Distributed Hypothesis Testing over a Noisy Channel: Error-exponents Trade-off

Sreejith Sreekumar and Deniz Gündüz
Dept. of Electrical and Electronic Engineering, Imperial College London
{s.sreekumar15,d.gunduz}@imperial.ac.uk

Abstract

A distributed hypothesis testing problem with two parties, one referred to as the observer and the other as the detector, is considered. The observer observes a discrete memoryless source and communicates its observations to the detector over a discrete memoryless noisy channel. The detector observes a side-information correlated with the observer's observations, and performs a binary hypothesis test on the joint probability distribution of its own observations with that of the observer. With the objective of characterizing the performance of the hypothesis test, we obtain two inner bounds on the trade-off between the exponents of the type I and type II error probabilities. The first inner bound is obtained using a combination of a type-based quantize-bin scheme and Borade et al.'s unequal error protection scheme, while the second inner bound is established using a type-based hybrid coding scheme. These bounds extend the achievability result of Han and Kobayashi obtained for the special case of a rate-limited noiseless channel to a noisy channel. For the special case of testing for the marginal distribution of the observer's observations with no side-information at the detector, we establish a single-letter characterization of the optimal trade-off between the type I and type II error-exponents. Our results imply that a "separation" holds in this case, in the sense that the optimal trade-off between the error-exponents is achieved by a scheme that performs independent hypothesis testing and channel coding.

I. Introduction

Consider the distributed hypothesis testing (HT) setting depicted in Fig. 1, in which, a node referred to as the *detector*, wants to ascertain the underlying joint probability distribution of its own observed data with the data observed at a remote node, referred to as the *observer*. The k data samples observed at the observer, denoted by u^k, are communicated to the detector over a noisy communication channel. Based on its own observations, denoted by v^k, and the data received from the observer, the detector performs a hypothesis test to determine the joint probability distribution that generated (u^k, v^k). Assuming that the data samples under each hypothesis are drawn (by nature) independent and identically according to a fixed distribution, the simplest case of such a test is a binary hypothesis test with the following null (H_0) and alternate (H_1) hypotheses :

$$H_0 : \prod_{i=1}^{k} P_{UV}(u,v), \ \forall \ (u,v) \in \mathcal{U} \times \mathcal{V}, \tag{1a}$$

$$H_1 : \prod_{i=1}^{k} P_{\bar{U}\bar{V}}(u,v), \ \forall \ (u,v) \in \mathcal{U} \times \mathcal{V}. \tag{1b}$$

Here, P_{UV} and $P_{\bar{U}\bar{V}}$ denote the joint probability distribution from which the data is generated under the null and alternate hypothesis, respectively. Our goal is to characterize the performance of the above hypothesis test as measured by the exponents of the type I and type II error probabilities (see Definition 1 below).

The above problem is an extension of the classical distributed HT problem with communication constraints considered by Ahlswede and Csiszár in [1], and extensively studied thereafter [2] [3][4]. While the centralized setting in which all the data is available at a single location is well understood, thanks to [5], [6], [7], [8], the optimal characterization of the error-exponents for the distributed HT problem remains open except for some special cases. In [1]-[4] and the follow up literature henceforth, the communication channel between the observer and the detector is assumed to be rate-limited and error-free. In [1]-[3], the maximum asymptotic value of the exponent of the type II error probability, known as the *type II error-exponent* (T2EE), subject to a fixed constraint on the type I error probability is studied under various settings. The fundamental results related to distributed HT is established in [1], which includes a lower bound on the optimal T2EE, a single-letter characterization of the optimal T2EE for a special case known as *testing against independence*, where $P_{\bar{U}\bar{V}} = P_U \times P_V$, and a *strong converse* which shows that the optimal T2EE is independent of the type I error probability constraint. Improved lower bounds on the T2EE are obtained in [2] and [9]. The trade-off between the exponents of both the type I and type II error probabilities is explored in [4], where the authors establish an inner bound using a type based quantization scheme. Recently, there is a renewed interest in the distributed HT problem and extensions of the problem has been studied in several different contexts, e.g. multi-terminal settings [10], [11],[12], [13], under security or privacy constraints [14], [15] [16] [17], along with lossy compression [18], etc. to name a few. The trade-off between the type I and type II error-exponents studied in [4] is revisited in [19], where an inner bound is obtained using the technique of structured binning and analogy to the channel detection problem. While the above works focus on the asymptotic performance in distributed HT, a Neyman-Pearson like test for zero-rate multiterminal HT is

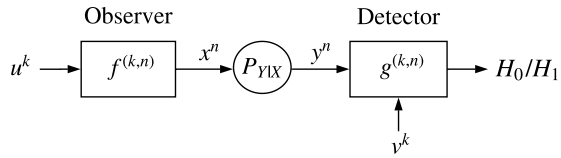

Fig. 1: Distributed hypothesis testing over a noisy channel.

proposed in [20], which in addition to achieving the optimal type II error-exponent, also achieves the optimal second order term of the exponent among the class of all type based testing schemes.

When the communication channel between the observer and the detector is noisy, then besides the type I and type II errors that arise due to the compression of the sequence u^k, additional errors occur because of the noisiness of the channel. Since the reliability of the transmitted message depends on the communication rate employed, there is a trade-off between transmitting less information more reliably versus transmitting more information less reliably, to the detector. In [21], we proved a single-letter characterization of the optimal T2EE for testing against independence in the setting in Fig. 1, which revealed the interesting fact that the optimal T2EE depends on the communication channel only through its capacity. Extensions of this problem to general HT is investigated in [22] and [23].

In this paper, we study the trade-off between both the type I and type II error-exponents for the distributed HT problem in Fig. 1. Our main contributions are as follows:

(i) In Theorem 3, we establish a single-letter characterization of the optimal trade-off between the type I and type II error-exponents[1] for the special case where the side-information v^k is absent and the hypothesis test is on the marginal distribution of u^k, referred to henceforth as the *non-distributed* setting.

(ii) We obtain an inner bound (Theorem 4) on the trade-off between the error-exponents in the distributed setting by using a *separate HT and channel coding* scheme (SHTCC) that is a combination of a type based quantize-bin strategy and unequal error-protection scheme in [25]. This result recovers the inner bound obtained in [4] for the case of a rate-limited noiseless channel, and a lower bound on the type II error-exponent for a fixed constraint on the type I error probability established in [22] for the case of a noisy channel.

(iii) We obtain a second inner bound (Theorem 5) on the error-exponents trade-off by using a *joint HT and channel coding scheme* (JHTCC) based on *hybrid coding*. This bound is at least as tight as the one in Theorem 4 when the type I error-exponent is zero.

The problem studied here has been investigated recently in [26], where an inner bound on the error-exponents is obtained using a combination of a type based quantization scheme and unequal error protection scheme of [27] with two special messages. Our schemes differ from that in [26] in the following aspects: (i) In the SHTCC scheme, the encoder employs binning subsequent to quantization and Borade et al.'s unequal error protection with a single special message (in place of [27]); (ii) In the JHTCC scheme, the encoder uses a hybrid coding scheme that transmits the channel codeword as a function of the quantization codeword as well as the sequence u^k.

Before proceeding to the results, we introduce the notations used in the paper.

A. Notations

Random variables (r.v.'s) and their realizations are denoted by upper and lower case letters (e.g., X and x), respectively. The distribution of a r.v. X is denoted by P_X. Sets are denoted by calligraphic letters, e.g., the alphabet of a r.v. X is denoted by $\mathcal{X}$. The cartesian product of sets $\mathcal{X}$ and $\mathcal{Y}$ is denoted by $\mathcal{X} \times \mathcal{Y}$. The n- fold cartesian product of a set $\mathcal{X}$ is represented by $\mathcal{X}^n$. The set of probability distributions on $\mathcal{X}$ is denoted by $\mathcal{P}_{\mathcal{X}}$. We will extensively use the method of types [28]. Accordingly, the *type* (or empirical distribution) of a sequence $x^n \in \mathcal{X}^n$ is denoted by P_{x^n} or $P_{\tilde{X}}$, where $\tilde{X}$ denotes a r.v. with distribution equal to the empirical distribution of x^n. The type class of $P_{\tilde{X}}$, i.e., the set of sequences of length n with type $P_{\tilde{X}}$ is denoted by $\mathcal{T}_n(P_{\tilde{X}})$ or $\mathcal{T}_n(\tilde{X})$. The set of all possible types of sequences of length n with alphabet $\mathcal{X}$ is denoted by $\mathcal{T}_n(\mathcal{X})$. Similar notations will be used for pairs and larger combinations of sequences, e.g., the joint type of (x^n, y^n) is denoted by $P_{x^n y^n}$ or $P_{\tilde{X}\tilde{Y}}$, where $\tilde{X}\tilde{Y}$ is a r.v. with distribution $P_{x^n y^n}$. By abuse of notation, $P_{\tilde{X}} \in \mathcal{F}$, $\mathcal{F} \subseteq \mathcal{P}_{\mathcal{X}}$, will also be denoted by $\tilde{X} \in \mathcal{F}$, e.g., $P_{\tilde{X}} \in \mathcal{T}_n(\mathcal{X})$ by $\tilde{X} \in \mathcal{T}_n(\mathcal{X})$. For a given $x^n \in \mathcal{T}_n(P_{\tilde{X}})$, the conditional type class of x^n for a conditional type $P_{\tilde{Y}|\tilde{X}}$, i.e., the set of $y^n \in \mathcal{T}_n(P_{\tilde{Y}})$ such that $(x^n, y^n) \in \mathcal{T}_n(P_{\tilde{X}\tilde{Y}})$, is denoted by $\mathcal{T}_n(P_{\tilde{Y}|\tilde{X}}, x^n)$. The Shannon entropy of X, the mutual information between X and Y, and the Kullback-Leibler (KL) divergence between X and $\hat{X}$ with same support $\mathcal{X}$ are denoted by $H(X)$, $I(X;Y)$ and $D(X||\hat{X})$ (or $D(P_X||P_{\hat{X}})$), respectively. The conditional divergence between two distributions $P_{X_1|X_2}$ and $P_{\bar{X}_1|\bar{X}_2}$ (defined on same alphabets) is denoted by $D\left(P_{X_1|X_2}||P_{\bar{X}_1|\bar{X}_2}\middle|P_{X_2}\right)$ or $D\left(X_1|X_2||\bar{X}_1|\bar{X}_2\middle|X_2\right)$ where,

[1]A corner point of this trade-off, namely the optimal T2EE for a fixed non-zero constraint on the type I error probability, is established in [24].

$$D\left(P_{X_1|X_2}||P_{\bar{X}_1|\bar{X}_2}\middle|P_{X_2}\right) := D\left(X_1|X_2||\bar{X}_1|\bar{X}_2\middle|X_2\right) = \sum_{x_2\in\mathcal{X}_2} P_{X_2}(x_2) D\left(P_{X_1|X_2=x_2}||P_{\bar{X}_1|\bar{X}_2=x_2}\right).$$

When $X_2 = \bar{X}_2$, the notation above is further simplified to $D\left(P_{X_1|X_2}||P_{\bar{X}_1|X_2}\right)$ or $D\left(X_1|X_2||\bar{X}_1|X_2\right)$. For any two sequences $x^n \in \mathcal{X}^n$ and $y^n \in \mathcal{Y}^n$ with joint type $P_{\tilde{X}\tilde{Y}}$, the empirical conditional entropy of y^n given x^n is defined by

$$H_e(y^n|x^n) := H(\tilde{Y}|\tilde{X}),$$

where $H(\cdot)$ denotes the standard Shannon conditional entropy. The set of r-divergent sequences from X is denoted by $\mathcal{J}_n^r(X)$, i.e.,

$$\mathcal{J}_n^r(X) = \{x^n \in \mathcal{X}^n : D(P_{x^n}||P_X) \le r\}.$$

For $a \in \mathbb{R}^+$, the set of integers $\{1, 2, \dots, \lceil a \rceil\}$ is denoted by $[a]$. The limiting inequalities $\lim_{k\to\infty} a_k = b$, $\lim_{k\to\infty} a_k \ge b$, etc. are denoted by $a_k \xrightarrow{(k)} b$, $a_k \overset{(k)}{\ge} b_k$, etc., respectively. Probabilistic events are denoted by calligraphic letters, e.g., $\mathcal{E}$, and its probability by $\mathbb{P}(\mathcal{E})$. The complement of $\mathcal{E}$ is denoted by $\mathcal{E}^c$. Finally, the indicator function is denoted by $\mathbb{1}(\cdot)$ and the standard asymptotic notations of Big-o, Big-omega and Little-o are represented by $O(\cdot)$, $\Omega(\cdot)$ and $o(\cdot)$, respectively.

B. Problem formulation

All r.v.'s considered henceforth are discrete with finite support unless specified otherwise, and all logarithms are with respect to the natural base e. Let $k, n \in \mathbb{Z}^+$. The encoder observes source sequence u^k, and transmits codeword $x^n = f^{(k,n)}(u^k)$, where $f^{(k,n)} : \mathcal{U}^k \to \mathcal{X}^n$ represents the encoding function (possibly stochastic) of the observer. The channel output y^n given input x^n is generated according to the probability law

$$P_{Y^n|X^n}(y^n|x^n) = \prod_{j=1}^{n} P_{Y|X}(y_j|x_j), \tag{2}$$

i.e., the channel $P_{Y^n|X^n}$ is memoryless. Depending on the received symbols y^n and side-information v^k observed at the detector, the detector makes a decision between the two hypotheses H_0 and H_1 given in (1). Let $P_{U^kV^kX^nY^n} := P_{U^kV^k}P_{X^n|U^k}P_{Y^n|X^n}$ and $P_{\bar{U}^k\bar{V}^k\bar{X}^n\bar{Y}^n} := P_{\bar{U}^k\bar{V}^k}P_{\bar{X}^n|\bar{U}^k}P_{\bar{Y}^n|\bar{X}^n}$ denote the probability distribution of the source sequence, side-information, channel input and channel output under the null and alternate hypothesis, respectively, where $P_{X^n|U^k}(x^n|u^k) = P_{\bar{X}^n|\bar{U}^k}(x^n|u^k) = \mathbb{P}(f^{(k,n)}(u^k) = x^n)$ for all $(u^k, x^n) \in \mathcal{U}^k \times \mathcal{X}^n$, and $P_{\bar{Y}^n|\bar{X}^n} := P_{Y^n|X^n}$. Let $H \in \{0, 1\}$ denote the actual hypothesis and $\hat{H} \in \{0, 1\}$ denote the output of the hypothesis test, where 0 and 1 denote H_0 and H_1, respectively. Let $\mathcal{A}_{k,n} \subseteq \mathcal{V}^k \times \mathcal{Y}^n$ denote the acceptance region for H_0. Then, the decision rule $g^{(k,n)} : \mathcal{V}^k \times \mathcal{Y}^n \to \{0, 1\}$ is given by

$$g^{(k,n)}\left(v^k, y^n\right) = 1 - \mathbb{1}\left((v^k, y^n) \in \mathcal{A}_{k,n}\right).$$

Let

$$\alpha\left(k, n, f^{(k,n)}, g^{(k,n)}\right) := 1 - P_{V^kY^n}\left(\mathcal{A}_{k,n}\right),$$
$$\text{and } \beta\left(k, n, f^{(k,n)}, g^{(k,n)}\right) := P_{\bar{V}^k\bar{Y}^n}\left(\mathcal{A}_{k,n}\right),$$

denote the type I and type II error probabilities for the encoding function $f^{(k,n)}$ and decision rule $g^{(k,n)}$, respectively. The following definition formally states the error-exponents trade-off we aim to characterize.

Definition 1. *Let $\tau \in (0, \infty)$. An exponent pair $(\kappa_\alpha, \kappa_\beta)$ is τ-achievable if there exists sequences of integers k and n_k, corresponding sequence of encoding functions $f^{(k,n_k)}$ and decoding functions $g^{(n_k)}$, and $k_0 \in \mathbb{Z}^+$ such that*

$$n_k \le \tau k, \ \forall\ k \ge k_0, \tag{3a}$$

$$\alpha\left(k, n_k, f^{(k,n_k)}, g^{(k,n_k)}\right) \le e^{-k\kappa_\alpha}, \ \forall\ k \ge k_0, \tag{3b}$$

$$\textit{and } \liminf_{k\to\infty} -\frac{1}{k}\log\left(\beta\left(k, n_k, f^{(k,n_k)}, g^{(k,n_k)}\right)\right) \ge \kappa_\beta. \tag{3c}$$

Let

$$\kappa(\tau, \kappa_\alpha) := \sup\{\kappa_\beta : (\kappa_\alpha, \kappa_\beta) \text{ is } \tau\text{-achievable}\}.$$

For a fixed $\tau \in (0,\infty)$, we are interested in characterizing the boundary of the set of all τ-achievable $(\kappa_\alpha, \kappa_\beta)$ tuples defined as

$$\mathcal{R} := \{(\kappa_\alpha, \kappa(\tau,\kappa_\alpha)) : \kappa_\alpha \in (0,\infty]\}.$$

Towards this, we will first obtain a single-letter characterization of $\mathcal{R}$ in the non-distributed setting. This characterization will be subsequently used to obtain an inner bound on $\mathcal{R}$ in the distributed setting.

II. HT: Error exponents trade-off

In the non-distributed setting, the hypothesis test in (1) specializes to the following test:

$$H_0 : \prod_{i=1}^{k} P_U(u),\ \forall\ u \in \mathcal{U}, \tag{4a}$$

$$H_1 : \prod_{i=1}^{k} P_{\bar{U}}(u),\ \forall\ u \in \mathcal{U}. \tag{4b}$$

For brevity, we will denote the r.v. with distribution $P_{Y|X=x}$ by Y_x and the corresponding probability distribution by P_{Y_x} for all $x \in \mathcal{X}$. Let us define

$$\kappa_0 := \kappa_0(\tau, P_U, P_{\bar{U}}, P_{Y|X}) := \min\left(D(P_{\bar{U}}||P_U), \tau E_c\right),$$

where,

$$E_c := E_c(P_{Y|X}) := D(P_{Y_a}||P_{Y_b}), \tag{5}$$

$$\text{and } (a,b) := \underset{(x,x')\in\mathcal{X}\times\mathcal{X}}{\arg\max}\ D(P_{Y_x}||P_{Y_{x'}}). \tag{6}$$

It follows by interchanging P_U and $P_{\bar{U}}$ in Theorem 2 [24] that, we may restrict the range of κ_α within the interval $(0,\kappa_0]$ since $\kappa(\tau,\kappa_\alpha) = 0$ for $\kappa_\alpha \geq \kappa_0$. Hence, $\mathcal{R}$ can be redefined as $\mathcal{R} = \{(\kappa_\alpha,\ \kappa(\tau,\kappa_\alpha)) : \kappa_\alpha \in (0,\kappa_0]\}$.

In order to state our single-letter characterization of $\mathcal{R}$, we need some concepts regarding the log moment generating function (Log-MGF) of a r.v., which we briefly review below. For a given function $f : \mathcal{Z} \to \mathbb{R}$ and a probability distribution P_Z on Z, the log-MGF of Z with respect to f, denoted by $\psi_{Z,f}(\lambda)$ is given by

$$\psi_{Z,f}(\lambda) := \psi_{P_Z,f}(\lambda) := \log\left(\mathbb{E}_{P_Z}\left(e^{\lambda f(Z)}\right)\right).$$

Let

$$\psi^*_{Z,f}(\theta) := \psi^*_{P_Z,f}(\theta) := \sup_{\lambda\in\mathbb{R}} \theta\lambda - \psi_{Z,f}(\lambda). \tag{7}$$

The following simple facts are straightforward to verify.

Lemma 1. *[29, Theorem 13.2, Theorem 13.3]*

(i) $\psi_{Z,f}(0) = 0$ *and* $\psi'_{Z,f}(0) = E_{P_Z}(f(Z))$, *where* $\psi'_{Z,f}(\lambda)$ *denotes the derivative of* $\psi_{Z,f}(\lambda)$ *with respect to* λ.

(ii) $\psi_{Z,f}(\lambda)$ *is a strictly convex function in* λ.

(iii) $\psi^*_{Z,f}(\theta)$ *is strictly convex and strictly positive in* θ *except* $\psi^*_{Z,f}(\mathbb{E}(Z)) = 0$.

We will assume[2] that

Assumption 1. $P_U \ll P_{\bar{U}}$, $P_{\bar{U}} \ll P_U$ *and* $P_{Y|X}$ *is such that* $P_{Y_x} \ll P_{Y_{x'}}$, $\forall\ (x,x') \in \mathcal{X}\times\mathcal{X}$.

When u^k is observed directly at the detector, a single-letter characterization of the optimal trade-off between the error-exponents is obtained in [8]. Below, we state an equivalent form of this characterization that is given in [29].

Theorem 1. *[29, Theorem 15.1] When u^k is observed directly at the detector, then for the HT given in* (4)*,*

$$\mathcal{R} = \left\{\left(\psi^*_{U,f_{\mathcal{U}}}(\theta), \psi^*_{U,f_{\mathcal{U}}}(\theta) - \theta\right) :\ \theta \in \mathcal{I}(U,\bar{U})\right\},$$

where, $f_{\mathcal{U}} : \mathcal{U} \to \mathbb{R}^+$ *is defined as*

$$f_{\mathcal{U}}(u) := \log\left(\frac{P_{\bar{U}}(u)}{P_U(u)}\right), \tag{8}$$

[2]This technical condition ensures that for functions f and distributions P that we consider below, $\psi_{P,f}(\lambda) < \infty,\ \forall\ \lambda \in \mathbb{R}$.

and $\mathcal{I}(P_U,P_{\bar{U}}) := \big(-D(P_U||P_{\bar{U}}), D(P_{\bar{U}}||P_U)\big]$. *The decision rule that achieves the exponent pair* $\left(\psi^*_{U,f_{\mathcal{U}}}(\theta), \psi^*_{U,f_{\mathcal{U}}}(\theta)-\theta\right)$ *is the Neyman-Pearson (NP) test [5] given by*

$$g^{(k)}_{\theta,\mathcal{U}}(u^k) = \mathbb{1}\left(\sum_{i=1}^{k} \log\left(\frac{P_{\bar{U}}(u_i)}{P_U(u_i)}\right) \geq k\theta\right). \tag{9}$$

To prove the main result, a strong converse result that follows from [29, Theorem 12.5] will turn out to be handy. We state it below for completeness. For the scenario where u^k is observed directly at the detector, let us denote the type I and type II error probabilities achieved by a decision rule (possibly stochastic) $\bar{g}^{(k)}: \mathcal{U}^k \mapsto \{0,1\}$ by $\alpha'(\bar{g}^{(k)})$ and $\beta'(\bar{g}^{(k)})$, respectively. The following single-shot result provides a lower bound on a weighted sum of the type I and type II error probabilities[3].

Theorem 2. *[29, Theorem 12.5] For any* $k \in \mathcal{Z}^+$ *and any decision rule* $\bar{g}^{(k)}$ *as defined above,*

$$\alpha'\left(\bar{g}^{(k)}\right) + \gamma\beta'\left(\bar{g}^{(k)}\right) \geq P_{U^k}\left(\log\left(\frac{P_{U^k}(U^k)}{P_{\bar{U}^k}(U^k)}\right) \leq \log\gamma\right), \ \forall\ \gamma > 0.$$

We will also require a slight generalization of Theorem 1 for the case when the data samples are drawn from a product of finite non-identical distributions, i.e., the samples are independent, but not necessarily identically distributed. For an arbitrary given joint distribution $P_{X_0X_1} \in \mathcal{P}(\mathcal{X}\times\mathcal{X})$, let $\{(x_0^n, x_1^n)\}_{n\in\mathbb{Z}^+}$ denote a given pair of sequences such that

$$P_{x_0^n x_1^n}(x,x') \xrightarrow{(n)} P_{X_0X_1}(x,x'),\ \forall\ (x,x') \in \mathcal{X}\times\mathcal{X}. \tag{10}$$

Consider the following HT:

$$H_0 : Y^n \sim \prod_{i=1}^{n} P_{Y_{x_{0i}}}, \tag{11a}$$

$$H_1 : Y^n \sim \prod_{i=1}^{n} P_{Y_{x_{1i}}}. \tag{11b}$$

For a given decision rule $g^{(n)}(y^n) = 1 - \mathbb{1}(Y^n \in \mathcal{A}_n)$ with acceptance region $\mathcal{A}_n \subseteq \mathcal{Y}^n$ for H_0, let $\bar{\alpha}\left(n, g^{(n)}, x_0^n, x_1^n\right)$ and $\bar{\beta}\left(n, g^{(n)}, x_0^n, x_1^n\right)$ denote the type I and type II error probabilities, respectively, where

$$\bar{\alpha}\left(n, g^{(n)}, x_0^n, x_1^n\right) := 1 - \prod_{i=1}^{n} P_{Y_{x_{0i}}}\left(\mathcal{A}_n\right),$$

$$\text{and}\ \ \bar{\beta}\left(n, g^{(n)}, x_0^n, x_1^n\right) := \prod_{i=1}^{n} P_{Y_{x_{1i}}}\left(\mathcal{A}_n\right).$$

Definition 2. *For a given joint distribution* $P_{X_0X_1}$ *and a pair of infinite sequences* $\{(x_0^n,x_1^n)\}_{n\in\mathbb{Z}^+}$ *such that* (10) *holds, an exponent pair* $(\kappa_\alpha,\kappa_\beta)$ *is achievable for the HT in* (11) *if there exists a sequence of decision rules* $\{g^{(n)}\}_{n\in\mathbb{Z}^+}$ *and* $n_0 \in \mathbb{Z}^+$ *such that*

$$\bar{\alpha}\left(n, g^{(n)}, x_0^n, x_1^n\right) \leq e^{-n\kappa_\alpha},\ \forall\ n \geq n_0, \tag{12a}$$

$$\textit{and}\ \liminf_{n\to\infty} -\frac{1}{n}\log\left(\bar{\beta}\left(n, g^{(n)}, x_0^n, x_1^n\right)\right) \geq \kappa_\beta. \tag{12b}$$

As will become evident later, the performance of the HT in (11) depends on $\{(x_0^n,x_1^n)\}_{n\in\mathbb{Z}^+}$ only through $P_{X_0X_1}$. Let

$$\bar{\kappa}_c(\kappa_\alpha, P_{X_0X_1}) := \sup\{\kappa_\beta : (\kappa_\alpha,\kappa_\beta) \text{ is achievable for HT in (11)}\}.$$
$$\text{and}\ \mathcal{R}_N(P_{X_0X_1}) := \{(\kappa_\alpha, \bar{\kappa}_c(\kappa_\alpha, P_{X_0X_1})) : \kappa_\alpha \in (0,\kappa_\alpha^*]\}.$$

where κ_α^* is the smallest number such that $\bar{\kappa}(\kappa_\alpha^*, P_{X_0X_1}) = 0$. The following proposition provides a single-letter characterization of $\mathcal{R}_N(P_{X_0X_1})$, and will be used later for obtaining a single-letter characterization of $\mathcal{R}$ in Theorem 3.

Proposition 1.

$$\mathcal{R}_N(P_{X_0X_1}) = \left\{\left(\mathbb{E}_{P_{X_0X_1}}\left(\psi^*_{Y_{X_0},\tilde{h}_{X_0,X_1}}(\theta)\right),\ \mathbb{E}_{P_{X_0X_1}}\left(\psi^*_{Y_{X_0},\tilde{h}_{X_0,X_1}}(\theta)\right) - \theta\right),\ \theta \in \mathcal{I}(P_{X_0X_1}, P_{Y|X})\right\},$$

where, for each $(x,x') \in \mathcal{X}\times\mathcal{X}$, $\tilde{h}_{x,x'} : \mathcal{Y} \to \mathbb{R}$ *is given by*

$$\tilde{h}_{x,x'}(y) := \log\left(\frac{P_{Y_{x'}}(y)}{P_{Y_x}(y)}\right), \tag{13}$$

[3]Note that α denotes the complement of the type I error probability in [29, Theorem 12.5], whereas we use α' to denote the type I error probability.

and

$$\mathcal{I}(P_{X_0X_1}, P_{Y|X}) := \big(-D(P_{Y_{X_0}}||P_{Y_{X_1}}|P_{X_0X_1}), D(P_{Y_{X_1}}||P_{Y_{X_0}}|P_{X_0X_1})\big],$$

$$D(P_{Y_{X_0}}||P_{Y_{X_1}}|P_{X_0X_1}) := \sum_{(x_0,x_1)\in\mathcal{X}\times\mathcal{X}} P_{X_0X_1}(x_0,x_1) D(P_{Y_{x_0}}||P_{Y_{x_1}}),$$

$$D(P_{Y_{X_1}}||P_{Y_{X_0}}|P_{X_0X_1}) := \sum_{(x_0,x_1)\in\mathcal{X}\times\mathcal{X}} P_{X_0X_1}(x_0,x_1) D(P_{Y_{x_1}}||P_{Y_{x_0}}).$$

The decision rule that achieves the exponent pair $\left(\mathbb{E}_{P_{X_0X_1}}\left(\psi^*_{Y_{X_0},\tilde{h}_{X_0,X_1}}(\theta)\right),\ \mathbb{E}_{P_{X_0X_1}}\left(\psi^*_{Y_{X_0},\tilde{h}_{X_0,X_1}}(\theta)\right)-\theta\right)$ *is the NP test given by*

$$g^{(n)}_{\theta,\mathcal{Y}}(y^n) = \mathbb{1}\left(\sum_{i=1}^{n} \log\left(\frac{P_{Y_{x_{1i}}}(y_i)}{P_{Y_{x_{0i}}}(y_i)}\right) \geq n\theta\right). \tag{14}$$

Proof: The proof is given in Appendix A. ■

The next theorem provides a single-letter characterization of $\kappa(\tau,\kappa_\alpha)$ and thereby of $\mathcal{R}$.

Theorem 3.

$$\kappa(\tau,\kappa_\alpha) = \sup\{\kappa_\beta : (\kappa_\alpha,\kappa_\beta)\in\mathcal{R}^*\}$$

where

$$\mathcal{R}^* := \bigcup_{\substack{P_{X_0X_1}\in\\ \mathcal{P}_{\mathcal{X}\times\mathcal{X}}}} \ \bigcup_{\substack{(\theta_0,\theta_1)\ \in\\ \mathcal{I}(P_U,P_{\bar{U}})\times\mathcal{I}(P_{X_0X_1},P_{Y|X})}} \left(\zeta_0(\theta_0,\theta_1,P_{X_0X_1}),\ \zeta_1(\theta_0,\theta_1,P_{X_0X_1})\right),$$

$$\zeta_0(\theta_0,\theta_1,P_{X_0X_1}) := \min\left(\psi^*_{U,f_{\mathcal{U}}}(\theta_0),\ \tau\mathbb{E}_{P_{X_0X_1}}\left(\psi^*_{Y_{X_0},\tilde{h}_{X_0,X_1}}(\theta_1)\right)\right),$$

$$\zeta_1(\theta_0,\theta_1,P_{X_0X_1}) := \min\left(\psi^*_{U,f_{\mathcal{U}}}(\theta_0)-\theta_0,\ \tau\left(\mathbb{E}_{P_{X_0X_1}}\left(\psi^*_{Y_{X_0},\tilde{h}_{X_0,X_1}}(\theta_1)\right)-\theta_1\right)\right),$$

where $f_{\mathcal{U}}$ *and* $\tilde{h}_{x,x'}$, $(x,x')\in\mathcal{X}\times\mathcal{X}$ *are defined in* (8) *and* (13)*, respectively.*

Proof: Achievability: Fix $P_{X_0X_1}\in\mathcal{P}_{\mathcal{X}\times\mathcal{X}}$. Let $n_k=\lfloor\tau k\rfloor$, and $x_0^{n_k}$ and $x_1^{n_k}$ be two arbitrary sequences from the set $\mathcal{X}^{n_k}$ such that (10) holds. Let $\theta_0\in\mathcal{I}(P_U,P_{\bar{U}})$ and $\theta_1\in\mathcal{I}(P_{X_0X_1},P_{Y|X})$. The achievability scheme is as follows:
The encoder first locally performs the NP test $g^{(k)}_{\theta,\mathcal{U}}$ given in (9) on the observed samples u^k with $\theta=\theta_0$, and outputs the channel input codeword $f^{(k,n_k)}(u^k)$ according to the following rule:

$$f^{(k,n_k)}(u^k) = \begin{cases} x_0^{n_k}, & \text{if } g^{(k)}_{\theta_0,\mathcal{U}}(u^k)=0,\\ x_1^{n_k}, & \text{otherwise.}\end{cases}$$

Based on the observed samples y^{n_k}, the detector outputs the decision of the HT according to the decision rule $g^{(n_k)}=g^{(n_k)}_{\theta_1,\mathcal{Y}}$ defined in (14). Let $\mathcal{A}^{(n_k)}_{\theta_1}=\left\{y^{n_k}\in\mathcal{Y}^{n_k}:\sum_{i=1}^{n_k}\log\left(\frac{P_{Y|X=x_{1i}}(y_i)}{P_{Y|X=x_{0i}}(y_i)}\right)<n_k\theta_1\right\}$. The type I error error probability can be upper bounded for sufficiently large k (and n_k) as follows:

$$\begin{aligned}
&\alpha\left(k,n_k,f^{(k,n_k)},g^{(n_k)}\right)\\
&\leq \mathbb{P}\left(g^{(k)}_{\theta_0,\mathcal{U}}(U^k)=1\right)P_{Y^{n_k}|X^{n_k}=x_1^{n_k}}\left(\mathcal{Y}^{n_k}\backslash\mathcal{A}^{(n_k)}_{\theta_1}\right)+\mathbb{P}\left(g^{(k)}_{\theta_0,\mathcal{U}}(U^k)=0\right)P_{Y^{n_k}|X^{n_k}=x_0^{n_k}}\left(\mathcal{Y}^{n_k}\backslash\mathcal{A}^{(n_k)}_{\theta_1}\right)\\
&\leq \mathbb{P}\left(g^{(k)}_{\theta_0,\mathcal{U}}(U^k)=1|H=0\right)+P_{Y^{n_k}|X^{n_k}=x_0^{n_k}}\left(\mathcal{Y}^{n_k}\backslash\mathcal{A}^{(n_k)}_{\theta_1}\right)\\
&\leq e^{-k\left(\psi^*_{U,f_{\mathcal{U}}}(\theta_0)-\delta\right)}+e^{-n_k\left(\mathbb{E}_{P_{X_0X_1}}\left(\psi^*_{Y_{X_0},\tilde{h}_{X_0,X_1}}(\theta_1)\right)-\delta\right)},
\end{aligned} \tag{15}$$

where $\delta>0$ is an arbitrary small, but fixed number. Similarly, the type II error probability can be upper bounded as follows:

$$\begin{aligned}
&\beta\left(k,n_k,f^{(k,n_k)},g^{(n_k)}\right)\\
&\leq \mathbb{P}\left(g^{(k)}_{\theta_0,\mathcal{U}}(\bar{U}^k)=0\right)P_{Y^{n_k}|X^{n_k}=x_0^{n_k}}\left(\mathcal{A}^{(n_k)}_{\theta_1}\right)+\mathbb{P}\left(g^{(k)}_{\theta_0,\mathcal{U}}(\bar{U}^k)=1\right)P_{Y^{n_k}|X^{n_k}=x_1^{n_k}}\left(\mathcal{A}^{(n_k)}_{\theta_1}\right)\\
&\leq \mathbb{P}\left(g^{(k)}_{\theta_0,\mathcal{U}}(\bar{U}^k)=0\right)+P_{Y^{n_k}|X^{n_k}=x_1^{n_k}}\left(\mathcal{A}^{(n_k)}_{\theta_1}\right)\\
&\leq e^{-k\left(\psi^*_{\bar{U},f_{\mathcal{U}}}(\theta_0)-\delta\right)}+e^{-n_k\left(\mathbb{E}_{P_{X_0X_1}}\left(\psi^*_{Y_{X_1},\tilde{h}_{X_0,X_1}}(\theta_1)\right)-\delta\right)}
\end{aligned}$$

$$= e^{-k\left(\psi^*_{U,f_{\mathcal{U}}}(\theta_0)-\theta_0-\delta\right)} + e^{-n_k\left(\mathbb{E}_{P_{X_0X_1}}\left(\psi^*_{Y_{X_0},\tilde{h}_{X_0,X_1}}(\theta_1)\right)-\theta_1-\delta\right)}. \tag{16}$$

It follows from (15) and (16), respectively, that,

$$\begin{aligned}\liminf_{k\to\infty} -\frac{1}{k}\log\left(\alpha\left(k,n_k,f^{(k,n_k)},g^{(n_k)}\right)\right) &\geq \min\left(\psi^*_{U,f_{\mathcal{U}}}(\theta_0),\ \mathbb{E}_{P_{X_0X_1}}\left(\psi^*_{Y_{X_0},\tilde{h}_{X_0,X_1}}(\theta_1)\right)\right)-\delta \\ &= \zeta_0(\theta_0,\theta_1,P_{X_0X_1})-\delta,\end{aligned}$$

and

$$\begin{aligned}\liminf_{k\to\infty} -\frac{1}{k}\log\left(\beta\left(k,n_k,f^{(k,n_k)},g^{(n_k)}\right)\right) &\geq \min\left(\psi^*_{U,f_{\mathcal{U}}}(\theta_0)-\theta_0,\ \mathbb{E}_{P_{X_0X_1}}\left(\psi^*_{Y_{X_0},\tilde{h}_{X_0,X_1}}(\theta_1)\right)-\theta_1\right)-\delta \\ &= \zeta_1(\theta_0,\theta_1,P_{X_0X_1})-\delta.\end{aligned}$$

Since δ is arbitrary, it follows by varying $P_{X_0X_1}\in\mathcal{P}_{\mathcal{X}\times\mathcal{X}}$, $\theta_0\in\mathcal{I}(P_U,P_{\bar{U}})$ and $\theta_1\in\mathcal{I}(P_{X_0X_1},P_{Y|X})$ that

$$\kappa(\tau,\kappa_\alpha)\geq\sup\{\kappa_\alpha:\ (\kappa_\alpha,\kappa_\beta)\in\mathcal{R}^*\}.$$

This completes the proof of achievability.

Converse: From the proof of the converse part of Theorem 1, it follows that for $\theta_0\in\mathcal{I}(P_U,P_{\bar{U}})$,

$$\mathcal{R}\subseteq\bigcup_{\theta_0\in\mathcal{I}(P_U,P_{\bar{U}})}\left(\psi^*_{U,f_{\mathcal{U}}}(\theta_0),\ \psi^*_{U,f_{\mathcal{U}}}(\theta_0)-\theta_0\right). \tag{17}$$

Also, note that for any encoding function $f^{(k,n_k)}$ and decoding function $g^{(n_k)}$ with decision region $\mathcal{A}^{(n_k)}\subseteq\mathcal{Y}^{n_k}$, we have

$$\begin{aligned}\alpha\left(k,n_k,f^{(k,n_k)},g^{(n_k)}\right) &= \sum_{u^k\in\mathcal{U}^k}P_U(u^k)\sum_{x^{n_k}\in\mathcal{X}^{n_k}}P_{X^{n_k}|U^k=u^k}(x^{n_k})P_{Y^{n_k}|X^{n_k}=x^{n_k}}(\mathcal{Y}^{n_k}\backslash\mathcal{A}^{n_k}) \\ &\geq P_{Y^{n_k}|X^{n_k}=\bar{x}_0^{n_k}}(\mathcal{Y}^{n_k}\backslash\mathcal{A}^{n_k}),\end{aligned} \tag{18}$$

for some $\bar{x}_0^{n_k}\in\mathcal{X}^{n_k}$ that depends on $\mathcal{A}^{n_k}$. Similarly,

$$\begin{aligned}\beta\left(k,n_k,f^{(k,n_k)},g^{(n_k)}\right) &= \sum_{u^k\in\mathcal{U}^k}P_{\bar{U}}(u^k)\sum_{x^{n_k}\in\mathcal{X}^{n_k}}P_{X^{n_k}|U^k=u^k}(x^{n_k})P_{Y^{n_k}|X^{n_k}=x^{n_k}}(\mathcal{A}^{n_k}) \\ &\geq P_{Y^{n_k}|X^{n_k}=\bar{x}_1^{n_k}}(\mathcal{A}^{n_k}),\end{aligned} \tag{19}$$

for some $\bar{x}_1^{n_k}\in\mathcal{X}^{n_k}$ (depends on $\mathcal{A}^{n_k}$). Let $\bar{P}_{X_0X_1}$ denote the joint type of the sequences $(\bar{x}_0^n,\bar{x}_1^n)$. Note that the R.H.S. of (18) and (19) correspond to the type I and type II error probabilities of the HT given in (11) with $n=n_k$, $x_0^{n_k}=\bar{x}_0^{n_k}$ and $x_1^{n_k}=\bar{x}_1^{n_k}$. Then, it follows from the converse part of the proof of Lemma 1 that if for some $\theta_1\in\mathcal{I}(\bar{P}_{X_0X_1},P_{Y|X})$ and all sufficiently large n_k, it holds that,

$$\alpha\left(k,n_k,f^{(k,n_k)},g^{(n_k)}\right)<e^{-n_k\left(\mathbb{E}_{\bar{P}_{X_0X_1}}\left(\psi^*_{Y_{X_0},\tilde{h}_{X_0,X_1}}(\theta_1)\right)\right)}, \tag{20}$$

then

$$\limsup_{k\to\infty}-\frac{1}{k}\log\left(\beta\left(k,n_k,f^{(k,n_k)},g^{(n_k)}\right)\right)\leq\tau\left(\mathbb{E}_{\bar{P}_{X_0X_1}}\left(\psi^*_{Y_{X_0},\tilde{h}_{X_0,X_1}}(\theta_1)-\theta_1\right)\right). \tag{21}$$

From (20) and (21), we have

$$\mathcal{R}\subseteq\bigcup_{P_{X_0X_1}\in P_{\mathcal{X}\times\mathcal{X}}}\ \bigcup_{\theta_1\in\mathcal{I}(P_{X_0X_1},P_{Y|X})}\left(\tau\mathbb{E}_{P_{X_0X_1}}\left(\psi^*_{P_{Y_{X_0}},\tilde{h}_{X_0,X_1}}(\theta_1)\right),\tau\left(\mathbb{E}_{P_{X_0X_1}}\left(\psi^*_{P_{Y_{X_0}},\tilde{h}_{X_0,X_1}}(\theta_1)\right)-\theta_1\right)\right). \tag{22}$$

It follows from (17) and (22) that $(\kappa_\alpha,\kappa_\beta)\in\mathcal{R}$ only if there exists some $P_{X_0X_1}$ and $(\theta_0,\theta_1)\in\ \mathcal{I}(P_U,P_{\bar{U}})\times\ \mathcal{I}(P_{X_0X_1},P_{Y|X})$ such that

$$\begin{aligned}\kappa_\alpha &\leq \min\left(\psi^*_{U,f_{\mathcal{U}}}(\theta_0),\ \tau\mathbb{E}_{P_{X_0X_1}}\left(\psi^*_{Y_{X_0},\tilde{h}_{X_0,X_1}}(\theta_1)\right)\right), \\ \text{and } \kappa_\beta &\leq \min\left(\psi^*_{U,f_{\mathcal{U}}}(\theta_0)-\theta_0,\ \tau\left(\mathbb{E}_{P_{X_0X_1}}\left(\psi^*_{Y_{X_0},\tilde{h}_{X_0,X_1}}(\theta_1)\right)-\theta_1\right)\right),\end{aligned}$$

from which it follows that $\kappa(\tau,\kappa_\alpha)\leq\sup\{\kappa_\beta:(\kappa_\alpha,\kappa_\beta)\in\mathcal{R}^*\}$. This completes the proof. ■

Remark 1. *The optimal T2EE for a fixed constraint on the type I error probability can be recovered by taking limit* $\theta_0\to$

$-D(P_U||P_{\bar{U}})$ *and* $\theta_1 \rightarrow -D(P_{Y_{X_0}}||P_{Y_{X_1}}|P_{X_0X_1})$. *In this case,*

$$\zeta_0\left(-D(P_U||P_{\bar{U}}), -D(P_{Y_{X_0}}||P_{Y_{X_1}}|P_{X_0X_1}), P_{X_0X_1}\right) = 0$$
$$\text{and } \zeta_1\left(-D(P_U||P_{\bar{U}}), -D(P_{Y_{X_0}}||P_{Y_{X_1}}|P_{X_0X_1}), P_{X_0X_1}\right) = \min\left(D(P_U||P_{\bar{U}}), D(P_{Y_{X_0}}||P_{Y_{X_1}}|P_{X_0X_1})\right).$$

Maximizing the second argument over all possible $P_{X_0X_1}$ *yields,* $\max_{P_{X_0X_1}\in\mathcal{P}_{\mathcal{X}\times\mathcal{X}}} D(P_{Y_{X_0}}||P_{Y_{X_1}}|P_{X_0X_1}) = E_c$. *Hence,* $\lim_{\kappa_\alpha\rightarrow 0}\kappa(\tau,\kappa_\alpha) = \kappa_0$, *which is equal to the optimal T2EE established in Theorem 2 [24]. Note that* $E_c < \infty$ *under the assumption* $P_{Y_x} << P_{Y_{x'}}$, $\forall\ (x,x') \in \mathcal{X}\times\mathcal{X}$.

III. Distributed HT: Error-exponents trade-off

In [4, Theorem 1], Han and Kobayashi obtained an inner bound on $\mathcal{R}$ in the distributed setting, where the communication channel is rate-limited and noiseless. At a high level, their coding scheme involved a type-based quantization of sequences u^k, whose type P_{u^k} lies within a distance (in terms of KL divergence) equal to the desired type I error-exponent κ_α from P_U. The index of the codeword within the quantization codebook is revealed to the detector which takes the decision on the hypothesis based on the received index and side-information v^k. On the other hand, it is well-known from [9] that by performing binning subsequent to quantization and decoding using the side-information v^k can help reduce the communication rate to the detector. This enables better quantization of u^k and higher error-exponent in channel coding. However, these benefits come at a cost, as binning introduces additional errors that affect the type I and type II error probability. More specifically, a *binning error* occurs when the detector decodes the incorrect quantization codeword from the correctly decoded bin-index.

In this section, we will obtain inner bounds on $\mathcal{R}$ using the SHTCC and JHTCC schemes mentioned in Section I. As already stated, the former scheme is a separation based scheme that performs independent hypothesis testing and channel coding, while the latter scheme is a joint hypothesis testing and channel coding scheme that uses hybrid coding for communication between the observer and the detector.

A. SHTCC scheme:

The SHTCC scheme is a combination of a generalization of the Shimokawa-Han-Amari (SHA) scheme [9] and the Borade-Nakiboğlu-Zheng unequal error-protection scheme [25]. More specifically, the scheme involves

(i) quantization and binning of sequences u^k whose type P_{u^k} is within a distance of κ_α (in terms of KL-divergence) from P_U (instead of just the dominant type P_U as is done in the SHA scheme), and using the side-information v^k to decode for the quantization codeword from the (decoded) bin-index at the detector.
(ii) unequal error-protection channel coding scheme in [25] for protecting a special message which informs the detector that P_{u^k} is at a distance greater than κ_α from P_U.

Before, we state the inner bound on $\mathcal{R}$ achieved by the SHTCC scheme, some definitions are required. Let S denote a r.v. with support $\mathcal{S} = \mathcal{X}$, such that $S - X - Y$ and $P_{SXY} = P_{SX}P_{Y|X}$. For $x \in \mathcal{X}$, we define

$$r_x(y) := \log\left(\frac{P_{Y|X=x}(y)}{P_{Y|S=x}(y)}\right), \tag{23}$$

and

$$E_m(P_{SX},\theta) = \sum_{s\in\mathcal{S}} P_S(s)\psi^*_{P_{Y|S=s},r_s}(\theta). \tag{24}$$

For a fixed P_{SX} and $R \geq 0$, let $E_x(R, P_{SX}, P_{Y|X})$ denote the expurgated exponent [30][28][22] defined as follows:

$$E_x(R,P_{SX}) := E_x(R,P_{SX},P_{Y|X})$$
$$:= \max_{\rho\geq 1}\left\{-\rho\ R - \rho\ \log\left(\sum_{s,x,\tilde{x}} P_S(s)P_{X|S}(x|s)P_{X|S}(\tilde{x}|s)\left(\sum_y \sqrt{P_{Y|X}(y|x)P_{Y|X}(y|\tilde{x})}\right)^{\frac{1}{\rho}}\right)\right\}. \tag{25}$$

Let Ω denote the set of all continuous mappings from $\mathcal{P}_{\mathcal{U}}$ to $\mathcal{P}_{\mathcal{W}|\mathcal{U}}$, where $\mathcal{W}$ is an arbitrary finite set. Let

$$\theta_L(P_{SX}) := \sum_{s\in\mathcal{S}} P_S(s)D\left(P_{Y|S=s}||P_{Y|X=s}\right), \tag{26}$$

$$\theta_U(P_{SX}) := \sum_{s\in\mathcal{S}} P_S(s)D\left(P_{Y|X=s}||P_{Y|S=s}\right), \tag{27}$$

$$\Theta(P_{SX}) := (-\theta_L(P_{SX}), \theta_U(P_{SX})), \tag{28}$$

$$\mathcal{L}(\kappa_\alpha,\tau) := \left\{ \begin{array}{l} (\omega, R, P_{SX}, \theta) \in \Omega \times \mathbb{R}^+ \times \mathcal{P}_{\mathcal{SX}} \times \Theta(P_{SX}) : \zeta_q(\kappa_\alpha,\omega) - \rho(\kappa_\alpha,\omega) \leq R < \tau I(X;Y|S), \\ \min\left(\tau E_m(P_{SX},\theta), \tau E_x\left(\frac{R}{\tau}, P_{SX}\right), E_b(\kappa_\alpha,\omega,R)\right) \geq \kappa_\alpha \end{array} \right\},$$

$$\hat{\mathcal{L}}(\kappa_\alpha,\omega) := \left\{ \hat{U}\hat{V}\hat{W} : D(\hat{U}\hat{V}\hat{W}||UVW) \leq \kappa_\alpha, P_{W|U} = P_{\hat{W}|\hat{U}} = \omega(P_{\hat{U}}), V - U - W \right\}, \tag{29}$$

$$E_b(\kappa_\alpha,\omega,R) := \begin{cases} R - \zeta_q(\kappa_\alpha,\omega) + \rho(\kappa_\alpha,\omega) & \text{if } 0 \leq R < \zeta_q(\kappa_\alpha,\omega), \\ \infty & \text{otherwise}, \end{cases}$$

$$\zeta_q(\kappa_\alpha,\omega) := \max_{\substack{\hat{U}\hat{W}:\ \exists \hat{V}, \\ \hat{U}\hat{V}\hat{W} \in \hat{\mathcal{L}}(\kappa_\alpha,\omega)}} I(\hat{U};\hat{W}),$$

$$\rho(\kappa_\alpha,\omega) := \min_{\substack{\hat{V}\hat{W}:\ \exists \hat{U}, \\ \hat{U}\hat{V}\hat{W} \in \hat{\mathcal{L}}(\kappa_\alpha,\omega)}} I(\hat{V};\hat{W}),$$

$$E_1(\kappa_\alpha,\omega) := \min_{\tilde{U}\tilde{V}\tilde{W} \in \mathcal{T}_1(\kappa_\alpha,\omega)} D(\tilde{U}\tilde{V}\tilde{W}||\bar{U}\bar{V}\bar{W}),$$

$$E_2(\kappa_\alpha,\omega,R) := \begin{cases} \min\limits_{\tilde{U}\tilde{V}\tilde{W} \in \mathcal{T}_2(\kappa_\alpha,\omega)} D(\tilde{U}\tilde{V}\tilde{W}||\bar{U}\bar{V}\bar{W}) + E_b(\kappa_\alpha,\omega,R), & \text{if } R < \zeta_q(\kappa_\alpha,\omega), \\ \infty, & \text{otherwise}, \end{cases}$$

$$E_3(\kappa_\alpha,\omega,R,P_{SX},\tau) := \begin{cases} \min\limits_{\tilde{U}\tilde{V}\tilde{W} \in \mathcal{T}_3(\kappa_\alpha,\omega)} D(\tilde{U}\tilde{V}\tilde{W}||\bar{U}\bar{V}\bar{W}) + E_b(\kappa_\alpha,\omega,R) + \tau E_x\left(\frac{R}{\tau}, P_{SX}\right), & \text{if } R < \zeta_q(\kappa_\alpha,\omega), \\ \min\limits_{\tilde{U}\tilde{V}\tilde{W} \in \mathcal{T}_3(\kappa_\alpha,\omega)} D(\tilde{U}\tilde{V}\tilde{W}||\bar{U}\bar{V}\bar{W}) + \rho(\kappa_\alpha,\omega) + \tau E_x\left(\frac{R}{\tau}, P_{SX}\right), & \text{otherwise}, \end{cases}$$

$$E_4(\kappa_\alpha,\omega,R,P_{SX},\theta,\tau) := \begin{cases} \min\limits_{\hat{V}:\hat{U}\hat{V}\hat{W} \in \hat{\mathcal{L}}(\kappa_\alpha,\omega)} D(\hat{V}||\bar{V}) + E_b(\kappa_\alpha,\omega,R) + \tau\left(E_m(P_{SX},\theta) - \theta\right), & \text{if } R < \zeta_q(\kappa_\alpha,\omega), \\ \min\limits_{\hat{V}:\hat{U}\hat{V}\hat{W} \in \hat{\mathcal{L}}(\kappa_\alpha,\omega)} D(\hat{V}||\bar{V}) + \rho(\kappa_\alpha,\omega) + \tau\left(E_m(P_{SX},\theta) - \theta\right), & \text{otherwise}, \end{cases}$$

where,

$$P_{\bar{W}|\bar{U}} := P_{\tilde{W}|\tilde{U}},\ \bar{V} - \bar{U} - \bar{W},$$

$$\mathcal{T}_1(\kappa_\alpha,\omega) := \left\{ \tilde{U}\tilde{V}\tilde{W} :\ P_{\tilde{U}\tilde{W}} = P_{\hat{U}\hat{W}},\ P_{\tilde{V}\tilde{W}} = P_{\hat{V}\hat{W}} \text{ for some } \hat{U}\hat{V}\hat{W} \in \hat{\mathcal{L}}(\kappa_\alpha,\omega) \right\},$$

$$\mathcal{T}_2(\kappa_\alpha,\omega) := \left\{ \tilde{U}\tilde{V}\tilde{W} :\ P_{\tilde{U}\tilde{W}} = P_{\hat{U}\hat{W}},\ P_{\tilde{V}} = P_{\hat{V}},\ H(\tilde{W}|\tilde{V}) \geq H(\hat{W}|\hat{V}) \text{ for some } \hat{U}\hat{V}\hat{W} \in \hat{\mathcal{L}}(\kappa_\alpha,\omega) \right\},$$

and $\mathcal{T}_3(\kappa_\alpha,\omega) := \left\{ \tilde{U}\tilde{V}\tilde{W} :\ P_{\tilde{U}\tilde{W}} = P_{\hat{U}\hat{W}},\ P_{\tilde{V}} = P_{\hat{V}} \text{ for some } \hat{U}\hat{V}\hat{W} \in \hat{\mathcal{L}}(\kappa_\alpha,\omega) \right\}.$

We have the following lower bound for $\kappa(\tau,\kappa_\alpha)$.

Theorem 4. *$\kappa(\tau,\kappa_\alpha) \geq \kappa_s^*(\tau,\kappa_\alpha)$, where*

$$\kappa_s^*(\tau,\kappa_\alpha) := \max_{\substack{(\omega,R,P_{SX},\theta) \\ \in \mathcal{L}(\kappa_\alpha,\tau)}} \min\left(E_1(\kappa_\alpha,\omega), E_2(\kappa_\alpha,\omega,R), E_3(\kappa_\alpha,\omega,R,P_{SX},\tau), E_4(\kappa_\alpha,\omega,R,P_{SX},\theta,\tau)\right). \tag{30}$$

The proof of Theorem 4 is presented in Appendix B. As a corollary, Theorem 4 recovers the lower bound for $\kappa(\tau,\kappa_\alpha)$ obtained in [4] for the case of a rate-limited noiseless channel by

1) setting $E_x\left(\frac{R}{\tau}, P_{SX}\right)$, $E_m(P_{SX},\theta)$ and $E_m(P_{SX},\theta) - \theta$ to ∞, which holds when the channel is noiseless.
2) maximizing over the set $\{(\omega,R,P_{SX},\theta) \in \Omega \times \mathbb{R}^+ \times \mathcal{P}_{\mathcal{SX}} \times \Theta(P_{SX}) : \zeta_q(\kappa_\alpha,\omega) \leq R < \tau I(X;Y|S)\} \subseteq \mathcal{L}(\kappa_\alpha,\tau,P_{Y|X})$ in (30).

Then, note that the terms $E_2(\kappa_\alpha,\omega,R)$, $E_3(\kappa_\alpha,\omega,R,P_{SX},\tau)$ and $E_4(\kappa_\alpha,\omega,R,P_{SX},\theta,\tau)$ all equal ∞, and thus the inner bound in Theorem 4 reduces to that given in [4, Theorem 1].

Remark 2. *Since the lower bound on $\kappa(\tau,\kappa_\alpha)$ in Theorem 4 is not necessarily concave, a tighter bound can be obtained using the technique of "time-sharing" similar to [4, Theorem 3]. We omit its description as it is cumbersome, although straightforward.*

Specializing the lower bound in Theorem 4 to the case of *testing against independence*, we obtain the following.

Corollary 1. *Let $P_{\bar{U}\bar{V}} = P_U P_V$. Then,*

$$\kappa(\tau,\kappa_\alpha) \geq \max_{\substack{(\omega,R,P_{SX},\theta) \in \\ \mathcal{L}^*(\kappa_\alpha,\tau)}} \min\left(E_1^I(\kappa_\alpha,\omega), E_2^I(\kappa_\alpha,\omega,R,P_{SX},\tau), E_3^I(\kappa_\alpha,\omega,R,P_{SX},\theta,\tau)\right),$$

where

$$\mathcal{L}^*(\kappa_\alpha,\tau) := \left\{ \begin{array}{l} (\omega, R, P_{SX}, \theta) \in \Omega \times \mathbb{R}^+ \times \mathcal{P}_{\mathcal{SX}} \times \Theta(P_{SX}) : \zeta_q(\kappa_\alpha,\omega) \leq R < \tau I(X;Y|S), \\ \min\left(\tau E_m(P_{SX},\theta), \tau E_x\left(\frac{R}{\tau}, P_{SX}\right)\right) \geq \kappa_\alpha \end{array} \right\},$$

$$E_1^I(\kappa_\alpha,\omega) := \min_{\hat{V}\hat{W}: \hat{U}\hat{V}\hat{W} \in \hat{\mathcal{L}}(\kappa_\alpha,\omega)} \left[I(\hat{V};\hat{W}) + D(\hat{V}||V) \right],$$

$$E_2^I(\kappa_\alpha,\omega,R,P_{SX},\tau) := \min_{\hat{V}: \hat{U}\hat{V}\hat{W} \in \hat{\mathcal{L}}(\kappa_\alpha,\omega)} D(\hat{V}||V) + \rho(\kappa_\alpha,\omega) + \tau E_x\left(\frac{R}{\tau}, P_{SX}\right),$$

$$E_3^I(\kappa_\alpha,\omega,P_{SX},\theta,\tau) := \min_{\hat{V}: \hat{U}\hat{V}\hat{W} \in \hat{\mathcal{L}}(\kappa_\alpha,\omega)} D(\hat{V}||V) + \rho(\kappa_\alpha,\omega) + \tau\left(E_m(P_{SX},\theta) - \theta\right),$$

and $\hat{\mathcal{L}}(\kappa_\alpha,\omega)$ *is as defined in* (29).

Proof: Note that $\mathcal{L}^*(\kappa_\alpha,\tau) \subseteq \mathcal{L}(\kappa_\alpha,\tau)$. Then, for any $\omega \in \mathcal{L}^*(\kappa_\alpha,\tau)$ and any $\tilde{U}\tilde{V}\tilde{W} \in \mathcal{T}_1(\kappa_\alpha,\omega)$,

$$\begin{aligned} D(\tilde{U}\tilde{V}\tilde{W}||\bar{U}\bar{V}\bar{W}) &= D(\tilde{U}\tilde{W}||\bar{U}\bar{W}) + D(\tilde{V}|\tilde{U}\tilde{W}||\bar{V}|\bar{U}\bar{W}) \\ &= D(\tilde{U}||\bar{U}) + D(\tilde{V}|\tilde{U}\tilde{W}||\bar{V}) \\ &\geq D(\tilde{U}||\bar{U}) + D(\tilde{V}|\tilde{W}||\bar{V}) \qquad (31) \\ &= D(\hat{U}||\bar{U}) + D(\hat{V}|\hat{W}||\bar{V}) \qquad (32) \\ &= D(\hat{U}||U) + I(\hat{V};\hat{W}) + D(\hat{V}||V), \end{aligned}$$

where, in (31), we used the data processing (DPI) inequality for KL-divergence, and in (32), we used the fact that for $\tilde{U}\tilde{V}\tilde{W} \in \mathcal{T}_1(\kappa_\alpha,\omega)$, $P_{\tilde{U}\tilde{W}} = P_{\hat{U}\hat{W}}$ and $P_{\tilde{V}\tilde{W}} = P_{\hat{V}\hat{W}}$ for some $\hat{U}\hat{V}\hat{W} \in \hat{\mathcal{L}}(\kappa_\alpha,\omega)$. Minimizing over all $\hat{U}\hat{V}\hat{W} \in \hat{\mathcal{L}}(\kappa_\alpha,\omega)$ yields that

$$E_1(\kappa_\alpha,\omega) = \min_{\tilde{U}\tilde{V}\tilde{W} \in \mathcal{T}_1(\kappa_\alpha,\omega)} D(\tilde{U}\tilde{V}\tilde{W}||\bar{U}\bar{V}\bar{W}) \geq \min_{\hat{U}\hat{V}\hat{W} \in \hat{\mathcal{L}}(\kappa_\alpha,\omega)} \left[I(\hat{V};\hat{W}) + D(\hat{V}||V) \right] = E_1^I(\kappa_\alpha,\omega).$$

Since $\zeta_q(\kappa_\alpha,\omega) \leq R$, we have that $E_2(\kappa_\alpha,\omega,R) = \infty$,

$$\begin{aligned} E_3(\kappa_\alpha,\omega,R,P_{SX},\tau) &= \min_{\tilde{U}\tilde{V}\tilde{W} \in \mathcal{T}_2(\kappa_\alpha,\omega)} D(\tilde{U}\tilde{V}\tilde{W}||\bar{U}\bar{V}\bar{W}) + \rho(\kappa_\alpha,\omega) + \tau E_x\left(\frac{R}{\tau}, P_{SX}\right) \\ &\geq \min_{\hat{V}: \hat{U}\hat{V}\hat{W} \in \hat{\mathcal{L}}(\kappa_\alpha,\omega)} D(\hat{V}||V) + \rho(\kappa_\alpha,\omega) + \tau E_x\left(\frac{R}{\tau}, P_{SX}\right) \qquad (33) \\ &= E_2^I(\kappa_\alpha,\omega,R,P_{SX},\tau), \end{aligned}$$

and

$$\begin{aligned} E_4(\kappa_\alpha,\omega,R,P_{SX},\theta,\tau) &:= \min_{\hat{V}: \hat{U}\hat{V}\hat{W} \in \hat{\mathcal{L}}(\kappa_\alpha,\omega)} D(\hat{V}||\bar{V}) + \rho(\kappa_\alpha,\omega) + \tau\left(E_m(P_{SX},\theta) - \theta\right) \\ &\geq \min_{\hat{V}: \hat{U}\hat{V}\hat{W} \in \hat{\mathcal{L}}(\kappa_\alpha,\omega)} D(\hat{V}||V) + \rho(\kappa_\alpha,\omega) + \tau\left(E_m(P_{SX},\theta) - \theta\right) \qquad (34) \\ &= E_3^I(\kappa_\alpha,\omega,P_{SX},\theta,\tau), \end{aligned}$$

where, to obtain (33) and (34), we used DPI for KL-divergence. This completes the proof. ■

Corollary 2.

$$\lim_{\kappa_\alpha \to 0} \kappa_s^*(\tau,\kappa_\alpha) = \kappa_s(\tau),$$

where $\kappa_s(\tau)$ *is the lower bound on the type II error-exponent for a fixed type I error probability constraint established in [22, Theorem 2].*

Proof: The result follows by noting that

$$\begin{aligned} \hat{\mathcal{L}}(0,\omega) &= \{UVW, P_{W|U} = \omega(P_U),\ V - U - W\}, \\ \zeta_q(0,\omega) &= I(U;W), \\ \rho(0,\omega) &= I(V;W), \end{aligned}$$

and the fact that $E_m(P_{SX},\theta), E_x\left(\frac{R}{\tau}, P_{SX}\right), E_b(\kappa_\alpha,\omega,R)$, and $E_m(P_{SX},\theta) - \theta$ for $\theta \in \Theta(P_{SX})$, are all greater than or equal to zero. ■

The optimal T2EE for testing against independence in the Stein's regime (i.e. when $\kappa_\alpha \to 0$) can be recovered from Corollary 1 by taking the limit $\kappa_\alpha \to 0$.

Corollary 3. *Let $P_{\bar{U}\bar{V}} = P_U P_V$. Then,*

$$\lim_{\kappa_\alpha \to 0} \kappa(\tau, \kappa_\alpha) = \max_{\substack{W: W - U - V \\ I(U;W) \le \tau C}} I(V;W).$$

Proof: Note that

$$\hat{\mathcal{L}}(0,\omega) := \{UVW : P_{W|U} = \omega(P_U), V - U - W\},$$

and

$$\mathcal{L}^*(0,\tau) := \{(\omega, R, P_{SX}, \theta) \in \Omega \times \mathbb{R}^+ \times \mathcal{P}_{\mathcal{S}\mathcal{X}} \times \Theta(P_{SX}) : \; I(U;W) \le R < \tau I(X;Y|S), \; P_{W|U} = \omega(P_U)\}.$$

Hence,

$$E_1^I(0,\omega) \ge \min_{\hat{U}\hat{V}\hat{W} \in \hat{\mathcal{L}}(0,\omega)} I(\hat{V};\hat{W}) = I(V;W), \tag{35}$$

for some $V - U - W$ such that $P_{W|U} = \omega(P_U)$. Also, we have

$$\rho(0,\omega) = I(V;W),$$

$$E_2^I(0,\omega,R,P_{SX},\tau) \ge \rho(0,\omega), \tag{36}$$

$$E_3^I(0,\omega,P_{SX},\theta,\tau) \ge \rho(0,\omega). \tag{37}$$

From (35)-(37), the result follows. ■

B. JHTCC scheme

It is well known that joint source channel coding schemes outperforms separation based coding schemes in the context of reliable communication over a noisy channel [31], [32], [33]. Recently, it is shown in [22] that a JHTCC scheme based on hybrid coding is atleast as good as a separation based scheme in general, when the objective is to maximize the type II error-exponent for a fixed type I error probability constraint. Here, we obtain an inner bound on $\mathcal{R}$ using a generalization of the JHTCC scheme in [22].

For simplicity, we will assume that $k = n$, i.e., $\tau = 1$. Let Ω' denote the set of all continuous mappings from $\mathcal{P}_{\mathcal{U}} \times \mathcal{P}_{\mathcal{S}}$ to $\mathcal{P}_{\mathcal{W}'|\mathcal{U}\mathcal{S}}$, where $\mathcal{W}'$ is an arbitrary finite set. Let

$$\mathcal{L}_h(\kappa_\alpha) := \left\{ \left(P_S, \omega'(\cdot, P_S), P_{X|USW'}, P_{X'|S}\right) \in \; \mathcal{P}_{\mathcal{S}} \times \Omega' \times \mathcal{P}_{\mathcal{X}|\mathcal{U}\mathcal{S}\mathcal{W}'} \times \mathcal{P}_{\mathcal{X}|\mathcal{S}} : \; E_b'(\kappa_\alpha, \omega', P_S, P_{X|USW'}) > \kappa_\alpha \right\},$$

$$\hat{\mathcal{L}}_h(\kappa_\alpha, \omega', P_S, P_{X|USW'})$$

$$:= \left\{ \begin{array}{l} \hat{U}\hat{V}\hat{W}\hat{Y}S : D(\hat{U}\hat{V}\hat{W}\hat{Y}||UVW'Y|S) \le \kappa_\alpha, \; P_{SUVW'XY} := P_S P_{UV} P_{W'|US} P_{X|USW'} P_{Y|X}, \\ \qquad\qquad P_{W'|US} = P_{\hat{W}|\hat{U}S} = \omega'(P_{\hat{U}}, P_S) \end{array} \right\},$$

$$E_b'(\kappa_\alpha, \omega', P_S, P_{X|USW'}) := \rho'(\kappa_\alpha, \omega', P_S, P_{X|USW'}) - \zeta_q'(\kappa_\alpha, \omega', P_S),$$

$$\zeta_q'(\kappa_\alpha, \omega', P_S) := \max_{\substack{\hat{U}\hat{W}S: \, \exists \, \hat{V}\hat{Y}, \\ \hat{U}\hat{V}\hat{W}\hat{Y}S \, \in \\ \hat{\mathcal{L}}_h(\kappa_\alpha, \omega', P_S, P_{X|USW'})}} I(\hat{U};\hat{W}|S),$$

$$\rho'(\kappa_\alpha, \omega', P_S, P_{X|USW'}) := \min_{\substack{\hat{V}\hat{W}\hat{Y}S: \, \exists \, \hat{U}, \\ \hat{U}\hat{V}\hat{W}\hat{Y}S \, \in \\ \hat{\mathcal{L}}_h(\kappa_\alpha, \omega', P_S, P_{X|USW'})}} I(\hat{Y},\hat{V};\hat{W}|S),$$

$$E_1'(\kappa_\alpha, \omega') := \min_{\tilde{U}\tilde{V}\tilde{W}\tilde{Y}S \, \in \, \mathcal{T}_1'(\kappa_\alpha, \omega')} D(\tilde{U}\tilde{V}\tilde{W}\tilde{Y}||\bar{U}\bar{V}\bar{W}'\bar{Y}|S),$$

$$E_2'(\kappa_\alpha, \omega', P_S, P_{X|USW'}) := \min_{\substack{\tilde{U}\tilde{V}\tilde{W}\tilde{Y}S \, \in \\ \mathcal{T}_2'(\kappa_\alpha, \omega', P_S, P_{X|USW'})}} D(\tilde{U}\tilde{V}\tilde{W}\tilde{Y}||\bar{U}\bar{V}\bar{W}'\bar{Y}|S) + E_b'(\kappa_\alpha, \omega', P_S, P_{X|USW'}),$$

$$E_3'(\kappa_\alpha, \omega', P_S, P_{X|USW'}, P_{X'|S}) := \min_{\substack{\hat{V}\hat{Y}S: \hat{U}\hat{V}\hat{W}\hat{Y}S \, \in \\ \hat{\mathcal{L}}_h(\kappa_\alpha, \omega', P_S, P_{X|USW'})}} D(\hat{V}\hat{Y}||\bar{V}\check{Y}|S) + E_b'(\kappa_\alpha, \omega', P_S, P_{X|USW'}),$$

$$P_{S\bar{U}\bar{V}\bar{W}'\bar{X}\bar{Y}} := P_S P_{\bar{U}\bar{V}} P_{\bar{W}'|\bar{U}S} P_{\bar{X}|\bar{U}S\bar{W}'} P_{\bar{Y}|\bar{X}}, \; P_{\bar{W}'|\bar{U}S} := P_{\tilde{W}|\tilde{U}S}, \; P_{\bar{X}|\bar{U}S\bar{W}'} := P_{X|USW'}, \; P_{\bar{Y}|\bar{X}} := P_{Y|X},$$

$$P_{S\bar{U}\bar{V}X'\check{Y}} := P_S P_{\bar{U}\bar{V}} P_{X'|S} P_{\check{Y}|X'},\ P_{\check{Y}|X'} := P_{Y|X},$$

$$\mathcal{T}_1'(\kappa_\alpha, \omega', P_S, P_{X|USW'}) := \left\{ \begin{array}{ll} \tilde{U}\tilde{V}\tilde{W}\tilde{Y}S: & P_{\tilde{U}\tilde{W}S} = P_{\hat{U}\hat{W}S},\ P_{\tilde{V}\tilde{W}\tilde{Y}S} = P_{\hat{V}\hat{W}\hat{Y}S}, \\ & \text{for some } \hat{U}\hat{V}\hat{W}\hat{Y}S \in \hat{\mathcal{L}}_h(\kappa_\alpha, \omega', P_S, P_{X|USW'}) \end{array} \right\},$$

$$\mathcal{T}_2'(\kappa_\alpha, \omega', P_S, P_{X|USW'}) := \left\{ \begin{array}{ll} \tilde{U}\tilde{V}\tilde{W}\tilde{Y}S: & P_{\tilde{U}\tilde{W}S} = P_{\hat{U}\hat{W}S},\ P_{\tilde{V}\tilde{Y}S} = P_{\hat{V}\hat{Y}S},\ H(\tilde{W}|\tilde{V},\tilde{Y},S) \geq H(\hat{W}|\hat{V},\hat{Y},S), \\ & \text{for some } \hat{U}\hat{V}\hat{W}\hat{Y}S \in \hat{\mathcal{L}}_h(\kappa_\alpha, \omega', P_S, P_{X|USW'}) \end{array} \right\}.$$

Then, we have the following result.

Theorem 5.

$\kappa(1, \kappa_\alpha) \geq \kappa_h^*(\kappa_\alpha)$, where

$$\kappa_h^*(\kappa_\alpha) := \max_{\substack{(P_S, \omega', P_{X|USW'}, P_{X'|S}) \\ \in \mathcal{L}_h(\kappa_\alpha)}} \min \left(E_1'(\kappa_\alpha, \omega'),\ E_2'(\kappa_\alpha, \omega', P_S, P_{X|USW'}),\ E_3'(\kappa_\alpha, \omega', P_S, P_{X|USW'}, P_{X'|S}) \right).$$

The proof of Theorem 5 is given in Appendix C. It is easy to see that Theorem 5 recovers the lower bound on the type II error-exponent for a given constraint on the type I error probability proved in [22, Theorem 5].

Corollary 4.

$$\lim_{\kappa_\alpha \to 0} \kappa_h^*(\kappa_\alpha) = \kappa_h,$$

where κ_h is as given in [22, Theorem 5].

Proof: The result follows by noting that

$$\hat{\mathcal{L}}_h(0, \omega', P_S, P_{X|USW'}) := \{UVW'YS:\ P_{SUVW'XY} := P_S P_{UV} P_{W'|US} P_{X|USW'} P_{Y|X},\ P_{W'|US} = \omega'(P_U, P_S)\}$$

$$\zeta_q'(0, \omega', P_S) := I(U; W'|S),$$

$$\rho(0, \omega', P_S, P_{X|USW'}) := I(Y, V; W'|S),$$

$$\mathcal{L}_h(0) := \{\left(P_S, \omega'(\cdot, P_S), P_{X|USW'}, P_{X'|S}\right) \in\ \mathcal{P}_{\mathcal{S}} \times \Omega' \times \mathcal{P}_{\mathcal{X}|\mathcal{USW}'} \times \mathcal{P}_{\mathcal{X}|\mathcal{S}} : I(U; W'|S) < I(Y, V; W'|S)\},$$

and $E_b'(0, \omega', P_S, P_{X|USW'}) := I(Y, V; W'|S) - I(U; W'|S)$.

■

It is shown in [22, Theorem 7] that $\kappa_h \geq \kappa_s(\tau)$. This implies that the lower bound on the type II error-exponent achieved by hybrid coding is at least as good as that achieved by the SHTCC scheme, when the type I error-exponent is zero.

IV. Conclusion

In this paper, we studied the trade-off between the exponents of the type I and type II error probabilities for distributed HT over a noisy channel with side-information at the detector. In the non-distributed setting, we obtained a single-letter characterization of the optimal trade-off between the error-exponents. The direct part of the proof shows that the optimal trade-off is achieved by a scheme, in which the observer performs an appropriate NP test locally and communicates the decision of the test to the detector using a suitable channel code, while the detector performs an appropriate NP test on the channel output. This implies that "separation" holds, in the sense that, there is no loss in optimality incurred by separating the tasks of HT and channel coding. For the distributed setting, we obtained inner bounds on the error-exponents trade-off using the SHTCC and JHTCC schemes. The latter bound is at least as good as the former when the type I error-exponent is zero. Exploring whether joint schemes offer strict advantage over separation based schemes is something worth investigating. It would also be interesting to explore the trade-off between the error-exponents in a related setting, where the side-information also needs to be communicated to the detector over a noisy communication channel.

Appendix A
Proof of Proposition 1

First, we show the proof of achievability, i.e., for $-D(P_{Y_{X_0}}||P_{Y_{X_1}}|P_{X_0X_1}) < \theta \leq D(P_{Y_{X_1}}||P_{Y_{X_0}}|P_{X_0X_1})$,

$$\kappa\left(\mathbb{E}_{P_{X_0X_1}}\left[\psi^*_{Y_{X_0}, \tilde{h}_{X_0,X_1}}(\theta)\right], x_0^n, x_1^n\right) \geq \mathbb{E}_{P_{X_0X_1}}\left[\psi^*_{Y_{X_0}, \tilde{h}_{X_0,X_1}}(\theta)\right] - \theta.$$

Let $\tilde{h}'_{\mathcal{Y}^n}: \mathcal{Y}^n \to \mathbb{R}$ defined as

$$\tilde{h}'_{\mathcal{Y}^n}(y^n) := \log\left(\frac{P_{Y^n|X^n=x_1^n}(y^n)}{P_{Y^n|X^n=x_0^n}(y^n)}\right).$$

For the decision rule $g^{(n)}_{\theta,\mathcal{Y}}$ defined in (14), the type I error probability can be upper bounded for $\theta > -D(P_{Y_{X_0}}||P_{Y_{X_1}}|P_{X_0X_1})$ as follows:

$$\begin{aligned}\alpha\left(n, g^{(n)}_{\theta,\mathcal{Y}}, x_0^n, x_1^n\right) &= P_{Y^n|X^n=x_0^n}\left(\log\left(\frac{P_{Y^n|X^n=x_1^n}(Y^n)}{P_{Y^n|X^n=x_0^n}(Y^n)}\right) \geq n\theta\right)\\ &\leq e^{-\sup_{\lambda\geq 0}\left(n\theta\lambda - \psi_{P_{Y^n|X^n=x_0^n},\tilde{h}'_{\mathcal{Y}^n}}(\lambda)\right)} &(38)\\ &= e^{-\sup_{\lambda\in\mathbb{R}}\left(n\left(\theta\lambda - \frac{1}{n}\psi_{P_{Y^n|X^n=x_0^n},\tilde{h}'_{\mathcal{Y}^n}}(\lambda)\right)\right)}. &(39)\end{aligned}$$

Here, (38) follows using the standard Chernoff bound. Eqn. (39) follows due to the fact that for $\theta > -D(P_{Y_{X_0}}||P_{Y_{X_1}}|P_{X_0X_1})$, the supremum in (38) is always achieved at $\lambda \geq 0$, which in turn follows from Lemma 1 (i) and (ii).

Simplifying the term within the exponent in (39), we obtain

$$\begin{aligned}\frac{1}{n}\psi_{P_{Y^n|X^n=x_0^n},\tilde{h}'_{\mathcal{Y}^n}}(\lambda) &:= \frac{1}{n}\log\left(\mathbb{E}_{P_{Y^n|X^n=x_0^n}}\left(\frac{P^{\lambda}_{Y^n|X^n=x_1^n}(Y^n)}{P^{\lambda}_{Y^n|X^n=x_0^n}(Y^n)}\right)\right)\\ &= \frac{1}{n}\log\left(\mathbb{E}_{P_{Y^n|X^n=x_0^n}}\left(\prod_{i=1}^{n}\frac{P^{\lambda}_{Y_i|X_i=x_{1i}}(Y_i)}{P^{\lambda}_{Y_i|X_i=x_{0i}}(Y_i)}\right)\right)\\ &= \frac{1}{n}\log\left(\prod_{i=1}^{n}\mathbb{E}_{P_{Y_i|X_i=x_{0i}}}\left(\frac{P^{\lambda}_{Y_i|X_i=x_{1i}}(Y_i)}{P^{\lambda}_{Y_i|X_i=x_{0i}}(Y_i)}\right)\right)\\ &= \frac{1}{n}\sum_{i=1}^{n}\log\left(\mathbb{E}_{P_{Y_i|X_i=x_{0i}}}\left(\frac{P^{\lambda}_{Y_i|X_i=x_{1i}}(Y_i)}{P^{\lambda}_{Y_i|X_i=x_{0i}}(Y_i)}\right)\right)\\ &= \sum_{x,x'}P_{x_0^n x_1^n}(x,x')\log\left(\mathbb{E}_{P_{Y_x}}\left(\frac{P^{\lambda}_{Y_{x'}}(Y)}{P^{\lambda}_{Y_x}(Y)}\right)\right) &(40)\\ &\xrightarrow{(n)} \mathbb{E}_{P_{X_0X_1}}\left[\log\left(\mathbb{E}_{P_{Y_{X_0}}}\left(e^{\lambda\tilde{h}_{X_0,X_1}(Y)}\right)\right)\right], &(41)\end{aligned}$$

where, (41) follows from (10) and Assumption 1. Substituting (41) in (39) and using (7), we obtain for arbitrarily small but fixed $\delta > 0$ and sufficiently large n that

$$\begin{aligned}\alpha\left(n, g^{(n)}_{\theta,\mathcal{Y}}, x_0^n, x_1^n\right) &\leq e^{-\sup_{\lambda\in\mathbb{R}}\left(n\left(\theta\lambda - \mathbb{E}_{P_{X_0X_1}}\left[\log\left(\mathbb{E}_{P_{Y_{X_0}}}\left(e^{\lambda\tilde{h}_{X_0,X_1}(Y)}\right)\right)\right]-\delta\right)\right)}\\ &= e^{-n\left(\mathbb{E}_{P_{X_0X_1}}\left[\sup_{\lambda\in\mathbb{R}}\left(\theta\lambda - \mathbb{E}_{P_{Y_{X_0}}}\left(e^{\lambda\tilde{h}_{X_0,X_1}(Y)}\right)\right)\right]-\delta\right)}\\ &= e^{-n\left(\mathbb{E}_{P_{X_0X_1}}\left[\psi^*_{Y_{X_0},\tilde{h}_{X_0,X_1}}(\theta)\right]-\delta\right)}. &(42)\end{aligned}$$

Similarly, we can show that for $\theta \leq D(P_{Y_{X_1}}||P_{Y_{X_0}}|P_{X_0X_1})$,

$$\beta\left(n, g^{(n)}_{\theta,\mathcal{Y}}, x_0^n, x_1^n\right) \leq e^{-n\left(\mathbb{E}_{P_{X_0X_1}}\left[\psi^*_{Y_{X_1},\tilde{h}_{X_0,X_1}}(\theta)\right]-\delta\right)}. \tag{43}$$

We also have

$$\psi_{Y_{x'},\tilde{h}_{x,x'}}(\lambda) = \sum_{y\in\mathcal{Y}}P_{Y_{x'}}\frac{P^{\lambda}_{Y_{x'}}}{P^{\lambda}_{Y_x}} = \sum_{y\in\mathcal{Y}}P_{Y_x}\frac{P^{\lambda+1}_{Y_{x'}}}{P^{\lambda+1}_{Y_x}} = \psi_{Y_x,\tilde{h}_{x,x'}}(\lambda+1).$$

It follows that

$$\psi^*_{Y_{x'},\tilde{h}_{x,x'}}(\theta) := \sup_{\lambda\in\mathbb{R}}\left(\lambda\theta - \psi_{Y_{x'},\tilde{h}_{x,x'}}(\lambda)\right) = \sup_{\lambda\in\mathbb{R}}\left(\lambda\theta - \psi_{Y_x,\tilde{h}_{x,x'}}(\lambda+1)\right) = \psi^*_{Y_x,\tilde{h}_{x,x'}}(\theta) - \theta.$$

Hence,

$$\mathbb{E}_{P_{X_0X_1}}\left[\psi^*_{Y_{X_1},\tilde{h}_{X_0,X_1}}(\theta)\right]=\mathbb{E}_{P_{X_0X_1}}\left[\psi^*_{Y_{X_0},\tilde{h}_{X_0,X_1}}(\theta)\right]-\theta. \tag{44}$$

From (42), (43) and (44), it follows that for $-D(P_{Y_{X_0}}||P_{Y_{X_1}}|P_{X_0X_1})<\theta\leq D(P_{Y_{X_1}}||P_{Y_{X_0}}|P_{X_0X_1})$,

$$\bar{\kappa}\left(\mathbb{E}_{P_{X_0X_1}}\left[\psi^*_{Y_{X_0},\tilde{h}_{X_0,X_1}}(\theta)\right]-\delta,P_{X_0X_1}\right)\geq\mathbb{E}_{P_{X_0X_1}}\left[\psi^*_{Y_{X_0},\tilde{h}_{X_0,X_1}}(\theta)\right]-\theta-\delta.$$

Noting that $\delta>0$ is arbitrary and $\bar{\kappa}(\kappa_\alpha,P_{X_0X_1})$ is a continuous function of κ_α for a fixed $P_{X_0X_1}$, the proof of achievability is complete.

Next, we prove the converse. Let

$$\mu_n(x,x'):=P_{x_0^nx_1^n}(x,x'),$$
$$\text{and } \mathcal{I}_n(x,x'):=\{i\in[n]\text{ s.t. } x_{0i}=x\text{ and } x_{1i}=x'\}.$$

Denoting $\alpha\left(n,g^{(k,n)},x_0^n,x_1^n\right)$ and $\beta\left(n,g^{(k,n)},x_0^n,x_1^n\right)$ by α_n and β_n, respectively, we obtain that for any $\theta\in\mathbb{R}$,

$$\begin{aligned}
\alpha_n+e^{-n\theta}\beta_n &\geq P_{Y^n|X^n=x_0^n}\left(\log\left(\frac{P_{Y^n|X^n=x_1^n}(Y^n)}{P_{Y^n|X^n=x_0^n}(Y^n)}\right)\geq n\theta\right) &(45)\\
&=P_{Y^n|X^n=x_0^n}\left(\sum_{i=1}^n\log\left(\frac{P_{Y_i|X=x_{1i}}(Y_i)}{P_{Y_i|X=x_{0i}}(Y_i)}\right)\geq n\theta\right)\\
&=P_{Y^n|X^n=x_0^n}\left(\sum_{x,x'}\sum_{i\in\mathcal{I}_n(x,x')}\log\left(\frac{P_{Y_i|X=x_{1i}}(Y_i)}{P_{Y_i|X=x_{0i}}(Y_i)}\right)\geq n\theta\right)\\
&=P_{Y^n|X^n=x_0^n}\left(\sum_{x,x'}\sum_{i\in\mathcal{I}_n(x,x')}\log\left(\frac{P_{Y_i|X=x_{1i}}(Y_i)}{P_{Y_i|X=x_{0i}}(Y_i)}\right)\geq\sum_{(x,x')\in\mathcal{X}\times\mathcal{X}}n\mu_n(x,x')\theta\right)\\
&\geq P_{Y^n|X^n=x_0^n}\left(\bigcap_{x,x'}\left(\sum_{i\in\mathcal{I}_n(x,x')}\log\left(\frac{P_{Y_i|X=x_{1i}}(Y_i)}{P_{Y_i|X=x_{0i}}(Y_i)}\right)\geq n\mu_n(x,x')\theta\right)\right)\\
&=\prod_{(x,x')\in\mathcal{X}\times\mathcal{X}}P_{Y^n|X^n=x_0^n}\left(\sum_{i\in\mathcal{I}_n(x,x')}\log\left(\frac{P_{Y_i|X=x_{1i}}(Y_i)}{P_{Y_i|X=x_{0i}}(Y_i)}\right)\geq n\mu_n(x,x')\theta\right).
\end{aligned}$$

Here, (45) follows by applying Theorem 2. Then, for arbitrary $\delta>0$ and sufficiently large n, we can write

$$\begin{aligned}
\alpha_n+e^{-n\theta}\beta_n &\geq\prod_{(x,x')\in\mathcal{X}\times\mathcal{X}}e^{-n\mu_n(x,x')\left(\left(\min_{\tilde{Q}_x:\mathbb{E}_{\tilde{Q}_x}(Y)\geq\theta}D\left(\tilde{Q}_x||P_{Y_x}\right)\right)+\delta\right)} &(46)\\
&\geq\prod_{(x,x')\in\mathcal{X}\times\mathcal{X}}e^{-n\mu_n(x,x')\left(\psi^*_{Y_x,\tilde{h}_{x,x'}}(\theta)+\delta\right)} &(47)\\
&=e^{-n\left(\mathbb{E}_{P_{X_0X_1}}\left(\psi^*_{Y_{X_0},\tilde{h}_{X_0,X_1}}(\theta)\right)+\delta'\right)}, &(48)
\end{aligned}$$

where, $\delta'>\delta$ is arbitrary. Here, (46) follows from [29, Theorem 14.1]; (47) follows from [29, Theorem 13.3] and [29, Theorem 14.3]; and (48) follows from (10). Note that (48) holds even if $\psi^*_{Y_x,\tilde{h}_{x,x'}}(\theta)=\infty$ for some $x,x'\in\mathcal{X}\times\mathcal{X}$ and $\theta>0$ since in this case, both (47) and (48) equal 0. Equation (48) implies that

$$\limsup_{n\to\infty}\min\left(-\frac{\log(\alpha_n)}{n},-\frac{\log(\beta_n)}{n}+\theta\right)\leq\mathbb{E}_{P_{X_0X_1}}\left(\psi^*_{Y_{X_0},\tilde{h}_{X_0,X_1}}(\theta)\right)+\delta'. \tag{49}$$

Hence, if it holds that for all sufficiently large n,

$$\alpha_n<e^{-n\left(\mathbb{E}_{P_{X_0X_1}}\left(\psi^*_{Y_{X_0},\tilde{h}_{X_0,X_1}}(\theta)\right)+\delta'\right)}, \tag{50}$$

then

$$\limsup_{n\to\infty}-\frac{\log(\beta_n)}{n}\leq\mathbb{E}_{P_{X_0X_1}}\left(\psi^*_{Y_{X_0},\tilde{h}_{X_0,X_1}}(\theta)\right)-\theta+\delta'. \tag{51}$$

Since δ (and δ') is arbitrary, this implies via the continuity of $\bar{\kappa}\left(\kappa_\alpha, P_{X_0X_1}\right)$ in κ_α that

$$\bar{\kappa}\left(\mathbb{E}_{P_{X_0X_1}}\left[\psi^*_{Y_{X_0},\tilde{h}_{X_0,X_1}}(\theta)\right], P_{X_0X_1}\right) \leq \mathbb{E}_{P_{X_0X_1}}\left[\psi^*_{Y_{X_0},\tilde{h}_{X_0,X_1}}(\theta)\right] - \theta.$$

To complete the proof, we need to show that θ can be restricted to lie in $\mathcal{I}(P_{X_0X_1}, P_{Y|X})$. To prove this, it suffices to show the following:

(i) $\mathbb{E}_{P_{X_0X_1}}\left(\psi^*_{Y_{X_0},\tilde{h}_{X_0,X_1}}\left(-D\left(P_{Y_{X_0}}||P_{Y_{X_1}}|P_{X_0X_1}\right)\right)\right) = 0.$

(ii) $\mathbb{E}_{P_{X_0X_1}}\left(\psi^*_{Y_{X_0},\tilde{h}_{X_0,X_1}}\left(D\left(P_{Y_{X_1}}||P_{Y_{X_0}}|P_{X_0X_1}\right)\right)\right) = D\left(P_{Y_{X_1}}||P_{Y_{X_0}}|P_{X_0X_1}\right).$

(iii) $\mathbb{E}_{P_{X_0X_1}}\left(\psi^*_{Y_{X_0},\tilde{h}_{X_0,X_1}}(\theta)\right)$ and $\mathbb{E}_{P_{X_0X_1}}\left(\psi^*_{Y_{X_0},\tilde{h}_{X_0,X_1}}(\theta)\right) - \theta$ are convex functions of θ.

We have,

$$\begin{aligned}
&\mathbb{E}_{P_{X_0X_1}}\left(\psi^*_{Y_{X_0},\tilde{h}_{X_0,X_1}}\left(-D\left(P_{Y_{X_0}}||P_{Y_{X_1}}|P_{X_0X_1}\right)\right)\right) \\
&= \sup_{\lambda\in\mathbb{R}}\left[-\lambda\ D\left(P_{Y_{X_0}}||P_{Y_{X_1}}|P_{X_0X_1}\right) - \mathbb{E}_{P_{X_0X_1}}\left(\psi_{Y_{X_0},\tilde{h}_{X_0,X_1}}(\lambda)\right)\right] \\
&\leq \sum_{x_0,x_1} P_{X_0X_1}(x_0,x_1)\ \left[\sup_{\lambda\in\mathbb{R}} -\lambda D\left(P_{Y_{X_0}}||P_{Y_{X_1}}\right) - \psi_{Y_{X_0},\tilde{h}_{x_0,x_1}}(\lambda)\right] && (52) \\
&= 0, && (53)
\end{aligned}$$

where, (53) follows since each term inside the square braces in (52) is zero, which in turn follows from Lemma 1 (iii). Also,

$$\begin{aligned}
&\mathbb{E}_{P_{X_0X_1}}\left(\psi^*_{Y_{X_0},\tilde{h}_{X_0,X_1}}\left(-D\left(P_{Y_{X_0}}||P_{Y_{X_1}}|P_{X_0X_1}\right)\right)\right) \\
&= \sum_{x_0,x_1} P_{X_0X_1}(x_0,x_1)\ \psi^*_{Y_{X_0},\tilde{h}_{X_0,X_1}}\left(-D\left(P_{Y_{X_0}}||P_{Y_{X_1}}|P_{X_0X_1}\right)\right) \\
&\geq 0, && (54)
\end{aligned}$$

where, (54) again follows from Lemma 1 (iii). Combining (53) and (54) proves (i). We also have that

$$\begin{aligned}
&\mathbb{E}_{P_{X_0X_1}}\left(\psi^*_{Y_{X_0},\tilde{h}_{X_0,X_1}}\left(D\left(P_{Y_{X_1}}||P_{Y_{X_0}}|P_{X_0X_1}\right)\right)\right) - D\left(P_{Y_{X_1}}||P_{Y_{X_0}}|P_{X_0X_1}\right) \\
&= \mathbb{E}_{P_{X_0X_1}}\left(\psi^*_{Y_{X_1},\tilde{h}_{X_0,X_1}}\left(D\left(P_{Y_{X_1}}||P_{Y_{X_0}}|P_{X_0X_1}\right)\right)\right) \\
&= 0, && (55)
\end{aligned}$$

where, (55) follows similarly to the proof of (i). This proves (ii). Finally, (iii) follows from Lemma 1 (iii) and the fact that a weighted sum of convex functions is convex provided the weights are non-negative. This completes the proof.

APPENDIX B
PROOF OF THEOREM 4

Fix $\kappa_\alpha > 0$ and $(\omega, R, P_{SX}, \theta) \in \mathcal{L}(\kappa_\alpha, \tau)$. Let $\eta > 0$ be a small number, and let $R' \geq 0$ and $R \geq 0$ be defined as

$$R' := \zeta_q(\kappa_\alpha, \omega), \tag{56}$$

$$\text{and } \zeta_q(\kappa_\alpha, \omega) - \rho(\kappa_\alpha, \omega) \leq R < \tau I(X;Y|S). \tag{57}$$

Encoding:

The encoder is composed of two stages, a source encoder followed by a channel encoder. The source encoding comprises of a quantization scheme followed by binning (if necessary). The details are as follows:

Quantization scheme: Let

$$\mathcal{D}_k^U(\eta) := \{\hat{U} \in \mathcal{T}_k(\mathcal{U}) : D(\hat{U}||U) \leq \kappa_\alpha + \eta\}. \tag{58}$$

Consider some ordering on the types in $\mathcal{D}_k^U(\eta)$ and denote the elements as $\hat{U}_1$, $\hat{U}_2$, ..., etc. For each type variable $\hat{U}_i \in \mathcal{D}_k^U(\eta)$, $1 \leq i \leq |\mathcal{D}_k^U(\eta)|$, choose a joint type variable $\hat{U}_i\hat{W}_i$, $\hat{W}_i \in \mathcal{T}_k(\mathcal{W})$, such that

$$D\left(\hat{W}_i|\hat{U}_i||W_i|U|\hat{U}_i\right) \leq \frac{\eta}{3}, \tag{59}$$

$$I(\hat{U}_i;\hat{W}_i) \leq R' + \frac{\eta}{3}, \tag{60}$$

where $P_{W_i|U} = \omega(P_{\hat{U}_i})$. This is always possible for k large enough due to (56) and the continuity of ω (see [4]). Let

$$\mathcal{D}_k^{UW}(\eta) := \{\hat{U}_i\hat{W}_i : 1 \leq i \leq |\mathcal{D}_k^U(\eta)|\}, \tag{61}$$

$$\text{and } R_i' := I(\hat{U}_i;\hat{W}_i) + \frac{\eta}{3}, 1 \leq i \leq |\mathcal{D}_k^U(\eta)|. \tag{62}$$

Let

$$\mathcal{C}_k = \left\{ w^k(j), j \in \left[1 : \sum_{i=1}^{|\mathcal{D}_k^U(\eta)|} e^{kR_i'}\right]\right\},$$

denote a quantization codebook such that each codeword $w^k(j)$, $j \in \mathcal{M}_i' := [1+\sum_{m=1}^{i-1} e^{kR_m'} : \sum_{m=1}^{i} e^{kR_m'}]$, $1 \leq i \leq |\mathcal{D}_k^U(\eta)|$, belongs to the set $\mathcal{T}_k(\hat{W}_i)$. For $u^k \in \mathcal{T}_k(\hat{U}_i)$ such that $\hat{U}_i \in \mathcal{D}_k^U(\eta)$ for some $1 \leq i \leq |\mathcal{D}_k^U(\eta)|$, let

$$\mu(u^k, \mathcal{C}_k) := \{j \in \mathcal{M}_i' : \ w^k(j) \in \mathcal{C}_k \text{ and } (u^k, w^k(j)) \in \mathcal{T}_k(\hat{U}_i\hat{W}_i), \ \hat{U}_i\hat{W}_i \in \mathcal{D}_k^{UW}(\eta)\}.$$

If $|\mu(u^k, \mathcal{C}_k)| \geq 1$, let $M'(u^k, \mathcal{C}_k)$ denote an index selected uniformly at random from the set $\mu(u^k, \mathcal{C}_k)$, otherwise, set $M'(u^k, \mathcal{C}_k) = 0$. Given $\mathcal{C}_k$ and $u^k \in \mathcal{U}^k$, the quantizer outputs $M' = M'(u^k, \mathcal{C}_k)$, where the support of M' is given by

$$\mathcal{M}' := \left[0 : \sum_{i=1}^{|\mathcal{D}_k^U(\eta)|} e^{kR_i'}\right].$$

Note that for sufficiently large k,

$$\begin{aligned} |\mathcal{M}'| &\leq 1 + \sum_{i=1}^{|\mathcal{D}_k^U(\eta)|} e^{kR_i'} \leq 1 + |\mathcal{D}_k^U(\eta)| e^{k\left(\max_{\hat{U}\hat{W} \in \mathcal{D}_k^{UW}(\eta)} I(\hat{U};\hat{W}) + \frac{\eta}{3}\right)} \\ &\leq 1 + |\mathcal{D}_k^U(\eta)| e^{k\left(R' + \frac{2\eta}{3}\right)} \leq e^{k(R'+\eta)}, \end{aligned} \tag{63}$$

where, in (63), we used the fact that $|\mathcal{D}_k^U(\eta)| \leq (k+1)^{|\mathcal{U}|}$.

Let

$$\begin{aligned} R_k &:= \log\left(\frac{e^{kR}}{|\mathcal{D}_k^U(\eta)|}\right), \\ \mathcal{M}_i &:= [1 + (i-1)R_k : iR_k], \ 1 \leq i \leq |\mathcal{D}_k^U(\eta)|, \\ \text{and } \mathcal{M} &:= \{0\} \bigcup \bigcup_{i=1}^{|\mathcal{D}_k^U(\eta)|} \mathcal{M}_i. \end{aligned}$$

Note that

$$e^{kR_k} \geq e^{k\left(R - \frac{|\mathcal{U}|}{k}\log(k+1)\right)}. \tag{64}$$

Let $f_b : \mathcal{M}' \mapsto \mathcal{M}$ denote a function such that $f_b(j) = 0$ iff $j = 0$, and for each index $j \in \mathcal{M}_i'$, $f_b(j) \in \mathcal{M}_i$, $1 \leq i \leq |\mathcal{D}_k^U(\eta)|$. Given f_b, the source encoder outputs $M = f_b(M')$. If $R' + \eta \leq R$, then f_b is taken to be the identity map, and in this case, $M = M'$.

Channel Encoding: Let $n = \lfloor \tau k \rfloor$. Each index in $\mathcal{M}$ is mapped to a codeword in the *channel codebook* $\mathcal{C}_X^n := \{X^n(j), j \in \mathcal{M}\}$, which is generated similar to the codebook used for the unequal error protection of a single message in [25]. Without loss of generality (w.l.o.g.), denote the elements of the set $\mathcal{S} = \mathcal{X}$ by $\{1, \ldots, |\mathcal{X}|\}$. The codeword length n is divided into $|\mathcal{S}| = |\mathcal{X}|$ blocks, where the length of the first block is $\lceil P_S(1)n \rceil$, the second block is $\lceil P_S(2)n \rceil$, so on so forth, and the length of the last block is chosen such that the total length is n. The codeword $X^n(0) = s^n$ corresponding to $M = 0$ is obtained by repeating i in block i for $1 \leq i \leq |\mathcal{X}|$. The remaining[4] $\lceil e^{kR} \rceil$ ordinary codewords $X^n(j), \ j \in \left[e^{kR}\right]$, are obtained by blockwise i.i.d. random coding, i.e., the symbols in the i^{th} block of each codeword are generated i.i.d. according to $P_{X|S=i}$. The sequence s^n is revealed to the detector.

Decoding:

The decoder consists of two parts, a channel decoder followed by a tester.

Channel decoding: At the detector, the channel decoder first performs a NP test on the channel output Y^n using the decision

[4]Actually, the number of codewords generated should be slightly higher (e.g. $e^{k(R+\delta)}$ for a small positive number δ), as an expurgation step is involved later.

rule $g_\theta : \mathcal{Y}^n \mapsto \{0,1\}$, where

$$g_\theta(y^n) := \mathbb{1}\left(\sum_{j=1}^{n} \log\left(\frac{P_{Y|X=s(j))}(y_j)}{P_{Y|S=s(j)}(y_j)}\right) \geq n\theta\right),$$

$$s(j) := i \text{ if } \sum_{l=1}^{i-1} \lceil P_S(l)n \rceil < j \leq \sum_{l=1}^{i} \lceil P_S(l)n \rceil. \tag{65}$$

In (65), the empty sum is defined to be equal to 0. If $g_\theta(y^n) = 1$, then $\hat{M} = 0$ and $\hat{H} = 1$ is declared. Else, maximum likelihood (ML) decoding is done on the remaining codewords $X^n(j), j \in \left[e^{kR}\right]$, and $\hat{M}$ is set equal to the ML estimate. Note that since the i^{th} block of each codeword $X^n(j),\ j \in \left[e^{kR}\right]$, is generated independently and i.i.d. according to distribution $P_{X|S=i}$, the channel outputs in the i^{th} block is distributed i.i.d. according to $P_{Y|S=i}$. It then follows similar to Proposition 1 that for k sufficiently large,

$$\mathbb{P}\left(\hat{M} = 0 | M \neq 0\right) \leq e^{-k\tau(E_m(P_{SX},\theta)-\eta)} \tag{66}$$

and

$$\mathbb{P}\left(\hat{M} \neq 0 | M = 0\right) \leq e^{-k\tau(E_m(P_{SX},\theta)-\theta-\eta)}. \tag{67}$$

Also, given $\hat{M} \neq 0$, it follows from the analysis based on random coding and expurgation (see [28, Exercise 10.18, 10.24] and [30]) that there exists a deterministic codebook $\mathcal{C}_X^n$ such that (66) and (67) holds, and the ML-decoding described above asymptotically yields

$$\mathbb{P}\left(\hat{M} \neq m | M = m \neq 0, \hat{M} \neq 0\right) \leq e^{-n\left(E_x\left(\frac{R}{\tau}, P_{SX}\right)-\eta\right)}. \tag{68}$$

$$\tag{69}$$

This deterministic codebook is used for channel coding.
Testing: The acceptance region for the hypothesis test is the same as that given in [4, Theorem 1]. More specifically, for a given codebook $\mathcal{C}_k$, let $\mathcal{O}_{m'}$ denote the set of u^k such that the source encoder outputs m', $m' \in \mathcal{M}'\backslash\{0\}$. For each $m' \in \mathcal{M}'\backslash\{0\}$ and $u^k \in \mathcal{O}_{m'}$, let

$$\mathcal{B}_{m'}(u^k) = \{v^k \in \mathcal{V}^k : (w_{m'}^k, u^k, v^k) \in \mathcal{J}_k^{\kappa_\alpha+\eta}(W_{m'}UV)\},$$

where $W_{m'}UV$ is uniquely specified by

$$W_{m'} - U - V \text{ and } P_{W_{m'}|U} = \omega(P_{u^k}). \tag{70}$$

For $m' \in \mathcal{M}'\backslash\{0\}$, we define

$$\mathcal{B}_{m'} := \{v^k : v^k \in \mathcal{B}_{m'}(u^k) \text{ for some } u^k \in \mathcal{O}_{m'}\}.$$

Define the acceptance region for H_0 at the detector as

$$\mathcal{A}_k := \underset{m' \in \mathcal{M}'\backslash 0}{\cup} m' \times \mathcal{B}_{m'}, \tag{71}$$

or equivalently as

$$\mathcal{A}_k^e := \underset{m' \in \mathcal{M}'\backslash 0}{\cup} \mathcal{O}_{m'} \times \mathcal{B}_{m'}. \tag{72}$$

The tester takes $\hat{M}$ as input, decodes for the quantization codeword $w^k(\hat{M}')$ (if required) using the empirical conditional entropy decoder (ECED), and declares the output of the hypothesis test based on $w^k(\hat{M}')$ and V^k. More specifically, if binning is not performed, i.e., if $R' + \eta \leq R$, set $\hat{M}' = \hat{M}$. Otherwise (if $R' + \eta > R$), given $\hat{M} = \hat{m}$ and $V^k = v^k$, set $\hat{M}' = \hat{m}'$, where

$$\hat{m}' := \begin{cases} 0, & \text{if } \hat{M} = 0, \\ \underset{j: f_b(j) = \hat{m}}{\arg\min}\, H_e(w^k(j)|v^k), & \text{otherwise.} \end{cases}$$

If $\hat{M}' = 0$, $\hat{H} = 1$ is declared. Otherwise, given $\hat{M}' = \hat{m}' \neq 0$ and $V^k = v^k$, $\hat{H} = 0$ or $\hat{H} = 1$ is declared depending on whether $(\hat{m}', v^k) \in \mathcal{A}_k$ or $(\hat{m}', v^k) \notin \mathcal{A}_k$, respectively.
Analysis of the type I and type II error probabilities:
Using the method of random coding, we will analyze the type I and type II error probabilities over an ensemble of randomly

generated quantization and binning codebooks. Then, the standard random coding argument followed by an expurgation technique [30] guarantees the existence of a deterministic quantization and binning codebook that achieves the lower bound given in Theorem 4. Let each codeword $w^k(j)$, $j \in \mathcal{M}'_i$, $1 \leq i \leq |\mathcal{D}_k^U(\eta)|$, be selected (with replacement) independently and uniformly at random from the set $\mathcal{T}_k(\hat{W}_i)$ (see quantization scheme above). Let f_B denote the random binning function such that for each index $j \in \mathcal{M}'_i$, an index $f_B(j)$ is selected (with replacement) independently and uniformly at random from the set $\mathcal{M}_i$. We proceed to analyze the type I and type II error probabilities averaged over these random codebooks. Note that a type I error can occur only under the following events:

(i) $\mathcal{E}_{EE} := \bigcup_{\hat{U} \in \mathcal{D}_k^U(\eta)} \bigcup_{u^k \in \mathcal{T}_k(\hat{U})} \mathcal{E}_{EE}(u^k)$, where

$$\mathcal{E}_{EE}(u^k) := \{\nexists\, W^k(j) \in \mathcal{C}_k,\ j \in [1 : |\mathcal{M}'|],\ \text{s.t. } (u^k, W^k(j)) \in \mathcal{T}_k(\hat{U}_i \hat{W}_i), P_{\hat{U}_i} = P_{u^k}, \hat{U}_i \hat{W}_i \in \mathcal{D}_k^{UW}(\eta)\}.$$

(ii) $\hat{M}' = M'$.
(iii) $M' \neq 0$ and $\hat{M} \neq M$.
(iv) $M' = M = 0$ and $\hat{M} \neq M$.
(v) $M' \neq 0$, $\hat{M} = M$ and $\hat{M}' \neq M'$.

Here, (i) corresponds to the event that there does not exist a quantization codeword corresponding to atleast one sequence u^k of type $P_{u^k} \in \mathcal{D}_k^U(\eta)$; (ii) corresponds to the event, in which, there is neither an error at the channel decoder nor at the ECED; (iii) and (iv) corresponds to the case, in which, there is an error at the channel decoder (hence also at the ECED); and, (v) corresponds to the case such that there is an error only at the ECED.

As we show later in (109), it follows by a generalization of the *type-covering lemma* [28, Lemma 9.1] that

$$\mathbb{P}(\mathcal{E}_{EE}) \leq e^{-e^{k\Omega(\eta)}}. \tag{73}$$

Since $\frac{e^{k\Omega(\eta)}}{k} \xrightarrow{(k)} \infty$ for $\eta > 0$, we may safely ignore this event from the analysis of the exponent of type I and type II error probability. Given $\mathcal{E}_{EE}^c$ and that event (ii) holds, it follows from [4, Equation 4.22] that for any given codebook $\mathcal{C}_k$, the type I error probability is asymptotically upper bounded by $e^{-k\kappa_\alpha}$, since the acceptance region is the same. Hence, it also holds when averaged over the random quantization codebooks such that $\mathcal{E}_{EE}^c$ holds, implying that

$$\mathbb{P}\left(\hat{H} = 1 | \mathcal{E}_{EE}^c, \hat{M}' = M'\right) \leq e^{-k\kappa_\alpha}. \tag{74}$$

Next, consider event (iii). By the design of the channel codebook $\mathcal{C}_X^n$, it holds asymptotically that

$$\begin{aligned}
\mathbb{P}\left(M' \neq 0, \hat{M} \neq M | H = 0\right) &= \mathbb{P}\left(M' \neq 0 | H = 0\right) \mathbb{P}\left(\hat{M} \neq M | M \neq 0\right) \\
&\leq \mathbb{P}\left(\hat{M} \neq M | M \neq 0\right) \\
&\leq \mathbb{P}\left(\hat{M} = 0 | M \neq 0\right) + \mathbb{P}\left(\hat{M} \neq M | M \neq 0, \hat{M} \neq 0\right) \\
&\leq e^{-k\tau(E_m(P_{SX}, \theta) - \eta)} + e^{-k\tau\left(E_x\left(\frac{R}{\tau}, P_{SX}\right) - \eta\right)} \qquad (75) \\
&= e^{-k\tau\left(\min\left(E_m(P_{SX}, \theta), E_x\left(\frac{R}{\tau}, P_{SX}\right)\right) - \eta\right)}, \qquad (76)
\end{aligned}$$

where, in (75), we used (66) and (68). Also, note by the definition of $\mathcal{D}_k(\hat{U})$ and (73) that the probability of event (iv) can be upper bounded as

$$\mathbb{P}\left(M = 0, \hat{M} \neq M | H = 0\right) \leq \mathbb{P}\left(M' = 0 | H = 0\right) \leq e^{-k\kappa_\alpha}. \tag{77}$$

Next, consider the event (v). Note that this event is impossible when $R' + \eta \leq R$, since there is no binning involved. Hence, assume that $R' + \eta > R$. Since $M = 0$ iff $M' = 0$, $M' \neq 0$ and $\hat{M} = M$ implies that $\hat{M} \neq 0$. Let

$$\mathcal{D}_k^{VW}(\eta) := \left\{\hat{V}\hat{W} :\ \exists\ (w^k, u^k, v^k) \in \bigcup_{m' \in \mathcal{M}' \setminus \{0\}} \mathcal{J}_k^{\kappa_\alpha + \eta}(W_{m'} UV), W_{m'} UV \text{ satisfies (70) and } P_{w^k u^k v^k} = P_{\hat{W}\hat{U}\hat{V}}\right\}.$$

We can write,

$$\begin{aligned}
\mathbb{P}\left(M' \neq 0, \hat{M} = M, \hat{M}' \neq M' | H = 0\right) = \mathbb{P}\left(M' \neq 0, \hat{M} = M, \hat{M}' \neq M', (M', V^k) \in \mathcal{A}_k | H = 0\right) \\
+ \mathbb{P}\left(M' \neq 0, \hat{M} = M, \hat{M}' \neq M', (M', V^k) \notin \mathcal{A}_k | H = 0\right). \qquad (78)
\end{aligned}$$

The second term in (78) can be upper-bounded as

$$\mathbb{P}\left(M' \neq 0, \hat{M} = M, \hat{M}' \neq M', (M', V^k) \notin \mathcal{A}_k | H = 0\right)$$

$$\begin{aligned}
&\leq \mathbb{P}\left((M',V^k)\notin \mathcal{A}_k, \mathcal{E}_{EE}|H=0\right)+\mathbb{P}\left((M',V^k)\notin \mathcal{A}_k, \mathcal{E}_{EE}^c|H=0\right)\\
&\leq e^{-e^{k\Omega(\eta)}}+\mathbb{P}\left((M',V^k)\notin \mathcal{A}_k|\mathcal{E}_{EE}^c, H=0\right)\\
&\leq e^{-e^{k\Omega(\eta)}}+\mathbb{P}\left((U^k,V^k)\notin \mathcal{A}_k^e\right)\\
&\leq e^{-e^{k\Omega(\eta)}}+e^{-k\kappa_\alpha}, \qquad (79)
\end{aligned}$$

where, the inequality in (79) follows from [4, Equation 4.22] for sufficiently large k, since the acceptance region is the same. Let

$$\mathcal{D}_k(V) := \{\hat{V} : \exists\ \hat{W} \text{ s.t. } \hat{V}\hat{W} \in \mathcal{D}_k^{VW}(\eta)\}.$$

The first term in (78) can be bounded as shown below:

$$\begin{aligned}
&\mathbb{P}\left(M'\neq 0, \hat{M}=M, \hat{M}'\neq M', (M',V^k)\in\mathcal{A}_k|H=0\right)\\
&\leq \sum_{\substack{v^k\in\mathcal{T}_k(\hat{V}):\\ \hat{V}\in\mathcal{D}_k(V)}} \mathbb{P}\left(V^k=v^k, \exists\ j\in f_B^{-1}(M),\ j\neq M' : H_e(W^k(j)|v^k)\leq H_e(W^k(M')|v^k)|M'\neq 0\right)\\
&= \sum_{\substack{v^k\in\mathcal{T}_k(\hat{V}):\\ \hat{V}\in\mathcal{D}_k(V)}} \mathbb{P}\left(V^k=v^k|M'\neq 0\right)\mathbb{P}\left(\exists\ j\in f_B^{-1}(M),\ j\neq M' : H_e(W^k(j)|v^k)\leq H_e(W^k(M')|v^k)|V^k=v^k, M'\neq 0\right) \qquad (80)
\end{aligned}$$

Defining the events

$$\begin{aligned}
&\mathcal{E}_1' := \{V^k=v^k, M'\neq 0\},\\
&\mathcal{E}_2' := \{V^k=v^k, M'=m'\neq 0, M=m\},
\end{aligned}$$

we can write

$$\begin{aligned}
&\mathbb{P}\left(\exists\ j\in f_B^{-1}(M),\ j\neq M' : H_e(W^k(j)|v^k)\leq H_e(W^k(M')|v^k)|\mathcal{E}_1'\right)\\
&= \sum_{\substack{m'\in\\ \mathcal{M}'\backslash\{0\}}}\sum_{\substack{m\in\\ \mathcal{M}\backslash\{0\}}} \mathbb{P}\left(M'=m', M=m|\mathcal{E}_1'\right)\ \mathbb{P}\left(\exists\ j\in f_B^{-1}(m),\ j\neq m' : H_e(W^k(j)|v^k)\leq H_e(W^k(m')|v^k)|\mathcal{E}_2'\right). \qquad (81)
\end{aligned}$$

Consider the second term in (81). Denoting the type of v^k by $\hat{V}$, it follows that

$$\begin{aligned}
&\mathbb{P}\left(\exists\ j\in f_B^{-1}(m),\ j\neq m' : H_e(W^k(j)|v^k)\leq H_e(W^k(m')|v^k)|\mathcal{E}_2'\right)\\
&= \sum_{j\in\mathcal{M}'\backslash\{0,m'\}} \mathbb{P}\left(f_B(j)=m : H_e(W^k(j)|v^k)\leq H_e(W^k(m')|v^k)|\mathcal{E}_2'\right)\\
&\leq \frac{1}{e^{kR_k}}\sum_{j\in\mathcal{M}'\backslash\{0,m'\}} \mathbb{P}\left(H_e(W^k(j)|v^k)\leq H_e(W^k(m')|v^k)|\mathcal{E}_2'\cup\{f_B(j)=m\}\right) \qquad (82)\\
&\leq \frac{1}{e^{kR_k}}\sum_{j\in\mathcal{M}'\backslash\{0,m'\}}\sum_{\substack{\hat{W}:\\ \hat{V}\hat{W}\in\mathcal{D}_k^{VW}(\eta)}}\sum_{\substack{w^k:\\ (v^k,w^k)\in\mathcal{T}_k(\hat{V}\hat{W})}} \mathbb{P}\left(W^k(m')=w^k|\mathcal{E}_2'\cup\{f_B(j)=m\}\right)\\
&\qquad \sum_{\substack{\tilde{w}^k\in\mathcal{T}_k(\hat{W})\\ H_e(\tilde{w}^k|v^k)\leq H(\hat{W}|\hat{V})}} \mathbb{P}\left(W^k(j)=\tilde{w}^k|\mathcal{E}_2'\cup\{f_B(j)=m\}\cup\{W^k(m')=w^k\}\right) \qquad (83)\\
&\leq \frac{1}{e^{kR_k}}\sum_{j\in\mathcal{M}'\backslash\{0,m'\}}\sum_{\substack{\hat{W}:\\ \hat{V}\hat{W}\in\mathcal{D}_k^{VW}(\eta)}}\sum_{\substack{w^k:\\ (v^k,w^k)\in\mathcal{T}_k(\hat{V}\hat{W})}} \mathbb{P}\left(W^k(m')=w^k|\mathcal{E}_2'\cup\{f_B(j)=m\}\right)\\
&\qquad \sum_{\substack{\tilde{w}^k\in\mathcal{T}_k(\hat{W}):\\ H_e(\tilde{w}^k|v^k)\leq H(\hat{W}|\hat{V})}} 2\ \mathbb{P}\left(W^k(j)=\tilde{w}^k\right). \qquad (84)
\end{aligned}$$

In (82), we used the fact that binning is done uniformly at random; in (83), we used the following: if $v^k\in\mathcal{T}_k(\hat{V})$ is such that $\hat{V}\in\mathcal{D}_k(V)$, then $M'\neq 0$ implies that $(W^k(M'),v^k)\in\mathcal{T}_k(\hat{V}\hat{W})$ for some $\hat{V}\hat{W}\in\mathcal{D}_k^{VW}(\eta)$. In (84), we used

$$\mathbb{P}\left(W^k(j)=\tilde{w}^k|\mathcal{E}_2'\cup\{f_B(j)=m\}\cup\{W^k(m')=w^k\}\right)\leq 2\ \mathbb{P}\left(W^k(j)=\tilde{w}^k\right), \qquad (85)$$

which will be shown later. Continuing, we can write (for sufficiently large k)

$$\mathbb{P}\left(\exists\ j \in f_B^{-1}(m),\ j \neq m' : H_e(W^k(j)|v^k) \leq H_e(W^k(m')|v^k)|\mathcal{E}_2'\right)$$

$$\leq \frac{1}{e^{kR_k}} \sum_{j \in \mathcal{M}' \backslash \{0,m'\}} \sum_{\substack{\hat{W}: \\ \hat{V}\hat{W} \in \mathcal{D}_k^{VW}(\eta)}} \sum_{\substack{w^k: \\ (v^k,w^k) \in \mathcal{T}_k(\hat{V}\hat{W})}} \mathbb{P}\left(W^k(m') = w^k|\mathcal{E}_2' \cup \{f_B(j) = m\}\right) \sum_{\substack{\tilde{w}^k \in T_{\hat{W}}: \\ H_e(\tilde{w}^k|v^k) \\ \leq H(\hat{W}|\hat{V})}} 2\ e^{-k(H(\hat{W})-\eta)} \tag{86}$$

$$\leq \frac{1}{e^{kR_k}} \sum_{j \in \mathcal{M}' \backslash \{0,m'\}} \sum_{\substack{\hat{W}: \\ \hat{V}\hat{W} \in \mathcal{D}_k^{VW}(\eta)}} \sum_{\substack{w^k: \\ (v^k,w^k) \in \mathcal{T}_k(\hat{V}\hat{W})}} \mathbb{P}\left(W^k(m') = w^k|\mathcal{E}_2' \cup \{f_B(j) = m\}\right)$$
$$(k+1)^{|\mathcal{V}||\mathcal{W}|} e^{kH(\hat{W}|\hat{V})} 2\ e^{-k(H(\hat{W})-\eta)} \tag{87}$$

$$\leq \frac{1}{e^{kR_k}} \sum_{j \in \mathcal{M}' \backslash \{0,m'\}} \sum_{\substack{\hat{W}: \\ \hat{V}\hat{W} \in \mathcal{D}_k^{VW}(\eta)}} 2\ (k+1)^{|\mathcal{V}||\mathcal{W}|}\ e^{-k\left(I(\hat{W};\hat{V})-\eta\right)}$$

$$\leq \frac{1}{e^{kR_k}} \sum_{j \in \mathcal{M}' \backslash \{0,m'\}} 2\ (k+1)^{|\mathcal{W}|}\ (k+1)^{|\mathcal{V}||\mathcal{W}|}\ e^{-k\left(\min_{\hat{V}\hat{W} \in \mathcal{D}_k^{VW}(\eta)} I(\hat{W};\hat{V})-\eta\right)} \tag{88}$$

$$\leq e^{-k(R-R'+\rho_k-\eta_k')}, \tag{89}$$

where,

$$\rho_k := \min_{\hat{V}\hat{W} \in \mathcal{D}_k^{VW}(\eta)} I(\hat{V};\hat{W})$$
$$\text{and } \eta_k' := 3\eta + \frac{|\mathcal{W}|(|\mathcal{V}|+1)\log(k+1)}{k} + \frac{\log(2)}{k} + \frac{|\mathcal{U}|\log(k+1)}{k}.$$

In (86), we used [28, Lemma 2.3] and the fact that codewords are chosen uniformly at random from $\mathcal{T}_k(\hat{W})$; in (87), we used that the total number of sequences $\tilde{w}^k \in \mathcal{T}_k(\tilde{W})$ such that $P_{\tilde{w}^k v^k} = P_{\tilde{W}\hat{V}}$ and $H(\tilde{W}|\hat{V}) \leq H(\hat{W}|\hat{V})$ is upper bounded by $e^{kH(\hat{W}|\hat{V})}$ and $|\mathcal{T}_n(\mathcal{W} \times \mathcal{V})| \leq (k+1)^{|\mathcal{V}||\mathcal{W}|}$; in (88), we used [28, Lemma 2.2]; and, in (89), we used (56), (57), (63) and (64). Thus, for sufficiently large k, since $\rho_k \to \rho(\kappa_\alpha, \omega) + O(\eta)$, we have from (79), (80), (81) and (89) that for sufficiently large k,

$$\mathbb{P}\left(M' \neq 0, \hat{M} = M, \hat{M}' \neq M'|H = 0\right) \leq e^{-k(\min(\kappa_\alpha, R-\zeta_q(\kappa_\alpha,\omega)+\rho(\kappa_\alpha,\omega)-O(\eta)))}, \tag{90}$$

By choice of $(\omega, P_{SX}, \theta) \in \mathcal{L}(\kappa_\alpha, \tau)$, it follows from (73), (74), (76), (77) and (90) that the type I error probability is upper bounded by $e^{-k(\kappa_\alpha - O(\eta))}$, asymptotically.

Next, we analyze the type II error probability averaged over the random codebooks. For a given codebook $\mathcal{C}_k$, let $\tilde{U}, \tilde{V}, \tilde{W}$ and $\tilde{W}_d$ denote the type variable for the realizations of $\bar{U}^k, \bar{V}^k, W^k(M')$ $(M' \neq 0)$ and $W^k(\hat{M}')$ $(\hat{M}' \neq 0)$, respectively. A type II error can occur only under the following events:

(a) $\mathcal{E}_a := \{\hat{M} = M, \hat{M}' = M' \neq 0, (\bar{U}^k, \bar{V}^k, W^k(M')) \in \mathcal{T}_k(\hat{U}\hat{V}\hat{W})$ s.t. $\hat{U}\hat{W} \in \mathcal{D}_k^{UW}(\eta)$ and $\hat{V}\hat{W} \in \mathcal{D}_k^{VW}(\eta)\}$.

(b)

$$\mathcal{E}_b := \left\{ \begin{array}{l} M' \neq 0, \hat{M} = M, \hat{M}' \neq M', f_B(\hat{M}') = f_B(M'), (\bar{U}^k, \bar{V}^k, W^k(M'), W^k(\hat{M}')) \in \mathcal{T}_k(\hat{U}\hat{V}\hat{W}\hat{W}_d) \text{ s.t.} \\ \hat{U}\hat{W} \in \mathcal{D}_k^{UW}(\eta), \hat{V}\hat{W}_d \in \mathcal{D}_k^{VW}(\eta), \text{ and } H_e\left(W^k(\hat{M}')|\bar{V}^k\right) \leq H_e\left(W^k(M')|\bar{V}^k\right) \end{array} \right\}.$$

(c)

$$\mathcal{E}_c := \left\{ \begin{array}{l} M' \neq 0, \hat{M} \neq M \text{ or } 0, (\bar{U}^k, \bar{V}^k, W^k(M'), W^k(\hat{M}')) \in \mathcal{T}_k(\hat{U}\hat{V}\hat{W}\hat{W}_d) \text{ s.t.} \\ \hat{U}\hat{W} \in \mathcal{D}_k^{UW}(\eta) \text{ and } \hat{V}\hat{W}_d \in \mathcal{D}_k^{VW}(\eta) \end{array} \right\}.$$

(d) $\mathcal{E}_d := \{M = M' = 0, \hat{M} \neq M, (\bar{V}^k, W^k(\hat{M}')) \in \mathcal{T}_k(\hat{V}\hat{W}_d)$ s.t. $\hat{V}\hat{W}_d \in \mathcal{D}_k^{VW}(\eta)\}$.

Since the exponent of probability of the event $\mathcal{E}_{EE}$ tends to ∞ with k by (73), we may assume that $\mathcal{E}_{EE}^c$ holds for the type II error-exponent analysis. It then follows from the analysis in [4, Eq. 4.23-4.27] that for sufficiently large k, we have

$$\mathbb{P}\left(\mathcal{E}_a|\mathcal{E}_{EE}^c\right) \leq e^{-k(E_1(\kappa_\alpha,\omega)-O(\eta))}. \tag{91}$$

When $R' + \eta \leq R$, note that $\mathcal{E}_b$ is impossible, and hence, the exponent of this event is ∞. Assume that $R' + \eta > R$. Let

$$\mathcal{F}_{2,k}(\eta) := \{\tilde{U}\tilde{V}\tilde{W}\tilde{W}_d \in \mathcal{T}_k(\mathcal{U} \times \mathcal{V} \times \mathcal{W} \times \mathcal{W}) : \tilde{U}\tilde{W} \in \mathcal{D}_k^{UW}(\eta), \tilde{V}\tilde{W}_d \in \mathcal{D}_k^{VW}(\eta) \text{ and } H(\tilde{W}_d|\tilde{V}) \leq H(\tilde{W}|\tilde{V})\}.$$

Then, we can write

$$
\begin{aligned}
&\mathbb{P}\left(\mathcal{E}_b|H=1\right)\\
&\leq \sum_{\substack{\tilde{U}\tilde{V}\tilde{W}\tilde{W}_d\\ \in\mathcal{F}_{2,k}(\eta)}} \sum_{\substack{(u^k,v^k,w^k,\bar{w}^k)\\ \in\mathcal{T}_k(\tilde{U}\tilde{V}\tilde{W}\tilde{W}_d)}} \sum_{m'\in\mathcal{M}'\backslash\{0\}} \mathbb{P}\left(\bar{U}^k=u^k,\bar{V}^k=v^k,M'=m',W^k(M')=w^k|H=1\right)\\
&\qquad\left[\sum_{\hat{m}'\in\mathcal{M}'\backslash\{0\}}\mathbb{P}\left(W^k(\hat{m}')=\bar{w}^k,f_B(m')=f_B(\hat{m}')|\bar{U}^k=u^k,\bar{V}^k=v^k,M'=m',W^k(m')=w^k\right)\right]
\end{aligned}
\tag{92}
$$

The first term in (92) can be written as

$$
\begin{aligned}
&\mathbb{P}\left(\bar{U}^k=u^k,\bar{V}^k=v^k,M'=m',W^k(M')=w^k|H=1\right)\\
&=\mathbb{P}\left(\bar{U}^k=u^k,\bar{V}^k=v^k,M'=m'|H=1\right)\ \mathbb{P}\left(W^k(m')=w^k|\bar{U}^k=u^k,\bar{V}^k=v^k,M'=m'\right)
\end{aligned}
\tag{93}
$$

Note that $M'\neq 0$ and $\bar{U}^k=u^k$ implies that $\tilde{U}\tilde{W}\in\mathcal{D}_k(UW)$. Hence, we can bound the second term in (92) for sufficiently large k as

$$
\mathbb{P}\left(W^k(m')=w^k|\bar{U}^k=u^k,\bar{V}^k=v^k,M'=m'\right)\leq\begin{cases}\frac{1}{e^{k\left(H(\tilde{W}|\tilde{U})-\eta\right)}}, & \text{if } w^k\in\mathcal{T}_k(\tilde{W}),\\ 0, & \text{otherwise,}\end{cases}
\tag{94}
$$

where, we used the fact that given $M'=m'$ and $\bar{U}^k=u^k$, $W^k(m')$ is uniformly distributed in the set $\mathcal{T}_k(P_{\tilde{W}|\tilde{U}},u^k)$ and that for sufficiently large k,

$$
|\mathcal{T}_k(P_{\tilde{W}|\tilde{U}},u^k)|\geq e^{k\left(H(\tilde{W}|\tilde{U})-\eta\right)}.
$$

On the other hand, the second term in (92) can be bounded as follows:

$$
\begin{aligned}
&\mathbb{P}\left(W^k(\hat{m}')=\bar{w}^k,f_B(m')=f_B(\hat{m}')|\bar{U}^k=u^k,\bar{V}^k=v^k,M'=m',W^k(m')=w^k\right)\\
&\leq\frac{1}{e^{kR_k}}\mathbb{P}\left(W^k(\hat{m}')=\bar{w}^k|\bar{U}^k=u^k,\bar{V}^k=v^k,M'=m',W^k(m')=w^k\right) &(95)\\
&\leq\frac{1}{e^{kR_k}}2\ \mathbb{P}\left(W^k(\hat{m}')=\bar{w}^k\right), &(96)
\end{aligned}
$$

where, in (95), we used the fact that the binning is uniformly distributed and independent of the codebook generation; in (96), we used

$$
\mathbb{P}\left(W^k(\hat{m}')=\bar{w}^k|\bar{U}^k=u^k,\bar{V}^k=v^k,M'=m',W^k(m')=w^k\right)\leq 2\ \mathbb{P}\left(W^k(\hat{m}')=\bar{w}^k\right).
\tag{97}
$$

which will be shown below. Thus, from (94) and (96), we can bound the term in (92) (for sufficiently large k) as

$$
\begin{aligned}
&\mathbb{P}\left(\mathcal{E}_b|H=1\right)\\
&\leq\frac{2}{e^{kR_k}}\sum_{\substack{\tilde{U}\tilde{V}\tilde{W}\tilde{W}_d\\ \in\mathcal{F}_{2,k}(\eta)}} \sum_{\substack{(u^k,v^k,w^k,\bar{w}^k)\\ \in\mathcal{T}_k(\tilde{U}\tilde{V}\tilde{W}\tilde{W}_d)}} \sum_{m'\in\mathcal{M}'\backslash\{0\}}\mathbb{P}\left(\bar{U}^k=u^k,\bar{V}^k=v^k,M'=m'|H=1\right)\frac{1}{e^{k\left(H(\tilde{W}|\tilde{U})-\eta\right)}}\\
&\qquad\qquad\sum_{\hat{m}'\in\mathcal{M}'\backslash\{0\}}\mathbb{P}\left(W^k(\hat{m}')=\bar{w}^k\right)\\
&\leq\frac{2}{e^{kR_k}}\sum_{\substack{\tilde{U}\tilde{V}\tilde{W}\tilde{W}_d\\ \in\mathcal{F}_{2,k}(\eta)}} \sum_{\substack{(u^k,v^k,w^k,\bar{w}^k)\\ \in\mathcal{T}_k(\tilde{U}\tilde{V}\tilde{W}\tilde{W}_d)}} e^{-k\left(H(\tilde{U}\tilde{V})+D(\tilde{U}\tilde{V}||\bar{U}\bar{V})\right)}\frac{1}{e^{k\left(H(\tilde{W}|\tilde{U})-\eta\right)}}\\
&\qquad\qquad\sum_{m'\in\mathcal{M}'\backslash\{0\}}\mathbb{P}\left(M'=m'|\bar{U}^k=u^k,\bar{V}^k=v^k\right)\sum_{\hat{m}'\in\mathcal{M}'\backslash\{0\}}\mathbb{P}\left(W^k(\hat{m}')=\bar{w}^k\right)\\
&\leq\frac{2}{e^{kR_k}}\sum_{\substack{\tilde{U}\tilde{V}\tilde{W}\tilde{W}_d\\ \in\mathcal{F}_{2,k}(\eta)}} \sum_{\substack{(u^k,v^k,w^k,\bar{w}^k)\\ \in\mathcal{T}_k(\tilde{U}\tilde{V}\tilde{W}\tilde{W}_d)}} e^{-k\left(H(\tilde{U}\tilde{V})+D(\tilde{U}\tilde{V}||\bar{U}\bar{V})\right)}\frac{1}{e^{k\left(H(\tilde{W}|\tilde{U})-\eta\right)}}\sum_{\hat{m}'\in\mathcal{M}'\backslash\{0\}}\mathbb{P}\left(W^k(\hat{m}')=\bar{w}^k\right)\\
&\leq\frac{2}{e^{kR_k}}\sum_{\substack{\tilde{U}\tilde{V}\tilde{W}\tilde{W}_d\\ \in\mathcal{F}_{2,k}(\eta)}} \sum_{\substack{(u^k,v^k,w^k,\bar{w}^k)\\ \in\mathcal{T}_k(\tilde{U}\tilde{V}\tilde{W}\tilde{W}_d)}} e^{-k\left(H(\tilde{U}\tilde{V})+D(\tilde{U}\tilde{V}||\bar{U}\bar{V})\right)}\frac{1}{e^{k\left(H(\tilde{W}|\tilde{U})-\eta\right)}}\ \frac{e^{k(R'+\eta)}}{e^{k\left(H(\tilde{W}_d)-\eta\right)}}
\end{aligned}
$$

$$\leq \frac{2}{e^{kR_k}} \sum_{\substack{\tilde{U}\tilde{V}\tilde{W}\tilde{W}_d \\ \in \mathcal{F}_{2,k}(\eta)}} \sum_{\substack{(u^k,v^k,w^k) \\ \in \mathcal{T}_k(\tilde{U}\tilde{V}\tilde{W})}} e^{-k\left(H(\tilde{U}\tilde{V})+D(\tilde{U}\tilde{V}||\bar{U}\bar{V})\right)} \frac{1}{e^{k\left(H(\tilde{W}|\tilde{U})-\eta\right)}} \frac{e^{k(R'+\eta)}}{e^{k\left(H(\tilde{W}_d)-\eta\right)}} e^{kH(\tilde{W}_d|\tilde{V})}$$

$$\leq \frac{2}{e^{kR_k}} \sum_{\substack{\tilde{U}\tilde{V}\tilde{W}\tilde{W}_d \\ \in \mathcal{F}_{2,k}(\eta)}} e^{kH(\tilde{U}\tilde{V}\tilde{W})} e^{-k\left(H(\tilde{U}\tilde{V})+D(\tilde{U}\tilde{V}||\bar{U}\bar{V})\right)} \frac{1}{e^{k\left(H(\tilde{W}|\tilde{U})-\eta\right)}} \frac{e^{k(R'+\eta)}}{e^{k\left(H(\tilde{W}_d)-\eta\right)}} e^{kH(\tilde{W}_d|\tilde{V})}$$

$$\leq e^{-kE_{2,k}},$$

where

$$E_{2,k} := \min_{\tilde{U}\tilde{V}\tilde{W}\tilde{W}_d \in \mathcal{F}_{2,k}(\eta)} -H(\tilde{U}\tilde{V}\tilde{W}) + H(\tilde{U}\tilde{V}) + D(\tilde{U}\tilde{V}||\bar{U}\bar{V}) + H(\tilde{W}|\tilde{U}) + I(\tilde{V};\tilde{W}_d) + R - R' - 3\eta - \delta_k',$$

$$\delta_k' := \frac{|\mathcal{U}||\mathcal{V}||\mathcal{W}|^2}{k}\log(k+1) + \frac{|\mathcal{U}|}{k}\log(k+1) + \frac{\log(2)}{k}. \tag{98}$$

Note that since $\tilde{U}\tilde{V}\tilde{W}\tilde{W}_d \in \mathcal{F}_{2,k}(\eta)$ implies that $\tilde{V}\tilde{W}_d \in \mathcal{D}_k^{VW}(\eta)$, we have

$$E_{2,k} \geq \min_{\tilde{U}\tilde{V}\tilde{W}\tilde{W}_d \in \mathcal{F}_{2,k}(\eta)} -H(\tilde{U}\tilde{V}\tilde{W}) + H(\tilde{U}\tilde{V}) + D(\tilde{U}\tilde{V}||\bar{U}\bar{V}) + H(\tilde{W}|\tilde{U}) + \rho_k + R - R' - 3\eta - \delta_k'. \tag{99}$$

Simplifying the terms in (99) and using $\rho_k \xrightarrow{(k)} \rho(\kappa_\alpha,\omega) + O(\eta)$, we obtain by the continuity of KL-divergence that

$$\frac{-1}{k}\log\left(\mathbb{P}\left(\mathcal{E}_b|H=1\right)\right)$$
$$\overset{(k)}{\geq} \begin{cases} \min_{\tilde{U}\tilde{V}\tilde{W} \in \mathcal{T}_2(\kappa_\alpha,\omega)} D(\tilde{U}\tilde{V}\tilde{W}||\bar{U}\bar{V}\bar{W}) + E_b(\kappa_\alpha,\omega,R) - O(\eta), & \text{if } R < \zeta_q(\kappa_\alpha,\omega)+\eta, \\ \infty, & \text{otherwise}, \end{cases}$$
$$= E_2(\kappa_\alpha,\omega,R) - O(\eta). \tag{100}$$

Next, consider the event $\mathcal{E}_c$. Assume that $R'+\eta > R$ (i.e., binning is required). Let

$$\mathcal{F}_{3,k}(\eta) := \{\tilde{U}\tilde{V}\tilde{W}\tilde{W}_d \in \mathcal{T}_k(\mathcal{U}\times\mathcal{V}\times\mathcal{W}\times\mathcal{W}) : \tilde{U}\tilde{W} \in \mathcal{D}_k^{UW}(\eta) \text{ and } \tilde{V}\tilde{W}_d \in \mathcal{D}_k^{VW}(\eta)\}.$$

Then, we can write (for sufficiently large k) that,

$$\mathbb{P}\left(\mathcal{E}_c|H=1\right)$$
$$\leq \sum_{\substack{\tilde{U}\tilde{V}\tilde{W}\tilde{W}_d \\ \in \mathcal{F}_{3,k}(\eta)}} \sum_{\substack{(u^k,v^k,w^k,\bar{w}^k) \\ \in \mathcal{T}_k(\tilde{U}\tilde{V}\tilde{W}\tilde{W}_d)}} \sum_{m' \in \mathcal{M}'\setminus\{0\}} \mathbb{P}\left(\bar{U}^k=u^k, \bar{V}^k=v^k, M'=m', W^k(M')=w^k|H=1\right)$$
$$\sum_{\substack{m\neq 0, \hat{m}\neq 0: \\ \hat{m}\neq m}} \mathbb{P}\left(M=m|H=1\right)\mathbb{P}\left(\hat{M}=\hat{m}|M=m\right)$$
$$\left[\sum_{\hat{m}' \in \mathcal{M}'\setminus\{0\}} \mathbb{P}\left(W^k(\hat{m}')=\bar{w}^k, f_B(\hat{m}')=\hat{m}|\bar{U}^k=u^k, \bar{V}^k=v^k, M'=m', W^k(m')=w^k\right)\right]$$
$$\leq \frac{2}{e^{kR_k}} \sum_{\substack{\tilde{U}\tilde{V}\tilde{W}\tilde{W}_d \\ \in \mathcal{F}_{3,k}(\eta)}} e^{kH(\tilde{U}\tilde{V}\tilde{W})} e^{-k\left(H(\tilde{U}\tilde{V})+D(\tilde{U}\tilde{V}||\bar{U}\bar{V})\right)} \frac{1}{e^{k\left(H(\tilde{W}|\tilde{U})-\eta\right)}} \frac{e^{k(R'+\eta)}}{e^{k\left(H(\tilde{W}_d)-\eta\right)}} e^{kH(\tilde{W}_d|\tilde{V})}\ e^{-k\tau\left(E_x\left(\frac{R}{\tau},P_{SX}\right)-\eta\right)} \tag{101}$$
$$\leq e^{-kE_{3,k}},$$

where,

$$E_{3,k} := \min_{\tilde{U}\tilde{V}\tilde{W}\tilde{W}_d \in \mathcal{F}_{3,k}(\eta)} -H(\tilde{U}\tilde{V}\tilde{W}) + H(\tilde{U}\tilde{V}) + D(\tilde{U}\tilde{V}||\bar{U}\bar{V}) + H(\tilde{W}|\tilde{U}) + \tau E_x\left(\frac{R}{\tau},P_{SX}\right)$$
$$+ \rho_k + R - R' - O(\eta) - \delta_k',$$

and δ_k' is as defined in (98). To obtain (101), we used (68), (94) and (96). On the other hand, if $R'+\eta \leq R$, it can be shown similarly that,

$$\mathbb{P}\left(\mathcal{E}_c|H=1\right) \leq e^{-kE_{3,k}'},$$

where

$$E'_{3,k} := \min_{\tilde{U}\tilde{V}\tilde{W}\tilde{W}_d \in \mathcal{F}_{3,k}(\eta)} -H(\tilde{U}\tilde{V}\tilde{W}) + H(\tilde{U}\tilde{V}) + D(\tilde{U}\tilde{V}||\bar{U}\bar{V}) + H(\tilde{W}|\tilde{U}) + \tau E_x\left(\frac{R}{\tau}, P_{SX}\right) + \rho_k - O(\eta) - \frac{|\mathcal{U}||\mathcal{V}||\mathcal{W}|^2}{k}\log(k+1) - \frac{\log(2)}{k}.$$

Hence, we obtain

$$\begin{aligned}
&\frac{-1}{k}\log\left(\mathbb{P}\left(\mathcal{E}_c|H=1\right)\right)\\
&\overset{(k)}{\geq} \begin{cases} \min\limits_{\tilde{U}\tilde{V}\tilde{W}\in\mathcal{T}_3(\kappa_\alpha,\omega)} D(\tilde{U}\tilde{V}\tilde{W}||\bar{U}\bar{V}\bar{W}) + E_b(\kappa_\alpha,\omega,R) + \tau E_x\left(\frac{R}{\tau},P_{SX}\right) - O(\eta), & \text{if } R < \zeta_q(\kappa_\alpha,\omega)+\eta,\\ \min\limits_{\tilde{U}\tilde{V}\tilde{W}\in\mathcal{T}_3(\kappa_\alpha,\omega)} D(\tilde{U}\tilde{V}\tilde{W}||\bar{U}\bar{V}\bar{W}) + \rho(\kappa_\alpha,\omega) + \tau E_x\left(\frac{R}{\tau},P_{SX}\right) - O(\eta), & \text{otherwise,}\end{cases}\\
&= E_3(\kappa_\alpha,\omega,R,P_{SX},\tau) - O(\eta).
\end{aligned} \tag{102}$$

Finally, we consider the event $\mathcal{E}_d$. Assume that $R'+\eta > R$. We have

$$\mathbb{P}\left(\mathcal{E}_d|H=1\right) = \sum_{\substack{u^k\in\mathcal{T}_k(\tilde{U}):\\ \tilde{U}\in\mathcal{D}_k^U(\eta)}} \mathbb{P}\left(\bar{U}^k=u^k,\mathcal{E}_{EE},\mathcal{E}_d|H=1\right) + \sum_{\substack{u^k\in\mathcal{T}_k(\tilde{U}):\\ \tilde{U}\notin\mathcal{D}_k^U(\eta)}} \mathbb{P}\left(\bar{U}^k=u^k,\mathcal{E}_d|H=1\right), \tag{103}$$

where, (103) follows from the fact that if $\tilde{U}\in\mathcal{D}_k^U(\eta)$, then $\mathcal{E}_d$ can occur only if $\mathcal{E}_{EE}$ occurs. From (73), for any $u^k\in\mathcal{T}_k(\tilde{U})$ such that $\tilde{U}\in\mathcal{D}_k^U(\eta)$, we have

$$\mathbb{P}\left(\bar{U}^k=u^k,\mathcal{E}_{EE},\mathcal{E}_d|H=1\right) \leq e^{-e^{k\Omega(\eta)}}.$$

Next, note that if $\tilde{U}\notin\mathcal{D}_k^U(\eta)$, then $M'=0$ is chosen with probability 1 independent of the codebook $\mathcal{C}_k$. Hence, we can write the second term in (103) as follows:

$$\begin{aligned}
&\sum_{\substack{u^k\in\mathcal{T}_k(\tilde{U}):\\ \tilde{U}\notin\mathcal{D}_k^U(\eta)}} \mathbb{P}\left(\bar{U}^k=u^k,\mathcal{E}_d|H=1\right)\\
&\leq \sum_{\substack{u^k\in\mathcal{T}_k(\tilde{U}):\\ \tilde{U}\notin\mathcal{D}_k^U(\eta)}} \sum_{\substack{(v^k,\bar{w}^k)\in\mathcal{T}_k(\tilde{V}\tilde{W}_d):\\ \tilde{V}\tilde{W}_d\in\mathcal{D}_k^{VW}(\eta)}} \sum_{\hat{m}\in\mathcal{M}\setminus\{0\}} \mathbb{P}\left(\bar{U}^k=u^k,\bar{V}^k=v^k,M'=M=0|H=1\right)\mathbb{P}\left(\hat{M}=\hat{m}|M=0\right)\\
&\qquad \sum_{\hat{m}'\in\mathcal{M}'\setminus\{0\}} \mathbb{P}\left(f_B(\hat{m}')=\hat{m},W^k(\hat{m}')=\bar{w}^k\right)\\
&\leq \sum_{\substack{u^k\in\mathcal{T}_k(\tilde{U}):\\ \tilde{U}\notin\mathcal{D}_k^U(\eta)}} \sum_{\substack{(v^k,\bar{w}^k)\in\mathcal{T}_k(\tilde{V}\tilde{W}_d):\\ \tilde{V}\tilde{W}_d\in\mathcal{D}_k^{VW}(\eta)}} \mathbb{P}\left(\bar{U}^k=u^k,\bar{V}^k=v^k,M'=M=0|H=1\right) \sum_{\hat{m}\in\mathcal{M}\setminus\{0\}} \mathbb{P}\left(\hat{M}=\hat{m}|M=0\right)\\
&\qquad \sum_{\hat{m}'\in\mathcal{M}'\setminus\{0\}} \frac{1}{e^{kR_k}}\,\frac{1}{e^{k\left(H(\tilde{W}_d)-\eta\right)}}\\
&\leq \sum_{\substack{u^k\in\mathcal{T}_k(\tilde{U}):\\ \tilde{U}\notin\mathcal{D}_k^U(\eta)}} \sum_{\substack{(v^k,\bar{w}^k)\in\mathcal{T}_k(\tilde{V}\tilde{W}_d):\\ \tilde{V}\tilde{W}_d\in\mathcal{D}_k^{VW}(\eta)}} \mathbb{P}\left(\bar{U}^k=u^k,\bar{V}^k=v^k,M'=M=0|H=1\right) \sum_{\hat{m}\in\mathcal{M}\setminus\{0\}} \mathbb{P}\left(\hat{M}=\hat{m}|M=0\right)\\
&\qquad \frac{e^{k(R'+\eta)}}{e^{kR_k}}\,\frac{1}{e^{k\left(H(\tilde{W}_d)-\eta\right)}}\\
&\leq \sum_{\substack{u^k\in\mathcal{T}_k(\tilde{U}):\\ \tilde{U}\notin\mathcal{D}_k^U(\eta)}} \sum_{\substack{(v^k,\bar{w}^k)\in\mathcal{T}_k(\tilde{V}\tilde{W}_d):\\ \tilde{V}\tilde{W}_d\in\mathcal{D}_k^{VW}(\eta)}} \mathbb{P}\left(\bar{U}^k=u^k,\bar{V}^k=v^k,M'=M=0|H=1\right)\, e^{-k\tau(E_m(P_{SX},\theta)-\theta-\eta)}\\
&\qquad \frac{e^{k(R'+\eta)}}{e^{kR_k}}\,\frac{1}{e^{k\left(H(\tilde{W}_d)-\eta\right)}} \qquad (104)\\
&\leq \sum_{\substack{\tilde{U}\tilde{V}\tilde{W}_d\\ \in\mathcal{D}_k^U(\eta)^c\times\mathcal{D}_k^{VW}(\eta)}} e^{kH(\tilde{U}\tilde{V})}e^{-k\left(H(\tilde{U}\tilde{V})+D(\tilde{U}\tilde{V}||\bar{U}\bar{V})\right)}\, e^{-k\tau(E_m(P_{SX},\theta)-\theta-\eta)}\,\frac{e^{k(R'+\eta)}}{e^{kR_k}}\,\frac{e^{kH(\tilde{W}_d|\tilde{V})}}{e^{k\left(H(\tilde{W}_d)-\eta\right)}}
\end{aligned}$$

$$\leq e^{-kE_{4,k}},$$

where,

$$E_{4,k} := \min_{\substack{\tilde{U}\tilde{V}\tilde{W}_d \\ \in \mathcal{D}_k^U(\eta)^c \times \mathcal{D}_k^{VW}(\eta)}} D(\tilde{U}\tilde{V}||\bar{U}\bar{V}) + \tau\left(E_m(P_{SX},\theta) - \theta\right) + \rho_k + R - R' - O(\eta) - \frac{|\mathcal{U}||\mathcal{V}||\mathcal{W}|}{k}\log(k+1)$$

$$\geq \min_{\substack{\tilde{V}:\exists \tilde{W}, \\ \tilde{V}\tilde{W} \in \mathcal{D}_k^{VW}(\eta)}} D(\tilde{V}||\bar{V}) + \tau\left(E_m(P_{SX},\theta) - \theta\right) + \rho_k + R - R' - O(\eta) - \frac{|\mathcal{U}||\mathcal{V}||\mathcal{W}|}{k}\log(k+1).$$

In (104), we used (67).

If $R' + \eta \leq R$, it can be shown that,

$$\mathbb{P}\left(\mathcal{E}_c|H=1\right) \leq e^{-kE'_{4,k}},$$

where

$$E'_{4,k} \geq \min_{\substack{\tilde{V}:\exists \tilde{W}, \\ \tilde{V}\tilde{W} \in \mathcal{D}_k^{VW}(\eta)}} D(\tilde{V}||\bar{V}) + \tau\left(E_m(P_{SX},\theta) - \theta\right) + \rho_k - O(\eta) - \frac{|\mathcal{U}||\mathcal{V}||\mathcal{W}|}{k}\log(k+1).$$

Hence, we obtain

$$\frac{-1}{k}\log\left(\mathbb{P}\left(\mathcal{E}_d|H=1\right)\right)$$
$$\overset{(k)}{\geq} \begin{cases} \min_{\substack{\tilde{V}:\exists \tilde{W}, \\ \tilde{V}\tilde{W} \in \mathcal{D}_k^{VW}(\eta)}} D(\tilde{V}||\bar{V}) + E_b(\kappa_\alpha,\omega,R) + \tau\left(E_m(P_{SX},\theta) - \theta\right) - O(\eta) & \text{if } R < \zeta_q(\kappa_\alpha,\omega) + \eta, \\ \min_{\substack{\tilde{V}:\exists \tilde{W}, \\ \tilde{V}\tilde{W} \in \mathcal{D}_k^{VW}(\eta)}} D(\tilde{V}||\bar{V}) + \rho(\kappa_\alpha,\omega) + \tau\left(E_m(P_{SX},\theta) - \theta\right) - O(\eta), & \text{otherwise}, \end{cases}$$
$$= E_4(\kappa_\alpha,\omega,R,P_{SX},\theta,\tau) - O(\eta). \tag{105}$$

Since the exponent of the type II error probability is lower bounded by the minimum of the exponent of the type II error causing events, it follows from (91), (100), (102) and (105) that for a fixed $(\omega,R,P_{SX},\theta) \in \mathcal{L}(\kappa_\alpha,\tau)$,

$$\kappa(\tau,\kappa_\alpha) \geq \min\left(E_1(\kappa_\alpha,\omega), E_2(\kappa_\alpha,\omega,R), E_3(\kappa_\alpha,\omega,R,P_{SX},\tau), E_4(\kappa_\alpha,\omega,R,P_{SX},\theta,\tau)\right) - O(\eta). \tag{106}$$

To complete the proof, we need to show (73), (85) and (97). Since $W^k(j), j \in \mathcal{M}'_i$, is selected uniformly at random from the set $\mathcal{T}_k(\hat{W}_i)$, we have from [28, Lemma 2.5] that, for any $u^k \in \mathcal{T}_k(\hat{U}_i)$ and sufficiently large k,

$$\mathbb{P}\left((u^k, W^k(j)) \notin \mathcal{T}_k(\hat{U}_i\hat{W}_i)\right) \leq \left(1 - \frac{e^{k\left(H(\hat{W}_i|\hat{U}_i) - \frac{\eta}{4}\right)}}{e^{kH(\hat{W}_i)}}\right). \tag{107}$$

Since the codewords are selected independently, we have by the union bound that

$$\mathbb{P}\left(\nexists (u^k, W^k(j)) \notin \mathcal{T}_k(\hat{U}_i\hat{W}_i),\ j \in \mathcal{M}'_i\right) \leq \left(1 - \frac{e^{k\left(H(\hat{W}_i|\hat{U}_i) - \frac{\eta}{4}\right)}}{e^{kH(\hat{W}_i)}}\right)^{e^{kR'_i}}$$
$$\leq e^{-e^{k\left(R'_i - I(\hat{U}_i;\hat{W}_i) - \frac{\eta}{4}\right)}}. \tag{108}$$

Hence, by the choice of R'_i in (62), we have for sufficiently large k that

$$\mathbb{P}(\mathcal{E}_{EE}) = \sum_{i=1}^{|\mathcal{D}_k^U(\eta)|} e^{-e^{k\frac{\eta}{12}}} \leq (k+1)^{|\mathcal{U}|} e^{-e^{k\frac{\eta}{12}}} \leq e^{-e^{k\frac{\eta}{15}}}. \tag{109}$$

This completes the proof of (73).

Next, we prove (85). Note that by the encoding procedure, $M' \neq 0$ and $w^k \in \mathcal{T}_k(\hat{W}_i)$ for some $1 \leq i \leq |\mathcal{D}_k^U(\eta)|$ implies that $U^k \in \mathcal{T}_k(P_{\hat{U}_i|\hat{W}_i}, w^k)$. Hence, we can write for $j \neq m'$, that

$$\mathbb{P}\left(W^k(j) = \tilde{w}^k|V^k = v^k, M' = m' \neq 0, M = m, f_B(j) = m, W^k(m') = w^k\right)$$

$$= \sum_{u^k \in \mathcal{T}_k(P_{\hat{U}_i|\hat{W}_i}, w^k)} \mathbb{P}\left(U^k = u^k|V^k = v^k, M' = m' \neq 0, M = m, f_B(j) = m, W^k(m') = w^k\right)$$
$$\mathbb{P}\left(W^k(j) = \tilde{w}^k|U^k = u^k, V^k = v^k, M' = m' \neq 0, M = m, f_B(j) = m, W^k(m') = w^k\right)$$

Let

$$\mathcal{C}^-_{m',j} := \mathcal{C}_k \backslash \{W^k(m'), W^k(j)\},$$
$$\mathcal{E} := \{U^k = u^k, V^k = v^k, M' = m' \neq 0, M = m, f_B(j) = m, W^k(m') = w^k\}.$$

Then, we can write,

$$\mathbb{P}\left(W^k(j) = \tilde{w}^k|\mathcal{E}\right) = \sum_{\mathcal{C}^-_{m',j}=c} \mathbb{P}(\mathcal{C}^-_{m',j} = c|\mathcal{E})\mathbb{P}(W^k(j) = \tilde{w}^k|\mathcal{E}, \mathcal{C}^-_{m',j} = c). \tag{110}$$

We can write the term within the summation in (110) as follows:

$$\begin{aligned}
&\mathbb{P}(W^k(j) = \tilde{w}^k|\mathcal{E}, \mathcal{C}^-_{m',j} = c) \\
&= \mathbb{P}(W^k(j) = \tilde{w}^k|U^k = u^k, V^k = v^k,\ \mathcal{C}^-_{m',j} = c) \\
&\quad \frac{\mathbb{P}(M' = m'|W^k(j) = \tilde{w}^k, W^k(m') = w^k, U^k = u^k, V^k = v^k, \mathcal{C}^-_{m',j} = c)}{\mathbb{P}(M' = m'|W^k(m') = w^k, U^k = u^k, V^k = v^k, \mathcal{C}^-_{m',j} = c)} \\
&\quad \frac{\mathbb{P}(M = m, f_B(j) = m|M' = m', W^k(j) = \tilde{w}^k, W^k(m') = w^k, U^k = u^k, V^k = v^k, \mathcal{C}^-_{m',j} = c)}{\mathbb{P}(M = m, f_B(j) = m|M' = m', W^k(m') = w^k, U^k = u^k, V^k = v^k, \mathcal{C}^-_{m',j} = c)} && (111) \\
&= \mathbb{P}(W^k(j) = \tilde{w}^k)\ \frac{\mathbb{P}(M' = m'|W^k(j) = \tilde{w}^k, W^k(m') = w^k, U^k = u^k, V^k = v^k, \mathcal{C}^-_{m',j} = c)}{\mathbb{P}(M' = m'|W^k(m') = w^k, U^k = u^k, V^k = v^k, \mathcal{C}^-_{m',j} = c)}. && (112)
\end{aligned}$$

In (112), we used

$$\begin{aligned}
&\mathbb{P}(M = m, f_B(j) = m|M' = m', W^k(j) = \tilde{w}^k, W^k(m') = w^k, U^k = u^k, V^k = v^k, \mathcal{C}^-_{m',j} = c) \\
&= \mathbb{P}(M = m, f_B(j) = m|M' = m', W^k(m') = w^k, U^k = u^k, V^k = v^k, \mathcal{C}^-_{m',j} = c) \\
&= \mathbb{P}(M = m, f_B(j) = m),
\end{aligned}$$

which in turn follows from the fact that the binning is performed independent of the $\mathcal{C}_k$, U^k and V^k. Let

$$N(u^k, \mathcal{C}^-_{m',j}) = |\{w^k(l) \in \mathcal{C}^-_{m',j} : l \neq m', j,\ (u^k, w^k(l)) \in \mathcal{T}_k(\hat{U}_i\hat{W}_i)\}|.$$

Recall that if there are multiple indices l in the codebook $\mathcal{C}_k$ such that $(u^k, w^k(l)) \in \mathcal{T}_k(\hat{U}_i\hat{W}_i)$, then the encoder selects one of them uniformly at random. Also, note that since $M' = m' \neq 0$, $(u^k, w^k(m')) \in \mathcal{T}_k(\hat{U}_i\hat{W}_i)$. Thus, if $(u^k, \tilde{w}^k) \in \mathcal{T}_k(\hat{U}_i\hat{W}_i)$, then

$$\begin{aligned}
&\frac{\mathbb{P}(M' = m'|W^k(j) = \tilde{w}^k, W^k(m') = w^k, U^k = u^k, V^k = v^k, \mathcal{C}^-_{m',j} = c)}{\mathbb{P}(M' = m'|W^k(m') = w^k, U^k = u^k, V^k = v^k, \mathcal{C}^-_{m',j} = c)} \\
&= \left[\frac{1}{N(u^k, \mathcal{C}^-_{m',j}) + 2}\right] \frac{1}{\mathbb{P}(M = m|U^k = u^k, V^k = v^k, \mathcal{C}^-_{m',j} = c)} \\
&\leq \frac{N(u^k, \mathcal{C}^-_{m',j}) + 2}{N(u^k, \mathcal{C}^-_{m',j}) + 2} = 1. && (113)
\end{aligned}$$

On the other hand, if $(u^k, \tilde{w}^k) \notin \mathcal{T}_k(\hat{U}_i\hat{W}_i)$, then

$$\begin{aligned}
&\frac{\mathbb{P}(M' = m'|W^k(j) = \tilde{w}^k, W^k(m') = w^k, U^k = u^k, V^k = v^k, \mathcal{C}^-_{m',j} = c)}{\mathbb{P}(M' = m'|W^k(m') = w^k, U^k = u^k, V^k = v^k, \mathcal{C}^-_{m',j} = c)} \\
&= \left[\frac{1}{N(u^k, \mathcal{C}^-_{m',j}) + 1}\right] \frac{1}{\mathbb{P}(M = m|U^k = u^k, V^k = v^k, \mathcal{C}^-_{m',j} = c)} \leq \frac{N(u^k, \mathcal{C}^-_{m',j}) + 2}{N(u^k, \mathcal{C}^-_{m',j}) + 1} \leq 2. && (114)
\end{aligned}$$

Substituting (113) and (114) in (110), we obtain (85). The proof of (97) is similar to that of (85), and hence, omitted.

Thus, we have shown that for a fixed $(\omega, R, P_{SX}, \theta) \in \mathcal{L}(\kappa_\alpha, \tau)$, the probability of type I and type II error probabilities averaged over the ensemble of randomly generated codebooks and binning functions satisfy

$$\mathbb{P}\left(\hat{H} = 1|H = 0\right) \leq e^{-k(\kappa_\alpha - O(\eta))}, \tag{115}$$
$$\text{and } \mathbb{P}\left(\hat{H} = 0|H = 1\right) \leq e^{-k(\kappa_s^*(\tau, \kappa_\alpha) - O(\eta))}, \tag{116}$$

for all sufficiently large k. By the random coding argument followed by an expurgation step [30], there exists a deterministic codebook $\mathcal{C}_k$ and deterministic binning function f_b such that (115) and (116) are satisfied. Maximizing over $(\omega, R, P_{SX}, \theta) \in$

$\mathcal{L}(\kappa_\alpha, \tau)$ and noting that $\eta > 0$ is arbitrary completes the proof.

APPENDIX C
PROOF OF THEOREM 5

Fix $\kappa_\alpha > 0$ and $\left(P_S, \omega'(\cdot, P_S), P_{X|USW'}, P_{X'|S}\right) \in \mathcal{L}_h(\kappa_\alpha)$. Let $\eta > 0$ be a small number, and choose a sequence $s^n \in \mathcal{T}_n(\hat{S}^*)$ which is revealed to both the encoder and the detector, where $\hat{S}^*$ satisfies $D(\hat{S}^*||S) \leq \eta$. Let $R' := \zeta'_q(\kappa_\alpha, \omega', P_{\hat{S}^*})$.

Encoding:
The encoder performs type based quantization followed by channel coding similar to that in hybrid coding [33]. The details are as follows:
Quantization scheme: Let

$$\mathcal{D}_n^U(\eta) := \{\hat{U} \in \mathcal{T}_n(\mathcal{U}) : D(\hat{U}||U) \leq \kappa_\alpha + \eta\}. \tag{117}$$

Consider some ordering on the types in $\mathcal{D}_n^U(\eta)$ and denote the elements as $\hat{U}_1$, $\hat{U}_2$, ..., etc. For each joint type variable $\hat{S}^*\hat{U}_i$, $\hat{U}_i \in \mathcal{D}_n^U(\eta)$, $1 \leq i \leq |\mathcal{D}_n^U(\eta)|$, such that $\hat{S}^* \perp \hat{U}_i$, choose a joint type variable $\hat{S}^*\hat{U}_i\hat{W}'_i$, $\hat{W}'_i \in \mathcal{T}_n(\mathcal{W}')$, such that

$$D\left(\hat{W}'_i|\hat{U}_i, \hat{S}^*||W'_i|U, \hat{S}^*\middle|\hat{U}_i, \hat{S}^*\right) \leq \frac{\eta}{3},$$
$$I(\hat{S}^*, \hat{U}_i; \hat{W}'_i) \leq R' + \frac{\eta}{3},$$

where $P_{W'_i|U,S} = \omega'(P_{\hat{U}_i}, P_{\hat{S}^*})$. Let

$$\mathcal{D}_n^{SUW'}(\eta) := \{\hat{S}^*\hat{U}_i\hat{W}'_i : 1 \leq i \leq |\mathcal{D}_n^U(\eta)|\},$$
$$\text{and } R'_i := I(\hat{S}^*, \hat{U}_i; \hat{W}'_i) + \frac{\eta}{3}, 1 \leq i \leq |\mathcal{D}_n^U(\eta)|. \tag{118}$$

Let

$$\mathcal{C}'_n = \left\{ w'^n(j), j \in \left[1 : \sum_{i=1}^{|\mathcal{D}_n^U(\eta)|} e^{nR'_i}\right]\right\},$$

denote a quantization codebook such that each codeword $w'^n(j)$, $j \in \mathcal{M}'_i := [1 + \sum_{m=1}^{i-1} e^{nR'_m} : \sum_{m=1}^{i} e^{nR'_m}]$, $1 \leq i \leq |\mathcal{D}_n^U(\eta)|$, belongs to the set $\mathcal{T}_n(\hat{W}'_i)$. For $u^n \in \mathcal{T}_n(\hat{U}_i)$ such that $\hat{U}_i \in \mathcal{D}_n(U)$ for some $1 \leq i \leq |\mathcal{D}_n^U(\eta)|$, let

$$\mu'(u^n, \mathcal{C}'_n) := \{j \in \mathcal{M}'_i : \ w'^n(j) \in \mathcal{C}'_n \text{ and } (s^n, u^n, w'^n(j)) \in \mathcal{T}_n(\hat{S}^*\hat{U}_i\hat{W}'_i), \ \hat{S}^*\hat{U}_i\hat{W}'_i \in \mathcal{D}_n^{SUW'}(\eta)\}.$$

If $|\mu'(u^n, \mathcal{C}'_n)| \geq 1$, let $M'(u^n, \mathcal{C}'_n)$ denote an index selected uniformly at random from the set $\mu'(u^n, \mathcal{C}'_n)$, otherwise, set $M'(u^n, \mathcal{C}'_n) = 0$. Given $\mathcal{C}'_n$ and $u^n \in \mathcal{U}^n$, the quantizer outputs $M' = M'(u^n, \mathcal{C}'_k)$, where the support of M' is given by

$$\mathcal{M}' := \left[0 : \sum_{i=1}^{|\mathcal{D}_n^U(\eta)|} e^{nR'_i}\right].$$

Note that for sufficiently large n, it follows similarly to (63) that

$$|\mathcal{M}'| \leq e^{n(R'+\eta)}.$$

If $M' = m' \neq 0$, the encoder transmits X^n over the channel, where $X^n = x^n$ is generated according to the distribution $\prod_{i=1}^n P_{X|USW'}(x_i|u_i, s_i, w'_i(m'))$. If $M' = 0$, the encoder transmits $X'^n = x'^n$ randomly according to $\prod_{i=1}^n P_{X'|S}(x'_i|s_i)$.

Decoding:
For a given codebook $\mathcal{C}'_n$ and $m' \in \mathcal{M}'\backslash\{0\}$, let $\mathcal{O}_{m'}$ denote the set of u^n such that $M'(u^n, \mathcal{C}'_n) = m'$. For each $m' \in \mathcal{M}'\backslash\{0\}$ and $u^n \in \mathcal{O}_{m'}$, let

$$\mathcal{B}'_{m'}(u^n) = \{(v^n, y^n) \in \mathcal{V}^n \times \mathcal{Y}^n : (s^n, u^n, w'^n_{m'}, v^n, y^n) \in \mathcal{J}_n^{\kappa_\alpha+\eta}(\hat{S}^*UW'_{m'}VY)\},$$

where $\hat{S}^*UW'_{m'}VY$ is uniquely specified by $\hat{S}^* \perp (U, V)$

$$W'_{m'} - (U, \hat{S}^*) - V, \ Y - (U, \hat{S}^*, W'_{m'}) - V, \ P_{W'_{m'}|U\hat{S}^*} = \omega'(P_{u^n}, P_{\hat{S}^*}), \tag{119}$$
$$P_{Y|U\hat{S}^*W'_{m'}}(y|u, s, w') = \sum\nolimits_{x \in \mathcal{X}} P_{X|U\hat{S}^*W'_{m'}}(x|u, s, w')P_{Y|X}(y|x), \ \forall \ (y, u, s, w') \in \mathcal{Y} \times \mathcal{U} \times \mathcal{S} \times \mathcal{W}'. \tag{120}$$

For $m' \in \mathcal{M}'\backslash\{0\}$, we define

$$\mathcal{B}'_{m'} := \{(v^n, y^n) : (v^n, y^n) \in B_{m'}(u^n) \text{ for some } u^n \in \mathcal{O}_{m'}\}.$$

Define the acceptance region for H_0 at the detector as

$$\mathcal{A}_n := \underset{m' \in \mathcal{M}' \backslash 0}{\cup} s^n \times m' \times \mathcal{B}'_{m'},$$

or equivalently as

$$\mathcal{A}_n^e := \underset{m' \in \mathcal{M}' \backslash 0}{\cup} s^n \times \mathcal{O}_{m'} \times \mathcal{B}'_{m'}.$$

Given $Y^n = y^n$ and $V^n = v^n$, if $(s^n, v^n, y^n) \in \{s^n\} \times \bigcup_{m' \in \mathcal{M}' \backslash \{0\}} \mathcal{B}'_{m'}$, then set $\hat{M}' = m'$, where

$$\hat{m}' := \underset{j \in \mathcal{M}' \backslash 0}{\arg\min} H_e(w'^n(j)|v^n, y^n, s^n).$$

Otherwise, set $\hat{M}' = 0$. If $\hat{M}' = 0$, $\hat{H} = 1$ is declared. Otherwise, $\hat{H} = 0$ or $\hat{H} = 1$ is declared depending on whether $(s^n, \hat{m}', v^n, y^n) \in \mathcal{A}_n$ or $(s^n, \hat{m}', v^n, y^n) \notin \mathcal{A}_n$, respectively.

Analysis of the type I and type II error probabilities:

Similar to Theorem 4, we will analyze the average type I and type II error probabilities over an ensemble of randomly generated quantization codebooks. Then, the standard random coding argument followed by the expurgation technique in [30] guarantees the existence of a deterministic quantization codebook that achieves the lower bound given in Theorem 5. Let each codeword $w'^n(j)$, $j \in \mathcal{M}'_i$, $1 \leq i \leq |\mathcal{D}_n^U(\eta)|$, be selected (with replacement) independently and uniformly at random from the set $\mathcal{T}_n(\hat{W}'_i)$ (see quantization scheme above). We proceed to analyze the type I and type II error probabilities averaged over these random codebooks. Note that a type I error can occur only under the following events:

(i) $\mathcal{E}'_{EE} := \bigcup_{\hat{U} \in \mathcal{D}'_n(SU)} \bigcup_{u^n \in \mathcal{T}_n(\hat{U})} \mathcal{E}'_{EE}(u^n)$, where

$$\mathcal{E}'_{EE}(u^n) := \left\{ \begin{array}{l} \nexists\, W'^n(j) \in \mathcal{C}'_n,\ j \in [1 : |\mathcal{M}'|],\ \text{s.t.}\ (s^n, u^n, W'^n(j)) \in \mathcal{T}_n(\hat{S}^* \hat{U}_i \hat{W}'_i),\ P_{\hat{S}^* \hat{U}_i} = P_{s^n u^n}, \\ \hat{S}^* \hat{U}_i \hat{W}'_i \in \mathcal{D}_n^{SUW'}(\eta) \end{array} \right\}.$$

(ii) $\hat{M}' = M'$.

(iii) $M' \neq 0$ and $\hat{M}' \neq M'$.

(iv) $M' = 0$ and $\hat{M}' \neq M'$.

Similar to (73), we have since R'_i satisfies (118), that

$$\mathbb{P}(\mathcal{E}'_{EE}) \leq e^{-e^{n\Omega(\eta)}}. \tag{121}$$

Next, consider event (ii). Due to (121), we can write

$$\mathbb{P}\left(\hat{H} = 1|\hat{M}' = M', H = 0\right) \leq e^{-e^{n\Omega(\eta)}} + \mathbb{P}\left(\hat{H} = 1|\hat{M}' = M', \mathcal{E}'^c_{EE}, H = 0\right). \tag{122}$$

The second term in (122) can be bounded as

$$\mathbb{P}\left(\hat{H} = 1|\hat{M}' = M', \mathcal{E}'^c_{EE}, H = 0\right) = \mathbb{P}\left((s^n, M', V^n, Y^n) \notin \mathcal{A}_n|\mathcal{E}'^c_{EE}, H = 0\right) \tag{123}$$

$$= 1 - \mathbb{P}\left((s^n, U^n, V^n, Y^n) \in \mathcal{A}_n^e|\mathcal{E}'^c_{EE}, H = 0\right) \tag{124}$$

We have similar to [4, Equation 4.17] that for $u^n \in \mathcal{O}_{m'}$ that

$$\mathbb{P}((V^n, Y^n) \in \mathcal{B}'_{m'}(u^n)|U^n = u^n, \mathcal{E}'^c_{EE}) = \mathbb{P}((V^n, Y^n) \in \mathcal{B}'_{m'}(u^n)|U^n = u^n, W'^n(m') = w'^n_{m'}, \mathcal{E}'^c_{EE}) \tag{125}$$

$$\geq 1 - e^{-n\left(\kappa_\alpha + \frac{\eta}{3} - D(P_{u^n}||P_U)\right)}. \tag{126}$$

Then, using (117) and (126), it follows similarly to [4, Equation 4.22] that

$$\mathbb{P}((s^n, U^n, V^n, Y^n) \in \mathcal{A}_e^n|\mathcal{E}'^c_{EE}) \geq 1 - e^{-n\kappa_\alpha}. \tag{127}$$

Substituting (127) in (124), it follows that

$$\mathbb{P}\left(\hat{H} = 1|\hat{M}' = M', \mathcal{E}'^c_{EE}, H = 0\right) \leq e^{-n\kappa_\alpha}. \tag{128}$$

The probability of event (iii) can be upper bounded as follows:

$$\mathbb{P}\left(M' \neq 0, \hat{M}' \neq M'|H = 0\right)$$
$$\leq \mathbb{P}\left(M' \neq 0, \hat{M}' \neq M', (s^n, M', V^n, Y^n) \in \mathcal{A}_n|H = 0\right) + \mathbb{P}\left(M' \neq 0, \hat{M}' \neq M', (s^n, M', V^n, Y^n) \notin \mathcal{A}_n|H = 0\right)$$

$$\leq \mathbb{P}\left(M' \neq 0, \hat{M}' \neq M', (s^n, M', V^n, Y^n) \in \mathcal{A}_n | H=0\right) + e^{-e^{n\Omega(\eta)}} + e^{-n\kappa_\alpha} \tag{129}$$

$$\leq \mathbb{P}\left(\hat{M}' \neq M' | M' \neq 0, (s^n, M', V^n, Y^n) \in \mathcal{A}_n, H=0\right) + e^{-e^{n\Omega(\eta)}} + e^{-n\kappa_\alpha}$$

$$\leq e^{-n\left(\rho'(\kappa_\alpha,\omega',P_S,P_{X|USW'}) - \zeta_q'(\kappa_\alpha,\omega',P_{\hat{S}^*}) - O(\eta)\right)} + e^{-e^{n\Omega(\eta)}} + e^{-n\kappa_\alpha} \tag{130}$$

where (129) follows similar to (79) using (121) and (127), and (130) follows similar to (89) by noting that $(s^n, M', V^n, Y^n) \in \mathcal{A}_n$ implies that $\hat{M}' \neq 0$. Also, from (121) and the definition of $\mathcal{D}_n^U(\eta)$, we can bound the probability of event (iv) as

$$\mathbb{P}\left(M' = 0, \hat{M}' \neq M' | H=0\right) \leq \mathbb{P}\left(M' = 0 | H=0\right) \leq e^{-n\kappa_\alpha}. \tag{131}$$

From (121), (128), (130) and (131), it follows that the type I error probability satisfies $e^{-k(\kappa_\alpha - O(\eta))}$, asymptotically.

Next, we analyze the type II error probability of the above scheme averaged over the random codebooks. For a given codebook $\mathcal{C}_n'$, let $\tilde{U}, \tilde{V}, \tilde{W}, \tilde{Y}$ and $\tilde{W}_d$ denote the type variable for the realizations of $\bar{U}^n, \bar{V}^n, W'^n(M')$ $(M' \neq 0)$, $\bar{Y}^n$ and $W'^n(\hat{M}')$ $(\hat{M}' \neq 0)$, respectively. Let

$$\mathcal{D}_n^{SVW'Y}(\eta) := \left\{ \begin{array}{ll} \hat{S}^*\hat{V}\hat{W}\hat{Y}: & \exists\, (s^n, u^n, v^n, w^n, y^n) \in \underset{m' \in \mathcal{M}' \setminus \{0\}}{\cup} \mathcal{J}_n^{\kappa_\alpha + \eta}(\hat{S}^* U V W'_{m'} Y),\ \hat{S}^* U V W'_{m'} Y \text{ satisfies (119)} \\ & \text{and (120), and } P_{s^n u^n v^n w^n y^n} = P_{\hat{S}^*\hat{U}\hat{V}\hat{W}\hat{Y}} \end{array} \right\}.$$

A type II error can occur only under the following events:

(a)

$$\mathcal{E}_a' := \left\{ \begin{array}{l} \hat{M}' = M' \neq 0, (s^n, \bar{U}^n, \bar{V}^n, W'^n(M'), \bar{Y}^n) \in \mathcal{T}_n(\hat{S}^*\hat{U}\hat{V}\hat{W}\hat{Y}) \text{ s.t. } \hat{U}\hat{W} \in \mathcal{D}_n^{SUW'}(\eta) \\ \text{and } \hat{S}^*\hat{V}\hat{W}\hat{Y} \in \mathcal{D}_n^{SVW'Y}(\eta) \end{array} \right\}.$$

(b)

$$\mathcal{E}_b' := \left\{ \begin{array}{l} M' \neq 0, \hat{M}' \neq M', (s^n, \bar{U}^n, \bar{V}^n, W'^n(M'), \bar{Y}^n, W'^n(\hat{M}')) \in \mathcal{T}_n(\hat{S}^*\hat{U}\hat{V}\hat{W}\hat{Y}\hat{W}_d) \text{ s.t.} \\ \hat{S}^*\hat{U}\hat{W} \in \mathcal{D}_n^{SUW'}(\eta), \hat{S}^*\hat{V}\hat{W}_d\hat{Y} \in \mathcal{D}_n^{SVW'Y}(\eta), \text{ and } H_e\left(W'^n(\hat{M}')|s^n, \bar{V}^n, \bar{Y}^n\right) \\ \leq H_e\left(W'^n(M')|s^n, \bar{V}^n, \bar{Y}^n\right) \end{array} \right\}.$$

(c) $\mathcal{E}_c' := \{M' = 0, \hat{M}' \neq M', (s^n, \bar{V}^n, \bar{Y}^n, W'^n(\hat{M}')) \in \mathcal{T}_n(\hat{S}^*\hat{V}\hat{Y}\hat{W}_d) \text{ s.t. } \hat{S}^*\hat{V}\hat{W}_d\hat{Y} \in \mathcal{D}_n^{SVW'Y}(\eta)\}$.

Let

$$\mathcal{F}_{1,n}'(\eta) := \{\hat{S}^*\tilde{U}\tilde{V}\tilde{W}\tilde{Y} \in \mathcal{T}_n(\mathcal{S} \times \mathcal{U} \times \mathcal{V} \times \mathcal{W}' \times \mathcal{Y}) : \hat{S}^*\tilde{U}\tilde{W} \in \mathcal{D}_n^{SUW'}(\eta), \hat{S}^*\tilde{V}\tilde{W}\tilde{Y} \in \mathcal{D}_n^{SVW'Y}(\eta)\}.$$

Then, we can write

$$\begin{aligned}
&\mathbb{P}\left(\mathcal{E}_a' | H=1\right) \\
&\leq \sum_{\substack{\hat{S}^*\tilde{U}\tilde{V}\tilde{W}\tilde{Y} \\ \in \mathcal{F}_{1,n}'(\eta)}} \sum_{\substack{(u^n, v^n, w'^n, y^n): \\ (s^n, u^n, v^n, w'^n, y^n) \\ \in \mathcal{T}_n(\hat{S}^*\tilde{U}\tilde{V}\tilde{W}\tilde{Y})}} \sum_{m' \in \mathcal{M}' \setminus \{0\}} \mathbb{P}\left(\bar{U}^n = u^n, \bar{V}^n = v^n, M' = m', W'^n(m') = w'^n, \bar{Y}^n = y^n | S^n = s^n\right) \\
&\leq \sum_{\substack{\hat{S}^*\tilde{U}\tilde{V}\tilde{W}\tilde{Y} \\ \in \mathcal{F}_{1,n}'(\eta)}} \sum_{\substack{(u^n, v^n, w'^n, y^n): \\ (s^n, u^n, v^n, w'^n, y^n) \\ \in \mathcal{T}_n(\hat{S}^*\tilde{U}\tilde{V}\tilde{W}\tilde{Y})}} \sum_{m' \in \mathcal{M}' \setminus \{0\}} \mathbb{P}\left(\bar{U}^n = u^n, \bar{V}^n = v^n, M' = m' | S^n = s^n\right) \\
&\qquad \mathbb{P}\left(W'^n(m') = w'^n | \bar{U}^n = u^n, \bar{V}^n = v^n, M' = m', S^n = s^n\right) \\
&\qquad \mathbb{P}\left(\bar{Y}^n = y^n | \bar{U}^n = u^n, \bar{V}^n = v^n, M' = m', W'^n(m') = w'^n, S^n = s^n\right) \\
&\leq \sum_{\substack{\hat{S}^*\tilde{U}\tilde{V}\tilde{W}\tilde{Y} \\ \in \mathcal{F}_{1,n}'(\eta)}} \sum_{\substack{(u^n, v^n, w'^n, y^n): \\ (s^n, u^n, v^n, w'^n, y^n) \\ \in \mathcal{T}_n(\hat{S}^*\tilde{U}\tilde{V}\tilde{W}\tilde{Y})}} \sum_{\substack{m' \in \\ \mathcal{M}' \setminus \{0\}}} e^{-n\left(H(\tilde{U}\tilde{V}) + D(\tilde{U}\tilde{V}||\bar{U}\bar{V})\right)}\, \mathbb{P}\left(M' = m' | \bar{U}^n = u^n, \bar{V}^n = v^n, S^n = s^n\right) \\
&\qquad \frac{1}{e^{n\left(H(\tilde{W}|\hat{S}^*\tilde{U}) - \eta\right)}}\; e^{-n\left(H(\tilde{Y}|\tilde{U}\hat{S}^*\tilde{W}) + D\left(\tilde{Y}|\tilde{U}\hat{S}^*\tilde{W}||Y|USW'\middle|\tilde{U}\hat{S}^*\tilde{W}\right)\right)}
\end{aligned} \tag{132}$$

$$\leq \sum_{\substack{\hat{S}^*\tilde{U}\tilde{V}\tilde{W}\tilde{Y}\\ \in \mathcal{F}'_{1,n}(\eta)}} e^{nH(\tilde{U}\tilde{V}\tilde{W}\tilde{Y}|\hat{S}^*)} e^{-n\left(H(\tilde{U}\tilde{V})+D(\tilde{U}\tilde{V}||\bar{U}\bar{V})\right)} \frac{1}{e^{n\left(H(\tilde{W}|\hat{S}^*\tilde{U})-\eta\right)}} e^{-n\left(H(\tilde{Y}|\tilde{U}\hat{S}^*\tilde{W})+D\left(\tilde{Y}|\tilde{U}\hat{S}^*\tilde{W}||Y|USW'\middle|\tilde{U}\hat{S}^*\tilde{W}\right)\right)}$$

$$\leq e^{-nE'_{1,n}}, \tag{133}$$

where

$$E'_{1,n} := \min_{\hat{S}^*\tilde{U}\tilde{V}\tilde{W}\tilde{Y}\in\mathcal{F}'_{1,n}(\eta)} H(\tilde{U}\tilde{V}) + D(\tilde{U}\tilde{V}||\bar{U}\bar{V}) + H(\tilde{W}|\hat{S}^*\tilde{U}) - \eta + H(\tilde{Y}|\tilde{U}\hat{S}^*\tilde{W}) + D\left(\tilde{Y}|\tilde{U}\hat{S}^*\tilde{W}||Y|USW'\middle|\tilde{U}\hat{S}^*\tilde{W}\right)$$
$$- H(\tilde{U}\tilde{V}\tilde{W}\tilde{Y}|\hat{S}^*) - \frac{1}{n}|\mathcal{U}||\mathcal{V}||\mathcal{W}'||\mathcal{Y}|\log(n+1)$$
$$\overset{(n)}{\geq} \min_{\tilde{U}\tilde{V}\tilde{W}\tilde{Y}S\in\mathcal{T}'_1(\kappa_\alpha,\omega',P_S,P_{X|USW'})} D(\tilde{U}\tilde{V}\tilde{W}\tilde{Y}||\bar{U}\bar{V}\bar{W}'\bar{Y}|S) - O(\eta) = E'_1(\kappa_\alpha,\omega') - O(\eta).$$

In (132), we used the fact that

$$\mathbb{P}\left(W'^n(m') = w'^n|\bar{U}^n = u^n, \bar{V}^n = v^n, S^n = s^n, M' = m'\right) \leq \begin{cases} \frac{1}{e^{n\left(H(\tilde{W}|\hat{S}^*\tilde{U})-\eta\right)}}, & \text{if } w'^n \in \mathcal{T}_n(\tilde{W}),\\ 0, & \text{otherwise,}\end{cases}$$

which in turn follows from the fact that given $M' = m'$ and $\bar{U}^n = u^n$, $W'^n(m')$ is uniformly distributed in the set $\mathcal{T}_n(P_{\tilde{W}|\hat{S}^*\tilde{U}},(s^n,u^n))$ and that for sufficiently large n,

$$|\mathcal{T}_n(P_{\tilde{W}|\hat{S}^*\tilde{U}},(s^n,u^n))| \geq e^{n\left(H(\tilde{W}|\hat{S}^*\tilde{U})-\eta\right)}.$$

Next, we analyze the probability of the event $\mathcal{E}'_b$. Let

$$\mathcal{F}'_{2,n}(\eta) := \left\{ \begin{array}{ll} \hat{S}^*\tilde{U}\tilde{V}\tilde{W}\tilde{Y}\tilde{W}_d \in \mathcal{T}_n(\mathcal{S}\times\mathcal{U}\times\mathcal{V}\times\mathcal{W}'\times\mathcal{Y}\times\mathcal{W}') : & \hat{S}^*\tilde{U}\tilde{W}\in\mathcal{D}_n^{SUW'}(\eta), \hat{S}^*\tilde{V}\tilde{W}_d\tilde{Y}\in\mathcal{D}_n^{SVW'Y}(\eta)\\ & \text{and } H\left(\tilde{W}_d|\hat{S}^*\tilde{V}\tilde{Y}\right) \leq H\left(\tilde{W}|\hat{S}^*\tilde{V}\tilde{Y}\right) \end{array}\right\}.$$

Then,

$$\mathbb{P}\left(\mathcal{E}'_b|H=1\right)$$
$$\leq \sum_{\substack{\hat{S}^*\tilde{U}\tilde{V}\tilde{W}\tilde{Y}\tilde{W}_d\\ \in\mathcal{F}'_{2,n}(\eta)}} \sum_{\substack{(u^n,v^n,w'^n,y^n,\bar{w}^n):\\ (s^n,u^n,v^n,w'^n,y^n,\bar{w}^n)\\ \in\mathcal{T}_n(\hat{S}^*\tilde{U}\tilde{V}\tilde{W}\tilde{Y}\tilde{W}_d)}} \sum_{m'\in\mathcal{M}'\backslash\{0\}} \mathbb{P}\left(\bar{U}^n = u^n, \bar{V}^n = v^n, M' = m', W'^n(m') = w'^n, \bar{Y}^n = y^n|S^n = s^n\right)$$
$$\sum_{\hat{m}'\in\mathcal{M}'\backslash\{0,m'\}} \mathbb{P}\left(W'^n(\hat{m}') = \bar{w}^n|\bar{U}^n = u^n, M' = m', W'^n(m') = w'^n, S^n = s^n\right)$$
$$\leq \sum_{\substack{\hat{S}^*\tilde{U}\tilde{V}\tilde{W}\tilde{Y}\tilde{W}_d\\ \in\mathcal{F}'_{2,n}(\eta)}} \sum_{\substack{(u^n,v^n,w'^n,y^n,\bar{w}^n):\\ (s^n,u^n,v^n,w'^n,y^n,\bar{w}^n)\\ \in\mathcal{T}_n(\hat{S}^*\tilde{U}\tilde{V}\tilde{W}\tilde{Y}\tilde{W}_d)}} \sum_{m'\in\mathcal{M}'\backslash\{0\}} \Bigg[e^{-n\left(H(\tilde{U}\tilde{V})+D(\tilde{U}\tilde{V}||\bar{U}\bar{V})\right)}\, \mathbb{P}\left(M' = m'|\bar{U}^n = u^n, \bar{V}^n = v^n, S^n = s^n\right)$$
$$\frac{1}{e^{n\left(H(\tilde{W}|\hat{S}^*\tilde{U})-\eta\right)}} e^{-n\left(H(\tilde{Y}|\tilde{U}\hat{S}^*\tilde{W})+D\left(\tilde{Y}|\tilde{U}\hat{S}^*\tilde{W}||Y|USW'\middle|\tilde{U}\hat{S}^*\tilde{W}\right)\right)}$$
$$\sum_{\hat{m}'\in\mathcal{M}'\backslash\{0,m'\}} \mathbb{P}\left(W'^n(\hat{m}') = \bar{w}^n|\bar{U}^n = u^n, M' = m', W'^n(m') = w'^n, S^n = s^n\right)\Bigg]$$
$$\leq \sum_{\substack{\hat{S}^*\tilde{U}\tilde{V}\tilde{W}\tilde{Y}\tilde{W}_d\\ \in\mathcal{F}'_{2,n}(\eta)}} \sum_{\substack{(u^n,v^n,w'^n,y^n):\\ (s^n,u^n,v^n,w'^n,y^n)\\ \in\mathcal{T}_n(\hat{S}^*\tilde{U}\tilde{V}\tilde{W}\tilde{Y})}} \sum_{m'\in\mathcal{M}'\backslash\{0\}} \Bigg[e^{-n\left(H(\tilde{U}\tilde{V})+D(\tilde{U}\tilde{V}||\bar{U}\bar{V})\right)}\, \mathbb{P}\left(M' = m'|\bar{U}^n = u^n, \bar{V}^n = v^n, S^n = s^n\right)$$
$$\frac{1}{e^{n\left(H(\tilde{W}|\hat{S}^*\tilde{U})-\eta\right)}} e^{-n\left(H(\tilde{Y}|\tilde{U}\hat{S}^*\tilde{W})+D\left(\tilde{Y}|\tilde{U}\hat{S}^*\tilde{W}||Y|USW'\middle|\tilde{U}\hat{S}^*\tilde{W}\right)\right)} \sum_{\hat{m}'\in\mathcal{M}'\backslash\{0,m'\}} \frac{2e^{nH(\tilde{W}_d|\hat{S}^*\tilde{V}\tilde{Y})}}{e^{n\left(H(\tilde{W}_d)-\eta\right)}}\Bigg]$$
$$\leq \sum_{\substack{\hat{S}^*\tilde{U}\tilde{V}\tilde{W}\tilde{Y}\tilde{W}_d\\ \in\mathcal{F}'_{2,n}(\eta)}} \sum_{\substack{(u^n,v^n,w'^n,y^n):\\ (s^n,u^n,v^n,w'^n,y^n)\\ \in\mathcal{T}_n(\hat{S}^*\tilde{U}\tilde{V}\tilde{W}\tilde{Y})}} \Bigg[e^{-n\left(H(\tilde{U}\tilde{V})+D(\tilde{U}\tilde{V}||\bar{U}\bar{V})\right)} \frac{1}{e^{n\left(H(\tilde{W}|\hat{S}^*\tilde{U})-\eta\right)}} e^{-n\left(H(\tilde{Y}|\tilde{U}\hat{S}^*\tilde{W})+D\left(\tilde{Y}|\tilde{U}\hat{S}^*\tilde{W}||Y|USW'\middle|\tilde{U}\hat{S}^*\tilde{W}\right)\right)}$$

$$e^{n\left(\zeta_q'(\kappa_\alpha,\omega',P_{\hat{S}^*})+\eta\right)} \frac{2e^{nH(\tilde{W}_d|\hat{S}^*\tilde{V}\tilde{Y})}}{e^{n\left(H(\tilde{W}_d)-\eta\right)}}\Bigg]$$

$$\leq e^{-nE_{2,n}'}, \tag{134}$$

where

$$\begin{aligned}
E_{2,n}' &:= \min_{\hat{S}^*\tilde{U}\tilde{V}\tilde{W}\tilde{Y}\tilde{W}_d\in\mathcal{F}_{2,n}'(\eta)} H(\tilde{U}\tilde{V}) + D(\tilde{U}\tilde{V}||\bar{U}\bar{V}) + H(\tilde{W}|\hat{S}^*\tilde{U}) - 2\eta + H(\tilde{Y}|\tilde{U}\hat{S}^*\tilde{W}) + D\left(\tilde{Y}|\tilde{U}\hat{S}^*\tilde{W}||Y|USW'\middle|\tilde{U}\hat{S}^*\tilde{W}\right) \\
&\quad + I(\tilde{W}_d;\hat{S}^*\tilde{V}\tilde{Y}) - H(\tilde{U}\tilde{V}\tilde{W}\tilde{Y}|\hat{S}^*) - \zeta_q'(\kappa_\alpha,\omega',P_{\hat{S}^*}) - \frac{\log 2}{n} - \frac{|\mathcal{S}||\mathcal{U}||\mathcal{V}||\mathcal{W}|^2|\mathcal{Y}|\log(n+1)}{n} \\
&\geq \min_{\hat{S}^*\tilde{U}\tilde{V}\tilde{W}\tilde{Y}\tilde{W}_d\in\mathcal{F}_{2,n}'(\eta)} D(\tilde{U}\tilde{V}\tilde{W}\tilde{Y}||\bar{U}\bar{V}\bar{W}'\bar{Y}|S) + I(\tilde{W}_d;\hat{S}^*\tilde{V}\tilde{Y}) - \zeta_q'(\kappa_\alpha,\omega',P_{\hat{S}^*}) - \frac{\log 2}{n} \\
&\qquad - \frac{|\mathcal{S}||\mathcal{U}||\mathcal{V}||\mathcal{W}|^2|\mathcal{Y}|\log(n+1)}{n} \\
&\overset{(n)}{\geq} \min_{\substack{\tilde{U}\tilde{V}\tilde{W}\tilde{Y}S\ \in \\ \mathcal{T}_2'(\kappa_\alpha,\omega',P_S,P_{X|USW'})}} D(\tilde{U}\tilde{V}\tilde{W}\tilde{Y}||\bar{U}\bar{V}\bar{W}'\bar{Y}|S) + \rho'(\kappa_\alpha,\omega',P_S,P_{X|USW'}) - \zeta_q'(\kappa_\alpha,\omega',P_S) - O(\eta) \\
&= E_2'(\kappa_\alpha,\omega',P_S,P_{X|USW'}) - O(\eta).
\end{aligned}$$

Similar to (103), we can write

$$\mathbb{P}\left(\mathcal{E}_c'|H=1\right) = \sum_{\substack{u^n\in\mathcal{T}_n(\tilde{U}):\\ \tilde{U}\in\mathcal{D}_n^U(\eta)}} \mathbb{P}\left(\bar{U}^n=u^n,\mathcal{E}_{EE}',\mathcal{E}_c'|S^n=s^n,H=1\right) + \sum_{\substack{u^n\in\mathcal{T}_n(\tilde{U}):\\ \tilde{U}\notin\mathcal{D}_n^U(\eta)}} \mathbb{P}\left(\bar{U}^n=u^n,\mathcal{E}_c'|S^n=s^n,H=1\right). \tag{135}$$

The first term in (135) decays double exponentially as $e^{-e^{n\Omega(\eta)}}$. The second term in (135) can be simplified as follows:

$$\begin{aligned}
&\sum_{\substack{u^n\in\mathcal{T}_n(\tilde{U}):\\ \tilde{U}\notin\mathcal{D}_n^U(\eta)}} \mathbb{P}\left(\bar{U}^n=u^n,\mathcal{E}_c'|S^n=s^n,H=1\right) \\
&\leq \sum_{\substack{u^n\in\mathcal{T}_n(\tilde{U}):\\ \tilde{U}\notin\mathcal{D}_n^U(\eta)}} \sum_{\substack{(v^n,y^n,\bar{w}^n):\\ (s^n,v^n,y^n,\bar{w}^n)\in\mathcal{T}_n(\hat{S}^*\tilde{V}\tilde{Y}\tilde{W}_d)\\ \hat{S}^*\tilde{V}\tilde{W}_d\tilde{Y}\in\mathcal{D}_n^{SVW'Y}(\eta)}} \sum_{\hat{m}\in\mathcal{M}\setminus\{0\}} \mathbb{P}\left(\bar{U}^n=u^n,\bar{V}^n=v^n,M'=0,\bar{Y}^n=y^n|S^n=s^n,H=1\right) \\
&\qquad\qquad \sum_{\hat{m}'\in\mathcal{M}'\setminus\{0\}} \mathbb{P}\left(W^k(\hat{m}')=\bar{w}^k\right) \\
&\leq \sum_{\substack{u^n\in\mathcal{T}_n(\tilde{U}):\\ \tilde{U}\notin\mathcal{D}_n^U(\eta)}} \sum_{\substack{(v^n,y^n,\bar{w}^n):\\ (s^n,v^n,y^n,\bar{w}^n)\in\mathcal{T}_n(\hat{S}^*\tilde{V}\tilde{Y}\tilde{W}_d)\\ \hat{S}^*\tilde{V}\tilde{W}_d\tilde{Y}\in\mathcal{D}_n^{SVW'Y}(\eta)}} \mathbb{P}\left(\bar{U}^n=u^n,\bar{V}^n=v^n,M'=0,\bar{Y}^n=y^n|S^n=s^n,H=1\right) \\
&\qquad\qquad \sum_{\hat{m}'\in\mathcal{M}'\setminus\{0\}} \frac{1}{e^{k\left(H(\tilde{W}_d)-\eta\right)}} \\
&\leq \sum_{\substack{u^n\in\mathcal{T}_n(\tilde{U}):\\ \tilde{U}\notin\mathcal{D}_n^U(\eta)}} \sum_{\substack{(v^n,y^n):\\ (s^n,v^n,y^n)\in\mathcal{T}_n(\hat{S}^*\tilde{V}\tilde{Y})\\ \hat{S}^*\tilde{V}\tilde{W}_d\tilde{Y}\in\mathcal{D}_n^{SVW'Y}(\eta)}} \mathbb{P}\left(\bar{U}^n=u^n,\bar{V}^n=v^n\right)\mathbb{P}\left(\bar{Y}^n=y^n|\bar{U}^n=u^n,\bar{V}^n=v^n,M'=0,S^n=s^n,H=1\right) \\
&\qquad\qquad \frac{e^{nH(\tilde{W}_d|\hat{S}^*\tilde{V}\tilde{Y})}e^{n(R'+\eta)}}{e^{n\left(H(\tilde{W}_d)-\eta\right)}} \\
&\leq \sum_{\substack{\tilde{U}\hat{S}^*\tilde{V}\tilde{W}_d\tilde{Y}\\ \in\mathcal{D}_n^U(\eta)^c\times\mathcal{D}_n^{SVW'Y}(\eta)}} e^{nH(\tilde{U}\tilde{V}\tilde{Y}|\hat{S}^*)}e^{-n\left(H(\tilde{U}\tilde{V}\tilde{Y}|\hat{S}^*)+D\left(\tilde{U}\tilde{V}\tilde{Y}||\bar{U}\bar{V}\bar{Y}\middle|\hat{S}^*\right)\right)} \frac{e^{nH(\tilde{W}_d|\hat{S}^*\tilde{V}\tilde{Y})}e^{n(R'+\eta)}}{e^{n\left(H(\tilde{W}_d)-\eta\right)}} \\
&\leq e^{-nE_{3,n}'},
\end{aligned} \tag{136}$$

where,

$$\begin{aligned}E'_{3,n} &:= \min_{\substack{\tilde{U}\hat{S}^*\tilde{V}\tilde{W}_d\tilde{Y} \\ \in \mathcal{D}_n^U(\eta)^c \times \mathcal{D}_n^{SVW'Y}(\eta)}} D\left(\tilde{U}\tilde{V}\tilde{Y}||\bar{U}\bar{V}\check{Y}\middle|\hat{S}^*\right) + I(\tilde{W}_d;\hat{S}^*,\tilde{V},\tilde{Y}) - R' - O(\eta) - \frac{|\mathcal{U}||\mathcal{V}||\mathcal{W}||\mathcal{Y}||\mathcal{S}|}{n}\log(n+1) \\ &\overset{(n)}{\geq} \min_{\substack{\hat{V}\hat{Y}S:\hat{U}\hat{V}\hat{W}\hat{Y}S \in \\ \hat{\mathcal{L}}_h(\kappa_\alpha,\omega',P_S,P_{X|USW'})}} D\left(\hat{V}\hat{Y}||\bar{V}\check{Y}|S\right) + \rho'(\kappa_\alpha,\omega',P_S,P_{X|USW'}) - \zeta'_q(\kappa_\alpha,\omega',P_S) - O(\eta) \\ &= E'_3(\kappa_\alpha,\omega',P_S,P_{X|USW'},P_{X'|S}) - O(\eta).\end{aligned}$$

Since the exponent of the type II error probability is lower bounded by the minimum of the exponent of the type II error causing events, it follows from (133), (134) and (136) that for a fixed $\left(P_S,\omega'(\cdot,P_S),P_{X|USW'},P_{X'|S}\right) \in \mathcal{L}_h(\kappa_\alpha)$,

$$\kappa(\tau,\kappa_\alpha) \geq \min\left(E'_1(\kappa_\alpha,\omega'),E'_2(\kappa_\alpha,\omega',P_S,P_{X|USW'}),E'_3(\kappa_\alpha,\omega',P_S,P_{X|USW'},P_{X'|S})\right) - O(\eta).$$

Maximizing over $\left(P_S,\omega'(\cdot,P_S),P_{X|USW'},P_{X'|S}\right) \in \mathcal{L}_h(\kappa_\alpha)$ and noting that $\eta > 0$ is arbitrary completes the proof.

REFERENCES

[1] R. Ahlswede and I. Csiszár, "Hypothesis testing with communication constraints," *IEEE Trans. Inf. Theory*, vol. 32, no. 4, pp. 533–542, Jul. 1986.
[2] T. S. Han, "Hypothesis testing with multiterminal data compression," *IEEE Trans. Inf. Theory*, vol. 33, no. 6, pp. 759–772, Nov. 1987.
[3] H. M. H. Shalaby and A. Papamarcou, "Multiterminal detection with zero-rate data compression," *IEEE Trans. Inf. Theory*, vol. 38, no. 2, pp. 254–267, Mar. 1992.
[4] T. S. Han and K. Kobayashi, "Exponential-type error probabilities for multiterminal hypothesis testing," *IEEE Trans. Inf. Theory*, vol. 35, no. 1, pp. 2–14, Jan. 1989.
[5] J. Neyman and E. Pearson, "On the problem of the most efficient tests of statistical hypotheses," *Philos. Trans. of the Royal Society of London*, vol. 231, pp. 289–337, Feb. 1933.
[6] H. Chernoff, "A measure of asymptotic efficiency for tests of a hypothesis based on a sum of observations," *Ann. Math. Statist.*, vol. 23, no. 4, pp. 493–507, 1952.
[7] W. Hoeffding, "Asymptotically optimal tests for multinominal distributions," *Ann. Math. Stat.*, vol. 36, no. 2, pp. 369–400, 1965.
[8] R. E. Blahut, "Hypothesis testing and information theory," *IEEE Trans. Inf. Theory*, vol. 20, no. 4, pp. 405–417, Jul. 1974.
[9] H. Shimokawa, T. S. Han, and S. Amari, "Error bound of hypothesis testing with data compression," in *Proc. IEEE Int. Symp. Inf. Theory*, Trondheim, Norway, 1994.
[10] 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.
[11] W. Zhao and L. Lai, "Distributed testing against independence with multiple terminals," in *52nd Annual Allerton Conference on Communication, Control and Computing*, Monticello (IL), USA, Oct. 2014.
[12] M. Wigger and R. Timo, "Testing against independence with multiple decision centers," in *Int. Conf. on Signal Processing and Communication*, Bengaluru, India, Jun. 2016.
[13] S. Salehkalaibar, M. Wigger, and L. Wang, "Hypothesis testing in multi-hop networks," *arXiv:1708.05198*.
[14] M. Mhanna and P. Piantanida, "On secure distributed hypothesis testing," in *Proc. IEEE Int. Symp. Inf. Theory*, Hong Kong, China, Jun. 2015.
[15] S. Sreekumar and D. Gúndúz, "Testing against conditional independence under security constraints," in *Proc. IEEE Int. Symp. Inf. Theory*, Vail, USA, Jun. 2018.
[16] S. Sreekumar, A. Cohen, and D. Gündüz, "Distributed hypothesis testing with a privacy constraint," *arXiv:1807.02764*.
[17] A. Gilani, S. B. Amor, S. Salehkalaibar, and V. Y. F. Tan, "Distributed hypothesis testing with privacy constraints," *arXiv:1806.02015*.
[18] G. Katz, P. Piantanida, and M. Debbah, "Distributed binary detection with lossy data compression," *IEEE Trans. Inf. Theory*, vol. 63, no. 8, pp. 5207–5227, Mar. 2017.
[19] 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, Apr. 2019.
[20] S. Watanabe, "Neyman-Pearson test for zero-rate multiterminal hypothesis testing," *IEEE Trans. Inf. Theory*, vol. 64, no. 7, pp. 4923–4939, Jul. 2018.
[21] S. Sreekumar and D. Gündüz, "Distributed hypothesis testing over noisy channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Aachen, Germany, Jun. 2017.
[22] ——, "Distributed hypothesis testing over discrete memoryless channels," *arXiv:1802.07665 [cs.IT]*.
[23] S. Salehkalaibar and M. Wigger, "Distributed hypothesis testing based on unequal-error protection codes," *arXiv:1806.05533*.
[24] S. Sreekumar and D. Gündüz, "Hypothesis testing over a noisy channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Paris, France, 2019.
[25] S. Borade, B. Nakiboğlu, and L. Zheng, "Unequal error protection: An information-theoretic perspective," *IEEE Trans. Inf. Theory*, vol. 55, no. 12, pp. 5511–5539, Dec 2009.
[26] N. Weinberger, Y. Kochman, and M. Wigger, "Exponent trade-off for hypothesis testing over noisy channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Paris, France, 2019.
[27] I. Csiszár, "On the error exponent of source-channel transmission with a distortion threshold," *IEEE Trans. Inf. Theory*, vol. 28, no. 6, pp. 823–828, Nov. 1982.
[28] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Cambridge University Press, 2011.
[29] Y. Polyanskiy and Y. Wu, *Lecture Notes on Information Theory*. http://people.lids.mit.edu/yp/homepage/data/itlectures_v5.pdf [Available Online], 2017.
[30] R. Gallager, "A simple derivation of the coding theorem and some applications," *IEEE Trans. Inf. Theory*, vol. 11, no. 1, pp. 3–18, Jan. 1965.
[31] I. Csiszár, "Joint source-channel error exponent," *Prob. of Control and Inf. Theory*, vol. 9, no. 5, pp. 315–328, Oct. 1980.
[32] T. Cover, A. E. Gamal, and M. Salehi, "Multiple access channels with arbitrarily correlated sources," *IEEE Trans. Inf. Theory*, vol. 26, no. 6, pp. 648–657, Nov. 1980.
[33] 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.