A New Upper Bound on the Capacity of a Class of Primitive Relay Channels

Ravi Tandon Sennur Ulukus
Department of Electrical and Computer Engineering
University of Maryland, College Park, MD 20742
ravit@umd.edu *ulukus@umd.edu*

***Abstract*— We obtain a new upper bound on the capacity of a class of discrete memoryless relay channels. For this class of relay channels, the relay observes an i.i.d. sequence T, which is independent of the channel input X. The channel is described by a set of probability transition functions $p(y|x,t)$ for all $(x,t,y) \in \mathcal{X} \times \mathcal{T} \times \mathcal{Y}$. Furthermore, a noiseless link of finite capacity R_0 exists from the relay to the receiver. Although the capacity for these channels is not known in general, the capacity of a subclass of these channels, namely when $T = g(X,Y)$, for some deterministic function g, was obtained in [1] and it was shown to be equal to the cut-set bound. Another instance where the capacity was obtained was in [2], where the channel output Y can be written as $Y = X \oplus Z$, where $\oplus$ denotes modulo-m addition, Z is independent of X, $|\mathcal{X}| = |\mathcal{Y}| = m$, and T is some stochastic function of Z. The compress-and-forward (CAF) achievability scheme [3] was shown to be capacity achieving in both cases.**

Using our upper bound we recover the capacity results of [1] and [2]. We also obtain the capacity of a class of channels which does not fall into either of the classes studied in [1] and [2]. For this class of channels, CAF scheme is shown to be optimal but capacity is strictly less than the cut-set bound for certain values of R_0. We further illustrate the usefulness of our bound by evaluating it for a particular relay channel with binary multiplicative states and binary additive noise for which the channel is given as $Y = TX + N$. We show that our upper bound is strictly better than the cut-set upper bound for certain values of R_0 but it lies strictly above the rates yielded by the CAF achievability scheme.

I. Introduction

The relay channel is one of the simplest, yet arguably among the least understood multi-user channels in information theory. A special class of discrete memoryless relay channel is the primitive relay channel [1]. For this class, the channel is defined by a channel input X, a channel output Y and a relay output T, and a set of probability functions $p(y,t|x)$ for all $x \in \mathcal{X}$. In this setting, the relay does not have an explicit coded input for the channel. Moreover, it is also assumed that there is an orthogonal link of finite capacity R_0, from the relay to the receiver. Zhang [4] considered this relay channel and obtained a partial converse for a degraded case. For a comprehensive survey on related work on primitive relay channels, see [5].

Recently, Kim [1] established the capacity of a class of semi-deterministic primitive relay channels, for which the relay output T can be expressed as a deterministic function of the channel input X and the channel output Y, i.e., $T = g(Y,X)$. The cut-set upper bound [6] was shown to be the capacity through an algebraic reduction of the compress-and-forward (CAF) achievable rate [3] to the cut-set upper bound. This was the first instance where the CAF achievability scheme was shown to be capacity achieving for any relay channel.

In this paper, we consider a subclass of the primitive relay channel. In this subclass, the relay observes an i.i.d. sequence T which is independent of the channel input X and the channel output Y is given by the set of probability transition functions $p(y|x,t)$ for all $(x,t,y) \in \mathcal{X}\times\mathcal{T}\times\mathcal{Y}$. Alternatively, this channel can be interpreted as a state dependent discrete memoryless channel with rate-limited state information available at the receiver (Figure 1). This channel was also studied in [7] with various modifications regarding the rate-limited knowledge of the channel state T at the transmitter and the receiver. A CAF achievability scheme for this state dependent channel was given by Ahlswede and Han in [8] and it was conjectured to be the capacity for this class of channels. In fact, the same achievable rates for this channel were obtained in [7] and can also be obtained via Theorem 6 of [3].

It follows from the result of [1] that this conjecture is true for the subclass when the state T can be expressed as a deterministic function of X and Y, i.e., $T = g(X,Y)$. An example of such a channel is the case when X, T and Y are all binary, $T \sim \text{Ber}(\delta)$ and independent of X, and the channel is given by $Y = X \oplus T$, where $\oplus$ denotes modulo-2 addition. Note that, in this case, T is a deterministic function of X and Y, since $T = X \oplus Y$. A capacity result following up on the aforementioned modulo-additive noise channel was obtained in [2], where it was assumed that the receiver observes $Y = X \oplus Z$ and the relay observes a noisy version of the forward noise, i.e., $T = Z \oplus \tilde{Z}$. Clearly, if $\tilde{Z} = 0$, then this channel reduces to the class studied in [1]. However, when $\tilde{Z} \neq 0$, T cannot be written as a deterministic function of X and Y, and this modulo-additive class lies outside of the class of channels considered in [1]. By proving a converse, it was shown in [2] that CAF scheme is capacity achieving for this modulo-additive case. The remarkable fact was that the capacity was shown to be strictly less than the cut-set upper bound for certain values of R_0. However, it is worth noting that the converse proved in [2] relied heavily on the modulo-additive nature of the forward channel.

This work was supported by NSF Grants CCF 04-47613, CCF 05-14846, CNS 07-16311 and CCF 07-29127.

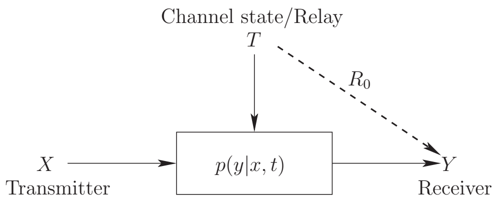

Figure 1: Channel with rate-limited state information.

In this paper, we obtain a new upper bound on the capacity of the state-dependent discrete memoryless channel, where the states are i.i.d. and the state information is available to the receiver through a noiseless link of finite capacity R_0. Our upper bound serves a dual purpose. Firstly, using our upper bound, we recover the capacity results obtained in [1] for the case where $T = g(X, Y)$ and the capacity result obtained in [2] for the modulo-additive noise case. Secondly, we confirm the validity of the conjecture due to Ahlswede-Han [8] for another class of channels which does not fall into any of the cases considered in [1] and [2].

To further illustrate the application of our upper bound, we consider a channel where X, T, N are binary and Y is ternary and the channel is given by $Y = TX + N$, i.e., when the state sequence is binary and multiplicative and there is additive binary noise at the receiver. This channel can be interpreted as the discrete analogue of a fast fading channel with fade information available in a rate-limited fashion at the receiver. We evaluate our upper bound for this channel and show that it is strictly less than the cut-set bound for certain values of R_0 although our upper bound is strictly larger than the rates yielded by the CAF scheme.

II. Relay Channel Model

We consider a relay channel with finite input alphabet $\mathcal{X}$, finite output alphabet $\mathcal{Y}$ and finite relay output alphabet $\mathcal{T}$. Moreover, the relay observes an i.i.d. state sequence $T^n \in \mathcal{T}^n$ with some given probability distribution $p(t)$. The relay channel is described by the set of transition probabilities $p(y|x, t)$ which are defined for all $(x, t, y) \in \mathcal{X} \times \mathcal{T} \times \mathcal{Y}$. Furthermore, there is a finite-capacity noiseless link of capacity R_0 from the relay to the receiver.

An (n, M, P_e) code for this relay channel consists of the set of integers $\mathcal{M} = \{1, 2, \dots, M\}$ and the following:

$$\begin{aligned} &f_t : \mathcal{M} \to \mathcal{X}^n, \quad f_r : \mathcal{T}^n \to \{1, 2, \dots, L\} \\ &\phi : \mathcal{Y}^n \times \{1, 2, \dots, L\} \to \mathcal{M} \end{aligned} \tag{1}$$

where f_t is the transmitter encoding function, f_r is the relay encoding function and g is the decoding function. Furthermore, as the relay to receiver link is of limited capacity R_0, we have $L \leq 2^{nR_0}$. For a distribution $p(w)$ on $\mathcal{M}$, the joint probability distribution on $\mathcal{M} \times \mathcal{X}^n \times \mathcal{T}^n \times \mathcal{Y}^n$ is given as

$$p(w, x^n, t^n, y^n) = p(w)p(x^n|w) \prod_{i=1}^{n} p(t_i) \prod_{i=1}^{n} p(y_i|x_i, t_i)$$

For a uniform distribution $p(w)$ on $\mathcal{M}$, the average probability of error is given as, $P_e = \Pr(\phi(Y^n, f_r(T^n)) \neq W)$. A rate R is achievable if for any $\epsilon > 0$ and all n sufficiently large, there exists an (n, M, P_e) code such that $P_e \leq \epsilon$ and $M \geq 2^{nR}$. The capacity of the relay channel is the supremum of the set of all achievable rates.

III. A New Upper Bound on the Capacity

We will denote by U^n as the output of the finite capacity link R_0, i.e., $U^n = f_r(T^n)$. We will now obtain an upper bound on the rate as follows,

$$\begin{aligned}
nR &= H(W) && (2)\\
&= I(W; Y^n, U^n) + H(W|Y^n, U^n) && (3)\\
&\leq I(W; Y^n, U^n) + n\epsilon_n && (4)\\
&\leq I(X^n; Y^n, U^n) + n\epsilon_n && (5)\\
&= I(X^n; Y^n|U^n) + n\epsilon_n && (6)\\
&= \sum_{i=1}^{n} I(X^n; Y_i|U^n, Y^{i-1}) + n\epsilon_n && (7)\\
&= \sum_{i=1}^{n} \Big[H(Y_i|U^n, Y^{i-1}) - H(Y_i|U^n, Y^{i-1}, X^n)\Big] + n\epsilon_n && (8)\\
&\leq \sum_{i=1}^{n} \Big[H(Y_i) - H(Y_i|U^n, Y^{i-1}, X^n)\Big] + n\epsilon_n && (9)\\
&\leq \sum_{i=1}^{n} \Big[H(Y_i) - H(Y_i|U^n, T^{i-1}, Y^{i-1}, X^n)\Big] + n\epsilon_n && (10)\\
&= \sum_{i=1}^{n} \Big[H(Y_i) - H(Y_i|U^n, T^{i-1}, X^n)\Big] + n\epsilon_n && (11)\\
&= \sum_{i=1}^{n} \Big[H(Y_i) - H(Y_i|U^n, T^{i-1}, X_i)\Big] + n\epsilon_n && (12)\\
&= \sum_{i=1}^{n} I(X_i, U^n, T^{i-1}; Y_i) + n\epsilon_n && (13)\\
&= \sum_{i=1}^{n} I(X_i, V_i; Y_i) + n\epsilon_n && (14)\\
&= nI(X, V; Y) + n\epsilon_n && (15)
\end{aligned}$$

where (4) follows by Fano's inequality [6], (5) follows from the data processing inequality, (6) follows from the fact that X^n is independent of T^n and is hence independent of U^n, (9) follows from the fact that conditioning reduces entropy and hence we upper bound by dropping (U^n, Y^{i-1}) from the first term. Next, (10) follows by adding T^{i-1} in the conditional entropy in the second term and obtaining an upper bound, (11) follows from the memoryless property of the channel, i.e., given (X^{i-1}, T^{i-1}), the channel output Y^{i-1} is independent of everything else and (12) follows from the following Markov chain, $X^n \setminus X_i \to (X_i, U^n, T^{i-1}) \to Y_i$. The proof of this Markov chain is given at the beginning of next page. Finally, (14) follows by defining $V_i = (U^n, T^{i-1})$, and we introduce a random variable Q, uniform on $\{1, 2, \dots, n\}$ to define $X = (X_i, Q)$, $Y = (Y_i, Q)$ and $V = (V_i, Q)$ to arrive at (15).

We define $X^{-i} \triangleq (X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X^n)$ and obtain the Markov chain by showing the following,

$$
\begin{aligned}
\Pr(Y_i, X^{-i}|X_i, U^n, T^{i-1}) &= \frac{\Pr(Y_i, X^{-i}, X_i, U^n, T^{i-1})}{\Pr(X_i, U^n, T^{i-1})} && (16)\\
&= \frac{\sum_{t_i} \mathrm{P}(t_i)\Pr(Y_i, X^{-i}, X_i, U^n, T^{i-1}|t_i)}{\Pr(X_i, U^n, T^{i-1})} && (17)\\
&= \frac{\sum_{t_i} \mathrm{P}(t_i)\Pr(X_i, U^n, T^{i-1}|t_i)\Pr(Y_i, X^{-i}|X_i, t_i, U^n, T^{i-1})}{\Pr(X_i, U^n, T^{i-1})} && (18)\\
&= \frac{\sum_{t_i} \mathrm{P}(t_i)\Pr(X_i, U^n, T^{i-1}|t_i)\Pr(X^{-i}|X_i, t_i, U^n, T^{i-1})\Pr(Y_i|X_i, t_i, U^n, T^{i-1}, X^{-i})}{\Pr(X_i, U^n, T^{i-1})} && (19)\\
&= \frac{\sum_{t_i} \mathrm{P}(t_i)\Pr(X_i, U^n, T^{i-1}|t_i)\Pr(X^{-i}|X_i)\Pr(Y_i|X_i, t_i, U^n, T^{i-1})}{\Pr(X_i, U^n, T^{i-1})} && (20)\\
&= \Pr(X^{-i}|X_i)\frac{\sum_{t_i} \mathrm{P}(t_i)\Pr(X_i, U^n, T^{i-1}|t_i)\Pr(Y_i|X_i, t_i, U^n, T^{i-1})}{\Pr(X_i, U^n, T^{i-1})} && (21)\\
&= \Pr(X^{-i}|X_i)\sum_{t_i} \mathrm{P}(t_i|X_i, U^n, T^{i-1})\Pr(Y_i|X_i, U^n, T^{i-1}, t_i) && (22)\\
&= \Pr(X^{-i}|X_i)\Pr(Y_i|X_i, U^n, T^{i-1}) && (23)
\end{aligned}
$$

In addition to (15), we also need the following trivial upper bound on the rate,

$$
\begin{aligned}
nR &\leq I(X^n; Y^n, T^n) + n\epsilon_n && (24)\\
&= I(X^n; Y^n|T^n) + n\epsilon_n && (25)\\
&= \sum_{i=1}^{n} I(X^n; Y_i|T^n, Y^{i-1}) + n\epsilon_n && (26)\\
&= \sum_{i=1}^{n} \Big[H(Y_i|T^n, Y^{i-1}) - H(Y_i|T^n, Y^{i-1}, X^n)\Big] + n\epsilon_n && (27)\\
&= \sum_{i=1}^{n} \Big[H(Y_i|T_i) - H(Y_i|T^n, Y^{i-1}, X^n)\Big] + n\epsilon_n && (28)\\
&= \sum_{i=1}^{n} \Big[H(Y_i|T_i) - H(Y_i|T_i, X_i)\Big] + n\epsilon_n && (29)\\
&= \sum_{i=1}^{n} I(X_i; Y_i|T_i) + n\epsilon_n && (30)\\
&= nI(X; Y|T) + n\epsilon_n && (31)
\end{aligned}
$$

where (24) follows by Fano's inequality, (25) follows because X^n is independent of T^n, (28) follows by dropping $(Y^{i-1}, T^n \setminus T_i)$ from the conditioning in the first term, (29) follows from the memoryless property of the channel, i.e., given (X_i, T_i), the channel output Y_i is independent of everything else.

We now obtain a bound on the allowable distributions of the involved random variables. Using the fact that the side information is limited by the rate R_0, we have that

$$
\begin{aligned}
nR_0 &\geq I(T^n; U^n) && (32)\\
&= \sum_{i=1}^{n} I(T_i; U^n|T^{i-1}) && (33)\\
&= \sum_{i=1}^{n} I(T_i; U^n, T^{i-1}) && (34)\\
&= nI(T; V) && (35)
\end{aligned}
$$

where (34) follows from the fact that T_i are i.i.d.

Combining (15), (31) and (35), we have an upper bound on the capacity of the relay channel as

$$
\begin{aligned}
\mathcal{UB} = \sup \min&\{I(X, V; Y), I(X; Y|T)\}\\
\text{s.t. } & R_0 \geq I(T; V)\\
\text{over } & p(x)p(t)p(v|t) && (36)
\end{aligned}
$$

where the supremum can be restricted over those V such that $|\mathcal{V}| \leq |\mathcal{T}| + 2$.

IV. Comparison with the Cut-set Bound

The best known upper bound for the relay channel is the cut-set bound [6], which reduces for the relay channel in consideration to [1], [5]

$$
\mathcal{CS} = \max_{p(x)} \min\{I(X; Y) + R_0, I(X; Y|T)\} \qquad (37)
$$

On comparing with the cut-set bound, it can be observed that our bound differs from the cut-set bound in the multiple access cut. We will show next that our upper bound is in general smaller than the cut-set bound.

We start by upper bounding the expression $I(X, V; Y)$ as follows,

$$
\begin{aligned}
I(X, V; Y) &= I(X; Y) + I(V; Y|X) && (38)\\
&= I(X; Y) + H(V|X) - H(V|Y, X) && (39)\\
&= I(X; Y) + H(V) - H(V|Y, X) && (40)\\
&\leq I(X; Y) + H(V) - H(V|T, Y, X) && (41)\\
&= I(X; Y) + H(V) - H(V|T) && (42)
\end{aligned}
$$

$$= I(X;Y) + I(T;V) \quad (43)$$
$$\leq I(X;Y) + R_0 \quad (44)$$

where (40) follows from the fact that V is independent of X, (41) follows from the fact that conditioning reduces entropy, (42) follows from the Markov chain $(X,Y) \rightarrow T \rightarrow V$ and (43) follows by using the fact that $I(T;V) \leq R_0$. Using (44) and (31), we have the following

$$\mathcal{UB} \leq \max_{p(x)} \min\{I(X;Y) + R_0, I(X;Y|T)\} \quad (45)$$

Thus, our upper bound is in general smaller than the cut-set bound given in (37). It was shown in [1] that the cut-set bound is tight for the case when $T = g(X,Y)$ and is achieved by the CAF achievability scheme. Note the fact that for this special subclass, the inequality in (41) is in fact an equality and our bound exactly equals the cut-set bound.

V. Recovering the Capacity of Modulo-Additive Relay Channel

A specific modulo-additive relay channel was considered in [2] for which the channel is given as,

$$Y = X \oplus Z \quad (46)$$
$$T = Z \oplus \tilde{Z} \quad (47)$$

where X, Y, T, Z and $\tilde{Z}$ are all binary and $Z \sim \text{Ber}(\delta)$, $\tilde{Z} \sim \text{Ber}(\tilde{\delta})$. Clearly this channel does not fall into the class of channels studied in [1], where T can be written as a deterministic function of X and Y. It was shown that the capacity of this channel is given by [2, Theorem 1]

$$\mathcal{C} = \max_{p(v|t):I(T;V)\leq R_0} 1 - H(Z|V) \quad (48)$$

We will show that our bound is equal to the capacity for this class of channels. First, note that

$$I(X,V;Y) = H(Y) - H(Y|X,V) \quad (49)$$
$$= H(Y) - H(Z|V) \quad (50)$$
$$\leq 1 - H(Z|V) \quad (51)$$

where (51) follows by the fact that the entropy of a binary random variable is upper bounded by 1. Next, consider the other cut,

$$I(X;Y|T) = H(Y|T) - H(Y|X,T) \quad (52)$$
$$= H(Y|T) - H(Z|T) \quad (53)$$
$$\leq 1 - H(Z|T) \quad (54)$$

We note that (51) and (54) are achieved with equality for a uniform X. Moreover, from (51) and (54), it can be observed that the bound $I(X;Y|T)$ is redundant since $V \rightarrow T \rightarrow Z$ implies $H(Z|T) \leq H(Z|V)$. Hence, our upper bound reduces to

$$\mathcal{UB} = \max_{p(v|t):I(T;V)\leq R_0} 1 - H(Z|V) \quad (55)$$

We should remark that the converse obtained in [2] for this channel utilized the modulo-additive nature of the channel. For such a channel, a uniform distribution on X makes the channel output Y independent of noise Z, thereby making the proceedings in the converse easier. Our upper bound does not rely on the nature of the channel and holds for any $p(y|x,t)$. We have thus shown that for all the cases where the capacity is established, our bound is tight. To illustrate the usefulness of our bound, we will consider a channel which does not fall into any of these classes.

VI. Capacity Result for a Symmetric Binary Erasure Channel with Two States

We will show that for a particular symmetric binary input erasure channel with two states, our upper bound yields the capacity which turns out to be strictly less than the cut-set bound. The state T is binary with $\Pr(T = 0) = \alpha$. The channel input X is binary and channel output Y is ternary. For channel states $T = 0, 1$, the transition matrices $p(y|x,t)$ are given as (Figure 2),

$$W_0 = \begin{bmatrix} 0 & 1-\epsilon & \epsilon \\ \epsilon & 1-\epsilon & 0 \end{bmatrix} \qquad W_1 = \begin{bmatrix} \epsilon & 1-\epsilon & 0 \\ 0 & 1-\epsilon & \epsilon \end{bmatrix}$$

It should be noted that this class of channels does not fall into the class of channels considered in [1] since T cannot be obtained as a deterministic function of X and Y. Moreover, the channel output Y cannot be expressed in the form as $Y = X \oplus Z$, for some $p(t|z)$, where $\oplus$ is modulo-2 addition, since the cardinality of Y is different from the cardinality of X. Hence, the converse technique developed in [2] for modulo-additive relay channels does not apply to this channel. However, our upper bound holds for any $p(y|x,t)$. We begin by evaluating the achievable rates given by the CAF scheme,

$$\begin{aligned} &\mathcal{C} \geq \sup I(X;Y|V) \\ &\quad \text{s.t. } I(T;V|Y) \leq R_0 \\ &\quad \text{for some } p(x,t,v) = p(x)p(t)p(v|t) \end{aligned} \quad (56)$$

Throughout this paper, we denote the entropy function as $h^{(k)}(s_1,\ldots,s_k) = -\sum_{i=1}^{k} s_i \log(s_i)$ where $s_i \geq 0$, $i = 1,\ldots,k$ and $\sum_i s_i = 1$. We will denote the binary entropy function as $h(s)$. We first define $\Pr(X = 0) = p$ and obtain the involved probabilities,

$$p(Y = 0) = \epsilon(\alpha * p) \quad (57)$$
$$p(Y = 1) = 1 - \epsilon \quad (58)$$
$$p(Y = 2) = \epsilon(1 - \alpha * p) \quad (59)$$

and

$$p(Y = 0|T = 0) = \epsilon(1 - p) \quad (60)$$
$$p(Y = 1|T = 0) = 1 - \epsilon \quad (61)$$
$$p(Y = 2|T = 0) = \epsilon p \quad (62)$$

and

$$p(Y = 0|T = 1) = \epsilon p \quad (63)$$
$$p(Y = 1|T = 1) = 1 - \epsilon \quad (64)$$
$$p(Y = 2|T = 1) = \epsilon(1 - p) \quad (65)$$

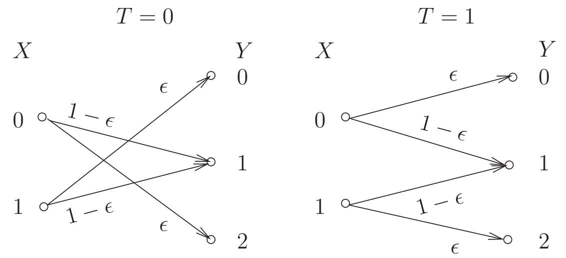

Figure 2: A symmetric binary erasure channel with two states.

where we have defined

$$a * b = a(1-b) + b(1-a) \tag{66}$$

Furthermore, we also note the following inequality,

$$\begin{aligned} h^{(3)}(a,b,c) &= \frac{1}{2}h^{(3)}(a,b,c) + \frac{1}{2}h^{(3)}(c,b,a) && (67)\\ &\le h^{(3)}\left(\frac{a+c}{2}, b, \frac{a+c}{2}\right) && (68)\\ &= h(b) + 1 - b && (69)\end{aligned}$$

Using this fact, we have

$$\begin{aligned} H(Y) &= h^{(3)}\left(\epsilon(\alpha * p), 1-\epsilon, \epsilon(1-\alpha * p)\right) && (70)\\ &\le h(\epsilon) + \epsilon && (71)\end{aligned}$$

Also, a uniform distribution on X, yields the maximum entropy for Y, and makes Y and T independent. Note that the maximum entropy of Y in this case is $h(\epsilon) + \epsilon$ which is strictly less than $\log(3)$ for all $\epsilon \in [0,1]$. Hence, for a uniform X, we have

$$\begin{aligned} H(Y|V) &= H(Y) && (72)\\ &= h(\epsilon) + \epsilon && (73)\end{aligned}$$

We also define,

$$\eta_v = \Pr(T = 1|V = v), \; v = 1, \ldots, |\mathcal{V}| \tag{74}$$

Using this definition, we can write $H(Y|X,V)$ for any distribution $p(x)$ on X as follows,

$$\begin{aligned} H(Y|X,V) &= \sum_v p(v) \sum_x p(x) H(Y|X=x, V=v) && (75)\\ &= \sum_v p(v) h^{(3)}\left(\eta_v \epsilon, 1-\epsilon, (1-\eta_v)\epsilon\right) && (76)\\ &= H(U|V) && (77)\end{aligned}$$

where we have defined a random variable U with $|\mathcal{U}| = 3$ and $p(u|t)$, expressed as a stochastic matrix W which is given as

$$W = \begin{pmatrix} \epsilon & 1-\epsilon & 0 \\ 0 & 1-\epsilon & \epsilon \end{pmatrix} \tag{78}$$

Thus, $H(Y|X,V)$ is invariant to the distribution of X. Moreover, by construction, the random variables (T,U,V) satisfy the Markov chain $V \to T \to U$.

We now return to the evaluation of the rates given by the CAF scheme given in (56). Using (73) and (77), we have for a uniform distribution on X,

$$\begin{aligned} I(X;Y|V) &= H(Y|V) - H(Y|X,V) && (79)\\ &= h(\epsilon) + \epsilon - H(U|V) && (80)\end{aligned}$$

Furthermore, for uniform X, we have $I(T;V|Y) = I(T;V)$, thus the constraint in (56) simplifies to $I(T;V) \le R_0$. For simplicity, define the set

$$\mathcal{L}(\gamma) = \{p(v|t) : H(T|V) \ge \gamma; \; V \to T \to U\} \tag{81}$$

Using (80) and (81), we obtain a lower bound on the capacity as

$$\mathcal{C} \ge h(\epsilon) + \epsilon - \inf_{p(v|t) \in \mathcal{L}(h(\alpha) - R_0)} H(U|V) \tag{82}$$

We now evaluate our upper bound. Using the following fact,

$$\min(I(X,V;Y), I(X;Y|T)) \le I(X,V;Y) \tag{83}$$

we obtain a weaker version of our upper bound in (36) as

$$\begin{aligned} \mathcal{C} &\le \sup I(X,V;Y) && (84)\\ &= \sup(H(Y) - H(Y|X,V)) && (85)\\ &\le \sup(h(\epsilon) + \epsilon - H(Y|X,V)) && (86)\\ &= h(\epsilon) + \epsilon - \inf H(Y|X,V) && (87)\\ &= h(\epsilon) + \epsilon - \inf_{p(v|t) \in \mathcal{L}(h(\alpha) - R_0)} H(U|V) && (88)\end{aligned}$$

where (86) follows from (71), and the $\sup$ in (84)-(86) is taken over all $p(x)$ and those $p(v|t)$ which satisfy $I(T;V) \le R_0$.

Hence, from (82) and (88), the capacity is given by

$$\mathcal{C} = h(\epsilon) + \epsilon - \inf_{p(v|t) \in \mathcal{L}(h(\alpha) - R_0)} H(U|V) \tag{89}$$

We will now explicitly evaluate the capacity expression obtained in (89) and compare it with the cut-set bound. For this purpose, we need a result on the conditional entropy of dependent random variables [9]. Let T, U be a pair of dependent random variables with a joint distribution $p(t,u)$. For $0 \le \gamma \le H(T)$, define the function $G(\gamma)$ as the infimum of $H(U|V)$, with respect to all discrete random variables V such that $H(T|V) = \gamma$ and the random variables V and U are conditionally independent given T. For the case when T is binary and $p(u|t)$, expressed as a stochastic matrix W, takes the form in (78), we have from [9],

$$\begin{aligned} G(\gamma) &= \inf_{p(v|t) \in \mathcal{L}(\gamma)} H(U|V) && (90)\\ &= h(\epsilon) + \epsilon\gamma && (91)\end{aligned}$$

We will use this result from [9] in explicitly evaluating the capacity in (89). First note that, if $R_0 \ge h(\alpha)$, then

$$G(h(\alpha) - R_0) = G(0) = h(\epsilon) \tag{92}$$

whereas, if $R_0 < h(\alpha)$, then

$$G(h(\alpha) - R_0) = h(\epsilon) + \epsilon(h(\alpha) - R_0) \tag{93}$$

Using (92) and (93), the capacity expression in (89) evaluates to,

$$\mathcal{C}(R_0) = \begin{cases} \epsilon, & R_0 \geq h(\alpha) \\ \epsilon(1-h(\alpha)) + \epsilon R_0, & R_0 < h(\alpha) \end{cases} \tag{94}$$

which can be written in a compact form as,

$$\mathcal{C}(R_0) = \min(\epsilon(1-h(\alpha)) + \epsilon R_0, \epsilon) \tag{95}$$

The cut-set bound is obtained by evaluating (37) for the channel in consideration. Evaluation of the cut-set bound is straightforward by noting that $I(X;Y)$ and $I(X;Y|T)$ are both maximized by a uniform $p(x)$. For a uniform distribution on X, we have $I(X;Y) = \epsilon(1-h(\alpha))$ and $I(X;Y|T) = \epsilon$. Hence, the cut-set bound is given as,

$$\mathcal{CS}(R_0) = \min(\epsilon(1-h(\alpha)) + R_0, \epsilon) \tag{96}$$

The difference between the capacity and the cut-set bound is evident from the first term in the min operation, i.e., the capacity expression in (95) has an ϵR_0 appearing in the minimum, as opposed to R_0 appearing in the cut-set bound at the corresponding place in (96). The cut-set bound and the capacity are shown in Figure 3 as functions of R_0 for $\alpha = 0.3$ and $\epsilon = 0.4$.

In conclusion, for this channel which does not fall into the classes of channels studied in [1] and [2], our upper bound equals the CAF achievable rate, thus yielding the capacity, which is strictly less than the cut-set bound for $R_0 < h(\alpha)$.

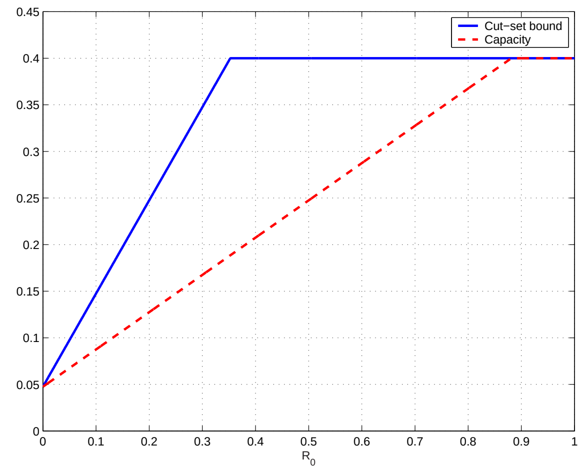

Figure 3: Capacity of the binary symmetric erasure channel for $\alpha = 0.3$ and $\epsilon = 0.4$.

VII. A Channel with Binary Multiplicative State and Binary Additive Noise

We will evaluate our upper bound and compare it with the cut-set bound for the case when X, T and N are binary and the channel is given as,

$$Y = TX + N \tag{97}$$

The channel output Y takes values in the set $\{0, 1, 2\}$. The random variables T and N are distributed as $T \sim \text{Ber}(\alpha)$ and $N \sim \text{Ber}(\delta)$. This relay channel does not fall into the subclass of channels considered in [1]. Moreover, the converse obtained in [2] does not apply to this channel since the output cannot be written as a modulo-sum.

To evaluate our upper bound, let us define

$$\Pr(X = 1) = p \tag{98}$$
$$\Pr(T = 1) = \alpha \tag{99}$$
$$\Pr(N = 1) = \delta \tag{100}$$

We then obtain $H(Y)$ as follows

$$H(Y) = h^{(3)}(P_Y(0), P_Y(1), P_Y(2)) \tag{101}$$

where

$$P_Y(0) = p(1-\alpha)(1-\delta) + (1-p)(1-\delta) \tag{102}$$
$$P_Y(1) = (1-p)\delta + p[(1-\alpha)\delta + \alpha(1-\delta)] \tag{103}$$
$$P_Y(2) = p\alpha\delta \tag{104}$$

and $H(Y|X)$ is obtained as,

$$H(Y|X) = (1-p)H(N) + pH(T+N) \tag{105}$$
$$= (1-p)h(\delta) + ph^{(3)}((1-\alpha)(1-\delta), \alpha * \delta, \alpha\delta) \tag{106}$$

The broadcast cut is obtained as,

$$I(X;Y|T) = H(Y|T) - H(Y|X,T) \tag{107}$$
$$= (1-\alpha)h(\delta) + \alpha h^{(3)}((1-p)(1-\delta), p * \delta, p\delta) - h(\delta) \tag{108}$$

The cut-set bound is given by,

$$\mathcal{CS} = \max_p \min\{I(X;Y) + R_0, I(X;Y|T)\} \tag{109}$$

We now evaluate our bound by first considering,

$$I(X,V;Y) = H(Y) - H(Y|X,V) \tag{110}$$

We have already evaluated $H(Y)$ in (101). Consider $H(Y|X,V)$:

$$H(Y|X,V) = \sum_{(x,v)} P_X(x)P_V(v)H(Y|X=x, V=v) \tag{111}$$
$$= \sum_v P_V(v)\big[(1-p)H(Y|X=0, V=v) + pH(Y|X=1, V=v)\big] \tag{112}$$
$$= \sum_v P_V(v)\big[(1-p)H(N) + pH(T+N|V=v)\big] \tag{113}$$
$$= \sum_v P_V(v)\big[(1-p)h(\delta) + pH(W|V=v)\big] \tag{114}$$
$$= (1-p)h(\delta) + pH(W|V) \tag{115}$$

where we have defined another random variable W as follows,

$$W = T + N \tag{116}$$

We are interested in lower bounding $H(W|V)$. We also know that any permissible conditional distribution $p(v|t)$ satisfies the constraint $I(T;V) \leq R_0$. Using this, we also have the following,

$$H(T|V) \geq h(\alpha) - R_0 \tag{117}$$

Let us also define,

$$P_{T|V}(T=1|V=v) = \eta_v,\ v \in 1,\ldots,|\mathcal{V}| \tag{118}$$

We now return to calculating $H(W|V)$

$$\begin{aligned}\Pr(W=w|V=v) &= \sum_t P_{T|V}(t|v)P_{W|T,V}(w|t,v) && (119)\\ &= (1-\eta_v)P(w|T=0,V=v) \\ &\quad + \eta_v P(w|T=1,V=v) && (120)\end{aligned}$$

Since the random variable W takes values in the set $\{0,1,2\}$, we obtain,

$$\begin{aligned}\Pr(W=0|V=v) &= (1-\eta_v)(1-\delta) && (121)\\ \Pr(W=1|V=v) &= \eta_v * \delta && (122)\\ \Pr(W=2|V=v) &= \eta_v\delta && (123)\end{aligned}$$

We finally obtain,

$$H(W|V) = \sum_v P_V(v)h^{(3)}((1-\eta_v)(1-\delta), \eta_v * \delta, \eta_v\delta) \tag{124}$$

For the special case when the additive noise is $N \sim \text{Ber}(1/2)$, the above expression simplifies to

$$\begin{aligned}H(W|V) &= \sum_v P_V(v)h^{(3)}\left(\frac{(1-\eta_v)}{2}, \frac{1}{2}, \frac{\eta_v}{2}\right) && (125)\\ &= \sum_v P_V(v)\left(\frac{1}{2}h(\eta_v) + 1\right) && (126)\\ &= \frac{1}{2}H(T|V) + 1 && (127)\\ &\geq \frac{1}{2}\left(h(\alpha) - R_0\right) + 1 && (128)\end{aligned}$$

where (128) follows from (117). Substituting (128) in (115) we obtain

$$\begin{aligned}H(Y|X,V) &= (1-p)h(\delta) + pH(W|V) && (129)\\ &\geq (1-p)h(\delta) + p\left(\frac{1}{2}\left(h(\alpha) - R_0\right) + 1\right) && (130)\end{aligned}$$

Continuing from (110), we obtain an upper bound on $I(X,V;Y)$ as follows,

$$\begin{aligned}I(X,V;Y) &= H(Y) - H(Y|X,V) && (131)\\ &\leq H(Y) - 1 - \frac{p}{2}(h(\alpha) - R_0) && (132)\end{aligned}$$

Moreover, the first term appearing in the cut-set bound simplifies to

$$\begin{aligned}I(X;Y) + R_0 &= H(Y) - H(Y|X) + R_0 && (133)\\ &= H(Y) - 1 - \frac{p}{2}h(\alpha) + R_0 && (134)\end{aligned}$$

We thus obtain our upper bound as,

$$\mathcal{UB} = \max_{p\in[0,1]} \min\left[H(Y) - 1 - \frac{p}{2}h(\alpha) + pR_0, I(X;Y|T)\right] \tag{135}$$

whereas the cut-set bound is,

$$\mathcal{CS} = \max_{p\in[0,1]} \min\left[H(Y) - 1 - \frac{p}{2}h(\alpha) + R_0, I(X;Y|T)\right] \tag{136}$$

The difference between the cut-set bound and our upper bound is evident from the first term in the $\min$ operation, i.e., our upper bound has a pR_0 term in (135), as opposed to R_0 at the corresponding place in (136).

Both these bounds along with the CAF rate are illustrated in Figure 4 as a function of R_0 for the case when $\alpha = 1/2$ and $\delta = 1/2$. We should remark here that although our bound is strictly smaller than the cut-set bound for certain values of R_0, it is strictly larger than the rates given by the CAF scheme. Here, the CAF rates are evaluated by restricting V to be binary, i.e., by considering all conditional distributions $p(v|t)$, such that, $|\mathcal{V}| = 2$. Therefore, the CAF rates plotted in Figure 4 are potentially suboptimal and can be potentially improved upon by increasing the cardinality of V.

VIII. Discussion

Let us recall our upper bound obtained in (36),

$$\begin{aligned}\mathcal{UB} = &\sup \min\{I(X,V;Y), I(X;Y|T)\}\\ &\text{s.t. } R_0 \geq I(T;V)\\ &\text{over } p(x)p(t)p(v|t) && (137)\end{aligned}$$

Using the fact that

$$\min(I(X,V;Y), I(X;Y|T)) \leq I(X,V;Y) \tag{138}$$

and observing that

$$I(X,V;Y) = I(V;Y) + I(X;Y|V) \tag{139}$$

it can be noted that our upper bound in (137) can be further upper bounded as

$$\begin{aligned}\mathcal{C} &\leq \sup I(V;Y) + I(X;Y|V) && (140)\\ &\text{s.t. } I(T;V) \leq R_0 && (141)\\ &\text{for some } p(x)p(v|t) && (142)\end{aligned}$$

On the other hand, the capacity is always lower bounded by the CAF rate,

$$\begin{aligned}\mathcal{C} &\geq \sup I(X;Y|V) && (143)\\ &\text{s.t. } I(T;V|Y) \leq R_0 && (144)\\ &\text{for some } p(x)p(v|t) && (145)\end{aligned}$$

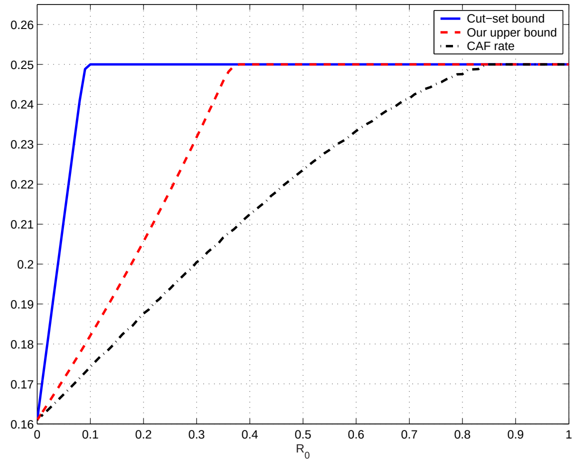

Figure 4: Comparison of our upper bound with the cut-set bound when $T \sim$ Ber$(1/2)$ and $N \sim$ Ber$(1/2)$.

Now using the following fact,

$$I(T;V|Y) = H(V|Y) - H(V|T) \tag{146}$$
$$= I(T;V) - I(V;Y) \tag{147}$$

we can rewrite the CAF lower bound on the capacity as

$$\mathcal{C} \geq \sup I(X;Y|V) \tag{148}$$
$$\text{s.t. } I(T;V) - I(V;Y) \leq R_0 \tag{149}$$
$$\text{for some } p(x)p(v|t) \tag{150}$$

We can see that the CAF lower bound on the capacity involves taking a supremum of $I(X;Y|V)$ subject to the constraint $I(T;V)-I(V;Y) \leq R_0$ whereas our upper bound involves taking a supremum of a larger quantity $I(V;Y) + I(X;Y|V)$ subject to a stricter constraint $I(T;V) \leq R_0$.

Although these two maximization problems are different, for the class of channels for which capacity was obtained, at the capacity achieving input distribution $p(x)$, we had $I(V;Y) = 0$. Moreover, the same input distribution $p(x)$ yielded the maximum for both maximization problems. Thus, for the class of channels considered in Section VI, these two maximization problems are equivalent. This observation yields a heuristic explanation as to why we were able to obtain the capacity results for these classes of channels.

IX. A New Lower Bound on Critical R_0

In [10], Cover posed a slightly different problem regarding the general primitive relay channel. Considering the capacity as a function of R_0, i.e., $\mathcal{C}(R_0)$, first observe the following facts,

$$\mathcal{C}(0) = \sup_{p(x)} I(X;Y) \tag{151}$$
$$\mathcal{C}(\infty) = \sup_{p(x)} I(X;Y|T) \tag{152}$$

Moreover, $\mathcal{C}(R_0)$ is a nondecreasing function of R_0. Cover posed the following question in [10]: what is the smallest value of R_0, say R_0^*, for which $\mathcal{C}(R_0^*) = \mathcal{C}(\infty)$? As an application of our upper bound, we implicitly provide a new lower bound on R_0^* for the class of primitive relay channels studied in this paper.

For the class of channels considered in Section VI, we obtained the capacity. As a consequence, we can explicitly characterize R_0^* for this class of channels as $h(\alpha)$. Furthermore, for the class of channels considered in Section VII, our upper bound on the capacity yields an improved lower bound on R_0^* than the one provided by the cut-set bound, which is clearly evident in Figure 4.

X. Conclusions

We obtained a new upper bound for a class of primitive relay channels. The primitive relay channel studied in this paper can also be considered as a state-dependent discrete memoryless channel, with rate-limited state information available at the receiver and no state information available at the transmitter.

Using our upper bound, we first recover all previously known capacity results for such channels. Furthermore, we explicitly characterize the capacity of a new subclass of these primitive relay channels which does not overlap with the classes previously studied in [1], [2]. In particular, for this class of channels, it is assumed that there are two channel states, and for each channel state, there is an erasure channel from X to Y. We show that the capacity for such channels is strictly smaller than the cut-set bound for certain values of R_0. This capacity result validates a conjecture due to Ahlswede and Han [8] for this class of channels.

Moreover, we also evaluated our upper bound for a case where $Y = TX + N$, where T, X and N are binary. This channel does not fall into any of the classes studied in [1], [2] and neither does it fall into the aforementioned class of channels. We show that our upper bound strictly improves upon the cut-set bound for certain values of R_0, although, our upper bound is strictly larger than the rates yielded by a potentially suboptimal evaluation of the CAF scheme.

References

[1] Y-H. Kim. Capacity of a class of deterministic relay channels. *IEEE Trans. on Information Theory*, 54(3):1328–1329, Mar. 2008.

[2] M. Aleksic, P. Razaghi, and W. Yu. Capacity of a class of modulo-sum relay channels. Submitted to *IEEE Trans. on Information Theory*, 2007.

[3] T. M. Cover and A. El Gamal. Capacity theorems for the relay channel. *IEEE Trans. on Information Theory*, 25(5):572–584, September 1979.

[4] Z. Zhang. Partial converse for the relay channel. *IEEE Trans. on Information Theory*, 34(5):1106–1110, September 1988.

[5] Y-H. Kim. Coding techniques for primitive relay channels. *Proc. Annual Allerton Conference on Communication, Control and Computing*, pages 129–135, 2007.

[6] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. New York:Wiley, 1991.

[7] C. Heegard and A. El Gamal. On the capacity of computer memory with defects. *IEEE Trans. on Information Theory*, 29(5):731–739, Sep. 1983.

[8] 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. on Information Theory*, 29(3):396–412, May 1983.

[9] H. S. Witsenhausen and A. D. Wyner. A conditional entropy bound for a pair of discrete random variables. *IEEE Trans. on Information Theory*, 21(5):493–501, September 1975.

[10] T. M. Cover. The capacity of the relay channel. In *Open Problems in Communication and Computation*, pages 72–73, 1987.