Mathematics and Engineering Communications Laboratory

Technical Report

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

Y. Zhong, *F. Alajaji*, and *L. L. Campbell*

March 2007

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

Yangfan Zhong Fady Alajaji L. Lorne Campbell

Abstract

Consider transmitting two discrete memoryless correlated sources, consisting of a common and a private source, over a discrete memoryless multi-terminal 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 2-user source-channel system. In this work, we derive a universally achievable joint source-channel coding (JSCC) error exponent pair for the 2-user system by using a technique which generalizes Csiszár's method [8] 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 the JSCC theorem with single letter characterization.

Index Terms: discrete memoryless correlated sources, broadcast channel, multiple access channel, common and private message, joint source-channel coding, error exponent, type packing lemma.

*Mathematics and Engineering Technical Report, March 2007. This research was supported by the Natural Sciences and Engineering Research Council of Canada. The authors are with the Dept. of Mathematics & Statistics, Queen's University, Kingston, ON K7L 3N6, Canada.

1 Introduction

Recently, the study of the error exponent (reliability function) for point-to-point (single-user) source-channel 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], [23], [24]). It is of natural interest to study the JSCC error exponent for multi-terminal source-channel systems.

In this work we address the asymmetric 2-user source-channel system depicted in Fig. 1. Two discrete memoryless correlated source messages $(\mathbf{s}, \mathbf{l}) \in \mathcal{S}^{\tau n} \times \mathcal{L}^{\tau n}$ drawn from a joint distribution $Q_{SL} : \mathcal{S} \times \mathcal{L}$, consisting of a common source messages $\mathbf{s}$ and a private source message $\mathbf{l}$ 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 set-up, the goal is to send the common information to both receivers, and send the private information to only one receiver (say, Decoder 1).

It is worthy to point out that the asymmetric 2-user system can be specialized to the following two classical asymmetric multi-terminal scenarios.

(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 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 [19], and the JSCC theorem for CS-ABC system with arbitrary transmission rate can also be analogously carried out (e.g., [16]). In this work, we study a refined version of the JSCC theorem for the general asymmetric 2-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 (1-dimension) type setting to a joint (2-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 2-user channel $W_{YZ|TUX}$ for a given triple of n-length sequences $(\mathbf{t}, \mathbf{u}, \mathbf{x})$; see Theorem 1. Here, the augmented channel $W_{YZ|TUX}$ is induced from the original 2-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 $(\mathbf{s}, \mathbf{l})$ to a codeword triplet $(\mathbf{t}, \mathbf{u}, \mathbf{x})$, where $\mathbf{t}$ is the the auxiliary superposition codeword. For the asymmetric 2-user system, since one of the encoders has full access to both sources, it knows the output of the other. 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 independent of the statistics of the source and the channel); this generalizes Körner and Sgarro's exponent pair for ABC coding (with uniform message sets) [20]. 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). 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.

At this point we pause to mention some related works in the literature on the multi-terminal JSCC of CS. The JSCC theorem for transmitting two CS over a (symmetric) multiple access channel (each encoder can only access one source) has been studied in [1, 7, 13, 17, 18, 22], and the JSCC theorem for transmitting two CS over a (symmetric) broadcast channel (each decoder needs to reconstruct one source) has been addressed in [5, 16]. 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, the 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 2, we introduce the notation and some basic facts regarding the method of types. A generalized joint type packing lemma is presented in Section 3. In Section 4 we establish a universally achievable error exponent pair for the 2-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 given. In Section 5, we apply our results for the CS-AMAC and CS-ABC systems. Finally, we partially address the computation for the bounds for the system JSCC error exponent in Section 6.

2 Preliminaries

The following notations and conventions are adopted from [8, 11]. 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 $\mathbf{x} \triangleq (x_1, x_2, \cdots, x_n) \in \mathcal{X}^n$ is the empirical probability distribution $P_{\mathbf{x}} \in \mathcal{P}(\mathcal{X})$ defined by

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

where $N(a|\mathbf{x})$ is the number of occurrences of a in $\mathbf{x}$. 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 $\mathbf{x} \in \mathcal{X}^n$ 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 $\mathbf{x} \in \mathcal{X}^n$ and $\mathbf{y} \triangleq (y_1, y_2, \cdots, y_n) \in \mathcal{Y}^n$ is the empirical joint probability distribution $P_{\mathbf{xy}} \in \mathcal{P}(\mathcal{X} \times \mathcal{Y})$ defined by

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

Let $\mathcal{P}_n(\mathcal{X} \times \mathcal{Y}) \subseteq \mathcal{P}_n(\mathcal{X} \times \mathcal{Y})$ be the collection of all joint types of sequences in $\mathcal{X}^n \times \mathcal{Y}^n$. The set of all $\mathbf{x} \in \mathcal{X}^n$ and $\mathbf{y} \in \mathcal{Y}^n$ 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 $\mathbf{y} \in \mathcal{Y}^n$ given $\mathbf{x} \in \mathbb{T}_{P_X}$ is the empirical conditional probability distribution $P_{\mathbf{y}|\mathbf{x}} \in \mathcal{P}(\mathcal{Y}|\mathcal{X})$ defined by

$$P_{\mathbf{y}|\mathbf{x}}(b|a) = \frac{N(a, b|\mathbf{x}, \mathbf{y})}{N(a|\mathbf{x})}, \quad (a, b) \in \mathcal{X} \times \mathcal{Y}.$$

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

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}. Conversely, $P_X P_{YZ|X}$ denotes a joint distribution on $\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ with marginal distribution P_X and conditional distribution $P_{YZ|X}$. Note that for a given joint type $P_{XY} \in \mathcal{P}_n(\mathcal{X} \times \mathcal{Y})$, $\mathbb{T}_{P_{Y|X}}(\mathbf{x}) = \{\mathbf{y} : (\mathbf{x}, \mathbf{y}) \in \mathbb{T}_{P_{XY}}\}$. Note also that

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

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 different distributions (or types) defined on the same alphabet, we use sub-subscript, 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 $\mathbf{x} \in \mathcal{X}^n$, $\mathbf{y} \in \mathcal{Y}^n$ and $\mathbf{z} \in \mathcal{Z}^n$, since the types $P_{\mathbf{x}}$, $P_{\mathbf{xy}}$ and $P_{\mathbf{xyz}}$ can also be represented as distributions of dummy RV's, we define the empirical entropy and mutual information by $H(\mathbf{x}) \triangleq H_{P_{\mathbf{x}}}(X)$, $I(\mathbf{x};\mathbf{y}) \triangleq I_{P_{\mathbf{xy}}}(X;Y)$ and $I(\mathbf{x};\mathbf{y}|\mathbf{z}) \triangleq I_{P_{\mathbf{xyz}}}(X;Y|Z)$. 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-dimension product distributions. All logarithms and exponentials throughout this paper are in base 2. The following facts will be widely used throughout this paper.

Lemma 1 [11]

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

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

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

$$W_{Y|X}^{(n)}(\mathbf{y}|\mathbf{x}) = 2^{-n[D(V_{Y|X} \| W_{Y|X}|P_X) + H_{P_X V_{Y|X}}(Y|X)]},$$

and hence

$$(n+1)^{-|\mathcal{X}||\mathcal{Y}|} 2^{-nD(V_{Y|X} \| W_{Y|X}|P_X)} \leq W_{Y|X}^{(n)}(\mathbb{T}_{V_{Y|X}}(\mathbf{x})|\mathbf{x}) \leq 2^{-nD(V_{Y|X} \| W_{Y|X}|P_X)}.$$

3 A Joint Type Packing Lemma

We extend Csiszár's type packing lemma [8, Theorem. 5] from a (1-dimensional) single-letter type setting to a (2-dimensional) joint type setting. This lemma plays a key role in deriving an exponentially achievable upper bound for the probability of erroneous transmission for the asymmetric 2-user channel.

Lemma 2 (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 $\{m'_{in}\}$ associated with every $i = 1, 2, ..., m_n$, 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}, $i = 1, 2, ..., m_n$ and $j = j(i) = 1, 2, ..., m'_{in}$ with

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

and

$$\frac{1}{n} \log_2 M_{ij} < H_{P_{A_i} P_{B_j|A_i}}(B|A) - \delta, \tag{2}$$

where

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

there exist m_n disjoint subsets

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

such that

$$|\mathbb{T}_{V_{A'|A}}(\mathbf{a}_p^{(i)}) \bigcap \Omega_k| \leq N_k 2^{-n\left[I_{P_{A_i} V_{A'|A}}(A;A') - \delta\right]}, \tag{3}$$

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 $\mathbf{u}_p^{(i)} \in \Omega_i$ and every i, there exist m'_{in} disjoint subsets

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

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

$$\left|\mathbb{T}_{V_{A'B'|AB}}(\mathbf{a}_p^{(i)}, \mathbf{b}_{p,q}^{(j)}) \bigcap \bigcup_{p'=1}^{N_k} \Omega_{kl}(\mathbf{a}_{p'}^{(k)})\right| \leq N_k M_{kl} 2^{-n\left[I_{P_{A_i B_j} V_{A'B'|AB}}(A,B;A',B') - \delta\right]}, \tag{4}$$

$$\left|\mathbb{T}_{V_{A'B'|AB}}(\mathbf{a}_p^{(i)}, \mathbf{b}_{p,q}^{(j)}) \bigcap \bigcup_{p'=1}^{N_i} \Omega_{il}(\mathbf{a}_{p'}^{(i)})\right| \leq M_{il} 2^{-n\left[I_{P_{A_i B_j} V_{A'B'|AB}}(B;B'|A) - \delta\right]}, \tag{5}$$

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. We remark that the assertion of (3) is Csiszár's type packing lemma [8, Theorem 5] for a single-letter type setting. Roughly and intuitively, if $(\mathbf{a}, \mathbf{b})$ is a pair of transmitted codewords, then the possible sequences decoded as $(\mathbf{a}, \mathbf{b})$ can be seen as elements in the "sphere" $\mathbb{T}_{V_{A'B'|AB}}(\mathbf{a}, \mathbf{b})$ "centered" at $(\mathbf{a}, \mathbf{b})$ for some $V_{A'B'|AB}$. Equation (4) in the packing lemma (similar to (3) and (5)) states that there exist disjoint sets $\Omega_{kl} = \bigcup_{p'=1}^{N_k} \Omega_{kl}(\mathbf{a}_{p'}^{(k)})$

with bounded cardinalities such that the size of intersection between the sphere $\mathbb{T}_{V_{A'B'|AB}}(\mathbf{a},\mathbf{b})$ for every $(\mathbf{a},\mathbf{b}) \in \Omega_{ij}$ and every set Ω_{kl} is "exponentially small" compared with the size of each Ω_{kl}. So the packing lemma can be used to prove the existence of good codes that have an exponentially small probability of error.

Note also that the above extended packing lemma is analogous to, but different from the one introduced by Körner and Sgarro in [20], which is used to prove a lower bound for the channel coding ABC exponent. Lemma 2 here is used for the JSCC problem.

4 Transmitting CS over the Asymmetric 2-User Channel

4.1 System

Let $\{W_{YZ|UX} : \mathcal{U} \times \mathcal{X} \to \mathcal{Y} \times \mathcal{Z}\}$ be a 2-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

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

where $u \in \mathcal{U}$, $x \in \mathcal{X}$, $y \in \mathcal{Y}$, $z \in \mathcal{Z}$, $\mathbf{u} \triangleq (u_1,...,u_n) \in \mathcal{U}^n$, $\mathbf{x} = (x_1,...,,x_n) \in \mathcal{X}^n$, $\mathbf{y} = (y_1,...,y_n) \in \mathcal{Y}^n$, and $\mathbf{z} \triangleq (z_1,...,z_n) \in \mathcal{Z}^n$. 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 $W^{(n)}_{YZ|UX}$ are denoted by $W^{(n)}_{Y|UX}$ and $W^{(n)}_{Z|UX}$, respectively.

Consider two discrete memoryless CS with a generic joint distribution $Q_{SL}(s,l)$ defined on the finite alphabet $\mathcal{S} \times \mathcal{L}$ such that the k-tuple joint distribution is $Q^{(k)}_{SL}(\mathbf{s},\mathbf{l}) = \prod_{i=1}^{k} Q_{SL}(s_i,l_i)$, where $(s,l) \in \mathcal{S} \times \mathcal{L}$, and $(\mathbf{s},\mathbf{l}) \triangleq ((s_1,l_1),...,(s_k,l_k)) \in \mathcal{S}^k \times \mathcal{L}^k$. For each pair of source messages $(\mathbf{s},\mathbf{l})$ drawn from the above joint distribution, we need to transmit the *common message* $\mathbf{s}$ over the channel $W_{YZ|UX}$ to Receivers Y and Z and transmit the *private message* $\mathbf{l}$ only to Receiver Y. A joint source-channel (JSC) code with block length n and transmission rate τ (source symbol/channel use) for transmitting Q_{SL} through $W_{YZ|UX}$ 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

$$P^{(n)}_{Ye}(Q_{SL}, W_{YZ|UX}, \tau) \triangleq \Pr(\{\varphi_n(\mathbf{y}) \neq (\mathbf{s},\mathbf{l})\}) = \sum_{\mathbf{s},\mathbf{l}} Q^{(\tau n)}_{SL}(\mathbf{s},\mathbf{l}) \sum_{\mathbf{y}:\varphi_n(\mathbf{y}) \neq (\mathbf{s},\mathbf{l})} W^{(n)}_{Y|UX}(\mathbf{y}|\mathbf{u},\mathbf{x}) \tag{6}$$

and

$$P^{(n)}_{Ze}(Q_{SL}, W_{YZ|UX}, \tau) \triangleq \Pr(\{\psi_n(\mathbf{z}) \neq \mathbf{s}\}) = \sum_{\mathbf{s},\mathbf{l}} Q^{(\tau n)}_{SL}(\mathbf{s},\mathbf{l}) \sum_{\mathbf{z}:\psi_n(\mathbf{z}) \neq \mathbf{s}} W^{(n)}_{Z|UX}(\mathbf{z}|\mathbf{u},\mathbf{x}) \tag{7}$$

where $\mathbf{x} \triangleq f_n(\mathbf{s}, \mathbf{l})$ and $\mathbf{u} \triangleq g_n(\mathbf{s})$ are the corresponding codewords of the source message pair $(\mathbf{s}, \mathbf{l})$ and the source message $\mathbf{s}$, and $\mathbf{y}$ and $\mathbf{z}$ 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

$$P_{ie}^{(n)}(Q_{SL}, W_{YZ|UX}, \tau) = \sum_{P_{SL} \in \mathcal{P}_{\tau n}(\mathcal{S} \times \mathcal{L})} Q_{SL}^{(\tau n)}(\mathbb{T}_{SL}) P_{ie}(\mathbb{T}_{SL}) \quad i = Y, Z, \tag{8}$$

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

$$P_{Ye}(\mathbb{T}_{SL}) = \frac{1}{|\mathbb{T}_{SL}|} \sum_{(\mathbf{s},\mathbf{l}) \in \mathbb{T}_{SL}} \sum_{\mathbf{y}: \varphi_n(\mathbf{y}) \neq (\mathbf{s},\mathbf{l})} W_{Y|UX}^{(n)}(\mathbf{y}|\mathbf{u}, \mathbf{x}) \tag{9}$$

and

$$P_{Ze}(\mathbb{T}_{SL}) = \frac{1}{|\mathbb{T}_{SL}|} \sum_{(\mathbf{s},\mathbf{l}) \in \mathbb{T}_{SL}} \sum_{\mathbf{z}: \psi_n(\mathbf{z}) \neq \mathbf{s}} W_{Z|UX}^{(n)}(\mathbf{z}|\mathbf{u}, \mathbf{x}). \tag{10}$$

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

$$P_{ie}^{(n)}(Q_{SL}, W_{YZ|UX}, \tau) \leq 2^{-n[E_{Ai} - \delta]}, \quad i = Y, Z \tag{11}$$

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

$$P_{e}^{(n)}(Q_{SL}, W_{YZ|UX}, \tau) \triangleq \Pr\left(\{\varphi_n(\mathbf{y}) \neq (\mathbf{s}, \mathbf{l})\} \cup \{\psi_n(\mathbf{z}) \neq \mathbf{s}\}\right),$$

where $(\mathbf{s}, \mathbf{l})$ are drawn according to $Q_{SL}^{(\tau n)}$.

Definition 1 Given Q_{SL}, $W_{YZ|UX}$ and $\tau > 0$, the system JSCC error exponent $E_J(Q_{SL}, W_{YZ|UX}, \tau)$ is defined as 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 blocklength n and transmission rate τ such that

$$E \leq \liminf_{n \to \infty} -\frac{1}{n} \log_2 P_e^{(n)}(Q_{SL}, W_{YZ|UX}, \tau). \tag{12}$$

Since the system probability of error must be larger than $P_{Ye}^{(n)}$ and $P_{Ze}^{(n)}$ defined by (6) and (7), and is also upper bounded by the sum of the two, it follows that for any sequence of JSC codes $(f_n, g_n, \varphi_n, \psi_n)$

$$\liminf_{n \to \infty} -\frac{1}{n} \log_2 P_e^{(n)}(Q_{SL}, W_{YZ|UX}, \tau) = \liminf_{n \to \infty} -\frac{1}{n} \log_2 \max\left(P_{Ye}^{(n)}, P_{Ze}^{(n)}\right). \tag{13}$$

4.2 Superposition Encoding for Asymmetric 2-User Channels

Given an asymmetric 2-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 $\mathbf{s}$ to a pair of n-length sequences $(\mathbf{t},\mathbf{u}) \in \mathcal{T}^n \times \mathcal{U}^n$ with a fixed type, say P_{TU}, and then sends the codeword $\mathbf{u}$ over the channel, i.e., $g_n(\mathbf{s}) = \mathbf{u}$. The encoder f_n first maps each pair $(\mathbf{s},\mathbf{l})$ to a triple of sequences $(\mathbf{t},\mathbf{u},\mathbf{x}) \in \mathcal{T}^n \times \mathcal{U}^n \times \mathcal{X}^n$ such that $\mathbf{x} \in \mathbb{T}_{P_{X|TU}}(\mathbf{t},\mathbf{u})$, then f_n sends the codeword $\mathbf{x}$ over the channel, i.e., $f_n(\mathbf{s},\mathbf{l}) = \mathbf{x}$. In other words, g_n and f_n map $(\mathbf{s},\mathbf{l})$ to a tuple of sequences $(\mathbf{t},\mathbf{u},\mathbf{x})$ with a joint type $P_{TU}P_{X|TU}$, although only $\mathbf{u}$ and $\mathbf{x}$ are sent to the channel, where $\mathbf{t}$ plays the role of a dummy codeword.

Since $W^{(n)}_{YZ|TUX}(\mathbf{y},\mathbf{z}|\mathbf{t},\mathbf{u},\mathbf{x})$ is equal to $W^{(n)}_{YZ|UX}(\mathbf{y},\mathbf{z}|\mathbf{u},\mathbf{x})$ and is independent of $\mathbf{t}$, transmitting the codewords $(\mathbf{u},\mathbf{x})$ through the channel $W_{YZ|UX}$ can be viewed as transmitting the codewords $(\mathbf{t},\mathbf{u},\mathbf{x})$ over the augmented channel $W_{YZ|TUX}$. Here, the common outputs of g_n and f_n, $(\mathbf{t},\mathbf{u})$'s, are called auxiliary *cloud centers* according to the traditional superposition coding notion [3], which convey the information of the common message $\mathbf{s}$, and the codewords $\mathbf{x}$'s corresponding to the same $(\mathbf{t},\mathbf{u})$ are called satellite codewords of $(\mathbf{t},\mathbf{u})$, which contain both the common and private information. At the decoding stage, Receiver Z only needs to figure out which cloud $(\mathbf{t},\mathbf{u})$ was transmitted, and Receiver Y needs to estimate not only the cloud but also the satellite codeword $\mathbf{x}$. We employ superposition encoding to derive the achievable error exponent pair and the lower bound of system JSCC error exponent in Section 4.3.

4.3 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

$$E_Y(R_1,R_2,W_{Y|TUX},P_{TUX}) \triangleq \min_{V_{Y|TUX}} \Big[D(V_{Y|TUX} \parallel W_{Y|TUX}|P_{TUX}) + \min\Big(\Big| I_{P_{TUX}V_{Y|TUX}}(T,U,X;Y) - (R_1+R_2) \Big|^+ , \Big| I_{P_{TUX}V_{Y|TUX}}(X;Y|T,U) - R_2 \Big|^+ \Big) \Big], \tag{14}$$

and

$$E_Z(R_1,R_2,W_{Z|TUX},P_{TUX}) \triangleq \min_{V_{Z|TUX}} \Big[D(V_{Z|TUX} \parallel W_{Z|TUX}|P_{TUX}) + \Big| I_{P_{TUX}V_{Z|TUX}}(T,U;Z) - R_1 \Big|^+ \Big], \tag{15}$$

where $|x|^+ = \max(0, x)$, and the outer minimum in (14) (respectively (15)) 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

$$\begin{aligned} R_1 + R_2 &\geq I_{P_{TUX}W_{Y|TUX}}(T, U, X; Y) = I(U, X; Y), && (16)\\ R_2 &\geq I_{P_{TUX}W_{Y|TUX}}(X; Y|T, U), && (17)\end{aligned}$$

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

$$R_1 \geq I_{P_{TUX}W_{Z|TUX}}(T, U; Z). \tag{18}$$

Using Lemma 2 and employing generalized maximum mutual information 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 $\{m'_{in}\}$ associated with every $i = 1, 2, ..., m_n$ with

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

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}, $i = 1, 2, ..., m_n$ and $j = j(i) = 1, 2, ..., m'_{in}$ 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 $\Omega_i = \left\{(\mathbf{t}, \mathbf{u})_p^{(i)}\right\}_{p=1}^{N_i} \subseteq \mathbb{T}_{(TU)_i}$, m'_{in} disjoint subsets

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

with $\mathbf{x}_{p,q}^{(j)} \in \mathbb{T}_{X_j|(TU)_i}((\mathbf{t}, \mathbf{u})_p^{(i)})$ for every $(\mathbf{t}, \mathbf{u})_p^{(i)} \in \Omega_i$ 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 $\Omega_{ij} = \bigcup_{p=1}^{N_i} \Omega_{ij}((\mathbf{t}, \mathbf{u})_p^{(i)})$, such that the probabilities of erroneous transmission of a triplet $(\mathbf{t}, \mathbf{u}, \mathbf{x}) \in \Omega$ over the augmented channel $W_{YZ|TUX}$ using decoders $(\varphi_n^{(0)}, \psi_n^{(0)})$ are *simultaneously* bounded by

$$\begin{aligned} P_{Ye}^{(n)}(\mathbf{t}, \mathbf{u}, \mathbf{x}) &\triangleq \sum_{\mathbf{y}:\varphi_n^{(0)}(\mathbf{y}) \neq ((\mathbf{t},\mathbf{u}),\mathbf{x})} W_{Y|TUX}^{(n)}(\mathbf{y}|\mathbf{t}, \mathbf{u}, \mathbf{x})\\ &\leq 2^{-n\left[E_Y\left(R_i, R_{ij}, W_{Y|TUX}, P_{(TU)_i}P_{X_j|(TU)_i}\right) - \delta\right]} \end{aligned} \tag{19}$$

and

$$\begin{aligned} P_{Ze}^{(n)}(\mathbf{t}, \mathbf{u}, \mathbf{x}) &\triangleq \sum_{\mathbf{z}:\psi_n^{(0)}(\mathbf{z}) = ((\mathbf{t},\mathbf{u})', \mathbf{x}') \text{ such that } (\mathbf{t},\mathbf{u})' \neq (\mathbf{t},\mathbf{u})} W_{Z|TUX}^{(n)}(\mathbf{z}|\mathbf{t}, \mathbf{u}, \mathbf{x})\\ &\leq 2^{-n\left[E_Z\left(R_i, R_{ij}, W_{Z|TUX}, P_{(TU)_i}P_{X_j|(TU)_i}\right) - \delta\right]} \end{aligned} \tag{20}$$

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

Proof: We apply the packing lemma (Lemma 2) 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 2. For the $\{m_n\}$, $\{m'_{in}\}$, $P_{(TU)_i}$, $P_{X_j|(TU)_i}$ given in Theorem 1, according to Lemma 2, there exist pairwise disjoint subsets Ω_i and $\Omega_{ij}((\mathbf{t},\mathbf{u})_p^{(i)})$ satisfying (3), (4), and (5) for every $1 \le i \le m_n$, $1 \le j \le m'_{in}$, $1 \le p \le N_i$, $V_{(TU)'|TU} \in \mathcal{P}_n(\mathcal{T} \times \mathcal{U}|\mathcal{T} \times \mathcal{U})$, 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

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

We shall show that for such Ω_{ij}, there exists a pair of mappings $(\varphi_n^{(0)}, \psi_n^{(0)})$ such that (19) and (20) are satisfied.

We first show that there exists a Y-decoder $\varphi_n^{(0)}$ such that (19) holds. For any $((\mathbf{t},\mathbf{u}),\mathbf{x}) \in \Omega$ and $\mathbf{y} \in \mathcal{Y}^n$, let

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

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

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

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

$$\begin{aligned} P_{Ye}^{(n)}((\mathbf{t},\mathbf{u}),\mathbf{x}) &= W_{Y|TUX}^{(n)}\left(\left\{\mathbf{y} : \varphi_n^{(0)}(\mathbf{y}) \ne ((\mathbf{t},\mathbf{u}),\mathbf{x})\right\} \Big| (\mathbf{t},\mathbf{u}),\mathbf{x}\right) \\ &\le \sum_{\widehat{V}_{Y|TUX} \in \mathcal{P}_n(\mathcal{Y}|P_{(TU)_i X_j})} W_{Y|TUX}^{(n)}\left(\mathbb{T}_{\widehat{V}_{Y|TUX}}((\mathbf{t},\mathbf{u}),\mathbf{x}) \bigcap \left\{\mathbf{y} : \varphi_n^{(0)}(\mathbf{y}) \ne ((\mathbf{t},\mathbf{u}),\mathbf{x})\right\} \Big| \mathbf{t},\mathbf{u},\mathbf{x}\right). \quad (21) \end{aligned}$$

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

$$\left\{\mathbf{y} : \varphi_n^{(0)}(\mathbf{y}) \ne ((\mathbf{t},\mathbf{u}),\mathbf{x})\right\} = \underbrace{\left\{\mathbf{y} : \varphi_n^{(0)}(\mathbf{y}) = ((\mathbf{t},\mathbf{u})',\mathbf{x}'), (\mathbf{t},\mathbf{u})' \ne (\mathbf{t},\mathbf{u})\right\}}_{\triangleq \mathcal{E}_1} \bigcup \underbrace{\left\{\mathbf{y} : \varphi_n^{(0)}(\mathbf{y}) = ((\mathbf{t},\mathbf{u}),\mathbf{x}'), \mathbf{x}' \ne \mathbf{x}\right\}}_{\triangleq \mathcal{E}_2},$$

we can upper bound

$$\begin{aligned} & W_{Y|TUX}^{(n)}\left(\mathbb{T}_{\widehat{V}_{Y|TUX}}((\mathbf{t},\mathbf{u}),\mathbf{x}) \bigcap \left\{\mathbf{y} : \varphi_n^{(0)}(\mathbf{y}) \ne ((\mathbf{t},\mathbf{u}),\mathbf{x})\right\} \Big| \mathbf{t},\mathbf{u},\mathbf{x}\right) \\ &\le \sum_{\mathbf{y} \in \mathbb{T}_{\widehat{V}_{Y|TUX}}((\mathbf{t},\mathbf{u}),\mathbf{x}) \bigcap \mathcal{E}_1} W_{Y|TUX}^{(n)}(\mathbf{y}|\mathbf{t},\mathbf{u},\mathbf{x}) + \sum_{\mathbf{y} \in \mathbb{T}_{\widehat{V}_{Y|TUX}}((\mathbf{t},\mathbf{u}),\mathbf{x}) \bigcap \mathcal{E}_2} W_{Y|TUX}^{(n)}(\mathbf{y}|\mathbf{t},\mathbf{u},\mathbf{x}). \quad (22) \end{aligned}$$

[1] Note that for the symmetric multiple access channel, it has been shown in [21] that the minimum conditional entropy (MCE) decoder leads to a larger channel error exponent than the MMI decoder; however, for the asymmetric 2-user channel with superposition coding, MMI decoding is equivalent to MCE decoding.

It can be shown by the type packing lemma (Lemma 2) and a standard counting argument (see Appendix B) that

$$\left|\mathbb{T}_{\widehat{V}_{Y|TUX}}((\mathbf{t},\mathbf{u}),\mathbf{x})\bigcap\mathcal{E}_1\right| \le m_n\left(\max_i m'_{in}\right)(n+1)^{|\mathcal{T}\times\mathcal{U}|^2|\mathcal{X}|^2|\mathcal{Y}|} \times 2^{n\left[H_{P_{(TU)_iX_j}\widehat{V}_{Y|TUX}}(Y|T,U,X)-\left|I_{P_{(TU)_iX_j}\widehat{V}_{Y|TUX}}(T,U,X;Y)-(R_i+R_{ij})\right|^+\right]}, \tag{23}$$

and

$$\left|\mathbb{T}_{\widehat{V}_{Y|TUX}}((\mathbf{t},\mathbf{u}),\mathbf{x})\bigcap\mathcal{E}_2\right| \le \left(\max_i m'_{in}\right)(n+1)^{|\mathcal{T}\times\mathcal{U}||\mathcal{X}|^2|\mathcal{Y}|} \times 2^{n\left[H_{P_{(TU)_iX_j}\widehat{V}_{Y|TUX}}(Y|T,U,X)-\left|I_{P_{(TU)_iX_j}\widehat{V}_{Y|TUX}}(X;Y|T,U)-R_{ij}\right|^+\right]}. \tag{24}$$

Using the identity (cf. Lemma 1) when $((\mathbf{t},\mathbf{u}),\mathbf{x})\in\Omega_{ij}\subseteq\mathbb{T}_{(TU)_iX_j}$ and $\mathbf{y}\in\mathbb{T}_{\widehat{V}_{Y|TUX}}((\mathbf{t},\mathbf{u}),\mathbf{x})$

$$W^{(n)}_{Y|TUX}(\mathbf{y}|(\mathbf{t},\mathbf{u}),\mathbf{x}) = 2^{-n\left[D\left(\widehat{V}_{Y|TUX}\|W_{Y|TUX}|P_{(TU)_iX_j}\right)+H_{P_{(TU)_iX_j}\widehat{V}_{Y|TUX}}(Y|T,U,X)\right]},$$

we obtain

$$\sum_{\mathbf{y}\in\mathbb{T}_{\widehat{V}_{Y|TUX}}((\mathbf{t},\mathbf{u}),\mathbf{x})\bigcap\mathcal{E}_1} W^{(n)}_{Y|TUX}(\mathbf{y}|((\mathbf{t},\mathbf{u}),\mathbf{x})\in\Omega_{ij}) \le m_n\left(\max_i m'_{in}\right)(n+1)^{|\mathcal{T}\times\mathcal{U}|^2|\mathcal{X}|^2|\mathcal{Y}|} \times 2^{-n\left[D\left(\widehat{V}_{Y|TUX}\|W_{Y|TUX}|P_{(TU)_iX_j}\right)+\left|I_{P_{(TU)_iX_j}\widehat{V}_{Y|TUX}}(T,U,X;Y)-(R_i+R_{ij})\right|^+\right]}, \tag{25}$$

and

$$\sum_{\mathbf{y}\in\mathbb{T}_{\widehat{V}_{Y|TUX}}((\mathbf{t},\mathbf{u}),\mathbf{x})\bigcap\mathcal{E}_2} W^{(n)}_{Y|TUX}(\mathbf{y}|((\mathbf{t},\mathbf{u}),\mathbf{x})\in\Omega_{ij}) \le \left(\max_i m'_{in}\right)(n+1)^{|\mathcal{T}\times\mathcal{U}||\mathcal{X}|^2|\mathcal{Y}|} \times 2^{-n\left[D\left(\widehat{V}_{Y|TUX}\|W_{Y|TUX}|P_{(TU)_iX_j}\right)+\left|I_{P_{(TU)_iX_j}\widehat{V}_{Y|TUX}}(X;Y|T,U)-R_{ij}\right|^+\right]}. \tag{26}$$

Substituting (25) and (26) back into (22) and (21) successively, noting that $|\mathcal{P}_n(\mathcal{Y}|P_{(TU)_iX_j})|$ is polynomial in n by Lemma 1, we obtain that, for any $\delta>0$, there exists a Y-decoder $\varphi_n^{(0)}$ such that, given $((\mathbf{t},\mathbf{u}),\mathbf{x})\in\Omega_{ij}$, the probability of Y-error is bounded by

$$P^{(n)}_{Ye}((\mathbf{t},\mathbf{u}),\mathbf{x}) \le 2^{-n\left[E_Y\left(R_i,R_{ij},W_{Y|TUX},P_{(TU)_i}P_{X_j|(TU)_i}\right)-\delta\right]} \tag{27}$$

for sufficiently large n.

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

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

where $R_i = \frac{1}{n}\log_2 N_i$ if $(\mathbf{t},\mathbf{u}) \in \Omega_i$. Note that $\beta((\mathbf{t},\mathbf{u}),\mathbf{x};\mathbf{z})$ is independent of $\mathbf{x}$. Let $\widetilde{\Omega} = \sum_{i=1}^{m_n} \Omega_i$. The Z-decoder $\psi_n^{(0)} : \mathcal{Z}^n \to \Omega$ is defined by

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

It can be shown in a similar manner by using (3) in Lemma 2 that, under the decoder $\psi_n^{(0)}$, the probability of the Z-error is bounded by

$$P_{Ze}^{(n)}((\mathbf{t},\mathbf{u}),\mathbf{x}) \leq 2^{-n\left[E_Z\left(R_i, R_{ij}, W_{Z|TUX}, P_{(TU)_i} P_{X_j|(TU)_i}\right)-\delta\right]} \tag{28}$$

for sufficiently large n. Finally, we remark that Lemma 2 ensures that there exist mappings $(\varphi_n^{(0)}, \psi_n^{(0)})$ such that (28) holds simultaneously with (27). ■

Theorem 1 is an auxiliary result for the channel coding problem for the 2-user asymmetric channel. To apply it to our 2-user source-channel system, we need to design the encoders which can map a pair of correlated source messages to a particular $(\mathbf{t},\mathbf{u},\mathbf{x})$ with a joint type, so that the total probabilities of error still vanish exponentially. We hence can establish the following bounds.

Theorem 2 Given 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,

$$E_{JY}(Q_{SL}, W_{YZ|TUX}, \widetilde{P}_{TUX}, \tau) \triangleq \min_{P_{SL}} \left[\tau D(P_{SL} \parallel Q_{SL}) + E_Y(\tau H_P(S), \tau H_P(L|S), W_{Y|TUX}, \widetilde{P}_{TUX})\right], \tag{29}$$

and

$$E_{JZ}(Q_{SL}, W_{YZ|TUX}, \widetilde{P}_{TUX}, \tau) \triangleq \min_{P_{SL}} \left[\tau D(P_{SL} \parallel Q_{SL}) + E_Z(\tau H_P(S), \tau H_P(L|S), W_{Z|TUX}, \widetilde{P}_{TUX})\right], \tag{30}$$

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

$$E_J(Q_{SL}, W_{YZ|UX}, \tau) \geq \min_{P_{SL}} \left[\tau D(P_{SL} \parallel Q_{SL}) + E_r(\tau H_P(S), \tau H_P(L|S), W_{YZ|UX})\right] \tag{31}$$

where

$$E_r(R_1, R_2, W_{YZ|UX}) \triangleq \sup_{\mathcal{T}} \max_{P_{TUX}} E_r(R_1, R_2, W_{YZ|TUX}, P_{TUX}), \tag{32}$$

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

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

where E_Y and E_Z are given by (14) and (15), respectively.

We remark that (29) and (30) can be achieved by a sequence of codes without the knowledge of Q_{SL} and $W_{YZ|UX}$, but the lower bound (31) 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 (29) and (30). We need to show that, for any given $\widetilde{P}_{TUX} \in \mathcal{P}(\mathcal{T} \times \mathcal{U} \times \mathcal{X})$ and $\delta > 0$, there exists a sequence of JSC codes such that both the probabilities of decoding error are upper bounded by

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

where E_{JY} and E_{JZ} are given by (29) and (30).

To apply Theorem 1, set $m_n \triangleq |\mathcal{P}_{\tau n}(\mathcal{S})|$. For each type $P_{S_i} \in \mathcal{P}_{\tau n}(\mathcal{S})$, $i = 1, 2, ..., m_n$, denote N_i be the cardinalities of these type classes, $N_i \triangleq |\mathbb{T}_{S_i}|$, and set $m'_{in} \triangleq |\mathcal{P}_{\tau n}(\mathcal{L}|P_{S_i})|$. For each conditional type $P_{L_j|S_i} \in \mathcal{P}_{\tau n}(\mathcal{L}|P_{S_i})$, $j = 1, 2, ..., m'_{in}$, denote M_{ij} be the cardinalities of these type classes, $M_{ij} \triangleq |\mathbb{T}_{L_j|S_i}(\mathbf{s})|$ where $\mathbf{s}$ is an arbitrary sequence in $\mathbb{T}_{S_i}$. Note that $|\mathbb{T}_{L_j|S_i}(\mathbf{s})|$ is constant for all $\mathbf{s} \in \mathbb{T}_{S_i}$. 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 $\widetilde{P}_{TUX}$ 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 $P_{TUX} \to \widetilde{P}_{TUX}$ uniformly[2] as $n \to \infty$. Thus, we can choose, by the continuity of $E_k(R_i, R_{ij}, W_{k|TUX}, \widetilde{P}_{TUX})$ with respect to $\widetilde{P}_{TUX}$, for each $i = 1, 2, ..., m_n$, and $j = j(i) = 1, 2, ..., m'_{in}$, the joint type $P_{(TU)_i X_j} = P_{TUX}$ such that the following are satisfied

$$\left| E_k(R_i, R_{ij}, W_{k|TUX}, P_{TUX}) - E_k(R_i, R_{ij}, W_{k|TUX}, \widetilde{P}_{TUX}) \right| < \frac{\delta}{4}, \qquad k = Y, Z$$

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 $i = 1, 2, ..., m_n$ 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 $i = 1, 2, ..., m_n$ and $j = 1, 2, ..., m'_{in}$, i.e., $P_{X|TU}(\mathbf{x}|\mathbf{t}, \mathbf{u}) = P_{TUX}(\mathbf{t}, \mathbf{u}, \mathbf{x})/P_{TU}(\mathbf{t}, \mathbf{u})$ for any $(\mathbf{t}, \mathbf{u}, \mathbf{x}) \in \mathbb{T}_{TUX}$.

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 $i = 1, 2, ..., \widehat{m}_n$, $j = 1, 2, ..., \widehat{m}'_{in}$,

$$R_i < H_{P_{(TU)_i}}(T, U) - \frac{\delta}{4}, \qquad i = 1, 2, ..., \widehat{m}_n \tag{33}$$

and

$$R_{ij} < H_{P_{(TU)_i X_j}}(X|T, U) - \frac{\delta}{4}, \qquad i = 1, 2, ..., \widehat{m}_n, \quad j = j(i) = 1, 2, ..., \widehat{m}'_{in}, \tag{34}$$

where $\widehat{m}_n \leq m_n$ and $\widehat{m}'_{in} \leq m'_n$. Then according to Theorem 1, there exist pairwise disjoint subsets $\Omega_{ij} \subseteq \mathbb{T}_{(TU)_i X_j}$ with $|\Omega_{ij}| = N_i M_{ij}$, $i = 1, 2, ..., \widehat{m}_n$, $j = 1, 2, ..., \widehat{m}'_{in}$, and a pair of mappings $\left(\varphi_n^{(0)}, \psi_n^{(0)}\right)$,

[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 [11] between P_{X_i} and P_X^* converges to zero as $n \to \infty$.

such that the probabilities of erroneous transmission of a $((\mathbf{t},\mathbf{u}),\mathbf{x}) \in \Omega_{ij}$ are *simultaneously* bounded for the channel $W_{YZ|TUX}$ as

$$\begin{aligned} P_{Ye}^{(n)}(\mathbf{t},\mathbf{u},\mathbf{x}) &\leq 2^{-n\left[E_Y\left(R_i,R_{ij},W_{Y|TUX},P_{(TU)_iX_j}\right)-\delta/4\right]} \\ &\leq 2^{-n[E_Y\left(R_i,R_{ij},W_{Y|TUX},\widetilde{P}_{TUX}\right)-\delta/2]} \end{aligned} \tag{35}$$

and

$$\begin{aligned} P_{Ze}^{(n)}(\mathbf{t},\mathbf{u},\mathbf{x}) &\leq 2^{-n\left[E_Z\left(R_i,R_{ij},W_{Z|TUX},P_{(TU)_iX_j}\right)-\delta/4\right]} \\ &\leq 2^{-n[E_Z\left(R_i,R_{ij},W_{Z|TUX},\widetilde{P}_{TUX}\right)-\delta/2]}. \end{aligned} \tag{36}$$

For the N_i, M_{ij}, $P_{(TU)_i}$, and $P_{X_j|(TU)_i}$ violating (33) or (34) (i.e., for $i > \widehat{m}_n$ or $j > \widehat{m}'_{in}$), (35) and (36) trivially hold for arbitrary choice of disjoint subsets Ω_{ij} since $E_Y\left(R_i,R_{ij},W_{Y|TUX},P_{(TU)_iX_j}\right)$ or $E_Z\left(R_i,R_{ij},W_{Z|TUX},P_{(TU)_iX_j}\right)$ 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,

$$\begin{aligned} E_Y\left(R_i,R_{ij},W_{Y|TUX},P_{(TU)_iX_j}\right) \leq \min\Big\{ &I_{P_{(TU)_iX_j}W_{Y|TUX}}(T,U,X;Y) - R_i - R_{ij}, \\ &I_{P_{(TU)_iX_j}W_{Y|TUX}}(X;Y|T,U) - R_{ij}\Big\} \end{aligned} \tag{37}$$

and

$$E_Z\left(R_i,R_{ij},W_{Z|TUX},P_{(TU)_iX_j}\right) \leq I_{P_{(TU)_iX_j}W_{Z|TUX}}(T,U;Z) - R_i. \tag{38}$$

If $R_i \geq H_{P_{(TU)_i}}(T,U)-\frac{\delta}{4} \geq I_{P_{(TU)_iX_j}W_{Z|TUX}}(T,U;Z)-\frac{\delta}{4}$, then by (38) $E_Z\left(R_i,R_{ij},W_{Z|TUX},P_{(TU)_iX_j}\right) \leq \frac{\delta}{4}$. Similarly, if $R_{ij} \geq H_{P_{(TU)_iX_j}}(X|T,U) - \frac{\delta}{4}$, then by (37) $E_Y\left(R_i,R_{ij},W_{Y|TUX},P_{(TU)_iX_j}\right) \leq \frac{\delta}{4}$.

Therefore, we may construct the JSC code $(f_n, g_n, \varphi_n, \psi_n)$ for CS Q_{SL} and the 2-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.

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

Encoder f_n: For the message pair $(\mathbf{s},\mathbf{l}) \in \mathbb{T}_{S_iL_j}$ such that $i > \widehat{m}_n$ or $j > \widehat{m}'_{in}$, let $f_n(\mathbf{s},\mathbf{l}) = \mathbf{0} \in \mathcal{X}^n$. For the $(\mathbf{s},\mathbf{l}) \in \mathbb{T}_{S_iL_j}$ such that $i \leq \widehat{m}_n$ and $j \leq \widehat{m}'_{in}$, noting that $|\mathbb{T}_{L_j|S_i}(\mathbf{s})| = |\Omega_{ij}(\varphi_n(\mathbf{s}))| = M_{ij}$ if $\mathbf{s} \in \mathbb{T}_{S_i}$, let $f_n^{(1)}(\mathbf{s},\cdot) : \mathbb{T}_{L_j|S_i}(\mathbf{s}) \to \Omega_{ij}(g_n(\mathbf{s}))$ be a bijection such that $f_n^{(1)}(\mathbf{s},\mathbf{l}) = (g_n^{(1)}(\mathbf{s}),\mathbf{x}) \in \Omega_{ij}$. Let $f_n(\mathbf{s},\mathbf{l})$ be the third component $\mathbf{x}$ of $f_n^{(1)}(\mathbf{s},\mathbf{l})$.

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

<u>*Y-Decoder* φ_n*:*</u> The $Y-$decoder is defined by

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

<u>*Z-Decoder* ψ_n*:*</u> The $Z-$decoder is defined by

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

For such JSC code (f_n, g_n, φ_n, ψ_n), the probabilities of Y-error and Z-error are bounded by

$$P_{Ye}^{(n)}(\mathbf{s}, \mathbf{l}) \le 2^{-n[E_Y\left(R_i, R_{ij}, W_{Y|TUX}, \widetilde{P}_{TUX}\right) - \delta/2]} \qquad \text{if } (\mathbf{s}, \mathbf{l}) \in \mathbb{T}_{S_i L_j} \tag{39}$$

and

$$P_{Ze}^{(n)}(\mathbf{s}, \mathbf{l}) \le 2^{-nE_Z\left(R_i, R_{ij}, W_{Z|TUX}, \widetilde{P}_{TUX}\right) - \delta/2]} \qquad \text{if } (\mathbf{s}, \mathbf{l}) \in \mathbb{T}_{S_i L_j}. \tag{40}$$

Substituting (39) and (40) into (8) 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,

$$\begin{aligned} & P_{Ye}^{(n)}(Q_{SL}, W_{YZ|UX}, \tau) \\ \le & \sum_{i,j} 2^{-n[\tau D(P_{S_i L_j} \| Q_{SL}) + E_Y(R_i, R_{ij}, W_{Y|TUX}, \widetilde{P}_{TUX}) - \delta/2]} \\ \le & \sum_{P_{SL}} 2^{-n\left[\tau D(P_{SL} \| Q_{SL}) + E_Y\left(\tau H_P(S) - o_1(n), \tau H_P(L|S) - o_2(n), W_{Y|TUX}, \widetilde{P}_{TUX}\right) - \delta/2\right]} \\ \le & \sum_{P_{SL}} 2^{-n\left[\tau D(P_{SL} \| Q_{SL}) + E_Y\left(\tau H_P(S), \tau H_P(L|S), W_{Y|TUX}, \widetilde{P}_{TUX}\right) - \delta\right]} \end{aligned} \tag{41}$$

and

$$\begin{aligned} & P_{Ze}^{(n)}(Q_{SL}, W_{YZ|UX}, \tau) \\ \le & \sum_{i,j} 2^{-n[\tau D(P_{S_i L_j} \| Q_{SL}) + E_Z(R_i, R_{ij}, W_{Z|TUX}, \widetilde{P}_{TUX}) - \delta/2]} \\ \le & \sum_{P_{SL}} 2^{-n\left[\tau D(P_{SL} \| Q_{SL}) + E_Z\left(\tau H_P(S) - o_1(n), \tau H_P(L|S) - o_2(n), W_{Z|TUX}, \widetilde{P}_{TUX}\right) - \delta/2\right]} \\ \le & \sum_{P_{SL}} 2^{-n\left[\tau D(P_{SL} \| Q_{SL}) + E_Z\left(\tau H_P(S), \tau H_P(L|S), W_{Z|TUX}, \widetilde{P}_{TUX}\right) - \delta\right]}, \end{aligned} \tag{42}$$

where $o_1(n) = \frac{|\mathcal{S}| \log_2(\tau n + 1)}{n}$ and $o_2(n) = \frac{|\mathcal{S}||\mathcal{L}| \log_2(\tau n + 1)}{n}$. Finally, the bounds (29) and (30) follow from (41) and (42), 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 (31), we slightly modify the above approach by choosing $P_{(TU)_i X_j} = \widetilde{P}^*_{(TU)_i X_j}$ which achieves the maximum and the supremum of $E_r(R_i, R_{ij}, W_{YZ|UX})$ in (32) for every R_i and R_{ij},

$i = 1, 2, ..., m_n,\ j = 1, 2, ..., m'_{in}$. Then the probabilities of Y-error and Z-error in (39) and (40) are bounded by

$$\begin{aligned} P_{Ye}^{(n)}(\mathbf{s}, \mathbf{l}) &\leq 2^{-n[E_Y\left(R_i, R_{ij}, W_{Y|TUX}, \widetilde{P}^*_{(TU)_i X_j}\right) - \delta/2]} \\ &\leq 2^{-n[E_r(R_i, R_{ij}, W_{YZ|UX}) - \delta/2]} \qquad \text{if } (\mathbf{s}, \mathbf{l}) \in \mathbb{T}_{S_i L_j} \end{aligned} \tag{43}$$

and

$$\begin{aligned} P_{Ze}^{(n)}(\mathbf{s}, \mathbf{l}) &\leq 2^{-nE_Z\left(R_i, R_{ij}, W_{Z|TUX}, \widetilde{P}^*_{TU_i X_j}\right) - \delta/2]} \\ &\leq 2^{-n[E_r(R_i, R_{ij}, W_{YZ|UX}) - \delta/2]} \qquad \text{if } (\mathbf{s}, \mathbf{l}) \in \mathbb{T}_{S_i L_j} \end{aligned} \tag{44}$$

for n sufficiently large. The left of the proof is the same as the one for (29) and (30). ■

By examining the positivity of the lower bound to E_J, we obtain a sufficient condition for reliable transmissibility for the asymmetric 2-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

$$\mathcal{R}(W_{YZ|UX}) \triangleq \bigcup_{\mathcal{T}:|\mathcal{T}| \leq |\mathcal{U}||\mathcal{X}|+1} \ \bigcup_{P_{TUX} \in \mathcal{P}(\mathcal{T} \times \mathcal{U} \times \mathcal{X})} \mathcal{R}(W_{YZ|TUX}, P_{TUX}) \tag{45}$$

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 $P_{TUXYZ} = P_{TUX} W_{YZ|UX}$.

Theorem 3 (JSCC Theorem) Given Q_{SL}, $W_{YZ|UX}$ and $\tau > 0$, the following statements hold.
(1) The sources can be transmitted over the channel with $P_e^{(n)} \to 0$ as $n \to \infty$ if $(\tau H_Q(S), \tau H_Q(L|S)) \in \mathcal{R}(W_{YZ|UX})$;
(2) Conversely, if the sources can be transmitted over the channel with an arbitrarily small probability of error $P_e^{(n)}$ as $n \to \infty$, then $(\tau H_Q(S), \tau H_Q(L|S)) \in \mathcal{R}(W_{YZ|UX})$ with $<$ replaced by $\leq$ in $\mathcal{R}(W_{YZ|UX})$.

Proof:
Forward Part (1): It follows from (16)-(18) 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 (31) 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 from a similar manner as the converse part of [15, Theorem 1] for a broadcast channel. For the sake of completeness, we also provide a full proof here since we deal with a 2-user channel. First, we remark that (as shown in [15, Theorem 2]) the region $\mathcal{R}(W_{YZ|TUX}, P_{TUX})$ can be equivalently rewritten by

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

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 a 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 $(\tau H_Q(S), \tau H_Q(L|S)) \in \mathcal{R}(W_{YZ|UX}, P_{TUX})$ with $<$ replaced by $\le$, 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

$$H(S^k, L^k|Y^n) \le P_{Ye}^{(n)} \log_2 |\mathcal{S}^k \times \mathcal{L}^k| + H\left(P_{Ye}^{(n)}\right) \triangleq n\epsilon_{1n} \tag{46}$$

$$H(S^k|Z^n) \le P_{Ze}^{(n)} \log_2 |\mathcal{S}^k| + H\left(P_{Ze}^{(n)}\right) \triangleq n\epsilon_{2n}, \tag{47}$$

where $S^k \triangleq (S_1, S_2, \cdots, S_k)$; similar definitions apply for the other tuples. It follows from (46)-(47) that

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

where $Y^{i-1} = (Y_1, Y_2, ..., Y_{i-1})$ and $\mathbf{Z}^{i+1} \triangleq (Z_{i+1}, Z_{i+2}, ..., Z_n)$. Substituting the identity [12, Lemma 7]

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

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

$$\begin{aligned} kH(S,L) &\leq \sum_{i=1}^{n} \Big[I(L^k; Y_i|T_i) + I(T_i; Z_i) \Big] + n(\epsilon_{1n} + \epsilon_{2n}) \\ &\stackrel{(a)}{=} \sum_{i=1}^{n} \Big[I(L^k; Y_i|T_i, U_i) + I(T_i, U_i; Z_i) \Big] + n(\epsilon_{1n} + \epsilon_{2n}) \\ &\stackrel{(b)}{\leq} \sum_{i=1}^{n} [I(X^n; Y_i|T_i, U_i) + I(T_i, U_i; Z_i)] + n(\epsilon_{1n} + \epsilon_{2n}) \\ &\stackrel{(c)}{=} \sum_{i=1}^{n} [I(X_i; Y_i|T_i, U_i) + I(T_i, U_i; Z_i)] + n(\epsilon_{1n} + \epsilon_{2n}), \end{aligned} \tag{48}$$

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

$$\begin{aligned} kH(S,L) &= H(S^k, L^k) \\ &= I(S^k, L^k; Y^n) + H(S^k, L^k|Y^n) \\ &\leq I(X^n, U^n; Y^n) + n\epsilon_{1n} \\ &= \sum_{i=1}^{n} I(U_i, X_i; Y_i) + n\epsilon_{1n}. \end{aligned} \tag{49}$$

Likewise, it follows from (47) that

$$\begin{aligned} kH(S) &= H(S^k) \\ &= I(S^k; Z^n) + H(S^k|Z^n) \\ &= \sum_{i=1}^{n} I(S^k; Z_i|\mathbf{Z}^{i+1}) + H(S^k|Z^n) \\ &\leq \sum_{i=1}^{n} I(S^k, \mathbf{Z}^{i+1}; Z_i) + n\epsilon_{2n} \\ &\leq \sum_{i=1}^{n} I(S^k, Y^{i-1}, \mathbf{Z}^{i+1}, U_i; Z_i) + n\epsilon_{2n} \\ &= \sum_{i=1}^{n} I(T_i, U_i; Z_i) + n\epsilon_{2n}. \end{aligned} \tag{50}$$

Note also that $T_i \longrightarrow (U_i, X_i) \longrightarrow (Y_i, Z_i)$ for all $1 \leq i \leq n$. According to (48), (49), and (50), and recalling that $k = \tau n$, it is easy to show (e.g., see [12]) that there exists an auxiliary RV T with $P_{TUXYZ} = P_T P_{UX|T} W_{YZ|UX}$ such that

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

which is equivalent to

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

Finally, by using the Carathéodory theorem (cf. [11, p. 311]) we can show that there exists a RV $\widehat{T}$ with $|\widehat{\mathcal{T}}| \le |\mathcal{U}||\mathcal{X}|+1$ such that $P_{\widehat{T}UXYZ} = P_{\widehat{T}} P_{UX|\widehat{T}} W_{YZ|UX}$ and

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

This completes the proof of the converse part. ■

4.4 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 2-user channel, it can be shown by using a similar approach based on the method of types that the following is true.

As a special case of the JSCC system, let the (common and private) message pair $(\mathbf{s},\mathbf{l})$ be uniformly drawn from the finite set $\mathcal{M}_s \times \mathcal{M}_l$, where $\mathcal{M}_s \triangleq \{1,2,...,M_s\}$ and $\mathcal{M}_l \triangleq \{1,2,...,M_l\}$. An asymmetric 2-user channel code with block length n for transmitting the uniform message set is a quadruple of mappings, $(f_{cn}, g_{cn}, \varphi_{cn}, \psi_{cn})$, where $f_{cn} : \mathcal{M}_s \times \mathcal{M}_l \to \mathcal{X}^n$ and $g_{cn} : \mathcal{M}_s \to \mathcal{U}^n$ are the channel encoders, and $\varphi_{cn} : \mathcal{Y}^n \to \mathcal{M}_s \times \mathcal{M}_l$ and $\psi_{cn} : \mathcal{Z}^n \to \mathcal{M}_s$ are respectively the Y-decoder and Z-decoder. Let $R_1 \triangleq \frac{1}{n}\log_2 M_s$ and $R_2 \triangleq \frac{1}{n}\log_2 M_l$ be the common and private rates of the code respectively. The probabilities of Y- and Z-error of the channel coding are respectively given by

$$P_{Ye}^{(n)}(R_1,R_2,W_{YY|UX}) \triangleq \Pr(\{\varphi_{cn}(\mathbf{y}) \ne (\mathbf{s},\mathbf{l})\}) = \frac{1}{2^{R_1+R_2}} \sum_{\mathbf{s},\mathbf{l}} \sum_{\mathbf{y}:\varphi_{cn}(\mathbf{y})\ne(\mathbf{s},\mathbf{l})} W_{Y|X}^{(n)}(\mathbf{y}|\mathbf{u},\mathbf{x}) \tag{51}$$

and

$$P_{Ze}^{(n)}(R_1,R_2,W_{YZ|UX}) \triangleq \Pr(\{\psi_{cn}(\mathbf{z}) \ne \mathbf{s}\}) = \frac{1}{2^{R_1+R_2}} \sum_{\mathbf{s},\mathbf{l}} \sum_{\mathbf{z}:\psi_{cn}(\mathbf{z})\ne\mathbf{s}} W_{Z|X}^{(n)}(\mathbf{z}|\mathbf{u},\mathbf{x}) \tag{52}$$

where $\mathbf{x} \triangleq f_{cn}(\mathbf{s}, \mathbf{l})$ and $\mathbf{u} \triangleq g_{cn}(\mathbf{s})$. Similarly, the probability of the overall asymmetric 2-user channel coding error is given by

$$P_e^{(n)}(R_1, R_2, W_{YZ|UX}) \triangleq \Pr\left(\{\varphi_{cn}(\mathbf{y}) \neq (\mathbf{s}, \mathbf{l})\} \cup \{\psi_{cn}(\mathbf{z}) \neq \mathbf{s}\}\right), \tag{53}$$

where $(\mathbf{s}, \mathbf{l})$ are uniformly drawn from $\mathcal{M}_s \times \mathcal{M}_l$.

Definition 2 The asymmetric 2-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 blocklength n, the common rate no less than R_1, and the private rate no less than R_2, such that

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

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 $P_{Ye}^{(n)}(R_1, R_2, W_{Y|UX})$ and $P_{Ze}^{(n)}(R_1, R_2, W_{Z|UX})$) but less than the sum of the two, so we have

$$\liminf_{n\to\infty} -\frac{1}{n} \log_2 P_e^{(n)}(R_1, R_2, W_{YZ|UX}) = \liminf_{n\to\infty} -\frac{1}{n} \log_2 \max\left(P_{Ye}^{(n)}(R_1, R_2, W_{Y|UX}), P_{Ze}^{(n)}(R_1, R_2, W_{Z|UX})\right). \tag{55}$$

Our upper bound for the system JSCC error exponent is stated as follows.

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

$$E_J(Q_{SL}, W_{YZ|UX}, \tau) \leq \inf_{P_{SL}} \left[\tau D(P_{SL} \parallel Q_{SL}) + E(\tau H_P(S), \tau H_P(L|S), W_{YZ|UX})\right], \tag{56}$$

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

Proof: First, from (8) we can write

$$P_{ie}^{(n)}(Q_{SL}, W_{YZ|UX}, \tau) \geq \max_{P_{SL}\in\mathcal{P}_{\tau n}(\mathcal{S}\times\mathcal{L})} Q_{SL}^{(\tau n)}(\mathbb{T}_{SL}) P_{ie}(\mathbb{T}_{SL}) \quad i = Y, Z, \tag{57}$$

where $P_{Ye}(\mathbb{T}_{SL})$ and $P_{Ze}(\mathbb{T}_{SL})$ are given by (9) and (10), respectively. Comparing (9) with (51), and comparing (10) with (52), 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 2-user channel coding with (common and private) message sets $\mathbb{T}_{SL}$, since $(\mathbf{s}, \mathbf{l})$ 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 $\mathbf{s} \in \mathbb{T}_S = \mathbb{T}_{P_S}$,

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

and that $\mathbb{T}_{L|S}(\mathbf{s})$ is the same set for all $\mathbf{s} \in \mathbb{T}_S$. Hence, we can write $\mathbb{T}_{SL}$ by the product of two sets $\mathbb{T}_{SL} = \mathbb{T}_S \times \mathbb{T}_{L|S}(\mathbf{s})$. Setting $R_1 = \frac{1}{n}\log_2 |\mathbb{T}_S|$ and $R_2 = \frac{1}{n}\log_2 |\mathbb{T}_{L|S}(\mathbf{s})|$, it follows that, by the definition of asymmetric 2-user channel coding error exponent and (55),

$$\begin{aligned}\liminf_{n\to\infty} -\frac{1}{n}\log_2 \max_{i=Y,Z} P_{ie}(\mathbb{T}_{SL}) &\le E(\liminf_{n\to\infty} R_1, \liminf_{n\to\infty} R_2, W_{YZ|UX}) \\ &= E(\tau H_P(S), \tau H_P(L|S), W_{YZ|UX}) \qquad (58)\end{aligned}$$

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

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

According to (13), we write

$$\begin{aligned}&\liminf_{n\to\infty} -\frac{1}{n}\log_2 P_e^{(n)}(Q_{SL}, W_{YZ|UX}, \tau) \\ &= \liminf_{n\to\infty} -\frac{1}{n}\log_2 \max\left(P_{Ye}^{(n)}(Q_{SL}, W_{Y|X}, \tau), P_{Ze}^{(n)}(Q_{SL}, W_{Z|X}, \tau)\right) \\ &\le \liminf_{n\to\infty} -\frac{1}{n}\log_2 \max_{i=Y,Z} \max_{P_{SL}\in\mathcal{P}_{\tau n}(\mathcal{S}\times\mathcal{L})} Q_{SL}^{(\tau n)}(\mathbb{T}_{SL}) P_{ie}(\mathbb{T}_{SL}) \\ &= \liminf_{n\to\infty} \min_{P_{SL}\in\mathcal{P}_{\tau n}(\mathcal{S}\times\mathcal{L})} -\frac{1}{n}\log_2 Q_{SL}^{(\tau n)}(\mathbb{T}_{SL}) \max_{i=Y,Z} P_{ie}(\mathbb{T}_{SL}) \\ &= \liminf_{n\to\infty} \min_{P_{SL}\in\mathcal{P}_{\tau n}(\mathcal{S}\times\mathcal{L})} \left[-\frac{1}{n}\log_2 Q_{SL}^{(\tau n)}(\mathbb{T}_{SL}) - \frac{1}{n}\log_2 \max_{i=Y,Z} P_{ie}(\mathbb{T}_{SL})\right]. \qquad (59)\end{aligned}$$

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

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

which implies

$$\limsup_{n\to\infty} -\frac{1}{n}\log_2 Q_{SL}^{(\tau n)}(\mathbb{T}_{SL}) \le \tau D(P_{SL} \parallel Q_{SL}). \qquad (60)$$

Now assume that

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

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

$\left\{\widehat{P}_{SL} \in \mathcal{P}_{\tau n}(\mathcal{S} \times \mathcal{L})\right\}_{n=n_o}^{\infty}$ such that $\widehat{P}_{SL} \to P^*_{SL}$ uniformly[3]. It then follows from (59), (58) and (60) that

$$\begin{aligned}
&\liminf_{n\to\infty} -\frac{1}{n}\log_2 P_e^{(n)}(Q_{SL}, W_{YZ|UX}, \tau) \\
&\leq \liminf_{n\to\infty}\left[-\frac{1}{n}\log_2 Q_{SL}^{(\tau n)}(\mathbb{T}_{\widehat{P}_{SL}}) - \frac{1}{n}\log_2 \max_{i=Y,Z} P_{ie}(\mathbb{T}_{\widehat{P}_{SL}})\right] \\
&\leq \tau D(P^*_{SL} \parallel Q_{SL}) + E(\tau H_{P^*}(S), \tau H_{P^*}(L|S), W_{YZ|UX}). \qquad (61)
\end{aligned}$$

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

5 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.

5.1 CS-AMAC System

Setting $|\mathcal{Z}| = 1$ and removing the decoder ψ_n, the 2-user asymmetric channel $W_{YZ|UX}$ reduces to an AMAC $W_{Y|UX}$. Since the CS-AMAC system is a special case of the 2-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 "Z-error" and achievable error exponent pair here. The first union in (45) 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 $I(X; Y|T, U) \leq I(X; Y|U)$. Thus Theorem 3 reduces to the same JSCC theorem established in [4] for the CS-AMAC system. Now if we choose 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

$$E_J(Q_{SL}, W_{Y|UX}, \tau) \geq \min_{P_{SL}} \left[\tau D(P_{SL} \parallel Q_{SL}) + E_r(\tau H_P(S), \tau H_P(L|S), W_{Y|UX})\right], \qquad (62)$$

and

$$E_J(Q_{SL}, W_{Y|UX}, \tau) \leq \inf_{P_{SL}} \left[\tau D(P_{SL} \parallel Q_{SL}) + E(\tau H_P(S), \tau H_P(L|S), W_{Y|UX})\right], \qquad (63)$$

where $E(\tau H_P(S), \tau H_P(L|S), W_{Y|UX})$ is the channel error exponent defined in (54), and

$$E_r(R_1, R_2, W_{Y|UX}) = \max_{P_{UX}} E_Y(R_1, R_2, W_{Y|UX}, P_{UX}) \qquad (64)$$

where $E_Y(R_1, R_2, W_{Y|UX}, P_{UX})$ is defined in (14) by setting $T = 1$.

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

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

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

where

$$E_{sp}(R_1, R_2, W_{Y|UX}) \triangleq \max_{P_{UX} \in \mathcal{P}(\mathcal{U} \times \mathcal{X})} \min D(V_{Y|UX} \parallel W_{Y|UX} | P_{UX}), \tag{65}$$

where the minimum is taken over $V_{Y|UX} \in \mathcal{P}(\mathcal{Y}|\mathcal{U} \times \mathcal{X})$ such that $I_{P_{UX}V_{Y|UX}}(U, X; Y) \leq R_1 + R_2$ or $I_{P_{UX}V_{Y|UX}}(X; Y|U) \leq R_2$.

As a consequence, we see that

$$E_J(Q_{SL}, W_{Y|UX}, \tau) \leq \inf_{P_{SL}} \left[\tau D(P_{SL} \parallel Q_{SL}) + E_{sp}(\tau H_P(S), \tau H_P(L|S), W_{Y|UX}) \right]. \tag{66}$$

In Section 6 we investigate the lower and upper bounds when the AMAC has a symmetric distribution.

5.2 CS-ABC System

Setting $|\mathcal{U}| = 1$ and removing the encoder g_n, the 2-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 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, 3 and 4 to the following corollaries.

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

$$\mathcal{R}(W_{YZ|X}) \triangleq \bigcup_{\mathcal{T}: |\mathcal{T}| \leq |\mathcal{X}|+1} \bigcup_{P_{TX} \in \mathcal{P}(\mathcal{T} \times \mathcal{X})} \mathcal{R}(W_{YZ|TX}, P_{TX}) \tag{67}$$

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 $\mathcal{R}(W_{YZ|X})$ is the capacity region of the ABC $W_{YZ|X}$ [19].

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

(1) The sources can be transmitted over the channel 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 can be transmitted over the channel with an arbitrarily small probability of error $P_e^{(n)}$ as $n \to \infty$, then $(\tau H_Q(S), \tau H_Q(L|S)) \in \mathcal{R}(W_{YZ|X}$ with $<$ replaced by $\leq$ in $\mathcal{R}(W_{YZ|X})$.

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

$$E_{JY}(Q_{SL}, W_{YZ|TX}, \widetilde{P}_{TX}, \tau) \triangleq \min_{P_{SL}} \left[\tau D(P_{SL} \parallel Q_{SL}) + E_Y(\tau H_P(S), \tau H_P(L|S), W_{Y|TX}, \widetilde{P}_{TX})\right], \tag{68}$$

and

$$E_{JZ}(Q_{SL}, W_{YZ|TX}, \widetilde{P}_{TX}, \tau) \triangleq \min_{P_{SL}} \left[\tau D(P_{SL} \parallel Q_{SL}) + E_Z(\tau H_P(S), \tau H_P(L|S), W_{Z|TX}, \widetilde{P}_{TX})\right], \tag{69}$$

where E_Y and E_Z are defined in (14) and (15) by setting $U = 1$. Furthermore, given Q_{SL}, $W_{YZ|X}$, and τ, the system JSCC error exponent satisfies

$$E_J(Q_{SL}, W_{YZ|X}, \tau) \geq \min_{P_{SL}} \left[\tau D(P_{SL} \parallel Q_{SL}) + E_r(\tau H_P(S), \tau H_P(L|S), W_{YZ|X})\right] \tag{70}$$

and

$$E_J(Q_{SL}, W_{YZ|X}, \tau) \leq \inf_{P_{SL}} \left[\tau D(P_{SL} \parallel Q_{SL}) + E(\tau H_P(S), \tau H_P(L|S), W_{YZ|X})\right] \tag{71}$$

where $E_r(R_1, R_2, W_{YZ|X})$ is the same as $E_r(R_1, R_2, W_{YZ|UX})$ defined by (32) by setting $U = 1$, and $E(R_1, R_2, W_{YZ|X})$ is the channel error exponent defined in (54).

6 Evaluation of the Bounds for E_J: CS over Symmetric AMAC

We established the lower and upper bounds for the system JSCC error exponent for asymmetric 2-user JSCC systems. However, we are not able to simplify these bounds for general 2-user JSCC systems (even for general CS-AMAC and CS-ABC systems) to a computable parametric form as we did for the point-to-point systems [23, 24]. 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 form of the functions $E_r(R_1, R_2, W_{Y|UX})$ defined in (64) and $E_{sp}(R_1, R_2, W_{Y|UX})$ defined in (65). For any $R_1, R_2 > 0$, rewrite

$$E_Y(R_1, R_2, W_{Y|UX}, P_{UX}) = \min\left\{E_r^{(1)}(R_1 + R_2, W_{Y|UX}, P_{UX}), E_r^{(2)}(R_2, W_{Y|UX}, P_{UX})\right\}$$

where

$$E_r^{(1)}(R, W_{Y|UX}, P_{UX}) \triangleq \min_{V_{Y|UX}} \left[D(V_{Y|UX} \parallel W_{Y|UX}|P_{UX}) + \left|I_{P_{UX}V_{Y|UX}}(U, X; Y) - R\right|^+\right] \tag{72}$$

and

$$E_r^{(2)}(R, W_{Y|UX}, P_{UX}) \triangleq \min_{V_{Y|UX}} \left[D(V_{Y|UX} \parallel W_{Y|UX}|P_{UX}) + \left|I_{P_{UX}V_{Y|UX}}(X; Y|U) - R\right|^+\right]. \tag{73}$$

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

$$E_{sp}(R_1, R_2, W_{Y|UX}, P_{UX}) = \min \left\{ E_{sp}^{(1)}(R_1 + R_2, W_{Y|UX}, P_{UX}), E_{sp}^{(2)}(R_2, W_{Y|UX}, P_{UX}) \right\}$$

where

$$E_{sp}^{(1)}(R, W_{Y|UX}, P_{UX}) \triangleq \min_{V_{Y|UX}} \Big(D(V_{Y|UX} \parallel W_{Y|UX}|P_{UX}) : \quad I_{P_{UX}V_{Y|UX}}(U, X; Y) \leq R \Big) \tag{74}$$

and

$$E_{sp}^{(2)}(R, W_{Y|UX}, P_{UX}) \triangleq \min_{V_{Y|UX}} \Big(D(V_{Y|UX} \parallel W_{Y|UX}|P_{UX}) : \quad I_{P_{UX}V_{Y|UX}}(X; Y|U) \leq R \Big). \tag{75}$$

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 [11]. In fact, it has been shown in [2] that

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

where

$$E_1(\rho_1, W_{Y|UX}, P_{UX}) \triangleq -\log_2 \sum_{y \in \mathcal{Y}} \left(\sum_{(u,x) \in \mathcal{U} \times \mathcal{X}} P_{UX}(u, x) W_{Y|UX}(y|u, x)^{\frac{1}{1+\rho_1}} \right)^{1+\rho_1}, \tag{76}$$

and

$$E_2(\rho_2, W_{Y|UX}, P_{UX}) = -\log_2 \sum_{u \in \mathcal{U}} P_U(u) \sum_{y \in \mathcal{Y}} \left(\sum_{x \in \mathcal{X}} P_{X|U}(x|u) W_{Y|UX}(y|u, x)^{\frac{1}{1+\rho_2}} \right)^{1+\rho_2}. \tag{77}$$

Analogously to [11, Lemma 5.4, Corollary 5.4, p. 168], we can prove the following results; some of them has been proved in [2].

Lemma 3 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 $R \geq R_{cr}^{(i)}(W_{Y|UX}, P_{UX})$ 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 $R \leq R_{cr}^{(i)}(W_{Y|UX}, P_{UX})$, i.e.

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

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 (76) and (77) respectively.

Therefore, we can write the functions $E_r(R_1, R_2, W_{Y|UX})$ defined in (64) and $E_{sp}(R_1, R_2, W_{Y|UX})$ defined in (65) as follows.

$$E_r(R_1, R_2, W_{Y|UX}) = \max_{P_{UX}} \min_{i=1,2} \max_{0 \le \rho \le 1} [E_i(\rho, W_{Y|UX}, P_{UX}) - \rho_i \widehat{R}_i] \tag{78}$$

and

$$E_{sp}(R_1, R_2, W_{Y|UX}) = \max_{P_{UX}} \min_{i=1,2} \max_{\rho \ge 0} [E_i(\rho_i, W_{Y|UX}, P_{UX}) - \rho \widehat{R}_i] \tag{79}$$

where $\widehat{R}_1 = R_1 + R_2$ and $\widehat{R}_2 = R_2$. Since it is very hard to find the optimized solution P_{UX} in general for E_r and E_{sp} in the above, we 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 $\mathcal{Y} = \mathcal{U} = \mathcal{X} = \mathcal{F} = \{0, 1, 2, ..., B-1\}$, and F_i is independent of X_i and U_i, $i = 1, 2, ..., n$.

It is shown in [2] that if the multiple access channel $W_{Y|UX}$ is U-symmetric, then the outer maximum of (78) and (79) would be achieved by a joint distribution with 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 (78) and (79) would be 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 (78) and (79) yields

$$E_r(R_1, R_2, W_{Y|UX}) = \min_{i=1,2} \max_{0 \le \rho \le 1} [\widetilde{E}_i(\rho, W_{Y|UX}) - \rho \widehat{R}_i] \tag{80}$$

and

$$E_{sp}(R_1, R_2, W_{Y|UX}) = \min_{i=1,2} \max_{\rho \geq 0} [\widetilde{E}_i(\rho, W_{Y|UX}) - \rho \widehat{R}_i] \tag{81}$$

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

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

and

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

We also can prove the following identities using a standard optimization method (cf. [23]).

Lemma 4

$$\min_{P_{SL}: H_P(S,L)=R} D(P_{SL} \| Q_{SL}) = \max_{\rho \geq 0} \left[\rho R - E_{s1}(\rho, Q_{SL})\right], \tag{82}$$

$$\min_{P_{SL}: H_P(L|S)=R} D(P_{SL} \| Q_{SL}) = \max_{\rho \geq 0} \left[\rho R - E_{s2}(\rho, Q_{SL})\right], \tag{83}$$

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

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

Clearly, if the marginal distribution $Q_S(s)$ is uniform, then (82) and (83) are equal. Using (80) and (82) we now can write

$$\begin{aligned}
& \min_{P_{SL}} \left[\tau D(P_{SL} \parallel Q_{SL}) + E_r(\tau H_P(S), \tau H_P(L|S), W_{Y|UX})\right] \\
= & \min\Bigg\{ \min_{P_{SL}} \left[\tau D(P_{SL} \parallel Q_{SL}) + \max_{0 \leq \rho_1 \leq 1}[\widetilde{E}_1(\rho_1, W_{Y|UX}) - \rho_1 \tau H_P(S,L)]\right], \\
& \quad \min_{P_{SL}} \left[\tau D(P_{SL} \parallel Q_{SL}) + \max_{0 \leq \rho_2 \leq 1}[\widetilde{E}_2(\rho_2, W_{Y|UX}) - \rho_2 \tau H_P(L|S)]\right] \Bigg\} \\
= & \min\Bigg\{ \min_{R} \left[\min_{P_{SL}: \tau H_P(S,L)=R} \tau D(P_{SL} \parallel Q_{SL}) + \max_{0 \leq \rho_1 \leq 1}[\widetilde{E}_1(\rho_1, W_{Y|UX}) - \rho_1 R]\right], \\
& \quad \min_{R} \left[\min_{P_{SL}: \tau H_P(L|S)=R} \tau D(P_{SL} \parallel Q_{SL}) + \max_{0 \leq \rho_2 \leq 1}[\widetilde{E}_2(\rho_2, W_{Y|UX}) - \rho_2 R]\right] \Bigg\}
\end{aligned} \tag{84}$$

and similarly using (81) we have

$$\begin{aligned}
& \inf_{P_{SL}} \left[\tau D(P_{SL} \parallel Q_{SL}) + E_{sp}(\tau H_P(S), \tau H_P(L|S), W_{Y|UX})\right] \\
= & \min\Bigg\{ \inf_{R} \left[\min_{P_{SL}: \tau H_P(S,L)=R} \tau D(P_{SL} \parallel Q_{SL}) + \max_{\rho_1 \geq 0}[\widetilde{E}_1(\rho_1, W_{Y|UX}) - \rho_1 R]\right], \\
& \quad \inf_{R} \left[\min_{P_{SL}: \tau H_P(L|S)=R} \tau D(P_{SL} \parallel Q_{SL}) + \max_{\rho_2 \geq 0}[\widetilde{E}_2(\rho_2, W_{Y|UX}) - \rho_2 R]\right] \Bigg\}.
\end{aligned} \tag{85}$$

Consequently, using an optimization technique based on Fenchel duality [23], we obtain the following theorem and corollary.

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 (62) and the upper bound given in (66) can be equivalently expressed as

$$\begin{aligned}\min_{i=1,2}\max_{0\le\rho\le 1}[\widetilde{E}_i(\rho,W_{Y|UX})-\tau E_{si}(\rho,Q_{SL})] &\le E_J(Q_{SL},W_{Y|UX},\tau)\\ &\le \min_{i=1,2}\max_{\rho\ge 0}[\widetilde{E}_i(\rho,W_{Y|UX})-\tau E_{si}(\rho,Q_{SL})]. \qquad (86)\end{aligned}$$

Example (Binary CS-AMAC System): Now consider binary CS Q_{SL} with distribution

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

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

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

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

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

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

It follows that

$$\widetilde{E}_1(\rho,W_{Y|UX})=\widetilde{E}_2(\rho,W_{Y|UX})=\rho\log_2 2-(1+\rho)\log_2\left(\epsilon^{\frac{1}{1+\rho}}+(1-\epsilon)^{\frac{1}{1+\rho}}\right).$$

In Fig. 2, we plot the lower and upper bounds for the JSCC error exponent E_J for different (q,ϵ) pairs with transmission rate $t=0.25, 0.35$. As illustrated, the upper and lower bounds coincide (this can be verified by checking whether the two outer minimums in (86) are achieved by the same i and 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.

A Proof of Lemma 2

Although the result (3) of Lemma 2 was already shown in [8], we include its proof here since we need to show that (3) holds simultaneously with (4) and (5). We employ a random selection argument as used in [8]. For each $i = 1, 2, ..., m_n$, we randomly generate a set of $2N_i$ sequences (according to a uniform distribution) from the type class $\mathbb{T}_{A_i} = \mathbb{T}_{P_{A_i}}$, $\mathcal{C}_i \triangleq \left\{\mathbf{a}_1^{(i)}, \mathbf{a}_2^{(i)}, ..., \mathbf{a}_{2N_i}^{(i)}\right\} \subseteq \mathbb{T}_{A_i}$, i.e., each $\mathbf{a}_p^{(i)}$ is randomly drawn from the type class $\mathbb{T}_{A_i}$ with probability $1/|\mathbb{T}_{A_i}|$, $p = 1, 2, ..., 2N_i$. Each set has $2N_i$ elements rather than N_i because an expurgation operation will be performed later. Also, we denote the set $\mathcal{C}_i^p \triangleq \mathcal{C}_i / \left\{\mathbf{a}_p^{(i)}\right\}$.

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

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

such that the set

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

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

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

and

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

Based on the above set-up, the following inequalities hold.

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

$$\begin{aligned}
& \mathbb{E}\left|\mathbb{T}_{V_{A'B'|AB}}\left(\mathbf{a}_p^{(i)}, \mathbf{b}_{p,q}^{(j)}\right) \bigcap \mathcal{C}_{kl}\right| \\
\leq \ & \mathbb{E}\left|\left\{(p',q') : \left(\mathbf{a}_{p'}^{(k)}, \mathbf{b}_{p',q'}^{(l)}\right) \in \mathbb{T}_{V_{A'B'|AB}}\left(\mathbf{a}_p^{(i)}, \mathbf{b}_{p,q}^{(j)}\right)\right\}\right| \\
= \ & 4N_k M_{kl} \Pr\left\{\left(\mathbf{a}_1^{(k)}, \mathbf{b}_{1,1}^{(l)}\right) \in \mathbb{T}_{V_{A'B'|AB}}\left(\mathbf{a}_p^{(i)}, \mathbf{b}_{p,q}^{(j)}\right)\right\} \\
= \ & 4N_k M_{kl} \frac{\left|\mathbb{T}_{V_{A'B'|AB}}\left(\mathbf{a}_p^{(i)}, \mathbf{b}_{p,q}^{(j)}\right)\right|}{|\mathbb{T}_{A_k B_l}|} \\
\leq \ & 4N_k M_{kl} (n+1)^{|\mathcal{A}||\mathcal{B}|} 2^{-nI_{P_{A_iB_j}V_{A'B'|AB}}(A',B';A,B)},
\end{aligned} \tag{87}$$

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

$$\widehat{P}_{k,l}\left(\mathbf{a}_{p'}^{(k)}, \mathbf{b}_{p',q'}^{(l)}\right) \triangleq \frac{1}{|\mathbb{T}_{A_kB_l}|} \quad \forall \quad 1 \leq k \leq m_n, \quad 1 \leq l \leq m'_{kn}, \quad 1 \leq p' \leq N_k, \quad 1 \leq q' \leq M_{kl}, \tag{88}$$

and (87) follows from the basic facts (Lemma 1) that

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

and that

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

noting that the marginal distribution of $P_{A_iB_j}V_{A'B'|AB}$ for RV's (A',B') is $P_{A_kB_l}$.

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

$$\mathbb{E}\left|\mathbb{T}_{V_{A'B'|AB}}\left(\mathbf{a}_p^{(i)}, \mathbf{b}_{p,q}^{(j)}\right) \bigcap \mathcal{C}_{ij}^{pq}\right| \leq 4N_i M_{ij} (n+1)^{|\mathcal{A}||\mathcal{B}|} 2^{-nI_{P_{A_iB_j}V_{A'B'|AB}}(A',B';A,B)}, \tag{89}$$

where the expectation is taken over the uniform distribution $\widehat{P}_{i,j}$ defined by (88).

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

$$\mathbb{E}\left|\mathbb{T}_{V_{A'B'|AB}}\left(\mathbf{a}_p^{(i)}, \mathbf{b}_{p,q}^{(j)}\right) \bigcap \mathcal{C}_{il}\right| \leq 4N_i M_{il} (n+1)^{|\mathcal{A}||\mathcal{B}|} 2^{-nI_{P_{A_iB_j}V_{AB'|AB}}(A,B';A,B)}.$$

Using the identity

$$I_{P_{A_iB_j}V_{AB'|AB}}(A,B';A,B) = H_{P_{A_i}}(A) + I_{P_{A_iB_j}V_{AB'|AB}}(B';B|A)$$

and assumption (1)

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

we obtain another bound

$$\mathbb{E}\left|\mathbb{T}_{V_{A'B'|AB}}\left(\mathbf{a}_p^{(i)}, \mathbf{b}_{p,q}^{(j)}\right) \bigcap \mathcal{C}_{il}\right| \leq 4M_{il} (n+1)^{|\mathcal{A}||\mathcal{B}|} 2^{-nI_{P_{A_iB_j}V_{A'B'|AB}}(B';B|A)}, \tag{90}$$

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

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

$$\mathbb{E}\left|\mathbb{T}_{V_{A'B'|AB}}\left(\mathbf{a}_p^{(i)}, \mathbf{b}_{p,q}^{(j)}\right) \bigcap \mathcal{C}_{ij}^{pq}\right| \leq 4M_{ij}(n+1)^{|\mathcal{A}||\mathcal{B}|}2^{-nI_{P_{A_i B_j} V_{A'B'|AB}}(B';B|A)}, \tag{91}$$

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

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

$$\begin{aligned}
\mathbb{E}\left|\mathbb{T}_{V_{A'|A}}\left(\mathbf{a}_p^{(i)}\right) \bigcap \mathcal{C}_k\right| &\leq \mathbb{E}\left|\left\{p' : \mathbf{a}_{p'}^{(k)} \in \mathbb{T}_{V_{A'|A}}\left(\mathbf{a}_p^{(i)}\right)\right\}\right| \\
&= 2N_k \Pr\left\{\mathbf{a}_1^{(i)} \in \mathbb{T}_{V_{A'|A}}\left(\mathbf{a}_p^{(i)}\right)\right\} \\
&= 2N_k \frac{\left|\mathbb{T}_{V_{A'|A}}\left(\mathbf{a}_p^{(i)}\right)\right|}{|\mathbb{T}_{A_k}|} \\
&\leq 2N_k(n+1)^{-|\mathcal{A}|}2^{-nI_{P_{A_i} V_{A'|A}}(A';A)},
\end{aligned} \tag{92}$$

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

$$\widetilde{P}_k(\mathbf{a}_{p'}^{(k)}) \triangleq \frac{1}{|\mathbb{T}_{A_k}|}, \quad \forall \quad 1 \leq k \leq m_n, \quad 1 \leq p' \leq N_k, \tag{93}$$

and (92) follows from the basic facts (Lemma 1) that

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

and that

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

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,

$$\mathbb{E}\left|\mathbb{T}_{V_{A'|A}}\left(\mathbf{a}_p^{(i)}\right) \bigcap \mathcal{C}_i^p\right| \leq 2N_k(n+1)^{-|\mathcal{A}|}2^{-nI_{P_{A_i} V_{A'|A}}(A';A)}, \tag{94}$$

where the expectation is taken over the uniform distribution $\widetilde{P}_i$ defined in (93).

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

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

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

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

Therefore, it follows from (87) and (89) that for any $V_{A'B'|AB} \in \mathcal{P}_n(\mathcal{A} \times \mathcal{B}|\mathcal{A} \times \mathcal{B})$,

$$\begin{aligned}
&\frac{\mathbb{E}\left|\mathbb{T}_{V_{A'B'|AB}}\left(\mathbf{a}_p^{(i)}, \mathbf{x}_{p,q}^{(j)}\right) \bigcap \mathcal{C}_{ij}^{pq}\right|}{4N_i M_{ij}} + \sum_{(k,l)\neq(i,j)} \frac{\mathbb{E}\left|\mathbb{T}_{V_{A'B'|AB}}\left(\mathbf{a}_p^{(i)}, \mathbf{b}_{p,q}^{(j)}\right) \bigcap \mathcal{C}_{kl}\right|}{4N_k M_{kl}} \\
\leq\ & m_n(\max_i m'_{in})(n+1)^{|\mathcal{A}||\mathcal{B}|}2^{-nI_{P_{A_i B_j} V_{A'B'|AB}}(A',B';A,B)}.
\end{aligned} \tag{95}$$

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)

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

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

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

where

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

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

$$\mathbb{E} \sum_{i=1}^{m_n} \sum_{j=1}^{m'_{in}} \frac{1}{4N_iM_{ij}} \sum_{p=1}^{2N_i} \sum_{q=1}^{2M_{ij}} S_{ij}^{pq} \leq (n+1)^{2|\mathcal{A}|^2|\mathcal{B}|^2} m_n^2 (\max_i m'_{in})^2. \tag{96}$$

Similarly, it follows from (90) and (91) that

$$\mathbb{E} \sum_{i=1}^{m_n} \sum_{j=1}^{m'_{in}} \frac{1}{4N_iM_{ij}} \sum_{p=1}^{2N_i} \sum_{q=1}^{2M_{ij}} K_{ij}^{pq} \leq (n+1)^{2|\mathcal{A}|^2|\mathcal{B}|^2} m_n (\max_i m'_{in})^2 \leq (n+1)^{2|\mathcal{A}|^2|\mathcal{B}|^2} m_n^2 (\max_i m'_{in})^2, \tag{97}$$

where

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

and it follows from (92) and (94) that

$$\mathbb{E} \sum_{i=1}^{m_n} \sum_{j=1}^{m'_{in}} \frac{1}{4N_iM_{ij}} \sum_{p=1}^{2N_i} \sum_{q=1}^{2M_{ij}} L_{ij}^{pq} \leq (n+1)^{2|\mathcal{A}|^2} m_n^2 (\max_i m'_{in}) \leq (n+1)^{2|\mathcal{A}|^2|\mathcal{B}|^2} m_n (\max_i m'_{in})^2, \tag{98}$$

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

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

Summing (96), (97) and (98) together, we obtain

$$\mathbb{E}\sum_{i=1}^{m_n}\sum_{j=1}^{m'_{in}}\frac{1}{4N_iM_{ij}}\sum_{p=1}^{2N_i}\sum_{q=1}^{2M_{ij}}\left(S_{ij}^{pq}+K_{ij}^{pq}+L_{ij}^{pq}\right)\leq 3(n+1)^{2|\mathcal{A}|^2|\mathcal{B}|^2}m_n^2(\max_i m'_{in})^2. \tag{99}$$

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

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

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

$$\frac{1}{4N_iM_{ij}}\sum_{p=1}^{2N_i}\sum_{q=1}^{2M_{ij}}\left(S_{ij}^{pq}+K_{ij}^{pq}+L_{ij}^{pq}\right)\leq 3(n+1)^{2|\mathcal{A}|^2|\mathcal{B}|^2}m_n^2(\max_i m'_{in})^2. \tag{100}$$

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

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

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

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

Similarly, suppose for each $p=1,2,...,N_i$,

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

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

$$S_{ij}^{pq}+K_{ij}^{pq}+L_{ij}^{pq}\leq 12(n+1)^{2|\mathcal{A}|^2|\mathcal{B}|^2}m_n^2(\max_i m'_{in})^2. \tag{101}$$

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

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

and denote also $\Omega_{ij}^{pq} \triangleq \Omega_{ij}/\left\{\left(\mathbf{a}_p^{(i)},\mathbf{b}_{p,q}^{(j)}\right)\right\} \subseteq \widehat{\mathcal{C}}_{ij}^{pq}$. Immediately, it follows from (101) that for every $i = 1,2,...,m_n,\ j = 1,2,...,m'_{in},\ k = 1,2,...,m_n,\ l = 1,2,...,m'_{kn},\ p = 1,2,...,N_i,\ q = 1,2,...,M_{ij}$, and every $V_{A'B'|AB} \in \mathcal{P}_n(\mathcal{A}\times\mathcal{B}|\mathcal{A}\times\mathcal{B})$ and $V_{A'|A} \in \mathcal{P}_n(\mathcal{A}|\mathcal{A})$

$$\frac{\left|\mathbb{T}_{V_{A'B'|AB}}\left(\mathbf{a}_p^{(i)},\mathbf{b}_{p,q}^{(j)}\right)\bigcap\Omega_{kl}\right|}{N_kM_{kl}} \leq 2^{-n\left[I_{P_{A_iB_j}V_{A'B'|AB}}(A',B';A,B)-\delta\right]}, \quad (k,l)\neq(i,j), \tag{102}$$

$$\frac{\left|\mathbb{T}_{V_{A'B'|AB}}\left(\mathbf{a}_p^{(i)},\mathbf{b}_{p,q}^{(j)}\right)\bigcap\Omega_{ij}^{pq}\right|}{N_iM_{ij}} \leq 2^{-n\left[I_{P_{A_iB_j}V_{A'B'|AB}}(A',B';A,B)-\delta\right]}, \tag{103}$$

$$\frac{\left|\mathbb{T}_{V_{A'B'|AB}}\left(\mathbf{a}_p^{(i)},\mathbf{b}_{p,q}^{(j)}\right)\bigcap\Omega_{il}\right|}{M_{il}} \leq 2^{-n\left[I_{P_{A_iB_j}V_{A'B'|AB}}(B';B|A)-\delta\right]}, \quad l\neq j, \tag{104}$$

$$\frac{\left|\mathbb{T}_{V_{A'B'|AB}}\left(\mathbf{a}_p^{(i)},\mathbf{b}_{p,q}^{(j)}\right)\bigcap\Omega_{ij}^{pq}\right|}{M_{ij}} \leq 2^{-n\left[I_{P_{A_iB_j}V_{A'B'|AB}}(B';B|A)-\delta\right]}, \tag{105}$$

$$\frac{\left|\mathbb{T}_{V_{A'|A}}\left(\mathbf{a}_p^{(i)}\right)\bigcap\Omega_{k}\right|}{N_k} \leq 2^{-n\left[I_{P_{A_i}V_{A'|A}}(A';A)-\delta\right]}, \quad k\neq i, \tag{106}$$

$$\frac{\left|\mathbb{T}_{V_{A'|A}}\left(\mathbf{a}_p^{(i)}\right)\bigcap\Omega_{i}^{p}\right|}{N_i} \leq 2^{-n\left[I_{P_{A_i}V_{A'|A}}(A';A)-\delta\right]}, \tag{107}$$

where

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

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 (102)–(107) for any $V_{A'|A}$ and $V_{A'B'|AB}$. It remains to show that these sets are disjoint and have distinct elements provided assumptions (1) and (2). Indeed, since (106) and (107) hold for every $V_{A'|A} \in \mathcal{P}_n(\mathcal{A}|\mathcal{A})$, they of course hold when $V_{A'|A}$ is a conditional distribution such that $V^*_{A'|A}(a'|a)$ is 1 if $a'=a$ and 0 otherwise. It then follows from (1)

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

that $\left|\mathbb{T}_{V^*_{A'|A}}\left(\mathbf{a}_p^{(i)}\right)\bigcap\Omega_k\right| = \left|\left\{\mathbf{a}_p^{(i)}\right\}\bigcap\Omega_k\right| < 1$ or equivalently, $\left|\left\{\mathbf{a}_p^{(i)}\right\}\bigcap\Omega_k\right| = 0$, 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 (1) in (107), we see that

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

which means that Ω_i has N_i disjoint elements. Similarly, setting $V_{A'B'|AB}$ be the conditional distribution such that $V^*_{A'B'|AB}(a',b'|a,b)$ is 1 if $a'=a$, $b'=b$ and 0 otherwise, and using (2)

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

we see that for any $\mathbf{a}_p^{(i)} \in \Omega_i$, $\Omega_{ij}(\mathbf{a}_p^{(i)})$'s are disjoint and the elements in $\Omega_{ij}(\mathbf{a}_p^{(i)})$ are all distinct, i.e., $|\Omega_{ij}(\mathbf{a}_p^{(i)})| = M_{ij}$ for every $\mathbf{a}^{(i)} \in \Omega_i$. Finally, when $V_{A'|A}$ is not the conditional distribution such that $V_{A'|A}(a'|a)$ is 1 if $a' = a$ and 0 otherwise, we can write (106) and (107) in the same way as (3), and when $V_{A'B'|AB}$ is not the conditional distribution such that $V_{A'B'|AB}(a', b'|a, b)$ is 1 if $a' = a$, $b' = b$ and 0 otherwise, we can write (102)–(103) as (4), and write (104)–(105) as (5), since

$$\begin{aligned}
\left|\mathbb{T}_{V_{A'|A}}\left(\mathbf{a}_p^{(i)}\right)\bigcap\Omega_i^p\right| &= \left|\mathbb{T}_{V_{A'|A}}\left(\mathbf{a}_p^{(i)}\right)\bigcap\Omega_i\right|,\\
\left|\mathbb{T}_{V_{A'B'|AB}}\left(\mathbf{a}_p^{(i)}, \mathbf{b}_{p,q}^{(j)}\right)\bigcap\Omega_{ij}^{pq}\right| &= \left|\mathbb{T}_{V_{A'B'|AB}}\left(\mathbf{a}_p^{(i)}, \mathbf{b}_{p,q}^{(j)}\right)\bigcap\Omega_{ij}\right|,\\
\left|\mathbb{T}_{V_{A'B'|AB}}\left(\mathbf{a}_p^{(i)}, \mathbf{b}_{p,q}^{(j)}\right)\bigcap\Omega_{ij}^{pq}\right| &= \left|\mathbb{T}_{V_{A'B'|AB}}\left(\mathbf{a}_p^{(i)}, \mathbf{b}_{p,q}^{(j)}\right)\bigcap\Omega_{ij}\right|.
\end{aligned}$$

■

B Proof of (23) and (24)

B.1 Upper Bound on $\left|\mathbb{T}_{\widehat{V}_{Y|TUX}}((\mathbf{t},\mathbf{u}),\mathbf{x})\bigcap\mathcal{E}_1\right|$

If we fix a $k = 1, 2, ..., m_n$ and a $l = 1, 2, ..., m'_{kn}$, then $\mathcal{E}_1$ is the set of all $\mathbf{y}$ such that there exist some $((\mathbf{t},\mathbf{u})', \mathbf{x}') \in \Omega_{kl}$, $(\mathbf{t},\mathbf{u})' \neq (\mathbf{t},\mathbf{u})$, $((\mathbf{t},\mathbf{u}), \mathbf{x}, (\mathbf{t},\mathbf{u})', \mathbf{x}', \mathbf{y})$ admits a joint type $P_{(\mathbf{t},\mathbf{u})\mathbf{x}(\mathbf{t},\mathbf{u})'\mathbf{x}'\mathbf{y}} \in \mathcal{P}_n(\mathcal{T}^2 \times \mathcal{U}^2 \times \mathcal{X}^2 \times \mathcal{Y})$ and

$$I((\mathbf{t},\mathbf{u})', \mathbf{x}'; \mathbf{y}) - (R_k + R_{kl}) \geq I((\mathbf{t},\mathbf{u}), \mathbf{x}; \mathbf{y}) - (R_i + R_{ij}). \tag{108}$$

Note that (108) can be represented as for dummy R.V.'s $(TU) \in \mathcal{T} \times \mathcal{U}$, $X \in \mathcal{X}$, $(TU)' \in \mathcal{T} \times \mathcal{U}$, $X' \in \mathcal{X}$, and $Y \in \mathcal{Y}$, the following holds under the joint distribution $P_{(TU)X(TU)'X'Y} = P_{(\mathbf{t},\mathbf{u})\mathbf{x}(\mathbf{t},\mathbf{u})'\mathbf{x}'\mathbf{y}}$,

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

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

$$\mathbb{T}_{\widehat{V}_{Y|(TU)X}}((\mathbf{t},\mathbf{u}),\mathbf{x})\bigcap\mathcal{E}_1 = \bigcup_{k=1}^{m_n}\bigcup_{l=1}^{m'_{kn}}\bigcup_{P_{(TU)X(TU)'X'Y}\in\mathcal{C}_{k,l}((\mathbf{t},\mathbf{u}),\mathbf{x})}\mathcal{F}_{k,l}((\mathbf{t},\mathbf{u}),\mathbf{x},P_{(TU)X(TU)'X'Y}) \tag{109}$$

where

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

where $P_{(TU)X}$, $P_{(TU)'X'}$ and $P_{Y|(TU)X}$, etc, are the corresponding marginal and conditional distributions induced by $P_{(TU)X(TU)'X'Y}$, and

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

where $\mathbb{T}_{(TU)X(TU)'X'Y} \triangleq \mathbb{T}_{P_{(TU)X(TU)'X'Y}}$. Clearly, given any k, l, and $P_{(TU)X(TU)'X'Y}$,

$$\begin{aligned} &\left|\mathcal{F}_{k,l}((\mathbf{t},\mathbf{u}),\mathbf{x},P_{(TU)X(TU)'X'Y})\right| \\ \leq\;& \left| \left\{ ((\mathbf{t},\mathbf{u})',\mathbf{x}',\mathbf{y}) : \begin{array}{l} ((\mathbf{t},\mathbf{u}),\mathbf{x},(\mathbf{t},\mathbf{u})',\mathbf{x}',\mathbf{y}) \in \mathbb{T}_{(TU)X(TU)'X'Y} \\ ((\mathbf{t},\mathbf{u})',\mathbf{x}') \in \Omega_{kl}, \quad (\mathbf{t},\mathbf{u})' \neq (\mathbf{t},\mathbf{u}) \end{array} \right\} \right| \\ =\;& \left| \left\{ ((\mathbf{t},\mathbf{u})',\mathbf{x}') : \begin{array}{l} ((\mathbf{t},\mathbf{u}),\mathbf{x},(\mathbf{t},\mathbf{u})',\mathbf{x}') \in \mathbb{T}_{(TU)X(TU)'X'} \\ ((\mathbf{t},\mathbf{u})',\mathbf{x}') \in \Omega_{kl}, \quad (\mathbf{t},\mathbf{u})' \neq (\mathbf{t},\mathbf{u}) \end{array} \right\} \right| \times \left|\mathbb{T}_{Y|(TU)X(TU)'X'}((\mathbf{t},\mathbf{u}),\mathbf{x},(\mathbf{t},\mathbf{u})',\mathbf{x}')\right| \\ \leq\;& N_k M_{kl} 2^{-n\left[I_{P_{(TU)X(TU)'X'}}((T,U),X;(T,U)',X')-\eta\right]} \times 2^{nH_{P_{(TU)X(TU)'X'Y}}(Y|(T,U),X,(T,U)',X')}, \end{aligned} \tag{110}$$

where the last inequality follows from Lemma 2. Meanwhile, when $((\mathbf{t},\mathbf{u}),\mathbf{x}) \in \Omega_{ij}$, the following simple bound also holds

$$\left|\mathcal{F}_{k,l}((\mathbf{t},\mathbf{u}),\mathbf{x},P_{(TU)X(TU)'X'Y})\right| \leq \left|\mathbb{T}_{Y|(TU)X}((\mathbf{t},\mathbf{u}),\mathbf{x})\right| \leq 2^{nH_{P_{(TU)XY}}(Y|(T,U),X)} = 2^{nH_{P_{((TU))_iX_j}\widehat{V}_{Y|(TU)X}}(Y|(T,U),X)} \tag{111}$$

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

$$\begin{aligned} & H_{P_{(TU)X(TU)'X'Y}}(Y|(T,U),X,(T,U)',X') - I_{P_{(TU)X(TU)'X'}}((T,U),X;(T,U)',X') \\ =\;& H_{P_{(TU)XY}}(Y|(T,U),X) - I_{P_{(TU)X(TU)'X'Y}}((T,U)',X';(T,U),X,Y) \\ \leq\;& H_{P_{(TU)XY}}(Y|(T,U),X) - I_{P_{(TU)'X'Y}}((T,U)',X';Y) \end{aligned} \tag{112}$$

into (110), combining with (111) together, we obtain

$$\left|\mathcal{F}_{k,l}((\mathbf{t},\mathbf{u}),\mathbf{x},P_{(TU)X(TU)'X'Y})\right| \leq 2^{n\left[H_{P_{((TU))_iX_j}\widehat{V}_{Y|(TU)X}}(Y|(T,U),X) - \left|I_{P_{(TU)'X'Y}}((T,U)',X';Y)-(R_k+R_{kl})\right|^+\right]}. \tag{113}$$

Again recall that for $P_{(TU)X(TU)'X'Y} \in \mathcal{C}_{k,l}((\mathbf{t},\mathbf{u}),\mathbf{x})$, $P_{(TU)XY} = P_{((TU))_iX_j}\widehat{V}_{Y|(TU)X}$, and note that

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

This implies when $P_{(TU)X(TU)'X'Y} \in \mathcal{C}_{k,l}((\mathbf{t},\mathbf{u}),\mathbf{x})$

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

and hence

$$\left|\mathbb{T}_{\widehat{V}_{Y|(TU)X}}((\mathbf{t},\mathbf{u}),\mathbf{x})\bigcap\mathcal{E}_1\right| \leq m_n\left(\max_i m'_{in}\right)(n+1)^{|\mathcal{T}\times\mathcal{U}|^2|\mathcal{X}|^2|\mathcal{Y}|}$$
$$\times 2^{n\left[H_{P_{((TU))_iX_j}\widehat{V}_{Y|(TU)X}}(Y|(T,U),X)-\left|(I_{P_{((TU))_iX_j}\widehat{V}_{Y|(TU)X}}((T,U),X;Y)-(R_i+R_{ij}))\right|^+\right]},$$

since by Lemma 1

$$|\mathcal{C}_{k,l}((\mathbf{t},\mathbf{u}),\mathbf{x})| \leq \left|\mathcal{P}_n(\mathcal{T}^2\times\mathcal{U}^2\times\mathcal{X}^2\times\mathcal{Y})\right| \leq (n+1)^{|\mathcal{T}|^2|\mathcal{U}|^2|\mathcal{X}|^2|\mathcal{Y}|}.$$

B.2 Upper Bound on $\left|\mathbb{T}_{\widehat{V}_{Y|(TU)X}}((\mathbf{t},\mathbf{u}),\mathbf{x})\bigcap\mathcal{E}_2\right|$

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

$$I((\mathbf{t},\mathbf{u}),\mathbf{x}';\mathbf{y})-(R_i+R_{il}) \geq I((\mathbf{t},\mathbf{u}),\mathbf{x};\mathbf{y})-(R_i+R_{ij}). \tag{114}$$

Using the identity

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

on both sides of (114) we see it is equivalent to

$$I(\mathbf{x}';\mathbf{y}|\mathbf{t},\mathbf{u})-R_{il} \geq I(\mathbf{x};\mathbf{y}|\mathbf{t},\mathbf{u})-R_{ij}. \tag{115}$$

Note that (115) can be represented as for dummy R.V.'s $(TU)\in\mathcal{T}\times\mathcal{U}$, $X\in\mathcal{X}$, $X'\in\mathcal{X}$, and $Y\in\mathcal{Y}$, the following holds under the joint distribution $P_{(TU)XX'Y}=P_{(\mathbf{t},\mathbf{u})\mathbf{x}\mathbf{x}'\mathbf{y}}$,

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

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

$$\mathbb{T}_{\widehat{V}_{Y|(TU)X}}((\mathbf{t},\mathbf{u}),\mathbf{x})\bigcap\mathcal{E}_2=\bigcup_{l=1}^{m'_{in}}\bigcup_{P_{(TU)XX'Y}\in\mathcal{C}_l((\mathbf{t},\mathbf{u}),\mathbf{x})}\mathcal{F}_l((\mathbf{t},\mathbf{u}),\mathbf{x},P_{(TU)XX'Y}) \tag{116}$$

where

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

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

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

where $\mathbb{T}_{(TU)XX'Y} = \mathbb{T}_{P_{(TU)XX'Y}}$. Using a similar counting argument, and applying Lemma 2, we can bound, for any $l = 1, 2, ..., m'_{in}$ and $P_{(TU)XX'Y} \in \mathcal{C}_l((\mathbf{t},\mathbf{u}),\mathbf{x})$,

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

and finally, we obtain,

$$\begin{aligned} \left|\mathbb{T}_{\widehat{V}_{Y|(TU)X}}((\mathbf{t},\mathbf{u}),\mathbf{x}) \bigcap \mathcal{E}_2\right| &\leq \left(\max_i m'_{in}\right) (n+1)^{|\mathcal{T}\times\mathcal{U}||\mathcal{X}|^2|\mathcal{Y}|} \\ &\times 2^{n\left[H_{P_{((TU))_i X_j}\widehat{V}_{Y|(TU)X}}(Y|(T,U),X) - \left|I_{P_{((TU))_i X_j}\widehat{V}_{Y|(TU)X}}(X;Y|T,U) - R_{ij}\right|^+\right]} \end{aligned}$$

since $|\mathcal{C}_l((\mathbf{t},\mathbf{u}),\mathbf{x})| \leq (n+1)^{|\mathcal{T}||\mathcal{U}||\mathcal{X}|^2|\mathcal{Y}|}$. ■

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. Inform. Theory*, vol. 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, April–June 1975.

[3] P. P. Bergmans, "Random coding theorem for broadcast channels with degraded components," *IEEE Trans. Inform. Theory*, vol. 19, 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. Inform. Theory*, vol. 33, pp. 716–718, Sep. 1987.

[5] S. Choi and S. S. Pradhan, "A graph-based framework for transmission of correlated sources over broadcast channels," submitted to *IEEE Trans. Inform. Theory*, Aug. 2006.

[6] T. M. Cover, "Broadcast channels," *IEEE Trans. Inform. Theory*, vol. 18, pp. 2–14, Jan. 1972.

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

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

[9] I. Csiszár, "On the error exponent of source-channel transmission with a distortion threshold," *IEEE Trans. Inform. Theory*, vol. 28, pp. 823–828, Nov. 1982.

[10] I. Csiszár, "The method of types," *IEEE Trans. Inform. Theory*, vol. 44, pp. 2505–2522, Nov. 1998.

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

[12] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inform. Theory*, vol. 24, pp. 339–348, May 1978.

[13] G. Dueck, "A note on the multiple access channel with correlated sources," *IEEE Trans. Inform. Theory*, vol. 27, no. 2, Mar. 1981.

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

[15] A. A. EL Gamal, "The capacity of a class of broadcast channels," *IEEE Trans. Inform. Theory*, vol. 25, pp. 166–169, March. 1979.

[16] T. S. Han and M. H. M. Costa, "Broadcast channels with arbitrary correlated sources," *IEEE Trans. Inform. Theory*, vol. 33, no. 5, pp. 641–650, Sep. 1987.

[17] 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.

[18] W. Kang and S. Ulukus, "An outer bound for multiple access channels with correlated sources," *Proceedings of the 2006 IEEE Int'l. Symp. Inform. Theory*, Seattle, USA, July 2006.

[19] J. Körner and K. Marton, "Genreal broadcast channels with degraded message sets," *IEEE Trans. Inform. Theory*, vol. 23, pp. 60–64, Jan. 1977.

[20] J. Körner and A. Sgarro, "Universally attainable error exponents for broadcast channels with degraded message sets," *IEEE Trans. Inform. Theory*, vol. 26, pp. 670–679, Nov. 1980.

[21] Y. S. Liu and B. L. Hughes, "A new universal random coding bound for the multiple-access channel," *IEEE Tran. Inform. Theory*, vol. 42, pp. 376–386, Mar. 1996.

[22] S. S. Pradhan, S. Choi and K. Ramchandran, “A graph-based framework for transmission of correlated sources over multiple access channels,” submitted to *IEEE Trans. Inform. Theory*, Jan. 2006.

[23] Y. Zhong, F. Alajaji, and L. L. Campbell, “On the joint source-channel coding error exponent for discrete memoryless systems,” *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1450–1468, April 2006.

[24] Y. Zhong, F. Alajaji, and L. L. Campbell, “On the joint source-channel coding error exponent for systems with memory,” in Proceedings *2005 IEEE Int’l. Symp. Inform. Theory*, Adelaide, Australia, pp. 1386–1390. Sept. 2005.

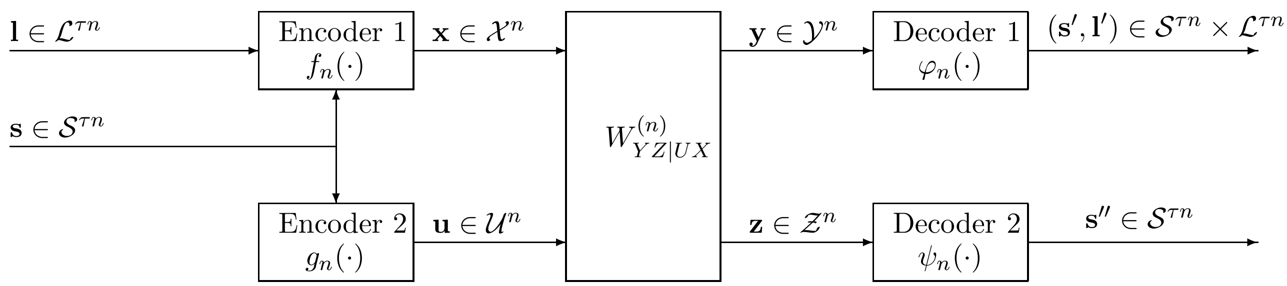

Figure 1: Transmitting two CS over the asymmetric 2-user communication channel.

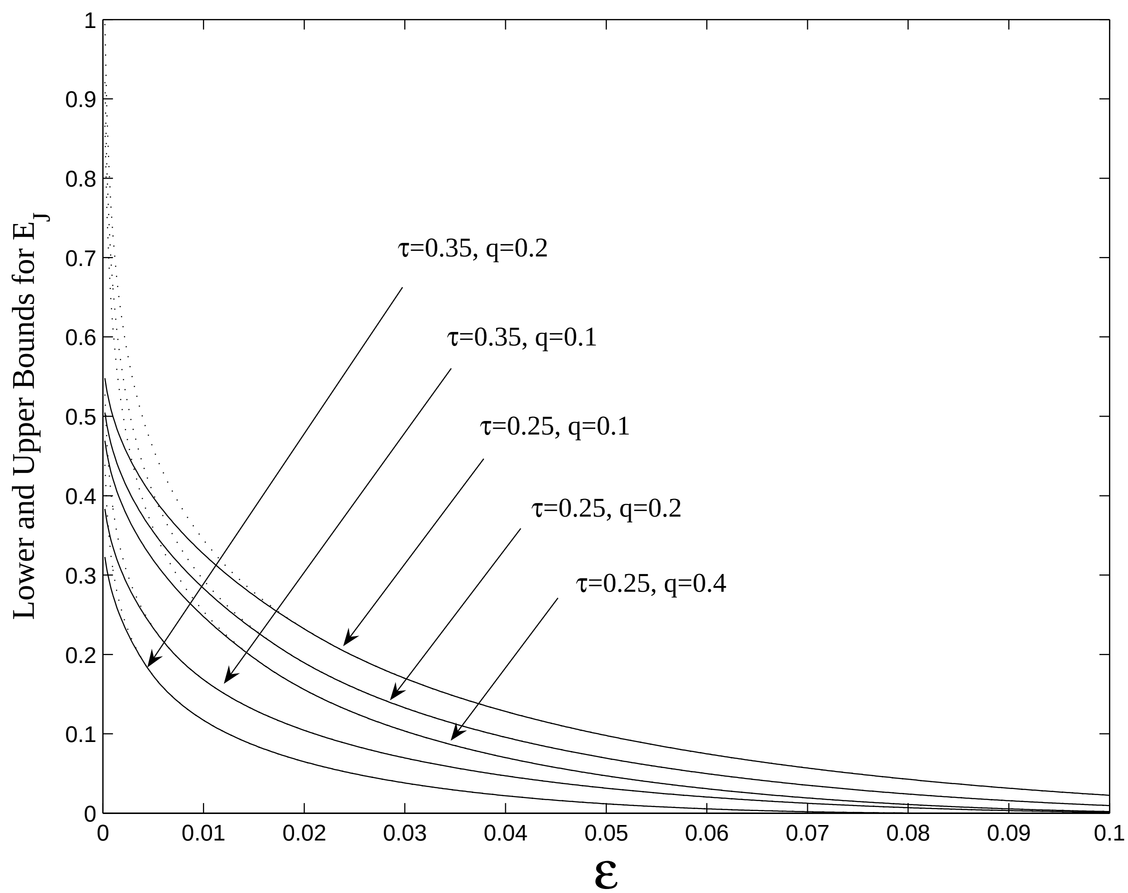

Figure 2: The lower bound (solid line) and the upper bound (dash line) for the system JSCC error exponent for transmitting binary CS over the binary AMAC with binary additive noise.