

# Capacity of a Class of Modulo-Sum Relay Channels

Marko Aleksic, Peyman Razaghi, and Wei Yu

Department of Electrical and Computer Engineering, University of Toronto, Canada

e-mails: {aleksicm,peyman,weiyu}@comm.utoronto.ca

**Abstract**—This paper characterizes the capacity of a class of modulo 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 conjecture by Ahlswede and Han about the capacity of channels with rate limited state information at the destination for this particular class of channels.

## I. INTRODUCTION

The relay channel is a fundamental building block in network information theory. Complete characterization of the relay channel capacity would be a first step toward finding the capacities of larger networks. 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, reversely degraded [1], orthogonal [2], semideterministic [3], and recently a special class of deterministic [4] relay channels. All the above relay channels for which capacities are characterized have 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 prove the looseness of the cut-set bound have proved to be quite difficult. Zhang's partial converse [5] demonstrated the latter; Zahedi [2] provided some justifications for why the cut-set bound cannot be tight in all cases.

In this paper we find the capacity for a non-trivial 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 [1, Theorem 6]. The quantize-and-forward strategy was previously only known to achieve the cut-set bound capacity of one class of deterministic relay channels [4]. The modulo-sum relay channel appears to be a first example of a channel where 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 modulo 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 contained in this paper crucially depends 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.

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 destination. The capacity result for modulo-sum relay channels coincides with a hypothesis by Ahlswede and Han [6] about the capacity of channels with rate limited state information to the destination.

## 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. 1. Here, the channel input  $X$  goes through a binary symmetric channel (BSC) with crossover probability  $p$  to reach  $Y$ , i.e.,  $Y = X + Z \pmod{2}$  with  $Z$  being an i.i.d.  $Ber(p)$  random variable. The relay gets to observe a noisy version of  $Z$ , namely  $Y_1 = Z + V$ , where  $V$  is an i.i.d.  $Ber(\delta)$  random variable. The relay also has a separate BSC to the destination  $S = X_1 + N$ , where  $N$  is an i.i.d.  $Ber(\epsilon)$  random variable.

Let us define

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

for future reference. If there were no corrupting variable  $V$ , then the capacity of this channel is as recently characterized

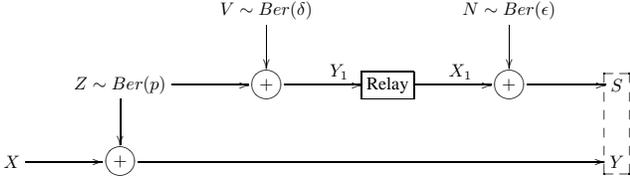


Fig. 1. A Binary Relay Channel

in [4]

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

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  $I(X; Y) + R_0$ . This bound is obtained by considering the achievable rate 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  $X$  is 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 knows 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 paper is that the above approach is capacity achieving for a class of modulo-sum relay channels including the channel in Fig. 1.

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

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

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

Fix the input distribution of  $X$  as  $Ber(\frac{1}{2})$ . The capacity can be achieved by a direct application of Theorem 6 in [1], if we identify  $U$  with  $\hat{Y}_1$ . A separate proof is provided here for completeness based on the theory of jointly strongly typical sequences [7].

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

*Codebook Generation:* Generate  $2^{nR}$  independently and identically distributed  $n$ -sequences,  $\mathbf{X}(w), w \in \{1 \dots 2^{nR}\}$

where each element is generated i.i.d.  $\sim \prod_{i=1}^n p(x_i)$ , and  $p(x_i)$  has the  $Ber(\frac{1}{2})$  distribution. Fix a  $p(u|y_1)$  such that it satisfies the constraint  $I(U; Y_1) \leq R_0$ . Generate  $2^{nI(U; Y_1)}$  i.i.d.  $n$ -sequences,  $\mathbf{U}(t), t \in \{1 \dots 2^{nI(U; Y_1)}\}$  where each element is generated i.i.d.  $\sim \prod_{i=1}^n p(u_i)$ .

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

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

*Analysis of the Probability of Error:* Because of the symmetry of the code construction we can perform the analysis assuming  $\mathbf{X}(1)$  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:

- $E_1$ :  $(\mathbf{X}(1), \mathbf{Y}, \mathbf{Y}_1)$  are not jointly strongly typical.
- $E_2$ :  $\nexists t, (\mathbf{U}(t), \mathbf{Y}_1)$  are jointly strongly typical.
- $E_3$ :  $(\mathbf{X}(1), \mathbf{Y}, \mathbf{U}(t))$  are not jointly strongly typical.
- $E_4$ : The destination makes an error decoding  $t$  in the next block.
- $E_5$ :  $\exists w \neq 1, (\mathbf{X}(w), \mathbf{Y}, \mathbf{U}(t))$  are jointly strongly typical.

For  $n$  sufficiently large we have  $P(E_1) < \frac{\epsilon}{5B}$ , and  $P(E_2 \cap E_1^c) < \frac{\epsilon}{5B}$ . By the Markov lemma [7, Lemma 14.8.1], since  $(\mathbf{X}(1), \mathbf{Y}) - \mathbf{Y}_1 - \mathbf{U}(t)$  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  $I(U; Y_1) \leq R_0$ , 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  $\mathbf{X}(w)$  is jointly strongly typical with both  $\mathbf{Y}$  and  $\mathbf{U}(t)$  is less than  $2^{-n(I(X; Y, U) - \gamma)}$ . Using the union bound, we have,  $P(E_5 \cap \bigcap_{i=1}^4 E_i^c) < 2^{nR} 2^{-n(I(X; Y, U) - \gamma)}$ . Thus, when

$$R < I(X; Y, U), \quad (4)$$

we have  $P(E_5 \cap \bigcap_{i=1}^4 E_i^c) < \frac{\epsilon}{5B}$ , for sufficiently large  $n$ . Now, since  $X$  and  $U$  are independent, we have

$$I(X; Y, U) = I(X; Y|U) \quad (5)$$

$$= H(Y|U) - H(Z|U) \quad (6)$$

$$= 1 - H(Z|U), \quad (7)$$

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

## B. Converse

The converse will be easy once we prove the following lemma.

*Lemma 1:* Let  $Z, V, N$  be independent Bernoulli random variables and let  $Y_1 = Z + V$ ,  $Y = X + Z$ , and  $S = X_1 + N$  as shown in Fig. 1. The following inequality holds for any encoding scheme at the relay,

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

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

*Proof:* The proof of the lemma is closely based on the proof of [7, Theorem 14.8.1]. Fixing an encoding scheme at the relay, our strategy is to show that there always exists a  $U$  for which  $H(Z^n|S^n) \geq nH(Z|U)$  and  $I(Y_1; U) \leq R_0$ . This would allow us to conclude that

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

We start by finding a lower bound for  $H(Z^n|S^n)$ :

$$H(Z^n|S^n) = \sum_{i=1}^n H(Z_i|S^n, Z_1, \dots, Z_{i-1}) \quad (9)$$

$$\geq \sum_{i=1}^n H(Z_i|S^n, Z^{i-1}, Y_1^{i-1}) \quad (10)$$

$$= \sum_{i=1}^n H(Z_i|S^n, Y_1^{i-1}) \quad (11)$$

where in the third line we use the fact that  $Z_i - S^n Y_1^{i-1} - S^n Y_1^{i-1} Z^{i-1}$  forms a Markov chain. The Markov chain follows because  $Z_i$ 's are i.i.d.,  $S^n$  is only a function of  $Y_1^n$ , and  $Z_i$  can only be affected by  $Z^{i-1}$  through  $S^n$ . Now define  $U_i = (S^n, Y_1^{i-1})$ , we get:

$$H(Z^n|S^n) \geq \sum_{i=1}^n H(Z_i|U_i). \quad (12)$$

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

$$I(X_1^n; S^n) \geq I(Y_1^n; S^n) \quad (13)$$

$$= \sum_{i=1}^n I(Y_{1i}; S^n | Y_{11}, \dots, Y_{1(i-1)}) \quad (14)$$

$$= \sum_{i=1}^n I(Y_{1i}; S^n, Y_1^{i-1}) \quad (15)$$

where in the third line we use the fact that  $Y_{1i}$  is independent of  $Y_1^{i-1}$  and consequently  $I(Y_{1i}; Y_1^{i-1}) = 0$ . Using our definition of  $U$  we get

$$I(X_1^n; S^n) \geq \sum_{i=1}^n I(Y_{1i}; U_i). \quad (16)$$

Recall that  $R_0 = \max_{p(x_1)} I(X_1; S)$ . Thus, we have shown the following inequalities:

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

$$\frac{1}{n} H(Z^n|S^n) \geq \frac{1}{n} \sum_{i=1}^n H(Z_i|U_i). \quad (18)$$

Introducing a standard timesharing random variable  $Q$ , the above equations can be rewritten as

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

$$\frac{1}{n} H(Z^n|S^n) \geq \frac{1}{n} \sum_{i=1}^n H(Z_i | U_i, Q = i) = H(Z_Q | U_Q, Q) \quad (20)$$

Now, since  $Q$  is independent of  $Y_{1Q}$ , we have

$$I(Y_{1Q}; U_Q | Q) = I(Y_{1Q}; U_Q, Q) - I(Y_{1Q}; Q) = I(Y_{1Q}; U_Q, Q). \quad (21)$$

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

$$R_0 \geq I(Y_1; U) \quad (22)$$

$$H(Z^n|S^n) \geq nH(Z|U) \quad (23)$$

for any particular 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 [7, Theorem 14.8.1]. ■

The converse can now be proved in a straightforward manner with:

$$nR = H(W) \quad (24)$$

$$= I(W; Y^n, S^n) + H(W | Y^n, S^n) \quad (25)$$

$$\stackrel{(a)}{\leq} I(W; Y^n, S^n) + n\epsilon_n \quad (26)$$

$$\leq I(X^n; Y^n, S^n) + n\epsilon_n \quad (27)$$

$$\stackrel{(b)}{=} I(X^n; Y^n | S^n) + n\epsilon_n \quad (28)$$

$$= H(Y^n | S^n) - H(Y^n | S^n, X^n) + n\epsilon_n \quad (29)$$

$$\stackrel{(c)}{\leq} n - H(Z^n | S^n, X^n) + n\epsilon_n \quad (30)$$

$$= n - H(Z^n | S^n) + n\epsilon_n \quad (31)$$

$$\stackrel{(d)}{\leq} \max_{p(u|y_1): I(U; Y_1) \leq R_0} n(1 - H(Z|U)) + n\epsilon_n \quad (32)$$

$$= nC + n\epsilon_n \quad (33)$$

where

- (a) follows from Fano's inequality,
- (b) follows from the fact that  $X^n$  is independent of  $S^n$ ,
- (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 high 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  $Ber(\frac{1}{2})$  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 to act 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 of Theorem 1 can be strictly below the cut-set bound, consider the case in which  $Z^n$  has an i.i.d.  $Ber(\frac{1}{2})$  distribution. The capacity can now be evaluated as

$$C = 1 - h(h^{-1}(1 - R_0) * \delta), \quad (34)$$

where  $h(p) = -p \log_2 p - (1 - p) \log_2 (1 - p)$ , 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 of Theorem 1 can be rewritten as

$$H(Y_1|U) \geq H(Y_1) - R_0. \quad (35)$$

Now we use Wyner and Ziv's version of the conditional entropy power inequality for binary random variables [8] to claim that if

$$H(Y_1|U) \geq \alpha, \quad (36)$$

then

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

with equality if  $Y_1$  given  $U$  is a  $Ber(h^{-1}(\alpha))$  random variable. Wyner and Ziv's inequality holds because when  $Z$  is  $Ber(\frac{1}{2})$  we can write  $Z = Y_1 + V$ , where  $V$  is  $Ber(\delta)$  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  $Ber(h^{-1}(H(Y_1) - R_0))$ , 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  $\nu$  must satisfy  $H(\nu) = H(Y_1|U) = H(Y_1) - R_0$  and  $Y_1$  given  $U$  must be  $Ber(\nu)$ . Further, as  $Y_1$  is  $Ber(\frac{1}{2})$ , the distribution of the optimal  $U$  is also  $Ber(\frac{1}{2})$ . 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 equals [1]

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

When  $Z$  is  $Ber(\frac{1}{2})$ , we have

$$I(X, X_1; Y, S) = H(Y, S) - H(Y, S|X, X_1) \quad (39)$$

$$\leq 2 - H(Z, N|X, X_1) \quad (40)$$

$$= 1 - H(Z) + 1 - H(N) \quad (41)$$

$$= R_0, \quad (42)$$

where the equality in (40) is achieved by letting  $X$  and  $X_1$  have independent and identical  $Ber(\frac{1}{2})$  distributions.

Similarly, for the broadcast bound we have

$$I(X; Y, S, Y_1|X_1) = I(X; Y|S, Y_1, X_1) \quad (43)$$

$$= H(Y|S, Y_1, X_1) - H(Z|S, Y_1, X, X_1) \quad (44)$$

$$\leq 1 - H(V|S, Y_1, X, X_1) \quad (45)$$

$$= 1 - H(\delta). \quad (46)$$

In the first line, we use the fact that  $X$  is independent of  $Y_1$  and  $S$  given  $X_1$ . In the third line, we again use the fact that  $Y_1 = Z + V$  and since  $Z$  is  $Ber(\frac{1}{2})$ , so is  $Y_1$ , thus  $Z = Y_1 + V$ , and  $Y_1$  and  $V$  are independent. The equality in (45) is achieved again with  $X$  and  $X_1$  as independent  $Ber(\frac{1}{2})$  distributed random variables. Since both (40) and (45) are achieved with equality with the same maximizing  $p(x, x_1)$ , we have shown that the cut-set bound for this particular channel is equal to

$$\min\{R_0, 1 - H(\delta)\}. \quad (47)$$

The capacity given by (34) is strictly below the cut-set bound for all values of  $R_0 \geq 1 - H(\delta)$ .

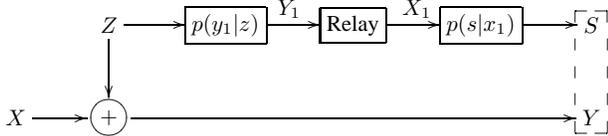


Fig. 2. The modulo-sum Relay Channel

### 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. 2. 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

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

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 and additive relay channel, in which the relay observes  $Y_1$ , with  $p(y_1|x, y, z) = p(y_1|z)$ , and the destination observes  $Y = X + Z \bmod m$  from the source and  $S$  from the relay through a separate channel with transition probabilities  $p(s|x_1)$ , is

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

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).

Achievability follows by applying a simple extension to the achievability proof of Theorem 1. 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 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. CONNECTION TO AHLWEDE-HAN CONJECTURE

The Ahlswede-Han [6] conjecture states that for channels with rate limited state information to the decoder as shown in Fig. 3, the capacity is given by,

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

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

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

and the auxiliary random variable  $\hat{S}'$  has cardinality  $|\hat{S}'| \leq |S'| + 1$ .

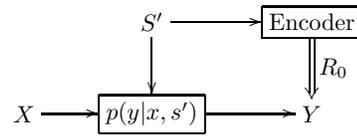


Fig. 3. Channel with rate limited state information to the decoder

For these channels, the output  $Y$  depends stochastically on both the input  $X$  and the particular channel state  $S'$ . The channel state is observed at another encoder that has a digital link to the destination with capacity  $R_0$ . The conjecture claims 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  $Y$ . By identifying  $S'$  with  $Y_1$ , and  $\hat{S}'$  with  $U$ , we observe that the class of relay channels described in Theorem 2 is a special case of the channel with rate limited state information to the decoder. We also note that the uniform distribution on  $X$  maximizes the capacity and makes  $Y$  independent of  $S'$ , so that the rates achievable by (50) and (49) are identical<sup>1</sup>, thus confirming the conjecture for the class of channels described in this paper.

### V. CONCLUSION

The capacity of a class of modular additive relay channels was found. The capacity was 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 can be strictly below the cut-set bound. It was proved that there is no advantage to 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 conjecture by Ahlswede and Han about the capacity of the rate limited channels with state information for this class of channels.

### REFERENCES

- [1] T. M. Cover and A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inform. Theory*, vol. 25, no. 5, pp. 572–584, Sept. 1979.
- [2] S. Zahedi, "On reliable communication over relay channels," Ph.D. dissertation, Stanford Univ., Stanford, CA, 2005.
- [3] A. El Gamal and M. Aref, "The capacity of the semideterministic relay channel," *IEEE Trans. Inform. Theory*, vol. 28, no. 3, p. 536, May 1982.
- [4] T. M. Cover and Y.-H. Kim, "Capacity of a class of deterministic relay channels," in *Proc. IEEE Int. Symp. Information Theory*, June 2007.
- [5] Z. Zhang, "Partial converse for a relay channel," *IEEE Trans. Inform. Theory*, vol. 34, no. 5, pp. 1106–1110, Sept. 1988.
- [6] R. Ahlswede and T. S. Han, "On source coding with side information via a multiple-access channel and related problems in multi-user information theory," *IEEE Trans. Inform. Theory*, vol. 29, no. 3, pp. 396–412, May 1983.
- [7] T. M. Cover and J. S. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
- [8] A. D. Wyner and J. Ziv, "A theorem on the entropy of certain binary sequences and applications: Part I," *IEEE Trans. Inform. Theory*, vol. 19, no. 6, pp. 769–772, Nov 1973.

<sup>1</sup>Allowing for the difference in cardinality bounds.