Capacity of a Class of Relay Channels With Orthogonal Components

Abbas El Gamal, *Fellow, IEEE,* and
Sina Zahedi, *Student Member, IEEE*

Abstract—The capacity of a class of discrete-memoryless relay channels with orthogonal channels from the sender to the relay receiver and from the sender and relay to the receiver is shown to be equal to the max-flow min-cut upper bound. The result is extended to additive white Gaussian noise (AWGN) relay channels where the channel from the sender to the relay uses a different frequency band from the channel from the sender and the relay to the receiver.

Index Terms—Additive white Gaussian noise (AWGN) relay channel, discrete memoryless relay channel.

I. INTRODUCTION

The discrete-memoryless relay channel denoted by $(\mathcal{X} \times \mathcal{X}_1, p(y, y_1|x, x_1), \mathcal{Y} \times \mathcal{Y}_1)$ consists of a sender $X \in \mathcal{X}$, a receiver $Y \in \mathcal{Y}$, a relay sender $X_1 \in \mathcal{X}_1$, a relay receiver $Y_1 \in \mathcal{Y}_1$, and a family of conditional probability mass functions $p(y, y_1|x, x_1)$ on $\mathcal{Y} \times \mathcal{Y}_1$, one for each $(x, x_1) \in \mathcal{X} \times \mathcal{X}_1$. A $(2^{nR}, n)$ code for the channel consists of: i) a set of messages $\{1, 2, \ldots, 2^{nR}\}$, ii) an encoding function that maps each message w into a codeword $x^n(w)$ of length n, iii) relay encoding functions $x_{1i} = f_i(y_{11}, y_{12}, \ldots, y_{1(i-1)})$, for $1 \leq i \leq n$, and iv) a decoding function that maps each received sequence y^n into an estimate $\hat{w}(y^n)$. A rate R is achievable if there exists a sequence of $(2^{nR}, n)$ codes with $P_e^{(n)} = \mathrm{P}\{\hat{W} \neq W\} \to 0$, as $n \to \infty$. Channel capacity C is defined as the supremum over the set of achievable rates.

The relay channel was introduced by van der Meulen in [1]. The capacity of degraded and reversely degraded relay channels as well as upper and lower bounds on the capacity of the general relay channel were established in [2]. The capacity of the general relay channel, however, is not known. Recent work on relay channel capacity has been motivated by wireless communication scenarios. Since in a practical wireless communication system, a node cannot transmit and receive at the same time or over the same frequency band, there has been much interest in channels with orthogonal components [4], [5]. In [4], a half-duplex additive white Gaussoan noise (AWGN) relay channel model is introduced where the relay transmits and receives at nonoverlapping time slots and the effect of multipath fading on the performance for several simple cooperative schemes is studied. In [5], a simple network of AWGN multiple-access channel with generalized feedback is considered where orthogonal channels carry partial feedback information and transmitters can partially cooperate through the feedback paths. Upper and lower bounds on the capacity region of this channel are presented and performance of several practical coding schemes are studied under fading conditions.

In this note, we consider the following class of discrete-memoryless relay channels with orthogonal channel components from the sender to the relay and from the sender and relay to the receiver.

Definition: A discrete-memoryless relay channel is said to have orthogonal components if the sender alphabet $\mathcal{X} = \mathcal{X}_D \times \mathcal{X}_R$ and the channel can be expressed as

$$p(y, y_1|x, x_1) = p(y|x_D, x_1)p(y_1|x_R, x_1)$$

for all

$$(x_D, x_R, x_1, y, y_1) \in \mathcal{X}_D \times \mathcal{X}_R \times \mathcal{X}_1 \times \mathcal{Y} \times \mathcal{Y}_1.$$

Our main result is to establish the capacity of this class of relay channels.

Theorem: The capacity of the relay channel with orthogonal components is given by

$$C = \max \min \{I(X_D, X_1; Y), I(X_R; Y_1|X_1) + I(X_D; Y|X_1)\}$$

where the maximum is taken over all joint probability mass functions of the form $p(x_D, x_R, x_1) = p(x_1)p(x_D|x_1)p(x_R|x_1)$.

We also extend this result to the class of AWGN relay channels where the channel from the sender to the relay uses a different frequency band from the channel from the sender and the relay to the receiver.

Note that if $p(y_1|x_R, x_1) = p(y_1|x_R)$ for all $(y_1, x_R, x_1) \in \mathcal{Y}_1 \times \mathcal{X}_R \times \mathcal{X}_1$, then our theorem reduces to a special case of the capacity region of the multiple-access channel with partially cooperating transmitters established in [6]. By setting the capacity of the link from one of the transmitters to the other to zero, the problem studied in [6] reduces to a special case of the relay channel with orthogonal components studied here. Therefore, the main contribution of our correspondence can be viewed as generalizing a special case of [6] and extending it to the AWGN case. It is also worth noting that using our AWGN extension, the capacity region for AWGN multiple-access channel with partially cooperating transmitters can be readily established.

II. PROOF OF THEOREM

To prove the theorem, we use two bounds on the capacity of the general discrete-memoryless relay channel from [2]. The first bound is the "max-flow min-cut" upper bound

$$C \leq \max_{p(x, x_1)} \min\{I(X, X_1; Y), I(X; Y, Y_1|X_1)\}. \qquad (1)$$

The second bound is a special case of Theorem 7 in [2], which yields the lower bound

$$C \geq \max_{p(u, x, x_1)} \min \{I(X, X_1; Y), I(U; Y_1|X_1) + I(X; Y|X_1, U)\}. \qquad (2)$$

This lower bound is achieved using a *generalized block-Markov* coding scheme, where in each block, the relay decodes part of the new message (represented by U) and cooperatively sends enough information with the sender (represented by X) to help the receiver decode the previous message (U then X). Note that in [7], this lower bound was shown to be optimal and equal to the max-flow min-cut upper bound for the class of semi-deterministic relay channels.

Achievability: We show that any $R < C$ is achievable using the generalized block-Markov encoding scheme. Substituting $X = (X_D, X_R)$ and $U = X_R$ in (2) and assuming joint probability mass function of the form $p(x_1)p(x_R|x_1)p(x_D|x_1)$, we obtain

$$\begin{aligned} I(X, X_1; Y) &= I(X_R, X_D, X_1; Y) \\ &= I(X_D, X_1; Y) + I(X_R; Y|X_D, X_1) \\ &= I(X_D, X_1; Y) \end{aligned}$$
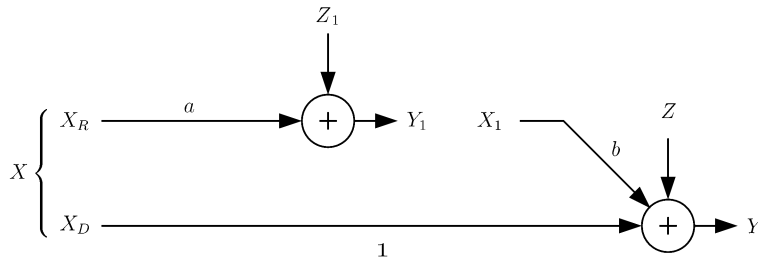
Fig. 1.   Frequency-division AWGN relay channel.

and

$$I(U; Y_1|X_1) + I(X; Y|X_1, U)$$
$$= I(X_R; Y_1|X_1) + I(X_D, X_R; Y|X_1, X_R)$$
$$\overset{(a)}{=} I(X_R; Y_1|X_1) + I(X_D; Y|X_1),$$

where (a) follows by the fact that $X_R \rightarrow X_1 \rightarrow Y$ form a Markov chain.

Converse: We show that C is equal to the max-flow min-cut upper bound. Clearly

$$I(X, X_1; Y) = I(X_D, X_1; Y).$$

Next, we consider the second term under the min in (1)

$$I(X; Y, Y_1|X_1)$$
$$= I(X_D, X_R; Y, Y_1|X_1)$$
$$= I(X_D, X_R; Y_1|X_1) + I(X_D, X_R; Y|X_1, Y_1)$$
$$= I(X_R; Y_1|X_1) + I(X_D; Y_1|X_1, X_R)$$
$$\quad + I(X_D, X_R; Y|X_1, Y_1)$$
$$\overset{(b)}{=} I(X_R; Y_1|X_1) + I(X_D, X_R; Y|X_1, Y_1)$$
$$= I(X_R; Y_1|X_1) + H(Y|X_1, Y_1)$$
$$\quad - H(Y|X_1, X_D, X_R, Y_1)$$
$$\overset{(c)}{=} I(X_R; Y_1|X_1) + H(Y|X_1, Y_1) - H(Y|X_1, X_D)$$
$$\leq I(X_R; Y_1|X_1) + H(Y|X_1) - H(Y|X_1, X_D)$$
$$= I(X_R; Y_1|X_1) + I(X_D; Y|X_1),$$

where (b) and (c) follow from the fact that $X_D \rightarrow (X_1, X_R) \rightarrow Y_1$ and $(X_R, Y_1) \rightarrow (X_1, X_D) \rightarrow Y$ each form a Markov chain, respectively. Thus, we have shown that

$$C \leq \max \min \{I(X_D, X_1; Y), I(X_R; Y_1|X_1) + I(X_D; Y|X_1)\}$$

where the maximization is over the choice of joint probability mass function $p(x_D, x_R, x_1)$. Without loss of generality, we can restrict the joint probability mass functions to be of the form

$$p(x_D, x_R, x_1) = p(x_1)p(x_D|x_1)p(x_R|x_1).$$

This completes the proof of the theorem.

III. EXTENSION TO AWGN RELAY CHANNEL

The result of the theorem can be extended to establish the capacity of the AWGN relay channel in Fig. 1, where the channel from the sender to the relay uses a different frequency band from the channel from the sender and relay sender to the receiver. The AWGN processes $\{Z_{1i}\}$

and $\{Z_i\}$ are independent each with power N, and the constants $a, b > 0$ represent the gain of the signal over the path from the sender to the relay and from the relay to the receiver, respectively, relative to the gain of the direct channel (which is assumed to be equal to one).

Assuming average power constraints P on $X = (X_D, X_R)$ and γP, on X_1, for $\gamma \geq 0$, the capacity of this channel is given by[1]

$$C(P, \gamma P) = \max_{0 \leq \alpha, \rho \leq 1} \min \left\{ \mathcal{C} \left(\frac{(\alpha + b^2\gamma + 2b\rho\sqrt{\alpha\gamma})P}{N} \right), \right.$$
$$\left. \mathcal{C} \left(\frac{a^2(1-\alpha)P}{N} \right) + \mathcal{C} \left(\frac{\alpha(1-\rho^2)P}{N} \right) \right\}$$

where $\mathcal{C}(x) = \frac{1}{2} \log_2(1 + x)$. Again achievability is established using the generalized block-Markov scheme, where we let $X_D \sim \mathcal{N}(0, \alpha P)$ and $X_R \sim \mathcal{N}(0, (1-\alpha)P)$ be independent, and $X_1 \sim \mathcal{N}(0, \gamma P)$ be independent of X_R but jointly Gaussian with X_D with $E(X_1 X_D) = \rho\sqrt{\alpha\gamma}P$. New information is sent to the relay through X_R and to the receiver through part of X_D (with power $\alpha(1 - \rho^2)P$), and the relay and the rest of X_D cooperatively send information to remove the uncertainty of the receiver about the previous message.

To prove the converse, first note that from the theorem, given any $(2^{nR}, n)$ sequence of codes with $P_e^{(n)} \rightarrow 0$

$$C \leq \min\{I(X_D, X_1; Y), I(X_R; Y_1|X_1) + I(X_D; Y|X_1)\} \quad (3)$$

for some joint probability distribution

$$p(x_D, x_R, x_1) = p(x_1)p(x_D|x_1)p(x_R|x_1).$$

Since the channel structure in this case has the special form

$$p(y, y_1|x, x_1) = p(y|x_D, x_1)p(y_1|x_R)$$

it follows that

$$I(X_R; Y_1|X_1) = H(Y_1|X_1) - H(Y_1|X_1, X_R)$$
$$= H(Y_1|X_1) - H(Y_1|X_R)$$
$$\leq H(Y_1) - H(Y_1|X_R) = I(X_R; Y_1).$$

Equation (3) then reduces to

$$C \leq \min\{I(X_D, X_1; Y), I(X_R; Y_1) + I(X_D; Y|X_1)\} \quad (4)$$

for some joint distribution $p(x_D, x_1)p(x_R)$. Note that this result can also be viewed as a special case of the capacity problem studied in [6].

Now the power cosntraints require that $E(X_D^2) + E(X_R^2) \leq P$ and $E(X_1^2) \leq \gamma P$. Thus, for some $0 \leq \alpha \leq 1$, $E(X_D^2) \leq \alpha P$ and

[1]This result was reported in [8].

$E(X_R^2) \leq (1-\alpha)P$. Define ρ to be the correlation coefficient between X_D and X_1. It is now straightforward to show that

$$I(X_D, X_1; Y) \leq \mathcal{C}\left(\frac{(\alpha + b^2\gamma + 2b\rho\sqrt{\alpha\gamma})P}{N}\right).$$

The second term under the minimum in (4) can be similarly upper-bounded to yield

$$I(X_R; Y_1) + I(X_D; Y|X_1) \leq \mathcal{C}\left(\frac{a^2(1-\alpha)P}{N}\right) + \mathcal{C}\left(\frac{\alpha(1-\rho^2)P}{N}\right).$$

This completes the proof of converse.

IV. CONCLUSION

This correspondence establishes the capacity of the class of relay channels with orthogonal channels from the sender to the relay and from the sender and relay to the receiver. As for all other classes of relay channels with known capacity, e.g., [2], [7], the capacity for this class is equal to the max-flow min-cut upper bound.

Our result points to the main difficulty in establishing the capacity of the general relay channel, which is finding the optimal broadcasting strategy from the sender to the relay Y_1 and the receiver Y. For the class of relay channels discussed here, the optimal strategy is to split the message into two parts, one is decoded by the relay and sent cooperatively with the sender to the receiver and the other is sent directly to the receiver. This strategy, however, is not optimal in general. In [9], it was shown that for AWGN relay channels, even when the channel from the sender to both the relay and receiver is orthogonal to the channel from the relay to the receiver, other strategies such as linear relaying and side information coding can achieve higher rates.

ACKNOWLEDGMENT

The authors wish to thank the reviewers for several helpful comments.

REFERENCES

[1] E. C. van der Meulen, "Three-terminal communication channels," *Adv. Appl. Probab.*, vol. 3, pp. 120–154, 1971.
[2] T. M. Cover and A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. IT-25, no. 5, pp. 572–584, Sep. 1979.
[3] G. Kramer, M. Gastpar, and P. Gupta, "Capacity theorems for wireless relay channels," in *Proc. 41st Annu. Allerton Conf. Communication, Control and Computation*, Allerton, IL, Oct. 2003, pp. 1074–1083.
[4] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.
[5] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity-part I: System description," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927–1938, Nov. 2003.
[6] F. M. J. Willems, "The discrete memoryless multiple access channel with partially cooperating encoders," *IEEE Trans. Inf. Theory*, vol. IT-29, no. 3, pp. 441–445, May 1983.
[7] A. El Gamal and M. Aref, "The capacity of semi-deterministic relay channel," *IEEE Trans. Inf. Theory*, vol. IT-28, no. 3, pp. 536–536, May 1982.
[8] A. El Gamal and S. Zahedi, "Minimum energy communication over a relay channel," in *Proc. IEEE Int. Symp. Information Theory*, Yokohama, Japan, Jun./Jul. 2003, p. 344.
[9] S. Zahedi, M. Mohseni, and A. El Gamal, "On the capacity of AWGN relay channels with linear relaying functions," in *Proc. IEEE Int. Sym. Information Theory*, Chicago, IL, Jun./Jul. 2004, p. 399.

On Channels With Partial Channel State Information at the Transmitter

Aviv Rosenzweig, Yossef Steinberg, *Member, IEEE*, and Shlomo Shamai (Shitz), *Fellow, IEEE*

Abstract—We consider an independent and identically distributed (i.i.d.) state-dependent channel with partial (rate-limited) channel state information (CSI) at the transmitter (CSIT) and full CSI at the receiver (CSIR). The CSIT comprises two parts, both subject to a rate constraint, and communicated over noiseless way-side channels. The first part is coded CSI, provided by a third party (a genie), and the second part is the output of a deterministic scalar quantizer of the state. A single-letter expression for the capacity of the channel is given. In case the CSIT comprises only the coded part, we show that the capacity previously suggested in the literature is too optimistic, and suggest a correction as part of our expression for the information capacity. For the general setting, an optimal coding scheme based upon multiplexing of several codebooks is presented. It is proved that the capacity of the channel is the same whether the quantizer's output is observed causally or noncausally by the encoder. In the rest of the correspondence, we focus on the special case where the system employs a stationary quantizer. Using a rate distortion approach we bound the alphabet's size of the auxiliary random variable (RV) of the information capacity. Next, we turn to the additive white Gaussian noise (AWGN) channel with fading, and show that the determination of the capacity region reduces to finding the optimal genie strategies and the optimal power allocation distribution along the product alphabet of the auxiliary RV and the quantizer's output alphabets. The suggested model can be applied, for example, to an orthogonal frequency-division multiplexing (OFDM) communication system. Here the fading across frequencies comprises the channel state sequence. Coded fading information is provided to the channel encoder via a way-side, rate-limited channel. In addition, since coding operation is expensive, a simpler scheme provides the channel encoder with quantized fading information, e.g., whether each coefficient is above/below a threshold.

Index Terms—Channel with state information, fading channels, power allocation, rate distortion with side information, rate-limited channel state information, scalar quantization.

I. INTRODUCTION

The rapid development of wireless communications systems over the last decade creates an increasing demand for a definition of corresponding realistic analytic models and for the exploration of the fundamental limits on reliable information transmission over these systems. A common model in the literature assumes a memoryless channel, whose conditional probability distribution is controlled by a time-varying state. Many studies have been devoted over the years to various scenarios related to a state-dependent channel model, where full or partial channel state information (CSI) is available at the transmitter (CSIT) and/or at the receiver (CSIR). In this correspondence, we suggest a model of a channel controlled by an independent and identically distributed (i.i.d.) state process, where the CSIT is subject to a rate constraint. Our channel model provides a unified framework for the models considered in [13, Theorem 2d], [4, Proposition 1], and in [15].