Capacity of a Class of Relay Channel with Orthogonal Components and Non-causal Channel State

Zhixiang Deng*, Fei Lang*, Bao-Yun Wang† and Sheng-mei Zhao*
*College of Telecommunications & Information Engineering
Nanjing University of Posts and Telecommunications, Nanjing, China, 210003
†College of Automation, Nanjing University of Posts and Telecommunications, Nanjing, China, 210003
Email: dengzhixiang@gmail.com, {langf, bywang, zhaosm}@njupt.edu.cn

***Abstract*—In this paper, a class of state-dependent relay channel with orthogonal channels from the source to the relay and from the source and the relay to the destination is studied. The two orthogonal channels are corrupted by common channel state which is known to both the source and the relay non-causally. The lower bound on the capacity for the channel is derived firstly. Then, we show that if the receiver output Y is a deterministic function of the relay input X_r, the channel state S and one of the source inputs X_D, i.e. $Y = f(X_r, X_D, S)$, the explicit capacity can be characterized.**

I. INTRODUCTION

State-dependent channels have brought wide attention in recent years [1]. Shannon first considered the single-user channel with channel state known causally at the transmitter [2], in which the capacity of this channel was characterized. When the channel state is known non-causally at the transmitter, Gel'fand and Pinsker [3] derived the capacity for the channel by a later-called Gelfand-Pinsker (GP) coding scheme. In [4], Costa studied the Gaussian channel with additive white Gaussian noise (AWGN) and additive Gaussian state known non-causally at the transmitter, and demonstrated that dirty paper coding (DPC) achieved the same capacity as the standard AWGN channel without channel state.

Extensions to the multiple user channels were performed by Gelfand-Pinsker in [5], where it was shown that interference cancellation was possible in Gaussian broadcast channel (BC) and Gaussian multiple-access channel (MAC). In [6], Y. Steinberg derived the inner and outer bounds for the degraded BC with non-causal side-information and characterized the capacity region when the side-information was provided to the transmitter in a causal manner.

The state-dependent two-user MAC with state non-causally known to the encoders was considered in [7]-[12]. Inner and outer bounds on the capacity of the state-dependent MAC were derived in [7] and [8] when the channel state was known to the transmitters fully or partially. For the finite-state-dependent MAC, inner and outer bounds on the capacity were also derived in [9]. Capacity regions are only characterized in some special cases, e.g. Gaussian MAC with additive state known at both encoders [10]. In some cases where cooperation between transmitters is allowed, capacity regions are also characterized, e.g. in [11] capacity region was characterized for the MAC where the informed encoder knew both its private message and the message from the uninformed encoder, and in [12] capacity region was derived for the two-user dirty paper Gaussian MAC with conferencing encoders.

The relay channels capture the characteristics of both MAC and BC. The state-dependent relay channels were studied in [13]-[18]. Lower and upper bounds were derived when the channel state was known only to the relay non-causally [13] or known only to the source non-causally [14]-[16]. The authors [16] obtained two corresponding lower bounds for the state-dependent relay channel with orthogonal components and with channel states known non-causally at the source. A similar orthogonal relay channel that was corrupted by an interference which was known non-causally at the source was considered in [17], in which several transmission strategies were proposed assuming the interference had structure.

However, capacity for the general relay channels hasn't been characterized yet even if the relay channel is state-independent. Capacity for the state-independent relay channels is characterized only in some special channels, e.g. physically degraded/reversely degraded relay channel [19], a class of deterministic relay channels [20] and a class of relay channels with orthogonal components [21]. To our knowledge, explicit capacity results for the state-dependent relay channel with state known to the transmitters were derived mainly for two cases: i) the physically degraded relay channel with state known *causally* to both the source and the relay; ii) the *Gaussian physically degraded* relay channel with state known non-causally to the source and the relay.

In this paper, we investigate the state-dependent relay channel with orthogonal components, in which the source communicates with the relay through the channel (say channel 1) orthogonal to the channel (say channel 2) by which the source and relay communicate with the destination. We assume that both channel 1 and channel 2 are corrupted by common channel state which is known non-causally to both the source and the relay. In this setup, we establish the lower bound on the capacity for the channel firstly. Further, we characterize the capacity exactly when the receiver output Y is a deterministic function of the relay input X_r, the channel state S and one of the source inputs X_D, i.e. $Y = f(X_r, X_D, S)$.

Notations: The shorthand notation $\boldsymbol{x}_i^j$ is used to abbreviate $(x_i, x_{i+1}, \ldots, x_j)$, $\boldsymbol{x}^i$ is used to abbreviate $(x_1, x_2, \ldots, x_i)$, and x_i is used to denote the *i*-th element of $\boldsymbol{x}^n$ where $1 \leq i \leq j \leq n$. The probability law of a random variable X will be denoted by P_X, and the conditional probability distribution of Y given X will be denoted by $P_{Y|X}$. The alphabet of a scalar random variable X is designated by the corresponding calligraphic letter $\mathcal{X}$. The cardinality of a set $\mathcal{A}$ will be denoted by $|\mathcal{A}|$. $\mathcal{T}_\varepsilon^n(X)$ denotes the set of ε-typical sequences $\boldsymbol{x}^n \in \mathcal{X}^n$.

II. System Model

As shown in Fig. 1, we consider the state-dependent relay channel with orthogonal components denoted by $P_{Y,Y_r|X_R,X_D,X_r,S}$, where, $Y \in \mathcal{Y}$ and $Y_r \in \mathcal{Y}_r$ are the channel outputs from the destination and the relay respectively, $X_R \in \mathcal{X}_R$ and $X_D \in \mathcal{X}_D$ are the orthogonal channel inputs from the source while $X_r \in \mathcal{X}_r$ is the channel input from the relay, and $S \in \mathcal{S}$ denotes the common random channel state that corrupts channel 1 and channel 2. The channel state at time instant i is independently drawn from the distribution Q_S and the state sequence S^n is known non-causally to both the source and the relay.

The message W is uniformly distributed over the set $\mathcal{W} = \{1, 2, \ldots, M\}$. The source transmits a message W to the destination with the help of a relay in *n* channel uses. Let $\boldsymbol{X}_R^n = (X_{R,1}, \ldots, X_{R,n})$, $\boldsymbol{X}_D^n = (X_{D,1}, \ldots, X_{D,n})$ and $\boldsymbol{X}_r^n = (X_{r,1}, \ldots, X_{r,n})$ be the channel inputs of the source and relay respectively, the relay channel is said to be memoryless and to have orthogonal components if

$$\begin{aligned} P(&\boldsymbol{y}_r^n, \boldsymbol{y}^n | \boldsymbol{x}_R^n, \boldsymbol{x}_D^n, \boldsymbol{x}_r^n, s^n) \\ &= \prod_i^n P(y_{r,i}|x_{r,i}, x_{R,i}, s_i) P(y_i|x_{r,i}, x_{D,i}, s_i) \end{aligned} \quad (1)$$

A (*M*, *n*)-code for the state-dependent relay channel with channel state known non-causally to the source and the relay consists of encoding function at the source

$$\varphi^n : \{1, 2, \ldots, M\} \times \boldsymbol{S}^n \to \mathcal{X}_D^n \times \mathcal{X}_R^n \quad (2)$$

a sequence of encoding functions at the relay

$$\varphi_{r,i}^n : \mathcal{Y}_r^{i-1} \times \boldsymbol{S}^n \to \mathcal{X}_r \quad (3)$$

for *i* = 1, 2, , *n*, and a decoding function at the destination

$$\phi^n : \mathcal{Y}^n \to \{1, 2, \ldots, M\} \quad (4)$$

The information rate R is defined as

$$R = \frac{1}{n}\log_2(M) \quad \text{bits per transmission}$$

An (ε_n, n, R)-code for the non-causal state-dependent relay channel with orthogonal components is a code having average probability of error smaller than ε_n, i.e.,

$$Pr\left(W \neq \phi^n(y^n)\right) \leq \varepsilon_n$$

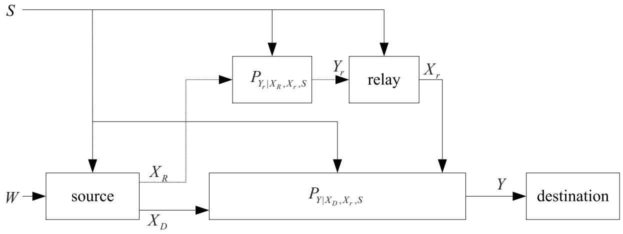

Fig. 1. Orthogonal relay channel with state information available at both the source and the relay

Rate *R* is said to be achievable if there exists a sequence of (ε_n, n, R)-codes with $\lim_{n\to\infty} \varepsilon_n = 0$. The capacity for the channel is defined as the supremum of the set of achievable rates.

III. Main Results

In this section, we first give the lower bound on the capacity for the channel shown in Fig. 1. Then, we characterized the capacity for a class of semi-deterministic relay channel whose output at the destination *Y* satisfies $Y = f(X_r, X_D, S)$.

The following theorem provides the lower bound on the capacity for the state-dependent relay channel shown in Fig.1.

Theorem 1: **(lower bound)** For the orthogonal relay channel with state non-causally known at both the source and relay, the following rate is achievable

$$\begin{aligned} R \leq \max\ \min\big\{ &I(U, U_r; Y) - I(U, U_r; S), \\ &I(X_R; Y_r|X_r, S) + I(U; Y|U_r) - I(U; S|U_r)\big\} \end{aligned} \quad (5)$$

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

$$P_{S,X_r,U_r,U,X_R,X_D} = Q_S P_{X_r,U_r|S} P_{X_R|U_r,X_r,S} P_{U,X_D|U_r,X_r,S} \quad (6)$$

U and U_r are auxiliary random variables with finite cardinality bounds.

Remark 1: Since both the source and the relay know the channel state non-causally, with partial-decode-and-forward (PDF) relaying, they can transmit the messages to the destination cooperatively with GP coding, namely cooperative GP coding. The source communicates with the relay treating $\boldsymbol{s}^n$ as a time-sharing sequence for the same reason that the channel state is known to both the source and the relay non-causally.

Proof: The achievable scheme is based on the combination of superposition coding at the source, PDF relaying at the relay and cooperative GP coding at the source and the relay. The message *W* is divided into two parts $W_D \in \{1, 2, \ldots, 2^{nR_D}\}$ and $W_R \in \{1, 2, \ldots, 2^{nR_R}\}$. Since both the source and the relay know the channel state non-causally, the achievability at the relay can be proved by treating $\boldsymbol{s}^n$ as a time-sharing sequence [22]. The encoding of the message W_R at the source and the decoding at the relay follows from [22]. Consider *B*+1 blocks, each of *n* symbols. A sequence of *B* messages $w(k) \in [1, 2^{nR}]$, *k* = 1, 2, , *B*, where, $w(k) = (w_D(k), w_R(k))$, $w_D(k) \in [1, 2^{nR_D}]$, $w_R(k) \in [1, 2^{nR_R}]$ and

$R = R_D + R_R$, will be sent over the channel in $n(B+1)$ transmissions. Let $\boldsymbol{s}^n(k)$ be the state sequence in block k and $\big(w_D(B+1), w_R(B+1)\big) = (1,1)$.

Codebook generation:

Fix a measure $P_{X_r,U_r|S}P_{X_R|U_r,X_r,S}P_{U,X_D|U_r,X_r,S}$.

(i) Generate $2^{n(R_R+R_{r,s})}$ independent and identically distributed (i.i.d.) codewords $\{\boldsymbol{u}_r^n(\tilde{w}_R, j_r)\}$, each with i.i.d. components drawn according to P_{U_r}.

(ii) For each $\{\boldsymbol{u}_r^n(\tilde{w}_R, j_r)\}$, generate $2^{n(R_D+R_{d,s})}$ i.i.d. codewords $\{\boldsymbol{u}^n(w_D, j_d|\tilde{w}_R, j_r)\}$, each with i.i.d. components drawn according to $P_{U|U_r}$.

(iii) For each $\{\boldsymbol{u}_r^n(\tilde{w}_R, j_r)\}$ and for each $s \in \mathcal{S}$, randomly and independently generate 2^{nR_s} sequences $\{\boldsymbol{x}_R^n(m_s|\tilde{w}_R, j_r, s)\}$, $m_s \in [1, 2^{nR_s}]$, each according to $P_{X_R|U_r,S}$. These sequences constitute the codebook $\mathcal{C}_s\,(s \in \mathcal{S})$. There are $|\mathcal{S}|$ such codebooks for each $\{\boldsymbol{u}_r^n(\tilde{w}_R, j_r)\}$. Set $R_R = \sum_{s\in\mathcal{S}} R_s$.

Encoding: We pick up the story in block k. Let $w(k) = (w_D(k), w_R(k)) \in \{1, 2, \ldots, 2^{nR}\}$, be the new message to be sent from the source node at the beginning of block k. The encoding at the beginning of block k is as follows.

(i) The relay knows $w_R(k-1)$ (this will be justified below), and searches the smallest $j_r(k) \in \{1, 2, \ldots, 2^{nR_{r,s}}\}$ such that $\boldsymbol{u}_r^n(w_R(k-1), j_r(k))$ is jointly typical with $\boldsymbol{s}^n(k)$. If no such $j_r(k)$ exists, an error is declared and $j_r(k)$ is set to 1. By covering lemma [23, lemma 3.3], this error probability tends to 0 as n approaches infinite, if

$$R_{r,s} \geq I(U_r; S) \tag{7}$$

Then, the relay sends a vector $\boldsymbol{x}_r^n(k)$ with i.i.d. components given $\boldsymbol{u}_r^n(w_R(k-1), j_r(k))$ and $\boldsymbol{s}^n(k)$, drawn according to the marginal $P_{X_r|U_r,S}$.

(ii) Similarly, knowing $\boldsymbol{u}_r^n(w_R(k-1), j_r(k))$, the source selects $\boldsymbol{u}^n(w_D(k), j_d(k)|w_R(k-1), j_r(k))$, and the probability that there exists such a $j_d(k) \in \{1, 2, \ldots, 2^{nR_{d,s}}\}$ tends to 1 if

$$R_{d,s} \geq I(U; S|U_r) \tag{8}$$

The source then sends a vector $\boldsymbol{x}_D^n(k)$ with i.i.d. components given $\boldsymbol{u}^n(w_D(k), j_d(k)|w_R(k-1), j_r(k))$ and $\boldsymbol{s}^n(k)$, drawn according to the marginal $P_{X_D|U,S}$.

(iii) Meanwhile, to send a message $w_R(k) \in [1, 2^{nR_R}]$, express it as a unique set of messages $\{m_s(k) : s \in \mathcal{S}\}$. Given $\boldsymbol{u}_r^n(w_R(k-1), j_r(k))$, consider the set of codewords $\{\boldsymbol{x}_R^n(m_s(k)|w_R(k-1), j_r(k), s) : s \in \mathcal{S}\}$. Store each codeword in a FIFO (first-in-first-out) buffer of length n. A multiplexer is used to choose a symbol at each transmission time $i \in [1, n]$ from one of the FIFO buffers according to the state $s_i(k)$. The chosen symbol is then transmitted.

Decoding: At the end of block k, the relay and the destination observe $\boldsymbol{y}_r^n(k)$ and $\boldsymbol{y}^n(k)$ respectively.

(i) Having decoded $w_R(k-1)$ successfully in block $k-1$ and knowing $\boldsymbol{x}_r^n(k)$, $\boldsymbol{u}_r^n(w_R(k-1), j_r(k))$ and $\boldsymbol{s}^n(k)$, the relay demultiplexes $\boldsymbol{y}_r^n(k)$ into subsequences $\{\boldsymbol{y}_r^{n_s(k)}(s)\}$, where $\sum_{s\in\mathcal{S}} n_s(k) = n$. Assuming $\boldsymbol{s}^n(k) \in \mathcal{T}_\varepsilon^n(s)$, and thus $n_s(k) \geq n(1-\varepsilon)p(s)$ for all $s \in \mathcal{S}$, it finds for each s, a unique $\hat{m}_s(k)$ such that the codeword subsequence $\boldsymbol{x}_R^{n(1-\varepsilon)p(s)}\big(\hat{m}_s(k)|w_R(k-1), j_r(k), s\big)$ is jointly typical with $\boldsymbol{y}_r^{n(1-\varepsilon)p(s)}(s)$ given $\boldsymbol{x}_r^{n(1-\varepsilon)p(s)}(k)$, where $p(s) = Q_S(S=s)$ By LLN (law of large number) and the packing lemma [23, lemma 3.1], the probability error of each decoding step approaches 0 as $n \to \infty$ if $R_s \leq p(s)I(X_R; Y_r|X_r, S=s)$. Therefore, the total probability error in decoding $\hat{w}_R(k)$ approaches 0 for sufficiently large n if the following condition is satisfied

$$R_R = \sum_{s\in\mathcal{S}} R_s \leq I(X_R; Y_r|X_r, S) \tag{9}$$

(ii) Observing $\boldsymbol{y}^n(k)$, the destination finds a quadruple $\Big(\hat{w}_R(k-1), \hat{j}_r(k), \hat{w}_D(k), \hat{j}_d(k)\Big)$ such that

$$\Big(\boldsymbol{u}^n\big(\hat{w}_D(k), \hat{j}_d(k)|\hat{w}_R(k-1), \hat{j}_r(k)\big), \boldsymbol{u}_r^n\big(\hat{w}_R(k-1), \hat{j}_r(k)\big), \boldsymbol{y}^n\Big) \in \mathcal{T}_\varepsilon^n(U, U_r, Y)$$

If there is no such quadruple or it is not unique, an error is declared. By packing lemma [23, lemma 3.1], it can be shown that for sufficiently large n, decoding is correct with high probability if

$$\begin{aligned} R_D + R_{d,s} &\leq I(U; Y|U_r) \\ R_D + R_{d,s} + R_R + R_{r,s} &\leq I(U, U_r; Y) \end{aligned} \tag{10}$$

Combining (7)-(10), $w(k-1) = (w_D(k-1), w_R(k-1))$ is decoded correctly with high probability at the end of block k, if

$$\begin{aligned} R &\leq I(X_R; Y_r|X_r, S) + I(U; Y|U_r) - I(U; S|U_r) \\ R &\leq I(U, U_r; Y) - I(U, U_r; S) \end{aligned} \tag{11}$$

This completes the proof.

The next theorem shows that the lower bound derived in Theorem 1 is tight for the channel whose output Y is a deterministic function of X_D, X_r and S, i.e. $Y = f(X_D, X_r, S)$.

Theorem 2: **(capacity)** For the state-dependent relay channel with orthogonal components and non-causal channel state known to the source and the relay, if $Y = f(X_D, X_r, S)$, the capacity for this channel is characterized as

$$\begin{aligned} C = \max\min\big\{&H(Y) - I(U_r, Y; S), \\ &I(X_R; Y_r|X_r, S) + H(Y|U_r, S)\big\} \end{aligned} \tag{12}$$

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

$$P_{S,X_r,U_r,X_R,X_D} = Q_S P_{X_r,U_r|S} P_{X_R|U_r,X_r,S} P_{X_D|U_r,X_r,S} \tag{13}$$

U_r is an auxiliary random variable with finite cardinality bound.

Proof: The achievability follows form Theorem 1 by taking $U = Y$ since $Y = f(X_D, X_r, S)$. The proof of the converse is as follows.

Consider a (ε_n, n, R)-code with an average error probability $P_e^{(n)} \leq \varepsilon_n$. By Fano's inequality, we have

$$H(W|\boldsymbol{Y}^n) \leq nRP_e^{(n)} + 1 = n\delta_n \tag{14}$$

where $\delta_n \to 0$ as $n \to +\infty$. Thus,

$$nR = H(W) \le I(W;\boldsymbol{Y}^n) + n\delta_n \tag{15}$$

Defining the auxiliary random variable $\bar{U}_{r,i} = \left(\boldsymbol{Y}^{i-1}, \boldsymbol{S}_{i+1}^n\right)$, we have

$$\begin{aligned}
I(W;\boldsymbol{Y}^n) &\le I(W;\boldsymbol{Y}^n,\boldsymbol{Y}_r^n)\\
&\le I(W;\boldsymbol{Y}^n,\boldsymbol{Y}_r^n|\boldsymbol{S}^n)\\
&= \sum_i I(W;Y_i,Y_{r,i}|\boldsymbol{Y}^{i-1},\boldsymbol{Y}_r^{i-1},\boldsymbol{S}^n)\\
&= \sum_i I(W;Y_{r,i}|\boldsymbol{Y}^{i-1},\boldsymbol{Y}_r^{i-1},\boldsymbol{S}^n)\\
&\quad +\sum_i I(W;Y_i|\boldsymbol{Y}^{i-1},\boldsymbol{Y}_r^{i-1},Y_{r,i},\boldsymbol{S}^n)
\end{aligned} \tag{16}$$

where, the second inequality follows from that $\boldsymbol{S}^n$ is independent of W.

Calculate the two items in (16) separately as follows:

$$\begin{aligned}
&\sum_i I(W;Y_{r,i}|\boldsymbol{Y}^{i-1},\boldsymbol{Y}_r^{i-1},\boldsymbol{S}^n)\\
&= \sum_i H(Y_{r,i}|\boldsymbol{Y}^{i-1},\boldsymbol{Y}_r^{i-1},\boldsymbol{S}^n)\\
&\quad -\sum_i H(Y_{r,i}|\boldsymbol{Y}^{i-1},\boldsymbol{Y}_r^{i-1},\boldsymbol{S}^n,W)\\
&\overset{(a)}{=} \sum_i H(Y_{r,i}|\boldsymbol{Y}^{i-1},\boldsymbol{Y}_r^{i-1},\boldsymbol{S}^n,X_{r,i})\\
&\quad -\sum_i H(Y_{r,i}|\boldsymbol{Y}^{i-1},\boldsymbol{Y}_r^{i-1},\boldsymbol{S}^n,W,X_{r,i},X_{R,i})\\
&\overset{(b)}{\le} \sum_i (Y_{r,i}|X_{r,i},S_i) - H(Y_{r,i}|X_{r,i},S_i,X_{R,i})\\
&= \sum_i I(X_{R,i};Y_{r,i}|X_{r,i},S_i)
\end{aligned} \tag{17}$$

where, (a) follows since $X_{r,i}$ is the function of $\left(\boldsymbol{Y}_r^{i-1},\boldsymbol{S}^n\right)$ and $X_{R,i}$ is the function of $(W,\boldsymbol{S}^n)$; (b) follows from the fact that conditioning reduces entropy and the Markov chain $\left(\boldsymbol{Y}^{i-1},\boldsymbol{Y}_r^{i-1},\boldsymbol{S}_{i+1}^n,\boldsymbol{S}^{i-1},W\right) \leftrightarrow (X_{r,i},X_{R,i},S_i) \leftrightarrow Y_{r,i}$.

$$\begin{aligned}
&\sum_i I(W;Y_i|\boldsymbol{Y}^{i-1},\boldsymbol{Y}_r^{i-1},Y_{r,i},\boldsymbol{S}^n)\\
&= \sum_i H(Y_i|\boldsymbol{Y}^{i-1},\boldsymbol{Y}_r^{i-1},Y_{r,i},\boldsymbol{S}^n)\\
&\quad -\sum_i H(Y_{r,i}|\boldsymbol{Y}^{i-1},\boldsymbol{Y}_r^{i-1},Y_{r,i},\boldsymbol{S}^n,W)\\
&= \sum_i H(Y_i|\boldsymbol{Y}^{i-1},\boldsymbol{Y}_r^{i-1},Y_{r,i},\boldsymbol{S}^n)\\
&\quad -\sum_i H(Y_i|\boldsymbol{Y}^{i-1},\boldsymbol{Y}_r^{i-1},Y_{r,i},\boldsymbol{S}^n,W,X_{D,i},X_{r,i})\\
&\overset{(a)}{=} \sum_i H(Y_i|\boldsymbol{Y}^{i-1},\boldsymbol{Y}_r^{i-1},Y_{r,i},\boldsymbol{S}^n)\\
&\overset{(b)}{\le} \sum_i H(Y_i|\boldsymbol{Y}^{i-1},\boldsymbol{S}_{i+1}^n,S_i)\\
&= \sum_i H(Y_i|\bar{U}_{r,i},S_i)
\end{aligned} \tag{18}$$

where, (a) follows since $Y_i = f\left(X_{D,i},X_{r,i},S_i\right)$; (b) follows that conditioning reduces entropy.

From (15)-(18), we have

$$R \le \frac{1}{n}\left(\sum_i I(X_{R,i};Y_{r,i}|X_{r,i},S_i) + H(Y_i|\bar{U}_{r,i},S_i)\right) + \delta_n \tag{19}$$

Meanwhile,

$$\begin{aligned}
I(W;\boldsymbol{Y}^n) &= \sum_i I(W;Y_i|\boldsymbol{Y}^{i-1})\\
&\le \sum_i I(W,\boldsymbol{Y}^{i-1};Y_i)\\
&= \sum_i I(W,\boldsymbol{Y}^{i-1},\boldsymbol{S}_{i+1}^n;Y_i) - I(\boldsymbol{S}_{i+1}^n;Y_i|W,\boldsymbol{Y}^{i-1})\\
&\overset{(a)}{=} \sum_i I(W,\boldsymbol{Y}^{i-1},\boldsymbol{S}_{i+1}^n;Y_i) - I(\boldsymbol{Y}^{i-1};S_i|W,\boldsymbol{S}_{i+1}^n)\\
&\overset{(b)}{=} \sum_i H(Y_i) - H(Y_i|W,\boldsymbol{Y}^{i-1},\boldsymbol{S}_{i+1}^n)\\
&\quad -\sum_i I(W,\boldsymbol{Y}^{i-1},\boldsymbol{S}_{i+1}^n;S_i)\\
&= \sum_i H(Y_i) - H(Y_i|W,\boldsymbol{Y}^{i-1},\boldsymbol{S}_{i+1}^n)\\
&\quad -I(W,\boldsymbol{Y}^{i-1},\boldsymbol{S}_{i+1}^n,Y_i;S_i) + I(Y_i;S_i|W,\boldsymbol{Y}^{i-1},\boldsymbol{S}_{i+1}^n)\\
&\overset{(c)}{\le} \sum_i H(Y_i) - I(W,\boldsymbol{Y}^{i-1},\boldsymbol{S}_{i+1}^n,Y_i;S_i)\\
&\le \sum_i H(Y_i) - I(\boldsymbol{Y}^{i-1},\boldsymbol{S}_{i+1}^n,Y_i;S_i)\\
&= \sum_i H(Y_i) - I(\bar{U}_{r,i},Y_i;S_i)
\end{aligned} \tag{20}$$

where, (a) follows by Csiszar and Korner's sum identity; (b) follows since S_i is independent of $\left(W,\boldsymbol{S}_{i+1}^n\right)$; (c) follows from the fact $H(Y_i|W,\boldsymbol{Y}^{i-1},\boldsymbol{S}_{i+1}^n) \ge I(Y_i;S_i|W,\boldsymbol{Y}^{i-1},\boldsymbol{S}_{i+1}^n)$.

By (15) and (20),

$$R \le \frac{1}{n}\sum_i \left(H(Y_i) - I(\bar{U}_{r,i},Y_i;S_i)\right) + \delta_n \tag{21}$$

From the above, we have

$$\begin{aligned}
R &\le \frac{1}{n}\left(\sum_i I(X_{R,i};Y_{r,i}|X_{r,i},S_i) + H(Y_i|\bar{U}_{r,i},S_i)\right) + \delta_n\\
R &\le \frac{1}{n}\sum_i \left(H(Y_i) - I(\bar{U}_{r,i},Y_i;S_i)\right) + \delta_n
\end{aligned} \tag{22}$$

Introduce a time-sharing random variable *T*, which is uniformly distributed over {1, 2, ..., *n*} and denote the collection of random variables

$$\left(X_R,X_r,Y_r,Y,\bar{U}_r,S\right) = \left(X_{R,T},X_{r,T},Y_{r,T},Y_T,\bar{U}_{r,T},S_T\right)$$

Considering the first bound in (22), we have

$$\begin{aligned}&\frac{1}{n}\left(\sum_i I(X_{R,i};Y_{r,i}|X_{r,i},S_i)+H(Y_i|\bar{U}_{r,i},S_i)\right)\\&\quad=I(X_R;Y_r|X_r,S,T)+H(Y|\bar{U}_r,S,T)\\&\quad=H(Y_r|X_r,S,T)-H(Y_r|X_R,X_r,S,T)\\&\qquad+H(Y|\bar{U}_r,S,T)\\&\quad\leq I(X_R;Y_r|X_r,S)+H(Y|\bar{U}_r,S,T)\end{aligned}\tag{23}$$

where, the last step follows from the fact that T is independent of all the other variables and the Markov chain $T\leftrightarrow(X_R,X_r,S)\leftrightarrow Y_r$.

Similarly, considering the second bound in (22), we have

$$\begin{aligned}&\frac{1}{n}\sum_i\big(H(Y_i)-I(\bar{U}_{r,i},Y_i;S_i)\big)\\&\quad=H(Y|T)-I(\bar{U}_r,Y;S|T)\\&\quad\leq H(Y)-I(\bar{U}_r,T,Y;S)+I(T;S)\\&\quad=H(Y)-I(\bar{U}_r,T,Y;S)\end{aligned}\tag{24}$$

where, the last step follows since T is independent of S.

Define $U_r=\big(\bar{U}_r,T\big)$, we have

$$\begin{aligned}&R\leq H(Y)-I(U_r,Y;S)+\delta_n\\&R\leq I(X_R;Y_r|X_r,S)+H(Y|U_r,S)+\delta_n\end{aligned}\tag{25}$$

Therefore, for a given sequence of (ε_n,n,R)-code with ε_n going to zero as *n* goes to infinity, there exists a measure of the form $P_{S,X_r,U_r,X_R,X_D}=Q_SP_{U_r|S}P_{X_r|U_r,S}P_{X_R,X_D|U_r,X_r,S}$ such that the rate *R* essentially satisfies (12). Without loss of generality, we can restrict the joint probability mass function to be of the form

$$\begin{aligned}&P_{S,X_r,U_r,X_R,X_D}\\&\quad=Q_SP_{U_r|S}P_{X_r|U_r,S}P_{X_R|U_r,X_r,S}P_{X_D|U_r,X_r,S}\end{aligned}$$

This concludes the proof.

IV. Conclusion

In this paper, we consider a state-dependent relay channel with orthogonal channels from the source to the relay and from the source and the relay to the destination. The orthogonal channels are corrupted by common channel state which is known to both the source and the relay non-causally. Lower bound on the capacity for the channel is established with superposition coding at the source, PDF relaying at the relay and cooperative GP coding at the source and the relay. We further show that if the output of the destination Y is a deterministic function of the relay input X_r, the channel state S and one of the source inputs X_D, i.e. $Y=f(X_r,X_D,S)$, the capacity can be characterized exactly.

Acknowledgment

The work was partly supported by the National Natural Science Foundation of China (Grant No. 61271232, 60972045, 61071089), the Priority Academic Program Development of Jiangsu Province, the University Postgraduate Research and Innovation Project in Jiangsu Province (No. CXZZ11_0395).

References

[1] G. Keshet , Y. Steinberg and N. Merhav, "Channel coding in the presence of side information: subject review," *Foundation and Trends in Communications and Information Theory*, Now publishers, 2008.

[2] C. E. Shannon, "Channels with side information at the transmitter," *IBM Journal Res. and Develop.*, vol. 2, pp. 289–293, Oct. 1958.

[3] S. I. Gelfand and M. S. Pinsker, "Coding for channel with random parameters," *Prob. of Contr. Inf. Theory*, vol. 9, pp.19–31, 1980.

[4] M. H. M. Costa, "Writing on dirty paper," *IEEE Trans. Inf. Theory*, vol. 29, pp.439–441, May 1983.

[5] S. I. Gelfand and M. S. Pinsker, "On Gaussian channel with random parameters," in *Proc. IEEE Int. Symp. Infor. Theory*, Tashkent, U.S.S.R., 1984, pp. 247–250.

[6] Y. Steinberg, "Coding for the degraded broadcast channel with random parameters, with causal and noncausal side information," *IEEE Trans. Inf. Theory*, vol. 51, no. 8, pp.2867–2877, Aug. 2005.

[7] S. Kotagiri and J. N. Laneman, "Multiple access channels with state information known to some encoders and independent messages," *EURASIP J. Wireless Commun. and Netw.*, vol. 2008, 2008.

[8] Y. Cemal and Y. Steinberg, "The multiple-access channel with partial state information at the encoders," *IEEE Trans. Inf. Theory*, vol. 51, no.11, pp. 3392–4003, Nov. 2005.

[9] N. Cen, G. Como, S. Yuksel and F. Alajaji, "On the capacity of memoryless finite-state multiple access channels with asymmetric noisy State information at the encoders," in *Proc. 49th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Sept. 2011, pp.1210–1215

[10] Y. Kim, A. Sutivong and S. Sigurjonsson, "Multiple user writing on dirty paper," in *Proc. IEEE Int. Symp. Infor. Theory*, Chicago, June. 2004, pp. 534.

[11] A. Somekh-Baruch, S. Shamai (Shitz), and S. Verd, "Cooperative multiple-access encoding with states available at one transmitter," *IEEE Trans. Inf. Theory*, vol. 54, no. 10, pp. 4448–4469, Oct. 2008.

[12] S. I. Bross, A. Lapidoth, and M. Wigger, "Dirty-paper coding for the Gaussian multiaccess channel with conferencing," *IEEE Trans. Inf. Theory*, vol. 58, no. 9, pp. 5640–5668, Sept. 2012.

[13] A. Zaidi, S. P. Kotagiri, J. Nicholas Laneman and L. Vandendorpe, "Cooperative relaying with state available noncausally at the relay," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2272–2298, May 2010.

[14] A. Zaidi and L. Vandendorpe, "Lower bounds on the capacity of the relay channel with states at the source," *EURASIP J. Wireless Commun. and Netw.*, vol. 2009, 2009.

[15] A. Zaidi, S. Shamai (Shitz), P. Piantanida and L. Vandendorpe, "Bounds on the capacity of the relay channel with noncausal state at source,"2011[online].available at http://arxiv.org/abs/1104.1057v1

[16] A. Zaidi, S. Shamai (Shitz), P. Piantanida and L. Vandendorpe, "On the Capacity of a Class of Relay Channels with Orthogonal Components and Noncausal State Information at Source," in *Proc. 2010 5th International Symposium on Wireless Pervasive Computing (ISWPC)*, 2010. pp. 528–533.

[17] K. Bakanoglu, E. Erkip, O. Simeone and S. Shamai (Shitz), "Relay channel with orthogonal components and structured interference known at the source,"2012[online].available at http://arxiv.org/abs/1202.4425

[18] A. Zaidi and L. Vandendorpe, "Rate regions for the partially-cooperative relay-broadcast channel with non-causal side information," in *Proc. IEEE Int. Symp. Infor. Theory (ISIT)*, Nice, France, June, 2007, pp. 1246–1250.

[19] T. M. Cover, A. EL Gamal, "Capacity Theorems for the Relay Channel," *IEEE Trans. Inf. Theory*, vol. IT-25, no. 5, pp. 572–584, 1979.

[20] Y. H. Kim, "Capacity of a Class of Relay Channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 3, pp. 1328–1329, 2008.

[21] A. El 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.

[22] A. J. Goldsmith and P. P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. Inf. Theory*, vol. 43, no. 6, pp. 1986–1992, 1997.

[23] A. El Gamal, Y.H. Kim, "Network information theory," *Cambridge University Press*, 2011.