Distributed Hypothesis Testing Based on Unequal-Error Protection Codes

Sadaf Salehkalaibar, *Member, IEEE*, and Michèle Wigger, *Senior Member, IEEE*

Abstract—Coding and testing schemes for binary hypothesis testing over noisy networks are proposed and their corresponding type-II error exponents are derived. When communication is over a discrete memoryless channel (DMC), our scheme combines Shimokawa-Han-Amari's hypothesis testing scheme with Borade-Nakiboglu-Zheng's unequal error protection (UEP) for channel coding where source and channel codewords are simultaneously decoded. The resulting exponent is optimal for the newly introduced class of *generalized testing against conditional independence*. When communication is over a multi-access channel (MAC), our scheme combines hybrid coding with UEP. The resulting error exponent over the MAC is optimal in the case of generalized testing against conditional independence with independent observations at the two sensors when the MAC decomposes into two individual DMCs. In this case, separate source-channel coding is sufficient and no UEP is required. This same conclusion holds also under arbitrarily correlated sensor observations when testing is against independence.

Index Terms—Hypothesis testing, hybrid coding, unequal error protection (UEP), discrete memoryless channels (DMCs).

I. INTRODUCTION

SENSOR networks are important parts of the future Internet of Things (IoT). In these networks, data collected at sensors is transmitted over a wireless medium to remote decision centers, which use this information to decide on one of multiple hypotheses. We follow previous works in the information theory community [1], [2] and assume that the terminals observe memoryless sequences that follow one of two possible joint distributions, depending on the underlying hypothesis $\mathcal{H} \in \{0, 1\}$. The performance of the decision system is characterized by two error probabilities: the probability of type-I error of deciding on $\mathcal{H} = 1$ when the true hypothesis is $\mathcal{H} = 0$, and the probability of type-II error of deciding on $\mathcal{H} = 0$ when the true hypothesis is $\mathcal{H} = 1$. We consider asymmetric scenarios where one of the two errors (typically the type-II error) is more harmful than the other, and therefore a more

stringent constraint on the asymptotic decay of this error probability is imposed. Specifically, the type-I error probability can decay to 0 arbitrarily slowly in the blocklength, whereas the type-II error probability is required to decay exponentially fast. The goal of our research is to find the largest possible type-II error exponent for a given distributed decision system.

This problem statement has first been considered for the setup with a single sensor and a single decision center when communication is over a noiseless link of given capacity [1], [2]. For this canonical problem, the optimal error exponent has been identified in the special cases of *testing against independence* [1] and *testing against conditional independence*. In the former case, the joint distribution of the two sources under $\mathcal{H} = 1$ equals the product of the two marginal distributions under $\mathcal{H} = 0$. In the latter case, this product structure holds only conditional on a second observation at the decision center, which has same marginal distribution under both hypotheses. The optimal exponent for testing against conditional independence is achieved by the *Shimokawa-Han-Amari (SHA)* scheme [3], which applies Wyner-Ziv source coding combined with two local joint typicality tests at the sensor (between the quantized sequence and the sensor's observation) and at the decision center (between the quantized sequence and the decision center's observation). The decision center declares the alternative hypothesis $\mathcal{H} = 1$ whenever one of the two joint typicality tests fails. To this end, the sensor sends a special 0-message over the noiseless link to the decision center whenever its local typicality test fails. The reason for sending this special 0-message is that given the more stringent constraint on the type-II error probability, the decision center should decide on $\mathcal{H} = 1$ in case of slightest doubt.

The SHA scheme yields an achievable error exponent for all distributed hypothesis testing problems (not only testing against conditional independence) [3], but it might not be optimal in general [4]. The SHA scheme has been extended to various more involved setups such as noiseless networks with multiple sensors and a single decision center [2], [5]–[7]; networks where the sensor and the decision center can communicate interactively [8], [9]; multi-hop networks [10], and networks with multiple decision centers [10], [11].

The main focus of this paper is to extend above works to *noisy channels*. In [12], it was shown that the optimal exponent for testing against conditional independence over a discrete memoryless channel (DMC) coincides with the optimal exponent for the same test over a noiseless link of rate equal to the capacity of the DMC. A similar result is obtained also for multi-access channels (MACs) with two individual DMCs

connecting the two transmitters to the single receiver [12]. In these previous works, the optimal error exponent is thus not degraded because channels are noisy. Only capacity matters.

In this paper, we build on our conference publications [13] and [14] and propose coding and testing schemes for general hypothesis testing over DMCs and (not necessarily orthogonal) MACs. Our schemes suggest that for general hypothesis tests, the transition law of the channel matters, not only its capacity.

For DMCs, we propose a scheme that combines the SHA hypothesis testing scheme with *Borade-Nakiboglu-Zheng*'s (BNZ) Unequal Error Protection (UEP) [15], [16] coding to specially protect the 0 source-coding message. Source and channel codewords are decoded jointly, similar to [17]. Notice that under separate sequential decoding or when the UEP code is replaced by a standard channel code, the error exponent of our scheme degrades in general.

The proposed scheme achieves the optimal exponent for a generalization of testing against conditional independence where the observations at the decision center can follow a different marginal distribution under the two hypotheses. This optimal exponent is shown to be achievable also with a simplified scheme that combines the SHA hypothesis testing scheme with a good channel code in a separate source-channel coding architecture. This contrasts the findings in [18] stating that for testing against conditional independence an operational separation does not hold.

The error exponent achieved by our DMC scheme consists of three competing exponents. Two of them coincide with that of the noiseless setup [3] when the rate of the noiseless link is replaced by the mutual information between the input and output of the channel. The third error exponent coincides with BNZ's *missed-detection* exponent [15]. Depending on the DMC and the type of hypothesis test to perform, this third error exponent can be active or not. It is in particular not active for above described generalized testing against conditional independence tests, illustrating why the optimal type-II error exponent in this setup only depends on the DMC's capacity but not on its other properties.

Using hybrid coding [19], above coding and testing scheme is extended to MACs. In this case, the error exponent achieved by our scheme is expressed in terms of nine competing exponents. One of them corresponds to that of [2]; three of them correspond to an incorrect decoding of the hybrid scheme; three of them correspond to the missed-detection exponents of the UEP scheme; and the other two correspond both to the UEP mechanism and incorrect decoding. The proposed coding scheme establishes the optimal error exponent of the generalized testing against conditional independence when the sources at the transmitters are independent under both hypotheses and the MAC decomposes into two individual DMCs. In this case, hybrid coding can be replaced by separate source-channel coding and no UEP is required. Separate source-channel coding is more generally shown to attain the optimal error exponent for arbitrary source correlations when testing against independence over two individual DMCs.

The last contribution of the paper is a study of a Gaussian example with jointly Gaussian sources (that do not correspond to testing against independence) and a Gaussian MAC. For this example, the error exponents achieved by our coding and testing scheme are evaluated and numerical simulations show that they are close to a new upper bound on the optimal exponent that we derive based on Witsenhausen's max-correlation argument [20].

Following [13], in a parallel work [12], Sreekumar and Gündüz proposed two coding and testing schemes for the general hypothesis testing problem over DMCs. Their schemes either employ hybrid coding (similar to our MAC scheme) or maximum likelihood decoding. Which one of the two achieves the better exponent might depend on the source distributions and the DMC. We show in this paper that in some cases the exponent of the hybrid coding scheme strictly improves over the exponent for the DMC presented in this paper. However, as we explained in some cases our exponent is optimal and thus our simpler scheme is sufficient. For certain hypothesis testing problems, an even simpler scheme based on separate channel coding and hypothesis testing achieves the optimal type-II error exponent.

We conclude this introduction with a summary of the main contributions of the paper and remarks on notation.

A. Contributions

The main contributions of the paper are as follows.

- A coding and testing scheme for DMCs is proposed (Theorem 1 in Section II). The scheme is based on SHA's hypothesis testing scheme and UEP channel coding where source and channel codewords are decoded jointly. A matching converse is derived for generalized testing against conditional independence (Theorem 2 in Section II), thus establishing the optimal exponent for this case. The employed UEP mechanism and the joint decoding of source and channel codewords are important ingredients of our scheme. Without them, the error exponent of our scheme degrades. (See Remark 1 and Fig. 3 in Section II-D.)

- A coding and testing scheme for MACs is proposed (Theorem 4 in Section III). The scheme is based on hybrid coding and UEP. A matching converse is derived for generalized testing against conditional independence over an orthogonal MAC when the sources are independent under both hypotheses (Theorem 5 in Section III). In this special case, separate source-channel coding is sufficient and no UEP is required. Separate source-channel coding is shown to be optimal also for testing against independence under arbitrarily correlated sensor observations when the MAC decomposes into two orthogonal DMCs from each of the sensors to the decision center (Proposition 2 in Section III). The results on the MAC are concluded with the study of a Gaussian example, where the error exponent achieved by our scheme numerically matches a newly derived upper bound on the optimal error exponent (Corollary 4 and Theorem 6 in Section III-D, see also Fig. 6).

B. Notation

We mostly follow the notation in [21]. Random variables are denoted by capital letters, e.g., X, Y, and their realizations by
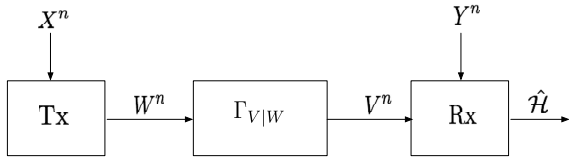
Fig. 1. Hypothesis testing over a DMC $\Gamma_{V|W}$.

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.

We write the probability mass function (pmf) of a discrete random variable X as P_X, which we also indicate as $X \sim P_X$. To distinguish the pmf of a random variable X under the two hypotheses, we also use Q_X for the pmf of X under the hypothesis $\mathcal{H} = 1$. The conditional pmf of X given Y is written as $P_{X|Y}$, or as $Q_{X|Y}$ when $\mathcal{H} = 1$.

The term $D(P\|Q)$ stands for the Kullback-Leibler (KL) divergence between two pmfs P and Q over the same alphabet. We use $\text{tp}(a^n, b^n)$ to denote the *joint type* of the pair of sequences (a^n, b^n), and $\text{ctp}(a^n|b^n)$ for the conditional type of a^n given b^n [22]. 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 $\mu > 0$, let $\mathcal{T}_\mu^n(P_X)$ be the set of *μ-typical sequences* in $\mathcal{X}^n$ as defined in [8, Sec. 2.4]. Similarly, $\mathcal{T}_\mu^n(P_{XY})$ stands for the set of *jointly μ-typical sequences*. The expectation operator is written as $\mathbb{E}[\cdot]$. We abbreviate *independent and identically distributed* by *i.i.d.*.

II. HYPOTHESIS TESTING OVER DISCRETE MEMORYLESS CHANNELS

A. System Model

Consider the distributed hypothesis testing problem in Fig. 1, where a transmitter observes source sequence X^n and a receiver source sequence Y^n. Under the null hypothesis:

$$\mathcal{H} = 0\colon (X^n, Y^n) \text{ i.i.d. } \sim P_{XY}, \tag{1}$$

and under the alternative hypothesis:

$$\mathcal{H} = 1\colon (X^n, Y^n) \text{ i.i.d. } \sim Q_{XY}. \tag{2}$$

for two given pmfs P_{XY} and Q_{XY}. The transmitter can communicate with the receiver over n uses of a discrete memory channel $(\mathcal{W}, \mathcal{V}, \Gamma_{V|W})$ where $\mathcal{W}$ denotes the finite channel input alphabet, $\mathcal{V}$ the finite channel output alphabet,

and $\Gamma_{V|W}$ the DMC transition law. Specifically, the transmitter feeds inputs

$$W^n = f^{(n)}(X^n) \tag{3}$$

to the channel, where $f^{(n)}$ denotes the chosen (possibly stochastic) encoding function

$$f^{(n)}\colon \mathcal{X}^n \to \mathcal{W}^n. \tag{4}$$

The receiver observes the ouputs V^n, where for a given input $W_t = w_t$,

$$V_t \sim \Gamma_{V|W}(\cdot|w_t), \qquad t \in \{1, \ldots, n\}. \tag{5}$$

Based on the sequence of channel outputs V^n and the source sequence Y^n, the receiver decides on the hypothesis $\mathcal{H}$. That means, it produces the guess

$$\hat{\mathcal{H}} = g^{(n)}(V^n, Y^n), \tag{6}$$

by means of a decoding function

$$g^{(n)}\colon \mathcal{V}^n \times \mathcal{Y}^n \to \{0, 1\}. \tag{7}$$

Definition 1: An exponent θ is said achievable, if for each $\epsilon > 0$ and sufficiently large blocklengths n, there exist encoding and decoding functions $(f^{(n)}, g^{(n)})$ such that the corresponding type-I and type-II error probabilities at the receiver

$$\alpha_n := \Pr[\hat{\mathcal{H}} = 1 | \mathcal{H} = 0], \tag{8}$$

$$\beta_n := \Pr[\hat{\mathcal{H}} = 0 | \mathcal{H} = 1], \tag{9}$$

satisfy

$$\alpha_n \le \epsilon, \tag{10}$$

and

$$-\varlimsup_{n\to\infty} \frac{1}{n} \log_2 \beta_n \ge \theta. \tag{11}$$

The goal is to maximize the achievable type-II error exponent θ.

B. Coding and Testing Scheme

Our coding and testing scheme combines SHA's hypothesis testing scheme for a noiseless link [3] with BNZ's UEP channel coding that protects the 0-message (which indicates that the transmitter decides on $\mathcal{H} = 1$) better than the other messages [15], [16]. In fact, since here we are only interested in the type-II error exponent, the receiver should decide on $\mathcal{H} = 0$ only if the transmitter also shares this opinion.

We describe the coding and testing scheme in detail. The analysis is presented in Appendix A.

Preparations: Choose a large positive integer n, an auxiliary distribution P_T over $\mathcal{W}$, a conditional channel input distribution $P_{W|T}$, and a conditional source distribution $P_{S|X}$ over a finite auxiliary alphabet $\mathcal{S}$ so that

$$I(S; X) < I(S; Y) + I(V; W|T), \tag{12}$$

where mutual informations in this section are calculated according to the following joint distribution

$$P_{SXYWVT} = P_{S|X} \cdot P_{XY} \cdot P_T \cdot P_{W|T} \cdot \Gamma_{V|W}. \tag{13}$$
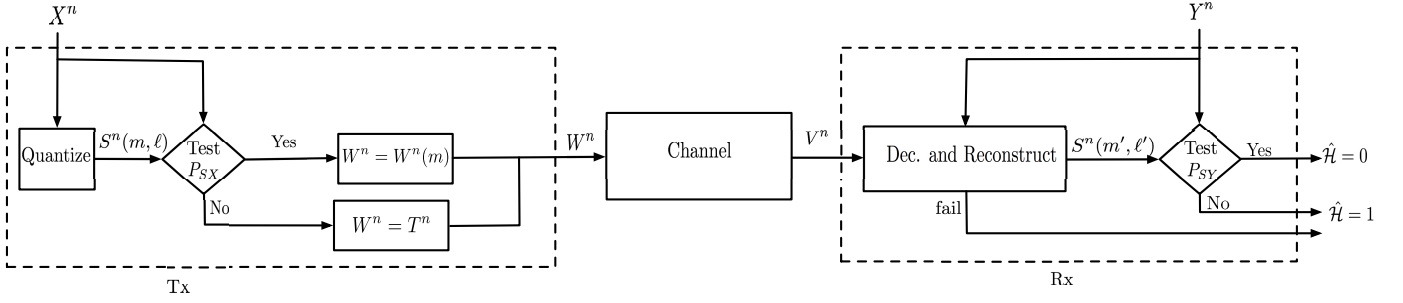
Fig. 2.   Coding and testing scheme for hypothesis testing over a DMC.

Then, choose a sufficiently small $\mu > 0$. If $I(S; X) < I(W; V|T)$, let

$$R = I(S; X) + \mu, \tag{14}$$
$$R' = 0. \tag{15}$$

If $I(S; X) \geq I(W; V|T)$, let

$$R = I(W; V|T) - \mu, \tag{16}$$
$$R' = I(S; X) - I(W; V|T) + 2\mu. \tag{17}$$

Code Construction: Construct a random codebook

$$\mathcal{C}_S = \left\{ S^n(m, \ell) \colon m \in \{1, ..., \lfloor 2^{nR} \rfloor\}, \ell \in \{1, ..., \lfloor 2^{nR'} \rfloor\} \right\}, \tag{18}$$

by independently drawing all codewords i.i.d. according to $P_S(s) = \sum_{x \in \mathcal{X}} P_X(x) P_{S|X}(s|x)$.

Generate a sequence T^n i.i.d. according to P_T. Construct a random codebook

$$\mathcal{C}_W = \left\{ W^n(m) \colon m \in \{1, ..., \lfloor 2^{nR} \rfloor\} \right\}$$

superpositioned on T^n where each codeword is drawn independently according to $P_{W|T}$ conditioned on T^n. Reveal the realizations of the codebooks and the realization of the time-sharing sequence $T^n = t^n$ to all terminals.

Transmitter: Given that it observes the source sequence $X^n = x^n$, the transmitter looks for a pair (m, ℓ) that satisfies

$$(s^n(m, \ell), x^n) \in \mathcal{T}_{\mu/2}^n(P_{SX}). \tag{19}$$

If successful, it picks one of these pairs uniformly at random and sends the codeword $w^n(m)$ over the channel. Otherwise it sends the sequence of inputs t^n over the channel.

Receiver: Assume that $V^n = v^n$ and $Y^n = y^n$. The receiver first looks for indices $m' \in \{1, \ldots, \lfloor 2^{nR} \rfloor\}$ and $\ell' \in \{1, \ldots, \lfloor 2^{nR'} \rfloor\}$ such that the following two conditions are simultaneously satisfied:

$$(t^n, w^n(m'), v^n) \in \mathcal{T}_{\mu}^n(P_{TWV}), \tag{20}$$

and

$$H_{\text{tp}(s^n(m', \ell'), y^n)}(S|Y) = \min_{\tilde{m}, \tilde{\ell}} H_{\text{tp}(s^n(\tilde{m}, \tilde{\ell}), y^n)}(S|Y). \tag{21}$$

If one or multiple such pairs exist, it chooses one of them uniformly at random and checks whether the chosen pair (m', ℓ') satisfies the following typicality check:

$$(s^n(m', \ell'), y^n) \in \mathcal{T}_{\mu}^n(P_{SY}). \tag{22}$$

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

C. Results on the Error Exponent

The coding and testing scheme described in the previous section allows to establish the following theorem.

Theorem 1: Every error exponent $\theta \geq 0$ that satisfies the following condition (23) is achievable:

$$\theta \leq \max_{\substack{P_{S|X}, P_T, P_{W|T}: \\ I(S; X|Y) \leq I(W; V|T)}} \min \left\{ \theta^{\text{standard}}, \; \theta^{\text{dec}}, \; \theta^{\text{miss}} \right\}, \tag{23}$$

where for given (conditional) pmfs $P_{S|X}$, P_T, and $P_{W|T}$ we define the joint pmf

$$P_{SXYWVT} := P_{S|X} \cdot P_{XY} \cdot P_T \cdot P_{W|T} \cdot \Gamma_{V|W}. \tag{24}$$

and the exponents

$$\theta^{\text{standard}} := \min_{\substack{\tilde{P}_{SXY}: \\ \tilde{P}_{SX} = P_{SX} \\ \tilde{P}_{SY} = P_{SY}}} D(\tilde{P}_{SXY} \| P_{S|X} Q_{XY}), \tag{25}$$

$$\theta^{\text{dec}} := \min_{\substack{\tilde{P}_{SXY}: \\ \tilde{P}_{SX} = P_{SX} \\ \tilde{P}_Y = P_Y \\ H(S|Y) \leq H_{\tilde{P}}(S|Y)}} D(\tilde{P}_{SXY} \| P_{S|X} Q_{XY})$$
$$+ I(V; W|T) - I(S; X|Y), \tag{26}$$

$$\theta^{\text{miss}} := D(P_Y \| Q_Y) + I(V; W|T) - I(S; X|Y)$$
$$+ \sum_{t \in \mathcal{W}} P_T(t) \cdot D(P_{V|T=t} \| \Gamma_{V|W=t}). \tag{27}$$

Here, mutual informations, the conditional entropy term $H(S|Y)$, and the conditional marginal pmf $P_{V|T}$ are calculated with respect to the joint distribution in (24).

Proof: See Appendix A. ∎

The expressions in Theorem 1 show three competing error exponents. In (25) and (26), we recognize the two competing error exponents of the SHA scheme for the noiseless setup: θ^{standard} is the exponent associated with the event that the receiver reconstructs the correct binned codeword and decides on $\hat{\mathcal{H}} = 0$ instead of $\mathcal{H} = 1$, and θ^{dec} is associated with the event that either the binning or the noisy channel introduces a decoding error followed by a wrong decision on the hypothesis. The exponent θ^{miss} in (27) is new and can be associated with the event that the specially protected 0-message is wrongly decoded followed by a wrong decision

on the hypothesis. We remark in particular that θ^{miss} contains the term

$$E_{\text{miss}} := \sum_{t \in \mathcal{W}} P_T(t) \cdot D(P_{V|T=t} \| P_{V|W=t}), \qquad (28)$$

which represents the largest possible *miss-detection exponent* for a single specially protected message at a rate $I(W; V|T)$ [15].

Which of the three exponents $\theta^{\text{standard}}, \theta^{\text{dec}}, \theta^{\text{miss}}$ is smallest depends on the source and channel parameters and of the choice of $P_{S|X}$, P_T, and $P_{W|T}$.

Remark 1 (Separate Versus Joint Source-Channel Decoding): The scheme in Subsection II-B is a joint source-channel coding scheme: it employs independent source and channel codebooks but the receiver jointly decodes the channel and source codewords.[1]

We compare our proposed error exponent with that of the corresponding separate source-channel coding scheme where the receiver first decodes the channel codeword using joint typicality decoding and then the source codeword using Han's conditional minimum-entropy decoder. The error exponent achieved by this separation-based scheme is generally smaller and given by:

$$\theta^{\text{dec}}_{\text{sep}} = \max_{\substack{P_{S|X}, P_T, P_{W|T}: \\ I(S;X|Y) \leq I(W;V|T)}} \min \left\{ \theta^{\text{standard}}, \ \theta^{\text{dec}}_{\text{sep}}, \ \theta^{\text{miss}} \right\}, \quad (29)$$

where θ^{standard} and θ^{miss} are defined as before and

$$\theta^{\text{dec}}_{\text{sep}} := \min_{\substack{\tilde{P}_{SXY}: \\ \tilde{P}_{SX}=P_{SX} \\ \tilde{P}_Y=P_Y}} D(\tilde{P}_{SXY} \| P_{S|X} Q_{XY}) + I(V;W|T) - I(S;X|Y).$$

$$(30)$$

We notice that $\theta^{\text{dec}}_{\text{sep}}$ is generally smaller than θ^{dec} because of the relaxed minimization, which does not include the conditional entropy constraint. We thus conclude that changing the joint source-channel coding scheme presented in Subsection II-B to separate source-channel coding can reduce the achieved error exponent if θ^{dec} is a minimizer of (23) and the entropy constraint in the minimization of θ^{dec} is stringent.

Remark 2 (UEP Versus Standard Channel Coding): If in the presented scheme, the UEP sequence t^n is set to a constant and the encoder sends a standard random codeword $w^n(0)$ instead of t^n, then the miss-detection error exponent θ^{miss} needs to be replaced by the exponent

$$\theta^{\text{miss}}_{\text{no-UEP}} := D(P_Y \| Q_Y) + I(W; V) - I(S; X|Y). \quad (31)$$

In this case, the error exponent achieved by our scheme depends on the DMC only through the mutual information $I(W; V)$, and if W is chosen the capacity-achieving input distribution, then only through its capacity C. Since the exponents $\theta^{\text{miss}}_{\text{no-UEP}}$ and θ^{dec} are increasing in $I(W; V)$, choosing a capacity-achieving W maximizes the error exponent achieved by our scheme without UEP.

[1] The error exponent of Theorem 1 was first given in [13] where it was mistakenly reported to be achievable with separate source-channel coding. A careful analysis reveals that the separate source-channel coding scheme in [13] achieves only the exponent in (29).

Notice further that the exponent $\theta^{\text{miss}}_{\text{no-UEP}}$ is smaller or equal to both $\theta^{\text{dec}}_{\text{sep}}$ and θ^{dec}, and therefore, without UEP, the separate and the joint source-channel coding schemes perform equally well.

From the previous two remarks, we conclude that generally the UEP mechanism and the joint decoding are necessary in our scheme to attain the error exponent in Theorem 1. The following corollary shows that for a certain class of source distributions they are not beneficial and the performance of Theorem 1 is also achieved by a simple scheme that combines SHA hypothesis testing with a random channel code in a separate souce-channel coding architecture.

Corollary 1: For source distributions P_{XY} and Q_{XY} where irrespective of the choice of the auxiliary distribution $P_{S|X}$:

$$\min_{\substack{\tilde{P}_{SXY}: \\ \tilde{P}_{SX}=P_{SX} \\ \tilde{P}_Y=P_Y \\ H(S|Y) \leq H_{\tilde{P}}(S|Y)}} \mathbb{E}_{P_Y}[D(\tilde{P}_{SX|Y} \| P_{S|X} Q_{X|Y})] = 0, \quad (32)$$

error exponent θ^{miss} is never smaller than θ^{dec}, and therefore non-active. In this case, it is best to choose T a constant and W a capacity-achieving input. So, when condition (32) is satisfied for all auxiliary distributions $P_{S|X}$, then Theorem 1 specializes to:

$$\theta \leq \max_{\substack{P_{S|X}: \\ I(S;X|Y) \leq C}} \min \left\{ \tilde{\theta}^{\text{standard}}, \tilde{\theta}^{\text{dec}} \right\}, \quad (33)$$

where

$$\tilde{\theta}^{\text{standard}} := \min_{\substack{\tilde{P}_{SXY}: \\ \tilde{P}_{SX}=P_{SX} \\ \tilde{P}_{SY}=P_{SY}}} D(\tilde{P}_{SXY} \| Q_{XY} P_{S|X}), \quad (34)$$

$$\tilde{\theta}^{\text{dec}} := D(P_Y \| Q_Y) + C - I(S; X|Y). \quad (35)$$

Notice that since W is chosen to achieve capacity, $\tilde{\theta}^{\text{dec}} = \theta^{\text{miss}}_{\text{no-UEP}}$ and moreover for the considered sources $\tilde{\theta}^{\text{dec}} = \theta^{\text{dec}}_{\text{sep}}$. Thus, the exponent in (33) is achievable by combining, in a separate source-channel coding architecture, SHA's hypothesis testing scheme with a random channel code $\mathcal{C}_W$ whose entries are drawn i.i.d. according to a capacity-achieving input distribution and decoding is performed using joint typicality or ML decoding.

We consider a special case where the expression in (33) can be further simplified and the resulting exponent can be proved to be optimal.

Theorem 2: If there exists a function f from $\mathcal{Y}$ to an auxiliary domain $\mathcal{Z}$ so that

$$\text{under } \mathcal{H} = 1: \quad X \to f(Y) \to Y, \quad (36)$$

and the pair $(X, f(Y))$ has the same distribution under both hypotheses, then the optimal error exponent is:

$$\theta^* = D(P_Y \| Q_Y) + \max_{\substack{P_{S|X}: \\ I(S;X|f(Y)) \leq C}} I(S; Y|f(Y)), \quad (37)$$

where C denotes the capacity of the DMC.

Proof: The achievability is based on Theorem 1 and observing that the setup under consideration satisfies Condition (32). See Appendix B. ∎

This theorem recovers the optimal error exponents for testing against conditional independence over a noisy channel [12, Lemma 5] or over a noiseless link [5, Theorem 1]. It is however more general, because here Y can have different marginals under both hypotheses. As we explained at the end of Corollary 1 above, the optimal exponent θ^* is achievable even with a basic separate source-channel coding scheme that combines the SHA scheme in [3] with a standard capacity-achieving random code with a joint-typicality (or ML) decoder. This contrasts the findings in [18] stating that for the problem of testing against conditional independence an operational separation between hypothesis testing and channel coding does not hold.

Proposition 1: The result of Theorem 2 remains valid when there is instantaneous noise-free feedback from the receiver to the transmitter.

Proof: A close inspection reveals that the converse proof of the theorem remains valid even with feedback. ∎

Theorem 2 was stated for discrete memoryless sources. It can be shown that it remains valid when the sources are memoryless and jointly Gaussian.

Example 1 (Theorem 2 for Gaussian Sources): For given $\rho_0 \in [0, 1]$, define the two covariance matrices

$$\mathbf{K}_{XY}^0 = \begin{bmatrix} 1 & \rho_0 \\ \rho_0 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{K}_{XY}^1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \quad (38)$$

Under the null hypothesis,

$$\mathcal{H} = 0: \quad (X, Y) \sim \mathcal{N}(0, \mathbf{K}_{XY}^0), \quad (39)$$

and under the alternative hypothesis,

$$\mathcal{H} = 1: \quad (X, Y) \sim \mathcal{N}(0, \mathbf{K}_{XY}^1). \quad (40)$$

Moreover, assume that the transmitter communicates to the receiver over a DMC of capacity C. This setup is a special case of Theorem 2. Appendix C shows that in this case, the optimal error exponent in (37) evaluates to:

$$\theta^* = \frac{1}{2} \log_2 \left(\frac{1}{1 - \rho_0^2 + \rho_0^2 \cdot 2^{-2C}} \right). \quad (41)$$

This result recovers as a special case the optimal exponent for testing against independence of Gaussian sources over a noiseless link in [5, Corollary 7].

D. Numerical Example to Theorem 1

We now present an example and evaluate the largest type-II error exponents attained by Theorem 1 for this example. We also show that depending on the parameters of the sources or the channel, a different error exponent θ^{standard}, θ^{dec}, or θ^{miss} is active. Let under the null hypothesis

$$\mathcal{H} = 0: \quad X \sim \text{Bern}(p_0), \quad Y = X \oplus N_0,$$
$$N_0 \sim \text{Bern}(q_0), \quad (42)$$

for N_0 independent of X. Under the alternative hypothesis:

$$\mathcal{H} = 1: \quad X \sim \text{Bern}(p_1), \quad Y \sim \text{Bern}(p_0 \star q_0), \quad (43)$$

with X and Y independent. Assume that $\Gamma_{V|W}$ is a binary symmetric channel (BSC) with cross-over probability $r \in [0, 1/2]$.

For this example, $Q_{XY} = Q_X P_Y$ and therefore θ^{standard} simplifies to:

$$\theta^{\text{standard}} = D(P_{SX} \| P_{S|X} Q_X)$$
$$+ \min_{\substack{\tilde{P}_{SXY}: \\ \tilde{P}_{SX} = P_{SX} \\ \tilde{P}_{SY} = P_{SY}}} \mathbb{E}_{P_{SX}} \left[D(\tilde{P}_{Y|SX} \| P_Y) \right] \quad (44)$$

$$= D(P_X \| Q_X) + I(S; Y), \quad (45)$$

Moreover, since $\tilde{P}_{SXY} = P_{SX} P_Y$ is a valid choice in the minimization of θ^{dec}, this latter exponent simplifies to:

$$\theta^{\text{dec}} = D(P_{SX} \| P_{S|X} Q_X)$$
$$+ \min_{\substack{\tilde{P}_{SXY}: \\ \tilde{P}_{SX} = P_{SX} \\ \tilde{P}_Y = P_Y \\ H(S|Y) \leq H_{\tilde{P}}(S|Y)}} D(\tilde{P}_{Y|SX} \| P_Y) + I(V; W|T) - I(S; X|Y)$$
$$(46)$$

$$= D(P_X \| Q_X) + I(V; W|T) - I(S; X|Y). \quad (47)$$

As a consequence, Theorem 1 simplifies to:

$$\theta \leq \max_{\substack{P_{S|X}, P_{TW}: \\ I(S;X|Y) \leq I(W;V|T)}} \min \left\{ \theta^{\text{standard}}, \theta^{\text{dec}}, \theta^{\text{miss}} \right\}, \quad (48)$$

where

$$\theta^{\text{standard}} \leq D(P_X \| Q_X) + I(S; Y), \quad (49)$$

$$\theta^{\text{dec}} \leq D(P_X \| Q_X) + I(V; W|T) - I(S; X|Y), \quad (50)$$

$$\theta^{\text{miss}} \leq \sum_{t \in \mathcal{W}} P_T(t) D(P_{V|T=t} \| P_{V|W=t})$$
$$+ I(V; W|T) - I(S; X|Y). \quad (51)$$

Depending on the parameters of the setup and the choice of the auxiliary distributions, either of the exponents $\theta^{\text{standard}}, \theta^{\text{dec}}$, or θ^{miss} is active. For example, when the cross-over probability of the BSC is large, $r \geq 0.4325$, then

$$D(P_X \| Q_X) \geq \sum_{t \in \mathcal{W}} P_T(t) D(P_{V|T=t} \| \Gamma_{V|W=t}) + I(V; W|T), \quad (52)$$

and irrespective of the choice of the random variables S, T, W the exponent θ^{miss} is smaller than θ^{standard} and θ^{dec}. Since $I(S; X|Y) \geq 0$, it is then optimal to choose S constant and (T, W) so as to maximize the sum

$$\sum_{t \in \mathcal{W}} P_T(t) D(P_{V|T=t} \| \Gamma_{V|W=t}) + I(V; W|T)$$
$$= \sum_{t, w \in \mathcal{W}} P_{TW}(t, w) D(\Gamma_{V|W=w} \| \Gamma_{V|W=t}). \quad (53)$$

That means, choose W and T deterministically equal to two maximally distinguishable inputs. Since on a BSC there are only two inputs (0 and 1) and the channel law is symmetric

TABLE I
THE SMALLEST ERROR EXPONENT AS A FUNCTION OF r

	$0 \leq r \leq 0.286$	$0.286 \leq r \leq 0.5$
$\hat{\theta}_{\text{NoUEP}}$	$\theta^{\text{miss}}_{\text{no-UEP}} = \theta^{\text{standard}}$	$\theta^{\text{miss}}_{\text{no-UEP}} \leq \theta^{\text{standard}}$

	$0 \leq r \leq 0.046$	$0.046 \leq r \leq 0.296$	$0.296 \leq r \leq 0.351$	$0.351 \leq r \leq 0.5$
$\hat{\theta}$	$\theta^{\text{standard}} \leq \min\{\theta^{\text{dec}}, \theta^{\text{miss}}\}$	$\theta^{\text{dec}} = \theta^{\text{standard}} \leq \theta^{\text{miss}}$	$\theta^{\text{standard}} = \theta^{\text{miss}} \leq \theta^{\text{dec}}$	$\theta^{\text{miss}} \leq \min\{\theta^{\text{standard}}, \theta^{\text{dec}}\}$

with respect to these inputs, for $r \in (0.4325, 0.5)$ the largest error exponent achieved by our scheme is:

$$\hat{\theta} := \max_{\substack{P_{S|X}, P_{TW}: \\ I(S;X|Y) \leq I(W;V|T)}} \min\{\theta^{\text{standard}}, \theta^{\text{dec}}, \theta^{\text{miss}}\}$$

$$= D(P_{V|W=0} \| P_{V|W=1}) = (1-2r) \log \frac{1-r}{r}. \quad (54)$$

For example, when $r = \frac{4}{9}$, one obtains $\hat{\theta} = 0.0358 = \frac{1}{9} \log \frac{5}{4}$.

In contrast, when the cross-over probability of the BSC is small, the miss-detection exponent (28) is large and the exponent θ^{miss} is never active irrespective of the choice of the auxiliary random variable S. The overall exponent is then determined by the smaller of θ^{standard} and θ^{dec}, and in particular by a choice S, X, W that makes the two equal. In this case, for a scenario with parameters $p_0 = 0.2, q_0 = 0.3, p_1 = 0.4$, and $r = 0.1$, the largest exponent achieved by our scheme is $\theta = 0.19$. Notice that it suffices to consider auxiliary random variables S over alphabets of size $|\mathcal{X}| + 1$. This can be proved using standard arguments as explained in [21, Appendix C].

In the following, we study the maximum error exponent achieved by our scheme $\hat{\theta}$ in function of the channel cross-over probability r. This dependency is shown in Figure 3, and Table I indicates which of the three exponents $\theta^{\text{standard}}, \theta^{\text{dec}}, \theta^{\text{miss}}$ is smallest. Notice that for $r \geq 0.296$, error exponent θ^{miss} is smallest, and for $r \leq 0.046$, error exponent θ^{standard} is smallest.

An important feature of our scheme is the UEP mechanism used to send the 0-message. As explained in Remark 2, without UEP, exponent θ^{miss} in (51) needs to be replaced by the smaller exponent

$$\theta^{\text{miss}}_{\text{no-UEP}} = D(P_Y \| Q_Y) + I(V;W) - I(S;X|Y). \quad (55)$$

Notice that $\theta^{\text{miss}}_{\text{no-UEP}} \leq \theta^{\text{dec}}$ and thus without UEP our coding and testing scheme would achieve only exponents that satisfy

$$\theta \leq \hat{\theta}_{\text{no-UEP}} := \max_{\substack{P_{S|X}, P_{TW}: \\ I(S;X|Y) \leq I(W;V|T)}} \min\{\theta^{\text{standard}}, \theta^{\text{miss}}_{\text{NoUEP}}\},$$

$$(56)$$

Figure 3 also shows the exponent in (56).

E. Comparison With the Parallel Work [12]

In our conference publication [13] we used UEP and the conditional entropy decoder for the problem of distributed hypothesis testing over a noisy channel. Sreekumar and Gündüz proposed in [12] to replace the separate source and channel codes with a single hybrid code, similarly to our hybrid scheme for the MAC presented in the following Section III. (See also the conference publication [14]). Additionally, unlike in our MAC scheme, [12] lets the UEP codeword depend on the encoder's source sequence X^n. As we will see at the end of this section, this dependence can enhance the miss-detection error exponent.

We recall here the exponent achieved with hybrid coding in [12].

Theorem 3 (Theorem 5 in [12]): Every error exponent $\theta \geq 0$ that satisfies the following condition (57) is achievable:

$$\theta \leq \max_{\substack{P_T, P_{S|XT}, P_{W|XST}, P_{W'|XT}: \\ I(S;X|T) \leq I(S;V,Y|T)}} \min\{\theta^{\text{standard}}_{\text{hyb}}, \theta^{\text{dec}}_{\text{hyb}}, \theta^{\text{miss}}_{\text{hyb}}\},$$

$$(57)$$

where for given (conditional) pmfs $P_T, P_{S|XT}, P_{W|SXT}, P_{W'|XT}$, we define the joint pmfs

$$P_{SXYWVT} = P_{XY} \cdot P_T \cdot P_{S|XT} \cdot P_{W|XST} \cdot \Gamma_{V|W}, \quad (58)$$

$$Q_{SXYWVT} = Q_{XY} \cdot P_T \cdot P_{S|XT} \cdot P_{W|XST} \cdot \Gamma_{V|W}, \quad (59)$$

$$Q'_{XYW'VT} = Q_{XY} \cdot P_T \cdot P_{W'|XT} \cdot \Gamma_{V|W}, \quad (60)$$

and the exponents

$$\theta^{\text{standard}}_{\text{hyb}} := \min_{\substack{\tilde{P}_{SXYVT}: \\ \tilde{P}_{SXT} = P_{SXT} \\ \tilde{P}_{SYVT} = P_{SYVT}}} D(\tilde{P}_{SXYVT} \| Q_{SXYVT}), \quad (61)$$

$$\theta^{\text{dec}}_{\text{hyb}} := \min_{\substack{\tilde{P}_{SXYVT}: \\ \tilde{P}_{SXT} = P_{SXT} \\ \tilde{P}_{YVT} = P_{YVT} \\ H(S|YVT) \leq H_{\tilde{P}}(S|YVT)}} D(\tilde{P}_{SXYVT} \| Q_{SXYVT})$$

$$+ I(S;Y,V|T) - I(S;X|T), \quad (62)$$

$$\theta^{\text{miss}}_{\text{hyb}} := D(P_{YVT} \| Q'_{YV'T}) + I(S;Y,V|T) - I(S;X|T). \quad (63)$$

Unless otherwise stated, mutual informations and entropies are with respect to the pmf defined in (58).

Theorem 3 recovers our Theorem 1 as a special case when $S = (S', W)$ and $W' = T$ with the pair (S', X) independent of the pair (T, W). (Here, the random variable S' takes on
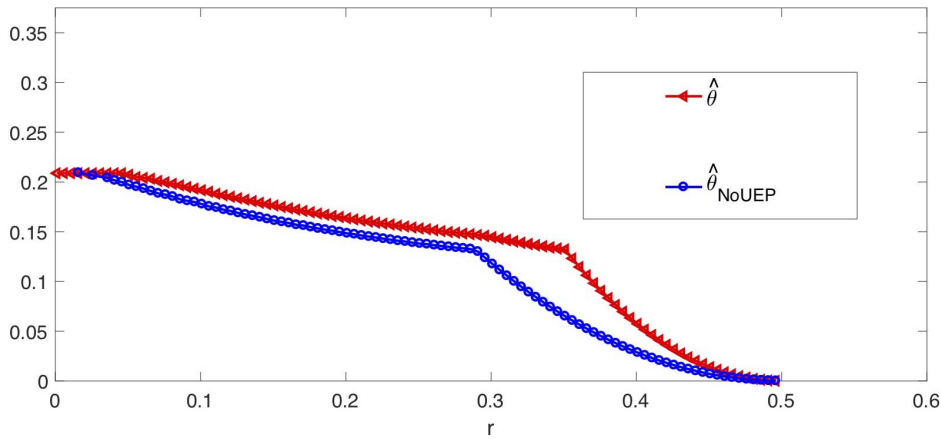
Fig. 3. The achievable error exponents with and without unequal error protection, $\hat{\theta}$ in (54) and $\hat{\theta}_{\text{NoUEP}}$ in (56), for the proposed example with $p_0 = 0.2$, $p_1 = 0.4$ and $q_0 = 0.3$.

the role of S in Theorem 1.) Inspired by [12, Example 1] we present a simple example where above exponent of [12, Theorem 5] outperforms our Theorem 1. In fact, simple uncoded transmission achieves a better exponent for this example than Theorem 1.

Example 2: Let both sources be Bernoulli-$1/2$ and let $X = Y$ under $\mathcal{H} = 0$ but $X \neq Y$ under $\mathcal{H} = 1$. The two source distributions P_{XY} and Q_{XY} are thus as indicated by the following two tables:

P_{XY}	$X = 0$	$X = 1$
$Y = 0$	1	0
$Y = 1$	0	1

and

Q_{XY}	$X = 0$	$X = 1$
$Y = 0$	0	1
$Y = 1$	1	0

We further consider a binary symmetric channel (BSC) of parameter $\alpha \in [0, 1]$, i.e.,

$$\Gamma_{V|W}(v|w) = \begin{cases} 1 - \alpha, & \text{if } v = w \\ \alpha, & \text{if } v \neq w. \end{cases} \quad (64)$$

We first upper bound the achievable error exponent of Theorem 1 and then show that this upper bound is outperformed by uncoded transmission. To upper bound the exponent of Theorem 1, we simply focus on the exponent θ^{dec}. We notice that the minimization in θ^{dec} evaluates to zero by the choice $\tilde{P}_{SXY} = P_{SX} \cdot Q_{Y|X}$, which satisfies the optimization constraints because $H_P(S|Y) = H_P(S|X) = H_{\tilde{P}}(S|X) = H_{\tilde{P}}(S|Y)$. Therefore,

$$\theta^{\text{dec}} = I(W; V|T) - I(S; X|Y)$$
$$\leq C_{\text{BSC}}(\alpha) = 1 - h_{\text{b}}(\alpha), \quad (65)$$

where $C_{\text{BSC}}(\alpha)$ denotes the capacity of a BSC of parameter α and $h_{\text{b}}(\alpha)$ denotes the binary entropy function. We conclude that the exponent θ achieved by Theorem 1 for this example is upper bounded by the capacity of the DMC $1 - h_{\text{b}}(\alpha)$.

The achievable error exponent of uncoded transmission without UEP can be obtained from above Theorem 3 by

choosing the pmfs P_T, $P_{S|XT}$, $P_{W|XST}$, and $P_{W'|XT}$ such that $W' = W = X$ and S and T are both deterministic, e.g., equal to 0. In this case, since $W' = W$ the two joint pmfs $Q'_{YW'V}$ and $Q'_{YW'V}$ coincide and since S is deterministic, the mutual information terms in Theorem 3 all vanish. Moreover, the choice $\tilde{P}_{XYV} = P_{YV} \cdot Q_{X|VY}$ is the optimal choice in both minimizations, which both evaluate to $D(P_{YV} \| Q_{YV})$. As a consequence, for the described choice, i.e., for uncoded transmission, we have:

$$\theta^{\text{standard}}_{\text{hyb}} = \theta^{\text{dec}}_{\text{hyb}} = \theta^{\text{miss}}_{\text{hyb}}$$
$$= D(P_{YV} \| Q_{YV}) \quad (66)$$
$$= \frac{1}{2} D(P_{V|Y=0} \| Q_{V|Y=0}) + \frac{1}{2} D(P_{V|Y=1} \| Q_{V|Y=1}) \quad (67)$$
$$= 2 \cdot \frac{1}{2} \left(\alpha \log_2 \frac{\alpha}{1 - \alpha} + (1 - \alpha) \log_2 \frac{1 - \alpha}{\alpha} \right) \quad (68)$$
$$= \alpha \log_2 \frac{1}{1 - \alpha} + (1 - \alpha) \log_2 \frac{1}{\alpha} - h_b(\alpha) \quad (69)$$
$$\geq 1 - h_b(\alpha), \quad (70)$$

where the inequality is strict unless $\alpha = 1/2$.

We conclude that for this example, the exponent achieved by Theorem 3 is strictly larger than the exponent achieved by our Theorem 1 whenever $\alpha \neq 1/2$.

Finally, notice that the desired performance of Theorem 3 was achieved by the choice $W' = X$, in which case the UEP codeword coincides with the source sequence X^n. If we let W' only depend on T, i.e., $W' = T = 0$, the miss-detection exponent would only evaluate to

$$D(P_{YV} \| Q_Y \Gamma_{V|W=T}) = \frac{1}{2} D(P_{V|Y=1} \| Q_{V|W=0}) \quad (71)$$
$$= \frac{1}{2} \left(\alpha \log_2 \frac{\alpha}{1 - \alpha} + (1 - \alpha) \log_2 \frac{1 - \alpha}{\alpha} \right), \quad (72)$$

and thus would be smaller.

The work [12] also proposed a separate channel-coding and hypothesis-testing scheme based on maximum likelihood decoding of the channel codewords. The corresponding error exponent in [12, Theorem 2] is generally larger than the one
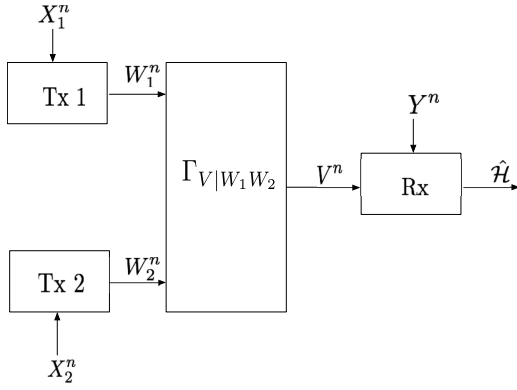
Fig. 4. Hypothesis testing over a noisy MAC.

discussed in Remark 1, which is based on joint typicality decoding. It is unclear at the moment whether in certain cases the error exponent in [12, Theorem 2] can even improve over the exponent in [12, Theorem 5] which applies the joint typicality decoding common to hybrid coding.

III. HYPOTHESIS TESTING OVER MULTI-ACCESS CHANNELS

A. System Model

Consider a setup with two sensors that communicate to a single decision center over a discrete memoryless multiple-access channel (MAC), see Fig. 4. The channel is described by the quadruple $(\mathcal{W}_1 \times \mathcal{W}_2, \mathcal{V}, \Gamma_{V|W_1,W_2})$, where $\mathcal{W}_1$ and $\mathcal{W}_2$ denote the finite channel input alphabets and $\mathcal{V}$ denotes the finite channel output alphabet. Each transmitter i ($i = 1, 2$) observes the sequence X_i^n and produces channel inputs W_i^n as

$$W_i^n = f_i^{(n)}(X_i^n) \qquad (73)$$

by means of a possibly stochastic encoding function

$$f_i^{(n)}: \mathcal{X}_i^n \to \mathcal{W}_i^n. \qquad (74)$$

The receiver observes the corresponding channel outputs V^n as well as the source sequence Y^n. Under the null hypothesis

$$\mathcal{H} = 0: \quad (X_1^n, X_2^n, Y^n) \text{ i.i.d. } \sim P_{X_1 X_2 Y}, \qquad (75)$$

and under the alternative hypothesis

$$\mathcal{H} = 1: \quad (X_1^n, X_2^n, Y^n) \text{ i.i.d. } \sim Q_{X_1 X_2 Y}, \qquad (76)$$

for two given pmfs $P_{X_1 X_2 Y}$ and $Q_{X_1 X_2 Y}$. The receiver should decide on the hypothesis $\mathcal{H}$. Besides Y^n, it also observes the MAC ouputs V^n, where for given inputs $W_{1,t} = w_{1,t}$ and $W_{2,t} = w_{2,t}$,

$$V_t \sim \Gamma_{V|W_1 W_2}(\cdot|w_{1,t}, w_{2,t}), \qquad t \in \{1, \ldots, n\}. \qquad (77)$$

It thus produces the guess

$$\hat{\mathcal{H}} = g^{(n)}(V^n, Y^n) \qquad (78)$$

using a decoding function

$$\mathcal{V}^n \times \mathcal{Y}^n \to \{0, 1\}. \qquad (79)$$

Definition 2: An exponent θ is said achievable, if for each $\epsilon > 0$ and sufficiently large blocklength n, there exist encoding and decoding functions $(f^{(n)}, g^{(n)})$ such that the corresponding type-I and type-II error probabilities at the receiver

$$\alpha_n := \Pr[\hat{\mathcal{H}} = 1|\mathcal{H} = 0], \qquad (80)$$
$$\beta_n := \Pr[\hat{\mathcal{H}} = 0|\mathcal{H} = 1], \qquad (81)$$

satisfy

$$\alpha_n \leq \epsilon, \qquad (82)$$

and

$$-\varlimsup_{n \to \infty} \frac{1}{n} \log_2 \beta_n \geq \theta. \qquad (83)$$

The goal is to maximize the type-II error exponent θ.

B. Coding and Testing Scheme

Our coding and testing scheme for the MAC combines local hypothesis tests at the transmitters with hybrid coding [19] and UEP codes. Details are as follows.

We describe a coding and testing scheme for distributed hypothesis testing over a noisy MAC, see Fig. 5.

Preparations: Choose a sufficiently large blocklength n, auxiliary alphabets $\mathcal{S}_1$ and $\mathcal{S}_2$, and functions

$$f_i: \mathcal{S}_i \times \mathcal{X}_i \to \mathcal{W}_i, \qquad i \in \{1, 2\}, \qquad (84)$$

and define the shorthand notation

$$\Gamma_{V|S_1 S_2 X_1 X_2}(v|s_1, s_2, x_1, x_2) :=$$
$$\Gamma_{V|W_1 W_2}(v|f_1(s_1, x_1), f_2(s_2, x_2)),$$
$$\forall s_1 \in \mathcal{S}_1, s_2 \in \mathcal{S}_2, x_1 \in \mathcal{X}_1, x_2 \in \mathcal{X}_2. \qquad (85)$$

Choose then a distribution $P_{T_1 T_2}$ over $\mathcal{W}_1 \times \mathcal{W}_2$, and for $i \in \{1, 2\}$, a conditional distribution $P_{S_i|X_i T_1 T_2}$ over $\mathcal{S}_i$ in a way that:

$$I(S_1; X_1|T_1, T_2) < I(S_1; S_2, Y, V|T_1, T_2), \qquad (86a)$$
$$I(S_2; X_2|T_1, T_2) < I(S_2; S_1, Y, V|T_1, T_2), \qquad (86b)$$
$$I(S_1, S_2; X_1, X_2|T_1, T_2) < I(S_1, S_2; Y, V|T_1, T_2) \qquad (86c)$$

when these mutual informations and all subsequent mutual informations in this section are evaluated according to the joint pmf

$$P_{S_1 S_2 X_1 X_2 Y V T_1 T_2} = P_{S_1|X_1 T_1 T_2} \cdot P_{S_2|X_2 T_1 T_2} \cdot P_{X_1 X_2 Y}$$
$$\cdot P_{T_1 T_2} \cdot \Gamma_{V|S_1 S_2 X_1 X_2}. \qquad (87)$$

Further, choose $\mu > 0$ and positive rates:

$$R_i = I(S_i; X_i|T_1, T_2) + \mu, \qquad i \in \{1, 2\}, \qquad (88)$$

so that the following three conditions hold:

$$R_1 < I(S_1; S_2, Y, V|T_1, T_2), \qquad (89a)$$
$$R_2 < I(S_2; S_1, Y, V|T_1, T_2), \qquad (89b)$$
$$R_1 + R_2 < I(S_1, S_2; Y, V|T_1, T_2) + I(S_1; S_2|T_1, T_2). \qquad (89c)$$

Code Construction: Generate a pair of sequences $T_1^n = (T_{1,1}, \ldots, T_{1,n})$ and $T_2^n = (T_{2,1}, \ldots, T_{2,n})$ by independently
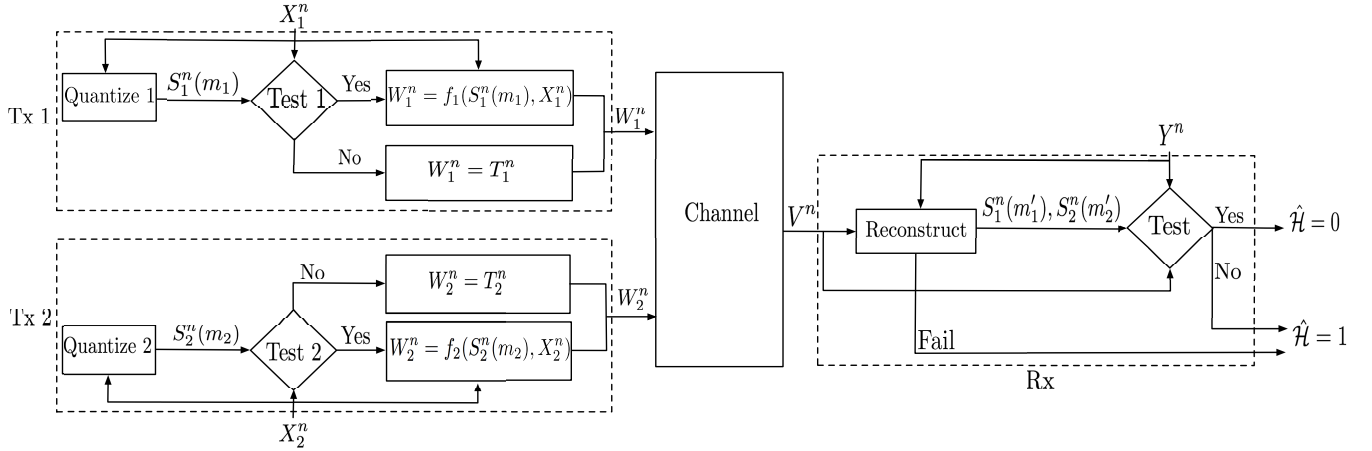
Fig. 5. Coding and testing scheme for hypothesis testing over a MAC.

drawing each pair $(T_{1,k}, T_{2,k})$ according to $P_{T_1 T_2}$. For $i \in \{1,2\}$, construct a random codebook

$$\mathcal{C}_{S_i} = \left\{ S_i^n(m_i) \colon m_i \in \{1, ..., \lfloor 2^{nR_i} \rfloor\} \right\}, \quad (90)$$

superpositioned on (T_1^n, T_2^n) by independently drawing the k-th component of each codeword according to the conditional law $P_{S_i|T_1 T_2}(\cdot|T_1, T_2)$ when $T_{1,k} = t_1$, and $T_{2,k} = t_2$. Reveal the realizations of the codebooks and the realizations (t_1^n, t_2^n) of (T_1^n, T_2^n) to all terminals.

Transmitter $i \in \{1,2\}$: Given source sequence $X_i^n = x_i^n$, Transmitter i looks for an index m_i that satisfies

$$(s_i^n(m_i), x_i^n, t_1^n, t_2^n) \in \mathcal{T}_{\mu/2}^n(P_{S_i X_i T_1 T_2}). \quad (91)$$

If successful, it picks one of these indices uniformly at random and sends the sequence w_i^n over the channel, where

$$w_{i,k} = f_i(s_{i,k}(m_i), x_{i,k}), \qquad k \in \{1, \ldots, n\},$$

and where $s_{i,k}(m_i)$ denotes the k-th component of codeword $s_i^n(m_i)$. Otherwise, Transmitter i sends t_i^n over the channel.

Receiver: Assume that the receiver observes the sequences $V^n = v^n$ and $Y^n = y^n$. It first searches for a pair of indices (m_1', m_2') that satisfies the condition:

$$H_{\mathrm{tp}(s_1^n(m_1'), s_2^n(m_2'), y^n, v^n, t_1^n, t_2^n)}(S_1, S_2|Y, V, T_1, T_2)$$
$$= \min_{\tilde{m}_1, \tilde{m}_2} H_{\mathrm{tp}(s_1^n(\tilde{m}_1), s_2^n(\tilde{m}_2), y^n, v^n, t_1^n, t_2^n)}(S_1, S_2|Y, V, T_1, T_2). \quad (92)$$

It picks one such pair at random and checks whether the chosen pair (m_1', m_2') satisfies

$$(s_1^n(m_1'), s_2^n(m_2'), y^n, v^n, t_1^n, t_2^n) \in \mathcal{T}_\mu^n(P_{S_1 S_2 Y V T_1 T_2}). \quad (93)$$

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

C. Results on the Error Exponent

The coding and testing scheme described in the previous section yields Theorem 4 ahead. For given (conditional) pmfs $P_{T_1 T_2}$, $P_{S_1|X_1 T_1 T_2}$, and $P_{S_2|X_2 T_1 T_2}$, and functions f_1 and f_2 as in (84), let the conditional and joint pmfs $\Gamma_{V|S_1 S_2 X_1 X_2}$ and $P_{S_1 S_2 X_1 X_2 Y V T_1 T_2}$ be as in (85) and (87). Define also for all $s_1 \in \mathcal{S}_1$, $s_2 \in \mathcal{S}_2$, $t_1 \in \mathcal{T}_1$, $t_2 \in \mathcal{T}_2$, $x_1 \in \mathcal{X}_1$, $x_2 \in \mathcal{X}_2$, and

$v \in \mathcal{V}$:

$$\Gamma_{V|T_1 S_2 X_2}^{(1)}(v|t_1, s_2, x_2) := \Gamma_{V|W_1 W_2}(v|t_1, f_2(s_2, x_2)), \quad (94)$$

$$\Gamma_{V|S_1 X_1 T_2}^{(2)}(v|s_1, x_1, t_2) := \Gamma_{V|W_1 W_2}(v|f_1(s_1, x_1), t_2), \quad (95)$$

$$\Gamma_{V|T_1 T_2}^{(12)}(v|T_1, T_2) := \Gamma_{V|W_1 W_2}(v|t_1, t_2), \quad (96)$$

and the following nine exponents:

$$\theta^{\mathrm{standard}} := \min_{\substack{\tilde{P}_{S_1 S_2 X_1 X_2 Y V T_1 T_2}: \\ \tilde{P}_{S_i X_i T_1 T_2} = P_{S_i X_i T_1 T_2},\, i \in \{1,2\}, \\ \tilde{P}_{S_1 S_2 Y V T_1 T_2} = P_{S_1 S_2 Y V T_1 T_2}}} D\Big(\tilde{P}_{S_1 S_2 X_1 X_2 Y V T_1 T_2} \big\| P_{S_1|X_1 T_1 T_2} P_{S_2|X_2 T_1 T_2}$$
$$\cdot Q_{X_1 X_2 Y} P_{T_1 T_2} \Gamma_{V|S_1 S_2 X_1 X_2} \Big), \quad (97)$$

$$\theta^{\mathrm{dec},1} := \min_{\substack{\tilde{P}_{S_1 S_2 X_1 X_2 Y V T_1 T_2}: \\ \tilde{P}_{S_i X_i T_1 T_2} = P_{S_i X_i T_1 T_2},\, i \in \{1,2\}, \\ \tilde{P}_{S_2 Y V T_1 T_2} = P_{S_2 Y V T_1 T_2} \\ H(S_1|S_2, Y, V, T_1, T_2) \le H_{\tilde{P}}(S_1|S_2, Y, V, T_1, T_2)}} D\Big(\tilde{P}_{S_1 S_2 X_1 X_2\ Y T_1 T_2 V} \big\| P_{S_1|X_1 T_1 T_2} P_{S_2|X_2 T_1 T_2}$$
$$\cdot Q_{X_1 X_2 Y} P_{T_1 T_2} \Gamma_{V|S_1 S_2 X_1 X_2} \Big)$$
$$+ I(S_1; Y, V|S_2, T_1, T_2) - I(S_1; X_1|S_2, T_1, T_2), \quad (98)$$

$$\theta^{\mathrm{dec},2} := \min_{\substack{\tilde{P}_{S_1 S_2 X_1 X_2 Y V T_1 T_2}: \\ \tilde{P}_{S_i X_i T_1 T_2} = P_{S_i X_i T_1 T_2},\, i \in \{1,2\}, \\ \tilde{P}_{S_1 Y V T_1 T_2} = P_{S_1 Y V T_1 T_2} \\ H(S_2|S_1, Y, V, T_1, T_2) \le H_{\tilde{P}}(S_2|S_1, Y, V, T_1, T_2)}} D\Big(\tilde{P}_{S_1 S_2 X_1 X_2\ Y V T_1 T_2} \big\| P_{S_1|X_1 T_1 T_2} P_{S_2|X_2 T_1 T_2}$$
$$\cdot Q_{X_1 X_2 Y} P_{T_1 T_2} \Gamma_{V|S_1 S_2 X_1 X_2} \Big)$$
$$+ I(S_2; Y, V|S_1, T_1, T_2) - I(S_2; X_2|S_1, T_1, T_2), \quad (99)$$

$$\theta^{\mathrm{dec},12} := \min_{\substack{\tilde{P}_{S_1 S_2 X_1 X_2 Y V T_1 T_2}: \\ \tilde{P}_{S_i X_i T_1 T_2} = P_{S_i X_i T_1 T_2},\, i \in \{1,2\}, \\ \tilde{P}_{Y V T_1 T_2} = P_{Y V T_1 T_2} \\ H(S_1, S_2|Y, V, T_1, T_2) \le H_{\tilde{P}}(S_1, S_2|Y, V, T_1, T_2)}} D\Big(\tilde{P}_{S_1 S_2 X_1 X_2\ Y V T_1 T_2} \big\| P_{S_1|X_1 T_1 T_2} P_{S_2|X_2 T_1 T_2} \quad (100)$$
$$\cdot Q_{X_1 X_2 Y} P_{T_1 T_2} \Gamma_{V|S_1 S_2 X_1 X_2} \Big)$$
$$+ I(S_1, S_2; Y, V|T_1, T_2) - I(S_1, S_2; X_1, X_2|T_1, T_2), \quad (101)$$

$$\theta^{\mathrm{miss},1a} := \min_{\substack{\tilde{P}_{S_2 X_2 Y V T_1 T_2}: \\ \tilde{P}_{S_2 X_2 T_1 T_2} = P_{S_2 X_2 T_1 T_2} \\ \tilde{P}_{Y V T_1 T_2} = P_{Y V T_1 T_2} \\ H(S_2|Y,V,T_1,T_2) \le H_{\tilde{P}}(S_2|Y,V,T_1,T_2)}}$$

$$D\Big(\tilde{P}_{S_2 X_2 Y V T_1 T_2} \big\| P_{S_2|X_2 T_1 T_2} Q_{X_2 Y} P_{T_1 T_2} \Gamma^{(1)}_{V|T_1 S_2 X_2} \Big)$$
$$+ I(S_1,S_2;V,Y|T_1,T_2) - I(S_1,S_2;X_1,X_2|T_1,T_2), \quad (102)$$

$$\theta^{\mathrm{miss},1b} := \min_{\substack{\tilde{P}_{S_1 S_2 X_2 Y V T_1 T_2}: \\ \tilde{P}_{S_2 X_2 T_1 T_2} = P_{S_2 X_2 T_1 T_2} \\ \tilde{P}_{S_2 Y V T_1 T_2} = P_{S_2 Y V T_1 T_2}}}$$

$$D\Big(\tilde{P}_{S_2 X_2 Y V T_1 T_2} \big\| P_{S_2|X_2 T_1 T_2} Q_{X_2 Y} P_{T_1 T_2} \Gamma^{(1)}_{V|T_1 S_2 X_2} \Big)$$
$$+ I(S_1;V,Y|S_2,T_1,T_2) - I(S_1;X_1|S_2,T_1,T_2), \quad (103)$$

$$\theta^{\mathrm{miss},2a} := \min_{\substack{\tilde{P}_{S_1 X_1 Y V T_1 T_2}: \\ \tilde{P}_{S_1 X_1 T_1 T_2} = P_{S_1 X_1 T_1 T_2} \\ \tilde{P}_{Y V T_1 T_2} = P_{Y V T_1 T_2} \\ H(S_1|Y,V,T_1,T_2) \le H_{\tilde{P}}(S_1|Y,V,T_1,T_2)}}$$

$$D\Big(\tilde{P}_{S_1 X_1 Y V T_1 T_2} \big\| P_{S_1|X_1 T_1 T_2} Q_{X_1 Y} P_{T_1 T_2} \Gamma^{(2)}_{V|S_1 X_1 T_2} \Big)$$
$$+ I(S_1,S_2;V,Y|T_1,T_2) - I(S_1,S_2;X_1,X_2|T_1,T_2), \quad (104)$$

$$\theta^{\mathrm{miss},2b} := \min_{\substack{\tilde{P}_{S_1 X_1 Y V T_1 T_2}: \\ \tilde{P}_{S_1 X_1 T_1 T_2} = P_{S_1 X_1 T_1 T_2} \\ \tilde{P}_{S_1 Y V T_1 T_2} = P_{S_1 Y V T_1 T_2}}}$$

$$D\Big(\tilde{P}_{S_1 X_1 Y V T_1 T_2} \big\| P_{S_1|X_1 T_1 T_2} Q_{X_1 Y} P_{T_1 T_2} \Gamma^{(2)}_{V|S_1 X_1 T_2} \Big)$$
$$+ I(S_2;V,Y|S_1,T_1,T_2) - I(S_2;X_2|S_1,T_1,T_2), \quad (105)$$

$$\theta^{\mathrm{miss},12} := \mathbb{E}_{P_{T_1 T_2}} \big[D(P_{Y V|T_1 T_2} \| Q_Y \Gamma^{(12)}_{V|T_1 T_2}) \big]$$
$$+ I(S_1,S_2;Y,V|T_1,T_2) - I(S_1,S_2;X_1,X_2|T_1,T_2),$$
$$(106)$$

where the conditional pmf $P_{Y V T_1 T_2}$, mutual informations, and—unless otherwise stated—conditional entropies are calculated according to the joint pmf $P_{S_1 S_2 X_1 X_2 Y V T_1 T_2}$ in (87).

Theorem 4: Error exponent $\theta \ge 0$ is achievable, if it satisfies
$$\theta \le \max \ \min\{\theta^{\mathrm{standard}}, \theta^{\mathrm{dec},1}, \theta^{\mathrm{dec},2}, \theta^{\mathrm{dec},12}, \theta^{\mathrm{miss},1a}, \theta^{\mathrm{miss},1b},$$
$$\theta^{\mathrm{miss},2a}, \theta^{\mathrm{miss},2b}, \theta^{\mathrm{miss},12}\},$$
$$(107)$$

where the maximization is over all (conditional) pmfs $P_{T_1 T_2}$, $P_{S_1|X_1 T_1 T_2}$, and $P_{S_2|X_2 T_1 T_2}$, and functions f_1 and f_2 as in (84) so that the conditions in (86) are satisfied with strict inequalities "<" replaced by non-strict inequalities "$\le$".

Proof: See Appendix D. ∎

Notice that the solution to the minimization problem in (102) is smaller than the solution to the minimization problem in (103). (In fact, the constraints are less stringent since $\tilde{P}_{S_2 Y V T_1 T_2} = P_{S_2 Y V T_1 T_2}$ implies $\tilde{P}_{Y V T_1 T_2} = P_{Y V T_1 T_2}$ and $H(S_2|Y,V,T_1,T_2) \le H_{\tilde{P}}(S_2|V,T_1,T_2)$.) In the same way, the solution to the minimization problem in (103) is smaller than the solution to the minimization in (105). However, since the difference of mutual informations in (102) is larger than the one in (103), and the one in (104) is larger than the one

in (105), it is à priori not clear which of these exponents is smallest.

A similar reasoning shows that the solution to the minimization problem in (101) is smaller than the solutions to the minimization problems in (97), (99), and (98), but the difference of mutual informations is larger. It is thus again unclear which of these exponents is smallest.

Remark 3: Theorem 4 recovers Theorem 1 as a special case when one of the sources is degenerate, e.g., X_2 is a constant, or the MAC degrades to a DMC, e.g., because $\Gamma_{V|W_1 W_2}$ does not depend on W_2. When X_2 is a constant or $\Gamma_{V|W_1 W_2}$ does not depend on W_2, there is no loss in optimality by choosing S_2 and T_2 constants. In this case, the exponent $\theta^{\mathrm{dec},2}$ and $\theta^{\mathrm{miss},2b}$ coincide both with the exponent $\theta^{\mathrm{standard}}$ and exponents $\theta^{\mathrm{dec},12}$ and $\theta^{\mathrm{miss},2a}$ coincide both with $\theta^{\mathrm{dec},1}$. Moreover, $\theta^{\mathrm{miss},1a}$ coincides with $\theta^{\mathrm{miss},1b}$ and $\theta^{\mathrm{miss},12}$. Choosing $S_1 = (\bar{S}_1, W_1)$ and W_1 independent of the pair $(X_1, \bar{S}_1)$ (i.e., choosing independent source and channel codebooks) finally specializes Theorem 4 to Theorem 1 under the described assumptions.

Remark 4: The error exponents in the preceding theorem are obtained by means of the hybrid coding scheme described in the previous subsection III-B. Choosing the auxiliary random variables S_1 and S_2 constant and $W_1 = f_1(X_1)$ and $W_2 = f_2(X_2)$, is equivalent to replacing hybrid coding with uncoded transmission. Choosing instead the auxiliary random variables $S_1 = (W_1, \bar{S}_1)$ and $S_2 = (W_2, \bar{S}_2)$ and the tuple (T_1, T_2, W_1, W_2) independent of the tuple $(\bar{S}_1, \bar{S}_2, X_1, X_2)$, is equivalent to replacing the single hybrid source-channel codebook by two independent source and channel codebooks. Specifically, $(\bar{S}_1, \bar{S}_2)$ then correspond to the source coding random variables and (T_1, T_2, W_1, W_2) to the channel coding random variables. Notice that even with independent codebooks, the proposed scheme performs joint source-channel coding because the source and channel codewords are decoded jointly. With separate decoding, the conditional entropy conditions in the minimizations in (99)–(101) are lost.

In this paper we restrict to hybrid coding. Similarly to standard joint source-channel coding, in some cases better schemes can be designed by using structured codes [23]–[25]. Moreover, as explained at the end of Section II-E at hand of the DMC, the achievable error exponent in Theorem 4 can be improved if throughout the conditional laws $\Gamma^{(1)}_{V|T_1 S_2 X_2}(v|t_1, s_2, x_2)$, $\Gamma^{(2)}_{V|S_1 X_1 T_2}(v|s_1, x_1, t_2)$, and $\Gamma^{(12)}_{V|T_1 T_2}(v|T_1, T_2)$ are replaced, respectively, by

$$\Gamma^{(1)}_{V|T_1 X_1 S_2 X_2}(v|t_1, x_1, s_2, x_2)$$
$$:= \Gamma_{V|W_1 W_2}(v|\eta_1(t_1, x_1), f_2(s_2, x_2)) \quad (108)$$

$$\Gamma^{(2)}_{V|S_1 X_1 T_2 X_2}(v|s_1, x_1, t_2, x_2)$$
$$:= \Gamma_{V|W_1 W_2}(v|f_1(s_1, x_1), \eta_2(t_2, x_2)) \quad (109)$$

$$\Gamma^{(12)}_{V|T_1 X_1 T_2 X_2}(v|t_1, x_1, t_2, x_2)$$
$$:= \Gamma_{V|W_1 W_2}(v|\eta_1(t_1, x_1), \eta_2(t_2, x_2)), \quad (110)$$

for some (randomized) functions $\eta_1(\cdot, \cdot)$ and $\eta_2(\cdot, \cdot)$.

To attain the desired performance, the scheme proposed in Section III-B needs to be modified so that in case of an UEP

event, Transmitter $i \in \{1, 2\}$ does not send $w_i^n = t_i^n$ but $w_{i,k} = \eta_i(t_{i,k}, x_{i,k})$, for $k \in \{1, \ldots, n\}$.

The introduced UEP mechanism with the special input sequences t_1^n and t_2^n and hybrid coding are in general important features of our scheme, without which the achieved exponent degrades. As we will see in the following, and similar to the DMC scenario, this is not the case for various special cases. We start by stating an achievable set of exponents for generalized testing against conditional independence by simplifying Theorem 4 for this setup. The simplified exponent requires hybrid coding but does not depend on the chosen UEP sequences t_1^n and t_2^n. In fact, it can even be obtained if in the scheme proposed in Subsection III-B the UEP sequences t_1^n and t_2^n are replaced with standard hybrid coding inputs.

Corollary 2: In the special case where $Y = (\bar{Y}, Z)$ and under the alternative hypothesis $\mathcal{H} = 1$:

$$Q_{X_1 X_2 \bar{Y} Z} = P_{X_1 X_2 Z} \cdot Q_{\bar{Y}|Z}, \qquad (111)$$

any error exponent $\theta \geq 0$ that satisfies

$$\theta \leq \max \left(\mathbb{E}_{P_{ZV}}\left[D(P_{\bar{Y}|ZV}\|Q_{\bar{Y}|Z})\right] + I(S_1, S_2; \bar{Y}|Z, V) \right) \tag{112}$$

is achievable, where the maximization is over all (conditional) pmfs $P_{S_1|X_1}$, and $P_{S_2|X_2}$, and functions f_1 and f_2 as in (84) that satisfy the following conditions:

$$I(S_1; X_1|S_2, Z) \leq I(S_1; V|S_2, Z), \qquad (113a)$$
$$I(S_2; X_2|S_1, Z) \leq I(S_2; V|S_1, Z), \qquad (113b)$$
$$I(S_1, S_2; X_1, X_2|Z) \leq I(S_1, S_2; V|Z), \qquad (113c)$$

and all mutual informations and the conditional pmf $P_{\bar{Y}|ZV}$ are calculated with respect to the joint pmf

$$P_{S_1 S_2 X_1 X_2 \bar{Y} ZV} = P_{S_1|X_1} P_{S_2|X_2} P_{X_1 X_2 \bar{Y} Z} \Gamma_{V|S_1 S_2 X_1 X_2}. \tag{114}$$

Proof: See Appendix E. ∎

For testing against conditional independence, i.e.,

$$Q_{\bar{Y}|Z} = P_{\bar{Y}|Z}, \qquad (115)$$

and when communication is over noiseless links of given rates, Corollary 2 recovers as a special case the result in [5, Theorem 1].

Similarly, for testing against independence, i.e., when

$$Q_{X_1 X_2 Y} = P_{X_1 X_2} P_Y, \qquad (116)$$

and when the MAC $\Gamma_{V|W_1 W_2}$ decomposes into two orthogonal DMCs $\Gamma_{V_1|W_1}$ and $\Gamma_{V_2|W_2}$, i.e.,

$$V = (V_1, V_2), \qquad (117a)$$
$$\Gamma_{V_1 V_2|W_1 W_2}(v_1, v_2|w_1, w_2) = \Gamma_{V_1|W_1}(v_1|w_1) \cdot \Gamma_{V_2|W_2}(v_2|w_2), \tag{117b}$$

then specializing Corollary 2 to two independent source and channel codebooks recovers the achievable error exponent in [12, Theorem 6]. The proposed joint source-channel coding scheme can then be replaced by a simple separate source-channel coding scheme without UEP.

Specializing Corollary 2 to separate source-channel coding without hybrid coding and UEP (but with a time-sharing sequence) results in the following corollary.

Corollary 3: Reconsider the setup in Corollary 2. Using separate source-channel coding without UEP, any error exponent $\theta \geq 0$ is achievable that satisfies

$$\theta \leq \mathbb{E}_{P_Z}\left[D(P_{\bar{Y}|Z}\|Q_{\bar{Y}|Z})\right] + \max\ I(\bar{S}_1, \bar{S}_2; \bar{Y}|Z), \quad (118)$$

where the maximization is over all (conditional) pmfs $P_{\bar{S}_1|X_1}$, $P_{\bar{S}_2|X_2}$, P_T, $P_{W_1|T}$, and $P_{W_2|T}$ that satisfy the following conditions:

$$I(\bar{S}_1; X_1|\bar{S}_2, Z) \leq I(W_1; V|W_2, T), \qquad (119a)$$
$$I(\bar{S}_2; X_2|\bar{S}_1, Z) \leq I(W_2; V|W_1, T), \qquad (119b)$$
$$I(\bar{S}_1, \bar{S}_2; X_1, X_2|Z) \leq I(W_1, W_2; V|T), \qquad (119c)$$

and where all mutual informations are calculated with respect to the joint pmf

$$P_{\bar{S}_1 \bar{S}_2 X_1 X_2 \bar{Y} ZTW_1 W_2 V} = P_{\bar{S}_1|X_1} \cdot P_{\bar{S}_2|X_2} \cdot P_{X_1 X_2 \bar{Y} Z} \cdot P_T$$
$$\cdot P_{W_1|T} \cdot P_{W_2|T} \cdot \Gamma_{V|W_1 W_2}. \tag{120}$$

This corollary recovers, for example, the optimal error exponent in [5, Corollary 4] for the Gaussian one-helper hypothesis testing against independence problem where communication takes place over two individual noiseless links. As shown in [5, Corollary 4], in this case the exponent of Corollary 3 is optimal. The following theorem proves that the exponent in Corollary 3 is also optimal for generalized testing against independence when the sources are independent under both hypotheses.

Theorem 5: Consider generalized testing against independence with independent sources, i.e.,

$$P_{X_1 X_2 Y} = P_{X_1} \cdot P_{X_2} \cdot P_{Y|X_1 X_2} \qquad (121)$$
$$Q_{X_1 X_2 Y} = P_{X_1} \cdot P_{X_2} \cdot Q_Y, \qquad (122)$$

and assume that communication from the sensors to the decision center takes place over two orthogonal DMCs $\Gamma_{V_1|W_1}$ and $\Gamma_{V_2|W_2}$ as defined in (117). Let C_1 and C_2 denote the capacities of the two DMCs $\Gamma_{V_1|W_1}$ and $\Gamma_{V_2|W_2}$. The optimal error exponent is:

$$\theta^* = D(P_Y\|Q_Y) + \max_{\substack{P_{\bar{S}_i|X_i}, P_{W_i}, i\in\{1,2\} \\ I(\bar{S}_1; X_1|\bar{S}_2) \leq C_1 \\ I(\bar{S}_2; X_2|\bar{S}_1) \leq C_2 \\ I(\bar{S}_1, \bar{S}_2; X_1, X_2) \leq C_1 + C_2}} I(\bar{S}_1, \bar{S}_2; Y). \tag{123}$$

Proof: Achievability follows directly by specializing Corollary 3 to Z a constant and thus $\bar{Y} = Y$. The converse is proved in Appendix F. ∎

It can be shown that the above theorem remains valid for the following example with independent Gaussian sources.

Example 3 (Theorem 5 for Gaussians): Let X_1 and X_2 be independent standard Gaussians under both hypotheses. Under the null hypothesis,

$$\mathcal{H} = 0: \qquad Y = X_1 + X_2 + N_0, \qquad N_0 \sim \mathcal{N}(0, \sigma_0^2), \tag{124}$$

for an N_0 independent of (X_1, X_2) and for a given nonnegative variance $\sigma_0^2 > 0$. Under the alternative hypothesis,

$$\mathcal{H} = 1: \quad Y \sim \mathcal{N}(0, \sigma_y^2), \quad \text{independent of } (X_1, X_2), \tag{125}$$

for a given nonnegative variance $\sigma_y^2 > 0$. Further assume an orthogonal MAC as in (117b) with the two individual DMCs of capacities C_1 and C_2.

The described setup is a special case of the setup considered in Theorem 5. Appendix H shows that in this case, the optimal exponent in (123) evaluates to:

$$\theta^* = \frac{1}{2} \log_2 \left(\frac{\sigma_y^2}{2^{-2C_1} + 2^{-2C_2} + \sigma_0^2} \right)$$
$$+ \left(\frac{2 + \sigma_0^2}{2\sigma_y^2} - \frac{1}{2} \right) \cdot \log_2 e. \tag{126}$$

Theorem 5 shows that separate source-channel coding is optimal for generalized testing against conditional independence over two orthogonal channels. The following proposition extends this result to all joint source distributions $P_{X_1 X_2}$. The proposition also provides a multi-letter characterization of the optimal error exponent in this case.

Proposition 2: Consider testing against independence over an orthogonal MAC, i.e., assume that (115)–(117) hold. Then, the optimal error exponent is given by

$$\theta^* = D(P_Y \| Q_Y) + \lim_{N \to \infty} \frac{1}{N} \max I(S_1^N, S_2^N; Y^N), \tag{127}$$

where the maximization is over all $P_{S_1^N | X_1^N}$ and $P_{S_2^N | X_2^N}$ satisfying:

$$\lim_{N \to \infty} \frac{1}{N} I\left(X_1^N; S_1^N | S_2^N \right) \le C_1, \tag{128}$$

$$\lim_{N \to \infty} \frac{1}{N} I\left(X_2^N; S_2^N | S_1^N \right) \le C_2, \tag{129}$$

$$\lim_{N \to \infty} \frac{1}{N} I\left(X_1^N, X_2^N; S_1^N, S_2^N \right) \le C_1 + C_2. \tag{130}$$

Proof: The achieved error exponent is a multi-letter extension of the exponent in Theorem 5, and can thus be achieved by applying the coding and testing scheme of Theorem 5 to n-tuples of source symbols and channel uses. The converse proof follows similar arguments as in [26, Theorem 2.4]. It is detailed out in Appendix G for completeness. ∎

D. Correlated Gaussian Sources Over a Gaussian MAC

In this last subsection of Section III, we focus on testing against independence over a Gaussian MAC when the sources are jointly Gaussian (but not necessarily independent as in Example 3). Consider a symmetric Gaussian setup where under both hypotheses:

$$(X_1, X_2) \sim \mathcal{N}(0, \mathbf{K}_{X_1 X_2}) \tag{131}$$

for a positive semidefinite covariance matrix

$$\mathbf{K}_{X_1 X_2} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}, \quad 0 \le \rho \le 1. \tag{132}$$

Assume as in Example 3 that under the null hypothesis,

$$\mathcal{H} = 0: \quad Y = X_1 + X_2 + N_0, \quad N_0 \sim \mathcal{N}(0, \sigma_0^2), \tag{133}$$

for N_0 independent of (X_1, X_2) and for $\sigma_0^2 > 0$, and under the alternative hypothesis,

$$\mathcal{H} = 1: \quad Y \sim \mathcal{N}(0, \sigma_y^2), \quad \text{independent of } (X_1, X_2), \tag{134}$$

for $\sigma_y^2 > 0$.

Communication takes place over the Gaussian MAC

$$V = W_1 + W_2 + N, \tag{135}$$

where the noise N is zero-mean Gaussian of variance $\sigma^2 > 0$, independent of the inputs (W_1, W_2). Each transmitter's input sequence is subject to an average block-power constraint P.

The described setup corresponds to generalized testing against conditional independence. Since Corollary 2 can be shown to hold also for this Gaussian setup, the following achievability result holds:

Corollary 4: For the described Gaussian setup any error exponent $\theta \ge 0$ is achievable that satisfies the following condition:

$$\theta \le \max \frac{1}{2} \log_2 \left(\frac{\sigma_y^2}{\frac{2\xi^2(1+\rho)\sigma^2}{2\xi^2(\alpha-\beta)^2 \cdot (1+\rho) + \sigma^2(1+\rho+\xi^2)} + \sigma_0^2} \right)$$
$$+ \frac{1}{2} \left(\frac{\sigma_0^2 + 2 + 2\rho}{\sigma_y^2} - 1 \right) \cdot \log_2 e, \tag{136}$$

where the maximization is over all $\xi^2, \alpha^2, \beta^2, \gamma^2 \ge 0$ satisfying

$$\gamma^2 + \alpha^2 + \beta^2 \xi^2 \le P, \tag{137}$$

and

$$\frac{(1+\xi^2)^2 - \rho^2}{(1+\xi^2) \cdot \xi^2} \le \frac{\sigma^2 + 2P - \gamma^2 + 2\alpha^2\rho - \frac{(\alpha \cdot (1+\rho) + \beta \cdot \xi^2)^2}{1+\xi^2}}{\sigma^2 + \frac{2(\alpha-\beta)^2 \cdot (1+\rho)\xi^2}{1+\rho+\xi^2}}, \tag{138a}$$

$$\frac{(1+\xi^2)^2 - \rho^2}{\xi^4} \le \frac{\sigma^2 + 2P + 2\alpha^2\rho}{\sigma^2 + \frac{2(\alpha-\beta)^2 \cdot (1+\rho)\xi^2}{1+\rho+\xi^2}}. \tag{138b}$$

Proof: See Appendix I. ∎

The following theorem provides an upper bound on the optimal error exponent.

Theorem 6: For the proposed Gaussian setup, the optimal error exponent θ^* satisfies

$$\theta^* \le \frac{1}{2} \cdot \log_2 \left(\frac{\sigma_y^2}{\frac{2(1+\rho)\sigma^2}{2P(1+\rho) + \sigma^2} + \sigma_0^2} \right)$$
$$+ \frac{1}{2} \left(\frac{\sigma_0^2 + 2 + 2\rho}{\sigma_y^2} - 1 \right) \cdot \log_2 e. \tag{139}$$

Proof: See Appendix J. ∎

Figure 6 compares the presented upper and lower bounds on the optimal error exponent θ^*. They are very close for the considered setup. For comparison, the figure also shows
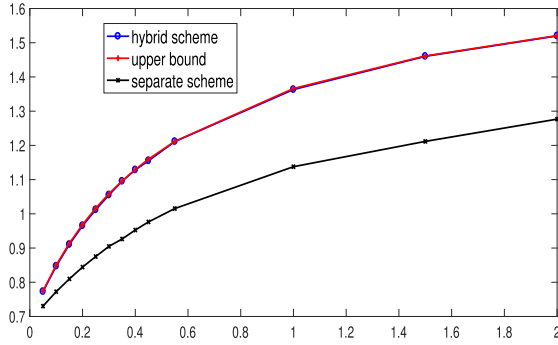
Fig. 6. Upper and lower bounds on the optimal exponent θ^* of the proposed Gaussian example as a function of P for $\rho = 0.8$, $\sigma_0^2 = 1$, $\sigma_y^2 = 1.5$ and $\sigma^2 = 1$.

the exponent that is achieved with the same choice of source variables but with separate source-channel coding. That means, by specializing the exponent in (136) to $\alpha = \beta = 0$.

IV. SUMMARY AND DISCUSSION

The paper proposes coding and testing schemes for distributed binary hypothesis testing over DMCs and MACs based on hybrid coding (in case of MAC) and UEP mechanisms that specially protect the transmission of single messages (typically the tentative guesses of the sensor nodes). These features can significantly improve the achieved error exponents. The schemes recover previous optimality results for testing against conditional independence when terminals are connected by noisefree links or DMCs, and they are shown to achieve the optimal exponents for a more general testing setup that we term *generalized testing against conditional independence*.

In this work, we have focused on the most basic communication channels: DMCs and MACs. Similar investigations can be performed for other networks. Another interesting line of research concerns the bandwidth-mismatched scenario where source and channel sequences have different blocklengths. Our coding scheme for the DMC extends in a straightforward manner to this more involved case and the corresponding error exponents are obtained by multiplying capacity and mutual informations and divergences involving channel random variables by the bandwidth mismatch factor. The proposed hybrid coding scheme for the MAC is more difficult to extend, similarly to classic joint source-channel coding [27], [28].

APPENDIX A
PROOF OF THEOREM 1

The proof of the theorem is based on the scheme in Section II-B. Fix a choice of the blocklength n, a small positive μ, and (conditional) pmfs P_T, $P_{W|T}$, and $P_{S|X}$ so that (12) holds. Assume that $I(S; X) \geq I(W; V|T)$, in which case the rates R and R' are chosen as in (16) and (17). Also, set for convenience of notation:

$$P_{S'}(s) = P_S(s), \qquad \forall s \in \mathcal{S}, \tag{140}$$

$$P_{W'|T}(w|t) = P_{W|T}(w|t), \qquad \forall t \in \mathcal{T}, \ w \in \mathcal{W}. \tag{141}$$

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

$(s, s', y) \in \mathcal{S} \times \mathcal{S} \times \mathcal{Y}$:

$$|\pi_{SY}(s, y) - P_{SY}(s, y)| \leq \mu, \tag{142}$$

$$|\pi_{S'}(s') - P_S(s)| \leq \mu, \tag{143}$$

$$H_{\pi_{S'Y}}(S'|Y) \leq H_{\pi_{SY}}(S|Y). \tag{144}$$

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

$$\mathcal{P}^*_{\text{type-I}} := \big\{ \tilde{P}_{SS'Y} \colon \tilde{P}_{SY} = P_{SY} \text{ and } \tilde{P}_{S'} = P_S \text{ and }$$
$$H_{\tilde{P}_{S'Y}}(S'|Y) \leq H_{\tilde{P}_{SY}}(S|Y) \big\}. \tag{145}$$

Consider now the type-I error probability averaged over the random code construction. Let (M, L) be the indices of the codeword chosen at the transmitter, if they exist, and define the following events:

$$\mathcal{E}_{\text{Tx}} \colon \{\nexists (m, \ell) \colon (S^n(m, \ell), X^n) \in \mathcal{T}^n_{\mu/2}(P_{SX})\} \tag{146}$$

$$\mathcal{E}^{(1)}_{\text{Rx}} \colon \{(S^n(M, L), Y^n) \notin \mathcal{T}^n_{\mu}(P_{SY})\} \tag{147}$$

$$\mathcal{E}^{(2)}_{\text{Rx}} \colon \{\exists m' \neq M \colon (T^n, W^n(m'), V^n) \in \mathcal{T}^n_{\mu}(P_{TWV})\} \tag{148}$$

$$\mathcal{E}^{(3)}_{\text{Rx}} \colon \{\exists \ell' \neq L \colon (S^n(M, \ell'), Y^n) \in \mathcal{T}^n_{\mu}(P_{SY}) \text{ and }$$
$$H_{\text{tp}(S^n(M,L),Y^n)}(S|Y) \geq H_{\text{tp}(S^n(M,\ell'),Y^n)}(S|Y)\}. \tag{149}$$

When the decoder decides on $\hat{\mathcal{H}} = 1$, then at least one of the events $\mathcal{E}_{\text{Tx}}$, $\mathcal{E}^{(1)}_{\text{Rx}}$, $\mathcal{E}^{(2)}_{\text{Rx}}$, or $\mathcal{E}^{(3)}_{\text{Rx}}$ has occurred. Therefore,

$$\mathbb{E}_{\mathcal{C}}[\alpha_n] \leq \Pr\left[\mathcal{E}_{\text{Tx}} \middle| \mathcal{H} = 0\right] + \Pr\left[\mathcal{E}^{(1)}_{\text{Rx}} \middle| \mathcal{E}^c_{\text{Tx}}, \mathcal{H} = 0\right]$$
$$+ \Pr\left[\mathcal{E}^{(2)}_{\text{Rx}} \middle| \mathcal{E}^{(1)c}_{\text{Rx}}, \mathcal{E}^c_{\text{Tx}}, \mathcal{H} = 0\right]$$
$$+ \Pr\left[\mathcal{E}^{(3)}_{\text{Rx}} \middle| \mathcal{E}^{(1)c}_{\text{Rx}}, \mathcal{E}^{(2)c}_{\text{Rx}}, \mathcal{E}^c_{\text{Tx}}, \mathcal{H} = 0\right] \tag{150}$$

$$\leq \epsilon/8 + \epsilon/8 + \epsilon/8 + \epsilon/8 \tag{151}$$

$$= \epsilon/2, \tag{152}$$

where the second inequality holds for all sufficiently large values of n. In fact the first summand of (150) can be upper bounded by means of the covering lemma [21] and using the rate constraints (16) and (17); the second by means of the Markov lemma [21] and the way the codeword $S^n(M, L)$ is picked; the third summand can be upper bounded by means of the packing lemma [21] and using the rate constraint (16); and the fourth summand can be upper bounded as described in the following. By the symmetry of the codebook construction, when bounding the probability $\Pr\left[\mathcal{E}^{(3)}_{\text{Rx}} \middle| \mathcal{E}^{(1)c}_{\text{Rx}}, \mathcal{E}^{(2)c}_{\text{Rx}}, \mathcal{E}^c_{\text{Tx}}, \mathcal{H} = 0\right]$, we can specify $M = L = 1$ and proceed as:

$$\Pr\left[\mathcal{E}^{(3)}_{\text{Rx}} \middle| \mathcal{E}^{(1)c}_{\text{Rx}}, \mathcal{E}^{(2)c}_{\text{Rx}}, \mathcal{E}^c_{\text{Tx}}, M = L = 1, \mathcal{H} = 0\right]$$

$$\overset{(a)}{\leq} \sum_{\ell'=2}^{\lfloor 2^{nR'} \rfloor} \Pr\Big[S^n(1, \ell') \in \mathcal{T}^n_{\mu}(P_S),$$

$$H_{\text{tp}(S^n(1,1),Y^n)}(S|Y) \geq H_{\text{tp}(S^n(1,\ell'),Y^n)}(S|Y) \,\Big|$$
$$(S^n(1,1), Y^n) \in \mathcal{T}^n_{\mu}(P_{SY}),$$
$$(S^n(1,1), X^n) \in \mathcal{T}^n_{\mu/2}(P_{SX}),$$
$$M = L = 1, \ \mathcal{H} = 0\Big] \tag{153}$$

$$\leq \sum_{\ell'=2}^{\lfloor 2^{nR'} \rfloor} \Pr\Big[H_{\text{tp}(S^n(1,1),Y^n)}(S|Y) \geq H_{\text{tp}(S^n(1,\ell'),Y^n)}(S|Y) \Big|$$
$$(S^n(1,1),Y^n) \in \mathcal{T}_\mu^n(P_{SY}),$$
$$(S^n(1,1),X^n) \in \mathcal{T}_{\mu/2}^n(P_{SX}),$$
$$S^n(1,\ell') \in \mathcal{T}_\mu^n(P_S), \ M = L = 1, \ \mathcal{H} = 0 \Big] \tag{154}$$

$$= \sum_{\substack{\pi_{SS'Y} \\ \in \mathcal{P}_{\mu,\text{type-I}}^n}} \sum_{\ell'=2}^{\lfloor 2^{nR'} \rfloor} \sum_{\substack{s^n,s'^n,y^n: \\ \text{tp}(s^n,s'^n,y^n) \\ =\pi_{SS'Y}}}$$
$$\Pr\Big[S^n(1,1) = s^n, S^n(1,\ell') = s'^n, Y^n = y^n \ \Big|$$
$$(S^n(1,1),Y^n) \in \mathcal{T}_\mu^n(P_{SY}),$$
$$(S^n(1,1),X^n) \in \mathcal{T}_{\mu/2}^n(P_{SX}),$$
$$S^n(1,\ell') \in \mathcal{T}_\mu^n(P_S), \ M = L = 1, \ \mathcal{H} = 0 \Big] \tag{155}$$

$$\overset{(b)}{=} \sum_{\substack{\pi_{SS'Y} \\ \in \mathcal{P}_{\mu,\text{type-I}}^n}} \sum_{\ell'=2}^{\lfloor 2^{nR'} \rfloor} \sum_{\substack{s^n,s'^n,y^n: \\ \text{tp}(s^n,s'^n,y^n) \\ =\pi_{SS'Y}}}$$
$$\Pr\Big[S^n(1,1) = s^n, Y^n = y^n \ \Big|$$
$$(S^n(1,1),Y^n) \in \mathcal{T}_\mu^n(P_{SY}),$$
$$(S^n(1,1),X^n) \in \mathcal{T}_{\mu/2}^n(P_{SX}),$$
$$S^n(1,\ell') \in \mathcal{T}_\mu^n(P_S), \ M = L = 1, \ \mathcal{H} = 0 \Big]$$
$$\cdot \Pr\Big[S^n(1,\ell') = s'^n \Big|$$
$$(S^n(1,1),Y^n) \in \mathcal{T}_\mu^n(P_{SY}),$$
$$(S^n(1,1),X^n) \in \mathcal{T}_{\mu/2}^n(P_{SX}),$$
$$S^n(1,\ell') \in \mathcal{T}_\mu^n(P_S), \ M = L = 1, \ \mathcal{H} = 0 \Big] \tag{156}$$

$$\overset{(c)}{\leq} (n+1)^{|\mathcal{S}|^2 \cdot |\mathcal{Y}|}$$
$$\sum_{\pi_{SS'Y} \in \mathcal{P}_{\mu,\text{type-I}}^n} \sum_{\ell'=2}^{\lfloor 2^{nR'} \rfloor} \sum_{\substack{s^n,s'^n,y^n: \\ \text{tp}(s^n,s'^n,y^n)=\pi_{SS'Y}}} 2^{-nH_\pi(S,Y)} \cdot 2^{-nH_\pi(S')} \tag{157}$$

$$\overset{(d)}{\leq} (n+1)^{|\mathcal{S}|^2 \cdot |\mathcal{Y}|}$$
$$\sum_{\pi_{SS'Y} \in \mathcal{P}_{\mu,\text{type-I}}^n} \sum_{\ell'=2}^{\lfloor 2^{nR'} \rfloor} 2^{nH_\pi(S,S',Y)} \cdot 2^{-nH_\pi(S,Y)} \cdot 2^{-nH_\pi(S')} \tag{158}$$

$$= (n+1)^{|\mathcal{S}|^2 \cdot |\mathcal{Y}|} \sum_{\pi_{SS'Y} \in \mathcal{P}_{\mu,\text{type-I}}^n} 2^{n(R'-I_\pi(S';Y,S))} \tag{159}$$

$$\leq (n+1)^{|\mathcal{S}|^2 \cdot |\mathcal{Y}|} \sum_{\pi_{SS'Y} \in \mathcal{P}_{\mu,\text{type-I}}^n} 2^{n(R'-I_\pi(S';Y))} \tag{160}$$

$$\overset{(e)}{\leq} (n+1)^{|\mathcal{S}|^4 \cdot |\mathcal{Y}|^2} \cdot \max_{\pi_{SS'Y} \in \mathcal{P}_{\mu,\text{type-I}}^n} 2^{n(R'-I(S;Y)+\delta_n(\mu))} \tag{161}$$

$$\overset{(f)}{\leq} \epsilon/8, \tag{162}$$

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

- (a) holds because event $\mathcal{E}_{\text{Rx}}^{(3)}$ can only hold if there exists at least one index $\ell' \neq 1$ such that both $H_{\text{tp}(S^n(1,1),Y^n)}(S|Y) \geq H_{\text{tp}(S^n(1,\ell'),Y^n)}(S|Y)$ and $(S^n(1,\ell'),Y^n) \in \mathcal{T}_\mu^n(P_{SY})$ hold, and because the latter condition implies $S^n(1,\ell') \in \mathcal{T}_\mu^n(P_S)$;

- (b) holds because conditioned on the events $(S^n(1,1),Y^n) \in \mathcal{T}_\mu^n(P_{SY})$, $(S^n(1,1),X^n) \in \mathcal{T}_{\mu/2}^n(P_{SX})$, $S^n(1,\ell') \in \mathcal{T}_\mu^n(P_S)$, $M = L = 1$ and $\mathcal{H} = 0$, the codeword $S^n(1,\ell')$ is independent of the pair $(S^n(1,1),Y^n)$;

- (c) holds because even conditioned on the events $(S^n(1,1),Y^n) \in \mathcal{T}_\mu^n(P_{SY})$, $(S^n(1,1),X^n) \in \mathcal{T}_{\mu/2}^n(P_{SX})$, $S^n(1,\ell') \in \mathcal{T}_\mu^n(P_S)$, $M = L = 1$, and $\mathcal{H} = 0$, 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} [22, Lemma 2.3] and at least $\frac{1}{(n+1)^{|\mathcal{S}|}} \cdot 2^{nH_{\pi_{S'}}(S')}$ sequences of type $\pi_{S'}$;

- (d) holds because there are at most $2^{nH_\pi(S,S',Y)}$ different n-length sequences of same joint type $\pi_{SS'Y}$;

- (e) holds because $|\mathcal{P}_{\mu,\text{type-I}}^n| \leq (n+1)^{|\mathcal{S}|^2 \cdot |\mathcal{Y}|}$, because $H_\pi(S'|Y) \leq H_\pi(S|Y)$, because each element of $\mathcal{P}_{\mu,\text{type-I}}^n$ 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

- (f) 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 $\mu \to 0$.

Now, consider the type-II error probability averaged over the random code construction. For all $m, m' \in \{1, \ldots, \lfloor 2^{nR} \rfloor\}$ and $\ell, \ell' \in \{1, \ldots, \lfloor 2^{nR'} \rfloor\}$ define events:

$$\mathcal{E}_{\text{Tx}}(m,\ell) \colon \{(S^n(m,\ell),X^n) \in \mathcal{T}_{\mu/2}^n(P_{SX}), \ W^n(m) \text{ is sent}\}, \tag{163}$$

and

$$\mathcal{E}_{\text{Rx}}(m',\ell') \colon$$
$$\{(S^n(m',\ell'),Y^n) \in \mathcal{T}_\mu^n(P_{SY}),$$
$$(T^n,W^n(m'),V^n) \in \mathcal{T}_\mu^n(P_{TWV}),$$
$$H_{\text{tp}(S^n(m',\ell'),Y^n)}(S|Y) = \min_{\tilde{m},\tilde{\ell}} H_{\text{tp}(S^n(\tilde{m},\tilde{\ell}),Y^n)}(S|Y)\}, \tag{164}$$

and notice that when averaged over all codebooks,

$$\mathbb{E}_{\mathcal{C}}[\beta_n] = \Pr[\hat{\mathcal{H}} = 0 | \mathcal{H} = 1]$$
$$= \Pr\left[\bigcup_{m',\ell'} \mathcal{E}_{\text{Rx}}(m',\ell') \Big| \mathcal{H} = 1 \right], \tag{165}$$

where the union is over all indices $(m', \ell') \in \{1, \ldots, \lfloor 2^{nR} \rfloor\} \times \{1, \ldots, \lfloor 2^{nR'} \rfloor\}$. By the union bound, above probability is upper bounded by the sum of the probabilities of the following

four events:

$$\mathcal{B}_1 \colon \{\exists \, (m, \ell) \quad \text{s.t.} \quad \mathcal{E}_{\text{Tx}}(m, \ell) \text{ and } \mathcal{E}_{\text{Rx}}(m, \ell)\}, \quad (166)$$

$$\mathcal{B}_2 \colon \{\exists \, (m, m', \ell, \ell') \text{ with } m \neq m'$$
$$\text{s.t.} \quad \mathcal{E}_{\text{Tx}}(m, \ell) \text{ and } \mathcal{E}_{\text{Rx}}(m', \ell')\}, \quad (167)$$

$$\mathcal{B}_3 \colon \{\exists \, (m, \ell, \ell') \text{ with } \ell \neq \ell'$$
$$\text{s.t.} \quad \mathcal{E}_{\text{Tx}}(m, \ell) \text{ and } \mathcal{E}_{\text{Rx}}(m, \ell')\}, \quad (168)$$

$$\mathcal{B}_4 \colon \{\forall \, (m, \ell) \quad \mathcal{E}_{\text{Tx}}^c(m, \ell) \text{ holds}$$
$$\text{and } \exists \, (m', \ell') \quad \text{s.t.} \quad \mathcal{E}_{\text{Rx}}(m', \ell')\}, \quad (169)$$

i.e.,

$$\mathbb{E}_{\mathcal{C}}[\beta_n] \leq \sum_{i=1}^{4} \Pr\left[\mathcal{B}_i | \mathcal{H} = 1\right]. \quad (170)$$

We will bound the four probabilities on the RHS of (170) individually. To simplify notation, we introduce the following sets of types

$$\mathcal{P}_{\mu, \text{standard}} = \{\pi_{SXY} \colon |\pi_{SX} - P_{SX}| < \mu/2,$$
$$|\pi_{SY} - P_{SY}| < \mu\}, \quad (171)$$

$$\mathcal{P}_{\mu, \text{decoding}} = \{\pi_{SS'XY} \colon |\pi_{SX} - P_{SX}| < \mu/2,$$
$$|\pi_{S'Y} - P_{SY}| < \mu, \ H_\pi(S'|Y) \leq H_\pi(S|Y)\}. \quad (172)$$

Consider the probability of the first event $\mathcal{B}_1$:

$$\Pr\left[\mathcal{B}_1 | \mathcal{H} = 1\right]$$
$$\leq \sum_{m, \ell} \Pr\left[S^n(m, \ell), X^n\right) \in \mathcal{T}_{\mu/2}^n(P_{SX}),$$
$$(S^n(m, \ell), Y^n) \in \mathcal{T}_\mu^n(P_{SY}),$$
$$(T^n, W^n(m), V^n) \in \mathcal{T}_\mu^n(P_{TWV}) \, \Big| \, \mathcal{H} = 1\right] \quad (173)$$

$$\leq \sum_{m, \ell} \Pr\left[(S^n(m, \ell), X^n) \in \mathcal{T}_{\mu/2}^n(P_{SX}),$$
$$(S^n(m, \ell), Y^n) \in \mathcal{T}_\mu^n(P_{SY}) \, \Big| \, \mathcal{H} = 1\right] \quad (174)$$

$$\overset{(g)}{\leq} 2^{n(R + R')} \cdot (n+1)^{|\mathcal{S}||\mathcal{X}||\mathcal{Y}|}$$
$$\cdot \max_{\pi_{SXY} \in \mathcal{P}_{\mu, \text{standard}}} 2^{-nD(\pi_{SXY} \| P_S Q_{XY})}, \quad (175)$$

where inequality (g) holds by Sanov's theorem and by the way the source sequences, the codewords, and the channel outputs are generated. Define now

$$\tilde{\theta}_\mu^{\text{standard}} := \min_{\pi \in \mathcal{P}_{\mu, \text{standard}}} D(\pi_{SXY} \| P_S Q_{XY}) - R - R' - \delta'_{1,n}, \quad (176)$$

where $\delta'_{1,n} := |\mathcal{S}||\mathcal{X}||\mathcal{Y}| \frac{\log(n+1)}{n}$, and observe that by (16) and (17):

$$\tilde{\theta}_\mu^{\text{standard}} = \min_{\pi \in \mathcal{P}_{\mu, \text{standard}}} D(\pi_{SXY} \| P_S Q_{XY}) - I(S; X) - \mu - \delta'_{1,n}$$
$$= \min_{\pi \in \mathcal{P}_{\mu, \text{standard}}} D(\pi_{SXY} \| P_{S|X} Q_{XY}) - \delta_{1,n}(\mu)$$
$$= \theta_\mu^{\text{standard}} - \delta_{1,n}(\mu), \quad (177)$$

for a sequence of functions $\delta_{1,n}(\mu)$ that goes to zero as $n \to \infty$ and $\mu \to 0$, and

$$\theta_\mu^{\text{standard}} := \min_{\pi_{SXY} \in \mathcal{P}_{\mu, \text{standard}}} D(\pi_{SXY} \| P_{S|X} Q_{XY}). \quad (178)$$

Combining (175)–(177), we obtain:

$$\Pr\left[\mathcal{B}_1 | \mathcal{H} = 1\right] \leq 2^{-n\left(\theta_\mu^{\text{standard}} - \delta_{1,n}(\mu)\right)}. \quad (179)$$

Consider next the probability of event $\mathcal{B}_2$:

$$\Pr\left[\mathcal{B}_2 | \mathcal{H} = 1\right]$$
$$\leq \sum_{\substack{m, m'; \\ m \neq m'}} \sum_{\ell, \ell'} \Pr\left[(S^n(m, \ell), X^n) \in \mathcal{T}_{\mu/2}^n(P_{SX}),\right.$$
$$W^n(m) \text{ is sent}, \ (S^n(m', \ell'), Y^n) \in \mathcal{T}_\mu^n(P_{SY}),$$
$$H_{\text{tp}(S^n(m', \ell'), Y^n)}(S|Y) = \min_{\tilde{m}, \tilde{\ell}} H_{\text{tp}(S^n(\tilde{m}, \tilde{\ell}), Y^n)}(S|Y),$$
$$\left. (T^n, W^n(m'), V^n) \in \mathcal{T}_\mu^n(P_{TWV}) \Big| \mathcal{H} = 1\right] \quad (180)$$

$$= \sum_{\substack{m, m'; \\ m \neq m'}} \sum_{\ell, \ell'} \Pr\left[(S^n(m, \ell), X^n) \in \mathcal{T}_{\mu/2}^n(P_{SX}),\right.$$
$$(S^n(m', \ell'), Y^n) \in \mathcal{T}_\mu^n(P_{SY}),$$
$$H_{\text{tp}(S^n(m', \ell'), Y^n)}(S|Y) = \min_{\tilde{m}, \tilde{\ell}} H_{\text{tp}(S^n(\tilde{m}, \tilde{\ell}), Y^n)}(S|Y)$$
$$\left. \Big| \mathcal{H} = 1\right]$$
$$\cdot \Pr\left[W^n(m) \text{ is sent},\right.$$
$$\left. (T^n, W^n(m'), V^n) \in \mathcal{T}_\mu^n(P_{TWV}) \ | \ \mathcal{H} = 1\right] \quad (181)$$

$$\leq \sum_{\substack{m, m'; \\ m \neq m'}} \sum_{\ell, \ell'} \Pr\left[(S^n(m, \ell), X^n) \in \mathcal{T}_{\mu/2}^n(P_{SX}),\right.$$
$$(S^n(m', \ell'), Y^n) \in \mathcal{T}_\mu^n(P_{SY}),$$
$$H_{\text{tp}(S^n(m', \ell'), Y^n)}(S|Y) \leq H_{\text{tp}(S^n(m, \ell), Y^n)}(S|Y)$$
$$\left. \Big| \mathcal{H} = 1\right]$$
$$\cdot \Pr\left[(T^n, W^n(m'), V^n) \in \mathcal{T}_\mu^n(P_{TWV})\right.$$
$$\left. \big| \ W^n(m) \text{ is sent}, \mathcal{H} = 1\right] \quad (182)$$

$$= \sum_{\substack{m, m'; \\ m \neq m'}} \sum_{\ell, \ell'} \Pr\left[\left|\text{tp}(S^n(m, \ell), X^n) - P_{SX}\right| < \mu/2,\right.$$
$$\left|\text{tp}(S^n(m', \ell'), Y^n) - P_{SY}\right| < \mu, \quad (183)$$
$$H_{\text{tp}(S^n(m', \ell'), Y^n)}(S|Y) \leq H_{\text{tp}(S^n(m, \ell), Y^n)}(S|Y)$$
$$\left. \Big| \mathcal{H} = 1\right] \quad (184)$$
$$\cdot \Pr\left[(T^n, W^n(m'), V^n) \in \mathcal{T}_\mu^n(P_{TWV})\right.$$
$$\left. \big| \ W^n(m) \text{ is sent}, \mathcal{H} = 1\right] \quad (185)$$

$$= \sum_{\substack{m, m'; \\ m \neq m'}} \sum_{\ell, \ell'} \sum_{\pi \in \mathcal{P}_{\mu, \text{decoding}}}$$
$$\Pr\left[\text{tp}(S^n(m, \ell), S^n(m', \ell'), X^n, Y^n) = \pi_{SS'XY}\right.$$
$$\left. \Big| \mathcal{H} = 1\right] \quad (186)$$

$$\cdot \Pr\Big[(T^n, W^n(m'), V^n) \in \mathcal{T}^n_\mu(P_{TWV})$$

$$\Big| W^n(m) \text{ is sent}, \mathcal{H} = 1\Big]$$

$$\leq 2^{n(2R+2R')} \max_{\pi \in \mathcal{P}_{\mu,\text{decoding}}} 2^{-n\big(D(\pi_{SS'XY}\|P_S P_{S'} Q_{XY}) - \delta'_{2,n}\big)}$$

$$\cdot \max_{\substack{\pi_{TW'V}: \\ |\pi_{TW'V} - P_{TWV}| \leq \mu}} 2^{-n\big(D(\pi_{TW'V}\|P_{TV} P_{W'|T}) - \delta''_{2,n}\big)},$$
$$(187)$$

where the last inequality holds by Sanov's theorem and by the way the codebooks and the channel outputs are generated and $\delta'_{2,n}, \delta''_{2,n}$ are sequences that tend to 0 as $n \to \infty$. Define

$$\tilde{\theta}^{\text{dec}}_\mu := \min_{\pi \in \mathcal{P}_{\mu,\text{decoding}}} D(\pi_{SS'XY}\|P_S P_{S'} Q_{XY})$$

$$+ \min_{\substack{\pi_{TW'V}: \\ |\pi_{TW'V} - P_{TWV}| \leq \mu}} D(\pi_{TW'V}\|P_{TV} P_{W'|T})$$

$$- 2R - 2R' - \delta'_{2,n} - \delta''_{2,n}, \qquad (188)$$

and observe that:

$$\tilde{\theta}^{\text{dec}}_\mu \overset{\text{(eq.(16)\&(17))}}{=} \min_{\pi \in \mathcal{P}_{\mu,\text{decoding}}} D(\pi_{SS'XY}\|P_S P_{S'} Q_{XY})$$

$$+ \min_{\substack{\pi_{TW'V}: \\ |\pi_{TW'V} - P_{TWV}| \leq \mu}} D(\pi_{TW'V}\|P_{TV} P_{W'|T})$$

$$- 2I(S;X) - 2\mu - \delta'_{2,n} - \delta''_{2,n} \qquad (189)$$

$$\overset{\text{(CR)}}{=} \min_{\pi \in \mathcal{P}_{\mu,\text{decoding}}} \Big[D(\pi_{SXY}\|P_S Q_{XY})$$

$$+ \mathbb{E}_{\pi_{SXY}} \big[D(\pi_{S'|SXY}\|P_{S'})\big]\Big]$$

$$+ \min_{\substack{\pi_{TW'V}: \\ |\pi_{TW'V} - P_{TWV}| \leq \mu}} D(\pi_{TW'V}\|P_{TV} P_{W'|T})$$

$$- 2I(S;X) - 2\mu - \delta'_{2,n} - \delta''_{2,n} \qquad (190)$$

$$\overset{\text{(DP)}}{\geq} \min_{\pi \in \mathcal{P}_{\mu,\text{decoding}}} \Big[D(\pi_{SXY}\|P_S Q_{XY})$$

$$+ \mathbb{E}_{\pi_Y} \big[D(\pi_{S'|Y}\|P_{S'})\big]\Big]$$

$$+ \min_{\substack{\pi_{TW'V}: \\ |\pi_{TW'V} - P_{TWV}| \leq \mu}} D(\pi_{TW'V}\|P_{TV} P_{W'|T})$$

$$- 2I(S;X) - 2\mu - \delta'_{2,n} - \delta''_{2,n} \qquad (191)$$

$$\overset{(h)}{=} \min_{\pi \in \mathcal{P}_{\mu,\text{decoding}}} D(\pi_{SXY}\|P_{S|X} Q_{XY}) + I(S;Y)$$

$$+ I(W;V|T) - I(S;X) - \delta_{2,n}(\mu) \qquad (192)$$

$$= \min_{\pi \in \mathcal{P}_{\mu,\text{decoding}}} D(\pi_{SXY}\|P_{S|X} Q_{XY}) + I(S;Y)$$

$$+ I(W;V|T) - I(S;X) - \delta_{2,n}(\mu) \qquad (193)$$

$$= \theta^{\text{dec}}_\mu - \delta_{2,n}(\mu), \qquad (194)$$

for a sequence of functions $\delta_{2,n}(\mu)$ that goes to zero as $n \to \infty$ and $\mu \to 0$, and

$$\theta^{\text{dec}}_\mu := \min_{\pi_{SS'XY} \in \mathcal{P}_{\mu,\text{decoding}}} D(\pi_{SXY}\|P_{S|X} Q_{XY})$$

$$+ I(S;Y) + I(W;V|T) - I(S;X). \quad (195)$$

Here, (CR) and (DP) refer to chain rule and data processing inequality arguments, (h) follows because

$|\pi_{TW'V} - P_{TWV}| \leq \mu$ and $P_{W'|T} = P_{W|T}$ and because $|\pi_{S'Y} - P_{SY}| \leq \mu$. (Notice that the DP-inequality can be shown to hold with equality.) Combining (187), (188), and (194), we have

$$\Pr\big[\mathcal{B}_2|\mathcal{H}=1\big] \leq 2^{-n\big(\theta^{\text{dec}}_\mu - \delta_{2,n}(\mu)\big)}. \qquad (196)$$

Consider next the third event $\mathcal{B}_3$:

$$\Pr\big[\mathcal{B}_3|\mathcal{H}=1\big]$$

$$\leq \sum_m \sum_{\ell,\ell'} \Pr\Big[(S^n(m,\ell), X^n) \in \mathcal{T}^n_{\mu/2}(P_{SX}),$$

$$W^n(m) \text{ is sent}, \ (S^n(m,\ell'), Y^n) \in \mathcal{T}^n_\mu(P_{SY}),$$

$$(T^n, W^n(m), V^n) \in \mathcal{T}^n_\mu(P_{TWV}),$$

$$H_{\text{tp}(S^n(m,\ell'),Y^n)}(S'|Y) = \min_{\tilde{m},\tilde{\ell}} H_{\text{tp}(S^n(\tilde{m},\tilde{\ell}),Y^n)}(S|Y)$$

$$\Big|\mathcal{H}=1\Big] \qquad (197)$$

$$\leq \sum_m \sum_{\ell,\ell'} \Pr\Big[(S^n(m,\ell), X^n) \in \mathcal{T}^n_{\mu/2}(P_{SX}),$$

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

$$H_{\text{tp}(S^n(m,\ell'),Y^n)}(S'|Y) = \min_{\tilde{m},\tilde{\ell}} H_{\text{tp}(S^n(\tilde{m},\tilde{\ell}),Y^n)}(S|Y)$$

$$\Big|\mathcal{H}=1\Big] \qquad (198)$$

$$\leq 2^{n(R+2R')} \cdot \max_{\pi_{SS'XY} \in \mathcal{P}_{\mu,\text{decoding}}} 2^{-n\big(D(\pi_{SS'XY}\|P_S P_S Q_{XY}) - \delta'_{3,n}\big)},$$
$$(199)$$

for some sequence $\delta'_{3,n}$ that tends to 0 as $n \to \infty$. Here, the last inequality holds by Sanov's theorem and the way the codebooks and the channel outputs are generated.

Define

$$\tilde{\theta}'^{\text{dec}}_\mu := \min_{\pi_{SS'XY} \in \mathcal{P}_{\mu,\text{decoding}}} D(\pi_{SS'XY}\|P_S P_{S'} Q_{XY})$$

$$- R - 2R' - \delta'_{3,n}, \qquad (200)$$

and notice that:

$$\tilde{\theta}'^{\text{dec}}_\mu \overset{\text{(eq.(16)\&(17))}}{=} \min_{\pi_{SS'XY} \in \mathcal{P}_{\mu,\text{decoding}}} D(\pi_{SS'XY}\|P_S P_{S'} Q_{XY})$$

$$+ I(W;V|T) - 2\,I(S;X) - 3\mu - \delta'_{3,n}$$

$$\overset{\text{(CR)\&(DP)}}{\geq} \min_{\pi_{SS'XY} \in \mathcal{P}_{\mu,\text{decoding}}} \Big[D(\pi_{SXY}\|P_S Q_{XY})$$

$$+ \mathbb{E}_{\pi_Y} \big[D(\pi_{S'|Y}\|P_{S'})\big]\Big]$$

$$+ I(W;V|T) - 2\,I(S;X) - 3\mu - \delta'_{3,n}$$

$$\overset{(i)}{=} \min_{\pi_{SS'XY} \in \mathcal{P}_{\mu,\text{decoding}}} D(\pi_{SXY}\|P_{S|X} Q_{XY})$$

$$+ I(S;Y) + I(W;V|T) - I(S;X) - \delta_{3,n}(\mu)$$

$$= \theta^{\text{dec}}_\mu - \delta_{3,n}(\mu), \qquad (201)$$

where $\delta_{3,n}(\mu)$ is a sequence of functions that goes to zero as $n \to \infty$ and $\mu \to 0$. Here, (i) holds because $|\pi_{S'Y} - P_{SY}| \leq \mu$.

(Notice that the DP-inequality can again be shown to hold with equality.) By (199)–(201), we conclude

$$\Pr\left[\mathcal{B}_3|\mathcal{H}=1\right] \leq 2^{-n\left(\theta_\mu^{\text{dec}}-\delta_{3,n}(\mu)\right)}. \qquad (202)$$

Finally, consider the probability of the fourth event $\mathcal{B}_4$. By the union bound:

$$\Pr\left[\mathcal{B}_4|\mathcal{H}=1\right]$$

$$\leq \sum_{m'}\sum_{\ell'}\Pr\left[\left(\bigcap_{(m,\ell)}\mathcal{E}_{\text{Tx}}^c(m,\ell)\right) \cap \mathcal{E}_{\text{Rx}}(m',\ell')\middle|\mathcal{H}=1\right] \qquad (203)$$

$$\overset{(j)}{\leq}\sum_{m'}\sum_{\ell'}\Pr\left[(S^n(m',\ell'),Y^n)\in\mathcal{T}_\mu^n(P_{SY})\middle|\mathcal{H}=1\right]$$

$$\cdot\Pr\left[(T^n,W^n(m'),V^n)\in\mathcal{T}_\mu^n(P_{TWV})\middle|\right.$$

$$\left.\left(\bigcap_{(m,\ell)}\mathcal{E}_{\text{Tx}}^c(m,\ell)\right),\mathcal{H}=1\right] \qquad (204)$$

$$\overset{(k)}{\leq}\sum_{m'}\sum_{\ell'}\left(\sum_{\substack{\pi_{S'Y}:\\|\pi_{S'Y}-P_{SY}|<\mu}}2^{-nD(\pi_{S'Y}\|P_{S'}Q_Y)}\right)$$

$$\cdot\left(\sum_{\substack{\pi_{TW'V}:\\|\pi_{TW'V}-P_{TWV}|<\mu}}2^{-nD(\pi_{TW'V}\|P_{TW'}\Gamma_{V|W=T})}\right) \qquad (205)$$

$$\leq 2^{n(R+R')}\cdot\max_{\substack{\pi_{S'Y}:\\|\pi_{S'Y}-P_{SY}|<\mu}}2^{-n(D(\pi_{S'Y}\|P_{S'}Q_Y)-\delta'_{4,n})}$$

$$\cdot\max_{\substack{\pi_{TW'V}:\\|\pi_{TW'V}-P_{TWV}|<\mu}}2^{-n(D(\pi_{TW'V}\|P_{TW'}\Gamma_{V|W=T})-\delta''_{4,n})}, \qquad (206)$$

where $\delta'_{4,n}$ and $\delta''_{4,n}$ are sequences that tend to 0 as $n\to\infty$. Here, (j) holds because the tuple $(T^n,W^n(m'),V^n)$ is generated independently of the pair $(S^n(m',\ell'),Y^n)$ and (k) holds by Sanov's theorem and the way the codebooks and the source sequences are generated.

Define now

$$\tilde{\theta}_\mu^{\text{miss}}:=\min_{\substack{\pi_{S'Y}:\\|\pi_{S'Y}-P_{SY}|<\mu}}D(\pi_{S'Y}\|P_{S'}Q_Y)$$

$$+\min_{\substack{\pi_{TW'V}:\\|\pi_{TW'V}-P_{TWV}|<\mu}}D(\pi_{TW'V}\|P_{TW'}\Gamma_{V|W=T})$$

$$-R-R'-\delta'_{4,n}-\delta''_{4,n}, \qquad (207)$$

and notice that by (16) and (17):

$$\tilde{\theta}_\mu^{\text{miss}}=\min_{\substack{\pi_{S'Y}:\\|\pi_{S'Y}-P_{SY}|<\mu}}D(\pi_{S'Y}\|P_{S'}Q_Y)$$

$$+\min_{\substack{\pi_{TW'V}:\\|\pi_{TW'V}-P_{TWV}|<\mu}}D(\pi_{TW'V}\|P_{TW'}\Gamma_{V|W=T})$$

$$-I(S;X)-\mu-\delta'_{4,n}-\delta''_{4,n} \qquad (208)$$

$$\overset{(l)}{=}D(P_{SY}\|P_SQ_Y)+D(P_{TWV}\|P_{TW}\Gamma_{V|W=T})$$

$$-I(S;X)-\delta_{4,n}(\mu) \qquad (209)$$

$$\overset{(\text{CR})}{=}D(P_Y\|Q_Y)+D(P_{TWV}\|P_{TW}\Gamma_{V|W=T})$$

$$+I(S;Y)-I(S;X)-\delta_{4,n}(\mu) \qquad (210)$$

$$=\theta_\mu^{\text{miss}}-\delta_{4,n}(\mu), \qquad (211)$$

for some sequence of functions $\delta_{4,n}(\mu)$ that tends to 0 as $n\to\infty$ and $\mu\to 0$ and

$$\theta_\mu^{\text{miss}}:=D(P_Y\|Q_Y)+I(S;Y)+D(P_{TWV}\|P_{TW}\Gamma_{V|W=T})$$

$$-I(S;X). \qquad (212)$$

Here, step (l) holds because $|\pi_{TW'V}-P_{TWV}|\leq\mu$, and $|\pi_{S'Y}-P_{SY}|\leq\mu$. By (205)–(211), we have

$$\Pr\left[\mathcal{B}_4|\mathcal{H}=1\right]\leq 2^{-n\left(\theta_\mu^{\text{miss}}-\delta_4(\mu)\right)}. \qquad (213)$$

Combining (170) with (179), (196), (202) and (213), proves that for sufficiently large blocklengths n and sufficiently small values of μ, the average type-II error probability satisfies

$$\mathbb{E}_\mathcal{C}[\beta_n]\leq 4\cdot\max\left\{2^{-n\theta_\mu^{\text{standard}}},\ 2^{-n\theta_\mu^{\text{dec}}},\ 2^{-n\theta_\mu^{\text{miss}}}\right\}. \qquad (214)$$

By standard arguments and successively eliminating the worst half of the codewords with respect to α_n and the exponents $\theta_\mu^{\text{standard}}$, θ_μ^{dec}, and θ_μ^{miss}, it can be shown that there exists at least one codebook for which

$$\alpha_n\leq\epsilon, \qquad (215)$$

$$\beta_n\leq 64\cdot\max\left\{2^{-n\theta_\mu^{\text{standard}}},\ 2^{-n\theta_\mu^{\text{dec}}},\ 2^{-n\theta_\mu^{\text{miss}}}\right\}. \qquad (216)$$

Letting $\mu\to 0$ and $n\to\infty$, we get $\theta_\mu^{\text{standard}}\to\theta^{\text{standard}}$, $\theta_\mu^{\text{dec}}\to\theta^{\text{dec}}$, $\theta_\mu^{\text{miss}}\to\theta^{\text{miss}}$.

This proves the theorem for $I(S;X)\geq I(W;V|T)$. When $I(S;X)<I(W;V|T)$, rates R and R' are chosen as in (14) and (15). The analysis is similar to above, but since $R'=0$, event $\mathcal{B}_3$ can be omitted.

APPENDIX B
PROOF OF THEOREM 2

Let f be a function satisfying the properties in the theorem. In this case, $Q_{X|Y}=P_{X|f(Y)}$ and for the choice

$$\tilde{P}_{SXY}=P_{S|X}P_{X|f(Y)}P_Y \qquad (217)$$

the expectation in (32) evaluates to 0. The proposed choice in (217) is a valid candidate for the minimization in (32) as we show in the following. It is straightforward to see from (217) that the marginal $\tilde{P}_Y$ coincides with P_Y. One can also observe that:

$$\tilde{P}_{SX}(s,x)$$

$$=\sum_y\tilde{P}_{SXY}(s,x,y) \qquad (218)$$

$$=P_{S|X}(s|x)\cdot\sum_y P_{X|f(Y)}(x|f(y))P_Y(y) \qquad (219)$$

$$=P_{S|X}(s|x)\cdot\sum_z P_{X|f(Y)}(x|z)\left(\sum_{y:\ z=f(y)}P_Y(y)\right) \qquad (220)$$

$$=P_{S|X}(s|x)\cdot\sum_z P_{X|f(Y)}(x|z)P_{f(Y)}(z) \qquad (221)$$

$$=P_{S|X}(s|x)\cdot P_X(x) \qquad (222)$$

$$=P_{SX}(s,x). \qquad (223)$$

To show $H(S|Y) \leq H_{\tilde{P}}(S|Y)$, we introduce the enhanced type

$$\tilde{P}_{SXYY'} = \tilde{P}_{SXY} \cdot \tilde{P}_{Y'|Xf(Y)}, \qquad (224)$$

with $\tilde{P}_{SXY}$ as chosen in (217) and $\tilde{P}_{Y'|Xf(Y)} = P_{Y|Xf(Y)}$. By (217), similarly to (223), we then have

$$\tilde{P}_{SXY'}(s, x, y')$$
$$= \sum_{y} \tilde{P}_{SXY'Y}(s, x, y', y) \qquad (225)$$

$$= P_{S|X}(s|x) \sum_{z} P_{X|f(Y)}(x|z) P_{Y|X,f(Y)}(y'|x, z)$$
$$\cdot \Big(\sum_{y:\, z=f(y)} P_Y(y) \Big) \qquad (226)$$

$$= P_{S|X}(s|x) \sum_{z} P_{X|f(Y)}(x|z) P_{Y|X,f(Y)}(y'|x, z) P_{f(Y)}(z) \qquad (227)$$

$$= P_{S|X}(s|x) \sum_{z} P_{X,f(Y),Y}(x, z, y') \qquad (228)$$

$$= P_{S|X}(s|x) P_{XY}(x, y') \qquad (229)$$
$$= P_{SXY}(s, x, y'), \qquad (230)$$

and we conclude

$$H(S|Y) = H_{\tilde{P}}(S|Y') \qquad (231)$$
$$\overset{(a)}{=} H_{\tilde{P}}(S|Y', f(Y)) \qquad (232)$$
$$\overset{(b)}{=} H_{\tilde{P}}(S|Y', f(Y), Y) \qquad (233)$$
$$\leq H_{\tilde{P}}(S|Y). \qquad (234)$$

where (a) holds because under $\tilde{P}$ we have $f(Y) = f(Y')$ with probability 1 and (b) holds because under $\tilde{P}$ we have the Markov chain $S \to (Y', f(Y)) \to Y$. We can thus conclude that we are in the case treated in Remark 1.

We continue to evaluate the right-hand side of (33). Let $P_{S|X}$ satisfy the stronger condition $I(S; X|f(Y)) \leq C$. Then,

$$\theta^{\text{dec}} \geq D(P_Y\|Q_Y) + I(S; X|f(Y)) - I(S; X|Y)$$
$$= D(P_Y\|Q_Y) + I(S; Y|f(Y)), \qquad (235)$$

where the second inequality holds by the Markov chain $S \to X \to (Y, f(Y))$. Now, notice that under the conditions of the corollary:

$$\sum_{x \in \mathcal{X}} P_{S|X}(s, x) Q_{XY}(x, y)$$
$$= \sum_{x \in \mathcal{X}} \big(P_{S|X}(s|x) P_{X|f(Y)}(x|f(y)) \big) \cdot Q_Y(y) \qquad (236)$$

$$= P_{S|f(Y)}(s|f(y)) \cdot Q_Y(y) \qquad (237)$$

and thus by (34):

$$\theta^{\text{standard}} \geq D(P_Y\|Q_Y) + \min_{\substack{\pi_{SY}:\\ \pi_{SY}=P_{SY}}} D(\pi_{S|Y}\|P_{S|f(Y)}) \qquad (238)$$

$$= D(P_Y\|Q_Y) + I(S; Y|f(Y)). \qquad (239)$$

We now prove the converse direction. Defining $Z_i := f(Y_i)$ and $\delta(\epsilon) := H(\epsilon)/n/(1-\epsilon)$ and following standard

arguments [8], we obtain

$$\theta \leq \frac{1}{(1-\epsilon)n} D(P_{V^n Y^n|\mathcal{H}=0}\|P_{V^n Y^n|\mathcal{H}=1}) + \delta(\epsilon) \quad (240)$$

$$\leq \frac{1}{(1-\epsilon)n} D(P_{V^n Y^n|\mathcal{H}=0}\|P_{V^n Y^n|\mathcal{H}=1}) + \delta(\epsilon) \quad (241)$$

$$= \frac{1}{(1-\epsilon)n} \mathbb{E}_{P_{Y^n}} \big[D(P_{V^n|Y^n,\mathcal{H}=0}\|P_{V^n|Z^n,\mathcal{H}=1}) \big]$$
$$+ \frac{1}{(1-\epsilon)} \cdot D(P_Y\|Q_Y) + \delta(\epsilon) \quad (242)$$

$$\leq \frac{1}{(1-\epsilon)n} I(V^n; Y^n|Z^n)$$
$$+ \frac{1}{(1-\epsilon)} \cdot D(P_Y\|Q_Y) + \delta(\epsilon) \quad (243)$$

$$= \frac{1}{(1-\epsilon)n} \sum_{i=1}^{n} I(V^n, Y^{i-1}, Z^{i-1}, Z_{i+1}^n; Y_i|Z_i) \quad (244)$$
$$+ \frac{1}{(1-\epsilon)} \cdot D(P_Y\|Q_Y) + \delta(\epsilon) \quad (245)$$

$$\leq \frac{1}{(1-\epsilon)n} \sum_{i=1}^{n} I(V^n, X^{i-1}, Z^{i-1}, Z_{i+1}^n; Y_i|Z_i) \quad (246)$$
$$+ \frac{1}{(1-\epsilon)} \cdot D(P_Y\|Q_Y) + \delta(\epsilon) \quad (247)$$

$$\leq \frac{1}{(1-\epsilon)} \big(I(S; Y|f(Y)) + D(P_Y\|Q_Y) \big) + \delta(\epsilon), \quad (248)$$

where the last inequality follows by introducing a time-sharing random variable Q that is uniform over $\{1, \ldots, n\}$ and by defining $S := (Q, V^n, X^{Q-1}, Z^{Q-1}, Z_{Q+1}^n)$ and $Y = Y_Q$.

We turn to the constraint on capacity:

$$C \geq I(W^n; V^n) \qquad (249)$$
$$\geq I(W^n; V^n|Z^n) \qquad (250)$$
$$\geq I(X^n; V^n|Z^n) \qquad (251)$$
$$\geq \sum_{i=1}^{n} I(X_i; V^n, X^{i-1}, Z^{i-1}, Z_{i+1}^n|Z_i) \qquad (252)$$
$$\geq I(X; S|f(Y)), \qquad (253)$$

where for the last inequality we defined $X = X_Q$. The proof is established by noticing the Markov chain

$$S - X - Y. \qquad (254)$$

APPENDIX C
PROOF OF EXAMPLE 1

Notice that X and Y are independent under $\mathcal{H} = 1$. Moreover, Y (and X) has same marginal under both hypotheses. Therefore, when applying Theorem 2, the term $D(P_Y\|Q_Y) = 0$ and the function f can be ignored.

Let now $S = X + G$ with G a zero-mean Gaussian random variable of variance ξ^2 and independent of X. For this choice:

$$I(S; Y) = \frac{1}{2} \log_2 \left(\frac{1}{1 - \frac{\rho_0^2}{1+\xi^2}} \right), \qquad (255)$$

and

$$I(S; X) = \frac{1}{2} \log_2 \left(\frac{1+\xi^2}{\xi^2} \right). \qquad (256)$$

Thus, by Theorem 2, the optimal exponent for the presented Gaussian setup is lower bounded as:

$$\theta^* \geq \max_{\xi^2: \frac{1}{2}\log_2\left(\frac{1+\xi^2}{\xi^2}\right) \leq C} \frac{1}{2}\log_2\left(\frac{1}{1 - \frac{\rho_0^2}{1+\xi^2}}\right) \tag{257}$$

$$= \frac{1}{2}\log_2\left(\frac{1}{1 - \rho_0^2 + \rho_0^2 \cdot 2^{-2C}}\right). \tag{258}$$

We now show that θ^* is also upper bounded by the right-hand side of (258). To this end, notice first that:

$$I(S; X) = \frac{1}{2}\log_2(2\pi e) - h(X|S), \tag{259}$$

and thus constraint $C \geq I(S; X)$ is equivalent to:

$$2^{2h(X|S)} \geq (2\pi e) \cdot 2^{-2C}. \tag{260}$$

Moreover, (under $\mathcal{H} = 0$) one can write $Y = \rho_0 X + F$, with F zero-mean Gaussian of variance $1 - \rho_0^2$ and independent of X. This implies that for any S forming the Markov chain $S-X-Y$, also the pair (S, X) is independent of F. By the EPI and because $h(\rho_0 X) = \log_2|\rho_0| + h(X)$, we then have:

$$h(Y|S) \geq \frac{1}{2}\log_2\left(2\pi e\left(\frac{1}{2\pi e}2^{2h(\rho_0 X|S)} + (1 - \rho_0^2)\right)\right) \tag{261}$$

$$= \frac{1}{2}\log_2\left(2\pi e\left(\frac{\rho_0^2}{2\pi e}2^{2h(X|S)} + (1 - \rho_0^2)\right)\right). \tag{262}$$

By Theorem 2, the optimal error exponent is upper bounded as:

$$\theta^* = \max_{\substack{S: \\ \text{s.t. (260)}}} I(S; Y) \tag{263}$$

$$= h(Y) - \min_{\substack{S: \\ \text{(260) holds}}} h(Y|S) \tag{264}$$

$$\overset{(a)}{\leq} \frac{1}{2}\log_2(2\pi e)$$
$$- \min_{\substack{S: \\ \text{(260) holds}}} \frac{1}{2}\log_2\left(2\pi e\left(\frac{\rho_0^2}{2\pi e}2^{2h(X|S)} + (1 - \rho_0^2)\right)\right) \tag{265}$$

$$\overset{(b)}{\leq} \frac{1}{2}\log_2(2\pi e) - \frac{1}{2}\log_2\left(2\pi e\left(\rho_0^2 \cdot 2^{-2C} + (1 - \rho_0^2)\right)\right) \tag{266}$$

$$= \frac{1}{2}\log_2\left(\frac{1}{1 - \rho_0^2 + \rho_0^2 \cdot 2^{-2C}}\right), \tag{267}$$

where (a) holds by (262) and (b) by (260). Combining this upper bound with the lower bound in (258), completes the proof.

APPENDIX D
PROOF OF THEOREM 4

The proof is based on the scheme of Section III-B. Fix a choice of the blocklength n, the small positive μ, the (conditional) pmfs $P_{T_1 T_2}$, $P_{S_1|X_1 T_1 T_2}$ and $P_{S_2|X_2 T_1 T_2}$, and the functions f_1 and f_2 so that (86) holds. Define the set $\mathcal{P}^n_{\mu,\text{type-I}}$ to be the subset of types $\pi_{S_1 S_1' S_2 S_2' V Y T_1 T_2}$ such that for all

$(s_1, s_1', s_2, s_2', v, y, t_1, t_2) \in \mathcal{S}_1 \times \mathcal{S}_1 \times \mathcal{S}_2 \times \mathcal{S}_2 \times \mathcal{V} \times \mathcal{Y} \times \mathcal{W}_1 \times \mathcal{W}_2$ the following four conditions hold:

$$|\pi_{S_i X_i T_1 T_2}(s_i, x_i, t_1, t_2) - P_{S_i X_i T_1 T_2}(s_i, x_i, t_1, t_2)| \leq \mu/2,$$
$$i \in \{1, 2\}, \tag{268a}$$

$$|\pi_{S_1 S_2 Y V T_1 T_2}(s_1, s_2, y, v, t_1, t_2)$$
$$- P_{S_1 S_2 Y V T_1 T_2}(s_1, s_2, y, v, t_1, t_2)| \leq \mu, \tag{268b}$$

$$H_{\pi_{S_1' S_2' Y V T_1 T_2}}(S_1', S_2'|Y, V, T_1, T_2)$$
$$\leq H_{\pi_{S_1 S_2 Y V T_1 T_2}}(S_1, S_2|Y, V, T_1, T_2), \tag{268c}$$

$$|\pi_{S_1' S_2' T_1 T_2}(s_1', s_2', t_1, t_2) - P_{S_1 S_2 T_1 T_2}(s_1', s_2', t_1, t_2)| \leq \mu. \tag{268d}$$

Also, set for convenience of notation:

$$P_{S_1'|T_1 T_2}(s_1|t_1, t_2) := P_{S_1|T_1 T_2}(s_1|t_1, t_2),$$
$$\forall(s_1, t_1, t_2) \in \mathcal{S}_1 \times \mathcal{T}_1 \times \mathcal{T}_2, \tag{269}$$

$$P_{S_2'|T_1 T_2}(s_2|t_1, t_2) := P_{S_2|T_1 T_2}(s_2|t_1, t_2),$$
$$\forall(s_2, t_1, t_2) \in \mathcal{S}_2 \times \mathcal{T}_1 \times \mathcal{T}_2. \tag{270}$$

In the following, for simplicity of presentation, we abbreviate the pair (T_1^n, T_2^n) by $\mathbf{T}^n$ and its realization (t_1^n, t_2^n) by $\mathbf{t}^n$.

We first analyze the type-I error probability averaged over the random code construction. Let (M_1, M_2) be the indices of the chosen codewords at the transmitters, if they exist, and define the following events:

$$\mathcal{E}_{\text{Tx}_i}: \left\{\nexists m_i: (S_i^n(m_i), X_i^n, \mathbf{T}^n) \in \mathcal{T}_{\mu/2}^n(P_{S_i X_i \mathbf{T}})\right\},$$
$$i \in \{1, 2\}, \tag{271}$$

$$\mathcal{E}_{\text{Rx}}^{(1)}: \left\{(S_1^n(M_1), S_2^n(M_2), Y^n, V^n, \mathbf{T}^n) \notin \mathcal{T}_\mu^n(P_{S_1 S_2 Y V \mathbf{T}})\right\}, \tag{272}$$

$$\mathcal{E}_{\text{Rx}}^{(2)}: \Big\{\exists m_1' \neq M_1, m_2' \neq M_2:$$
$$H_{\text{tp}(s_1^n(m_1'), s_2^n(m_2'), y^n, v^n, \mathbf{t}^n)}(S_1, S_2|Y, V, \mathbf{T})$$
$$= \min_{\tilde{m}_1, \tilde{m}_2} H_{\text{tp}(s_1^n(\tilde{m}_1), s_2^n(\tilde{m}_2), y^n, v^n, \mathbf{t}^n)}(S_1, S_2|Y, V, \mathbf{T})\Big\}, \tag{273}$$

$$\mathcal{E}_{\text{Rx}}^{(3)}: \Big\{\exists m_2' \neq M_2:$$
$$H_{\text{tp}(s_1^n(M_1), s_2^n(m_2'), y^n, v^n, \mathbf{t}^n)}(S_1, S_2|Y, V, \mathbf{T})$$
$$\min_{\tilde{m}_2} = H_{\text{tp}(s_1^n(M_1), s_2^n(\tilde{m}_2), y^n, v^n, \mathbf{t}^n)}(S_1, S_2|Y, V, \mathbf{T})\Big\}, \tag{274}$$

$$\mathcal{E}_{\text{Rx}}^{(4)}: \Big\{\exists m_1' \neq M_1:$$
$$H_{\text{tp}(s_1^n(m_1'), s_2^n(M_2), y^n, v^n, \mathbf{t}^n)}(S_1, S_2|Y, V, \mathbf{T})$$
$$= \min_{\tilde{m}_1} H_{\text{tp}(s_1^n(\tilde{m}_1), s_2^n(M_2), y^n, v^n, \mathbf{t}^n)}(S_1, S_2|Y, V, \mathbf{T})\Big\}. \tag{275}$$

Notice that the event $\left(\mathcal{E}_{\text{Tx}_1} \cup \mathcal{E}_{\text{Tx}_2} \cup \mathcal{E}_{\text{Rx}}^{(1)} \cup \mathcal{E}_{\text{Rx}}^{(2)} \cup \mathcal{E}_{\text{Rx}}^{(3)} \cup \mathcal{E}_{\text{Rx}}^{(4)}\right)^c$ implies that the receiver decides on $\hat{\mathcal{H}} = 0$. Thus, we obtain

$$\mathbb{E}_\mathcal{C}[\alpha_n] \leq \Pr[\mathcal{E}_{\text{Tx}_1}] + \Pr[\mathcal{E}_{\text{Tx}_2}] + \Pr\left[\mathcal{E}_{\text{Rx}}^{(1)}\Big|\mathcal{E}_{\text{Tx}_1}^c, \mathcal{E}_{\text{Tx}_2}^c\right]$$
$$+ \Pr\left[\mathcal{E}_{\text{Rx}}^{(2)}\Big|\mathcal{E}_{\text{Tx}_1}^c, \mathcal{E}_{\text{Tx}_2}^c, \mathcal{E}_{\text{Rx}}^{(1)c}\right]$$
$$+ \Pr\left[\mathcal{E}_{\text{Rx}}^{(3)}\Big|\mathcal{E}_{\text{Tx}_1}^c, \mathcal{E}_{\text{Tx}_2}^c, \mathcal{E}_{\text{Rx}}^{(1)c}\right]$$

$$+ \Pr\left[\mathcal{E}_{\text{Rx}}^{(4)}\middle|\mathcal{E}_{\text{Tx}_1}^c, \mathcal{E}_{\text{Tx}_2}^c, \mathcal{E}_{\text{Rx}}^{(1)c}\right] \tag{276}$$

$$\leq \epsilon/6 + \epsilon/6 + \epsilon/6 + \epsilon/6 + \epsilon/6 + \epsilon/6 \tag{277}$$

$$= \epsilon, \tag{278}$$

where the second inequality holds for all sufficiently small values of μ and sufficiently large blocklengths n and can be proved as follows. The first and second summands of (276) can be upper bounded by means of the covering lemma [21] and the rate constraint (88); the third by means of the Markov lemma [21]. To prove the upper bound on the fourth term, consider the following set of inequalities

$$\Pr\left[\mathcal{E}_{\text{Rx}}^{(2)}\middle|\mathcal{E}_{\text{Tx}_1}^c, \mathcal{E}_{\text{Tx}_2}^c, \mathcal{E}_{\text{Rx}}^{(1)c}, \mathcal{H} = 0\right]$$

$$= \Pr\left[H_{\text{tp}(s_1^n(M_1), s_2^n(M_2), y^n, v^n, \mathbf{t}^n)}(S_1', S_2'|Y, V, \mathbf{T})\right.$$

$$\geq \min_{\substack{\tilde{m}_1 \neq M_1 \\ \tilde{m}_2 \neq M_2}} H_{\text{tp}(s_1^n(\tilde{m}_1), s_2^n(\tilde{m}_2), y^n, v^n, \mathbf{t}^n)}(S_1, S_2|Y, V, \mathbf{T})\Bigg|$$

$$(S_i^n(M_i), X_i^n, \mathbf{T}^n) \in \mathcal{T}_{\mu/2}^n(P_{S_iX_i\mathbf{T}}), \ i \in \{1, 2\},$$

$$(S_1^n(M_1), S_2^n(M_2), Y^n, V^n, \mathbf{T}^n) \in \mathcal{T}_\mu^n(P_{S_1S_2YV\mathbf{T}}),$$

$$\mathcal{H} = 0\Bigg] \tag{279}$$

$$\overset{(a)}{=} \Pr\left[H_{\text{tp}(s_1^n(1), s_2^n(1), y^n, v^n, \mathbf{t}^n)}(S_1', S_2'|Y, V, \mathbf{T})\right.$$

$$\geq \min_{\substack{\tilde{m}_1 \neq 1 \\ \tilde{m}_2 \neq 1}} H_{\text{tp}(s_1^n(\tilde{m}_1), s_2^n(\tilde{m}_2)|y^n, v^n, \mathbf{t}^n)}(S_1, S_2|Y, V, \mathbf{T})\Bigg|$$

$$(S_i^n(1), X_i^n, \mathbf{T}^n) \in \mathcal{T}_{\mu/2}^n(P_{S_iX_i\mathbf{T}}), \ i \in \{1, 2\},$$

$$(S_1^n(1), S_2^n(1), Y^n, V^n, \mathbf{T}^n) \in \mathcal{T}_\mu^n(P_{S_1S_2YV\mathbf{T}}),$$

$$M_1 = M_2 = 1, \ \mathcal{H} = 0\Bigg] \tag{280}$$

$$= \Pr\left[\bigcup_{\substack{\tilde{m}_1 \neq 1 \\ \tilde{m}_2 \neq 1}} \left\{H_{\text{tp}(s_1^n(1), s_2^n(1), y^n, v^n, \mathbf{t}^n)}(S_1', S_2'|Y, V, \mathbf{T})\right.\right.$$

$$\geq H_{\text{tp}(s_1^n(\tilde{m}_1), s_2^n(\tilde{m}_2), y^n, v^n, \mathbf{t}^n)}(S_1, S_2|Y, V, \mathbf{T})\Bigg\}\Bigg|$$

$$(S_i^n(1), X_i^n, \mathbf{T}^n) \in \mathcal{T}_{\mu/2}^n(P_{S_iX_i\mathbf{T}}), \ i \in \{1, 2\},$$

$$(S_1^n(1), S_2^n(1), Y^n, V^n, \mathbf{T}^n) \in \mathcal{T}_\mu^n(P_{S_1S_2YV\mathbf{T}}),$$

$$M_1 = M_2 = 1, \ \mathcal{H} = 0\Bigg] \tag{281}$$

$$\overset{(b)}{\leq} \sum_{\tilde{m}_1=2}^{2^{nR_1}} \sum_{\tilde{m}_2=2}^{2^{nR_2}} \sum_{\pi \in \mathcal{P}_{\mu,\text{type-I}}^n} \sum_{\substack{s_1^n, s_1'^n, s_2^n, s_2'^n, v^n, y^n, \mathbf{t}^n: \\ \text{tp}(s_1^n, s_1'^n, s_2^n, s_2'^n, v^n, y^n, \mathbf{t}^n) = \pi}}$$

$$\Pr\left[S_1^n(1) = s_1^n, S_2^n(1) = s_2^n, V^n = v^n, Y^n = y^n, \mathbf{T}^n = \mathbf{t}^n\right|$$

$$(S_i^n(1), X_i^n, \mathbf{T}^n) \in \mathcal{T}_{\mu/2}^n(P_{S_iX_i\mathbf{T}}), i \in \{1, 2\},$$

$$(S_1^n(1), S_2^n(1), Y^n, V^n, \mathbf{T}^n) \in \mathcal{T}_\mu^n(P_{S_1S_2YV\mathbf{T}}),$$

$$M_1 = M_2 = 1, \ \mathcal{H} = 0\Bigg]$$

$$\cdot \Pr\left[S_1^n(\tilde{m}_1) = s_1'^n\middle|\mathbf{T}^n = \mathbf{t}^n,\right.$$

$$(S_i^n(1), X_i^n, \mathbf{T}^n) \in \mathcal{T}_{\mu/2}^n(P_{S_iX_i\mathbf{T}}), i \in \{1, 2\},$$

$$M_1 = M_2 = 1, \ \mathcal{H} = 0\Bigg]$$

$$\cdot \Pr\left[S_2^n(\tilde{m}_2) = s_2'^n\middle|\mathbf{T}^n = \mathbf{t}^n,\right.$$

$$(S_i^n(1), X_i^n, \mathbf{T}^n) \in \mathcal{T}_{\mu/2}^n(P_{S_iX_i\mathbf{T}}), i \in \{1, 2\},$$

$$M_1 = M_2 = 1, \ \mathcal{H} = 0\Bigg] \tag{282}$$

$$\overset{(c)}{\leq} \sum_{\pi \in \mathcal{P}_{\mu,\text{type-I}}^n} 2^{nR_1} \cdot 2^{nR_2} \cdot 2^{nH_\pi(S_1, S_1', S_2, S_2', Y, V, \mathbf{T})}$$

$$\cdot 2^{-nH_\pi(S_1, S_2, Y, V, \mathbf{T})} \cdot 2^{-nH_\pi(S_1'|\mathbf{T})}$$

$$\cdot 2^{-nH_\pi(S_2'|\mathbf{T})} \tag{283}$$

$$\leq \sum_{\pi \in \mathcal{P}_{\mu,\text{type-I}}^n} 2^{n(R_1+R_2-H_\pi(S_1', S_2'|Y, V, \mathbf{T}))}$$

$$\cdot 2^{-nH_\pi(S_1'|\mathbf{T})} \cdot 2^{-nH_\pi(S_2'|\mathbf{T})} \tag{284}$$

$$\overset{(d)}{\leq} (n+1)^{|\mathcal{S}_1|^2|\mathcal{S}_2|^2|\mathcal{Y}||\mathcal{V}||\mathcal{W}_1||\mathcal{W}_2|}$$

$$\cdot 2^{n(R_1+R_2-I(S_1,S_2;V,Y|\mathbf{T})-I(S_1;S_2|\mathbf{T})+\delta(\mu))} \tag{285}$$

$$\overset{(e)}{\leq} \epsilon/6, \tag{286}$$

where $\delta(\mu)$ is a function that tends to 0 as $\mu \to 0$, and

- (a) holds by the symmetry of the code construction and the encoding;
- (b) holds by the union bound and because conditioned on $\mathbf{T}^n$ and $M_1 = M_2 = 1$, the sequences $S_1^n(\tilde{m}_1)$ and $S_2^n(\tilde{m}_2)$ are generated independently of each other and of all other sequences;
- (c) holds because all $2^{nH_\pi(S_1, S_2, Y, V, \mathbf{T})}$ tuples $(s_1^n, s_2^n, v^n, y^n, \mathbf{t}^n)$ of the same type π have same conditional probability and similarly all $2^{nH_\pi(S_i'|\mathbf{T})}$ sequences s_i^n, for $i \in \{1, 2\}$, of same joint type with $\mathbf{t}^n$ have same conditional probability;
- (d) holds because for all π in $\mathcal{P}_{\mu,\text{type-I}}^n$, $H_\pi(S_1', S_2'|Y, V, \mathbf{T}) \leq H_\pi(S_1, S_2|Y, V, \mathbf{T})$, because $|\pi_{S_1'S_2'\mathbf{T}} - P_{S_1S_2\mathbf{T}}| \leq \mu$ and $|\pi_{S_1S_2YV\mathbf{T}} - P_{S_1S_2YV\mathbf{T}}| \leq \mu$ and the continuity considerations, and by the standard upper bound on the number of types;
- (e) holds by the rate constraint in (89c).

That also the fifth and sixth summands of (276) are upper bounded by $\epsilon/6$, can be shown in a similar way.

Next, we analyze the type-II error probability averaged over the random code construction. Define events:

$$\mathcal{E}_{\text{Tx}_i}(m_i): \left\{\left(S_i^n(m_i), X_i^n, \mathbf{T}^n\right) \in \mathcal{T}_{\mu/2}^n(P_{S_iX_i\mathbf{T}})\right.$$

$$\left.\text{and} \quad W_i^n = f_i(S_i^n(m_i), X_i^n) \text{ is sent}\right\} \tag{287}$$

and

$$\mathcal{E}_{\text{Rx}}(m_1', m_2'):$$

$$\left\{\left(S_1^n(m_1'), S_2^n(m_2'), Y^n, V^n, \mathbf{T}^n\right) \in \mathcal{T}_\mu^n(P_{S_1S_2YV\mathbf{T}})\right.$$

and

$$H_{\text{tp}(S_1^n(m_1'), S_2^n(m_2'), Y^n, V^n, \mathbf{T}^n)}(S_1', S_2'|Y, V, \mathbf{T}) =$$

$$\min_{\tilde{m}_1, \tilde{m}_2} H_{\text{tp}(S_1^n(\tilde{m}_1), S_2^n(\tilde{m}_2), Y^n, V^n, \mathbf{T}^n)}(S_1, S_2|Y, V, \mathbf{T})\},$$

(288)

Notice that

$$\mathbb{E}_{\mathcal{C}}[\beta_n] = \Pr\left[\hat{\mathcal{H}} = 0 | \mathcal{H} = 1\right]$$

$$= \Pr\left[\bigcup_{m_1', m_2'} \mathcal{E}_{\text{Rx}}(m_1', m_2') \Big| \mathcal{H} = 1\right], \quad (289)$$

where the union is over indices $(m_1', m_2') \in \{1, \ldots, \lfloor 2^{nR_1} \rfloor\} \times \{1, \ldots, \lfloor 2^{nR_2} \rfloor\}$. Notice further that the above probability is upper bounded by the sum of the probabilities of the following nine events:

$\mathcal{B}_1$: $\{\exists (m_1, m_2) \quad \text{s.t.}$
$\quad \mathcal{E}_{\text{Tx}_1}(m_1) \text{ and } \mathcal{E}_{\text{Tx}_2}(m_2) \text{ and } \mathcal{E}_{\text{Rx}}(m_1, m_2)\}$

$\mathcal{B}_2$: $\{\exists (m_1, m_1', m_2) \quad \text{with} \quad m_1 \neq m_1' \quad \text{s.t.}$
$\quad \mathcal{E}_{\text{Tx}_1}(m_1) \text{ and } \mathcal{E}_{\text{Tx}_2}(m_2) \text{ and } \mathcal{E}_{\text{Rx}}(m_1', m_2)\}$

$\mathcal{B}_3$: $\{\exists (m_1, m_2, m_2') \quad \text{with} \quad m_2 \neq m_2' \quad \text{s.t.}$
$\quad \mathcal{E}_{\text{Tx}_1}(m_1) \text{ and } \mathcal{E}_{\text{Tx}_2}(m_2) \text{ and } \mathcal{E}_{\text{Rx}}(m_1, m_2')\}$

$\mathcal{B}_4$: $\{\exists (m_1, m_1', m_2, m_2') \quad \text{with} \quad m_1 \neq m_1' \text{ and } m_2 \neq m_2'$
$\quad \text{s.t.} \quad \mathcal{E}_{\text{Tx}_1}(m_1) \text{ and } \mathcal{E}_{\text{Tx}_2}(m_2) \text{ and } \mathcal{E}_{\text{Rx}}(m_1', m_2')\}$

$\mathcal{B}_5$: $\{\forall m_1 \quad \mathcal{E}_{\text{Tx}_1}^c(m_1) \text{ holds and } \exists (m_1', m_2, m_2') \quad \text{with}$
$\quad m_2 \neq m_2' \quad \text{s.t.} \quad \mathcal{E}_{\text{Tx}_2}(m_2) \text{ and } \quad \mathcal{E}_{\text{Rx}}(m_1', m_2')\}$

$\mathcal{B}_6$: $\{\forall m_1 \quad \mathcal{E}_{\text{Tx}_1}^c(m_1) \text{ holds}\} \cup \{\exists (m_1', m_2) \quad \text{s.t.}$
$\quad \mathcal{E}_{\text{Tx}_2}(m_2) \text{ and } \mathcal{E}_{\text{Rx}}(m_1', m_2)\}$

$\mathcal{B}_7$: $\{\forall m_2 \quad \mathcal{E}_{\text{Tx}_2}^c(m_2) \text{ holds}\} \cup \{\exists (m_1, m_1', m_2') \quad \text{with}$
$\quad m_1 \neq m_1' \quad \text{s.t.} \quad (\mathcal{E}_{\text{Tx}_1}(m_1) \text{ and } \mathcal{E}_{\text{Rx}}(m_1', m_2'))\}$

$\mathcal{B}_8$: $\{\forall m_2 \quad \mathcal{E}_{\text{Tx}_2}^c(m_2) \text{ holds}\} \cup \{\exists (m_1, m_2') \quad \text{s.t.}$
$\quad (\mathcal{E}_{\text{Tx}_1}(m_1) \text{ and } \mathcal{E}_{\text{Rx}}(m_1, m_2'))\}$

$\mathcal{B}_9$: $\{\forall (m_1, m_2) \quad (\mathcal{E}_{\text{Tx}_1}^c(m_1) \quad \cup \quad \mathcal{E}_{\text{Tx}_2}^c(m_2)) \text{ hold}\} \quad \cup$
$\quad \{\exists (m_1', m_2') \quad \text{s.t.} \quad \mathcal{E}_{\text{Rx}}(m_1', m_2')\}$

So, we have

$$\mathbb{E}_{\mathcal{C}}[\beta_n] \leq \sum_{\ell=1}^{9} \Pr\left[\mathcal{B}_\ell | \mathcal{H} = 1\right]. \quad (290)$$

We will bound the nine probabilities on the right-hand side of (290) individually. To simplify the notation, we introduce the following sets of types:

$$\mathcal{P}_{\mu, \text{standard}} := \{\pi_{S_1 S_2 X_1 X_2 YV\mathbf{T}}:$$
$$|\pi_{S_i X_i \mathbf{T}} - P_{S_i X_i \mathbf{T}}| < \mu/2, \ i \in \{1, 2\},$$
$$|\pi_{S_1 S_2 YV\mathbf{T}} - P_{S_1 S_2 YV\mathbf{T}}| < \mu\},$$
(291)

$$\mathcal{P}_{\mu, \text{dec}, 1} := \{\pi_{S_1 S_1' S_2 X_1 X_2 YV\mathbf{T}}:$$
$$|\pi_{S_i X_i \mathbf{T}} - P_{S_i X_i \mathbf{T}}| < \mu/2, \ i \in \{1, 2\},$$

$$|\pi_{S_1' S_2 YV\mathbf{T}} - P_{S_1 S_2 YV\mathbf{T}}| < \mu,$$
$$H_\pi(S_1'|S_2, Y, V, \mathbf{T}) \leq H_\pi(S_1|S_2, Y, V, \mathbf{T})\},$$
(292)

$$\mathcal{P}_{\mu, \text{dec}, 2} := \{\pi_{S_1 S_2 S_2' X_1 X_2 YV\mathbf{T}}:$$
$$|\pi_{S_i X_i \mathbf{T}} - P_{S_i X_i \mathbf{T}}| < \mu/2, \ i \in \{1, 2\},$$
$$|\pi_{S_1 S_2' YV\mathbf{T}} - P_{S_1 S_2 YV\mathbf{T}}| < \mu,$$
$$H_\pi(S_2'|S_1, Y, V, \mathbf{T}) \leq H_\pi(S_2|S_1, Y, V, \mathbf{T})\},$$
(293)

$$\mathcal{P}_{\mu, \text{dec}, 12} := \{\pi_{S_1 S_1' S_2 S_2' X_1 X_2 YV\mathbf{T}}:$$
$$|\pi_{S_i X_i \mathbf{T}} - P_{S_i X_i \mathbf{T}}| < \mu/2, \ i \in \{1, 2\},$$
$$|\pi_{S_1' S_2' YV\mathbf{T}} - P_{S_1 S_2 YV\mathbf{T}}| < \mu,$$
$$H_\pi(S_1', S_2'|Y, V, \mathbf{T}) \leq H_\pi(S_1, S_2|Y, V, \mathbf{T})\},$$
(294)

$$\mathcal{P}_{\mu, \text{miss}, 1a} := \{\pi_{S_1' S_2 S_2' X_2 YV\mathbf{T}}:$$
$$|\pi_{S_2 X_2 \mathbf{T}} - P_{S_2 X_2 \mathbf{T}}| < \mu/2,$$
$$|\pi_{S_1' S_2' YV\mathbf{T}} - P_{S_1 S_2 YV\mathbf{T}}| < \mu,$$
$$H(S_1', S_2'|Y, V, \mathbf{T}) \leq H_\pi(S_1', S_2|Y, V, \mathbf{T})\},$$
(295)

$$\mathcal{P}_{\mu, \text{miss}, 1b} := \{\pi_{S_1' S_2 X_2 YV\mathbf{T}}:$$
$$|\pi_{S_2 X_2 \mathbf{T}} - P_{S_2 X_2 \mathbf{T}}| < \mu/2,$$
$$|\pi_{S_1' S_2 YV\mathbf{T}} - P_{S_1 S_2 YV\mathbf{T}}| < \mu\}, \quad (296)$$

$$\mathcal{P}_{\mu, \text{miss}, 2a} := \{\pi_{S_1 S_1' S_2' X_1 YV\mathbf{T}Y}:$$
$$|\pi_{S_1 X_1 \mathbf{T}} - P_{S_1 X_1 \mathbf{T}}| < \mu/2,$$
$$|\pi_{S_1' S_2' YV\mathbf{T}} - P_{S_1' S_2' YV\mathbf{T}}| < \mu,$$
$$H(S_1', S_2'|Y, V, \mathbf{T}) \leq H_\pi(S_1, S_2'|Y, V, \mathbf{T})\},$$
(297)

$$\mathcal{P}_{\mu, \text{miss}, 2b} := \{\pi_{S_1 S_2' X_1 YV\mathbf{T}}:$$
$$|\pi_{S_1 X_1 \mathbf{T}} - P_{S_1 X_1 \mathbf{T}}| < \mu/2,$$
$$|\pi_{S_1 S_2' YV\mathbf{T}} - P_{S_1 S_2 YV\mathbf{T}}| < \mu\}. \quad (298)$$

Consider the probability of event $\mathcal{B}_1$. By Sanov's theorem [29] and the way the source sequences and the codebooks are generated, we have

$$\Pr\left[\mathcal{B}_1 | \mathcal{H} = 1\right] \leq \sum_{m_1, m_2}$$

$$\Pr\Big[(S_i^n(m_i), X_i^n, \mathbf{T}^n) \in \mathcal{T}_{\mu/2}^n(P_{S_i X_i \mathbf{T}}) \text{ and}$$

$$W_i^n = f_i(S_i^n(m_i), X_i^n) \text{ is sent for } i \in \{1, 2\}, \text{ and}$$

$$(S_1^n(m_1), S_2^n(m_2), Y^n, V^n, T_1^n, T_2^n) \in \mathcal{T}_\mu^n(P_{S_1 S_2 YV\mathbf{T}})$$
$$\Big| \mathcal{H} = 1\Big]$$
(299)

$$\leq 2^{n(R_1 + R_2)} \cdot \max_{\pi \in \mathcal{P}_{\mu, \text{standard}}}$$

$$2^{-n\left(D\left(\pi_{S_1 S_2 X_1 X_2 YV\mathbf{T}} \left\| \begin{array}{c} P_{S_1|\mathbf{T}} P_{S_2|\mathbf{T}} Q_{X_1 X_2 Y} \\ \cdot P_{\mathbf{T}} \Gamma_{V|S_1 S_2 X_1 X_2} \end{array}\right.\right) - \delta_{1,n}'\right)},$$
(300)

for some sequence $\delta'_{1,n}$ that tends to 0 as $n \to \infty$. Define now:

$$\tilde{\theta}_\mu^{\text{standard}} := \min_{\pi \in \mathcal{P}_{\mu,\text{standard}}}$$
$$D\Big(\pi_{S_1 S_2 X_1 X_2 Y V \mathbf{T}} \| P_{S_1|\mathbf{T}} P_{S_2|\mathbf{T}} Q_{X_1 X_2 Y} P_{\mathbf{T}}$$
$$\cdot \Gamma_{V|S_1 S_2 X_1 X_2}\Big)$$
$$- R_1 - R_2 - \delta'_{1,n}, \tag{301}$$

and observe that:

$$\tilde{\theta}_\mu^{\text{standard}} \overset{\text{(eq. (88))}}{=} \min_{\pi \in \mathcal{P}_{\mu,\text{standard}}}$$
$$D\Big(\pi_{S_1 S_2 X_1 X_2 Y V \mathbf{T}} \| P_{S_1|\mathbf{T}} P_{S_2|\mathbf{T}} Q_{X_1 X_2 Y} P_{\mathbf{T}}$$
$$\cdot \Gamma_{V|S_1 S_2 X_1 X_2}\Big)$$
$$- I(S_1; X_1|\mathbf{T}) - I(X_2; S_2|\mathbf{T}) - 2\mu - \delta'_{1,n} \tag{302}$$
$$= \min_{\pi \in \mathcal{P}_{\mu,\text{standard}}}$$
$$D\Big(\pi_{S_1 S_2 X_1 X_2 Y V \mathbf{T}} \| P_{S_1|X_1 \mathbf{T}} P_{S_2|X_2 \mathbf{T}} Q_{X_1 X_2 Y}$$
$$\cdot P_{\mathbf{T}} \Gamma_{V|S_1 S_2 X_1 X_2}\Big)$$
$$- \delta_{1,n}(\mu) \tag{303}$$
$$= \theta_\mu^{\text{standard}} - \delta_{1,n}(\mu), \tag{304}$$

for some sequence of functions $\delta_{1,n}(\mu)$ that goes to zero as $n \to \infty$ and $\mu \to 0$ and

$$\theta_\mu^{\text{standard}} := \min_{\pi \in \mathcal{P}_{\mu,\text{standard}}} D\Big(\pi_{S_1 S_2 X_1 X_2 Y V \mathbf{T}} \| P_{S_1|X_1 \mathbf{T}} P_{S_2|X_2 \mathbf{T}}$$
$$\cdot Q_{X_1 X_2 Y} P_{\mathbf{T}} \Gamma_{V|S_1 S_2 X_1 X_2}\Big) \tag{305}$$

Combining (300)–(304), we have:

$$\Pr\big[\mathcal{B}_1 | \mathcal{H} = 1\big] \leq 2^{-n\big(\theta_\mu^{\text{standard}} - \delta_{1,n}(\mu)\big)}. \tag{306}$$

Consider next event $\mathcal{B}_2$. Recall that $\text{ctp}(a^n|b^n)$ denotes the conditional type of sequence a^n given b^n. Its probability can be upper bounded as:

$$\Pr\big[\mathcal{B}_2 | \mathcal{H} = 1\big]$$
$$\leq \sum_{m_1, m'_1, m_2} \Pr\Big[\big(\mathcal{E}_{\text{Tx}_1}(m_1) \cap \mathcal{E}_{\text{Tx}_2}(m_2) \cap \mathcal{E}_{\text{Rx}}(m'_1, m_2)\big)\Big|\mathcal{H}=1\Big] \tag{307}$$
$$\leq \sum_{m_1, m'_1, m_2} \Pr\Big[(S_i^n(m_i), X_i^n, \mathbf{T}^n) \in \mathcal{T}_{\mu/2}^n(P_{S_i X_i \mathbf{T}}) \text{ and}$$

$$W_i^n = f_i(S_i^n(m_i), X_i^n) \text{ is sent for } i \in \{1,2\},$$
$$(S_1^n(m'_1), S_2^n(m_2), Y^n, V^n, \mathbf{T}^n) \in \mathcal{T}_\mu^n(P_{S_1 S_2 Y V \mathbf{T}}),$$
$$H_{\text{tp}(S_1^n(m'_1), S_2^n(m_2), Y^n, V^n, \mathbf{T}^n)}(S'_1|S_2, Y, V, \mathbf{T})$$
$$\leq H_{\text{tp}(S_1^n(m_1), S_2^n(m_2), Y^n, V^n, \mathbf{T}^n)}(S_1|S_2, Y, V, \mathbf{T})$$
$$\Big|\mathcal{H}=1\Big] \tag{308}$$

$$= \sum_{m_1, m'_1, m_2} \sum_{\pi \in \mathcal{P}_{\mu,\text{dec},1}} \Pr\Big[W_i^n = f_i(S_i^n(m_i), X_i^n),$$
$$\text{tp}\big(S_1^n(m_1), S_1^n(m'_1), S_2^n(m_2), X_1^n, X_2^n, Y^n, V^n, \mathbf{T}^n\big)$$
$$= \pi_{S_1 S'_1 S_2 X_1 X_2 Y V \mathbf{T}} \Big|\mathcal{H} = 1\Big] \tag{309}$$

$$\leq \sum_{m_1, m'_1, m_2} \sum_{\pi \in \mathcal{P}_{\mu,\text{dec},1}}$$
$$\Pr\Big[\text{tp}\big(S_1^n(m_1), S_1^n(m'_1), S_2^n(m_2), X_1^n, X_2^n, Y^n, \mathbf{T}^n\big)$$
$$= \pi_{S_1 S'_1 S_2 X_1 X_2 Y \mathbf{T}} \Big|\mathcal{H} = 1\Big]$$
$$\cdot \Pr\Big[\text{ctp}\big(V^n|S_1^n(m_1), S_1^n(m'_1), S_2^n(m_2), X_1^n, X_2^n, Y^n, \mathbf{T}^n\big)$$
$$= \pi_{V|S_1 S'_1 S_2 X_1 X_2 Y \mathbf{T}}\Big|$$
$$\text{tp}\big(S_1^n(m_1), S_1^n(m'_1), S_2^n(m_2), X_1^n, X_2^n, Y^n, \mathbf{T}^n\big)$$
$$= \pi_{S_1 S'_1 S_2 X_1 X_2 Y \mathbf{T}}, \quad W_i^n = f_i(S_i^n(m_i), X_i^n), \quad \mathcal{H} = 1\Big] \tag{310}$$

$$\leq \sum_{m_1, m'_1, m_2} (n+1)^{|\mathcal{S}_1|^2 |\mathcal{S}_2||\mathcal{X}_1||\mathcal{X}_2||\mathcal{Y}||\mathcal{W}_1||\mathcal{W}_2|}$$
$$\cdot \min_{\pi \in \mathcal{P}_{\mu,\text{dec},1}} \Big[2^{-n D(\pi_{S_1 S'_1 S_2 X_1 X_2 Y \mathbf{T}} \| P_{S_1|\mathbf{T}} P_{S'_1|\mathbf{T}} P_{S_2|\mathbf{T}} Q_{X_1 X_2 Y} P_{\mathbf{T}})}$$
$$\cdot 2^{-n \mathbb{E}_\pi[D(\pi_{V|S_1 S'_1 S_2 X_1 X_2 Y \mathbf{T}} \| \Gamma_{V|S_1 S_2 X_1 X_2})]} \Big] \tag{311}$$

$$\leq 2^{-n\tilde{\theta}_\mu^{\text{dec},1}}, \tag{312}$$

where we define:

$$\tilde{\theta}_\mu^{\text{dec},1} := \min_{\pi \in \mathcal{P}_{\mu,\text{dec},1}} D\Big(\pi_{S_1 S'_1 S_2 X_1 X_2 Y V \mathbf{T}} \| P_{S_1|\mathbf{T}} P_{S'_1|\mathbf{T}} P_{S_2|\mathbf{T}}$$
$$\cdot Q_{X_1 X_2 Y} P_{\mathbf{T}} \Gamma_{V|S_1 S_2 X_1 X_2}\Big)$$
$$- 2R_1 - R_2 - \delta'_{2,n} \tag{313}$$

for some sequence $\delta'_{2,n}$ that tends to 0 as $n \to \infty$. Notice the following set of inequalities. By (88):

$$\tilde{\theta}_\mu^{\text{dec},1} = \min_{\pi \in \mathcal{P}_{\mu,\text{dec},1}} D\big(\pi_{S_1 S'_1 S_2 X_1 X_2 Y V \mathbf{T}} \| P_{S_1|\mathbf{T}} P_{S'_1|\mathbf{T}} P_{S_2|\mathbf{T}}$$
$$\cdot Q_{X_1 X_2 Y} P_{\mathbf{T}} \Gamma_{V|S_1 S_2 X_1 X_2}\big)$$
$$- 2I(S_1; X_1|\mathbf{T}) - I(S_2; X_2|\mathbf{T}) - 3\mu - \delta'_{2,n} \tag{314}$$

$$\overset{\text{(CR)}}{=} \min_{\pi \in \mathcal{P}_{\mu,\text{dec},1}} \Big[D\big(\pi_{S_1 S_2 X_1 X_2 Y V \mathbf{T}} \| P_{S_1|\mathbf{T}} P_{S_2|\mathbf{T}}$$
$$\cdot Q_{X_1 X_2 Y} P_{\mathbf{T}} \Gamma_{V|S_1 S_2 X_1 X_2}\big)$$
$$+ \mathbb{E}_{\pi_{S_1 S_2 X_1 X_2 Y V \mathbf{T}}}\big[D(\pi_{S'_1|S_1 S_2 X_1 X_2 Y V \mathbf{T}} \| P_{S'_1|\mathbf{T}})\big]\Big]$$
$$- 2I(S_1; X_1|\mathbf{T}) - I(S_2; X_2|\mathbf{T}) - 3\mu - \delta'_{2,n} \tag{315}$$

$$\overset{\text{(DP)}}{\geq} \min_{\pi \in \mathcal{P}_{\mu,\text{dec},1}} \Big[D\big(\pi_{S_1 S_2 X_1 X_2 Y V \mathbf{T}} \| P_{S_1|\mathbf{T}} P_{S_2|\mathbf{T}}$$
$$\cdot Q_{X_1 X_2 Y} P_{\mathbf{T}} \Gamma_{V|S_1 S_2 X_1 X_2}\big)$$
$$+ \mathbb{E}_{\pi_{S'_1 Y V \mathbf{T}}}\big[D(\pi_{S'_1|S_2 Y V \mathbf{T}} \| P_{S'_1|\mathbf{T}})\big]\Big]$$
$$- 2I(S_1; X_1|\mathbf{T}) - I(S_2; X_2|\mathbf{T}) - 3\mu - \delta'_{2,n} \tag{316}$$

$$\overset{(g)}{=} \min_{\pi \in \mathcal{P}_{\mu,\text{dec},1}} D(\pi_{S_1 S_2 X_1 X_2 YV\mathbf{T}} \| P_{S_1|X_1\mathbf{T}} P_{S_2|X_2\mathbf{T}}$$
$$\cdot Q_{X_1 X_2 Y} P_{\mathbf{T}} \Gamma_{V|S_1 S_2 X_1 X_2})$$
$$+ I(S_1; S_2, Y, V|\mathbf{T}) - I(S_1; X_1|\mathbf{T})$$
$$- \delta_{2,n}(\mu) \tag{317}$$

$$\overset{(h)}{=} \min_{\pi \in \mathcal{P}_{\mu,\text{dec},1}} D(\pi_{S_1 S_2 X_1 X_2 YV\mathbf{T}} \| P_{S_1|X_1\mathbf{T}} P_{S_2|X_2\mathbf{T}}$$
$$\cdot Q_{X_1 X_2 Y} P_{\mathbf{T}} \Gamma_{V|S_1 S_2 X_1 X_2})$$
$$+ I(S_1; Y, V|S_2, \mathbf{T}) - I(S_1; X_1|S_2, \mathbf{T})$$
$$- \delta_{2,n}(\mu) \tag{318}$$

$$= \theta_{\mu}^{\text{dec},1} - \delta_{2,n}(\mu), \tag{319}$$

where $\delta_{2,n}(\mu)$ is a sequence of functions that goes to zero as $n \to \infty$ and $\mu \to 0$; (g) holds because $|\pi_{S_1 S_2 YV\mathbf{T}} - P_{S_1 S_2 YV\mathbf{T}}| \leq \mu$; and (h) holds by the Markov chain $S_2 \to (X_1, \mathbf{T}) \to S_1$, and we defined

$$\theta_{\mu}^{\text{dec},1} := \min_{\pi \in \mathcal{P}_{\mu,\text{dec},1}} D(\pi_{S_1 S_2 X_1 X_2 YV\mathbf{T}} \| P_{S_1|X_1\mathbf{T}} P_{S_2|X_2\mathbf{T}}$$
$$\cdot Q_{X_1 X_2 Y} P_{\mathbf{T}} \Gamma_{V|S_1 S_2 X_1 X_2})$$
$$+ I(S_1; S_2, Y, V|\mathbf{T}) - I(S_1; X_1|\mathbf{T}). \tag{320}$$

Combining (308)–(319), one then obtains:

$$\Pr\left[\mathcal{B}_2 \big| \mathcal{H} = 1\right] \leq 2^{-n\left(\theta_{\mu}^{\text{dec},1} - \delta_{2,n}(\mu)\right)}. \tag{321}$$

In a similar way, one can also derive the upper bound

$$\Pr\left[\mathcal{B}_3 \big| \mathcal{H} = 1\right] \leq 2^{-n\left(\theta_{\mu}^{\text{dec},2} - \delta_{3,n}(\mu)\right)}, \tag{322}$$

where

$$\theta_{\mu}^{\text{dec},2} := \min_{\pi \in \mathcal{P}_{\mu,\text{dec},2}}$$
$$D(\pi_{S_1 S_2 X_1 X_2 YV\mathbf{T}} \| P_{S_1|X_1\mathbf{T}} P_{S_2|X_2\mathbf{T}}$$
$$\cdot Q_{X_1 X_2 Y} P_{\mathbf{T}} \Gamma_{V|S_1 S_2 X_1 X_2})$$
$$+ I(S_2; Y, V|S_1, \mathbf{T}) - I(S_2; X_2|S_1, \mathbf{T}), \tag{323}$$

and $\delta_{3,n}(\mu)$ is a sequence of functions that goes to zero as $n \to \infty$ and $\mu \to 0$.

Next, consider event $\mathcal{B}_4$. Using similar steps as before, its probability can be upper bounded as

$$\Pr\left[\mathcal{B}_4 \big| \mathcal{H} = 1\right] \leq 2^{-n\tilde{\theta}_{\mu}^{\text{dec},12}}, \tag{324}$$

where

$$\tilde{\theta}_{\mu}^{\text{dec},12} := \min_{\pi \in \mathcal{P}_{\mu,\text{dec},12}}$$
$$D(\pi_{S_1 S_1' S_2 S_2' X_1 X_2 YV\mathbf{T}} \| P_{S_1|\mathbf{T}} P_{S_1'|\mathbf{T}} P_{S_2|\mathbf{T}} P_{S_2'|\mathbf{T}}$$
$$\cdot Q_{X_1 X_2 Y} P_{\mathbf{T}} \Gamma_{V|S_1 S_2 X_1 X_2})$$
$$- 2R_1 - 2R_2 - \delta_{4,n}', \tag{325}$$

and $\delta_{4,n}'$ is a sequence that tends to 0 as $n \to \infty$ and $\mu \to 0$. We have the following set of inequalities. By (88):

$$\tilde{\theta}_{\mu}^{\text{dec},12} = \min_{\pi \in \mathcal{P}_{\mu,\text{dec},12}} D(\pi_{S_1 S_1' S_2 S_2' X_1 X_2 YV\mathbf{T}} \| P_{S_1|\mathbf{T}} P_{S_1'|\mathbf{T}}$$
$$\cdot P_{S_2|\mathbf{T}} P_{S_2'|\mathbf{T}} Q_{X_1 X_2 Y} P_{\mathbf{T}} \Gamma_{V|S_1 S_2 X_1 X_2})$$
$$- 2I(S_1; X_1|\mathbf{T}) - 2I(S_2; X_2|\mathbf{T}) - 4\mu - \delta_{4,n}' \tag{326}$$

$$\overset{(\text{DP})}{\geq} \min_{\pi \in \mathcal{P}_{\mu,\text{dec},12}} \Big[D(\pi_{S_1 S_2 X_1 X_2 YV\mathbf{T}} \| P_{S_1|\mathbf{T}} P_{S_2|\mathbf{T}}$$
$$\cdot Q_{X_1 X_2 Y} P_{\mathbf{T}} \Gamma_{V|S_1 S_2 X_1 X_2})$$
$$+ \mathbb{E}_{\pi_{YV\mathbf{T}}}\big[D(\pi_{S_1' S_2'|YV\mathbf{T}} \| P_{S_1'|\mathbf{T}} P_{S_2'|\mathbf{T}}) \big] \Big]$$
$$- 2I(S_1; X_1|\mathbf{T}) - 2I(S_2; X_2|\mathbf{T}) - 4\mu - \delta_{4,n}' \tag{327}$$

$$\overset{(i)}{=} \min_{\pi \in \mathcal{P}_{\mu,\text{dec},12}} D(\pi_{S_1 S_2 X_1 X_2 YV\mathbf{T}} \| P_{S_1|X_1\mathbf{T}} P_{S_2|X_2\mathbf{T}}$$
$$\cdot Q_{X_1 X_2 Y} P_{\mathbf{T}} \Gamma_{V|S_1 S_2 X_1 X_2})$$
$$+ I(S_1, S_2; Y, V|\mathbf{T}) + I(S_1; S_2|\mathbf{T})$$
$$- I(S_1; X_1|\mathbf{T}) - I(S_2; X_2|\mathbf{T}) - \delta_{4,n}(\mu) \tag{328}$$

$$\overset{(j)}{=} \min_{\pi \in \mathcal{P}_{\mu,\text{dec},12}} D(\pi_{S_1 S_2 X_1 X_2 YV\mathbf{T}} \| P_{S_1|X_1\mathbf{T}} P_{S_2|X_2\mathbf{T}}$$
$$\cdot Q_{X_1 X_2 Y} P_{\mathbf{T}} \Gamma_{V|S_1 S_2 X_1 X_2})$$
$$+ I(S_1, S_2; Y, V|\mathbf{T}) - I(S_1, S_2; X_1, X_2|\mathbf{T})$$
$$- \delta_{4,n}(\mu) \tag{329}$$

$$= \theta_{\mu}^{\text{dec},12} - \delta_{4,n}(\mu), \tag{330}$$

where $\delta_{4,n}(\mu)$ is a sequence of functions that goes to zero as $n \to \infty$ and $\mu \to 0$ and

$$\theta_{\mu}^{\text{dec},12} := \min_{\pi \in \mathcal{P}_{\mu,\text{dec},12}} D(\pi_{S_1 S_2 X_1 X_2 YV\mathbf{T}} \| P_{S_1|X_1\mathbf{T}} P_{S_2|X_2\mathbf{T}}$$
$$\cdot Q_{X_1 X_2 Y} P_{\mathbf{T}} \Gamma_{V|S_1 S_2 X_1 X_2})$$
$$+ I(S_1, S_2; Y, V|\mathbf{T}) - I(S_1, S_2; X_1, X_2|\mathbf{T}). \tag{331}$$

Here, (i) holds by $|\pi_{S_1' S_2' YV\mathbf{T}} - P_{S_1 S_2 YV\mathbf{T}}| \leq \mu$, by rearranging terms, and by the continuity of KL-divergence; and (j) holds by the Markov chains $(S_2, X_2) \to (X_1, \mathbf{T}) \to S_1$ and $(S_1, X_1) \to (X_2, \mathbf{T}) \to S_2$. Combining (324)–(330), one then obtains:

$$\Pr\left[\mathcal{B}_4 \big| \mathcal{H} = 1\right] \leq 2^{-n\left(\theta_{\mu}^{\text{dec},12} - \delta_{4,n}(\mu)\right)}. \tag{332}$$

We upper bound the probability of event $\mathcal{B}_5$. We have:

$$\Pr\left[\mathcal{B}_5 \big| \mathcal{H} = 1\right]$$
$$\leq \sum_{m_1', m_2, m_2'}$$
$$\Pr\Big[W_1^n = T_1^n, \quad (S_2^n(m_2), X_2^n, \mathbf{T}^n) \in \mathcal{T}_{\mu/2}^n(P_{S_2 X_2 \mathbf{T}}),$$
$$(S_1^n(m_1'), S_2^n(m_2'), Y^n, V^n, \mathbf{T}^n) \in \mathcal{T}_{\mu}^n(P_{S_2 YV\mathbf{T}})$$

$$H_{\text{tp}(S_1^n(m_1'),S_2^n(m_2')),Y^n,V^n,\mathbf{T}^n)}(S_1',S_2'|Y,V,\mathbf{T})$$

$$\leq H_{\text{tp}(S_1^n(m_1'),S_2^n(m_2)),Y^n,V^n,\mathbf{T}^n)}(S_1,S_2|Y,V,\mathbf{T})$$
$$\Big| \mathcal{H}=1 \Big] \tag{333}$$

$$= \sum_{m_1',m_2,m_2'} \sum_{\pi\in\mathcal{P}_{\mu,\text{miss,1a}}}$$

$$\Pr\Big[\text{tp}\big(S_1^n(m_1'),S_2^n(m_2),S_2^n(m_2'),X_2^n,Y^n,V^n,\mathbf{T}^n\big)$$
$$= \pi_{S_1'S_2S_2'XYV\mathbf{T}}, \quad W_1^n = T_1^n, \Big|\mathcal{H}=1\Big] \tag{334}$$

$$\leq \sum_{m_1',m_2,m_2'} \sum_{\pi\in\mathcal{P}_{\mu,\text{miss,1a}}}$$

$$\Pr\Big[\text{tp}\big(S_1^n(m_1'),S_2^n(m_2),S_2^n(m_2'),X_2^n,Y^n,\mathbf{T}^n\big)$$
$$= \pi_{S_1'S_2S_2'XY\mathbf{T}}\Big|\mathcal{H}=1\Big]$$
$$\cdot\Pr\Big[\text{ctp}\big(V^n|S_1^n(m_1'),S_2^n(m_2),S_2^n(m_2'),X_2^n,Y^n,$$
$$T_1^n,T_2^n\big) = \pi_{V|S_1'S_2S_2'XY\mathbf{T}}\Big|$$
$$\text{tp}\big(S_1^n(m_1'),S_2^n(m_2),S_2^n(m_2'),X_2^n,Y^n,\mathbf{T}^n\big)$$
$$= \pi_{S_1'S_2S_2'XY\mathbf{T}}, W_1^n = T_1^n, \mathcal{H}=1\Big] \tag{335}$$

$$\leq \sum_{m_1',m_2,m_2'} \sum_{\pi\in\mathcal{P}_{\mu,\text{miss,1a}}}$$

$$2^{-n\big(D(\pi_{S_1'S_2S_2'X_2Y\mathbf{T}}\|P_{S_1'|\mathbf{T}}P_{S_2|\mathbf{T}}P_{S_2'|\mathbf{T}}Q_{X_2Y}P_{\mathbf{T}})-\delta_{5,n}'\big)}$$
$$\cdot 2^{-n\big(\mathbb{E}_{\pi_{S_1'S_2S_2'X_2Y\mathbf{T}}}[D(\pi_{V|S_1'S_2S_2'X_2Y\mathbf{T}}\|\Gamma_{V|T_1S_2X_2}^{(1)})]-\delta_{5,n}''\big)} \tag{336}$$

$$\leq 2^{-n\tilde{\theta}_\mu^{\text{miss,1a}}}, \tag{337}$$

where $\delta_{5,n}', \delta_{5,n}''$ are sequences that tend to 0 as $n\to\infty$ and

$$\tilde{\theta}_\mu^{\text{miss,1a}} := \min_{\pi\in\mathcal{P}_{\mu,\text{miss,1a}}} D\big(\pi_{S_1'S_2S_2'X_1X_2YV\mathbf{T}}\|P_{S_1'|\mathbf{T}}P_{S_2|\mathbf{T}}P_{S_2'|\mathbf{T}}$$
$$\cdot Q_{X_2Y}P_{\mathbf{T}}\Gamma_{V|T_1S_2X_2}^{(1)}\big)$$
$$-R_1-2R_2-\delta_{5,n}'-\delta_{5,n}''. \tag{338}$$

We have the following set of inequalities. By (88):

$$\tilde{\theta}_\mu^{\text{miss,1a}} = \min_{\pi\in\mathcal{P}_{\mu,\text{miss,1a}}} D\big(\pi_{S_1'S_2S_2'X_1X_2YV\mathbf{T}}\|P_{S_1'|\mathbf{T}}P_{S_2|\mathbf{T}}P_{S_2'|\mathbf{T}}$$
$$\cdot Q_{X_2Y}P_{\mathbf{T}}\Gamma_{V|T_1S_2X_2}^{(1)}\big)$$
$$-I(S_1;X_1|\mathbf{T})-2I(S_2;X_2|\mathbf{T})$$
$$-3\mu-\delta_{5,n}'-\delta_{5,n}'' \tag{339}$$

$$\overset{\text{(DP)}}{\geq} \min_{\pi\in\mathcal{P}_{\mu,\text{miss,1a}}} \Big[D\big(\pi_{S_2X_1X_2YV\mathbf{T}}\|P_{S_2|\mathbf{T}}Q_{X_2Y}P_{\mathbf{T}}$$
$$\cdot\Gamma_{V|T_1S_2X_2}^{(1)}\big)$$

$$+\mathbb{E}_{\pi_{YV\mathbf{T}}}\big[D(\pi_{S_1'S_2'|YV\mathbf{T}}\|P_{S_1'|\mathbf{T}}P_{S_2'|\mathbf{T}})\big]\Big]$$
$$-I(S_1;X_1|\mathbf{T})-2I(S_2;X_2|\mathbf{T})-3\mu-\delta_{5,n}'-\delta_{5,n}'' \tag{340}$$

$$= \min_{\pi\in\mathcal{P}_{\mu,\text{miss,1a}}} D(\pi_{S_2X_2YV\mathbf{T}}\|P_{S_2|X_2\mathbf{T}}Q_{X_2Y}P_{\mathbf{T}}$$
$$\cdot\Gamma_{V|T_1S_2X_2}^{(1)})$$
$$+I(S_1,S_2;Y,V|\mathbf{T})+I(S_1;S_2|\mathbf{T})$$
$$-I(S_1;X_1|\mathbf{T})-I(S_2;X_2|\mathbf{T})-\delta_{5,n}(\mu) \tag{341}$$

$$= \min_{\pi\in\mathcal{P}_{\mu,\text{miss,1a}}} D(\pi_{S_2X_2YV\mathbf{T}}\|P_{S_2|X_2\mathbf{T}}Q_{X_2Y}P_{\mathbf{T}}$$
$$\cdot\Gamma_{V|T_1S_2X_2}^{(1)})$$
$$+I(S_1,S_2;Y,V|\mathbf{T})-I(S_1,S_2;X_1,X_2|\mathbf{T})$$
$$-\delta_{5,n}(\mu) \tag{342}$$

$$= \theta_\mu^{\text{miss,1a}}-\delta_5(\mu), \tag{343}$$

where $\delta_5(\mu)$ is a sequence of functions that goes to zero as $n\to\infty$ and $\mu\to 0$ and

$$\theta_\mu^{\text{miss,1a}} :=$$
$$\min_{\pi\in\mathcal{P}_{\mu,\text{miss,1a}}} D(\pi_{S_2X_2YV\mathbf{T}}\|P_{S_2|X_2\mathbf{T}}Q_{X_2Y}P_{\mathbf{T}}\Gamma_{V|T_1S_2X_2}^{(1)})$$
$$+I(S_1,S_2;Y,V|\mathbf{T})-I(S_1,S_2;X_1,X_2|\mathbf{T}). \tag{344}$$

Combining (337)–(343) leads to:

$$\Pr\big[\mathcal{B}_5\big|\mathcal{H}=1\big] \leq 2^{-n\big(\theta_\mu^{\text{miss,1a}}-\delta_5(\mu)\big)}. \tag{345}$$

The probability of event $\mathcal{B}_6$ can be upper bounded in a similar way to obtain:

$$\Pr\big[\mathcal{B}_6\big|\mathcal{H}=1\big] \leq 2^{-n\tilde{\theta}_\mu^{\text{miss,1b}}}, \tag{346}$$

where

$$\tilde{\theta}_\mu^{\text{miss,1b}} :=$$
$$\min_{\pi\in\mathcal{P}_{\mu,\text{miss,1b}}} D\big(\pi_{S_1'S_2X_2YV\mathbf{T}}\|P_{S_1'|\mathbf{T}}P_{S_2|\mathbf{T}}Q_{X_2Y}P_{\mathbf{T}}\Gamma_{V|T_1S_2X_2}^{(1)}\big)$$
$$-R_1-R_2-\delta_{6,n}' \tag{347}$$

for some sequence $\delta_{6,n}'$ that tends to 0 as $n\to\infty$. We have the following set of inequalities. By (88):

$$\tilde{\theta}_\mu^{\text{miss,1b}} = \min_{\pi\in\mathcal{P}_{\mu,\text{miss,1b}}} D\big(\pi_{S_1'S_2X_2YV\mathbf{T}}\|P_{S_1'|\mathbf{T}}P_{S_2|\mathbf{T}}Q_{X_2Y}$$
$$\cdot P_{\mathbf{T}}\Gamma_{V|T_1S_2X_2}^{(1)}\big)$$
$$-I(S_1;X_1|\mathbf{T})-I(S_2;X_2|\mathbf{T})-2\mu-\delta_{6,n}' \tag{348}$$

$$\overset{\text{(CR)}}{=} \min_{\pi\in\mathcal{P}_{\mu,\text{miss,1b}}} \Big[D(\pi_{S_2X_2YV\mathbf{T}}\|P_{S_2|\mathbf{T}}Q_{X_2Y}P_{\mathbf{T}}$$
$$\cdot\Gamma_{V|T_1S_2X_2}^{(1)})$$
$$+\mathbb{E}_{\pi_{S_2X_2YV\mathbf{T}}}\big[D(\pi_{S_1'|S_2X_2YV\mathbf{T}}\|P_{S_1'|\mathbf{T}})\big]\Big]$$
$$-I(S_1;X_1|\mathbf{T})-I(S_2;X_2|\mathbf{T})-2\mu-\delta_{6,n}' \tag{349}$$

$$\overset{\text{(DP)}}{\geq} \min_{\pi\in\mathcal{P}_{\mu,\text{miss,1b}}} \Big[D(\pi_{S_2X_2YV\mathbf{T}}\|P_{S_2|\mathbf{T}}Q_{X_2Y}P_{\mathbf{T}}$$
$$\cdot\Gamma_{V|T_1S_2X_2}^{(1)})$$
$$+\mathbb{E}_{\pi_{S_2YV\mathbf{T}}}\big[D(\pi_{S_1'|S_2YV\mathbf{T}}\|P_{S_1'|\mathbf{T}})\big]\Big]$$
$$-I(S_1;X_1|\mathbf{T})-I(S_2;X_2|\mathbf{T})-2\mu-\delta_{6,n}' \tag{350}$$

$$\overset{(k)}{=} \min_{\pi \in \mathcal{P}_{\mu,\mathrm{miss,1b}}} D(\pi_{S_2 X_2 Y V \mathbf{T}} \| P_{S_2|\mathbf{T}} Q_{X_2 Y} P_{\mathbf{T}}$$
$$\cdot \Gamma^{(1)}_{V|T_1 S_2 X_2})$$
$$+ I(S_1; S_2, Y, V|\mathbf{T}) - I(S_1; X_1|\mathbf{T})$$
$$- I(S_2; X_2|\mathbf{T}) - \delta_{6,n}(\mu) \tag{351}$$
$$= \min_{\pi \in \mathcal{P}_{\mu,\mathrm{miss,1b}}} D(\pi_{S_2 X_2 Y V \mathbf{T}} \| P_{S_2|X_2 \mathbf{T}} Q_{X_2 Y} P_{\mathbf{T}}$$
$$\cdot \Gamma^{(1)}_{V|T_1 S_2 X_2})$$
$$+ I(S_1; S_2, Y, V|\mathbf{T}) - I(S_1; X_1|\mathbf{T}) - \delta_{6,n}(\mu) \tag{352}$$
$$\overset{(l)}{=} \min_{\pi \in \mathcal{P}_{\mu,\mathrm{miss,1b}}} D(\pi_{S_2 X_2 Y V \mathbf{T}} \| P_{S_2|X_2 \mathbf{T}} Q_{X_2 Y} P_{\mathbf{T}}$$
$$\cdot \Gamma^{(1)}_{V|T_1 S_2 X_2})$$
$$+ I(S_1; Y, V|S_2, \mathbf{T}) - I(S_1; X_1|S_2, \mathbf{T}) - \delta_{6,n}(\mu) \tag{353}$$
$$= \theta^{\mathrm{miss,1b}}_\mu - \delta_{6,n}(\mu), \tag{354}$$

where $\delta_{6,n}(\mu)$ is a sequence of functions that goes to zero as $n \to \infty$ and $\mu \to 0$ and

$$\theta^{\mathrm{miss,1b}}_\mu :=$$
$$\min_{\pi \in \mathcal{P}_{\mu,\mathrm{miss,1b}}} D(\pi_{S_2 X_2 Y V \mathbf{T}} \| P_{S_2|X_2 \mathbf{T}} Q_{X_2 Y} P_{\mathbf{T}} \Gamma^{(1)}_{V|T_1 S_2 X_2})$$
$$+ I(S_1; Y, V|S_2, \mathbf{T}) - I(S_1; X_1|S_2, \mathbf{T}). \tag{355}$$

Here, (k) holds because $|\pi_{S'_1 Y V \mathbf{T}} - P_{S_1 Y V \mathbf{T}}| \leq \mu$ and (l) holds because of the Markov chain $S_1 \to (X_1, \mathbf{T}) \to S_2$. From (337)–(354), we obtain

$$\Pr[\mathcal{B}_6|\mathcal{H} = 1] \leq 2^{-n(\theta^{\mathrm{miss,1b}}_\mu - \delta_{6,n}(\mu))}. \tag{356}$$

Following similar steps to above, one can show that

$$\Pr[\mathcal{B}_7|\mathcal{H} = 1] \leq 2^{-n(\theta^{\mathrm{miss,2a}}_\mu - \delta_{7,n}(\mu))}, \tag{357}$$
$$\Pr[\mathcal{B}_8|\mathcal{H} = 1] \leq 2^{-n(\theta^{\mathrm{miss,2b}}_\mu - \delta_{8,n}(\mu))}, \tag{358}$$

where $\delta_{7,n}(\mu)$ and $\delta_{8,n}(\mu)$ are sequences of functions that go to zero as $n \to \infty$ and $\mu \to 0$ and

$$\theta^{\mathrm{miss,2a}}_\mu :=$$
$$\min_{\pi \in \mathcal{P}_{\mu,\mathrm{miss,2a}}} D(\pi_{S_1 X_1 Y V \mathbf{T}} \| P_{S_1|X_1 \mathbf{T}} Q_{X_1 Y} P_{\mathbf{T}} \Gamma^{(2)}_{V|S_1 X_1 T_2})$$
$$+ I(S_1, S_2; Y, V|\mathbf{T}) - I(S_1, S_2; X_1, X_2|\mathbf{T}), \tag{359}$$
$$\theta^{\mathrm{miss,2b}}_\mu :=$$
$$\min_{\pi \in \mathcal{P}_{\mathrm{miss,2b}}} D(\pi_{S_1 X_1 Y V \mathbf{T}} \| P_{S_1|X_1 \mathbf{T}} Q_{X_1 Y} P_{\mathbf{T}} \Gamma^{(2)}_{V|S_1 X_1 T_2})$$
$$+ I(S_2; Y, V|S_1, \mathbf{T}) - I(S_2; X_2|S_1, \mathbf{T}). \tag{360}$$

Finally, the probability of event $\mathcal{B}_9$ can be upper bounded as:

$$\Pr[\mathcal{B}_9|\mathcal{H} = 1] \leq 2^{-n\tilde{\theta}^{\mathrm{miss,12}}_\mu}, \tag{361}$$

where

$$\tilde{\theta}^{\mathrm{miss,12}}_\mu := \min_{\substack{\pi: \\ |\pi_{S'_1 S'_2 Y V \mathbf{T}} - P_{S_1 S_2 Y V \mathbf{T}}| < \mu}}$$
$$D(\pi_{S'_1 S'_2 Y V \mathbf{T}} \| P_{S'_1|\mathbf{T}} P_{S'_2|\mathbf{T}} Q_Y P_{\mathbf{T}} \Gamma^{(12)}_{V|\mathbf{T}})$$
$$- R_1 - R_2 - \delta'_{9,n}, \tag{362}$$

for some sequence $\delta'_{9,n}$ that tends to 0 as $n \to \infty$. We have the following set of inequalities. By (88)

$$\tilde{\theta}^{\mathrm{miss,12}}_\mu = \min_{\substack{\pi: \\ |\pi_{S'_1 S'_2 Y V \mathbf{T}} - P_{S_1 S_2 Y V \mathbf{T}}| < \mu}}$$
$$D(\pi_{S'_1 S'_2 Y V \mathbf{T}} \| P_{S'_1|\mathbf{T}} P_{S'_2|\mathbf{T}} Q_Y P_{\mathbf{T}} \Gamma^{(12)}_{V|\mathbf{T}})$$
$$- I(S_1; X_1|\mathbf{T}) - I(S_2; X_2|\mathbf{T}) - 2\mu - \delta'_{9,n} \tag{363}$$

$$= \min_{\substack{\pi: \\ |\pi_{S'_1 S'_2 Y V \mathbf{T}} - P_{S_1 S_2 Y V \mathbf{T}}| < \mu}} \Big[D(\pi_{Y V \mathbf{T}} \| Q_Y P_{\mathbf{T}} \Gamma^{(12)}_{V|\mathbf{T}})$$
$$+ \mathbb{E}_{\pi_{Y V \mathbf{T}}} \big[D(\pi_{S'_1 S'_2|Y V \mathbf{T}} \| P_{S'_1|\mathbf{T}} P_{S'_2|\mathbf{T}}) \big] \Big]$$
$$- I(S_1; X_1|\mathbf{T}) - I(S_2; X_2|\mathbf{T}) - 2\mu - \delta'_{9,n} \tag{364}$$

$$\overset{(m)}{=} \mathbb{E}_{P_{\mathbf{T}}} \big[D(P_{Y V|\mathbf{T}} \| Q_Y \Gamma^{(12)}_{V|\mathbf{T}}) \big] + I(S_1, S_2; Y, V|\mathbf{T})$$
$$- I(S_1, S_2; X_1, X_2|\mathbf{T}) - \delta_{9,n}(\mu) \tag{365}$$

$$= \theta^{\mathrm{miss,12}}_\mu - \delta_{9,n}(\mu), \tag{366}$$

where $\delta_{9,n}(\mu)$ is a sequence of functions that goes to zero as $n \to \infty$ and $\mu \to 0$ and

$$\theta^{\mathrm{miss,12}}_\mu := \mathbb{E}_{P_{\mathbf{T}}} \big[D(P_{Y V|\mathbf{T}} \| Q_Y \Gamma^{(12)}_{V|\mathbf{T}}) \big]$$
$$+ I(S_1, S_2; Y, V|\mathbf{T}) I(S_1, S_2; X_1, X_2|\mathbf{T}). \tag{367}$$

Here, (m) holds because $|\pi_{Y V \mathbf{T}} - P_{Y V \mathbf{T}}| \leq \mu$ and $|\pi_{S'_1 S'_2 Y V \mathbf{T}} - P_{S_1 S_2 Y V \mathbf{T}}| \leq \mu$ and by the Markov chains $S_1 \to (X_1, \mathbf{T}) \to S_2$ and $S_2 \to (X_2, \mathbf{T}) \to S_1$. Combining (361)–(366) yields:

$$\Pr[\mathcal{B}_9|\mathcal{H} = 1] \leq 2^{-n(\theta^{\mathrm{miss,12}}_\mu - \delta_{9,n}(\mu))}. \tag{368}$$

Combining now (306), (321), (322), (332), (345), (356), (357), (358), and (368), for sufficiently large values of n and small values of μ, the average type-II error probability can be bounded as:

$$\mathbb{E}_{\mathcal{C}}[\beta_n] \leq 2 \max \Big\{ 2^{-n\theta^{\mathrm{standard}}_\mu}, 2^{-n\theta^{\mathrm{dec,1}}_\mu}, 2^{-n\theta^{\mathrm{dec,2}}_\mu}, 2^{-n\theta^{\mathrm{dec,12}}_\mu},$$
$$2^{-n\theta^{\mathrm{miss,1a}}_\mu}, 2^{-n\theta^{\mathrm{miss,1b}}_\mu}, 2^{-n\theta^{\mathrm{miss,2a}}_\mu}, 2^{-n\theta^{\mathrm{miss,2b}}_\mu}, 2^{-n\theta^{\mathrm{miss,12}}_\mu} \Big\}. \tag{369}$$

By standard arguments and successively eliminating the worst half of the codebooks, it can be shown that there exists at least one codebook for which:

$$\alpha_n \leq \epsilon, \tag{370}$$
$$\beta_n \leq 1024 \cdot \max \Big\{ 2^{-n\theta^{\mathrm{standard}}_\mu}, 2^{-n\theta^{\mathrm{dec,1}}_\mu}, 2^{-n\theta^{\mathrm{dec,2}}_\mu}, 2^{-n\theta^{\mathrm{dec,12}}_\mu},$$
$$2^{-n\theta^{\mathrm{miss,1a}}_\mu}, 2^{-n\theta^{\mathrm{miss,1b}}_\mu}, 2^{-n\theta^{\mathrm{miss,2a}}_\mu}, 2^{-n\theta^{\mathrm{miss,2b}}_\mu}, 2^{-n\theta^{\mathrm{miss,12}}_\mu} \Big\}. \tag{371}$$

Letting $n \to \infty$ and $\mu \to 0$ for $i \in \{1, 2\}$, we get $\theta^{\mathrm{standard}}_\mu \to \theta^{\mathrm{standard}}$, $\theta^{\mathrm{dec,i}}_\mu \to \theta^{\mathrm{dec,i}}$, $\theta^{\mathrm{dec,12}}_\mu \to \theta^{\mathrm{dec,12}}$, $\theta^{\mathrm{miss,ia}}_\mu \to \theta^{\mathrm{miss,ia}}$, $\theta^{\mathrm{miss,ib}}_\mu \to \theta^{\mathrm{miss,ib}}$ and $\theta^{\mathrm{miss,12}}_\mu \to \theta^{\mathrm{miss,12}}$, which concludes the proof of the theorem.

APPENDIX E
PROOF OF COROLLARY 2

Choose for each $i \in \{1, 2\}$ the random variable $T_i = t_i$ deterministically, and a (conditional) pmf $P_{S_i|X_i}$ and function $f_i : \mathcal{S}_i \times \mathcal{X}_i \to \mathcal{W}_i$ so that conditions (113) are satisfied. It can be noticed that for the present setup such a choice also satisfies the conditions (86). Therefore, by Theorem 4, any error exponent satisfying (107) for the chosen pmfs and functions is achievable. In the following, we simplify the conditions in (107) for our choice. For convenience of notation, we define the following conditional laws:

$$
\Gamma_{V|S_2Z}(v|s_2, z)
$$
$$
\triangleq \sum_{s_1, x_1, x_2} P_{S_1|X_1}(s_1|x_1) \cdot P_{S_2|X_2}(s_2|P_{X_1X_2|Z}(x_1, x_2|z)x_2) \cdot
$$
$$
\cdot \Gamma_{V|S_1S_2X_1X_2}(v|s_1, s_2, x_1, x_2), \tag{372}
$$

$$
\Gamma^{(1)}_{V|Z}(v|z) \triangleq \sum_{s_2, x_2} P_{S_2|X_2}(s_2|x_2) \cdot P_{X_2|Z}(x_2|z)
$$
$$
\cdot \Gamma^{(1)}_{V|T_1S_2X_2}(v|t_1, s_2, x_2), \tag{373}
$$

$$
\Gamma^{(1)}_{V|S_2Z}(v|z) \triangleq \sum_{x_2} P_{X_2|Z}(x_2|z) \cdot \Gamma^{(1)}_{V|T_1S_2X_2}(v|t_1, s_2, x_2), \tag{374}
$$

$$
\Gamma^{(2)}_{V|Z}(v|z) \triangleq \sum_{s_1, x_1} P_{S_1|X_1}(s_1|x_1) \cdot P_{X_1|Z}(x_1|z)
$$
$$
\cdot \Gamma^{(2)}_{V|S_1X_1T_2}(v|s_1, x_1, t_2), \tag{375}
$$

$$
\Gamma^{(2)}_{V|S_2Z}(v|z) \triangleq \sum_{x_1} P_{X_1|Z}(x_1|z) \Gamma^{(2)}_{V|S_1X_1T_2}(v|s_1, x_1, t_2). \tag{376}
$$

We start by simplifying the decoding-error exponents:

$$
\theta^{\mathrm{dec},1} := \min_{\substack{\tilde{P}_{S_1S_2X_1X_2YV}: \\ \tilde{P}_{S_iX_i} = P_{S_iX_i}, \, i \in \{1,2\}, \\ \tilde{P}_{S_2\bar{Y}ZV} = P_{S_2\bar{Y}ZV} \\ H(S_1|S_2,\bar{Y},Z,V) \le H_{\tilde{P}}(S_1|S_2,\bar{Y},Z,V)}} D\Big(\tilde{P}_{S_1S_2X_1X_2\bar{Y}ZV} \| P_{S_1|X_1} P_{S_2|X_2} P_{X_1X_2Z}
$$
$$
Q_{\bar{Y}|Z} \Gamma_{V|S_1S_2X_1X_2}\Big)
$$
$$
+ I(S_1; \bar{Y}, Z, V|S_2) - I(S_1; X_1|S_2) \tag{377}
$$

$$
\overset{(\mathrm{CR})}{=} \min_{\substack{\tilde{P}_{S_1S_2X_1X_2}: \\ \tilde{P}_{S_iX_i} = P_{S_iX_i}, \, i \in \{1,2\}, \\ \tilde{P}_{S_2\bar{Y}ZV} = P_{S_2\bar{Y}ZV} \\ H(S_1|S_2,\bar{Y},Z,V) \le H_{\tilde{P}}(S_1|S_2,\bar{Y},Z,V)}} \Big[D\Big(\tilde{P}_{S_1S_2X_1X_2\ Z} \| P_{S_1|X_1} P_{S_2|X_2} P_{X_1X_2Z}\Big)
$$
$$
+ \mathbb{E}_{\tilde{P}_{S_1S_2X_1X_2\ Z}} [D(\tilde{P}_{\bar{Y}V|S_1S_2X_1X_2\ Z} \| Q_{\bar{Y}|Z} \Gamma_{V|S_1S_2X_1X_2})] \Big]
$$
$$
+ I(S_1; \bar{Y}, Z, V|S_2) - I(S_1; X_1|S_2) \tag{378}
$$

$$
\overset{(\mathrm{DP})}{\ge} \min_{\substack{\tilde{P}_{S_1S_2X_1X_2}: \\ \tilde{P}_{S_iX_i} = P_{S_iX_i}, \, i \in \{1,2\}, \\ \tilde{P}_{S_2\bar{Y}ZV} = P_{S_2\bar{Y}Z} \\ H(S_1|S_2,\bar{Y},Z,V) \le H_{\tilde{P}}(S_1|S_2,\bar{Y},Z,V)}} \Big[D\Big(\tilde{P}_{S_1S_2X_1X_2\ Z} \| P_{S_1|X_1} P_{S_2|X_2} P_{X_1X_2Z}\Big)
$$
$$
+ \mathbb{E}_{\tilde{P}_{S_2\ Z}} [D(\tilde{P}_{\bar{Y}V|S_2\ Z} \| Q_{\bar{Y}|Z} \Gamma_{V|S_2Z})] \Big]
$$
$$
+ I(S_1; \bar{Y}, Z, V|S_2) - I(S_1; X_1|S_2) \tag{379}
$$

$$
\overset{(a)}{=} \min_{\substack{\tilde{P}_{S_1S_2X_1X_2}: \\ \tilde{P}_{S_iX_i} = P_{S_iX_i}, \, i \in \{1,2\}, \\ \tilde{P}_{S_2,\bar{Y},ZV} = P_{S_2\bar{Y}ZV} \\ H(S_1|S_2,\bar{Y},Z,V) \le H_{\tilde{P}}(S_1|S_2,\bar{Y},Z,V)}} D\Big(\tilde{P}_{S_1S_2X_1X_2\ Z} \| P_{S_1|X_1} P_{S_2|X_2} P_{X_1X_2Z}\Big)
$$
$$
+ \mathbb{E}_{P_{S_2\ ZV}} [D(P_{\bar{Y}|S_2\ ZV} \| Q_{\bar{Y}|Z})]
$$
$$
+ I(S_1; \bar{Y}, Z, V|S_2) - I(S_1; X_1|S_2) \tag{380}
$$

$$
\overset{(b)}{=} \mathbb{E}_{P_{S_2\ ZV}} [D(P_{\bar{Y}|S_2\ ZV} \| Q_{\bar{Y}|Z})]
$$
$$
+ I(S_1; \bar{Y}, Z, V|S_2) - I(S_1; X_1|S_2), \tag{381}
$$

where (a) holds by the second constraint in the minimization and (b) holds because KL-divergence is nonnegative and $\tilde{P}_{S_1S_2X_2Z} = P_{S_1|X_1} P_{S_2|X_2} P_{X_1X_2Z}$ is a valid choice in the minimization and because $P_{V|S_2Z} = \Gamma_{V|S_2Z}$.

Moreover, above inequality $\overset{(\mathrm{DP})}{\ge}$ holds with equality, because evaluating $D\big(\tilde{P}_{S_1S_2X_1X_2\bar{Y}ZV} \| P_{S_1|X_1} P_{S_2|X_2} P_{X_1X_2Z} Q_{\bar{Y}|Z} \Gamma_{V|S_1S_2X_1X_2}\big)$ for the choice

$$
\tilde{P}_{S_1S_2X_1X_2\bar{Y}ZV} = P_{S_1|X_1} P_{S_2|X_2} P_{X_1X_2Z} P_{\bar{Y}V|S_2Z} \tag{382}
$$

(which is a valid candidate for the minimization) results in the KL-divergence on the right-hand side of (381). So, we conclude that

$$
\theta^{\mathrm{dec},1} = \mathbb{E}_{P_{S_2\ ZV}} [D(P_{\bar{Y}|S_2\ ZV} \| Q_{\bar{Y}|Z})]
$$
$$
+ I(S_1; \bar{Y}, Z, V|S_2) - I(S_1; X_1|S_2), \tag{383a}
$$

and in an analogous way it can be shown that also

$$
\theta^{\mathrm{dec},2} = \mathbb{E}_{P_{S_1\ ZV}} [D(P_{\bar{Y}|S_1\ ZV} \| Q_{\bar{Y}|Z})]
$$
$$
+ I(S_2; \bar{Y}, Z, V|S_1) - I(S_2; X_1|S_1), \tag{383b}
$$

and

$$
\theta^{\mathrm{dec},12} = \mathbb{E}_{P_{ZV}} [D(P_{\bar{Y}|ZV} \| Q_{\bar{Y}|Z})]
$$
$$
+ I(S_1, S_2; \bar{Y}, Z, V) - I(S_1, S_2; X_1, X_2). \tag{383c}
$$

Following similar steps, we obtain:

$$
\theta^{\mathrm{miss},1a}
$$
$$
= \min_{\substack{\tilde{P}_{S_2X_2\bar{Y}ZV}: \\ \tilde{P}_{S_2X_2} = P_{S_2X_2} \\ \tilde{P}_{\bar{Y}ZV} = P_{\bar{Y}ZV} \\ H(S_2|\bar{Y},Z,V) \le H_{\tilde{P}}(S_2|\bar{Y},Z,V)}} D\big(\tilde{P}_{S_2X_2\bar{Y}ZV} \| P_{S_2|X_2} P_{X_2Z} Q_{\bar{Y}|Z} \Gamma^{(1)}_{V|T_1S_2X_2}\big)
$$
$$
+ I(S_1, S_2; V, \bar{Y}, Z) - I(S_1, S_2; X_1, X_2) \tag{384}
$$

(CR)&(DP)
$$\geq \min_{\substack{\tilde{P}_{S_2 X_2 \bar{Y} Z V}: \\ \tilde{P}_{S_2 X_2} = P_{S_2 X_2} \\ \tilde{P}_{\bar{Y} Z V} = P_{\bar{Y} Z V} \\ H(S_2|\bar{Y}, Z, V) \leq H_{\tilde{P}}(S_2|\bar{Y}, Z, V)}} \Big[D\big(\tilde{P}_{S_2 X_2 Z} \| P_{S_2|X_2} P_{X_2 Z}\big)$$

$$+ \mathbb{E}_{\tilde{P}_Z} \big[D\big(\tilde{P}_{\bar{Y} V|Z} \| Q_{\bar{Y}|Z} \Gamma_{V|Z}^{(1)}\big)\big]\Big]$$

$$+ I(S_1, S_2; V, \bar{Y}, Z) - I(S_1, S_2; X_1, X_2) \quad (385)$$

$$= \mathbb{E}_{P_Z} \big[D(P_{\bar{Y} V|Z} \| Q_{\bar{Y}|Z} \Gamma_{V|Z}^{(1)})\big]$$

$$+ I(S_1, S_2; V, \bar{Y}, Z) - I(S_1, S_2; X_1, X_2), \quad (386)$$

and

$\theta^{\text{miss},1b}$

$$= \min_{\substack{\tilde{P}_{S_2 X_2 \bar{Y} Z V}: \\ \tilde{P}_{S_2 X_2} = P_{S_2 X_2} \\ \tilde{P}_{S_2 \bar{Y} Z V} = P_{S_2 \bar{Y} Z V}}}$$

$$D\big(\tilde{P}_{S_2 X_2 \bar{Y} Z V} \| P_{S_2|X_2} P_{X_2 Z} Q_{\bar{Y}|Z} \Gamma_{V|T_1 S_2 X_2}^{(1)}\big)$$

$$+ I(S_1; V, \bar{Y}, Z|S_2) - I(S_1; X_1|S_2) \quad (387)$$

$$\overset{\text{(CR)\&(DP)}}{\geq} \min_{\substack{\tilde{P}_{S_2 X_2 \bar{Y} Z V}: \\ \tilde{P}_{S_2 X_2} = P_{S_2 X_2} \\ \tilde{P}_{S_2 \bar{Y} Z V} = P_{S_2 \bar{Y} Z V}}} \Big[D\big(\tilde{P}_{S_2 X_2 Z} \| P_{S_2|X_2} P_{X_2 Z}\big)$$

$$+ \mathbb{E}_{\tilde{P}_{S_2 Z}} \big[D\big(\tilde{P}_{\bar{Y} V|S_2 Z} \| Q_{\bar{Y}|Z} \Gamma_{V|S_2 Z}^{(1)}\big)\big]\Big]$$

$$+ I(S_1; V, \bar{Y}, Z|S_2) - I(S_1; X_1|S_2) \quad (388)$$

$$\geq \mathbb{E}_{P_{S_2 Z}} \big[D(P_{\bar{Y} V|S_2 Z} \| Q_{\bar{Y}|Z} \Gamma_{V|S_2 Z}^{(1)})\big]$$

$$+ I(S_1; V, \bar{Y}, Z|S_2) - I(S_1; X_1|S_2). \quad (389)$$

Therefore:

$$\theta^{\text{miss},1a} = \mathbb{E}_{P_Z} \big[D(P_{\bar{Y} V|Z} \| Q_{\bar{Y}|Z} \Gamma_{V|Z}^{(1)})\big]$$

$$+ I(S_1, S_2; V, \bar{Y}, Z) - I(S_1, S_2; X_1, X_2) \quad (390a)$$

$$\theta^{\text{miss},1b} \geq \mathbb{E}_{P_{S_2 Z}} \big[D(P_{\bar{Y} V|S_2 Z} \| Q_{\bar{Y}|Z} \Gamma_{V|S_2 Z}^{(1)})\big]$$

$$+ I(S_1; V, \bar{Y}, Z|S_2) - I(S_1; X_1|S_2). \quad (390b)$$

By similar arguments, also

$$\theta^{\text{miss},2a} = \mathbb{E}_{P_Z} \big[D(P_{\bar{Y} V|Z} \| Q_{\bar{Y}|Z} \Gamma_{V|Z}^{(2)})\big]$$

$$+ I(S_1, S_2; V, \bar{Y}, Z) - I(S_1, S_2; X_1, X_2) \quad (390c)$$

$$\theta^{\text{miss},2b} \geq \mathbb{E}_{P_{S_1 Z}} \big[D(P_{\bar{Y} V|S_1 Z} \| Q_{\bar{Y}|Z} \Gamma_{V|S_2 Z}^{(2)})\big]$$

$$+ I(S_2; V, \bar{Y}, Z|S_1) - I(S_2; X_1|S_1). \quad (390d)$$

Finally, it is straightforward to see that

$$\theta^{\text{miss},12} = \mathbb{E}_{P_Z} \big[D(P_{\bar{Y} V|Z} \| Q_{\bar{Y}|Z} \Gamma_V^{(12)})\big]$$

$$+ I(S_1, S_2; V, \bar{Y}, Z) - I(S_1, S_2; X_1, X_2). \quad (390e)$$

Comparing (383a) with (390b), by the nonnegativity and the chain rule of KL-divergence, we see that exponent $\theta^{\text{miss},1b}$

is redundant in view of exponent $\theta^{\text{dec},1}$. In the same way, exponent $\theta^{\text{miss},2b}$ is redundant in view of $\theta^{\text{dec},2}$ and the three exponents $\theta^{\text{miss},1a}, \theta^{\text{miss},2a}, \theta^{\text{miss},12}$ are redundant in view of $\theta^{\text{dec},12}$.

We thus conclude that in this example and for any choice of the pmfs $P_{S_1|X_1}$ and $P_{S_2|X_2}$ and functions f_1 and f_2 satisfying (113), the following exponents are achievable:

$$\theta \leq \max \min \{\theta^{\text{standard}}, \theta^{\text{dec},1}, \theta^{\text{dec},2}, \theta^{\text{dec},12}\}, \quad (391)$$

where $\theta^{\text{dec},1}, \theta^{\text{dec},2}, \theta^{\text{dec},12}$ are given in (383) and θ^{standard} can be simplified to:

$$\theta^{\text{standard}} = \mathbb{E}_{P_{S_1 S_2 Z V}} \big[D(P_{\bar{Y}|S_1 S_2 Z V} \| Q_{\bar{Y}|Z})\big]. \quad (392)$$

We next show that if the pmfs $P_{S_1|X_1}$ and $P_{S_2|X_2}$ and the functions f_1 and f_2 are chosen to satisfy inequalities (113), then the minimum in (391) is attained by θ^{standard}. In fact, we can write the following set of inequalities:

$$\theta^{\text{dec},1} = \mathbb{E}_{P_{S_2\ ZV}} \big[D(P_{\bar{Y}|S_2\ ZV} \| Q_{\bar{Y}|Z})$$

$$+ I(S_1; \bar{Y}, Z, V|S_2) - I(S_1; X_1|S_2) \quad (393)$$

$$= \mathbb{E}_{P_{S_1 S_2\ ZV}} \big[D(P_{\bar{Y}|S_1 S_2\ ZV} \| Q_{\bar{Y}|Z})$$

$$+ I(S_1; Z, V|S_2) - I(S_1; X_1|S_2) \quad (394)$$

$$\overset{(d)}{\geq} \mathbb{E}_{P_{S_1 S_2\ ZV}} \big[D(P_{\bar{Y}|S_1 S_2\ ZV} \| Q_{\bar{Y}|Z})\big] \quad (395)$$

$$= \theta^{\text{standard}}, \quad (396)$$

where (d) holds by (113a); and by similar arguments also

$$\theta^{\text{dec},2} \geq \theta^{\text{standard}} \quad (397)$$

$$\theta^{\text{dec},12} \geq \theta^{\text{standard}}. \quad (398)$$

This concludes the proof.

APPENDIX F
PROOF OF CONVERSE FOR THEOREM 5

All mutual informations are calculated with respect to the pmfs under $\mathcal{H} = 0$. Define $\bar{S}_{1,t} := (V_1^n, X_1^{t-1})$ and $\bar{S}_{2,t} := (V_2^n, X_2^{t-1})$ and notice the Markov chains $\bar{S}_{1,t} \to X_{1,t} \to \bar{S}_{2,t}$ and $\bar{S}_{2,t} \to X_{2,t} \to S_{1,t}$. Define $\delta(\epsilon) := H_{\text{b}}(\epsilon)/n/(1-\epsilon)$. Then, by [8]:

$$\theta \leq \frac{1}{n(1-\epsilon)} D(P_{V^n Y^n|\mathcal{H}=0} \| P_{V^n Y^n|\mathcal{H}=1}) + \delta(\epsilon) \quad (399)$$

$$= \frac{1}{n(1-\epsilon)} \mathbb{E}_{P_{Y^n}} \big[D(P_{V^n|Y^n, \mathcal{H}=0} \| P_{V^n|Y^n, \mathcal{H}=1})\big]$$

$$+ \frac{1}{1-\epsilon} \cdot D(P_Y \| Q_Y) + \delta(\epsilon) \quad (400)$$

$$= \frac{1}{n(1-\epsilon)} \mathbb{E}_{P_{Y^n}} \big[D(P_{V^n|Y^n, \mathcal{H}=0} \| P_{V^n|\mathcal{H}=1})\big]$$

$$+ \frac{1}{1-\epsilon} \cdot D(P_Y \| Q_Y) + \delta(\epsilon) \quad (401)$$

$$= \frac{1}{n(1-\epsilon)} I(V^n; Y^n) + \frac{1}{1-\epsilon} \cdot D(P_Y \| Q_Y) + \delta(\epsilon) \quad (402)$$

$$= \frac{1}{n(1-\epsilon)} \sum_{t=1}^{n} I(V^n, Y^{t-1}; Y_t)$$

$$+ \frac{1}{1-\epsilon} \cdot D(P_Y \| Q_Y) + \delta(\epsilon) \quad (403)$$

$$= \frac{1}{n(1-\epsilon)} \sum_{t=1}^{n} I(V^n, Y^{t-1}; Y_t)$$

$$+ \frac{1}{1-\epsilon} \cdot D(P_Y \| Q_Y) + \delta(\epsilon) \tag{404}$$

$$\overset{(a)}{\leq} \frac{1}{n(1-\epsilon)} \sum_{t=1}^{n} I(V^n, X_1^{t-1}, X_2^{t-1}; Y_t)$$

$$+ \frac{1}{1-\epsilon} \cdot D(P_Y \| Q_Y) + \delta(\epsilon) \tag{405}$$

$$= \frac{1}{n(1-\epsilon)} \sum_{t=1}^{n} I(\bar{S}_{1,t}, \bar{S}_{2,t}; Y_t)$$

$$+ \frac{1}{1-\epsilon} \cdot D(P_Y \| Q_Y) + \delta(\epsilon), \tag{406}$$

$$= \frac{1}{1-\epsilon} I(\bar{S}_1, \bar{S}_2; Y) + \frac{1}{1-\epsilon} \cdot D(P_Y \| Q_Y) + \delta(\epsilon), \tag{407}$$

where (a) follows from the Markov chain $Y^{t-1} \to (V^n, X_1^{t-1}, X_2^{t-1}) \to Y_t$. The last equality holds by defining a time-sharing random variable Q that is uniform over $\{1, \ldots, n\}$ and $\bar{S}_i := (Q, V_i^n, X_i^{Q-1})$, for $i \in \{1, 2\}$, and $Y := Y_Q$.

Next, consider the following term,

$$I(X_1^n; V_1^n | V_2^n)$$

$$= \sum_{t=1}^{n} I(X_{1,t}; V_1^n | X_1^{t-1}, V_2^n) \tag{408}$$

$$\overset{(b)}{=} \sum_{t=1}^{n} I(X_{1,t}; X_1^{t-1}, V_1^n | V_2^n) \tag{409}$$

$$\overset{(c)}{=} \sum_{t=1}^{n} I(X_{1,t}; X_1^{t-1}, V_1^n, X_2^{t-1} | V_2^n) \tag{410}$$

$$\geq \sum_{t=1}^{n} I(X_{1,t}; X_1^{t-1}, V_1^n | X_2^{t-1}, V_2^n) \tag{411}$$

$$= \sum_{t=1}^{n} I(X_{1,t}; \bar{S}_{1,t} | \bar{S}_{2,t}) \tag{412}$$

$$= nI(X_1; \bar{S}_1 | \bar{S}_2) \tag{413}$$

where (b) and (c) follow from the Markov chains $X_{1,t} \to V_2^n \to X_1^{t-1}$ and $X_{1,t} \to (V_1^n, V_2^n, X_1^{t-1}) \to X_2^{t-1}$, respectively. Both Markov chains hold because X_1^n and X_2^n are independent under both hypotheses and by the orthogonality of the MAC. The last equality holds by defining $X_i := (Q, X_{i,Q})$, for $i \in \{1, 2\}$. Notice that $\bar{S}_i \to X_i \to S_i$.

Similarly, we get

$$I(X_2^n; V_2^n | V_1^n) \geq nI(X_2; \bar{S}_2 | \bar{S}_1), \tag{414}$$

$$I(X_1^n, X_2^n; V_1^n, V_2^n) \geq nI(X_1, X_2; \bar{S}_1, \bar{S}_2). \tag{415}$$

On the other hand, we have

$$I(X_1^n; V_1^n | V_2^n)$$

$$\leq I(W_1^n; V_1^n | V_2^n) \tag{416}$$

$$= H(V_1^n | V_2^n) - H(V_1^n | W_1^n, V_2^n) \tag{417}$$

$$\leq H(V_1^n) - H(V_1^n | W_1^n, V_2^n) \tag{418}$$

$$\overset{(d)}{=} H(V_1^n) - H(V_1^n | W_1^n) \tag{419}$$

$$= I(W_1^n; V_1^n) \tag{420}$$

$$\leq \sum_{t=1}^{n} I(W_{1,t}; V_{1,t}) \tag{421}$$

$$= nI(W_1; V_1) \tag{422}$$

$$\leq nC_1, \tag{423}$$

where (d) follows from the Markov chain $V_1^n \to W_1^n \to V_2^n$ and the orthogonality assumption. The last equality holds by defining $W_i := (Q, W_{i,Q})$ and $V_i = V_{i,Q}$ for $i \in \{1, 2\}$. Similarly, we have

$$I(X_2^n; V_2^n | V_1^n) \leq nC_2, \tag{424}$$

$$I(X_2^n, X_1^n; V_1^n, V_2^n) \leq nC_1 + nC_2. \tag{425}$$

Appropriately combining the derived inequalities concludes the proof of the converse.

APPENDIX G
CONVERSE PROOF FOR PROPOSITION 2

The proof follows similar steps to [26]. Define $\delta(\epsilon) := H_b(\epsilon)/n/(1-\epsilon)$. Then, by [8]:

$$\theta \leq \frac{1}{n(1-\epsilon)} D(P_{V^n Y^n | \mathcal{H}=0} \| P_{V^n Y^n | \mathcal{H}=1}) + \delta(\epsilon) \tag{426}$$

$$= \frac{1}{n(1-\epsilon)} \mathbb{E}_{P_{Y^n}} \left[D(P_{V^n | Y^n, \mathcal{H}=0} \| P_{V^n | Y^n, \mathcal{H}=1}) \right]$$

$$+ \frac{1}{1-\epsilon} \cdot D(P_Y \| Q_Y) + \delta(\epsilon) \tag{427}$$

$$= \frac{1}{n(1-\epsilon)} \mathbb{E}_{P_{Y^n}} \left[D(P_{V^n | Y^n, \mathcal{H}=0} \| P_{V^n | \mathcal{H}=1}) \right]$$

$$+ \frac{1}{1-\epsilon} \cdot D(P_Y \| Q_Y) + \delta(\epsilon) \tag{428}$$

$$= \frac{1}{n(1-\epsilon)} I(V^n; Y^n) + \frac{1}{1-\epsilon} \cdot D(P_Y \| Q_Y) + \delta(\epsilon) \tag{429}$$

$$= \frac{1}{n(1-\epsilon)} I(V_1^n, V_2^n; Y^n) + \frac{1}{1-\epsilon} \cdot D(P_Y \| Q_Y) + \delta(\epsilon). \tag{430}$$

Next, consider the following set of inequalities:

$$I(X_1^n; V_1^n | V_2^n) \leq I(W_1^n; V_1^n | V_2^n) \tag{431}$$

$$= H(V_1^n | V_2^n) - H(V_1^n | W_1^n, V_2^n) \tag{432}$$

$$\leq H(V_1^n) - H(V_1^n | W_1^n, V_2^n) \tag{433}$$

$$\overset{(a)}{=} H(V_1^n) - H(V_1^n | W_1^n) \tag{434}$$

$$= I(W_1^n; V_1^n) \tag{435}$$

$$\leq \sum_{t=1}^{n} I(W_{1,t}; V_{1,t}) \tag{436}$$

$$= nI(W_1; V_1) \tag{437}$$

$$\leq nC_1, \tag{438}$$

where (a) follows from the Markov chain $V_2^n \to W_1^n \to V_1^n$. Similarly, we have

$$I(X_2^n; V_2^n | V_1^n) \leq nC_2, \tag{439}$$

$$I(X_1^n, X_2^n; V_1^n, V_2^n) \leq n(C_1 + C_2). \tag{440}$$

Defining the auxiliaries S_1^n and S_2^n to be V_1^n and V_2^n, respectively, considering the Markov chains $V_1^n \to X_1^n \to V_2^n$,

$V_2^n \to X_2^n \to V_1^n$ and letting $\epsilon \to 0$ completes the proof of the theorem.

APPENDIX H
PROOF OF EXAMPLE 3

We simplify the result of Theorem 5 for the proposed Gaussian setup. Notice first that since X_1 and X_2 are independent and because of the Markov chains $\bar{S}_1 \to X_1 \to X_2$ and $\bar{S}_2 \to X_2 \to X_1$, the pair $(X_1, \bar{S}_1)$ is independent of $(X_2, \bar{S}_2)$. As a consequence,

$$I(\bar{S}_1; X_1 | \bar{S}_2) = I(\bar{S}_1; X_1) \tag{441}$$

$$I(\bar{S}_2; X_2 | \bar{S}_1) = I(\bar{S}_2; X_2) \tag{442}$$

$$I(\bar{S}_1, \bar{S}_2; X_1, X_2) = I(\bar{S}_1; X_1) + I(\bar{S}_2; X_2), \tag{443}$$

and the three constraints in the maximization of (123) simplify to the two constraints:

$$I(\bar{S}_1; X_1) \le C_1, \tag{444a}$$

$$I(\bar{S}_2; X_2) \le C_2. \tag{444b}$$

Choose now the auxiliary random variables $\bar{S}_1$ and $\bar{S}_2$ as

$$\bar{S}_i = X_i + F_i, \quad F_i \sim \mathcal{N}(0, \xi_i^2), \quad i \in \{1, 2\}, \tag{445}$$

where

$$\xi_i^2 := \frac{1}{2^{2C_i} - 1}, \quad i \in \{1, 2\}. \tag{446}$$

It is easily checked that this choice satisfies constraints (444). Moreover, the mutual information term in the achievable error exponent evaluates to:

$$I(\bar{S}_1, \bar{S}_2; Y) = \frac{1}{2} \log_2 \frac{2 + \sigma_0^2}{\sigma_0^2 + \frac{\xi_1^2}{1+\xi_1^2} + \frac{\xi_2^2}{1+\xi_2^2}} \tag{447}$$

$$= \frac{1}{2} \log_2 \frac{2 + \sigma_0^2}{2^{-2C_1} + 2^{-2C_2} + \sigma_0^2}, \tag{448}$$

and the KL-divergence term to:

$$D(P_Y \| Q_Y)$$

$$= -h(Y) + \mathbb{E}_{P_Y}\left[\log_2 \frac{1}{Q_Y}\right] \tag{449}$$

$$= -h(Y) + \mathbb{E}_{P_Y}\left[\log_2 \left(\sqrt{2\pi\sigma_y^2} e^{\frac{Y^2}{2\sigma_y^2}}\right)\right] \tag{450}$$

$$= -h(Y) + \frac{1}{2}\log_2\left(2\pi\sigma_y^2\right) + \mathbb{E}_{P_Y}\left[\frac{Y^2}{2\sigma_y^2}\right] \cdot \log_2 e \tag{451}$$

$$= -h(Y) + \frac{1}{2}\log_2\left(2\pi\sigma_y^2\right) + \frac{2 + \sigma_0^2}{2\sigma_y^2} \cdot \log_2 e \tag{452}$$

$$= -\frac{1}{2}\log_2\left(2\pi e(2 + \sigma_0^2)\right) + \frac{1}{2}\log_2\left(2\pi\sigma_y^2\right)$$
$$+ \frac{2 + \sigma_0^2}{2\sigma_y^2} \cdot \log_2 e \tag{453}$$

$$= \frac{1}{2}\log_2\left(\frac{\sigma_y^2}{2 + \sigma_0^2}\right) + \left(\frac{2 + \sigma_0^2}{2\sigma_y^2} - \frac{1}{2}\right) \cdot \log_2 e. \tag{454}$$

Combining (448) and (454), by Theorem 5, any error exponent $\theta \ge 0$ is achievable if it satisfies:

$$\theta \le \frac{1}{2}\log_2 \frac{\sigma_y^2}{2^{-2C_1} + 2^{-2C_2} + \sigma_0^2} + \left(\frac{2 + \sigma_0^2}{2\sigma_y^2} - \frac{1}{2}\right) \cdot \log_2 e. \tag{455}$$

We now show that by Theorem 5 no larger exponent is achievable. Notice first that since each X_i is standard Gaussian, constraints (444) are equivalent to

$$2^{2h(X_i | \bar{S}_i)} \ge 2\pi e \cdot 2^{-2C_i}, \quad i \in \{1, 2\}. \tag{456}$$

Then, by Theorem 5, any exponent has to satisfy:

$$\theta \le D(P_Y \| Q_Y) + \max_{\substack{\bar{S}_1, \bar{S}_2 \\ \text{s.t. (456)}}} I(\bar{S}_1, \bar{S}_2; Y) \tag{457}$$

$$= D(P_Y \| Q_Y) + h(Y) - \min_{\substack{\bar{S}_1, \bar{S}_2 \\ \text{s.t. (456)}}} h(Y | \bar{S}_1, \bar{S}_2) \tag{458}$$

$$= D(P_Y \| Q_Y) + h(Y) - \min_{\substack{\bar{S}_1, \bar{S}_2 \\ \text{s.t. (456)}}} h(Y | \bar{S}_1, \bar{S}_2) \tag{459}$$

$$\stackrel{(a)}{\le} D(P_Y \| Q_Y) + h(Y)$$
$$- \min_{\substack{\bar{S}_1, \bar{S}_2 \\ \text{s.t. (456)}}} \frac{1}{2}\log_2 \left(2\pi e\left(\frac{1}{2\pi e} 2^{2h(X_1 | \bar{S}_1, \bar{S}_2)}\right.\right.$$
$$+ \frac{1}{2\pi e} 2^{2h(X_2 | \bar{S}_1, \bar{S}_2)}$$
$$\left.\left.+ \frac{1}{2\pi e} 2^{2h(N_0 | \bar{S}_1, \bar{S}_2)}\right)\right) \tag{460}$$

$$\stackrel{(b)}{=} D(P_Y \| Q_Y) + h(Y)$$
$$- \min_{\substack{\bar{S}_1, \bar{S}_2 \\ \text{s.t. (456)}}} \frac{1}{2}\log_2 \left(2\pi e\left(\frac{1}{2\pi e} 2^{2h(X_1 | \bar{S}_1)}\right.\right.$$
$$\left.\left.+ \frac{1}{2\pi e} 2^{2h(X_2 | \bar{S}_2)} + \sigma_0^2\right)\right) \tag{461}$$

$$\stackrel{(c)}{\le} D(P_Y \| Q_Y) + h(Y)$$
$$- \frac{1}{2}\log_2 \left(2\pi e\left(2^{-2C_1} + 2^{-2C_2} + \sigma_0^2\right)\right) \tag{462}$$

$$= \frac{1}{2}\log_2 \left(\frac{\sigma_y^2}{2^{-2C_1} + 2^{-2C_2} + \sigma_0^2}\right)$$
$$+ \left(\frac{2 + \sigma_0^2}{2\sigma_y^2} - \frac{1}{2}\right) \cdot \log_2 e, \tag{463}$$

where (a) follows from the conditional EPI and the fact that given $(\bar{S}_1, \bar{S}_2)$, the three random variables X_1, X_2, and N_0 are independent; (b) follows because X_1 is independent of $\bar{S}_2$, X_2 is independent of $\bar{S}_1$ and N_0 is independent of both $(\bar{S}_1, \bar{S}_2)$; and (c) follows by (456). This concludes the proof.

APPENDIX I
PROOF OF COROLLARY 4

We evaluate the exponent in Corollary 2 for the following choice of *Gaussian* auxiliary random variables.

Let F_1, F_2, G_1, G_2 be independent zero-mean Gaussian random variables of variances $\xi^2, \xi^2, \gamma^2, \gamma^2$ and independent of the source variables (X_1, X_2, Y). Then, define

$$\bar{S}_i := X_i + G_i, \qquad i \in \{1, 2\}, \tag{464}$$

and

$$S_i = (\bar{S}_i, F_i), \qquad i \in \{1, 2\}. \tag{465}$$

We apply hybrid coding with channel inputs:

$$W_i = \alpha X_i + \beta G_i + F_i, \tag{466}$$

for some real numbers α and β such that

$$\gamma^2 + \alpha^2 + \beta^2 \cdot \xi^2 = P. \tag{467}$$

We first investigate for which parameters $\alpha, \beta, \gamma, \xi$, the presented choice of random variables satisfies the three constraints in the corollary. Notice first that:

$$I(S_1; V | S_2)$$
$$= \frac{1}{2} \log_2 \left(\sigma^2 + 2P - \gamma^2 + 2\alpha^2 \rho - \frac{(\alpha \cdot (1 + \rho) + \beta \cdot \xi^2)^2}{1 + \xi^2} \right)$$
$$- \frac{1}{2} \log_2 \left(\sigma^2 + \frac{2(\alpha - \beta)^2 \cdot (1 + \rho)\xi^2}{1 + \rho + \xi^2} \right), \tag{468a}$$

$$I(S_2; V | S_1)$$
$$= \frac{1}{2} \log_2 \left(\sigma^2 + 2P + 2\alpha^2 \rho - \frac{(\alpha \cdot (1 + \rho) + \beta \cdot \xi^2)^2}{1 + \xi^2} \right)$$
$$- \frac{1}{2} \log_2 \left(\sigma^2 + \frac{2(\alpha - \beta)^2 \cdot (1 + \rho)\xi^2}{1 + \rho + \xi^2} \right), \tag{468b}$$

and

$$I(S_1, S_2; V) = \frac{1}{2} \log_2 \left(\frac{\sigma^2 + 2P + 2\alpha^2 \rho}{\sigma^2 + \frac{2(\alpha-\beta)^2 \cdot (1+\rho)\xi^2}{1+\rho+\xi^2}} \right). \tag{468c}$$

Moreover,

$$I(\bar{S}_1; X_1 | \bar{S}_2) = \frac{1}{2} \log_2 \left(\frac{(1 + \xi^2)^2 - \rho^2}{(1 + \xi^2)\xi^2} \right), \tag{469a}$$

$$I(\bar{S}_2; X_2 | \bar{S}_1) = \frac{1}{2} \log_2 \left(\frac{(1 + \xi^2)^2 - \rho^2}{(1 + \xi^2)\xi^2} \right), \tag{469b}$$

and

$$I(\bar{S}_1, \bar{S}_2; X_1, X_2) = \frac{1}{2} \log_2 \left(\frac{(1 + \xi^2)^2 - \rho^2}{\xi^4} \right). \tag{469c}$$

Combining (468) and (469), shows that the presented choice of auxiliaries satisfies the three constraints (113a)–(113c) in Corollary 2, whenever the two conditions (138) are satisfied.

We now evaluate the error exponent (136) for the proposed choice of auxiliaries. To this end, notice that

$$\mathbb{E}_{P_V} \left[D(P_{Y|V} \| Q_Y) \right] + I(S_1, S_2; Y | V)$$
$$= D(P_Y \| Q_Y) + I(V; Y) + I(S_1, S_2; Y) \tag{470}$$
$$= D(P_Y \| Q_Y) + I(S_1, S_2, V; Y). \tag{471}$$

Moreover,

$$I(S_1, S_2, V; Y)$$
$$= \frac{1}{2} \log_2 \left(\sigma_0^2 + 2 + 2\rho \right)$$
$$- \frac{1}{2} \log_2 \left(\sigma_0^2 + \frac{2\xi^2(1 + \rho)\sigma^2}{2\xi^2(\alpha - \beta)^2 \cdot (1 + \rho) + \sigma^2(1 + \rho + \xi^2)} \right) \tag{472}$$

and (by similar steps as in (454)):

$$D(P_Y \| Q_Y)$$
$$= -h(Y) + \mathbb{E}_{P_Y} \left[\log_2 \frac{1}{Q_Y} \right] \tag{473}$$
$$= \frac{1}{2} \log_2 \left(\frac{\sigma_y^2}{2 + 2\rho + \sigma_0^2} \right) + \left(\frac{2 + 2\rho + \sigma_0^2}{2\sigma_y^2} - \frac{1}{2} \right) \cdot \log_2 e. \tag{474}$$

Combining (472) and (474) yields the error exponent in (136). This concludes the proof.

APPENDIX J
PROOF OF THEOREM 6

Fix a blocklength n and encoding and decoding/testing functions. Then, notice that by Witsenhausen's max-correlation argument [20], see also [30],

$$\frac{1}{2} \log_2 \left(1 + \frac{2P(1 + \rho)}{\sigma^2} \right)$$
$$\geq \frac{1}{n} I(W_1^n, W_2^n; V^n) \tag{475}$$
$$\overset{(a)}{\geq} \frac{1}{n} I(X_1^n, X_2^n; V^n) \tag{476}$$
$$= \frac{1}{n} h(X_1^n, X_2^n) - \frac{1}{n} h(X_1^n, X_2^n | V^n) \tag{477}$$
$$\overset{(b)}{=} \frac{1}{n} h(X_1^n, X_2^n) - \frac{1}{n} h(X_1^n + X_2^n, X_1^n - X_2^n | V^n) + 1 \tag{478}$$
$$= \frac{1}{n} h(X_1^n, X_2^n) - \frac{1}{n} h(X_1^n + X_2^n | V^n)$$
$$\quad - \frac{1}{n} h(X_1^n - X_2^n | X_1^n + X_2^n, V^n) + 1 \tag{479}$$
$$\overset{(c)}{\geq} \frac{1}{n} h(X_1^n, X_2^n) - \frac{1}{n} h(X_1^n + X_2^n | V^n)$$
$$\quad - \frac{1}{n} h(X_1^n - X_2^n) + 1 \tag{480}$$
$$= \frac{1}{2} \log_2 \left((2\pi e) \cdot (2 + 2\rho) \right) - \frac{1}{n} h(X_1^n + X_2^n | V^n), \tag{481}$$

where (a) holds by the Markov chain $(X_1^n, X_2^n) \to (W_1^n, W_2^n) \to V^n$; (b) holds because for each t the vector $(X_{1,t} + X_{2,t}, X_{1,t} - X_{2,t})$ is obtained from (X_1, X_2) by rotating it with the matrix

$$A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \tag{482}$$

and because for any bivariate vector $\mathbf{X}$ differential entropy satisfies $h(A\mathbf{X}) = h(\mathbf{X}) + \log |A| = h(\mathbf{X}) + 1$; and (c) holds because conditioning cannot increase differential entropy. Inequality (481) is equivalent to:

$$2^{\frac{2}{n} h(X_1^n + X_2^n | V^n)} \geq 2\pi e \cdot \frac{2(1 + \rho)\sigma^2}{2P(1 + \rho) + \sigma^2}. \tag{483}$$

We proceed to upper bound the error exponent. Define $\delta(\epsilon) := H(\epsilon)/n/(1-\epsilon)$. Then,

$$\theta \leq \frac{1}{n(1-\epsilon)} D(P_{V^n Y^n | \mathcal{H}=0} \| P_{V^n Y^n | \mathcal{H}=1}) + \delta(\epsilon) \quad (484)$$

$$= \frac{1}{1-\epsilon} \cdot D(P_Y \| Q_Y)$$
$$+ \frac{1}{n(1-\epsilon)} \mathbb{E}_{P_{Y^n}} \left[D(P_{V^n | Y^n, \mathcal{H}=0} \| P_{V^n | Y^n, \mathcal{H}=1}) \right] + \delta(\epsilon) \quad (485)$$

$$= \frac{1}{1-\epsilon} \cdot D(P_Y \| Q_Y)$$
$$+ \frac{1}{n(1-\epsilon)} \mathbb{E}_{P_{Y^n}} \left[D(P_{V^n | Y^n, \mathcal{H}=0} \| P_{V^n | \mathcal{H}=1}) \right] + \delta(\epsilon) \quad (486)$$

$$= \frac{1}{1-\epsilon} \cdot D(P_Y \| Q_Y) + \frac{1}{n(1-\epsilon)} I(V^n; Y^n) + \delta(\epsilon) \quad (487)$$

$$= \frac{1}{1-\epsilon} \cdot D(P_Y \| Q_Y)$$
$$+ \frac{1}{n(1-\epsilon)} \left[h(Y^n) - h(Y^n | V^n) \right] + \delta(\epsilon) \quad (488)$$

$$= \frac{1}{1-\epsilon} \cdot \left[D(P_Y \| Q_Y) + h(Y) \right]$$
$$- \frac{1}{n(1-\epsilon)} h(Y^n | V^n) + \delta(\epsilon) \quad (489)$$

$$\overset{(d)}{=} \frac{1}{1-\epsilon} \cdot \left[D(P_Y \| Q_Y) + h(Y) \right]$$
$$- \frac{1}{n(1-\epsilon)} h(X_1^n + X_2^n + N_0^n | V^n) + \delta(\epsilon) \quad (490)$$

$$\overset{(e)}{\leq} \frac{1}{1-\epsilon} \cdot \left[D(P_Y \| Q_Y) + h(Y) \right]$$
$$- \frac{1}{2(1-\epsilon)} \log_2 \left(2\pi e \left(\frac{1}{2\pi e} 2^{\frac{2}{n} h(X_1^n + X_2^n | V^n)} + \sigma_0^2 \right) \right)$$
$$+ \delta(\epsilon) \quad (491)$$

$$\overset{(f)}{\leq} \frac{1}{1-\epsilon} \cdot \left[D(P_Y \| Q_Y) + h(Y) \right]$$
$$- \frac{1}{2(1-\epsilon)} \log_2 \left(2\pi e \left(\frac{2(1+\rho)\sigma^2}{2P(1+\rho) + \sigma^2} + \sigma_0^2 \right) \right) + \delta(\epsilon) \quad (492)$$

$$\overset{(g)}{=} \frac{1}{2(1-\epsilon)} \cdot \log_2 \left(\frac{\sigma_y^2}{\frac{2(1+\rho)\sigma^2}{2P(1+\rho)+\sigma^2} + \sigma_0^2} \right)$$
$$+ \frac{1}{2(1-\epsilon)} \cdot \left(\frac{2 + 2\rho + \sigma_0^2}{\sigma_y^2} - 1 \right) \cdot \log_2 e + \delta(\epsilon), \quad (493)$$

where (d) follows from the definition of Y^n in (133); (e) follows from the conditional EPI and noting that for given V^n, the two random variables N_0^n and $X_1^n + X_2^n$ are independent; (f) follows from (483); (g) follows from (474). The proof is concluded by letting $\epsilon \to 0$.

REFERENCES

[1] R. Ahlswede and I. Csiszàr, "Hypothesis testing with communication constraints," *IEEE Trans. Inf. Theory*, vol. IT-32, no. 4, pp. 533–542, Jul. 1986.

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

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

[4] N. Weinberger and Y. Kochman, "On the reliability function of distributed hypothesis testing under optimal detection," *IEEE Trans. Inf. Theory*, vol. 65, no. 8, pp. 4940–4965, Aug. 2019.

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

[6] W. Zhao and L. Lai, "Distributed testing with cascaded encoders," *IEEE Trans. Inf. Theory*, vol. 64, no. 11, pp. 7339–7348, Nov. 2018.

[7] Y. Ugur, I. E. Aguerri, and A. Zaidi, "Vector Gaussian CEO problem under logarithmic loss and applications," *IEEE Trans. Inf. Theory*, early access, Feb. 7, 2020, doi: 10.1109/TIT.2020.2972348.

[8] Y. Xiang and Y.-H. Kim, "Interactive hypothesis testing against independence," in *Proc. IEEE Int. Symp. Inf. Theory*, Istanbul, Turkey, Jul. 2013, pp. 2840–2844.

[9] G. Katz, P. Piantanida, and M. Debbah, "Collaborative distributed hypothesis testing with general hypotheses," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2016, pp. 1705–1709.

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

[11] S. Salehkalaibar, M. Wigger, and L. Wang, "Hypothesis testing over the two-hop relay network," *IEEE Trans. Inf. Theory*, vol. 65, no. 7, pp. 4411–4433, Jul. 2019.

[12] S. Sreekumar and D. Gündüz, "Distributed hypothesis testing over discrete memoryless channels," *IEEE Trans. Inf. Theory*, vol. 66, no. 4, pp. 2044–2066, Apr. 2020.

[13] S. Salehkalaibar and M. Wigger, "Distributed hypothesis testing over a noisy channel," in *Proc. Int. Zurich Seminar (IZS)*, Zurich, Switzerland, Feb. 2018, pp. 25–29.

[14] S. Salehkalaibar and M. Wigger, "Distributed hypothesis testing over multi-access channels," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Abu Dhabi, United Arab Emirates, Dec. 2018, pp. 1–6.

[15] S. Borade, B. Nakiboglu, and L. Zheng, "Unequal error protection: An information-theoretic perspective," *IEEE Trans. Inf. Theory*, vol. 55, no. 12, pp. 5511–5539, Dec. 2009.

[16] D. Wang, V. Chandar, S.-Y. Chung, and G. W. Wornell, "On reliability functions for single-message unequal error protection," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2012, pp. 2934–2938.

[17] E. Tuncel, "Slepian-wolf coding over broadcast channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1469–1482, Apr. 2006.

[18] S. Sreekumar and D. Gündüz, "Distributed hypothesis testing over noisy channels," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2017, pp. 983–987.

[19] P. Minero, S. H. Lim, and Y.-H. Kim, "A unified approach to hybrid coding," *IEEE Trans. Inf. Theory*, vol. 61, no. 4, pp. 1509–1523, Apr. 2015.

[20] H. S. Witsenhausen, "On sequences of pairs of dependent random variables," *SIAM J. Appl. Math.*, vol. 28, no. 1, pp. 100–113, Jan. 1975.

[21] A. El Gamal and Y. H. Kim, *Network Information Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[22] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[23] E. Haim and Y. Kochman, "Binary distributed hypothesis testing via Körner-Marton coding," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Sep. 2016, pp. 146–150.

[24] B. Nazer and M. Gastpar, "Computation over multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3498–3516, Oct. 2007.

[25] J. Körner and K. Marton, "How to encode the modulo-two sum of binary sources (corresp.)," *IEEE Trans. Inf. Theory*, vol. IT-25, no. 2, pp. 219–221, Mar. 1979.

[26] J. J. Xiao and Z. Q. Luo, "Multiterminal source–channel communication over an orthogonal multiple-access channel," *IEEE Trans. Inf. Theory*, vol. 53, no. 9, pp. 3255–3264, Sep. 2007.

[27] E. Köken and E. Tuncel, "Joint source–channel coding for broadcasting correlated sources," *IEEE Trans. Commun.*, vol. 65, no. 7, pp. 3012–3022, Jul. 2017.

[28] H. Behroozi, F. Alajaji, and T. Linder, "Hybrid digital-analog joint source-channel coding for broadcasting correlated Gaussian sources," in *Proc. IEEE Int. Symp. Inf. Theory*, Seoul, South Korea, Jun. 2009, pp. 2785–2789.

[29] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Hoboken, NJ, USA: Wiley, 2006.

[30] A. Lapidoth and S. Tinguely, "Sending a bivariate Gaussian over a Gaussian multiple access channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2714–2752, Jun. 2010.

Sadaf Salehkalaibar (Member, IEEE) received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from the Sharif University of Technology, Tehran, Iran, in 2008, 2010, and 2014, respectively. She was a Post-Doctoral Fellow with Telecom ParisTech, Paris, France, in 2015 and 2017. She is currently an Assistant Professor with the Department of Electrical and Computer Engineering, University of Tehran, Tehran. Her special fields of interest include network information theory, hypothesis testing, and fundamental limits of secure communication with emphasis on information-theoretic security.

Michèle Wigger (Senior Member, IEEE) received the M.Sc. (Hons.) and Ph.D. degrees in electrical engineering from ETH Zurich in 2003 and 2008, respectively. In 2009, she was a Post-Doctoral Fellow with the University of California at San Diego, USA, and then joined Telecom Paris, France, where she is currently a Full Professor. She has held visiting professor appointments at the Technion–Israel Institute of Technology and ETH Zurich. Her research interests lie in various areas of information theory including channel coding, compression, coordination, secrecy problems, and applications in statistics. From 2016 to 2019, she has also served on the Board of Governors for the IEEE Information Theory Society. She has previously served as an Associate Editor for the IEEE COMMUNICATION LETTERS and an Associate Editor for Shannon Theory for the IEEE TRANSACTIONS ON INFORMATION THEORY.