The Capacity of a Class of Mixture Semi-Deterministic Relay Channels

Hon-Fah Chong and Mehul Motani
Electrical and Computer Engineering Department,
National University of Singapore, Singapore 117576.
✉: {chong.hon.fah,motani}@nus.edu.sg

Abstract

The relay channel consists of a transmitter input x_1, a relay input x_2, a relay output y_2 and a receiver output y_3. In this report, we combine a degraded semi-deterministic relay channel with a semi-deterministic orthogonal relay channel to obtain a new mixture relay channel. For the new mixture relay channel, a combination of the compress-and-forward strategy and the partial decode-and-forward strategy achieves the capacity.

I. Introduction

The discrete-memoryless relay channel consists of four sets—$\mathcal{X}_1$, $\mathcal{X}_2$, $\mathcal{Y}_2$, $\mathcal{Y}_3$—and a collection of conditional probability mass functions $p(.,.|x_1,x_2)$ on $\mathcal{Y}_2 \times \mathcal{Y}_3$, one for each $(x_1,x_2) \in \mathcal{X}_1 \times \mathcal{X}_2$. The transmitter input is denoted by $x_1 \in \mathcal{X}_1$, the relay input by $x_2 \in \mathcal{X}_2$, the relay output by $y_2 \in \mathcal{Y}_2$ and the receiver output by $y_3 \in \mathcal{Y}_3$.

A $\left(2^{NR}, N\right)$ code for a relay channel without feedback consists of a set of integers $\mathcal{W} = \left\{1, 2, ..., \left\lfloor 2^{NR} \right\rfloor\right\}$, an encoding function

$$e : \left\{1, 2, ..., \left\lfloor 2^{NR} \right\rfloor\right\} \to \mathcal{X}_1^N$$

a set of relay functions $\{\Psi_n\}_{n=1}^{n=N}$ such that

$$\Psi_n : \mathcal{Y}_2^{n-1} \to \mathcal{X}_2, \ \ 1 \leq n \leq N$$

and a decoding function

$$d : \mathcal{Y}_3^N \to \left\{1, 2, ..., \left\lfloor 2^{NR} \right\rfloor\right\}.$$

The relay is causal in nature. Hence, the input of the relay x_{2n} is allowed to depend only on the past outputs of the relay $y_{21}, y_{22}, ..., y_{2n-1}$. If the message $w \in \mathcal{W}$ is sent, let

$$\lambda(w) = \Pr\left\{d\left(Y_3^N\right) \neq w | w \text{ sent}\right\}$$

denote the conditional probability of error. The average probability of error is defined by

$$P_e^{(N)} = \frac{1}{\left\lfloor 2^{NR} \right\rfloor} \sum_w \lambda(w).$$

The probability of error is calculated under the uniform distribution over the codewords $w \in \mathcal{W}$. The rate R is said to be achievable by the relay channel if there exists a sequence of $\left(2^{NR}, N\right)$ codes with $P_e^{(N)} \to 0$ as $N \to \infty$. The capacity C of a relay channel is the supremum of the set of achievable rates.

For a relay channel with causal and noiseless relay-transmitter feedback, the only difference is that the transmitter consists of a set of encoding functions $\{\Xi_n\}_{n=1}^{n=N}$ such that

$$\Xi_n : \mathcal{W} \times \mathcal{Y}_2^{n-1} \to \mathcal{X}_1, \quad 1 \leq n \leq N.$$

The relay channel was first introduced by van der Meulen in [1], [2]. Cover & El Gamal established two fundamental coding theorems for the relay channel in an important paper [3]. In addition, these two coding theorems were combined in the same paper to give the best lower bound for the capacity of a general relay channel [3, Theorem 7]. Recently, Chong et al. determined a potentially larger achievable rate in [4, Theorem 2]. In particular, they determined that the following rate is achievable for any relay channel:

$$R_{\text{CMG}} = \sup \min \left\{ \begin{array}{c} I\left(X_1; \hat{Y}_2 Y_3 | U X_2\right) + I\left(U; Y_2 | V X_2\right) \\ I\left(X_1 X_2; Y_3\right) - I\left(Y_2; \hat{Y}_2 | U X_1 X_2 Y_3\right) \end{array} \right\} \tag{1}$$

where the supremum is taken over all joint probability mass functions of the form

$$p\left(v, u, x_1, x_2, y_2, \hat{y}_2, y_3\right) = p\left(v\right) p\left(u|v\right) p\left(x_1|u\right) p\left(x_2|v\right) p\left(y_2, y_3|x_1, x_2\right) p\left(\hat{y}_2|x_2, y_2, u\right) \tag{2}$$

and subject to the constraint

$$I\left(X_2; Y_3|UV\right) \geq I\left(\hat{Y}_2; Y_2|U X_2 Y_3\right). \tag{3}$$

The capacity of the relay channel has been determined for the following special cases:

1) the degraded relay channel, the reversely degraded relay channel and the relay channel with causal noiseless feedback from the receiver to the relay [3];
2) the semideterministic relay channel [5];
3) a class of relay channels with orthogonal components [6];
4) a class of modulo-sum relay channels [7];
5) a class of deterministic relay channels [8].

However, the capacity of the general relay channel remains unknown. The achievability of the above classes of relay channels follows directly from appropriate substitutions for the auxiliary random variables in [3, Th. 7]. Moreover, except for the class of modulo-sum relay channels [7], the capacity of all the other classes of relay channels meet the cut-set upper bound. The question remains as to whether there exists other classes of relay channels where the lower bound given by [3, Th. 7] meets the cut-set upper bound.

In [9], we answered the question affirmatively by determining the capacity of three new classes of relay channels:

- The first class of relay channels corresponds closely to the degraded relay channels considered by Cover and El Gamal. However, for the first class of relay channels, the output of the relay y_2 depends probabilistically only on s, x_2 and y_3 where s is a deterministic function of x_1, i.e., $s = f_1\left(x_1\right)$. In addition, we require that $S \to \left(X_2, Y_2\right) \to Y_3$ forms a Markov chain for all input probability distributions $p\left(x_1, x_2\right)$. If the deterministic function f_1 is not a one-to-one mapping, the class of degraded semi-deterministic relay channels strictly contains the class of degraded relay channels. The capacity for the degraded semi-deterministic relay channel is achieved by the partial decode-and-forward strategy.

- The second class of relay channels that we consider is the class of semi-deterministic orthogonal relay channels. There are orthogonal channels from the transmitter/relay to the receiver and the output of the relay y_2 is a deterministic function of x_1, x_2 and y_3. Furthermore, the capacity cut-set upper bound for this class of relay channels is maximized by independent input probability distributions and can be achieved by the compress-and-forward strategy. The class of semi-deterministic orthogonal relay channels strictly contains the class of deterministic relay channels considered by Kim.
- The third class of relay channels also corresponds closely to the deterministic relay channel considered by Kim. However, instead of having a noiseless relay receiver link, we have a causal and noiseless feedback from the relay to the transmitter. Furthermore, instead of requiring the output of the relay y_2 to be a deterministic function of x_2 and y_3, we require the output of the relay y_2 to be a deterministic function of x_1, x_2 and y_3. The capacity for the class of semi-deterministic relay channels with causal and noiseless relay-transmitter feedback can be achieved by the hash-and-forward strategy.

In this report, we combine the first two classes of semi-deterministic relay channels together to obtain new channels whose capacity can also be determined.

II. Capacity region of a new class of mixture relay channels

Let us first describe the class of mixture relay channels below.

***Definition** 1:* The mixture relay channel has two relay outputs—y_{21} and y_{22} and also two receiver outputs—y_{31} and y_{32}. Furthermore, the class of mixture relay channels satisfies the following conditions:

- The conditional probability mass function describing the channel is given by

$$p\left(y_{21},y_{22},y_{31},y_{32}|x_1,x_2\right)=p\left(y_{31}|x_1\right)p\left(y_{21}|x_1,t,y_{31}\right)p\left(y_{22}|s,t,y_{31},y_{21}\right)p\left(y_{32}|x_2,t\right)p\left(t\right) \tag{4}$$

  where s is a deterministic function of x_1, i.e., $s=f_1\left(x_1\right)$ and T is independent of the inputs X_1 and X_2,
- s is a deterministic function x_2, y_{21} and y_{22}, i.e., $s=f_2\left(x_2,y_{21},y_{22}\right)$,
- t is a deterministic function of x_2 and y_{32}, i.e., $t=f_3\left(x_2,y_{32}\right)$,
- and that y_{21} is a deterministic function of x_1, x_2, y_{31} and y_{32}, i.e., $y_{21}=f_4\left(x_1,x_2,y_{31},y_{32}\right)$.

We note that this is strictly neither a degraded semi-deterministic relay channel ($X_1\rightarrow\left(S,X_2,Y_3\right)\rightarrow Y_{21}$ does not form a Markov chain) nor a semi-deterministic orthogonal relay channel ($y_{22}\neq f_4\left(x_1,x_2,y_{31},y_{32}\right)$). In fact, this is a mixture of the degraded semi-deterministic relay channel and the semi-deterministic orthogonal relay channel.

The capacity of the class of mixture relay channels described in Definition 1 is given by the following theorem:

***Theorem** 1:* The capacity of the mixture relay channel is given by

$$C_4=\sup_{P_{X_1X_2}}\left\{I\left(X_1;Y_{21}Y_{31}|ST\right)+H\left(S\right),I\left(X_1;Y_{31}\right)+I\left(X_2;Y_{32}\right)\right\} \tag{5}$$

where the supremum is taken over all independent input probability distributions.

We first reproduce [9, Property 1-Property 6] for the class of semi-deterministic orthogonal relay channels below for convenience.

***Property** 1:* For all input probability distributions $p(x_1, x_2)$, $(T, X_2, Y_{32}) \rightarrow X_1 \rightarrow Y_{31}$ forms a Markov chain.

***Property** 2:* For all input probability distributions $p(x_1, x_2)$, $(X_1, Y_{31}) \rightarrow X_2 \rightarrow Y_{32}$ forms a Markov chain.

***Property** 3:* For all input probability distributions $p(x_1, x_2)$, $(X_2, Y_{32}) \rightarrow (X_1, T) \rightarrow (Y_2, Y_{31})$ forms a Markov chain.

***Property** 4:* For independent input probability distributions, i.e., $p(x_1, x_2) = p(x_1)\, p(x_2)$, $(X_2, Y_{32}) \rightarrow T \rightarrow (Y_2, Y_{31})$ forms a Markov chain.

***Property** 5:* For independent input probability distributions, i.e., $p(x_1, x_2) = p(x_1)\, p(x_2)$, Y_{31} and Y_{32} are independent.

***Property** 6:* The terms $I(X_1; Y_2 Y_{31}|T)$ and $I(X_1; Y_{31})$ for the semi-deterministic orthogonal relay channel is maximized by the marginal input probability distribution $p(x_1)$. The term $I(X_2; Y_{32})$ for the semi-deterministic orthogonal relay channel is maximized by the marginal input probability distribution $p(x_2)$.

The proofs are given in Appendix I.

Next, we note that that for the mixture relay channel, $p(y_{21}, y_{31}, y_{32}|x_1, x_2)$ is given by

$$\begin{aligned}\sum_{y_{22} \in \mathcal{Y}_{22}} p(y_{21}, y_{22}, y_{31}, y_{32}|x_1, x_2) &\overset{(a)}{=} \sum_{y_{22} \in \mathcal{Y}_{22}} p(y_{31}|x_1)\, p(y_{21}|x_1, t, y_{31})\, p(y_{22}|s, t, y_{31}, y_{21})\, p(y_{32}|x_2, t)\, p(t) \\ &= p(y_{31}|x_1)\, p(y_{21}|x_1, t, y_{31})\, p(y_{32}|x_2, t)\, p(t) \sum_{y_{22} \in \mathcal{Y}_{22}} p(y_{22}|s, t, y_{31}, y_{21}) \\ &= p(y_{31}|x_1)\, p(y_{21}|x_1, t, y_{31})\, p(y_{32}|x_2, t)\, p(t) \qquad (6)\end{aligned}$$

where

(a) follows from the conditional probability mass function describing the mixture channel.

Since (6) is of the same form as the conditional probability mass function describing the semi-deterministic orthogonal relay channel, Property 1-Property 6 applies to the mixture relay channel as well, with Y_2 replaced by Y_{21}.

In addition, we note the following three additional properties that will be useful to our proof later on.

***Property** 7:* For all input probability distributions $p(x_1, x_2)$, $(X_1, X_2, Y_{32}) \rightarrow (S, T, Y_{31}, Y_{21}) \rightarrow Y_{22}$ forms a Markov chain.

***Property** 8:* For independent input probability distributions, i.e., $p(x_1, x_2) = p(x_1)\, p(x_2)$, $(X_2, Y_{32}) \rightarrow (S, T) \rightarrow (Y_{21}, Y_{31})$ forms a Markov chain.

***Property** 9:* The term $I(X_1; Y_{21} Y_{31}|ST)$ for the mixture relay channel is maximized by the marginal input probability distribution $p(x_1)$.

The proofs are given in Appendix II.

We now consider the proof of Theorem 1 below:

Proof: 1) Achievability: This follows from a mixture of the compress-and-forward strategy and the partial decode-and-forward strategy. Substituting $U \triangleq S$, $V \triangleq \emptyset$ and $\hat{Y}_2 \triangleq \hat{Y}_{21}$ in [4, Theorem 2] gives us the achievable rate

$$R = \sup_{P_{X_1 X_2 Y_{21} Y_{22} \hat{Y}_{21} Y_{31} Y_{32}}} \min \left\{ I\left(X_1; \hat{Y}_{21} Y_{31} Y_{32} | S X_2\right) + I\left(S; Y_{21} Y_{22} | X_2\right), I\left(X_1 X_2; Y_{31} Y_{32}\right) \right\} \quad (7)$$

subject to the constraint

$$I\left(X_2; Y_{31} Y_{32} | S\right) \geq I\left(Y_{21} Y_{22}; \hat{Y}_{21} | S X_2 Y_{31} Y_{32}\right) \quad (8)$$

where the supremum is taken over all joint probability mass functions of the form

$$p\left(x_1, x_2, y_{21}, y_{22}, \hat{y}_{21}, y_{31}, y_{32}\right) = p\left(x_1\right) p\left(x_2\right) p\left(y_{21}, y_{22}, y_{31}, y_{32} | x_1, x_2\right) p\left(\hat{y}_{21} | x_2, y_{21}, y_{22}, s\right). \quad (9)$$

We set $p\left(\hat{y}_{21} | x_2, y_{21}, y_{22}, s\right)$ to be $p\left(\hat{y}_{21} | x_2, y_{21}, s\right)$ in (9). Hence, the supremum is taken over all joint probability mass functions of the form

$$p\left(x_1, x_2, y_{21}, y_{22}, \hat{y}_{21}, y_{31}, y_{32}\right) = p\left(x_1\right) p\left(x_2\right) p\left(y_{21}, y_{22}, y_{31}, y_{32} | x_1, x_2\right) p\left(\hat{y}_{21} | x_2, y_{21}, s\right). \quad (10)$$

Since $(Y_{22}, Y_{31}, Y_{32}) \rightarrow (X_2, Y_{21}, S) \rightarrow \hat{Y}_{21}$ forms a Markov chain (this can be seen by inspection from (10)), the constraint (8) is now given by

$$\begin{aligned} I\left(X_2; Y_{31} Y_{32} | S\right) &\geq I\left(Y_{21} Y_{22}; \hat{Y}_{21} | S X_2 Y_{31} Y_{32}\right) \\ &= I\left(Y_{21}; \hat{Y}_{21} | S X_2 Y_{31} Y_{32}\right) + I\left(Y_{22}; \hat{Y}_{21} | S X_2 Y_{21} Y_{31} Y_{32}\right) \\ &= I\left(Y_{21}; \hat{Y}_{21} | S X_2 Y_{31} Y_{32}\right). \end{aligned} \quad (11)$$

If $H\left(Y_{21} | S X_2 Y_{31} Y_{32}\right) \leq I\left(X_2; Y_{31} Y_{32} | S\right)$, we obtain

$$R = \min\left\{ I\left(X_1; Y_{21} Y_{31} Y_{32} | S X_2\right) + I\left(S; Y_{21} Y_{22} | X_2\right), I\left(X_1 X_2; Y_{31} Y_{32}\right) \right\}. \quad (12)$$

If $H\left(Y_{21} | S X_2 Y_{31} Y_{32}\right) > I\left(X_2; Y_{31} Y_{32} | S\right)$, we have for the first term of (7)

$$\begin{aligned} & I\left(X_1; \hat{Y}_{21} Y_{31} Y_{32} | S X_2\right) + I\left(S; Y_{21} Y_{22} | X_2\right) \\ &= I\left(X_1; Y_{31} Y_{32} | S X_2\right) + I\left(X_1; \hat{Y}_{21} | S X_2 Y_{31} Y_{32}\right) + I\left(S; Y_{21} Y_{22} | X_2\right) \\ &= I\left(X_1; Y_{31} Y_{32} | S X_2\right) + H\left(\hat{Y}_{21} | S X_2 Y_{31} Y_{32}\right) - H\left(\hat{Y}_{21} | S X_1 X_2 Y_{31} Y_{32}\right) + I\left(S; Y_{21} Y_{22} | X_2\right) \\ &\overset{(a)}{=} I\left(X_1; Y_{31} Y_{32} | S X_2\right) + H\left(\hat{Y}_{21} | S X_2 Y_{31} Y_{32}\right) - H\left(\hat{Y}_{21} | S X_1 X_2 Y_{21} Y_{31} Y_{32}\right) + I\left(S; Y_{21} Y_{22} | X_2\right) \\ &= I\left(X_1; Y_{31} Y_{32} | S X_2\right) + H\left(\hat{Y}_{21} | S X_2 Y_{31} Y_{32}\right) - H\left(\hat{Y}_{21} | S X_2 Y_{21} Y_{31} Y_{32}\right) + I\left(S; Y_{21} Y_{22} | X_2\right) \\ &= I\left(X_1; Y_{31} Y_{32} | S X_2\right) + I\left(Y_{21}; \hat{Y}_{21} | S X_2 Y_{31} Y_{32}\right) + I\left(S; Y_{21} Y_{22} | X_2\right) \\ &\overset{(b)}{=} I\left(X_1; Y_{31} Y_{32} | S X_2\right) + I\left(X_2; Y_{31} Y_{32} | S\right) + I\left(S; Y_{21} Y_{22} | X_2\right) \\ &= I\left(X_1 X_2; Y_{31} Y_{32} | S\right) + I\left(S; Y_{21} Y_{22} | X_2\right) \\ &\overset{(c)}{=} I\left(X_1 X_2; Y_{31} Y_{32} | S\right) + I\left(S; X_2 Y_{21} Y_{22}\right) \\ &\overset{(d)}{\geq} I\left(X_1 X_2; Y_{31} Y_{32} | S\right) + I\left(S; Y_{31} Y_{32}\right) \\ &= I\left(X_1 X_2; Y_{31} Y_{32}\right) \end{aligned} \quad (13)$$

where

(a) follows from the fact that y_{21} is a deterministic function of x_1, x_2, y_{31} and y_{32},

(b) follows from the fact that if $H(Y_{21}|SX_2Y_{31}Y_{32}) > I(X_2;Y_{31}Y_{32}|S)$, there exists probability mass functions of the form (10) such that $I(X_2;Y_{31}Y_{32}|S) = I\left(Y_{21};\hat{Y}_{21}|SX_2Y_{31}Y_{32}\right)$ and

(c) follows from the fact that the input probability distributions are independent and

(d) follows from the fact that $S \to (X_2, Y_{21}, Y_{22}) \to (Y_{31}, Y_{32})$ forms a Markov chain since s is a deterministic function of x_2, y_{21} and y_{22}.

We note that $I(X_1X_2;Y_{31}Y_{32}) = I(X_1;Y_{31}) + I(X_2;Y_{32})$. This follows from the proof of [9, Thm. 3] (see [9, (37)]). We also note that

$$
\begin{aligned}
&I(X_1;Y_{21}Y_{31}Y_{32}|SX_2) + I(S;Y_{21}Y_{22}|X_2) \\
&\quad = I(X_1;Y_{32}|SX_2) + I(X_1;Y_{21}Y_{31}|SX_2Y_{32}) + I(S;Y_{21}Y_{22}|X_2) \\
&\quad \overset{(a)}{=} I(X_1;Y_{21}Y_{31}|SX_2Y_{32}) + I(S;Y_{21}Y_{22}|X_2) \\
&\quad \overset{(b)}{=} I(X_1;Y_{21}Y_{31}|STX_2Y_{32}) + I(S;Y_{21}Y_{22}|X_2) \\
&\quad \overset{(c)}{=} I(X_1;Y_{21}Y_{31}|ST) + I(S;Y_{21}Y_{22}|X_2) \\
&\quad = I(X_1;Y_{21}Y_{31}|ST) + H(S|X_2) - H(S|X_2Y_{21}Y_{22}) \\
&\quad \overset{(d)}{=} I(X_1;Y_{21}Y_{31}|ST) + H(S)
\end{aligned} \tag{14}
$$

where

(a) follows from the fact that $(S, X_1) \to X_2 \to Y_{32}$ forms a Markov chain (see Property 2) and hence, $X_1 \to (S, X_2) \to Y_{32}$ forms a Markov chain (follows from the weak union property of Markov chains, see [10, Sec. 1.1.5]),

(b) follows from the fact that t is a deterministic function of x_2 and y_{32},

(c) follows from the fact that for independent input probability distributions $(X_2, Y_{32}) \to (S, T) \to (Y_{21}, Y_{31})$ forms a Markov chain (refer to Property 8) and the fact that $(X_2, Y_{32}) \to (X_1, T) \to (Y_{21}, Y_{31})$ forms a Markov chain for all input probability distributions (refer to Property 3) and

(d) follows from the fact that the input probability distributions are independent and that s is a deterministic function of x_2, y_{21} and y_{22}.

Hence, a combination of the partial decode-and forward strategy and the compress-and-forward strategy achieves the rate (5) for the mixture relay channel and where the supremum is taken over all independent input probability distributions.

2) Converse: The converse follows from the cut-set upper bound. For the first term in [9, (10)], we obtain

$$
\begin{aligned}
&I(X_1;Y_{21}Y_{22}Y_{31}Y_{32}|X_2) \\
&\quad \overset{(a)}{=} I(SX_1;Y_{21}Y_{22}Y_{31}Y_{32}|X_2) \\
&\quad = I(X_1;Y_{21}Y_{22}Y_{31}Y_{32}|SX_2) + I(S;Y_{21}Y_{22}Y_{31}Y_{32}|X_2) \\
&\quad = I(X_1;Y_{22}|Y_{21}Y_{31}Y_{32}SX_2) + I(X_1;Y_{21}Y_{31}Y_{32}|SX_2) + I(S;Y_{21}Y_{22}Y_{31}Y_{32}|X_2) \\
&\quad \overset{(b)}{=} I(X_1;Y_{22}|Y_{21}Y_{31}Y_{32}STX_2) + I(X_1;Y_{21}Y_{31}Y_{32}|SX_2) + I(S;Y_{21}Y_{22}Y_{31}Y_{32}|X_2)
\end{aligned}
$$

$$\stackrel{(c)}{=} I\left(X_1; Y_{21}Y_{31}Y_{32}|SX_2\right) + I\left(S; Y_{21}Y_{22}|X_2\right) + I\left(S; Y_{31}Y_{32}|X_2Y_{21}Y_{22}\right)$$

$$\stackrel{(d)}{=} I\left(X_1; Y_{21}Y_{31}Y_{32}|SX_2\right) + I\left(S; Y_{21}Y_{22}|X_2\right)$$

$$= I\left(X_1; Y_{32}|SX_2\right) + I\left(X_1; Y_{21}Y_{31}|SX_2Y_{32}\right) + I\left(S; Y_{21}Y_{22}|X_2\right)$$

$$\stackrel{(e)}{=} I\left(X_1; Y_{21}Y_{31}|SX_2Y_{32}\right) + I\left(S; Y_{21}Y_{22}|X_2\right)$$

$$\stackrel{(b)}{=} I\left(X_1; Y_{21}Y_{31}|STX_2Y_{32}\right) + I\left(S; Y_{21}Y_{22}|X_2\right)$$

$$= H\left(Y_{21}Y_{31}|STX_2Y_{32}\right) - H\left(Y_{21}Y_{31}|STX_1X_2Y_{32}\right) + I\left(S; Y_{21}Y_{22}|X_2\right)$$

$$\stackrel{(f)}{\leq} H\left(Y_{21}Y_{31}|ST\right) - H\left(Y_{21}Y_{31}|STX_1X_2Y_{32}\right) + I\left(S; Y_{21}Y_{22}|X_2\right)$$

$$\stackrel{(g)}{=} H\left(Y_{21}Y_{31}|ST\right) - H\left(Y_{21}Y_{31}|STX_1\right) + I\left(S; Y_{21}Y_{22}|X_2\right)$$

$$= I\left(X_1; Y_{21}Y_{31}|ST\right) + I\left(S; Y_{21}Y_{22}|X_2\right)$$

$$= I\left(X_1; Y_{21}Y_{31}|ST\right) + H\left(S|X_2\right) - H\left(S|X_2Y_{21}Y_{22}\right)$$

$$\stackrel{(d)}{=} I\left(X_1; Y_{21}Y_{31}|ST\right) + H\left(S|X_2\right)$$

$$\stackrel{(f)}{\leq} I\left(X_1; Y_{21}Y_{31}|ST\right) + H\left(S\right). \tag{15}$$

where

(a) follows from the fact that s is a deterministic function of x_1,

(b) follows from the fact that t is a deterministic function of x_2 and y_{32},

(c) follows from the fact that $(X_1, X_2, Y_{32}) \rightarrow (S, T, Y_{31}, Y_{21}) \rightarrow Y_{22}$ forms a Markov chain (see Property 7),

(d) follows from the fact that $S \rightarrow (X_2, Y_{21}, Y_{22}) \rightarrow (Y_{31}, Y_{32})$ forms a Markov chain since s is a deterministic function of x_2, y_{21} and y_{22},

(e) follows from the fact that $(S, X_1) \rightarrow X_2 \rightarrow Y_{32}$ forms a Markov chain (see Property 2) and hence, $X_1 \rightarrow (S, X_2) \rightarrow Y_{32}$ forms a Markov chain (follows from the weak union property of Markov chains, see [10, Sec. 1.1.5]),

(f) follows from the fact that conditioning reduces entropy and

(g) follows from the fact that $(X_2, Y_{32}) \rightarrow (X_1, T) \rightarrow (Y_{21}, Y_{31})$ forms a Markov chain (Property 3) and hence, $(X_2, Y_{32}) \rightarrow (S, X_1, T) \rightarrow (Y_{21}, Y_{31})$ forms a Markov chain since s is a deterministic function of x_1.

We note that $I\left(X_1; Y_{21}Y_{31}|ST\right) + H\left(S\right)$ is maximized by the marginal probability distribution $p\left(x_1\right)$ (Property 9 and the fact that s is a deterministic function of x_1). We also note that (f) may be replaced by an equality as long as the input probability distributions are independent (Property 8 and from the fact that s is a deterministic function of x_1). Hence, independent input probability distributions maximize the first term in ([9, (10)]). We may also show that independent input probability distributions maximizes the second term in ([9, (10)]) (refer to [9, (40)]) and is given by the second term in (5). ■

APPENDIX I
PROOF OF PROPERTIES 1-6

We first prove the following property that will be useful in the rest of the proofs.

***Property** 10:* For all input probability distributions $p(x_1, x_2)$, $X_1 \to X_2 \to Y_{32}$ forms a Markov chain.

Proof: This follows from the conditional probability mass function [9, (14)] describing the channel as seen from the following equalities:

$$\begin{aligned}
p(y_{32}|x_1,x_2) &= \sum_{y_2\in\mathcal{Y}_2}\sum_{y_{31}\in\mathcal{Y}_{31}} p(y_2,y_{31},y_{32}|x_1,x_2) \\
&= \sum_{y_2\in\mathcal{Y}_2}\sum_{y_{31}\in\mathcal{Y}_{31}} p(y_{31}|x_1)\,p(y_2|x_1,t,y_{31})\,p(y_{32}|x_2,t)\,p(t) \\
&= p(y_{32}|x_2,t)\,p(t)\sum_{y_{31}\in\mathcal{Y}_{31}} p(y_{31}|x_1)\sum_{y_2\in\mathcal{Y}_2} p(y_2|x_1,t,y_{31}) \\
&= p(y_{32}|x_2,t)\,p(t) \\
&\overset{(a)}{=} p(y_{32},t|x_2) \\
&\overset{(b)}{=} p(y_{32}|x_2)\,.
\end{aligned} \tag{16}$$

where

(a) follows from the fact that T is independent of X_2 and

(b) follows from the fact that t is a deterministic function of x_2 and y_{32} [9, (13)].

■

Property 1 can be shown from the following equalities:

$$\begin{aligned}
p(y_{31}|x_1,t,x_2,y_{32}) &\overset{(a)}{=} p(y_{31}|x_1,x_2,y_{32}) \\
&= \frac{p(y_{31}|x_1,x_2,y_{32})\,p(y_{32}|x_1,x_2)}{p(y_{32}|x_1,x_2)} \\
&= \frac{p(y_{31},y_{32}|x_1,x_2)}{p(y_{32}|x_1,x_2)} \\
&= \frac{\sum_{y_2\in\mathcal{Y}_2} p(y_2,y_{31},y_{32}|x_1,x_2)}{p(y_{32}|x_1,x_2)} \\
&= \frac{\sum_{y_2\in\mathcal{Y}_2} p(y_{31}|x_1)\,p(y_2|x_1,t,y_{31})\,p(y_{32}|x_2,t)\,p(t)}{p(y_{32}|x_1,x_2)} \\
&= \frac{p(y_{31}|x_1)\,p(y_{32}|x_2,t)\,p(t)\sum_{y_2\in\mathcal{Y}_2} p(y_2|x_1,t,y_{31})}{p(y_{32}|x_1,x_2)} \\
&= \frac{p(y_{31}|x_1)\,p(y_{32}|x_2,t)\,p(t)}{p(y_{32}|x_1,x_2)} \\
&\overset{(b)}{=} \frac{p(y_{31}|x_1)\,p(y_{32},t|x_2)}{p(y_{32}|x_1,x_2)} \\
&\overset{(a)(c)}{=} \frac{p(y_{31}|x_1)\,p(y_{32}|x_2)}{p(y_{32}|x_2)} \\
&= p(y_{31}|x_1)
\end{aligned} \tag{17}$$

where

(a) follows from the fact that t is a deterministic function of x_2 and y_{32} ([9, (13)]),

(b) follows from the fact that T is independent of X_2 and

(c) follows from the fact that $X_1 \to X_2 \to Y_{32}$ forms a Markov chain (Property 10).

Property 2 can be shown from the following equalities:

$$\begin{aligned}
p(y_{32}|x_1,x_2,y_{31}) &= \frac{p(y_{31},y_{32}|x_1,x_2)}{p(y_{31}|x_1,x_2)} \\
&= \frac{\sum_{y_2\in\mathcal{Y}_2} p(y_2,y_{31},y_{32}|x_1,x_2)}{p(y_{31}|x_1,x_2)} \\
&= \frac{\sum_{y_2\in\mathcal{Y}_2} p(y_{31}|x_1)\,p(y_2|x_1,t,y_{31})\,p(y_{32}|x_2,t)\,p(t)}{p(y_{31}|x_1,x_2)} \\
&= \frac{p(y_{32}|x_2,t)\,p(t)\,p(y_{31}|x_1)\sum_{y_2\in\mathcal{Y}_2} p(y_2|x_1,t,y_{31})}{p(y_{31}|x_1,x_2)} \\
&= \frac{p(y_{32}|x_2,t)\,p(t)\,p(y_{31}|x_1)}{p(y_{31}|x_1,x_2)} \\
&\overset{(a)}{=} \frac{p(y_{32},t|x_2)\,p(y_{31}|x_1)}{p(y_{31}|x_1,x_2)} \\
&\overset{(b)}{=} \frac{p(y_{32}|x_2)\,p(y_{31}|x_1)}{p(y_{31}|x_1,x_2)} \\
&\overset{(c)}{=} \frac{p(y_{32}|x_2)\,p(y_{31}|x_1)}{p(y_{31}|x_1)} \\
&= p(y_{32}|x_2)
\end{aligned} \tag{18}$$

where

(a) follows from the fact that T is independent of X_2,

(b) follows from the fact that t is a deterministic function of x_2 and y_{32} [9, (1)] and

(c) follows from the fact that $(T, X_2, Y_{32}) \to X_1 \to Y_{31}$ forms a Markov chain (Property 1).

Property 3 can be shown from the following equalities:

$$\begin{aligned}
p(y_2,y_{31}|x_1,t,x_2,y_{32}) &\overset{(a)}{=} p(y_2,y_{31}|x_1,x_2,y_{32}) \\
&= \frac{p(y_2,y_{31}|x_1,x_2,y_{32})\,p(y_{32}|x_1,x_2)}{p(y_{32}|x_1,x_2)} \\
&= \frac{p(y_2,y_{31},y_{32}|x_1,x_2)}{p(y_{32}|x_1,x_2)} \\
&= \frac{p(y_{31}|x_1)\,p(y_2|x_1,t,y_{31})\,p(y_{32}|x_2,t)\,p(t)}{p(y_{32}|x_1,x_2)} \\
&\overset{(b)}{=} \frac{p(y_{31}|x_1)\,p(y_2|x_1,t,y_{31})\,p(y_{32},t|x_2)}{p(y_{32}|x_1,x_2)} \\
&\overset{(a)}{=} \frac{p(y_{31}|x_1)\,p(y_2|x_1,t,y_{31})\,p(y_{32}|x_2)}{p(y_{32}|x_1,x_2)} \\
&\overset{(c)}{=} \frac{p(y_{31}|x_1)\,p(y_2|x_1,t,y_{31})\,p(y_{32}|x_2)}{p(y_{32}|x_2)} \\
&= p(y_{31}|x_1)\,p(y_2|x_1,t,y_{31})
\end{aligned} \tag{19}$$

where

(a) follows from the fact that t is a deterministic function of x_2 and y_{32} [9, (13)],

(b) follows from the fact that T is independent of X_2 and

(c) follows from the fact that $X_1 \to X_2 \to Y_{32}$ forms a Markov chain (Property 10).

Property 4 can be shown from the following equalities:

$$
\begin{aligned}
p(y_2, y_{31}|t, x_2, y_{32}) &\overset{(a)}{=} p(y_2, y_{31}|x_2, y_{32}) \\
&= \frac{p(y_2, y_{31}, x_2, y_{32})}{p(x_2, y_{32})} \\
&= \frac{\sum_{x_1 \in \mathcal{X}_1} p(x_1, x_2, y_2, y_{31}, y_{32})}{p(x_2, y_{32})} \\
&= \frac{\sum_{x_1 \in \mathcal{X}_1} p(y_2, y_{31}, y_{32}|x_1, x_2) p(x_1) p(x_2)}{p(x_2, y_{32})} \\
&= \frac{\sum_{x_1 \in \mathcal{X}_1} p(y_{31}|x_1) p(y_2|x_1, t, y_{31}) p(y_{32}|x_2, t) p(t) p(x_1) p(x_2)}{p(x_2, y_{32})} \\
&\overset{(b)}{=} \frac{\sum_{x_1 \in \mathcal{X}_1} p(y_{31}|x_1) p(y_2|x_1, t, y_{31}) p(y_{32}, t|x_2) p(x_1) p(x_2)}{p(x_2, y_{32})} \\
&\overset{(a)}{=} \frac{\sum_{x_1 \in \mathcal{X}_1} p(y_{31}|x_1) p(y_2|x_1, t, y_{31}) p(y_{32}|x_2) p(x_1) p(x_2)}{p(x_2, y_{32})} \\
&= \frac{\sum_{x_1 \in \mathcal{X}_1} p(y_{31}|x_1) p(y_2|x_1, t, y_{31}) p(x_2, y_{32}) p(x_1)}{p(x_2, y_{32})} \\
&= \sum_{x_1 \in \mathcal{X}_1} p(y_{31}|x_1) p(y_2|x_1, t, y_{31}) p(x_1) \\
&= \sum_{x_1 \in \mathcal{X}_1} p(x_1, y_{31}) p(y_2|x_1, t, y_{31}) \\
&\overset{(c)}{=} \sum_{x_1 \in \mathcal{X}_1} p(x_1, y_2, y_{31}|t) \\
&= p(y_2, y_{31}|t)
\end{aligned}
\tag{20}
$$

where

(a) follows from the fact that t is a deterministic function of x_2 and y_{32} [9, (13)],

(b) follows from the fact that T is independent of X_2 and

(c) follows from the fact that $T \to X_1 \to Y_{31}$ forms a Markov chain (Property 1) and the fact that T is independent of X_1.

Property 5 can be shown from the following equalities:

$$
\begin{aligned}
p(y_{31}, y_{32}) &= \sum_{x_1, x_2 \in \mathcal{X}_1 \times \mathcal{X}_2} p(y_{31}, y_{32}|x_1, x_2) p(x_1) p(x_2) \\
&= \sum_{x_1, x_2 \in \mathcal{X}_1 \times \mathcal{X}_2} \sum_{y_2 \times \mathcal{Y}_2} p(y_2, y_{31}, y_{32}|x_1, x_2) p(x_1) p(x_2) \\
&= \sum_{x_1, x_2 \in \mathcal{X}_1 \times \mathcal{X}_2} \sum_{y_2 \times \mathcal{Y}_2} p(y_{31}|x_1) p(y_2|x_1, t, y_{31}) p(y_{32}|x_2, t) p(t) p(x_1) p(x_2) \\
&= \sum_{x_1, x_2 \in \mathcal{X}_1 \times \mathcal{X}_2} p(y_{31}|x_1) p(y_{32}|x_2, t) p(t) p(x_1) p(x_2) \sum_{y_2 \times \mathcal{Y}_2} p(y_2|x_1, t, y_{31}) \\
&= \sum_{x_1, x_2 \in \mathcal{X}_1 \times \mathcal{X}_2} p(y_{31}|x_1) p(y_{32}|x_2, t) p(t) p(x_1) p(x_2) \\
&\overset{(a)}{=} \sum_{x_1, x_2 \in \mathcal{X}_1 \times \mathcal{X}_2} p(y_{31}|x_1) p(y_{32}, t|x_2) p(x_1) p(x_2) \\
&\overset{(b)}{=} \sum_{x_1, x_2 \in \mathcal{X}_1 \times \mathcal{X}_2} p(y_{31}|x_1) p(y_{32}|x_2) p(x_1) p(x_2)
\end{aligned}
$$

$$= p(y_{31})\, p(y_{32}) \tag{21}$$

where

(a) follows from the fact that T is independent of X_2 and

(b) follows from the fact that t is a deterministic function of x_2 and y_{32} [9, (13)].

To prove Property 6, we first consider the term $H(Y_2Y_{31}|T)$ which depends only on the probability distribution $p(t)$ (which is independent of X_1 and X_2) and the conditional probability distribution $p(y_2, y_{31}|t)$. We note that the conditional probability distribution $p(y_2, y_{31}|t)$ depends on the marginal probability distribution $p(x_1)$ and not on the joint probability distribution $p(x_1, x_2)$ from the following equalities:

$$\begin{aligned} p(y_2, y_{31}|t) &\overset{(a)}{=} \sum_{x_1\in\mathcal{X}_1}\sum_{x_2\in\mathcal{X}_2} p(y_2, y_{31}|t, x_1, x_2)\, p(x_1, x_2) \\ &\overset{(b)}{=} \sum_{x_1\in\mathcal{X}_1}\sum_{x_2\in\mathcal{X}_2} p(y_2, y_{31}|t, x_1)\, p(x_1, x_2) \\ &= \sum_{x_1\in\mathcal{X}_1} p(y_2, y_{31}|t, x_1)\, p(x_1) \sum_{x_2\in\mathcal{X}_2} p(x_2|x_1) \\ &= \sum_{x_1\in\mathcal{X}_1} p(y_2, y_{31}|t, x_1)\, p(x_1). \end{aligned} \tag{22}$$

where

(a) follows from the fact that T is independent of X_1 and X_2 and

(b) follows from the fact that $X_2 \to (X_1, T) \to (Y_2, Y_{31})$ forms a Markov chain (Property 3).

Hence, the term $H(Y_2Y_{31}|T)$ is maximized by the marginal input probability distribution $p(x_1)$. Similarly, the term $H(Y_2Y_{31}|TX_1)$ is maximized by the marginal input probability distribution $p(x_1)$. Therefore, $I(X_1; Y_2Y_{31}|T)$ is maximized by the marginal input probability distribution $p(x_1)$. We can likewise prove that $I(X_1; Y_{31})$ is maximized by the marginal input probability distribution $p(x_1)$ from the fact that $X_2 \to X_1 \to Y_{31}$ forms a Markov chain (Property 1). We can also prove that $I(X_2; Y_{32})$ is maximized by the marginal input probability distribution $p(x_2)$ from the fact that $X_1 \to X_2 \to Y_{32}$ forms a Markov chain (Property 2).

APPENDIX II
PROOF OF PROPERTIES 7-9

Property 7 can be shown from the following equalities:

$$\begin{aligned} p(y_{22}|s, t, x_1, x_2, y_{21}, y_{31}, y_{32}) &= \frac{p(s, t, x_1, x_2, y_{21}, y_{22}, y_{31}, y_{32})}{p(s, t, x_1, x_2, y_{21}, y_{31}, y_{32})} \\ &\overset{(a)}{=} \frac{p(x_1, x_2, y_{21}, y_{22}, y_{31}, y_{32})}{p(x_1, x_2, y_{21}, y_{31}, y_{32})} \\ &\overset{(b)}{=} \frac{p(y_{31}|x_1)\, p(y_{21}|x_1, t, y_{31})\, p(y_{22}|s, t, y_{31}, y_{21})\, p(y_{32}|x_2, t)\, p(t)\, p(x_1, x_2)}{p(y_{31}|x_1)\, p(y_{21}|x_1, t, y_{31})\, p(y_{32}|x_2, t)\, p(t)\, p(x_1, x_2)} \\ &= p(y_{22}|s, t, y_{31}, y_{21}) \end{aligned} \tag{23}$$

where

(a) follows from the fact that s is a deterministic function of x_1 and t is a deterministic function of x_2 and y_{32} and

(b) follows from the conditional probability mass function describing the mixture channel and from (6).

Property 8 can be shown from the following equalities:

$$
\begin{aligned}
p(y_2, y_{31}|s,t,x_2,y_{32}) &\overset{(a)}{=} p(y_2, y_{31}|s,x_2,y_{32}) \\
&= \frac{p(y_2,y_{31},s,x_2,y_{32})}{p(s,x_2,y_{32})} \\
&= \frac{\sum_{x_1\in f_1^{-1}(s)} p(s,x_1,x_2,y_2,y_{31},y_{32})}{p(s,x_2,y_{32})} \\
&\overset{(b)}{=} \frac{\sum_{x_1\in f_1^{-1}(s)} p(x_1,x_2,y_2,y_{31},y_{32})}{p(s,x_2,y_{32})} \\
&= \frac{\sum_{x_1\in f_1^{-1}(s)} p(y_2,y_{31},y_{32}|x_1,x_2)\, p(x_1)\, p(x_2)}{p(s,x_2,y_{32})} \\
&= \frac{\sum_{x_1\in f_1^{-1}(s)} p(y_{31}|x_1)\, p(y_2|x_1,t,y_{31})\, p(y_{32}|x_2,t)\, p(t)\, p(x_1)\, p(x_2)}{p(s,x_2,y_{32})} \\
&\overset{(c)}{=} \frac{\sum_{x_1\in f_1^{-1}(s)} p(y_{31}|x_1)\, p(y_2|x_1,t,y_{31})\, p(y_{32},t|x_2)\, p(x_1)\, p(x_2)}{p(s,x_2,y_{32})} \\
&\overset{(a)}{=} \frac{\sum_{x_1\in f_1^{-1}(s)} p(y_{31}|x_1)\, p(y_2|x_1,t,y_{31})\, p(y_{32}|x_2)\, p(x_1)\, p(x_2)}{p(s,x_2,y_{32})} \\
&= \frac{\sum_{x_1\in f_1^{-1}(s)} p(y_{31}|x_1)\, p(y_2|x_1,t,y_{31})\, p(x_2,y_{32})\, p(x_1)}{p(s,x_2,y_{32})} \\
&\overset{(d)}{=} \frac{\sum_{x_1\in f_1^{-1}(s)} p(y_{31}|x_1)\, p(y_2|x_1,t,y_{31})\, p(x_2,y_{32})\, p(x_1)}{p(s)\, p(x_2,y_{32})} \\
&= \frac{\sum_{x_1\in f_1^{-1}(s)} p(y_{31}|x_1)\, p(y_2|x_1,t,y_{31})\, p(x_1)}{p(s)} \\
&\overset{(b)}{=} \frac{\sum_{x_1\in f_1^{-1}(s)} p(y_{31}|x_1,s)\, p(y_2|x_1,s,t,y_{31})\, p(x_1,s)}{p(s)} \\
&\overset{(e)}{=} \frac{\sum_{x_1\in f_1^{-1}(s)} p(y_{31}|x_1,s,t)\, p(y_2|x_1,s,t,y_{31})\, p(x_1,s|t)}{p(s)} \\
&= \frac{\sum_{x_1\in f_1^{-1}(s)} p(s,x_1,y_2,y_{31}|t)}{p(s)} \\
&= \frac{p(s,y_2,y_{31}|t)}{p(s)} \\
&\overset{(e)}{=} p(y_2,y_{31}|s,t)
\end{aligned}
\tag{24}
$$

where

(a) follows from the fact that t is a deterministic function of x_2 and y_{32},

(b) follows from the fact that s is a deterministic function of x_1,

(c) follows from the fact that T is independent of X_2,

(d) follows from the fact that the input probability distributions are independent and from the fact that $X_1 \to X_2 \to Y_{32}$ forms a Markov chain (Property 2) and

(e) follows from the fact that $T \to X_1 \to Y_{31}$ forms a Markov chain (Property 1) and from the fact that T is independent of X_1 and hence is independent of S.

To prove Property 9, we first consider the term $H(Y_{21}Y_{31}|ST)$ which depends only on the probability distribution $p(t)\,p(s)$ (T is independent of X_1 and X_2) and the conditional probability distribution $p(y_{21}, y_{31}|s,t)$. We note that the conditional probability distribution $p\left(y_{21}, y_{31}|s, t\right)$ depends on the marginal probability distribution $p\left(x_1\right)$ and not on the joint probability distribution $p\left(x_1, x_2\right)$ from the following equalities:

$$\begin{aligned} p\left(y_{21}, y_{31}|s, t\right) &= \sum_{x_1 \in \mathcal{X}_1} \sum_{x_2 \in \mathcal{X}_2} p\left(y_{21}, y_{31}|s, t, x_1, x_2\right) p\left(x_1, x_2|s, t\right) \\ &\overset{(a)(b)}{=} \sum_{x_1 \in \mathcal{X}_1} \sum_{x_2 \in \mathcal{X}_2} p\left(y_{21}, y_{31}|s, t, x_1\right) p\left(x_1, x_2|s\right) \\ &= \sum_{x_1 \in \mathcal{X}_1} p\left(y_{21}, y_{31}|s, t, x_1\right) p\left(x_1|s\right) \sum_{x_2 \in \mathcal{X}_2} p\left(x_2|x_1, s\right) \\ &= \sum_{x_1 \in \mathcal{X}_1} p\left(y_2, y_{31}|s, t, x_1\right) p\left(x_1|s\right). \end{aligned} \tag{25}$$

where

(a) follows from the fact that T is independent of X_1 and X_2 and

(b) follows from the fact that $X_2 \to \left(X_1, T\right) \to \left(Y_{21}, Y_{31}\right)$ forms a Markov chain (Property 3) and the fact that s is a deterministic function of x_1.

Hence, the term $H\left(Y_{21}Y_{31}|ST\right)$ is maximized by the marginal input probability distribution $p\left(x_1\right)$. Similarly, the term $H\left(Y_{21}Y_{31}|STX_1\right)$ is maximized by the marginal input probability distribution $p\left(x_1\right)$.

REFERENCES

[1] E. C. van der Meulen, "Transmission of information in a T-terminal discrete memoryless channel," Ph.D. dissertation, Univ. of California, Berkeley, CA, Jun. 1968.

[2] ——, "Three-terminal communication channels," *Adv. Appl. Prob.*, vol. 3, pp. 120–154, 1971.

[3] T. Cover and A. E. Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 5, pp. 572–584, Sept. 1979.

[4] H. F. Chong, M. Motani, and H. K. Garg, "Generalized backward decoding strategies for the relay channel," *IEEE Trans. Inf. Theory*, vol. 53, no. 1, pp. 394–401, Jan 2007.

[5] A. E. Gamal and M. Aref, "The capacity of the semideterministic relay channel," *IEEE Trans. Inf. Theory*, vol. 28, no. 3, p. 536, May 1982.

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

[7] M. Aleksic, P. Razaghi, and W. Yu, "Capacity of a class of modulo-sum relay channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 3, pp. 921–930, Mar. 2009.

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

[9] H. F. Chong and M. Motani, "The capacity of several new classes of semi-deterministic relay channels," *IEEE Trans. Inf. Theory*, vol. submitted.

[10] J. Pearl, *Causality: Models, Reasoning, and Inference*, 1st ed. Cambridge University Press, 2000, ch. 1.