CAPACITY OF A MODULO-SUM SIMPLE RELAY NETWORK

A Thesis by

Youvaraj T. Sagar

B.E., Sri Jayachamarajendra College of Engineering, VTU, 2006

Submitted to the Department of Electrical and Computer Engineering
and the faculty of the Graduate School of
Wichita State University
in partial fulfillment of
the requirements for the degree of
Master of Science

December 2009

CAPACITY OF A MODULO-SUM SIMPLE RELAY NETWORK

The following faculty members have examined the final copy of this thesis for form and content, and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science with a major in Electrical Engineering.

_____

Hyuck M. Kwon, Committee Chair

_____

Yanwu Ding, Committee Member

_____

Xiaomi Hu, Committee Member

DEDICATION

To my Dad, Thayappa, and my, Mom Neelamma

ACKNOWLEDGEMENTS

Foremost, I would like to thank my advisor Dr. Hyuck M. Kwon for his continuous support of my research, patience, motivation, enthusiasm, and immense knowledge. His guidance helped me throughout the entire time of research and writing this thesis. Also, I would like to thank Dr. Yanwu Ding and Dr. Xiaomi Hu for their valuable suggestions, comments, and support.

I thank my fellow laboratory colleagues in the Wireless Research and Development Group (WiReD): Amitav Mukherjee, Andy Zerngast, Bi Yu, Chandana Jayasooriya, Jayesh Sonawane, Jie Yang, Paul Okokhere, Phillip Shaw, Shane Hodges, Shuang Feng, Tze Wong, Xiong Wenhao, and Zuojun Wang, for their stimulating discussions, for the sleepless nights we spent working together to meet deadlines, and for all the fun we had during my tenure.

Further, I would like to thank my friends Nahusha, Manju, Jay, Prakash, Santosh, Praveen, Ganesh, Jagadeesh, Vinay, Shreyas, Guru, Ashwin, Girish, Harsha, Damodar, Ashok, Pooja, Prashant, Rakesh, Nagendra, Padmaraj, Ravi, Pradeep, Harish, Siddharth, Anirudh, Sumanth, Gangadhar, Shilpa, and Sandesh, for their encouragement and moral support.

Finally, I would like to thank NASA and the U.S. Army Research Office for their support of our WiReD group.

ABSTRACT

This thesis proposes the capacity of a modulo-sum simple relay network. In previous work related to relay channels, capacity was characterized in the case where noise was transmitted to the relay, and the closed-form capacity was derived only for the noise with a Bernoulli-$(1/2)$ distribution. However, in this work, the source is transmitted to the relay, and a more general case of noise with an arbitrary Bernoulli-(p) distribution, $p \in [0,0.5]$, is considered. The relay observes a corrupted version of the source, uses a quantize-and-forward strategy, and transmits the encoded codeword through a separate dedicated channel to the destination. The destination receives the codeword from both the relay and source. For the relay channel, it is assumed that the channel is discrete and memoryless. After deriving the achievable capacity theorem (i.e., the forward theorem) for the binary symmetric simple relay network, it is proven that the capacity is strictly below the cut-set bound. In addition, this thesis presents the proof of the converse theorem. Finally, the capacity of the binary symmetric simple relay network is extended to that of an *m*-ary modulo-sum relay network.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.1 Motivation and Objective

The relay network is a fundamental building block in network information theory, whereby any arbitrary network can be characterized. First, a discussion of the relay network is necessary. The relay network is a channel that has one sender and one receiver, with a number of intermediate nodes acting as relays to assist with the communication between sender and receiver. This thesis exchanges the terminology of the relay channel proposed by Cover and Thomas [1] frequently with the relay network, because here a network is defined as consisting of more than two nodes [2], whereas a channel is for communication between two nodes. The simplest relay network or channel has one sender, one receiver, and one relay node. Figure 1 shows this type of relay network, which is called a "simple" relay network. To the author's knowledge, the capacity of a general relay network has not yet been found. This thesis characterizes the capacity of a modulo-sum simple relay network.

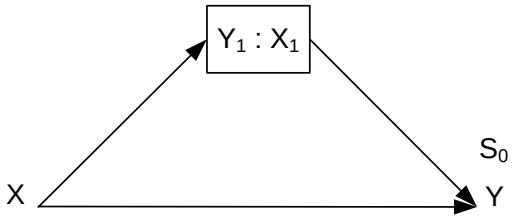

Figure 1. Simple relay network.

1.2 Technical Background and Previous Work

The first original model of a relay network was introduced by Van Der Meulen in 1971 [3]. After that, extensive research was done to find the upper bounds, cut-set bounds, and exact capacity for this network. In 1979, Cover and Gamal obtained the capacity for a special class of

channels called physically degraded relay channels [4]. In that paper, they discussed the capacity of the relay channel with feedback and found an upper bound for a simple relay network, which is shown in Figure 1. Later, Gamal and Aref found the capacity for a special class of relay channels called "semideterministic relay channels" [5]. Then, Kim found the capacity for a class of deterministic relay channels [6], where he modeled the simple relay network as a noiseless channel between the relay and the destination. Also, Van Der Meulen and Vanroose corrected Van Der Meulen's previous upper bound on the capacity of the simple relay network with and without delay [7].

Using Kim's results [6], Aleksic et al. modeled the channel between the relay and the destination as a modular sum noise channel [8]. Here, binary side information or channel state information is transmitted to the relay [8]. These authors mentioned that the capacity of the simple relay network is not yet known. Recently, however, Tandon and Ulukus found a new upper bound for the simple relay network with general noise, obtained the capacity for a symmetric binary erasure relay channel, and compared them with the cut-set bound [9].

Aleksic et al. [8] introduced a corrupting variable to the noiseless channel [6], whereby the noise in the direct channel between the source and the destination is transmitted to the relay. The relay observes a corrupted version of the noise and has a separate dedicated channel to the destination. For this case, the capacity was characterized [8]. However, the closed-form capacity was derived only for the noise with a Bernoulli-$(p = 1/2)$. distribution.

Also, there are many relay protocols: for example, amplify-and-forward, decode-and-forward, compress-and-forward, estimate-and-forward, and quantize-and-forward. In the amplify-and-forward strategy, the relay nodes scale the noisy version of the received signal based on the amplification coefficient and forward it to the destination. In the decode-and-

forward strategy, the relay nodes decode the source message first, and then encode and forward it to the destination. The compress-and-forward and estimate-and-forward strategies are similar. However, the source coding is applied at the relay in the compress-and-forward strategy. The quantize-and-forward strategy is based on theoretic information and is used at the relay. In this strategy, the relay nodes quantizes the received signal and forwards it to the destination.

1.3 Thesis Overview

From Figure 1, it can be seen that the source input is transmitted to both the relay and destination through the binary symmetric channel. The relay observes a corrupted version of the source input, uses a quantize-and-forward strategy, and transmits the encoded codeword through a separate dedicated channel to the destination. The destination receives the codeword from both the relay and source. The objective of this thesis is to find the capacity of the simple modular sum relay network and show that its capacity is strictly below the cut-set bound [4]. This thesis work also presents a closed-form capacity for a general case, such as for any p where the source is transmitted to both the relay and the destination.

This work considers all noisy channels, i.e., from the source to the destination, from the source to the relay, and from the relay to the destination, as shown in Figure 1, where all noisy channels are binary symmetric channels (BSCs) with a certain crossover probability, e.g., p. Also, this work derives the capacity for this class of relay channels. In other words, the capacity of a modulo-sum simple relay network is presented here. The capacity proof for the binary symmetric simple relay network and the proof for the converse depend crucially on the input distribution. For the BSC, a uniform input distribution at the source is assumed because this distribution maximizes the entropy of the output (or the capacity) regardless of additive noise. Furthermore, because of the uniform input distribution, the output of a binary symmetric relay

network is independent of additive noise. After presenting the proof for the capacity of a binary symmetric simple relay network, this thesis work proves that the capacity obtained is strictly below the cut-set bound by using the results in [4]. Finally, our work shows the converse theorem for this class of networks.

1.4 Thesis Outline

Chapter 2 describes the system model and presents the capacity of the binary symmetric simple relay network. Proofs of the converse and achievability for this theorem are provided in Appendices A and B. Chapter 3 discusses the cut-set bound for the binary symmetric simple relay network and presents the numerical analysis results. Chapter 4 extends the capacity to the *m*-ary modular additive case. Finally, Chapter 5 concludes with comments.

Note: I have submitted this work to *IEEE Transactions on Information Theory* [12].

CHAPTER 2

SYSTEM MODEL AND NETWORK CAPACITY

Figure 2 shows a realistic binary phase-shift keying (BPSK) system under additive white Gaussian noise (AWGN), where X and Y are the binary input and output signal, respectively. Here, Y is obtained with a hard decision on the demodulated signal.

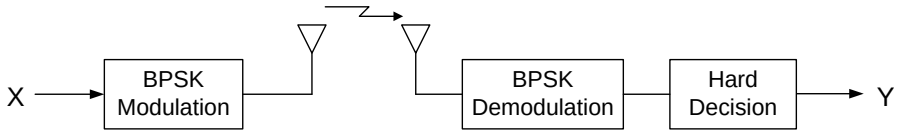

Figure 2. Realistic BPSK communication system under AWGN.

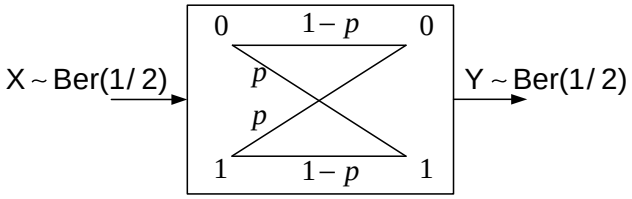

Figure 3. BSC equivalent to Figure 2.

Figure 3 shows a BSC with the crossover probability p equivalent to the realistic communication system, as shown in Figure 2. Here, the crossover probability p is equal to $Q\left[\sqrt{2E_b/N_0}\right]$, where $Q(\alpha) = \int_{\alpha}^{\infty}(1/\sqrt{2\pi})\, e^{-t^2/2}dt$, and E_b and N_0 denote the bit energy and the one-side AWGN power spectral density, respectively.

This paper models a channel between any adjacent nodes in Figure 1 as a BSC that has one sender, one receiver (or destination), and one relay node [1]. The random variable Y represents the received signal through the direct channel and is written as $Y = X \oplus Z$, where X and Z denote the transmitted and noise random variable with distribution $Ber(1/2)$ and $Ber(p)$,

respectively, and $\oplus$ denotes the binary modulo-sum, i.e., Z = 1 with probability p, and Z = 0 with probability $(1-p)$.

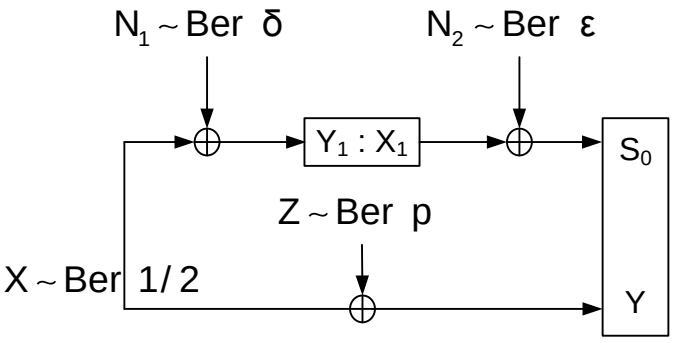

Figure 4. Binary symmetric simple relay network.

The simple relay network in Figure 1 can be redrawn as Figure 4. Here, the relay node has an input Y_1 and an output X_1. The relay node observes the corrupted version of X, i.e., $Y_1 = X \oplus N_1$, encodes it using a codebook $\mathcal{U}^n$ of jointly typical strong sequences [1], and transmits the code symbol X_1 through another separate BSC to the destination node, where $\mathcal{U}$, n, and N_1 denote the alphabet of code symbols, the codeword length, and the noise random variable at the relay with distribution $Ber(\delta)$, respectively. The destination receives Y, through the direct channel and $S_0 = X_1 \oplus N_2$, through the relay node, where $N_2 \sim Ber(\varepsilon)$ represents the noise at the destination for the relay network. Note that the binary modulo-sum and the BSC can be extended to an m–ary modulo-sum and an m–ary symmetric channel (MSC).

To the authors' knowledge, there is no network capacity expression in the literature, even for the simple relay network shown in Figure 4. Only the capacity of a deterministic relay channel, i.e., the case of $N_2 = 0$ in Figure 4, is presented [6]. The capacity of a relay network by replacing X with Z, i.e., the case where the relay observes a corrupted version of the direct channel noise Z, is presented [8]. This paper presents the capacity of the simple relay network shown in Figure 4 in the following theorem.

***Theorem 1**:* The capacity C of the binary symmetric simple relay network shown in Figure 4 is

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

where the maximization is over the U's conditional probability density function (p.d.f.) given Y_1; the cardinality of the alphabet $\mathcal{U}$, is bounded by $|\mathcal{U}| \leq |\mathcal{Y}_1| + 2$; and R_0 is the capacity for the channel between X_1 and S_0, which can be written as

$$R_0 = \max_{p(x_1)} I(X_1; S_0). \quad (2)$$

The closed-form network capacity for the simple relay network shown in Figure 4 can be written as

$$C = 1 + \mathcal{H}(\{\varepsilon * \delta\} * p) - \mathcal{H}(p) - \mathcal{H}(\varepsilon * \delta). \quad (3)$$

Here, $H(X)$ and $I(X;Y)$ are the entropy of X and the mutual information between X and Y, respectively [1]; $\mathcal{H}(\alpha)$ is the binary entropy function written as $\mathcal{H}(\alpha) = -\alpha\log_2\alpha - (1-\alpha)\log_2(1-\alpha)$; and $\alpha * \beta = \alpha(1-\beta) + (1-\alpha)\beta$ [10].

Proof: Formal proofs for the achievability and the converse of Theorem 1 are presented in Appendices A and B, respectively.

Note that if the direct channel noise Z is transmitted through the relay rather than X, then equation (1) becomes equation (3) [8] or equation (4), written as

$$C = \max_{p(u|y_1):I(U;Y_1)\leq R_0} 1 - H(Z|U). \quad (4)$$

This is because $H(Y|U)$ and $H(Z)$ in equation (1) will become 1 and $H(X) = 1$, respectively.

CHAPTER 3

CUT-SET BOUND AND ANALYTICAL RESULTS

This chapter proves that the capacity of the binary symmetric simple relay network in Figure 4 is strictly below the cut-set bound, except for the two trivial points at $R_0 = 0$ and $R_0 = 1$. The capacity in equation (1) can be upper-bounded by the cut-set bound as

$$C \leq \max_{p(x,\, x_1)} \min\{I(X, X_1; Y, S_0),\ I(X; Y, Y_1)\} \tag{5}$$

where the Ford-Fulkerson theorem [11], [4] is applied to the simple relay network in Figure 4. Using equation (5), Theorem 2 can be established.

Theorem 2: The cut-set bound for the capacity of the binary symmetric simple relay network shown in Figure 4 can be written as

$$\begin{aligned} C &\leq \min\{1 - H(Z) + R_0, 1 - H(Z) + 1 - H(N_1)\} \\ &= \min\{1 - \mathcal{H}(p) + R_0,\ 1 - \mathcal{H}(p) + 1 - \mathcal{H}(\delta)\}. \end{aligned} \tag{6}$$

Proof: A formal proof for Theorem 2 is presented in Appendix C.

Figures 5(a) and 5(b) show the capacity in bits per transmission versus R_0 bits for $\delta = 0.1$, when $p = 0.1$ and $p = 0.5$, respectively. If $p = 0.5$, then the results are the same as those found by Aleksic et al. [8]. Only the closed form of the capacity for the special case of $p = 0.5$ was analyzed and presented [8], where the capacity C of the binary simple relay network was obtained by replacing X with Z at the relay input shown in Figure 4 and written as [8]

$$C = 1 - \mathcal{H}(\mathcal{H}^{-1}\{1 - R_0\} * \delta). \tag{7}$$

Here $\mathcal{H}^{-1}(\cdot)$ is the inverse of $\mathcal{H}(p)$ in the domain $p \in [0,0.5]$. Note that the capacity in equation (3) of this paper is valid for a general p between 0 and 0.5, whereas the one in equation (34) in the work of Aleksic et al. [8] or equation (7) is valid for only $p = 0.5$. Also, in Figure 5(b) the

red dotted line is the multiple-access cut applied in Figure 4, and similarly, the green dotted line is the broadcast cut. Since for Figure 5(b), $\delta = 0.1$ and $p = 0.5$ are considered, the capacity for the broadcast cut is ≈ 0.53 for R_0 ranging from 0 to 1.

Note that the capacity in equation (3) is strictly below the cut-set bound in equation (6), as shown in Figure 5(b).

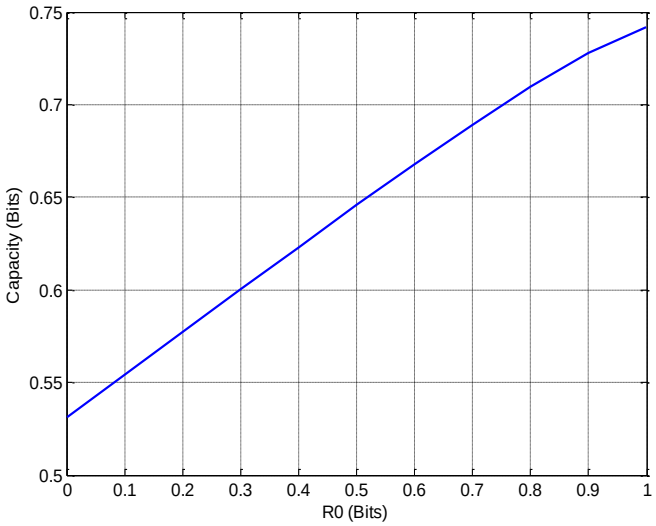

Figure 5(a). Capacity of a binary symmetric simple relay network shown in Figure 4 for $\delta = 0.1$ and $p = 0.1$.

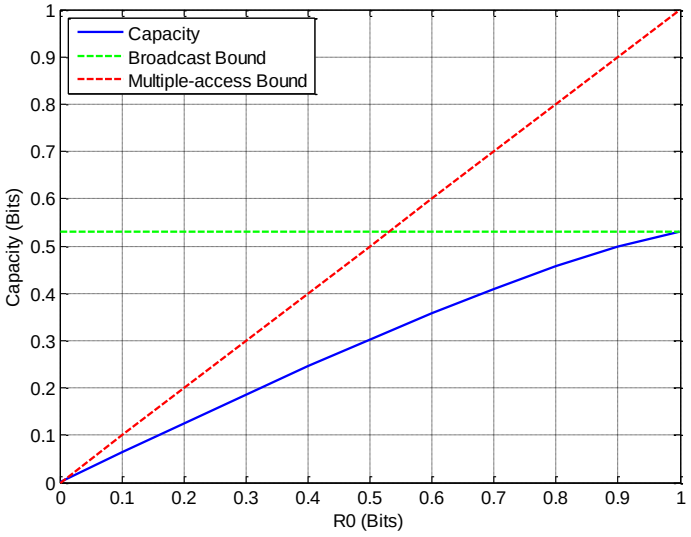

Figure 5(b). Capacity of a binary symmetric simple relay network shown in Figure 4 for $\delta = 0.1$ and $p = 0.5$.

CHAPTER 4

CAPACITY FOR M-ARY MODULO-SUM RELAY NETWORK

This section extends the capacity derived for the binary symmetric simple relay network to the *m*-ary modular additive relay network. The received signal at the destination node can be written as $Y = X + Z\ (mod\ m)$. The relay observes the corrupted version of X, i.e., $Y_1 = X + N_1\ (mod\ m)$, and the relay also has a separate channel to the destination: $S_0 = X_1 + N_2\ (mod\ m)$ with a capacity $R_0 = \max\limits_{p(x_1)} I(X_1; S_0)$. Therefore, equation (1) becomes equation (8) in Theorem 3.

Theorem 3*:* The capacity C of the symmetric *m*-ary modulo-sum simple relay network is

$$C = \max_{p(u|y_1):I(U;Y_1)\leq R_0} \left\{m + H(Y\,|\,U) - H(Z) - H(X\,|\,U)\right\} \tag{8}$$

where maximization is over the conditional U's p.d.f. given Y_1 with $|\mathcal{U}| \leq |\mathcal{Y}_1| + 2$, and R_0 is defined in equation (2).

Proof: The achievability for Theorem 3 follows the same steps as Theorem 1 by changing the binary to the *m*-ary case. Also, the uniform input distribution at the source maximizes the entropy of the output, regardless of the additive noise. Furthermore, because of the uniform input distribution, the output of an *m*-ary modulo-sum relay network is independent of the additive noise. Therefore, equation (8) holds true. The converse for Theorem 3 also holds true using the same steps of Theorem 1 by changing the binary modulo-sum to the *m*-ary modulo-sum.

CHAPTER 5

CONCLUSIONS

It has been an open problem to find the capacity of the simple relay network. This paper presented the closed form capacity of the binary symmetric simple relay network. Also, this paper extended the capacity for the binary to the m-ary modulo-sum symmetric simple relay network. Two conditions are necessary for the derivation of this capacity: (1) a uniform Bernoulli-$(1/2)$ input distribution, and (2) a modular additive channel between the two adjacent nodes. Using these conditions, both proofs for the achievability and the converse of the capacity theorem were presented. Furthermore, this paper derived the cut-set bound and presented the numerical results for this network. Finally, this paper determined that the capacity is strictly below the cut-set bound and achievable using a quantize-and-forward strategy at the relay.

REFERENCES

LIST OF REFERENCES

1. T. M. Cover and J. A. Thomas, *Elements of Information Theory*, second edition, New York: Wiley, 2006.

2. M. Schwartz, *Telecommunication Networks: Protocols, Modeling and Analysis*, Menlo Park, CA: Addison-Wesely Publising Co., 1988.

3. E. C. Van Der Meulen. "Three-Terminal Communication Channels," *Adv. Appl. Prob*., vol. 3, pp. 120–154, 1971.

4. T. M. Cover and El Gamal, "Capacity Theorems for the Relay Channel," *IEEE Trans. Inf. Theory*, vol. IT-25, no. 5, pp. 572–584, Sept. 1979.

5. A. El Gamal and M. Aref, "The Capacity of the Semideterministic Relay Channel," *IEEE Trans. on Inf. Theory*, vol. IT-28, no. 3, p. 536, May 1982.

6. Y. H. Kim, "Capacity of a Class of Deterministic Relay Channels," *IEEE Trans. on Inf. Theory*, vol. IT-54, no. 3, pp. 1328–1329, Mar. 2008.

7. E. C. Van Der Meulen and P. Vanroose, "The Capacity of a Relay Channel, Both with and without Delay," *IEEE Trans. on Inf. Theory*, vol. 53, no.10, pp. 3774–3776, Oct. 2007.

8. M. Aleksic, P. Razaghi, and W. Yu, "Capacity of a Class of Modulo-Sum Relay Channels," *IEEE Trans. on Inf. Theory*, vol. 55, no. 3, pp. 921–930, Mar. 2009.

9. R. Tandon and S. Ulukus, "A New Upper Bound on the Capacity of a Class of Primitive Relay Channels," *Communication, Control, and Computing, 2008 46th Annual Allerton Conference*, pp. 1562–1569, Sept. 2008.

10. A. D. Wyner and J. Ziv, "A Theorem on the Entropy of Certain Binary Sequences and Applications" *IEEE Trans. Inf. Theory*, vol. IT-19, no. 6, pp. 769–777, Nov. 1973.

11. L. R. Ford and D. R. Fulkerson, *Flows in Networks*, Princeton, NJ: Princeton University Press, 1962.

12. Youvaraj T. Sagar, Hyuck M. Kwon, and Yanwu Ding, "Capacity of a Modulo-Sum Simple Relay Network," submitted to *IEEE Trans. on Inf. Theory* in September 2009; available at website: http://webs.wichita.edu/?u=ECE&p=/Wireless/Publications/.

APPENDICES

APPENDIX A

PROOF: ACHIEVABILITY OF THEOREM 1

This paper uses the same steps as those used by Aleksic et al. [8] and modifies them for Figure 4. This appendix proves the achievability of Theorem 1 as follows:

1. *Generation of Codebook.* With the best input distribution of $X \sim Ber(1/2)$, 2^{nR} number of codewords $X^n(w)$ are independently and identically generated, indexed with $w \in \{1, 2, \ldots\ldots\ldots, 2^{nR}\}$, and used at the source. Similarly, with the best $p(u|y_1)$ distribution satisfying $I(U; Y_1) \leq R_0$, $2^{nI(U;Y_1)}$ number of codewords $U^n(v)$ are independently and identically generated, indexed with $v \in \{1, 2, \ldots\ldots\ldots, 2^{nI(U;Y_1)}\}$, and used at the relay.
2. *Encoding*. The two generated codebooks, $\mathcal{X}^n$ and $\mathcal{U}^n$, are revealed to (source and destination) and (relay and destination), respectively. Now, to send the ith block of the source, the channel encoder at the source sends codeword $X^n(w_i)$ with index w_i to both the destination and the relay node. The relay finds the jointly strongly typical sequence $U^n(v_i)$ from its codebook $\mathcal{U}^n$, using the previously observed ${Y_1}^n(i-1)$ word, which is a corrupted version of $X^n(w_{i-1})$. And the relay encodes the index v_i, and transmits it to the destination.
3. *Decoding.* The destination, upon receiving the two sequence information at different time slots from the source and the relay, decodes separately. After decoding the received words for index v_i, the destination looks for the jointly strongly typical sequence $X^n(w_{i-1})$ with $U^n(v_i)$ and $Y^n(i-1)$.

4. *Analysis of Probability of Error.* The decoding of the received word $Y^n(w_i)$ and $S_0^n(v_i)$ are done at different time slots, i.e., when $X^n(w_i)$ is transmitted, the destination receives $Y^n(w_i)$, but not from the relay. For the next time slot, the relay transmits index v_i to the destination, and the transmitter sends the codeword $X^n(w_{i+1})$ to the destination simultaneously. Now, the destination for the second time slot receives the two vectors $Y^n(w_{i+1})$ and $S_0^n(v_i)$. Therefore, for the first block transmission, the destination receives only from the direct channel. And for the last block, no message is transmitted from the source but rather only from the relay to the destination. Here, the last block from the source can be neglected for the rate calculation as the number of blocks $\rightarrow \infty$.

For the error probability analysis, this paper considers the average error probability over all codewords generated. Since the codebook construction is symmetric, the average error probability does not depend on a particular index sent. Therefore, it is assumed that $X^n(1)$ was transmitted. Then, this paper applies a jointly strongly typical decoding strategy. An error event can happen when either the triple $(X^n(1),\ Y^n(w_i)\ , U^n(v_i))$ is not jointly strongly typical or the received words $(Y^n(w_i),\ U^n(v_i))$ are jointly strongly typical with an incorrect codeword $X^n(w_i \neq 1)$. The probability of mapping a correct codeword into an incorrect codeword is $2^{-nI(X;Y,U)}$ [1, p. 327]. This event is considered an "undetected codeword error event," denoted by E_i. The probability of event E_i is given as [1, Lemma 10.6.2]

$$P(E_i) < 2^{nR}2^{-n(I(X;Y,U)-\epsilon_1)}. \tag{9}$$

If n is sufficiently large and $R < I(X;Y,U) - \epsilon_1$, then the average error probability can be made arbitrarily small, i.e., $P(E) < \varepsilon$. Hence, it requires

$$R < I(X;Y,U). \tag{10}$$

Since $X \sim \mathrm{Ber}(1/2)$, the mutual information is already maximized over $p(x)$. From the chain rule [1, p. 34],

$$I(X;Y,U) = I(X;Y|U) + I(X;U). \tag{11}$$

This can be rewritten as

$$I(X;Y,U) = H(Y|U) - H(Y|X,U) + H(X) - H(X|U). \tag{12}$$

Even if $X \sim \mathrm{Ber}(1/2)$, it can be shown that Y and Y_1 are not independent, and hence Y and U are dependent. Therefore, $H(Y|U)$ in equation (12) cannot be simplified to $(Y) = 1$, as done by Aleksic et al. [8]. Equation (12) can be rewritten as

$$I(X;Y,U) = H(Y|U) - H(Y|X,U) + 1 - H(X|U) \tag{13}$$

$$= H(Y|U) - H(X+Z|X,U) + 1 - H(X|U) \tag{14}$$

$$= H(Y|U) - H(Z|X,U) + 1 - H(X|U) \tag{15}$$

$$\overset{(a)}{=} 1 + H(Y|U) - H(Z|X) - H(X|U) \tag{16}$$

$$\overset{(b)}{=} 1 + H(Y|U) - H(Z) - H(X|U) \tag{17}$$

where

(a) follows from the fact that Z and U are independent; and

(b) follows from the fact that Z and X are independent.

Hence,

$$C = \max_{p(u|y_1):I(U;Y_1)\leq R_0} \{1 + H(Y|U) - H(Z) - H(X|U)\}. \tag{18}$$

The capacity can be expressed in a closed-form through which it is simple to calculate the capacity for given input and noise distributions. The capacity can now be evaluated as

$$C = 1 + \mathcal{H}(\{\mathcal{H}^{-1}(1-R_0) * \delta\} * p) - H(Z) - \mathcal{H}(\mathcal{H}^{-1}(1-R_0) * \delta). \tag{19}$$

Proof of equation (19) is presented as follows: The mutual information between U and Y_1 in Figure 4 is given as

$$I(U;Y_1) = H(Y_1) - H(Y_1|U) \tag{20}$$

Applying the constraint $I(U;Y_1) \leq R_0$ in equation (18) to equation (20), it can be written as

$$H(Y_1|U) \geq H(Y_1) - R_0 \tag{21}$$

From Corollary 4 [10], called "Mrs. Gerber's Lemma," if $H(X_0|U) \geq v$, then $H(Y_0|U) \geq \mathcal{H}\left(p_0 * \mathcal{H}^{-1}(v)\right)$, where X_0 and Y_0 are the binary random input and output random variables of a BSC with crossover probability p_0, respectively [9]. Equality holds if and only if code symbols X_0 are independent of each other. Now, this Corollary 4 is applied for the case of input Y_1 and output X in Figure 4 as

$$X = Y_1 + N_1 \ (mod\ 2) = Y_1 \oplus N_1. \tag{22}$$

Since $X \sim \text{Ber}(1/2)$ and $N_1 \sim \text{Ber}(\delta)$, Y_1 and N_1 are independent, if

$$H(Y_1|U) \geq \alpha, \tag{23}$$

then

$$H(X|U) \geq \mathcal{H}(\mathcal{H}^{-1}(\alpha) * \delta) \tag{24}$$

Equality in equation (24) holds if and only if $Y_1|U \sim \text{Ber}(\beta = \mathcal{H}^{-1}(\alpha))$. Here, $\alpha \triangleq H(Y_1) - R_0$ from equation (21). And the standard rate-distortion theory [8, p. 925] is applied, which says that "$Y_1|U \sim Ber(\mathcal{H}^{-1}(H(Y_1) - R_0))$, *is precisely the U that minimizes the Hamming distortion of* Y_1 *under a rate constraint* R_0." With this, the minimum achievable average distortion β under rate constraint R_0 must satisfy $\mathcal{H}(\beta) = H(Y_1|U) = H(Y_1) - R_0$, and hence, $Y_1|U$ must be $\text{Ber}(\beta)$ [8, p. 925]. Also, since $X \sim \text{Ber}(1/2)$, then Y_1 is $\text{Ber}(1/2)$. Therefore, the optimal distribution of U is also $\text{Ber}(1/2)$. Hence, $H(Y_1) = 1$ and $\alpha = 1 - R_0$. By substituting α into equation (24), it can be written for the equality case as

$$H(X|U) = \mathcal{H}(\mathcal{H}^{-1}(1 - R_0) * \delta) \tag{25}$$

By substituting equation (25) into equation (1), (equation 1) can be rewritten as

$$C = 1 + H(Y|U) - H(Z) - \mathcal{H}(\mathcal{H}^{-1}(1 - R_0) * \delta) \tag{26}$$

Again, "Mrs. Gerber's Lemma" is applied for $H(Y|U)$ in equation (26). In other words, if $H(X|U) \geq v$, then $H(Y|U) \geq \mathcal{H}(p * \mathcal{H}^{-1}(v))$ with equality, if and only if X given U is a Ber$(\mathcal{H}^{-1}(v))$, i.e., $H(X|U) = \mathcal{H}(\mathcal{H}^{-1}(1 - R_0) * \delta)$. Therefore,

$$H(Y|U) = \mathcal{H}(p * \{\mathcal{H}^{-1}(1 - R_0) * \delta\}) \tag{27}$$

By substituting equation (27) into equation (26), equation (26) can be rewritten as

$$C = 1 + \mathcal{H}(p * \{\mathcal{H}^{-1}(1 - R_0) * \delta\}) - H(Z) - \mathcal{H}(\mathcal{H}^{-1}(1 - R_0) * \delta) \tag{28}$$

Equation (28) can be simplified further as follows: From equation (2), R_0 is the capacity of a BSC between X_1 and S_0 with crossover probability ε. Using the uniform distribution for input X_1, equation (2) can be rewritten as

$$R_0 = 1 - \mathcal{H}(\varepsilon) \tag{29}$$

where $N_2 \sim \text{Ber}(\varepsilon)$. Now, by substituting equation (29) into equation (28), equation (28) can be rewritten as

$$C = 1 + \mathcal{H}(\{\mathcal{H}^{-1}(1 - [1 - \mathcal{H}(\varepsilon)]) * \delta\} * p) - H(Z) - \mathcal{H}(\mathcal{H}^{-1}(1 - [1 - \mathcal{H}(\varepsilon)]) * \delta) \tag{30}$$

$$= 1 + \mathcal{H}(\{\varepsilon * \delta\} * p) - H(Z) - \mathcal{H}(\mathcal{H}^{-1}\big(\mathcal{H}(\varepsilon)\big) * \delta) \tag{31}$$

$$= 1 + \mathcal{H}(\{\varepsilon * \delta\} * p) - \mathcal{H}(p) - \mathcal{H}(\varepsilon * \delta) \tag{32}$$

This completes the proof of the achievability for Theorem 1.

APPENDIX B

PROOF: CONVERSE (REVERSE) OF THEOREM 1

The proof of the converse is close to that found by Cover and Thomas [1, Theorem 15.8.1] or Aleksic et al. [8, Theorem 1]. However, a detailed proof is provided here again because Figure 4 in this current paper has different characteristics from those in Figure 15.32 of the former [1] or Figure 2 of the latter [8]. First, this paper derives Lemma 1.

Lemma 1: Let (X^n, Z^n, N_1^n, N_2^n) are be independently and identically distributed (i.i.d.) random variables of each other, with distributions of $X \sim Ber(1/2)$, $Z \sim Ber(p)$, $N_1 \sim Ber(\delta)$, and $N_2 \sim Ber(\varepsilon)$, as shown in Figure 4. And let the input at the relay and the received word at the destination be written as $Y_1^n = X^n + N_1^n$ and $S_0^n = X_1^n + N_2^n$, respectively. Then, the following inequality holds for any encoding scheme applied at the relay with the constraint $I(U; Y_1) \leq R_0$:

$$H(X^n|S_0^n) \geq \min_{p(u|y_1): I(U;Y_1) \leq R_0} nH(X|U) \tag{33}$$

where $U = (U_Q, Q)$ is an auxiliary random vector satisfying $|\mathcal{U}| \leq |\mathcal{Y}_1| + 2$, and Q is a time sharing random variable with i.i.d. over $\{1, 2, \dots, n\}$.

Proof for Lemma 1: The proof for Lemma 1 follows the same steps as done by Aleksic et al. [8, Lemma 1]. To prove Lemma 1 in this current paper for any encoding scheme with the constraint, it is necessary to show that $H(X^n|S_0^n) \geq nH(X|U)$ and $I(U; Y_1) \leq R_0$, which conclude equation (33).

From the chain rule, the left part of equation (33) can be written as

$$H(X^n|S_0^n) = \sum_{i=1}^n H(X_i|S_0^n, X_1, X_2, \dots\dots, X_{i-1}) \tag{34}$$

$$\stackrel{(a)}{=} \sum_{i=1}^n H\left(X_i\middle|S_0^n, X^{i-1}\right) \tag{35}$$

$$\stackrel{(b)}{\geq} \sum_{i=1}^n H\left(X_i\middle|S_0^n, X^{i-1}, Y_1^{i-1}\right) \tag{36}$$

$$\overset{(c)}{\geq} \sum_{i=1}^{n} H\left(X_i \middle| S_0^n, Y_1^{i-1}\right) \tag{37}$$

where

(a) follows from the definition of $X^{i-1} \triangleq X_1, X_2, \ldots, X_{i-1}$;

(b) follows from the conditional entropy property, i.e., $H(X) \geq H(X|Y)$; and

(c) follows from the fact that $X^{i-1} \rightarrow (S_0^n, Y_1^{i-1}) \rightarrow X_i$ forms a Markov chain because X_i and X^{i-1} are independent, and hence, $X^{i-1}$and X_i are conditionally independent for any given encoding scheme $U_i \triangleq (S_0^n, Y_1^{i-1})$. Hence, X_i is only affected by X^{i-1} through $\left(S_0^n, Y_1^{i-1}\right)$.

Then, from equation (37),

$$H(X^n|S_0^n) \geq \sum_{i=1}^{n} H(X_i|U_i) \tag{38}$$

From Figure 4, $X^n \rightarrow Y_1^n \rightarrow X_1^n \rightarrow S_0^n$ forms a Markov chain. Using the data processing inequality,

$$I(X_1^n; S_0^n) \geq I(Y_1^n; S_0^n) \tag{39}$$

$$\overset{(a)}{=} \sum_{i=1}^{n} I\left(Y_{1i}; S_0^n \middle| Y_1^{i-1}\right) \tag{40}$$

$$\overset{(b)}{=} \sum_{i=1}^{n} I(Y_{1i}; S_0^n, Y_1^{i-1}) - \sum_{i=1}^{n} I(Y_{1i}; Y_1^{i-1}) \tag{41}$$

$$\overset{(c)}{=} \sum_{i=1}^{n} I(Y_{1i}; S_0^n, Y_1^{i-1}) \tag{42}$$

$$\geq \sum_{i=1}^{n} I(Y_{1i}; U_i) \tag{43}$$

where

(a) and (b) follow from the chain rule and from the definition of $Y_1^{i-1} \triangleq Y_{11}, Y_{12}, \ldots, Y_{1i-1}$; and

(c) follows from the fact that Y_{1i} and Y_1^{i-1} are independent, i.e., $I\left(Y_{1i}; Y_1^{i-1}\right) = 0$ because $X \sim Ber(1/2)$,; hence, $Y_{1i} \sim Ber(1/2)$ too, and (Y_{1i}, N_1) is an independent pair, regardless of the noise distribution $N_1 \sim Ber(\delta)$.

The lemma by Cover and Thomas [1, §7.9.2] says that the capacity of a discrete memoryless channel is lower bounded by $I(X^n; Y^n) \leq nC$. Applying this to the channel between the relay and the destination, i.e., X_1 and S_0, the capacity can be lower-bounded as $I(X_1^n; S_0^n) \leq nR_0$. Substituting this inequality into equation (43),

$$nR_0 \geq I(X_1^n; S_0^n) \geq \sum_{i=1}^{n} I(Y_{1i}; U_i) \tag{44}$$

or

$$R_0 \geq \frac{1}{n} \sum_{i=1}^{n} I(Y_{1i}; U_i). \tag{45}$$

From equation (38),

$$\frac{1}{n} H(X^n | S_0^n) \geq \frac{1}{n} \sum_{i=1}^{n} H(X_i | U_i) . \tag{46}$$

From the work of Cover and Thomas[1, p. 578], a time-sharing uniform random variable Q distributed over $\{1, 2, \ldots., n\}$ is introduced, which is independent of (X_i, Y_{1i}, U_i). Therefore, $I(Y_{1i}; U_i) = I(Y_{1i}; U_i | Q = i)$ and $H(X_i | U_i) = H(X_i | U_i, Q = i)$. These are applied to equations (45) and (46), which can be rewritten as

$$R_0 \geq \frac{1}{n} \sum_{i=1}^{n} I(Y_{1i}; U_i | Q = i) = I\left(Y_{1Q}; U_Q \middle| Q\right) \tag{47}$$

$$\frac{1}{n} H(X^n | S_0^n) \geq \frac{1}{n} \sum_{i=1}^{n} H(X_i | U_i, Q = i) = H\left(X_Q \middle| U_Q, Q\right) \tag{48}$$

From the chain rule, the right-hand side of equation (47) can be rewritten as

$$I\left(Y_{1Q}; U_Q \middle| Q\right) = I\left(Y_{1Q}; U_Q, Q\right) - I(Y_{1Q}; Q) \tag{49}$$

$$= I\left(Y_{1Q}; U_Q, Q\right) \tag{50}$$

where $I(Y_{1Q}; Q) = 0$ in equation (49) because Y_{1Q} and Q are independent. The pair (Y_{1Q}, X_Q) have the same joint distribution as (Y_1, X). Define $X \triangleq X_Q$, $Y_1 \triangleq Y_{1Q}$, and $U \triangleq (U_Q, Q)$. Then, substitute these and equation (50) into equations (47) and (48), which can be rewritten as

$$R_0 \geq I(Y_1; U) \tag{51}$$

$$H(X^n|S_0^n) \geq nH(X|U). \tag{52}$$

This completes the proof of Lemma 1.

To prove the converse theorem, it is necessary to show that $R \leq C$, if the average codeword error probability $P^{(n)}(E) \to 0$ for sufficiently large n. First, construct a codebook $\mathcal{U}^n$, where each codeword is i.i.d. and satisfies the constraint $I(U; Y_1) \leq R_0$. The cardinality bound is the same as the one used by Cover and Thomas [1, Theorem15.8.1], which can be proved using the Fano's inequality [1]. Let $W \in \{1, 2, \ldots, 2^{nR}\}$ be the source message random variable, which carries the input message. Because W is a uniform random variable,

$$nR = H(W) \tag{53}$$

$$\overset{(a)}{=} I(W; Y^n, S_0^n) + H(W|Y^n, S_0^n) \tag{54}$$

$$\overset{(b)}{\leq} I(W; Y^n, S_0^n) + n\varepsilon_n \tag{55}$$

$$\overset{(c)}{\leq} I(X^n; Y^n, S_0^n) + n\varepsilon_n \tag{56}$$

$$= I(X^n; Y^n|S_0^n) + I(X^n; S_0^n) + n\varepsilon_n \tag{57}$$

$$= H(Y^n|S_0^n) - H(Y^n|X^n, S_0^n) + H(X^n) - H(X^n|S_0^n) + n\varepsilon_n \tag{58}$$

$$= H(Y^n|S_0^n) - H(Y^n = X^n + Z^n|X^n, S_0^n) + H(X^n) - H(X^n|S_0^n) + n\varepsilon_n \tag{59}$$

$$\overset{(d)}{=} H(Y^n|S_0^n) - H(Z^n|X^n, S_0^n) + H(X^n) - H(X^n|S_0^n) + n\varepsilon_n \tag{60}$$

$$\overset{(e)}{=} H(Y^n|S_0^n) - H(Z^n) + H(X^n) - H(X^n|S_0^n) + n\varepsilon_n \tag{61}$$

$$\overset{(f)}{\leq} nH(Y|U) - nH(Z) + n - nH(X|U) + n\varepsilon_n \tag{62}$$

$$= n\{H(Y|U) - H(Z) + 1 - H(X|U)\} + n\varepsilon_n \tag{63}$$

$$\overset{(g)}{\leq} nC + n\varepsilon_n \tag{64}$$

$$R \leq C + \varepsilon_n \tag{65}$$

where

(a) follows from the definition of mutual information;

(b) follows from the Fano's inequality, i.e., $H(W|Y^n, S_0^n) \leq n\varepsilon_n$ because $P^{(n)}(E) \to 0$;

(c) follows from the data processing inequality;

(d) follows from the fact that X^n is given;

(e) follows from the fact that Z^n is independent of (X^n, S_0^n);

(f) follows from the fact that $H(Y^n|S_0^n) = nH(Y|U)$ because Y^n is i.i.d., and the inequality $H(X^n|S_0^n) \geq nH(X|U)$ holds from Lemma 1; and

(g) follows by substituting the capacity in equation (1) into equation (63).

This completes the proof for the converse of Theorem 1.

APPENDIX C

PROOF OF THEOREM 2: CUT-SET BOUND

From the multiple-access cut in Figure 4,

$$I(X, X_1; Y, S_0) \overset{(a)}{=} H(Y, S_0) - H(Y, S_0|X, X_1) \tag{66}$$

$$\overset{(b)}{=} H(S_0) + H(Y|S_0) - H(Y = X + Z, S_0 = X_1 + N_2|X, X_1) \tag{67}$$

$$\overset{(c)}{=} H(S_0) + H(Y|S_0) - H(Z, N_2) \tag{68}$$

$$\overset{(d)}{\leq} 1 + 1 - H(Z) - H(N_2) \tag{69}$$

$$= 1 - H(Z) + 1 - H(N_2) \tag{70}$$

$$\overset{(e)}{\leq} 1 - \mathcal{H}(p) + R_0 \tag{71}$$

where

(a) follows from the definition of mutual information;

(b) follows from the chain rule;

(c) follows from the fact that Y and S_0 are functions of (X, Z) and (X_1, N_2), respectively;

(d) follows from the fact that $H(S_0) \leq 1$ and $H(Y|S_0) \leq 1$ because the entropy of a binary random variable is upper-bounded by 1; $H(Z, N_2) = H(Z) + H(N_2)$ because the pair (Z, N_2) are independent; and

(e) follows from (29) and $Z \sim Ber(p)$.

From the broadcast cut in Figure 4,

$$I(X; Y, Y_1) \overset{(a)}{=} I(X; Y) + I(X; Y_1|Y) \tag{72}$$

$$\overset{(b)}{=} I(X; Y) + H(Y_1|Y) - H(Y_1|Y, X) \tag{73}$$

$$= I(X; Y) + H(Y_1|Y) - H(Y_1 = X + N_1|Y, X) \tag{74}$$

$$\overset{(c)}{=} I(X;Y) + H(Y_1|Y) - H(N_1|Y,X) \tag{75}$$

$$\overset{(d)}{=} I(X;Y) + H(Y_1|Y) - H(N_1) \tag{76}$$

$$\overset{(e)}{\leq} 1 - \mathcal{H}(p) + 1 - \mathcal{H}(\delta) \tag{77}$$

where

(a) follows from the chain rule;

(b) follows from the definition of mutual information;

(c) follows from $H(Y_1 = X + N_1|X) = H(N_1)$;

(d) follows from the fact that N_1 is independent of the pair (Y, X);

(e) follows from the fact that $I(X;Y)$=$1 - \mathcal{H}(p)$, and $H(Y_1|Y) \leq 1$, and $H(N_1) = 1 - \mathcal{H}(\delta)$.

From (5), (71), and (77), the final cut-set bound for this particular channel is equal to

$$\min\{1 - \mathcal{H}(p) + R_0,\ 1 - \mathcal{H}(p) + 1 - \mathcal{H}(\delta)\}. \tag{78}$$

Therefore, this completes the proof of Theorem 2 for the cut-set bound.

APPENDIX D

MATLAB CODE

```
clc;
clear;
r0 = 0;
delt = 0.1;
 temphin = zeros(1,11);
zp=0.5
hz = -zp*log2(zp)-(1-zp)*log2(1-zp);
for i=1:11
he = 1-r0;
p1 =0;
 e1 = 0.0002;
    for j=1:10000
        hp1(j) = -p1*log2(p1)-(1-p1)*log2(1-p1);

        if ((he) <= hp1(j)+e1)&&((he) >= hp1(j)-e1);
            temphin(i) = p1;
            break
         end
        p1 = p1+0.00005;
    end


    edelta=(temphin(i)*(1-delt))+((1-temphin(i))*delt);

    edeltp=(edelta*(1-zp))+((1-edelta)*zp);
hedeltp(i) = -edeltp*log2(edeltp)-(1-edeltp)*log2(1-edeltp);
hp(i) = -edelta*log2(edelta)-(1-edelta)*log2(1-edelta);

c(i) = 1-hp(i)-hz+hedeltp(i);
%===============CUT-SET BOUND=======================
hdelta(i) = -delt*log2(delt)-(1-delt)*log2(1-delt);
cbroad = 1-hz+1-hdelta;

cmultiacc(i) = 1-hz+r0;
%=================================================
plotr0(i) = r0;
r0=r0+0.1;
end
plot(plotr0,c,'-');
grid on
hold on
plot(plotr0,cbroad,'--g');
plot(plotr0,cmultiacc,'--r');
xlabel('R0 (Bits)')
ylabel('Capacity (Bits)')
legend('Capacity', 'Broadcast Bound', 'Multiple-access Bound',2)
```