

Interactive Hypothesis Testing Against Independence

Yu Xiang and Young-Han Kim
 Department of Electrical and Computer Engineering
 University of California, San Diego
 La Jolla, CA 92093, USA
 Email: {yxiang,yhk}@ucsd.edu

Abstract—A hypothesis testing problem with communication constraints is studied, in which two nodes interactively communicate with each other in q rounds to perform a simple binary hypothesis testing on whether the observed random sequences at the nodes are generated independently or not. The optimal tradeoff between the communication rates in q rounds interaction and the testing performance measured by the type II error exponent such that the type I error probability asymptotically vanishes. An example is provided that shows that a two-way test outperforms the optimal one-way test and thus that interaction helps for hypothesis testing.

I. INTRODUCTION

Berger [1], in an inspiring attempt at combining information theory and statistical inference, formulated the problem of hypothesis testing with communication constraints as depicted in Fig. 1. Let $(X_1^n, X_2^n) \sim \prod_{i=1}^n p_{X_1, X_2}(x_{1i}, x_{2i})$ be a pair of independent and identically distributed (i.i.d.) n -sequences generated by a two-component discrete memoryless source (2-DMS) (X_1, X_2) . Suppose that there are two hypotheses on the joint distribution of (X_1, X_2) , namely,

$$\begin{aligned} H_0 : (X_1, X_2) &\sim p_0(x_1, x_2), \\ H_1 : (X_1, X_2) &\sim p_1(x_1, x_2). \end{aligned}$$

In order to decide which hypothesis is true, nodes 1 and 2 that observe X_1^n and X_2^n , respectively, compress their observed sequences into indices of rates R_1 and R_2 , and communicate them over noiseless links to node 3, which then makes a decision $\hat{H} \in \{H_0, H_1\}$ based on the received compression indices. What is the optimal tradeoff between the communication rates and the type I and II error probabilities? Despite many natural applications, however, theoretical understanding of this problem is far from complete and a simple characterization of this rate–exponent tradeoff remains open in general.

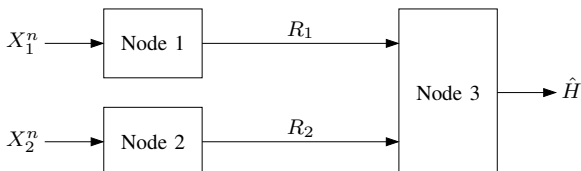


Fig. 1. Multiterminal hypothesis testing with communication constraints.

In their celebrated paper [2], Ahlswede and Csiszár studied the special case in which the sequence X_2^n is fully available at the destination node, i.e., $R_2 = \infty$. They established single-letter inner and outer bounds on the optimal tradeoff and showed that these bounds are tight for testing *against independence*, i.e., the alternative hypothesis H_1 is $p_1(x_1, x_2) = p_0(x_1)p_0(x_2)$. Later, Han [3] and Shimokawa, Han, and Amari [4] provided a new coding scheme that improves upon the Ahlswede and Csiszar inner bound for the general hypothesis testing problem. A more comprehensive survey on the earlier literature can be found in [5]. Several variations of this setup have been studied, including successive refinement hypothesis testing [6] and testing against conditional independence [7].

This paper studies an interactive version of hypothesis testing with communication constraints. Two nodes communicate with each other in q rounds through noiseless links and one of the nodes is to perform hypothesis testing at the end of interactive communication. For the special case of hypothesis testing *against independence*, we establish a single-letter characterization of the optimal tradeoff between the communication rates and the type II error probability when the type I error probability is arbitrarily small. The $q = 2$ case has been reported in [8].

The rest of the paper is organized as follows. In Section II, we review the problem of one-way hypothesis testing with communication constraints. In Section III, we formulate the problem of interactive hypothesis testing with communication constraints and present our main theorem. In Section IV, we compare the interactive hypothesis testing problem with the interactive lossy source coding problem by Kaspi [9].

Throughout the paper, we closely follow the notation in [10]. In particular, for $X \sim p(x)$ and $\epsilon \in (0, 1)$, we define the set of ϵ -typical n -sequences x^n (or the typical set in short) [11] as $\mathcal{T}_\epsilon^{(n)}(X) = \{x^n : |\#\{i : x_i = x\}/n - p(x)| \leq \epsilon p(x) \text{ for all } x \in \mathcal{X}\}$. We say that $X \rightarrow Y \rightarrow Z$ form a Markov chain if $p(x, y, z) = p(x)p(y|x)p(z|y)$, that is, X and Z are conditionally independent of each other given X_2 .

II. ONE-WAY HYPOTHESIS TESTING WITH COMMUNICATION CONSTRAINTS

As before, let $(X_1^n, X_2^n) \sim \prod_{i=1}^n p_{X_1, X_2}(x_{1i}, x_{2i})$ be a pair of i.i.d. sequences generated by a 2-DMS (X_1, X_2) and

consider hypothesis testing against independence

$$\begin{aligned} H_0 : (X_1, X_2) &\sim p_0(x_1, x_2), \\ H_1 : (X_1, X_2) &\sim p_1(x_1, x_2) = p_0(x_1)p_0(x_2). \end{aligned}$$

Here $p_0(x_1)$ and $p_0(x_2)$ are marginal distributions of $p_0(x_1, x_2)$. We consider the special case of the problem depicted in Fig. 1, in which $R_2 = \infty$; see Fig. 2.

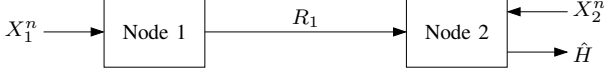


Fig. 2. One-way hypothesis testing with communication constraint.

A $(2^{nR_1}, n)$ hypothesis test consists of

- an encoder that assigns an index $m_1(x_1^n) \in [1 : 2^{nR_1}]$ to each sequence $x_1^n \in \mathcal{X}_1^n$, and
- a tester that assigns $\hat{h}(m_1, x_2^n) \in \{H_0, H_1\}$ to each $(m_1, x_2^n) \in [1 : 2^{nR_1}] \times \mathcal{X}_2^n$.

The acceptance region is defined as

$$\mathcal{A}_n := \{(m_1, x_2^n) \in [1 : 2^{nR_1}] \times \mathcal{X}_2^n : \hat{h}(m_1, x_2^n) = H_0\}.$$

Then the type I error probability is

$$P_0(\mathcal{A}_n^c) = \sum_{(x_1^n, x_2^n) : (m_1(x_1^n), x_2^n) \in \mathcal{A}_n^c} p_0(x_1^n, x_2^n)$$

and the type II error probability is

$$P_1(\mathcal{A}_n) = \sum_{(x_1^n, x_2^n) : (m_1(x_1^n), x_2^n) \in \mathcal{A}_n} p_1(x_1^n, x_2^n).$$

For $\epsilon \in (0, 1)$, define the optimal type II error probability as

$$\beta_n^*(R_1, \epsilon) := \min P_1(\mathcal{A}_n),$$

where the minimum is over all $(2^{nR_1}, n)$ tests such that $P_0(\mathcal{A}_n^c) \leq \epsilon$. Further define the optimal type II error exponent as

$$\theta_1(R_1, \epsilon) := \lim_{n \rightarrow \infty} -\frac{1}{n} \log \beta_n^*(R_1, \epsilon).$$

Theorem 1 (Ahlsvede and Csiszár [2]): For every $\epsilon \in (0, 1)$,

$$\theta_1(R_1, \epsilon) = \max_{p(u_1|x_1) : R_1 \geq I(U_1; X_1)} I(U_1; X_2), \quad (1)$$

where the cardinality bound for U_1 is $|\mathcal{U}_1| \leq |\mathcal{X}_1| + 1$.

We illustrate the theorem with the following.

Example 1 ([8]): Consider the following *forward* Z binary sources (X_1, X_2) depicted in Fig. 3(a), where X_2 is the output of X_1 through a Z channel and

$$\begin{aligned} p_{X_1, X_2}(0, 0) &= 1/2, & p_{X_1, X_2}(0, 1) &= 0, \\ p_{X_1, X_2}(1, 0) &= 1/4, & p_{X_1, X_2}(1, 1) &= 1/4. \end{aligned}$$

The entire curve of the optimal type II error exponent, denoted by $\theta_1^{\rightarrow}(R_1, \epsilon)$, is plotted in Fig. 3(c).

Example 2: Now consider the following *backward* Z binary sources (X_1, X_2) depicted in Fig. 3(b), where X_1 is the output

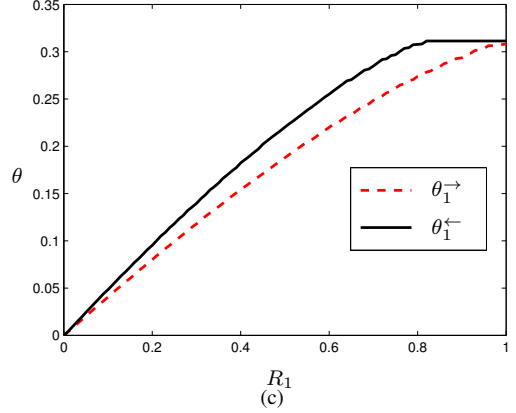
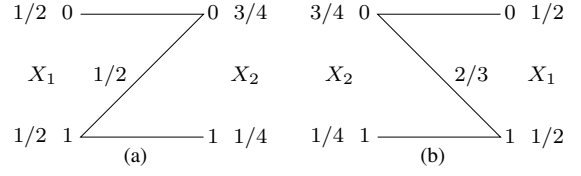


Fig. 3. Comparison of error exponents for (a) forward Z binary sources and (b) backward Z binary sources. The solid black curve corresponds to $\theta_1^-(R_1, \epsilon)$ and the dotted red curve corresponds to $\theta_1^+(R_1, \epsilon)$.

of X_2 through an inverted Z channel. The entire curve of the optimal type II error exponent, denoted by $\theta_1^{\leftarrow}(R_1, \epsilon)$ this time, is plotted in Fig. 3(c). Observe that for every $R_1 \in (0, 1)$,

$$\theta_1^{\leftarrow}(R_1, \epsilon) > \theta_1^{\rightarrow}(R_1, \epsilon). \quad (2)$$

III. INTERACTIVE HYPOTHESIS TESTING WITH COMMUNICATION CONSTRAINTS

Suppose now that instead of making an immediate decision based on one round of communication, the two nodes can interactively communicate over a noiseless bidirectional link before one of the nodes performs hypothesis testing. We wish to characterize the optimal tradeoff between the communication rates and the performance of hypothesis testing.

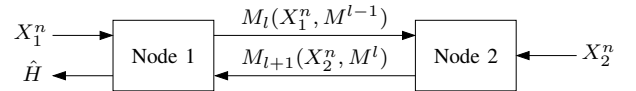


Fig. 4. Interactive hypothesis testing with communication constraints.

As before, we consider testing against independence. Assume without loss of generality that node 1 sends the first index and that the number of rounds of communication q is even. A $(2^{nR_1}, \dots, 2^{nR_q}, n)$ hypothesis test consists of

- two encoders, one for each node, where in round $l_j \in \{j, j+2, \dots, q-2+j\}$, encoder $j \in \{1, 2\}$ sends an index $m_{l_j}(x_j^n, m^{l_j-1}) \in [1 : 2^{nR_{l_j}}]$, that is, a function of its sequence and all previously transmitted indices, and
- a tester that assigns $\hat{h}(m^q, x_1^n) \in \{H_0, H_1\}$ to each $(m^q, x_1^n) \in [1 : 2^{nR_1}] \times \dots \times [1 : 2^{nR_q}] \times \mathcal{X}_1^n$.

The type I and II error probabilities are defined similarly as in the one-way case. In particular, the optimal type II error exponent is

$$\theta_q(R_1, \dots, R_q, \epsilon) := \lim_{n \rightarrow \infty} -\frac{1}{n} \log \beta_n^*(R_1, \dots, R_q, \epsilon).$$

We establish the optimal tradeoff between the rate constraints and the testing performance by characterizing $\theta_q(R_1, \dots, R_q, \epsilon)$ in the limit.

Theorem 2:

$$\lim_{\epsilon \rightarrow 0} \theta_q(R_1, \dots, R_q, \epsilon) = \max \sum_{l=1}^q I(U_l; X_{j_l} | U^{l-1}), \quad (3)$$

where the maximum is over all $\prod_{l=1}^q p(u_l | u^{l-1}, x_{j_l})$ with $|\mathcal{U}_l| \leq |\mathcal{X}_{j_l}| \cdot \prod_{j=1}^{l-1} |\mathcal{U}_j| + 1$ such that

$$R_l \geq I(U_l; X_{j_l} | U^{l-1})$$

for $l \in [1 : q]$ and $j_l = 1$ if l is odd and $j_l = 2$ if l is even.

Remark 1: By setting $U_l = \emptyset$ and $R_l = 0$ for $l = 2, \dots, q$, Theorem 2 recovers the optimal one-way type II error exponent in Theorem 1.

Remark 2: We recover the following result [8] for $q = 2$:

$$\lim_{\epsilon \rightarrow 0} \theta_2(R_1, R_2, \epsilon) = \max(I(U_1; X_2) + I(U_2; X_1 | U_1)),$$

where the maximum is over all $p(u_1 | x_1)p(u_2 | u_1, x_2)$ with $|\mathcal{U}_1| \leq |\mathcal{X}_1| + 1$ and $|\mathcal{U}_2| \leq |\mathcal{X}_2| \cdot |\mathcal{U}_1| + 1$ such that

$$\begin{aligned} R_1 &\geq I(U_1; X_1), \\ R_2 &\geq I(U_2; X_2 | U_1). \end{aligned}$$

Remark 3: We can express the optimal tradeoff between communication constraints and the type II error exponent by the *rate–exponent region* that consists of all rate–exponent tuples $(R_1, \dots, R_q, \theta)$ such that

$$\begin{aligned} R_l &\geq I(U_l; X_{j_l} | U^{l-1}), \quad l \in [1 : q], \\ \theta &\leq \sum_{l=1}^q I(U_l; X_{j_l} | U^{l-1}) \end{aligned}$$

for some pmfs $\prod_{l=1}^q p(u_l | u^{l-1}, x_{j_l})$.

Example 3 (Interaction helps): This example is motivated by [12]. We revisit the Z binary sources in Examples 1 and 2. Recall that $\theta_1^{\rightarrow}(R_1, \epsilon)$ and $\theta_1^{\leftarrow}(R_1, \epsilon)$ denote the optimal type II error exponents for the forward and backward Z binary sources, respectively. Now consider the following *double Z binary sources* as depicted in Fig. 5, where (X_1, X_2) is independent of (Y_1, Y_2) . Let

$$\theta_2(R, \epsilon) := \max_{R_1, R_2: R_1 + R_2 = R} \theta_2(R_1, R_2, \epsilon).$$

It can be easily verified that if $R \in (R^*, 2R^*)$, where $R^* = \min\{R : \theta^{\leftarrow}(R, \epsilon) = I(X_1; X_2) = 0.3113\}$, then

$$\theta_2(R, \epsilon) \geq 2\theta^{\leftarrow}(R/2, \epsilon),$$

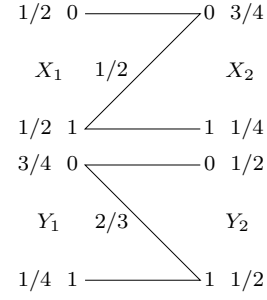


Fig. 5. Double Z binary sources.

while

$$\begin{aligned} \theta_1(R, \epsilon) &= \theta^{\leftarrow}(R^*, \epsilon) + \theta^{\rightarrow}(R - R^*, \epsilon) \\ &\stackrel{(a)}{<} \theta^{\leftarrow}(R^*, \epsilon) + \theta^{\leftarrow}(R - R^*, \epsilon) \\ &\stackrel{(b)}{\leq} 2\theta^{\leftarrow}(R/2, \epsilon), \end{aligned}$$

where (a) follows by (2) and (b) follows by the concavity of $\theta_1^{\leftarrow}(R)$ over $[0, R^*]$ (see, for example, [2, Lemma 1]). For example, when $R = 3/2$, we have $\theta_1(3/2, \epsilon) \approx 0.5548$ and $\theta_2(3/2, \epsilon) \geq 0.5934$. Thus there is strict improvement by using interaction.

In the following we prove the converse for Theorem 2. The proof of achievability is similar for the two-round case [8] and omitted for brevity.

Proof of the Converse for Theorem 2: Consider q is odd and let $j_l = 1$ if l is odd and $j_l = 2$ if l is even. Given a $(2^{nR_1}, \dots, 2^{nR_q}, n)$ test characterized by the encoding functions m_l , $l = 1, \dots, q$, and the acceptance region \mathcal{A}_n , we can apply the data processing inequality for relative entropy as for the two-round case [8] to have the multiletter upper bound

$$\lim_{\epsilon \rightarrow 0} \theta(R_1, \dots, R_q, \epsilon) \leq \lim_{n \rightarrow \infty} \frac{1}{n} D(p_0(x_1^n, m^q) || p_1(x_1^n, m^q)),$$

where

$$\begin{aligned} p_0(x_1^n, m^q) &:= \sum_{x_2^n} p_0(x_1^n, x_2^n) \prod_{l=1}^q p(m_l | m^{l-1}, x_{j_l}^n) \\ &= \left(p_0(x_1^n) \prod_{\substack{l=1 \\ l \text{ odd}}}^q p(m_l | m^{l-1}, x_1^n) \right) \\ &\quad \cdot \sum_{x_2^n} \left(p_0(x_2^n | x_1^n) \prod_{\substack{l=1 \\ l \text{ even}}}^q p(m_l | m^{l-1}, x_2^n) \right) \end{aligned}$$

and

$$\begin{aligned} p_1(x_1^n, m^q) &:= \sum_{x_2^n} p_0(x_1^n) p_0(x_2^n) \prod_{l=1}^q p(m_l | m^{l-1}, x_{j_l}^n) \\ &= \left(p_0(x_1^n) \prod_{\substack{l=1 \\ l \text{ odd}}}^q p(m_l | m^{l-1}, x_1^n) \right) \\ &\quad \cdot \sum_{x_2^n} \left(p_0(x_2^n) \prod_{\substack{l=1 \\ l \text{ even}}}^q p(m_l | m^{l-1}, x_2^n) \right). \end{aligned}$$

It is easy to verify that

$$\begin{aligned}
& \prod_{\substack{l=1 \\ l \text{ even}}}^q p_0(m_l | m^{l-1}, x_1^n) \\
&= \sum_{x_2^n} \left(p_0(x_2^n | x_1^n) \prod_{\substack{l=1 \\ l \text{ even}}}^q p(m_l | m^{l-1}, x_2^n) \right), \\
& \prod_{\substack{l=1 \\ l \text{ even}}}^q p_1(m_l | m^{l-1}, x_1^n) \\
&= \sum_{x_2^n} \left(p_0(x_2^n) \prod_{\substack{l=1 \\ l \text{ even}}}^q p(m_l | m^{l-1}, x_2^n) \right).
\end{aligned}$$

We now prove that

$$D(p_0(x_1^n, m^q) || p_1(x_1^n, m^q)) \leq \sum_{l=1}^q I(M_l; X_{j_{l+1}}^n | M^{l-1}). \quad (4)$$

To show this, we expand the relative entropy term in (4) as (5) at the bottom of this page. The second term in (5) can be upper bounded as (6) at the bottom of the page, which establishes (4).

To complete the proof, we single-letterize the upper bound in (4) as

$$\begin{aligned}
& I(M_l; X_1^n | M^{l-1}) \\
&= \sum_{i=1}^n I(M_l; X_{1i} | M^{l-1}, X_1^{i-1})
\end{aligned}$$

$$\begin{aligned}
&= \sum_{i=1}^n I(M_l; X_{1i} | M^{l-1}, X_1^{i-1}, X_2^{i-1}) \\
&\quad + I(M_l; X_2^{i-1} | M^{l-1}, X_1^{i-1}) - I(M_l; X_2^{i-1} | M^{l-1}, X_1^i) \\
&\stackrel{(a)}{\leq} \sum_{i=1}^n I(M_l; X_{1i} | M^{l-1}, X_1^{i-1}, X_2^{i-1}),
\end{aligned}$$

where (a) follows from

$$\begin{aligned}
& I(M_l; X_2^{i-1} | M^{l-1}, X_1^{i-1}) - I(M_l; X_2^{i-1} | M^{l-1}, X_1^i) \\
&= H(X_2^{i-1} | M^{l-1}, X_1^{i-1}) - H(X_2^{i-1} | M^l, X_1^{i-1}) \\
&\quad - H(X_2^{i-1} | M^{l-1}, X_1^i) + H(X_2^{i-1} | M^l, X_1^i) \\
&= I(X_2^{i-1}; X_{1i} | M^{l-1}, X_1^{i-1}) - I(X_2^{i-1}; X_{1i} | M^l, X_1^{i-1}) \\
&\leq 0.
\end{aligned}$$

To bound the rate constraints, consider for l even,

$$\begin{aligned}
nR_l &\geq H(M_l) \\
&\geq I(M_l; X_1^n, X_2^n | M^{l-1}) \\
&= \sum_{i=1}^n I(M_l; X_{1i}, X_{2i} | M^{l-1}, X_1^{i-1}, X_2^{i-1}) \\
&\geq \sum_{i=1}^n I(M_l; X_{2i} | M^{l-1}, X_1^{i-1}, X_2^{i-1}).
\end{aligned}$$

When $l > 1$ is odd, the rate constraints and the terms in (4) can be bounded similarly. The case of $l = 1$ needs to be

$$\begin{aligned}
& D(p_0(x_1^n, m^q) || p_1(x_1^n, m^q)) \\
&= \sum_{x_1^n, m^q} p_0(x_1^n, m^q) \log \frac{p_0(m_q | m^{q-1}, x_1^n) p_0(m_{q-2} | m^{q-3}, x_1^n) \cdots p_0(m_1 | x_1^n)}{p_1(m_q | m^{q-1}) p_1(m_{q-2} | m^{q-3}) \cdots p_1(m_1)} \\
&= \sum_{x_1^n, m^q} p_0(x_1^n, m^q) \log \left(\frac{p_0(m_q | m^{q-1}, x_1^n)}{p_0(m_q | m^{q-1})} \frac{p_0(m_{q-2} | m^{q-3}, x_1^n)}{p_1(m_{q-2} | m^{q-3})} \cdots \frac{p_0(m_1 | x_1^n)}{p_0(m_1)} \frac{p_0(m_1)}{p_1(m_1)} \right) \\
&= \sum_{\substack{l=1 \\ l \text{ even}}}^q I(M_l; X_1^n | M^{l-1}) + \sum_{\substack{m^l \\ l \text{ even}}} p_0(m^l) \log \frac{p_0(m_l | m^{l-1})}{p_1(m_l | m^{l-1})}, \tag{5}
\end{aligned}$$

$$\begin{aligned}
& \sum_{m^l} p_0(m^l) \log \frac{p_0(m_l | m^{l-1})}{p_1(m_l | m^{l-1})} \\
&= \sum_{m^{l-2}} p_0(m^{l-2}) \sum_{m_l, m_{l-1}} p_0(m_l, m_{l-1} | m^{l-2}) \log \frac{p_0(m_l | m^{l-1})}{p_1(m_l | m^{l-1})} \\
&= \sum_{m^{l-2}} p_0(m^{l-2}) D(p_0(m_l | m^{l-1}) p_0(m_{l-1} | m^{l-2}) || p_1(m_l | m^{l-1}) p_0(m_{l-1} | m^{l-2})) \\
&= \sum_{m^{l-2}} p_0(m^{l-2}) D(p_0(m_{l-1} | m^{l-2}) \sum_{x_2^n} p_0(m_l | m^{l-1}, x_2^n) p_0(x_2^n | m^{l-1}) || p_0(m_{l-1} | m^{l-2}) \sum_{x_2^n} p_0(m_l | m^{l-1}, x_2^n) p_0(x_2^n | m^{l-2})) \\
&\leq \sum_{m^{l-2}} p_0(m^{l-2}) D(p_0(m_{l-1} | m^{l-2}) p_0(m_l | m^{l-1}, x_2^n) p_0(x_2^n | m^{l-1}) || p_0(m_{l-1} | m^{l-2}) p_0(m_l | m^{l-1}, x_2^n) p_0(x_2^n | m^{l-2})) \\
&= I(M_{l-1}; X_2^n | M^{l-2}). \tag{6}
\end{aligned}$$

considered separately and it can be easily verified that

$$nR_1 \geq \sum_{i=1}^n I(M_1, X_1^{i-1}, X_2^{i-1}; X_{2i})$$

$$I(M_1; X_1^n) \leq \sum_{i=1}^n I(M_1, X_2^{i-1}, X_1^{i-1}; X_{1i}).$$

Identify $U_{1i} = (M_1, X_1^{i-1}, X_2^{i-1})$ and $U_{li} = M_l$ for $l \geq 2$. Define the time-sharing random variable Q to be uniformly distributed over $[1 : n]$ and independent of (M^q, X_1^n, X_2^n) , and let $U_l = (Q, U_{lQ})$, $X_1 = X_{1Q}$, and $X_2 = X_{2Q}$. Clearly, $U_l \rightarrow (U^{l-1}, X_{j_l}) \rightarrow X_{j_{l+1}}$ form Markov chains. Finally, the cardinality bounds on U_l follow the standard technique, in particular, the one used in the 2-round interactive lossy source coding problem [9]. This completes the converse proof. ■

IV. RELATIONSHIP TO INTERACTIVE LOSSY COMPRESSION

In this section, we compare the two-round interactive hypothesis testing problem with the two-round interactive lossy source coding problem. Consider the interactive lossy source coding problem depicted in Fig. 6. Here two nodes interactively communicate with each other so that each node can reconstruct the source observed by the other node with prescribed distortions. Kaspi [9] established the optimal tradeoff between communication constraints and the distortion pair (D_1, D_2) . (See also Ma and Ishwar [13] for an ingenious example demonstrating that interactive lossy compression can strictly outperform one-way lossy compression.)

The optimal tradeoff between communication and distortion is characterized by the rate-distortion region, the formal definition of which can be found in [9] or [10, Section 20.3].

Theorem 3 (Kaspi [9]): The two-round rate-distortion region is the set of all rate pairs (R_1, R_2) such that

$$R_1 \geq I(X_1; U^q | X_2),$$

$$R_2 \geq I(X_2; U^q | X_1)$$

for some conditional pmf $\prod_{l=1}^q p(u_l | u^{l-1}, x_{j_l})$ with $|U_l| \leq |\mathcal{X}_{j_l}| \cdot (\prod_{j=1}^l |U_j|) + 1$ and functions $\hat{x}_1(u^q, x_2)$ and $\hat{x}_2(u^q, x_1)$ that satisfy $E(d_j(X_j, \hat{X}_j)) \leq D_j$, $j = 1, 2$, where $j_l = 1$ if l is odd and $j_l = 2$ if l is even.

Achievability is established by performing Wyner-Ziv coding [14] in each round, i.e., joint typicality encoding followed by binning. By contrast, the scheme we used for the interactive hypothesis testing problem is joint typicality encoding in each round (without binning). It turns out, however, that this distinction between binning and no binning is not fundamental. By using Wyner-Ziv coding in the interactive hypothesis testing

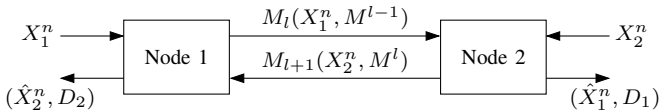


Fig. 6. Interactive lossy compression.

problem, we can establish the following tradeoff between communication constraints and the testing performance.

Proposition 1: The rate-exponent region for q -round interactive hypothesis testing is the set of rate-exponent triples $(R_1, \dots, R_q, \theta)$ such that

$$\theta \leq \sum_{l=1}^q I(U_l; X_{j_l} | U^{l-1}),$$

$$R_l \geq I(U_l; X_{j_l} | U^{l-1}) - I(U_l; X_{j_{l+1}} | U^{l-1}),$$

$$l \in [1 : q],$$

$$\sum_{l=1}^q R_l - \theta \geq \sum_{l=1}^q (I(U_l; X_{j_l} | U^{l-1}) - I(U_l; X_{j_l} | U^{l-1}))$$

for some $\prod_{l=1}^q p(u_l | u^{l-1}, x_{j_l})$, where $j_l = 1$ if l is odd and $j_l = 2$ if l is even.

It can be shown that the rate-exponent region in Proposition 1 is equivalent to the region in Remark 3. As pointed out by Rahman and Wagner [7] in the one-way setup, binning never hurts. Therefore, the coding scheme for q -round interactive lossy source coding leads to an essentially identical scheme for q -round interactive hypothesis testing. It is refreshing to note that the same scheme is optimal for both problems.

REFERENCES

- [1] T. Berger, "Decentralized estimation and decision theory," in *Proc. IEEE Inf. Theory Workshop*, Mt. Kisco, NY, Sep. 1979.
- [2] R. Ahlswede and I. Csiszár, "Hypothesis testing with communication constraints," *IEEE Trans. Inf. Theory*, vol. 32, no. 4, pp. 533–542, 1986.
- [3] T. S. Han, "Hypothesis testing with multiterminal data compression," *IEEE Trans. Inf. Theory*, vol. 33, no. 6, pp. 759–772, 1987.
- [4] H. Shimokawa, T. S. Han, and S. Amari, "Error bound of hypothesis testing with data compression," in *Proc. IEEE Internat. Symp. Inf. Theory*, Jun. 1994, p. 29.
- [5] T. S. Han and S. Amari, "Statistical inference under multiterminal data compression," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2300–2324, Oct. 1998.
- [6] C. Tian and J. Chen, "Successive refinement for hypothesis testing and lossless one-helper problem," *IEEE Trans. Inf. Theory*, vol. 54, no. 10, pp. 4666–4681, Oct. 2008.
- [7] M. S. Rahman and A. Wagner, "On the optimality of binning for distributed hypothesis testing," *IEEE Trans. Inf. Theory*, vol. 58, no. 10, pp. 6282–6303, Oct. 2012.
- [8] Y. Xiang and Y.-H. Kim, "Interactive hypothesis testing with communication constraints," in *Proc. 50th Ann. Allerton Conf. Commun. Control Comput.*, Sep. 2012.
- [9] A. H. Kaspi, "Two-way source coding with a fidelity criterion," *IEEE Trans. Inf. Theory*, vol. 31, no. 6, pp. 735–740, 1985.
- [10] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge: Cambridge University Press, 2011.
- [11] A. Orlitsky and J. R. Roche, "Coding for computing," *IEEE Trans. Inf. Theory*, vol. 47, no. 3, pp. 903–917, 2001.
- [12] A. A. Gohari and V. Anantharam, "Information-theoretic key agreement of multiple terminals—I: Source model," *IEEE Trans. Inf. Theory*, vol. 56, no. 8, pp. 3973–3996, Aug. 2010.
- [13] N. Ma and P. Ishwar, "Some results on distributed source coding for interactive function computation," *IEEE Trans. Inf. Theory*, vol. 57, no. 9, pp. 6180–6195, Sep. 2011.
- [14] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inf. Theory*, vol. 22, no. 1, pp. 1–10, 1976.