Capacity of a Class of Modulo-Sum Relay Channels

Marko Aleksic, Peyman Razaghi, *Student Member, IEEE*, and Wei Yu, *Senior Member, IEEE*

***Abstract*—This paper characterizes the capacity of a class of modular additive noise relay channels, in which the relay observes a corrupted version of the noise and has a separate channel to the destination. The capacity is shown to be strictly below the cut-set bound in general and achievable using a quantize-and-forward strategy at the relay. This result confirms a previous conjecture on the capacity of channels with rate-limited side information at the receiver for this particular class of modulo-sum channels. This paper also considers a more general setting in which the relay is capable of conveying noncausal rate-limited side information about the noise to both the transmitter and the receiver. The capacity is characterized for the case where the channel is binary symmetric with a crossover probability $\frac{1}{2}$. In this case, the rates available for conveying side information to the transmitter and to the receiver can be traded with each other arbitrarily—the capacity is a function of the sum of the two rates.**

***Index Terms*—Channel with side information, cut-set bound, modulo-sum channel, quantize-and-forward, relay channel.**

I. INTRODUCTION

THE relay channel is a fundamental building block in network information theory. Although the capacity of the general relay channel is not yet known, the capacities of many specific classes of relay channels have been found. These special classes include the degraded and reversely degraded [1], the semideterministic [2], a class of orthogonal [3], a class of deterministic [4] relay channels, as well as specific examples of Gaussian relay channels with phase fading [5] and Gaussian orthogonal relay channels [6]. All the above relay channels for which capacities are characterized share one thing in common: they achieve their respective cut-set bounds. This makes converses straightforward. Unfortunately it appears that the cut-set bound cannot be achieved for many practical relay channels. Efforts to find different bounds, or to prove the looseness of the cut-set bound have proved to be quite difficult. Zhang's partial converse [7] demonstrated the latter for a relay channel with a noiseless relay link (without characterizing its actual capacity); Zahedi [8] provided some justifications for why the cut-set bound cannot be tight in all cases.

Manuscript received June 13, 2007; revised October 15, 2008. Current version published February 25, 2009. This work was supported by the Natural Science and Engineering Research Council (NSERC) of Canada under a discovery grant and under the Canada Research Chairs program. The material in this paper was presented in part at the IEEE International Symposium on Information Theory, Nice, France, June 2007.

The authors are with the Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON M5S 3G4, Canada (e-mail: marko.aleksic@utoronto.ca; peyman@comm.utoronto.ca; weiyu@comm.utoronto.ca).

Communicated by G. Kramer, Associate Editor for Shannon Theory.

Color version of Figure 3 in this paper is available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2008.2011518

In the first part of this paper, we find the capacity of a nontrivial class of modulo-sum relay channels. In these channels, the relay observes a correlated version of the noise between the source and the destination, and has a dedicated channel to the destination. We show that the capacity can be strictly below the cut-set bound, and is achievable by a quantize-and-forward strategy (also known as estimate-and-forward and compress-and-forward) due to Cover and El Gamal [1, Theorem 6]. The quantize-and-forward strategy was previously only known to achieve the cut-set bound capacities of one class of deterministic relay channels [4] and the trivial case where the relay link can support a sufficiently large rate to allow for a complete description of the relay's observation at the destination. The modulo-sum relay channel appears to be a first example of a channel for which this strategy achieves a capacity strictly below the cut-set bound.

The quantize-and-forward strategy was designed for use in channels where the relay has a poor-quality channel from the source. In this strategy, the relay quantizes its received signal, and transmits the quantized signal to the destination. The destination first decodes the quantized signal from the relay, then uses this signal to help decode the source message. The destination may also use its own received signal to help the decoding of the quantized signal from the relay, because the two signals may be correlated in a general relay channel. This technique is known as Wyner–Ziv coding. Quantize-and-forward is a natural strategy for the modular channel considered in this paper where the relay observes only the noise. This is because there is no message for the relay to decode; all the relay can do is to describe the noise to the destination.

The converse result for this class of relay channels depends crucially on two properties of modulo-sum channels. In these channels, a uniform distribution on the input alphabet achieves the maximum possible entropy of the output, regardless of the statistics of the additive noise. Further, under a uniform input distribution, the output of a modulo-sum channel is also independent of the additive noise. This has the consequence of simplifying the converse: the side information in Wyner–Ziv coding is not useful since the destination's observation is independent of the relay's output.

In the second part of this paper, we explore the connection between the class of relay channels considered above and the class of channels with rate-limited noncausal state information available to the transmitter and the receiver. In a channel with noncausal side information, the channel transition probability depends on a state variable S, which is assumed to be observed *noncausally* by a relay. The relay encodes the entire sequence S^n under rate constraints R_1 to the transmitter and R_2 to the receiver, respectively, in order to facilitate the communication between the two. The general channel model is depicted in Fig. 1.

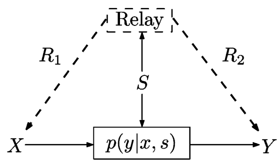

Fig. 1. Channel with rate-limited side information.

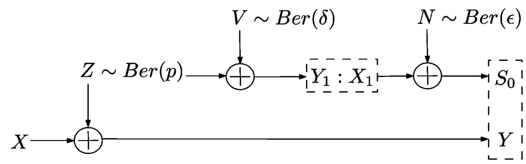

Fig. 2. A binary relay channel where the relay observes a corrupted version of the channel noise.

Channels with side information have been studied extensively in the literature. Gel'fand and Pinsker [9] derived the capacity for the case in which complete side information is available at the transmitter but no side information is available at the receiver (i.e., $R_1 = H(S)$, $R_2 = 0$). Ahlswede and Han [10] gave an achievable rate for the case where $R_1 = 0$ but R_2 can be arbitrary, and conjectured that their expression is indeed the capacity. Their achievability result was further generalized by Heegard and El Gamal [11, Theorem 1] for the case where R_1 and R_2 can both be arbitrary. Further, for the special case in which R_1 is arbitrary, but R_2 is sufficiently large to allow for a complete description of the state at receiver, the capacity was characterized by Rosenzweig, Steinberg, and Shamai [12]. In general, the capacities of channels with arbitrary R_1 and R_2 are still open.

A relay channel where the relay only gets to observe some possibly stochastic function of the noise and has a dedicated finite-capacity channel to the destination can be viewed as a channel with rate-limited state information available to the receiver and with no side information at the transmitter (i.e., $R_1 = 0$). The capacity result for modulo-sum relay channels presented in the first part of the paper confirms the conjecture by Ahlswede and Han [10] for this special class of modulo-sum channels.

The converse used in the first part of this paper also allows us to make progress on the capacity of channels with rate-limited side information to both the transmitter and the receiver (i.e., with arbitrary R_1 and R_2). In particular, we characterize the capacity for a case where the channel between the transmitter and receiver is binary symmetric with a Bernoulli-$\frac{1}{2}$ noise distribution and the relay observes the channel noise through another binary symmetric channel (BSC) with an arbitrary crossover probability. The capacity coincides with the achievable rate provided by [11, Theorem 1]. Further, the resulting capacity expression in this particular case has the property that it is only a function of the sum of R_1 and R_2. Thus, the rates available for the transfer of state information to the transmitter and to the receiver not only are interchangeable, but also can be traded with each other arbitrarily. This appears to be a first example of a nontrivial channel with rate-limited state information available to both the transmitter and receiver for which the capacity is characterized.

The rest of the paper is organized as follows. Section II characterizes the capacity of a binary symmetric relay channel and shows that the capacity can be strictly below the cut-set bound. The extension to the modulo-sum case is presented in Section III. The connection with channel with rate-limited side information is presented in Section IV. The paper concludes with Section V.

II. A Binary Symmetric Relay Channel

We begin by deriving the capacity of a particular binary symmetric relay channel. The derivation will be directly applicable to a broader class of modulo-sum relay channels. The simple binary symmetric case is used to distil the essential steps and ideas.

Consider the relay channel as shown in Fig. 2. Here, the channel input X goes through a memoryless BSC with crossover probability p to reach Y, i.e., $Y = X + Z \pmod 2$ with Z being a Bernoulli-p random variable denoted as $\mathrm{Ber}(p)$ (i.e., $Z = 1$ with probability p and $Z = 0$ with probability $1 - p$). The relay gets to observe a noisy version of Z through a memoryless BSC with crossover probability δ, i.e., $Y_1 = Z + V \pmod 2$, where $V \sim Ber(\delta)$. The relay also has a separate memoryless BSC to the destination $S_0 = X_1 + N \pmod 2$, where $N \sim \mathrm{Ber}(\epsilon)$.

Let us define

$$R_0 = \max_{p(x_1)} \quad I(X_1; S_0) \tag{1}$$

for future reference. If there were no corrupting variable V, then the capacity of this channel is as recently characterized by Kim [4]

$$C = \max_{p(x)} \min\{I(X;Y) + R_0, I(X;Y,Y_1)\}. \tag{2}$$

Both hash-and-forward [4], a strategy where the relay simply hashes Y_1 into equal sized bins, and the classic quantize-and-forward are capacity achieving. The multiple-access cut-set bound is $I(X;Y) + R_0$, which is obtained by assuming that the relay already knew the message the source would like to transmit. One way to interpret the achievability of the multiple-access cut-set bound is that if V were absent, decoding X would be the same as decoding Z. So, the relay, by sending parity information about Z, can be interpreted to be performing a version of decode-and-forward, as if it already knew the message; random parities for Z turn into random parities for X. This interpretation would fail if the relay's observation of Z is corrupted by V. To the best of the authors' knowledge, the capacity of this class of relay channels when V is present has not been characterized previously.

The following is a reasonable strategy for this channel. The relay tries to quantize Y_1 in such a way as to minimize the uncertainty about Z at the destination. The main result of this section is that the above approach is capacity achieving for a class of modulo-sum relay channels including the channel in Fig. 2.

Theorem 1: The capacity C of the binary relay channel in Fig. 2 is

$$C = \max_{p(u|y_1): I(U;Y_1) \le R_0} 1 - H(Z|U) \tag{3}$$

where the maximization may be restricted to U's with $|\mathcal{U}| \leq |\mathcal{Y}_1| + 2$, and R_0 is as defined in (1).

A. Proof of Achievability for Theorem 1

The achievability can be proved by a direct application of [1, Theorem 6] if we fix the input distribution of X as $\mathrm{Ber}(\frac{1}{2})$ and identify U with $\hat{Y}_1$ in [1]. A separate proof is provided here for completeness based on the theory of jointly strongly typical sequences [13, p. 326].

We transmit at rate R over $B-1$ blocks, each of length n. For the last block, no message is transmitted. As $B \to \infty$, $\frac{R(B-1)}{B}$ becomes arbitrarily close to R.

Codebook Generation: Generate 2^{nR} independent random sequences $X^n(w)$ of length n, $w \in \{1 \ldots 2^{nR}\}$, where each sequence is generated with an independent and identical distribution (i.i.d.) $\mathrm{Ber}(\frac{1}{2})$. Fix a $p(u|y_1)$ such that it satisfies the constraint $I(U;Y_1) < R_0$. Generate $2^{nI(U;Y_1)}$ independent random sequences $U^n(t)$ of length n, $t \in \{1 \ldots 2^{nI(U;Y_1)}\}$, where each sequence is generated with i.i.d. $p(u)$.

Encoding: We describe the encoding for block i. To send message $w_i, w_i \in \{1 \ldots 2^{nR}\}$, the transmitter simply sends $X^n(w_i)$. The relay, having observed the entire corrupted noise sequence from the previous block $Y_1^n(i-1)$, looks in its U^n codebook and finds a sequence $U^n(t_i)$ that is jointly strongly typical with $Y_1^n(i-1)$. It then encodes and sends the index t_i across the separate channel to the destination. Only the relay transmits to the destination in the last block.

Decoding: The destination, upon decoding t_i, looks for a w_{i-1} such that $X^n(w_{i-1})$ is jointly strongly typical with $U^n(t_i)$ and $Y^n(i-1)$.

Analysis of the Probability of Error: Because of the symmetry of the codebook construction we can perform the analysis assuming $X^n(1)$ was sent over all the blocks. Since the decodings of different blocks are independent, we can focus on the probability of error over the first block and drop the time indices. The error events are as follows:

- E_1: $(X^n(1), Y^n, Y_1^n)$ are not jointly strongly typical;
- E_2: $\nexists\, t$, $(U^n(t), Y_1^n)$ are jointly strongly typical;
- E_3: $(X^n(1), Y^n, U^n(t))$ are not jointly strongly typical;
- E_4: The destination makes an error decoding t sent by the relay across the separate channel;
- E_5: $\exists\, w \neq 1$, $(X^n(w), Y^n, U^n(t))$ are jointly strongly typical.

For n sufficiently large, we have $P(E_1) < \frac{\epsilon}{5B}$. Following the argument of [13, Sec. 10.6], $P(E_2 \cap E_1^c) < \frac{\epsilon}{5B}$ for sufficiently large n. By the Markov lemma [13, Lemma 15.8.1], since (X,Y)–Y_1–U forms a Markov chain, $P(E_3 \cap E_1^c \cap E_2^c) < \frac{\epsilon}{5B}$ for n sufficiently large. Since by construction $I(U;Y_1) < R_0$, the index t can be sent to the destination with an arbitrarily small probability of error, so $P(E_4) < \frac{\epsilon}{5B}$. Finally, the probability that another randomly generated $X^n(w)$ is jointly strongly typical with Y^n and $U^n(t)$ is less than $2^{-n(I(X;Y,U)-\gamma)}$ for any $\gamma > 0$ and sufficiently large n [13, Lemma 10.6.2]. Using the union bound, we have $P\left(E_5 \cap \bigcap_{i=1}^4 E_i^c\right) < 2^{nR} 2^{-n(I(X;Y,U)-\gamma)}$. Thus, when

$$R < I(X;Y,U), \tag{4}$$

we can find a sufficiently large n such that

$$P\left(E_5 \cap \bigcap_{i=1}^4 E_i^c\right) < \frac{\epsilon}{5B}.$$

Now, since X and U are independent, we have

$$\begin{aligned} I(X;Y,U) &= I(X;Y|U) && (5)\\ &= H(Y|U) - H(Z|U) && (6)\\ &= 1 - H(Z|U) && (7) \end{aligned}$$

where $H(Y|U) = 1$ because for BSCs under the uniform input distribution $\mathrm{Ber}\left(\frac{1}{2}\right)$, the output Y is independent of the additive noise Z, and hence U. Collecting terms we see that $P\left(\bigcup_{i=1}^5 E_i\right) < \frac{\epsilon}{B}$, so that using the union bound again we can make the probability of error over all of the B blocks less than ϵ as long as $R < 1 - H(Z|U)$.

B. Proof of Converse for Theorem 1

The converse depends on the following lemma.

Lemma 1: Let Z^n, V^n, N^n be independent Bernoulli i.i.d. sequences. Let $Y_1^n = Z^n + V^n$ and $S_0^n = X_1^n + N^n$, where Y_1^n and X_1^n are the input and output sequences of the relay as shown in Fig. 2. The following inequality holds for any encoding scheme at the relay:

$$H\left(Z^n | S_0^n\right) \geq \min_{p(u|y_1): I(U;Y_1) \leq R_0} nH(Z|U) \tag{8}$$

where the minimization on the right-hand side may be restricted to U's with $|\mathcal{U}| \leq |\mathcal{Y}_1| + 2$.

Proof: The proof of the lemma is closely based on the proof of Theorem 15.8.1 in [13]. Fixing an encoding scheme at the relay, our strategy is to show that there always exists a U for which $H\left(Z^n|S_0^n\right) \geq nH(Z|U)$ and $I(Y_1;U) \leq R_0$. This would allow us to conclude that

$$H\left(Z^n | S_0^n\right) \geq \min_{p(u|y_1): I(U;Y_1) \leq R_0} nH(Z|U).$$

We start by finding a lower bound for $H\left(Z^n|S_0^n\right)$

$$\begin{aligned} H\left(Z^n|S_0^n\right) &= \sum_{i=1}^n H\left(Z_i | S_0^n, Z_1, \ldots, Z_{i-1}\right) && (9)\\ &\geq \sum_{i=1}^n H\left(Z_i | S_0^n, Z^{i-1}, Y_1^{i-1}\right) && (10)\\ &= \sum_{i=1}^n H\left(Z_i | S_0^n, Y_1^{i-1}\right) && (11) \end{aligned}$$

where the third line follows from the fact that $S_0^n Y_1^{i-1} Z^{i-1} - S_0^n Y_1^{i-1} - Z_i$ forms a Markov chain. The Markov chain follows because Z_i's are i.i.d., the channel from Z to Y_1 is memoryless, and S_0^n is only a function of Y_1^n, hence Z_i can only be affected by Z^{i-1} through S_0^n and Y_1^{i-1}. Now define $U_i = \left(S_0^n, Y_1^{i-1}\right)$, we get

$$H\left(Z^n|S_0^n\right) \geq \sum_{i=1}^{n} H(Z_i|U_i). \tag{12}$$

Next, note that $Z^n - Y_1^n - X_1^n - S_0^n$ forms a Markov chain. As a result

$$\begin{aligned} I\left(X_1^n; S_0^n\right) &\geq I\left(Y_1^n; S_0^n\right) & (13)\\ &= \sum_{i=1}^{n} I\left(Y_{1i}; S_0^n | Y_{11}, \ldots, Y_{1(i-1)}\right) & (14)\\ &= \sum_{i=1}^{n} I\left(Y_{1i}; S_0^n, Y_1^{i-1}\right) & (15)\end{aligned}$$

where (13) follows from the data processing inequality, and (15) follows from the fact that Y_{1i} is independent of Y_1^{i-1} and consequently $I\left(Y_{1i}; Y_1^{i-1}\right) = 0$. Using the definition of U, we get

$$I\left(X_1^n; S_0^n\right) \geq \sum_{i=1}^{n} I(Y_{1i}; U_i). \tag{16}$$

Recall that $R_0 = \max_{p(x_1)} I(X_1; S_0)$. Since the channel between X_1 and S_0 is memoryless, by [13, Lemma 7.9.2], $R_0 \geq \frac{1}{n} I\left(X_1^n; S_0^n\right)$. Thus, we have shown the following inequalities:

$$R_0 \geq \frac{1}{n} \sum_{i=1}^{n} I(Y_{1i}; U_i) \tag{17}$$

$$\frac{1}{n} H\left(Z^n|S_0^n\right) \geq \frac{1}{n} \sum_{i=1}^{n} H(Z_i|U_i). \tag{18}$$

Introducing an independent timesharing random variable Q uniformly distributed over $\{1, \ldots, n\}$ (see [13, Theorem 15.8.1]), the above equations can be rewritten as

$$R_0 \geq \frac{1}{n} \sum_{i=1}^{n} I(Y_{1i}; U_i | Q=i) = I(Y_{1Q}; U_Q | Q) \tag{19}$$

$$\frac{1}{n} H\left(Z^n|S_0^n\right) \geq \frac{1}{n} \sum_{i=1}^{n} H(Z_i|U_i, Q=i) = H(Z_Q|U_Q, Q). \tag{20}$$

Now since Q is independent of Y_{1Q} (as the distribution of Y_{1i} does not depend on i), we have

$$\begin{aligned} I(Y_{1Q}; U_Q|Q) &= I(Y_{1Q}; U_Q, Q) - I(Y_{1Q}; Q)\\ &= I(Y_{1Q}; U_Q, Q). & (21)\end{aligned}$$

Finally, Y_{1Q} and Z_Q have the same joint distribution as Y_1 and Z, so defining $U = (U_Q, Q)$, $Z = Z_Q$, and $Y_1 = Y_{1Q}$, we have shown the existence of a random variable U such that

$$R_0 \geq I(Y_1; U) \tag{22}$$

$$H\left(Z^n|S_0^n\right) \geq nH(Z|U) \tag{23}$$

for any encoding scheme at the relay. Since for every possible encoding scheme at the relay we can construct an i.i.d. U satisfying the above equations, the minimum over all U's satisfying $I(U; Y) \leq R_0$ must satisfy (8). The cardinality bound is the same as in [13, Theorem 15.8.1]. □

The converse can now be proved as follows. Let $W \in \{1, 2, \ldots, 2^{nR}\}$ be the source message.

$$\begin{aligned} nR &= H(W) & (24)\\ &= I\left(W; Y^n, S_0^n\right) + H\left(W|Y^n, S_0^n\right) & (25)\\ &\overset{(a)}{\leq} I\left(W; Y^n, S_0^n\right) + n\epsilon_n & (26)\\ &\leq I\left(X^n; Y^n, S_0^n\right) + n\epsilon_n & (27)\\ &\overset{(b)}{=} I\left(X^n; Y^n|S_0^n\right) + n\epsilon_n & (28)\\ &= H\left(Y^n|S_0^n\right) - H\left(Y^n|S_0^n, X^n\right) + n\epsilon_n & (29)\\ &\overset{(c)}{\leq} n - H\left(Z^n|S_0^n, X^n\right) + n\epsilon_n & (30)\\ &= n - H\left(Z^n|S_0^n\right) + n\epsilon_n & (31)\\ &\overset{(d)}{\leq} \max_{p(u|y_1): I(U;Y_1) \leq R_0} n(1 - H(Z|U)) + n\epsilon_n & (32)\\ &= nC + n\epsilon_n & (33)\end{aligned}$$

where
(a) follows from Fano's inequality,
(b) follows from the fact that X^n is independent of S_0^n,
(c) follows from the fact that the maximum entropy of a binary random variable of length n is n,
(d) follows from Lemma 1.

Thus, we have shown that for any relaying scheme with a low probability of error, $R \leq C$.

C. *Comments on Theorem 1*

The capacity of the binary symmetric relay channel considered above is achieved essentially by digitizing the separate channel between the relay and destination. All that matters is that the capacity of the separate channel is sufficiently large to support the relay's description of U, the quantization variable. There is no advantage in joint source–channel coding at the relay. The input codebook for X is drawn from the uniform $\text{Ber}\left(\frac{1}{2}\right)$ distribution, identical to the capacity achieving distribution if the relay were absent; the source merely increases its rate once the relay is introduced.

There are two conditions which are important for the converse to work. The channel between the source and destination should be additive and modular. These two conditions allow for two crucial simplifications in the converse. First, a uniform input distribution maximizes the output entropy, regardless of any information that the relay may convey about the noise; this was used in (30). Second, the linear nature of the channel, combined with the expansion in (29), reduces the role of the relay to essentially source coding with a distortion metric being the conditional entropy of Z. This is in contrast to a general relay channel where the relay observes a combination of the source message and noise, so there is an opportunity for the destination to use its received signal as side information in the decoding of the relay's quantized message. For the binary symmetric relay channel, the uniform input distribution completely eliminates any aid the destination's output can provide in the decoding of the relay's message; this makes the converse easier to prove.

D. Capacity Can Be Below the Cut-Set Bound

To see that the capacity as stated in Theorem 1 can be strictly below the cut-set bound, consider the case in which Z^n has an i.i.d. $\text{Ber}\left(\frac{1}{2}\right)$ distribution. The capacity can now be evaluated as

$$C = 1 - h\left(h^{-1}(1 - R_0) * \delta\right) \tag{34}$$

where $h(p) = -p\log_2 p - (1-p)\log_2(1-p)$, $h^{-1}(\cdot)$ is the inverse of $h(p)$ in the domain $p \in [0, 0.5]$, and $\alpha * \beta = \alpha(1-\beta) + (1-\alpha)\beta$. This capacity expression follows by noting that $I(U;Y_1) = H(Y_1) - H(Y_1|U)$, so that the constraint in the maximization in (3) can be rewritten as

$$H(Y_1|U) \geq H(Y_1) - R_0. \tag{35}$$

Now we use a lemma due to Wyner and Ziv on the conditional entropy of binary random variables, known as "Mrs. Gerber's Lemma" [14], to claim that if

$$H(Y_1|U) \geq \alpha \tag{36}$$

then

$$H(Z|U) \geq h(h^{-1}(\alpha) * \delta) \tag{37}$$

with equality if and only if Y_1 given U is a $\text{Ber}(h^{-1}(\alpha))$ random variable. Wyner and Ziv's inequality holds here because when Z is $\text{Ber}\left(\frac{1}{2}\right)$, we can write $Z = Y_1 + V$, where V is $\text{Ber}(\delta)$ and Y_1 and V are independent.

Now, let $\alpha = H(Y_1) - R_0$. Observe that the U that achieves equality in (37), i.e., the U that gives rise to Y_1 given U as $\text{Ber}\left(h^{-1}(H(Y_1) - R_0)\right)$, is precisely the U that minimizes the Hamming distortion of Y_1 under a rate constraint R_0 in standard rate–distortion theory. This is because rate–distortion theory states that for binary random variables, under a rate constraint R_0, the minimum achievable average distortion ν must satisfy $h(\nu) = H(Y_1|U) = H(Y_1) - R_0$ and Y_1 given U must be $\text{Ber}(\nu)$. Further, as Y_1 is $\text{Ber}\left(\frac{1}{2}\right)$, the distribution of the optimal U is also $\text{Ber}\left(\frac{1}{2}\right)$. The capacity (34) follows by using this U in (3) and by substituting $H(Y_1) = 1$ and $\alpha = 1 - R_0$ in (37).

We now show that the capacity as given in (34) is strictly below the cut-set bound. The cut-set bound for the relay channel is [1]

$$\max_{p(x,x_1)} \min\{I(X, X_1; Y, S_0), I(X; Y, S_0, Y_1|X_1)\}. \tag{38}$$

When Z is $\text{Ber}\left(\frac{1}{2}\right)$, for the multiple-access cut we have

$$\begin{aligned} I(X, X_1; Y, S_0) &= H(Y, S_0) - H(Y, S_0|X, X_1) && (39)\\ &\leq 2 - H(Z, N|X, X_1) && (40)\\ &= 1 - H(Z) + 1 - H(N) && (41)\\ &= R_0 && (42)\end{aligned}$$

where the equality in (40) is achieved by letting X and X_1 have independent $\text{Ber}\left(\frac{1}{2}\right)$ distributions. Essentially, the source–destination channel has zero capacity when Z is $\text{Ber}\left(\frac{1}{2}\right)$; the multiple-access bound is just the capacity of the relay–destination link. Similarly, for the broadcast cut we have

$$\begin{aligned} I(X; Y, S_0, Y_1|X_1) &= I(X; Y|S_0, Y_1, X_1) && (43)\\ &= H(Y|S_0, Y_1, X_1) \\ &\quad - H(Z|S_0, Y_1, X, X_1) && (44)\\ &\leq 1 - H(V|S_0, Y_1, X, X_1) && (45)\\ &= 1 - h(\delta). && (46)\end{aligned}$$

In the first line, we use the fact that X is independent of Y_1 and S_0 given X_1. In the third line, we use the fact that the entropy of a binary random variable is bounded by 1, and for modulo-2 addition $Y_1 = Z + V$ can be equivalently written as $Z = Y_1 + V$. In the fourth line, we use the fact that Z is $\text{Ber}\left(\frac{1}{2}\right)$ and independent of V, so Y_1 is independent of V. Consequently, V is independent of both X_1 and S_0 as well. Thus, (46) follows. Note that the equality in (45) is achieved when $H(Y|S_0, Y_1, X_1) = 1$, which happens if we choose X and X_1 to be independent, and let $X \sim \text{Ber}\left(\frac{1}{2}\right)$. Intuitively, (46) is the capacity of a binary channel with X as the input and $Y + Y_1 \pmod 2$ as the output.

Now, if we set X and X_1 to be independent $\text{Ber}\left(\frac{1}{2}\right)$ random variables, the equalities in both the multiple-access and broadcast bounds (40) and (45) are simultaneously achieved with the same $p(x, x_1)$. Therefore, the cut-set bound for this particular channel is equal to

$$\min\{R_0, 1 - h(\delta)\}. \tag{47}$$

The capacity given by (34) is in general strictly below the cut-set bound as illustrated by an example in Fig. 3.

III. Extension to Modular Relay Channels

We now extend the capacity results in Section II to include the general modulo-sum relay channel depicted in Fig. 4. The source and the destination are related by a modulo-sum channel. The relay observes Y_1, which is a correlated version of the noise Z with a conditional distribution $p(y_1|z)$. The relay also has a dedicated channel to the destination with a capacity

$$R_0 = \max_{p(x_1)} I(X_1; S_0). \tag{48}$$

The binary symmetric relay channel considered in Section II is a specific instance of the modulo-sum relay channel. The capacity proof for the binary case can be augmented to give the capacity of the modulo-sum relay channel.

Theorem 2: The capacity of a modular additive noise relay channel with $X \in \mathcal{X}$ and $|\mathcal{X}| = m$, where the relay observes Y_1 with $p(y_1|x, y, z) = p(y_1|z)$, and where the destination observes $Y = X + Z \pmod m$ from the source and observes S_0 from the relay through a separate channel with transition probabilities $p(s_0|x_1)$, is

$$C = \max_{p(u|y_1): I(U;Y_1) \leq R_0} m - H(Z|U) \tag{49}$$

where the maximization may be restricted to U's with $|\mathcal{U}| \leq |\mathcal{Y}_1| + 2$, and R_0 is as defined in (48).

Proof: Achievability follows by applying a simple extension to the achievability proof of Theorem 1. In addition, the binary symmetric relay channel converse appropriately modified to reflect the different alphabet sizes remains valid. This is because all the necessary conditions for the converse to work

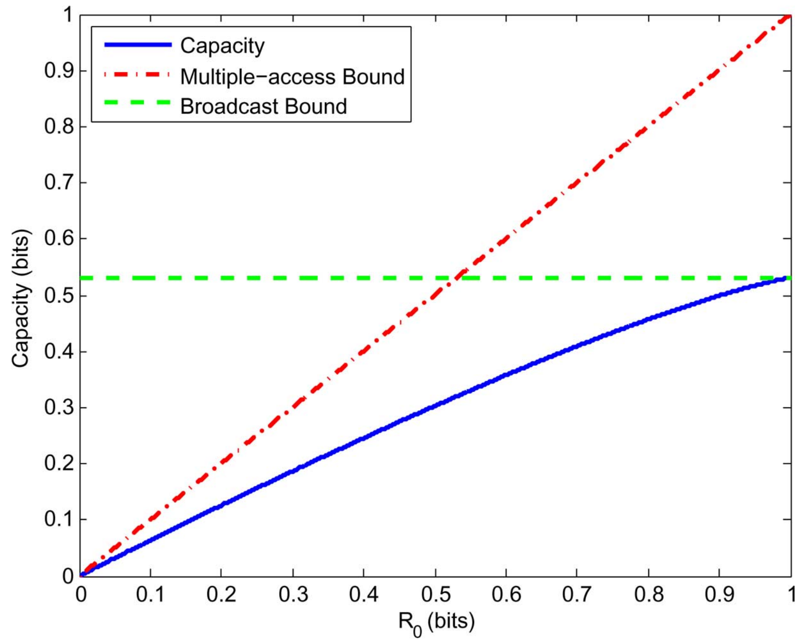

Fig. 3. The capacity and the cut-set bound for a binary modulo-sum relay channel, with $p = \frac{1}{2}$ and $\delta = 0.1$, as functions of R_0.

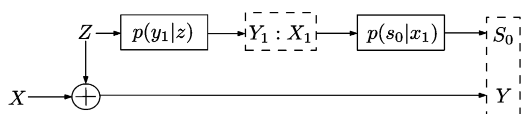

Fig. 4. A modulo-sum relay channel where $Y = X + Z \pmod{m}$ and the relay observes a corrupted version of the channel noise.

are satisfied. The modulo-sum channel is linear, and the uniform distribution applied at the input maximizes the output channel entropy regardless of how much is known about the additive noise, so (30) holds. □

IV. Channels With Rate-Limited Side Information at the Transmitter and Receiver

This section explores how the above result on the relay channel relates to and expands upon the body of work on channels with rate-limited state information to the transmitter and to the receiver. We show that the relay channel result confirms a conjecture due to Ahlswede and Han in this particular modular channel case. Further, we show that our converse technique gives a new upper bound for modular channels with rate-limited side information and a new capacity result for a specific binary case.

A. Connection With the Ahlswede–Han Conjecture

The modulo-sum relay channel model presented in Fig. 4 fits into the framework of a channel with side information presented in Fig. 1, once we set $R_1 = 0$, identify Y_1 with S, and digitize the relay–destination link to be of rate R_0 and set $R_0 = R_2$. The relay observation Y_1 can be considered as a channel state, because Y_1 depends on Z through a discrete memoryless channel and one can always factor out $p(y|x, y_1)$ from the overall transition probability $p(y|x, z)p(y_1|z)p(z)$. With this identification in mind, Theorem 2 confirms a conjecture due to Ahlswede and Han [10] on the capacity of channels with rate-limited side information to the receiver for this specific class of modulo-sum channels.

The Ahlswede–Han conjecture states that for channels with rate-limited state information to the receiver as shown in Fig. 5, the capacity is given by

$$C = \max\ I(X; Y|\hat{S}) \tag{50}$$

where the maximization is over all probability distributions of the form $p(x)p(s)p(y|x, s)p(\hat{s}|s)$ such that

$$I(\hat{S}; S|Y) \leq R_0 \tag{51}$$

and the auxiliary random variable $\hat{S}$ has cardinality $|\hat{\mathcal{S}}| \leq |\mathcal{S}| + 1$. The conjecture states that the state variable S should be quantized at rate R_0 in such a way as to maximize the resulting mutual information between X and Y. Ahlswede and Han proved in [10] that all rates below the capacity are achievable.

Now, consider the modulo-sum relay channel. Identify S with Y_1 and $\hat{S}$ with U. We claim that a uniform distribution on X maximizes the Ahlswede–Han rate in this case. This is because the uniform distribution on X makes Y independent of S, so that both the rate expression (50) and the constraint (51) in the Ahlswede–Han conjecture, respectively, coincides with the rate expression and constraint in (49).[1] By the converse part of Theorem 2, this achievable rate is the capacity, thus confirming the conjecture for the class of modulo-sum channels described in this paper.

[1]Allowing for the difference in cardinality bounds.

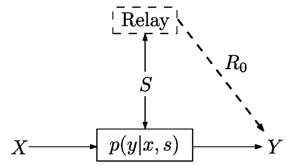

Fig. 5. Channel with rate-limited state information at the receiver.

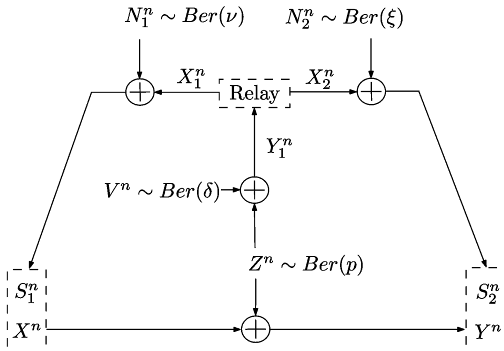

Fig. 6. A binary channel with rate-limited state information available to both the transmitter and the receiver.

B. An Upper Bound for Modulo-Sum Channels With Rate-Limited Side Information at the Transmitter and the Receiver

We now augment the modulo-sum channel considered in Section II by allowing the relay to convey noncausal side information to both the transmitter and the receiver at limited rates. The binary case is depicted in Fig. 6, where the channel input X^n goes through a BSC with crossover probability p to reach Y^n, i.e., $Y^n = X^n + Z^n \pmod 2$ where Z^n is i.i.d. $\text{Ber}(p)$. The relay observes the entire sequence Y_1^n which is a noisy version of Z^n, i.e., $Y_1^n = Z^n + V^n$, where V^n is i.i.d. $\text{Ber}(\delta)$. The relay also has a separate memoryless BSC to the transmitter $S_1 = X_1 + N_1$, where N_1 is $\text{Ber}(\nu)$, and a separate memoryless BSC to the destination $S_2 = X_2 + N_2$, where N_2 is $\text{Ber}(\xi)$. Let us define

$$R_1 = \max_{p(x_1)} I(X_1; S_1) \tag{52}$$

$$R_2 = \max_{p(x_2)} I(X_2; S_2) \tag{53}$$

for future reference. The main result of this section is that the converse proof for the class of relay channels considered in the first part of this paper can be modified to yield an upper bound for the channel in Fig. 6.

Theorem 3: The capacity of the channel in Fig. 6 is bounded above by

$$C \le \max_{p(u|y_1): I(U;Y_1) \le R_1 + R_2} 1 - H(Z|U) \tag{54}$$

where the maximization may be restricted to U's with $|\mathcal{U}| \le |\mathcal{Y}_1| + 2$, and R_1, R_2 are as defined in (52) and (53).

The bound depends on an extension of Lemma 1.

Lemma 2: Let Z^n, V^n, N_1^n and N_2^n be independent Bernoulli i.i.d. sequences. Let $Y_1^n = Z^n + V^n$, $S_1^n = X_1^n + N_1^n$, and $S_2^n = X_2^n + N_2^n$, where Y_1^n, X_1^n, and X_2^n are the input and the two output sequences of the relay as shown in Fig. 6. The following inequality holds for any encoding scheme at the relay:

$$H\left(Z^n | S_1^n, S_2^n\right) \ge \min_{p(u|y_1): I(U;Y_1) \le R_1 + R_2} nH(Z|U) \tag{55}$$

where the minimization on the right-hand side may be restricted to U's with $|\mathcal{U}| \le |\mathcal{Y}_1| + 2$.

Proof: Setting $S_0 = (S_1, S_2)$, letting (X_1, X_2) play the role of X_1, and allowing for a slight abuse of notation, Lemma 2 follows directly from Lemma 1. □

Proof of Theorem 3: The upper bound can now be proved as follows. Let $W \in \{1, 2, \ldots, 2^{nR}\}$ be the source message.

$$\begin{aligned}
nR &= H(W) && (56)\\
&\overset{(a)}{=} H\left(W | S_1^n\right) && (57)\\
&= I\left(W; Y^n, S_2^n | S_1^n\right) + H\left(W | Y^n, S_1^n, S_2^n\right) && (58)\\
&\overset{(b)}{\le} I\left(W; Y^n, S_2^n | S_1^n\right) + n\epsilon_n && (59)\\
&\overset{(c)}{\le} I\left(X^n; Y^n, S_2^n | S_1^n\right) + n\epsilon_n && (60)\\
&\overset{(d)}{=} I\left(X^n; Y^n | S_1^n, S_2^n\right) + n\epsilon_n && (61)\\
&= H\left(Y^n | S_1^n, S_2^n\right) - H\left(Y^n | S_1^n, S_2^n, X^n\right) + n\epsilon_n && (62)\\
&\overset{(e)}{\le} n - H\left(Z^n | S_1^n, S_2^n\right) + n\epsilon_n && (63)\\
&\overset{(f)}{\le} \max_{p(u|y_1): I(U;Y_1) \le R_1 + R_2} n(1 - H(Z|U)) + n\epsilon_n && (64)
\end{aligned}$$

where

(a) follows from the independence of the message W and S_1,
(b) follows from Fano's inequality,
(c) follows from the fact that $W - X^n - (Y^n, S_2^n)$ forms a Markov chain when S_1^n is known,
(d) follows from the fact that X^n is independent of S_2^n given S_1^n,
(e) follows from the fact that the maximum entropy of a binary random variable of length n is n,
(f) follows from Lemma 2. □

Just as in Section III, the above upper bound can be trivially modified to hold for a more general class of modular channels with rate-limited side information.

At this point, it is natural to ask whether this upper bound is achievable in general for the modular channel. Unfortunately, the evaluation of the optimal U in (54) via exhaustive search gives some indication that the upper bound may not always be achievable.

There is, however, at least one special case where the optimal U is analytically tractable and the upper bound is achievable. This is when Z^n is i.i.d. $\text{Ber}\left(\frac{1}{2}\right)$. In Section II-D, a particular binary relay channel with Z^n being i.i.d. $\text{Ber}\left(\frac{1}{2}\right)$ was considered, and its capacity was characterized to be strictly below the cut-set bound. This capacity is achieved by letting the relay quantize its observation with the goal of minimizing the Hamming distortion for Z at the destination. The resulting optimal U has an additive structure: the quantized signal, once conveyed to the destination via a separate channel, is merely to be added to the destination's received signal from the source. Since no Wyner–Ziv decoding is performed at the destination while decoding the quantized signal from the relay, decoding the quantized signal and adding it to the channel could just as easily have been performed at the transmitter with no loss in the overall rate.

C. Capacity of a Binary Channel With Rate-Limited Side Information at the Transmitter and the Receiver

Consider the binary channel with rate-limited side information as shown in Fig. 6 with Z^n having an i.i.d. $\mathrm{Ber}\left(\frac{1}{2}\right)$ distribution, i.e., $p = \frac{1}{2}$. The main result of this section is that the capacity of this channel is precisely (54) and is achieved by a simple quantization scheme. In this case, the capacity depends only on the sum of rates $R_1 + R_2$ available for state description at the relay.

Theorem 4: The capacity of the channel in Fig. 6 with Z^n being i.i.d. $\mathrm{Ber}\left(\frac{1}{2}\right)$ is

$$C = 1 - h\left(h^{-1}(1 - R_1 - R_2)) * \delta\right). \tag{65}$$

Proof: The converse is based on the evaluation of the expression on the right side of (54). As in Section II-D, this can be done using Wyner and Ziv's conditional binary entropy inequality, "Mrs. Gerber's Lemma" [14]. Wyner and Ziv's inequality holds because, as mentioned in Section II-D, when Z is $\mathrm{Ber}\left(\frac{1}{2}\right)$ we can write $Z = Y_1 + V$, where V is $\mathrm{Ber}(\delta)$ and Y_1 and V are independent. Observing that the constraint $I(U; Y_1) \le R_1 + R_2$ is equivalent to the constraint $H(Y_1|U) \ge 1 - R_1 - R_2$ and applying Wyner and Ziv's inequality yields $R \le C$.

The achievability can be established based on an evaluation of the achievable result of Heegard and El Gamal [11, Theorem 1]. A formal proof is presented in the Appendix. □

In the following, we give justification for the achievability of the rate given by (65) by outlining a simple quantization strategy for the relay. In this strategy, the relay first quantizes Y_1 with U_1 at rate R_1 with a goal of minimizing the Hamming distortion of Z. Let D_1 represent the resulting Hamming distortion. Next, the relay subtracts off U_1 from Y_1 and quantizes $(Y_1 - U_1)$ using U_2 at rate R_2 with a goal of further minimizing the Hamming distortion of Z. Let the resulting distortion be D_2. The relay sends U_1 to the transmitter and U_2 to the receiver through its two separate channels. The transmitter simply adds U_1 to its transmitted signal while the receiver adds U_2 to the channel output. Recalling that since Z is $\mathrm{Ber}\left(\frac{1}{2}\right)$ we can write $Z = Y_1 + V$, where V is $\mathrm{Ber}(\delta)$ and Y_1 and V are independent. The channel from the transmitter to receiver now looks like

$$Y = X + Z + U_1 + U_2 \tag{66}$$
$$= X + Y_1 + V + U_1 + U_2 \tag{67}$$
$$= X + V + Z' \tag{68}$$

where all additions are modulo 2, and $Z' = Y_1 + U_1 + U_2 \pmod 2$ is independent of X and V and distributed like a $\mathrm{Ber}\left(h^{-1}(1 - R_1 - R_2)\right)$ random variable. The distribution of Z' follows from the rate distortion theory for binary random variables, notably the fact that a binary random variable is successively refinable under the Hamming distortion measure [15] and the fact that the refining quantization variables have an additive structure (which can be verified using the method in [16]). In other words, quantizing Y_1 first at rate R_1, then subtracting off the quantization variable and further quantizing with R_2 yields a final distortion as if Y_1 were originally quantized at a rate of $R_1 + R_2$. More precisely, since

$$h(D_1) = H(Y_1) - R_1 \tag{69}$$
$$h(D_2) = h(D_1) - R_2 \tag{70}$$

adding the two equations together, we get $h(D_2) = H(Y_1) - R_1 - R_2$, implying $D_2 = h^{-1}(1 - R_1 - R_2)$ and Z' is distributed as $\mathrm{Ber}\left(h^{-1}(1 - R_1 - R_2)\right)$. Now, the channel between the transmitter and the receiver becomes equivalent to $Y = X + Z''$, where $Z'' = Z' + V$ has a $\mathrm{Ber}\left(h^{-1}(1 - R_1 - R_2) * \delta\right)$ distribution. This yields (65).

D. Comments on Theorem 4

In general, when Z is $\mathrm{Ber}(1/2)$, the rates available to the relay for communicating side information to the transmitter and to the receiver may be traded with each other arbitrarily without incurring any penalty to the overall rate. State information at the transmitter and at the receiver is equally valuable. Note that by setting R_1 to zero in the capacity expression of Theorem 4 we recover our previous result for the binary relay channel, while setting R_2 to zero yields a new capacity result for a channel with rate-limited side information to the transmitter only.

Note also that at the relay, the order of quantization is irrelevant; (U_1, U_2) may be paired for transmission with the transmitter and the receiver indiscriminately. Finally, we observe that digitizing the links to the transmitter and the receiver is optimal. No joint source–channel coding is necessary.

V. Conclusion

The paper identifies the capacity of a class of modular relay channels in which the relay observes only the additive noise. The capacity is shown to be strictly below the cut-set bound and achievable using a quantize-and-forward scheme where quantization is performed with a new metric, the conditional entropy of the noise at the destination. This is the first example of a relay channel for which the capacity is characterized to be strictly below the cut-set bound. It is proved that there is no advantage in performing joint source–channel coding of the relay's message over its dedicated link to the destination—digitizing the link is capacity achieving. The capacity derived here confirms a previous conjecture on the capacity of channels with rate-limited side information at the receiver for this modular channel case.

This paper also characterizes the capacity of a binary channel with rate-limited side information at both the transmitter and the receiver. It is shown that the capacity is achieved via a successive refining quantization scheme for the binary random variable. The capacity is shown to be a function of the sum of the rates available for communicating to the transmitter and to the receiver. This is the first example of a nontrivial channel with rate-limited state information available to the transmitter and the receiver whose capacity is characterized.

APPENDIX
ACHIEVABILITY PROOF OF THEOREM 4

For the channel in Fig. 1, Heegard and El Gamal [11, Theorem 1] proved that the following set of rates is achievable. Fixing $p(s)p(y|x,s)$ and auxiliary random variables U, S_o, S_e, and S_d, all rates (R, R_1, R_2) in the interior of the convex hull of the set

$$\begin{aligned}\{(R,R_1,R_2)|&R_1 > I(S_o,S_e;S)\\ &R_2 > I(S_o,S_d;S) - I(S_o,S_d;Y)\\ &R_2 > I(S_d;S|S_o) - I(S_d;Y|S_o)\\ &R_1+R_2 > I(S_o,S_e,S_d;S) - I(S_o,S_d;Y) + I(S_e;S_d|S_o)\\ &R_1+R_2 > I(S_e,S_d;S|S_o) - I(S_d;Y|S_o) + I(S_e;S_d|S_o)\\ &R < I(U;Y,S_d|S_o) - I(U;S_e|S_o)\}\end{aligned}$$

where the mutual information expressions are evaluated with probability distribution function of the form

$$p(s,s_o,s_e,s_d,u,x) = p(s)p(s_o,s_e,s_d|s)p(x,u|s_o,s_e)$$

are achievable.

To prove Theorem 4 , we identify Y_1 with S, U_1 with S_e, U_2 with S_d, and set $S_o = \emptyset$. Then by a direct substitution into [11, Theorem 1], we have the following result. Fixing $p(y_1)p(y|x,y_1)$ and auxiliary random variables (U,U_1,U_2), all rates (R,R_1,R_2) in the interior of the convex hull of the set

$$\begin{aligned}\{(R,R_1,R_2)|&R_1 > I(U_1;Y_1)\\ &R_2 > I(U_2;Y_1) - I(U_2;Y)\\ &R_1+R_2 > I(U_1,U_2;Y_1) - I(U_2;Y) + I(U_1;U_2)\\ &R < I(U;Y,U_2) - I(U;U_1)\}\end{aligned} \tag{71}$$

where the mutual information expressions are evaluated with probability distribution function of the form

$$p(y_1,u_1,u_2,u,x) = p(y_1)p(u_1,u_2|y_1)p(x,u|u_1)$$

are achievable.

To show that (65) is achievable, we identify the joint distribution $p(u_1,u_2|y_1)p(x,u|u_1)$ as follows. Let

$$U_1 \sim \mathrm{Ber}\left(\frac{1}{2}\right) \tag{72}$$

$$U_2 \sim \mathrm{Ber}\left(\frac{h^{-1}(1-R_1) - h^{-1}(1-R_1-R_2)}{1-2h^{-1}(1-R_1-R_2)}\right) \tag{73}$$

be independent. The idea is to use U_1 for a coarse quantization of Y_1 at a rate R_1 with distortion $D_1 = h^{-1}(1-R_1)$, and U_2 for a successive refinement at an additional rate R_2 and with a final distortion $D_2 = h^{-1}(1-R_1-R_2)$. By standard rate–distortion theory, the quantization variable U_1 is $\mathrm{Ber}\left(\frac{1}{2}\right)$ and is related to Y_1 via a BSC test channel with crossover probability D_1. Now, to obtain a successive refinement, U_2 should be a binary random variable representing the quantization of $(Y_1 - U_1)$, which is $\mathrm{Ber}(D_1)$. In this case, standard rate–distortion theory states that U_2 must be $\mathrm{Ber}\left(\frac{D_1-D_2}{1-2D_2}\right)$, yielding the expression (73). Note that because Y_1 is $\mathrm{Ber}\left(\frac{1}{2}\right)$, U_1 and U_2 are independent. Further, the joint distribution $p(u_1,u_2,y_1)$ is given by

$$Y_1 = U_1 + U_2 + Z'$$

where $Z' \sim \mathrm{Ber}\left(h^{-1}(1-R_1-R_2)\right)$ is independent of (U_1,U_2).

Next, we specify $p(x,u|u_1)$ by setting $U \sim \mathrm{Ber}\left(\frac{1}{2}\right)$ to be independent of U_1 and let

$$X = U + U_1.$$

Finally, the channel model gives $Y = X + Z$, $Z = Y_1 + V$, where V is $\mathrm{Ber}(\delta)$ independent of (U,U_1,U_2,Z'). All additions here are performed modulo 2.

Let us consider the first three constraints of (71) in turn

$$\begin{aligned}I(U_1;Y_1) &= H(Y_1) - H(Y_1|U_1)\\ &= 1 - H(U_2 + Z'|U_1)\\ &= 1 - 1 + R_1\\ &= R_1.\end{aligned} \tag{74}$$

Similarly, with the second constraint

$$I(U_2;Y_1) - I(U_2;Y) = 0 - 0 < R_2 \tag{75}$$

where $I(U_2;Y_1) = 0$ follows because $Y_1 = U_1 + U_2 + Z'$ and $U_1 \sim \mathrm{Ber}\left(\frac{1}{2}\right)$ is independent of U_2, and similarly $I(U_2;Y) = 0$ holds because $Y = U_1 + U_2 + Z$. Finally, with the last constraint

$$\begin{aligned}I(U_1,&U_2;Y_1) - I(U_2;Y) + I(U_1;U_2)\\ &= I(U_1,U_2;Y_1) + 0 - 0\\ &= H(Y_1) - H(Y_1|U_1,U_2)\\ &= 1 - H(Z')\\ &= 1 - 1 + R_1 + R_2\\ &= R_1 + R_2.\end{aligned} \tag{76}$$

Since all the constraints are satisfied, the following rate between the transmitter and the receiver is on the boundary of the Heegard and El Gamal's region:

$$\begin{aligned}I(U;&Y,U_2) - I(U;U_1)\\ &= I(U;Y,U_2) - 0\\ &= H(U) - H(U|Y,U_2)\\ &= 1 - H(U+Y+U_2|Y,U_2)\\ &= 1 - H(Z'+V|Y,U_2)\\ &= 1 - H(Z'+V)\\ &= 1 - h\left(h^{-1}(1-R_1-R_2) * \delta\right).\end{aligned} \tag{77}$$

Therefore, any rate $R < 1 - h\left(h^{-1}(1-R_1-R_2) * \delta\right)$ is achievable. □

REFERENCES

[1] T. M. Cover and A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. IT-25, no. 5, pp. 572–584, Sep. 1979.

[2] A. El Gamal and M. Aref, "The capacity of the semideterministic relay channel," *IEEE Trans. Inf. Theory*, vol. IT-28, no. 3, pp. 536–536, May 1982.

[3] A. El Gamal and S. Zahedi, "Capacity of a class of relay channels with orthogonal components," *IEEE Trans. Inf. Theory*, vol. 51, no. 5, pp. 1815–1817, May 2005.

[4] Y. H. Kim, "Capacity of a class of deterministic relay channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 3, pp. 1328–1329, Mar. 2008.

[5] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3037–3062, Sep. 2005.

[6] Y. Liang and V. V. Veeravalli, "Gaussian orthogonal relay channels: Optimal resource allocation and capacity," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3284–3289, Sep. 2005.

[7] Z. Zhang, "Partial converse for a relay channel," *IEEE Trans. Inf. Theory*, vol. 34, no. 5, pp. 1106–1110, Sep. 1988.

[8] S. Zahedi, "On Reliable Communication Over Relay Channels," Ph.D. dissertation, Stanford Univ., Stanford, CA, 2005.

[9] S. I. Gel'fand and S. M. Pinsker, "Coding for the channel with random parameters," *Probl. Contr. and Inf. Theory*, vol. 9, no. 1, pp. 19–31, 1980.

[10] R. Ahlswede and T. S. Han, "On souce coding with side information via a multiple-access channel and related problems in multi-user information theory," *IEEE Trans. Inf. Theory*, vol. IT-29, no. 3, pp. 396–412, May 1983.

[11] C. Heegard and A. El Gamal, "On the capacity of computer memory with defects," *IEEE Trans. Inf. Theory*, vol. IT-29, no. 5, pp. 731–739, Sep. 1983.

[12] A. Rosenzweig, Y. Steinberg, and S. Shamai (Shitz), "On channels with partial channel state information at the transmitter," *IEEE Trans. Inf. Theory*, vol. 51, no. 5, pp. 1817–1830, May 2005.

[13] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, second ed. New York: Wiley, 2006.

[14] A. D. Wyner and J. Ziv, "A theorem on the entropy of certain binary sequences and applications: Part I," *IEEE Trans. Inf. Theory*, vol. IT-19, no. 6, pp. 769–772, Nov. 1973.

[15] W. H. R. Equitz and T. M. Cover, "Successive refinement of information," *IEEE Trans. Inf. Theory*, vol. 37, no. 2, pp. 269–275, Mar. 1991.

[16] E. Tuncel and K. Rose, "Additive successive refinement," *IEEE Trans. Inf. Theory*, vol. 49, no. 8, pp. 1983–1991, Aug. 2003.

Marko Aleksic received the B.A.Sc. degree in engineering science in 2004 and the M.A.Sc. degree in electrical and computer engineering in 2007, both from the University of Toronto, Toronto, ON, Canada.

He is now with McKinsey & Company.

Peyman Razaghi (S'99) was born in Isfahan, Iran, 1980. He received the B.Sc. degree in electrical engineering and the M.Sc. degree in communication systems from Sharif University of Technology, Tehran, Iran, in 2002 and 2004, respectively.

He is currently working towards the Ph.D. degree at the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada. His research interests are in the areas of multiuser information theory, wireless communications, coding, and signal processing. During his Master and Ph.D. studies, he has been working on complexity manageable space–time codes, coding for the relay channel, relaying protocols for multirelay networks, and characterization of the capacity of relay channels.

Wei Yu (S'97–M'02–SM'08) received the B.A.Sc. degree in computer engineering and mathematics from the University of Waterloo, Waterloo, ON, Canada, in 1997 and M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, in 1998 and 2002, respectively.

He has been with the Electrical and Computer Engineering Department at the University of Toronto, Toronto, ON, Canada, since 2002, where he is now an Associate Professor. His research interests include information theory, optimization, wireless communications and broadband access networks.

Prof. Wei Yu holds a Canada Research Chair in Information Theory and Digital Communications. He is currently an Editor for IEEE TRANSACTIONS ON COMMUNICATIONS. He was an Editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2004 to 2007, was a Guest Editor of IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS for a special issue on "Nonlinear Optimization of Communications Systems" in 2006, and a Guest Editor of *EURASIP Journal on Applied Signal Processing* for a special issue on "Advanced Signal Processing for Digital Subscriber Lines" in 2005. He received the Early Researcher Award from the Province of Ontario in 2006, the Early Career Teaching Award from the Faculty of Applied Science and Engineering, University of Toronto in 2007, the McCharles Prize for Early Career Research Distinction in 2008, and the IEEE Signal Processing Society 2008 Best Paper Award.