Distributed Hypothesis Testing Over Multi-Access Channels

[1]Sadaf Salehkalaibar and [2]Michèle Wigger

[1]ECE Department, College of Engineering, University of Tehran, Tehran, Iran, s.saleh@ut.ac.ir

[2]LTCI, Telecom ParisTech, 75013 Paris, France, michele.wigger@telecom-paristech.fr

Abstract—Consider distributed hypothesis testing over multiple-access channels (MACs), where the receiver wishes to maximize the type-II error exponent under a constrained type-I error probability. For this setup, we propose a scheme that combines hybrid coding with a MAC-version of Borades unequal error protection. It achieves the optimal type-II error exponent for a generalization of testing against independence over an orthogonal MAC when the transmitters' sources are independent. In this case, hybrid coding can be replaced by the simpler separate source-channel coding. The paper also presents upper and lower bounds on the optimal type-II error exponent for generalized testing against independence of Gaussian sources over a Gaussian MAC. The bounds are close and significantly larger than a type-II error exponent that is achievable using separate source-channel coding.

I. INTRODUCTION

Consider the distributed hypothesis testing problem in Fig. 1. Both transmitters and the receiver observe individual source sequences and the transmitters communicate to the receiver over a discrete memoryless multiple-access channel (MAC). The joint distribution underlying the three sources depends on the hypothesis $\mathcal{H} = 0$ or $\mathcal{H} = 1$. The goal of the communication is that the receiver can decide on whether $\mathcal{H} = 0$ or $\mathcal{H} = 1$. In this paper, the focus is on maximizing the exponential decay of the probability of type-II error (declaring $\mathcal{H} = 0$ when $\mathcal{H} = 1$) under a constrained probability of type-I error (declaring $\mathcal{H} = 1$ when $\mathcal{H} = 0$).

Han and Amari [1] studied this problem when the MAC consists of two individual noise-free bit-pipes from each of the transmitters to the receiver. Rahman and Wagner [2] improved the scheme in [1] for the special case known as *testing against conditional independence*. This latter result has been extended to *orthogonal noisy MACs* where each of the two transmitters communicates to the receiver over an individual discrete memoryless channel (DMC) [3].

The authors studied the general hypothesis testing problem over a noisy point-to-point channel, i.e., a setup with only a single transmitter and general joint distributions under the two hypotheses [4]. That work presents an achievable error exponent based on a scheme that combines Shimokawa-Han-Amari's hypothesis testing scheme [5] with Borade's unequal error protection (UEP) channel coding [6] by means of a separation-based architecture. The resulting achievable error exponent consists of three competing exponents where two coincide with the two Shimokawa-Han-Amari exponents and the third includes Borade's miss-detection exponent.
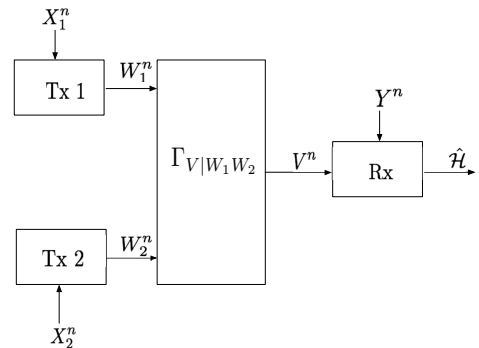

Fig. 1. Hypothesis testing over a noisy MAC

In this paper, we present a coding and testing scheme for general distributed hypothesis testing over a MAC. Our scheme combines hybrid coding [7] with a MAC-extension of Borade's unequal error protection idea [6]. The resulting achievable error exponent consists of nine competing exponents: one of them is the exponent in [1], three are associated with an incorrect decoding of the hybrid coding scheme, another three exponents are associated with the UEP mechanism, and two exponents with incorrect decoding *and* UEP.

The proposed coding scheme is optimal when the transmitters' sources are independent, communication is over an orthogonal MAC, and hypothesis testing is *generalized testing against independence*. In generalized testing against independence, the transmitters' sources have same joint distribution under both hypotheses and under $\mathcal{H} = 1$ they are independent of the receiver's source.[1] For this scenario, separate source-channel coding suffices. This statement is shown to hold more generally even when the transmitters' sources are not independent.

The paper also specializes the coding scheme to generalized testing against independence of jointly Gaussian sources over a Gaussian MAC. Numerical simulations show that the error exponent achieved by our hybrid coding scheme is close to a new upper bound on the optimal exponent. It significantly improves over a similar scheme based on separate source-channel coding.

---

[1]The test is thus a relaxed version of *testing against independence* where the receiver's source distribution is not required to be the same under both hypotheses.

Notation: We mostly follow the notation in [8]. Moreover, we use tp($\cdot$) to denote the *joint type* of a tuple. For a joint type π_{AB} over alphabets $\mathcal{A} \times \mathcal{B}$, we denote by $I_{\pi_{AB}}(A; B)$ the mutual information of a pair of random variables (A, B) with probability mass function (pmf) π_{AB} and by $H_{\pi_A}(A)$ and $H_{\pi_{AB}}(A|B)$ the entropy and conditional entropy of A given B. When it is unambiguous, we may abbreviate π_{AB} by π. The notation $|P_X - Q_X|$ denotes the total variation distance between two distributions P_X and Q_X on the random variable X. We also abbreviate *independent and identically distributed* by i.i.d. Logarithms are taken with respect to base 2.

II. SYSTEM MODEL

Consider a setup with two sensors that communicate to a single decision center over a discrete memoryless multiple-access channel (MAC), see Fig. 1. The channel is described by the quadruple $(\mathcal{W}_1 \times \mathcal{W}_2, \mathcal{V}, \Gamma_{V|W_1,W_2})$, where $\mathcal{W}_1$ and $\mathcal{W}_2$ denote the finite channel input alphabets, and $\mathcal{V}$ denotes the finite channel output alphabet. Each transmitter i ($i = 1, 2$) observes the sequence X_i^n and produces channel inputs $W_i^n := (W_{i,1}, \ldots, W_{i,n})$ as $W_i^n = f_i^{(n)}(X_i^n)$ by means of a possibly stochastic encoding function $f_i^{(n)} : \mathcal{X}_i^n \to \mathcal{W}^n$. The receiver observes the corresponding channel outputs $V^n := (V_1, \ldots, V_n)$ as well as the source sequence $Y^n := (Y_1, \ldots, Y_n)$. Under the null hypothesis

$$\mathcal{H} = 0: \quad (X_1^n, X_2^n, Y^n) \text{ i.i.d.} \sim P_{X_1 X_2 Y}, \quad (1)$$

and under the alternative hypothesis

$$\mathcal{H} = 1: \quad (X_1^n, X_2^n, Y^n) \text{ i.i.d.} \sim Q_{X_1 X_2 Y}, \quad (2)$$

for two given pmfs $P_{X_1 X_2 Y}$ and $Q_{X_1 X_2 Y}$. The receiver should decide on the hypothesis $\mathcal{H}$. It thus produces the guess $\hat{\mathcal{H}} = g^{(n)}(V^n, Y^n)$ using a decoding function $\mathcal{V}^n \times \mathcal{Y}^n \to \{0, 1\}$.

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

$$\alpha_n \triangleq \Pr[\hat{\mathcal{H}} = 1 | \mathcal{H} = 0], \quad (3)$$
$$\beta_n \triangleq \Pr[\hat{\mathcal{H}} = 0 | \mathcal{H} = 1], \quad (4)$$

satisfy

$$\alpha_n \leq \epsilon, \quad (5)$$

and

$$-\varlimsup_{n \to \infty} \frac{1}{n} \log \beta_n \geq \theta. \quad (6)$$

The goal is to characterize the set of achievable type-II error exponents θ.

III. CODING AND TESTING SCHEME

We describe a coding and testing scheme for distributed hypothesis testing over a noisy MAC.

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

$$f_i : \mathcal{S}_i \times \mathcal{X}_i \to \mathcal{W}_i, \qquad i \in \{1, 2\}. \quad (7)$$

Define for each tuple $(v, s_1, s_2, x_1, x_2) \in \mathcal{V} \times \mathcal{S}_1 \times \mathcal{S}_2 \times \mathcal{X}_1 \times \mathcal{X}_2$ the shorthand notation

$$\Gamma_{V|S_1 S_2 X_1 X_2}(v|s_1, s_2, x_1, x_2)$$
$$:= \Gamma_{V|W_1, W_2}(v|f_1(s_1, x_1), f_2(s_2, x_2)). \quad (8)$$

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

$$I(S_1; X_1 | T_1, T_2) < I(S_1; S_2, Y, V | T_1, T_2), \quad (9a)$$
$$I(S_2; X_2 | T_1, T_2) < I(S_2; S_1, Y, V | T_1, T_2), \quad (9b)$$
$$I(S_1, S_2; X_1, X_2 | T_1, T_2) < I(S_1, S_2; Y, V | T_1, T_2) \quad (9c)$$

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

$$P_{S_1 S_2 X_1 X_2 Y T_1 T_2 V} = P_{T_1 T_2} \cdot P_{X_1 X_2 Y} \cdot P_{S_1 | X_1 T_1 T_2}$$
$$\cdot P_{S_2 | X_2 T_1 T_2} \cdot \Gamma_{V | S_1 S_2 X_1 X_2}. \quad (10)$$

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

$$R_i = I(S_i; X_i | T_1, T_2) + \mu, \qquad i \in \{1, 2\}, \quad (11)$$

so that the following three conditions hold:

$$R_1 < I(S_1; S_2, Y, V | T_1, T_2), \quad (12a)$$
$$R_2 < I(S_2; S_1, Y, V | T_2, T_2), \quad (12b)$$
$$R_1 + R_2 < I(S_1, S_2; Y, V | T_1, T_2) + I(S_1; S_2 | T_1, T_2). \quad (12c)$$

Code Construction: Generate a pair of sequences $T_1^n = (T_{1,1}, \ldots, T_{1,n})$ and $T_2^n = (T_{2,1}, \ldots, T_{2,n})$ by independently drawing each pair $(T_{1,k}, T_{2,k})$ according to $P_{T_1 T_2}(.,.)$. For $i \in \{1, 2\}$, construct a random codebook

$$\mathcal{C}_{S_i} = \{S_i^n(m_i) : m_i \in \{1, \ldots, \lfloor 2^{nR_i} \rfloor\}\}, \quad (13)$$

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

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

$$(s_i^n(m_i), x_i^n, t_1^n, t_2^n) \in \mathcal{T}_{\mu/2}^n(P_{S_i X_i T_1 T_2}). \quad (14)$$

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

$$w_{i,k} = f_i(s_{i,k}(m_i), x_{i,k}), \quad 1 \leq k \leq n,$$

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

<u>*Receiver:*</u> Assume that the receiver observes the sequences $V^n = v^n$ and $Y^n = y^n$. It first searches for a pair of indices (m'_1, m'_2) that satisfies the condition:

$$H_{\mathrm{tp}(s_1^n(m'_1), s_2^n(m'_2), y^n, l_1^n, l_2^n, v^n)}(S_1, S_2|Y, T_1, T_2, V)$$
$$= \min_{\tilde{m}_1, \tilde{m}_2} H_{\mathrm{tp}(s_1^n(\tilde{m}_1), s_2^n(\tilde{m}_2), y^n, l_1^n, l_2^n, v^n)}(S_1, S_2|Y, T_1, T_2, V).$$
$$\tag{15}$$

It picks one such pair at random and checks whether

$$(s_1^n(m'_1), s_2^n(m'_2), y^n, t_1^n, t_2^n, v^n) \in \mathcal{T}_\mu^n(P_{S_1 S_2 Y T_1 T_2 V}). \tag{16}$$

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

IV. RESULTS ON THE ERROR EXPONENT

The coding and testing scheme described in the previous section yields Theorem 1 ahead. For given (conditional) pmfs $P_{T_1 T_2}$, $P_{S_1|X_1 T_1 T_2}$, and $P_{S_2|X_2 T_1 T_2}$, and functions f_1 and f_2 as in (7), let the conditional and joint pmfs $\Gamma_{V|S_1 S_2 X_1 X_2}$ and $P_{S_1 S_2 X_1 X_2 Y W_1 W_2 V T_1 T_2}$ be as in (8) and (10). Define also

$$\Gamma^{(1)}_{V|T_1 S_2 X_2}(v|t_1, s_2, x_2) := \Gamma_{V|W_1 W_2}(v|t_1, f_2(s_2, x_2)), \tag{17}$$
$$\Gamma^{(2)}_{V|S_1 X_1 T_2}(v|s_1, x_1, t_2) := \Gamma_{V|W_1 W_2}(v|f_1(s_1, x_1), t_2), \tag{18}$$
$$\Gamma^{(12)}_{V|T_1 T_2}(v|t_1, t_2) := \Gamma_{V|W_1 W_2}(v|t_1, t_2), \tag{19}$$

and the following nine exponents:

$$\theta^{\mathrm{standard}} := \min_{\substack{\tilde{P}_{S_1 S_2 X_1 X_2 Y T_1 T_2 V}: \\ \tilde{P}_{S_i X_i T_1 T_2} = P_{S_i X_i T_1 T_2},\, i \in \{1,2\}, \\ \tilde{P}_{S_1 S_2 Y T_1 T_2 V} = P_{S_1 S_2 Y T_1 T_2 V}}} D\Big(\tilde{P}_{S_1 S_2 X_1 X_2 Y T_1 T_2 V} \| P_{S_1|X_1 T_1 T_2} P_{S_2|X_2 T_1 T_2}$$
$$\cdot Q_{X_1 X_2 Y} P_{T_1 T_2} \Gamma_{V|S_1 S_2 X_1 X_2}\Big), \tag{20}$$

$$\theta^{\mathrm{dec},1} := \min_{\substack{\tilde{P}_{S_1 S_2 X_1 X_2 Y T_1 T_2 V}: \\ \tilde{P}_{S_i X_i T_1 T_2} = P_{S_i X_i T_1 T_2},\, i \in \{1,2\}, \\ \tilde{P}_{S_2 Y T_1 T_2 V} = P_{S_2 Y T_1 T_2 V} \\ H(S_1|S_2, Y, T_1, T_2, V) \leq H_{\tilde{P}}(S_1|S_2, Y, T_1, T_2, V)}} D\Big(\tilde{P}_{S_1 S_2 X_1 X_2 Y T_1 T_2 V} \| P_{S_1|X_1 T_1 T_2} P_{S_2|X_2 T_1 T_2}$$
$$\cdot Q_{X_1 X_2 Y} P_{T_1 T_2} \Gamma_{V|S_1 S_2 X_1 X_2}\Big)$$
$$+ I(S_1; Y, V|S_2, T_1, T_2) - I(S_1; X_1|S_2, T_1, T_2), \tag{21}$$

$$\theta^{\mathrm{dec},2} := \min_{\substack{\tilde{P}_{S_1 S_2 X_1 X_2 Y T_1 T_2 V}: \\ \tilde{P}_{S_i X_i T_1 T_2} = P_{S_i X_i T_1 T_2},\, i \in \{1,2\}, \\ \tilde{P}_{S_1 Y T_1 T_2 V} = P_{S_1 Y T_1 T_2 V} \\ H(S_2|S_1, Y, T_1, T_2, V) \leq H_{\tilde{P}}(S_2|S_1, Y, T_1, T_2, V)}} D\Big(\tilde{P}_{S_1 S_2 X_1 X_2 Y T_1 T_2 V} \| P_{S_1|X_1 T_1 T_2} P_{S_2|X_2 T_1 T_2}$$
$$\cdot Q_{X_1 X_2 Y} P_{T_1 T_2} \Gamma_{V|S_1 S_2 X_1 X_2}\Big)$$
$$+ I(S_2; Y, V|S_1, T_1, T_2) - I(S_2; X_2|S_1, T_1, T_2), \tag{22}$$

$$\theta^{\mathrm{dec},12} := \min_{\substack{\tilde{P}_{S_1 S_2 X_1 X_2 Y T_1 T_2 V}: \\ \tilde{P}_{S_i X_i T_1 T_2} = P_{S_i X_i T_1 T_2},\, i \in \{1,2\}, \\ \tilde{P}_{Y T_1 T_2 V} = P_{Y T_1 T_2 V} \\ H(S_1, S_2|Y, T_1, T_2, V) \leq H_{\tilde{P}}(S_1, S_2|Y, T_1, T_2, V)}} D\Big(\tilde{P}_{S_1 S_2 X_1 X_2 Y T_1 T_2 V} \| P_{S_1|X_1 T_1 T_2} P_{S_2|X_2 T_1 T_2}$$

$$\cdot Q_{X_1 X_2 Y} P_{T_1 T_2} \Gamma_{V|S_1 S_2 X_1 X_2}\Big)$$
$$+ I(S_1, S_2; Y, V|T_1, T_2) - I(S_1, S_2; X_1, X_2|T_1, T_2), \tag{23}$$

$$\theta^{\mathrm{miss},1a} := \min_{\substack{\tilde{P}_{S_2 X_2 Y T_1 T_2 V}: \\ \tilde{P}_{S_2 X_2 T_1 T_2} = P_{S_2 X_2 T_1 T_2} \\ \tilde{P}_{Y T_1 T_2 V} = P_{Y T_1 T_2 V} \\ H(S_2|Y, T_1, T_2, V) \leq H_{\tilde{P}}(S_2|Y, T_1, T_2, V)}} D\Big(\tilde{P}_{S_2 X_2 Y T_1 T_2 V} \| P_{S_2|X_2 T_1 T_2} Q_{X_2 Y} P_{T_1 T_2} \Gamma^{(1)}_{V|T_1 S_2 X_2}\Big)$$
$$+ I(S_1, S_2; V, Y|T_1, T_2) - I(S_1, S_2; X_1, X_2|T_1, T_2), \tag{24}$$

$$\theta^{\mathrm{miss},1b} := \min_{\substack{\tilde{P}_{S_1 S_2 X_2 Y T_1 T_2 V}: \\ \tilde{P}_{S_2 X_2 T_1 T_2} = P_{S_2 X_2 T_1 T_2} \\ \tilde{P}_{S_2 Y T_1 T_2 V} = P_{S_2 Y T_1 T_2 V}}} D\Big(\tilde{P}_{S_2 X_2 Y T_1 T_2 V} \| P_{S_2|X_2 T_1 T_2} Q_{X_2 Y} P_{T_1 T_2} \Gamma^{(1)}_{V|T_1 S_2 X_2}\Big)$$
$$+ I(S_1; V, Y|S_2, T_1, T_2) - I(S_1; X_1|S_2, T_1, T_2), \tag{25}$$

$$\theta^{\mathrm{miss},2a} := \min_{\substack{\tilde{P}_{S_1 X_1 Y T_1 T_2 V}: \\ \tilde{P}_{S_1 X_1 T_1 T_2} = P_{S_1 X_1 T_1 T_2} \\ \tilde{P}_{Y T_1 T_2 V} = P_{Y T_1 T_2 V} \\ H(S_1|Y, T_1, T_2, V) \leq H_{\tilde{P}}(S_1|V, Y, T_1, T_2)}} D\Big(\tilde{P}_{S_1 X_1 Y V T_1 T_2} \| P_{S_1|X_1 T_1 T_2} Q_{X_1 Y} P_{T_1 T_2} \Gamma^{(2)}_{V|S_1 X_1 T_2}\Big)$$
$$+ I(S_1, S_2; V, Y|T_1, T_2) - I(S_1, S_2; X_1, X_2|T_1, T_2), \tag{26}$$

$$\theta^{\mathrm{miss},2b} := \min_{\substack{\tilde{P}_{S_1 X_1 Y T_1 T_2 V}: \\ \tilde{P}_{S_1 X_1 T_1 T_2} = P_{S_1 X_1 T_1 T_2} \\ \tilde{P}_{S_1 Y T_1 T_2 V} = P_{S_1 Y T_1 T_2 V}}} D\Big(\tilde{P}_{S_1 X_1 Y T_1 T_2 V} \| P_{S_1|X_1 T_1 T_2} Q_{X_1 Y} P_{T_1 T_2} \Gamma^{(2)}_{V|S_1 X_1 T_2}\Big)$$
$$+ I(S_2; V, Y|S_1, T_1, T_2) - I(S_2; X_2|S_1, T_1, T_2), \tag{27}$$

$$\theta^{\mathrm{miss},12} := \mathbb{E}_{P_{T_1 T_2}}\big[D(P_{YV|T_1 T_2} \| Q_Y \Gamma^{(12)}_{V|T_1 T_2})\big]$$
$$+ I(S_1, S_2; Y, V|T_1, T_2) - I(S_1, S_2; X_1, X_2|T_1, T_2), \tag{28}$$

where mutual informations and the conditional pmf $P_{VY|T_1 T_2}$ are calculated according to the joint pmf $P_{S_1 S_2 X_1 X_2 Y V T_1 T_2}$ in (10).

Theorem 1: Error exponent $\theta \geq 0$ is achievable, if it satisfies

$$\theta \leq \max\ \min\big\{\theta^{\mathrm{standard}}, \theta^{\mathrm{dec},1}, \theta^{\mathrm{dec},2}, \theta^{\mathrm{dec},12},$$
$$\theta^{\mathrm{miss},1a}, \theta^{\mathrm{miss},1b}, \theta^{\mathrm{miss},2a}, \theta^{\mathrm{miss},2b}, \theta^{\mathrm{miss},12}\big\}, \tag{29}$$

where the maximization is over all (conditional) pmfs $P_{T_1 T_2}$, $P_{S_1|X_1 T_1 T_2}$, and $P_{S_2|X_2 T_1 T_2}$, and functions f_1 and f_2 as in (7) so that the conditions in (9) are satisfied with strict inequalities "$<$" replaced with non-strict inequalities "$\leq$".

Proof: See [10]. ∎

Remark 1: The error exponents in the preceding theorem are obtained by means of the hybrid coding scheme described in the previous section III. As usual, choosing the auxiliary random variables $S_1 = (W_1, \bar{S}_1)$ and $S_2 = (W_2, \bar{S}_2)$ with the tuple (S_1, S_2, T_1, T_2) independent of the tuple $(\bar{S}_1, \bar{S}_2, X_1, X_2)$,

is equivalent to replacing the hybrid coding scheme by a separate source-channel coding scheme. Specifically, $\bar{S}_1, \bar{S}_2$ then correspond to the source coding random variables and (T_1, T_2, W_1, W_2) to the channel coding random variables.

Similarly, choosing the auxiliary random variables S_1 and S_2 constant and $W_1 = f_1(X_1)$ and $W_2 = f_2(X_2)$, corresponds to uncoded transmission.

It can be shown that in the MAC setup, the missed-detection exponents are sometimes not active. This is in particular the case for the following case of *generalized testing against conditional independence.*

Corollary 1: Consider the *generalized testing against conditional independence scenario* where $Y^n = (\bar{Y}^n, Z^n)$ and

$$\mathcal{H} = 0: \quad (X_1^n, X_2^n, Z^n, \bar{Y}^n) \text{ i.i.d.} \sim P_{X_1 X_2 Z \bar{Y}}, \quad (30)$$
$$\mathcal{H} = 1: \quad (X_1^n, X_2^n, Z^n, \bar{Y}^n) \text{ i.i.d.} \sim P_{X_1 X_2 Z} \cdot Q_{\bar{Y}|Z}. \quad (31)$$

In this case, any error exponent $\theta \geq 0$ is achievable that satisfies

$$\theta \leq \max \left(\mathbb{E}_{P_{ZV}}[D(P_{\bar{Y}|ZV}\|Q_{\bar{Y}|Z})] + I(S_1, S_2; \bar{Y}|Z, V) \right). \quad (32)$$

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

$$I(S_1; X_1|S_2, Z) \leq I(S_1; V|S_2, Z), \quad (33a)$$
$$I(S_2; X_2|S_1, Z) \leq I(S_2; V|S_1, Z), \quad (33b)$$
$$I(S_1, S_2; X_1, X_2|Z) \leq I(S_1, S_2; V|Z), \quad (33c)$$

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

$$P_{S_1 S_2 X_1 X_2 \bar{Y} Z V} = P_{S_1|X_1} \cdot P_{S_2|X_2} \cdot P_{X_1 X_2 \bar{Y} Z} \cdot \Gamma_{V|S_1 S_2 X_1 X_2}. \quad (34)$$

Proof: See [10]. ∎

For testing against conditional independence, i.e.,

$$Q_{\bar{Y}|Z} = P_{\bar{Y}|Z}, \quad (35)$$

with communication over noiseless bit-pipes of given rates, Corollary 1 recovers as a special case the result in [2]. When specialized to separate source-channel coding, it also recovers the achievable exponent in [3] for testing against independence over an *orthogonal MAC*, i.e., over a channel where $\Gamma_{V|W_1 W_2}$ decomposes into two individual DMCs:

$$V = (V_1, V_2) \quad (36a)$$
$$\Gamma_{V_1 V_2|W_1 W_2} = \Gamma_{V_1|W_1} \cdot \Gamma_{V_2|W_2}, \quad (36b)$$

for some choice of the DMCs $\Gamma_{V_1|W_1}$ and $\Gamma_{V_2|W_2}$. Testing against independence means that (35) holds as well as:

$$Z = \emptyset, \quad (37a)$$
$$Q_{\bar{Y}} = P_{\bar{Y}}. \quad (37b)$$

The following theorem proves optimality of the exponent in Corollary 1 for generalized testing against independent when the sources are independent under both hypotheses:

$$P_{X_1 X_2} = P_{X_1} \cdot P_{X_2} \quad (38)$$

Theorem 2: Consider generalized testing against independence with independent sources when the communication from the sensors to the decision center takes place over two orthogonal DMCs. Thus, assume that (35)–(38) hold. If C_1 and C_2 denote the capacities of the two DMCs $\Gamma_{V_1|W_1}$ and $\Gamma_{V_2|W_2}$, then, the optimal error exponent is:

$$\theta^* = D(P_Y\|Q_Y) + \max_{\substack{P_{\bar{S}_i|X_i}, P_{W_i}, i \in \{1,2\} \\ I(\bar{S}_1; X_1|\bar{S}_2) \leq C_1 \\ I(\bar{S}_2; X_2|\bar{S}_1) \leq C_2 \\ I(\bar{S}_1, \bar{S}_2; X_1, X_2) \leq C_1 + C_2}} I(\bar{S}_1, \bar{S}_2; Y). \quad (39)$$

Proof: See Appendix A. ∎

We specialize above Theorem to an example with independent Gaussian sources.[2]

Example 1: Let X_1 and X_2 be independent standard Gaussians under both hypotheses. Under the null hypothesis,

$$\mathcal{H} = 0: \qquad Y = X_1 + X_2 + N_0, \qquad N_0 \sim \mathcal{N}(0, \sigma_0^2), \quad (40)$$

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

$$\mathcal{H} = 1: \qquad Y \sim \mathcal{N}(0, \sigma_y^2), \text{ independent of } (X_1, X_2), \quad (41)$$

for a given nonnegative variance $\sigma_y^2 > 0$. Further assume an orthogonal MAC as in (36) with the two individual DMCs $\Gamma_{V_1|W_1}$ and $\Gamma_{V_2|W_2}$ having capacities C_1 and C_2.

The described setup is a special case of the setup considered in Theorem 2. The error exponent in (39) evaluates to (for the proof, see [10]):

$$\theta^* = \frac{1}{2} \log \left(\frac{\sigma_y^2}{2^{-2C_1} + 2^{-2C_2} + \sigma_0^2} \right) + \left(\frac{2 + \sigma_0^2}{2\sigma_y^2} - \frac{1}{2} \right) \cdot \log e. \quad (42)$$

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

Proposition 1: Consider testing against independence over an orthogonal MAC, i.e., assume that (35)–(36) hold. Then, the optimal error exponent is given by

$$\theta^* = D(P_Y\|Q_Y) + \lim_{n \to \infty} \frac{1}{n} \max I(S_1^n, S_2^n; Y^n), \quad (43)$$

[2]Strictly speaking, the results first have to be extended to continuous alphabets. But this is a technicality that we omit here.

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

$$\lim_{n\to\infty} \frac{1}{n} I(X_1^n; S_1^n | S_2^n) \leq C_1, \tag{44}$$

$$\lim_{n\to\infty} \frac{1}{n} I(X_2^n; S_2^n | S_1^n) \leq C_2, \tag{45}$$

$$\lim_{n\to\infty} \frac{1}{n} I(X_1^n, X_2^n; S_1^n, S_2^n) \leq C_1 + C_2. \tag{46}$$

Proof: See [10]. ∎

V. A GAUSSIAN EXAMPLE WITH A CONVERSE BOUND

Consider a symmetric Gaussian setup where under both hypotheses:

$$(X_1, X_2) \sim \mathcal{N}(0, \mathbf{K}_{X_1 X_2}) \tag{47}$$

for a positive semidefinite covariance matrix

$$\mathbf{K}_{X_1 X_2} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}, \qquad 0 \leq \rho \leq 1. \tag{48}$$

Assume as in Example 1 that under the null hypothesis,

$$\mathcal{H} = 0: \qquad Y = X_1 + X_2 + N_0, \qquad N_0 \sim \mathcal{N}(0, \sigma_0^2), \tag{49}$$

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

$$\mathcal{H} = 1: \qquad Y \sim \mathcal{N}(0, \sigma_y^2), \text{ independent of } (X_1, X_2), \tag{50}$$

for $\sigma_y^2 > 0$.

Communication takes place over the Gaussian MAC

$$V = W_1 + W_2 + N, \tag{51}$$

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

The described setup corresponds to generalized testing against independence. We can thus use Corollary 1 to obtain an achievable error exponent for this problem. The above choice of random variables yields the following result on the achievable error exponent.

Corollary 2: For the described Gaussian setup any error exponent $\theta \geq 0$ is achievable that satisfies the following condition:

$$\theta \leq \left(\frac{\sigma_0^2 + 2 + 2\rho}{2\sigma_y^2} - \frac{1}{2} \right) \cdot \log e$$

$$+ \max \frac{1}{2} \log \frac{\sigma_y^2}{\frac{2\xi^2(1+\rho)\sigma^2}{2\xi^2(\alpha-\beta)^2 \cdot (1+\rho) + \sigma^2(1+\rho+\xi^2)} + \sigma_0^2}, \tag{52}$$

where the maximization is over all $\xi^2, \alpha^2, \beta^2, \gamma^2 \geq 0$ such that

$$\gamma^2 + \alpha^2 + \beta^2 \xi^2 \leq P, \tag{53}$$

and

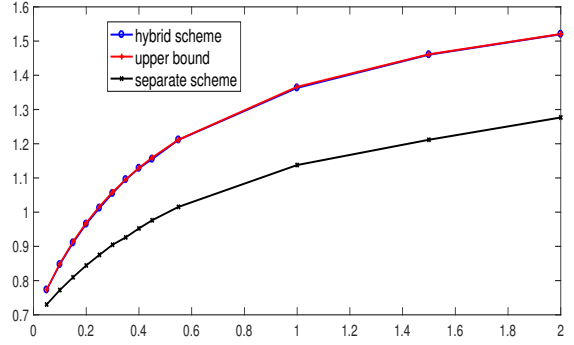$$\frac{(1+\xi^2)^2 - \rho^2}{(1+\xi^2) \cdot \xi^2} \leq$$

Fig. 2. Upper and lower bounds on the optimal exponent θ^* of the proposed Gaussian example for $\rho = 0.8$, $\sigma_0^2 = 1$, $\sigma_y^2 = 1.5$ and $\sigma^2 = 1$.

$$\frac{\sigma^2 + 2P - \gamma^2 + 2\alpha^2\rho - \frac{(\alpha\cdot(1+\rho)+\beta\cdot\xi^2)^2}{1+\xi^2}}{\sigma^2 + \frac{2(\alpha-\beta)^2\cdot(1+\rho)\xi^2}{1+\rho+\xi^2}}, \tag{54a}$$

$$\frac{(1+\xi^2)^2 - \rho^2}{\xi^4} \leq \frac{\sigma^2 + 2P + 2\alpha^2\rho}{\sigma^2 + \frac{2(\alpha-\beta)^2\cdot(1+\rho)\xi^2}{1+\rho+\xi^2}}. \tag{54b}$$

Proof: See [10]. ∎

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

Theorem 3: For the proposed Gaussian setup, the optimal error exponent θ^* satisfies

$$\theta^* \leq \frac{1}{2} \log \left(\frac{\sigma_y^2}{\frac{2(1+\rho)\sigma^2}{2P(1+\rho)+\sigma^2} + \sigma_0^2} \right)$$

$$+ \frac{1}{2} \left(\frac{2 + 2\rho + \sigma_0^2}{\sigma_y^2} - 1 \right) \cdot \log e. \tag{55}$$

Proof: See [10]. ∎

Figure 2 compares the presented upper and lower bounds on the optimal error exponent θ^*. They are very close for the considered setup. For comparison, the figure also shows the exponent that is achieved with the same choice of source variables but with separate source-channel coding. That means, by specializing the exponent in (52) to $\alpha = \beta = 0$.

VI. CONCLUSION

In this work, distributed hypothesis testing over a MAC is studied. A coding and testing scheme is proposed which combines hybrid coding with Borade's unequal error protection. Optimality of the achievable error exponent is shown for some special cases of generalized testing against independence. A future work is to extend the proposed achievable scheme for other communication scenarios such as broadcast channels.

REFERENCES

[1] T. S. Han and S. Amari, "Statistical inference under multiterminal data compression," *IEEE Trans. on Info. Theory*, vol. 44, no. 6, pp. 2300–2324, Oct. 1998.

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

[3] S. Sreekumar and D. Gündüz, "Distributed hypothesis testing over orthogonal discrete memoryless channels," available at https://arxiv.org/abs/1802.07665.

[4] S. Salehkalaibar and M. Wigger, "Distributed hypothesis testing over a noisy channel," *in Proc. International Zurich Seminar (IZS)*, Zurich, Switzerland, Feb. 2018.

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

[6] S. P. Borade, "When all information is not created equal," Thesis, Massachusetts Institute of Technology, 2008.

[7] P. Minero, S. H. Lim, and Y.-H. Kim, "A unified approach to hybrid coding," *IEEE Trans. on Info. Theory*, vol. 61, no. 4, pp. 1509–1523, Apr. 2015.

[8] A. El Gamal and Y. H. Kim, *Network information theory*, Cambridge Univ. Press, 2011.

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

[10] S. Salehkalaibar and M. Wigger, "Distributed hypothesis testing based on unequal-error protection codes," Avaiable at https://arxiv.org/abs/1806.05533.

APPENDIX A
PROOF OF CONVERSE FOR THEOREM 2

All mutual informations are with respect to the distribution under $\mathcal{H} = 0$. Define $\bar{S}_{1,t} \triangleq (V_1^n, X_1^{t-1})$ and $\bar{S}_{2,t} \triangleq (V_2^n, X_2^{t-1})$. Notice that the Markov chains $\bar{S}_{1,t} \to X_{1,t} \to \bar{S}_{2,t}$ and $\bar{S}_{2,t} \to X_{2,t} \to S_{1,t}$ hold. Define $\delta(\epsilon) := H(\epsilon)/n/(1-\epsilon)$ as in [9]. Then:

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

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

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

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

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

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

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

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

$$= \frac{1}{n(1-\epsilon)} \sum_{t=1}^{n} I(\bar{S}_{1,t}, \bar{S}_{2,t}; Y_t)$$

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

$$= \frac{1}{1-\epsilon} I(\bar{S}_1, \bar{S}_2; Y) + \frac{1}{1-\epsilon} \cdot D(P_Y \| Q_Y) + \delta(\epsilon), \quad (56)$$

where (a) follows from the Markov chain $Y^{t-1} \to (V^n, X_1^{t-1}, X_2^{t-1}) \to Y_t$. The last equality holds by defining a time-sharing random variable Q that is uniform over $\{1, \ldots, n\}$ and $\bar{S}_i := (Q, V_i^n, X_i^{Q-1})$ for $i \in \{1, 2\}$ and $Y := Y_Q$. Next, consider the following term,

$$I(X_1^n; V_1^n | V_2^n) = \sum_{t=1}^{n} I(X_{1,t}; V_1^n | X_1^{t-1}, V_2^n)$$

$$\overset{(b)}{=} \sum_{t=1}^{n} I(X_{1,t}; X_1^{t-1}, V_1^n | V_2^n)$$

$$\overset{(c)}{=} \sum_{t=1}^{n} I(X_{1,t}; X_1^{t-1}, V_1^n, X_2^{t-1} | V_2^n)$$

$$\geq \sum_{t=1}^{n} I(X_{1,t}; X_1^{t-1}, V_1^n | X_2^{t-1}, V_2^n)$$

$$= \sum_{t=1}^{n} I(X_{1,t}; \bar{S}_{1,t} | \bar{S}_{2,t})$$

$$= n I(X_1; \bar{S}_1 | \bar{S}_2) \quad (57)$$

where (b) and (c) follow from the Markov chains $X_{1,t} \to V_2^n \to X_1^{t-1}$ and $X_{1,t} \to (V_1^n, V_2^n, X_1^{t-1}) \to X_2^{t-1}$, respectively. Both Markov chains hold because X_1^n and X_2^n are independent under both hypotheses and also the orthogonality of the MAC. The last equality holds by defining $X_i := (Q, X_{i,Q})$ for $i \in \{1, 2\}$. Similarly, we get

$$I(X_2^n; V_2^n | V_1^n) \geq n I(X_2; \bar{S}_2 | \bar{S}_1), \quad (58)$$

$$I(X_1^n, X_2^n; V_1^n, V_2^n) \geq n I(X_1, X_2; \bar{S}_1, \bar{S}_2). \quad (59)$$

On the other hand, we have

$$I(X_1^n; V_1^n | V_2^n) \leq H(V_1^n) - H(V_1^n | W_1^n, V_2^n)$$

$$\overset{(d)}{=} H(V_1^n) - H(V_1^n | W_1^n)$$

$$\leq \sum_{t=1}^{n} I(W_{1,t}; V_{1,t})$$

$$= n I(W_1; V_1)$$

$$\leq n C_1, \quad (60)$$

where (d) follows from the Markov chain $V_1^n \to W_1^n \to V_2^n$ and the orthogonality assumption. The last equality holds by defining $W_i := (Q, W_{i,Q})$ and $V_i = V_{i,Q}$ for $i \in \{1, 2\}$. Similarly, we have

$$I(X_2^n; V_2^n | V_1^n) \leq n C_2, \quad (61)$$

$$I(X_2^n, X_1^n; V_1^n, V_2^n) \leq n C_1 + n C_2. \quad (62)$$

Combining these findings, concludes the proof of the converse.