

# Hypothesis Testing over Cascade Channels

<sup>1,2</sup>Sadaf Salehkalaibar, <sup>1</sup>Michèle Wigger, and <sup>3</sup>Ligong Wang

<sup>1</sup>LTCI, Telecom ParisTech, Université Paris-Saclay, 75013 Paris, France, michele.wigger@telecom-paristech.fr

<sup>2</sup>ECE Department, College of Engineering, University of Tehran, Tehran, Iran, s.saleh@ut.ac.ir

<sup>3</sup>ETIS – Université Paris Seine, Université de Cergy-Pontoise, ENSEA, CNRS, ligong.wang@ensea.fr

**Abstract**—Binary hypothesis testing over single and parallel cascade channels is considered where sensors communicate with dedicated relays, and these relays with a single final receiver. All relays as well as the final receiver decide on the binary hypothesis governing the joint probability distribution of the observations at the sensors, relays, and final receiver. The quantity of interest is the set of feasible type-II error exponents that allow for the type-I error probabilities to vanish asymptotically as the observation length increases. A coding scheme is proposed and the corresponding set of feasible type-II error exponents is analyzed by means of a modified Han-type analysis that can account for distributed decisions based on different codebooks and for nodes forwarding their decisions to other nodes. The obtained exponent region is optimal in some special cases.

## I. INTRODUCTION

Consider a multi-hop hypothesis testing problem where sensors communicate with intermediate relays, which then communicate with a single final decision center. This setup models scenarios where intermediate relays collect information from nearby sensors and send a sort of summary information to a final decision center that is located farther away. Compared to traditional architectures where the sensors directly have to communicate with the final decision center, this new architecture allows to reduce the energy consumption at the sensors, because they only have to communicate with the closer relays. Moreover, in this setup, intermediate relays can also take decisions, which can be exploited to faster detect certain events, though typically with reduced reliability compared to the decisions taken at the final decision center.

In this work, we start with the simplest scenario with a single sensor, a single relay, and a final receiver, and with both the relay and the receiver deciding on a binary hypothesis ( $\mathcal{H} = 0$  or  $\mathcal{H} = 1$ ) underlying the joint probability distribution of the observations at the three terminals. Communication takes place over individual noise-free bit pipes of given rates. We present a coding scheme for this scenario and analyze its type-II error (deciding on  $\hat{\mathcal{H}} = 0$  when  $\mathcal{H} = 1$ ) exponent so that the type-I error (deciding on  $\hat{\mathcal{H}} = 1$  when  $\mathcal{H} = 0$ ) probability vanishes as the observation length grows to infinity. For some special cases, the proposed type-II error exponent is optimal under the rate constraints imposed on the bit pipes.

We further propose a coding scheme for an extended setup with two sensors, two dedicated relays, and a single common final receiver, and analyze the type-II error exponents given that the type-I error probabilities vanish. Again, the proposed type-II error exponents are optimal in some special cases.

Ahlsvede and Csiszár [1], Han [2], and Shimokawa, Han, and Amari [3] studied a related setup where one or two

observers directly communicate with a single final decision center. Upper and lower bounds were derived on the maximum type-II error exponent under the requirement of vanishing type-I error probability. The exact maximum type-II error exponent was derived for “testing against independence,” where under the alternative hypothesis  $\mathcal{H} = 1$ , the joint distribution of the observations is given by the product of the marginal distributions under  $\mathcal{H} = 0$ . A similar result for “testing against conditional independence” was derived in [4]. Recently, also more involved communication scenarios were considered, for example, with noisy communication channels [5], [6] or interaction between two terminals [7]–[9]. Multiple decision centers were considered in [10].

The results in this work are obtained by refining Han’s construction and analysis [2]. We divide the transmitter’s message into two parts, one for the relay’s decision, and the other to be forwarded to the receiver. The relay’s message is also in two parts, one coming directly from the transmitter, and the other chosen based on both its own observation and the message it receives. For this scheme to perform well, we require that no joint type occurs “too often” in the codebooks.

*Notation:* We mostly follow the notation in [11]. Moreover, we use  $\text{tp}(\cdot)$  to denote the *joint type* of a tuple. For a joint type  $\pi_{ABC}$ , we let  $I_{\pi_{ABC}}(A; B|C)$  denote the conditional mutual information assuming  $(A, B, C)$  has probability mass function (pmf)  $\pi_{ABC}$ . Similarly for the entropy  $H_{\pi_{ABC}}(A)$  and the conditional entropy  $H_{\pi_{ABC}}(A|B)$ . When  $\pi_{ABC}$  is clear from the context, we write  $\pi_A$  and  $\pi_{AB}$  for the corresponding subtypes and sometimes even abbreviate  $\pi_{ABC}$  by  $\pi$ .

## II. HYPOTHESIS TESTING OVER A SINGLE CASCADE CHANNEL

### A. Setup

Consider the cascade hypothesis testing problem with three terminals in Fig. 1. The first terminal in the system, the *transmitter*, observes the sequence  $X^n = (X_1, \dots, X_n)$ ; the second terminal, the *relay*, observes the sequence  $Y^n = (Y_1, \dots, Y_n)$ ; and the third terminal, the *receiver*, observes the sequence  $Z^n = (Z_1, \dots, Z_n)$ . Under the null hypothesis

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

whereas under the alternative hypothesis

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

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

There is a noise-free bit pipe of rate  $R$  from the transmitter to the relay, and a noise-free bit pipe of rate  $T$  from the relay to

the receiver. After observing  $X^n$ , the transmitter computes the message  $M = \phi^{(n)}(X^n)$  using a possibly stochastic encoding function  $\phi^{(n)} : \mathcal{X}^n \rightarrow \{0, \dots, 2^{nR}\}$  and sends it over the bit pipe to the relay.<sup>1</sup> The relay, after observing  $Y^n$  and receiving  $M$ , computes the message  $B = \phi_y^{(n)}(M, Y^n)$  using a possibly stochastic encoding function  $\phi_y^{(n)} : \mathcal{Y}^n \times \{0, \dots, 2^{nR}\} \rightarrow \{0, \dots, 2^{nT}\}$  and sends it over the bit pipe to the receiver.

The goal of the above communication is that, based on their own observations and based on the received messages, the relay and the receiver should be able to produce their respective guesses of  $\mathcal{H}$ . The relay produces the guess

$$\hat{\mathcal{H}}_y = g_y^{(n)}(Y^n, M) \quad (3)$$

using a decoding function  $g_y^{(n)} : \mathcal{Y}^n \times \{0, \dots, 2^{nR}\} \rightarrow \{0, 1\}$ , and the receiver produces the guess

$$\hat{\mathcal{H}}_z = g_z^{(n)}(Z^n, B) \quad (4)$$

using a decoding function  $g_z^{(n)} : \mathcal{Z}^n \times \{0, \dots, 2^{nT}\} \rightarrow \{0, 1\}$ .

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

$$\gamma_n \triangleq \Pr[\hat{\mathcal{H}}_y = 1 | \mathcal{H} = 0], \quad (5)$$

$$\zeta_n \triangleq \Pr[\hat{\mathcal{H}}_y = 0 | \mathcal{H} = 1], \quad (6)$$

and at the receiver

$$\alpha_n \triangleq \Pr[\hat{\mathcal{H}}_z = 1 | \mathcal{H} = 0], \quad (7)$$

$$\beta_n \triangleq \Pr[\hat{\mathcal{H}}_z = 0 | \mathcal{H} = 1], \quad (8)$$

satisfy, for sufficiently large  $n$ ,

$$\alpha_n \leq \epsilon, \quad (9)$$

$$\gamma_n \leq \epsilon, \quad (10)$$

and

$$-\overline{\lim}_{n \rightarrow \infty} \frac{1}{n} \log \beta_n \geq \theta, \quad (11)$$

$$-\overline{\lim}_{n \rightarrow \infty} \frac{1}{n} \log \zeta_n \geq \eta. \quad (12)$$

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

### B. Results on the Exponent Region

Our main achievability result is the next theorem.

*Theorem 1:* The above scheme achieves all nonnegative eponent pairs  $(\eta, \theta)$  that satisfy

$$\eta < \min_{\substack{\tilde{P}_{SUVXYZ}: \\ \tilde{P}_{SUX} = P_{SUX} \\ \tilde{P}_{SVY} = P_{SVY}}} D(\tilde{P}_{SUVXYZ} || P_{SU|X} Q_{XYZ}), \quad (13)$$

<sup>1</sup>For notational convenience, we allow the total number of messages to be  $2^{nR} + 1$ . Clearly, this does not change the message rate as  $n$  tends to infinity.

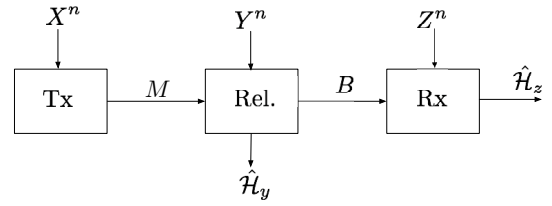


Fig. 1. Hypothesis testing over a cascade channel

$$\theta < \min_{\substack{\tilde{P}_{SUVXYZ}: \\ \tilde{P}_{SUX} = P_{SUX} \\ \tilde{P}_{SVUY} = P_{SVUY} \\ \tilde{P}_{SVZ} = P_{SVZ}}} D(\tilde{P}_{SUVXYZ} || P_{SU|X} P_{V|SUY} Q_{XYZ}), \quad (14)$$

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

$$R \geq I(U, S; X), \quad (15)$$

$$T \geq I(X; S) + I(V; Y, U|S). \quad (16)$$

We shall prove Theorem 1 by describing and analyzing a coding scheme. We start with some preliminaries.

**Preliminaries:** Fix  $\mu > 0$ , an arbitrary blocklength  $n$ , and joint conditional pmfs  $P_{SU|X}$  and  $P_{V|SUY}$  over finite auxiliary alphabets  $\mathcal{S}$ ,  $\mathcal{U}$ , and  $\mathcal{V}$ . Define the joint pmf  $P_{SUVXYZ} = P_{XYZ} P_{SU|X} P_{V|SUY}$  and the following nonnegative rates, which are calculated according to  $P_{SUVXYZ}$  for some parameter  $\mu > 0$ ,

$$R_s \triangleq I(X; S) + \mu, \quad (17)$$

$$R_u \triangleq I(U; X|S) + \mu, \quad (18)$$

$$R_v \triangleq I(V; Y, U|S) + \mu. \quad (19)$$

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

The following lemma is key to our main result (Theorem 1). An important aspect of the lemma is that it bounds the number of codeword tuples with a given joint type, see point 3). Without insisting on 3), the scheme's error exponent would be decreased by  $I(U; V|S)$ .

*Lemma 1:* For every sufficiently large blocklength  $n$ , it is possible to create a codebook

$$\mathcal{C}_S = \{s^n(i) : i \in \{1, \dots, 2^{nR_s}\}\}, \quad (20)$$

and for each  $i \in \{1, \dots, 2^{nR_s}\}$  the sub-codebooks

$$\mathcal{C}_U(i) = \{u^n(j|i) : j \in \{1, \dots, 2^{nR_u}\}\} \quad (21)$$

$$\mathcal{C}_V(i) = \{v^n(k|i) : k \in \{1, \dots, 2^{nR_v}\}\}, \quad (22)$$

such that the following three properties hold:

- 1) Codebooks  $\mathcal{C}_S$  and  $\mathcal{C}_U$  cover the source sequence  $X^n$  with high probability:

$$\Pr[\exists(i, j) : (s^n(i), u^n(j|i), X^n) \in \mathcal{T}_{\mu/4}^n(P_{SUX})] > 1 - \epsilon/4. \quad (23)$$

- 2) If codebooks  $\mathcal{C}_S$  and  $\mathcal{C}_U(i)$  cover  $Y^n$  for some  $i$ , then also codebooks  $\mathcal{C}_S$ ,  $\mathcal{C}_U(i)$ , and  $\mathcal{C}_V(i)$  cover  $Y^n$  with high probability:

$$\Pr \left[ \exists k: (s^n(i), u^n(j|i), v^n(k|i), Y^n) \in \mathcal{T}_{\mu/2}^n(P_{SUVY}) \mid (s^n(i), u^n(j|i), Y^n) \in \mathcal{T}_{3\mu/8}^n(P_{SUY}) \right] > 1 - \epsilon/4. \quad (24)$$

- 3) No joint type  $\pi_{SUV}$  is “over-represented” in the triple of codebooks  $\mathcal{C}_S$ ,  $\mathcal{C}_U$ , and  $\mathcal{C}_V$  in the sense that, for all  $\pi_{SUV}$ ,

$$\left| \{(i, j, k): \text{tp}(s^n(i), u^n(j|i), v^n(k|i)) = \pi_{SUV}\} \right| \leq 2^{n(R_s + R_u + R_v - I_{\pi_{SUV}}(U;V|S) + \mu)}. \quad (25)$$

*Proof:* See [12]. ■

**Code Construction:** Choose a codebook satisfying the three conditions in Lemma 1.

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

$$(s^n(i), u^n(j|i), x^n) \in \mathcal{T}_{\mu/4}^n(P_{SUX}). \quad (26)$$

If successful, it picks one such pair uniformly at random and sends  $m = (i, j)$  over the bit pipe to the relay. Otherwise, it sends  $m = 0$ .

**Relay:** Assume that the relay observes the sequence  $y^n$  and receives the message  $m$ . If  $m = 0$ , it declares  $\hat{\mathcal{H}}_y = 1$  and sends  $b = 0$  over the bit pipe to the receiver. Otherwise, it looks for an index  $k$  such that

$$(s^n(i), u^n(j|i), v^n(k|i), y^n) \in \mathcal{T}_{\mu/2}^n(P_{SUVY}). \quad (27)$$

If such an index  $k$  exists, the relay sends the pair  $b = (i, k)$  to the receiver. Otherwise, it sends  $b = 0$ .

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

$$(s^n(i), v^n(k|i), z^n) \in \mathcal{T}_{\mu}^n(P_{SVZ}). \quad (28)$$

If (28) is true, the receiver declares  $\hat{\mathcal{H}}_z = 0$ . Otherwise, it declares  $\hat{\mathcal{H}}_z = 1$ .

**Analysis:** If  $M \neq 0$  and  $B \neq 0$ , let  $I, J, K$  be the random indices sent over the noise-free bit pipes. We first analyze the type-I error probability at the receiver. For the case  $M \neq 0$  and  $B \neq 0$ , define the following events:

$$\begin{aligned} \mathcal{E}_{\text{Relay}} &: (s^n(I), u^n(J|I), Y^n) \notin \mathcal{T}_{3\mu/8}^n(P_{SUY}), \\ \mathcal{E}_{\text{Rx}} &: (s^n(I), u^n(J|I), v^n(K|I), Z^n) \notin \mathcal{T}_{\mu}^n(P_{SVZ}). \end{aligned}$$

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

$$\begin{aligned} \alpha_n &\leq \Pr[M = 0 \text{ or } B = 0 \text{ or } \mathcal{E}_{\text{Relay}} \text{ or } \mathcal{E}_{\text{Rx}}] \\ &\leq \Pr[M = 0] + \Pr[B = 0 \text{ or } \mathcal{E}_{\text{Relay}} | M \neq 0] \\ &\quad + \Pr[\mathcal{E}_{\text{Rx}} | M \neq 0, B \neq 0] \end{aligned} \quad (29)$$

$$\begin{aligned} &\stackrel{(a)}{\leq} \epsilon/4 + \Pr[\mathcal{E}_{\text{Relay}} | M \neq 0] \\ &\quad + \Pr[B = 0 | M \neq 0, \mathcal{E}_{\text{Relay}}^c] + \epsilon/4 \end{aligned} \quad (30)$$

$$\stackrel{(b)}{\leq} \epsilon/4 + \epsilon/4 + \epsilon/4 + \epsilon/4 \quad (31)$$

$$= \epsilon, \quad (32)$$

where (a) holds because the chosen code construction satisfies (23) and by the Markov lemma, and (b) holds because the code construction satisfies (24) and by the Markov lemma.

We now bound the probability of type-II error at the receiver. Let  $\mathcal{P}^n$  be the set of all types over the product alphabet  $\mathcal{S}^n \times \mathcal{U}^n \times \mathcal{V}^n \times \mathcal{X}^n \times \mathcal{Y}^n \times \mathcal{Z}^n$ . Also, let  $\mathcal{P}_{\mu}^n$ ,  $\mu > 0$ , be the subset of types  $\pi_{SUVXYZ} \in \mathcal{P}^n$  that simultaneously satisfy the following conditions: for all  $(s, u, v, x, y, z)$ ,

$$|\pi_{SUX}(s, u, x) - P_{SUX}(s, u, x)| \leq \mu/4, \quad (33)$$

$$|\pi_{SUVY}(s, u, v, y) - P_{SUVY}(s, u, v, y)| \leq \mu/2, \quad (34)$$

$$|\pi_{SVZ}(s, v, z) - P_{SVZ}(s, v, z)| \leq \mu. \quad (35)$$

The type-II error probability can then be bounded as

$$\beta_n \leq \Pr[(X^n, Y^n, Z^n) \in \mathcal{A}_{\text{Rx},n} | \mathcal{H} = 1], \quad (36)$$

where

$$\begin{aligned} \mathcal{A}_{\text{Rx},n} &\triangleq \bigcup_{i,j,k} \left\{ (x^n, y^n, z^n): \right. \\ &\quad \left. \text{tp}(s^n(i), u^n(j|i), v^n(k|i), x^n, y^n, z^n) \in \mathcal{P}_{\mu}^n \right\}. \end{aligned} \quad (37)$$

We thus have

$$\begin{aligned} \beta_n &\leq \sum_{\pi_{SUVXYZ} \in \mathcal{P}_{\mu}^n} \sum_{\substack{(i,j,k): \\ \text{tp}(s^n(i), u^n(j|i), v^n(k|i)) \\ = \pi_{SUV}}} \\ &\quad \sum_{\substack{(x^n, y^n, z^n): \\ \text{tp}(s^n(i), u^n(j|i), v^n(k|i), x^n, y^n, z^n) \\ = \pi_{SUVXYZ}}} \\ &\quad \Pr[X^n = x^n, Y^n = y^n, Z^n = z^n | \mathcal{H} = 1]. \end{aligned} \quad (38)$$

Notice that for a triple  $(x^n, y^n, z^n)$  of type  $\pi_{XYZ}$ ,

$$\begin{aligned} \Pr[X^n = x^n, Y^n = y^n, Z^n = z^n | \mathcal{H} = 1] \\ = 2^{-n(H_{\pi}(X,Y,Z) + D(\pi_{XYZ} || Q_{XYZ}))}. \end{aligned} \quad (39)$$

Moreover, by standard arguments, for any joint type  $\pi_{SUVXYZ}$  and any triple of sequences  $(s^n, u^n, v^n)$  of matching subtype  $\pi_{SUV}$ :

$$\begin{aligned} \left| \{(x^n, y^n, z^n): \text{tp}(s^n, u^n, v^n, x^n, y^n, z^n) = \pi_{SUVXYZ}\} \right| \\ \leq 2^{nH_{\pi}(X,Y,Z|S,U,V)}. \end{aligned} \quad (40)$$

Also,

$$|\mathcal{P}_{\mu}^n| \leq |\mathcal{P}^n| \leq (n+1)^{|\mathcal{S}| \cdot |\mathcal{U}| \cdot |\mathcal{V}| \cdot |\mathcal{X}| \cdot |\mathcal{Y}| \cdot |\mathcal{Z}|}. \quad (41)$$

Combining (38)–(41) with Property (25) in Lemma 1 yields the following upper bound:

$$\begin{aligned} \beta_n &\leq (n+1)^{|\mathcal{S}| \cdot |\mathcal{U}| \cdot |\mathcal{V}| \cdot |\mathcal{X}| \cdot |\mathcal{Y}| \cdot |\mathcal{Z}|} \\ &\quad \times \max_{\pi_{SUVXYZ} \in \mathcal{P}_{\mu}^n} \left[ 2^{n(R_s + R_u + R_v - I_{\pi}(U;V|S) + \mu)} \right. \\ &\quad \left. \cdot 2^{nH_{\pi}(X,Y,Z|S,U,V)} \right] \end{aligned}$$

$$\cdot 2^{-n(H_\pi(XYZ) + D(\pi_{XYZ} \| Q_{XYZ}))} \Big]. \quad (42)$$

Plugging the rate expressions (17)–(19) into (42), and simplifying, result in the following upper bound:

$$\beta_n \leq (n+1)^{|S| \cdot |U| \cdot |V| \cdot |X| \cdot |Y| \cdot |Z|} \cdot 2^{-n\theta_\mu}, \quad (43)$$

where

$$\begin{aligned} \theta_\mu \triangleq & \min_{\pi_{SUVXYZ} \in \mathcal{P}_\mu^n} [H_\pi(XYZ) + D(\pi_{XYZ} \| Q_{XYZ}) \\ & - H_\pi(XYZ|SUV) + I_\pi(U; V|S)] \\ & - I(X; SU) - I(YU; V|S). \end{aligned} \quad (44)$$

Now we let  $n \rightarrow \infty$  and  $\mu \rightarrow 0$ . By continuity properties of the relative entropy and some simple manipulations, we obtain that  $-\frac{1}{n} \log \beta_n$  tends to

$$\begin{aligned} \theta = & \min_{\substack{\tilde{P}_{SUVXYZ}: \\ \tilde{P}_{SUX} = P_{SUX} \\ \tilde{P}_{SVUY} = P_{SVUY} \\ \tilde{P}_{SVZ} = P_{SVZ}}} D(\tilde{P}_{SUVXYZ} \| P_{SUX} P_{SVUY} Q_{XYZ}). \end{aligned} \quad (45)$$

Similar steps can be used also to analyze the probability of type-I and type-II errors at the relay.

### C. Optimality Results on Exponent Region

The achievable exponent region of Theorem 1 is optimal in some special cases. We discuss three of them in the following. For the first special case, consider a setup where the pmfs  $P_{XYZ}$  and  $Q_{XYZ}$  decompose as

$$P_{XYZ} = P_X \cdot P_{Y|X} \cdot P_{Z|Y}, \quad (46)$$

$$Q_{XYZ} = P_X \cdot P_Y \cdot P_Z, \quad (47)$$

and where  $Y^n$  can be losslessly described to the receiver:

$$T \geq H(Y). \quad (48)$$

*Proposition 1:* If (46)–(48) hold, then the exponent region  $\mathcal{E}(R, T)$  is the set of all nonnegative pairs  $(\eta, \theta)$  that satisfy

$$\eta \leq I(U; Y), \quad (49)$$

$$\theta \leq I(U; Y) + I(Y; Z), \quad (50)$$

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

$$R \geq I(U; X). \quad (51)$$

*Proof:* Achievability follows by specializing Theorem 1 to  $S = U$  and  $V = Y$ . For the converse, see [12]. ■

Next consider a setup where (48) holds and

$$P_{XYZ} = P_{XZ} \cdot P_{Y|Z}, \quad (52)$$

$$Q_{XYZ} = P_{XZ} \cdot P_Y. \quad (53)$$

*Proposition 2:* If (48), (52), and (53) hold, then  $\mathcal{E}(R, T)$  is the set of all nonnegative pairs  $(\eta, \theta)$  satisfying

$$\eta \leq I(U; Y), \quad (54)$$

$$\theta \leq I(Y; Z), \quad (55)$$

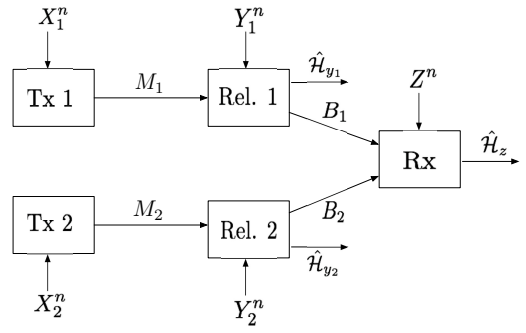


Fig. 2. Hypothesis testing over a parallel cascade network.

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

$$R \geq I(U; X). \quad (56)$$

*Proof:* Achievability follows again by specializing Theorem 1 to  $S = U$  and  $V = Y$ . For the converse, see [12]. ■

Our third special case is where

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

$$Q_{XYZ} = P_X \cdot P_{YZ}, \quad (58)$$

$$T \geq R. \quad (59)$$

*Proposition 3:* Assuming (57)–(59), the exponent region  $\mathcal{E}(R, T)$  is given by the set of all nonnegative pairs  $(\eta, \theta)$  such that

$$\eta \leq I(S; Y), \quad (60)$$

$$\theta \leq I(S; Z), \quad (61)$$

for some auxiliary random variable  $S$  satisfying  $S \rightarrow X \rightarrow (Y, Z)$  such that

$$R \geq I(S; X). \quad (62)$$

*Proof:* For achievability, let  $S = U = V$  in Theorem 1. For the converse, see [12]. ■

*Remark 1:* In achieving Propositions 1 and 2, the transmitter's message is only used by the relay to produce its guess on  $\mathcal{H}$ , and the relay's optimal encoding strategy is to inform the receiver its own guess and describe  $Y^n$ . For Proposition 3, the relay does the opposite: it simply forwards the message from the transmitter to the receiver. Both schemes are special cases of the more general scheme used in Theorem 1.

### III. HYPOTHESIS TESTING OVER A PARALLEL CASCADE NETWORK

Consider the two-transmitter, two-relay setup in Fig. 2. The observations at Transmitters 1 and 2, Relays 1 and 2, and the receiver are, respectively,  $X_1^n \triangleq (X_{1,1}, \dots, X_{1,n})$ ,  $X_2^n \triangleq (X_{2,1}, \dots, X_{2,n})$ ,  $Y_1^n \triangleq (Y_{1,1}, \dots, Y_{1,n})$ ,  $Y_2^n \triangleq (Y_{2,1}, \dots, Y_{2,n})$ , and  $Z^n \triangleq (Z_1, \dots, Z_n)$ . Depending on the value of the hypothesis  $\mathcal{H}$ , the tuple  $(X_1^n, X_2^n, Y_1^n, Y_2^n, Z^n)$  is i.i.d. according to one of the following two distributions:

$$\text{under } \mathcal{H} = 0: P_{X_1 X_2 Y_1 Y_2 Z}, \quad (63)$$

$$\text{under } \mathcal{H} = 1: Q_{X_1 X_2 Y_1 Y_2 Z}. \quad (64)$$

Transmitter  $i$ ,  $i \in \{1, 2\}$ , can communicate with its corresponding Relay  $i$  over a noise-free bit pipe of rate  $R_i$ , and Relay  $i$  in turn can communicate with the final receiver over a noise-free bit pipe of rate  $T_i$ . Encoding and detection functions, and probabilities of type-I and type-II errors are defined in a similar way as for the single cascade channel in Section II. Let  $\eta_1$  and  $\eta_2$  denote the type-II error exponents at Relays 1 and 2, respectively, and  $\theta$  that at the receiver. As in Section II, our interest is in identifying the set of all achievable exponent triples  $(\eta_1, \eta_2, \theta)$  under the condition that the type-I error probabilities at all three terminals must vanish as the observation length tends to infinity. The exponent region  $\mathcal{E}_{\text{par-2}}(R_1, R_2, T_1, T_2)$  for this setup is defined analogously as in Definition 2.

*Theorem 2:* The exponent region  $\mathcal{E}_{\text{par-2}}(R_1, R_2, T_1, T_2)$  includes all nonnegative triples  $(\eta_1, \eta_2, \theta)$  such that

$$\eta_i \leq \min_{\substack{\tilde{P}_{S_i U_i X_i Y_i} \\ \tilde{P}_{S_i U_i X_i} = P_{S_i U_i X_i} \\ \tilde{P}_{S_i U_i Y_i} = P_{S_i U_i Y_i}}} D(\tilde{P}_{S_i U_i X_i Y_i} \| P_{S_i U_i | X_i} Q_{X_i Y_i}), \quad i \in \{1, 2\} \quad (65)$$

$$\theta \leq \min_{\substack{\tilde{P}_{S_1 U_1 V_1 S_2 U_2 V_2 X_1 Y_1 X_2 Y_2 Z} \\ \tilde{P}_{S_i U_i X_i} = P_{S_i U_i X_i} \\ \tilde{P}_{S_i V_i U_i Y_i} = P_{S_i V_i U_i Y_i} \\ \tilde{P}_{S_1 S_2 V_1 V_2 Z} = P_{S_1 S_2 V_1 V_2 Z}}} D(\tilde{P}_{S_1 U_1 V_1 S_2 U_2 V_2 X_1 Y_1 X_2 Y_2 Z} \| P_{S_1 U_1 | X_1} P_{S_2 U_2 | X_2} P_{V_1 | S_1 U_1 Y_1} P_{V_2 | S_2 U_2 Y_2} Q_{X_1 X_2 Y_1 Y_2 Z}) \quad (66)$$

for some auxiliary random variables  $(S_1, S_2, U_1, U_2, V_1, V_2)$  satisfying, for  $i \in \{1, 2\}$ , the Markov chains  $(U_i, S_i) \rightarrow X_i \rightarrow (Y_i, Z)$  and  $V_i \rightarrow (S_i, Y_i) \rightarrow Z$  and the rate constraints

$$R_i \geq I(U_i, S_i; X_i), \quad (67)$$

$$T_i \geq I(X_i; S_i) + I(V_i; Y_i, U_i | S_i). \quad (68)$$

*Proof:* See [12]. ■

*Remark 2:* The main challenge in the analysis of the exponent region  $\mathcal{E}_{\text{par-2}}(R_1, R_2, T_1, T_2)$  is to bound the number of codewords with a same joint type, similarly to property 3) in Lemma 1. Without this step, the error exponent in (66) is reduced by  $I_{\tilde{P}}(U_1; V_1 | S_1) + I_{\tilde{P}}(U_2; V_2 | S_2) + I_{\tilde{P}}(S_1, U_1, V_1; S_2, U_2, V_2)$ .

Consider the special case where the pmfs  $P_{X_1 X_2 Y_1 Y_2 Z}$  and  $Q_{X_1 X_2 Y_1 Y_2 Z}$  decompose as

$$P_{X_1 X_2 Y_1 Y_2 Z} = P_{X_1 | Y_1} \cdot P_{X_2 | Y_2} \cdot P_{Y_1 Y_2} \cdot P_{Z | Y_1 Y_2}, \quad (69)$$

$$Q_{X_1 X_2 Y_1 Y_2 Z} = P_{X_1} \cdot P_{X_2} \cdot P_{Y_1 Y_2} \cdot P_Z, \quad (70)$$

and where  $(Y_1^n, Y_2^n)$  can be described losslessly to the receiver:

$$T_1 \geq H(Y_1), \quad (71)$$

$$T_2 \geq H(Y_2). \quad (72)$$

*Proposition 4:* If (69)–(72) hold, then the exponent region  $\mathcal{E}_{\text{par-2}}(R_1, R_2, T_1, T_2)$  is given by the set of all nonnegative triples  $(\eta_1, \eta_2, \theta)$  such that

$$\eta_i \leq I(U_i; Y_i), \quad i \in \{1, 2\}, \quad (73)$$

$$\theta \leq I(U_1; Y_1) + I(U_2; Y_2) + I(Y_1, Y_2; Z) \quad (74)$$

for some auxiliary random variables  $(U_1, U_2)$  satisfying, for  $i \in \{1, 2\}$ , the Markov chains  $U_i \rightarrow X_i \rightarrow (Y_i, Z)$  and the rate constraints

$$R_i \geq I(U_i; X_i). \quad (75)$$

*Proof:* For achievability, let  $U_i = S_i$  and  $V_i = Y_i$  for  $i = 1, 2$  in Theorem 2. For the converse, see [12]. ■

#### IV. CONCLUDING REMARKS

This paper extends Hans distributed hypothesis testing scheme to two cascade scenarios with relay terminals that also perform the hypothesis testing. Key components of the new schemes are cascade source-coding techniques and a *decision-forwarding strategy* where the final receiver only decides on the null-hypothesis if all terminals also have taken this decision. The schemes attain the optimal type-II error exponent for some special cases of testing against independence. Examples are provided where this optimal type-II error exponent is achieved when the cascade source-code reduces to a simple forwarding strategy at the relay or when it reduces to having the relay sending only local data to the receiver. An extension of the proposed coding scheme to include binning can be found in [12].

#### REFERENCES

- [1] A. Ahlswede and I. Csiszár, "Hypothesis testing with communication constraints," *IEEE Trans. on Info. Theory*, vol. 32, no. 4, pp. 533–542, Jul. 1986.
- [2] T. S. Han, "Hypothesis testing with multiterminal data compression," *IEEE Trans. on Info. Theory*, vol. 33, no. 6, pp. 759–772, Nov. 1987.
- [3] H. Shimokawa, T. Han and S. I. Amari, "Error bound for hypothesis testing with data compression," in *Proc. IEEE Int. Symp. on Info. Theory*, Jul. 1994, p. 114.
- [4] M. S. Rahman and A. B. Wagner, "On the Optimality of binning for distributed hypothesis testing," *IEEE Trans. on Info. Theory*, vol. 58, no. 10, pp. 6282–6303, Oct. 2012.
- [5] S. Salehkalaibar, M. Wigger and R. Timo, "On hypothesis testing against independence with multiple decision centers," available at: <https://arxiv.org/abs/1708.03941>.
- [6] S. Sreekuma and D. Gündüz, "Distributed hypothesis testing over noisy channels," available at: <https://arxiv.org/abs/1704.01535>.
- [7] W. Zhao and L. Lai, "Distributed testing against independence with multiple terminals," in *Proc. 52nd Allerton Conf. Comm. Cont. and Comp.*, IL, USA, pp. 1246–1251, Oct. 2014.
- [8] 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.
- [9] G. Katz, P. Piantanida and M. Debbah, "Collaborative distributed hypothesis testing," available at: <https://arxiv.org/abs/1604.01292>.
- [10] M. Wigger and R. Timo, "Testing against independence with multiple decision centers," in *Proc. of SPCOM*, Bangalore, India, Jun. 2016.
- [11] A. El Gamal and Y. H. Kim, *Network information theory*, Cambridge Univ. Press, 2011.
- [12] S. Salehkalaibar, M. Wigger and L. Wang, "Hypothesis testing in multi-hop networks," available at: <https://arxiv.org/abs/1708.05198>.