Adaptive Mode Selection and Power Allocation in Bidirectional Buffer-aided Relay Networks

Vahid Jamali†, Nikola Zlatanov†, Aissa Ikhlef‡, and Robert Schober†
† Friedrich-Alexander University (FAU), Erlangen, Germany
‡ University of British Columbia (UBC), Vancouver, Canada

Abstract—In this paper, we consider the problem of sum rate maximization in a bidirectional relay network with fading. Hereby, user 1 and user 2 communicate with each other only through a relay, i.e., a direct link between user 1 and user 2 is not present. In this network, there exist six possible transmission modes: four point-to-point modes (user 1-to-relay, user 2-to-relay, relay-to-user 1, relay-to-user 2), a multiple access mode (both users to the relay), and a broadcast mode (the relay to both users). Most existing protocols assume a fixed schedule of using a subset of the aforementioned transmission modes, as a result, the sum rate is limited by the capacity of the weakest link associated with the relay in each time slot. Motivated by this limitation, we develop a protocol which is not restricted to adhere to a predefined schedule for using the transmission modes. Therefore, all transmission modes of the bidirectional relay network can be used adaptively based on the instantaneous channel state information (CSI) of the involved links. To this end, the relay has to be equipped with two buffers for the storage of the information received from users 1 and 2, respectively. For the considered network, given a total average power budget for all nodes, we jointly optimize the transmission mode selection and power allocation based on the instantaneous CSI in each time slot for sum rate maximization. Simulation results show that the proposed protocol outperforms existing protocols for all signal-to-noise ratios (SNRs). Specifically, we obtain a considerable gain at low SNRs due to the adaptive power allocation and at high SNRs due to the adaptive mode selection.

I. INTRODUCTION

In a bidirectional relay network, two users exchange information via a relay node [1]. Several protocols have been proposed for such a network under the practical half-duplex constraint, i.e., a node cannot transmit and receive at the same time and in the same frequency band. The simplest protocol is the traditional two-way relaying protocol in which the transmission is accomplished in four successive point-to-point phases: user 1-to-relay, relay-to-user 2, user 2-to-relay, and relay-to-user 1. In contrast, the time division broadcast (TDBC) protocol exploits the broadcast capability of the wireless medium and combines the relay-to-user 1 and relay-to-user 2 phases into one phase, the broadcast phase [2]. Thereby, the relay broadcasts a superimposed codeword, carrying information for both user 1 and user 2, such that each user is able to recover its intended information by self-interference cancellation. Another existing protocol is the multiple access broadcast (MABC) protocol in which the user 1-to-relay and user 2-to-relay phases are also combined into one phase, the multiple-access phase [3]. In the multiple-access phase, both user 1 and user 2 simultaneously transmit to the relay which is able to decode both messages. Generally, for the bidirectional relay network without a direct link between user 1 and user 2, six transmission modes are possible: four point-to-point modes (user 1-to-relay, user 2-to-relay, relay-to-user 1, relay-to-user 2), a multiple access mode (both users to the relay), and a broadcast mode (the relay to both users), where the capacity region of each transmission mode is known [4], [5]. Using this knowledge, a significant research effort has been dedicated to obtaining the achievable rate region of the bidirectional relay network [1]-[8]. Specifically, the achievable rates of most existing protocols for two-hop relay transmission are limited by the instantaneous capacity of the weakest link associated with the relay. The reason for this is the fixed schedule of using the transmission modes which is adopted in all existing protocols, and does not exploit the instantaneous channel state information (CSI) of the involved links. For one-way relaying, an adaptive link selection protocol was proposed in [9] where based on the instantaneous CSI, in each time slot, either the source-relay or relay-destination links are selected for transmission. To this end, the relay has to have a buffer for data storage. This strategy was shown to achieve the capacity of the one-way relay channel with fading [10].

Moreover, in fading AWGN channels, power control is necessary for rate maximization. The highest degree of freedom that is offered by power control is obtained for a joint average power constraint for all nodes. Any other power constraint with the same total power budget is more restrictive than the joint power constraint and results in a lower sum rate. Therefore, motivated by the protocols in [9] and [10], our goal is to utilize all available degrees of freedom of the three-node half-duplex bidirectional relay network with fading, via an adaptive mode selection and power allocation policy. In particular, given a joint power budget for all nodes, we find a policy which in each time slot selects the optimal transmission mode from the six possible modes and allocates the optimal powers to the nodes transmitting in the selected mode, such that the sum rate is maximized.

Adaptive mode selection for bidirectional relaying was also considered in [8] and [11]. However, the selection policy in [8] does not use all possible modes, i.e., it only selects from two point-to-point modes and the broadcast mode, and assumes that the transmit powers of all three nodes are fixed and identical. Although the selection policy in [11] considers all possible transmission modes for adaptive mode selection, the transmit powers of the nodes are assumed to be fixed, i.e., power allocation is not possible. Interestingly, mode selection and power allocation are mutually coupled and the modes selected with the protocol in [11] for a given channel are different from the modes selected with the proposed protocol. Power allocation can considerably improve the sum rate by optimally allocating the powers to the nodes based on the
patterns. Simulation results confirm that the proposed protocols can be used as a reference for other low complexity suboptimal protocols. Hence, the sum rate achieved with the proposed protocol can be bounded, as shown in [9], which causes only a small loss in the achieved rate. The delay analysis of the proposed protocol is beyond the scope of the current work and is left for future research.

II. SYSTEM MODEL

In this section, we first describe the channel model. Then, we provide the achievable rates for the six possible transmission modes.

A. Channel Model

We consider a simple network in which user 1 and user 2 exchange information with the help of a relay node as shown in Fig. 1. We assume that there is no direct link between user 1 and user 2, and thus, user 1 and user 2 communicate with each other only through the relay node. We assume that all three nodes in the network are half-duplex. Furthermore, we assume that time is divided into slots of equal length and that each node transmits codewords which span one time slot or a fraction of a time slot as will be explained later. We assume that the user-to-relay and relay-to-user channels are impaired by AWGN with unit variance and block fading, i.e., the channel coefficients are constant during one time slot and change from one time slot to the next. Moreover, in each time slot, the channel coefficients are assumed to be reciprocal such that the user 1-to-relay and the user 2-to-relay channels are identical to the relay-to-user 1 and relay-to-user 2 channels, respectively. Let \( h_1(t) \) and \( h_2(t) \) denote the channel coefficients between user 1 and the relay and between user 2 and the relay in the \( i \)-th time slot, respectively. Furthermore, let \( S_1(t) = |h_1(t)|^2 \) and \( S_2(t) = |h_2(t)|^2 \) denote the squares of the channel coefficient amplitudes in the \( i \)-th time slot. \( S_1(t) \) and \( S_2(t) \) are assumed to be ergodic and stationary random processes with means \( \Omega_1 = E\{S_1\} \) and \( \Omega_2 = E\{S_2\} \), respectively, where \( E\{\cdot\} \) denotes expectation. Since the noise is AWGN, in order to achieve the capacity of each mode, nodes have to transmit Gaussian distributed codewords. Therefore, the transmitted codewords of user 1, user 2, and the relay are comprised of symbols which are Gaussian distributed random variables with variances \( P_1(t) \), \( P_2(t) \), and \( P_r(t) \), respectively, where \( P_j(t) \) is the transmit power of node \( j \in \{1, 2, r\} \) in the \( i \)-th time slot. For ease of notation, we define \( C(x) \triangleq \log_2(1 + x) \). In the following, we describe the transmission modes and their achievable rates.

B. Transmission Modes and Their Achievable Rates

In the considered bidirectional relay network only six transmission modes are possible, cf. Fig. 2. The six possible transmission modes are denoted by \( M_1, \ldots, M_6 \), and \( R_{j,j'}(t) \geq 0 \), \( j, j' \in \{1, 2, r\} \), denotes the transmission rate from node \( j \) to node \( j' \) in the \( i \)-th time slot. Let \( B_1 \) and \( B_2 \) denote two infinite-size buffers at the relay in which the received information from user 1 and user 2 is stored, respectively. Moreover, \( Q_i(t), j \in \{1, 2\}, \) denotes the amount of normalized information in bits/symbol available in buffer \( B_j \) in the \( i \)-th time slot. Using this notation, the transmission modes and their respective rates are presented in the following:

\( M_1 \): User 1 transmits to the relay and user 2 is silent. In this mode, the maximum rate from user 1 to the relay in the \( i \)-th time slot is given by \( R_{1r}(t) = C_{1r}(t) \), where \( C_{1r}(t) = C(P_1(t), S_1(t)) \). The relay decodes this information and stores it in buffer \( B_1 \). Therefore, the amount of information in buffer \( B_1 \) increases to \( Q_1(t) = Q_1(t - 1) + R_{1r}(t) \).

\( M_2 \): User 2 transmits to the relay and user 1 is silent. In this mode, the maximum rate from user 2 to the relay in the \( i \)-th time slot is given by \( R_{2r}(t) = C_{2r}(t) \), where \( C_{2r}(t) = C(P_2(t), S_2(t)) \). The relay decodes this information and stores it in buffer \( B_2 \). Therefore, the amount of information in buffer \( B_2 \) increases to \( Q_2(t) = Q_2(t - 1) + R_{2r}(t) \).

\( M_3 \): Both users 1 and 2 transmit to the relay simultaneously. For this mode, we assume that multiple access transmission is used, see [5]. Thereby, the maximum achievable sum rate in the \( i \)-th time slot is given by \( R_{1r}(t) + R_{2r}(t) = C_{1r}(t) \), where \( C_{1r}(t) = C(P_1(t)S_1(t) + P_2(t)S_2(t)) \). Since user 1 and user 2 transmit independent messages, the sum rate, \( C_{1r}(t) \), can be decomposed into two rates, one from user 1 to the relay and the other one from user 2 to the relay. Moreover, these two capacity rates can be achieved via time sharing and successive interference cancelation. Thereby, in the first \( 0 \leq t(t) \leq 1 \) fraction of the \( i \)-th time slot, the relay first decodes the codeword received from user 2 and considers the signal from user 1 as noise. Then, the relay subtracts the signal received from user 2 from the received signal and decodes the codeword received from user 1. A similar procedure is performed in the remaining \( 1 - t(t) \) fraction of the \( i \)-th time slot but now the relay first decodes the codeword received from user 1 and treats the signal of user 2 as noise, and then decodes the codeword received from user 2. Therefore, for a given \( t(t) \), we decompose \( C_{1r}(t) = C_{1r}(t) = C_{12r}(t) + C_{21r}(t) \) and the maximum rates from users 1 and 2 to the relay in the

1In this paper, we drop time index \( t \) in expectations for notational simplicity.
The relay decodes the information received from user 1 and user 2 and stores it in its buffers $B_1$ and $B_2$, respectively. Therefore, the amounts of information stored in buffer $B_1$ and $B_2$ increase to $Q_1(i) = Q_1(i-1) + R_{1r}(i)$ and $Q_2(i) = Q_2(i-1) + R_{2r}(i)$, respectively.

$\mathcal{M}_4$: The relay transmits the information received from buffer $B_2$ to user 1. Specifically, the relay extracts the information from buffer $B_2(i)$, encodes it into a codeword, and transmits it to user 1. Therefore, the transmission rate from the relay to user 1 in the $i$-th time slot is limited by both the capacity of the relay-to-user 1 channel and the amount of information stored in buffer $B_2$. Thus, the maximum transmission rate from the relay to user 1 is given by $R_{1r}(i) = \min\{C_1(i), Q_2(i-1)\}$, where $C_1(i) = C_2(i) = C(P_2(i), S_2(i))$. Therefore, the amount of information in buffer $B_2$ decreases to $Q_2(i) = Q_2(i-1) - R_{2r}(i)$.

$\mathcal{M}_5$: The relay transmits the information received from buffer $B_2$ to both user 1 and 2 switching places. The maximum transmission rate from the relay to user 2 is given by $R_{2r}(i) = \min\{C_2(i), Q_1(i-1)\}$, where $C_2(i) = C(P_2(i), S_2(i))$ and the amount of information in buffer $B_2$ decreases to $Q_2(i) = Q_2(i-1) - R_{2r}(i)$.

$\mathcal{M}_6$: This mode is identical to $\mathcal{M}_4$ with user 1 and 2 switching places. Specifically, the relay extracts the information intended for user 2 from buffer $B_2(i)$ and the information intended for user 1 from buffer $B_1(i)$. Then, based on the scheme in [2], it constructs a superimposed codeword which contains the information from both users and broadcasts it to both users. Thus, in the $i$-th time slot, the maximum rate from the relay to users 1 and 2 is given by $R_{1r}(i) = \min\{C_1(i), Q_2(i-1)\}$ and $R_{2r}(i) = \min\{C_2(i), Q_1(i-1)\}$, respectively. Therefore, the amounts of information in buffers $B_1$ and $B_2$ decrease to $Q_1(i) = Q_1(i-1) - R_{1r}(i)$ and $Q_2(i) = Q_2(i-1) - R_{2r}(i)$, respectively.

Our aim is to develop an optimal mode selection and power allocation policy which selects one of the six transmission modes, $\mathcal{M}_1, \ldots, \mathcal{M}_6$, and allocates the optimal power to the transmitting nodes of the selected mode such that the average sum rate of both users is maximized. To this end, we introduce six binary variables, $q_k(i) \in \{0, 1\}$, $k = 1, \ldots, 6$, where $q_k(i)$ indicates whether or not transmission mode $\mathcal{M}_k$ is selected in the $i$-th time slot. In particular, $q_k(i) = 1$ if mode $\mathcal{M}_k$ is selected and $q_k(i) = 0$ if it is not selected in the $i$-th time slot. Furthermore, since in each time slot only one of the six transmission modes can be selected, only one of the mode selection variables is equal to one and the others are zero, i.e., $\sum_{k=1}^6 q_k(i) = 1$ holds.

In the proposed framework, we assume that all nodes have full knowledge of the CSI of both links. Thus, based on the CSI and the proposed protocol, each node is able to individually decide which transmission mode is selected and adapt its transmission strategy accordingly.

III. JOINT MODE SELECTION AND POWER ALLOCATION

In this section, we first investigate the achievable average sum rate of the network. Then, we formulate a maximization problem whose solution is the sum rate maximizing protocol.

A. Achievable Average Sum Rate

We assume that user 1 and user 2 always have enough information to send in all time slots and that the number of time slots, $N$, satisfies $N \to \infty$. Therefore, using $q_k(i)$, the user 1-to-relay, user 2-to-relay, relay-to-user 1, and relay-to-user 2 average transmission rates, denoted by $\bar{R}_{1r}$, $\bar{R}_{2r}$, $\bar{R}_{r1}$, and $\bar{R}_{r2}$, respectively, are obtained as

\[
\begin{align*}
\bar{R}_{1r} &= \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} [q_1(i)C_1(i) + q_2(i)C_2(i)] \\
\bar{R}_{2r} &= \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} [q_2(i)C_2(i) + q_2(i)C_2(i)] \\
\bar{R}_{r1} &= \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} [q_1(i) + q_1(i)] \min\{C_1(i), Q_2(i-1)\} \\
\bar{R}_{r2} &= \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} [q_1(i) + q_1(i)] \min\{C_2(i), Q_1(i-1)\}
\end{align*}
\]

The average rate from user 1 to user 2 is the average rate that user 2 receives from the relay, i.e., $\bar{R}_{2r}$. Similarly, the average rate from user 2 to user 1 is the average rate that user 1 receives from the relay, i.e., $\bar{R}_{r1}$. In the following theorem, we introduce a useful condition for the queues in the buffers of the relay leading to the optimal mode selection and power allocation policy.

Theorem 1 (Optimal Queue Condition): The maximum average sum rate, $\bar{R}_{1r} + \bar{R}_{2r}$, for the considered bidirectional relay network is obtained when the queues in the buffers $B_1$ and $B_2$ at the relay are at the edge of non-absorption. More precisely, the following conditions must hold for the maximum sum rate

\[
\begin{align*}
\bar{R}_{1r} &= \frac{1}{N} \sum_{i=1}^{N} [q_1(i) + q_1(i)] C_1(i) \\
\bar{R}_{2r} &= \frac{1}{N} \sum_{i=1}^{N} [q_2(i) + q_2(i)] C_2(i)
\end{align*}
\]

where $\bar{R}_{1r}$ and $\bar{R}_{2r}$ are given by (2a) and (2b), respectively.

Proof: Please refer to [11, Appendix A].

The above lemma indicates the following points: 1) As $N \to \infty$, the effect of queues at the relay is negligible, i.e., the relay has on average enough information to send to both users, 2) The average arriving data to the buffers is equal to the average departing data. Using this theorem, in the following, we derive the optimal transmission mode selection and power allocation policy.

B. Optimal Protocol

The available degrees of freedom in the considered network in each time slot are the mode selection variables, the transmit
powers of the nodes, and the time sharing variable for multiple access. Herein, we formulate an optimization problem which gives the optimal values of $q_k(i)$, $P_j(i)$, and $t(i)$, for $k = 1, \ldots, 6$, $j = 1, 2$, and $\forall i$, such that the average sum rate of the users is maximized. The optimization problem is as follows

$$\begin{align*}
\text{maximize} & \quad R_{1r} + R_{2r} \\
\text{subject to} & \quad R_{1r} = R_{12}r \\
& \quad R_{2r} = R_{21}r \\
& \quad P_1 + P_2 + P_r \leq P_t \\
& \quad \sum_{k=1}^{6} q_k(i) = 1, \forall i \\
& \quad q_k(i)(1 - q_k(i)) = 0, \forall i, k \\
& \quad P_j(i) \geq 0, \forall i, j \\
& \quad 0 \leq t(i) \leq 1, \forall i
\end{align*}$$

(4)

where $P_t$ is the total average power constraint of the nodes and $P_1, P_2$, and $P_r$ denote the average powers consumed by user 1, user 2, and the relay, respectively, and are given by

$$\begin{align*}
P_1 &= \frac{1}{N} \sum_{i=1}^{N} (q_1(i) + q_3(i)) P_1(i) \\
P_2 &= \frac{1}{N} \sum_{i=1}^{N} (q_2(i) + q_3(i)) P_2(i) \\
P_r &= \frac{1}{N} \sum_{i=1}^{N} (q_4(i) + q_5(i) + q_6(i)) P_r(i).
\end{align*}$$

(5a) (5b) (5c)

In the optimization problem given in (4), constraints C1 and C2 are the conditions for sum rate maximization introduced in Theorem 1. Constraints C3 and C6 are the average total transmit power constraint and the power non-negativity constraint, respectively. Moreover, constraints C4 and C5 guarantee that only one of the transmission modes is selected in each time slot, and constraint C7 specifies the acceptable interval for the time sharing variable $t(i)$. Furthermore, we maximize $R_{1r} + R_{2r}$ since, according to Theorem 1 (and constraints C1 and C2), $R_{1r} = R_{12}r$ and $R_{2r} = R_{21}r$ hold.

In the following Theorem, we introduce a protocol which achieves the maximum sum rate.

Theorem 2 (Mode Selection and Power Allocation Policy): Assuming $N \to \infty$, the optimal mode selection and power allocation policy which maximizes the sum rate of the considered three-node half-duplex bidirectional relay network with AWGN and block fading is given by

$$q_k(i) = \begin{cases} 
1, & \text{if } k^* = \arg \max_{k=1,2,3,6} \{\Lambda_k(i)\} \\
0, & \text{otherwise}
\end{cases}$$

(6)

where $\Lambda_k(i)$ is referred to as selection metric and is given by

$$\begin{align*}
\Lambda_1(i) &= (1 - \mu_1)C_{1r}(i) - \gamma P_1(i) \bigg|_{P_1(i) = P_1^{M1}(i)} \\
\Lambda_2(i) &= (1 - \mu_2)C_{2r}(i) - \gamma P_2(i) \bigg|_{P_2(i) = P_2^{M2}(i)} \\
\Lambda_3(i) &= (1 - \mu_1)C_{12r}(i) + (1 - \mu_2)C_{21r}(i) - \gamma(P_1(i) + P_2(i)) \bigg|_{P_1(i) = P_1^{M3}(i), P_2(i) = P_2^{M3}(i)} \\
\Lambda_4(i) &= \mu_1 C_{1r2}(i) + \mu_2 C_{2r1}(i) - \gamma P_r(i) \bigg|_{P_r(i) = P_r^{M4}(i)}
\end{align*}$$

(7a) (7b) (7c) (7d)

where $P_j^{M_k}(i)$ denotes the optimal transmit power of node $j$ for transmission mode $\mathcal{M}_k$, in the $i$-th time slot and is given by

$$\begin{align*}
P_1^{M1}(i) &= \left[1 - \frac{\mu_1}{\gamma \ln 2} \right] + \\
P_2^{M2}(i) &= \left[1 - \frac{\mu_2}{\gamma \ln 2} \right] + \\
P_1^{M3}(i) &= \left[1 - \frac{\mu_1}{\gamma \ln 2} - \frac{\mu_2}{\gamma \ln 2} + \frac{1}{\ln 2} \right] + , \text{if } \Omega_1 \geq \Omega_2 \\
P_2^{M3}(i) &= \left[1 - \frac{\mu_1}{\gamma \ln 2} - \frac{\mu_2}{\gamma \ln 2} + \frac{1}{\ln 2} \right] + , \text{otherwise} \\
P_r^{M4}(i) &= \left[-b + \sqrt{b^2 - 4ac} \right] 2a
\end{align*}$$

(8a) (8b) (8c) (8d) (8e)

where $[x]^+ = \max\{x, 0\}$, $a = \gamma \ln 2 \times (S_1(i) + S_2(i)) - \mu_1 S_2(i) S_2(i), b = \gamma \ln 2 \times (S_1(i) + S_2(i)) - (\mu_1 + \mu_2) S_1(i) S_2(i)$, and $c = \gamma \ln 2 - \mu_1 S_1(i) - \mu_2 S_2(i)$.

The thresholds $\mu_1$ and $\mu_2$ are chosen such that constraints C1 and C2 in (6) hold and threshold $\gamma$ is chosen such that the total average transmit power satisfies C3 in (4). The optimal value of $t(i)$ in C12r(i) and C21r(i) is given by

$$t^*(i) = \begin{cases} 
0, & \Omega_1 \geq \Omega_2 \\
1, & \Omega_1 < \Omega_2
\end{cases}$$

(9)

Proof: Please refer to Appendix A.

We note that the optimal solution utilizes neither modes $\mathcal{M}_4$ and $\mathcal{M}_5$, nor time sharing for any channel statistics and channel realizations.

Remark 1: The mode selection metric $\Lambda_k(i)$ introduced in (7) has two parts. The first part is the instantaneous capacity of mode $\mathcal{M}_k$, and the second part is the allocated power with negative sign. The capacity and the power terms are linked via thresholds $\mu_1$ and $\mu_2$ and $\gamma$. We note that these thresholds depend only on the channel statistics of the channels. Hence, these thresholds can be obtained offline and used as long as the channel statistics remain unchanged. To find the optimal values for the thresholds $\mu_1$, $\mu_2$, and $\gamma$, we need a three-dimensional search, where $\mu_1, \mu_2 \in \{0 \cup 1\}$ and $\gamma > 0$.

Remark 2: Adaptive mode selection for bidirectional relay networks under the assumption that the powers of the nodes are fixed is considered in [11]. Based on the average and instantaneous qualities of the links, all of the six possible transmission modes are selected in the protocol in [11]. However, in the proposed protocol, modes $\mathcal{M}_4$ and $\mathcal{M}_5$ are not selected at all. Moreover, the protocol in [11] utilizes a coin flip for implementation. Therefore, a central node must decide which transmission mode is selected in the next time slot. However, in the proposed protocol, all nodes can find the
optimal mode and powers based on the full CSI.

IV. SIMULATION RESULTS

In this section, we evaluate the average sum rate achievable with the proposed protocol in the considered bidirectional relay network in Rayleigh fading. Thus, channel gains $S_1(i)$ and $S_2(i)$ follow exponential distributions with means $\Omega_1$ and $\Omega_2$, respectively. All of the presented results were obtained for $\Omega_2 = 1$ and $N = 10^4$ time slots.

In Fig. 3, we illustrate the maximum achievable sum rate obtained with the proposed protocol as a function of the total average transmit power $P_t$. In this figure, to have a better resolution for the sum rate at low and high $P_t$, we show the sum rate for both log scale and linear scale $y$-axes, respectively. The lines without markers in Fig. 3 represent the achieved sum rates with the proposed protocol for $\Omega_1 = 1, 2, 5$. We observe that as the quality of the user 1-to-relay link increases (i.e., $\Omega_1$ increases), the sum rate increases too. However, for large $\Omega_1$, the bottleneck link is the relay-to-user 2 link, and since it is fixed, the sum rate saturates.

As performance benchmarks, we consider in Fig. 3 the sum rates of the TDBC protocol with and without power allocation (PA) [2] and the buffer-aided protocols presented in [8] and [11], respectively. For clarity, for the benchmark schemes, we only show the sum rates for $\Omega_1 = 1, 2, 5$. For the TDBC protocol without power allocation and the protocol in [8], all nodes transmit with equal powers, i.e., $P_1 = P_2 = P_r = P_t$. For the buffer-aided protocol in [11], we adopt $P_1 = P_2 = P_r = P$ and $P$ is chosen such that the average total power consumed by all nodes transmit with equal powers, i.e., $P_1 = P_2 = P_r = P_t$. For the buffer-aided protocol in [11], we adopt $P_1 = P_2 = P_r = P$ and $P$ is chosen such that the average total power consumed by all nodes transmit with equal powers, i.e., $P_1 = P_2 = P_r = P_t$. The protocol which achieves the maximum sum rate jointly optimizes the selection of the transmission mode and the transmit powers of the nodes. The proposed optimal mode selection and power allocation protocol requires the instantaneous CSI of the involved links in each time slot and their long-term statistics. Simulation results confirmed that the proposed selection policy outperforms existing protocols in terms of average sum rate.

V. CONCLUSION

We have derived the maximum sum rate of the three-node half-duplex bidirectional buffer-aided relay network with fading links. The protocol which achieves the maximum sum rate jointly optimizes the selection of the transmission mode and the transmit powers of the nodes. The proposed optimal mode selection and power allocation protocol requires the instantaneous CSI of the involved links in each time slot and their long-term statistics. Simulation results confirmed that the proposed selection policy outperforms existing protocols in terms of average sum rate.

APPENDIX A

PROOF OF THEOREM 2 (MODE SELECTION PROTOCOL)

Because of space constraints, a detailed proof of Theorem 2 is provided in [12, Appendix A] which is an extended version of the submitted paper. Herein, we only highlight the main steps in obtaining the solution. We first relax the binary condition for $q_k(i)$, i.e., $q_k(i) = 0, 1$, to $0 \leq q_k(i) \leq 1$, and in [12, Appendix B] we prove that this relaxation does not affect the maximum average sum rate. In the following, we investigate the Karush-Kuhn-Tucker (KKT) necessary conditions for the relaxed optimization problem.

To simplify the usage of the KKT conditions, we change the maximization of $R_{tr} + R_{2r}$ to the minimization of $-(R_{tr} + R_{2r})$ and, without loss of generality, we write all inequalities in the form $f(x) \leq 0$. The Lagrangian function for the relaxed optimization problem is given by

$$
\mathcal{L}(q_k(i), t(i), \mu_1, \gamma, \lambda(i), \alpha_k(i), \beta_k(i), \phi_1(i)) = -(R_{tr} + R_{2r}) + \mu_1(R_{tr} - \epsilon) + \mu_2(R_{2r} - \epsilon)
+ \gamma (P_1 + P_2 + P_t - P_1) + \sum_{i=1}^{N} \lambda(i) \left(\sum_{k=1}^{E} q_k(i) - 1\right)
+ \sum_{i=1}^{N} \sum_{k=1}^{E} \alpha_k(i) q_k(i) - 1
- \sum_{i=1}^{N} \sum_{k=1}^{E} \beta_k(i) q_k(i)
+ \sum_{i=1}^{N} \phi_1(i) t(i) - 1
- \sum_{i=1}^{N} \phi_0(i) t(i)
$$

(10)

We note that although the protocol in [8] outperforms the protocol in [11] if the sum rate is plotted as a function of total transmit power as is done in Fig. 3, the protocol in [11] is optimal for given fixed node transmit powers.

[2] We note that although the protocol in [8] outperforms the protocol in [11] if the sum rate is plotted as a function of total transmit power as is done in Fig. 3, the protocol in [11] is optimal for given fixed node transmit powers.
where $\mu_1, \mu_2, \gamma, \lambda(i)$, and $\nu_2(i)$ are the Lagrange multipliers corresponding to constraints C1, C2, C3, C4, and C6 in (4), respectively. Furthermore, $\alpha_k(i)$ and $\beta_k(i)$ are the Lagrange multipliers corresponding to the upper and lower bounds of constraint $0 \leq q_k(i) \leq 1$, respectively, and $\phi_1(i)$, and $\phi_2(i)$ are the Lagrange multipliers corresponding to the upper and lower bounds of constraint $0 \leq t(i) \leq 1$, respectively.

The KKT conditions include the following: 1) stationary condition: the differentiation of the Lagrangian function with respect to the primal variables, $q_k(i)$, $P_j(i)$, and $t(i)$ for $\forall i, j, k$, is zero for the optimal solution, 2) primal feasibility condition: the optimal solution has to satisfy the constraints of the relaxed optimization problem, 3) dual feasibility condition: the Lagrange multipliers for the inequality constraints have to be non-negative, and 4) complementary slackness: if an inequality does not hold with equality for the optimal solution, the corresponding Lagrange multiplier is zero.

Next, we calculate the derivative of the Lagrangian function with respect to $q_k(i)$. This leads to
\begin{align*}
-\frac{1}{N}(1-\mu_1)C_{i},i + \lambda(i) + \alpha_k(i) - \beta_k(i) + \frac{1}{N} \gamma P_i(i) = 0 \quad (11a) \\
-\frac{1}{N}(1-\mu_2)C_{2,i} + \lambda(i) + \alpha_k(i) - \beta_k(i) + \frac{1}{N} \gamma P_i(i) = 0 \quad (11b) \\
-\frac{1}{N}(1-\mu_1)C_{12,i} + \lambda(i) + \alpha_k(i) - \beta_k(i) + \frac{1}{N} \gamma P_i(i) = 0 \quad (11c) \end{align*}

where (11a)-(11f) are the derivatives of the Lagrangian function with respect to $q_1(i), \ldots, q_6(i)$, respectively. Without loss of generality, we first obtain the necessary condition for $q_1(i) = 1$ and then generalize the result to $q_k(i) = 1$ for $k = 2, \ldots, 6$. If $q_1(i) = 1$, then, from constraint C4 in (4), the other selection variables are zero, i.e., $q_k(i) = 0$, for $k = 2, \ldots, 6$. Furthermore, from the complementary slackness condition, we obtain $\alpha_k(i) = 0$, for $k = 2, \ldots, 6$ and $\beta_k(i) = 0$.

Then, by substituting these values into (11), we obtain
\begin{align*}
\lambda(i) + \alpha_1(i) = (1-\mu_1)C_{1,i} - \gamma P_i(i) \equiv \Lambda_1(i) \quad (12a) \\
\lambda(i) + \alpha_2(i) = (1-\mu_2)C_{2,i} - \gamma P_i(i) \equiv \Lambda_2(i) \quad (12b) \\
\lambda(i) - \beta_1(i) = (1-\mu_1)C_{12,i} + (1-\mu_2)C_{21,i} - \gamma P_i(i) \equiv \Lambda_3(i) \quad (12c) \\
\lambda(i) - \beta_2(i) = \mu_2C_{11,i} - \gamma P_i(i) \equiv \Lambda_4(i) \quad (12d) \\
\lambda(i) - \beta_3(i) = \mu_1C_{22,i} - \gamma P_i(i) \equiv \Lambda_5(i) \quad (12e) \\
\lambda(i) - \beta_4(i) = \mu_2C_{21,i} + \mu_2C_{11,i} - \gamma P_i(i) \equiv \Lambda_6(i) \quad (12f) \end{align*}

where $\Lambda_k(i)$ is referred to as selection metric. By subtracting (12a) from the rest of the equations in (12), we obtain
\begin{equation}
\Lambda_1(i) - \Lambda_k(i) = \alpha_k(i) + \beta_k(i), \quad k = 2, 3, 4, 5, 6. \quad (13)\end{equation}

From the dual feasibility conditions, we have $\alpha_k(i), \beta_k(i) \geq 0$.

By inserting $\alpha_k(i), \beta_k(i) \geq 0$ in (13), we obtain the necessary condition for $q_k^*(i) = 1$ as
\begin{equation}
\Lambda_1(i) \geq \max \{\Lambda_2(i), \Lambda_3(i), \Lambda_4(i), \Lambda_5(i), \Lambda_6(i)\}. \quad (14)\end{equation}

Repeating the same procedure for $q_k^*(i) = 1, k = 2, \ldots, 6$, we obtain a necessary condition for selecting transmission mode $M_k$- in the $i$-th time slot as follows
\begin{equation}
\Lambda_{k^*}(i) \geq \max_{k \in \{1, \ldots, 6\}} \{\Lambda_k(i)\}, \quad (15)\end{equation}

where the Lagrange multipliers $\mu_1, \mu_2$, and $\gamma$ are chosen such that C1, C2, and C3 in (4) hold. It is proved in [12, Appendix B] that the condition in (15) is necessary and sufficient. Moreover, in [12, Appendix A], it is proved that $\Lambda_k(\geq \Lambda_4(i)$ and $\Lambda_k(\geq \Lambda_5)$ hold for $\forall i$ and the probability that the inequalities hold with equality is zero. Therefore, modes $M_4$ and $M_5$ are not selected in the optimal selection policy.

Furthermore, in [12, Appendix A], we also calculate the derivatives of the Lagrangian function with respect to $P_j(i)$ and $t(i)$. Then, using the dual feasibility condition, complementary slackness, and by considering the constraints of (4), we find the optimal values for $P_j(i)$ and $t(i)$. Therefore, by obtaining the optimal optimization variables $q_k(i), P_j(i)$, and $t(i)$ and thresholds $\mu_1, \mu_2$, and $\gamma$, we can construct the protocol in Theorem 2. This completes the proof.

**REFERENCES**


