Hypothesis Testing over the Two-Hop Relay Network

Sadaf Salehkalaibar, *IEEE Member*, Michèle Wigger, *IEEE Senior Member*, Ligong Wang, *IEEE Member*

***Abstract*—Coding and testing schemes and the corresponding achievable type-II error exponents are presented for binary hypothesis testing over two-hop relay networks. The schemes are based on cascade source coding techniques and unanimous decision-forwarding, the latter meaning that a terminal decides on the null hypothesis only if all previous terminals have decided on the null hypothesis. If the observations at the transmitter, the relay, and the receiver form a Markov chain in this order, then, without loss in performance, the proposed cascade source code can be replaced by two independent point-to-point source codes, one for each hop. The decoupled scheme (combined with decision-forwarding) is shown to attain the optimal type-II error exponents for various instances of "testing against conditional independence." The same decoupling is shown to be optimal also for some instances of "testing against independence," when the observations at the transmitter, the receiver, and the relay form a Markov chain in this order, and when the relay-to-receiver link is of sufficiently high rate. For completeness, the paper also presents an analysis of the Shimokawa-Han-Amari binning scheme for the point-to-point hypothesis testing setup.**

I. Introduction

As part of the Internet of Things (IoT), sensor applications are rapidly increasing, thanks to lower cost and better performance of sensors. One of the major theoretical challenges in this respect is sensor networks with multiple sensors collecting correlated data, which they communicate to one or multiple decision centers. Of special practical and theoretical interest is to study the tradeoff between the quality of the decisions taken at the centers and the required communication resources. In this work, following the approach in [1], [2], we consider problems where decision centers have to decide on a binary hypothesis $\mathcal{H} = 0$ or $\mathcal{H} = 1$ that determines the underlying joint probability mass function (pmf) of all the terminals' observations. Our goal is to characterize the set of possible *type-II error exponents* (i.e., the error exponent in deciding $\hat{\mathcal{H}} = 0$ when in fact $\mathcal{H} = 1$) as a function of the available communication rates such that the *type-I error probabilities* (i.e., error probabilities of deciding $\hat{\mathcal{H}} = 1$ when in fact $\mathcal{H} = 0$) vanish as the lengths of the observations grow. Previous works on this *exponent-rate region* considered communication scenarios over dedicated noise-free links from one or many transmitters to a single decision center [1], [3], [4] or from a single transmitter to two decision centers [5]–[7]. The hypothesis testing problem from a signal processing perspective has been studied in several works [8]–[12]. Recently, simple interactive communication scenarios were also considered [13]–[15], as well as hypothesis testing over *noisy* communication channels [5], [16], [17]. All these distributed hypothesis testing problems are open in the general case; exact solutions have only been found for instances of "testing against independence" [1] and of "testing against conditional independence" [4]. "Testing against independence" refers to a scenario where the observations' joint pmf under $\mathcal{H} = 1$ is the product of the marginal pmfs under $\mathcal{H} = 0$, and "testing against conditional independence" refers to a scenario where this independence holds only conditional on some subsequence that is observed at the receiver and that has the same joint distribution with the sensor's observations under both hypotheses.

The focus of this paper is on the *two-hop network* depicted in Fig. 1. We model a situation with three sensors and two decision centers. The first terminal (the transmitter) models a simple sensor that observes an n-length sequence X^n. The second terminal (the relay) includes both a sensor observing the n-length sequence Y^n and a decision center which produces the guess $\hat{\mathcal{H}}_y \in \{0, 1\}$. Similarly, the third terminal (the receiver) includes a sensor observing Z^n and a decision center producing the guess $\hat{\mathcal{H}}_z \in \{0, 1\}$. Communication is directed and in two stages. The transmitter communicates directly with the relay over a noise-free link of rate $R > 0$, but it cannot directly communicate with the receiver, e.g., because the receiver is too far away. Such a restriction is particularly relevant for modern IoT applications where sensors are desired to consume very little energy so as to last for decades without the batteries being replaced. On the other hand, the receiver is assumed to be sufficiently close to the relay so that the relay can communicate directly with it over a noise-free link of rate $T > 0$. The task of the relay is not only to communicate information about its own observation to the receiver but also to *process and forward* information that it receives from the transmitter. Two-hop networks have previously been studied in information theory for source coding or coordination. These problems are open in general. Solutions to special cases were presented in [18]–[25].

In this paper, we propose two coding and testing schemes for binary hypothesis testing over the two-hop relay network. The two schemes apply two different source coding schemes for the two-hop relay network to convey quantization information about the distributed observations to the relay and the receiver, and combine these schemes with a *unanimous*

S. Salehkalaibar is with the Department of Electrical and Computer Engineering, College of Engineering, University of Tehran, Tehran, Iran, s.saleh@ut.ac.ir,

M. Wigger is with LTCI, Telecom ParisTech, 75013 Paris, michele.wigger@telecom-paristech.fr,

L. Wang is with ETIS, Université Paris Seine, Université de Cergy-Pontoise, ENSEA, CNRS, ligong.wang@ensea.fr.

Parts of the material in this paper were presented at the *IEEE Information Theory Workshop (ITW), Kaohsiung, Taiwan, November 2017.*

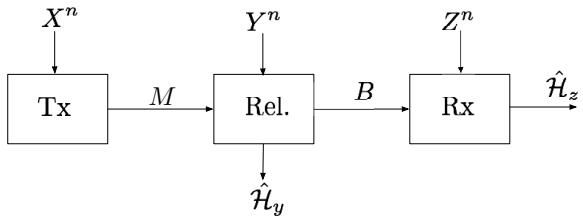

Fig. 1. Hypothesis testing over a single-relay multi-hop channel

decision-forwarding strategy. In this latter strategy, each of the terminals tests whether its reconstructed source sequences are jointly typical with its own observation under the null hypothesis $\mathcal{H} = 0$. If the test is positive and the preceding terminals have also decided on $\hat{\mathcal{H}} = 0$, then the terminal declares this null hypothesis $\hat{\mathcal{H}} = 0$. Otherwise it declares the alternative hypothesis $\hat{\mathcal{H}} = 1$. In both cases, it forwards its decision to the next terminal.

We characterize the relay and the receiver type-II error exponents achieved by our schemes. Our first scheme employs source coding without binning, which allows for a relatively simple characterization of the achieved exponents. Our second scheme employs source coding with binning and achieves larger exponents in some cases. However, with binning, the error exponent for $\hat{\mathcal{H}}_y$ is characterized by two competing exponents and the exponent for $\hat{\mathcal{H}}_z$ by four competing exponents. They are thus more complicated to evaluate.

In the second part of the manuscript, we focus on two cases: the first is where $X^n \to Y^n \to Z^n$ forms a Markov chain under both hypotheses, and the second is where $X^n \to Z^n \to Y^n$ forms a Markov chain under both hypotheses. The first case models an extreme situation where the relay lies in between the transmitter and the receiver, and thus the signals at the sensor and the receiver are conditionally independent given the signal at the relay. In such a situation, the two-hop network models, for example, short-range wireless communication where the sensor's signal only reaches the relay but not the more distant receiver. The second case models a situation where the receiver lies in between the sensor and the relay, and thus the signals at the transmitter and the relay are conditionally independent given the signal at the receiver. In such a situation, the two-hop network models, for example, communication in a cellular system where the relay is a powerful base station and all communication goes through this base station.

We show that, in the first case where $X^n \to Y^n \to Z^n$, our schemes simplify considerably in the sense that the source coding scheme for the two-hop relay network decouples into two independent point-to-point source coding schemes. In other words, it suffices to send quantization information about X^n from the transmitter to the relay and, independently thereof, to send quantization information about Y^n from the relay to the receiver (while also employing unanimous decision-forwarding) This contrasts the general scheme where the relay combines the quantization information about X^n with its own observation Y^n to create some kind of jointly processed quantization information to send to the receiver. The receiver error exponent achieved by the simplified scheme equals the sum of the exponent at the relay and the exponent achieved over the point-to-point link from the relay to the receiver, but, to compute the second exponent, we modify the pmf of the relay's observation under $\mathcal{H} = 1$ to being the same as its pmf under $\mathcal{H} = 0$. These simplified expressions are proved to be optimal in different special cases of testing against independence (achieved without binning) and testing against conditional independence (with binning). The focus of this paper is on *weak converses* where the type-I errors are also required to vanish asymptotically as $n \to \infty$. The existence of a *strong converse* for one of these special cases, i.e., a proof that the same exponents are optimal also when type-I error probability $\epsilon > 0$ is tolerated, was recently proved in [26].

For the second case where $X^n \to Z^n \to Y^n$, we present optimality results (in the weak converse sense) for two special cases. In the first special case, P_{YZ} is same under both hypothesis, so Y^n by itself is of no interest to the receiver. For rates $T \geq R$, the optimal strategy is for the relay to ignore its own observation and simply forward the transmitter's message to the receiver. Interestingly, this simple forwarding strategy can become suboptimal when $T < R$, because then the relay can act as a "coordinator" to reduce the communication rate T to the receiver. We present an example where the relay's own observation Y^n allows the relay to extract the relevant portion of X^n, and thus to reduce the required rate to the receiver T. In the second special case, P_{XZ} is same under both hypothesis, and for sufficiently large T the optimal strategy for the relay is to ignore all communication from the transmitter. Again, using an example, we show that for small T the transmitter can be useful by playing the role of a coordinator who reveals to the relay which portions of Y^n are relevant to the receiver.

Lastly, as a side-result, we also present a detailed analysis of the Shimokawa-Han-Amari [3] coding and testing scheme with binning for the point-to-point hypothesis testing problem. Previously this analysis has only appeared in Japanese [27].

We conclude this introduction with remarks on notation and an outline of the paper.

A. Notation

We mostly follow the notation in [28]. Random variables are denoted by capital letters, e.g., X, Y, and their realizations by lower-case letters, e.g., x, y. Script symbols such as $\mathcal{X}$ and $\mathcal{Y}$ stand for alphabets of random variables, and $\mathcal{X}^n$ and $\mathcal{Y}^n$ for the corresponding n-fold Cartesian products. Sequences of random variables $(X_i, ..., X_j)$ and realizations $(x_i, \ldots, x_j)$ are abbreviated by X_i^j and x_i^j. When $i = 1$, then we also use X^j and x^j instead of X_1^j and x_1^j.

Generally, we write the probability mass function (pmf) of a discrete random variable X as P_X; but we use Q_X to indicate the pmf under hypothesis $\mathcal{H} = 1$ when it is different from the pmf under $\mathcal{H} = 0$. The conditional pmf of X given Y is written as $P_{X|Y}$ (or as $Q_{X|Y}$ when $\mathcal{H} = 1$). The distributions of X^n, Y^n and (X^n, Y^n) under the same hypothesis are denoted by P_{X^n}, P_{Y^n} and $P_{X^nY^n}$, respectively. The notation

P_{XY}^n denotes the n-fold product distribution, i.e., for every $(x^n, y^n) \in \mathcal{X}^n \times \mathcal{Y}^n$, we have:

$$P_{XY}^n(x^n, y^n) = \prod_{i=1}^{n} P_{XY}(x_i, y_i). \tag{1}$$

The term $D(P\|Q)$ stands for the Kullback-Leibler (KL) divergence between two pmfs P and Q over the same alphabet. We use $\mathrm{tp}(\cdot)$ to denote the *joint type* of a tuple. For a joint type π_{ABC} over alphabet $\mathcal{A} \times \mathcal{B} \times \mathcal{C}$, we denote by $I_{\pi_{ABC}}(A; B|C)$ the mutual information assuming that the random triple (A, B, C) has pmf π_{ABC}; similarly for the entropy $H_{\pi_{ABC}}(A)$ and the conditional entropy $H_{\pi_{ABC}}(A|B)$. Sometimes we abbreviate π_{ABC} by π. Also, when π_{ABC} has been defined and is clear from the context, we write π_A or π_{AB} for the corresponding subtypes. When the type π_{ABC} coincides with the actual pmf of a triple (A, B, C), we omit the subscript and simply write $H(A)$, $H(A|B)$, and $I(A; B|C)$.

For a given P_X and a constant $\mu > 0$, the set of sequences with the same type P_X is denoted by $\mathcal{T}^n(P_X)$. We use $\mathcal{T}_\mu^n(P_X)$ to denote the set of *μ-typical sequences* in $\mathcal{X}^n$:

$$\mathcal{T}_\mu^n(P_X) = \left\{ x^n \colon \left| \frac{|\{i \colon x_i = x\}|}{n} - P_X(x) \right| \le \mu P_X(x), \ \ \forall x \in \mathcal{X} \right\}, \tag{2}$$

where $|\{i \colon x_i = x\}|$ is the number of positions where the sequence x^n equals x. Similarly, $\mathcal{T}_\mu^n(P_{XY})$ stands for the set of *jointly μ-typical sequences* whose definition is as in (2) with x replaced by (x, y).

The expectation operator is written as $\mathbb{E}[\cdot]$. The notation $\mathcal{U}\{a, \ldots, b\}$ is used to indicate a uniform distribution over the set $\{a, \ldots, b\}$; for the uniform distribution over $\{0, 1\}$ we also use $\mathcal{B}(1/2)$. The $\log$ function is taken with base 2. Finally, we abbreviate left-hand side and right-hand side by LHS and RHS.

B. Paper Outline

The remainder of the paper is organized as follows. Section II presents the problem description. Section III presents a coding and testing scheme without binning and the exponent region it achieves. Section IV presents an improved scheme employing binning and the corresponding achievable exponent region. Sections V and VI study the proposed achievable regions when the Markov chains $X^n \to Y^n \to Z^n$ and $X^n \to Z^n \to Y^n$ hold, respectively. The paper is concluded in Section VII and by technical appendices.

II. Detailed Problem Description

Consider the multi-hop hypothesis testing problem with three terminals in Fig. 1. The first terminal in the system, the *transmitter*, observes the sequence X^n, the second terminal, the *relay*, observes the sequence Y^n, and the third terminal, the *receiver*, observes the sequence Z^n. Under the null hypothesis

$$\mathcal{H} = 0 \colon \quad (X^n, Y^n, Z^n) \sim \text{i.i.d. } P_{XYZ}, \tag{3}$$

whereas under the alternative hypothesis

$$\mathcal{H} = 1 \colon \quad (X^n, Y^n, Z^n) \sim \text{i.i.d. } Q_{XYZ}, \tag{4}$$

for two given pmfs P_{XYZ} and Q_{XYZ}.

The problem encompasses a noise-free bit-pipe of rate R from the transmitter to the relay and a noise-free bit pipe of rate T from the relay to the receiver. That means, after observing X^n, the transmitter computes the message $M = \phi^{(n)}(X^n)$ using a possibly stochastic encoding function $\phi^{(n)} \colon \mathcal{X}^n \to \{0, ..., \lfloor 2^{nR} \rfloor\}$ and sends it to the relay. The relay, after observing Y^n and receiving M, computes the message $B = \phi_y^{(n)}(M, Y^n)$ using a possibly stochastic encoding function $\phi_y^{(n)} \colon \mathcal{Y}^n \times \{0, ..., \lfloor 2^{nR} \rfloor\} \to \{0, ..., \lfloor 2^{nT} \rfloor\}$ and sends it to the receiver.

The goal of the communication is that, based on their own observations and the received messages, the relay and the receiver can decide on the hypothesis $\mathcal{H}$. The relay thus produces the guess

$$\hat{\mathcal{H}}_y = g_y^{(n)}(Y^n, M) \tag{5}$$

using a guessing function $g_y^{(n)} \colon \mathcal{Y}^n \times \{0, ..., \lfloor 2^{nR} \rfloor\} \to \{0, 1\}$, and the receiver produces the guess

$$\hat{\mathcal{H}}_z = g_z^{(n)}(Z^n, B) \tag{6}$$

using a guessing function $g_z^{(n)} \colon \mathcal{Z}^n \times \{0, ..., \lfloor 2^{nT} \rfloor\} \to \{0, 1\}$.

Definition 1: For each $\epsilon \in (0, 1)$, we say that the exponent-rate tuple (η, θ, R, T) is ϵ-achievable if there exists a sequence of encoding and decoding functions $(\phi^{(n)}, \phi_y^{(n)}, g_y^{(n)}, g_z^{(n)})$, $n = 1, 2, \ldots$, such that the corresponding sequences of type-I and type-II error probabilities at the relay

$$\alpha_{y,n} := \Pr[\hat{\mathcal{H}}_y = 1 | \mathcal{H} = 0], \tag{7}$$

$$\beta_{y,n} := \Pr[\hat{\mathcal{H}}_y = 0 | \mathcal{H} = 1], \tag{8}$$

and at the receiver

$$\alpha_{z,n} := \Pr[\hat{\mathcal{H}}_z = 1 | \mathcal{H} = 0], \tag{9}$$

$$\beta_{z,n} := \Pr[\hat{\mathcal{H}}_z = 0 | \mathcal{H} = 1], \tag{10}$$

satisfy

$$\alpha_{y,n} \le \epsilon, \tag{11}$$

$$\alpha_{z,n} \le \epsilon, \tag{12}$$

and

$$-\varlimsup_{n\to\infty} \frac{1}{n} \log \beta_{y,n} \ge \theta_y, \tag{13}$$

$$-\varlimsup_{n\to\infty} \frac{1}{n} \log \beta_{z,n} \ge \theta_z. \tag{14}$$

Definition 2: For given rates (R, T), we define the *exponent-rate region* $\mathcal{E}^*(R, T)$ as the closure of all non-negative pairs (θ_y, θ_z) for which $(\theta_y, \theta_z, R, T)$ is ϵ-achievable for every $\epsilon \in (0, 1)$.

Remark 1: In this paper we do not attempt to prove any "strong converse." A strong converse in hypothesis testing would claim that the best achievable type-II error exponents for a given type-I error probability $\epsilon \in (0, 1)$ does not depend on the value of ϵ. For some special cases of the setting in Fig. 1, a strong converse has recently been studied in [26].

III. A Coding and Testing Scheme without Binning

In this section we present a first coding and testing scheme and characterize the achieved exponent-rate region using a relatively simple expression. The scheme is improved in the next section; the exponent-rate region achieved by the improved scheme is however more involved and includes multiple competing exponents.

A. The Coding and Testing Scheme

Fix $\mu > 0$, an arbitrary blocklength n, and joint conditional pmfs $P_{SU|X}$ and $P_{V|SUY}$ over finite auxiliary alphabets $\mathcal{S}, \mathcal{U}$, and $\mathcal{V}$. Define the joint pmf

$$P_{SUVXYZ} = P_{XYZ} P_{SU|X} P_{V|SUY} \tag{15}$$

and the following nonnegative rates, which are calculated according to the distribution in (15) and μ:

$$R_s := I(X;S) + \mu, \tag{16}$$

$$R_u := I(U;X|S) + \mu, \tag{17}$$

$$R_v := I(V;Y,U|S) + \mu. \tag{18}$$

Later, we shall choose the joint distributions in such a way that $R \geq R_s + R_u$ and $T \geq R_s + R_v$.

Code Construction: First, we randomly generate codewords

$$\mathcal{C}_S := \big\{S^n(i)\colon\ i \in \{1, \ldots, \lfloor 2^{nR_s} \rfloor\}\big\} \tag{19}$$

by picking all entries i.i.d. according to P_S. Then, for each $i \in \{1, \ldots, \lfloor 2^{nR_s} \rfloor\}$, we randomly generate codewords

$$\mathcal{C}_U(i) := \big\{U^n(j|i)\colon\ j \in \{1, \ldots, \lfloor 2^{nR_u} \rfloor\}\big\} \tag{20}$$

by choosing for each $t \in \{1, \ldots, n\}$ and $j \in \{1, \ldots, \lfloor 2^{nR_u} \rfloor\}$, the t-th component $U_t(j|i)$ of codeword $U^n(j|i)$ independently according to the conditional distribution $P_{U|S}(\cdot|S_t(i))$, where $S_t(i)$ denotes the t-th component of the codeword $S^n(i)$. For each $i \in \{1, \ldots, \lfloor 2^{nR_s} \rfloor\}$, generate also codewords

$$\mathcal{C}_V(i) := \big\{V^n(k|i)\colon k \in \{1, \ldots, \lfloor 2^{nR_v} \rfloor\}\big\} \tag{21}$$

by choosing for each $t \in \{1, \ldots, n\}$ and $k \in \{1, \ldots, \lfloor 2^{nR_v} \rfloor\}$ the t-th component $V_t(k|i)$ of codeword $V^n(k|i)$ independently according to the conditional distribution $P_{V|S}(\cdot|S_t(i))$.

Reveal the realizations $\{s^n(i)\}$, $\{u^n(j|i)\}$, and $\{v^n(k|i)\}$ of the random code constructions to all terminals.

Transmitter: Given that it observes the sequence $X^n = x^n$, the transmitter looks for a pair of indices (i, j) such that

$$(s^n(i), u^n(j|i), x^n) \in \mathcal{T}^n_{\mu/4}(P_{SUX}). \tag{22}$$

If successful, it picks one such pair uniformly at random and sends

$$M = (i, j) \tag{23}$$

to the relay. Otherwise, it sends $M = 0$.

Relay: Assume that the relay observes the sequence $Y^n = y^n$ and receives the message $M = m$. If $m = 0$, it declares $\hat{\mathcal{H}}_y = 1$ and sends $b = 0$ to the receiver. Otherwise, it decomposes $m = (i, j)$ and looks for an index k such that

$$(s^n(i), u^n(j|i), v^n(k|i), y^n) \in \mathcal{T}^n_{\mu/2}(P_{SUVY}). \tag{24}$$

If such an index k exists, the relay declares $\hat{\mathcal{H}}_y = 1$ and sends the pair

$$B = (i, k) \tag{25}$$

to the receiver. Otherwise, it declares $\hat{\mathcal{H}}_y = 1$ and sends the message $B = 0$.

Receiver: Assume that the receiver observes $Z^n = z^n$ and receives message $B = b$ from the relay. If $b = 0$, the receiver declares $\hat{\mathcal{H}}_z = 1$. Otherwise, it decomposes $b = (i, k)$ and checks whether

$$(s^n(i), v^n(k|i), z^n) \in \mathcal{T}^n_{\mu}(P_{SVZ}). \tag{26}$$

If the typicality check is successful, the receiver declares $\hat{\mathcal{H}}_z = 0$. Otherwise, it declares $\hat{\mathcal{H}}_z = 1$.

B. Achievable Exponent-Rate Region

We present the exponent region achieved by the preceding scheme.

Given two conditional pmfs $P_{SU|X}$ and $P_{V|SUY}$, define $\mathcal{E}_{\text{nobin}}(P_{SU|X}, P_{V|SUY})$ as the set of all pairs (θ_y, θ_z) that satisfy

$$\theta_y \leq \min_{\substack{\tilde{P}_{SUXY}:\\ \tilde{P}_{SUX}=P_{SUX}\\ \tilde{P}_{SUY}=P_{SUY}}} D(\tilde{P}_{SUXY} \| P_{SU|X} Q_{XY}), \tag{27}$$

$$\theta_z \leq \min_{\substack{\tilde{P}_{SUVXYZ}:\\ \tilde{P}_{SUX}=P_{SUX}\\ \tilde{P}_{SUVY}=P_{SUVY}\\ \tilde{P}_{SVZ}=P_{SVZ}}} D(\tilde{P}_{SUVXYZ} \| P_{SU|X} P_{V|SUY} Q_{XYZ}), \tag{28}$$

where the joint pmf P_{SUVXYZ} is defined as in (15) and P_{SUX}, P_{SUY}, P_{SUVY} and P_{SVZ} are marginals of this pmf.

Define further the exponent region

$$\mathcal{E}_{\text{nobin}}(R, T) := \bigcup_{P_{SU|X}, P_{V|SUY}} \mathcal{E}_{\text{nobin}}(P_{SU|X}, P_{V|SUY}) \tag{29}$$

where the union is over all pairs of conditional pmfs $(P_{SU|X}, P_{V|SUY})$ satisfying

$$R \geq I(S, U; X), \tag{30}$$

$$T \geq I(X; S) + I(V; Y, U|S) \tag{31}$$

and the mutual informations are again calculated according to the joint pmf defined in (15).

Theorem 1 (Achievable Region without Binning): For any pair of nonnegative rates $R, T \geq 0$, the set $\mathcal{E}_{\text{nobin}}(R, T)$ is achievable:

$$\mathcal{E}_{\text{nobin}}(R, T) \subseteq \mathcal{E}^*(R, T) \tag{32}$$

Proof: See Appendix A. ∎

In the above theorem, it suffices to consider auxiliary random variables S, U, and V over alphabets $\mathcal{S}$, $\mathcal{U}$, and $\mathcal{V}$ whose sizes satisfy: $|\mathcal{S}| \leq |\mathcal{X}| + 4$, $|\mathcal{U}| \leq |\mathcal{X}| \cdot |\mathcal{S}| + 3$ and $|\mathcal{V}| \leq |\mathcal{U}| \cdot |\mathcal{S}| \cdot |\mathcal{Y}| + 2$. This follows by simple applications of Caratheodory's theorem.

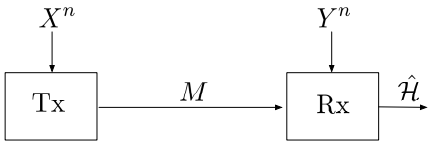

Fig. 2. Hypothesis testing over a point-to-point channel

IV. An Improved Scheme with Binning

In source coding, it is well known that binning can decrease the required rate of communication when observations at different terminals are correlated. The same holds for hypothesis testing. Before extending our coding and testing scheme from the previous section to include binning, for completeness, we provide a detailed proof of the Shimokawa, Han, and Amari error exponent [3] achieved over a point-to-point link when using binning. So far, a detailed proof was available only in Japanese [27].

A. Point-to-Point Link

Consider the network in Fig. 2, which can be obtained as a special case from the previously introduced two-hop relay network by setting $T = 0$ and Z a constant that is the same under both hypotheses. In this case, the exponent θ_z cannot be positive and is uninteresting. The system performance is then characterized by the exponent θ_y, and for the purpose of this subsection, the relay can be regarded as the final receiver. Therefore, in the remainder of this subsection, we call the terminal that observes Y^n "the receiver". We make the following definition:

Definition 3: Consider a single-hop system with only transmitter and receiver. The exponent-rate function $\theta^*(R)$ is the supremum of all ϵ-achievable error exponents for a given rate R, i.e.,

$$\theta^*(R) := \sup\,\{\theta_y \geq 0\colon (\theta_y, 0, R, 0) \text{ is } \epsilon\text{-achievable } \forall \epsilon > 0\}. \tag{33}$$

We recall the lower bound on $\theta^*(R)$ in [3], after presenting a coding and testing scheme that achieves this exponent. (The presented scheme slightly deviates from the scheme in [3].)

1) Coding and Testing Scheme: Fix $\mu > 0$, a sufficiently large blocklength n, and the conditional pmf $P_{S|X}$ over a finite auxiliary alphabet $\mathcal{S}$. Define the joint pmf

$$P_{SXY} = P_{XY} P_{S|X} \tag{34}$$

and, if $R > I(S;X)$ define the nonnegative rate $R' = 0$ and otherwise choose R' such that

$$R + R' \geq I(X;S) + \mu, \tag{35}$$

$$R' < I(Y;S). \tag{36}$$

In the following coding scheme, when $R \leq I(S;X)$, we distribute the s^n-codewords in bins. Instead of sending the complete index of the chosen s^n, the transmitter sends only its bin number to the receiver. The receiver then selects the s^n codeword from the indicated bin that is "most-compatible" with its local observation Y^n, and makes its decision based on this selected codeword. By performing binning, the transmitter and the receiver can use a smaller communication rate, but the error probabilities may be higher.

Code Construction: Construct a random codebook

$$\mathcal{C}_S := \{S^n(m,\ell)\colon\ m \in \{1,...,\lfloor 2^{nR}\rfloor\}, \ell \in \{1,...,\lfloor 2^{nR'}\rfloor\}\}, \tag{37}$$

by drawing all entries of all codewords i.i.d. according to the chosen distribution P_S.

Reveal the realization $\{s^n(m,\ell)\}$ of the random codebook to both terminals.

Transmitter: Given that it observes the sequence $X^n = x^n$, the transmitter looks for indices (m,ℓ) such that

$$(s^n(m,\ell), x^n) \in \mathcal{T}^n_{\mu/2}(P_{SX}). \tag{38}$$

If successful, it picks one of these indices uniformly at random and sends the index $M = m$ to the relay. Otherwise, it sends $M = 0$.

Receiver: Assume that the receiver observes $Y^n = y^n$ and receives the message $M = m$ from the transmitter. If $m = 0$, the receiver declares $\hat{\mathcal{H}} = 1$. Otherwise, it looks for an index $\ell' \in \{1,...,\lfloor 2^{nR'}\rfloor\}$ that minimizes $H_{\text{tp}(s^n(m,\ell''),y^n)}(S|Y)$ among all ℓ'' satisfying $s^n(m,\ell'') \in \mathcal{T}^n_\mu(P_S)$.[1] Then it checks whether

$$(s^n(m,\ell'), y^n) \in \mathcal{T}^n_\mu(P_{SY}),$$

and declares $\hat{\mathcal{H}} = 0$ if this typicality check is successful and $\hat{\mathcal{H}} = 1$ otherwise.

2) Result on the Error Exponent: The scheme described in the previous subsection yields the following lower bound on the exponent-rate function.

Theorem 2 ([3]): For every choice of the conditional distribution $P_{S|X}$ satisfying that $R \geq I(S;X|Y)$, the exponent-rate function $\theta^*(R)$ is lower-bounded as

$$\theta^*(R) \geq \min\Bigg\{ \min_{\substack{\tilde{P}_{SXY}\colon \\ \tilde{P}_{SX}=P_{SX} \\ \tilde{P}_{SY}=P_{SY}}} D(\tilde{P}_{SXY}\|P_{S|X}Q_{XY}),$$

$$\min_{\substack{\tilde{P}_{SXY}\colon \\ \tilde{P}_{SX}=P_{SX} \\ \tilde{P}_{Y}=P_{Y} \\ H(S|Y)\leq H_{\tilde{P}_{SY}}(S|Y)}} D(\tilde{P}_{SXY}\|P_{S|X}Q_{XY}) + R - I(S;X|Y)\Bigg\}. \tag{39}$$

For a choice of $P_{S|X}$ such that $R \geq I(S;X)$, the bound can be tightened to

$$\theta^*(R) \geq \min_{\substack{\tilde{P}_{SXY}\colon \\ \tilde{P}_{SX}=P_{SX} \\ \tilde{P}_{SY}=P_{SY}}} D(\tilde{P}_{SXY}\|P_{S|X}Q_{XY}), \tag{40}$$

Here mutual informations and the entropy $H(S|Y)$ in the miniminization constraint are calculated according to the joint pmf in (34) and the chosen conditional pmf $P_{S|X}$.

[1]Notice that because $m \neq 0$, there exists at least one codeword $s^n(m,\ell') \in \mathcal{T}^n_\mu(P_S)$ in bin m.

Proof: When $R \geq I(S;X)$, our scheme does not use binning and an analysis similar to Appendix A (the analysis of the multi-hop scheme without binning) yields the desired result. When $R < I(S;X)$, our scheme uses binning and is analyzed in Appendix B. ∎

In the above theorem, it suffices to consider auxiliary random variables S over alphabets $\mathcal{S}$ whose sizes satisfy: $|\mathcal{S}| \leq |\mathcal{X}| + 2$.

The inequality in Theorem 2 holds with equality in the special cases of *testing against independence* [1], where $Q_{XY} = P_X \cdot P_Y$,[2] and of *testing against conditional independence* [4], where Y decomposes as $Y = (Y_{\mathrm{C}}, Y_{\mathrm{H}})$ and $Q_{XY_{\mathrm{C}}Y_{\mathrm{H}}} = P_{XY_{\mathrm{C}}} P_{Y_{\mathrm{H}}|Y_{\mathrm{C}}}$.

B. The Two-Hop Relay Network

We turn back to the two-hop relay network and propose an improved coding and testing scheme employing binning.

1) Coding and Testing Scheme: Fix $\mu > 0$, an arbitrary blocklength n, and joint conditional pmfs $P_{SU|X}$ and $P_{V|SUY}$ over finite auxiliary alphabets $\mathcal{S}$, $\mathcal{U}$, and $\mathcal{V}$. Define the joint pmf $P_{SUVXYZ} = P_{XYZ} P_{SU|X} P_{V|SUY}$ and the following nonnegative rates, which are calculated according to the chosen distribution,

$$R_s = I(X;S) + \mu, \tag{41a}$$
$$R_u + R'_u = I(U;X|S) + \mu, \tag{41b}$$
$$R_v + R'_v = I(V;Y,U|S) + \mu, \tag{41c}$$
$$R'_u \leq I(U;Y|S), \tag{41d}$$
$$R'_v \leq I(V;Z|S). \tag{41e}$$

The joint distributions are chosen in such a way that

$$R \geq R_s + R_u \tag{42}$$
$$T \geq R_s + R_v. \tag{43}$$

In the following coding scheme, the transmitter distributes the u^n codewords in bins, and sends the bin number of the chosen u^n to the relay. The relay looks in that bin for the u^n codeword that is "most compatible" with its local observation Y^n. Similarly, the relay and the receiver perform binning on v^n. Note that for simplicity and ease of exposition, we do not bin the s^n codewords.

Code Construction: Construct a random codebook

$$\mathcal{C}_S = \{S^n(i)\colon \ \ i \in \{1, ..., \lfloor 2^{nR_s} \rfloor\}\}$$

by selecting each entry of the n-length codeword $s^n(i)$ in an i.i.d. manner according to the pmf P_S. Then, for each i, generate random codebooks

$$\mathcal{C}_U(i) = \{U^n(j,e|i)\colon j \in \{1, ..., \lfloor 2^{nR_u} \rfloor\}, e \in \{1, ..., \lfloor 2^{nR'_u} \rfloor\}\}$$

and

$$\mathcal{C}_V(i) = \{V^n(k,f|i)\colon k \in \{1, ..., \lfloor 2^{nR_v} \rfloor\}, f \in \{1, ..., \lfloor 2^{nR'_v} \rfloor\}\}$$

by selecting for each $t \in \{1, \ldots, n\}$, the t-th components $U_t(j,e|i)$ and $V_t(k,f|i)$ of the codewords $U^n(j,e|i)$ and $V^n(k,f|i)$ independently using the conditional pmfs $P_{U|S}(\cdot|S_t(i))$ and $P_{V|S}(\cdot|S_t(i))$, where $S_t(i)$ denotes the t-th component of codeword $S^n(i)$.

Reveal the realizations $\{s^n(i)\}$, $\{u^n(j,e|i)\}$, and $\{v^n(k,f|i)\}$ of the random codebooks to all terminals.

Transmitter: Given that the transmitter observes the sequence $X^n = x^n$, it looks for indices (i,j,e) such that

$$(s^n(i), u^n(j,e|i), x^n) \in \mathcal{T}^n_{\mu/4}(P_{SUX}). \tag{44}$$

If successful, it picks one such triple uniformly at random, and sends the first two indices of the triple:

$$M = (i,j) \tag{45}$$

to the relay. Otherwise, it sends $M = 0$.

Relay: Assume that the relay observes the sequence $Y^n = y^n$ and receives the message $M = m$. If $m = 0$, it declares $\hat{\mathcal{H}}_y = 1$ and sends $B = 0$ to the receiver. Otherwise, it looks for an index e' which minimizes $H_{\mathrm{tp}(s^n(i),u^n(j,e''|i),y^n)}(U|S,Y)$ among all e'' satisfying $(s^n(i), u^n(j,e''|i)) \in \mathcal{T}^n_{\mu/2}(P_{SU})$. It then looks for indices (k,f) such that

$$(s^n(i), u^n(j,e'|i), v^n(k,f|i), y^n) \in \mathcal{T}^n_{\mu/2}(P_{SUVY}). \tag{46}$$

If successful, it declares $\hat{\mathcal{H}}_y = 0$ and picks one of these index pairs uniformly at random. It then sends the corresponding indices

$$B = (i,k) \tag{47}$$

to the receiver. Otherwise, it declares $\hat{\mathcal{H}}_y = 1$ and sends the message $B = 0$ to the receiver.

Receiver: Assume that the receiver observes $Z^n = z^n$ and receives message $B = b$ from the relay. If $b = 0$, the receiver declares $\hat{\mathcal{H}}_z = 1$. Otherwise, it looks for an index f' which minimizes $H_{\mathrm{tp}(s^n(i),v^n(k,f''|i),z^n)}(V|S,Z)$ among all f'' satisfying $(s^n(i), v^n(k,f''|i)) \in \mathcal{T}^n_{\mu/2}(P_{SV})$. Then, it checks whether

$$(s^n(i), v^n(k,f'|i), z^n) \in \mathcal{T}^n_{\mu}(P_{SVZ}). \tag{48}$$

If successful, the receiver declares $\hat{\mathcal{H}}_z = 0$. Otherwise, it declares $\hat{\mathcal{H}}_z = 1$.

2) Result on the Exponent-Rate Region: The coding scheme in the previous subsection establishes the following theorem.

For any pair of conditional pmfs $P_{US|X}$ and $P_{V|SUY}$, let $\mathcal{E}_{\mathrm{bin}}(P_{US|X}, P_{V|SUY})$ denote the set of all exponent-pairs (θ_y, θ_z) that satisfy

$$\theta_y \leq \min\{\theta_y^{(1)}, \theta_y^{(2)}\}, \tag{49}$$
$$\theta_z \leq \min\{\theta_z^{(i)}\colon i = 1, ..., 4\}, \tag{50}$$

where

$$\theta_y^{(1)} := \min_{\substack{\tilde{P}_{SUXY}\colon \\ \tilde{P}_{SUX} = P_{SUX} \\ \tilde{P}_{SUY} = P_{SUY}}} D(\tilde{P}_{SUXY} \| P_{SU|X} Q_{XY}), \tag{51a}$$

[2] There is no need to apply the coding scheme with binning to attain the optimal error exponent in this case, see [1].

$$\theta_y^{(2)} := \min_{\substack{\tilde{P}_{SUXY}: \\ \tilde{P}_{SUX}=P_{SUX} \\ \tilde{P}_{SY}=P_{SY} \\ H(U|S,Y)\le H_{\tilde{P}}(U|S,Y)}} D(\tilde{P}_{SUXY}\|P_{SU|X}Q_{XY}) + R - I(S,U;X) + I(U;Y|S), \quad (51b)$$

$$\theta_z^{(1)} := \min_{\substack{\tilde{P}_{SUVXYZ}: \\ \tilde{P}_{SUX}=P_{SUX} \\ \tilde{P}_{SUVY}=P_{SUVY} \\ \tilde{P}_{SVZ}=P_{SVZ}}} D(\tilde{P}_{SUVXYZ}\|P_{SU|X}P_{V|SUY}Q_{XYZ}), \quad (51c)$$

$$\theta_z^{(2)} := \min_{\substack{\tilde{P}_{SUVXYZ}: \\ \tilde{P}_{SUX}=P_{SUX} \\ \tilde{P}_{SVY}=P_{SVY} \\ \tilde{P}_{SVZ}=P_{SVZ} \\ H(U|S,Y)\le H_{\tilde{P}}(U|S,Y)}} D(\tilde{P}_{SUVXYZ}\|P_{SU|X}P_{V|SY}Q_{XYZ}) + R - I(S,U;X) + I(U;Y|S), \quad (51d)$$

$$\theta_z^{(3)} := \min_{\substack{\tilde{P}_{SUVXYZ}: \\ \tilde{P}_{SUX}=P_{SUX} \\ \tilde{P}_{SUVY}=P_{SUVY} \\ \tilde{P}_{SZ}=P_{SZ} \\ H(V|S,Z)\le H_{\tilde{P}}(V|S,Z)}} D(\tilde{P}_{SUVXYZ}\|P_{SU|X}P_{V|SUY}Q_{XYZ}) + T - I(S;X) - I(V;Y,U|S) + I(V;Z|S), \quad (51e)$$

$$\theta_z^{(4)} := \min_{\substack{\tilde{P}_{SUVXYZ}: \\ \tilde{P}_{SUX}=P_{SUX} \\ \tilde{P}_{SVY}=P_{SVY} \\ \tilde{P}_{SZ}=P_{SZ} \\ H(U|S,Y)\le H_{\tilde{P}}(U|S,Y) \\ H(V|S,Z)\le H_{\tilde{P}}(V|S,Z)}} D(\tilde{P}_{SUVXYZ}\|P_{SU|X}P_{V|SY}Q_{XYZ}) + R + T - I(S,U;X) - I(X;S) - I(V;U,Y|S) + I(U;Y|S) + I(V;Z|S), \quad (51f)$$

where the mutual information and entropy terms, as well as the marginals P_{SUX}, P_{SUVY}, P_{SVY}, P_{SVZ}, and P_{SZ} are calculated with respect to the joint pmf

$$P_{SUVXYZ} = P_{US|X}P_{V|USY}P_{XYZ}. \quad (52)$$

Define then the exponent-rate region

$$\mathcal{E}_{\text{bin}}(R,T) := \bigcup_{(P_{US|X},P_{V|USY})} \mathcal{E}_{\text{bin}}(P_{US|X},P_{V|USY}) \quad (53)$$

where the union is over all pairs of conditional distributions so that the rate constraints

$$R \ge I(S,U;X) - I(U;Y|S), \quad (54)$$
$$T \ge I(V;Y,U|S) - I(V;Z|S) + I(S;X), \quad (55)$$

are satisfied when the mutual informations are calculated according to the joint pmf in (52).

Theorem 3 (Achievable Region with Binning): For any positive rate-pair (R,T):

$$\mathcal{E}_{\text{bin}}(R,T) \subseteq \mathcal{E}^*(R,T). \quad (56)$$

Proof: See Appendix C. ∎

For each choice of conditional pmfs $P_{US|X}$ and $P_{V|USY}$, the achievable exponents-region $\mathcal{E}_{\text{bin}}(P_{US|X},P_{V|USY})$ is characterized through two competing exponents at the relay and four competing exponents at the receiver, see (49) and (50). Extending our scheme by binning also the s^n codewords achieves an exponents region that is characterized by three competing exponents at the relay and ten competing exponents at the receiver. Details are omitted for brevity.

V. The Special Case "$X^n \to Y^n \to Z^n$ Under Both Hypotheses"

Consider a situation where the relay lies in between the transmitter and the receiver, and thus the signals at the sensor and the receiver are conditionally independent given the signal at the relay. In this situation, the two-hop relay network seems particularly adequate for modelling short-range wireless communication.

Assume that the pmfs P_{XYZ} and Q_{XYZ} decompose as

$$P_{XYZ} = P_X \cdot P_{Y|X} \cdot P_{Z|Y}, \quad (57)$$
$$Q_{XYZ} = Q_X \cdot Q_{Y|X} \cdot Q_{Z|Y}. \quad (58)$$

We start by showing that in this special case the compression mechanisms in the previously-presented coding and testing schemes can be simplified. There is no need to send compression information from the transmitter to the receiver. Hence, the message sent from the relay to the receiver consists only of the relay's own guess and compression information of the relay's observation. Technically, this means that the expressions for $\mathcal{E}_{\text{nobin}}(R,T)$ and $\mathcal{E}_{\text{bin}}(R,T)$ can be simplified for this special case by setting S to be a constant, and choosing V to be conditionally independent of U given Y. In the following, we use the subscript "dcpled" to refer to the region of this special case, which stands for "decoupled". Here, the transmitter-relay and relay-receiver links are basically decoupled from each other thanks to the Markov chain $X \to Y \to Z$.

A. Simplified Exponent Regions

Given two conditional pmfs $P_{U|X}$ and $P_{V|Y}$, define the exponents region $\mathcal{E}_{\text{dcpled}}(P_{U|X},P_{V|Y})$ as the set of all pairs (θ_y,θ_z) that satisfy

$$\theta_y \le \min_{\substack{\tilde{P}_{UXY}: \\ \tilde{P}_{UX}=P_{UX} \\ \tilde{P}_{UY}=P_{UY}}} D(\tilde{P}_{UXY}\|P_{U|X}Q_{XY}), \quad (59)$$

$$\theta_z \le \min_{\substack{\tilde{P}_{UXY}: \\ \tilde{P}_{UX}=P_{UX} \\ \tilde{P}_{UY}=P_{UY}}} D(\tilde{P}_{UXY}\|P_{U|X}Q_{XY}) + \min_{\substack{\tilde{P}_{VYZ}: \\ \tilde{P}_{VY}=P_{VY} \\ \tilde{P}_{VZ}=P_{VZ}}} \mathbb{E}_{P_Y}\Big[D(\tilde{P}_{VZ|Y}\|P_{V|Y}Q_{Z|Y})\Big], \quad (60)$$

where P_{UY} and P_{VZ} indicate the marginals of the joint pmfs $P_{U|X}P_{XY}$ and $P_{V|Y}P_{YZ}$, and further define

$$\mathcal{E}_{\text{dcpled}}(R,T) := \bigcup_{P_{U|X},P_{V|Y}} \mathcal{E}_{\text{dcpled}}(P_{U|X},P_{V|Y}) \quad (61)$$

where the union is over all pairs of conditional pmfs $(P_{U|X},P_{V|Y})$ satisfying

$$R \ge I(U;X), \quad (62)$$
$$T \ge I(V;Y) \quad (63)$$

for mutual informations that are calculated according to the joint pmfs $P_{UX} = P_{U|X}P_X$ and $P_{VY} = P_{V|Y}P_Y$.

Proposition 1 (Simplified Achievable Region Without Binnning): If (57) and (58) hold, then

$$\mathcal{E}_{\text{dcpled}}(R,T) = \mathcal{E}_{\text{nobin}}(R,T). \tag{64}$$

Proof: See Appendix D. ∎

In the above proposition, it suffices to consider auxiliary random variables U and V over alphabets $\mathcal{U}$ and $\mathcal{V}$ whose sizes satisfy: $|\mathcal{U}| \le |\mathcal{X}| + 1$ and $|\mathcal{V}| \le |\mathcal{Y}| + 1$.

Similarly, given two conditional pmfs $P_{U|X}$ and $P_{V|Y}$, let $\mathcal{E}_{\text{bin,dcpled}}(P_{U|X}, P_{V|Y})$ denote the set of all exponent-pairs (θ_y, θ_z) that satisfy

$$\begin{aligned}\theta_y \le \min\Big\{ &\min_{\substack{\tilde{P}_{UXY}:\\ \tilde{P}_{UX}=P_{UX}\\ \tilde{P}_{UY}=P_{UY}}} D(\tilde{P}_{UXY}\|P_{U|X}Q_{XY}),\\ &\min_{\substack{\tilde{P}_{UXY}:\\ \tilde{P}_{UX}=P_{UX}\\ \tilde{P}_Y=P_Y\\ H(U|Y)\le H_{\tilde{P}}(U|Y)}} D(\tilde{P}_{UXY}\|P_{U|X}Q_{XY}) + R - I(U;X|Y)\Big\},\end{aligned} \tag{65a}$$

and

$$\begin{aligned}\theta_z \le \min\Big\{ &\min_{\substack{\tilde{P}_{UXY}:\\ \tilde{P}_{UX}=P_{UX}\\ \tilde{P}_{UY}=P_{UY}}} D(\tilde{P}_{UXY}\|P_{U|X}Q_{XY}),\\ &\min_{\substack{\tilde{P}_{UXY}:\\ \tilde{P}_{UX}=P_{UX}\\ \tilde{P}_Y=P_Y\\ H(U|Y)\le H_{\tilde{P}}(U|Y)}} D(\tilde{P}_{UXY}\|P_{U|X}Q_{XY})\\ &\qquad\qquad +R - I(U;X|Y)\Big\},\\ +\min\Big\{ &\min_{\substack{\tilde{P}_{VZ|Y}:\\ \tilde{P}_{VY}=P_{VY}\\ \tilde{P}_{VZ}=P_{VZ}}} \mathbb{E}_{P_Y}\Big[D(\tilde{P}_{VZ|Y}\|P_{V|Y}Q_{Z|Y})\Big],\\ &\min_{\substack{\tilde{P}_{VZ|Y}:\\ \tilde{P}_{VY}=P_{VY}\\ \tilde{P}_Z=P_Z\\ H(V|Z)\le H_{\tilde{P}}(V|Z)}} \mathbb{E}_{P_Y}\Big[D(\tilde{P}_{VZ|Y}\|P_{V|Y}Q_{Z|Y})\Big]\\ &\qquad\qquad +T - I(V;Y|Z)\Big\},\end{aligned} \tag{65b}$$

where the mutual information and entropy terms, as well as the marginals P_{SUX}, P_{SUVY}, P_{SVY}, P_{SVZ}, and P_{SZ} are calculated with respect to the joint pmf

$$P_{SUVXYZ} = P_{US|X}P_{V|USY}P_{XYZ}. \tag{66}$$

Further define

$$\mathcal{E}_{\text{bin,dcpled}}(R,T) := \bigcup_{P_{U|X}, P_{V|Y}} \mathcal{E}_{\text{bin,dcpled}}(P_{U|X}, P_{V|Y}) \tag{67}$$

where the union is over all pairs of conditional distributions for which the rate constraints

$$R \ge I(U;X|Y), \tag{68}$$
$$T \ge I(V;Y|Z), \tag{69}$$

are satisfied when the mutual informations are calculated according to the joint pmf in (66).

Proposition 2 (Simplified Achievable Region With Binnning): If (57) and (58) hold, then

$$\mathcal{E}_{\text{bin,dcpled}}(R,T) = \mathcal{E}_{\text{bin}}(R,T). \tag{70}$$

Proof: The inclusion $\mathcal{E}_{\text{bin,dcpled}}(R,T) \subseteq \mathcal{E}_{\text{bin}}(R,T)$ follows by restricting to U and V to be conditionally independent given Y and S to be a constant. The proof of inclusion $\mathcal{E}_{\text{bin,dcpled}}(R,T) \supseteq \mathcal{E}_{\text{bin}}(R,T)$ is sketched in Appendix E. ∎

Remark 2: For both Propositions 1 and 2, the exponent at the receiver equals the sum of two exponents: the first is the exponent at the relay (i.e., the exponent attained over the transmitter-relay link), and the second is the exponent on the isolated relay-receiver link, but with Q_{YZ} replaced by $P_Y Q_{Z|Y}$.

B. Optimality Results

In the following, we prove optimality of the achievable region in Proposition 2 for some cases of "testing against conditional independence" under the Markov conditions (57) and (58). In the following examples, if the random variables Y_{C} and Z_{C} are constants, then the setups reduce to "testing against independence". For "testing against independence", achievability can also be established using the simpler Proposition 1. In other words, the optimal exponents can also be achieved without binning.

1) Special Case 1: Assume that the relay's and the receiver's observations decompose as

$$Y = (Y_{\text{C}}, Y_{\text{H}}) \tag{71}$$
$$Z = (Y_{\text{C}}, Z_{\text{C}}, Z_{\text{H}}) \tag{72}$$

and

$$\begin{aligned}\text{under } \mathcal{H}=0\text{:}\quad &\big(X^n, Y_{\text{C}}^n, Y_{\text{H}}^n, Z_{\text{C}}^n, Z_{\text{H}}^n\big) \text{ i.i.d.}\\ &\sim P_{X|Y_{\text{C}}Y_{\text{H}}} \cdot P_{Y_{\text{C}}Y_{\text{H}}Z_{\text{C}}Z_{\text{H}}},\end{aligned} \tag{73}$$

$$\begin{aligned}\text{under } \mathcal{H}=1\text{:}\quad &\big(X^n, Y_{\text{C}}^n, Y_{\text{H}}^n, Z_{\text{C}}^n, Z_{\text{H}}^n\big) \text{ i.i.d.}\\ &\sim P_{X|Y_{\text{C}}} \cdot P_{Y_{\text{H}}|Y_{\text{C}}Z_{\text{C}}} \cdot P_{Y_{\text{C}}Z_{\text{C}}Z_{\text{H}}}.\end{aligned} \tag{74}$$

The following corollary shows that in this case, the receiver's optimal error exponent equals the *sum* of the optimal error exponent at the relay and the optimal error exponent achieved over the isolated relay-receiver link.

Corollary 1: If (71)–(74) hold, the exponent-rate region $\mathcal{E}^*(R,T)$ is the set of all nonnegative pairs (θ_y, θ_z) that satisfy

$$\theta_y \le I(U;Y|Y_{\text{C}}), \tag{75}$$
$$\theta_z \le I(U;Y|Y_{\text{C}}) + I(V;Z|Z_{\text{C}}, Y_{\text{C}}), \tag{76}$$

for some auxiliary random variables (U,V) satisfying the Markov chains $U \to X \to Y$ and $V \to Y \to Z$ and the rate constraints

$$R \ge I(U;X|Y_{\text{C}}), \tag{77}$$
$$T \ge I(V;Y|Y_{\text{C}}, Z_{\text{C}}), \tag{78}$$

and where $Y = (Y_{\rm C}, Y_{\rm H})$, $Z = (Z_{\rm C}, Z_{\rm H})$, and $(X, Y_{\rm C}, Y_{\rm H}, Z_{\rm C}, Z_{\rm H}) \sim P_{X|Y_{\rm C}Y_{\rm H}} \cdot P_{Y_{\rm C}Y_{\rm H}Z_{\rm C}Z_{\rm H}}$.

Proof: Achievability follows by simplifying Proposition 2. For this special case, since $R \geq I(U;X|Y_{\rm C})$ and $T \geq I(V;Y|Y_{\rm C}, Z_{\rm C})$, exponents $\theta_z^{(2)}, \theta_z^{(3)}, \theta_z^{(4)}$ become inactive in view of $\theta_z^{(1)}$. The converse is proved in Appendix F. ∎

In the above theorem it suffices to consider auxiliary random variables U and V over alphabets $\mathcal{U}$ and $\mathcal{V}$ whose sizes satisfy: $|\mathcal{U}| \leq |\mathcal{X}| + 2$ and $|\mathcal{V}| \leq |\mathcal{Y}| + 1$.

Remark 3: If we set $Y_{\rm C}$ and $Z_{\rm C}$ to constants, then this special case reduces to one where

$$P_{XYZ} = P_{XY} \cdot P_{Z|Y} \tag{79}$$
$$Q_{XYZ} = P_X \cdot P_Y \cdot P_Z. \tag{80}$$

The exponent-region then becomes the set of all nonnegative pairs (θ_y, θ_z) that satisfy

$$\theta_y \leq I(U;Y), \tag{81}$$
$$\theta_z \leq I(U;Y) + I(V;Z), \tag{82}$$

for a pair of auxiliary random variables U and V satisfying the Markov chains $U \to X \to Y$ and $V \to Y \to Z$ and the rate constraints

$$R \geq I(U;X) \tag{83}$$
$$T \geq I(V;Y). \tag{84}$$

Furthermore, the exponent-rate region can be obtained using Proposition 1.

2) Special Case 2: Assume that the receiver's observation decomposes as

$$Z = (Z_{\rm C}, Z_{\rm H}) \tag{85}$$

and

$$\text{under } \mathcal{H} = 0\text{:} \quad \left(X^n, Y^n, Z_{\rm C}^n, Z_{\rm H}^n\right) \text{ i.i.d. } \sim P_{XYZ_{\rm C}Z_{\rm H}}, \tag{86}$$
$$\text{under } \mathcal{H} = 1\text{:} \quad \left(X^n, Y^n, Z_{\rm C}^n, Z_{\rm H}^n\right) \text{ i.i.d. } \sim P_{XYZ_{\rm C}} \cdot P_{Z_{\rm H}|Z_{\rm C}}. \tag{87}$$

In this case, the relay cannot obtain a positive exponent since $(X^n, Y^n) \sim P_{XY}$ under both hypotheses. Moreover, as the following corollary shows, the relay can completely ignore the message from the transmitter and act as if it was the transmitter of a simple point-to-point setup [2].

Corollary 2: Assume (85)–(87). The exponent-rate region $\mathcal{E}^*(R,T)$ is the set of all nonnegative pairs (θ_y, θ_z) that satisfy

$$\theta_y = 0 \tag{88}$$
$$\theta_z \leq I(V;Z_{\rm H}|Z_{\rm C}) \tag{89}$$

for an auxiliary random variable V satisfying the Markov chain $V \to Y \to Z$ and the rate constraint

$$T \geq I(V;Y|Z_{\rm C}), \tag{90}$$

where $Z = (Z_{\rm C}, Z_{\rm H})$, and $(X, Y, Z_{\rm C}, Z_{\rm H}) \sim P_{XYZ_{\rm C}Z_{\rm H}}$. (No constraint involves the rate R.)

Proof: Achievability follows by specializing Proposition 2 to $U = 0$ (deterministically) and then simplifying the expressions. In particular, notice that, since $T \geq I(V;Y|Z_{\rm C})$, exponents $\theta_z^{(2)}, \theta_z^{(3)}, \theta_z^{(4)}$ become inactive in view of $\theta_z^{(1)}$. The converse is standard; details can be found in Appendix G. ∎

Remark 4: If we set $Z_{\rm C}$ to a constant, then the problem reduces to one where

$$P_{XYZ} = P_{XY} \cdot P_{Z|Y} \tag{91}$$
$$Q_{XYZ} = P_{XY} \cdot P_Z. \tag{92}$$

The exponent-rate region then becomes the set of all nonnegative pairs (θ_y, θ_z) that satisfy

$$\theta_y = 0, \tag{93}$$
$$\theta_z \leq I(V;Z), \tag{94}$$

for an auxiliary random variable V satisfying the Markov chain $V \to Y \to Z$ and the rate constraint

$$T \geq I(V;Y). \tag{95}$$

The region is again achievable using Proposition 1.

3) Special Case 3: Assume that the relay's observation decomposes as

$$Y = (Y_{\rm C}, Y_{\rm H}), \tag{96}$$

and

$$\text{under } \mathcal{H} = 0\text{:} \quad \left(X^n, Y_{\rm C}^n, Y_{\rm H}^n, Z^n\right) \text{ i.i.d. } \sim P_{XY_{\rm C}Y_{\rm H}Z}, \tag{97}$$
$$\text{under } \mathcal{H} = 1\text{:} \quad \left(X^n, Y_{\rm C}^n, Y_{\rm H}^n, Z^n\right) \text{ i.i.d. } \sim P_{X|Y_{\rm C}} \cdot P_{Y_{\rm C}Y_{\rm H}Z}. \tag{98}$$

As the following corollary shows, in this case the optimal strategy is to let the relay decide on the hypothesis, and let the receiver simply follow this decision. It thus suffices that the relay forwards its decision to the receiver. No quantization information is needed at the receiver.

Corollary 3: Assume (96)–(98) hold. The exponent-rate region $\mathcal{E}^*(R,T)$ is the set of all nonnegative pairs (θ_y, θ_z) that satisfy

$$\theta_y \leq I(U;Y_{\rm H}|Y_{\rm C}) \tag{99}$$
$$\theta_z \leq I(U;Y_{\rm H}|Y_{\rm C}), \tag{100}$$

for an auxiliary random variable U satisfying the Markov chain $U \to X \to (Y, Z)$ and the rate constraint

$$R \geq I(U;X|Y_{\rm C}), \tag{101}$$

where $Y = (Y_{\rm C}, Y_{\rm H})$ and $(X, Y_{\rm C}, Y_{\rm H}, Z_{\rm C}, Z_{\rm H}) \sim P_{XY_{\rm C}Y_{\rm H}Z}$. (No constraint involves the rate T.)

Proof: Achievability follows by specializing Proposition 2 to V being a constant and simplifying the expressions. The converse is similar to the proof of the converse to Corollary 2. ∎

Remark 5: If we set $Y_{\rm C}$ to a constant, then the problem becomes one where

$$P_{XYZ} = P_{X|Y} \cdot P_{YZ} \tag{102}$$

$$Q_{XYZ} = P_X \cdot P_{YZ}. \tag{103}$$

The exponent-rate region then becomes the set of all nonnegative pairs (θ_y, θ_z) that satisfy

$$\theta_y \le I(U;Y) \tag{104}$$
$$\theta_z \le I(U;Y), \tag{105}$$

for an auxiliary random variable U satisfying the Markov chain $U \to X \to (Y,Z)$ and the rate constraint

$$R \ge I(U;X). \tag{106}$$

The region is achievable using Proposition 1.

VI. The Special Case "$X^n \to Z^n \to Y^n$ Under Both Hypotheses"

We consider a setup where $X^n \to Z^n \to Y^n$ forms a Markov chain under both hypotheses. This setting models a situation where the receiver lies in between the transmitter and the relay, and thus the signals at the sensor and the relay are conditionally independent given the signal at the receiver (decision center). The two-hop network can still be an adequate communication model if all the communication from the transmitter to the receiver needs to be directed through the relay. This is for example the case in cellular systems where the relay is associated with a base station.

We treat two special cases: 1) same P_{YZ} under both hypotheses, and 2) same P_{XZ} under both hypotheses. Combined with the Markov chain $X \to Z \to Y$, these assumptions seem to suggest that the receiver cannot improve its error exponent by learning information about the observations at the relay (for case 1) or about the observations at the transmitter (for case 2). As we shall see, this holds only if the rates of communication are sufficiently high. Otherwise, information about observations at both the transmitter and the relay can be combined to reduce the required rate of communication and thus also improve the performance of the system. In this section we shall not employ binning, i.e., all achievability results below follow from Theorem 1.

A. Special Case 1: Same P_{YZ} under both Hypotheses

Consider first the setup where the pmfs P_{XYZ} and Q_{XYZ} decompose as

$$P_{XYZ} = P_{X|Z} \cdot P_{YZ}, \tag{107}$$
$$Q_{XYZ} = P_X \cdot P_{YZ}. \tag{108}$$

Since the pair of sequences (Y^n, Z^n) has the same joint distribution under both hypotheses, no positive error exponent θ_z is possible when the message B sent from the relay to the receiver is only a function of Y^n but not of the incoming message M. The structure of (107) and (108) might even suggest that Y^n was not useful at the receiver and that the relay should simply forward a function of its incoming message M. Proposition 3 shows that this strategy is optimal when $T \ge R$, i.e., when the relay can forward the entire message to the receiver. On the other hand, Example 1 shows that it can be suboptimal when $T < R$.

Proposition 3: Assume conditions (107) and (108) and

$$T \ge R. \tag{109}$$

Then the exponent-rate region $\mathcal{E}(R,T)$ is the set of all nonnegative pairs (θ_y, θ_z) that satisfy

$$\theta_y \le I(S;Y) \tag{110}$$
$$\theta_z \le I(S;Z) \tag{111}$$

for some auxiliary random variable S satisfying the Markov chain $S \to X \to (Y,Z)$ and the rate constraint

$$R \ge I(S;X), \tag{112}$$

where $(X,Y,Z) \sim P_{X|Z} \cdot P_{YZ}$.

Proof: For achievability, specialize Theorem 1 to $S = U = V$. The converse is proved in Appendix H. ∎

We next consider an example that satisfies assumptions (107) and (108), but not (109). We assume $R \ge H(X)$, so the transmitter can reliably describe the sequence X^n to the relay. When $T \ge R$, by Proposition 3, the optimal strategy at the relay is to forward the incoming message $B = M$, i.e., to describe the entire X^n to the receiver. In this example, to achieve the same exponent, it suffices that the relay describes only part of X^n, the choice of which depends on the relay's own observations Y^n. Thus, the relay only requires a rate T that is smaller than R.

Example 1: Let under both hypotheses $\mathcal{H} = 0$ and $\mathcal{H} = 1$:

$$X \sim \mathcal{B}(1/2) \quad \text{and} \quad Y \sim \mathcal{B}(1/2)$$

be independent of each other. Also, let $N \sim \mathcal{B}(1/2)$ be independent of the pair (X,Y), and

$$Z = (Z', Y) \qquad \text{where} \quad Z' = \begin{cases} X & \text{if } Y = 0 \text{ and } \mathcal{H} = 0 \\ N & \text{otherwise.} \end{cases}.$$

Let P_{XYZ} denote the joint pmf under $\mathcal{H} = 0$ and Q_{XYZ} the joint pmf under $\mathcal{H} = 1$.

Notice that the triple (X,Y,Z) satisfies conditions (107) and (108). Moreover, since $P_{XY} = Q_{XY}$, the error exponent θ_y cannot be larger than zero, and we focus on the error exponent θ_z achievable at the receiver. Notice that the conditional pmf

$$P_{XZ|Y=1} = Q_{XZ|Y=1}. \tag{113}$$

The idea of our scheme is thus that the relay describes only the symbols

$$\{X_t \colon t \in \{1,\dots,n\}, Y_t = 0\} \tag{114}$$

to the receiver. All other symbols are useless for distinguishing the two hypotheses. Specifically, we specialize the scheme in Subsection III-A to the choice of random variables

$$S \quad \text{a constant} \tag{115a}$$
$$U = X \tag{115b}$$
$$V = \begin{cases} U & \text{if } Y = 0 \\ U' & \text{otherwise,} \end{cases} \tag{115c}$$

where $U' \sim \mathcal{B}(1/2)$ is independent of all other random variables. Evaluating Theorem 1 for this choice proves achievability of the following error exponent at the receiver:

$$\begin{aligned}
&\min_{\substack{\tilde{P}_{VXYZ}:\\ \tilde{P}_{VXY}=P_{VXY}\\ \tilde{P}_{VZ}=P_{VZ}}} D(\tilde{P}_{VXYZ}\|P_{V|XY}Q_{XYZ}) \\
&\quad \overset{(a)}{\geq} D(P_{VZ}\|Q_{VZ}) && (116)\\
&\quad = D(P_{VYZ'}\|Q_{VYZ'}) && (117)\\
&\quad \overset{(b)}{=} \mathbb{E}_{P_Y}[D(P_{VZ'|Y}\|Q_{VZ'|Y})] && (118)\\
&\quad \overset{(c)}{=} P_Y(0)\cdot D(P_{VZ'|Y=0}\|Q_{VZ'|Y=0}) && (119)\\
&\quad = P_Y(0)\cdot I(Z';V|Y=0) && (120)\\
&\quad = P_Y(0)\cdot I(X;V|Y=0) && (121)\\
&\quad = 1/2H(X) = 1/2, && (122)
\end{aligned}$$

where the pmfs P_{VXY}, P_{VZ}, $P_{VYZ'}$ and the pmfs Q_{VZ}, $Q_{VYZ'}$ are obtained from the definitions in (115) and the pmfs P_{XYZ} and Q_{XYZ}, respectively, and mutual informations are calculated according to the joint pmf $P_{VXYZ'}$ defined through (115) and P_{XYZ}. In the above, (a) holds by the data-processing inequality, and by the second condition in the minimization; (b) holds by the chain rule of KL-divergence and because $P_Y = Q_Y$; and (c) holds because $Q_{VZ'|Y=0} = P_{V|Y=0}\cdot P_{Z'|Y=0}$ whereas $Q_{VZ'|Y=1} = P_{VZ'|Y=1}$.

The scheme requires rates

$$R = H(X) = 1$$

and

$$T = I(V;Y,U) \overset{(d)}{=} I(V;X|Y) \overset{(e)}{=} P_Y(0)\cdot I(V;X|Y=0) = 1/2,$$

where (d) holds because V is independent of Y and (e) holds because V is also independent of X unless $Y = 0$.

The error exponent in (116) coincides with the largest exponent $D(P_{XYZ}\|Q_{XYZ})$ that is possible even in a fully centralized setup. We argue in the following that, provided $R = 1$ and $T < 1$, this error exponent cannot be achieved when the relay simply sends a function of the message M to the receiver. Notice that the setup incorporating only the transmitter and the receiver is a standard "testing against independence" two-terminal setup [1] with largest possible exponent equal to:

$$\begin{aligned}
&\max_{P_{S|X}\,:\,T\geq I(S;X)} I(S;Z)\\
&\quad \overset{(f)}{=} \max_{P_{S|X}\,:\,T\geq I(S;X)} I(S;Z'|Y)\\
&\quad = 1 - \min_{P_{S|X}\,:\,T\geq I(S;X)} H(Z'|Y,S)\\
&\quad = 1 - \min_{P_{S|X}\,:\,H(X|S)\geq 1-T} \frac{1}{2}\cdot H(X|S) - \frac{1}{2}\\
&\quad \leq \frac{1}{2}T, && (123)
\end{aligned}$$

where (f) holds because $Z = (Z', Y)$ and because (X, S) are independent of Y. This shows that the optimal exponent $1/2$ cannot be achieved if the relay simply sends a function of the incoming message whenever $T < 1$.

B. Special Case 2: Same P_{XZ} under both Hypotheses

Consider next a setup where

$$P_{XYZ} = P_{XZ}\cdot P_{Y|Z}, \quad (124)$$

$$Q_{XYZ} = P_{XZ}\cdot P_Y. \quad (125)$$

Notice that the pair of sequences (X^n, Z^n) has the same joint pmf under both hypotheses. Thus, when the relay simply forwards the incoming message M without conveying additional information about its observation Y^n to the receiver, no positive error exponent θ_z is possible. On the contrary, as the following proposition shows, if

$$T \geq H(Y), \quad (126)$$

then forwarding message M to the receiver is useless, and it suffices that the message B sent from the relay to the receiver describes Y^n. In other words, under constraint (126), the optimal error exponent θ_z coincides with the optimal error exponent of a point-to-point system that consists only of the relay and the receiver. The three-terminal multi-hop setup with a transmitter observing X^n can however achieve larger error exponent θ_z than the point-to-point system when (126) does not hold. This is shown through Example 2 ahead.

Proposition 4: Assume (124)–(126). Under these assumptions, the exponent-rate region $\mathcal{E}^*(R,T)$ is the set of all nonnegative pairs (θ_y, θ_z) that satisfy

$$\theta_y \leq I(U;Y), \quad (127)$$

$$\theta_z \leq I(Y;Z), \quad (128)$$

for some auxiliary random variable U satisfying the Markov chain $U \to X \to (Y, Z)$ and the rate constraint

$$R \geq I(U;X), \quad (129)$$

where $(X, Y, Z) \sim P_{XZ}\cdot P_{Y|Z}$.

Proof: Achievability follows by specializing Theorem 1 to $S = U$ and $V = Y$. The converse for (127) is the same as in the point-to-point setting (without receiver). The converse for (128) follows by Stein's lemma (without communication constraints) [29]. ■

We next consider an example where assumptions (124) and (125) hold, but not (126).

Example 2: Let under both hypotheses $\mathcal{H} = 0$ and $\mathcal{H} = 1$:

$$X \sim \mathcal{B}(1/2) \quad \text{and} \quad Y \sim \mathcal{B}(1/2)$$

be independent of each other. Also, let $N \sim \mathcal{B}(1/2)$ be independent of the pair (X, Y), and

$$Z = (Z', X) \qquad \text{where} \quad Z' = \begin{cases} Y & \text{if } X = 0 \text{ and } \mathcal{H} = 0\\ N & \text{otherwise.}\end{cases}$$

The described triple (X, Y, Z) satisfies conditions (124) and (125). Moreover, since the pmf of the sequences (X^n, Y^n) is the same under both hypotheses, the best error exponent θ_y is zero, so we focus on the receiver's error exponent θ_z. By

Proposition 4, the largest error exponent θ_z that is achievable is

$$\theta_z^\star = I(Y;Z) = I(Y;Z'|X) = 1/2. \tag{130}$$

As we show in the following, $\theta_z^\star$ is achievable with $T = 1/2$. To see this, notice that

$$P_{YZ|X=1} = Q_{YZ|X=1}. \tag{131}$$

It thus suffices that the relay conveys the values of its observations $\{Y_t \colon t \in \{1,\dots,n\}, X_t = 0\}$ to the receiver. This is achieved by specializing the coding and testing scheme of Subsection III-A to the choice of S being a constant and

$$U = \begin{cases} 0 & \text{if } X = 0 \\ 1 & \text{otherwise} \end{cases}$$

$$V = \begin{cases} Y & \text{if } U = 0 \\ Y' & \text{otherwise,} \end{cases}$$

where $Y' \sim \mathcal{B}(1/2)$ is independent of (X,Y,Z). By Theorem 1, the scheme requires rates

$$R = I(U;X) = H(U) = 1$$

and

$$T = I(V;Y,U) = P_U(0) \cdot I(V;Y|U=0) = 1/2.$$

It achieves the optimal error exponent $\theta_z^\star$ in (130):

$$\begin{aligned} \min_{\substack{\tilde{P}_{UVXYZ}\colon \\ \tilde{P}_{UX}=P_{UX} \\ \tilde{P}_{UVY}=P_{UVY} \\ \tilde{P}_{VZ}=P_{VZ}}} & D(\tilde{P}_{UVXYZ} \| P_{U|X} P_{V|UY} Q_{XYZ}) \\ &\overset{(a)}{\geq} D(P_{VZ} \| Q_{VZ}) \qquad (132) \\ &= \mathbb{E}_{P_X}\left[D(P_{VZ'|X} \| Q_{VZ'|X}) \right] \\ &\overset{(b)}{=} P_X(0) D(P_{VZ'|X=0} \| P_{V|X=0} P_{Z'|X=0}) \\ &= P_X(0) I(V;Z'|X=0) \\ &= 1/2, \qquad (133) \end{aligned}$$

where (a) holds by the data-processing inequality for KL-divergences and by defining Q_{VZ} to be the marginal of the joint pmf $P_{U|X}P_{V|UY}Q_{XYZ}$; and (b) holds because $Q_{VZ'|X=0} = P_{V|X=0}P_{Z'|X=0}$ whereas $Q_{VZ'|X=1} = P_{VZ'|X=1}$.

Using similar arguments as in Example 1, it can be shown that the optimal error exponent $\theta_z^\star$ in (130) cannot be achieved without the transmitter's help when $T < 1$.

VII. Concluding Remarks

The paper presents coding and testing schemes for a two-hop relay network, and the corresponding exponent-rate region. The schemes combine cascade source coding with a unanimous decision-forwarding strategy where the receiver decides on the null hypothesis only if both the transmitter and relay have decided on it. The schemes are shown to attain the entire exponent-rate region for some cases of testing against independence or testing against conditional independence when the Markov chain $X^n \to Y^n \to Z^n$ holds. In these cases, the source coding part of our coding schemes simplifies to independent source codes for the transmitter-to-relay link and for the relay-to-receiver link. The proposed schemes are also shown to be optimal in some special cases when the Markov chain $X^n \to Z^n \to Y^n$ holds. For large enough communication rates and when testing against independence, it is again optimal to employ independent source codes for the two links. But, when the rate on the relay-to-receiver link is small, this simplification can be suboptimal.

One of our coding schemes employs binning to decrease the required rates of communication. Binning makes the characterization of the achievable exponent region much more involved. For the proposed scheme we have two competing exponents for the error exponent at the relay and four competing exponents for the error exponent at the receiver. Notice that, in our scheme, we only bin the satellite codebooks but not the cloud-center codebooks. Further performance improvement might be obtained by binning also the cloud center; this would however lead to an expression with ten competing exponents at the receiver.

VIII. Acknowledgement

The authors would like to thank Pierre Escamilla and Abdellatif Zaidi for helpful discussions.

References

[1] R. Ahlswede and I. Csiszàr, "Hypothesis testing with communication constraints," *IEEE Trans. on Info. Theory*, vol. 32, pp. 533–542, Jul. 1986.

[2] T. S. Han, "Hypothesis testing with multiterminal data compression," *IEEE Trans. on Info. Theory*, vol. 33, no. 6, pp. 759–772, Nov. 1987.

[3] H. Shimokawa, T. Han, and S. I. Amari, "Error bound for hypothesis testing with data compression," in *Proc. IEEE Int. Symp. on Info. Theory*, Jul. 1994, p. 114.

[4] M. S. Rahman and A. B. Wagner, "On the optimality of binning for distributed hypothesis testing," *IEEE Trans. on Info. Theory*, vol. 58, no. 10, pp. 6282–6303, Oct. 2012.

[5] S. Salehkalaibar, M. Wigger, and R. Timo, "On hypothesis testing against independence with multiple decision centers," *IEEE Trans. on Communications*, vol. 66, no. 6, pp. 2409–2420, Jan. 2018.

[6] M. Wigger and R. Timo, "Testing against independence with multiple decision centers," in *Proc. of SPCOM*, Banglore, India, Jun. 2016.

[7] P. Escamilla, M. Wigger, and A. Zaidi, "Distributed hypothesis testing with concurrent detection," in *Proc. IEEE Int. Symp. on Info. Theory*, Jun. 2018, pp. 166–170.

[8] K. R. Varshney and L. R. Varshney, "Quantization of prior probabilities for hypothesis testing," *IEEE Trans. on Sig. Proc.*, vol. 56, no. 10, pp. 4553–4562, Oct. 2008.

[9] Y. Li, S. Nitinawarat, and V. V. Veeravalli, "Universal outlier hypothesis testing," in *Proc. IEEE Int. Symp. on Info. Theory*, Istanbul, Turkey, Jun. 2013, pp. 2666–2670.

[10] M. Naghshvar and T. Javidi, "Active m-ary sequential hypothesis testing," in *Proc. IEEE Int. Symp. on Info. Theory*, Austin, Texas, Jun. 2010, pp. 1623–1627.

[11] H. V. Poor, *An introduction to signal detection and estimation*. Springer, 1994.

[12] K. G. Nagananda and C. R. Murthy, "A hypothesis test for topology change detection in wireless sensor networks," in *Proc. IEEE Global Comm. Conf.*, Dec. 2017, pp. 1–6.

[13] W. Zhao and L. Lai, "Distributed testing against independence with multiple terminals," in *Proc. 52nd Allerton Conf. Comm, Cont. and Comp.*, Monticello, IL, USA, Oct. 2014, pp. 1246–1251.

[14] ——, "Distributed testing with cascaded encoders," *IEEE Trans. on Info. Theory*, 2018.

[15] Y. Xiang and Y. H. Kim, "Interactive hypothesis testing against independence," in *Proc. IEEE Int. Symp. on Info. Theory*, Istanbul, Turkey, Jun. 2013, pp. 2840–2844.

[16] S. Sreekuma and D. Gündüz, "Distributed hypothesis testing over discrete memoryless channels," 2018. [Online]. Available: https://arxiv.org/abs/1802.07665

[17] S. Salehkalaibar and M. Wigger, "Distributed hypothesis testing based on unequal-error protection codes," 2018. [Online]. Available: https://arxiv.org/abs/1806.05533

[18] H. Yamamoto, "Source coding theory for cascade and branching communication systems," *IEEE Trans. on Info. Theory*, vol. 27, no. 3, pp. 299–308, May 1981.

[19] P. Cuff, H. I. Su, and A. El Gamal, "Cascade multiterminal source coding," in *Proc. IEEE Int. Symp. on Info. Theory*, Seoul, Korea, Jun. 2009, pp. 1199–1203.

[20] D. Vasudevan, C. Tian, and S. Diggavi, "Lossy source coding for a cascade communication system with side information," in *Proc. Allerton Conf. Comm, Cont. and Comp.*, Monticello, IL, USA, Sept. 2006.

[21] P. Cuff, H. H. Permuter, and T. M. Cover, "Coordination capacity," *IEEE Trans. on Info. Theory*, vol. 56, no. 9, pp. 4181–4206, Sept. 2010.

[22] R. Tandon, S. Mohajer, and H. V. Poor, "Cascade source coding with erased side information," in *Proc. IEEE Int. Symp. on Info. Theory*, St. Petersburg, Russia, Jul.-Aug. 2011, pp. 2944–2948.

[23] Y. K. Chia, H. H. Permuter, and T. Weissman, "Cascade, triangular, and two-way source coding with degraded side information at the second user," *IEEE Trans. on Info. Theory*, vol. 58, no. 1, pp. 189–206, Jan. 2012.

[24] H. H. Permuter and T. Weissman, "Cascade and triangular source coding with side information at the first two nodes," *IEEE Trans. on Info. Theory*, vol. 58, no. 6, pp. 3339–3349, Feb. 2012.

[25] B. Ahmadi, R. Tandon, O. Simeone, and H. V. Poor, "Heegard-berger and cascade source coding problems with common reconstruction constraints," *IEEE Trans. on Info. Theory*, vol. 59, no. 3, pp. 1458–1474, Nov. 2012.

[26] D. Cao, L. Zhou, and V. Y. F. Tan, "Strong converse for hypothesis testing against independence over a two-hop network," 2018. [Online]. Available: https://arxiv.org/abs/1808.05366

[27] H. Shimokawa, T. Han, and S. I. Amari, "Asymptotic error bound of hypothesis testing under multiterminal data compression," in *Proc. 16th Symposium on Information Theory and Its Applications*, Oct. 1993, pp. 87–90.

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

[29] T. M. Cover and J. A. Thomas, *Elements of Information Theory, 2nd Ed.* Wiley, 2006.

[30] I. Csiszar and J. Korner, *Information theory: coding theorems for discrete memoryless systems*. Cambridge University Press, 2011.

APPENDIX A
PROOF OF THEOREM 1

We bound the probabilities of error of the scheme averaged over the random code construction $\mathcal{C}$. The analysis of the error probabilities at the relay is standard. We therefore focus on the error probabilities at the receiver.

If $M \neq 0$ and $B \neq 0$, let I, J, K be the random indices sent over the bit pipes and define the following events:

$$\begin{aligned}\mathcal{E}_{\text{Relay}}: &\quad \left\{(S^n(I), U^n(J|I), Y^n) \notin \mathcal{T}^n_{\mu/2}(P_{SUY})\right\},\\ \mathcal{E}_{\text{Rx}}: &\quad \left\{(S^n(I), U^n(J|I), V^n(K|I), Z^n) \notin \mathcal{T}^n_{\mu}(P_{SUVZ})\right\}.\end{aligned}$$

The type-I error probability at the receiver averaged over the random code construction can be bounded, for large enough n, as follows

$$\begin{aligned}\mathbb{E}_{\mathcal{C}}\left[\alpha_{z,n}\right] &\leq \Pr[M=0 \text{ or } B=0 \text{ or } \mathcal{E}_{\text{Relay}} \text{ or } \mathcal{E}_{\text{Rx}}|\mathcal{H}=0]\\ &\leq \Pr[M=0|\mathcal{H}=0]\\ &\quad + \Pr[B=0 \text{ or } \mathcal{E}_{\text{Relay}}|M\neq 0, \mathcal{H}=0]\\ &\quad + \Pr[\mathcal{E}_{\text{Rx}}|M\neq 0, B\neq 0, \mathcal{H}=0] \qquad (134)\\ &\overset{(a)}{\leq} \epsilon/32 + \Pr[\mathcal{E}_{\text{Relay}}|M\neq 0, \mathcal{H}=0]\\ &\quad + \Pr[B=0|M\neq 0, \mathcal{E}^c_{\text{Relay}}, \mathcal{H}=0] + \epsilon/32 \qquad (135)\\ &\overset{(b)}{\leq} \epsilon/32 + \epsilon/32 + \epsilon/32 + \epsilon/32 \qquad (136)\\ &= \epsilon/8, \qquad (137)\end{aligned}$$

where (a) holds by the covering lemma and the rate constraint (16), and both (a) and (b) hold by the Markov lemma [28].

We now bound the probability of type-II error at the receiver. Let $\mathcal{P}^n$ be the set of all types over the product alphabets $\mathcal{S}^n \times \mathcal{U}^n \times \mathcal{V}^n \times \mathcal{X}^n \times \mathcal{Y}^n \times \mathcal{Z}^n$. Also, let $\mathcal{P}^n_\mu$ be the subset of types $\pi_{SUVXYZ} \in \mathcal{P}^n$ that simultaneously satisfy the following three conditions:

$$|\pi_{SUX} - P_{SUX}| \leq \mu/4, \qquad (138)$$

$$|\pi_{SUVY} - P_{SUVY}| \leq \mu/2, \qquad (139)$$

$$|\pi_{SVZ} - P_{SVZ}| \leq \mu. \qquad (140)$$

Now, consider the type-II error probability averaged over the random code construction. For all $(i,j,k) \in \{1,\dots,\lfloor 2^{nR_s}\rfloor\} \times \{1,\dots,\lfloor 2^{nR_u}\rfloor\} \times \{1,\dots,\lfloor 2^{nR_v}\rfloor\}$, define the events:

$$\mathcal{E}_{\text{Tx}}(i,j) = \left\{(S^n(i), U^n(j|i), X^n) \in \mathcal{T}^n_{\mu/4}(P_{SUX})\right\}, \qquad (141)$$

$$\mathcal{E}_{\text{Rel}}(i,j,k) = \left\{(S^n(i), U^n(j|i), V^n(k|i), Y^n) \in \mathcal{T}^n_{\mu/2}(P_{SUVY})\right\}, \qquad (142)$$

$$\mathcal{E}_{\text{Rx}}(i,k) = \left\{(S^n(i), V^n(k|i), Z^n) \in \mathcal{T}^n_{\mu}(P_{SVZ})\right\}. \qquad (143)$$

We have

$$\begin{aligned}\mathbb{E}_{\mathcal{C}}\left[\beta_{z,n}\right] &= \Pr\left[\hat{\mathcal{H}}_z = 0 \middle| \mathcal{H}=1\right]\\ &\leq \Pr\left[\cup_{i,j,k}\ (\mathcal{E}_{\text{Rx}}(i,k)\ \cap\ \mathcal{E}_{\text{Rel}}(i,j,k)\ \cap \mathcal{E}_{\text{Tx}}(i,j))\ \middle|\ \mathcal{H}=1\right]. \qquad (144)\end{aligned}$$

(The inequality in (144) comes from the fact that the transmitter chooses the pair of indices (i,j) uniformly at random over all pairs for which event $\mathcal{E}_{\text{Tx}}(i,j)$ holds. There can thus exist a triple (i',j',k') satisfying $(\mathcal{E}_{\text{Rx}}(i',k')\ \cap\ \mathcal{E}_{\text{Rel}}(i',j',k')\ \cap \mathcal{E}_{\text{Tx}}(i',j'))$ but the receiver still decides on $\hat{\mathcal{H}}_z = 1$ because the transmitter chose a pair (i,j) for which $(\mathcal{E}_{\text{Rel}}(i,j,k)\ \cap \mathcal{E}_{\text{Tx}}(i,j))$ is violated for all values of k.)

We continue by applying the union bound:

$$\begin{aligned}&\Pr\left[\cup_{i,j,k}\ (\mathcal{E}_{\text{Rx}}(i,k)\ \cap\ \mathcal{E}_{\text{Rel}}(i,j,k)\ \cap \mathcal{E}_{\text{Tx}}(i,j))\ \middle|\ \mathcal{H}=1\right]\\ &\leq \sum_{i,j,k} \Pr\left[\mathcal{E}_{\text{Rx}}(i,k)\ \cap\ \mathcal{E}_{\text{Rel}}(i,j,k)\ \cap \mathcal{E}_{\text{Tx}}(i,j)\ \middle|\ \mathcal{H}=1\right]\\ &= \sum_{i,j,k} \Pr\Bigg[(S^n(i), V^n(k|i), Z^n) \in \mathcal{T}^n_{\mu}(P_{SVZ}),\\ &\qquad (S^n(i), U^n(j|i), V^n(k|i), Y^n) \in \mathcal{T}^n_{\mu/2}(P_{SUVY}),\\ &\qquad (S^n(i), U^n(j|i), X^n) \in \mathcal{T}^n_{\mu/4}(P_{SUX})\ \Big|\ \mathcal{H}=1\Bigg]\end{aligned}$$

$$= \sum_{i,j,k} \sum_{\substack{\pi_{SUVXYZ} \\ \in \mathcal{P}_\mu^n}} \Pr\Bigg[\mathrm{tp}\left(S^n(i), U^n(j|i), V^n(k|i), X^n, Y^n, Z^n\right) = \pi_{SUVXYZ} \Big| \mathcal{H}=1\Bigg]$$

$$\leq 2^{n(R_s+R_u+R_v)} \cdot \ (n+1)^{|\mathcal{S}|\cdot|\mathcal{U}|\cdot|\mathcal{V}|\cdot|\mathcal{X}|\cdot|\mathcal{Y}|\cdot|\mathcal{Z}|} \cdot \max_{\pi_{USVXYZ} \in \mathcal{P}_\mu^n} 2^{-n(D(\pi_{SUVXYZ}\|P_{SU}P_{V|S}Q_{XYZ})-\mu)}, \quad (145)$$

where the last inequality holds by Sanov's theorem [29]. Indeed, by the code construction, the three codewords $(S^n(i), U^n(j|i),\ V^n(k|i))$ are drawn i.i.d. according to $P_{SU}P_{V|S}$. Furthermore, they are independent of (X^n, Y^n, Z^n), which, under $\mathcal{H}=1$, are drawn i.i.d. according to Q_{XYZ}. Therefore,

$$\mathbb{E}_{\mathcal{C}}\left[\beta_{z,n}\right] \leq (n+1)^{|\mathcal{S}|\cdot|\mathcal{U}|\cdot|\mathcal{V}|\cdot|\mathcal{X}|\cdot|\mathcal{Y}|\cdot|\mathcal{Z}|} \times \max_{\pi_{USVXYZ} \in \mathcal{P}_\mu^n} \left[2^{n(R_s+R_u+R_v-D(\pi_{SUVXYZ}\|P_{SU}P_{V|S}Q_{XYZ})+\mu)}\right]. \quad (146)$$

Plugging the rate expressions (16)–(18) into (146) results in the following upper bound:

$$\mathbb{E}_{\mathcal{C}}\left[\beta_{z,n}\right] \leq (n+1)^{|\mathcal{S}|\cdot|\mathcal{U}|\cdot|\mathcal{V}|\cdot|\mathcal{X}|\cdot|\mathcal{Y}|\cdot|\mathcal{Z}|} \cdot 2^{-n\theta_{z,\mu}}, \quad (147)$$

where

$$\theta_{z,\mu} := \min_{\pi_{USVXYZ} \in \mathcal{P}_\mu^n} \big[D(\pi_{SUVXYZ}\|P_{SU}P_{V|S}Q_{XYZ}) - I(X;S,U) - I(Y,U;V|S) - \mu\big]. \quad (148)$$

Now, by simplifying terms and employing the continuity of KL-divergence, we conclude that

$$\mathbb{E}_{\mathcal{C}}[\beta_{z,n}] \leq 2^{-n(\theta_z - \delta_n(\mu))}, \quad (149)$$

for some function $\delta_n(\mu)$ that tends to 0 as $n \to \infty$ and $\mu \to 0$, and

$$\begin{aligned}
\theta_z :=& \min_{\substack{\tilde{P}_{SUVXYZ}: \\ \tilde{P}_{SUX}=P_{SUX} \\ \tilde{P}_{SVUY}=P_{SVUY} \\ \tilde{P}_{SVZ}=P_{SVZ}}} D(\tilde{P}_{SUVXYZ}\|P_{SU}P_{V|S}Q_{XYZ}) \\
& \qquad -I(X;S,U) - I(Y,U;V|S) \\
=& \min_{\substack{\tilde{P}_{SUVXYZ}: \\ \tilde{P}_{SUX}=P_{SUX} \\ \tilde{P}_{SVUY}=P_{SVUY} \\ \tilde{P}_{SVZ}=P_{SVZ}}} \sum_{s,u,v,x,y,z} \Bigg[\tilde{P}_{SUVXYZ}(s,u,v,x,y,z)\times \\
& \qquad \log \frac{\tilde{P}_{SUVXYZ}(s,u,v,x,y,z)}{P_{SU}(s,u)P_{V|S}(v|s)Q_{XYZ}(x,y,z)} \\
& \qquad -P_{SUX}(s,u,x)\log\frac{P_{SU|X}(s,u|x)}{P_{SU}(s,u)} \\
& \qquad -P_{SUVY}(s,u,v,y)\log\frac{P_{V|SUY}(v|s,u,y)}{P_{V|S}(v|s)}\Bigg] \\
\overset{(c)}{=}& \min_{\substack{\tilde{P}_{SUVXYZ}: \\ \tilde{P}_{SUX}=P_{SUX} \\ \tilde{P}_{SVUY}=P_{SVUY} \\ \tilde{P}_{SVZ}=P_{SVZ}}} \sum_{s,u,v,x,y,z} \tilde{P}_{SUVXYZ}(s,u,v,x,y,z)\times \\
& \qquad \log \frac{\tilde{P}_{SUVXYZ}(s,u,v,x,y,z)}{P_{SU|X}(s,u|x)P_{V|SUY}(v|s,u,y)Q_{XYZ}(x,y,z)} \\
=& \min_{\substack{\tilde{P}_{SUVXYZ}: \\ \tilde{P}_{SUX}=P_{SUX} \\ \tilde{P}_{SVUY}=P_{SVUY} \\ \tilde{P}_{SVZ}=P_{SVZ}}} D\left(\tilde{P}_{SUVXYZ}\|P_{SU|X}P_{V|SUY}Q_{XYZ}\right),
\end{aligned} \quad (150)$$

where (c) follows from the first and second constraints on the minimization and by re-arranging terms.

To summarize, we showed that on average (over the random codebook constructions $\mathcal{C}$) and for sufficiently large n:

$$\mathbb{E}_{\mathcal{C}}[\alpha_{z,n}] \leq \frac{\epsilon}{8} \quad (151)$$

$$\mathbb{E}_{\mathcal{C}}[\beta_{z,n}] \leq 2^{-n(\theta_z - \delta_n(\mu))}. \quad (152)$$

Similar arguments can be employed to show that also

$$\mathbb{E}_{\mathcal{C}}[\alpha_{y,n}] \leq \frac{\epsilon}{4} \quad (153)$$

$$\mathbb{E}_{\mathcal{C}}[\beta_{y,n}] \leq 2^{-n(\theta_y - \tilde{\delta}_n(\mu))}, \quad (154)$$

for some function $\tilde{\delta}_n(\mu)$ that tends to 0 as $n \to \infty$ and as $\mu \to 0$, and for

$$\theta_y := \min_{\substack{\tilde{P}_{SUXY}: \\ \tilde{P}_{SUX}=P_{SUX} \\ \tilde{P}_{SUY}=P_{SUY}}} D\left(\tilde{P}_{SUXY}\|P_{SU|X}Q_{XY}\right). \quad (155)$$

We now argue that for all sufficiently large blocklengths n there must exist at least one deterministic code construction $\mathcal{C}_n^*$ and a function $\hat{\delta}_n(\mu)$ that tends to 0 as $n \to \infty$ and as $\mu \to 0$, such that for this code:

$$\alpha_{y,n} \leq \epsilon \quad (156a)$$

$$\alpha_{z,n} \leq \epsilon \quad (156b)$$

$$\beta_{y,n} \leq 2^{-n(\theta_y - \hat{\delta}_n(\mu))} \quad (156c)$$

$$\beta_{z,n} \leq 2^{-n(\theta_z - \hat{\delta}_n(\mu))}. \quad (156d)$$

To this end, we start by eliminating a set of code constructions that yield largest $\alpha_{y,n}$. The size of the set is chosen such that its total probability is at least $1/2$ and at most $3/4$. (Instead of $3/4$, one can choose a value that is arbitrarily close to $1/2$. Such a choice is always feasible for sufficiently large blocklengths n, because the maximum probability of a single code construction tends to 0 as $n \to \infty$ unless all random variables are constants, but this latter case is not interesting.) Each of the code constructions in the remaining set $\mathcal{C}_1$ then has probability of type-I error

$$\alpha_{y,n} \leq \frac{\epsilon}{4} \cdot \frac{4}{3} = \frac{\epsilon}{3} \quad (157)$$

and on average these code constructions have probability of type-I error and type-II errors

$$\mathbb{E}_{\mathcal{C}_1}[\alpha_{z,n}] \leq \frac{\epsilon}{8} \cdot \frac{1}{1-\frac{3}{4}} = \frac{\epsilon}{2} \quad (158)$$

$$\mathbb{E}_{\mathcal{C}_1}[\beta_{z,n}] \leq 2^{-n(\theta_z - \delta_n(\mu))} \cdot \frac{1}{1-\frac{3}{4}} \quad (159)$$

$$\mathbb{E}_{\mathcal{C}_1}[\beta_{y,n}] \leq 2^{-n(\theta_y - \delta_n(\mu))} \cdot \frac{1}{1-\frac{3}{4}}. \quad (160)$$

In the same way we continue to eliminate a subset of $\mathcal{C}_1$ containing the code constructions with largest $\alpha_{z,n}$ such that the probability of this subset is at least $1/2$ and at most $3/4$ the probability of $\mathcal{C}_1$. Call the remaining set $\mathcal{C}_2$. From $\mathcal{C}_2$, we then eliminate code constructions that yield largest $\beta_{y,n}$, such that all the eliminated code constructions (in this step) constitute at least $1/2$ and at most $3/4$ the probability of $\mathcal{C}_2$. Finally, from the code constructions that survive all eliminations, we pick the one with the smallest $\beta_{z,n}$. This finally selected code $\mathcal{C}^*$ then satisfies

$$\alpha_{y,n} \leq \frac{\epsilon}{3} \tag{161}$$

$$\alpha_{z,n} \leq \frac{\epsilon}{2} \cdot \frac{4}{3} = \frac{2}{3}\epsilon \tag{162}$$

$$\beta_{y,n} \leq 2^{-n(\theta_y - \delta_n(\mu))} \cdot \left(\frac{1}{1-\frac{3}{4}}\right)^2 \cdot \frac{4}{3} = 2^{-n(\theta_y - \delta_n(\mu))} \cdot \frac{64}{3} \tag{163}$$

$$\beta_{z,n} \leq 2^{-n(\theta_z - \tilde{\delta}_n(\mu))} \cdot \left(\frac{1}{1-\frac{3}{4}}\right)^3 = 2^{-n(\theta_y - \tilde{\delta}_n(\mu))} \cdot 64. \tag{164}$$

If we set $\hat{\delta}_n(\mu) = \max\{\delta_n(\mu), \tilde{\delta}_n(\mu)\} + \frac{6}{n}$, then all inequalities (156) are satisfied.

Appendix B
Proof of Theorem 2

It only remains to prove (39). We analyze the probabilities of error of the coding and testing scheme described in Subsection IV-A1 averaged over the random code construction. By standard arguments (successively eliminating the worst half of the codebooks as described at the end of Appendix A) the desired result can be proved for a set of deterministic codebooks.

Fix an arbitrary $\epsilon > 0$ and the scheme's parameter $\mu > 0$. For a fixed blocklength n, let $\mathcal{P}^n_{\mu,\text{type-I}}$ be the subset of types over the product alphabet $\mathcal{S}^n \times \mathcal{S}^n \times \mathcal{Y}^n$ that satisfy the following conditions for all $(s, s', y) \in \mathcal{S} \times \mathcal{S} \times \mathcal{Y}$:

$$|\pi_{SY}(s,y) - P_{SY}(s,y)| \leq \mu, \tag{165}$$

$$|\pi_{S'}(s') - P_S(s)| \leq \mu, \tag{166}$$

$$H_{\pi_{S'Y}}(S'|Y) \leq H_{\pi_{SY}}(S|Y). \tag{167}$$

Notice that, when we let $n \to \infty$ and then $\mu \to 0$, each element in $\mathcal{P}^n_{\mu,\text{type-I}}$ will approach an element of

$$\mathcal{P}^*_{\text{type-I}} := \big\{\tilde{P}_{SS'Y} \colon \tilde{P}_{SY} = P_{SY} \text{ and } \tilde{P}_{S'} = P_S \text{ and } H_{\tilde{P}_{S'Y}}(S'|Y) \leq H_{\tilde{P}_{SY}}(S|Y)\big\}. \tag{168}$$

We first analyze the type-I error probability $\alpha_{y,n}$. For the case of $M \neq 0$, let L be the index chosen at the transmitter. Define events

$$\mathcal{E}^{(0)}_{\text{Tx}} := \big\{(S^n(m,\ell), X^n) \notin \mathcal{T}^n_{\mu/2}(P_{SX}),\ \forall (m,\ell)\big\}, \tag{169}$$

$$\mathcal{E}^{(1)}_{\text{Rx}} := \big\{(S^n(M,L), Y^n) \notin \mathcal{T}^n_{\mu}(P_{SY})\big\}, \tag{170}$$

$$\mathcal{E}^{(2)}_{\text{Rx}} := \big\{\exists \ell' \neq L\colon \ \ S^n(M,\ell') \in \mathcal{T}^n_\mu(P_S) \text{ and } H_{\text{tp}(S^n(M,L),Y^n)}(S|Y) \geq H_{\text{tp}(S^n(M,\ell'),Y^n)}(S|Y)\big\}. \tag{171}$$

For all sufficiently large n, the average type-I error probability can be bounded as:

$$\mathbb{E}_{\mathcal{C}}[\alpha_{y,n}] = \Pr\left[\hat{\mathcal{H}}_y = 1 \middle| \mathcal{H} = 0\right] \tag{172}$$

$$\leq \Pr\left[\mathcal{E}^{(0)}_{\text{Tx}} \cup \mathcal{E}^{(1)}_{\text{Rx}} \cup \mathcal{E}^{(2)}_{\text{Rx}} \middle| \mathcal{H} = 0\right] \tag{173}$$

$$\leq \Pr\left[\mathcal{E}^{(0)}_{\text{Tx}} \middle| \mathcal{H} = 0\right] + \Pr\left[\mathcal{E}^{(1)}_{\text{Rx}} \middle| \mathcal{E}^{(0)c}_{\text{Tx}}, \mathcal{H} = 0\right] + \Pr\left[\mathcal{E}^{(2)}_{\text{Rx}} \middle| \mathcal{E}^{(1)c}_{\text{Rx}}, \mathcal{E}^{(0)c}_{\text{Tx}}, \mathcal{H} = 0\right] \tag{174}$$

$$\overset{(a)}{\leq} \epsilon/6 + \Pr\left[\mathcal{E}^{(1)}_{\text{Rx}} \middle| \mathcal{E}^{(0)c}_{\text{Tx}}, \mathcal{H} = 0\right] + \Pr\left[\mathcal{E}^{(2)}_{\text{Rx}} \middle| \mathcal{E}^{(1)c}_{\text{Rx}}, \mathcal{E}^{(0)c}_{\text{Tx}}, \mathcal{H} = 0\right] \tag{175}$$

$$\overset{(b)}{\leq} \epsilon/6 + \epsilon/6 + \Pr\left[\mathcal{E}^{(2)}_{\text{Rx}} \middle| \mathcal{E}^{(1)c}_{\text{Rx}}, \mathcal{E}^{(0)c}_{\text{Tx}}, \mathcal{H} = 0\right] \tag{176}$$

$$\overset{(c)}{\leq} \epsilon/6 + \epsilon/6 + \epsilon/6 \tag{177}$$

$$= \epsilon/2, \tag{178}$$

where inequality (a) follows from the code construction; (b) follows from the Markov lemma [28]; and (c) is justified in what follows. Notice first that by the symmetry of the codebook construction, when bounding the probability $\Pr\left[\mathcal{E}^{(2)}_{\text{Rx}} \middle| \mathcal{E}^{(1)c}_{\text{Rx}}, \mathcal{E}^{(0)c}_{\text{Tx}}, \mathcal{H} = 0\right]$, we can specify $M = L = 1$ and proceed as:

$$\Pr\left[\mathcal{E}^{(2)}_{\text{Rx}} \middle| \mathcal{E}^{(1)c}_{\text{Rx}}, \mathcal{E}^{(0)c}_{\text{Tx}}, M = L = 1, \mathcal{H} = 0\right] \tag{179}$$

$$\leq \sum_{\ell'=2}^{\lfloor 2^{nR'} \rfloor} \Pr\Big[S^n(1,\ell') \in \mathcal{T}^n_\mu(P_S)\ , \ H_{\text{tp}(S^n(1,1),Y^n)}(S|Y) \geq H_{\text{tp}(S^n(1,\ell'),Y^n)}(S|Y)\ \Big| \ (S^n(1,1),Y^n) \in \mathcal{T}^n_\mu(P_{SY}), \ (S^n(1,1),X^n) \in \mathcal{T}^n_{\mu/2}(P_{SX}), \ M = L = 1,\ \mathcal{H} = 0\Big] \tag{180}$$

$$\leq \sum_{\ell'=2}^{\lfloor 2^{nR'} \rfloor} \Pr\Big[H_{\text{tp}(S^n(1,1),Y^n)}(S|Y) \geq H_{\text{tp}(S^n(1,\ell'),Y^n)}(S|Y)\ \Big| \ (S^n(1,1),Y^n) \in \mathcal{T}^n_\mu(P_{SY}), \ (S^n(1,1),X^n) \in \mathcal{T}^n_{\mu/2}(P_{SX}), \ S^n(1,\ell') \in \mathcal{T}^n_\mu(P_S)),\ M = L = 1,\ \mathcal{H} = 0\Big] \tag{181}$$

$$= \sum_{\substack{\pi_{SS'Y} \\ \in \mathcal{P}^n_{\mu,\text{type-I}}}} \sum_{\ell'=2}^{\lfloor 2^{nR'} \rfloor} \sum_{\substack{s^n, s'^n, y^n: \\ \text{tp}(s^n,s'^n,y^n) \\ = \pi_{SS'Y}}} \Pr\Big[S^n(1,1) = s^n, S^n(1,\ell') = s'^n, Y^n = y^n\ \Big| \ (S^n(1,1),Y^n) \in \mathcal{T}^n_\mu(P_{SY}), \ (S^n(1,1),X^n) \in \mathcal{T}^n_{\mu/2}(P_{SX}), \ S^n(1,\ell') \in \mathcal{T}^n_\mu(P_S)),\ \ M = L = 1,\ \ \mathcal{H} = 0\Big] \tag{182}$$

$$\stackrel{(d)}{=} \sum_{\substack{\pi_{SS'Y}\\ \in\mathcal{P}^n_{\mu,\text{type-I}}}} \sum_{\ell'=2}^{\lfloor 2^{nR'}\rfloor} \sum_{\substack{s^n,s'^n,y^n:\\ \text{tp}(s^n,s'^n,y^n)\\ =\pi_{SS'Y}}} \Pr\Big[S^n(1,1)=s^n, Y^n=y^n \,\Big|\, (S^n(1,1),Y^n)\in\mathcal{T}^n_\mu(P_{SY}), (S^n(1,1),X^n)\in\mathcal{T}^n_{\mu/2}(P_{SX}), S^n(1,\ell')\in\mathcal{T}^n_\mu(P_S),\ M=L=1,\ \mathcal{H}=0\Big] \tag{183}$$

$$\cdot \Pr\Big[S^n(1,\ell')=s'^n\Big| (S^n(1,1),Y^n)\in\mathcal{T}^n_\mu(P_{SY}), (S^n(1,1),X^n)\in\mathcal{T}^n_{\mu/2}(P_{SX}), S^n(1,\ell')\in\mathcal{T}^n_\mu(P_S),\ M=L=1,\ \mathcal{H}=0\Big] \tag{184}$$

$$\stackrel{(e)}{\le} (n+1)^{|\mathcal{S}|^2\cdot|\mathcal{Y}|} \sum_{\pi_{SS'Y}\in\mathcal{P}^n_{\mu,\text{type-I}}} \sum_{\ell'=2}^{\lfloor 2^{nR'}\rfloor} \sum_{\substack{s^n,y^n,s'^n:\\ \text{tp}(s^n,s'^n,y^n)=\pi_{SS'Y}}} 2^{-nH_\pi(S,Y)}\cdot 2^{-nH_\pi(S')} \tag{185}$$

$$\stackrel{(f)}{\le} (n+1)^{|\mathcal{S}|^2\cdot|\mathcal{Y}|} \sum_{\pi_{SS'Y}\in\mathcal{P}^n_{\mu,\text{type-I}}} \sum_{\ell'=2}^{\lfloor 2^{nR'}\rfloor} 2^{nH_\pi(S,S',Y)}\cdot 2^{-nH_\pi(S,Y)}\cdot 2^{-nH_\pi(S')} \tag{186}$$

$$= (n+1)^{|\mathcal{S}|^2\cdot|\mathcal{Y}|} \sum_{\pi_{SS'Y}\in\mathcal{P}^n_{\mu,\text{type-I}}} 2^{n(R'-I_\pi(S';Y,S))} \tag{187}$$

$$\le (n+1)^{|\mathcal{S}|^2\cdot|\mathcal{Y}|} \sum_{\pi_{SS'Y}\in\mathcal{P}^n_{\mu,\text{type-I}}} 2^{n(R'-I_\pi(S';Y))} \tag{188}$$

$$\stackrel{(g)}{\le} (n+1)^{|\mathcal{S}|^4\cdot|\mathcal{Y}|^2}\cdot \max_{\pi_{SS'Y}\in\mathcal{P}^n_{\mu,\text{type-I}}} 2^{n(R'-I(S;Y)+\delta_n(\mu))} \tag{189}$$

$$\stackrel{(h)}{\le} \epsilon/6, \tag{190}$$

where $\delta_n(\mu)$ tends to 0 as $n\to\infty$ and then $\mu\to 0$. The steps are justified as follows:

- (d) holds because conditioned on the events $(S^n(1,1),Y^n)\in\mathcal{T}^n_\mu(P_{SY})$, $(S^n(1,1),X^n)\in\mathcal{T}^n_{\mu/2}(P_{SX})$, $S^n(1,\ell')\in\mathcal{T}^n_\mu(P_S)$, $M=L=1$, $\mathcal{H}=0$, the codeword $S^n(1,\ell')$ is independent of the pair $(S^n(1,1),Y^n)$;
- (e) holds because even conditioned on the events $(S^n(1,1),Y^n)\in\mathcal{T}^n_\mu(P_{SY})$, $(S^n(1,1),X^n)\in\mathcal{T}^n_{\mu/2}(P_{SX})$, $S^n(1,\ell')\in\mathcal{T}^n_\mu(P_S))$, $M=L=1$, $\mathcal{H}=0$, all pairs (s^n,y^n) of same joint type have the same probability and all sequences s'^n of same type have the same probability, and because there are at least $\frac{1}{(n+1)^{|\mathcal{S}|\cdot|\mathcal{Y}|}}\cdot 2^{nH_{\pi_{SY}}(S,Y)}$ sequences of joint type π_{SY} [30, Lemma 2.3] and at least $\frac{1}{(n+1)^{|\mathcal{S}|}}\cdot 2^{nH_{\pi_{S'}}(S')}$ sequences of joint type $\pi_{S'}$;
- (f) because there are at most $2^{nH_\pi(S,S',Y)}$ different n-length sequences of same joint type $\pi_{SS'Y}$;
- (g) holds because $|\mathcal{P}^n_{\mu,\text{type-I}}|\le(n+1)^{|\mathcal{S}|^2\cdot|\mathcal{Y}|}$, because $H_\pi(S'|Y)\le H_\pi(S|Y)$, because each element of $\mathcal{P}^n_{\mu,\text{type-I}}$ must approach an element of $\mathcal{P}^*_{\text{type-I}}$ when $n\to\infty$ and $\mu\to 0$, and by the continuity of the entropy function; and
- (h) holds for all sufficiently large n and small μ because $R'<I(S;Y)$ and $\delta_n(\mu)\to 0$ as $n\to\infty$ and then $\mu\to 0$.

We now bound the probability of type-II error at the receiver (averaged over the random code construction). For all $m\in\{1,\dots,\lfloor 2^{nR}\rfloor\}$ and $\ell,\ell'\in\{1,\dots,\lfloor 2^{nR'}\rfloor\}$, define the following events:

$$\mathcal{E}_{\text{Tx}}(m,\ell) := \{(S^n(m,\ell),X^n)\in\mathcal{T}^n_{\mu/2}(P_{SX})\}, \tag{191}$$

$$\mathcal{E}_{\text{Rx}}(m,\ell') := \Big\{(S^n(m,\ell'),Y^n)\in\mathcal{T}^n_\mu(P_{SY}), H_{\text{tp}(S^n(m,\ell'),Y^n)}(S'|Y) = \min_{\tilde\ell} H_{\text{tp}(S^n(m,\tilde\ell),Y^n)}(S|Y)\Big\}. \tag{192}$$

Define

$$\mathcal{B}_1 := \{\exists\,(m,\ell)\colon \mathcal{E}_{\text{Tx}}(m,\ell) \text{ and } \mathcal{E}_{\text{Rx}}(m,\ell)\}, \tag{193}$$

$$\mathcal{B}_2 := \{\exists\,(m,\ell,\ell'),\ell\neq\ell'\colon \mathcal{E}_{\text{Tx}}(m,\ell) \text{ and } \mathcal{E}_{\text{Rx}}(m,\ell')\}. \tag{194}$$

Then we have:

$$\mathbb{E}_{\mathcal{C}}[\beta_{y,n}] \le \sum_{i=1}^{2}\Pr\left[\mathcal{B}_i\middle|\mathcal{H}=1\right]. \tag{195}$$

We bound each of the probabilities on the right-hand side of (195). We introduce the following type classes:

$$\mathcal{P}_{\mu,1} := \{\pi_{SXY}\colon |\pi_{SX}-P_{SX}|<\mu/2,\quad |\pi_{SY}-P_{SY}|<\mu\}, \tag{196}$$

$$\mathcal{P}_{\mu,2} := \Big\{\pi_{SS'XY}\colon |\pi_{SX}-P_{SX}|<\mu/2, |\pi_{S'Y}-P_{SY}|<\mu,\quad H_\pi(S'|Y)\le H_\pi(S|Y)\Big\}. \tag{197}$$

Consider $\mathcal{B}_1$ as follows:

$$\begin{aligned}\Pr[\mathcal{B}_1|\mathcal{H}=1] &\le \sum_{m,\ell}\Pr\left[\mathcal{E}_{\text{Tx}}(m,\ell)\cap\mathcal{E}_{\text{Rx}}(m,\ell)\middle|\mathcal{H}=1\right]\\ &\le \sum_{m,\ell}\Pr\Big[(S^n(m,\ell),X^n)\in\mathcal{T}^n_{\mu/2}(P_{SX}), (S^n(m,\ell),Y^n)\in\mathcal{T}^n_\mu(P_{SY})\Big|\mathcal{H}=1\Big]\\ &= \sum_{m,\ell}\sum_{\substack{\pi_{SXY}:\\ |\pi_{SX}-P_{SX}|<\mu/2,\\ |\pi_{SY}-P_{SY}|<\mu}}\Pr\left[\text{tp}\left(S^n(m,\ell),X^n,Y^n\right)=\pi_{SXY}\middle|\mathcal{H}=1\right]\\ &\le 2^{n(R+R')}\cdot\ (n+1)^{|\mathcal{S}|\cdot|\mathcal{X}|\cdot|\mathcal{Y}|}\\ &\quad\cdot\max_{\substack{\pi_{SXY}:\\ |\pi_{SX}-P_{SX}|<\mu/2,\\ |\pi_{SY}-P_{SY}|<\mu}} 2^{-n(D(\pi_{SXY}||P_SQ_{XY})-\mu)},\end{aligned} \tag{198}$$

where the last inequality follows from Sanov's theorem and the i.i.d. codebook construction. Define now:

$$\tilde{\theta}_{\mu,1} := \min_{\substack{\pi_{SXY}: \\ |\pi_{SX}-P_{SX}|<\mu/2, \\ |\pi_{SY}-P_{SY}|<\mu}} D(\pi_{SXY}||P_S Q_{XY}) - R - R' - \mu, \tag{199}$$

and notice that

$$\begin{aligned}
\tilde{\theta}_{\mu,1} &\overset{\text{Eq. (35)}}{=} \min_{\substack{\pi_{SXY}: \\ |\pi_{SX}-P_{SX}|<\mu/2, \\ |\pi_{SY}-P_{SY}|<\mu}} D(\pi_{SXY}||P_S Q_{XY}) - I(S;X) - 2\mu \\
&= \min_{\pi_{SXY}\in\mathcal{P}_{\mu,1}} \sum_{s,x,y} \Bigg[\pi_{SXY}(s,x,y)\log\frac{\pi_{SXY}(s,x,y)}{P_S(s)Q_{XY}(x,y)} \\
&\qquad\qquad -P_{SX}(s,x)\log\frac{P_{S|X}(s,x)}{P_S(s)}\Bigg] - 2\mu \\
&\overset{(j)}{=} \min_{\pi_{SXY}\in\mathcal{P}_{\mu,1}} \sum_{s,x,y} \Bigg[\pi_{SXY}(s,x,y)\log\frac{\pi_{SXY}(s,x,y)}{P_S(s)Q_{XY}(x,y)} \\
&\qquad\qquad -\pi_{SX}(s,x)\log\frac{P_{S|X}(s,x)}{P_S(s)}\Bigg] - \delta_1(\mu) \\
&\overset{(k)}{=} \min_{\pi_{SXY}\in\mathcal{P}_{\mu,1}} \sum_{s,x,y} \pi_{SXY}(s,x,y)\log\frac{\pi_{SXY}(s,x,y)}{P_{S|X}(s|x)Q_{XY}(x,y)} \\
&\qquad\qquad -\delta_1(\mu) \\
&= \min_{\pi_{SXY}\in\mathcal{P}_{\mu,1}} D(\pi_{SXY}\|P_{S|X}Q_{XY}) - \delta_1(\mu) \\
&= \theta_{\mu,1} - \delta_1(\mu),
\end{aligned} \tag{200}$$

for a function $\delta_1(\mu)$ that goes to zero as $\mu \to 0$ and

$$\theta_{\mu,1} := \min_{\pi_{SXY}\in\mathcal{P}_{\mu,1}} D(\pi_{SXY}\|P_{S|X}Q_{XY}). \tag{201}$$

Here, (j) holds because $|\pi_{SX} - P_{SX}| < \mu/2$ and by the continuity of the KL-divergence; (k) follows by re-arranging terms. Considering (198) and (200) yields the following:

$$\Pr\left[\mathcal{B}_1|\mathcal{H}=1\right] \leq (n+1)^{|\mathcal{S}|\cdot|\mathcal{X}|\cdot|\mathcal{Y}|} \cdot\ 2^{-n(\theta_{\mu,1}-\delta_1(\mu))}. \tag{202}$$

Next, consider $\mathcal{B}_2$ as follows:

$$\begin{aligned}
&\Pr\left[\mathcal{B}_2\middle|\mathcal{H}=1\right] \\
&\leq \sum_m \sum_{\ell\neq\ell'} \Pr\left[\mathcal{E}_{\text{Tx}}(m,\ell)\ \cap\ \mathcal{E}_{\text{Rx}}(m,\ell')\middle|\mathcal{H}=1\right] \\
&= \sum_m \sum_{\ell\neq\ell'} \Pr\Bigg[(S^n(m,\ell),X^n)\in\mathcal{T}^n_{\mu/2}(P_{SX}), \\
&\qquad (S^n(m,\ell'),Y^n)\in\mathcal{T}^n_{\mu}(P_{SY}), \\
&\qquad H_{\text{tp}(S^n(m,\ell'),Y^n)}(S'|Y) \\
&\qquad\quad = \min_{\tilde{\ell}} H_{\text{tp}(S^n(m,\tilde{\ell}),Y^n)}(S|Y)\Bigg|\mathcal{H}=1\Bigg] \\
&= \sum_m \sum_{\ell\neq\ell'} \sum_{\substack{\pi_{SS'XY}: \\ |\pi_{SX}-P_{SX}|<\mu/2, \\ |\pi_{S'Y}-P_{SY}|<\mu \\ H_\pi(S'|Y)\leq H_\pi(S|Y)}} \\
&\qquad \Pr\left[\text{tp}\left(S^n(m,\ell),S^n(m,\ell'),X^n,Y^n\right)=\pi_{SS'XY}\middle|\mathcal{H}=1\right] \\
&\leq 2^{n(R+2R')}\cdot\ (n+1)^{|\mathcal{S}|^2\cdot|\mathcal{X}|\cdot|\mathcal{Y}|} \\
&\qquad\cdot \max_{\substack{\pi_{SS'XY}: \\ |\pi_{SX}-P_{SX}|<\mu/2, \\ |\pi_{S'Y}-P_{SY}|<\mu \\ H_\pi(S'|Y)\leq H_\pi(S|Y)}} 2^{-n(D(\pi_{SS'XY}||P_SP_SQ_{XY})-\mu)},
\end{aligned} \tag{203}$$

where the last inequality follows from Sanov's theorem. Now, define:

$$\tilde{\theta}_{\mu,2} := \min_{\substack{\pi_{SS'XY}: \\ |\pi_{SX}-P_{SX}|<\mu/2, \\ |\pi_{S'Y}-P_{SY}|<\mu \\ H_\pi(S'|Y)\leq H_\pi(S|Y)}} D(\pi_{SS'XY}||P_SP_SQ_{XY}) - R - 2R' - \mu. \tag{204}$$

Consider the following chain of inequalities:

$$\begin{aligned}
\tilde{\theta}_{\mu,2} &\overset{\text{Eq. (35)}}{=} \min_{\substack{\pi_{SS'XY}: \\ |\pi_{SX}-P_{SX}|<\mu/2, \\ |\pi_{S'Y}-P_{SY}|<\mu \\ H_\pi(S'|Y)\leq H_\pi(S|Y)}} D(\pi_{SS'XY}||P_SP_SQ_{XY}) \\
&\qquad\qquad -2I(S;X) + R - 3\mu \\
&= \min_{\pi_{SS'XY}\in\mathcal{P}^n_{\mu,2}} D(\pi_{SS'XY}||P_SP_SQ_{XY}) \\
&\qquad\qquad -2I(S;X) + R - 3\mu \\
&\overset{(l)}{=} \min_{\pi_{SS'XY}\in\mathcal{P}^n_{\mu,2}} \Big[D(\pi_{SXY}||P_SQ_{XY}) \\
&\qquad\qquad +\mathbb{E}_{\pi_{SXY}}\left[D(\pi_{S'|SXY}\|P_S)\right]\Big] \\
&\qquad\qquad -2I(S;X) + R - 3\mu \\
&\overset{(m)}{\geq} \min_{\pi_{SS'XY}\in\mathcal{P}^n_{\mu,2}} \Big[D(\pi_{SXY}||P_SQ_{XY}) \\
&\qquad\qquad +\mathbb{E}_{\pi_Y}\left[D(\pi_{S'|Y}\|P_S)\right]\Big] \\
&\qquad\qquad -2I(S;X) + R - 3\mu \\
&\overset{(n)}{=} \min_{\pi_{SS'XY}\in\mathcal{P}^n_{\mu,2}} D(\pi_{SXY}||P_SQ_{XY}) \\
&\qquad\qquad +I(S;Y) - 2I(S;X) + R - \delta_2'(\mu) \\
&\overset{(o)}{=} \min_{\pi_{SS'XY}\in\mathcal{P}^n_{\mu,2}} D(\pi_{SXY}||P_{S|X}Q_{XY}) \\
&\qquad\qquad +I(S;Y) - I(S;X) + R - \delta_2(\mu) \\
&= \theta_{\mu,2} - \delta_2(\mu),
\end{aligned} \tag{205}$$

for functions $\delta_2'(\mu), \delta_2(\mu)$ that go to zero as $\mu \to 0$ and

$$\begin{aligned}
\theta_{\mu,2} := \min_{\pi_{SS'XY}\in\mathcal{P}^n_{\mu,2}} &D(\pi_{SXY}||P_{S|X}Q_{XY}) \\
&+I(S;Y) - I(S;X) + R.
\end{aligned} \tag{206}$$

Here, (l) follows from the chain rule for KL-divergence; (m) follows from the convexity of the KL-divergence and Jensen's inequality; (n) follows because $|\pi_{S'Y} - P_{SY}| < \mu$ and by the continuity of KL-divergence; (o) follows by re-arranging terms and employing similar steps leading to (200). Combining (203) and (205) yields the following:

$$\Pr\left[\mathcal{B}_2|\mathcal{H}=1\right] \leq (n+1)^{|\mathcal{S}|^2\cdot|\mathcal{X}|\cdot|\mathcal{Y}|} \cdot\ 2^{-n(\theta_{\mu,2}-\delta_2(\mu))}. \tag{207}$$

Combining (195), (202), and (207) proves that for large blocklengths n:

$$\mathbb{E}_{\mathcal{C}}\left[\beta_{y,n}\right] \leq (n+1)^{|\mathcal{S}|\cdot|\mathcal{X}|\cdot|\mathcal{Y}|} \cdot\ 2^{-n(\theta_{\mu,1}-\delta_1(\mu))}$$

$$+ (n+1)^{|\mathcal{S}|^2\cdot|\mathcal{X}|\cdot|\mathcal{Y}|} \cdot\ 2^{-n(\theta_{\mu,2}-\delta_2(\mu))}. \quad (208)$$

Letting $n\to\infty$ and then $\mu\to 0$, we get that $\theta_{\mu,1}\to\theta_1$ and $\theta_{\mu,2}\to\theta_2$, where we define:

$$\theta_1 := \min_{\substack{\tilde{P}_{SXY}:\\ \tilde{P}_{SX}=P_{SX}\\ \tilde{P}_{SY}=P_{SY}}} D(\tilde{P}_{SXY}\|P_{S|X}Q_{XY}), \quad (209a)$$

$$\theta_2 := \min_{\substack{\tilde{P}_{SXY}:\\ \tilde{P}_{SX}=P_{SX}\\ \tilde{P}_{Y}=P_{Y}\\ H(S|Y)\le H_{\tilde{P}_{SY}}(S|Y)}} D(\tilde{P}_{SXY}\|P_{S|X}Q_{XY}) + R - I(S;X) + I(S;Y), \quad (209b)$$

where P_{SY} in the minimization constraint is the marginal pmf of $P_{SXY}=P_{S|X}P_{XY}$ and the conditional entropy term $H(S|Y)$ is calculated according to this marginal.

The theorem then follows immediately by (209) and $I(S;X)-I(S;Y)=I(S;X|Y)$ (which holds by the Markov chain $S-X-Y$), and from the fact that $P_{S|X}$ can be chosen arbitrary.

Appendix C
Proof of Theorem 3

We analyze the probabilities of error of the coding and testing scheme described in Subsection IV-B averaged over the random code constructions. By successively eliminating the worst half of the codebooks, as sketched for example at the end of Appendix A), the desired result can be proved for a set of deterministic codebooks.

Fix an arbitrary $\epsilon>0$ and the parameter of the scheme μ sufficiently close to 0 as will become clear in the sequel. Fix also a blocklength n. If $M\neq 0$, let I,J be the indices sent from the transmitter to the relay. If both $B\neq 0$ and $M\neq 0$, let K denote the second index sent from the relay to the receiver.

We first analyze the type-I error probability at the receiver. Define events:

$$\mathcal{E}_{\text{Rx}}^{(1)}\colon \Big\{\exists f'\neq F\colon H_{\text{tp}(S^n(I),V^n(K,f'|I),Z^n)}(V|S,Z) = \min_{\tilde{f}} H_{\text{tp}(S^n(I),V^n(K,\tilde{f}|I),Z^n)}(V|S,Z)\Big\}, \quad (210)$$

$$\mathcal{E}_{\text{Rx}}^{(2)}\colon \left\{(X^n(I),V^n(K,F|I),Z^n)\notin\mathcal{T}_\mu^n(P_{SVZ})\right\}. \quad (211)$$

The type-I error probability can then be bounded as follows:

$$\begin{aligned}
\mathbb{E}_{\mathcal{C}}[\alpha_{z,n}] &\le \Pr\left[M=0\ \cup\ B=0\ \cup\ \mathcal{E}_{\text{Rx}}^{(1)}\ \cup\ \mathcal{E}_{\text{Rx}}^{(2)}\right]\\
&\le \Pr[M=0]+\Pr[B=0|M\neq 0]\\
&\quad+\Pr\left[\mathcal{E}_{\text{Rx}}^{(1)}\middle|M\neq 0,B\neq 0\right]\\
&\quad+\Pr\left[\mathcal{E}_{\text{Rx}}^{(2)}\middle|M\neq 0,B\neq 0,\mathcal{E}_{\text{Rx}}^{(1)c}\right]\\
&\overset{(a)}{\le} \epsilon/16+\Pr[B=0|M\neq 0]\\
&\quad+\Pr\left[\mathcal{E}_{\text{Rx}}^{(2)}\middle|M\neq 0,B\neq 0\right]\\
&\quad+\Pr\left[\mathcal{E}_{\text{Rx}}^{(1)}\middle|M\neq 0,B\neq 0,\mathcal{E}_{\text{Rx}}^{(1)c}\right]\\
&\overset{(b)}{\le} \epsilon/16+\epsilon/16+\Pr\left[\mathcal{E}_{\text{Rx}}^{(2)}\middle|M\neq 0,B\neq 0\right]\\
&\quad+\Pr\left[\mathcal{E}_{\text{Rx}}^{(1)}\middle|M\neq 0,B\neq 0,\mathcal{E}_{\text{Rx}}^{(1)c}\right]\\
&\overset{(c)}{\le} \epsilon/16+\epsilon/16+\epsilon/16\\
&\quad+\Pr\left[\mathcal{E}_{\text{Rx}}^{(1)}\middle|M\neq 0,B\neq 0,\mathcal{E}_{\text{Rx}}^{(2)c}\right]\\
&\overset{(d)}{\le} \epsilon/4.
\end{aligned} \quad (212)$$

where (a) holds by the covering lemma and the rate-constraints in (41); (b) and (d) can be proved following similar lines as the type-I error analysis in Appendix B; and (c) holds by the Markov lemma.

We now bound the probability of type-II error at the receiver. Define the following events:

$$\mathcal{E}_{\text{Tx}}(i,j,e)=\left\{(S^n(i),U^n(j,e|i),X^n)\in\mathcal{T}_{\mu/4}^n(P_{SUX})\right\} \quad (213)$$

$$\begin{aligned}
\mathcal{E}_{\text{Rel}}(i,j,e',k,f) = \Big\{&(S^n(i),U^n(j,e'|i),V^n(k,f|i),Y^n)\in\mathcal{T}_{\mu/2}^n(P_{SUVY}),\\
&H_{\text{tp}(S^n(i),U^n(j,e'|i),Y^n)}(U|S,Y) = \min_{\tilde{e}} H_{\text{tp}(S^n(i),U^n(j,\tilde{e}|i),Y^n)}(U|S,Y)\Big\},
\end{aligned} \quad (214)$$

$$\begin{aligned}
\mathcal{E}_{\text{Rx}}(i,k,f') = \Big\{&(S^n(i),V^n(k,f'|i),Z^n)\in\mathcal{T}_{\mu}^n(P_{SVZ}),\\
&H_{\text{tp}(S^n(i),V^n(k,f'|i),Z^n)}(V|S,Z) = \min_{\tilde{f}} H_{\text{tp}\left(S^n(i),V^n(k,\tilde{f}|i),Z^n\right)}(V|S,Z)\Big\}.
\end{aligned} \quad (215)$$

We then have:

$$\mathbb{E}_{\mathcal{C}}[\beta_{z,n}] = \Pr\left[\hat{\mathcal{H}}_z=0\middle|\mathcal{H}=1\right] \quad (216)$$

$$\le \Pr\left[B\neq 0,\ \cup_{i,k,f'}\mathcal{E}_{\text{Rx}}(i,k,f')\middle|\mathcal{H}=1\right] \quad (217)$$

$$\le \Pr\Big[\cup_{i,j,e,e',k,f,f'}\ \mathcal{E}_{\text{Tx}}(i,j,e) \text{ and } \mathcal{E}_{\text{Rel}}(i,j,e',k,f) \text{ and } \mathcal{E}_{\text{Rx}}(i,k,f')\Big|\mathcal{H}=1\Big] \quad (218)$$

We can further upper bound this last probability with the union bound to obtain:

$$\mathbb{E}_{\mathcal{C}}[\beta_{z,n}] \le \sum_{i=1}^{4}\Pr\left[\mathcal{B}_i\middle|\mathcal{H}=1\right], \quad (219)$$

where the four events $\mathcal{B}_1,\mathcal{B}_2,\mathcal{B}_3,\mathcal{B}_4$ are defined as:

$$\mathcal{B}_1\colon \Big\{\exists\,(i,j,e,k,f)\colon\ \mathcal{E}_{\text{Tx}}(i,j,e)\ \text{ and }\ \mathcal{E}_{\text{Rel}}(i,j,e,k,f)\ \text{ and }\mathcal{E}_{\text{Rx}}(i,k,f)\Big\}, \quad (220)$$

$$\mathcal{B}_2\colon \Big\{\exists\,(i,j,e,e',k,f)\colon\ e\neq e'\ \text{ and }\ \mathcal{E}_{\text{Tx}}(i,j,e)\ \text{ and }\ \mathcal{E}_{\text{Rel}}(i,j,e',k,f)\ \text{ and }\ \mathcal{E}_{\text{Rx}}(i,k,f)\Big\},$$

(221)

$$\mathcal{B}_3\colon \Big\{\exists\,(i,j,e,k,f,f')\colon\ \ f\neq f'\ \text{ and }\ \mathcal{E}_{\text{Tx}}(i,j,e)\ \text{ and }\ \mathcal{E}_{\text{Rel}}(i,j,e,k,f)\ \text{ and }\ \mathcal{E}_{\text{Rx}}(i,k,f')\Big\}, \tag{222}$$

$$\mathcal{B}_4\colon \Big\{\exists\,(i,j,e,e',k,f,f')\colon\ \ e\neq e'\ \text{ and }\ f\neq f'\ \text{ and }\ \mathcal{E}_{\text{Tx}}(i,j,e)\ \text{ and }\ \mathcal{E}_{\text{Rel}}(i,j,e',k,f)\ \text{ and }\ \mathcal{E}_{\text{Rx}}(i,k,f')\Big\}. \tag{223}$$

The summands in (219) can be analyzed by now standard arguments as used in Appendices A and B.

For each $i=1,2,3,4$, this yields an exponential bound of the form

$$\Pr[\mathcal{B}_i]\leq 2^{-n(\theta_i+\delta_i(\mu))}, \tag{224}$$

where $\delta_1(\mu),\delta_2(\mu),\delta_3(\mu),\delta_4(\mu)$ are functions that tend to 0 as $\mu\to 0$ and where

$$\theta_1:=\min_{\substack{\pi_{SUVXYZ}:\\ \pi_{SUX}=P_{SUX}\\ \pi_{SUVY}=P_{SUVY}\\ \pi_{SVZ}=P_{SVZ}}} D(\pi_{SUVXYZ}\|P_{US|X}P_{V|SUY}Q_{XYZ}), \tag{225}$$

$$\theta_2:=\min_{\substack{\pi_{SUU'VXYZ}:\\ \pi_{SUX}=P_{SUX}\\ H(U|S,Y)\leq H_\pi(U|S,Y)\\ \pi_{SU'VY}=P_{SU'VY}\\ \pi_{SVZ}=P_{SVZ}}} D(\pi_{SUU'VXYZ}\|P_{SU|X}P_{U'|S}P_{V|SU'Y}Q_{XYZ}) + R_u - I(U;X|S), \tag{226}$$

$$\theta_3:=\min_{\substack{\pi_{SUVV'XYZ}:\\ \pi_{SUX}=P_{SUX}\\ \pi_{SUVY}=P_{SUVY}\\ H(V|S,Z)\leq H_\pi(V|S,Z)\\ \pi_{SV'Z}=P_{SV'Z}}} D(\pi_{SUVV'XYZ}\|P_{SU|X}P_{V|SUY}P_{V'|S}Q_{XYZ}) + R_v - I(V;U,Y|S), \tag{227}$$

$$\theta_4:=\min_{\substack{\pi_{SUU'VV'XYZ}:\\ \pi_{SUX}=P_{SUX}\\ H(U|S,Y)\leq H_\pi(U|S,Y)\\ \pi_{SU'VY}=P_{SU'VY}\\ H(V|S,Z)\leq H_\pi(V|S,Z)\\ \pi_{SV'Z}=P_{SV'Z}}} D(\pi_{SUU'VV'XYZ}\|P_{SU|X}P_{U'|S}P_{V|SU'Y}P_{V'|S}Q_{XYZ}) + R_u+R_v - I(U;X|S) - I(V;U,Y|S). \tag{228}$$

Plugging the exponential bounds (224) into (219), extracting the term $I(U';Y|S)=I(U;Y|S)$ from (226) and (228) and the term $I(V';Z|S)=I(V;Z|S)$ from (227) and (228), and bounding R_u and R_v by $R-I(S;X)$ and $T-I(S;X)$, we obtain the result in the theorem.

Appendix D
Proof of Proposition 1

The inclusion

$$\mathcal{E}_{\text{dcpled}}(R,T)\subseteq\mathcal{E}_{\text{nobin}}(R,T), \tag{229}$$

is straightforward. It suffices to note that restricting the union in (29) to choices of the conditional pmfs $P_{SU|X}$ and $P_{V|SUY}$ where S is a constant and V is conditionally independent of U given Y, results in $\mathcal{E}_{\text{dcpled}}(R,T)$.

We now prove the reverse inclusion

$$\mathcal{E}_{\text{dcpled}}(R,T)\supseteq\mathcal{E}_{\text{nobin}}(R,T). \tag{230}$$

Fix an arbitrary pair $P_{SU|X}$ and $P_{V|SUY}$ satisfying the rate-constraints (30)–(31). Then, notice the following sequence of equalities:

$$\begin{aligned}
&\min_{\substack{\tilde{P}_{SUVXYZ}:\\ \tilde{P}_{SUX}=P_{SUX}\\ \tilde{P}_{SUVY}=P_{SUVY}\\ \tilde{P}_{SVZ}=P_{SVZ}}} D(\tilde{P}_{SUVXYZ}\|P_{SU|X}P_{V|SUY}Q_{XY}Q_{Z|Y})\\
&\quad=\min_{\substack{\tilde{P}_{SUXY}:\\ \tilde{P}_{SUX}=P_{SUX}\\ \tilde{P}_{SUY}=P_{SUY}}}\Bigg[D(\tilde{P}_{SUXY}\|P_{SU|X}Q_{XY})\\
&\qquad+\mathbb{E}_{\tilde{P}_{SUXY}}\Big[\min_{\substack{\tilde{P}_{VZ|SUXY}:\\ \tilde{P}_{V|SUY}=P_{V|SUY}\\ \tilde{P}_{Z|SV}=P_{Z|SV}}} D(\tilde{P}_{VZ|SUXY}\|P_{V|SUY}Q_{Z|Y})\Big]\Bigg]\\
&\quad\overset{(a)}{=}\min_{\substack{\tilde{P}_{SUXY}:\\ \tilde{P}_{SUX}=P_{SUX}\\ \tilde{P}_{SUY}=P_{SUY}}}\Bigg[D(\tilde{P}_{SUXY}\|P_{SU|X}Q_{XY})\\
&\qquad+\mathbb{E}_{P_{SUY}}\Big[\min_{\substack{\tilde{P}_{VZ|SUY}:\\ \tilde{P}_{V|SUY}=P_{V|SUY}\\ \tilde{P}_{Z|SV}=P_{Z|SV}}} D(\tilde{P}_{VZ|SUY}\|P_{V|SUY}Q_{Z|Y})\Big]\Bigg]\\
&\quad\overset{(b)}{=}\min_{\substack{\tilde{P}_{SUXY}:\\ \tilde{P}_{SUX}=P_{SUX}\\ \tilde{P}_{SUY}=P_{SUY}}}\Bigg[D(\tilde{P}_{SUXY}\|P_{SU|X}Q_{XY})\\
&\qquad+\mathbb{E}_{P_{SUVY}}\Big[\min_{\substack{\tilde{P}_{Z|SUVY}:\\ \tilde{P}_{Z|SV}=P_{Z|SV}}} D(\tilde{P}_{Z|SUVY}\|Q_{Z|Y})\Big]\Bigg]\\
&\quad\overset{(c)}{=}\min_{\substack{\tilde{P}_{SUXY}:\\ \tilde{P}_{SUX}=P_{SUX}\\ \tilde{P}_{SUY}=P_{SUY}}}\Bigg[D(\tilde{P}_{SUXY}\|P_{SU|X}Q_{XY})\\
&\qquad+\mathbb{E}_{P_{SVY}}\Big[\min_{\substack{\tilde{P}_{Z|SVY}:\\ \tilde{P}_{Z|SV}=P_{Z|SV}}} D(\tilde{P}_{Z|SVY}\|Q_{Z|Y})\Big]\Bigg],
\end{aligned} \tag{231}$$

where the steps are justified as follows:

- (a) follows because, by the convexity of the KL-divergence, the LHS is larger than or equal to the RHS; the reverse direction holds because the minimization on the LHS can only increase if one restricts pmfs to be of the form $\tilde{P}_{VZ|SUXY}=\tilde{P}_{VZ|SUY}$;
- (b) holds because $\tilde{P}_{V|SUY}=P_{V|SUY}$; and
- (c) follows because, by the convexity of the KL-divergence, the LHS is larger than or equal to the RHS; the reverse direction holds because the minimization on the LHS can only increase if one restricts pmfs to be of the form $\tilde{P}_{Z|SUVY}=\tilde{P}_{Z|SVY}$.

Defining now $\bar{U} := (U, S)$ and $\bar{V} := (V, S)$, we conclude that

$$\begin{aligned}&\min_{\substack{\tilde{P}_{SUXY}:\\ \tilde{P}_{SUX}=P_{SUX}\\ \tilde{P}_{SUY}=P_{SUY}}} D(\tilde{P}_{SUXY}\|P_{SU|X}Q_{XY})\\ &= \min_{\substack{\tilde{P}_{\bar{U}XY}:\\ \tilde{P}_{\bar{U}X}=P_{\bar{U}X}\\ \tilde{P}_{\bar{U}Y}=P_{\bar{U}Y}}} D(\tilde{P}_{\bar{U}XY}\|P_{\bar{U}|X}Q_{XY}) \end{aligned} \tag{232}$$

and

$$\begin{aligned}&\min_{\substack{\tilde{P}_{SUVXYZ}:\\ \tilde{P}_{SUX}=P_{SUX}\\ \tilde{P}_{SUVY}=P_{SUVY}\\ \tilde{P}_{SVZ}=P_{SVZ}}} D(\tilde{P}_{SUVXYZ}\|P_{SU|X}P_{V|SUY}Q_{XY}Q_{Z|Y})\\ &= \min_{\substack{\tilde{P}_{\bar{U}XY}:\\ \tilde{P}_{\bar{U}X}=P_{\bar{U}X}\\ \tilde{P}_{\bar{U}Y}=P_{\bar{U}Y}}} D(\tilde{P}_{\bar{U}XY}\|P_{\bar{U}|X}Q_{XY})\\ &\quad+\mathbb{E}_{P_Y}\Big[\min_{\substack{\tilde{P}_{\bar{V}Z|Y}:\\ \tilde{P}_{\bar{V}Y}=P_{\bar{V}Y}\\ \tilde{P}_{\bar{V}Z}=P_{\bar{V}Z}}} D(\tilde{P}_{\bar{V}Z|Y}\|P_{\bar{V}|Y}Q_{Z|Y})\Big]. \end{aligned} \tag{233}$$

Notice further the Markov chains $\bar{U} \to X \to Y$ and $\bar{V} \to Y \to Z$ and that the choice $(\bar{U}, \bar{V})$ satisfies the rate constraints

$$I(\bar{U}; X) = I(S, U; X) \le R \tag{234}$$

and

$$\begin{aligned} I(\bar{V}; Y) &= I(S; Y) + I(V; Y|S)\\ &\le I(S; X) + I(V; Y, U|S)\\ &\le T. \end{aligned} \tag{235}$$

From all these steps, we conclude that the choice $P_{\bar{U}|X} = P_{US|X}$ and $P_{\bar{V}|Y} = P_{SV|Y}$ satisfies the following three conditions:

$$I(\bar{U}; X) \le R \tag{236}$$

$$I(\bar{V}; Y) \le T \tag{237}$$

$$\mathcal{E}_{\text{dcpled}}(P_{\bar{U}|X}, P_{\bar{V}|Y}) \supseteq \mathcal{E}_{\text{nobin}}(P_{SU|X}, P_{V|SUY}). \tag{238}$$

This proves inclusion (230).

Appendix E
Proof of Inclusion $\mathcal{E}_{\text{bin,dcpled}}(R, T) \supseteq \mathcal{E}_{\text{bin}}(R, T)$

Fix a pair of conditional pmfs $P_{SU|X}$ and $P_{V|SUY}$ and define $\bar{U} := (U, S)$ and $\bar{V} := (V, S)$. Notice first that, since $\tilde{P}_{SY} = P_{SY}$ and $\tilde{P}_{SZ} = P_{SZ}$, the following hold:

- Condition $H(U|S, Y) \le H_{\tilde{P}}(U|S, Y)$ is equivalent to $H(U, S|Y) \le H_{\tilde{P}}(U, S|Y)$ and hence also equivalent to $H(\bar{U}|Y) \le H_{\tilde{P}}(\bar{U}|Y)$;
- Condition $H(V|S, Z) \le H_{\tilde{P}}(V|S, Z)$ is equivalent to $H(V, S|Z) \le H_{\tilde{P}}(V, S|Z)$ and hence also equivalent to $H(\bar{V}|Z) \le H_{\tilde{P}}(\bar{V}|Z)$.

Using these equivalences and following similar steps as in the proof of Proposition 1 in Appendix D, it can be shown that

$$\begin{aligned}&\min_{\substack{\tilde{P}_{\bar{U}XY}:\\ \tilde{P}_{\bar{U}X}=P_{UX}\\ \tilde{P}_{\bar{U}Y}=P_{\bar{U}Y}}} D(\tilde{P}_{\bar{U}XY}\|P_{\bar{U}|X}Q_{XY})\\ &\quad+\min_{\substack{\tilde{P}_{\bar{V}Z|Y}:\\ \tilde{P}_{\bar{V}Y}=P_{\bar{V}Y}\\ \tilde{P}_{\bar{V}Z}=P_{\bar{V}Z}}} \mathbb{E}_{P_Y}\Big[D(\tilde{P}_{\bar{V}Z|Y}\|P_{\bar{V}|Y}Q_{Z|Y})\Big]\\ &\ge \min_{\substack{\tilde{P}_{SUVXYZ}:\\ \tilde{P}_{SUX}=P_{SUX}\\ \tilde{P}_{SUVY}=P_{SUVY}\\ \tilde{P}_{SVZ}=P_{SVZ}}} D(\tilde{P}_{SUVXYZ}\|P_{SU|X}P_{V|SUY}Q_{XYZ}); \end{aligned} \tag{239}$$

$$\begin{aligned}&\min_{\substack{\tilde{P}_{\bar{U}XY}:\\ \tilde{P}_{\bar{U}X}=P_{\bar{U}X}\\ \tilde{P}_{Y}=P_{Y}\\ H(\bar{U}|Y)\le H_{\tilde{P}}(\bar{U}|Y)}} D(\tilde{P}_{\bar{U}XY}\|P_{\bar{U}|X}Q_{XY})\\ &\quad+\min_{\substack{\tilde{P}_{\bar{V}Z|Y}:\\ \tilde{P}_{\bar{V}Y}=P_{\bar{V}Y}\\ \tilde{P}_{\bar{V}Z}=P_{\bar{V}Z}}} \mathbb{E}_{P_Y}\Big[D(\tilde{P}_{\bar{V}Z|Y}\|P_{\bar{V}|Y}Q_{Z|Y})\Big]\\ &\ge \min_{\substack{\tilde{P}_{SUU'VXYZ}:\\ \tilde{P}_{SUX}=P_{SUX}\\ \tilde{P}_{SVY}=P_{SVY}\\ \tilde{P}_{SVZ}=P_{SVZ}\\ H(U|S,Y)\le H_{\tilde{P}}(U|S,Y)}} D(\tilde{P}_{SUVXYZ}\|P_{SU|X}P_{V|SY}Q_{XYZ}); \end{aligned} \tag{240}$$

$$\begin{aligned}&\min_{\substack{\tilde{P}_{\bar{U}XY}:\\ \tilde{P}_{\bar{U}X}=P_{\bar{U}X}\\ \tilde{P}_{Y}=P_{Y}\\ H(\bar{U}|Y)\le H_{\tilde{P}}(\bar{U}|Y)}} D(\tilde{P}_{\bar{U}XY}\|P_{\bar{U}|X}Q_{XY})\\ &\quad+\min_{\substack{\tilde{P}_{\bar{V}Z|Y}:\\ \tilde{P}_{\bar{V}Y}=P_{\bar{V}Y}\\ \tilde{P}_{Z}=P_{Z}\\ H(\bar{V}|Z)\le H_{\tilde{P}}(\bar{V}|Z)}} \mathbb{E}_{P_Y}\Big[D(\tilde{P}_{\bar{V}Z|Y}\|P_{\bar{V}|Y}Q_{Z|Y})\Big]\\ &\ge \min_{\substack{\tilde{P}_{SUVXYZ}:\\ \tilde{P}_{SUX}=P_{SUX}\\ \tilde{P}_{SUVY}=P_{SUVY}\\ \tilde{P}_{SZ}=P_{SZ}\\ H(V|S,Z)\le H_{\tilde{P}}(V|S,Z)}} D(\tilde{P}_{SUVXYZ}\|P_{SU|X}P_{V|SUY}Q_{XYZ}); \end{aligned} \tag{241}$$

$$\begin{aligned}&\min_{\substack{\tilde{P}_{\bar{U}XY}:\\ \tilde{P}_{\bar{U}X}=P_{\bar{U}X}\\ \tilde{P}_{Y}=P_{Y}\\ H(\bar{U}|Y)\le H_{\tilde{P}}(\bar{U}|Y)}} D(\tilde{P}_{\bar{U}XY}\|P_{\bar{U}|X}Q_{XY})\\ &\quad+\min_{\substack{\tilde{P}_{\bar{V}Z|Y}:\\ \tilde{P}_{\bar{V}Y}=P_{\bar{V}Y}\\ \tilde{P}_{Z}=P_{Z}\\ H(\bar{V}|Z)\le H_{\tilde{P}}(\bar{V}|Z)}} \mathbb{E}_{P_Y}\Big[D(\tilde{P}_{\bar{V}Z|Y}\|P_{\bar{V}|Y}Q_{Z|Y})\Big]\\ &\ge \min_{\substack{\tilde{P}_{SUVXYZ}:\\ \tilde{P}_{SUX}=P_{SUX}\\ \tilde{P}_{SVY}=P_{SVY}\\ \tilde{P}_{SZ}=P_{SZ}\\ H(U|S,Y)\le H_{\tilde{P}}(U|S,Y)\\ H(V|S,Z)\le H_{\tilde{P}}(V|S,Z)}} D(\tilde{P}_{SUVXYZ}\|P_{SU|X}P_{V|SY}Q_{XYZ}). \end{aligned} \tag{242}$$

Since moreover

$$\begin{aligned}-I(\bar{U};X|Y) &= -I(S,U;X|Y)\\ &= -I(S,U;X)+I(S,U;Y)\\ &\geq -I(S,U;X)+I(U;Y|S) \qquad (243)\\ -I(\bar{V};Y|Z) &= -I(S,V;Y)+I(S,V;Z)\\ &\geq -I(S;Y)-I(V;Y|S)+I(V;Z|S)\\ &\geq -I(S;X)-I(V;U,Y|S)+I(V;Z|S), \qquad (244)\end{aligned}$$

we can conclude that

$$\mathcal{E}_{\text{bin,dcpled}}(P_{\bar{U}|X},P_{\bar{V}|Y}) \supseteq \mathcal{E}_{\text{bin}}(P_{SU|X},P_{V|SUY}). \qquad (245)$$

This establishes the desired proof.

Appendix F
Proof of Converse to Corollary 1

Fix a sequence of encoding and decoding functions $\{\phi^{(n)},\phi_y^{(n)},g_y^{(n)},g_z^{(n)}\}$ so that the inequalities of Definition 1 hold for sufficiently large blocklengths n. Fix also such a sufficiently large n and define for each $t\in\{1,\ldots,n\}$:

$$\begin{aligned}U_t :=&(M,X^{t-1},Y_{\text{C}}^{t-1},Y_{\text{C},t+1}^n)\\ V_t :=&(B,Y_{\text{H}}^{t-1},Z_{\text{C}}^{t-1},Z_{\text{C},t+1}^n,Y_{\text{C}}^{t-1},Y_{\text{C},t+1}^n). \qquad (246)\end{aligned}$$

Define further $U:=(U_T,T)$; $V:=(V_T,T)$; $X:=X_T$; $Y:=Y_T$; $W:=W_T$; and $Z:=Z_T$; for $T\sim\mathcal{U}\{1,\ldots,n\}$ independent of the tuples (U^n,V^n,X^n,Y^n,Z^n). Notice the Markov chains $U\to X\to Y$ and $V\to Y\to Z$. Let $\delta(\epsilon):=H_{\text{b}}(\epsilon)/(n\cdot(1-\epsilon))$ where $H_{\text{b}}(\epsilon)$ denotes the entropy of the binary random variable with parameter ϵ.

First, consider the rate R:

$$\begin{aligned}R &= \frac{1}{n}H(M)\\ &\geq \frac{1}{n}I(M;X^n|Y_{\text{C}}^n)\\ &= \frac{1}{n}\sum_{t=1}^n I(M;X_t|X^{t-1},Y_{\text{C}}^n)\\ &\overset{(a)}{=} \frac{1}{n}\sum_{t=1}^n I(M,X^{t-1},Y_{\text{C}}^{t-1},Y_{\text{C},t+1}^n;X_t|Y_{\text{C},t})\\ &= \frac{1}{n}\sum_{t=1}^n I(U_t;X_t|Y_{\text{C},t})\\ &= I(U;X|Y_{\text{C}}), \qquad (247)\end{aligned}$$

where (a) follows from the memoryless property of the sources. Similarly,

$$\begin{aligned}T &= \frac{1}{n}H(B)\\ &\geq \frac{1}{n}I(B;Y^n|Y_{\text{C}}^n,Z_{\text{C}}^n)\\ &= \frac{1}{n}\sum_{t=1}^n I(B;Y_t|Y^{t-1},Y_{\text{C}}^n,Z_{\text{C}}^n)\\ &= \frac{1}{n}\sum_{t=1}^n I(B,Y^{t-1},Z_{\text{C}}^{t-1},Z_{\text{C},t+1}^n,Y_{\text{C}}^{t-1},Y_{\text{C},t+1}^n;Y_t|Y_{\text{C},t},Z_{\text{C},t})\\ &= \frac{1}{n}\sum_{t=1}^n I(V_t;Y_t|Y_{\text{C},t},Z_{\text{C},t})\\ &= I(V;Y|Y_{\text{C}},Z_{\text{C}}). \qquad (248)\end{aligned}$$

The type-II error probability at the relay can be bounded as

$$\begin{aligned}&-\frac{1}{n}\log\beta_{y,n}\\ &\leq \frac{1}{(1-\epsilon)n}D(P_{MY_{\text{H}}^nY_{\text{C}}^n|\mathcal{H}=0}\|P_{MY_{\text{H}}^nY_{\text{C}}^n|\mathcal{H}=1})+\delta(\epsilon)\\ &\overset{(b)}{=} \frac{1}{(1-\epsilon)n}D(P_{MY_{\text{H}}^nY_{\text{C}}^n}\|P_{M|Y_{\text{C}}^n}P_{Y_{\text{H}}^n|Y_{\text{C}}^n}P_{Y_{\text{C}}^n})+\delta(\epsilon)\\ &= \frac{1}{(1-\epsilon)n}I(M;Y_{\text{H}}^n|Y_{\text{C}}^n)+\delta(\epsilon)\\ &= \frac{1}{(1-\epsilon)n}\sum_{t=1}^n I(M;Y_{\text{H},t}|Y_{\text{H}}^{t-1},Y_{\text{C}}^n)+\delta(\epsilon)\\ &= \frac{1}{(1-\epsilon)n}\sum_{t=1}^n I(M,Y_{\text{H}}^{t-1},Y_{\text{C}}^{t-1},Y_{\text{C},t+1}^n;Y_{\text{H},t}|Y_{\text{C},t})+\delta(\epsilon)\\ &\overset{(c)}{\leq} \frac{1}{(1-\epsilon)n}\sum_{t=1}^n I(M,X^{t-1},Y_{\text{C}}^{t-1},Y_{\text{C},t+1}^n;Y_{\text{H},t}|Y_{\text{C},t})+\delta(\epsilon)\\ &= \frac{1}{(1-\epsilon)n}\sum_{t=1}^n I(U_t;Y_{\text{H},t}|Y_{\text{C},t})+\delta(\epsilon)\\ &= \frac{1}{1-\epsilon}I(U;Y|Y_{\text{C}})+\delta(\epsilon), \qquad (249)\end{aligned}$$

where (b) holds by the assumption on the distributions $P_{XY_{\text{C}}Y_{\text{H}}Z_{\text{C}}Z_{\text{H}}}$ and $Q_{XY_{\text{C}}Y_{\text{H}}Z_{\text{C}}Z_{\text{H}}}$ in (73)–(74) and the fact that M is a function of X^n; and (c) holds by the Markov chain $Y_{\text{H}}^{t-1}\to(M,X^{t-1},Y_{\text{C}}^n)\to Y_{\text{H},t}$. Finally, consider the type-II error probability at the receiver:

$$\begin{aligned}&-\frac{1}{n}\log\beta_{z,n}\\ &\leq \frac{1}{(1-\epsilon)n}D(P_{BZ_{\text{H}}^nZ_{\text{C}}^nY_{\text{C}}^n|\mathcal{H}=0}\|P_{BZ_{\text{H}}^nZ_{\text{C}}^nY_{\text{C}}^n|\mathcal{H}=1})+\delta(\epsilon)\\ &\overset{(d)}{=} \frac{1}{(1-\epsilon)n}\mathbb{E}_{Z_{\text{C}}^nY_{\text{C}}^n}\left[D(P_{BZ_{\text{H}}^n|Z_{\text{C}}^nY_{\text{C}}^n,\mathcal{H}=0}\|P_{BZ_{\text{H}}^n|Z_{\text{C}}^nY_{\text{C}}^n,\mathcal{H}=1})\right]\\ &\quad+\delta(\epsilon)\\ &\overset{(e)}{=} \frac{1}{(1-\epsilon)n}\mathbb{E}_{Z_{\text{C}}^nY_{\text{C}}^n}\left[D(P_{B|Z_{\text{C}}^nY_{\text{C}}^n,\mathcal{H}=0}\|P_{B|Z_{\text{C}}^nY_{\text{C}}^n,\mathcal{H}=1})\right]\\ &\quad+\frac{1}{(1-\epsilon)n}\mathbb{E}_{BZ_{\text{C}}^nY_{\text{C}}^n}\left[D(P_{Z_{\text{H}}^n|BZ_{\text{C}}^nY_{\text{C}}^n,\mathcal{H}=0}\|P_{Z_{\text{H}}^n|BZ_{\text{C}}^nY_{\text{C}}^n,\mathcal{H}=1})\right]\\ &\quad+\delta(\epsilon)\\ &\overset{(f)}{\leq} \frac{1}{(1-\epsilon)n}\mathbb{E}_{Z_{\text{C}}^nY_{\text{C}}^n}\left[D(P_{MY_{\text{H}}^n|Z_{\text{C}}^nY_{\text{C}}^n,\mathcal{H}=0}\|P_{MY_{\text{H}}^n|Z_{\text{C}}^nY_{\text{C}}^n,\mathcal{H}=1})\right]\\ &\quad+\frac{1}{(1-\epsilon)n}\mathbb{E}_{BZ_{\text{C}}^nY_{\text{C}}^n}\left[D(P_{Z_{\text{H}}^n|BZ_{\text{C}}^nY_{\text{C}}^n,\mathcal{H}=0}\|P_{Z_{\text{H}}^n|Z_{\text{C}}^nY_{\text{C}}^n,\mathcal{H}=1})\right]\\ &\quad+\delta(\epsilon)\\ &\overset{(g)}{=} \frac{1}{(1-\epsilon)n}\mathbb{E}_{Y_{\text{H}}^nY_{\text{C}}^nZ_{\text{C}}^n}\left[D(P_{M|Y_{\text{C}}^nY_{\text{H}}^nZ_{\text{C}}^n,\mathcal{H}=0}\|P_{M|Y_{\text{C}}^nY_{\text{H}}^nZ_{\text{C}}^n,\mathcal{H}=1})\right]\\ &\quad+\frac{1}{(1-\epsilon)n}I(B;Z_{\text{H}}^n|Z_{\text{C}}^n,Y_{\text{C}}^n)+\delta(\epsilon)\\ &\overset{(h)}{=} \frac{1}{(1-\epsilon)n}\mathbb{E}_{Y_{\text{H}}^nY_{\text{C}}^n}\left[D(P_{M|Y_{\text{C}}^nY_{\text{H}}^n,\mathcal{H}=0}\|P_{M|Y_{\text{C}}^nY_{\text{H}}^n,\mathcal{H}=1})\right]\end{aligned}$$

$$
\begin{aligned}
&+\frac{1}{(1-\epsilon)n}I(B;Z_{\mathrm{H}}^n|Z_{\mathrm{C}}^n,Y_{\mathrm{C}}^n)+\delta(\epsilon)\\
&\overset{(i)}{=}\frac{1}{(1-\epsilon)n}I(M;Y_{\mathrm{H}}^n|Y_{\mathrm{C}}^n)+\frac{1}{(1-\epsilon)n}I(B;Z_{\mathrm{H}}^n|Z_{\mathrm{C}}^n,Y_{\mathrm{C}}^n)\\
&\quad+\delta(\epsilon)\\
&\overset{(j)}{=}\frac{1}{(1-\epsilon)n}\sum_{t=1}^{n}\Big[I(M,Y_{\mathrm{H}}^{t-1},Y_{\mathrm{C}}^{t-1},Y_{\mathrm{C},t+1}^n;Y_{\mathrm{H},t}|Y_{\mathrm{C},t})\\
&\quad+I(B,Z_{\mathrm{H}}^{t-1},Z_{\mathrm{C}}^{t-1},Z_{\mathrm{C},t+1}^n,Y_{\mathrm{C}}^{t-1},Y_{\mathrm{C},t+1}^n;Z_{\mathrm{H},t}|Z_{\mathrm{C},t},Y_{\mathrm{C},t})\Big]\\
&\quad+\delta(\epsilon)\\
&\overset{(k)}{\leq}\frac{1}{(1-\epsilon)n}\sum_{t=1}^{n}I(U_t;Y_{\mathrm{H},t}|Y_{\mathrm{C},t})\\
&\quad+\frac{1}{(1-\epsilon)n}\sum_{t=1}^{n}I(V_t;Z_{\mathrm{H},t}|Z_{\mathrm{C},t},Y_{\mathrm{C},t})+\delta(\epsilon)\\
&=\frac{1}{1-\epsilon}I(U;Y|Y_{\mathrm{C}})+\frac{1}{1-\epsilon}I(V;Z|Z_{\mathrm{C}},Y_{\mathrm{C}})+\delta(\epsilon),
\end{aligned}
\tag{250}
$$

where (d) holds because the pair $(Y_{\mathrm{C}}^n,Z_{\mathrm{C}}^n)$ has the same distribution under both hypotheses; (e) holds by the chain rule for KL-divergence; (f) holds by the data-processing inequality and the fact that B is a function of $(M,Y_{\mathrm{H}}^n,Y_{\mathrm{C}}^n)$, and because under $\mathcal{H}=1$ and given $(Y_{\mathrm{C}}^n,Z_{\mathrm{C}}^n)$, the message B is independent of the observation Z_{H}^n; (g) holds because the two triples $(Y_{\mathrm{H}}^n,Y_{\mathrm{C}}^n,Z_{\mathrm{C}}^n)$ and $(Y_{\mathrm{C}}^n,Z_{\mathrm{H}}^n,Z_{\mathrm{C}}^n)$ have the same distribution under both hypotheses; (h) holds because under both hypotheses M is independent of Z_{H}^n given the pair $(Y_{\mathrm{H}}^n,Y_{\mathrm{C}}^n)$; (i) holds because the triple $(M,Y_{\mathrm{H}}^n,Y_{\mathrm{C}}^n)$ has same distribution under both hypotheses; (j) holds by the memoryless property of the sources; and (k) holds by the definitions of U_t and V_t and the Markov chain $Z_{\mathrm{H}}^{t-1}\to(B,Y_{\mathrm{H}}^{t-1},Z_{\mathrm{C}}^n,Y_{\mathrm{C}}^n)\to Z_{\mathrm{H},t}$.

Appendix G
Proof of the Converse to corollary 2

Fix sequences of encoding and decoding functions $\{\phi^{(n)},\phi_y^{(n)},g_y^{(n)},g_z^{(n)}\}$, and notice that there exists a function $\delta(\epsilon)$ which tends to zero when $\epsilon\to 0$ such that, for any $\epsilon>0$ and sufficiently large n:

$$
\begin{aligned}
-\frac{1}{n}\log\beta_{y,n}&\leq\frac{1}{(1-\epsilon)n}D(P_{MY^n|\mathcal{H}=0}\|P_{MY^n|\mathcal{H}=1})+\delta(\epsilon)\\
&=\delta(\epsilon)\\
-\frac{1}{n}\log\beta_{z,n}&\leq\frac{1}{(1-\epsilon)n}D(P_{BZ_{\mathrm{C}}^nZ_{\mathrm{H}}^n|\mathcal{H}=0}\|P_{BZ_{\mathrm{C}}^nZ_{\mathrm{H}}^n|\mathcal{H}=1})\\
&\quad+\delta(\epsilon)\\
&\overset{(a)}{=}\frac{1}{(1-\epsilon)n}D(P_{BZ_{\mathrm{H}}^n|Z_{\mathrm{C}}^n,\mathcal{H}=0}\|P_{BZ_{\mathrm{H}}^n|Z_{\mathrm{C}}^n,\mathcal{H}=1})\\
&\quad+\delta(\epsilon)\\
&\overset{(b)}{=}\frac{1}{(1-\epsilon)n}I(B;Z_{\mathrm{H}}^n|Z_{\mathrm{C}}^n)+\delta(\epsilon)\\
&=\frac{1}{(1-\epsilon)n}\sum_{t=1}^{n}I(B,Z_{\mathrm{H}}^{t-1};Z_{\mathrm{H},t}|Z_{\mathrm{C}}^n)+\delta(\epsilon)\\
&=\frac{1}{(1-\epsilon)n}\sum_{t=1}^{n}I(B,Z_{\mathrm{H}}^{t-1},Z_{\mathrm{C}}^{t-1},Z_{\mathrm{C},t+1}^n;Z_{\mathrm{H},t}|Z_{\mathrm{C},t})\\
&\quad+\delta(\epsilon)\\
&\overset{(c)}{\leq}\frac{1}{(1-\epsilon)n}\sum_{t=1}^{n}I(B,Y^{t-1},Z_{\mathrm{C},t+1}^n;Z_{\mathrm{H},t}|Z_{\mathrm{C},t})\\
&\quad+\delta(\epsilon),
\end{aligned}
$$

where (a) holds because Z_{C}^n has the same distribution under both hypotheses; (b) holds because, conditional on Z_{C}^n, the two random variables B and Z_{H}^n have the same marginals under both hypothesis, while being dependent under $\mathcal{H}=0$ and independent under $\mathcal{H}=1$; (c) holds by the Markov chain $Z_{\mathrm{H},t}\to(B,Y^{t-1},Z_{\mathrm{C},t}^n)\to(Z_{\mathrm{C}}^{t-1},Z_{\mathrm{H}}^{t-1})$. Moreover,

$$
\begin{aligned}
T=\frac{1}{n}H(B)&\geq\frac{1}{n}I(B;Y^n|Z_{\mathrm{C}}^n)\\
&=\frac{1}{n}\sum_{t=1}^{n}I(B,Y^{t-1},Z_{\mathrm{C}}^{t-1},Z_{\mathrm{C},t+1}^n;Y_t|Z_{\mathrm{C},t})\\
&\overset{(d)}{=}\frac{1}{n}\sum_{t=1}^{n}I(B,Y^{t-1},Z_{\mathrm{C},t+1}^n;Y_t|Z_{\mathrm{C},t}),
\end{aligned}
\tag{251}
$$

where (d) holds by the Markov chain $Y_t\to(B,Z_{\mathrm{C},t}^n,Y^{t-1})\to Z_{\mathrm{C}}^{t-1}$. The proof is finalized by introducing auxiliary random variables $V_t:=(B,Y^{t-1},Z_{\mathrm{C},t+1}^n)$, $t\in\{1,\dots,n\}$, relabeling the random variables, and taking $\epsilon\to 0$.

Appendix H
Proof of Proposition 3

We fix a sufficiently large n and a sequence of encoding and decoding functions such that the properties of Definition 1 hold. Also, define $S_t:=(M,X^{t-1},Z^{t-1})$. Notice the Markov chain $S_t\to X_t\to(Y_t,Z_t)$. First, consider the rate R:

$$
\begin{aligned}
nR&=H(M)\\
&\geq I(M;X^n,Z^n)\\
&=\sum_{t=1}^{n}I(M;X_t,Z_t|X^{t-1},Z^{t-1})\\
&\overset{(a)}{=}\sum_{t=1}^{n}I(M,X^{t-1},Z^{t-1};X_t,Z_t)\\
&\geq\sum_{t=1}^{n}I(M,X^{t-1},Z^{t-1};X_t)\\
&=\sum_{t=1}^{n}I(S_t;X_t),
\end{aligned}
$$

where (a) holds by the memoryless property of the sources. Now, consider the error exponent at the relay. We have:

$$
\begin{aligned}
-\frac{1}{n}\log\beta_{y,n}&\leq\frac{1}{(1-\epsilon)n}D(P_{MY^n|\mathcal{H}=0}\|P_{MY^n|\mathcal{H}=1})+\delta(\epsilon)\\
&\overset{(b)}{=}\frac{1}{(1-\epsilon)n}I(M;Y^n)+\delta(\epsilon)\\
&=\frac{1}{(1-\epsilon)n}\sum_{t=1}^{n}I(M;Y_t|Y^{t-1})+\delta(\epsilon)\\
&\overset{(c)}{=}\frac{1}{(1-\epsilon)n}\sum_{t=1}^{n}I(M,Y^{t-1};Y_t)+\delta(\epsilon)
\end{aligned}
$$

$$\leq \frac{1}{(1-\epsilon)n}\sum_{t=1}^{n} I(M, X^{t-1}, Y^{t-1}, Z^{t-1}; Y_t) + \delta(\epsilon)$$
$$\overset{(d)}{=} \frac{1}{(1-\epsilon)n}\sum_{t=1}^{n} I(M, X^{t-1}, Z^{t-1}; Y_t) + \delta(\epsilon)$$
$$= \frac{1}{(1-\epsilon)n}\sum_{t=1}^{n} I(S_t; Y_t) + \delta(\epsilon)$$

where (b) holds because under hypothesis $\mathcal{H} = 1$, the message M and the observation Y^n are independent; (c) holds by the memoryless property of the sources; and (d) by the Markov chain $(Y^{t-1}, Z^{t-1}) \to (M, X^{t-1}) \to Y_t$. Next, consider the error exponent at the receiver:

$$-\frac{1}{n}\log \beta_{z,n} \leq \frac{1}{(1-\epsilon)n} D(P_{BZ^n|\mathcal{H}=0} \| P_{BZ^n|\mathcal{H}=1}) + \delta(\epsilon)$$
$$\overset{(e)}{\leq} \frac{1}{(1-\epsilon)n} D(P_{MY^nZ^n|\mathcal{H}=0} \| P_{MY^nZ^n|\mathcal{H}=1}) + \delta(\epsilon)$$
$$\overset{(f)}{=} \frac{1}{(1-\epsilon)n} I(M; Y^n, Z^n) + \delta(\epsilon)$$
$$\overset{(g)}{=} \frac{1}{(1-\epsilon)n} I(M; Z^n) + \delta(\epsilon)$$
$$= \frac{1}{(1-\epsilon)n}\sum_{t=1}^{n} I(M; Z_t|Z^{t-1}) + \delta(\epsilon)$$
$$= \frac{1}{(1-\epsilon)n}\sum_{t=1}^{n} I(M, Z^{t-1}; Z_t) + \delta(\epsilon)$$
$$\leq \frac{1}{(1-\epsilon)n}\sum_{t=1}^{n} I(M, X^{t-1}, Z^{t-1}; Z_t) + \delta(\epsilon)$$
$$= \frac{1}{(1-\epsilon)n}\sum_{t=1}^{n} I(S_t; Z_t) + \delta(\epsilon)$$

where (e) holds by the data processing inequality and because B is a function of M and Y^n; (f) holds because M and (Y^n, Z^n) are independent under hypothesis $\mathcal{H} = 1$ with same marginals as under $\mathcal{H} = 0$; and (g) holds by the Markov chain $M \to X^n \to Z^n \to Y^n$. The proof of the converse is finally concluded by defining a time-sharing random variable $Q \sim \mathcal{U}\{1, ..., n\}$ and $S := (S_Q, Q)$, $X := X_Q$, $Y := Y_Q$ and $Z := Z_Q$ and letting $\epsilon \to 0$ and $n \to \infty$.

PLACE PHOTO HERE

Sadaf Salehkalaibar (S'10–M'14) received the B.Sc., M.Sc. and Ph.D. degrees in Electrical Engineering from Sharif University of Technology, Tehran, Iran in 2008, 2010 and 2014, respectively. She was a postdoctoral fellow at Telecom ParisTech, Paris, France in 2015 and 2017. She is currently an assistant professor at Electrical and Computer Engineering Department of University of Tehran, Tehran, Iran. Her special fields of interest include network information theory, hypothesis testing and fundamental limits of secure communication with emphasis on information-theoretic security.

PLACE PHOTO HERE

Michèle Wigger (S'05, M'09, SM'14) received the M.Sc. degree in electrical engineering, with distinction, and the Ph.D. degree in electrical engineering both from ETH Zurich in 2003 and 2008, respectively. In 2009, she was first a post-doctoral fellow at the University of California, San Diego, USA, and then joined Telecom Paris Tech, Paris, France, where she is currently a Full Professor. Dr. Wigger has held visiting professor appointments at the Technion-Israel Institute of Technology and ETH Zurich. Dr. Wigger has previously served as an Associate Editor of the IEEE Communication Letters, and is now Associate Editor for Shannon Theory of the IEEE Transactions on Information Theory. She is currently also serving on the Board of Governors of the IEEE Information Theory Society. Dr. Wigger's research interests are in multi-terminal information theory, in particular in distributed source coding and in capacities of networks with states, feedback, user cooperation, or caching.

PLACE PHOTO HERE

Ligong Wang (S'08–M'12) received the B.E. degree in electronic engineering from Tsinghua University, Beijing, China, in 2004, and the M.Sc. and Dr.Sc. degrees in electrical engineering from ETH Zurich, Switzerland, in 2006 and 2011, respectively. In the years 2011–2014 he was a Postdoctoral Associate at the Department of Electrical Engineering and Computer Science at the Massachusetts Institute of Technology, Cambridge, MA, USA. He is now a researcher (charg de recherche) with CNRS, France, and is affiliated with ETIS laboratory in Cergy-Pontoise. His research interests include classical and quantum information theory, physical-layer security, and digital, in particular optical communication.