The Capacity Region of the Parallel Partially Cooperative Relay Broadcast Channel with Unmatched Degraded Subchannels

Reza Khosravi-Farsani, Mahtab Mirmohseni, Bahareh Akhbari, Hamed Firouzi, and Mohammad Reza Aref

Information Systems and Security Lab (ISSL)
Department of Electrical Engineering, Sharif University of Technology, Tehran, Iran
Email:{reza.khosravi, mirmohseni, b_akhbari, h_firouzi}@ee.sharif.edu, and aref@sharif.edu

Abstract—We investigate the parallel partially cooperative relay broadcast channel with unmatched degraded subchannels. We characterize the capacity region of this channel in the general setting that, transmitter sends a common message for both users and a private message for each of them. Our achievability result is based on the choosing appropriate parameters for the general achievable rate region of partially cooperative relay broadcast channel. We provide a converse proof for this achievable region as well, which establishes the capacity region.

I. INTRODUCTION

In a classical Relay Channel, first introduced by Van der Meulen in [1], a relay assists the transmission in a single user channel between a sender-receiver pair. Cooperative relaying is a powerful technique which improves the throughput and the reliability in networks that include broadcast transmissions, such as wireless downlink communication systems [2]. In order to assume relaying and user-cooperation in downlink communications, the Relay Broadcast Channel (RBC) has been studied widely in [2], [3], where Partially and Fully cooperative RBC models are considered.

RBC is a communication channel which consists of a source node and two destination nodes where the source transmits both common and private messages to both destinations. Also, destinations are equipped with a transmitter which allows them to increase the rate via relaying information to each other. In Partially Cooperative Relay Broadcast Channel (PC-RBC), only one of the transmitters has relaying ability and could assist the other destination. The situation is like a broadcast channel, where one of the decoders could send information to the other one, in order to partially cooperate with. So, the building blocks of a PC-RBC are the broadcast and the relay channels, and the coding strategies which are used in [2], [3], are combination of those strategies for these two channels. The Fully Cooperative RBC (FC-RBC) is a more general model, where both destinations relay information to each other.

In [2], for PC-RBC and FC-RBC, inner and outer bounds on the capacity region are obtained, where encoding scheme is based on superposition coding at the source node and Decode-And-Forward (DF) strategy at the relay node [4, Theorem 1]. The inner bound derived in [3] for PC-RBC uses partial DF scheme [4, Theorem 7], [5], [6] instead of DF strategy. Also, in [2] the capacity region for the special case, where the PC-RBC is degraded, is established. These channels are also studied in [7] and [8]. The notion of sum and product of channels was first introduced by Shannon in [9]. In [10], El Gamal presented the sum and the product of unmatched Degraded Broadcast Channels (DBC) and derived the capacity region of these channels in the discrete memoryless case. He, also established the capacity region for the spectral Gaussian broadcast channels, which is a motivating example for studying the product of unmatched DBC.

In this paper we investigate the parallel PC-RBC with two unmatched degraded subchannels. This channel involves two PC-RBCs (subchannels I and II) with a source and two destinations, where each subchannel has orthogonal links with respect to the other one. Unmatched degradedness means that the output of the second destination in the first subchannel (subchannel I), is a degraded version of the output of the first destination (relay node) [3]. But, in the second subchannel (subchannel II), the output of the first destination (relay node) is a degraded version of the output of the second destination. Liang et al [3], have studied four classes of PC-RBCs which are: PC-RBC with degraded message sets, semi-deterministic PC-RBC, orthogonal PC-RBC and parallel PC-RBC with unmatched degraded subchannels. For the last case, the inner and the outer bounds on the capacity region were derived [3, Theorem 10], but these bounds do not match and the capacity, in general, is unknown for this channel. The inner bound is obtained using the achievable rate region established for the general PC-RBC in [3, Theorem 3] and the outer bound is based on the cut-set bound. Using the aforementioned inner bound, the capacity region is obtained in [3] for two special cases: 1) Degraded message sets, where the source has a common message for both destinations and a private message for the first destination only. 2) Parallel relay channel with unmatched degraded subchannels, i.e., source has only a private message for the second destination.

In this paper, we establish the capacity region of the parallel RBC.
PC-RBC with two unmatched degraded subchannels, in the general setting that, transmitter sends a common message for both users and a private message for each of them. We prove the achievability, using the achievable region given in [3, Theorem 2] for general PC-RBC. Note that the achievable region of [3, Theorem 3] is included in the achievable region given in [3, Theorem 2] for general PC-RBC. Hence, our inner bound contains the one derived in [3, Theorem 10] for parallel PC-RBC with unmatched degraded subchannels. Moreover, we provide a converse proof for our achievable region, which establishes the capacity region.

The rest of the paper is organized as follows. Section II gives the definition of the PC-RBC and the parallel PC-RBC with unmatched degraded subchannels and introduces the notations. In section III we state the capacity region for parallel PC-RBC with unmatched degraded subchannels. Achievability and converse proofs are also provided in this section. Finally, section IV concludes the paper.

II. PRELIMINARIES AND DEFINITIONS

The following notations are used throughout the paper. Random variables are denoted by upper case letters (e.g., $X$), and their realizations are shown with lower case letters (e.g., $x$). Designate alphabet sets are shown with $X$, ... The probability mass function (p.m.f) of a random variable $X$ is denoted by $p_X(x)$, where occasionally subscript $X$ is omitted. The notation $X^j_i$ indicates a sequence of random variables $(X_i, X_{i+1}, \ldots, X_j)$. For brevity, $X^j_i$ is used instead of $X^j_i$.

Definition 1: A discrete memoryless PC-RBC is denoted by $(X \times A_1, p(y_1, y_2|X, x_1), (Y_1 \times Y_2), p(y_1, y_2|x, x_1))$ is the probability distribution which characterizes the channel, $Y_1 \times Y_2$ are the outputs at destination 1 (the relay node) and 2. $X \times A_1$ are the source input and the relay (destination 1) input, respectively. The source sends a common message $W_0$ to both destinations and private messages $W_1$ and $W_2$ to destinations 1 and 2, respectively.

Definition 2: A parallel PC-RBC with unmatched degraded subchannels, shown in Fig. 1, consists of a vector source input alphabet $(X_{a1}, X_{a2})$, a vector relay input alphabet $(X_{r1}, X_{r2})$, a vector output alphabet at destination 1 (the relay node), $(Y_{d1}, Y_{d2})$, and a vector output alphabet at destination 2, $(Y_{a1}, Y_{a2})$. Note that, all alphabet are finite. A probability distribution $p(y_{a1}, y_{a2}|x_{a1}, x_{a2})p(y_{d1}, y_{d2}|x_{a1}, x_{a2})$ is defined, where to consider the unmatched physical degradedness property satisfies the following conditions [3]:

$$
p(y_{a1}, y_{a2}|x_{a1}, x_{a2}) = p(y_{a1}|x_{a1}, x_{a2})p(y_{a2}|y_{a1}, x_{a1})
$$

$$
p(y_{d1}, y_{d2}|x_{d1}, x_{d2}) = p(y_{d1}|x_{d1}, x_{d2})p(y_{d2}|y_{d1}, x_{d1})
$$

(1)

III. MAIN RESULTS

In this section, we characterize the capacity region of the parallel PC-RBC with unmatched degraded subchannels, defined in Definition 2. Before the main theorem we provide a useful lemma which is used in our achievability proof.

Lemma 1: ([3, Theorem 2]) The following, is an achievable rate region for the general PC-RBC:

$$
\mathcal{R} = \bigcup_{p(t, u, v, x, y, z)} \left\{ (R_0, R_1, R_2) : \begin{align*}
R_0 + R_1 &< I(T, U_1; Y_1 | X_1) \\
R_0 + R_2 &< I(T, X_1, U_2; Y_1) \\
R_0 + R_1 + R_2 &< I(T, U_1; Y_1 | X_1) \\
&+ I(2R_0 + R_1 + R_2 + I(T, U_1; Y_1 | X_1) \\
&+ I(T, X_1, U_2; Y_2) - I(U_1; U_2 | T, X_1) \\
&+ I(T, U_1; U_2 | X_1) \\
&+ I(T, U_1; U_2 | T, X_1)
\end{align*} \right\}
$$

(2)

Remark: Note that, the authors in [3] also derived another achievable region ([3, Theorem 3]) for the general PC-RBC, which is included in the above region as a subset. The achievable region obtained in [3, Theorem 10] for the parallel PC-RBC with unmatched degraded subchannels was based on the region of [3, Theorem 3], which is insufficient in achieving the capacity region of this channel. The capacity of special cases of this channel which were established in [3], was characterized using the region of [3, Theorem 3]. However, we utilize [3, Theorem 2] in our achievability proof of the capacity region of parallel PC-RBC with unmatched degraded subchannels.

Now, our main result is stated in Theorem 1.

Theorem 1: For the parallel PC-RBC with unmatched degraded subchannels, the capacity region is given by:

$$
\mathcal{C} = \bigcup_{p(u, v, x, y, z)} \left\{ (R_0, R_1, R_2) : \begin{align*}
R_0 + R_1 &\leq I(X_{a1}, Y_{d1}|X_{a1}) + I(V; Y_{a1}|X_{a1}) \\
R_0 + R_2 &\leq I(U, X_{a1}, Y_{d2}|X_{a1}) + I(X_{a1}, X_{a2}; Y_{a2}) \\
R_0 &+ R_1 + R_2 \leq I(V; Y_{a1} | X_{a1}) + I(V; Y_{a2} | X_{a1}) \\
R_0 + R_1 &+ R_2 \leq I(U, X_{a1}, Y_{d2}|X_{a1}) + I(U; Y_{a1} | X_{a1}) + I(U; X_{a1}; Y_{a2}) + I(X_{a1}, X_{a2}; Y_{a2})
\end{align*} \right\}
$$

(3)

for all joint p.m.f of the form:

$$
p(u, v, x_1, x_2, y_1, y_2) = p(u, x_1, x_2) p(v, x_1, x_2) p(y_1, y_2|x_1, x_2) p(y_1, y_2|x_1, x_2)
$$

(4)
where $U \in \mathcal{U}$ and $V \in \mathcal{V}$ are auxiliary random variables with:

\[
|U| \leq |X_a||X_{1a}| + 2 \\
|V| \leq |X_a||X_{1a}| + 2
\]  

(5)

**Proof:**

**A. Achievability**

By setting $X = (X_a, X_b)$, $X_1 = (X_{1a}, X_{1b})$, $Y_1 = (Y_{1a}, Y_{1b})$, $T = (U, V)$, $U_1 = (X_a, V)$, $U_2 = (U, X_a)$ in the rate region given in (2), and computing it for joint p.m.f (4), we obtain the following achievable region:

\[
\begin{align*}
R_0 + R_1 &\leq I(U, V, X_a; Y_{1a}, Y_{1b}|X_{1a}, X_{1b}) \\
R_0 + R_2 &\leq I(U, V, X_a, X_{1b}, X_b; Y_{2a}, Y_{2b}) \\
R_0 + R_1 + R_2 &\leq I(U, V, X_a; Y_{1a}, Y_{1b}|X_{1a}, X_{1b}) \\
R_0 + R_1 + R_2 &\leq I(X_a, V; Y_{1a}, Y_{1b}|U, V, X_{1a}, X_{1b}) \\
&\quad + I(U, V, X_{1a}, X_{1b}, X_b; Y_{2a}, Y_{2b}) \tag{6}
\end{align*}
\]

Note that for the joint p.m.f (4), $(U, X_a, X_b, Y_{1a}, Y_{1b})$ is independent of $(V, X_{1b}, X_b, Y_{2a}, Y_{2b})$, so we have:

\[
I(U_1; U_2|T, X_1) = I(X_a, V; X_a, U|U, V) = I(X_a; U|U) + I(V; X_a|V) = 0 \tag{7}
\]

Hence, the last bound in (2) is not advantageous. Because due to (7), it is obtained by adding the first two bounds in (2). Now, we compute the bounds in (6). Considering the joint p.m.f (4) and the Markov chain $U \rightarrow (X_a, X_{1a}) \rightarrow Y_{1a}$, we obtain:

\[
\begin{align*}
I(U, V, X_a; Y_{1a}, Y_{1b}|X_{1a}, X_{1b}) \\
&= I(U, X_a; Y_{1a}|X_{1a}) + I(V; Y_{1b}|X_{1b}) \tag{8}
\end{align*}
\]

Also, we have:

\[
\begin{align*}
I(U, V, X_{1a}, X_{1b}, X_b; Y_{2a}, Y_{2b}) \\
&= I(U, X_{1a}; Y_{2a}) + I(V, X_{1b}, X_b; Y_{2b}) \tag{9}
\end{align*}
\]

where the last equality is due to Markov chain $V \rightarrow (X_b, X_{1b}) \rightarrow Y_{2b}$, and the last two bounds in (6) are:

\[
\begin{align*}
I(U, V; Y_{2a}, Y_{2b}|U, V, X_{1a}, X_{1b}) \\
&= I(U; Y_{2a}|U, X_{1a}) + I(V; Y_{2b}|V, X_{1b}) \\
&= I(X_a; Y_{1a}|X_{1a}) + I(V; Y_{1b}|X_{1b}) \tag{10}
\end{align*}
\]

\[
\begin{align*}
I(X_a; Y_{1a}|U, V, X_{1a}, X_{1b}) \\
&= I(X_a; Y_{1a}|X_{1a}, U) + I(V; Y_{1b}|X_{1b}) \\
&= I(X_a; Y_{1a}|X_{1a}, U) \tag{11}
\end{align*}
\]

Now substituting (8) - (11) in (6) completes the proof of achievability.

**B. Converse**

Define random variables $V_i$ and $U_i$ as:

\[
\begin{align*}
U_i &\triangleq (W_0, W_2, Y_{1a}^{i-1}, Y_{1b}^{i-1}, Y_{2a}^{i-1}, Y_{2b}^{i-1}) \\
V_i &\triangleq (W_0, W_1, Y_{1a}^{i-1}, Y_{1b}^{i-1}, Y_{2a}^{i-1}, Y_{2b}^{i-1}) \tag{12}
\end{align*}
\]

Now using Fano’s inequality, we derive four bounds in (3). For the first bound, we have:

\[
\sum_{i=1}^{n} I(W_0, W_1; Y_{1a}^{i-1}, Y_{2a}^{i-1}) - n\delta_{1,n} \leq I(W_0, W_1; Y_{1a}^{n}, Y_{2a}^{n}) \tag{A}
\]

\[
= \sum_{i=1}^{n} I(W_0, W_1; Y_{1a}^{i-1}, Y_{1b}^{i-1}, Y_{1a}^{i-1}, Y_{1b}^{i-1}) \tag{A}
\]

\[
= \sum_{i=1}^{n} I(W_0, W_1; Y_{1a}^{i-1}, Y_{1b}^{i-1}, Y_{1a}^{i-1}, Y_{1b}^{i-1}) \tag{B}
\]

\[
= \sum_{i=1}^{n} I(W_0, W_1; Y_{1a}^{i-1}, Y_{1b}^{i-1}, Y_{1a}^{i-1}, Y_{1b}^{i-1}) \tag{C}
\]

\[
= \sum_{i=1}^{n} I(W_0, W_1; Y_{1a}^{i-1}, Y_{1b}^{i-1}, Y_{1a}^{i-1}, Y_{1b}^{i-1}) \tag{D}
\]

in which (A) is obtained from chain rule, (B) is due to the fact that $X_{1a}$ and $X_{1b}$ are deterministic functions of $Y_{1a}^{i-1}$ and $Y_{1b}^{i-1}$, respectively, (C) is due to the fact that conditioning does not increase entropy, and (D) follows from the fact that, conditioning on $(X_{1a}, X_{1b}), Y_{1a}$ is independent of $(W_0, W_1, Y_{1a}^{i-1}, Y_{1b}^{i-1}, Y_{1a})$. To derive the second bound in (3), again we utilize Fano’s inequality.

\[
n(R_0 + R_2) - n\delta_{2,n} = I(W_0, W_2; Y_{2a}^{n}, Y_{2b}^{n}) \tag{13}
\]

\[
= \sum_{i=1}^{n} I(W_0, W_2; Y_{2a}^{i-1}, Y_{2b}^{i-1}) \tag{A}
\]

\[
= \sum_{i=1}^{n} I(W_0, W_2; Y_{2a}^{i-1}, Y_{2b}^{i-1}) \tag{B}
\]

\[
= \sum_{i=1}^{n} I(W_0, W_2; Y_{2a}^{i-1}, Y_{2b}^{i-1}) \tag{C}
\]

\[
= \sum_{i=1}^{n} I(W_0, W_2; Y_{2a}^{i-1}, Y_{2b}^{i-1}) \tag{D}
\]

(14)
where (A) follows from chain rule. Noting that conditioning does not increase entropy we have:

\[ n(R_0 + R_2) - n\delta_{2,n} \leq \sum_{t=1}^{n} n(H(Y_{2at}) \]

\[ = \sum_{t=1}^{n} H(Y_{2at} | Y_{2a}^{t-1}, X_{1at}) + \sum_{t=1}^{n} n(H(Y_{2at}) \]

\[ \leq \sum_{t=1}^{n} H(Y_{2at} | Y_{2a}^{t-1}, X_{1at}) + \sum_{t=1}^{n} n(H(Y_{2at}) \]

\[ = \sum_{t=1}^{n} \{I(U_t; Y_{2at}) + I(X_{1at}; Y_{2at})\} \]

(15)

where (A) is true since \( Y_{2at} \) conditioned on \( X_{1at}, X_{1i} \) is independent of \( (W_0, W_1, Y_{2a}^{t-1}, Y_{2b}^{t-1}, X_{2i}) \). Now we compute the third bound in (3), using Fano’s inequality.

\[ n(R_0 + R_1 + R_2) - n\delta_{3,n} \leq \sum_{t=1}^{n} I(W_{2i}; Y_{2at} | V_i, X_{1at}) \]

\[ = \sum_{t=1}^{n} I(W_{2i}; Y_{2at} | V_i, X_{1at}) \]

\[ \leq \sum_{t=1}^{n} I(W_{2i}; Y_{2at}) \]

(16)

where (A) holds since conditioning does not increase entropy, (B) holds because \( (W_0, W_1) \) is independent of \( W_2 \), and (C) follows from (12). Using chain rule we have:

\[ n(R_0 + R_1 + R_2) - n\delta_{3,n} \leq \sum_{t=1}^{n} n(H(Y_{2at} | Y_{1a}^{t-1}, Y_{1b}^{t-1}) \]

\[ = \sum_{t=1}^{n} I(W_{2i}; Y_{2at}) \]

(17)

Using the degradedness conditions in (1), the forth and the sixth terms on the right hand side of (17) vanish. The first term in right hand side of (17) can be bounded as:

\[ \sum_{t=1}^{n} I(W_{2i}; Y_{2at}) \]

\[ \leq \sum_{t=1}^{n} I(W_{2i}; Y_{1at} | X_{1at}) \]

\[ \leq \sum_{t=1}^{n} I(X_{1at}; X_{1at}) \]

(18)

where (A) follows from the fact that \( X_{1at} \) is a deterministic function of \( Y_{1a}^{t-1} \). Then, we bound the third term in right hand side of (17) in the following:

\[ \sum_{t=1}^{n} I(W_{2i}; Y_{2at} | V_i, X_{1at}) \]

\[ = \sum_{t=1}^{n} I(W_{2i}; Y_{1at} | X_{1at}) \]

(19)

where (A) holds since \( X_{1at} \) is a deterministic function of \( Y_{1a}^{t-1} \) and also of \( V_i \) due to (12), and (B) follows from two facts: \( X_{1at} \) is a deterministic function of \( (W_0, W_1, Y_2, V_i) \), and \((V_i, W_2) \rightarrow X_{1at} \) is a Markov chain. The sum of the second and the fifth terms in the right hand side of (17) is bounded as:

\[ \sum_{t=1}^{n} I(W_{2i}; Y_{1at} | X_{1at}) \]

\[ \leq \sum_{t=1}^{n} I(W_{2i}; Y_{1at} | X_{1at}) \]

(20)

where (A) is due to the fact that conditioning on \( (X_{1at}, X_{1at}) \), \( Y_{1at} \) is independent of \( (W_0, W_1, Y_2, Y_{1a}^{t-1}, Y_{1b}^{t-1}, X_{1at}) \), and \( X_{1at} \) is a deterministic function of \( Y_{1a}^{t-1} \). Furthermore, \( X_{1at} \) is a deterministic function of \( (W_0, W_1, W_2) \). Now substituting (18)-(20) in (17) results in:

\[ n(R_0 + R_1 + R_2) - n\delta_{3,n} \leq \sum_{t=1}^{n} I(V_i; | X_{1at}) \]

(21)

Finally the last bound in (3), utilizing Fano’s inequality is obtained as:

\[ n(R_0 + R_1 + R_2) - n\delta_{3,n} \leq \sum_{t=1}^{n} I(V_i; | X_{1at}) \]

\[ = \sum_{t=1}^{n} I(X_{at}; X_{1at}) \]

(22)

in which (A) is true since conditioning does not increase entropy. Using chain rule and noting that \( (X_{1at}, X_{1at}) \) is a
deterministic function of \((Y_{1a}^{t-1}, Y_{1b}^{t-1})\), we have:
\[
\begin{align*}
n(R_0 + R_1 + R_2) - n\delta_{4,n} &\leq \sum_{t=1}^{n} I(W_1; Y_{1at}|U_t, X_{1at}) \\
&+ \sum_{t=1}^{n} I(W_1; Y_{2at}|U_t, Y_{1at}, X_{1at}) \\
&+ \sum_{t=1}^{n} I(W_1; Y_{2bt}|U_t, Y_{1at}, X_{1at}, X_{1bt}) \\
&\leq \sum_{t=1}^{n} I(W_0, W_1, Y_{1a}^{t-1}, Y_{1b}^{t-1}, Y_{2a}^{t-1}, Y_{2b}^{t-1}, Y_{2bt}^{t-1}, X_{1bt}^{t-1})
\end{align*}
\] (23)

(28)

where \((A)\) holds since, \(Y_{2bt}\) conditioned on \((X_{1bt}, X_{1bt}^{t-1})\) is independent of \((W_0, W_1, W_2, Y_{1a}^{t-1}, Y_{1b}^{t-1}, Y_{2a}^{t-1}, Y_{2b}^{t-1})\). So, substituting (25)-(28) in (24), the last bound becomes:
\[
\begin{align*}
n(R_0 + R_1 + R_2) - n\delta_{4,n} &\leq \sum_{t=1}^{n} \{I(X_{at}; Y_{1at}|U_t, X_{1at}) + I(U_t, X_{1at}; Y_{2at})\} \\
&+ \sum_{t=1}^{n} I(X_{at}, X_{1bt}; Y_{2bt})
\end{align*}
\] (29)

Finally, note that all terms in the bounds (13), (15), (21), and (29) are determined either by the distribution \(p(u_t, x_{at}, x_{1at}|y_{1at}, y_{2at})\) or by the distribution \(p(v_t, x_{at}, x_{1bt}|y_{1bt}, y_{2bt})\). Hence there is no loss of generality to consider only joint distributions of the form \(p(u_t, x_{at}, x_{1at})p(v_t, x_{at}, x_{1bt})\) for \((U_t, V_t, X_{at}, X_{1at}, X_{bt}, X_{1bt})\). The remaining part of the converse follows by applying a standard time-sharing argument, which completes the proof. ■

IV. CONCLUSION

In this paper, the parallel PC-RBC with unmatched degraded subchannels was considered. We obtained the achievable rate region for the general PC-RBC, and we provided a converse proof for this achievable region, so the capacity region was established in the general case. In future work we intend to study the Gaussian degraded setting.

REFERENCES