liminf_{n\to\infty} \tilde{R}_1, \liminf_{n\to\infty} \tilde{R}_2, W_{YZ\,|\,UX}) \\ &\quad = E(\tau H_P(S), \tau H_P(L\,|\,S), W_{YZ\,|\,UX}) \end{aligned} \tag{60}$$

for any sequence of JSC codes (f_n, φ_n, ψ_n), recalling from Lemma 1 that

$$(\tau n + 1)^{-|\mathcal{S}|} 2^{n\tau H_P(S)} \le |\mathbb{T}_S| \le 2^{n\tau H_P(S)}$$

and

$$(\tau n + 1)^{-|\mathcal{S}||\mathcal{L}|} 2^{n\tau H_P(L\,|\,S)} \le |\mathbb{T}_{L|S}(\boldsymbol{s})| \le 2^{n\tau H_P(L\,|\,S)}.$$

According to (16), we write

$$\begin{aligned} &\liminf_{n\to\infty} -\frac{1}{n}\log_2 P_e^{(n)}(Q_{SL}, W_{YZ\,|\,UX}, \tau) \\ &\quad = \liminf_{n\to\infty} -\frac{1}{n}\log_2 \max\Big(P_{Ye}^{(n)}(Q_{SL}, W_{Y|X}, \tau), \\ &\qquad\qquad P_{Ze}^{(n)}(Q_{SL}, W_{Z|X}, \tau)\Big) \\ &\quad \le \liminf_{n\to\infty} -\frac{1}{n}\log_2 \max_{i=Y,Z} \max_{P_{SL}\in\mathcal{P}_{\tau n}(\mathcal{S}\times\mathcal{L})} Q_{SL}^{(\tau n)} \\ &\qquad (\mathbb{T}_{SL}) P_{ie}(\mathbb{T}_{SL}) \end{aligned}$$

$$= \liminf_{n\to\infty} \min_{P_{SL}\in\mathcal{P}_{\tau n}(\mathcal{S}\times\mathcal{L})} -\frac{1}{n}\log_2 Q_{SL}^{(\tau n)} (\mathbb{T}_{SL}) \max_{i=Y,Z} P_{ie}(\mathbb{T}_{SL})$$
$$= \liminf_{n\to\infty} \min_{P_{SL}\in\mathcal{P}_{\tau n}(\mathcal{S}\times\mathcal{L})} \left[-\frac{1}{n}\log_2 Q_{SL}^{(\tau n)}(\mathbb{T}_{SL}) - \frac{1}{n}\log_2 \max_{i=Y,Z} P_{ie}(\mathbb{T}_{SL})\right]. \quad (61)$$

By Lemma 1, for any $P_{SL}\in\mathcal{P}_{\tau n}(\mathcal{S}\times\mathcal{L})$

$$-\frac{1}{\tau n}\log_2 Q_{SL}^{(\tau n)}(\mathbb{T}_{SL}) \le D(P_{SL}\,\|\,Q_{SL}) + |\mathcal{S}||\mathcal{L}|\frac{1}{\tau n}\log_2(1+\tau n)$$

which implies

$$\limsup_{n\to\infty} -\frac{1}{n}\log_2 Q_{SL}^{(\tau n)}(\mathbb{T}_{SL}) \le \tau D(P_{SL}\,\|\,Q_{SL}). \quad (62)$$

Now assume that

$$\inf_{P_{SL}\in\mathcal{P}(\mathcal{S}\times\mathcal{L})} [\tau D(P_{SL}\,\|\,Q_{SL}) + E(\tau H_P(S), \tau H_P(L\,|\,S), W_{YZ\,|\,UX})]$$

is finite (the upper bound is trivial if it is infinity) and the infimum actually becomes a minimum. Let the minimum be achieved by distribution $P_{SL}^{*}\in\mathcal{P}(\mathcal{S}\times\mathcal{L})$, then there must exists a sequence of types $\{\hat{P}_{SL}\in\mathcal{P}_{\tau n}(\mathcal{S}\times\mathcal{L})\}_{n=n_o}^{\infty}$ such that $\hat{P}_{SL}\to P_{SL}^{*}$ uniformly.[3] It then follows from (61)—(62) that

$$\liminf_{n\to\infty} -\frac{1}{n}\log_2 P_e^{(n)}(Q_{SL}, W_{YZ\,|\,UX}, \tau) \le \liminf_{n\to\infty}\left[-\frac{1}{n}\log_2 Q_{SL}^{(\tau n)}(\mathbb{T}_{\hat{P}_{SL}}) - \frac{1}{n}\log_2 \max_{i=Y,Z} P_{ie}(\mathbb{T}_{\hat{P}_{SL}})\right]$$
$$\le \tau D(P_{SL}^{*}\,\|\,Q_{SL}) + E(\tau H_{P^*}(S), \tau H_{P^*}(L\,|\,S), W_{YZ\,|\,UX}). \quad (63)$$

Since the above bound holds for any sequence of JSC codes, we complete the Proof of Theorem 4. □

V. Applications to CS-AMAC and CS-ABC Systems

As pointed out in the Introduction, our results obtained in the previous section can be directly applied to the CS-AMAC and CS-ABC source–channel systems.

A. CS-AMAC System

Setting $|\mathcal{Z}|=1$ and removing the decoder ψ_n, the two-user asymmetric channel $W_{YZ\,|\,UX}$ reduces to an AMAC $W_{Y\,|\,UX}$. Since the CS-AMAC system is a special case of the two-user system, the quantities defined before, including the system (overall) probability of error, the system JSCC error exponent, and the channel error exponent still hold for the CS-AMAC system. Note that there is only one decoder, so we do not have the Z-error probability (nor exponent) here. The first union in (48) can be removed since the largest region is given by $|\mathcal{T}|=1$. In fact, for any $T\to(U,X)\to Y$, $I(T,U,X;Y)=I(U,X;Y)$ and $I(X;Y\,|\,T,U)\le I(X;Y\,|\,U)$. Thus, Theorem 3 reduces to the same JSCC theorem established in [4] for the CS-AMAC system. Choosing the auxiliary alphabet $|\mathcal{T}|=1$, we specialize Theorems 2 and 4 to the following corollary.

Corollary 1: Given Q_{SL}, $W_{Y\,|\,UX}$, and τ, the system JSCC error exponent satisfies

$$E_J(Q_{SL}, W_{Y\,|\,UX}, \tau) \ge \min_{P_{SL}} [\tau D(P_{SL}\,\|\,Q_{SL}) + E_r(\tau H_P(S), \tau H_P(L\,|\,S), W_{Y\,|\,UX})] \quad (64)$$

and

$$E_J(Q_{SL}, W_{Y\,|\,UX}, \tau) \le \inf_{P_{SL}} [\tau D(P_{SL}\,\|\,Q_{SL}) + E(\tau H_P(S), \tau H_P(L\,|\,S), W_{Y\,|\,UX})] \quad (65)$$

where $E(\tau H_P(S), \tau H_P(L\,|\,S), W_{Y\,|\,UX})$ is the channel error exponent of the AMAC $W_{Y\,|\,UX}$ defined in (54) with $|\mathcal{Z}|=1$, and

$$E_r(R_1, R_2, W_{Y\,|\,UX}) = \max_{P_{UX}} E_Y(R_1, R_2, W_{Y\,|\,UX}, P_{UX}) \quad (66)$$

where $E_Y(R_1, R_2, W_{Y\,|\,UX}, P_{UX})$ is defined in (17) with $|\mathcal{T}|=1$.

It has been shown in [2] that for any $R_1>0$ and $R_2>0$, the channel exponent for AMAC $W_{Y\,|\,UX}$ satisfies

$$E(R_1, R_2, W_{YZ\,|\,X}) \le E_{sp}(R_1, R_2, W_{Y\,|\,UX})$$

where

$$E_{sp}(R_1, R_2, W_{Y\,|\,UX}) \triangleq \max_{P_{UX}\in\mathcal{P}(\mathcal{U}\times\mathcal{X})} \min D(V_{Y\,|\,UX}\,\|\,W_{Y\,|\,UX}\,|\,P_{UX}) \quad (67)$$

where the minimum is taken over $V_{Y\,|\,UX}\in\mathcal{P}(\mathcal{Y}\,|\,\mathcal{U}\times\mathcal{X})$ such that

$$I_{P_{UX}V_{Y\,|\,UX}}(U,X;Y)\le R_1+R_2$$

or

$$I_{P_{UX}V_{Y\,|\,UX}}(X;Y\,|\,U)\le R_2.$$

As a consequence, we obtain that

$$E_J(Q_{SL}, W_{Y\,|\,UX}, \tau) \le \inf_{P_{SL}} [\tau D(P_{SL}\,\|\,Q_{SL}) + E_{sp}(\tau H_P(S), \tau H_P(L\,|\,S), W_{Y\,|\,UX})]. \quad (68)$$

In Section VI, we investigate the evaluation of lower bound (64) and upper bound (68) when the AMAC has a symmetric distribution.

[3] The sequence $\{\hat{P}_{SL}\in\mathcal{P}_{\tau n}(\mathcal{S}\times\mathcal{L})\}_{n=n_o}^{\infty}$ here denotes a sequence for $n=n_o, 2n_o, 3n_o, \ldots$, where n_o is the smallest integer such that τn_o is also an integer.

B. CS-ABC System

Setting $|\mathcal{U}| = 1$ and removing the encoder g_n, the two-user asymmetric channel $W_{YZ\,|\,UX}$ reduces to an ABC $W_{YZ\,|\,X}$. The quantities defined before, including the probabilities of error at Y-decoder and Z-decoder, the achievable error exponent pair, system (overall) probability of error, the system JSCC error exponent, and the channel error exponent still hold for the CS-ABC system. Given an arbitrary and finite auxiliary alphabet $\mathcal{T}$, we augment the channel $W_{YZ\,|\,X}$ to $W_{YZ|TX}$ by introducing an RV $T \in \mathcal{T}$ such that $T \to X \to (YZ)$. Similarly, the marginal distributions of the augmented channel are denoted by $W_{Y\,|\,TX}$ and $W_{Z|TX}$. We then specialize Theorems 2–4 to the following corollaries.

Given $W_{YZ\,|\,X}$, $\mathcal{R}(W_{YZ\,|\,UX})$ of (48) reduces to $\mathcal{R}(W_{YZ\,|\,X})$ given by

$$\mathcal{R}(W_{YZ\,|\,X}) \triangleq \bigcup_{\mathcal{T}:|\mathcal{T}|\le|\mathcal{X}|+1} \bigcup_{P_{TX}\in\mathcal{P}(\mathcal{T}\times\mathcal{X})} \mathcal{R}(W_{YZ|TX}, P_{TX}) \tag{69}$$

where

$$\mathcal{R}(W_{YZ|TX}, P_{TX}) = \left\{ (R_1,R_2):\ \begin{array}{l} R_1+R_2 < I(T,X;Y) = I(X;Y) \\ R_1 < I(T;Z) \\ R_2 < I(X;Y\,|\,T) \end{array} \right\}$$

where the mutual informations are taken under the joint distribution $P_{TXYZ} = P_{TX}W_{YZ\,|\,X}$. We remark that the closure of $\mathcal{R}(W_{YZ\,|\,X})$, denoted by $\bar{\mathcal{R}}(W_{YZ\,|\,X})$, is the capacity region of the ABC $W_{YZ\,|\,X}$ [21].

Corollary 2: (JSCC Theorem for CS-ABC system) Given $Q_{SL}, W_{YZ\,|\,X}$, and $\tau > 0$, the following statements hold.

1) The sources Q_{SL} can be transmitted over the ABC $W_{YZ\,|\,X}$ with $P_e^{(n)} \to 0$ as $n \to \infty$ if

$$(\tau H_Q(S), \tau H_Q(L\,|\,S)) \in \mathcal{R}(W_{YZ\,|\,X}).$$

2) Conversely, if the sources Q_{SL} can be transmitted over the ABC $W_{YZ\,|\,X}$ with an arbitrarily small probability of error $P_e^{(n)}$ as $n \to \infty$, then $(\tau H_Q(S), \tau H_Q(L\,|\,S)) \in \bar{\mathcal{R}}(W_{YZ\,|\,X})$.

Corollary 3: Given an arbitrary and finite alphabet $\mathcal{T}$, for any $\tilde{P}_{TX} \in \mathcal{P}(\mathcal{T}\times\mathcal{X})$, the following exponent pair is universally achievable:

$$\begin{aligned} &E_{JY}(Q_{SL}, W_{YZ|TX}, \tilde{P}_{TX}, \tau) \\ &\quad\triangleq \min_{P_{SL}}[\tau D(P_{SL}\,\|\,Q_{SL}) + E_Y(\tau H_P(S), \\ &\qquad\qquad \tau H_P(L\,|\,S), W_{Y\,|\,TX}, \tilde{P}_{TX})] \end{aligned} \tag{70}$$

and

$$\begin{aligned} &E_{JZ}(Q_{SL}, W_{YZ|TX}, \tilde{P}_{TX}, \tau) \\ &\quad\triangleq \min_{P_{SL}}[\tau D(P_{SL}\,\|\,Q_{SL}) + E_Z(\tau H_P(S), \\ &\qquad\qquad \tau H_P(L\,|\,S), W_{Z|TX}, \tilde{P}_{TX})] \end{aligned} \tag{71}$$

where E_Y and E_Z are defined in (17) and (18) by setting $|\mathcal{U}| = 1$. Furthermore, given Q_{SL}, $W_{YZ\,|\,X}$, and τ, the system JSCC error exponent satisfies

$$\begin{aligned} &E_J(Q_{SL}, W_{YZ\,|\,X}, \tau) \\ &\quad\ge \min_{P_{SL}}[\tau D(P_{SL}\,\|\,Q_{SL}) + E_r(\tau H_P(S), \\ &\qquad\qquad \tau H_P(L\,|\,S), W_{YZ\,|\,X})] \end{aligned} \tag{72}$$

and

$$\begin{aligned} &E_J(Q_{SL}, W_{YZ\,|\,X}, \tau) \\ &\quad\le \inf_{P_{SL}}[\tau D(P_{SL}\,\|\,Q_{SL}) + E(\tau H_P(S), \\ &\qquad\qquad \tau H_P(L\,|\,S), W_{YZ\,|\,X})] \end{aligned} \tag{73}$$

where $E_r(R_1,R_2,W_{YZ\,|\,X})$ is given by $E_r(R_1,R_2,W_{YZ\,|\,UX})$ in (35) with $|\mathcal{U}| = 1$, and $E(R_1,R_2,W_{YZ\,|\,X})$ is the channel error exponent for the ABC $W_{YZ\,|\,X}$.

VI. Evaluation of the Bounds for E_J: CS Over Symmetric AMAC

We established the lower and upper bounds for the JSCC error exponent of the asymmetric two-user JSCC system. However, we are not able to simplify these bounds for general two-user JSCC systems (not even for general CS-AMAC and CS-ABC systems) into computable parametric forms as we did for the point-to-point systems [29], [30]. In the following, we only address a special case of CS-AMAC systems where the channel admits a symmetric transition probability distribution. We first introduce the parametric forms of functions $E_r(R_1,R_2,W_{Y\,|\,UX})$ and $E_{sp}(R_1,R_2,W_{Y\,|\,UX})$ defined in (66) and (67), respectively. For any $R_1, R_2 > 0$, rewrite

$$\begin{aligned} &E_Y(R_1,R_2,W_{Y\,|\,UX},P_{UX}) \\ &\quad= \min\Big\{E_r^{(1)}(R_1+R_2, W_{Y\,|\,UX}, P_{UX}), \\ &\qquad\qquad E_r^{(2)}(R_2, W_{Y\,|\,UX}, P_{UX})\Big\} \end{aligned}$$

where

$$\begin{aligned} &E_r^{(1)}(R, W_{Y\,|\,UX}, P_{UX}) \\ &\quad\triangleq \min_{V_{Y\,|\,UX}}[D(V_{Y\,|\,UX}\,\|\,W_{Y\,|\,UX}|P_{UX}) \\ &\qquad\qquad + |I_{P_{UX}V_{Y\,|\,UX}}(U,X;Y) - R|^+] \end{aligned} \tag{74}$$

and

$$\begin{aligned} &E_r^{(2)}(R, W_{Y\,|\,UX}, P_{UX}) \\ &\quad\triangleq \min_{V_{Y\,|\,UX}}[D(V_{Y\,|\,UX}\,\|\,W_{Y\,|\,UX}|P_{UX}) \\ &\qquad\qquad + |I_{P_{UX}V_{Y\,|\,UX}}(X;Y\,|\,U) - R|^+]. \end{aligned} \tag{75}$$

Also, rewrite

$$E_{sp}(R_1,R_2,W_{Y\,|\,UX}) = \max_{P_{UX}} E_{sp}(R_1,R_2,W_{Y\,|\,UX},P_{UX})$$

where

$$\begin{aligned} &E_{sp}(R_1,R_2,W_{Y\,|\,UX},P_{UX}) \\ &\quad= \min\Big\{E_{sp}^{(1)}(R_1+R_2, W_{Y\,|\,UX}, P_{UX}), \\ &\qquad\qquad E_{sp}^{(2)}(R_2, W_{Y\,|\,UX}, P_{UX})\Big\} \end{aligned}$$

where

$$E_{sp}^{(1)}(R, W_{Y\,|\,UX}, P_{UX}) \triangleq \min_{V_{Y\,|\,UX}} \big(D(V_{Y\,|\,UX} \parallel W_{Y\,|\,UX}|P_{UX}) : I_{P_{UX}V_{Y\,|\,UX}}(U, X; Y) \le R\big) \quad (76)$$

and

$$E_{sp}^{(2)}(R, W_{Y\,|\,UX}, P_{UX}) \triangleq \min_{V_{Y\,|\,UX}} \big(D(V_{Y\,|\,UX} \parallel W_{Y\,|\,UX}|P_{UX}) : I_{P_{UX}V_{Y\,|\,UX}}(X; Y\,|\,U) \le R\big). \quad (77)$$

Note that $E_r^{(1)}$ and $E_r^{(2)}$ (respectively, $E_{sp}^{(1)}$ and $E_{sp}^{(2)}$) are the random-coding (respectively, sphere-packing) type exponents expressed in terms of constrained Kullback–Leibler divergences and mutual informations [10]. In fact, it has been shown in [2] that

$$E_{sp}^{(i)}(R, W_{Y\,|\,UX}, P_{UX}) = \max_{\rho \ge 0}[E_i(\rho, W_{Y\,|\,UX}, P_{UX}) - \rho R], \quad i = 1, 2$$

where

$$E_1(\rho_1, W_{Y\,|\,UX}, P_{UX}) \triangleq -\log_2 \sum_{y \in \mathcal{Y}} \left(\sum_{(u,x) \in \mathcal{U} \times \mathcal{X}} P_{UX}(u, x) \times W_{Y\,|\,UX}(y|u, x)^{\frac{1}{1+\rho_1}} \right)^{1+\rho_1} \quad (78)$$

and

$$E_2(\rho_2, W_{Y\,|\,UX}, P_{UX}) = -\log_2 \sum_{u \in \mathcal{U}} P_U(u) \sum_{y \in \mathcal{Y}} \left(\sum_{x \in \mathcal{X}} P_{X\,|\,U}(x|u) \times W_{Y\,|\,UX}(y|u, x)^{\frac{1}{1+\rho_2}} \right)^{1+\rho_2}. \quad (79)$$

Analogously to [10, Lemma 5.4, Corollary 5.4, p. 168], we can prove the following results; some of them has been proved in [2].

Lemma 4: Let $i = 1, 2$. $E_r^{(i)}(R, W_{Y\,|\,UX}, P_{UX})$ coincides with $E_{sp}^{(i)}(R, W_{Y\,|\,UX}, P_{UX})$ if $R \ge R_{cr}^{(i)}(W_{Y\,|\,UX}, P_{UX})$ where

$$R_{cr}^{(i)}(W_{Y\,|\,UX}, P_{UX}) = \left. \frac{\partial E_i(\rho, W_{Y\,|\,UX}, P_{UX})}{\partial \rho} \right|_{\rho=1}$$

and is a straight-line tangent on $E_{sp}^{(i)}(R, W_{Y\,|\,UX}, P_{UX})$ with slope -1 if $R \le R_{cr}^{(i)}(W_{Y\,|\,UX}, P_{UX})$, i.e., $E_r^{(i)}(R, W_{Y\,|\,UX}, P_{UX})$ is given in the expression at the bottom of the page. Furthermore, $E_r^{(i)}(R, W_{Y\,|\,UX}, P_{UX})$ has the parametric form

$$E_r^{(i)}(R, W_{Y\,|\,UX}, P_{UX}) = \max_{0 \le \rho \le 1}[E_i(\rho, W_{Y\,|\,UX}, P_{UX}) - \rho R]$$

where $E_1(\rho, W_{Y\,|\,UX}, P_{UX})$ and $E_2(\rho, W_{Y\,|\,UX}, P_{UX})$ are given in (78) and (79), respectively.

Therefore, we can write the functions $E_r(R_1, R_2, W_{Y\,|\,UX})$ in (66) and $E_{sp}(R_1, R_2, W_{Y\,|\,UX})$ in (67) as follows.

$$E_r(R_1, R_2, W_{Y\,|\,UX}) = \max_{P_{UX}} \min_{i=1,2} \max_{0 \le \rho \le 1}[E_i(\rho, W_{Y\,|\,UX}, P_{UX}) - \rho_i \hat{R}_i] \quad (80)$$

and

$$E_{sp}(R_1, R_2, W_{Y\,|\,UX}) = \max_{P_{UX}} \min_{i=1,2} \max_{\rho \ge 0}[E_i(\rho_i, W_{Y\,|\,UX}, P_{UX}) - \rho \hat{R}_i] \quad (81)$$

where $\hat{R}_1 = R_1 + R_2$ and $\hat{R}_2 = R_2$. Since it is in general hard to find the optimizing solution P_{UX} for E_r and E_{sp} above, we next confine our attention to multiple-access channels with some symmetric distributions.

Definition 3 [2]: We say that the multiple-access channel $W_{Y\,|\,UX}$ is U-symmetric if for every $u \in \mathcal{U}$ the transition matrix $W_{Y\,|\,UX}(\cdot\,|\,u, \cdot)$ is symmetric in the sense that the rows (respectively, columns) are permutations of each other. An X-symmetric multiple-access channel is defined similarly. We then say that $W_{Y\,|\,UX}$ is symmetric if it is both U-symmetric and X-symmetric.

It follows that the multiple-access channel with additive noise is symmetric (e.g., see the example below), where a multiple-access channel $W_{Y\,|\,UX}$ with (modulo B) additive noise $\{P_F : \mathcal{F}\}$ is described as

$$Y_i = U_i \oplus X_i \oplus F_i \pmod{B}$$

where $Y_i \in \mathcal{Y}$, $X_i \in \mathcal{X}$, $U_i \in \mathcal{U}$, and $F_i \in \mathcal{F}$ are the channel's output, two input and noise symbols at time i such that $\mathcal{Y} = \mathcal{U} = \mathcal{X} = \mathcal{F} = \{0, 1, 2, \ldots, B-1\}$, and F_i is independent of X_i and U_i, $i = 1, 2, \ldots, n$.

It is shown in [2] that if the multiple-access channel $W_{Y\,|\,UX}$ is U-symmetric, then the outer maximum of (80) and (81) is achieved by a joint distribution of the form $P_{UX}(u, x) = P_U(u)/|\mathcal{X}|$ for every x and u. It then follows that for the symmetric multiple-access channel, the maximum of (80) and (81) is achieved by a uniform joint distribution

$$P_{UX}^*(u, x) = \frac{1}{|\mathcal{U}||\mathcal{X}|}$$

which is independent of ρ. Substituting P_{UX}^* in (80) and (81) yields

$$E_r(R_1, R_2, W_{Y\,|\,UX}) = \min_{i=1,2} \max_{0 \le \rho \le 1}[\tilde{E}_i(\rho, W_{Y\,|\,UX}) - \rho \hat{R}_i] \quad (82)$$

$$E_r^{(i)}(R, W_{Y\,|\,UX}, P_{UX}) = \begin{cases} E_{sp}^{(i)}(R, W_{Y\,|\,UX}, P_{UX}), & \text{if } R \ge R_{cr}^{(i)}(W_{Y\,|\,UX}, P_{UX}) \\ E_{sp}^{(i)}\left(R_{cr}^{(i)}(W_{Y\,|\,UX}, P_{UX}), W_{Y\,|\,UX}, P_{UX}\right) + R_{cr}^{(i)}(W_{Y\,|\,UX}, P_{UX}) - R, & \text{if } 0 < R \le R_{cr}^{(i)}(W_{Y\,|\,UX}, P_{UX}). \end{cases}$$

and

$$E_{sp}(R_1,R_2,W_{Y\,|\,UX}) = \min_{i=1,2}\max_{\rho\ge 0}[\tilde{E}_i(\rho,W_{Y\,|\,UX}) - \rho\hat{R}_i] \tag{83}$$

where $\hat{R}_1 = R_1 + R_2$, $\hat{R}_2 = R_2$

$$\tilde{E}_1(\rho,W_{Y\,|\,UX}) = (1+\rho)\log_2(|\mathcal{U}||\mathcal{X}|) - \log_2\sum_{y\in\mathcal{Y}}\left(\sum_{(u,x)\in\mathcal{U}\times\mathcal{X}} W_{Y\,|\,UX}(y|u,x)^{\frac{1}{1+\rho}}\right)^{1+\rho}$$

and

$$\tilde{E}_2(\rho,W_{Y\,|\,UX}) = (1+\rho)\log_2|\mathcal{X}| + \log_2|\mathcal{U}| - \log_2\sum_{(u,y)\in\mathcal{U}\times\mathcal{Y}}\left(\sum_{x\in\mathcal{X}} W_{Y\,|\,UX}(y|u,x)^{\frac{1}{1+\rho}}\right)^{1+\rho}.$$

We also can prove the following identities using a standard optimization method (cf. [29]).

Lemma 5:

$$\min_{P_{SL}:H_P(S,L)=R} D(P_{SL}\|Q_{SL}) = \max_{\rho\ge 0}[\rho R - E_{s1}(\rho,Q_{SL})] \tag{84}$$

$$\min_{P_{SL}:H_P(L\,|\,S)=R} D(P_{SL}\|Q_{SL}) = \max_{\rho\ge 0}[\rho R - E_{s2}(\rho,Q_{SL})] \tag{85}$$

where

$$E_{s1}(\rho,Q_{SL}) = (1+\rho)\log_2\sum_{(s,l)\in\mathcal{S}\times\mathcal{L}} Q_{SL}(s,l)^{\frac{1}{1+\rho}}$$

and

$$E_{s2}(\rho,Q_{SL}) = (1+\rho)\sum_{s\in\mathcal{S}} Q_S(s)\log_2\sum_{l\in\mathcal{L}} Q_{L|S}(l|s)^{\frac{1}{1+\rho}}.$$

Note that $E_{s1}(\rho,Q_{SL})$ and $E_{s2}(\rho,Q_{SL})$ are both concave in ρ. Clearly, if the marginal distribution $Q_S(s)$ is uniform, then (84) and (85) are equal. Using (82) we now can write (64) as

$$\begin{aligned}
&\min_{P_{SL}}[\tau D(P_{SL}\,\|\,Q_{SL}) + E_r(\tau H_P(S),\tau H_P(L\,|\,S),W_{Y\,|\,UX})]\\
&\quad= \min\{\min_{P_{SL}}[\tau D(P_{SL}\,\|\,Q_{SL})\\
&\qquad\qquad + \max_{0\le\rho_1\le 1}[\tilde{E}_1(\rho_1,W_{Y\,|\,UX}) - \rho_1\tau H_P(S,L)]],\\
&\qquad \min_{P_{SL}}[\tau D(P_{SL}\,\|\,Q_{SL})\\
&\qquad\qquad + \max_{0\le\rho_2\le 1}[\tilde{E}_2(\rho_2,W_{Y\,|\,UX}) - \rho_2\tau H_P(L\,|\,S)]]\}\\
&\quad= \min\{\min_R[\min_{P_{SL}:\tau H_P(S,L)=R}\tau D(P_{SL}\,\|\,Q_{SL})\\
&\qquad\qquad + \max_{0\le\rho_1\le 1}[\tilde{E}_1(\rho_1,W_{Y\,|\,UX}) - \rho_1 R]],\\
&\qquad \min_R[\min_{P_{SL}:\tau H_P(L\,|\,S)=R}\tau D(P_{SL}\,\|\,Q_{SL})\\
&\qquad\qquad + \max_{0\le\rho_2\le 1}[\tilde{E}_2(\rho_2,W_{Y\,|\,UX}) - \rho_2 R]]\}
\end{aligned} \tag{86}$$

and, similarly, using (83) we can write (65) as

$$\begin{aligned}
&\inf_{P_{SL}}[\tau D(P_{SL}\,\|\,Q_{SL}) + E_{sp}(\tau H_P(S),\tau H_P(L\,|\,S),W_{Y\,|\,UX})]\\
&\quad= \min\{\inf_R[\min_{P_{SL}:\tau H_P(S,L)=R}\tau D(P_{SL}\,\|\,Q_{SL})\\
&\qquad\qquad + \max_{\rho_1\ge 0}[\tilde{E}_1(\rho_1,W_{Y\,|\,UX}) - \rho_1 R]],\\
&\qquad \inf_R[\min_{P_{SL}:\tau H_P(L\,|\,S)=R}\tau D(P_{SL}\,\|\,Q_{SL})\\
&\qquad\qquad + \max_{\rho_2\ge 0}[\tilde{E}_2(\rho_2,W_{Y\,|\,UX}) - \rho_2 R]]\}.
\end{aligned} \tag{87}$$

Consequently, using an optimization technique based on Fenchel duality [29] and (84) and (85), we obtain the following.

Theorem 5: Given Q_{SL}, a symmetric $W_{Y\,|\,UX}$, and the transmission rate τ, the lower bound of the JSCC error exponent given in (64) and the upper bound given in (68) can be equivalently expressed as

$$\begin{aligned}
&\min_{i=1,2}\max_{0\le\rho\le 1}[\tilde{E}_i(\rho,W_{Y\,|\,UX}) - \tau E_{si}(\rho,Q_{SL})]\\
&\quad\le E_J(Q_{SL},W_{Y\,|\,UX},\tau)\\
&\quad\le \min_{i=1,2}\max_{\rho\ge 0}[\tilde{E}_i(\rho,W_{Y\,|\,UX}) - \tau E_{si}(\rho,Q_{SL})].
\end{aligned} \tag{88}$$

Example 1: Now consider binary CS Q_{SL} with distribution

$$Q_{SL}(S=0,L=0) = \frac{2(1-q)}{3},\quad Q_{SL}(S=1,L=0) = \frac{q}{2}$$
$$Q_{SL}(S=0,L=1) = \frac{q}{2},\quad Q_{SL}(S=1,L=1) = \frac{1-q}{3}$$

where $0 < q < 1/2$. Then

$$\begin{aligned}
&E_{s1}(\rho,Q_{SL})\\
&\quad= (1+\rho)\log_2\left\{\left[\left(\frac{2}{3}\right)^{\frac{1}{1+\rho}} + \left(\frac{1}{3}\right)^{\frac{1}{1+\rho}}\right] \times (1-q)^{\frac{1}{1+\rho}} + 2\left(\frac{q}{2}\right)^{\frac{1}{1+\rho}}\right\}
\end{aligned}$$

$$\begin{aligned}
&E_{s2}(\rho,Q_{SL})\\
&\quad= (1+\rho)\left(\frac{2(1-q)}{3} + \frac{q}{2}\right)\\
&\qquad\times\log_2\left[\left(\frac{\frac{2(1-q)}{3}}{\frac{2(1-q)}{3}+\frac{q}{2}}\right)^{\frac{1}{1+\rho}} + \left(\frac{\frac{q}{2}}{\frac{2(1-q)}{3}+\frac{q}{2}}\right)^{\frac{1}{1+\rho}}\right]\\
&\qquad + (1+\rho)\left(\frac{1-q}{3} + \frac{q}{2}\right)\\
&\qquad\times\log_2\left[\left(\frac{\frac{1-q}{3}}{\frac{1-q}{3}+\frac{q}{2}}\right)^{\frac{1}{1+\rho}} + \left(\frac{\frac{q}{2}}{\frac{1-q}{3}+\frac{q}{2}}\right)^{\frac{1}{1+\rho}}\right].
\end{aligned}$$

Consider a binary multiple-access channel $W_{Y\,|\,UX}$ with binary additive noise $P_F(F=1) = \epsilon(0 < \epsilon < 1/2)$. That is, the transition probabilities are given by

$$P_{Y\,|\,UX}(Y=0\,|\,U=0,X=0) = 1-\epsilon$$
$$P_{Y\,|\,UX}(Y=1\,|\,U=0,X=0) = \epsilon$$
$$P_{Y\,|\,UX}(Y=0\,|\,U=0,X=1) = \epsilon$$

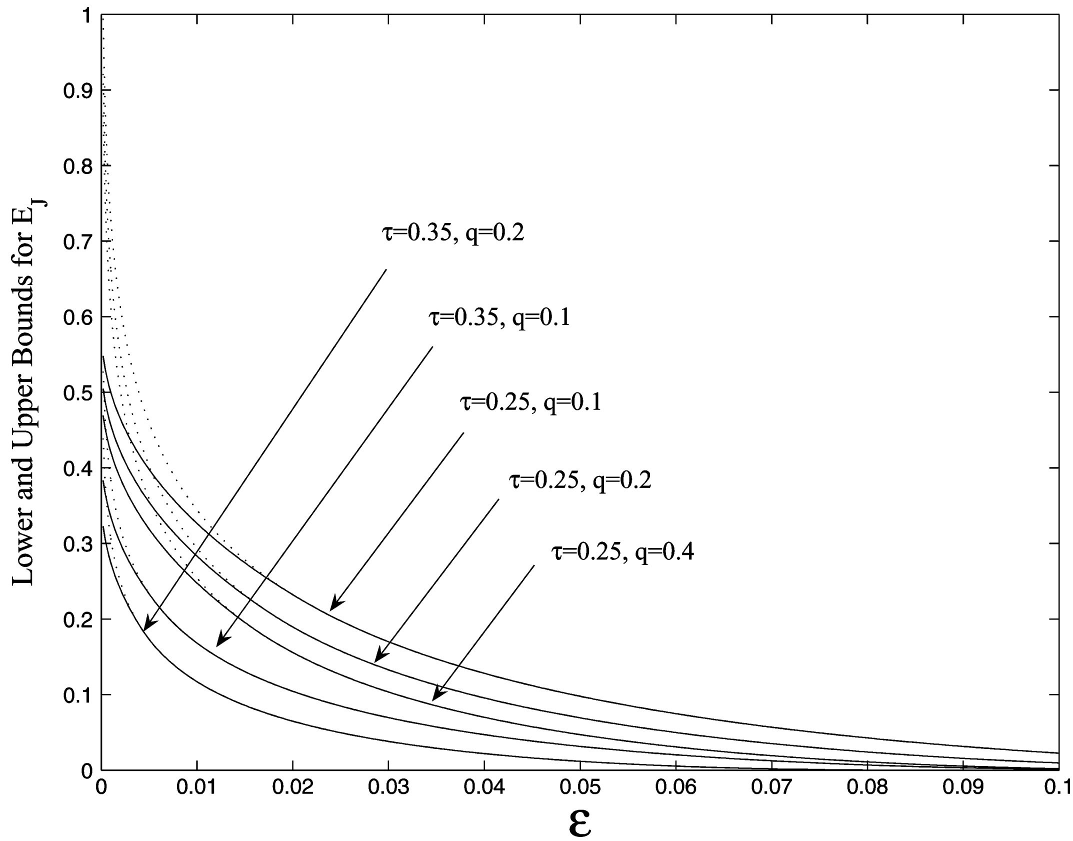

Fig. 6. The lower bound (solid line) and the upper bound (dash line) for the system JSCC error exponent for transmitting binary CS over the binary AMAC with binary additive noise.

$$P_{Y\,|\,UX}(Y=1\,|\,U=0,X=1)=1-\epsilon$$
$$P_{Y\,|\,UX}(Y=0\,|\,U=1,X=0)=\epsilon$$
$$P_{Y\,|\,UX}(Y=1\,|\,U=1,X=0)=1-\epsilon$$
$$P_{Y\,|\,UX}(Y=0\,|\,U=1,X=1)=1-\epsilon$$
$$P_{Y\,|\,UX}(Y=1\,|\,U=1,X=1)=\epsilon.$$

It follows that

$$\begin{aligned}\tilde{E}_1(\rho, W_{Y\,|\,UX}) &= \tilde{E}_2(\rho, W_{Y\,|\,UX})\\ &= \rho-(1+\rho)\log_2\left(\epsilon^{\frac{1}{1+\rho}}+(1-\epsilon)^{\frac{1}{1+\rho}}\right).\end{aligned}$$

In Fig. 6, we plot the lower and upper bounds for the JSCC error exponent E_J for different (q,ϵ) pairs with transmission rate $t=0.25$ and 0.35. As illustrated, the upper and lower bounds coincide (this can also be proved by checking that the two outer minima in (88) are achieved by the same i and that the inner maximum in the upper bound is achieved by $\rho\le 1$) for many (q,ϵ) pairs (e.g., when $\tau=0.25, q=0.1, \epsilon\ge 0.0205$, and when $\tau=0.35, q=0.1$, $\epsilon\ge 0.0056$), and hence exactly determine the exponent.

VII. Tandem Coding Error Exponent for the Asymmetric Two-User System

A. Tandem System With Common Randomization

In Section IV-D, we showed that the reliable transmissibility condition $(\tau H_Q(S), \tau H_Q(L\,|\,S))\in\mathcal{R}(W_{YZ\,|\,UX})$ in Theorem 3 can be achieved by a tandem coding system where separately designed source and channel coding operations are sequentially applied; see Figs. 4 and 5 with π_f and π_g being identity mappings. By "separately designed" we mean that the source code is designed without the knowledge of the channel statistics and the channel code is designed without the knowledge of the source statistics. Note, however, that as long as the source encoder is directly concatenated by a channel encoder (i.e., if π_f and π_g are identity mappings), the source statistics would be automatically brought into the channel coding stage. Thus, the performance of the channel code is affected by that of the source code (since the compressed messages (indices) fed into the channel encoders are not necessarily uniformly distributed). To statistically decouple the source and channel coding operations, we need to employ common randomization between the source and channel coding components (e.g., [18]). This results in a "complete" tandem coding system with fully separate source and channel coding operations, and for which we can establish an expression for its error exponent in terms of the source coding and channel coding exponents.

The tandem coding system is depicted in Figs. 4 and 5. As in Section IV-D, the encoder f_n is composed of two source encoders f_{sn} and g_{sn} and one channel encoder f_{cn}. The difference is that the indices $i=f_{sn}(\boldsymbol{l})$ and $j=g_{sn}(\boldsymbol{s})$ are separately mapped to channel indices through permutation functions $\pi_f:\{1,2,\ldots,M_l\}\to\{1,2,\ldots,M_l\}$ and $\pi_g:\{1,2,\ldots,M_s\}\to\{1,2,\ldots,M_s\}$, which are usually called index assignments (π_f and π_g are assumed to be known at both the transmitter and the receiver). Furthermore, the choice of π_f (π_g, respectively) is assumed random (independent from the source and the channel) and equally likely from all $M_l!$ ($M_s!$, respectively) different possible index assignments, so that the indices fed into the channel encoder have a uniform distribution and are mutually independent

$$\begin{aligned}
&\Pr(\pi_f(f_{sn}(L^{\tau n})) = a)\\
&\quad= \sum_{i=1}^{M_l} \Pr(f_{sn}(L^{\tau n}) = i)\Pr(\pi_f(i) = a | f_{sn}(L^{\tau n}) = i)\\
&\quad= \sum_{i=1}^{M_l} \Pr(f_{sn}(L^{\tau n}) = i)\frac{(M_l-1)!}{M_l!} = \frac{1}{M_l}\\
&\Pr(\pi_g(g_{sn}(S^{\tau n})) = b)\\
&\quad= \frac{1}{M_s}\\
&\Pr(\pi_g(f_{sn}(L^{\tau n})) = a, \pi_g(g_{sn}(S^{\tau n})) = b)\\
&\quad= \Pr(\pi_f(f_{sn}(L^{\tau n})) = a)\Pr(\pi_g(g_{sn}(S^{\tau n})) = b)
\end{aligned}$$

for any $(a,b) \in \{1,2,\ldots,M_l\} \times \{1,2,\ldots,M_s\}$. Hence, common randomization achieves statistical separation between the source and channel coding operations.

Similarly, the encoder g_n is independently composed of a source encoder g_{sn}, an index mapping $\pi_g : \{1,2,\ldots,M_s\} \to \{1,2,\ldots,M_s\}$, and a channel encoder $g_{cn} : \{1,2,\ldots,M_s\} \to \mathcal{U}^n$.

At the receiver side, the decoder φ_n is composed of a channel decoder φ_{cn}, a pair of index mappings (π_f^{-1}, π_g^{-1}) which maps every channel index pair $(\pi_f(\hat{i}), \pi_g(\hat{j}))$ back to a source index pair $(\hat{i}, \hat{j})$, and a source decoder φ_{sn} which outputs the approximation of the source messages $\boldsymbol{s}'$ and $\boldsymbol{l}'$. Similarly, the decoder ψ_n is composed of a channel decoder $\psi_{cn} : \mathcal{Z}^n \to \{1,2,\ldots,M_s\}$, an index mapping π_g^{-1}, and a source decoder $\psi_{sn} : \{1,2,\ldots,M_s\} \to \mathcal{S}^{\tau n}$.

B. A Formula for the Tandem Coding Error Exponent

We now can study the error performance and exponent of tandem source–channel coding (with common randomization) for the asymmetric two-user system. Since the tandem code consists of a source code $(f_{sn}, g_{sn}, \varphi_{sn}, \psi_{sn})$ and a channel code $(f_{cn}, g_{cn}, \varphi_{cn}, \psi_{cn})$, we first define the corresponding source coding error exponent (note that the corresponding channel coding error exponent for the asymmetric two-user channel was defined in Section IV-E).

Definition 4: The two-user source coding error exponent $E(R_1, R_2, Q_{SL})$, for any $R_1 > 0$ and $R_2 > 0$, is defined by the supremum of the set of all numbers E_s for which there exists a sequence of source codes $(f_{sn}, g_{sn}, \varphi_{sn}, \psi_{sn})$ with block length n, common rate no larger than R_1, and private rate no larger than R_2, such that

$$E_s \le \liminf_{n\to\infty} -\frac{1}{n}\log_2 P_{es}^{(n)}(R_1, R_2, Q_{SL}) \tag{89}$$

where $P_{es}^{(n)}(R_1, R_2, Q_{SL})$ is the source coding probability of error defined in (51).

Denote the probabilities of Y- and Z-error for the source coding by

$$\begin{aligned}
&P_{Yes}^{(n)}(\hat{R}_s, \hat{R}_l, Q_{SL})\\
&\quad\triangleq \Pr(\{\varphi_{sn}(i,j) \neq (S^{\tau n}, L^{\tau n})\})\\
&\quad= \sum_{(\boldsymbol{s},\boldsymbol{l}):\psi_{sn}(i,j)\neq(\boldsymbol{s},\boldsymbol{l})} Q_{SL}^{(n)}(\boldsymbol{s},\boldsymbol{l})
\end{aligned}\tag{90}$$

and

$$\begin{aligned}
&P_{Zes}^{(n)}(\hat{R}_s, \hat{R}_l, Q_{SL})\\
&\quad= P_{Ze}^{(n)}(\hat{R}_s, Q_S) \triangleq \Pr(\{\psi_{sn}(j) \neq S^{\tau n}\})\\
&\quad= \sum_{\boldsymbol{s}:\psi_{sn}(i)\neq\boldsymbol{s}} Q_S^{(n)}(\boldsymbol{s})
\end{aligned}\tag{91}$$

where $i \triangleq f_{sn}(\boldsymbol{l})$ and $j \triangleq g_{sn}(\boldsymbol{s})$. Clearly, for any sequence of source codes $(f_{sn}, g_{sn}, \varphi_{sn}, \psi_{sn})$, the error probability $P_{es}^{(n)}(R_1, R_2, Q_{SL})$ must be larger than $P_{Yes}^{(n)}(R_1, R_2, Q_{SL})$ and $P_{Zes}^{(n)}(R_1, R_2, Q_{SL}))$ but less than the sum of the two; so we have

$$\begin{aligned}
&\liminf_{n\to\infty} -\frac{1}{n}\log_2 P_{es}^{(n)}(R_1, R_2, Q_{SL})\\
&\quad= \liminf_{n\to\infty} -\frac{1}{n}\log_2 \max\Big(P_{Yes}^{(n)}(R_1, R_2, Q_{SL}),\\
&\qquad\qquad P_{Zes}^{(n)}(R_1, R_2, Q_{SL})\Big).
\end{aligned}\tag{92}$$

In what follows, we need to make three assumptions in order to analyze the probability of error of the overall tandem system. The first two assumptions (referred to as (A1) and (A2)) are regarding the source codes. Let the source codebook for (g_{sn}, ψ_{sn}) (Receiver Z) be $\mathcal{C}^{(g)} = \{\boldsymbol{c}_1^{(g)}, \ldots, \boldsymbol{c}_{M_l}^{(g)}\} \subseteq \mathcal{S}^{\tau n}$, and let the source codebook for $(f_{sn}, g_{sn}, \varphi_{sn})$ (Receiver Y) be $\mathcal{C}^{(f)} \times \mathcal{C}^{(g)}$ where $\mathcal{C}^{(f)} = \{\boldsymbol{c}_1^{(f)}, \ldots, \boldsymbol{c}_{M_s}^{(f)}\} \subseteq \mathcal{L}^{\tau n}$.

- We assume that (A1) the source encoder f_{sn} satisfies the condition (for every n): $Q_L^{\tau n}(f_{sn}^{-1}(i)) > 0$ and $\boldsymbol{c}_i^{(f)} \in f_{sn}^{-1}(i)$ for every $i = 1,2,\ldots,M_l$, where $f_{sn}^{-1}(i) \triangleq \{\boldsymbol{l} \in \mathcal{L}^{\tau n} : f_{sn}(\boldsymbol{l}) = i\}$. Clearly, the assumption has practical meaning. If $Q_L^{\tau n}(f_{sn}^{-1}(i)) = 0$ for some i, then the codeword $\boldsymbol{c}_i^{(f)}$ is redundant, and we can remove it from the codebook $\mathcal{C}^{(f)}$. If $\boldsymbol{c}_i^{(f)} \notin f_{sn}^{-1}(i)$, we can map the index i to some source message $\hat{\boldsymbol{l}}$ such that $Q_L^{\tau n}(\hat{\boldsymbol{l}}) > 0$ and $f_{sn}(\hat{\boldsymbol{l}}) = i$, so that the source coding probability of error $P_{Yes}^{(n)}(\hat{R}_s, \hat{R}_l, Q_{SL})$ is strictly reduced by setting $\hat{\boldsymbol{l}}$ as the codeword $\boldsymbol{c}_i^{(f)}$ (note that $P_{Zes}^{(n)}(\hat{R}_s, \hat{R}_l, Q_{SL})$ is independent of f_{sn}).
- Similarly, we assume that (A2) the source code g_{sn} satisfies the condition (for every n): $Q_S^{\tau n}(g_{sn}^{-1}(j)) > 0$ and $\boldsymbol{c}_j^{(g)} \in g_{sn}^{-1}(j)$ for every $j = 1,2,\ldots,M_s$, where $g_{sn}^{-1}(j) \triangleq \{\boldsymbol{s} \in \mathcal{S}^{\tau n} : g_{sn}(\boldsymbol{s}) = j\}$. If $Q_S^{\tau n}(g_{sn}^{-1}(j)) = 0$ for some j, then the codeword $\boldsymbol{c}_j^{(g)}$ is redundant, and we can remove it from the codebook $\mathcal{C}^{(g)}$. If $\boldsymbol{c}_j^{(g)} \notin g_{sn}^{-1}(j)$, we can map the index j to some source message $\hat{\boldsymbol{s}}$ such that $Q_S^{\tau n}(\hat{\boldsymbol{s}}) > 0$ and $g_{sn}(\hat{\boldsymbol{s}}) = j$, so that the source coding error probabilities $P_{Yes}^{(n)}(\hat{R}_s, \hat{R}_l, Q_{SL})$ and $P_{Zes}^{(n)}(\hat{R}_s, \hat{R}_l, Q_{SL})$ are strictly reduced by setting $\hat{\boldsymbol{s}}$ as the codeword $\boldsymbol{c}_j^{(g)}$.
- We assume that the limits $\lim_{n\to\infty} \frac{1}{n}\log_2 M_l$ and $\lim_{n\to\infty} \frac{1}{n}\log_2 M_s$ exist, i.e., $\liminf_{n\to\infty} \frac{1}{n}\log_2 M_l = \limsup_{n\to\infty} \frac{1}{n}\log_2 M_l$ and $\liminf_{n\to\infty} \frac{1}{n}\log_2 M_s = \limsup_{n\to\infty} \frac{1}{n}\log_2 M_s$. This assumption is used later to upper-bound the tandem coding error exponent in Theorem 6.

We remark that the source code satisfying (A1) and (A2) does not lose optimality in the sense of achieving the source error exponent.

Denote $\pi^{-1}(i,j) \triangleq (\pi_f^{-1}(i), \pi_g^{-1}(j))$. By introducing (A1) and (A2), the error probability of the tandem code $(f_n^*, \varphi_n^*) \triangleq (f_{sn}, g_{sn}, \varphi_{sn}, \psi_{sn}, f_{cn}, g_{cn}, \varphi_{cn}, \psi_{cn})$ is given by

$$\begin{aligned}
&P_{e^*}^{(n)}(Q_{SL}, W_{YZ|UX}, \tau) \\
&\triangleq \Pr\Big(\{\varphi_{sn}\left[\pi^{-1}(\varphi_{cn}(Y^n))\right] \neq (S^{\tau n}, L^{\tau n})\} \\
&\quad \bigcup \{\psi_{sn}\left[\pi_g^{-1}(\psi_{cn}(Z^n))\right] \neq S^{\tau n}\}\Big) \\
&= \sum_{a=1}^{M_l}\sum_{b=1}^{M_s} \underbrace{\Pr(\pi_f[f_{sn}(L^{\tau n})] = a)}_{=1/M_l}\underbrace{\Pr(\pi_g[g_{sn}(S^{\tau n})] = b)}_{=1/M_s} \\
&\quad \times \Big[\Pr\Big(\{\varphi_{cn}(Y^n) \neq (a,b)\}\bigcup\{\psi_{cn}(Z^n) \neq b\}\Big| \\
&\qquad \pi_f[f_{sn}(L^{\tau n})] = a, \pi_g[g_{sn}(S^{\tau n})] = b\Big) \\
&\qquad + \Pr\big(\{\varphi_{cn}(Y^n) = (a,b) \quad \text{and} \quad \psi_{cn}(Z^n) = b\} \\
&\qquad \bigcap\{\varphi_{sn}[\pi^{-1}(a,b)] \neq (S^{\tau n}, L^{\tau n}) \quad \text{or} \\
&\qquad\qquad \psi_{sn}[\pi_g^{-1}(b)] \neq S^{\tau n}\}\Big| \\
&\qquad \pi_f[f_{sn}(L^{\tau n})] = a, \pi_g[g_{sn}(S^{\tau n})] = b\big)\Big] \qquad (93)\\
&= \sum_{a=1}^{M_l}\sum_{b=1}^{M_s} \frac{1}{M_l M_s}\Pr\Big(\{\varphi_{cn}(Y^n) \neq (a,b)\} \\
&\quad \bigcup \{\psi_{cn}(Z^n) \neq b\} | (a,b) \text{ is sent}\Big) \\
&\quad + \Pr\Big(\{\varphi_{sn}[S^{\tau n}, L^{\tau n}] \neq (S^{\tau n}, L^{\tau n})\} \\
&\quad \bigcup\{\psi_{sn}[S^{\tau n}] \neq S^{\tau n}\}\Big) \\
&\quad \times \sum_{a=1}^{M_l}\sum_{b=1}^{M_s} \frac{1}{M_l M_s}\Pr\Big(\{\varphi_{cn}(Y^n) = (a,b)\} \\
&\quad \bigcap \{\psi_{cn}(Z^n) = b\} | (a,b) \text{ is sent}\Big) \qquad (94)\\
&= P_{ec}^{(n)}(\tau\hat{R}_s, \tau\hat{R}_l, W_{YZ|UX}) \\
&\quad + [1 - P_{ec}^{(n)}(\tau\hat{R}_s, \tau\hat{R}_l, W_{YZ|UX})] \\
&\quad \times P_{es}^{(\tau n)}(\hat{R}_s, \hat{R}_l, Q_{SL}) \qquad (95)
\end{aligned}$$

where (93) follows from assumptions (A1) and (A2), which imply that a channel decoding error must cause an overall system decoding error, (94) holds due to the statistical decoupling of source and channel coding.

Definition 5: The tandem coding error exponent $E_T(Q_{SL}, W_{YZ|UX}, \tau)$ for source Q_{SL} and channel $W_{YZ|UX}$ is defined as the supremum of the set of all numbers $\hat{E}$ for which there exists a sequence of tandem codes (f_n^*, φ_n^*) satisfying (A1) and (A2) with transmission rate τ such that

$$\hat{E} \leq \liminf_{n\to\infty} -\frac{1}{n}\log_2 P_{e^*}^{(n)}(Q_{SL}, W_{YZ|UX}, \tau).$$

When there is no possibility of confusion, $E_T(Q_{SL}, W_{YZ|UX}, \tau)$ will often be written as E_T. The following lemma illustrates the relation between E_T and E_J.

Lemma 6:

$$E_J(Q_{SL}, W_{YZ|UX}, \tau) \geq E_T(Q_{SL}, W_{YZ|UX}, \tau).$$

Proof: By definition, for any sequence of rate τ tandem codes $\{(f_n^*, \varphi_n^*)\}_{n=1}^\infty$ composed of a sequence of source codes $\{(f_{sn}, g_{sn}, \varphi_{sn}, \psi_{sn})\}_{n=1}^\infty$ and a sequence of channel codes $\{(f_{cn}, g_{cn}, \varphi_{cn}, \psi_{cn})\}_{n=1}^\infty$, we have

$$\begin{aligned}
&P_{e^*}^{(n)}(Q_{SL}, W_{YZ|UX}, \tau) \\
&= \mathbb{E}_{\pi_f,\pi_g}\Pr\Big(\{\varphi_{sn}[\pi^{-1}(\varphi_{cn}(Y^n))] \neq (S^{\tau n}, L^{\tau n})\} \\
&\quad \bigcup\{\psi_{sn}\left[\pi_g^{-1}(\psi_{cn}(Z^n))\right] \neq S^{\tau n}\}\Big| \pi_f \text{ and } \pi_g \text{ are fixed}\Big) \\
&\geq \min_{\pi_f,\pi_g}\Pr\Big(\{\varphi_{sn}\left[\pi^{-1}(\varphi_{cn}(Y^n))\right] \neq (S^{\tau n}, L^{\tau n})\} \\
&\quad \bigcup\{\psi_{sn}\left[\pi_g^{-1}(\psi_{cn}(Z^n))\right] \neq S^{\tau n}\}\Big| \pi_f \text{ and } \pi_g \text{ are fixed}\Big).
\end{aligned}$$

Let the above minimum be achieved by $\pi_f^* = \pi_f^*(n)$ and $\pi_g^* = \pi_f^*(n)$. Obviously, there exists a sequence of JSC codes $\{(f_n, g_n, \varphi_n, \psi_n)\}_{n=1}^\infty$ where f_n is composed of f_{sn}, g_{sn}, π_f^*, π_g^*, f_{cn}, g_n is composed of g_{sn}, π_g^*, and g_{cn}, φ_n is composed of φ_{cn}, $(\pi_f^{*-1}, \pi_g^{*-1})$, and φ_{sn}, and finally, ψ_n is composed of ψ_{cn}, π_g^{*-1}, and ψ_{sn} (cf. Figs. 4 and 5), such that

$$\begin{aligned}
&P_{e^*}^{(n)}(Q_{SL}, W_{YZ|UX}, \tau) \\
&\qquad\qquad \geq P_e^{(n)}(Q_{SL}, W_{YZ|UX}, \tau) \quad \text{for any } n \geq 1
\end{aligned}$$

where $P_e^{(n)}(Q_{SL}, W_{YZ|UX}, \tau)$ is the probability of error induced by the JSC codes $\{(f_{cn}, g_{cn}, \varphi_{cn}, \psi_{cn})\}$. Since this holds for any sequence of tandem codes (satisfying (A1) and (A2)), it then follows from the definition of joint and tandem exponents that $E_J \geq E_T$. □

We next derive a formula for E_T in terms of the corresponding source and channel error exponents.

Theorem 6:

$$\begin{aligned}
&E_T(Q_{SL}, W_{YZ|UX}, \tau) \\
&\quad = \sup_{R_1>0, R_2>0} \min\left\{\tau e\left(\frac{R_1}{\tau}, \frac{R_2}{\tau}, Q_{SL}\right),\right. \\
&\qquad\qquad\qquad \left. E(R_1, R_2, W_{YZ|UX})\right\}
\end{aligned}$$

where $e(R_1, R_2, Q_{SL})$ is the two-user source coding error exponent defined in (4) and $E(R_1, R_2, W_{YZ|UX})$ is the asymmetric two-user channel coding error exponent defined in (2).

Remark 1: As can be seen from the proof below, the common randomization setup together with the assumptions regarding the source and channel codes are essentially needed to prove the converse part of the tandem coding error exponent; the forward part (the proof of the lower bound on the exponent) is still valid for tandem systems without these assumptions.

Proof: Forward Part: we show that there exists a sequence of tandem codes (f_n^*, φ_n^*) satisfying (A1) and (A2) such that

$$\liminf_{n\to\infty} -\frac{1}{n}\log_2 P_{e^*}^{(n)}(Q_{SL}, W_{YZ|UX}, \tau)$$

$$> \sup_{R_1>0,R_2>0} \min\left\{\tau e\left(\frac{R_1}{\tau},\frac{R_2}{\tau},Q_{SL}\right),\right. \left. E(R_1,R_2,W_{YZ|UX})\right\} - \delta$$

for any $\delta > 0$. It follows from (95) that

$$P_{e^*}^{(n)}(Q_{SL},W_{YZ|UX},\tau) \le 2\max\left\{P_{es}^{(\tau n)}(\hat{R}_s,\hat{R}_l,Q_{SL}),\right. \left. P_{ec}^{(n)}(\tau\hat{R}_s,\tau\hat{R}_l,W_{YZ|UX})\right\}$$

or equivalently

$$\liminf_{n\to\infty} -\frac{1}{n}\log_2 P_{e^*}^{(n)}(Q_{SL},W_{YZ|UX},\tau) \ge \min\left\{\liminf_{n\to\infty} -\frac{1}{n}P_{es}^{(\tau n)}(\hat{R}_s,\hat{R}_l,Q_{SL})\right. \left. \liminf_{n\to\infty} -\frac{1}{n}P_{ec}^{(n)}(\tau\hat{R}_s,\tau\hat{R}_l,W_{YZ|UX})\right\} \quad (96)$$

Fix $\delta > 0$ and let $R_1 = \tau\lim_{n\to\infty}\hat{R}_s$ and $R_2 = \tau\lim_{n\to\infty}\hat{R}_l$. According to the definition of the two-user source coding error exponent, there exists a sequence of source codes $(\tilde{f}_{sn},\tilde{g}_{sn},\tilde{\varphi}_{sn},\tilde{\psi}_{sn})$ satisfying (A1) and (A2) (since (A1) and (A2) do not lose optimality) with common source rate $\hat{R}_s$ and private source rate $\hat{R}_l$ such that

$$\liminf_{n\to\infty} -\frac{1}{n}P_{es}^{(\tau n)}(\hat{R}_s,\hat{R}_l,Q_{SL}) \ge e\left(\frac{R_1}{\tau},\frac{R_2}{\tau},Q_{SL}\right) - \delta.$$

On the other hand, according to the definition of the asymmetric two-user channel coding error exponent, there exists a sequence of channel codes $(f_{cn},g_{cn},\varphi_{cn},\psi_{cn})$ with common rate $\tau\hat{R}_s$ and private rate $\tau\hat{R}_l$ such that

$$\liminf_{n\to\infty} -\frac{1}{n}P_{ec}^{(n)}(\tau\hat{R}_s,\tau\hat{R}_l,W_{YZ|UX}) \ge E(\tau R_1,\tau R_2,W_{YZ|UX}) - \delta.$$

Finally, since the sequences of rates $\hat{R}_s$ and $\hat{R}_l$ can be arbitrarily chosen, and so are R_1 and R_2, we can take the supremum of R_1 and R_2, completing the proof of the forward part.

Converse Part: We show that for any sequence of tandem codes (f_n^*,φ_n^*) with rate τ composed by source codes $\{(f_{sn},g_{sn},\varphi_{sn},\psi_{sn})\}_{n=1}^{\infty}$ satisfying assumptions (A1) and (A2) and channel codes $\{(f_{cn},g_{cn},\varphi_{cn},\psi_{cn})\}_{n=1}^{\infty}$

$$\liminf_{n\to\infty} -\frac{1}{n}\log_2 P_{e^*}^{(n)}(Q_{SL},W_{YZ|UX},\tau) \le \sup_{R_1>0,R_2>0}\min\left\{\tau e\left(\frac{R_1}{\tau},\frac{R_2}{\tau},Q_{SL}\right),\right. \left. E(R_1,R_2,W_{YZ|UX})\right\}. \quad (97)$$

Let the private index set for the tandem system be $\{1,2,\ldots,M_l\}$ (cf. Figs. 4 and 5). Thus, the private source and channel code rates are given by $\hat{R}_l = \frac{1}{\tau n}\log_2 M_l$ and $R_l = \tau\hat{R}_l$, respectively. Let the common index set be $\{1,2,\ldots,M_s\}$. Thus, the common source code rate and channel code rate are given by $\hat{R}_s = \frac{1}{\tau n}\log_2 M_s$ and $R_s = \tau\hat{R}_s$, respectively.

We first assume that

$$\limsup_{n\to\infty} -\frac{1}{n}\log_2[1-P_{ec}^{(n)}(\tau\hat{R}_s,\tau\hat{R}_l,W_{YZ|UX})] \ge \delta$$

for some positive δ independent of n, which implies that there exists a sequence $n_0 \le n_1 \le n_2 \le \cdots \le \infty$ such that

$$\lim_{i\to\infty} P_{ec}^{(n_i)}(\tau\hat{R}_s,\tau\hat{R}_l,W_{YZ|UX}) \ge 1 - \lim_{i\to\infty} 2^{-n_i\delta} = 1.$$

In this trivial case

$$\begin{aligned}\liminf_{n\to\infty} -\frac{1}{n}\log_2 P_{e^*}^{(n)}(Q_{SL},W_{YZ|UX},\tau) &\le \liminf_{n\to\infty} -\frac{1}{n}\log_2 P_{ec}^{(n)}(\tau\hat{R}_s,\tau\hat{R}_l,W_{YZ|UX}) \\ &\le \lim_{i\to\infty} -\frac{1}{n_i}\log_2 P_{ec}^{(n_i)}(\tau\hat{R}_s,\tau\hat{R}_l,W_{YZ|UX}) \\ &= 0\end{aligned}$$

and (97) holds. Next we assume that

$$\limsup_{n\to\infty} -\frac{1}{n}\log_2[1-P_{ec}^{(n)}(\tau\hat{R}_s,\tau\hat{R}_l,W_{YZ|UX})] = 0.$$

It then follows from (95) that

$$\begin{aligned}\liminf_{n\to\infty} -\frac{1}{n}\log_2 P_{e^*}^{(n)}(Q_{SL},W_{YZ|UX},\tau) &\le \liminf_{n\to\infty} -\frac{1}{n}\log_2\left[\left(1-P_{ec}^{(n)}(\tau\hat{R}_s,\tau\hat{R}_l,W_{YZ|UX}\right)\right. \\ &\qquad \left.\times\, P_{es}^{(\tau n)}(\hat{R}_s,\hat{R}_l,Q_{SL})\right] \\ &\le \liminf_{n\to\infty} -\frac{1}{n}\log_2 P_{es}^{(\tau n)}(\hat{R}_s,\hat{R}_l,Q_{SL}) \end{aligned} \quad (98)$$

and

$$\begin{aligned}\liminf_{n\to\infty} -\frac{1}{n}\log_2 P_{e^*}^{(n)}(Q_{SL},W_{YZ|UX},\tau) &\le \liminf_{n\to\infty} -\frac{1}{n}\log_2 P_{ec}^{(n)}(\tau\hat{R}_s,\tau\hat{R}_l,W_{YZ|UX}). \end{aligned} \quad (99)$$

Let

$$R_1 = \lim_{n\to\infty}\tau\hat{R}_s = \lim_{n\to\infty}\frac{\log_2 M_s}{n} \quad (100)$$

and

$$R_2 = \lim_{n\to\infty}\tau\hat{R}_l = \lim_{n\to\infty}\frac{\log_2 M_l}{n}. \quad (101)$$

By definition, the source error exponent $e(\frac{R_1}{\tau},\frac{R_2}{\tau},Q_{SL})$ is the largest (supremum) number E_s such that there exists a sequence of source codes $(\tilde{f}_{sn},\tilde{g}_{sn},\tilde{\varphi}_{sn},\tilde{\psi}_{sn})$ with message sets $\{1,2,\ldots,\tilde{M}_s\}$ and $\{1,2,\ldots,\tilde{M}_l\}$ satisfying

$$\limsup_{n\to\infty}\frac{\log_2\tilde{M}_s}{\tau n} \le \frac{R_1}{\tau}$$

$$\limsup_{n\to\infty}\frac{\log_2\tilde{M}_l}{\tau n} \le \frac{R_2}{\tau}$$

and

$$\liminf_{n\to\infty} -\frac{1}{\tau n}\log_2\Pr\left(\{\tilde{\varphi}_{sn}(\tilde{f}_{sn}(L^{\tau n}),\tilde{g}_{sn}(S^{\tau n})) \ne (S^{\tau n},L^{\tau n})\}\bigcup\{\tilde{\psi}_{sn}(\tilde{g}_{sn}(S^{\tau n}))\ne S^{\tau n}\}\right) \ge E_s.$$

This means that

$$\liminf_{n\to\infty} -\frac{1}{\tau n}\log_2 \Pr\Big(\{\tilde{\varphi}_{sn}(\tilde{f}_{sn}(L^{\tau n}),\tilde{g}_{sn}(S^{\tau n})) \neq (S^{\tau n},L^{\tau n})\}\bigcup\{\tilde{\psi}_{sn}(\tilde{g}_{sn}(S^{\tau n}))\neq S^{\tau n}\}\Big) \le e\left(\frac{R_1}{\tau},\frac{R_2}{\tau},Q_{SL}\right)$$

holds for all source codes $(\tilde{f}_{sn},\tilde{g}_{sn},\tilde{\varphi}_{sn},\tilde{\psi}_{sn})$ with

$$\limsup_{n\to\infty}\frac{\log_2\tilde{M}_s}{\tau n}\le\frac{R_1}{\tau} \quad\text{and}\quad \limsup_{n\to\infty}\frac{\log_2\tilde{M}_l}{\tau n}\le\frac{R_2}{\tau}$$

and hence holds for the sequence of block codes $(f_{sn},g_{sn},\varphi_{sn},\psi_{sn})$ with rates $(\hat{R}_s,\hat{R}_l)$ satisfying (A1) and (A2).

Similarly, by the definition of the asymmetric two-user channel error exponent and on account of (100) and (101) we have

$$\liminf_{n\to\infty} -\frac{1}{n}\log_2 \Pr\Big(\{\hat{\varphi}_{cn}(Y^n)\neq(J,I)\}\bigcup\{\hat{\psi}_{cn}(Z^n)\neq J\}\Big) \le E(R_1,R_2,W_{YZ\,|\,UX})$$

holds for all channel codes $(\hat{f}_{cn},\hat{g}_{cn},\hat{\varphi}_{cn},\hat{\psi}_{cn})$ with common and private message sets $\{1,2,\dots,\hat{M}_s\}$ and $\{1,2,\dots,\hat{M}_l\}$ such that

$$\liminf_{n\to\infty}\frac{\log_2\hat{M}_s}{n}\ge R_1 \quad\text{and}\quad \liminf_{n\to\infty}\frac{\log_2\hat{M}_l}{n}\ge R_2$$

and of course holds for the sequence of channel codes $(f_{cn},g_{cn},\varphi_{cn},\psi_{cn})$ with common rate $R_s=\tau\hat{R}_s$ and private rate $R_l=\tau\hat{R}_l$.

Putting things together, (98) and (99) yield

$$\liminf_{n\to\infty} -\frac{1}{n}\log_2 P_{e^*}^{(n)}(Q_{SL},W_{YZ\,|\,UX},\tau) \le \min\left\{\tau e\left(\frac{R_1}{\tau},\frac{R_2}{\tau},Q_{SL}\right),E(R_1,R_2,W_{YZ\,|\,UX})\right\}$$

holds for all the source codes satisfying (A1) and (A2) and all the channel codes with

$$\lim_{n\to\infty}\frac{\log_2 M_s}{n}=R_1 \quad\text{and}\quad \lim_{n\to\infty}\frac{\log_2 M_l}{n}=R_2.$$

Since the above is satisfied for any sequences of $M_s>0$ and $M_l>0$, and hence for all $R_1>0$ and $R_2>0$, we take the supremum over $R_1>0, R_2>0$ and obtain (97). □

C. Comparison of Joint and Tandem Coding Error Exponents

Although tandem source–channel coding can achieve reliable transmissibility, it might not achieve the system JSCC error exponent. In the following, we consider the tandem system consisting of CS Q_{SL} and AMAC $W_{Y\,|\,UX}$. For the CS-AMAC tandem system, we have only one receiver, Receiver Y, and the source decoder (cf. Fig. 5) φ_{sn} becomes a Slepian–Wolf decoder [6]. Furthermore

$$P_{es}^{(n)}(R_1,R_2,Q_{SL}) = P_{Yes}^{(n)}(R_1,R_2,Q_{SL}) = \sum_{(\boldsymbol{s},\boldsymbol{l}):\psi_{sn}(i,j)\neq(\boldsymbol{s},\boldsymbol{l})} Q_{SL}^{(n)}(\boldsymbol{s},\boldsymbol{l})$$

and

$$\begin{aligned}P_{ec}^{(n)}(R_1,R_2,W_{Y\,|\,UX}) &= P_{Yec}^{(n)}(R_1,R_2,W_{Y\,|\,UX})\\ &=\frac{1}{2^{R_1+R_2}}\sum_{\mathcal{M}_s\times\mathcal{M}_l}\ \sum_{\boldsymbol{y}:\varphi_{cn}(\boldsymbol{y})\neq(j,i)} W_{Y|X}^{(n)}(\boldsymbol{y}|\boldsymbol{u},\boldsymbol{x}).\end{aligned}$$

In this case, we can upper-bound the source error exponent by

$$\begin{aligned}e\left(\frac{R_1}{\tau},\frac{R_2}{\tau},Q_{SL}\right) &\le \min_{P_{SL}:\tau H_P(S,L)=R_1+R_2} D(P_{SL}\|Q_{SL})\\ &=\max_{\rho\ge 0}\left[\rho\frac{R_1+R_2}{\tau}-E_{s1}(\rho,Q_{SL})\right]\end{aligned} \tag{102}$$

which is obtained by viewing the two source encoders f_{sn} and g_{sn} as a joint encoder [11], where $E_{s1}(\rho,Q_{SL})$ is given by Lemma 5. Therefore, we can upper-bound the tandem coding error exponent for the CS-AMAC system by

$$\begin{aligned}E_T(Q_{SL},W_{YZ\,|\,UX},\tau) \le \sup_{R_1>0,R_2>0}\min\{&\max_{\rho\ge0}[\rho(R_1+R_2)-\tau E_{s1}(\rho,Q_{SL})],\\ &E_{sp}(R_1,R_2,W_{Y\,|\,UX})\}\end{aligned} \tag{103}$$

where $E_{sp}(R_1,R_2,W_{Y\,|\,UX})$ is an upper bound for the channel error exponent and is given by (81).

Example 2: Now consider the same binary CS Q_{SL} given in Example 1 such that

$$E_{s1}(\rho,Q_{SL})=(1+\rho)\log_2\left\{\left[\left(\frac{2}{3}\right)^{\frac{1}{1+\rho}}+\left(\frac{1}{3}\right)^{\frac{1}{1+\rho}}\right]\times(1-q)^{\frac{1}{1+\rho}}+2\left(\frac{q}{2}\right)^{\frac{1}{1+\rho}}\right\}$$

and consider the same binary multiple-access channel $W_{Y\,|\,UX}$ as in Example 1 with binary additive noise $P_F(F=1)=\epsilon(0<\epsilon<1/2)$ such that

$$\begin{aligned}E_{sp}(R_1,R_2,W_{Y\,|\,UX}) &= \min_{i=1,2}\max_{\rho\ge0}[\tilde{E}_i(\rho,W_{Y\,|\,UX})-\rho\hat{R}_i]\\ &=\max_{\rho\ge0}[\tilde{E}_i(\rho,W_{Y\,|\,UX})-\rho(R_1+R_2)]\end{aligned}$$

where $\hat{R}_1=R_1+R_2$, $\hat{R}_2=R_2$, and

$$\begin{aligned}\tilde{E}_1(\rho,W_{Y\,|\,UX}) &= \tilde{E}_2(\rho,W_{Y\,|\,UX})\\ &=\rho-(1+\rho)\log_2\left(\epsilon^{\frac{1}{1+\rho}}+(1-\epsilon)^{\frac{1}{1+\rho}}\right).\end{aligned}$$

It follows from (103) that the upper bound for E_T only depends on the sum rate R_1+R_2 and hence the upper bound can be reduced to

$$\begin{aligned}E_T(Q_{SL},W_{YZ\,|\,UX},\tau) \le \sup_{R>0}\min\{&\max_{\rho\ge0}[\rho R-\tau E_{s1}(\rho,Q_{SL})],\\ &\max_{\rho\ge0}[\tilde{E}_1(\rho,W_{Y\,|\,UX})-\rho R\}.\end{aligned}$$

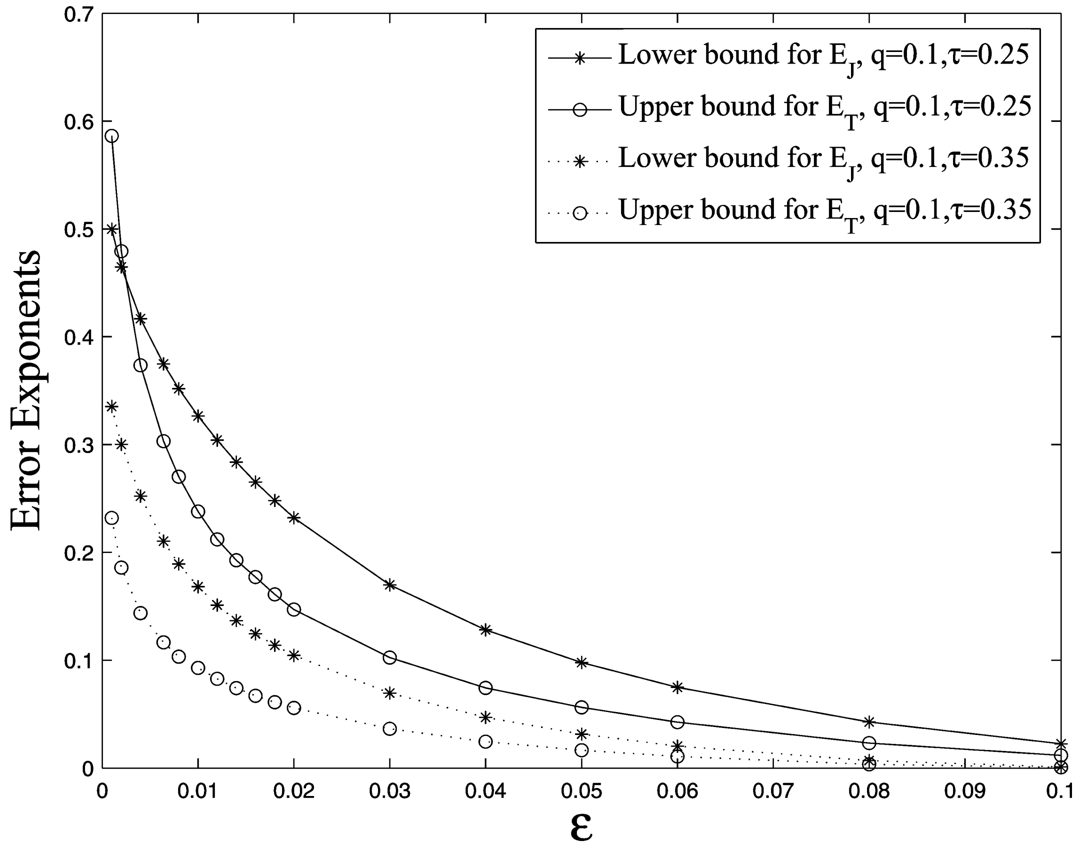

Fig. 7. The lower bound of E_J versus the upper bound of E_T.

In Fig. 7, we plot the lower bound for E_J from (88), and the above upper bound for E_T for different source and channel parameters. It is seen that for a large class of (q, ϵ) pairs with the same transmission rate τ, there is a considerable gap between the upper bound for E_T and the lower bound for E_J, which implies that JSCC can substantially outperform tandem coding in terms of error exponent for many binary CS-AMAC systems with additive noise. In fact, from Fig. 7, we see that E_J almost doubles E_T for many (q, ϵ) pairs. When $E_J \approx 2E_T$ holds, it can be equivalently interpreted that, to achieve the same system error performance, JSCC only requires around half delay of the tandem coding, provided that the coding length is sufficiently large.

VIII. Conclusion

In this paper, we study the error performance and exponents of JSCC for a class of discrete memoryless communication systems which transmit two correlated sources over a two-transmitter two-receiver channel in an "asymmetric" way. For such systems, we derive universally achievable error exponent pairs for the two receivers by employing a generalized type-packing lemma. We also establish a lower and an upper bound for the system JSCC error exponent. We next specialize these results to CS-AMAC and CS-ABC systems. As a special case, we study the analytical computation of the lower and upper bounds for CS-AMAC systems for which the channel admits a symmetric conditional distribution. We show that the lower and upper bounds coincide for many binary CS-AMAC source–channel pairs with additive noise, and hence exactly determine the JSCC error exponent.

As a consequence of our lower bound for the JSCC error exponent, we prove a JSCC theorem for the asymmetric two-user system, i.e., a sufficient and necessary condition for the reliable transmissibility of the two CS over the asymmetric channel is provided. It is demonstrated that the condition can actually be achieved by a tandem coding scheme, which combines separate source and channel coding. This means that tandem coding does not lose optimality from the point of view of reliable transmissibility. Nevertheless, tandem coding might not be optimal in terms of the error exponent. To exploit the advantage of JSCC over tandem coding for the two-user system, we show that the tandem coding exponent can never be larger than the JSCC exponent and we derive a formula for the tandem exponent in terms of the source and channel coding exponents. The formula holds under two basic assumptions on the source code and the assumption that common randomization is used at the transmitter and receiver sides to render the source and channel coding operations statistically decoupled from one another. By numerically comparing the upper bound for the tandem exponent and the lower bound for the JSCC exponent, we note that there is a considerable gain of the JSCC error exponent over the tandem coding error exponent for a large class of binary CS-AMAC systems with additive noise. Note that this prospective benefit of JSCC over tandem coding can also translate into substantial reductions in system complexity and coding delay.

Appendix A
Proof of Lemma 3

Although the result (5) of Lemma 3 was already shown in [8], we include its proof here since we need to show that (5) holds simultaneously with (6) and (7). We employ a random selection argument as used in [8]. For each $i = 1, 2, \ldots, m_n$, we randomly generate a set of $2N_i$ sequences (according to a uniform distribution) from the type class $\mathbb{T}_{A_i} = \mathbb{T}_{P_{A_i}}$,

$$\mathcal{C}_i \triangleq \{\boldsymbol{a}_1^{(i)}, \boldsymbol{a}_2^{(i)}, \ldots, \boldsymbol{a}_{2N_i}^{(i)}\} \subseteq \mathbb{T}_{A_i}$$

i.e., each $\boldsymbol{a}_p^{(i)}$ is randomly drawn from the type class $\mathbb{T}_{A_i}$ with probability $1/|\mathbb{T}_{A_i}|$, $p = 1, 2, \ldots, 2N_i$. Each set has $2N_i$ elements rather than N_i because an expurgation operation will be performed later. Also, we denote the set $\mathcal{C}_i^p \triangleq \mathcal{C}_i / \left\{\boldsymbol{a}_p^{(i)}\right\}$.

Now for each i with associated $j = j(i) = 1, 2, \ldots, m'_{\text{in}}$, we randomly generate $4N_iM_{ij}$ sequences (according to a uniform distribution)

$$\left\{\boldsymbol{b}_{11}^{(j)}, \boldsymbol{b}_{12}^{(j)}, \ldots, \boldsymbol{b}_{1,2M_{ij}}^{(j)}, \boldsymbol{b}_{21}^{(j)}, \boldsymbol{b}_{22}^{(j)}, \ldots, \boldsymbol{b}_{2,2M_{ij}}^{(j)}, \ldots, \boldsymbol{b}_{2N_i,1}^{(j)}, \boldsymbol{b}_{2N_i,2}^{(j)}, \ldots, \boldsymbol{b}_{2N_i,2M_{ij}}^{(j)}\right\}$$

such that the set

$$\begin{aligned}\mathcal{C}_{ij} \triangleq \Big\{ &\left(\boldsymbol{a}_1^{(i)}, \boldsymbol{b}_{11}^{(j)}\right), \left(\boldsymbol{a}_1^{(i)}, \boldsymbol{b}_{12}^{(j)}\right), \ldots, \left(\boldsymbol{a}_1^{(i)}, \boldsymbol{b}_{1,2M_{ij}}^{(j)}\right), \\ &\left(\boldsymbol{a}_2^{(i)}, \boldsymbol{b}_{21}^{(j)}\right), \left(\boldsymbol{a}_2^{(i)}, \boldsymbol{b}_{22}^{(j)}\right), \ldots, \left(\boldsymbol{a}_2^{(i)}, \boldsymbol{b}_{2,2M_{ij}}^{(j)}\right), \\ &\cdots\cdots \\ &\left(\boldsymbol{a}_{2N_i}^{(i)}, \boldsymbol{b}_{2N_i,1}^{(j)}\right), \left(\boldsymbol{a}_{2N_i}^{(i)}, \boldsymbol{b}_{2N_i,2}^{(j)}\right), \ldots, \left(\boldsymbol{a}_{2N_i}^{(i)}, \boldsymbol{b}_{2N_i,2M_{ij}}^{(j)}\right)\Big\} \\ &\subseteq \mathbb{T}_{A_iB_j} = \mathbb{T}_{P_{A_i}P_{B_j|A_i}}.\end{aligned}$$

In other words, each $\boldsymbol{b}_{p,q}^{(j)}$ is drawn from $\mathbb{T}_{B_j|A_i}(\boldsymbol{a}_p^{(i)})$ with probability $1/|\mathbb{T}_{B_j|A_i}(\boldsymbol{a}_p^{(i)})|$, $q = 1, 2, \ldots, M_{ij}$, and hence each pair $(\boldsymbol{a}_p^{(i)}, \boldsymbol{b}_{p,q}^{(j)})$ is drawn from $\mathbb{T}_{A_iB_j}$ with probability $1/\left|\mathbb{T}_{A_iB_j}\right|$. Furthermore, we denote the set $\mathcal{C}_{ij}^{pq} \triangleq \mathcal{C}_{ij}/\{(\boldsymbol{a}_p^{(i)}, \boldsymbol{b}_{p,q}^{(j)})\}$. For any $1 \le i, k \le m_n$, $1 \le j \le m'_{\text{in}}$, and $1 \le l \le m'_{kn}$, define

$$\mathcal{V}_{i,k} \triangleq \left\{ V_{A' \mid A} \in \mathcal{P}_n(\mathcal{A} \mid P_{A_i}) : \sum_{a \in \mathcal{A}} P_{A_i}(a) V_{A' \mid A}(a'|a) = P_{A_k}(a') \right\}$$

and

$$\mathcal{V}_{ij,kl} \triangleq \left\{ V_{A'B' \mid AB} \in \mathcal{P}_n(\mathcal{A} \times \mathcal{B}|P_{A_iB_j}) : \sum_{(a,b) \in \mathcal{A} \times \mathcal{B}} P_{A_iB_j}(a,b) V_{A'B' \mid AB}(a',b'|a,b) = P_{A_kB_l}(a',b') \right\}.$$

Based on the above setup, the following inequalities hold.

i) For any $(i,j) \ne (k,l)$ and any $V_{A'B' \mid AB} \in \mathcal{V}_{ij,kl}$

$$\begin{aligned}&\mathbb{E}\left|\mathbb{T}_{V_{A'B' \mid AB}}\left(\boldsymbol{a}_p^{(i)}, \boldsymbol{b}_{p,q}^{(j)}\right) \bigcap \mathcal{C}_{kl}\right| \\ &\le \mathbb{E}\left|\left\{(p',q') : \left(\boldsymbol{a}_{p'}^{(k)}, \boldsymbol{b}_{p',q'}^{(l)}\right) \in \mathbb{T}_{V_{A'B' \mid AB}}\left(\boldsymbol{a}_p^{(i)}, \boldsymbol{b}_{p,q}^{(j)}\right)\right\}\right| \\ &= 4N_kM_{kl}\Pr\left\{\left(\boldsymbol{a}_1^{(k)}, \boldsymbol{b}_{1,1}^{(l)}\right) \in \mathbb{T}_{V_{A'B' \mid AB}}\left(\boldsymbol{a}_p^{(i)}, \boldsymbol{b}_{p,q}^{(j)}\right)\right\} \\ &= 4N_kM_{kl}\frac{\left|\mathbb{T}_{V_{A'B' \mid AB}}\left(\boldsymbol{a}_p^{(i)}, \boldsymbol{b}_{p,q}^{(j)}\right)\right|}{|\mathbb{T}_{A_kB_l}|} \\ &\le 4N_kM_{kl}(n+1)^{|\mathcal{A}||\mathcal{B}|}2^{-nI_{P_{A_iB_j}V_{A'B' \mid AB}}(A',B';A,B)}\end{aligned} \tag{104}$$

where the above expectation and probability are taken over the uniform distribution

$$\hat{P}_{k,l}\left(\boldsymbol{a}_{p'}^{(k)}, \boldsymbol{b}_{p',q'}^{(l)}\right) \triangleq \frac{1}{|\mathbb{T}_{A_kB_l}|}, \quad \forall\, 1 \le k \le m_n, \quad 1 \le l \le m'_{kn}, \ 1 \le p' \le N_k, \ 1 \le q' \le M_{kl} \tag{105}$$

and (104) follows from the basic facts (Lemma 1) that

$$\left|\mathbb{T}_{V_{A'B' \mid AB}}\left(\boldsymbol{a}_p^{(i)}, \boldsymbol{b}_{p,q}^{(j)}\right)\right| \le 2^{nH_{P_{A_iB_j}V_{A'B' \mid AB}}(A',B'|A,B)}$$

and that

$$|\mathbb{T}_{A_kB_l}| \ge (n+1)^{-|\mathcal{A}||\mathcal{B}|}2^{nH_{P_{A_kB_l}}(A',B')}$$

noting that the marginal distribution of $P_{A_iB_j}V_{A'B' \mid AB}$ for RVs (A',B') is $P_{A_kB_l}$.

ii) For any $(i,j) = (k,l)$ and any $V_{A'B' \mid AB} \in \mathcal{V}_{ij,ij}$, likewise

$$\begin{aligned}&\mathbb{E}\left|\mathbb{T}_{V_{A'B' \mid AB}}\left(\boldsymbol{a}_p^{(i)}, \boldsymbol{b}_{p,q}^{(j)}\right) \bigcap \mathcal{C}_{ij}^{pq}\right| \\ &\le 4N_iM_{ij}(n+1)^{|\mathcal{A}||\mathcal{B}|} \\ &\quad \times 2^{-nI_{P_{A_iB_j}V_{A'B' \mid AB}}(A',B';A,B)}\end{aligned} \tag{106}$$

where the expectation is taken over the uniform distribution $\hat{P}_{i,j}$ defined by (105).

iii) For any i and $j \ne l$, and any $V_{AB',\mid AB} \in \mathcal{V}_{ij,il}$, similarly we have

$$\begin{aligned}&\mathbb{E}\left|\mathbb{T}_{V_{A'B' \mid AB}}\left(\boldsymbol{a}_p^{(i)}, \boldsymbol{b}_{p,q}^{(j)}\right) \bigcap \mathcal{C}_{il}\right| \\ &\le 4N_iM_{il}(n+1)^{|\mathcal{A}||\mathcal{B}|} \\ &\quad \times 2^{-nI_{P_{A_iB_j}V_{AB',\mid AB}}(A,B';A,B)}.\end{aligned}$$

Using the identity

$$\begin{aligned}&I_{P_{A_iB_j}V_{AB',\mid AB}}(A,B';A,B) \\ &\quad = H_{P_{A_i}}(A) + I_{P_{A_iB_j}V_{AB',\mid AB}}(B';B \mid A)\end{aligned}$$

and assumption (3)

$$\frac{1}{n}\log_2 N_i < H_{P_{A_i}}(A) - \delta$$

we obtain another bound

$$\mathbb{E}\left|\mathbb{T}_{V_{A'B'\,|\,AB}}\left(\boldsymbol{a}_p^{(i)},\boldsymbol{b}_{p,q}^{(j)}\right)\bigcap\mathcal{C}_{il}\right| \le 4M_{il}(n+1)^{|\mathcal{A}||\mathcal{B}|} \times 2^{-nI_{P_{A_i B_j}V_{A'B'\,|\,AB}}(B';B\,|\,A)} \quad (107)$$

where the expectation is taken over the uniform distribution $\hat{P}_{i,l}$.

iv) For any i and $j=l$, and any $V_{A'B'\,|\,AB}\in\mathcal{V}_{ij,il}$, likewise

$$\mathbb{E}\left|\mathbb{T}_{V_{A'B'\,|\,AB}}\left(\boldsymbol{a}_p^{(i)},\boldsymbol{b}_{p,q}^{(j)}\right)\bigcap\mathcal{C}_{ij}^{pq}\right| \le 4M_{ij}(n+1)^{|\mathcal{A}||\mathcal{B}|} \times 2^{-nI_{P_{A_i B_j}V_{A'B'\,|\,AB}}(B';B\,|\,A)} \quad (108)$$

where the expectation is taken over the uniform distribution $\hat{P}_{i,j}$.

v) For any $i\neq k$ and any $V_{A'\,|\,A}\in\mathcal{V}_{i,k}$

$$\begin{aligned}\mathbb{E}\left|\mathbb{T}_{V_{A'\,|\,A}}\left(\boldsymbol{a}_p^{(i)}\right)\bigcap\mathcal{C}_k\right| &\le \mathbb{E}\left|\left\{p':\boldsymbol{a}_{p'}^{(k)}\in\mathbb{T}_{V_{A'\,|\,A}}\left(\boldsymbol{a}_p^{(i)}\right)\right\}\right| \\ &= 2N_k\Pr\left\{\boldsymbol{a}_1^{(i)}\in\mathbb{T}_{V_{A'\,|\,A}}\left(\boldsymbol{a}_p^{(i)}\right)\right\} \\ &= 2N_k\frac{\left|\mathbb{T}_{V_{A'\,|\,A}}\left(\boldsymbol{a}_p^{(i)}\right)\right|}{|\mathbb{T}_{A_k}|} \\ &\le 2N_k(n+1)^{-|\mathcal{A}|}2^{-nI_{P_{A_i}V_{A'\,|\,A}}(A';A)}\end{aligned} \quad (109)$$

where the above expectation and probability are taken over the uniform distribution

$$\tilde{P}_k(\boldsymbol{a}_{p'}^{(k)})\triangleq\frac{1}{|\mathbb{T}_{A_k}|},\quad \forall\,1\le k\le m_n,\quad 1\le p'\le N_k \quad (110)$$

and (109) follows from the basic facts (Lemma 1) that

$$\left|\mathbb{T}_{V_{A'\,|\,A}}\left(\boldsymbol{a}_1^{(i)}\right)\right|\le 2^{nH_{P_{A_i}V_{A'\,|\,A}}(A'|A)}$$

and that

$$|\mathbb{T}_{A_k}|\ge(n+1)^{|\mathcal{A}|}2^{nH_{P_{A_k}}(A')}$$

noting that the marginal distribution of $P_{A_i}V_{A'\,|\,A}$ for the RV A' is P_{A_k}.

vi) For any $i=k$ and any $V_{A'\,|\,A}\in\mathcal{V}_{i,k}$, likewise

$$\mathbb{E}\left|\mathbb{T}_{V_{A'\,|\,A}}\left(\boldsymbol{a}_p^{(i)}\right)\bigcap\mathcal{C}_i^{p}\right| \le 2N_k(n+1)^{-|\mathcal{A}|}2^{-nI_{P_{A_i}V_{A'\,|\,A}}(A';A)} \quad (111)$$

where the expectation is taken over the uniform distribution $\tilde{P}_i$ defined in (110).

Note also if $V_{A'B'\,|\,AB}\notin\mathcal{V}_{ij,kl}$

$$\left|\mathbb{T}_{V_{A'B'\,|\,AB}}\left(\boldsymbol{a}_p^{(i)},\boldsymbol{b}_{p,q}^{(j)}\right)\bigcap\mathcal{C}_{kl}\right|=0$$

and if $V_{A'B'\,|\,AB}\notin\mathcal{V}_{ij,ij}$

$$\left|\mathbb{T}_{V_{A'B'\,|\,AB}}\left(\boldsymbol{a}_p^{(i)},\boldsymbol{b}_{p,q}^{(j)}\right)\bigcap\mathcal{C}_{ij}^{pq}\right|=0.$$

Therefore, it follows from (104) and (106) that for any $V_{A'B'\,|\,AB}\in\mathcal{P}_n(\mathcal{A}\times\mathcal{B}|\mathcal{A}\times\mathcal{B})$

$$\begin{aligned}&\frac{\mathbb{E}\left|\mathbb{T}_{V_{A'B'\,|\,AB}}\left(\boldsymbol{a}_p^{(i)},\boldsymbol{x}_{p,q}^{(j)}\right)\bigcap\mathcal{C}_{ij}^{pq}\right|}{4N_iM_{ij}} + \sum_{(k,l)\neq(i,j)}\frac{\mathbb{E}\left|\mathbb{T}_{V_{A'B'\,|\,AB}}\left(\boldsymbol{a}_p^{(i)},\boldsymbol{b}_{p,q}^{(j)}\right)\bigcap\mathcal{C}_{kl}\right|}{4N_kM_{kl}} \\ &\le m_n(\max_i m'_{\text{in}})(n+1)^{|\mathcal{A}||\mathcal{B}|} \times 2^{-nI_{P_{A_iB_j}V_{A'B'\,|\,AB}}(A',B';A,B)}.\end{aligned} \quad (112)$$

Taking the sum over all $V_{A'B'\,|\,AB}\in\mathcal{P}_n(\mathcal{A}\times\mathcal{B}|\mathcal{A}\times\mathcal{B})$, and using the fact (Lemma 1)

$$|\mathcal{P}_n(\mathcal{A}\times\mathcal{B}|\mathcal{A}\times\mathcal{B})|\le(n+1)^{|\mathcal{A}|^2|\mathcal{B}|^2}$$

and $|\mathcal{A}|^2|\mathcal{B}|^2+|\mathcal{A}||\mathcal{B}|\le 2|\mathcal{A}|^2|\mathcal{B}|^2$, we obtain

$$\mathbb{E}S_{ij}^{pq}\le(n+1)^{2|\mathcal{A}|^2|\mathcal{B}|^2}m_n(\max_i m'_{\text{in}})$$

where

$$\begin{aligned}S_{ij}^{pq}\triangleq&\sum_{V_{A'B'\,|\,AB}\in\mathcal{P}_n(\mathcal{A}\times\mathcal{B}|\mathcal{A}\times\mathcal{B})}2^{nI_{P_{A_iB_j}V_{A'B'\,|\,AB}}(A',B';A,B)} \\ &\times\left[\frac{\left|\mathbb{T}_{V_{A'B'\,|\,AB}}\left(\boldsymbol{a}_p^{(i)},\boldsymbol{b}_{p,q}^{(j)}\right)\bigcap\mathcal{C}_{ij}^{pq}\right|}{4N_iM_{ij}} + \sum_{(k,l)\neq(i,j)}\frac{\left|\mathbb{T}_{V_{A'B'\,|\,AB}}\left(\boldsymbol{a}_p^{(i)},\boldsymbol{b}_{p,q}^{(j)}\right)\bigcap\mathcal{C}_{kl}\right|}{4N_kM_{kl}}\right].\end{aligned}$$

Immediately, normalizing by $4N_iM_{ij}$ and taking the sum over $1\le i\le m_n$, $1\le j\le m'_{\text{in}}$, $1\le p\le N_i$,$1\le q\le M_{ij}$ yields

$$\mathbb{E}\sum_{i=1}^{m_n}\sum_{j=1}^{m'_{\text{in}}}\frac{1}{4N_iM_{ij}}\sum_{p=1}^{2N_i}\sum_{q=1}^{2M_{ij}}S_{ij}^{pq} \le(n+1)^{2|\mathcal{A}|^2|\mathcal{B}|^2}m_n^2(\max_i m'_{\text{in}})^2. \quad (113)$$

Similarly, it follows from (107) and (108) that

$$\begin{aligned}\mathbb{E}\sum_{i=1}^{m_n}\sum_{j=1}^{m'_{\text{in}}}\frac{1}{4N_iM_{ij}}\sum_{p=1}^{2N_i}\sum_{q=1}^{2M_{ij}}K_{ij}^{pq} &\le(n+1)^{2|\mathcal{A}|^2|\mathcal{B}|^2}m_n(\max_i m'_{\text{in}})^2 \\ &\le(n+1)^{2|\mathcal{A}|^2|\mathcal{B}|^2}m_n^2(\max_i m'_{\text{in}})^2\end{aligned} \quad (114)$$

where

$$K_{ij}^{pq}\triangleq\sum_{V_{A'B'\,|\,AB}\in\mathcal{P}_n(\mathcal{A}\times\mathcal{B}|\mathcal{A}\times\mathcal{B})}2^{nI_{P_{A_iB_j}V_{A'B'\,|\,AB}}(B';B\,|\,A)}$$

$$\times \left[\frac{\left| \mathbb{T}_{V_{A'B' \mid AB}} \left(\boldsymbol{a}_p^{(i)}, \boldsymbol{b}_{p,q}^{(j)} \right) \bigcap \mathcal{C}_{ij}^{pq} \right|}{4M_{ij}} + \sum_{l \neq j} \frac{\left| \mathbb{T}_{V_{A'B' \mid AB}} \left(\boldsymbol{a}_p^{(i)}, \boldsymbol{b}_{p,q}^{(j)} \right) \bigcap \mathcal{C}_{il} \right|}{4M_{il}} \right]$$

and it follows from (109) and (111) that

$$\mathbb{E} \sum_{i=1}^{m_n} \sum_{j=1}^{m'_{\text{in}}} \frac{1}{4N_i M_{ij}} \sum_{p=1}^{2N_i} \sum_{q=1}^{2M_{ij}} L_{ij}^{pq} \leq (n+1)^{2|\mathcal{A}|^2} m_n^2 (\max_i m'_{\text{in}}) \leq (n+1)^{2|\mathcal{A}|^2 |\mathcal{B}|^2} m_n (\max_i m'_{\text{in}})^2 \tag{115}$$

where L_{ij}^{pq} is actually independent of j and q and is given by

$$L_{ij}^{pq} = L_i^p \triangleq \sum_{V_{A' \mid A} \in \mathcal{P}_n(\mathcal{A}|\mathcal{A})} 2^{nI_{P_{A_i} V_{A' \mid A}}(A';A)} \times \left[\frac{\left| \mathbb{T}_{V_{A' \mid A}} \left(\boldsymbol{a}_p^{(i)} \right) \bigcap \mathcal{C}_i^p \right|}{2N_i} + \sum_{k \neq i} \frac{\left| \mathbb{T}_{V_{A' \mid A}} \left(\boldsymbol{a}_p^{(i)} \right) \bigcap \mathcal{C}_k \right|}{2N_k} \right].$$

Summing (113), (114), and (115) together, we obtain

$$\mathbb{E} \sum_{i=1}^{m_n} \sum_{j=1}^{m'_{\text{in}}} \frac{1}{4N_i M_{ij}} \sum_{p=1}^{2N_i} \sum_{q=1}^{2M_{ij}} \left(S_{ij}^{pq} + K_{ij}^{pq} + L_{ij}^{pq} \right) \leq 3(n+1)^{2|\mathcal{A}|^2 |\mathcal{B}|^2} m_n^2 (\max_i m'_{\text{in}})^2. \tag{116}$$

Therefore, there exists at least a selection of these sets $\{\hat{\mathcal{C}}_i\}_{i=1}^{m_n}$ and $\{\hat{\mathcal{C}}_{ij}\}_{i=1,j=1}^{i=m_n,j=m'_{\text{in}}}$ such that

$$\sum_{i=1}^{m_n} \sum_{j=1}^{m'_{\text{in}}} \frac{1}{4N_i M_{ij}} \sum_{p=1}^{2N_i} \sum_{q=1}^{2M_{ij}} \left(S_{ij}^{pq} + K_{ij}^{pq} + L_{ij}^{pq} \right) \leq 3(n+1)^{2|\mathcal{A}|^2 |\mathcal{B}|^2} m_n^2 (\max_i m'_{\text{in}})^2$$

which implies that for all $i = 1, 2, \ldots, m_n$ and $j = 1, 2, \ldots, m'_{\text{in}}$ the following is satisfied:

$$\frac{1}{4N_i M_{ij}} \sum_{p=1}^{2N_i} \sum_{q=1}^{2M_{ij}} \left(S_{ij}^{pq} + K_{ij}^{pq} + L_{ij}^{pq} \right) \leq 3(n+1)^{2|\mathcal{A}|^2 |\mathcal{B}|^2} m_n^2 (\max_i m'_{\text{in}})^2. \tag{117}$$

We next proceed with an expurgation argument. Without loss of generality, we assume

$$\frac{1}{2M_{ij}} \sum_{q=1}^{2M_{ij}} \left(S_{ij}^{1q} + K_{ij}^{1q} + L_{ij}^{1q} \right) \leq \frac{1}{2M_{ij}} \sum_{q=1}^{2M_{ij}} \left(S_{ij}^{2q} + K_{ij}^{2q} + L_{ij}^{2q} \right) \leq \cdots \leq \frac{1}{2M_{ij}} \sum_{q=1}^{2M_{ij}} \left(S_{ij}^{2N_i,q} + K_{ij}^{2N_i,q} + L_{ij}^{2N_i,q} \right)$$

then we must have, for every $1 \leq p \leq N_i$

$$\frac{1}{2M_{ij}} \sum_{q=1}^{2M_{ij}} S_{ij}^{pq} + K_{ij}^{pq} + L_{ij}^{pq} \leq 6(n+1)^{2|\mathcal{A}|^2 |\mathcal{B}|^2} m_n^2 (\max_i m'_{\text{in}})^2.$$

Similarly, suppose for each $p = 1, 2, \ldots, N_i$

$$S_{ij}^{p1} + K_{ij}^{p1} + L_{ij}^{p1} \leq S_{ij}^{p2} + K_{ij}^{p2} + L_{ij}^{p2} \leq \cdots \leq S_{ij}^{p,2M_{ij}} + K_{ij}^{p,2M_{ij}} + L_{ij}^{p,2M_{ij}}$$

the above implies that for each $p = 1, 2, \ldots, N_i$ and each $q = 1, 2, \ldots, M_{ij}$

$$S_{ij}^{pq} + K_{ij}^{pq} + L_{ij}^{pq} \leq 12(n+1)^{2|\mathcal{A}|^2 |\mathcal{B}|^2} m_n^2 (\max_i m'_{\text{in}})^2. \tag{118}$$

We now let for $i = 1, 2, \ldots, m_n$, $p = 1, 2, \ldots, N_i$, $\Omega_i \triangleq \{\boldsymbol{a}_1^{(i)}, \boldsymbol{a}_2^{(i)}, \ldots, \boldsymbol{a}_{N_i}^{(i)}\} \subseteq \hat{\mathcal{C}}_i$, $\Omega_i^p \triangleq \Omega_i / \{\boldsymbol{a}_p^{(i)}\} \subseteq \hat{\mathcal{C}}_i^p$, and for $j = 1, 2, \ldots, m'_{\text{in}}$, $q = 1, 2, \ldots, M_{ij}$, let $\Omega_{ij}(\boldsymbol{a}_p^{(i)}) = \{(\boldsymbol{a}_p^{(i)}, \boldsymbol{b}_{p,q}^{(j)})\}_{q=1}^{M_{ij}}$ such that

$$\begin{aligned} \Omega_{ij} &\triangleq \bigcup_{p=1}^{N_i} \Omega_{ij}(\boldsymbol{a}_p^{(i)}) \\ &= \Big\{ \left(\boldsymbol{a}_1^{(i)}, \boldsymbol{b}_{11}^{(j)} \right), \left(\boldsymbol{a}_1^{(i)}, \boldsymbol{b}_{12}^{(j)} \right), \ldots, \left(\boldsymbol{a}_1^{(i)}, \boldsymbol{b}_{1,M_{ij}}^{(j)} \right), \\ &\qquad \left(\boldsymbol{a}_2^{(i)}, \boldsymbol{b}_{21}^{(j)} \right), \left(\boldsymbol{a}_2^{(i)}, \boldsymbol{b}_{22}^{(j)} \right), \ldots, \left(\boldsymbol{a}_2^{(i)}, \boldsymbol{b}_{2,M_{ij}}^{(j)} \right), \\ &\qquad \cdots\cdots \\ &\qquad \left(\boldsymbol{a}_{N_i}^{(i)}, \boldsymbol{b}_{N_i,1}^{(j)} \right), \left(\boldsymbol{a}_{N_i}^{(i)}, \boldsymbol{b}_{N_i,2}^{(j)} \right), \ldots, \left(\boldsymbol{a}_{N_i}^{(i)}, \boldsymbol{b}_{N_i,M_{ij}}^{(j)} \right) \Big\} \\ &\subseteq \hat{\mathcal{C}}_{ij} \end{aligned}$$

and denote also $\Omega_{ij}^{pq} \triangleq \Omega_{ij} / \{(\boldsymbol{a}_p^{(i)}, \boldsymbol{b}_{p,q}^{(j)})\} \subseteq \hat{\mathcal{C}}_{ij}^{pq}$. Immediately, it follows from (118) that for every $i = 1, 2, \ldots, m_n$, $j = 1, 2, \ldots, m'_{\text{in}}$, $k = 1, 2, \ldots, m_n$, $l = 1, 2, \ldots, m'_{kn}$, $p = 1, 2, \ldots, N_i$, $q = 1, 2, \ldots, M_{ij}$, and every $V_{A'B' \mid AB} \in \mathcal{P}_n(\mathcal{A} \times \mathcal{B} | \mathcal{A} \times \mathcal{B})$ and $V_{A' \mid A} \in \mathcal{P}_n(\mathcal{A}|\mathcal{A})$

$$\frac{\left| \mathbb{T}_{V_{A'B' \mid AB}} \left(\boldsymbol{a}_p^{(i)}, \boldsymbol{b}_{p,q}^{(j)} \right) \bigcap \Omega_{kl} \right|}{N_k M_{kl}}$$

$$\leq 2^{-n[I_{P_{A_iB_j}V_{A'B'\mid AB}}(A',B';A,B)-\delta]}, \quad (k,l)\neq(i,j) \tag{119}$$

$$\frac{\left|\mathbb{T}_{V_{A'B'\mid AB}}\left(\boldsymbol{a}_p^{(i)},\boldsymbol{b}_{p,q}^{(j)}\right)\bigcap\Omega_{ij}^{pq}\right|}{N_iM_{ij}} \leq 2^{-n[I_{P_{A_iB_j}V_{A'B'\mid AB}}(A',B';A,B)-\delta]} \tag{120}$$

$$\frac{\left|\mathbb{T}_{V_{A'B'\mid AB}}\left(\boldsymbol{a}_p^{(i)},\boldsymbol{b}_{p,q}^{(j)}\right)\bigcap\Omega_{il}\right|}{M_{il}} \leq 2^{-n[I_{P_{A_iB_j}V_{A'B'\mid AB}}(B';B\mid A)-\delta]}, \quad l\neq j \tag{121}$$

$$\frac{\left|\mathbb{T}_{V_{A'B'\mid AB}}\left(\boldsymbol{a}_p^{(i)},\boldsymbol{b}_{p,q}^{(j)}\right)\bigcap\Omega_{ij}^{pq}\right|}{M_{ij}} \leq 2^{-n[I_{P_{A_iB_j}V_{A'B'\mid AB}}(B';B\mid A)-\delta]} \tag{122}$$

$$\frac{\left|\mathbb{T}_{V_{A'\mid A}}\left(\boldsymbol{a}_p^{(i)}\right)\bigcap\Omega_{k}\right|}{N_k} \leq 2^{-n[I_{P_{A_i}V_{A'\mid A}}(A';A)-\delta]}, \quad k\neq i \tag{123}$$

$$\frac{\left|\mathbb{T}_{V_{A'\mid A}}\left(\boldsymbol{a}_p^{(i)}\right)\bigcap\Omega_{i}^{p}\right|}{N_i} \leq 2^{-n[I_{P_{A_i}V_{A'\mid A}}(A';A)-\delta]} \tag{124}$$

where

$$\delta=\frac{2}{n}[|\mathcal{A}|^2|\mathcal{B}|^2\log_2(n+1)+\log_2 m_n + \log_2(\max_i m'_{\text{in}})+\log_2 12].$$

Thus far, we proved the existence of the sets Ω_i and Ω_{ij} with elements selected uniformly from each $\mathbb{T}_{A_i}$ and $\mathbb{T}_{A_iB_j}$ satisfying the inequalities (119)–(124) for any $V_{A'\mid A}$ and $V_{A'B'\mid AB}$. It remains to show that these sets are disjoint and have distinct elements provided assumptions (3) and (4). Indeed, since (123) and (124) hold for every $V_{A'\mid A}\in\mathcal{P}_n(\mathcal{A}|\mathcal{A})$, they of course hold when $V_{A'\mid A}$ is a conditional distribution such that $V^*_{A'\mid A}(a'|a)$ is 1 if $a'=a$ and 0 otherwise. It then follows from (3)

$$\frac{1}{n}\log_2 N_i < H_{P_{A_i}}(A)-\delta = I_{P_{A_i}V^*_{A'\mid A}}(A';A)-\delta$$

that $|\mathbb{T}_{V^*_{A'\mid A}}(\boldsymbol{a}_p^{(i)})\bigcap\Omega_k| = |\{\boldsymbol{a}_p^{(i)}\}\bigcap\Omega_k|<1$ or, equivalently, $|\{\boldsymbol{a}_p^{(i)}\}\bigcap\Omega_k|=0$, which means any elements in Ω_i does not belong to Ω_k for $i\neq k$, i.e., Ω_i and Ω_k are disjoint. Likewise, using assumption (3) in (124), we see that

$$\left|\mathbb{T}_{V^*_{A'\mid A}}\left(\boldsymbol{a}_p^{(i)}\right)\bigcap\Omega_i^p\right| = \left|\left\{\boldsymbol{a}_p^{(i)}\right\}\bigcap\Omega_i^p\right|=0$$

which means that Ω_i has N_i disjoint elements. Similarly, setting $V_{A'B'\mid AB}$ to be the conditional distribution such that $V^*_{A'B'\mid AB}(a',b'|a,b)$ is 1 if $a'=a$, $b'=b$, and 0 otherwise, and using (4)

$$\frac{1}{n}\log_2 M_{ij} < H_{P_{A_i}P_{B_j|A_i}}(B\mid A)-\delta$$

we see that for any $\boldsymbol{a}_p^{(i)}\in\Omega_i$, $\Omega_{ij}(\boldsymbol{a}_p^{(i)})$'s are disjoint and the elements in $\Omega_{ij}(\boldsymbol{a}_p^{(i)})$ are all distinct, i.e., $|\Omega_{ij}(\boldsymbol{a}_p^{(i)})| = M_{ij}$ for every $\boldsymbol{a}^{(i)}\in\Omega_i$. Finally, when $V_{A'\mid A}$ is not the conditional distribution such that $V_{A'\mid A}(a'|a)$ is 1 if $a'=a$ and 0 otherwise, we can write (123) and (124) in the same way as (5), and when $V_{A'B'\mid AB}$ is not the conditional distribution such that $V_{A'B'\mid AB}(a',b'|a,b)$ is 1 if $a'=a$, $b'=b$, and 0 otherwise, we can write (119)–(120) as (6), and write (121)–(122) as (7), since

$$\left|\mathbb{T}_{V_{A'\mid A}}\left(\boldsymbol{a}_p^{(i)}\right)\bigcap\Omega_i^p\right| = \left|\mathbb{T}_{V_{A'\mid A}}\left(\boldsymbol{a}_p^{(i)}\right)\bigcap\Omega_i\right|$$
$$\left|\mathbb{T}_{V_{A'B'\mid AB}}\left(\boldsymbol{a}_p^{(i)},\boldsymbol{b}_{p,q}^{(j)}\right)\bigcap\Omega_{ij}^{pq}\right| = \left|\mathbb{T}_{V_{A'B'\mid AB}}\left(\boldsymbol{a}_p^{(i)},\boldsymbol{b}_{p,q}^{(j)}\right)\bigcap\Omega_{ij}\right|$$
$$\left|\mathbb{T}_{V_{A'B'\mid AB}}\left(\boldsymbol{a}_p^{(i)},\boldsymbol{b}_{p,q}^{(j)}\right)\bigcap\Omega_{ij}^{pq}\right| = \left|\mathbb{T}_{V_{A'B'\mid AB}}\left(\boldsymbol{a}_p^{(i)},\boldsymbol{b}_{p,q}^{(j)}\right)\bigcap\Omega_{ij}\right|. \qquad \square$$

APPENDIX B
PROOF OF (26) AND (27)

A. Upper Bound on $|\mathbb{T}_{\hat{V}_{Y\mid TUX}}((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x})\bigcap\mathcal{E}_1|$

If we fix a $k=1,2,\ldots,m_n$ and an $l=1,2,\ldots,m'_{kn}$, then $\mathcal{E}_1$ is the set of all $\boldsymbol{y}$ such that there exist some $((\boldsymbol{t},\boldsymbol{u})',\boldsymbol{x}')\in\Omega_{kl}$, $(\boldsymbol{t},\boldsymbol{u})'\neq(\boldsymbol{t},\boldsymbol{u})$, $((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x},(\boldsymbol{t},\boldsymbol{u})',\boldsymbol{x}',\boldsymbol{y})$ admits a joint type

$$P_{(\boldsymbol{t},\boldsymbol{u})\boldsymbol{x}(\boldsymbol{t},\boldsymbol{u})'\boldsymbol{x}'\boldsymbol{y}}\in\mathcal{P}_n(\mathcal{T}^2\times\mathcal{U}^2\times\mathcal{X}^2\times\mathcal{Y})$$

and

$$I((\boldsymbol{t},\boldsymbol{u})',\boldsymbol{x}';\boldsymbol{y})-(R_k+R_{kl}) \geq I((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x};\boldsymbol{y})-(R_i+R_{ij}). \tag{125}$$

Note that (125) can be represented as for dummy RVs $(TU)\in\mathcal{T}\times\mathcal{U}$, $X\in\mathcal{X}$, $(TU)'\in\mathcal{T}\times\mathcal{U}$, $X'\in\mathcal{X}$, and $Y\in\mathcal{Y}$, the following holds under the joint distribution $P_{(TU)X(TU)'X'Y}=P_{(\boldsymbol{t},\boldsymbol{u})\boldsymbol{x}(\boldsymbol{t},\boldsymbol{u})'\boldsymbol{x}'\boldsymbol{y}}$:

$$I_{P_{(TU)'X'Y}}((T,U)',X';Y)-(R_k+R_{kl}) \geq I_{P_{TUXY}}((T,U),X;Y)-(R_i+R_{ij}),$$

where $P_{(TU)'X'Y}$ and P_{TUXY} are the corresponding marginal distributions induced by $P_{(TU)X(TU)'X'Y}$. Thus, $\mathbb{T}_{\hat{V}_{Y\mid TUX}}((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x})\bigcap\mathcal{E}_1$ can be written as a union of subsets

$$\mathbb{T}_{\hat{V}_{Y\mid (TU)X}}((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x})\bigcap\mathcal{E}_1 = \bigcup_{k=1}^{m_n}\bigcup_{l=1}^{m'_{kn}}\bigcup_{P_{(TU)X(TU)'X'Y}\in\mathcal{C}_{k,l}((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x})}\mathcal{F}_{k,l}((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x},P_{(TU)X(TU)'X'Y}) \tag{126}$$

where $\mathcal{C}_{k,l}((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x})$ is fiven in the first equation at the bottom of the page, where $P_{(TU)X}$, $P_{(TU)'X'}$, and $P_{Y\mid (TU)X}$, etc., are the corresponding marginal and conditional distributions induced by $P_{(TU)X(TU)'X'Y}$, and $\mathcal{F}_{k,l}((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x},P_{(TU)X(TU)'X'Y})$ is defined in the second expression at the bottom of the page, where $\mathbb{T}_{(TU)X(TU)'X'Y}\triangleq\mathbb{T}_{P_{(TU)X(TU)'X'Y}}$. Clearly, given any k, l, and $P_{(TU)X(TU)'X'Y}$, we get (127) at the top of the following page, where the last inequality follows from

$$\begin{aligned}\left|\mathcal{F}_{k,l}((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x},P_{(TU)X(TU)'X'Y})\right| &\le \left|\left\{((\boldsymbol{t},\boldsymbol{u})',\boldsymbol{x}',\boldsymbol{y}) : \begin{array}{l} ((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x},(\boldsymbol{t},\boldsymbol{u})',\boldsymbol{x}',\boldsymbol{y}) \in \mathbb{T}_{(TU)X(TU)'X'Y} \\ ((\boldsymbol{t},\boldsymbol{u})',\boldsymbol{x}')\in\Omega_{kl}, \quad (\boldsymbol{t},\boldsymbol{u})'\neq(\boldsymbol{t},\boldsymbol{u})\end{array}\right\}\right| \\ &= \left|\left\{((\boldsymbol{t},\boldsymbol{u})',\boldsymbol{x}') : \begin{array}{l} ((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x},(\boldsymbol{t},\boldsymbol{u})',\boldsymbol{x}') \in \mathbb{T}_{(TU)X(TU)'X'} \\ ((\boldsymbol{t},\boldsymbol{u})',\boldsymbol{x}')\in\Omega_{kl}, \quad (\boldsymbol{t},\boldsymbol{u})'\neq(\boldsymbol{t},\boldsymbol{u})\end{array}\right\}\right| \\ &\quad \times \left|\mathbb{T}_{Y|(TU)X(TU)'X'}((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x},(\boldsymbol{t},\boldsymbol{u})',\boldsymbol{x}')\right| \\ &\le N_k M_{kl} 2^{-n[I_{P_{(TU)X(TU)'X'}}((T,U),X;(T,U)',X')-\eta]} 2^{nH_{P_{(TU)X(TU)'X'Y}}(Y|(T,U),X,(T,U)',X')}\end{aligned} \tag{127}$$

Lemma 3. Meanwhile, when $((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x}) \in \Omega_{ij}$, the following simple bound also holds:

$$\begin{aligned}\left|\mathcal{F}_{k,l}((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x},P_{(TU)X(TU)'X'Y})\right| &\le \left|\mathbb{T}_{Y\,|\,(TU)X}((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x})\right| \\ &\le 2^{nH_{P_{(TU)XY}}(Y\,|\,(T,U),X)} \\ &= 2^{nH_{P_{((TU))_iX_j}\hat{V}_{Y\,|\,(TU)X}}(Y\,|\,(T,U),X)}\end{aligned} \tag{128}$$

since for each $\mathbb{T}_{(TU)X(TU)'X'Y} \in \mathcal{C}_{k,l}((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x})$, we have $P_{(TU)X} = P_{((TU))_iX_j}$, $P_{Y\,|\,(TU)X} = \hat{V}_{Y\,|\,(TU)X}$, and hence $P_{(TU)XY} = P_{((TU))_iX_j}\hat{V}_{Y\,|\,(TU)X}$. Now substituting the following inequality (cf. [8, Eq. (28)])

$$\begin{aligned}&H_{P_{(TU)X(TU)'X'Y}}(Y|(T,U),X,(T,U)',X') \\ &\quad - I_{P_{(TU)X(TU)'X'}}((T,U),X;(T,U)',X') \\ &\qquad = H_{P_{(TU)XY}}(Y\,|\,(T,U),X) \\ &\qquad\quad - I_{P_{(TU)X(TU)'X'Y}}((T,U)',X';(T,U),X,Y) \\ &\qquad \le H_{P_{(TU)XY}}(Y\,|\,(T,U),X) \\ &\qquad\quad - I_{P_{(TU)'X'Y}}((T,U)',X';Y)\end{aligned} \tag{129}$$

into (127), combining with (128) together, we obtain (130) shown at the bottom of the page. Again recall that for $P_{(TU)X(TU)'X'Y} \in \mathcal{C}_{k,l}((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x})$, $P_{(TU)XY} = P_{((TU))_iX_j}\hat{V}_{Y\,|\,(TU)X}$, and note that

$$\begin{aligned}&I_{P_{(TU)'X'Y}}((T,U)',X';Y) - (R_k+R_{kl}) \\ &\qquad\qquad \ge I_{P_{(TU)XY}}((T,U),X;Y) - (R_i+R_{ij}).\end{aligned}$$

This implies when $P_{(TU)X(TU)'X'Y} \in \mathcal{C}_{k,l}((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x})$, as shown by the second expression at the bottom of the page,and hence the third expression at the bottom of the page, since by Lemma 1

$$\begin{aligned}|\mathcal{C}_{k,l}((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x})| &\le \left|\mathcal{P}_n(\mathcal{T}^2\times\mathcal{U}^2\times\mathcal{X}^2\times\mathcal{Y})\right| \\ &\le (n+1)^{|\mathcal{T}|^2|\mathcal{U}|^2|\mathcal{X}|^2|\mathcal{Y}|}.\end{aligned}$$

B. *Upper Bound on* $|\mathbb{T}_{\hat{V}_{Y\,|\,(TU)X}}((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x})\bigcap\mathcal{E}_2|$

If we fix an $i=1,2,\ldots,m_n$ and an $l=1,2,\ldots,m'_{\rm in}$, then $\mathcal{E}_2$ is the set of all $\boldsymbol{y}$ such that there exist some $((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x}') \in \Omega_{il}$, $\boldsymbol{x}'\neq\boldsymbol{x}$, $((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x},\boldsymbol{x}',\boldsymbol{y})$ admits a joint type $P_{(\boldsymbol{t},\boldsymbol{u})\boldsymbol{x}\boldsymbol{x}'\boldsymbol{y}} \in \mathcal{P}_n(\mathcal{T}\times\mathcal{U}\times\mathcal{X}^2\times\mathcal{Y})$ and

$$I((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x}';\boldsymbol{y}) - (R_i+R_{il}) \ge I((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x};\boldsymbol{y}) - (R_i+R_{ij}). \tag{131}$$

Using the identity

$$I((T,U),X;Y) = I(T,U;Y) + I(X;Y\,|\,T,U)$$

on both sides of (131) we see it is equivalent to

$$I(\boldsymbol{x}';\boldsymbol{y}\,|\,\boldsymbol{t},\boldsymbol{u}) - R_{il} \ge I(\boldsymbol{x};\boldsymbol{y}|\boldsymbol{t},\boldsymbol{u}) - R_{ij}. \tag{132}$$

$$\mathcal{C}_{k,l}((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x}) \triangleq \left\{\begin{array}{ll} P_{(TU)X(TU)'X'Y} & P_{(TU)X} = P_{(\boldsymbol{t},\boldsymbol{u})\boldsymbol{x}} = P_{(TU)_iX_j}, \\ \in\mathcal{P}_n(\mathcal{T}^2\times\mathcal{U}^2\times\mathcal{X}^2\times\mathcal{Y}): & P_{(TU)'X'} = P_{(TU)_kX_l}, \quad P_{Y\,|\,(TU)X} = \hat{V}_{Y\,|\,(TU)X}, \\ & I_{P_{(TU)'X'Y}}((T,U)',X';Y) - (R_k+R_{kl}) \\ & \qquad \ge I_{P_{(TU)XY}}((T,U),X;Y) - (R_i+R_{ij}) \end{array}\right\}$$

$$\mathcal{F}_{k,l}((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x},P_{(TU)X(TU)'X'Y}) \triangleq \left\{\boldsymbol{y} : \begin{array}{ll} \exists \quad ((\boldsymbol{t},\boldsymbol{u})',\boldsymbol{x}') & ((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x},(\boldsymbol{t},\boldsymbol{u})',\boldsymbol{x}',\boldsymbol{y}) \in \mathbb{T}_{(TU)X(TU)'X'Y} \\ \text{such that} & ((\boldsymbol{t},\boldsymbol{u})',\boldsymbol{x}')\in\Omega_{kl}, \quad (\boldsymbol{t},\boldsymbol{u})'\neq(\boldsymbol{t},\boldsymbol{u}) \end{array}\right\}$$

$$\left|\mathcal{F}_{k,l}((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x},P_{(TU)X(TU)'X'Y})\right| \le 2^{n[H_{P_{((TU))_iX_j}\hat{V}_{Y\,|\,(TU)X}}(Y\,|\,(T,U),X)-|I_{P_{(TU)'X'Y}}((T,U)',X';Y)-(R_k+R_{kl})|^+]}. \tag{130}$$

$$\left|\mathcal{F}_{k,l}((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x},P_{(TU)X(TU)'X'Y})\right| \le 2^{n[H_{P_{((TU))_iX_j}\hat{V}_{Y\,|\,(TU)X}}(Y\,|\,(T,U),X)-|I_{P_{((TU))_iX_j}\hat{V}_{Y\,|\,(TU)X}}((T,U),X;Y)-(R_i+R_{ij})|^+]}$$

$$\begin{aligned}\left|\mathbb{T}_{\hat{V}_{Y\,|\,(TU)X}}((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x})\bigcap\mathcal{E}_1\right| &\le m_n\left(\max_i m'_{\rm in}\right)(n+1)^{|\mathcal{T}\times\mathcal{U}|^2|\mathcal{X}|^2|\mathcal{Y}|} \\ &\quad \times 2^{n[H_{P_{((TU))_iX_j}\hat{V}_{Y\,|\,(TU)X}}(Y\,|\,(T,U),X)-|(I_{P_{((TU))_iX_j}\hat{V}_{Y\,|\,(TU)X}}((T,U),X;Y)-(R_i+R_{ij}))|^+]}\end{aligned}$$

Note that (132) can be represented as for dummy RVs $(TU) \in \mathcal{T}\times\mathcal{U}$, $X \in \mathcal{X}$, $X' \in \mathcal{X}$, and $Y \in \mathcal{Y}$, the following holds under the joint distribution $P_{(TU)XX'Y} = P_{(\boldsymbol{t},\boldsymbol{u})\boldsymbol{x}\boldsymbol{x}'\boldsymbol{y}}$:

$$I_{P_{(TU)X'Y}}(X';Y\,|\,T,U)-R_{il} \geq I_{P_{(TU)XY}}(X;Y\,|\,T,U)-R_{ij}$$

where $P_{(TU)XY}$ and $P_{(TU)X'Y}$ are the corresponding marginal distributions induced by $P_{(TU)XX'Y}$. Thus, $\mathbb{T}_{\hat{V}_{Y\,|\,(TU)X}}((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x})\bigcap\mathcal{E}_2$ can be written as a union of subsets

$$\mathbb{T}_{\hat{V}_{Y\,|\,(TU)X}}((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x})\bigcap\mathcal{E}_2 = \bigcup_{l=1}^{m'_{\text{in}}} \bigcup_{P_{(TU)XX'Y}\in\mathcal{C}_l((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x})} \mathcal{F}_l((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x},P_{(TU)XX'Y}) \tag{133}$$

where $\mathcal{C}_l((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x})$ is given in the last expression at the bottom of the page, where $P_{(TU)X}$, $P_{(TU)X'}$, and $P_{Y\,|\,(TU)X}$, etc., are the corresponding marginal and conditional distributions induced by $P_{(TU)XX'Y}$, and

$$\mathcal{F}_l((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x},P_{(TU)XX'Y}) \triangleq \left\{\boldsymbol{y}: \begin{array}{ll} \exists\ \ ((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x}') & ((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x},\boldsymbol{x}',\boldsymbol{y})\in\mathbb{T}_{(TU)XX'Y} \\ \text{such that} & ((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x}')\in\Omega_{il},\quad \boldsymbol{x}'\neq\boldsymbol{x} \end{array}\right\}$$

where $\mathbb{T}_{(TU)XX'Y} = \mathbb{T}_{P_{(TU)XX'Y}}$. Using a similar counting argument, and applying Lemma 3, we can bound, for any $l = 1,2,\ldots,m'_{\text{in}}$ and $P_{(TU)XX'Y} \in \mathcal{C}_l((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x})$ we have the first expression at the bottom of the page, and, finally, we obtain the second expression at the bottom of the page, since $|\mathcal{C}_l((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x})| \leq (n+1)^{|\mathcal{T}||\mathcal{U}||\mathcal{X}|^2|\mathcal{Y}|}$. □

APPENDIX C
PROOF OF THEOREM 3

Forward Part (1): It follows from (19)–(21) that $E_r(R_1,R_2,W_{YZ\,|\,TUX},P_{TUX}) > 0$ if and only if $(R_1,R_2) \in \mathcal{R}(W_{YZ\,|\,TUX},P_{TUX})$. It then follows that $E_r(R_1,R_2,W_{YZ\,|\,UX}) > 0$ if $(R_1,R_2) \in \mathcal{R}(W_{YZ\,|\,UX})$. According to Theorem 2 and the definition of the system JSCC error exponent, $P_e^{(n)} \to 0$ if the lower bound (34) is positive, which needs

$$E_r(\tau H_P(S),\tau H_P(L\,|\,S),W_{YZ\,|\,UX}) > 0.$$

This means $P_e^{(n)} \to 0$ if the pair $(\tau H_Q(S),\tau H_Q(L\,|\,S)) \in \mathcal{R}(W_{YZ\,|\,UX})$.

Converse Part (2): The proof follows in a similar manner as the converse part of [16, Theorem 1] for a broadcast channel. For the sake of completeness, we also provide a full proof here since we deal with a two-user channel. We first prove the following lemma.

Lemma 7:

$$\bar{\mathcal{R}}(W_{YZ\,|\,UX}) = \mathcal{R}'(W_{YZ\,|\,UX})$$

where $\mathcal{R}'(W_{YZ\,|\,UX})$ is defined in (49).

Proof: It is straightforward to see that $\bar{\mathcal{R}}(W_{YZ\,|\,UX}) \subseteq \mathcal{R}'(W_{YZ\,|\,UX})$. To complete the proof, it suffices to show $\mathcal{R}'(W_{YZ\,|\,UX}) \subseteq \bar{\mathcal{R}}(W_{YZ\,|\,UX})$. We note that both $\bar{\mathcal{R}}(W_{YZ\,|\,UX})$ and $\mathcal{R}'(W_{YZ\,|\,UX})$ are convex and closed. Therefore, instead of verifying that all (R_1,R_2)'s in $\mathcal{R}'(W_{YZ\,|\,UX})$ belong to $\bar{\mathcal{R}}(W_{YZ\,|\,UX})$, we show that all the boundary points of $\mathcal{R}'(W_{YZ\,|\,UX})$ are in $\bar{\mathcal{R}}(W_{YZ\,|\,UX})$. By the definition of $\mathcal{R}'(W_{YZ\,|\,UX})$, we note that any boundary point (R_1,R_2) of $\mathcal{R}'(W_{YZ\,|\,UX})$ has to satisfy at least one of the following conditions.

- Case 1: there exist RV T and P_{TUX} such that

$$\begin{aligned} I(U,X;Y) &\leq I(X;Y\,|\,T,U) + I(T,U;Z) \\ R_1+R_2 &= I(U,X;Y) \\ R_1 &= I(T,U;Z). \end{aligned}$$

  This is true since if $R_1+R_2 < I(U,X;Y)$ or $R_1 < I(T,U;Z)$, we can increase R_1 or R_2 which contradicts the boundary point assumption on (R_1,R_2).
- Case 2: there exist RV T and P_{TUX} such that

$$\begin{aligned} I(U,X;Y) &\geq I(X;Y\,|\,T,U) + I(T,U;Z) \\ R_1+R_2 &= I(X;Y\,|\,T,U) + I(T,U;Z) \\ R_1 &= I(T,U;Z). \end{aligned}$$

Now if the boundary point (R_1,R_2) satisfies Case 1, clearly, for the same T and P_{TUX}, we have

$$\begin{aligned} R_1+R_2 &= I(U,X;Y) \\ R_1 &= I(T,U;Z) \\ R_2 &\leq I(X;Y\,|\,T,U). \end{aligned}$$

$$\mathcal{C}_l((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x}) \triangleq \left\{ \begin{array}{ll} P_{(TU)XX'Y} & P_{(TU)X} = P_{(\boldsymbol{t},\boldsymbol{u})\boldsymbol{x}} = P_{(TU)_iX_j}, \\ \in\mathcal{P}_n(\mathcal{T}\times\mathcal{U}\times\mathcal{X}^2\times\mathcal{Y}): & P_{(TU)X'} = P_{(TU)_iX_l},\quad P_{Y\,|\,(TU)X} = \hat{V}_{Y\,|\,TUX} \\ & I_{P_{(TU)X'Y}}(X';Y\,|\,T,U) - R_{il} \\ & \quad \geq I_{P_{(TU)XY}}(X;Y\,|\,T,U) - R_{ij} \end{array} \right\}$$

$$\left|\mathcal{F}_l((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x},P_{(TU)XX'Y})\right| \leq 2^{n[H_{P_{((TU))_iX_j}\hat{V}_{Y\,|\,(TU)X}}(Y\,|\,(T,U),X) - |I_{P_{((TU))_iX_j}\hat{V}_{Y\,|\,(TU)X}}(X;Y\,|\,T,U)-R_{ij}|^+]}$$

$$\left|\mathbb{T}_{\hat{V}_{Y\,|\,(TU)X}}((\boldsymbol{t},\boldsymbol{u}),\boldsymbol{x})\bigcap\mathcal{E}_2\right| \leq \left(\max_i m'_{\text{in}}\right)(n+1)^{|\mathcal{T}\times\mathcal{U}||\mathcal{X}|^2|\mathcal{Y}|} \times 2^{n[H_{P_{((TU))_iX_j}\hat{V}_{Y\,|\,(TU)X}}(Y\,|\,(T,U),X) - |I_{P_{((TU))_iX_j}\hat{V}_{Y\,|\,(TU)X}}(X;Y\,|\,T,U)-R_{ij}|^+]}$$

This shows that $(R_1,R_2) \in \bar{\mathcal{R}}(W_{YZ\,|\,UX})$. Similarly, if the boundary point (R_1,R_2) satisfies Case 2, for such T and P_{TUX}, we have

$$\begin{aligned} R_2 &= I(X;Y\,|\,T,U) \\ R_1 &= I(T,U;Z) \\ R_1+R_2 &\le I(U,X;Y) \end{aligned}$$

and thus, $(R_1,R_2) \in \bar{\mathcal{R}}(W_{YZ\,|\,UX})$. Since the boundary points of $\mathcal{R}'(W_{YZ\,|\,UX})$ are in $\bar{\mathcal{R}}(W_{YZ\,|\,UX})$, we conclude that the entire region of $\mathcal{R}'(W_{YZ\,|\,UX})$ is in $\bar{\mathcal{R}}(W_{YZ\,|\,UX})$, and hence Lemma 7 is proved. □

By lemma 7, it suffices to show that, for any $\epsilon > 0$, if

$$\max\left\{P_{Ye}^{(n)}(Q_{SL}, W_{YZ|XU}, \tau), P_{Ze}^{(n)}(Q_{SL}, W_{YZ\,|\,UX}, \tau)\right\} \le \epsilon_n \to 0$$

as n goes to infinity, then there exists an RV T satisfying $T \to (U,X) \to (Y,Z)$, i.e., the joint distribution P_{TUXYZ} can be factorized as $P_T P_{UX\,|\,T} W_{YZ\,|\,UX}$, such that

$$(\tau H_Q(S), \tau H_Q(L\,|\,S)) \in \mathcal{R}'(W_{YZ\,|\,UX}, P_{TUX})$$

i.e.,

$$\begin{aligned} \tau H_Q(S,L) &\le \min\{I(U,X;Y), I(X;Y\,|\,T,U) + I(T,U;Z)\} \\ \tau H_Q(S) &\le I(T,U;Z). \end{aligned}$$

Fix $k = \tau n$. Fano's inequality gives

$$H(S^k, L^k | Y^n) \le P_{Ye}^{(n)} \log_2 |\mathcal{S}^k \times \mathcal{L}^k| + H\left(P_{Ye}^{(n)}\right) \triangleq n\epsilon_{1n} \tag{134}$$

$$H(S^k | Z^n) \le P_{Ze}^{(n)} \log_2 |\mathcal{S}^k| + H\left(P_{Ze}^{(n)}\right) \triangleq n\epsilon_{2n} \tag{135}$$

where $S^k \triangleq (S_1, S_2, \ldots, S_k)$; similar definitions apply for the other tuples. It follows from (134)–(135) that

$$\begin{aligned} &kH(S,L) \\ &\quad = H(L^k|S^k) + H(S^k) \\ &\quad = I(L^k;Y^n|S^k) + H(L^k|S^k,Y^n) + I(S^k;Z^n) \\ &\qquad + H(S^k|Z^n) \\ &\quad \le \sum_{i=1}^{n} [I(L^k;Y_i|S^k,Y^{i-1}) + I(S^k;Z_i|\boldsymbol{Z}^{i+1})] \\ &\qquad + H(S^k,L^k|Y^n) + n\epsilon_{2n} \\ &\quad \le \sum_{i=1}^{n} [I(L^k,\boldsymbol{Z}^{i+1};Y_i|S^k,Y^{i-1}) \\ &\qquad + I(S^k,Y^{i-1};Z_i|\boldsymbol{Z}^{i+1}) \\ &\qquad - I(Y^{i-1};Z_i|S^k,\boldsymbol{Z}^{i+1})] + n(\epsilon_{1n}+\epsilon_{2n}), \\ &\quad \le \sum_{i=1}^{n} [I(L^k;Y_i|S^k,Y^{i-1},\boldsymbol{Z}^{i+1}) \\ &\qquad + I(\boldsymbol{Z}^{i+1};Y_i|S^k,Y^{i-1}) \\ &\qquad + I(S^k,\boldsymbol{Z}^{i+1},Y^{i-1};Z_i) - I(Y^{i-1};Z_i|S^k,\boldsymbol{Z}^{i+1})] \\ &\qquad + n(\epsilon_{1n}+\epsilon_{2n}) \end{aligned}$$

where

$Y^{i-1} = (Y_1, Y_2, \ldots, Y_{i-1})$ and $\boldsymbol{Z}^{i+1} \triangleq (Z_{i+1}, Z_{i+2}, \ldots, Z_n)$.

Substituting the identity [11, Lemma 7]

$$\sum_{i=1}^{n} I(\boldsymbol{Z}^{i+1};Y_i|S^k,Y^{i-1}) = \sum_{i=1}^{n} I(Y^{i-1};Z_i|S^k,\boldsymbol{Z}^{i+1})$$

into the above, and setting $T_i = (S^k, Y^{i-1}, \boldsymbol{Z}^{i+1})$ for $1 \le i \le n$ yields

$$\begin{aligned} &kH(S,L) \\ &\quad \le \sum_{i=1}^{n} [I(L^k;Y_i|T_i) + I(T_i;Z_i)] + n(\epsilon_{1n}+\epsilon_{2n}) \\ &\quad \overset{\text{(a)}}{=} \sum_{i=1}^{n} [I(L^k;Y_i|T_i,U_i) + I(T_i,U_i;Z_i)] + n(\epsilon_{1n}+\epsilon_{2n}) \\ &\quad \overset{\text{(b)}}{\le} \sum_{i=1}^{n} [I(X^n;Y_i|T_i,U_i) + I(T_i,U_i;Z_i)] + n(\epsilon_{1n}+\epsilon_{2n}) \\ &\quad \overset{\text{(c)}}{=} \sum_{i=1}^{n} [I(X_i;Y_i|T_i,U_i) + I(T_i,U_i;Z_i)] + n(\epsilon_{1n}+\epsilon_{2n}) \end{aligned} \tag{136}$$

where (a) holds since U_i is a deterministic function of S^k and hence of T_i, (b) follows from the data processing inequality, and (c) holds since Y_i is only determined by U_i and X_i due to the memoryless property of the channel. On the other hand, $kH(S,L)$ can also be bounded by

$$\begin{aligned} kH(S,L) &= H(S^k,L^k) \\ &= I(S^k,L^k;Y^n) + H(S^k,L^k|Y^n) \\ &\le I(X^n,U^n;Y^n) + n\epsilon_{1n} \\ &= \sum_{i=1}^{n} I(U_i,X_i;Y_i) + n\epsilon_{1n}. \end{aligned} \tag{137}$$

Likewise, it follows from (135) that

$$\begin{aligned} kH(S) &= H(S^k) \\ &= I(S^k;Z^n) + H(S^k|Z^n) \\ &= \sum_{i=1}^{n} I(S^k;Z_i\,|\,\boldsymbol{Z}^{i+1}) + H(S^k\,|\,Z^n) \\ &\le \sum_{i=1}^{n} I(S^k,\boldsymbol{Z}^{i+1};Z_i) + n\epsilon_{2n} \\ &\le \sum_{i=1}^{n} I(S^k,Y^{i-1},\boldsymbol{Z}^{i+1},U_i;Z_i) + n\epsilon_{2n} \\ &= \sum_{i=1}^{n} I(T_i,U_i;Z_i) + n\epsilon_{2n}. \end{aligned} \tag{138}$$

Note also that $T_i \longrightarrow (U_i,X_i) \longrightarrow (Y_i,Z_i)$ for all $1 \le i \le n$. According to (136), (137), and (138), and recalling that $k = \tau n$, it is easy to show (e.g., see [11]) that there exists an auxiliary RV T with $P_{TUXYZ} = P_T P_{UX\,|\,T} W_{YZ\,|\,UX}$ such that

$$\begin{aligned}\tau H(S,L) &\le \min\{I_{P_{UXYZ}}(U,X;Y),\\ &\qquad I_{P_{TUXYZ}}(X;Y\,|\,T,U) + I_{P_{TUXYZ}}(T,U;Z)\}\\ \tau H(S) &\le I_{P_{TUXYZ}}(T,U;Z).\end{aligned}$$

It remains to show that the alphabet of the RV T can be limited by $|\mathcal{T}| \le |\mathcal{U}||\mathcal{X}|+1$; i.e., we will show by applying the support lemma below, which is based on the Carathéodory theorem (cf. [10, p. 311]) that there exists an RV $\hat{T}$ with $|\hat{\mathcal{T}}| \le |\mathcal{U}||\mathcal{X}|+1$ such that $P_{\hat{T}UXYZ} = P_{\hat{T}} P_{UX|\hat{T}} W_{YZ\,|\,UX}$ and

$$\begin{aligned}&(I_{P_{UXYZ}}(U,X;Y), I_{P_{TUXYZ}}(T,U;Z), I_{P_{TUXYZ}}(X;Y\,|\,T,U))\\ &\quad = (I_{P_{UXYZ}}(U,X;Y), I_{P_{\hat{T}UXYZ}}(\hat{T},U;Z),\\ &\qquad\qquad I_{P_{\hat{T}UXYZ}}(X;Y|\hat{T},U)). \quad (139)\end{aligned}$$

Lemma 8: ([10, Support lemma, p. 311]) Let f_j, $j = 1, 2, \ldots, k$ be real-valued continuous functions on $\mathcal{P}(\mathcal{X})$. For any probability measure μ on the Borel σ-algebra of $\mathcal{P}(\mathcal{X})$, there exist k elements $P_1, P_2, \ldots, P_k$ of $\mathcal{P}(\mathcal{X})$ and k nonnegative reals $\alpha_1, \alpha_2, \ldots \alpha_k$ with $\sum_{i=1}^k \alpha_i = 1$ such that for every $j = 1, 2, \ldots, k$

$$\int_{\mathcal{P}(\mathcal{X})} f_j(P)\mu(dP) = \sum_{i=1}^k \alpha_i f_j(P_i).$$

We first rewrite

$$I_{P_{TUXYZ}}(T,U;Z) = H(Z) - H(Z|T,U)$$

and

$$\begin{aligned}I_{P_{TUXYZ}}(X;Y\,|\,T,U)) &= H(Y\,|\,T,U) - H(Y|X,T,U)\\ &= H(Y\,|\,T,U) - H(Y|X,U)\end{aligned}$$

where the last equality follows since $T \to (U,X) \to Y$ forms a Markov chain. To apply the support lemma, we define the following real-valued continuous functions of distribution $P_{UX\,|\,T}(\cdot,\cdot\,|\,t)$ on $\mathcal{P}(\mathcal{U}\times\mathcal{X})$ for fixed $t \in \mathcal{T}$:

$$f_i(P_{UX\,|\,T}(u,x|t)) \triangleq P_{UX\,|\,T}(u,x|t)$$

for all $(u,x) \in \mathcal{U}\times\mathcal{X}$ except one pair (u,x), so there are $m-1 = |\mathcal{U}||\mathcal{X}|-1$ functions; i.e., i ranges from 1 to m. Furthermore, we define real-valued continuous functions

$$f_m(P_{UX\,|\,T}(u,x|t)) \triangleq H(Z|T=t,U)$$

and

$$f_{m+1}(P_{UX\,|\,T}(u,x|t)) \triangleq H(Y\,|\,T=t,U).$$

According to the support lemma, there must exist a new RV $\hat{T}$ (jointly distributed with (U,X)) with alphabet size $|\hat{\mathcal{T}}| = m+1 = |\mathcal{U}||\mathcal{X}|+1$ such that the expectation of f_i with respect to P_T, $i = 1, 2, \ldots, m+1$, can be expressed in terms of the convex combination of $m+1$ points, i.e.,

$$\begin{aligned}P_{UX}(u,x) &= \sum_{\mathcal{T}} P_T(t) f_i(P_{UX\,|\,T}(\cdot,\cdot\,|\,t))\\ &= \sum_{\hat{\mathcal{T}}} P_{\hat{T}}(\hat{t}) f_i(P_{UX\,|\,t}(\cdot,\cdot\,|\,\hat{t})),\\ &\qquad\qquad i = 1, 2, \ldots, m-1 \quad (140)\\ H(Z\,|\,T,U) &= \sum_{\mathcal{T}} P_T(t) f_m(P_{UX\,|\,T}(\cdot,\cdot\,|\,t))\\ &= \sum_{\hat{\mathcal{T}}} P_{\hat{T}}(\hat{t}) f_m(P_{UX\,|\,\hat{T}}(\cdot,\cdot\,|\,\hat{t}))\\ &= H(Z\,|\,\hat{T},U) \quad (141)\end{aligned}$$

and

$$\begin{aligned}H(Y\,|\,T,U) &= \sum_{\mathcal{T}} P_T(t) f_{m+1}(P_{UX\,|\,T}(\cdot,\cdot\,|\,t))\\ &= \sum_{\hat{\mathcal{T}}} P_{\hat{T}}(\hat{t}) f_{m+1}(P_{UX\,|\,\hat{T}}(\cdot,\cdot\,|\,\hat{t}))\\ &= H(Y\,|\,\hat{T},U). \quad (142)\end{aligned}$$

Clearly, $\hat{T} \to (U,X) \to (Y,Z)$ forms a Markov chain and (139) holds. The proof for the converse part is complete. □

REFERENCES

[1] R. Ahlswede and T. S. Han, "On source coding with side information via a multiple-access channel and related problems in multi-user information theory," *IEEE Trans. Inf. Theory*, vol. IT-29, no. 3, May 1983.

[2] E. A. Arutyunyan (Haroutunian), "Lower bound for the error probability of multiple-access channels," *Probl. Pered. Inform.*, vol. 11, pp. 23–36, Apr.–June 1975.

[3] P. P. Bergmans, "Random coding theorem for broadcast channels with degraded components," *IEEE Trans. Inf. Theory*, vol. IT-19, no. 2, pp. 197–206, Mar. 1973.

[4] K. De Bruyn, V. V. Prelov, and E. Van Der Meulen, "Reliable transmission of two correlated sources over an asymmetric multiple-acess channel," *IEEE Trans. Inf. Theory*, vol. IT-33, no. 5, pp. 716–718, Sep. 1987.

[5] S. Choi and S. S. Pradhan, "A graph-based framework for transmission of correlated sources over broadcast channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 7, pp. 2841–2856, Jul. 2008.

[6] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New York: Wiley, 2006.

[7] T. M. Cover, A. El Gamal, and M. Salehi, "Multiple access channels with arbitrarily correlated sources," *IEEE Trans. Inf. Theory*, vol. IT–26, no. 6, pp. 648–657, Nov. 1980.

[8] I. Csiszár, "Joint source-channel error exponent," *Probl. Contr. Inf. Theory*, vol. 9, pp. 315–328, 1980.

[9] I. Csiszár, "The method of types," *IEEE Trans. Inf. Theory*, vol. 44, no. 7, pp. 2505–2522, Nov. 1998.

[10] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. New York: Academic, 1981.

[11] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 3, pp. 339–348, May 1978.

[12] N. Devroye, P. Mitran, and V. Tarokh, "Achievable rates in cognitive radio channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 1813–1827, May 2006.

[13] N. Devroye, P. Mitran, and V. Tarokh, "Limits on communications in a cognitive radio channel," *IEEE Commun. Mag.*, vol. 44, no. 6, pp. 44–49, Jun. 2006.

[14] G. Dueck, "A note on the multiple access channel with correlated sources," *IEEE Trans. Inf. Theory*, vol. IT-27, no. 2, pp. 232–235, Mar. 1981.

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

[16] A. A. El Gamal, "The capacity of a class of broadcast channels," *IEEE Trans. Inf. Theory*, vol. IT-25, no. 2, pp. 166–169, Mar. 1979.
[17] T. S. Han and M. H. M. Costa, "Broadcast channels with arbitrary correlated sources," *IEEE Trans. Inf. Theory*, vol. IT-33, no. 5, pp. 641–650, Sep. 1987.
[18] B. Hochwald and K. Zeger, "Tradeoff between source and channel coding," *IEEE Trans. Inf. Theory*, vol. 43, no. 5, pp. 1412–1424, Sep. 1997.
[19] K. Iwata and Y. Oohama, "Information-spectrum characterization of multiple-access channels with correlated sources," *IEICE Trans. Fundamentals*, vol. E88-A, no. 11, pp. 3196–3202, Nov. 2005.
[20] W. Kang and S. Ulukus, "An outer bound for the multi-terminal rate-distortion region," in *Proc. 2006 IEEE Int. Symp. Information Theory (ISIT)*, Seattle, WA, Jul. 2006, pp. 1419–1423.
[21] J. Körner and K. Marton, "General broadcast channels with degraded message sets," *IEEE Trans. Inf. Theory*, vol. IT–23, no. 1, pp. 60–64, Jan. 1977.
[22] J. Körner and A. Sgarro, "Universally attainable error exponents for broadcast channels with degraded message sets," *IEEE Trans. Inf. Theory*, vol. IT-26, no. 6, pp. 670–679, Nov. 1980.
[23] Y. Liang, A. Somekh-Baruch, H. V. Poor, S. Shamai (Shitz), and S. Verdú, "Cognitive interference channels with confidential messages," in *Proc. 45th Annu. Allerton Conf. Communication, Control, and Computing*, Monticello, IL, Sep. 2007, pp. 1–6.
[24] Y. S. Liu and B. L. Hughes, "A new universal random coding bound for the multiple-access channel," *IEEE Trans. Inf. Theory*, vol. 42, no. 2, pp. 376–386, Mar. 1996.
[25] J. Mitola, III, "The software radio architecture," *IEEE Commun. Mag.*, vol. 33, no. 5, pp. 26–38, May 1995.
[26] J. Mitola, III, "Cognitive Radio: An Integrated Agent Architecture for Software Defined Radio," Ph.D. dissertation, Royal Institute of Technology (KTH), Stockholm, Sweden, May 2000.
[27] J. Mitola, III*, Cognitive Radio Architecture: The Engineering Foundations of Radio XML*. New York: Wiley, 2006.
[28] S. S. Pradhan, S. Choi, and K. Ramchandran, "A graph-based framework for transmission of correlated sources over multiple access channels," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4583–4604, Dec. 2007.
[29] Y. Zhong, F. Alajaji, and L. L. Campbell, "On the joint source-channel coding error exponent for discrete memoryless systems," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1450–1468, Apr. 2006.
[30] Y. Zhong, F. Alajaji, and L. L. Campbell, "On the joint source-channel coding error exponent of discrete communication systems with Markovian memory," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4457–4472, Dec. 2007.

Yangfan Zhong (S'04) was born in Chengdu, China, on May 19, 1980. He received the B.E. degree in communication and information engineering from the University of Electronic Science and Technology of China, Chengdu, in 2002 and the M.Sc.E. and Ph.D. degrees in mathematics and engineering from Queen's University, Kingston, ON, Canada, in 2003 and 2008, respectively.

He is currently a Senior Business Analyst with the Bank of Montreal Financial Group, Toronto, ON, Canada. His research interests lie in the general area of single- and multiuser information theory, in particular, joint source–channel coding and error control coding.

Fady Alajaji (S'90–M'94–SM'00) was born in Beirut, Lebanon, on May 1, 1966. He received the B.E. degree with distinction from the American University of Beirut, and the M.Sc. and Ph.D. degrees from the University of Maryland, College Park, all in electrical engineering, in 1988, 1990, and 1994, respectively.

He held a postdoctoral appointment in 1994 at the Institute for Systems Research, University of Maryland. In 1995, he joined the Department of Mathematics and Statistics at Queen's University, Kingston, ON, Canada, where he is currently a Professor of Mathematics and Engineering. Since 1997, he has also been cross-appointed in the Department of Electrical and Computer Engineering at the same university. His research interests include information theory, digital communications, error control coding, joint source–channel coding, and data compression.

Dr. Alajaji currently serves as Area Editor and Editor for Source and Source-Channel Coding for the IEEE TRANSACTIONS ON COMMUNICATIONS. He served as Co-Chair of the 1999 Canadian Workshop on Information Theory, as Co-Chair of the Technical Program Committee (TPC) of the 2004 Biennial Symposium on Communications, and as a TPC member of several international conferences and workshops. He received the Premier's Research Excellence Award from the Province of Ontario.

L. Lorne Campbell (M'56–SM'81–F'90–LF'94) was born in Winnipeg, MB, Canada in 1928. He received the B.Sc. degree in mathematics and physics in 1950 from the University of Manitoba, Winnipeg, the M.S. degree in physics in 1951 from Iowa State University, Ames, and the Ph.D. degree in mathematics in 1955 from the University of Toronto, Toronto, ON, Canada.

From 1954 to 1958, he was at the Defence Research Telecommunications Establishment in Ottawa, ON, Canada. He taught mathematics at the University of Windsor, Windsor, ON, from 1958 to 1963 and at Queen's University, Kingston, ON, from 1963 to 1996. He is now Professor Emeritus at Queen's University. He was Head of the Department of Mathematics and Statistics from 1980 to 1990 at Queen's University.

Dr. Campbell was Chairman of the Program Committee for ISIT'83. He is a member of the Information Theory and Communications Societies and of several mathematical and statistical societies.