

Cooperative Strategies for the Half-Duplex Gaussian Parallel Relay Channel: Simultaneous Relaying versus Successive Relaying

Seyed Saeed Changiz Rezaei, Shahab Oveis Gharan, and Amir K. Khandani

Coding & Signal Transmission Laboratory
Department of Electrical & Computer Engineering
University of Waterloo
Waterloo, ON, N2L 3G1
sschangi, shahab, khandani@cst.uwaterloo.ca

Abstract

This study investigates the problem of communication for a network composed of two half-duplex parallel relays with additive white Gaussian noise. Two protocols, i.e., *Simultaneous* and *Successive* relaying, associated with two possible relay orderings are proposed. The simultaneous relaying protocol is based on *Dynamic Decode and Forward (DDF)* scheme. For the successive relaying protocol: (i) a *Non-Cooperative* scheme based on the *Dirty Paper Coding (DPC)*, and (ii) a *Cooperative* scheme based on the *Block Markov Encoding (BME)* are considered. Furthermore, the composite scheme of employing BME at one relay and DPC at another always achieves a better rate when compared to the *Cooperative* scheme. A “*Simultaneous-Successive Relaying based on Dirty paper coding scheme*” (*SSRD*) is also proposed. The optimum ordering of the relays and hence the capacity of the half-duplex Gaussian parallel relay channel in the low and high signal-to-noise ratio (SNR) scenarios is derived. In the low SNR scenario, it is revealed that under certain conditions for the channel coefficients, the ratio of the achievable rate of the simultaneous relaying based on DDF to the cut-set bound tends to be 1. On the other hand, as SNR goes to infinity, it is proved that successive relaying, based on the DPC, asymptotically achieves the capacity of the network.

I. INTRODUCTION

A. Motivation

The continuous growth in wireless communication has motivated information theoretists to extend Shannon's information theoretic arguments for a single user channel to the scenarios that involve communication among multiple users.

Financial supports provided by Nortel, and the corresponding matching funds by the Federal government: Natural Sciences and Engineering Research Council of Canada (NSERC) and Province of Ontario: Ontario Centres of Excellence (OCE) are gratefully acknowledged.

In this regard, cooperative wireless communication has been the focus of attention during recent years. Due to rapid decrease of the transmitted signal power with distance, the idea of multi-hopped communication has been proposed. In multi-hopped communication, some intermediate nodes as relays are exploited to facilitate data transmission from the source to the destination. Using this technique leads to saving battery power as well as increasing the physical coverage area. Moreover, relays by emulating distributed transmit antenna, can form spatial diversity and combat the multi-path fading effect of the wireless media.

Motivated by practical constraints, half-duplex relays which cannot transmit and receive at the same time and in the same frequency band are of great importance. Here, our goal is to study and analyze the performance limits of a half-duplex parallel relay channel.

B. History

Relay channel is a three terminal network which was introduced for the first time by Van der Meulen in 1971 [1]. The most important capacity results of the relay channel were reported by Cover and El Gamal [2]. Two relaying strategies are proposed in [2]. In one strategy, the relay decodes the transmitted message and forwards the re-encoded version to the destination, while in another one the relay does not decode the message, but sends the quantized received values to the destination.

Moreover, several works on multi-relay channels exist in the literature (See [3]–[11], [23], [29]–[36]). Schein in [3], [4] establishes upper and lower bounds on the capacity of a full-duplex parallel relay channel in which the channel consists of a source, two relays and a destination, where there is no direct link between the source and the destination, and also between the two relays. Generally, the best rate reported for the full-duplex Gaussian parallel relay channel is based on the Decode-Forward (DF) or Amplify-Forward (AF) schemes, with time sharing [3], [4]. Xie and Kumar generalize the block Markov encoding scheme in [2] for a network of multiple relays [5]. Gastpar, Kramer, and Gupta extend compress and forward scheme to a multiple relay channel by introducing the concept of antenna polling in [6]–[8]. In [9], Amichai, Shamai, Steinberg and Kramer consider a parallel relay setup, in which a nomadic source sends its information to a remote destination via some relays with lossless links to the destination. They investigate the case that these relays do not have any decoding capability, so signals received at the relays must be compressed. The authors also fully characterize the capacity of this case for the Gaussian channel. In [10], Maric and Yates investigate DF and AF schemes in a parallel-relay network. Motivated by applications in sensor networks, they assume large bandwidth resources allowing orthogonal transmissions at different nodes. They characterize optimum resource allocation for AF and DF and show that the wide-band regime minimizes the energy cost per information bit in DF, while AF should work in the band-limited regime to achieve the best rate. Razaghi and Yu in [11] propose a parity-forwarding scheme for full-duplex multiple relay. They show that parity-forwarding can achieve the capacity in a new form of degraded relay networks.

Radios that can receive and transmit simultaneously in the same frequency band require complex and expensive components [18]. Hence, Khojastepour and Aazhang in [13], [14] call the half-duplex relay as “*Cheap Relay*”.

Recently, half-duplex relaying has drawn a great deal of attention (See [13]–[19], [23], [29]–[36]). Zahedi and

El Gamal consider two different cases of frequency division Gaussian relay channel, deriving lower and upper bounds on the capacity [15]. They also derive single letter characterization of the capacity of frequency division additive white Gaussian noise (AWGN) relay channel with simple linear relaying scheme [16], [17]. The problem of time division relaying is also considered by Host-Madsen and Zhang [18]. By considering fading scenarios, and assuming channel state information (CSI), they study upper and lower bounds on the outage capacity and the Ergodic capacity. In [19], Liang and Veeralli present a Gaussian orthogonal relay model, in which the relay-to-destination channel is orthogonal to the source-to-relay and source-to-destination channel. They show that when the source-to-relay channel is better than the source-to-destination channel and the signal-to-noise ratio (SNR) of the relay-to-destination is less than a given threshold, optimizing resource allocation causes the lower and the upper bounds to coincide with each other.

C. Contributions and Relation to Previous Works

In this paper, we study transmission strategies for a network with a source, a destination, and two half-duplex relays with additive white Gaussian noise which cooperate with each other to facilitate data transmission from the source to the destination. Furthermore, it is assumed that no direct link exists between the source and the destination.

Half-duplex relaying, in multiple relay networks, is studied in [23], [29]–[36]. Gastpar in [23] shows that in a Gaussian parallel relay channel with infinite number of relays, the optimum coding scheme is AF. Rankov and Wittneben in [29], [30] further study the problem of half-duplex relaying in a two-hop communication scenario. In their study, they also consider a parallel relay setup with two relays where there is no direct link between the source and the destination, while there exists a link between the relays. Their relaying protocols are based on either AF or DF, in which the relays successively forward their messages from the source to the destination. We call this protocol “*Successive Relaying*” in the sequel. Xue and Sandhu in [31] further study different half-duplex relaying protocols for the Gaussian parallel relay channel. Since they assume that there is no link between the relays, they refer to their parallel channel as a *Diamond Relay Channel*.

In this work, our primary objective is to find the best ordering of the relays in the intended set-up. We consider two relaying protocols, i.e., simultaneous relaying versus successive relaying, associated with two possible relay orderings. For simultaneous relaying, each relay exploits “Dynamic DF (DDF)”. It should be noted that the DDF scheme considered here is slightly different from the DDF introduced in [34] and [35]. In those works, the DDF scheme is applied to the set-up of the multiple relay network in which the nodes only have the CSI of their receiving channel. In the DDF scheme described in [34], the source is broadcasting the message to all the network nodes during whole period of transmission and each relay, listens to the transmitted signal of the source and other relays until it can decode the transmitted message. Consequently, it transmits its signal coherently with the source and other active relays in the remaining time. However, in our set-up, all the nodes are assumed to have all the channel coefficients. Therefore, in a fixed pre-assigned portion of the time, the relays receive the signal transmitted from the source, and in the remaining time slot they transmit the re-encoded version of the decoded message together. In other words, the relays operate in a synchronous manner.

For successive relaying, we study a *Non-Cooperative* scheme based on “Dirty Paper Coding (DPC)” and also a *Cooperative* scheme based on “Block Markov Encoding (BME)”. It is worth noting that the authors in [36] also propose successive relaying protocol for the set up with two parallel relays and direct links between the relays and between the source and the destination. They propose a simple repetition coding at the relays, and show that their scheme can recover the loss in the multiplexing gain, while achieving diversity gain of 2.

We derive the optimum relay ordering in low and high SNR scenarios. In low SNR scenarios and under certain channel conditions, we show that the ratio of the achievable rate of DDF for simultaneous relaying to the cut-set bound tends to one. On the other hand, in high SNR scenarios, we prove that the proposed DPC for successive relaying asymptotically achieves the capacity.

After this work was completed, we became aware of [32] which has independently proposed an achievable rate based on the combination of superposition coding, BME and DPC. In their scheme, the intended message “ w ” is split into a message which is transmitted to the destination by exploiting cooperation between the relays “ w_r ” and a message which is transmitted to the destination without using any cooperation between the relays “ w_d ”. Hence, the signal associated with “ w_d ”, transmitted by one relay, can be considered as interference on the other relay. “ w_r ” is transmitted by using BME and “ w_d ” is transmitted by employing DPC. Therefore, in their general scheme, the associated signals with these two messages are superimposed and transmitted. As the channel between the two relays become strong, their proposed scheme is converted to BME. On the other hand, as the channel becomes weak, their proposed scheme becomes DPC.

Unlike [32], in which the authors only consider successive relaying and propose a combined BME and DPC, as the main result of this paper, simultaneous and successive relaying protocols are combined and a “Simultaneous-Successive Relaying based on Dirty paper coding” (SSRD) scheme with a new achievable rate is proposed. It is shown that in the low SNR scenario and under certain channel conditions, SSRD scheme is converted to simultaneous relaying based on DDF, while in the high SNR scenarios, when the ratio of the relay powers to the source power remain constant, it becomes successive relaying based on DPC (to achieve the capacity).

Besides this main result, some other results obtained in this paper are as follows:

- Two different types of decoding, i.e., *successive* and *backward* decoding, at the destination for the BME scheme are proposed. We prove that the achievable rate of BME with backward decoding is greater than that of BME with successive decoding, i.e., $C_{BME_{back}}^{low} \geq C_{BME_{succ}}^{low}$.
- It is proved that BME with backward decoding leads to a simple strategy in which at most, one of the relays is required to cooperate with the other relay in sending the bin index of the other relay’s message. Accordingly, in the Gaussian case, the combination of BME at one relay and DPC at the other relay always achieves a better rate than the simple BME.
- In the degraded case, where the destination receives a degraded version of the received signals at the relays, BME with backward decoding achieves the successive cut-set bound.

The rest of the paper is organized as follows: In section II, the system model is introduced. In section III, the achievable rates and coding schemes for a half-duplex relay network are derived. Optimality results are discussed

in section IV. Simulation results are presented in section V. Finally, section VI concludes the paper.

D. Notation

Throughout the paper, the superscript H stands for matrix operation of conjugate transposition. Lowercase bold letters and regular letters represent vectors and scalars, respectively. For any two functions $f(n)$ and $g(n)$, $f(n) = O(g(n))$ is equivalent to $\lim_{n \rightarrow \infty} \left| \frac{f(n)}{g(n)} \right| < \infty$, and $f(n) = \Theta(g(n))$ is equivalent to $\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = c$, where $0 < c < \infty$. And $C(x) \triangleq \frac{1}{2} \log_2(1+x)$. Furthermore, for the sake of brevity, $A_\epsilon^{(n)}$ denotes the set of weakly jointly typical sequences for any intended set of random variables.

II. SYSTEM MODEL

We consider a Gaussian network which consists of a source, two half-duplex relays, and a destination, and there is no direct link between the source and the destination. Here we define four time slots according to the transmitting and receiving mode of each relay (See Fig. 1), where t_b denotes the duration of time slot b ($\sum_{b=1}^4 t_b = 1$). Nodes 0, 1, 2, and 3 represent the source, relay 1, relay 2, and the destination, respectively. Moreover, the transmitting and receiving signals at node a during time slot b are represented by $\mathbf{x}_a^{(b)}$ and $\mathbf{y}_a^{(b)}$, respectively. Hence, at each node $c \in \{1, 2, 3\}$, we have

$$\mathbf{y}_c^{(b)} = \sum_{a \in \{0,1,2\}} h_{ac} \mathbf{x}_a^{(b)} + \mathbf{z}_c^{(b)}. \quad (1)$$

where h_{ac} 's denote channel coefficients from node a to node c , and $\mathbf{z}_c^{(b)}$ is the AWGN term with zero mean and variance of "1" per dimension.

Noting the transmission strategies in Fig. 1, we have

$$\mathbf{y}_1^{(1)} = h_{01} \mathbf{x}_0^{(1)} + h_{21} \mathbf{x}_2^{(1)} + \mathbf{z}_1^{(1)}, \quad (2)$$

$$\mathbf{y}_3^{(1)} = h_{23} \mathbf{x}_2^{(1)} + \mathbf{z}_3^{(1)}, \quad (3)$$

$$\mathbf{y}_2^{(2)} = h_{02} \mathbf{x}_0^{(2)} + h_{12} \mathbf{x}_1^{(2)} + \mathbf{z}_2^{(2)}, \quad (4)$$

$$\mathbf{y}_3^{(2)} = h_{13} \mathbf{x}_1^{(2)} + \mathbf{z}_3^{(2)}, \quad (5)$$

$$\mathbf{y}_k^{(3)} = h_{0k} \mathbf{x}_0^{(3)} + \mathbf{z}_k^{(3)}, k \in \{1, 2\}, \quad (6)$$

$$\mathbf{y}_3^{(4)} = \sum_{k=1}^2 h_{k3} \mathbf{x}_k^{(4)} + \mathbf{z}_3^{(4)}. \quad (7)$$

Throughout the paper, we assume that $h_{01} \geq h_{02}$ unless specified otherwise, and from reciprocity assumption, we have $h_{12} = h_{21}$. Furthermore, the power constraints P_0 , P_1 , and P_2 should be satisfied for the source, the first relay, and the second relay, respectively. Hence, denoting the power consumption of node a at time slot b by $P_a^{(b)} = E \left[\mathbf{x}_a^{(b)H} \mathbf{x}_a^{(b)} \right]$, we have

$$P_0^{(1)} + P_0^{(2)} + P_0^{(3)} = P_0, \quad (8)$$

$$P_1^{(2)} + P_1^{(4)} = P_1,$$

$$P_2^{(1)} + P_2^{(4)} = P_2.$$

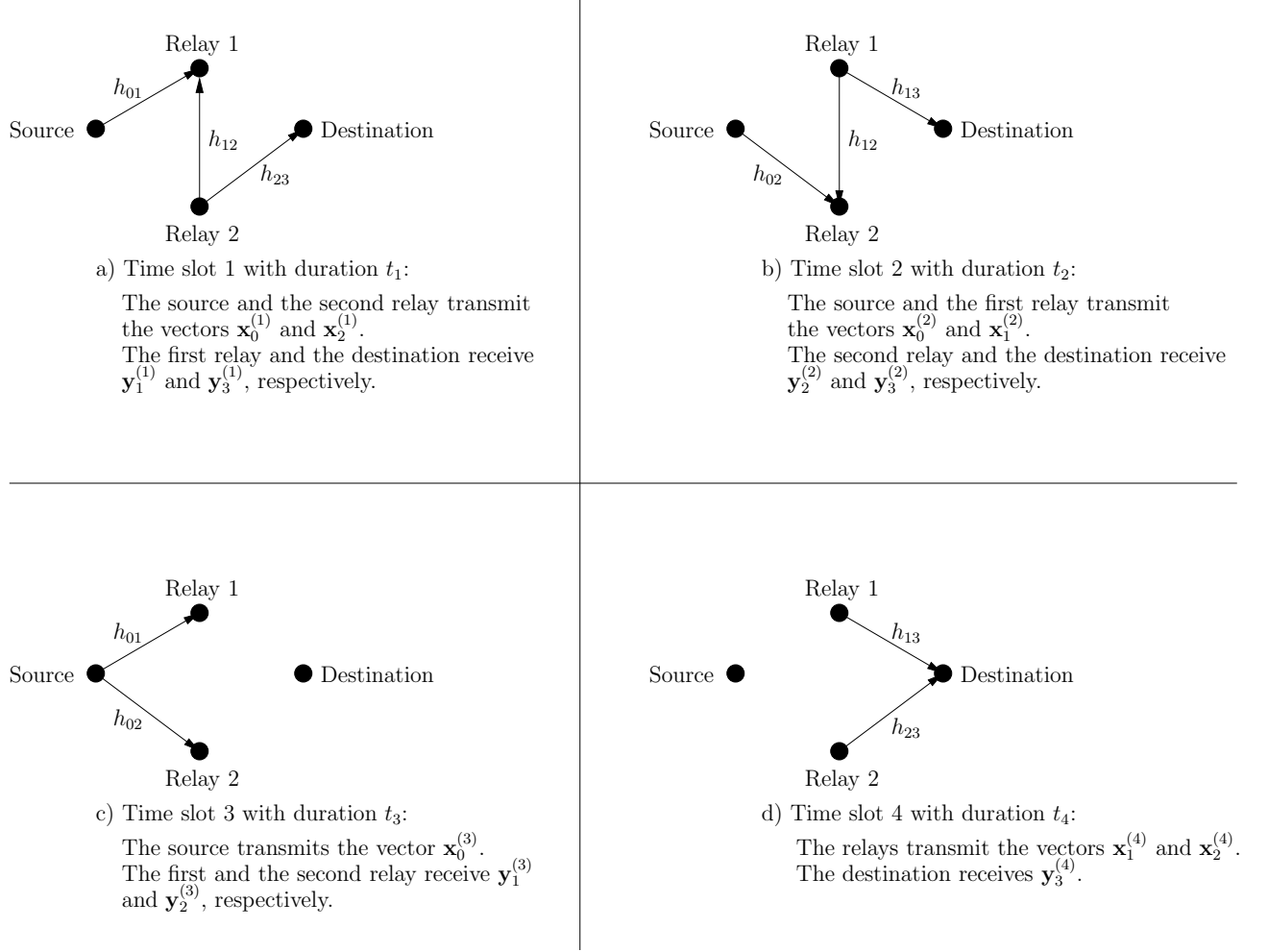


Fig. 1. System Model.

III. ACHIEVABLE RATES AND CODING SCHEMES

In this section, we propose two cooperative protocols, i.e. *Successive* and *Simultaneous* relaying protocols, for a half-duplex Gaussian parallel relay channel.

A. Successive Relaying Protocol

In *Successive* relaying protocol, the relays are not allowed to receive and transmit simultaneously, i.e. $t_3 = t_4 = 0$, and the relations between the transmitted and the received signals at the relays and at the destination follow from (2)-(5). For the successive relaying protocol, we propose a *Non-Cooperative* and a *Cooperative Coding* scheme in the sequel. In the proposed schemes, the time is divided into odd and even time slots with the duration t_1 and t_2 , respectively. Accordingly, at each odd and even time slots, the source transmits a new message to one of the relays, and the destination receives a new message from the other relay, successively (See Fig. 2).

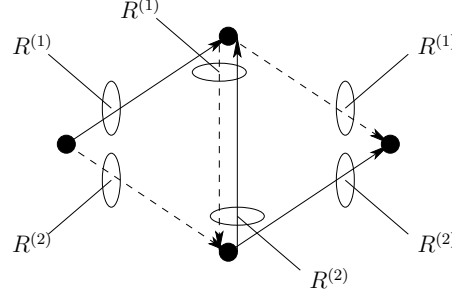


Fig. 2. Information flow transfer for successive relaying protocol for two relays.

1) *Non-Cooperative Coding*: In the Non-Cooperative Coding scheme, each relay considers the other's signal as interference. Since the source knows each relay's message, it can apply the Gelfand-Pinsker's coding scheme to transmit its message to the other relay. The following Theorem gives the achievable rate of this scheme.

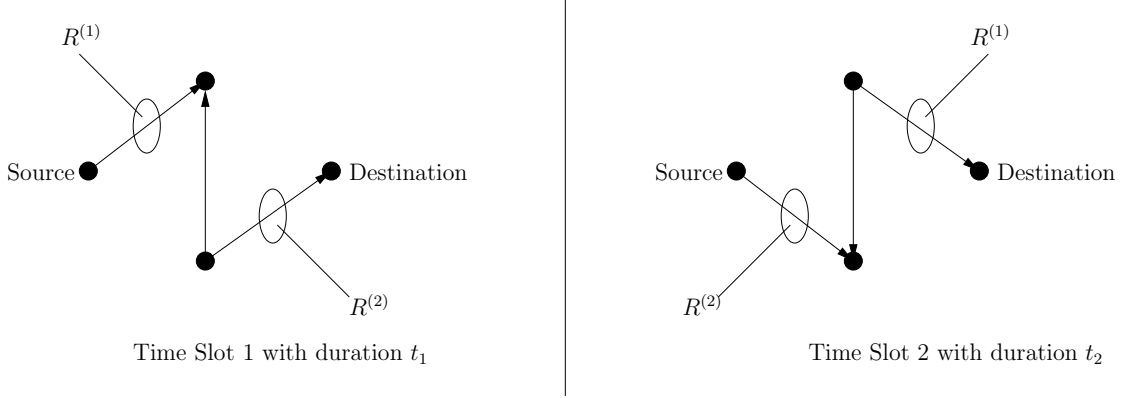


Fig. 3. Successive relaying protocol based on Non-Cooperative Coding.

Theorem 1 For the half-duplex parallel relay channel, assuming successive relaying, the following rate C_{DPC}^{low} is achievable:

$$C_{DPC}^{low} = \max_{0 \leq t_1, t_2, t_1+t_2=1} R^{(1)} + R^{(2)}, \quad (9)$$

subject to:

$$R^{(1)} \leq \min \left(t_1 (I(U_0^{(1)}; Y_1^{(1)}) - I(U_0^{(1)}; X_2^{(1)})), t_2 I(X_1^{(2)}; Y_3^{(2)}) \right), \quad (10)$$

$$R^{(2)} \leq \min \left(t_2 (I(U_0^{(2)}; Y_2^{(2)}) - I(U_0^{(2)}; X_1^{(2)})), t_1 I(X_2^{(1)}; Y_3^{(1)}) \right). \quad (11)$$

with probabilities:

$$p(x_2^{(1)}, u_0^{(1)}, x_0^{(1)}) = p(x_2^{(1)})p(u_0^{(1)}|x_2^{(1)})p(x_0^{(1)}|u_0^{(1)}, x_2^{(1)}),$$

$$p(x_1^{(2)}, u_0^{(2)}, x_0^{(2)}) = p(x_1^{(2)})p(u_0^{(2)}|x_1^{(2)})p(x_0^{(2)}|u_0^{(2)}, x_1^{(2)}).$$

Proof: See Appendix A. ■

From Theorem 1, the achievable rate of the proposed scheme for the Gaussian case can be obtained as follows.

corollary 1 *For the half-duplex Gaussian parallel relay channel, assuming successive relaying protocol with power constraint at the source and at each relay, DPC achieves the following rate:*

$$C_{DPC}^{low} = \max \left(R^{(1)} + R^{(2)} \right), \quad (12)$$

subject to:

$$R^{(1)} \leq \min \left(t_1 C \left(\frac{h_{01}^2 P_0^{(1)}}{t_1} \right), t_2 C \left(\frac{h_{13}^2 P_1}{t_2} \right) \right),$$

$$R^{(2)} \leq \min \left(t_2 C \left(\frac{h_{02}^2 P_0^{(2)}}{t_2} \right), t_1 C \left(\frac{h_{23}^2 P_2}{t_1} \right) \right),$$

$$P_0^{(1)} + P_0^{(2)} = P_0,$$

$$t_1 + t_2 = 1,$$

$$0 \leq t_1, t_2, P_0^{(1)}, P_0^{(2)}.$$

Proof: From Costa's Dirty Paper Coding [28], by having

$$U_0^{(1)} = X_0^{(1)} + \frac{h_{01} h_{12} P_0^{(1)}}{h_{01}^2 P_0^{(1)} + t_1} X_2^{(1)}, \quad (13)$$

$$U_0^{(2)} = X_0^{(2)} + \frac{h_{02} h_{12} P_0^{(2)}}{h_{02}^2 P_0^{(2)} + t_2} X_1^{(2)}. \quad (14)$$

where $X_0^{(1)} \sim \mathcal{N}(0, P_0^{(1)})$, $X_0^{(2)} \sim \mathcal{N}(0, P_0^{(2)})$, $X_2^{(1)} \sim \mathcal{N}(0, P_2)$, and $X_1^{(2)} \sim \mathcal{N}(0, P_1)$, and applying them to Theorem 1, we obtain corollary 1. ■

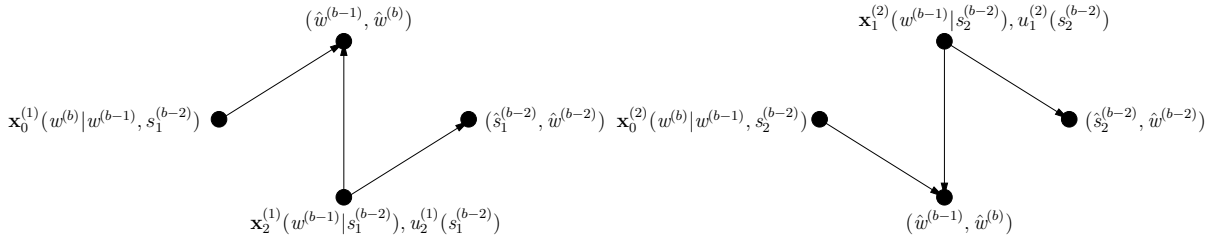


Fig. 4. Successive relaying protocol based on Cooperative Coding.

2) *Cooperative Coding:* In this type of coding scheme, we assume that, at each time slot, the receiving relay decodes not only the new transmitted message from the source, but also the previous message transmitted from the transmitting relay (See Figs. 2 and 4). Our proposed coding scheme is based on binning, superposition coding, and Block Markov Encoding. The source sends B messages $w^{(1)}, w^{(2)}, \dots, w^{(B)}$ in $B + 2$ time slots.

Generally, this scheme can be described as follows (See Figs. 4 and 5). In time slot b , the relay $(b + 1) \bmod 2 + 1$ decodes the transmitted messages $w^{(b)}$ and $w^{(b-1)}$ from the source and the other relay, respectively. In time slot

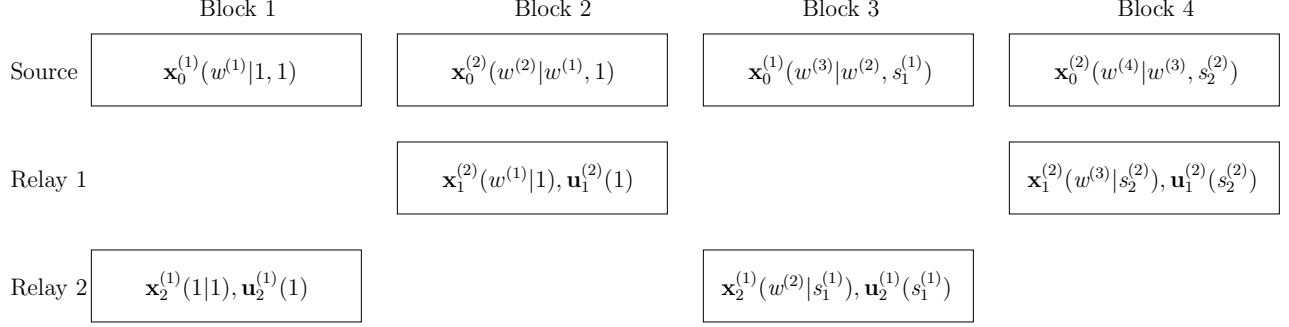


Fig. 5. Decode-and-forward for successive relaying protocol.

$b + 1$, it broadcasts $w^{(b)}$ and the bin index of $w^{(b-1)}$, $s_{(b+2) \bmod 2+1}^{(b-1)}$, to the destination using the binning function defined next.

Definition (The Binning Function): The binning function $f_{Bin}^{((b+1) \bmod 2+1)}(w^{(b-2)}) : W = \{1, 2, \dots, 2^{nR^{((b+1) \bmod 2+1)}}\} \longrightarrow \{1, 2, \dots, 2^{nr_{Bin}^{((b+1) \bmod 2+1)}}\}$ is defined by $f_{Bin}^{((b+1) \bmod 2+1)}(w^{(b-2)}) = s_{(b+1) \bmod 2+1}^{(b-2)}$, where $f_{Bin}^{((b+1) \bmod 2+1)}(\cdot)$ assigns a randomly uniform distributed integer between 1 and $2^{nr_{Bin}^{((b+1) \bmod 2+1)}}$ independently to each member of W .

As indicated in Fig. 5, in the first time slot, the source transmits the codeword $\mathbf{x}_0^{(1)}(w^{(1)}|1, 1)$ to the first relay, while the second relay transmits a doubly indexed codeword $\mathbf{x}_2^{(1)}(1|1)$ and the codeword $\mathbf{u}_2^{(1)}(1)$ to the first relay and to the destination. In the second time slot, the source transmits the codeword $\mathbf{x}_0^{(2)}(w^{(2)}|w^{(1)}, 1)$ to the second relay, and having decoded the message $w^{(1)}$, the first relay broadcasts the codewords $\mathbf{x}_1^{(2)}(w^{(1)}|1)$ and $\mathbf{u}_1^{(2)}(1)$ to the second relay and to the destination. It should be noted that the destination cannot decode the message $w^{(1)}$ at the end of this time slot; however, the second relay decodes $w^{(1)}$ and $w^{(2)}$ messages. Using the binning function, it finds the bin index of $w^{(1)}$ according to $s_1^{(1)} = f_{Bin}^{(1)}(w^{(1)})$. In the third time slot, the source transmits the codeword $\mathbf{x}_0^{(1)}(w^{(3)}|w^{(2)}, s_1^{(1)})$ to the first relay, and the second relay broadcasts the codewords $\mathbf{x}_2^{(1)}(w^{(2)}|s_1^{(1)})$ and $\mathbf{u}_2^{(1)}(s_1^{(1)})$ to the first relay and to the destination.

Two types of decoding can be used at the destination: successive decoding and backward decoding. Successive decoding at the destination can be described as follows. At the end of the b th time slot, the destination cannot decode the message $w^{(b-1)}$; however, having decoded the bin index $s_{(b+1) \bmod 2+1}^{(b-2)}$ from the received vector of the b th time slot, it can decode the message $w^{(b-2)}$ from $s_{(b+1) \bmod 2+1}^{(b-2)}$ and the received vector of the $(b-1)$ th time slot. On the other hand, backward decoding can be explained as follows. Having received the sequence of the $B+2$ 'th time slot, the final destination starts decoding the intended messages. In the time slot $B+2$, one of the relays transmits the dummy message "1" along with the bin index of the message $w^{(B)}$ to the destination. Having received this bin index, the destination decodes it, and then backwardly decodes messages $w^{(b)}$, $b = B, B-1, \dots, 1$ and their bin indices. The following Theorem gives the achievable rate of the proposed scheme.

Theorem 2 *For the half-duplex parallel relay channel, assuming successive relaying, the BME scheme achieves*

the rates $C_{BME_{succ}}^{low}$ and $C_{BME_{back}}^{low}$ using successive and backward decoding, respectively:

$$\begin{aligned}
C_{BME_{succ}}^{low} = R^{(1)} + R^{(2)} \leq & \max_{0 \leq t_1, t_2, t_1+t_2=1} \min (\\
& \min \left(t_1 I \left(X_0^{(1)}; Y_1^{(1)} \mid X_2^{(1)}, U_2^{(1)} \right), t_2 I \left(X_1^{(2)}; Y_3^{(2)} \mid U_1^{(2)} \right) + t_1 I \left(U_2^{(1)}; Y_3^{(1)} \right) \right) + \\
& \min \left(t_1 I \left(X_2^{(1)}; Y_3^{(1)} \mid U_2^{(1)} \right) + t_2 I \left(U_1^{(2)}; Y_3^{(2)} \right), t_2 I \left(X_0^{(2)}; Y_2^{(2)} \mid X_1^{(2)}, U_1^{(2)} \right) \right), \\
& t_1 I \left(X_0^{(1)}, X_2^{(1)}; Y_1^{(1)} \mid U_2^{(1)} \right), t_2 I \left(X_0^{(2)}, X_1^{(2)}; Y_2^{(2)} \mid U_1^{(2)} \right) \Big). \tag{15}
\end{aligned}$$

with probabilities

$$\begin{aligned}
p(x_0^{(1)}, x_2^{(1)}, u_2^{(1)}) &= p(u_2^{(1)})p(x_2^{(1)}|u_2^{(1)})p(x_0^{(1)}|x_2^{(1)}, u_2^{(1)}), \\
p(x_0^{(2)}, x_1^{(2)}, u_1^{(2)}) &= p(u_1^{(2)})p(x_1^{(2)}|u_1^{(2)})p(x_0^{(2)}|x_1^{(2)}, u_1^{(2)}), \\
p(x_2^{(1)}, u_2^{(1)}) &= p(u_2^{(1)})p(x_2^{(1)}|u_2^{(1)}), \\
p(x_1^{(2)}, u_1^{(2)}) &= p(u_1^{(2)})p(x_1^{(2)}|u_1^{(2)}).
\end{aligned}$$

$$\begin{aligned}
C_{BME_{back}}^{low} = R^{(1)} + R^{(2)} \leq & \max_{0 \leq t_1, t_2, t_1+t_2=1} \min \left(t_1 I \left(X_0^{(1)}, X_2^{(1)}; Y_1^{(1)} \right), t_2 I \left(X_0^{(2)}, X_1^{(2)}; Y_2^{(2)} \right), \right. \\
& t_1 I \left(X_0^{(1)}; Y_1^{(1)} \mid X_2^{(1)} \right) + t_2 I \left(X_0^{(2)}; Y_2^{(2)} \mid X_1^{(2)} \right), \\
& \left. t_1 I \left(X_2^{(1)}; Y_3^{(1)} \right) + t_2 I \left(X_1^{(2)}; Y_3^{(2)} \right) \right). \tag{16}
\end{aligned}$$

with probabilities

$$\begin{aligned}
p(x_0^{(1)}, x_2^{(1)}) &= p(x_2^{(1)})p(x_0^{(1)}|x_2^{(1)}), \\
p(x_0^{(2)}, x_1^{(2)}) &= p(x_1^{(2)})p(x_0^{(2)}|x_1^{(2)}).
\end{aligned}$$

Proof: See Appendix B. ■

Now, the following set of propositions and corollaries investigate the Non-Cooperative and Cooperative schemes and compare them with each other.

Proposition 1 *The BME with backward decoding achieves a better rate than the one with successive decoding,*

i.e., $C_{BME_{back}}^{low} \geq C_{BME_{succ}}^{low}$.

Proof: For the first term of minimization (15), we have

$$\begin{aligned}
& \min \left(t_1 I \left(X_0^{(1)}; Y_1^{(1)} \mid X_2^{(1)}, U_2^{(1)} \right), t_2 I \left(X_1^{(2)}; Y_3^{(2)} \mid U_1^{(2)} \right) + t_1 I \left(U_2^{(1)}; Y_3^{(1)} \right) \right) + \\
& \min \left(t_1 I \left(X_2^{(1)}; Y_3^{(1)} \mid U_2^{(1)} \right) + t_2 I \left(U_1^{(2)}; Y_3^{(2)} \right), t_2 I \left(X_0^{(2)}; Y_2^{(2)} \mid X_1^{(2)}, U_1^{(2)} \right) \right) \leq \\
& \min \left(t_1 I \left(X_0^{(1)}; Y_1^{(1)} \mid X_2^{(1)}, U_2^{(1)} \right) + t_2 I \left(X_0^{(2)}; Y_2^{(2)} \mid X_1^{(2)}, U_1^{(2)} \right), \right. \\
& \left. t_1 I \left(X_2^{(1)}, U_2^{(1)}; Y_3^{(1)} \right) + t_2 I \left(X_1^{(2)}, U_1^{(2)}; Y_3^{(2)} \right) \right). \tag{17}
\end{aligned}$$

Let us focus on $t_1 I(X_0^{(1)}; Y_1^{(1)} | X_2^{(1)}, U_2^{(1)}) + t_2 I(X_0^{(2)}; Y_2^{(2)} | X_1^{(2)}, U_1^{(2)})$:

$$\begin{aligned}
& t_1 I(X_0^{(1)}; Y_1^{(1)} | X_2^{(1)}, U_2^{(1)}) + t_2 I(X_0^{(2)}; Y_2^{(2)} | X_1^{(2)}, U_1^{(2)}) \stackrel{(a)}{=} \\
& t_1 H(Y_1^{(1)} | X_2^{(1)}, U_2^{(1)}) - t_1 H(Y_1^{(1)} | X_0^{(1)}, X_2^{(1)}) + \\
& t_2 H(Y_2^{(2)} | X_1^{(2)}, U_1^{(2)}) - t_2 H(Y_2^{(2)} | X_0^{(2)}, X_1^{(2)}) \stackrel{(b)}{\leq} \\
& t_1 H(Y_1^{(1)} | X_2^{(1)}) - t_1 H(Y_1^{(1)} | X_0^{(1)}, X_2^{(1)}) + \\
& t_2 H(Y_2^{(2)} | X_1^{(2)}) - t_2 H(Y_2^{(2)} | X_0^{(2)}, X_1^{(2)}) \stackrel{(c)}{=} \\
& t_1 I(X_0^{(1)}; Y_1^{(1)} | X_2^{(1)}) + t_2 I(X_0^{(2)}; Y_2^{(2)} | X_1^{(2)}). \tag{18}
\end{aligned}$$

(a) and (c) follow from the definition of mutual information, the fact that $U_2^{(1)} \rightarrow (X_0^{(1)}, X_2^{(1)}) \rightarrow Y_1^{(1)}$ and $U_1^{(2)} \rightarrow (X_0^{(2)}, X_1^{(2)}) \rightarrow Y_2^{(2)}$ form Markov chain, and (b) follows from the fact that conditioning reduces entropy. Inequality (b) becomes equality if $p(x_0^{(1)}, x_2^{(1)}, u_2^{(1)}) = p(u_2^{(1)})p(x_2^{(1)})p(x_0^{(1)}|x_2^{(1)})$ and $p(x_0^{(2)}, x_1^{(2)}, u_1^{(2)}) = p(u_1^{(2)})p(x_1^{(2)})p(x_0^{(2)}|x_1^{(2)})$. Using the similar argument for $t_1 I(X_2^{(1)}, U_2^{(1)}; Y_3^{(1)}) + t_2 I(X_1^{(2)}, U_1^{(2)}; Y_3^{(2)})$, $t_1 I(X_0^{(1)}, X_2^{(1)}; Y_1^{(1)} | U_2^{(1)})$, and $t_2 I(X_0^{(2)}, X_1^{(2)}; Y_2^{(2)} | U_1^{(2)})$ in (15) and (17), and the fact $U_2^{(1)} \rightarrow X_2^{(1)} \rightarrow Y_3^{(1)}$, $U_1^{(2)} \rightarrow X_1^{(2)} \rightarrow Y_3^{(2)}$, $U_2^{(1)} \rightarrow (X_0^{(1)}, X_2^{(1)}) \rightarrow Y_1^{(1)}$, $U_1^{(2)} \rightarrow (X_0^{(2)}, X_1^{(2)}) \rightarrow Y_2^{(2)}$ form Markov chain, and Appendix B, along with comparing $C_{BME_{succ}}^{low}$ and $C_{BME_{back}}^{low}$ in Theorem 2, we have $C_{BME_{back}}^{low} \geq C_{BME_{succ}}^{low}$. ■

From Theorem 2, we have the following corollary for the Gaussian case.

corollary 2 *For the half-duplex Gaussian parallel relay channel, assuming successive relaying protocol with power constraint at the source and each relay, BME achieves the following rates*

$$\begin{aligned}
C_{BME_{succ}}^{low} = & \max \min (C_{BME_1}^{low} + C_{BME_2}^{low}, \\
& t_1 C \left(\frac{h_{01}^2 P_0^{(1)} + h_{12}^2 \theta_2 P_2 + 2h_{01}h_{12}\sqrt{\bar{\alpha}_1 \theta_2 P_0^{(1)} P_2}}{t_1} \right), \\
& t_2 C \left(\frac{h_{02}^2 P_0^{(2)} + h_{12}^2 \theta_1 P_1 + 2h_{02}h_{12}\sqrt{\bar{\alpha}_2 \theta_1 P_0^{(2)} P_1}}{t_2} \right) \Bigg). \tag{19}
\end{aligned}$$

$$\begin{aligned}
C_{BME_{back}}^{low} = & \max \min \left(t_1 C \left(\frac{h_{01}^2 P_0^{(1)} + h_{12}^2 P_2 + 2h_{01}h_{12}\sqrt{\bar{\beta}_1 P_0^{(1)} P_2}}{t_1} \right), \right. \\
& t_2 C \left(\frac{h_{02}^2 P_0^{(2)} + h_{12}^2 P_1 + 2h_{02}h_{12}\sqrt{\bar{\beta}_2 P_0^{(2)} P_1}}{t_2} \right), \\
& t_1 C \left(\frac{h_{01}^2 \beta_1 P_0^{(1)}}{t_1} \right) + t_2 C \left(\frac{h_{02}^2 \beta_2 P_0^{(2)}}{t_2} \right), \\
& \left. t_1 C \left(\frac{h_{23}^2 P_2}{t_1} \right) + t_2 C \left(\frac{h_{13}^2 P_1}{t_2} \right) \right). \tag{20}
\end{aligned}$$

subject to:

$$C_{BME_1}^{low} = \min \left(t_1 C \left(\frac{h_{01}^2 \alpha_1 P_0^{(1)}}{t_1} \right), t_1 C \left(\frac{h_{23}^2 \bar{\theta}_2 P_2}{h_{23}^2 \theta_2 P_2 + t_1} \right) + t_2 C \left(\frac{h_{13}^2 \theta_1 P_1}{t_2} \right) \right), \quad (21)$$

$$C_{BME_2}^{low} = \min \left(t_2 C \left(\frac{h_{02}^2 \alpha_2 P_0^{(2)}}{t_2} \right), t_2 C \left(\frac{h_{13}^2 \bar{\theta}_1 P_1}{h_{13}^2 \theta_1 P_1 + t_2} \right) + t_1 C \left(\frac{h_{23}^2 \theta_2 P_2}{t_1} \right) \right), \quad (22)$$

$$P_0^{(1)} + P_0^{(2)} = P_0,$$

$$t_1 + t_2 = 1,$$

$$0 \leq \alpha_1, \alpha_2 \leq 1,$$

$$0 \leq \beta_1, \beta_2 \leq 1,$$

$$0 \leq \theta_1, \theta_2 \leq 1.$$

where $\bar{\theta}_i = 1 - \theta_i$, $\bar{\alpha}_i = 1 - \alpha_i$, and $\bar{\beta}_i = 1 - \beta_i$ for $i = 1, 2$.

Proof: Let $V_0^{(1)} \sim \mathcal{N}(0, \alpha_1 P_0^{(1)})$, $V_0^{(2)} \sim \mathcal{N}(0, \alpha_2 P_0^{(2)})$, $V_2^{(1)} \sim \mathcal{N}(0, \theta_2 P_2)$, $V_1^{(2)} \sim \mathcal{N}(0, \theta_1 P_1)$, $U_2^{(1)} \sim \mathcal{N}(0, \bar{\theta}_2 P_2)$ and $U_1^{(2)} \sim \mathcal{N}(0, \bar{\theta}_1 P_1)$, which are independent of each other.

Letting $X_0^{(1)} = V_0^{(1)} + \sqrt{\frac{\bar{\alpha}_1 P_0^{(1)}}{\theta_2 P_2}} V_2^{(1)}$, $X_0^{(2)} = V_0^{(2)} + \sqrt{\frac{\bar{\alpha}_2 P_0^{(2)}}{\theta_1 P_1}} V_1^{(2)}$, $X_2^{(1)} = V_2^{(1)} + U_2^{(1)}$, $X_1^{(2)} = V_1^{(2)} + U_1^{(2)}$ and using the result in the expression for the achievable rate obtained in Theorem 1, we obtain $C_{BME_{succ}}^{low}$ for the Gaussian case, as given in [32] and (19), (21), and (22), respectively.

For backward decoding, let $V_0^{(1)} \sim \mathcal{N}(0, \beta_1 P_0^{(1)})$, $V_0^{(2)} \sim \mathcal{N}(0, \beta_2 P_0^{(2)})$, $X_2^{(1)} \sim \mathcal{N}(0, P_2)$, and $X_1^{(2)} \sim \mathcal{N}(0, P_1)$, which are independent of each other. By setting $X_0^{(1)} = V_0^{(1)} + \sqrt{\frac{\bar{\beta}_1 P_0^{(1)}}{P_2}} X_2^{(1)}$, $X_0^{(2)} = V_0^{(2)} + \sqrt{\frac{\bar{\beta}_2 P_0^{(2)}}{P_1}} X_1^{(2)}$ and using the result in the expression for the achievable rate obtained in Theorem 1, we obtain $C_{BME_{back}}^{low}$ for the Gaussian case, as given in (20). ■

Proposition 2 In symmetric scenarios, where $h_{01} = h_{02}$, $h_{13} = h_{23}$, and $P_1 = P_2$, Non-Cooperative DPC scheme outperforms Cooperative BME scheme, i.e. $C_{BME_{back}}^{low} \leq C_{DPC}^{low}$.

Proof: Due to the symmetric assumption, we have $t_1 = t_2 = \frac{1}{2}$, $P_0^{(1)} = P_0^{(2)} = \frac{P_0}{2}$, and $\beta_1 = \beta_2 = \frac{1}{2}$. Hence, from (20), we have

$$C_{BME_{back}}^{low} \leq \min \left(C \left(\frac{h_{01}^2 P_0}{2} \right), C \left(2h_{13}^2 P_1 \right) \right). \quad (23)$$

And also C_{DPC}^{low} in (12) becomes

$$C_{DPC}^{low} = \min \left(C \left(h_{01}^2 P_0 \right), \frac{1}{2} C \left(h_{01}^2 P_0 \right) + \frac{1}{2} C \left(2h_{13}^2 P_1 \right), C \left(2h_{13}^2 P_1 \right) \right). \quad (24)$$

Comparing (23) and (24), we have $C_{BME_{back}}^{low} \leq C_{DPC}^{low}$. ■

According to the discussion in Appendix B, $r_{Bin}^{(1)} = 0$ or $r_{Bin}^{(2)} = 0$. In other words, in the Cooperative BME scheme based on backward decoding, at most one relay is necessary to use binning function for the message it

receives from another, and the other relay is not necessary to cooperate with this relay. Therefore, we propose a composite BME-DPC scheme. In this scheme, one of the relays decodes the other relay's message. Having decoded that, it then uses the binning function to cooperate with the other relay. On the other hand, using the Gelfand-Pinsker's result the source cancels the interference due to one relay on the other. Hence, we have the following Theorem.

Theorem 3 *The composite BME-DPC scheme, achieves the following rate:*

$$C_{BME-DPC}^{low} = \max_{0 \leq t_1, t_2, t_1+t_2=1} \min \left(t_1 I \left(X_0^{(1)}, X_2^{(1)}; Y_1^{(1)} \right), t_1 I \left(X_0^{(1)}; Y_1^{(1)} \mid X_2^{(1)} \right) + \right. \\ \left. t_2 \left(I \left(U_0^{(2)}; Y_2^{(2)} \right) - I \left(U_0^{(2)}; X_1^{(2)} \right) \right), t_1 I \left(X_2^{(1)}; Y_3^{(1)} \right) + t_2 I \left(X_1^{(2)}; Y_3^{(2)} \right), \right. \\ \left. t_2 \left(I \left(U_0^{(2)}; Y_2^{(2)} \right) - I \left(U_0^{(2)}; X_1^{(2)} \right) \right) + t_2 I \left(X_1^{(2)}; Y_3^{(2)} \right) \right). \quad (25)$$

Proof: Assuming $r_{Bin}^{(1)} = 0$, and using Theorem 1 and Theorem 2 along with a similar argument as in Appendix B, Theorem 3 is immediate. ■

corollary 3 *For the Gaussian case, the composite BME-DPC scheme achieves the following rate $C_{BME-DPC}^{low}$. Furthermore, $C_{BME-DPC}^{low} \geq C_{BMEback}^{low}$. In other words, the composite BME-DPC scheme always achieves a better rate than the BME scheme for the Gaussian scenario.*

$$C_{BME-DPC}^{low} = R^{(1)} + R^{(2)} \leq \\ \max \min \left(t_1 C \left(\frac{h_{01}^2 P_0^{(1)} + h_{12}^2 P_2 + 2h_{01}h_{12}\sqrt{\bar{\alpha}P_0^{(1)}P_2}}{t_1} \right), \right. \\ \left. t_1 C \left(\frac{h_{01}^2 \alpha P_0^{(1)}}{t_1} \right) + t_2 C \left(\frac{h_{02}^2 P_0^{(2)}}{t_2} \right), \right. \\ \left. t_1 C \left(\frac{h_{23}^2 P_2}{t_1} \right) + t_2 C \left(\frac{h_{13}^2 P_1}{t_2} \right), t_2 C \left(\frac{h_{02}^2 P_0^{(2)}}{t_2} \right) + t_2 C \left(\frac{h_{13}^2 P_1}{t_2} \right) \right). \quad (26)$$

subject to:

$$P_0^{(1)} + P_0^{(2)} = P_0,$$

$$t_1 + t_2 = 1,$$

$$0 \leq t_1, t_2, P_0^{(1)}, P_0^{(2)},$$

$$0 \leq \alpha \leq 1.$$

where $\bar{\alpha} = 1 - \alpha$.

Proof: As in Theorem 3, we assume that $r_{Bin}^{(1)} = 0$. Now, we show that every rate pairs $(R^{(1)}, R^{(2)})$ satisfying (101)-(107) satisfy (26). After specializing (101)-(107) for the Gaussian case and comparing with (26), one observes that the second term in minimization (101) does not exist. Substituting $r_{Bin}^{(1)} = 0$ in (102)-(107),

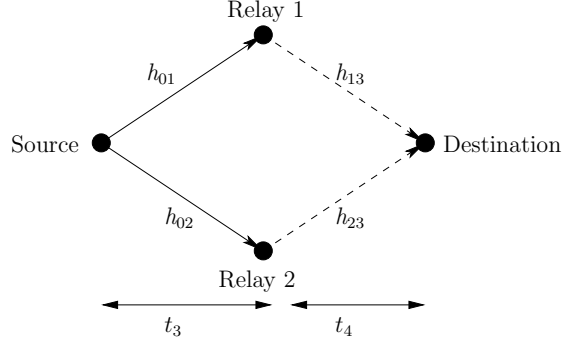


Fig. 6. Simultaneous relaying protocol for two relays.

one can obtain the other three corresponding terms. Comparing those terms with (26), it can be readily seen that $C_{BME-DPC}^{low} \geq C_{BME_{back}}^{low}$. ■

Remark 1 Assuming $r_{Bin}^{(1)} = 0$, as in Theorem 3 and corollary 3, the destination jointly decodes the current message and the bin index of the next message at the end of even time slots and then it can decode the next message at the end of odd time slots. Therefore, using backward decoding is not necessary in the BME-DPC scheme.

B. Simultaneous Relaying Protocol

Figure 6 shows simultaneous relaying protocol. In simultaneous relaying, in one time slot of duration t_3 the source transmits its signal simultaneously to the two relays. In the next time slot of duration t_4 , two relays transmit their signal coherently to the destination. Hence, in this protocol, $t_1 = t_2 = 0$ and our system model follows from (6) and (7).

1) *Dynamic Decode-and-Forward (DDF)*: In DDF scheme each relay decodes the transmitted message from the source in time slot t_3 (Broadcast (BC) State), and forwards its re-encoded version in time slot t_4 (Multiple Access (MAC) State). The following Theorem gives the achievable rate of the DDF scheme for the general discrete memoryless channels.

Theorem 4 For the half-duplex parallel relay channel, assuming simultaneous relaying and the fact that what the second relay receives is a degraded version of what the first relay receives, the following rate C_{DDF}^{low} is achievable:

$$C_{DDF}^{low} = \max_{0 \leq t_3, t_4, t_3+t_4=1} R_p + R_c, \quad (27)$$

subject to:

$$R_p \leq \min \left(t_3 I(X_0^{(3)}; Y_1^{(3)} | U_0^{(3)}), t_4 I(X_1^{(4)}; Y_3^{(4)} | X_2^{(4)}) \right), \quad (28)$$

$$R_c \leq t_3 I(U_0^{(3)}; Y_2^{(3)}), \quad (29)$$

$$R_p + R_c \leq t_4 I(X_1^{(4)}, X_2^{(4)}; Y_3^{(4)}). \quad (30)$$

with probabilities:

$$\begin{aligned} p(u_0^{(3)}, x_0^{(3)}) &= p(u_0^{(3)})p(x_0^{(3)}|u_0^{(3)}), \\ p(x_1^{(4)}, x_2^{(4)}) &= p(x_1^{(4)})p(x_2^{(4)}|x_1^{(4)}). \end{aligned}$$

Proof: The achievable rate of DDF is equal to $C_{DDF}^{low} = R_p + R_c$, where (R_p, R_c) should be both in the capacity region of BC (corresponding to the BC state) and MAC (corresponding to the MAC state). Applying the superposition coding of the degraded BC [12] the following rates are achievable for the first hop:

$$\begin{aligned} R_p &\leq t_3 I(X_0^{(3)}; Y_1^{(3)} | U_0^{(3)}), \\ R_c &\leq t_3 I(U_0^{(3)}; Y_2^{(3)}). \end{aligned} \tag{31}$$

with probability $p(u_0^{(3)}, x_0^{(3)}) = p(u_0^{(3)})p(x_0^{(3)}|u_0^{(3)})$.

And using the superposition coding of the extended MAC (See [25], [26]) the following rates are achievable for the second hop:

$$\begin{aligned} R_p &\leq t_4 I(X_1^{(4)}; Y_3^{(4)} | X_2^{(4)}), \\ R_p + R_c &\leq t_4 I(X_1^{(4)}, X_2^{(4)}; Y_3^{(4)}). \end{aligned} \tag{32}$$

with probability $p(x_1^{(4)}, x_2^{(4)}) = p(x_1^{(4)})p(x_2^{(4)}|x_1^{(4)})$. ■

In the Gaussian case (assuming $h_{01} \geq h_{02}$), the source splits its total available power P_0 to $P_{0,p}^{(3)}$ and $P_{0,c}^{(3)}$ associated with the “Private” and the “Common” messages, respectively. Letting $X_0^{(3)} \sim \mathcal{N}(0, P_0)$, $U_0^{(3)} \sim \mathcal{N}(0, P_{0,c}^{(3)})$, and $X_1^{(4)} \sim \mathcal{N}(0, P_1)$, assuming that relay 1 and relay 2 transmit their codewords associated with the common message with $\mathcal{N}(0, P_{1,c}^{(4)})$ and $\mathcal{N}(0, P_2)$, and using (31) and (32) we have the following corollary.

corollary 4 For the half-duplex Gaussian parallel relay channel, assuming simultaneous relaying protocol with power constraints at the source and at each relay, DDF achieves the following rate

$$\begin{aligned}
C_{DDF}^{low} &= R_p + R_c, \\
\text{subject to: } R_p &\leq \min \left(t_3 C \left(\frac{h_{01}^2 P_{0,p}^{(3)}}{t_3} \right), t_4 C \left(\frac{h_{13}^2 P_{1,p}^{(4)}}{t_4} \right) \right), \\
R_c &\leq t_3 C \left(\frac{h_{02}^2 P_{0,c}^{(3)}}{t_3 + h_{02}^2 P_{0,p}^{(3)}} \right), \\
R_p + R_c &\leq t_4 C \left(\frac{h_{13}^2 P_{1,p}^{(4)} + \left(h_{13} \sqrt{P_{1,c}^{(4)}} + h_{23} \sqrt{P_2} \right)^2}{t_4} \right), \\
P_{0,p}^{(3)} + P_{0,c}^{(3)} &= P_0, \quad P_{1,p}^{(4)} + P_{1,c}^{(4)} = P_1, \quad t_3 + t_4 = 1, \\
0 &\leq t_3, t_4, P_{0,p}^{(3)}, P_{0,c}^{(3)}, P_{1,p}^{(4)}, P_{1,c}^{(4)}.
\end{aligned} \tag{33}$$

Interestingly, successive decoding at the destination does not degrade the performance of DDF scheme in the Gaussian scenario as shown in the following Proposition.

Proposition 3 The rate of DDF scheme is achievable by successive decoding of the common and private messages at the destination.

Proof: Consider the sum rate for both the common message and the private message for the extended multiple access channel from relays to the destination,

$$R_p + R_c \leq t_4 C \left(\frac{h_{13}^2 P_{1,p}^{(4)} + (h_{13} \sqrt{P_{1,c}^{(4)}} + h_{23} \sqrt{P_2})^2}{t_4} \right). \tag{34}$$

It can be readily verified that subject to the constraint $P_{1,p}^{(4)} + P_{1,c}^{(4)} = P_1$, the right-hand side of (34) is a decreasing function of $P_{1,p}^{(4)}$ or equivalently an increasing function of $P_{1,c}^{(4)}$. Now, let us equate R_p in (34) with the private rate \dot{R}_p of another MAC which is achieved by successive decoding of common and private messages. Therefore, we have

$$R_p = \dot{R}_p = t_4 C \left(\frac{h_{13}^2 \dot{P}_{1,p}^{(4)}}{t_4} \right) \leq t_4 C \left(\frac{h_{13}^2 P_{1,p}^{(4)}}{t_4} \right). \tag{35}$$

According to (35), we have (See Fig. 7)

$$\begin{aligned}
\dot{P}_{1,p}^{(4)} &\leq P_{1,p}^{(4)} \implies \\
R_p + R_c &\leq \dot{R}_p + \dot{R}_c, \\
R_c &\leq \dot{R}_c.
\end{aligned}$$

Hence, (R_p, R_c) lies in the corner point of the extended MAC with parameters $(\dot{P}_{1,p}^{(4)}, \dot{P}_{1,c}^{(4)})$, i.e. successive decoding of common and private messages achieves the DF rate. ■

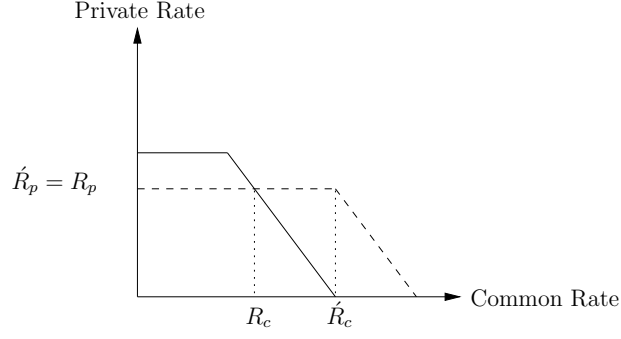


Fig. 7. The order of decoding “Common” and “Private” messages.

C. Simultaneous-Successive Relaying Protocol based on Dirty paper coding (SSRD)

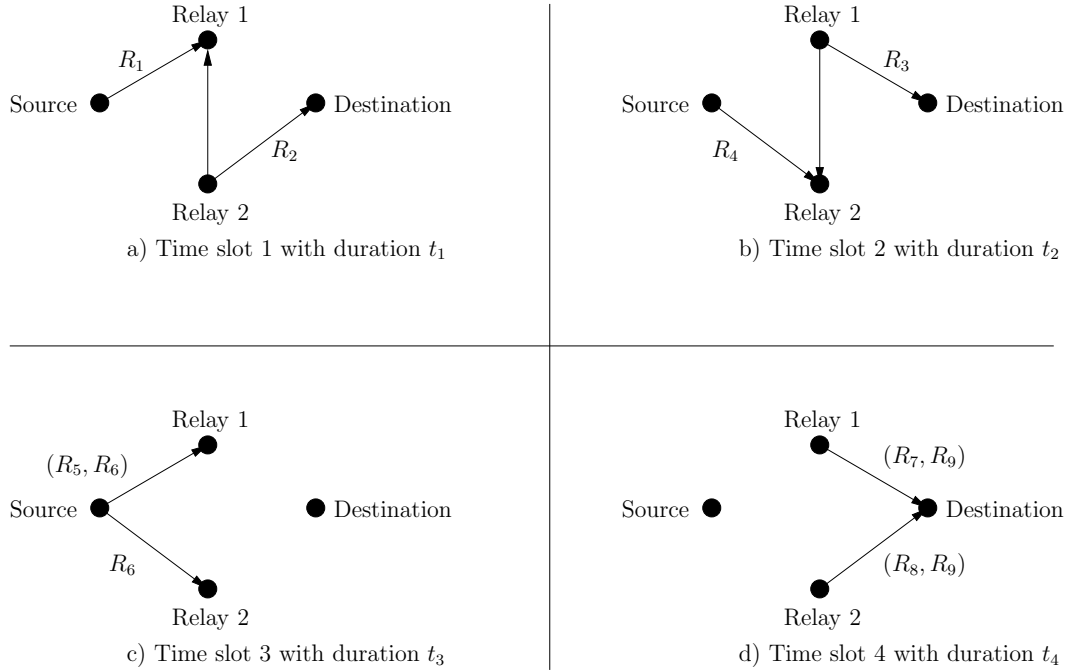


Fig. 8. SSRD Scheme for the Half-Duplex Parallel Relay Channel.

In this section, we propose an achievable rate for the half-duplex parallel relay channel. Our achievable scheme is based on the combination of the successive relaying protocol based on DPC scheme and simultaneous relaying protocol based on DDF (SSRD scheme). Hence, we have the following Theorem.

Theorem 5 Considering Fig. 8, for the half-duplex parallel relay channel, SSRD scheme achieves the following

rate C_{SSRD}^{low} :

$$C_{SSRD}^{low} = \min(R_1 + R_4 + R_5 + R_6, R_2 + R_3 + R_7 + R_8 + R_9), \quad (36)$$

subject to:

$$R_9 \leq R_6, R_1 + R_5 \leq R_3 + R_7, R_4 \leq R_2 + R_8. \quad (37)$$

Proof: SSRD scheme is illustrated in Fig. 8. As indicated in the figure, transmission is performed in 4 time slots. Relay 1 transmits its private message which was received in time slots t_1 and t_3 (corresponding to rates R_1 and R_5) in time slots t_2 and t_4 (corresponding to rates R_3 and R_7). On the other hand, relay 2 transmits its private message which has been received in time slot t_2 (corresponding to rate R_4) in time slots t_1 and t_4 (corresponding to rates R_2 and R_8). Furthermore, the two relays send the common message they have already received in time slot t_3 (corresponding to rate R_6) coherently in time slot t_4 (corresponding to rate R_9). As observed, here we consider the private rate for both relays in the MAC state, i.e. time slot t_4 . This is due to the reason that relay 2 also receives the private message in time slot t_2 . Hence, from the above description and Fig. 8, we have

$$C_{SSRD}^{low} = \min(R_1 + R_4 + R_5 + R_6, R_2 + R_3 + R_7 + R_8 + R_9), \quad (38)$$

subject to:

$$R_9 \leq R_6, R_1 + R_5 \leq R_3 + R_7, R_4 \leq R_2 + R_8. \quad (39)$$

■

Using corollaries 1, 4, and Proposition 3, for the Gaussian case we have

$$\begin{aligned}
C_{SSRD}^{low} = & \min \left(t_1 C \left(\frac{h_{01}^2 P_0^{(1)}}{t_1} \right) + t_2 C \left(\frac{h_{02}^2 P_0^{(2)}}{t_2} \right) + t_3 C \left(\frac{h_{01}^2 P_{0,p}^{(3)}}{t_3} \right) + t_3 C \left(\frac{h_{02}^2 P_{0,c}^{(3)}}{t_3 + h_{02}^2 P_{0,p}^{(3)}} \right), \right. \\
& t_1 C \left(\frac{h_{23}^2 P_2^{(1)}}{t_1} \right) + t_2 C \left(\frac{h_{13}^2 P_1^{(2)}}{t_2} \right) + \\
& \left. t_4 C \left(\frac{h_{13}^2 P_{1,p}^{(4)} + h_{23}^2 P_{2,p}^{(4)}}{t_4} \right) + t_4 C \left(\frac{\left(h_{13} \sqrt{P_{1,c}^{(4)}} + h_{23} \sqrt{P_{2,c}^{(4)}} \right)^2}{t_4 + h_{13}^2 P_{1,p}^{(4)} + h_{23}^2 P_{2,p}^{(4)}} \right) \right), \tag{40}
\end{aligned}$$

subject to:

$$\begin{aligned}
& t_4 C \left(\frac{\left(h_{13} \sqrt{P_{1,c}^{(4)}} + h_{23} \sqrt{P_{2,c}^{(4)}} \right)^2}{t_4 + h_{13}^2 P_{1,p}^{(4)} + h_{23}^2 P_{2,p}^{(4)}} \right) \leq t_3 C \left(\frac{h_{02}^2 P_{0,c}^{(3)}}{t_3 + h_{02}^2 P_{0,p}^{(3)}} \right), \\
& t_1 C \left(\frac{h_{01}^2 P_0^{(1)}}{t_1} \right) + t_3 C \left(\frac{h_{01}^2 P_{0,p}^{(3)}}{t_3} \right) \leq t_2 C \left(\frac{h_{13}^2 P_1^{(2)}}{t_2} \right) + t_4 C \left(\frac{h_{13}^2 P_{1,p}^{(4)}}{t_4} \right), \\
& t_2 C \left(\frac{h_{02}^2 P_0^{(2)}}{t_2} \right) \leq t_1 C \left(\frac{h_{23}^2 P_2^{(1)}}{t_1} \right) + t_4 C \left(\frac{h_{23}^2 P_{2,p}^{(4)}}{t_4} \right), \\
& P_0^{(1)} + P_0^{(2)} + P_{0,p}^{(3)} + P_{0,c}^{(3)} = P_0, \\
& P_1^{(2)} + P_{1,p}^{(4)} + P_{1,c}^{(4)} = P_1, \\
& P_2^{(1)} + P_{2,p}^{(4)} + P_{2,c}^{(4)} = P_2, \\
& t_1 + t_2 + t_3 + t_4 = 1, \\
& 0 \leq t_1, t_2, t_3, t_4, P_0^{(1)}, P_0^{(2)}, P_{0,p}^{(3)}, P_{0,c}^{(3)}, P_1^{(2)}, P_{1,p}^{(4)}, P_{1,c}^{(4)}, P_2^{(1)}, P_{2,p}^{(4)}, P_{2,c}^{(4)}.
\end{aligned}$$

According to corollary 3, another combined simultaneous-successive relaying protocol based on BME is not necessary. However, a “Simultaneous-Successive Relaying protocol based on BME-DPC”, can be easily derived. Assuming the first relay decodes the second one’s message, the achievable rate of this new scheme would be the same as C_{SSRD}^{low} . However, since the messages for the second relay are common, R_8 in the expression of the achievable rate is zero. Furthermore, the following constraints instead of (39) should be satisfied:

$$R_9 \leq R_4 + R_6, R_1 + R_5 \leq R_3 + R_7, R_1 + R_4 \leq t_1 I \left(X_0^{(1)}, X_2^{(1)}; Y_1^{(1)} \right). \tag{41}$$

IV. OPTIMALITY RESULTS

In this section, an upper bound for the half-duplex parallel relay channel is derived and investigated. The authors in [27] proposed some upper bounds on the achievable rate for general half-duplex multi-terminal networks. Here, we explain their results briefly and apply them to our half-duplex parallel relay network.

Authors in [27] define the concept of *state* for a half-duplex network with N nodes. The state of the network is a *valid partitioning of its nodes into two sets of the “sender nodes” and the “receiver nodes” such that there is no*

active link that arrives at a sender node, and \hat{t}_m is the portion of the time that network is used in state m where $m \in \{1, 2, \dots, M\}$. The following Theorem for the upper bound of the information flow from the subset S_1 to the subset S_2 of the nodes, where S_1 and S_2 are disjoint is proved in [27].

Theorem 6 For a general half-duplex network with N nodes and a finite number of states, M , the maximum achievable information rates $\{R^{ij}\}$ from a node set S_1 to a disjoint node set S_2 , $S_1, S_2 \subset \{0, 1, \dots, N-1\}$, is bounded by

$$\sum_{i \in S_1, j \in S_2} R^{ij} \leq \sup_{p(x_0^{(m)}, x_2^{(m)}, \dots, x_{N-1}^{(m)}), \hat{t}_m} \min_S \sum_{m=1}^M \hat{t}_m I(X_S^{(m)}; Y_S^{(m)} | X_{S^c}^{(m)}). \quad (42)$$

for some joint probability distribution $p(x_0^{(m)}, x_2^{(m)}, \dots, x_{N-1}^{(m)})$ when the minimization is over all the sets $S \subset \{0, 1, \dots, N-1\}$ subject to $S \cap S_1 = S_1$, $S \cap S_2 = \emptyset$ and the supremum is over all the non-negative \hat{t}_m subject to $\sum_{m=1}^M \hat{t}_m = 1$. Here, $x_S^{(m)}$, $y_S^{(m)}$, and $x_{S^c}^{(m)}$ denote the signals transmitted and received by nodes in set S , and transmitted by nodes in set S^c , during state m , respectively.

From Theorem 6, the maximum achievable rate C^{low} is upper bounded as

$$\begin{aligned} C^{low} \leq C^{up} \triangleq & \min \left(\hat{t}_1 I(X_0^{(1)}; Y_1^{(1)} | X_2^{(1)}) + \hat{t}_2 I(X_0^{(2)}; Y_2^{(2)} | X_1^{(2)}) + \hat{t}_3 I(X_0^{(3)}; Y_1^{(3)}, Y_2^{(3)}), \right. \\ & \hat{t}_2 I(X_0^{(2)}, X_1^{(2)}; Y_2^{(2)}, Y_3^{(2)}) + \hat{t}_3 I(X_0^{(3)}; Y_2^{(3)}) + \hat{t}_4 I(X_1^{(4)}; Y_3^{(4)} | X_2^{(4)}), \\ & \hat{t}_1 I(X_0^{(1)}, X_2^{(1)}; Y_1^{(1)}, Y_3^{(1)}) + \hat{t}_3 I(X_0^{(3)}; Y_1^{(3)}) + \hat{t}_4 I(X_2^{(4)}; Y_3^{(4)} | X_1^{(4)}), \\ & \left. \hat{t}_1 I(X_2^{(1)}; Y_3^{(1)}) + \hat{t}_2 I(X_1^{(2)}; Y_3^{(2)}) + \hat{t}_4 I(X_1^{(4)}, X_2^{(4)}; Y_3^{(4)}) \right), \end{aligned} \quad (43)$$

subject to

$$\hat{t}_1 + \hat{t}_2 + \hat{t}_3 + \hat{t}_4 = 1.$$

By setting $\hat{t}_3 = \hat{t}_4 = 0$ in (43), we obtain an upper bound on the successive relaying protocol which we call it *successive cut-set bound* in the sequel.

Theorem 7 In a degraded half-duplex parallel relay channel where the destination receives a degraded version of the received signals at relays, i.e. $X_2^{(1)} \longrightarrow Y_1^{(1)} \longrightarrow Y_3^{(1)}$ and $X_1^{(2)} \longrightarrow Y_2^{(2)} \longrightarrow Y_3^{(2)}$, BME based on backward decoding achieves the successive cut-set bound.

Proof: Setting $\hat{t}_3 = \hat{t}_4 = 0$ in (43) and comparing the result with (16) the Theorem is proved. ■

In high SNR scenarios, we have the following Theorem.

Theorem 8 In high SNR scenarios, assuming non-zero source-relay and relay-destination links, when power available for the source and each relay tends to infinity, time slots \hat{t}_3 and \hat{t}_4 in (43) tend to zero as $O\left(\frac{1}{\log P_0}\right)$. Furthermore, the upper bound on the capacity of the half-duplex parallel relay channel in high SNR scenarios is

$$C^{up} = C_{DPC}^{low} + O\left(\frac{1}{\log P_0}\right).$$

In other words, DPC achieves the capacity of a half-duplex Gaussian parallel relay channel as SNR goes to infinity.

Proof: Throughout the proof, we assume the power of the relays goes to infinity as $P_1 = \gamma_1 P_0$, $P_2 = \gamma_2 P_0$ where γ_1, γ_2 are constants independent of the SNR. Substituting $X_0^{(1)} \sim \mathcal{N}(0, \hat{P}_0^{(1)})$, $X_0^{(2)} \sim \mathcal{N}(0, \hat{P}_0^{(2)})$, $X_0^{(3)} \sim \mathcal{N}(0, \hat{P}_0^{(3)})$, $X_1^{(2)} \sim \mathcal{N}(0, \hat{P}_1^{(2)})$, $X_1^{(4)} \sim \mathcal{N}(0, \hat{P}_1^{(4)})$, $X_2^{(1)} \sim \mathcal{N}(0, \hat{P}_2^{(1)})$, and $X_2^{(4)} \sim \mathcal{N}(0, \hat{P}_2^{(4)})$ in (43), and assuming complete cooperation between the transmitting and receiving nodes for each cut in (43), we have

$$\begin{aligned}
C^{up} \leq & \min \left(\hat{t}_1 C \left(\frac{h_{01}^2 \hat{P}_0^{(1)}}{\hat{t}_1} \right) + \hat{t}_2 C \left(\frac{h_{02}^2 \hat{P}_0^{(2)}}{\hat{t}_2} \right) + \hat{t}_3 C \left(\frac{(h_{01}^2 + h_{02}^2) \hat{P}_0^{(3)}}{\hat{t}_3} \right), \right. \\
& \hat{t}_2 C \left(\frac{h_{02}^2 \hat{P}_0^{(2)}}{\hat{t}_2} + \frac{(h_{12}^2 + h_{13}^2) \hat{P}_1^{(2)}}{\hat{t}_2} + \frac{2h_{02}h_{12}\sqrt{\hat{P}_0^{(2)}\hat{P}_1^{(2)}}}{\hat{t}_2} + \frac{h_{02}^2 h_{13}^2 \hat{P}_0^{(2)} \hat{P}_1^{(2)}}{\hat{t}_2^2} \right) + \\
& \hat{t}_3 C \left(\frac{h_{02}^2 \hat{P}_0^{(3)}}{\hat{t}_3} \right) + \hat{t}_4 C \left(\frac{h_{13}^2 \hat{P}_1^{(4)}}{\hat{t}_4} \right), \\
& \hat{t}_1 C \left(\frac{h_{01}^2 \hat{P}_0^{(1)}}{\hat{t}_1} + \frac{(h_{12}^2 + h_{23}^2) \hat{P}_2^{(1)}}{\hat{t}_1} + \frac{2h_{01}h_{12}\sqrt{\hat{P}_0^{(1)}\hat{P}_2^{(1)}}}{\hat{t}_1} + \frac{h_{01}^2 h_{23}^2 \hat{P}_0^{(1)} \hat{P}_2^{(1)}}{\hat{t}_1^2} \right) + \\
& \hat{t}_3 C \left(\frac{h_{01}^2 \hat{P}_0^{(3)}}{\hat{t}_3} \right) + \hat{t}_4 C \left(\frac{h_{23}^2 \hat{P}_2^{(4)}}{\hat{t}_4} \right), \\
& \hat{t}_1 C \left(\frac{h_{23}^2 \hat{P}_2^{(1)}}{\hat{t}_1} \right) + \hat{t}_2 C \left(\frac{h_{13}^2 \hat{P}_1^{(2)}}{\hat{t}_2} \right) + \\
& \left. \hat{t}_4 C \left(\frac{h_{13}^2 \hat{P}_1^{(4)} + h_{23}^2 \hat{P}_2^{(4)} + 2h_{13}h_{23}\sqrt{\hat{P}_1^{(4)}\hat{P}_2^{(4)}}}{\hat{t}_4} \right) \right). \tag{44}
\end{aligned}$$

subject to:

$$\begin{aligned}
& \hat{P}_0^{(1)} + \hat{P}_0^{(2)} + \hat{P}_0^{(3)} = P_0, \\
& \hat{P}_1^{(2)} + \hat{P}_1^{(4)} = P_1, \\
& \hat{P}_2^{(1)} + \hat{P}_2^{(4)} = P_2, \\
& \hat{t}_1 + \hat{t}_2 + \hat{t}_3 + \hat{t}_4 = 1, \\
& 0 \leq \hat{t}_1, \hat{t}_2, \hat{t}_3, \hat{t}_4, \hat{P}_0^{(1)}, \hat{P}_0^{(2)}, \hat{P}_0^{(3)}, \hat{P}_1^{(2)}, \hat{P}_1^{(4)}, \hat{P}_2^{(1)}, \hat{P}_2^{(4)}.
\end{aligned}$$

Furthermore, from corollary 1, the achievable rate of the DPC scheme can be expressed as

$$\begin{aligned}
C_{DPC}^{low} = & \min \left(t_1 C \left(\frac{h_{01}^2 P_0^{(1)}}{t_1} \right) + t_2 C \left(\frac{h_{02}^2 P_0^{(2)}}{t_2} \right), \right. \\
& t_2 C \left(\frac{h_{02}^2 P_0^{(2)}}{t_2} \right) + t_2 C \left(\frac{h_{13}^2 P_1}{t_2} \right), \\
& t_1 C \left(\frac{h_{01}^2 P_0^{(1)}}{t_1} \right) + t_1 C \left(\frac{h_{23}^2 P_2}{t_1} \right), \\
& \left. t_1 C \left(\frac{h_{23}^2 P_2}{t_1} \right) + t_2 C \left(\frac{h_{13}^2 P_1}{t_2} \right) \right). \tag{45}
\end{aligned}$$

By setting $P_0^{(1)} = P_0^{(2)} = \frac{P_0}{2}$ and $t_1 = t_2 = 0.5$ in (45), expression (45) can be simplified as

$$C_{DPC}^{low} \geq \frac{1}{2} \ln P_0 + c. \quad (46)$$

where c is some constant which depends on channel coefficients. Knowing that the term corresponding to each cut-set in (44) for the optimum values of $\hat{t}_1, \dots, \hat{t}_4$ is indeed an upper-bound for C_{DPC}^{low} , and by setting $\hat{P}_0^{(1)} = \hat{P}_0^{(2)} = \hat{P}_0^{(3)} = P_0$ in (44), we have the following inequality between (46) and the first cut of (44).

$$\begin{aligned} \frac{1}{2} \ln P_0 + c &\leq \frac{\hat{t}_1}{2} \ln \left(\frac{h_{01}^2 P_0}{\hat{t}_1} \right) + \frac{\hat{t}_2}{2} \ln \left(\frac{h_{02}^2 P_0}{\hat{t}_2} \right) + \frac{\hat{t}_3}{2} \ln \left(\frac{(h_{01}^2 + h_{02}^2) P_0}{\hat{t}_3} \right) + \\ &\quad \frac{\hat{t}_1^2}{2h_{01}^2 P_0} + \frac{\hat{t}_2^2}{2h_{02}^2 P_0} + \frac{\hat{t}_3^2}{2(h_{01}^2 + h_{02}^2) P_0} \\ &= \frac{(1 - \hat{t}_4)}{2} \ln P_0 + \frac{\hat{t}_1}{2} \ln h_{01}^2 + \frac{\hat{t}_2}{2} \ln h_{02}^2 + \frac{\hat{t}_3}{2} \ln (h_{01}^2 + h_{02}^2) \\ &\quad - \frac{\hat{t}_1}{2} \ln \hat{t}_1 - \frac{\hat{t}_2}{2} \ln \hat{t}_2 - \frac{\hat{t}_3}{2} \ln \hat{t}_3 + \frac{\hat{t}_1^2}{2h_{01}^2 P_0} + \frac{\hat{t}_2^2}{2h_{02}^2 P_0} + \frac{\hat{t}_3^2}{2(h_{01}^2 + h_{02}^2) P_0}. \end{aligned} \quad (47)$$

Note that in deriving (46) and (47), the following inequality is applied to lower/upper-bound the corresponding terms:

$$\ln(x) \leq \ln(1+x) \leq \ln(x) + \frac{1}{x}, \forall x > 0. \quad (48)$$

Consequently, we have

$$\begin{aligned} \hat{t}_4 &\leq \frac{1}{\ln P_0} (2c + \hat{t}_1 \ln h_{01}^2 + \hat{t}_2 \ln h_{02}^2 + \hat{t}_3 \ln (h_{01}^2 + h_{02}^2) - \hat{t}_1 \ln \hat{t}_1 - \hat{t}_2 \ln \hat{t}_2 - \hat{t}_3 \ln \hat{t}_3) \\ &\quad + \frac{1}{\ln P_0} \left(\frac{\hat{t}_1^2}{h_{01}^2 P_0} + \frac{\hat{t}_2^2}{h_{02}^2 P_0} + \frac{\hat{t}_3^2}{(h_{01}^2 + h_{02}^2) P_0} \right). \end{aligned}$$

Hence, we can bound the optimum value of \hat{t}_4 in (44) as

$$0 \leq \hat{t}_4 \leq O \left(\frac{1}{\log P_0} \right). \quad (49)$$

Similarly, by considering the fourth cut in (44), we can derive another bound on the optimum value of \hat{t}_3 as follows:

$$0 \leq \hat{t}_3 \leq O \left(\frac{1}{\log P_0} \right). \quad (50)$$

Applying the inequality between (46) and the term corresponding to the second cut in (44), knowing (from (49) and (50)) the fact that $\hat{t}_3 \leq \frac{c_3}{\ln P_0}$, and $\hat{t}_4 \leq \frac{c_4}{\ln P_0}$ (where c_3 and c_4 are constants), and using inequalities (48), and

$$\ln(1+x) \leq x, \forall x \geq 0, \quad (51)$$

we obtain

$$\begin{aligned}
& \frac{1}{2} \ln P_0 + c \leq \\
& \frac{\hat{t}_2}{2} \ln \left(\frac{h_{02}^2 h_{13}^2 \gamma_1 P_0^2}{\hat{t}_2^2} \left(1 + \frac{\hat{t}_2}{\gamma_1 h_{13}^2 P_0} + \frac{\hat{t}_2 (h_{12}^2 + h_{13}^2)}{h_{02}^2 h_{13}^2 P_0} + \frac{\hat{t}_2 h_{12}}{h_{13}^2 h_{02} \sqrt{\gamma_1} P_0} \right) \right) + \\
& \frac{\hat{t}_3}{2} \ln \left(\frac{h_{02}^2 P_0}{\hat{t}_3} \right) + \frac{\hat{t}_4}{2} \ln \left(\frac{h_{13}^2 \gamma_1 P_0}{\hat{t}_4} \right) + \\
& \frac{\hat{t}_2^3}{2 (\hat{t}_2 h_{02}^2 P_0 + \hat{t}_2 \gamma_1 (h_{12}^2 + h_{13}^2) P_0 + 2 \hat{t}_2 h_{02} h_{12} \sqrt{\gamma_1} P_0 + h_{02}^2 h_{13}^2 \gamma_1 P_0^2)} + \\
& \frac{\hat{t}_3^2}{2 h_{02}^2 P_0} + \frac{\hat{t}_4^2}{2 \gamma_1 h_{13}^2 P_0} \\
& \leq \hat{t}_2 \ln P_0 + \frac{\hat{t}_2}{2} \ln \left(\frac{h_{02}^2 h_{13}^2 \gamma_1}{\hat{t}_2^2} \right) + \frac{\hat{t}_2^2}{2 \gamma_1 h_{13}^2 P_0} + \frac{\hat{t}_2^2 (h_{12}^2 + h_{13}^2)}{2 h_{02}^2 h_{13}^2 P_0} + \frac{\hat{t}_2^2 h_{12}}{2 h_{13}^2 h_{02} \sqrt{\gamma_1} P_0} + \\
& \frac{c_3}{2 \ln P_0} \ln h_{02}^2 - \frac{c_3}{2 \ln P_0} \ln \hat{t}_3 + \frac{c_3}{2} + \frac{c_4}{2 \ln P_0} \ln \gamma_1 h_{13}^2 - \frac{c_4}{2 \ln P_0} \ln \hat{t}_4 + \frac{c_4}{2} + \\
& \frac{\hat{t}_2^3}{2 (\hat{t}_2 h_{02}^2 P_0 + \hat{t}_2 \gamma_1 (h_{12}^2 + h_{13}^2) P_0 + 2 \hat{t}_2 h_{02} h_{12} \sqrt{\gamma_1} P_0 + h_{02}^2 h_{13}^2 \gamma_1 P_0^2)} + \\
& \frac{\hat{t}_3^2}{2 h_{02}^2 P_0} + \frac{\hat{t}_4^2}{2 \gamma_1 h_{13}^2 P_0}
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
& \frac{1}{2} \ln P_0 + c \leq \hat{t}_2 \ln P_0 + \epsilon \\
& + O\left(\frac{1}{\ln P_0}\right) + O\left(\frac{1}{P_0}\right).
\end{aligned}$$

Hence,

$$\frac{1}{2} - \frac{c_2}{\log P_0} \leq \hat{t}_2. \quad (52)$$

Similarly, from the third cut of (44), for \hat{t}_1 we have

$$\frac{1}{2} - \frac{c_1}{\log P_0} \leq \hat{t}_1. \quad (53)$$

From (52) and (53), and also the fact that $\hat{t}_1 + \hat{t}_2 + \hat{t}_3 + \hat{t}_4 = 1$, we obtain

$$\frac{1}{2} - \frac{c_2}{\log P_0} \leq \hat{t}_2 \leq \frac{1}{2} + \frac{c_1}{\log P_0}, \quad (54)$$

$$\frac{1}{2} - \frac{c_1}{\log P_0} \leq \hat{t}_1 \leq \frac{1}{2} + \frac{c_2}{\log P_0}. \quad (55)$$

Hence, from (49), (50), (54), and (55) as $P_0 \rightarrow \infty$, $\hat{t}_3, \hat{t}_4 \rightarrow 0$ and $\hat{t}_1, \hat{t}_2 \rightarrow 0.5$. This proves the first part of the Theorem.

Moreover, knowing that each term corresponding to the four cuts in (44) is greater than $0.5 \ln(P_0) + c$ and as \hat{t}_1, \hat{t}_2 are strictly above zero (approaching 0.5), we can easily conclude that

$$\hat{P}_0^{(1)}, \hat{P}_0^{(2)}, \hat{P}_1^{(2)}, \hat{P}_2^{(1)} \sim \Theta(P_0). \quad (56)$$

Now, we prove that the DPC scheme with the parameters $t_1 = \hat{t}_1 + \frac{\hat{t}_3 + \hat{t}_4}{2}$, $t_2 = \hat{t}_2 + \frac{\hat{t}_3 + \hat{t}_4}{2}$, $P_0^{(1)} = \hat{P}_0^{(1)}$ and $P_0^{(2)} = \hat{P}_0^{(2)}$, where $\hat{t}_1, \dots, \hat{t}_4, \hat{P}_0^{(1)}, \hat{P}_0^{(2)}$ are the parameters corresponding to the maximum value of (44), achieves the capacity with a gap no more than $O\left(\frac{1}{\log P_0}\right)$. To prove this, we show that each of the four terms in (45) is no more than $O\left(\frac{1}{\log P_0}\right)$ below the corresponding term (from the same cut) in (44). To show this, for the first cut we have

$$\begin{aligned}
& \hat{t}_1 C\left(\frac{h_{01}^2 \hat{P}_0^{(1)}}{\hat{t}_1}\right) + \hat{t}_2 C\left(\frac{h_{02}^2 \hat{P}_0^{(2)}}{\hat{t}_2}\right) + \hat{t}_3 C\left(\frac{(h_{01}^2 + h_{02}^2) \hat{P}_0^{(3)}}{\hat{t}_3}\right) - t_1 C\left(\frac{h_{01}^2 P_0^{(1)}}{t_1}\right) - t_2 C\left(\frac{h_{02}^2 P_0^{(2)}}{t_2}\right) \stackrel{(a)}{\leq} \\
& \frac{\hat{t}_1}{2} \ln\left(\frac{h_{01}^2 \hat{P}_0^{(1)}}{\hat{t}_1}\right) + \frac{\hat{t}_2}{2} \ln\left(\frac{h_{02}^2 \hat{P}_0^{(2)}}{\hat{t}_2}\right) + \hat{t}_3 C\left(\frac{(h_{01}^2 + h_{02}^2) \hat{P}_0^{(3)}}{\hat{t}_3}\right) - \left(\frac{\hat{t}_1}{2} + \frac{\hat{t}_3 + \hat{t}_4}{4}\right) \ln\left(\frac{h_{01}^2 \hat{P}_0^{(1)}}{t_1}\right) \\
& - \left(\frac{\hat{t}_2}{2} + \frac{\hat{t}_3 + \hat{t}_4}{4}\right) \ln\left(\frac{h_{02}^2 \hat{P}_0^{(2)}}{t_2}\right) + \frac{\hat{t}_1^2}{2h_{01}^2 \hat{P}_0^{(1)}} + \frac{\hat{t}_2^2}{2h_{02}^2 \hat{P}_0^{(2)}} \stackrel{(b)}{\lesssim} \\
& \frac{\hat{t}_1}{2} \ln\left(\frac{h_{01}^2 \hat{P}_0^{(1)}}{\hat{t}_1}\right) + \frac{\hat{t}_2}{2} \ln\left(\frac{h_{02}^2 \hat{P}_0^{(2)}}{\hat{t}_2}\right) + \frac{\hat{t}_3}{2} \ln\left(\frac{(h_{01}^2 + h_{02}^2) P_0}{\hat{t}_3 + \hat{t}_1}\right) - \left(\frac{\hat{t}_1}{2} + \frac{\hat{t}_3 + \hat{t}_4}{4}\right) \ln\left(\frac{h_{01}^2 \hat{P}_0^{(1)}}{t_1}\right) \\
& - \left(\frac{\hat{t}_2}{2} + \frac{\hat{t}_3 + \hat{t}_4}{4}\right) \ln\left(\frac{h_{02}^2 \hat{P}_0^{(2)}}{t_2}\right) + O\left(\frac{1}{\log P_0}\right) \stackrel{(c)}{\lesssim} \\
& \frac{\hat{t}_3}{2} \ln\left(\frac{P_0}{\sqrt{\hat{P}_0^{(1)} \hat{P}_0^{(2)}}}\right) - \frac{\hat{t}_4}{4} \ln(\hat{P}_0^{(1)} \hat{P}_0^{(2)}) + O\left(\frac{1}{\log P_0}\right) \stackrel{(d)}{\lesssim} O\left(\frac{1}{\log P_0}\right). \tag{57}
\end{aligned}$$

Here, (a) follows from (48), noting the function $\hat{t}_1 \ln(P_0 - x - y) + \hat{t}_2 \ln(y) + \hat{t}_3 \ln(\hat{t}_3 + (h_{01}^2 + h_{02}^2)x)$ takes its maximum value at $x \leq \frac{\hat{t}_3}{\hat{t}_3 + \hat{t}_1} P_0$ and hence substituting $\hat{P}_0^{(3)} = \frac{\hat{t}_3}{\hat{t}_3 + \hat{t}_1} P_0$ and finally noting $\hat{P}_0^{(1)}, \hat{P}_0^{(2)} \sim \Theta(P_0)$ result in (b), (c) follows from $\hat{t}_3, \hat{t}_4 \sim O\left(\frac{1}{\log P_0}\right)$ and $\ln\left(\frac{\hat{t}_1}{\hat{t}_1}\right) \sim O\left(\frac{1}{\log P_0}\right)$, and finally (d) follows from $\hat{P}_0^{(1)}, \hat{P}_0^{(2)} \sim \Theta(P_0)$.

Next, we bound the difference between the terms in the fourth cut of (44) and the fourth term in C_{DPC}^{low}

$$\begin{aligned}
& \hat{t}_1 C \left(\frac{h_{23}^2 \hat{P}_2^{(1)}}{\hat{t}_1} \right) + \hat{t}_2 C \left(\frac{h_{13}^2 \hat{P}_1^{(2)}}{\hat{t}_2} \right) + \hat{t}_4 C \left(\frac{h_{13}^2 \hat{P}_1^{(4)} + h_{23}^2 \hat{P}_2^{(4)} + 2h_{13}h_{23}\sqrt{\hat{P}_1^{(4)}\hat{P}_2^{(4)}}}{\hat{t}_4} \right) \\
& - t_1 C \left(\frac{h_{23}^2 P_2}{t_1} \right) - t_2 C \left(\frac{h_{13}^2 P_1}{t_2} \right) \stackrel{(a)}{\lesssim} \\
& \frac{\hat{t}_1}{2} \ln \left(\frac{h_{23}^2 \hat{P}_2^{(1)}}{\hat{t}_1} \right) + \frac{\hat{t}_2}{2} \ln \left(\frac{h_{13}^2 \hat{P}_1^{(2)}}{\hat{t}_2} \right) + \hat{t}_4 C \left(\frac{h_{13}^2 \hat{P}_1^{(4)} + h_{23}^2 \hat{P}_2^{(4)} + 2h_{13}h_{23}\sqrt{\hat{P}_1^{(4)}\hat{P}_2^{(4)}}}{\hat{t}_4} \right) \\
& - \left(\frac{\hat{t}_1}{2} + \frac{\hat{t}_3 + \hat{t}_4}{4} \right) \ln \left(\frac{h_{23}^2 P_2}{t_1} \right) - \left(\frac{\hat{t}_2}{2} + \frac{\hat{t}_3 + \hat{t}_4}{4} \right) \ln \left(\frac{h_{13}^2 P_1}{t_2} \right) + O \left(\frac{1}{P_0} \right) \stackrel{(b)}{\lesssim} \\
& \frac{\hat{t}_1}{2} \ln \left(\frac{h_{23}^2 P_2}{\hat{t}_1} \right) + \frac{\hat{t}_2}{2} \ln \left(\frac{h_{13}^2 P_1}{\hat{t}_2} \right) + \hat{t}_4 \ln \left(h_{13} \sqrt{\frac{P_1}{\hat{t}_2 + \hat{t}_4}} + h_{23} \sqrt{\frac{P_2}{\hat{t}_1 + \hat{t}_4}} \right) \\
& - \left(\frac{\hat{t}_1}{2} + \frac{\hat{t}_3 + \hat{t}_4}{4} \right) \ln \left(\frac{h_{23}^2 P_2}{t_1} \right) - \left(\frac{\hat{t}_2}{2} + \frac{\hat{t}_3 + \hat{t}_4}{4} \right) \ln \left(\frac{h_{13}^2 P_1}{t_2} \right) + O \left(\frac{1}{P_0} \right) \stackrel{(c)}{\lesssim} \\
& \frac{\hat{t}_4}{2} \ln \left(\frac{2}{\sqrt{(\hat{t}_1 + \hat{t}_4)(\hat{t}_2 + \hat{t}_4)}} + \frac{h_{13}}{h_{23}(\hat{t}_2 + \hat{t}_4)} \sqrt{\frac{P_1}{P_2}} + \frac{h_{23}}{(\hat{t}_1 + \hat{t}_4)h_{13}} \sqrt{\frac{P_2}{P_1}} \right) - \frac{\hat{t}_3}{4} \ln(P_1 P_2) + O \left(\frac{1}{\log P_0} \right) \stackrel{(d)}{\lesssim} \\
& O \left(\frac{1}{\log P_0} \right). \tag{58}
\end{aligned}$$

Here, (a) follows from (48) and noting $\hat{P}_1^{(2)}, \hat{P}_2^{(1)} \sim \Theta(P_0)$, noting the function $\hat{t}_1 \ln(P_2 - y) + \hat{t}_2 \ln(P_1 - x) + \hat{t}_4 \ln \left(\hat{t}_4 + (h_{13}\sqrt{x} + h_{23}\sqrt{y})^2 \right)$ takes its maximum value at $x \leq \frac{\hat{t}_4}{\hat{t}_4 + \hat{t}_2} P_1, y \leq \frac{\hat{t}_4}{\hat{t}_4 + \hat{t}_1} P_2$ and hence substituting $\hat{P}_1^{(4)} = \frac{\hat{t}_4}{\hat{t}_4 + \hat{t}_2} P_1$ and $\hat{P}_2^{(4)} = \frac{\hat{t}_4}{\hat{t}_4 + \hat{t}_1} P_2$ result in (b), (c) follows from $\hat{t}_3, \hat{t}_4 \sim O \left(\frac{1}{\log P_0} \right)$ and $\hat{t}_1, \hat{t}_2 \sim 0.5 + O \left(\frac{1}{\log P_0} \right)$, and finally (d) follows from the facts that $\frac{P_1}{P_2} \sim \Theta(1), \hat{t}_1 + \hat{t}_4, \hat{t}_2 + \hat{t}_4 \sim \Theta(1)$, and $\hat{t}_4 \sim O \left(\frac{1}{\log P_0} \right)$.

Next, we bound the difference between the terms in the second cut of (44) and the second term in C_{DPC}^{low}

$$\begin{aligned}
& \hat{t}_2 C \left(\frac{h_{02}^2 \hat{P}_0^{(2)}}{\hat{t}_2} + \frac{(h_{12}^2 + h_{13}^2) \hat{P}_1^{(2)}}{\hat{t}_2} + \frac{2h_{02}h_{12}\sqrt{\hat{P}_0^{(2)}\hat{P}_1^{(2)}}}{\hat{t}_2} + \frac{h_{02}^2 h_{13}^2 \hat{P}_0^{(2)} \hat{P}_1^{(2)}}{\hat{t}_2^2} \right) + \hat{t}_3 C \left(\frac{h_{02}^2 \hat{P}_0^{(3)}}{\hat{t}_3} \right) \\
& + \hat{t}_4 C \left(\frac{h_{13}^2 \hat{P}_1^{(4)}}{\hat{t}_4} \right) - t_2 C \left(\frac{h_{02}^2 P_0^{(2)}}{t_2} \right) - t_2 C \left(\frac{h_{13}^2 P_1}{t_2} \right) \stackrel{(a)}{\lesssim} \\
& \frac{\hat{t}_2}{2} \ln \left(\frac{h_{02}^2 h_{13}^2 \hat{P}_0^{(2)} \hat{P}_1^{(2)}}{\hat{t}_2^2} \right) + \hat{t}_3 C \left(\frac{h_{02}^2 \hat{P}_0^{(3)}}{\hat{t}_3} \right) + \hat{t}_4 C \left(\frac{h_{13}^2 \hat{P}_1^{(4)}}{\hat{t}_4} \right) - \left(\frac{\hat{t}_2}{2} + \frac{\hat{t}_3 + \hat{t}_4}{4} \right) \ln \left(\frac{h_{02}^2 h_{13}^2 \hat{P}_0^{(2)} P_1}{\hat{t}_2^2} \right) + O \left(\frac{1}{P_0} \right) \stackrel{(b)}{\lesssim} \\
& \frac{\hat{t}_2}{2} \ln \left(\frac{h_{02}^2 h_{13}^2 \hat{P}_0^{(2)} P_1}{\hat{t}_2^2} \right) + \frac{\hat{t}_3}{2} \ln \left(\frac{h_{02}^2 P_0}{\hat{t}_3 + \hat{t}_2} \right) + \frac{\hat{t}_4}{2} \ln \left(\frac{h_{13}^2 P_1}{\hat{t}_4 + \hat{t}_2} \right) \\
& - \left(\frac{\hat{t}_2}{2} + \frac{\hat{t}_3 + \hat{t}_4}{4} \right) \ln \left(\frac{h_{02}^2 \hat{P}_0^{(2)}}{t_2} \right) - \left(\frac{\hat{t}_2}{2} + \frac{\hat{t}_3 + \hat{t}_4}{4} \right) \ln \left(\frac{h_{13}^2 P_1}{t_2} \right) + O \left(\frac{1}{P_0} \right) \stackrel{(c)}{\lesssim} \\
& \frac{\hat{t}_3}{4} \ln \left(\frac{P_0^2}{\hat{P}_0^{(2)} P_1} \right) + \frac{\hat{t}_4}{4} \ln \left(\frac{P_1}{\hat{P}_1^{(2)}} \right) + O \left(\frac{1}{\log P_0} \right) \stackrel{(d)}{\lesssim} O \left(\frac{1}{\log P_0} \right). \tag{59}
\end{aligned}$$

Here, (a) follows from (48), the fact that $P_0^{(2)} = \hat{P}_0^{(2)} \sim \Theta(P_0)$ and upper-bounding $\hat{P}_0^{(3)} \leq P_0, \hat{P}_1^{(4)} \leq P_1$, noting the facts that $\hat{P}_0^{(2)} + \hat{P}_0^{(3)} \leq P_0$ and $\hat{P}_1^{(2)} + \hat{P}_1^{(4)} = P_1$, the functions $\hat{t}_2 \ln(P_0 - x) + \hat{t}_3 \ln(\hat{t}_3 + h_{02}^2 x)$ and

$\hat{t}_2 \ln(P_1 - y) + \hat{t}_4 \ln(\hat{t}_4 + h_{13}^2 y)$ are maximized at $x \leq \frac{\hat{t}_3}{\hat{t}_2 + \hat{t}_3} P_0$ and $y \leq \frac{\hat{t}_4}{\hat{t}_2 + \hat{t}_4} P_1$, hence, substituting $\hat{P}_0^{(3)} = \frac{\hat{t}_3}{\hat{t}_2 + \hat{t}_3} P_0$ and $\hat{P}_1^{(4)} = \frac{\hat{t}_4}{\hat{t}_2 + \hat{t}_4} P_1$ upper-bounds the expression which results in (b), (c) follows from $\hat{t}_3, \hat{t}_4 \sim O\left(\frac{1}{\log P_0}\right)$, $\hat{t}_1, \hat{t}_2 \sim 0.5 + O\left(\frac{1}{\log P_0}\right)$, and finally (d) follows from the fact that $\hat{P}_0^{(2)}, P_1 \sim \Theta(P_0)$ and also $\hat{t}_3, \hat{t}_4 \sim O\left(\frac{1}{\log P_0}\right)$.

Noting that the second and the third cuts are the same, and using the same argument as in (59), we can bound the difference between the terms in the third cut of (44) and the third term in C_{DPC}^{low} as

$$\begin{aligned} & \hat{t}_1 C \left(\frac{h_{01}^2 \hat{P}_0^{(1)}}{\hat{t}_1} + \frac{(h_{12}^2 + h_{23}^2) \hat{P}_2^{(1)}}{\hat{t}_1} + \frac{2h_{01}h_{12}\sqrt{\hat{P}_0^{(1)}\hat{P}_2^{(1)}}}{\hat{t}_1} + \frac{h_{01}^2 h_{23}^2 \hat{P}_0^{(1)} \hat{P}_2^{(1)}}{\hat{t}_1^2} \right) \\ & + \hat{t}_3 C \left(\frac{h_{01}^2 \hat{P}_0^{(3)}}{\hat{t}_3} \right) + \hat{t}_4 C \left(\frac{h_{23}^2 \hat{P}_2^{(4)}}{\hat{t}_4} \right) - t_1 C \left(\frac{h_{01}^2 P_0^{(1)}}{t_1} \right) - t_1 C \left(\frac{h_{23}^2 P_2}{t_1} \right) \leq O\left(\frac{1}{\log P_0}\right). \end{aligned} \quad (60)$$

Observing (57), (58), (59) and (60), completes the proof of the Theorem. \blacksquare

Theorem 9 In low SNR scenarios, assuming $P_1 = \gamma_1 P_0$, $P_2 = \gamma_2 P_0$ with γ_1, γ_2 constants independent of the SNR, when the power available for the source and each relay tends to zero and $(h_{13}\sqrt{\gamma_1} + h_{23}\sqrt{\gamma_2})^2 \leq \min(h_{01}^2, h_{02}^2)$, the ratio of the achievable rate of the simultaneous relaying protocol based on DDF to cut-set upper bound goes to 1. In this scenario $t_3 = t_4 = \frac{1}{2}$, and no private messages should be transmitted.

Proof: By the same argument as in Theorem 8 and considering only the fourth cut, we obtain another upper bound on the capacity. By the following inequality

$$\ln(1+x) \leq x. \quad (61)$$

we can bound the upper bound on the capacity as

$$C^{up} \leq \frac{(h_{13}\sqrt{\gamma_1} + h_{23}\sqrt{\gamma_2})^2 P_0}{2 \ln 2}. \quad (62)$$

Now, assuming $t_1 = t_2 = 0, t_3 = t_4 = \frac{1}{2}$, and transmitting just the common message, we can achieve the following rate C_{DDF}^{low} :

$$C_{DDF}^{low} = \min \left(\frac{1}{2} C(2h_{02}^2 P_0), \frac{1}{2} C(2(h_{13}\sqrt{\gamma_1} + h_{23}\sqrt{\gamma_2})^2 P_0) \right). \quad (63)$$

According to the Taylor expansion of $\ln(1+x)$ at $x=0$, we have

$$x - \frac{x^2}{2} \leq \ln(1+x), \quad (64)$$

Hence,

$$\frac{1}{\ln 2} \min \left(\frac{h_{02}^2 P_0}{2} - \frac{h_{02}^4 P_0^2}{2}, \frac{(h_{13}\sqrt{\gamma_1} + h_{23}\sqrt{\gamma_2})^2 P_0}{2} - \frac{(h_{13}\sqrt{\gamma_1} + h_{23}\sqrt{\gamma_2})^4 P_0^2}{2} \right) \leq C_{DDF}^{low}. \quad (65)$$

By (62), (65), and $(h_{13}\sqrt{\gamma_1} + h_{23}\sqrt{\gamma_2})^2 \leq \min(h_{01}^2, h_{02}^2)$, we have

$$\lim_{P_0 \rightarrow 0} \frac{C_{DDF}^{low}}{C^{up}} \rightarrow 1. \quad (66)$$

V. SIMULATION RESULT

In this section, the achievable rate of different proposed schemes, i.e., SSRD, DPC, BME, and BME-DPC are compared with each other and with the upper bound in different channel conditions.

Figure 9 compares the achievable rate of the SSRD scheme with that of the DPC scheme for successive relaying and the DDF scheme for simultaneous relaying protocols. Here the symmetric scenario in which $P_1 = P_2$ and $h_{01} = h_{02} = h_{12} = h_{13} = h_{23} = 1$ is considered. The upper bound is also included in the figure.

In order to satisfy the condition in Theorem 9, i.e., $(h_{13}\sqrt{\gamma_1} + h_{23}\sqrt{\gamma_2})^2 \leq \min(h_{01}^2, h_{02}^2)$, in Figs. 9a and b, we also assume $P_0 = P_1 + 10(\text{dB}) = P_2 + 10(\text{dB})$ and $P_0 = P_1 + 5(\text{dB}) = P_2 + 5(\text{dB})$, respectively. As the Figs. 9a and b show, SSRD achievable rate almost coincides with the upper bound over all ranges of SNR. As proved in the previous section, in high SNR scenario, SSRD scheme coincides with DPC and the successive relaying protocol becomes optimum, while in low SNR scenario it coincides with DDF and the simultaneous relaying protocol is optimum.

On the other hand, in Figs. 9c and d we assume that $P_0 = P_1 = P_2$ and $P_0 = P_1 - 5(\text{dB}) = P_2 - 5(\text{dB})$. In this situation, the condition in Theorem 9 is no longer satisfied. Therefore, as these figures show, the ratio of the achievable rate of the SSRD scheme to the cut-set bound, i.e., $\frac{C_{SSRD}^{low}}{C_{up}}$ does not tend to one. Furthermore, the achievable rates of the SSRD, DPC, and DDF schemes coincide with each other.

Figure 10 compares the achievable rate of different successive schemes with each other and the successive cut-set bound. It shows as the inter relay channel becomes stronger, BME scheme can achieve the successive cut-set bound, while the achievable rate of the DPC is independent of that channel. Furthermore, this figure indicates BME-DPC gives a better achievable rate with respect to BME with successive decoding which was proposed in [32].

VI. CONCLUSION

In this paper, we investigated the problem of cooperative strategies for a half-duplex parallel relay channel with two relays. We derived the optimum relay ordering and hence the asymptotic capacity of the half-duplex Gaussian parallel relay channel in low and high SNR scenarios.

Simultaneous and *Successive* relaying protocols, associated with two possible relay orderings were proposed. For simultaneous relaying, each relay employs *DDF*. On the other hand, for successive relaying, we proposed a *Non-Cooperative Coding* scheme based on DPC and a *Cooperative Coding* scheme based on BME. Moreover, a coding scheme based on the combination of DPC and BME, in which one of the relays uses DPC while the other one employs BME was proposed. We showed that this composite scheme achieves a better rate with respect to cooperative coding based on BME with backward or successive decoding in the Gaussian case.

We also proposed the SSRD scheme as a combination of the simultaneous and successive protocols based on DPC. In high SNR scenarios, we proved that our *Non-Cooperative Coding* scheme based on DPC asymptotically achieves the capacity. Hence, in the high SNR scenario, the optimum relay ordering is *Successive*. On the other hand, in low SNR where $(h_{13}\gamma_1 + h_{23}\gamma_2)^2 \leq \min(h_{01}^2, h_{02}^2)$, *DDF* achieves the capacity. Hence, in low SNR scenario and under the condition specified above for the channel coefficients, the optimum relay ordering is *Simultaneous*.

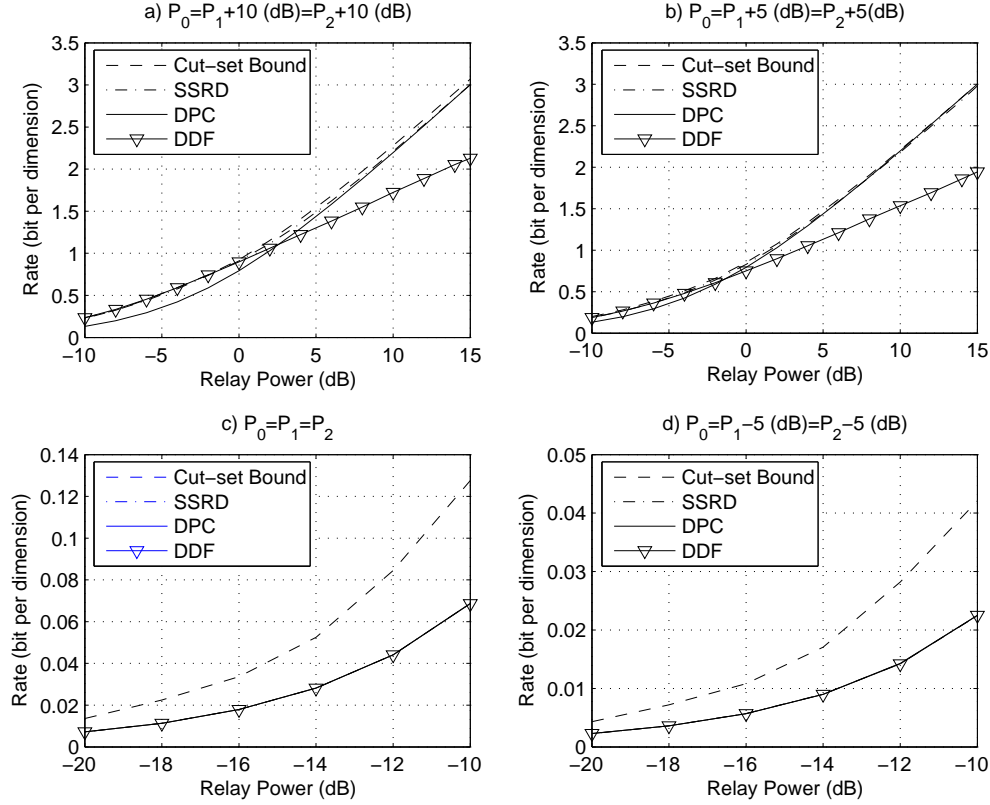


Fig. 9. Rate versus relay power.

APPENDIX A

*Proof of Theorem 1*Codebook Construction:

Let us divide time slot number b , $b = 1, 2, \dots, B + 1$ into odd and even numbers. At odd and even time slots, source generates $2^{nr_{AUX}^{(1)}}$ and $2^{nr_{AUX}^{(2)}}$ sequences $\mathbf{u}_0^{(1)}(q_1)$ and $\mathbf{u}_0^{(2)}(q_2)$ according to $\prod_{i=1}^{t_1 n} p(u_{0,i}^{(1)})$ and $\prod_{i=1}^{t_2 n} p(u_{0,i}^{(2)})$, respectively. Then, source throws $\mathbf{u}_0^{(1)}$ and $\mathbf{u}_0^{(2)}$ sequences uniformly into $2^{nR^{(1)}}$ and $2^{nR^{(2)}}$ bins, respectively. Let us denote $\mathcal{B}_1(w^{(b)})$ and $\mathcal{B}_2(w^{(b)})$ as the set of sequences at the odd or even time slot that belong to the $w^{(b)}$ 'th bin, respectively (for odd time slots, $w^{(b)} \leq 2^{nR^{(1)}}$, and for the even time slots, $w^{(b)} \leq 2^{nR^{(2)}}$).

Relay 1 and relay 2 generate $2^{nR^{(1)}}$ and $2^{nR^{(2)}}$ i.i.d $\mathbf{x}_1^{(2)}$ and $\mathbf{x}_2^{(1)}$ sequences according to probabilities $\prod_{i=1}^{t_2 n} p(x_{1,i}^{(2)})$ and $\prod_{i=1}^{t_1 n} p(x_{2,i}^{(1)})$. Furthermore, for all q_1 and q_2 , the source generates double indexed codebooks $\mathbf{x}_0^{(1)}(w^{(b)}|w^{(b-1)}, q_1)$ and $\mathbf{x}_0^{(2)}(w^{(b)}|w^{(b-1)}, q_2)$ according to $\prod_{i=1}^{t_1 n} p(x_{0,i}^{(1)} | x_{2,i}^{(1)}, u_{0,i}^{(1)})$ and $\prod_{i=1}^{t_2 n} p(x_{0,i}^{(2)} | x_{1,i}^{(2)}, u_{0,i}^{(2)})$, respectively.

Encoding:*Encoding at the source:*

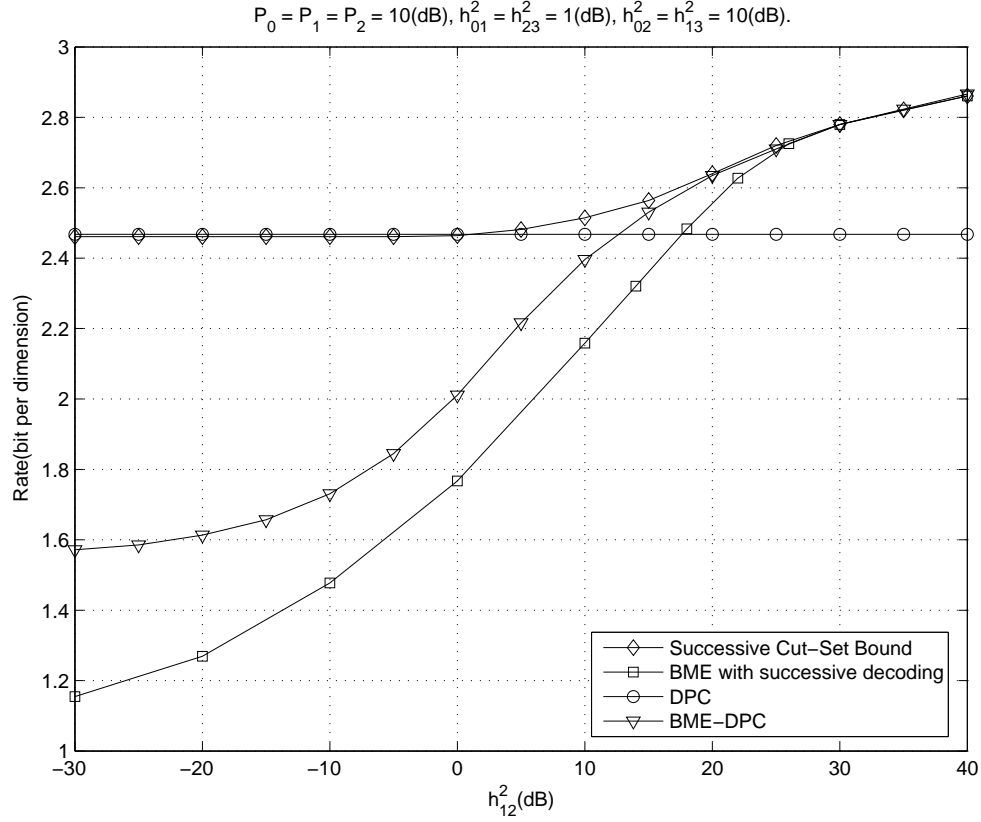


Fig. 10. Rate versus inter relay gain.

At the odd time slot b , the source intends to send the message $w^{(b)}$ to the first relay. In order to do that, since source knows what it has transmitted during the last time slot to the second relay, it chooses a codeword $\mathbf{u}_0^{(1)}(q_1)$ such that $\mathbf{u}_0^{(1)}(q_1) \in \mathcal{B}_1(w^{(b)})$ and $(\mathbf{u}_0^{(1)}(q_1), \mathbf{x}_2^{(1)}(w^{(b-1)})) \in A_\epsilon^{(n)}$. Such a task can be done almost surely, if $r_{AUX}^{(1)} - R^{(1)} \geq t_1 I(U_0^{(1)}; X_2^{(1)})$ (See [12]). Following that it sends $\mathbf{x}_0^{(1)}(\mathbf{u}_0^{(1)}, \mathbf{x}_2^{(1)})$.

At the even time slot b , the source sends the message $w^{(b)}$ to the second relay in the similar manner. Such a task can be done almost surely if $r_{AUX}^{(2)} - R^{(2)} \geq t_2 I(U_0^{(2)}; X_1^{(2)})$.

Encoding at relay 1:

At the even time slot b , relay 1 encodes $w^{(b-1)} \in \{1, \dots, 2^{nR^{(1)}}\}$ to $\mathbf{x}_1^{(2)}(w^{(b-1)})$.

Encoding at relay 2:

At the odd time slot b , relay 2 encodes $w^{(b-1)} \in \{1, \dots, 2^{nR^{(2)}}\}$ to $\mathbf{x}_2^{(1)}(w^{(b-1)})$.

Decoding:

Decoding at relay 1:

At the odd time slot b , relay 1 declares $\hat{w}^{(b)} = w^{(b)}$ iff all the sequences $\mathbf{u}_0^{(1)}(q_1)$ which are jointly typical with

$\mathbf{y}_1^{(1)}$ belong to a unique bin $\mathcal{B}_1(\hat{w}^{(b)})$. Therefore, in order to make the probability of error zero, from [12], we have

$$r_{AUX}^{(1)} \leq t_1 I(U_0^{(1)}; Y_1^{(1)}). \quad (67)$$

According to (67) and the encoding condition at source, we have

$$R^{(1)} \leq t_1 \left(I(U_0^{(1)}; Y_1^{(1)}) - I(U_0^{(1)}; X_2^{(1)}) \right). \quad (68)$$

Decoding at relay 2:

At the even time slot b , relay 2 declares $\hat{w}^{(b)} = w^{(b)}$ iff all the sequences $\mathbf{u}_0^{(2)}(q_2)$ which are jointly typical with $\mathbf{y}_2^{(2)}$ belong to a unique bin $\mathcal{B}_2(\hat{w}^{(b)})$. Therefore, in order to make the probability of error zero, from [12], we have

$$r_{AUX}^{(2)} \leq t_2 I(U_0^{(2)}; Y_2^{(2)}). \quad (69)$$

According to (69) and the encoding condition at source, we have

$$R^{(2)} \leq t_2 \left(I(U_0^{(2)}; Y_2^{(2)}) - I(U_0^{(2)}; X_1^{(2)}) \right). \quad (70)$$

Decoding at the final destination:

At the odd time slot b , destination declares $\hat{w}^{(b-1)} = w^{(b-1)}$ iff $(\mathbf{x}_2^{(1)}(\hat{w}^{(b-1)}), \mathbf{y}_3^{(1)}) \in A_\epsilon^{(n)}$. Hence, in order to make the probability of error zero, from [12], we have

$$R^{(1)} \leq t_1 I(X_2^{(1)}; Y_3^{(1)}). \quad (71)$$

Similarly, at the even time slot b , we have

$$R^{(2)} \leq t_2 I(X_1^{(2)}; Y_3^{(2)}). \quad (72)$$

From the encoding at the source and (67)-(72), we obtain (9)-(11).

APPENDIX B

Proof of Theorem 2

Codebook Construction:

Let us divide the time slots b , $b = 1, 2, \dots, B+2$ into odd and even time slots. The source generates two codebooks $\mathbf{x}_0^{(1)}(w^{(b)}|w^{(b-1)}, s_1^{(b-2)})$ and $\mathbf{x}_0^{(2)}(w^{(b)}|w^{(b-1)}, s_2^{(b-2)})$ of size $2^{nR^{(1)}}$ and $2^{nR^{(2)}}$ corresponding to even and odd time slots, respectively. The first codebook is generated according to the probability $p(\mathbf{x}_0^{(1)}, \mathbf{x}_2^{(1)}, \mathbf{u}_2^{(1)}) = \prod_{i=1}^{t_1 n} p(u_{2,i}^{(1)})p(x_{2,i}^{(1)}|u_{2,i}^{(1)})p(x_{0,i}^{(1)}|x_{2,i}^{(1)}, u_{2,i}^{(1)})$, and the second codebook is generated according to the probability $p(\mathbf{x}_0^{(2)}, \mathbf{x}_1^{(2)}, \mathbf{u}_1^{(2)}) = \prod_{i=1}^{t_2 n} p(u_{1,i}^{(2)})p(x_{1,i}^{(2)}|u_{1,i}^{(2)})p(x_{0,i}^{(2)}|x_{1,i}^{(2)}, u_{1,i}^{(2)})$.

On the other hand, relay 2 generates $2^{nr_{B+n}^{(1)}}$ i.i.d codewords $\mathbf{u}_2^{(1)}$ and $2^{nR^{(2)}}$ i.i.d codewords $\mathbf{x}_2^{(1)}$ according to the probabilities $p(\mathbf{u}_2^{(1)}) = \prod_{i=1}^{t_1 n} p(u_{2,i}^{(1)})$ and $p(\mathbf{x}_2^{(1)} | \mathbf{u}_2^{(1)}) = \prod_{i=1}^{t_1 n} p(x_{2,i}^{(1)} | u_{2,i}^{(1)})$ at each odd time slot and relay 1 generates $2^{nr_{B+n}^{(2)}}$ i.i.d codewords $\mathbf{u}_1^{(2)}$ and $2^{nR^{(1)}}$ i.i.d codewords $\mathbf{x}_1^{(2)}$ according to the probabilities $p(\mathbf{u}_1^{(2)}) = \prod_{i=1}^{t_2 n} p(u_{1,i}^{(2)})$ and $p(\mathbf{x}_1^{(2)} | \mathbf{u}_1^{(2)}) = \prod_{i=1}^{t_2 n} p(x_{1,i}^{(2)} | u_{1,i}^{(2)})$ at each even time slot, respectively.

Encoding:

Encoding at the source:

At the odd time slot b , source encodes $w^{(b)} \in \{1, \dots, 2^{nR^{(1)}}\}$ to $\mathbf{x}_0^{(1)}(w^{(b)}|w^{(b-1)}, s_1^{(b-2)})$ and at the even time slot b , it encodes $w^{(b)} \in \{1, \dots, 2^{nR^{(2)}}\}$ to $\mathbf{x}_0^{(2)}(w^{(b)}|w^{(b-1)}, s_2^{(b-2)})$ and sends them in odd and even time slots, respectively.

Encoding at relay 1:

At the even time slot b , relay 1 encodes the bin index $s_2^{(b-2)}$ of the message $w^{(b-2)}$ it has received from relay 2 in the previous time slot to $\mathbf{u}_1^{(2)}(s_2^{(b-2)})$. Following that, it encodes $w^{(b-1)}$ which was received from the source in time slot $b-1$ to $\mathbf{x}_1^{(2)}(w^{(b-1)}|s_2^{(b-2)})$ and sends it.

Encoding at relay 2:

At the odd time slot b , relay 2 encodes the bin index $s_1^{(b-2)}$ of the message $w^{(b-2)}$ it has received from relay 1 in the previous time slot to $\mathbf{u}_2^{(1)}(s_1^{(b-2)})$. Following that, it encodes $w^{(b-1)}$ which was received from the source in time slot $b-1$ to $\mathbf{x}_2^{(1)}(w^{(b-1)}|s_1^{(b-2)})$ and sends it.

Decoding:

Decoding at relay 1:

Knowing $w^{(b-2)}$ and consequently $s_1^{(b-2)}$, at time slot b , relay 1 declares $(\hat{w}^{(b-1)}, \hat{w}^{(b)}) = (w^{(b-1)}, w^{(b)})$ iff there exists a unique $(\hat{w}^{(b-1)}, \hat{w}^{(b)})$ such that

$$\left(\mathbf{x}_0^{(1)}(\hat{w}^{(b)}|\hat{w}^{(b-1)}, s_1^{(b-2)}), \mathbf{x}_2^{(1)}(\hat{w}^{(b-1)}|s_1^{(b-2)}), \mathbf{u}_2^{(1)}(s_1^{(b-2)}), \mathbf{y}_1^{(1)} \right) \in A_\epsilon^{(n)}.$$

Hence, in order to make probability of error zero, from the Extended MAC capacity region (See [12], [24], [25], and [26]), we have

$$R^{(1)} \leq t_1 I(X_0^{(1)}; Y_1^{(1)} | X_2^{(1)}, U_2^{(1)}), \quad (73)$$

$$R^{(1)} + R^{(2)} \leq t_1 I(X_0^{(1)}, X_2^{(1)}; Y_1^{(1)} | U_2^{(1)}). \quad (74)$$

Decoding at relay 2:

Knowing $w^{(b-2)}$ and consequently $s_2^{(b-2)}$, at time slot b , relay 2 declares $(\hat{w}^{(b-1)}, \hat{w}^{(b)}) = (w^{(b-1)}, w^{(b)})$ iff there exists a unique $(\hat{w}^{(b-1)}, \hat{w}^{(b)})$ such that

$$\left(\mathbf{x}_0^{(2)}(\hat{w}^{(b)}|\hat{w}^{(b-1)}, s_2^{(b-2)}), \mathbf{x}_1^{(2)}(\hat{w}^{(b-1)}|s_2^{(b-2)}), \mathbf{u}_1^{(2)}(s_2^{(b-2)}), \mathbf{y}_2^{(2)} \right) \in A_\epsilon^{(n)}.$$

Hence, in order to make the probability of error zero, from Extended MAC capacity region (See [12], [24], [25], and [26]), we have

$$R^{(2)} \leq t_2 I(X_0^{(2)}; Y_2^{(2)} | X_1^{(2)}, U_1^{(2)}), \quad (75)$$

$$R^{(1)} + R^{(2)} \leq t_2 I(X_0^{(2)}, X_1^{(2)}; Y_2^{(2)} | U_1^{(2)}). \quad (76)$$

Decoding at the final destination:

Decoding at the final destination can be done either *Successively* or *Backwardly* as follows.

1) Successive Decoding:

At the end of odd time slot b , destination first declares the bin index $\hat{s}_1^{(b-2)} = s_1^{(b-2)}$ of the message $w^{(b-2)}$ iff there exists a unique $\hat{s}_1^{(b-2)}$ such that $(\mathbf{u}_2^{(1)}(\hat{s}_1^{(b-2)}), \mathbf{y}_3^{(1)}) \in A_\epsilon^{(n)}$. Hence, in order to make the probability of error zero, from [12] we have

$$r_{Bin}^{(1)} \leq t_1 I(U_2^{(1)}; Y_3^{(1)}). \quad (77)$$

Having decoded the bin index $\hat{s}_1^{(b-2)}$ of the message $w^{(b-2)}$, destination can resolve its uncertainty about the message $w^{(b-2)}$ and declares $\hat{w}^{(b-2)} = w^{(b-2)}$ iff there exists a unique $\hat{w}^{(b-2)}$ such that $(\mathbf{x}_1^{(2)}(\hat{w}^{(b-2)} | s_2^{(b-3)}), \mathbf{u}_1^{(2)}(s_2^{(b-3)}), \mathbf{y}_3^{(2)}) \in A_\epsilon^{(n)}$. Hence, in order to make the probability of error zero, from [12] we have

$$R^{(1)} - r_{Bin}^{(1)} \leq t_2 I(X_1^{(2)}; Y_3^{(2)} | U_1^{(2)}). \quad (78)$$

Using the same argument for the even time slot b , we have

$$r_{Bin}^{(2)} \leq t_2 I(U_1^{(2)}; Y_3^{(2)}), \quad (79)$$

$$R^{(2)} - r_{Bin}^{(2)} \leq t_1 I(X_2^{(1)}; Y_3^{(1)} | U_2^{(1)}). \quad (80)$$

From (77), (78), (79), and (80), $R^{(1)}$ and $R^{(2)}$ are bounded as follows

$$R^{(1)} \leq t_2 I(X_1^{(2)}; Y_3^{(2)} | U_1^{(2)}) + t_1 I(U_2^{(1)}; Y_3^{(1)}), \quad (81)$$

$$R^{(2)} \leq t_1 I(X_2^{(1)}; Y_3^{(1)} | U_2^{(1)}) + t_2 I(U_1^{(2)}; Y_3^{(2)}). \quad (82)$$

From (73)-(76), (81), and (82), the achievable rate of BME scheme based on successive decoding is equal to

$$\begin{aligned} C_{BM_{succ}}^{low} = R^{(1)} + R^{(2)} &\leq \max_{0 \leq t_1, t_2, t_1+t_2=1} \min (\\ &\min \left(t_1 I(X_0^{(1)}; Y_1^{(1)} | X_2^{(1)}, U_2^{(1)}), t_2 I(X_1^{(2)}; Y_3^{(2)} | U_1^{(2)}) + t_1 I(U_2^{(1)}; Y_3^{(1)}) \right) + \\ &\min \left(t_1 I(X_2^{(1)}; Y_3^{(1)} | U_2^{(1)}) + t_2 I(U_1^{(2)}; Y_3^{(2)}) , t_2 I(X_0^{(2)}; Y_2^{(2)} | X_1^{(2)}, U_1^{(2)}) \right), \\ &t_1 I(X_0^{(1)}, X_2^{(1)}; Y_1^{(1)} | U_2^{(1)}), t_2 I(X_0^{(2)}, X_1^{(2)}; Y_2^{(2)} | U_1^{(2)}) \Big). \end{aligned} \quad (83)$$

2) Backward Decoding:

Following receiving the sequence corresponding to the $B+2$ 'th time slot, destination starts decoding the messages in a backward manner, i.e. from $w^{(B)}$ back to $w^{(1)}$. At the end of odd time slot b , knowing the value $s_2^{(b-1)}$ from the received signal in time slot $b+1$, destination declares $(\hat{w}^{(b-1)}, \hat{s}_1^{(b-2)}) = (w^{(b-1)}, s_1^{(b-2)})$ iff there exists a unique pair $(\hat{w}^{(b-1)}, \hat{s}_1^{(b-2)})$ such that $f_{Bin}^{(2)}(\hat{w}^{(b-1)}) = s_2^{(b-1)}$ and $(\mathbf{x}_2^{(1)}(\hat{w}^{(b-1)}, \hat{s}_1^{(b-2)}), \mathbf{u}_2^{(1)}(\hat{s}_1^{(b-2)}), \mathbf{y}_3^{(1)}) \in A_\epsilon^{(n)}$. Similarly, at the end of even time slot b , knowing the value $s_1^{(b-1)}$ for the received signal in time slot $b+1$, destination declares $(\hat{w}^{(b-1)}, \hat{s}_2^{(b-2)}) = (w^{(b-1)}, s_2^{(b-2)})$ iff there exists a unique pair $(\hat{w}^{(b-1)}, \hat{s}_2^{(b-2)})$ such that $f_{Bin}^{(1)}(\hat{w}^{(b-1)}) = s_1^{(b-1)}$ and $(\mathbf{x}_1^{(2)}(\hat{w}^{(b-1)}, \hat{s}_1^{(b-2)}), \mathbf{u}_1^{(2)}(\hat{s}_2^{(b-2)}), \mathbf{y}_3^{(2)}) \in A_\epsilon^{(n)}$. Hence, in order to make the

probability of error zero, from [12] we have

$$r_{Bin}^{(1)} \leq R^{(1)}, \quad (84)$$

$$r_{Bin}^{(2)} \leq R^{(2)}, \quad (85)$$

$$R^{(2)} - r_{Bin}^{(2)} \leq t_1 I \left(X_2^{(1)}; Y_3^{(1)} \mid U_2^{(1)} \right), \quad (86)$$

$$R^{(2)} - r_{Bin}^{(2)} + r_{Bin}^{(1)} \leq t_1 I \left(X_2^{(1)}, U_2^{(1)}; Y_3^{(1)} \right), \quad (87)$$

$$R^{(1)} - r_{Bin}^{(1)} \leq t_2 I \left(X_1^{(2)}; Y_3^{(2)} \mid U_1^{(2)} \right), \quad (88)$$

$$R^{(1)} - r_{Bin}^{(1)} + r_{Bin}^{(2)} \leq t_2 I \left(X_1^{(2)}, U_1^{(2)}; Y_3^{(2)} \right). \quad (89)$$

Hence, by employing BME and Backward decoding, the following rate is achievable subject to (73)-(76) and (84)-(89) constraints.

$$C_{BME_{back}}^{low} = R^{(1)} + R^{(2)}. \quad (90)$$

Optimum input distributions

Now, we prove there exists input probability distributions $(p(x_0^{(1)}, x_2^{(1)}, u_2^{(1)}))$ and $(p(x_0^{(2)}, x_1^{(2)}, u_1^{(2)}))$ which maximize (90) and have the following property: $u_2^{(1)}$ is independent from $(x_0^{(1)}, x_2^{(1)})$ and $u_1^{(2)}$ is independent from $(x_0^{(2)}, x_1^{(2)})$. To prove this, consider $p(x_0^{(1)}, x_2^{(1)}, u_2^{(1)})$ and $p(x_0^{(2)}, x_1^{(2)}, u_1^{(2)})$ along with t_1, t_2 which maximize (90) subject to the required constraints. Let us define $\hat{p}(x_0^{(1)}, x_2^{(1)}, u_2^{(1)})$ and $\hat{p}(x_0^{(2)}, x_1^{(2)}, u_1^{(2)})$ as

$$\hat{p}(x_0^{(1)}, x_2^{(1)}, u_2^{(1)}) = p(u_2^{(1)})p(x_0^{(1)}, x_2^{(1)}), \quad (91)$$

$$\hat{p}(x_0^{(2)}, x_1^{(2)}, u_1^{(2)}) = p(u_1^{(2)})p(x_0^{(2)}, x_1^{(2)}), \quad (92)$$

Now, we show that $\hat{p}(x_0^{(1)}, x_2^{(1)}, u_2^{(1)})$ and $\hat{p}(x_0^{(2)}, x_1^{(2)}, u_1^{(2)})$ along with t_1, t_2 achieve at least the same rate as the optimum one. Let us denote the values of mutual information and entropy with respect to the input distributions p, \hat{p} by I_p, H_p and $I_{\hat{p}}, H_{\hat{p}}$, respectively. The right-hand sides of (86)-(89) with respect to p can be upper-bounded by the ones corresponding to \hat{p} as follows

$$t_1 I_p \left(X_2^{(1)}; Y_3^{(1)} \mid U_2^{(1)} \right) \stackrel{(a)}{\leq} t_1 I_p \left(X_2^{(1)}; Y_3^{(1)} \right) = t_1 I_{\hat{p}} \left(X_2^{(1)}; Y_3^{(1)} \right), \quad (93)$$

$$t_1 I_p \left(X_2^{(1)}, U_2^{(1)}; Y_3^{(1)} \right) \stackrel{(a)}{=} t_1 I_p \left(X_2^{(1)}; Y_3^{(1)} \right) = t_1 I_{\hat{p}} \left(X_2^{(1)}; Y_3^{(1)} \right), \quad (94)$$

$$t_2 I_p \left(X_1^{(2)}; Y_3^{(2)} \mid U_1^{(2)} \right) \stackrel{(b)}{\leq} t_2 I_p \left(X_1^{(2)}; Y_3^{(2)} \right) = t_2 I_{\hat{p}} \left(X_1^{(2)}; Y_3^{(2)} \right), \quad (95)$$

$$t_2 I_p \left(X_1^{(2)}, U_1^{(2)}; Y_3^{(2)} \right) \stackrel{(b)}{=} t_2 I_p \left(X_1^{(2)}; Y_3^{(2)} \right) = t_2 I_{\hat{p}} \left(X_1^{(2)}; Y_3^{(2)} \right). \quad (96)$$

where (a) follows from the fact that $U_2^{(1)} \rightarrow X_2^{(1)} \rightarrow Y_3^{(1)}$ forms a Markov chain and (b) follows from the fact that $U_1^{(2)} \rightarrow X_1^{(2)} \rightarrow Y_3^{(2)}$ forms a Markov chain. Moreover as in distribution \hat{p} , $u_2^{(1)}$ and $u_1^{(2)}$ are independent from $(x_0^{(1)}, x_2^{(1)})$ and $(x_0^{(2)}, x_1^{(2)})$, it can be easily verified that the right-hand sides of (93)-(96) are equal to the right-hand sides of (86)-(89) with the input distribution \hat{p} , respectively. Hence, by utilizing \hat{p} instead of p , the region

that satisfies (86)-(89) is enlarged. Now, let us consider the right-hand sides of (73)-(76).

$$t_1 I_p \left(X_0^{(1)}; Y_1^{(1)} \mid X_2^{(1)}, U_2^{(1)} \right) \stackrel{(a)}{\leq} t_1 I_p \left(X_0^{(1)}; Y_1^{(1)} \mid X_2^{(1)} \right) = t_1 I_{\hat{p}} \left(X_0^{(1)}; Y_1^{(1)} \mid X_2^{(1)} \right) \quad (97)$$

$$t_1 I_p \left(X_0^{(1)}, X_2^{(1)}; Y_1^{(1)} \mid U_2^{(1)} \right) \stackrel{(a)}{\leq} t_1 I_p \left(X_0^{(1)}, X_2^{(1)}; Y_1^{(1)} \right) = t_1 I_{\hat{p}} \left(X_0^{(1)}, X_2^{(1)}; Y_1^{(1)} \right) \quad (98)$$

$$t_2 I_p \left(X_0^{(2)}; Y_2^{(2)} \mid X_1^{(2)}, U_1^{(2)} \right) \stackrel{(b)}{\leq} t_2 I_p \left(X_0^{(2)}; Y_2^{(2)} \mid X_1^{(2)} \right) = t_2 I_{\hat{p}} \left(X_0^{(2)}; Y_2^{(2)} \mid X_1^{(2)} \right) \quad (99)$$

$$t_2 I_p \left(X_0^{(2)}, X_1^{(2)}; Y_2^{(2)} \mid U_1^{(2)} \right) \stackrel{(b)}{\leq} t_2 I_p \left(X_0^{(2)}, X_1^{(2)}; Y_2^{(2)} \right) = t_2 I_{\hat{p}} \left(X_0^{(2)}, X_1^{(2)}; Y_2^{(2)} \right) \quad (100)$$

where (a) follows from the fact that $U_2^{(1)} \longrightarrow (X_2^{(1)}, X_0^{(1)}) \longrightarrow Y_1^{(1)}$ form a Markov chain and (b) follows from the fact that $U_1^{(2)} \longrightarrow (X_1^{(2)}, X_0^{(2)}) \longrightarrow Y_2^{(2)}$ form a Markov chain. Similarly, we observe that the right-hand sides of (97)-(100) represent the right-hand sides of inequalities (73)-(76) with the input distribution \hat{p} . Hence, the region of $(R^{(1)}, R^{(2)})$ that satisfies (73)-(76) and (84)-(89) is enlarged by utilizing the input distribution \hat{p} instead of p . This proves the independency of input distributions with $u^{(1)}$ and $u^{(2)}$ in the optimum distribution.

Simplifying the achievable rate

As we can assume that the input distributions are of the form (91) and (92), the achievable rate can be simplified as follows.

$$C_{BME_{back}}^{low} = R^{(1)} + R^{(2)} \leq \max_{0 \leq t_1, t_2, t_1+t_2=1} \min \left(t_1 I \left(X_0^{(1)}, X_2^{(1)}; Y_1^{(1)} \right), t_2 I \left(X_0^{(2)}, X_1^{(2)}; Y_2^{(2)} \right) \right), \quad (101)$$

subject to

$$r_{Bin}^{(1)} \leq R^{(1)}, \quad (102)$$

$$r_{Bin}^{(2)} \leq R^{(2)}, \quad (103)$$

$$R^{(1)} \leq t_1 I \left(X_0^{(1)}; Y_1^{(1)} \mid X_2^{(1)} \right), \quad (104)$$

$$R^{(2)} \leq t_2 I \left(X_0^{(2)}; Y_2^{(2)} \mid X_1^{(2)} \right), \quad (105)$$

$$R^{(2)} - r_{Bin}^{(2)} + r_{Bin}^{(1)} \leq t_1 I \left(X_2^{(1)}; Y_3^{(1)} \right), \quad (106)$$

$$R^{(1)} - r_{Bin}^{(1)} + r_{Bin}^{(2)} \leq t_2 I \left(X_1^{(2)}; Y_3^{(2)} \right). \quad (107)$$

with input distributions

$$p(x_0^{(1)}, x_2^{(1)}) = p(x_2^{(1)})p(x_0^{(1)}|x_2^{(1)}),$$

$$p(x_0^{(2)}, x_1^{(2)}) = p(x_1^{(2)})p(x_0^{(2)}|x_1^{(2)}).$$

Now, we show that (101)-(107) is equivalent to

$$\begin{aligned} C_{BME_{back}}^{low} &\leq \max_{0 \leq t_1, t_2, t_1+t_2=1} \min \left(t_1 I \left(X_0^{(1)}, X_2^{(1)}; Y_1^{(1)} \right), t_2 I \left(X_0^{(2)}, X_1^{(2)}; Y_2^{(2)} \right), \right. \\ &\quad \left. t_1 I \left(X_0^{(1)}; Y_1^{(1)} \mid X_2^{(1)} \right) + t_2 I \left(X_0^{(2)}; Y_2^{(2)} \mid X_1^{(2)} \right), \right. \\ &\quad \left. t_1 I \left(X_2^{(1)}; Y_3^{(1)} \right) + t_2 I \left(X_1^{(2)}; Y_3^{(2)} \right) \right). \end{aligned} \quad (108)$$

First, it is easy to verify that (101)-(107) imply (108). Now, in order to prove that the converse is also true, we show that for every possible rate r satisfying (108), there exists a quad-tuple $(R^{(1)}, R^{(2)}, r_{Bin}^{(1)}, r_{Bin}^{(2)})$ such that $R^{(1)} + R^{(2)} = r$, $(R^{(1)}, R^{(2)}, r_{Bin}^{(1)}, r_{Bin}^{(2)})$ satisfies (101)-(107), and moreover at least one of bin rates is equal to zero, i.e. $r_{Bin}^{(1)} = 0$ or $r_{Bin}^{(2)} = 0$.

Let us define $R^{(1)} \triangleq \min \left(r, t_1 I \left(X_0^{(1)}; Y_1^{(1)} \mid X_2^{(1)} \right) \right)$, $R^{(2)} \triangleq r - R^{(1)}$. As r satisfies (108), we conclude that $(R^{(1)}, R^{(2)})$ satisfies (101), (104), and (105). Furthermore, as $R^{(1)} + R^{(2)} = r \leq t_1 I \left(X_2^{(1)}; Y_3^{(1)} \right) + t_2 I \left(X_1^{(2)}; Y_3^{(2)} \right)$, we conclude that either $R^{(1)} \leq t_2 I \left(X_1^{(2)}; Y_3^{(2)} \right)$ or $R^{(2)} \leq t_1 I \left(X_2^{(1)}; Y_3^{(1)} \right)$. For the sake of symmetry, let us assume that the first case has occurred, i.e. $R^{(1)} \leq t_2 I \left(X_1^{(2)}; Y_3^{(2)} \right)$. Now, we define $r_{Bin}^{(1)} \triangleq 0$ and $r_{Bin}^{(2)} \triangleq \max \left(0, R^{(2)} - t_1 I \left(X_2^{(1)}; Y_3^{(1)} \right) \right)$. Obviously, (102), (103), and (106) are valid. Considering (107), we have

$$R^{(1)} - r_{Bin}^{(1)} + r_{Bin}^{(2)} = R^{(1)} + \max \left(0, r - R^{(1)} - t_1 I \left(X_2^{(1)}; Y_3^{(1)} \right) \right) \stackrel{(a)}{\leq} t_2 I \left(X_1^{(2)}; Y_3^{(2)} \right) \quad (109)$$

where (a) follows from the facts that $r \leq t_1 I \left(X_2^{(1)}; Y_3^{(1)} \right) + t_2 I \left(X_1^{(2)}; Y_3^{(2)} \right)$ and $R^{(1)} \leq t_2 I \left(X_1^{(2)}; Y_3^{(2)} \right)$. Hence, (107) is also valid. The second case in which $R^{(2)} \leq t_1 I \left(X_2^{(1)}; Y_3^{(1)} \right)$ can be dealt with in a similar manner.

Hence, from the above argument, the achievable rate of BME scheme with backward decoding can be simplified as follows:

$$\begin{aligned} C_{BME_{back}}^{low} &\leq \max_{0 \leq t_1, t_2, t_1+t_2=1} \min \left(t_1 I \left(X_0^{(1)}, X_2^{(1)}; Y_1^{(1)} \right), t_2 I \left(X_0^{(2)}, X_1^{(2)}; Y_2^{(2)} \right), \right. \\ &\quad \left. t_1 I \left(X_0^{(1)}; Y_1^{(1)} \mid X_2^{(1)} \right) + t_2 I \left(X_0^{(2)}; Y_2^{(2)} \mid X_1^{(2)} \right), \right. \\ &\quad \left. t_1 I \left(X_2^{(1)}; Y_3^{(1)} \right) + t_2 I \left(X_1^{(2)}; Y_3^{(2)} \right) \right), \end{aligned} \quad (110)$$

with probabilities

$$\begin{aligned} p(x_0^{(1)}, x_2^{(1)}) &= p(x_2^{(1)})p(x_0^{(1)} \mid x_2^{(1)}), \\ p(x_0^{(2)}, x_1^{(2)}) &= p(x_1^{(2)})p(x_0^{(2)} \mid x_1^{(2)}). \end{aligned}$$

REFERENCES

- [1] E. C. van-der Meulen, "Three-terminal communication channels," *Adv. Appl. Prob.*, vol. 3, pp. 120-154, 1971.
- [2] T. M. Cover and A. El Gamal, "Capacity Theorems for the Relay Channel," *IEEE Transactions on Information Theory*, Vol. 25, No. 5, pp. 572-584, September 1979.
- [3] B. Schein and R. G. Gallager, "The Gaussian parallel relay network," in *Proc IEEE Int. Symp. Information Theory*, Sorrento, Italy, Jun. 2000, p. 22.

- [4] B. E. Schein, Distributed coordination in network information theory, in *Ph.D thesis, Massachusetts Institute of Technology*, Sept. 2001.
- [5] L. L. Xie and P. R. Kumar, "An achievable rate for the multiple-level relay channel," *IEEE Transactions on Information Theory*, Vol. 51, No. 4, pp. 1348-1358, Apr. 2005.
- [6] M. Gastpar, G. Kramer, and P. Gupta, "The multiple-relay channel: Coding and antenna-clustering capacity," in *Proc. of International Symposium on Information Theory (ISIT02)*, June 2002.
- [7] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative Strategies and Capacity Theorems for Relay Networks," *IEEE Transactions on Information Theory*, Vol. 51, Issue 9, pp. 3037-3063, Sept. 2005.
- [8] G. Kramer, M. Gastpar, and P. Gupta, "Capacity theorems for wireless relay channels," in *Proc. of the Allerton Conference on Communications, Control and Computing, Monticello, IL*, Oct. 2003.
- [9] A. Sanderovich, S. Shamai, Y. Steinberg, and G. Kramer, "Communication via decentralized processing," in *Proc. of International Symposium on Information Theory (ISIT05)*, Sept. 2005.
- [10] Ivana Maric, Roy D. Yates, "Forwarding Strategies for Gaussian Parallel-Relay Networks," in *Proc. 38th Annu. Conf. Information Sciences and Systems (CISS'04), Princeton, NJ*, Mar. 2004.
- [11] P. Razaghi and W. Yu, "Parity forwarding for multiple-relay networks," in *Proc. of International Symposium on Information Theory (ISIT06)*, July 2006, pp. 1678-1682.
- [12] T. Cover, and J. Thomas, *Elements of Information Theory*, Wiley, New York, first edition, 1991.
- [13] M.A. Khojastepour, A. Sabharwal, B. Aazhang, "On Capacity of Gaussian Cheap Relay Channel," *IEEE 2003 Global Communications Conference (GlobeCom-2003)*, December 1-5, San Francisco, CA, 2003.
- [14] M.A. Khojastepour, B. Aazhang, "Cheap Relay Channels: A Unifying Approach to Time and Frequency Division Relaying," Forty-second Annual Allerton Conference on Communication, Control, and Computing, September 29-October 1, 2004.
- [15] S. Zahedi, and A. El Gamal, "Minimum energy communication over a relay channel," in *Proc. of International Symposium on Information Theory (ISIT03)*, pp. 344, 29 June-4 July 2003.
- [16] S. Zahedi, M. Mohseni, and A. El Gamal, "On the capacity of AWGN relay channels with linear relaying functions," in *Proc. of International Symposium on Information Theory (ISIT04)*, 27 June-2 July 2004.
- [17] A. E. Gamal, and S. Zahedi, "Capacity of a class of relay channels with orthogonal components," *IEEE Transactions on Information Theory*, Vol. 51, Issue 5, pp. 1815-1817, May 2005.
- [18] Anders Host-Madsen, and Junshan Zhang, "Capacity Bounds and Power Allocation for Wireless Relay Channels," *IEEE Transactions on Information Theory*, Vol. 51, Issue 6, pp. 2020-2040, June 2005.
- [19] Y. Liang and V. V. Veeravalli, "Gaussian orthogonal relay channels: Optimal resource allocation and capacity," *IEEE Transactions on Information Theory*, Vol. 51, No. 9, pp. 3284-3289, September 2005.
- [20] L. L. Xie and P. R. Kumar, "A Network Information Theory for Wireless Communication: Scaling Laws and Optimal Operation," *IEEE Transactions on Information Theory*, Vol. IT-50, No. 5, pp. 748-767, May 2004.
- [21] P. Gupta and P. R. Kumar, "Toward an information theory of large networks: An achievable rate region," *IEEE Transactions on Information Theory*, Vol. 49, No. 8, pp. 1877-1894, Aug. 2003.
- [22] M. Gastpar, To code or not to code, in *Ph.D. Thesis, Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland*, Nov. 2002.
- [23] M. Gastpar and M. Vetterli, "On asymptotic capacity of gaussian relay networks," in *Proc. of International Symposium on Information Theory (ISIT02)*, June 2002.
- [24] D. Slepian and J. K. Wolf, "A coding theorem for multiple access channels with correlated sources," *Bell Syst. Tech. J.*, vol. 52, pp. 1037-1076, 1973.
- [25] T. S. Han, "The capacity region of a general multiple-access channel with certain correlated sources," *Inform. Contr.*, vol. 40, no. 1, pp. 37-60, 1979.
- [26] V. Prelov, "Transmission over a multiple-access channel with a special source hierarchy," *Problemy Peredachi Informatsii*, 20:3-10, 1984. English translation pp. 233-9, 1985.
- [27] Mohammad Ali Khojastepour, Ashutosh Sabharwal and Behnaam Aazhang, "Bounds on Achievable Rates for General Multi-terminal Networks with Practical Constraints," in *Proc. of Information Processing in Sensor Networks: Second International Workshop, IPSN 2003, Palo Alto, CA, USA, April 22-23, 2003*.
- [28] Max H. M. Costa, "Writing on dirty paper," *IEEE Transactions on Information Theory*, vol. 29, no. 3, pp. 439-441, May 1983.

- [29] B. Rankov and A. Wittneben, "Spectral efficient protocols for nonregenerative half-duplex relaying," in in *Proc. of 43rd Allerton Conference on Communication, Control and Computing*, Monticello (IL), USA, October 2005.
- [30] B. Rankov and A. Wittneben, "Spectral efficient signaling for half-duplex relay channels," in *Proc. of Asilomar Conf. Signals, syst., comput.*, Pacific Grove, CA, Nov.,2005.
- [31] Feng Xue, and Sumeet Sandhu, "Cooperation in a Half-Duplex Gaussian Diamond Relay Channel," *IEEE Transactions on Information Theory*, Vol. 53, Issue 10, pp. 3806-3814, Oct. 2007.
- [32] Woohyuk Chang, Sae-Young Chung, and Yong H. Lee, "Capacity Bounds for Alternating Two-Path Relay Channels," in *Proc. of the Allerton Conference on Communications, Control and Computing*, Monticello, IL, Oct. 2007.
- [33] Yang Sheng, J.-C. Belfiore, "Towards the Optimal Amplify-and-Forward Cooperative Diversity Scheme," *IEEE Transactions on Information Theory*, Vol. 53, Issue 9, pp. 3114-3126, Sept. 2007.
- [34] K. Azarian, H. El Gamal, and Ph. Schniter, "On the Achievable Diversity-Multiplexing Tradeoff in Half-Duplex Cooperative Channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 12, pp. 4152-4172, Dec. 2005.
- [35] P. Mitran, H. Ochiai, and V. Tarokh, "Space-time diversity enhancements using collaborative communications," *IEEE Trans. Inform. Theory*, vol. 51, no. 6, pp. 2041-2057, June 2005.
- [36] Y. Fan, C. Wang, J. S. Thompson and H. V. Poor, "Recovering multiplexing loss through successive relaying using simple repetition coding," *IEEE Transactions on Wireless Communications*, Vol. 6, No. 12, pp. 4484-4493, Dec., 2007.