Achievable Rate Region of the Bidirectional Buffer-Aided Relay Channel with Block Fading

Vahid Jamali, Nikola Zlatanov, Student Member, IEEE, Aissa Ikhlef, Member, IEEE, and Robert Schober, Fellow, IEEE

Abstract—The bidirectional relay channel, in which two users communicate with each other through a relay node, is a simple but fundamental and practical network architecture. In this paper, we consider the block fading bidirectional relay channel with a decode-and-forward relay and propose efficient transmission strategies that exploit the block fading property of the channel. We assume that a direct link between the two users is not present and consider two transmission modes: the multiple-access mode (both users transmit to the relay) and the broadcast mode (the relay transmits to both users). Most existing relaying protocols assume a fixed schedule for using these transmission modes. In contrast, we abandon the restriction of having a fixed and predefined schedule and propose to optimize the selection of the transmission modes and the associated transmission rates based on the instantaneous channel state information (CSI) of the involved links. Thereby, we consider two different types of transmit power constraints: i) a fixed transmit power for each node, and ii) a per-node long-term power constraint. To enable the use of a non-predefined schedule for transmission mode selection, the relay has to be equipped with two buffers for storage of the information received from both users. We develop new relaying protocols based on adaptive mode selection and provide the corresponding achievable long-term rate regions. In particular, based on the CSI of the involved links, the optimal transmission mode as well as the optimal transmission rates and/or the transmit powers of the nodes are chosen in each time slot to maximize the weighted sum rate of both users. By varying the weights assigned to the users, the boundary surface of the achievable long-term rate region of the proposed protocol can be obtained. In addition, we discuss and address two practical challenges for the implementation of the proposed protocols, namely the availability of the knowledge of the channel statistics required for the implementation of the optimal protocols, and the increase of the end-to-end delay due to the data buffering. Numerical results confirm the superiority of the proposed buffer-aided protocols compared to existing bidirectional relaying protocols.

Index Terms—Bidirectional transmission, rate region, adaptive mode selection, power allocation, buffer-aided relaying.

I. INTRODUCTION

In the bidirectional relay channel, two users exchange information via a relay node [1]. The communication in several practical applications such as satellite communication and cellular communication with a base station can be modeled via the bidirectional relay channel. In this paper, we focus on decode-and-forward processing at the relay and assume that a direct link between user 1 and user 2 does not exist or is not exploited which is a practical assumption when the users are located far from each other. Different protocols have been proposed for this channel under the practical half-duplex constraint, i.e., a node cannot transmit and receive at the same time and in the same frequency band [2]. The simplest such protocol is the traditional two-way relaying protocol in which the transmission is accomplished via four successive point-to-point transmission phases, namely the user 1-to-relay, relay-to-user 2, user 2-to-relay, and relay-to-user 1 transmission phases. An alternative to this is the time division broadcast (TDBC) protocol which exploits the broadcast nature of the wireless medium and utilizes a relay-to-users broadcast transmission phase along with the user 1-to-relay and user 2-to-relay point-to-point transmission phases [3], [4]. Another existing protocol is the multiple-access broadcast (MABC) protocol in which a users-to-relay multiple-access transmission phase is used in addition to a relay-to-users broadcast transmission phase [5], [6]. For a comprehensive overview of existing protocols for the bidirectional relay channel, we refer to [2], [7], and for the case of decode-and-forward processing at the relay, we refer to [9], [10].

If the nodes in the network are mobile, environmental effects such as path loss, shadowing, and multipath fading cause time varying fluctuations of the received signals at the nodes. A general strategy to combat the time-varying nature of the channel is to utilize dynamic resource allocation such as adaptive power allocation and/or adaptive rate and coding scheme selection at the transmitter based on instantaneous channel state information (CSI) [11], [12]. For instance, Goldsmith and Varaiya [11] obtained the ergodic capacity of a point-to-point Gaussian channel with fading by using water-filling power allocation in time and adapting the rate to the instantaneous channel capacity. For general (multi-hop) relay channels, the instantaneous capacity is limited by the minimum of the capacities of the involved links/hops (i.e., the bottleneck capacity). To overcome this limitation, adaptive power alloca-

\footnote{We note that for the bidirectional relay channel without a direct link, the rate regions of the traditional two-way protocol and the TDBC protocol are subsets of the rate region of the MABC protocol.}

\footnote{We note that there are protocols in the literature which generalize the MABC and TDBC protocols to the case when the direct link between the users is exploited [2], [7], [8]. However, if the direct link is not present in the channel model, the achievable rates of the protocols in [2], [7], [8] coincide with those of the TDBC and MABC protocols.}
tion can be used to allocate more power to the weaker channels such that the bottleneck capacity increases [13]. Moreover, for one-way relaying, in [14], a buffer was used at the relay to enable the relay to receive for a fixed number of time slots before retransmitting the information to the destination. The buffering capability allows the protocol in [14] to overcome the limitation imposed by the instantaneous bottleneck capacity. In particular, assuming an infinite-size buffer, the achievable rate depends only on the ergodic capacities of the links and not on their instantaneous capacities. Building upon the idea of using relays with buffers, recently, adaptive link selection was proposed for one-way relaying in [15]. Thereby, based on the instantaneous CSI, in each time slot, either the source-to-relay link or the relay-to-destination link was selected for transmission. Buffer-aided relaying has been also considered for other network architectures in the literature, e.g., the two-way relay network [16]–[18], the multihop relay network [19], and the diamond relay network [20], [21].

For the bidirectional relay channel, in this paper, we develop protocols which employ two transmission modes: the multiple-access mode (both users transmit to the relay) and the broadcast mode (the relay transmits to both users). Motivated by the performance gains reported in [15], we extend the idea of adaptive link selection from one-way relaying to bidirectional relaying where, based on the instantaneous CSI, the optimal transmission mode is selected in each time slot. In order to be able to apply adaptive mode selection for the bidirectional relay channel, the relay has to be equipped with two buffers for storage of the information received from each of the users. The proposed protocols adopt the optimal transmission mode and the corresponding optimal transmission rates based on the instantaneous CSI of the involved links in each time slot such that the long-term weighted sum rate of both users is maximized. By varying the weights assigned to the users, the boundary surface of the achievable long-term rate region of the proposed protocol can be obtained. Moreover, depending on the type of power constraint imposed, the powers of the nodes might also be optimized. In this paper, we consider two types of transmit power constraints: i) a fixed transmit power for each node, and ii) a per-node long-term power constraint. While, for a given power budget, the long-term power constraint leads to a larger achievable rate region than the fixed node transmit powers, the fixed transmit powers are preferable from an implementation point of view, e.g., the complexity of calculating the optimal powers is avoided and the transmitters can be simpler as the transmit signals fluctuate less, facilitating the application of low-cost power amplifiers. Because of the buffers at the relay, the transmission from the users to the relay is not only limited by the capacity of the multiple-access channel but also by the amount of information that can be stored in the buffers. Similarly, the transmission from the relay to the users is not only limited by the capacity of the broadcast channel but also by the amount of information that is stored in the queues of the buffers. Neglecting the limiting effect of the queues on the transmission rates, we first obtain outer bounds for the achievable long-term rate regions of the proposed protocols. Subsequently, these outer bounds are proved to be achievable by the proposed protocols if the relay is equipped with infinite-size buffers.

We also discuss and address two practical challenges for the implementation of the proposed protocols, namely the availability of the knowledge of the channel statistics and delay-constrained transmission. In particular, the proposed protocols are obtained assuming that the statistics of the involved channels are known a priori. For the case when the channel statistics are not known a priori, we propose an adaptive algorithm based on tools from stochastic optimization [22]–[24] which requires only knowledge of the history of the previous channel realizations. Moreover, the assumption of infinite-size buffers at the relay for the optimal protocol results in delay-unconstrained transmission, and leads to performance upper bound in terms of the achievable rate region for the delay-constrained case. For the case of delay-constrained transmission, we propose a heuristic but efficient modification of the proposed protocols which is based on the use of finite-size buffers at the relay and limits the end-to-end delay. Numerical results reveal that, for the considered sets of parameters, even in cases where only a small delay is permitted, the protocols proposed for delay-constrained transmission can approach the performance upper bound provided by the delay-unconstrained case. Furthermore, the proposed protocols outperform all available protocols [2], [4], [5], [7], [9], [10] by a considerable margin.

A buffer-aided protocol for bidirectional relaying was also proposed in [18]. However, the protocol in [18] 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. Moreover, this protocol was designed for long-term sum rate maximization. Therefore, the rate pair achievable with the protocol in [18] is a point which lies in the interior of the achievable rate region of the protocols proposed in this paper.

The remainder of this paper is organized as follows. In Section II, the system model and the problem formulation are presented. In Section III, the optimal protocols for both considered power constraints are provided. In Section IV, some challenges for the implementation of the proposed protocols are discussed and addressed. Numerical results are provided in Section V, and conclusions are drawn in Section VI.

Notations: We use the following notations throughout this paper: $E\{\cdot\}$ denotes expectation; $|\cdot|$ represents the cardinality of a set and the absolute value of a scalar; $\wedge$ and $\vee$ denote the logical “and” and “or” operators, respectively. Moreover, bold capital and small letters are used to denote vectors. Furthermore, $a = [a_i]$ denotes a vector with elements $a_i$, $\forall i$, and $a \geq 0$ denotes a vector with non-negative elements, i.e., $a_i \geq 0$, $\forall i$. For notational convenience, we use the definitions $C(x) \triangleq \log_2(1 + x)$, $[x]^+ \triangleq \max\{0, x\}$, and $[x]_a^b \triangleq \min\{\max\{a, x\}, b\}$, where $a \leq b$.

II. System Model and Problem Formulation

In this section, we introduce the considered channel model, review the achievable rates of the multiple-access and broadcast modes in each time slot, and discuss the variables available for optimization. Furthermore, we formulate the weighted
sum rate maximization problem for characterization of the achievable long-term rate regions of the developed protocols and provide outer bounds for these rate regions.

A. Channel Model

We consider the bidirectional relay channel in which user 1 and user 2 exchange information with the help of a decode-and-forward 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 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 indexed by \( i = 1, \ldots, N \) and that each node transmits codewords which span one time slot. We assume that the user-to-relay and relay-to-user channels are impaired by additive white Gaussian noise (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. The channel reciprocity assumption is valid for time-division-duplex (TDD) systems where the user-to-relay and relay-to-user links utilize the same frequency band. The received codewords for the considered channel can be modelled as

\[
Y_1(i) = h_1(i)X_r(i) + Z_1(i), \quad \text{if relay transmits} \tag{1a}
\]
\[
Y_2(i) = h_2(i)X_r(i) + Z_2(i), \quad \text{if relay transmits} \tag{1b}
\]
\[
Y_r(i) = h_1(i)X_1(i) + h_2(i)X_2(i) + Z_r(i), \quad \text{otherwise,} \tag{1c}
\]

where \( X_j(i), Y_j(i), \) and \( Z_j(i), j \in \{1,2,r\} \), denote the transmitted codeword of node \( j \), the received codeword at node \( j \), and the noise at node \( j \) in the \( i \)-th time slot, respectively. We assume that the noises at the nodes are mutually independent and independent from the transmitted codewords. Furthermore, \( h_1(i) \) and \( h_2(i) \) denote the complex-valued channel coefficients between user 1 and the relay and between user 2 and the relay in the \( i \)-th time slot, respectively. The squares of the channel coefficient amplitudes in the \( i \)-th time slot are denoted by \( s_1(i) = |h_1(i)|^2 \) and \( s_2(i) = |h_2(i)|^2 \).

B. Transmission Modes and Their Achievable Rates

For the bidirectional relay channel, in this paper, we consider two transmission modes: the multiple-access mode (both users transmit to the relay) and the broadcast mode (the relay transmits to both users). In particular, the multiple-access mode constitutes the classical multiple-access channel [1] but the broadcast mode creates a special broadcast channel where the receivers have side information about the message intended for the other user [3]. Let \( R_{jj'}(s) \geq 0, j,j' \in \{1r, 2r, r1, r2\} \), denote the transmission rate from node \( j \) to node \( j' \) in the \( i \)-th time slot. Let \( B_1 \) and \( B_2 \) denote two buffers of sizes \( Q_{1\text{max}} \) and \( Q_{2\text{max}} \) at the relay in which the information received from user 1 and user 2 is stored, respectively. Moreover, \( Q_j(i), j \in \{1,2\} \), denotes the amount of normalized information in bits/symbol available in buffer \( B_j \) at the end of the \( i \)-th time slot. Using these notations, the coding schemes, the
transmission rate constraints, and the dynamics of the queues at the buffers for both transmission modes are presented in the following:

**Multiple-access mode:** Users 1 and 2 encode messages $W_{12}(i)$ and $W_{21}(i)$ into codewords $X_1(i)$ and $X_2(i)$, respectively, and transmit them simultaneously to the relay. The relay receives $Y_r(i)$ as shown in (1c) and employs successive decoding [1]. Thereby, the relay decodes the information received from users 1 and users 2 to $\hat{W}_{12}(i)$ and $\hat{W}_{21}(i)$ and stores it in buffer $B_1$ and $B_2$, respectively. For this mode, the transmission rates of the users in each time slot are limited by the capacity region of the multiple-access channel [1], see Fig. 2 a), and the space available in buffers $B_1$ and $B_2$ to store information. Therefore, for a successful transmission in the $i$-th time slot, i.e., $\hat{W}_{12}(i) = W_{12}(i)$ and $\hat{W}_{21}(i) = W_{21}(i)$, the transmission rates from user 1 and user 2 to the relay must satisfy

$$R_{1r}(s) \leq \min\{C(P_1(s)s_1), Q_1^{max} - Q_1(i-1)\}$$  \hspace{1em} (2a)
$$R_{2r}(s) \leq \min\{C(P_2(s)s_2), Q_2^{max} - Q_2(i-1)\}$$  \hspace{1em} (2b)
$$R_{1r}(s) + R_{2r}(s) \leq C(P_1(s)s_1 + P_2(s)s_2).$$  \hspace{1em} (2c)

Moreover, after the relay has received the messages transmitted by user 1 and user 2 in the $i$-th time slot, the amounts of information in buffers $B_1$ and $B_2$ increase to $Q_1(i) = Q_1(i-1) + R_{1r}(s)$ and $Q_2(i) = Q_2(i-1) + R_{2r}(s)$, respectively.

**Broadcast mode:** For this mode, the relay extracts the information intended for user 2, i.e., message $\hat{W}_{12}(i)$, from buffer $B_1$ and the information intended for user 1, i.e., message $\hat{W}_{21}(i)$, from buffer $B_2$. Then, based on the scheme in [3], it constructs superimposed codeword $X_r(i)$ which contains the information of both users and broadcasts it to the users. Therefore, user 1 and user 2 receive $Y_1(i)$ and $Y_2(i)$ as shown in (1a) and (1b), and using the side information $W_{12}$ and $W_{21}$, decode them to $\hat{W}_{21}(i)$ and $\hat{W}_{12}(i)$, respectively. For this mode, the transmission rates from the relay to the users in the $i$-th time slot are limited by both the capacity region of the broadcast channel with side information [3], see Fig. 2 b), and the amount of information stored in buffers $B_1$ and $B_2$. Thus, for a successful transmission in the $i$-th time slot, i.e., $\hat{W}_{12}(i) = W_{12}(i)$ and $\hat{W}_{21}(i) = W_{21}(i)$, the transmission rates from the relay to users 1 and 2 must satisfy

$$R_{1r}(s) = \min\{C(P_r(s)s_1), Q_2(i-1)\}$$  \hspace{1em} (3a)
$$R_{2r}(s) = \min\{C(P_r(s)s_2), Q_1(i-1)\}. \hspace{1em} (3b)$$

After the transmission in the $i$-th time slot, the amounts of information in buffers $B_1$ and $B_2$ decrease to $Q_1(i) = Q_1(i-1) - R_{1r}(s)$ and $Q_2(i) = Q_2(i-1) - R_{2r}(s)$, respectively.

### C. Optimization Variables

Most existing protocols for bidirectional relaying assume a fixed schedule for choosing the available transmission modes. As the main contribution of this paper, we abandon the assumption of using a fixed and predefined schedule for choosing the transmission modes. In particular, the proposed protocol optimizes: \textit{i)} the transmission mode selection in each time slot, \textit{ii)} the transmission rates of the transmitting nodes, and \textit{iii)} the transmit powers of the transmitting nodes given the adopted power constraint for the maximization of the achievable rate region.

For transmission mode selection, we introduce binary variables $q_k(s) \in \{0,1\}, k \in \{MAC,BC\}$ where $q_{MAC}(s) = 1$ if the multiple-access mode is selected in the $i$-th time slot and $q_{MAC}(s) = 0$ if it is not selected. Similarly, $q_{BC}(s) = 1$ if the broadcast mode is selected in the $i$-th time slot and $q_{BC}(s) = 0$ if it is not selected. Furthermore, since in each time slot only one of the multiple-access and broadcast modes can be selected, only one of the mode selection variables is equal to one and the other one is zero, i.e., $q_{MAC}(s) + q_{BC}(s) = 1, \forall s$ holds.

The transmit powers of the nodes can also be optimized for a given power constraint. Here, we consider two different types of transmit power constraints: \textit{i)} a fixed transmit power for each node, and \textit{ii)} a per-node long-term power constraint. Under the former power constraint, the transmit power of each node is fixed regardless of the fading state and a priori given, i.e., $P_j = P_j^0, \forall s$. In contrast, under the latter power constraint, the transmit power of the nodes depends on the fading state but the average consumed power of each node is limited, i.e.,

$$\bar{P}_j = E(P_j(s)) \leq P_j^{max}, \hspace{1em} j \in \{1,2,r\},$$

where $P_j^{max}, j \in \{1,2,r\}$, is the average power budget of node $j$. We note that the optimal transmission strategy fundamentally depends on the adopted power constraint, see Section III. This motivates the consideration of both power constraints in this paper.

### D. Long-Term Rate Region as the Solution of an Optimization Problem

In this paper, we focus on the long-term achievable rate region, averaged over all fading states, of the half-duplex bidirectional relay channel. The long-term achievable rate region is denoted by $R$, and the average transmission rates from user 1 to user 2 and from user 2 to user 1 are denoted by $R_{12}$ and $R_{21}$, respectively. 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 $N \rightarrow \infty$. Moreover, the user 1-to-relay, user 2-to-relay, relay-to-user 1, and relay-to-user 2 average transmission rates are denoted by $R_{1r} = E\{q_{MAC}(s)R_{1r}(s)\}$, $R_{2r} = E\{q_{MAC}(s)R_{2r}(s)\}$, $R_{1t} = E\{q_{BC}(s)R_{1t}(s)\}$, and $R_{2t} = E\{q_{BC}(s)R_{2t}(s)\}$, respectively, where transmission rate pair $(R_{1r}(s), R_{2r}(s))$ has to satisfy (2) and transmission rate pair $(R_{1t}(s), R_{2t}(s))$ has to satisfy (3). Furthermore, the average rate from user 1 to user 2 is the average rate that user 1 receives from the relay, i.e., $R_{12} = R_{2r}$. Similarly, the average rate from user 2 to user 1 is the average rate that user 1 receives from the relay, i.e., $R_{21} = R_{1r}$. To obtain the rate region of the proposed protocols, we first define the boundary surface of the rate region.

**Definition 1:** The boundary surface, $R^{\text{bound}}$, of rate region $R$ is the set of all points $(R_{12}, R_{21}) \in R$ such that if one of the rates is fixed and the other rate is increased, the resulting new point is no longer in $R$, i.e.,

$$R^{\text{bound}} = \left\{ (R_{12}, R_{21}) \in R \mid (R_{12}, R_{21} + \epsilon, R_{21}) \notin R \right. \ \wedge \left. (R_{12}, R_{21} + \epsilon, R_{21}) \notin R, \ \forall \epsilon > 0 \right\}. \quad (5)$$

Note that a transmission strategy that can achieve all points on the boundary surface of the achievable rate region can also achieve all the points inside the achievable rate region. In particular, all other points of $R$ can be achieved by deliberately decreasing the average transmission rates of user 1 and/or user 2. Hence, the boundary surface specifies the set of optimal operating points of the rate region.

In the following, we use the time sharing argument to introduce an alternative representation of the boundary surface of the rate region. In particular, the time sharing argument indicates that if rate pairs $(R_{12}, R_{21})$ and $(R_{12}', R_{21}')$ are achievable, rate pair $(\theta R_{12} + (1-\theta)R_{21}, \theta R_{21} + (1-\theta)R_{21}')$ is also achievable for any $\theta \in [0, 1]$. One direct result of the time sharing argument is that the rate region is convex [25]. Hence, the boundary surface of the rate region can be characterized by the closure of all points $(R_{12}, R_{21}) \in R^{\text{bound}}$ such that $(R_{12}, R_{21})$ is obtained by maximizing the weighted long-term sum rate [7], [12], [24]

$$\begin{align*}
\max_{F} & \quad \eta \bar{R}_{12} + (1 - \eta) \bar{R}_{21} \\
\text{subject to} & \quad C1 : R_{1r} = R_{2r} \\
& \quad C2 : R_{2r} = \bar{R}_{1r} \quad (6)
\end{align*}$$

where $F$ is the set of the optimization variables. For the problem at hand, in general, $F$ is comprised of the following variables. i) The mode selection variable $q = [q_k(s)] \in Q \ \forall s, k$, where $Q$ is the feasible set which imposes constraints $q_k(s) \in [0, 1]$, $k \in \{\text{MAC, BC}\}$ and $q_{MAC}(s) + q_{BC}(s) = 1$, $\forall s$. ii) The transmission rates $R = [R_{jj'}(s)]$, $\forall s, j, j' \in \{1, 2, r, 1, r, 2\}$, where rate pair $(R_{1r}(s), R_{2r}(s))$ has to satisfy (2) and rate pair $(R_{1t}(s), R_{2t}(s))$ has to satisfy (3). iii) The transmit powers $P = [P_j(s)] \in P$, $\forall s, j \in \{1, 2, r\}$, where $P$ is the feasible set of the adopted power constraint. Moreover, $\bar{R}_{1r} < R_{2r}$ and $\bar{R}_{2r} < \bar{R}_{1r}$ are not possible due to the conservation of flow. Furthermore, $R_{1r} > R_{2r}$ and $R_{2r} > \bar{R}_{1r}$ are also not permitted, since all information bits transmitted by the sources have to arrive at the intended destinations. Therefore, constraints C1 and C2 must hold for any long-term achievable rate pair. The complete boundary surface of the rate region can be obtained by maximizing the weighted long-term sum rate for all weights $\eta \in [0, 1]$.

**Remark 1:** In Fig. 3, the concept of using the optimization problem in (6) to obtain different points on the boundary surface of the rate region is illustrated. In particular, the boundary surface of the rate region in Fig. 3 is the dashed line. Moreover, the corner points of the boundary surface are obtained by setting $\eta = 0, 1$ in the optimization problem (6). We note that all points on the solid lines yield the same value for the weighted sum rate with $\eta = 0, 1$, however, only the corner points, where the dashed and solid lines meet, constitute the boundary surface of the rate region.

Note that throughout this paper, we assume that all nodes have full knowledge of the instantaneous CSI of both links.4 Thus, based on the CSI and the proposed protocols, cf. Theorem 1, Theorem 2, and Lemma 2, each node can individually determine which transmission mode is selected and adapt its transmission strategy accordingly. Furthermore, the proposed optimal protocols may require a coin flip. Thus, we assume that one node (e.g., the relay node) is responsible for performing the coin flip and conveying the result to the other nodes. Moreover, we assume that the channel states change slow enough such that the signaling overhead caused by channel estimation and feedback is negligible compared to the amount of transmitted information.

**E. Outer Bounds on the Achievable Long-Term Rate Regions of the Proposed Protocols**

Finding the optimal value of the optimization variables from (6) is quite involved due to the recursive dynamics of the queues, i.e., transmission rates $R_{jj'}(s)$, $\forall s, j, j' \in \{1, 2, r, 1, r, 2\}$ have to satisfy (2) and (3) which involve $Q_j(i)$, $j = 1, 2$. However, we can obtain outer bounds for the achievable long-term rate regions of the proposed protocols using optimal coding and decoding in the broadcast mode.

---

4We note that in order to achieve the broadcast capacity of the bidirectional relay channel as proposed in [3], both users have to know the transmission rates from the relay to both users. Therefore, knowledge of the global CSI at all nodes is necessary for any conventional/buffer-aided bidirectional relaying protocol using optimal coding and decoding in the broadcast mode.
protocols by neglecting the effect of the queues on the transmission rates in (2) and (3). In the rest of this paper, we formally refer to these outer bounds as $R^\text{out}$. In order to formally state these outer bounds and for clarity of presentation, we define $C_{1r}(s) = C(P_1(s)s_1)$, $C_{2r}(s) = C(P_2(s)s_2)$, $C_{1c}(s) = C(P_1(s)s_1)$, $C_{2c}(s) = C(P_2(s)s_2)$, and $C_{r}(s) = C(P_1(s)s_1 + P_2(s)s_2)$, $\forall s$. Moreover, we define the average capacity rates $\bar{C}_{1r} = E\{q_{\text{MAC}}(s)C_{1r}(s)\}$, $\bar{C}_{2r} = E\{q_{\text{BC}}(s)C_{2r}(s)\}$, $\bar{C}_{1c} = E\{q_{\text{BC}}(s)C_{1c}(s)\}$, $\bar{C}_{2c} = E\{q_{\text{MAC}}(s)C_{2c}(s)\}$, and $\bar{C}_{r} = E\{q_{\text{MAC}}(s)C_{r}(s)\}$.

**Outer bound $R^\text{out}$**: The achievable long-term rate region of the proposed protocol with adaptive mode selection for the bidirectional relay network is outer bounded by the closure of the convex hull of all rate pairs $(\bar{R}_{12}, \bar{R}_{21})$ obtained for all possible $\eta \in [0, 1]$ from the following optimization problem

$$
\max_{q \in \mathcal{Q}, \mu \in \mathcal{P}, R \geq 0} \eta \bar{R}_{12} + (1 - \eta) \bar{R}_{21}
$$

subject to $C_1: \bar{R}_{12} \leq \min\{\bar{C}_{1r}, \bar{C}_{2r}\}$

$C_2: \bar{R}_{21} \leq \min\{\bar{C}_{2r}, \bar{C}_{1r}\}$

$C_3: \bar{R}_{12} + \bar{R}_{21} \leq \bar{C}_{r}$, \hspace{1cm} (7)

where $R = [\bar{R}_{12}, \bar{R}_{21}]$ is an auxiliary optimization variable in (7).

We note that the optimization problems in (6) and (7) are equivalent if we assume that the effect of the queues on the transmission rates in (2) and (3) can be neglected. In particular, the average transmission rates $\bar{R}_{1r}$ and $\bar{R}_{2r}$ are bounded by constraints $\bar{R}_{1r} \leq C_{1r}$, $\bar{R}_{2r} \leq C_{2r}$, and $\bar{R}_{1r} + \bar{R}_{2r} \leq C_{r}$, see (2). Similarly, the average transmission rates $\bar{R}_{1c}$ and $\bar{R}_{2c}$ are bounded by $\bar{R}_{1c} \leq C_{1c}$ and $\bar{R}_{2c} \leq C_{2c}$, see (3). These bounds are tight if the effect of the queues in (2) and (3) can be neglected. Since the user 1-to-user 2 and user 2-to-user 1 transmission rates are given by $\bar{R}_{12} = \min\{\bar{R}_{1r}, \bar{R}_{2r}\}$ and $\bar{R}_{21} = \min\{\bar{R}_{2r}, \bar{R}_{1r}\}$, respectively, we obtain the constraints $C_1, C_2,$ and $C_3$ in (7). Solving the optimization problem for the outer bound in (7) is significantly simpler than solving the optimization problem in (6) as the recursive nature of the dynamics of the queues is avoided. In the following section, we employ this outer bound to find the achievable long-term rate regions of the proposed protocols.

### III. Achievable Long-Term Rate Regions for Bidirectional Relay Channel

In this section, we obtain the outer bounds for the considered power constraints. Subsequently, we specify the conditions and the optimal rate selection policy for which these outer bounds can be achieved.

#### A. The Outer Bound for Fixed Node Transmit Powers

In this subsection, we assume that each node transmits with a fixed power. Thus, the node powers are predefined and given. In the following, our goal is to find the optimal $q^*$ and $R^*$ as the solution of (7). Due to the binary constraints $q_k(s) \in \{0, 1\}, \forall s, k$, the optimization problem in (7) is an integer programming problem which belongs to the category of nondeterministic polynomial-time hard (NP hard) problems. To make the problem tractable, we relax the binary constraints to $0 \leq q_k(s) \leq 1$ which in general implies that the solution of the relaxed problem might not be in the feasible set of the original problem as the relaxed problem has a larger feasible set than the original problem. However, for the optimization problem in (7), one of the solutions always lies at the boundaries of $0 \leq q_k(s) \leq 1$, thus, this solution of the relaxed problem solves the original problem as well. For future reference, let $\mathcal{Q}$ denote the feasible set of $q$ for the relaxed constraint.

We adopt the dual formulation of relaxed problem (7) to solve the problem. As the relaxed problem is linear in the optimization variables, the duality gap is zero and the solution of the primal problem can be found from the solution of the dual problem [26]. Before formally stating the optimal mode selection policy, we introduce some variables that we require in the protocol. First, as we will see later, the optimal protocol may require a coin flip. Therefore, we define $\zeta(s) \in \{0, 1\}$ as the outcome of a coin flip with equiprobable faces in the $i$-th time slot. Second, in the optimal solution, the instantaneous capacities, i.e., $C_{1r}(s), C_{2r}(s), C_{1c}(s), C_{2c}(s),$ and $C_{r}(s)$ are weighted via constants $\mu_1, \mu_2,$ and $\mu_r$ which are referred to as selection weights$^3$. The value of $\mu = [\mu_1, \mu_2, \mu_r]$ depends on the channel statistics, the powers of the nodes, and the value of $\eta$. For given channel statistics, node powers, and $\eta$, the optimal value of $\mu$ is unique.

**Theorem 1**: Under the assumption of a fixed transmit power for each node, the optimal mode selection policy which leads to a point on the boundary surface of the outer bound $R^\text{out}$ is given by

$$
q^*_\text{MAC}(s) = \begin{cases} 1, & \text{if } \Lambda_{\text{MAC}}(s) > \Lambda_{\text{BC}}(s) \\ 0, & \text{if } \Lambda_{\text{MAC}}(s) < \Lambda_{\text{BC}}(s) \\ \zeta(s), & \text{if } \Lambda_{\text{MAC}}(s) = \Lambda_{\text{BC}}(s) \end{cases} (8)
$$

and $q^*_\text{BC}(s) = 1 - q^*_\text{MAC}(s)$ where $\Lambda_{\text{MAC}}(s)$ and $\Lambda_{\text{BC}}(s)$ are referred to as selection metrics and are given by

$$
\Lambda_{\text{MAC}}(s) = \mu_1 C_{1r}(s) + \mu_2 C_{2r}(s) + \mu_r C_{r}(s) \hspace{1cm} (9a)
$$

$$
\Lambda_{\text{BC}}(s) = (1 - \eta - \mu_2 - \mu_r) C_{r}(s) + (\eta - \mu_1 - \mu_r) C_{2c}(s). \hspace{1cm} (9b)
$$

Moreover, $\mu_1, \mu_2,$ and $\mu_r$ are constants which are specified in Proposition 1.

**Proof**: Please refer to Appendix A.

Theorem 1 specifies the optimal transmission mode in each time slot. Moreover, $\mu_1, \mu_2,$ and $\mu_r$ are long-term variables and depend only on the statistics of the channels, the powers of the nodes, and $\eta$. Hence, for fixed node powers and a given $\eta$, these constants can be obtained offline and used as long as the channel statistics remain unchanged.

**Proposition 1**: The optimal values of the selection weights $\mu_1, \mu_2,$ and $\mu_r$ used in Theorem 1 can be obtained iteratively with Algorithm 1 using the following update equations

$$
\mu_1[m+1] = \left[\mu_1[m] + \lambda_1[m] (C_{r}^*[m] - C_{1r}^*[m])\right]^{\eta} \hspace{1cm} (10a)
$$

$$
\mu_2[m+1] = \left[\mu_2[m] + \lambda_2[m] (C_{r}^*[m] - C_{2r}^*[m])\right]^{1-\eta} \hspace{1cm} (10b)
$$

$^3$Appendix A reveals that the selection weights $\mu_1, \mu_2,$ and $\mu_r$ are in fact Lagrange multipliers (dual variables) corresponding to the constraints $\bar{R}_{12} \leq C_{1r}, \bar{R}_{21} \leq C_{2r},$ and $\bar{R}_{12} + \bar{R}_{21} \leq \bar{C}_r$ in (7), respectively.
Algorithm 1 Gradient algorithm for $\mu^*$

initialize $m = 0$ and $\mu[0] = [\mu_1[0], \mu_2[0], \mu_r[0]]$
repeat
1. Compute $\tilde{C}^*[m]$ from (11) for $\mu[m]$
2. Update $\mu[m+1]$ based on (10)
3. Set $m = m + 1$
until convergence to $\mu^*

$$
\mu_r[m+1] = \left[\mu_r[m] + \lambda_r[m] \left(\tilde{C}_{r1}^*[m] + \tilde{C}_{r2}^*[m] - \tilde{C}_r^*[m]\right)\right]_{\min(\eta,1-\eta)},
$$
(10c)

where $m$ is the iteration index and $\lambda_j[m]$, $j \in \{1,2,r\}$, are appropriately chosen step size parameters. Moreover, for a given $\mu[m] = [\mu_1[m], \mu_2[m], \mu_r[m]]$ in the $m$th iteration, $q_k^*(s)$, $k \in \{\text{MAC, BC}\}$, is obtained from Theorem 1. Therefore, using the distribution of the fading gains $f_j(s)$, $j = 1,2$, the elements of $\mathbf{C}^*[m] = [\mathbf{C}_{r1}^*[m], \mathbf{C}_{r2}^*[m], \mathbf{C}_{r3}^*[m], \mathbf{C}_{r4}^*[m], \mathbf{C}_r^*[m]]$ are computed analytically from the following expressions:

$$
\tilde{C}_{r1}^*[m] = \int_{s_1 \times s_2} q_{1r}^*(s) C_{1r}(s) f_1(s_1) f_2(s_2) ds_1 ds_2
$$
(11a)

$$
\tilde{C}_{r2}^*[m] = \int_{s_1 \times s_2} q_{2r}^*(s) C_{2r}(s) f_1(s_1) f_2(s_2) ds_1 ds_2
$$
(11b)

$$
\tilde{C}_{r3}^*[m] = \int_{s_1 \times s_2} q_{rC}^*(s) C_{rC}(s) f_1(s_1) f_2(s_2) ds_1 ds_2
$$
(11c)

$$
\tilde{C}_{r4}^*[m] = \int_{s_1 \times s_2} q_{2C}^*(s) C_{2C}(s) f_1(s_1) f_2(s_2) ds_1 ds_2
$$
(11d)

$$
\tilde{C}_r^*[m] = \int_{s_1 \times s_2} q_{rC}^*(s) C_r(s) f_1(s_1) f_2(s_2) ds_1 ds_2.
$$
(11e)

Proof: Please refer to Appendix B.

Remark 2: The mode selection metric $A_k(i)$ introduced in (9) is a weighted sum of the capacity terms in each time slot where the weights, $\mu$, are constant. In each time slot, the mode with the highest value of the selection metric is selected. Since the fading states have continuous probability density functions, the probability that $A_{\text{MAC}}(s) = A_{\text{BC}}(s)$ is zero except for the following two cases: 1) if $\mu_1 = \mu_2 = 0$, $\eta = \frac{1}{2}$, and $P_2 = P_r$ hold, and 2) if $\mu_2 = \mu_r = 0$, $\eta = \frac{1}{2}$, and $P_1 = P_r$ hold. The former case corresponds to channels when $\Omega_1 \gg \Omega_2$ and the latter case corresponds to channels where $\Omega_1 \ll \Omega_2$. Since one of the links is the bottleneck in the network and $\eta = \frac{1}{2}$ holds for these cases, the optimal selection policy is to select the multiple-access and broadcast modes with equal probability. Apart from these special cases, the selection policy in (8) indicates that, for any fading state $s = (s_1, s_2)$, the choice of the optimal transmission mode is unique. In other words, for a given fading state, it is sub-optimal to share the resources between the two transmission modes and only one of the transmission modes should be used. Hence, adaptive mode selection is the key to improve the long-term rate region of the bidirectional relay channel with half-duplex nodes and block fading.

Remark 3: The gradient method in Algorithm 1 is guaranteed to converge to the optimal dual variable $\mu^*$ provided that the step sizes $\lambda_j[m]$, $\forall j, m$ are chosen sufficiently small [27], [28]. Common choices of the step sizes include: i) Constant step size, i.e., $\lambda_j[m] = \lambda_j$, $\forall j, m$, ii) constant step length, i.e., setting $\lambda_j[m]$ such that $|\mu_j[m+1] - \mu_j[m]| = \bar{\lambda}_j$, $\forall j, m$, iii) square summable but not absolutely summable step size, i.e., $\sum_{m=1}^{\infty} \lambda_j[m] < \infty$ and $\sum_{m=1}^{\infty} \lambda_j[m] \to \infty$, $\forall j$, and iv) nonsummable diminishing step size, i.e., $\lim_{m \to \infty} \lambda_j[m] = 0$ and $\sum_{m=1}^{\infty} \lambda_j[m] \to \infty$, $\forall j$. For a comprehensive study of different choices of step sizes and the resulting convergence properties, we refer to [27].

Remark 4: The protocol in Theorem 1 contains the following two special cases:

- One-way relaying: The one-way relaying solutions for the user 1-to-user 2 and user 2-to-user 1 information flows are obtained by setting $\eta = 1$ and $\eta = 0$, respectively. Note that for fixed node transmit powers, we can achieve the maximum rate of one-way relaying for one direction, and still provide a non-zero rate in the other direction. For instance, if we assume $\eta = 1$, the long-term variables are obtained as $\mu_2 = \mu_r = 0$. Therefore, $R_{12}$ is identical to the maximum rate obtained in [15], but still $R_{21} > 0$ holds.

- Sum rate: For sum rate maximization, we have to set $\eta = \frac{1}{2}$. We note that the sum rate maximization problem was also considered in [17].

In order to calculate the average rates $\mathbf{R}^* = [\bar{R}_{12}, \bar{R}_{21}]$ from optimization problem (7), we have to specify which of the five bounds in constraints C1, C2, and C3 in (7) are active for the optimal $\mathbf{C}^*$. In general, we can distinguish $2^5$ cases for the five bounds to be active or not. In this paper, instead of investigating these $2^5$ cases, without loss of generality, we distinguish seven mutually exclusive cases for the relation between $(\bar{C}_{r2}^*, \bar{C}_{r1}^*)$ and $(\bar{C}_{1r}^*, \bar{C}_{2r}^*, \bar{C}_r^*)$. In particular, given $(\bar{C}_{1r}^*, \bar{C}_{2r}^*, \bar{C}_r^*)$, the capacity pair $(\bar{C}_{r2}^*, \bar{C}_{r1}^*)$ falls into one of the seven mutually exclusive regions in Fig. 4, i.e., $(\bar{C}_{r2}^*, \bar{C}_{r1}^*) \in A_k$, $k = 1, \ldots, 7$. Therefore, if we know which region $A_k$, the capacity pair $(\bar{C}_{r2}^*, \bar{C}_{r1}^*)$ belongs to, we also know which bounds in constraints C1, C2, and C3 in (7) are active. In fact, which region $A_k$ the optimal $(\bar{C}_{r2}^*, \bar{C}_{r1}^*)$ belongs to, depends on the channel statistics, node powers, and the value of $\eta$. In the following corollary, the optimal rate pairs $(\bar{R}_{12}, \bar{R}_{21})$ for these seven possible mutually exclusive regions $(\bar{C}_{r2}^*, \bar{C}_{r1}^*) \in A_k$, $k = 1, \ldots, 7$ are provided.

Corollary 1: The boundary surface of the outer bound $\mathbf{R}^{\text{out}}$ is comprised of rate pairs $(\bar{R}_{12}, \bar{R}_{21})$ which are obtained for any given $\eta \in [0,1]$ as:

$$
(\bar{R}_{12}, \bar{R}_{21}) =
$$
(12)

$$
\begin{align*}
& (\bar{C}_{r2}^* - \bar{C}_{r1}^*, \bar{C}_{r2}^*), & \text{if } (\bar{C}_{r2}^*, \bar{C}_{r1}^*) \in A_1 \cup A_6 & & \land & & \eta \leq \frac{1}{2} \land \eta \geq \frac{1}{2}
& (\bar{C}_{1r}^*, \bar{C}_{r1}^* - \bar{C}_{r2}^*), & \text{if } (\bar{C}_{r2}^*, \bar{C}_{r1}^*) \in A_1 & & \land & & \eta \geq \frac{1}{2}
& (\bar{C}_{r1}^* - \bar{C}_{r2}^*, \bar{C}_{r1}^*), & \text{if } (\bar{C}_{r2}^*, \bar{C}_{r1}^*) \in A_2 \cup A_7 & & \land & & \eta \leq \frac{1}{2}
& (\bar{C}_{r2}^* - \bar{C}_{r1}^*, \bar{C}_{r2}^*), & \text{if } (\bar{C}_{r2}^*, \bar{C}_{r1}^*) \in A_3 & & \land & & \eta \geq \frac{1}{2}
& (\bar{C}_{r2}^*, \bar{C}_{r1}^*), & \text{if } (\bar{C}_{r2}^*, \bar{C}_{r1}^*) \in A_4
& (\bar{C}_{1r}^*, \bar{C}_{r1}^*), & \text{if } (\bar{C}_{r2}^*, \bar{C}_{r1}^*) \in A_5
\end{align*}
$$
where for each value of $\eta$, the elements of $C^*$ can be computed analytically from (11) by employing the optimal $q^*$ from Theorem 1 and the optimal $\mu^*$ from Proposition 1.

**Proof:** By obtaining $C^*$ for a given $\eta$, we can determine which set $A_k$, $k = 1, \ldots, 7$, the optimal $(C^*_{22}, C^*_{11})$ belongs to and which of the five bounds in constraints $C1$, $C2$, and $C3$ in (7) are active, see Fig. 4. Knowing the active bounds, it is straightforward to obtain the rate pair $(R_{12}^*, R_{21}^*)$ which maximizes $\eta R_{12} + (1 - \eta) R_{21}$. For example, if $(C^*_{22}, C^*_{11}) \in A_2$ holds, from constraints $C1$, $C2$, and $C3$ in (7), the active constraints are $R_{12} \leq C^*_{22}$, $R_{21} \leq C^*_{11}$, and $R_{12} + R_{21} \leq C^*$, see Fig. 4. Therefore, in order to maximize the weighted sum rate, we have to choose $(R_{12}, R_{21}) = (C^*_{22}, C^*_{11})$ and $(R_{12}, R_{21}) = (C^*_1, C^*_2 - C^*_3)$ if $\eta \leq \frac{1}{2}$ and $\eta \geq \frac{1}{2}$ hold, respectively. Note that for $\eta = \frac{1}{2}$, both aforementioned points and all the rate pairs that are achieved by time sharing from these two points are solutions, i.e., they lead to the same value for $\eta R_{12} + (1 - \eta) R_{21}$. In a similar manner, the optimal rate pairs $(R_{12}^*, R_{21}^*)$ for the remaining regions are obtained and given in (12). This completes the proof.

Note that we provide two expressions for $(R_{12}^*, R_{21}^*)$ in (12) for $(C^*_{22}, C^*_{11}) \in A_k$, $k = 1, 2, 6, 7$ if $\eta = \frac{1}{2}$. In particular, for these cases, both values provided in (12) and all the rate pairs that are achieved by time sharing from these two points are solutions of (7), i.e., they lead to the same value of $\eta R_{12} + (1 - \eta) R_{21}^*$.

**B. The Outer Bound for Per-Node Long-Term Power Constraints**

In this subsection, we consider a long-term power constraint for each node. Before we provide the optimal protocol, we note that solving (7) with the relaxed constraint on the mode selection variables requires additional considerations here. In particular, if $0 < q_k(s) < 1$, $k \in \{\text{MAC}, \text{BC}\}$ holds, only for a fraction $q_{\text{MAC}}(s)$ of the $i$-th time slot, user 1 and user 2 transmit and similarly for a fraction $q_{\text{BC}}(s)$ of the $i$-th time slot, the relay transmits. Therefore, the transmit powers of the nodes during the active time fraction are given by $P_1(s)/q_{\text{MAC}}(s)$, $P_2(s)/q_{\text{MAC}}(s)$, and $P_r(s)/q_{\text{BC}}(s)$ such that the powers of the nodes for the total duration of the $i$-th time slot remain $P_1(s)$, $P_2(s)$, and $P_r(s)$, respectively. Note that this problem formulation with scaling the powers has been widely employed in the literature [24], [29].

Similar to the previous subsection, we employ the dual formulation of the relaxed optimization problem in (7) to find the optimal mode selection and power allocation policies. However, unlike for the case of a fixed transmit power for each node, the optimization problem with a long-term power constraint for each node is not a linear program. In order to guarantee the optimality of the proposed solution, which is based on the dual formulation, we formally state the convexity of the optimization problem in (7) in the following lemma.

**Lemma 1:** The optimization problem in (7) with the relaxed constraint, $0 \leq q_k(s) \leq 1$, $k \in \{\text{MAC}, \text{BC}\}$, and the per-node long-term power constraint is convex in the optimization variables $q$, $P$, and $R$.

**Proof:** The cost function of (7) and the power constraint in (4) are affine in the optimization variables $R$ and $P$, respectively. Moreover, the average capacity terms in constraints $C1$, $C2$, and $C3$ are given by $C_{1r} = E[q_{\text{MAC}}(s)C(P_1(s)s_1)/q_{\text{MAC}}(s))]$, $C_{2r} = E[q_{\text{MAC}}(s)C(P_2(s)s_2)/q_{\text{MAC}}(s))]$, $\bar{C}_{11} = E[q_{\text{BC}}(s)C(P_r(s)s_1)/q_{\text{BC}}(s))]$, $\bar{C}_{22} = E[q_{\text{BC}}(s)C(P_r(s)s_2)/q_{\text{BC}}(s))]$, and $C_r = E[q_{\text{MAC}}(s)C(P_r(s)s_1 + P_2(s)s_2)/q_{\text{MAC}}(s))]$. These terms can be rewritten as a sum of functions of the general form $f(x,y) = x \log(1 + \frac{y}{x})$. For instance, we can rewrite $C_{1r}$ as

$$\bar{C}_{1r} = \frac{1}{N} \sum_{i=1}^{N} q_{\text{MAC}}(s) \log_2\left(1 + P_1(s)s_1/q_{\text{MAC}}(s)\right).$$

Furthermore, $f(x,y)$ was shown to be jointly concave in $(x,y)$ in [24, Proposition 1]. Therefore, the optimization problem in (7) with the relaxed constraint on $q_k(s)$, $k \in \{\text{MAC}, \text{BC}\}$, and a per-node long-term power constraint is convex in optimization variables $q$, $P$, and $R$. This completes the proof.

In the following, we provide the optimal $q^*$ and $P^*$. However, before formally stating the optimal mode selection and power allocation policies, we introduce constants $\mu_j$ and $\gamma_j$, $j \in \{1, 2, r\}$ which are referred to as selection weights and power weights in the optimal protocol, respectively. The values of $\mu$ = $[\mu_1, \mu_2, \mu_r]$ and $\gamma = [\gamma_1, \gamma_2, \gamma_r]$ depend on the channel statistics, power budgets, and the value of $\eta$. For given channel statistics, node power budgets, and $\eta$, the optimal values of $\mu$ and $\gamma$ are unique.

**Theorem 2:** Under a per-node long-term power constraint, the optimal mode selection and power allocation policies which lead to a point on the boundary surface of the outer bound $R_{\text{out}}$ are provided in the following. In particular, the
optimal mode selection policy is given by
\[ q_{\text{MAC}}^*(s) = \begin{cases} 1, & \text{if } \Lambda_{\text{MAC}}(s) \geq \Lambda_{\text{BC}}(s) \\ 0, & \text{if } \Lambda_{\text{MAC}}(s) < \Lambda_{\text{BC}}(s) \end{cases} \tag{14} \]

and \( q_{\text{BC}}^*(s) = 1 - q_{\text{MAC}}^*(s) \) where \( \Lambda_{\text{MAC}}(s) \) and \( \Lambda_{\text{BC}}(s) \) are given by
\[ \Lambda_{\text{MAC}}(s) = \mu_1 C_1(s) + \mu_2 C_2(s) + \mu_r C_r(s) - \frac{1}{\ln 2} \left[ \mu_1 \cdot P_1(s) s_1 + \mu_2 \cdot P_2(s) s_2 + \mu_r \cdot P_r(s) s_r \right] + \frac{\mu_r}{1 + P_1(s) s_1 + P_2(s) s_2} \tag{15a} \]
\[ \Lambda_{\text{BC}}(s) = (1 - \eta - \mu_2 - \mu_r) C_1(s) + (\eta - \mu_1 - \mu_r) C_2(s) - \frac{1}{\ln 2} \left[ (1 - \eta - \mu_2 - \mu_r) \cdot P_1(s) s_1 + (\eta - \mu_1 - \mu_r) \cdot P_2(s) s_2 \right] + \frac{\mu_r}{1 + P_1(s) s_1 + P_2(s) s_2} \tag{15b} \]

Moreover, the optimal power allocation policy is given by
\[ P_1^*(s) = \begin{cases} P_1(s), & \text{if } q_{\text{MAC}}^*(s) \neq 0 \\ 0, & \text{otherwise} \end{cases} \tag{16a} \]
\[ P_2^*(s) = \begin{cases} P_2(s), & \text{if } q_{\text{MAC}}^*(s) \neq 0 \\ 0, & \text{otherwise} \end{cases} \tag{16b} \]
\[ P_r^*(s) = \begin{cases} P_r(s), & \text{if } q_{\text{BC}}^*(s) \neq 0 \\ 0, & \text{otherwise} \end{cases} \tag{16c} \]

where
\[ P_1(s) = \begin{cases} 0, & \text{if } \left\{ s_1 < \frac{\gamma_1 \ln 2}{\mu_1 + \mu_r}, \frac{\gamma_2 \ln 2}{\mu_2 + \mu_r} \right\} \\ \frac{\mu_1 + \mu_r}{\gamma_1 \ln 2 - s_1 d}, & \text{if } \frac{\gamma_1 \ln 2}{\mu_1 + \mu_r} < s_1 < \frac{\gamma_2 \ln 2}{\mu_2 + \mu_r} \right\} \\ \frac{\mu_1 + \mu_r}{\gamma_1 \ln 2 - s_1 d}, & \text{otherwise} \end{cases} \tag{17a} \]
\[ P_2(s) = \begin{cases} 0, & \text{if } \left\{ s_1 < \frac{\gamma_1 \ln 2}{\mu_1 + \mu_r}, \frac{\gamma_2 \ln 2}{\mu_2 + \mu_r} \right\} \\ \frac{\mu_2 + \mu_r}{\gamma_1 \ln 2 - s_2 d}, & \text{if } \frac{\gamma_2 \ln 2}{\mu_2 + \mu_r} < s_2 < \frac{\gamma_2 \ln 2}{\mu_2 + \mu_r} \right\} \\ \frac{\mu_2 + \mu_r}{\gamma_1 \ln 2 - s_2 d}, & \text{otherwise} \end{cases} \tag{17b} \]
\[ P_r(s) = \left[ \frac{-b + \sqrt{b^2 - 4ac}}{2a} \right]^+, \tag{17c} \]

and \( a = \gamma_r \ln 2 s_1 s_2, b = \gamma_r \ln 2(s_1 + s_2) - (1 - \mu_1 - \mu_2 - 2\mu_r) s_1 s_2, c = \gamma_r \ln 2 - (1 - \eta - \mu_1 - \mu_r) s_1 - (\eta - \mu_1 - \mu_r) s_2, \) and \( d \) is chosen to satisfy
\[ \frac{\mu_1 s_1}{\gamma_1 \ln 2 - s_1 d} + \frac{\mu_2 s_2}{\gamma_2 \ln 2 - s_2 d} = \frac{\mu_r}{d} + 1, \tag{18} \]

**Algorithm 2 Gradient algorithm for \( \mu^* \) and \( \gamma^* \)**

Initialize \( m = 0, \mu[0] = [\mu_1[0], \mu_2[0], \mu_r[0]], \) and \( \gamma[0] = [\gamma_1[0], \gamma_2[0], \gamma_r[0]] \)

repeat

1. Compute \( \bar{C}^*[m] \) from (11) and \( \bar{P}^*[m] \) from (20) for \( \mu[m] \) and \( \gamma[m] \)
2. Update \( \mu[m+1] \) based on (10) and \( \gamma[m+1] \) based on (19)
3. Set \( m = m + 1 \)

until convergence to \( \mu^* \) and \( \gamma^* \)

Furthermore, \( \mu \) and \( \gamma \) are constants which are given in Proposition 2.

**Proof:** Please refer to Appendix C.

Theorem 2 specifies the optimal transmission mode and the transmit powers of the nodes in each time slot. The optimal values of constants \( \mu \) and \( \gamma \) are provided in the following proposition.

**Proposition 2:** The optimal values of the long-term variables \( \mu \) and \( \gamma \) used in Theorem 2 are obtained iteratively with Algorithm 2 where \( \mu[m] \) is updated as in (10) and \( \gamma[m] \) is updated as
\[ \gamma_1[m+1] = \gamma_1[m] + \kappa_1[m] (\bar{P}_1^*[m] - P_1^{\text{max}}) \left[ \frac{\mu_1[m] + \mu_2[m]}{\bar{P}_1^*[m]} \right] \tag{19a} \]
\[ \gamma_2[m+1] = \gamma_2[m] + \kappa_2[m] (\bar{P}_2^*[m] - P_2^{\text{max}}) \left[ \frac{\mu_2[m] + \mu_2[m]}{\bar{P}_2^*[m]} \right] \tag{19b} \]
\[ \gamma_r[m+1] = \gamma_r[m] + \kappa_r[m] \left( \bar{P}_r^*[m] - P_r^{\text{max}} \right) \tag{19c} \]

where \( m \) is the iteration number and \( \lambda_j[m], \kappa_j[m], j \in \{1,2,r\} \) are the step sizes in the \( m \)-th iteration. Moreover, for given \( \mu[m] \) and \( \gamma[m] \) in the \( m \)-th iteration, \( q_{\text{bc}}^*(s), k \in \{\text{MAC}, \text{BC}\} \) and \( P_j^*[s], j \in \{1,2,r\} \) are obtained from Theorem 2. Therefore, using the distribution of the fading gains \( f_j(s_j), j = 1,2 \), the elements of \( C^*[m] \) are computed analytically based on (11) and the elements of \( P^*[m] = [P_1^*[m], P_2^*[m], P_r^*[m]] \) are computed as
\[ P_j^*[m] = \frac{1}{s \in [S_1 \times S_2]} \int P_j^*(s) f_1(s_1) f_2(s_2) ds_1 ds_2, j \in \{1,2,r\} \tag{20} \]

**Proof:** Please refer to Appendix D.

**Remark 5:** Different options for choosing the step sizes \( \lambda_j[m], j \in \{1,2,r\} \) are introduced in Remark 3. The same options are also available for the step sizes \( \kappa_j[m], j \in \{1,2,r\} \) in Algorithm 2.

**Remark 6:** In order to gain some intuition regarding the optimal power allocation policy in Theorem 2, in Fig. 5, we illustrate for what fading states \( s = (s_1, s_2) \) the values of \( P_j(s) \) in (17) are positive and zero, respectively. Intuitively, from Fig. 5 a), we conclude that if the qualities of both channels are high, both users transmit with non-zero powers, if the quality of one of the channels is low, the corresponding user remains silent, and if the qualities of both channels are low, both users remain silent. In other words, due to the long-term power constraint,
the users save their powers and do not transmit if the channels are weak. Similarly, we observe from Fig. 5 b) for which fading states the relay transmits and for which fading states it remains silent.

Remark 7: The protocol in Theorem 2 contains the following two special cases:

**One-way relaying:** The one-way relaying solutions for the user 1-to-user 2 and user 2-to-user 1 information flows are obtained by setting \( \eta = 1 \) and \( \eta = 0 \), respectively. Since each node has its own power budget, we can achieve the maximum rate of one-way relaying in one direction, and still provide a non-zero rate for the other direction. For instance, if we assume \( \eta = 1 \), the long-term variables are obtained as \( \mu_2 = \mu_r = 0 \). Therefore, \( R_{12} \) is identical to the maximum rate obtained in [15], but still \( R_{21} > 0 \) holds.

**Sum rate:** For sum rate maximization, we set \( \bar{\mu} = 1 \). In order to develop some intuition for this case, we assume symmetric channels, i.e., \( \Omega_1 = \Omega_2 \) and \( P_{i}^{\text{max}} = P_2^{\text{max}} \). For this case, we obtain \( \mu_1 = \mu_2 = 0 \). Thereby, for the multiple-access mode, in each time slot, the optimal power allocation policy activates only one of the users, i.e., the multiple-access channel reduces to two orthogonal point-to-point channels. In particular, the optimal powers are given by

\[
P_1^*(s) = \begin{cases} \frac{\mu}{\gamma} & \text{if } s_1 > \frac{\gamma \ln 2}{\mu} \cap s_2 < \frac{\gamma}{\mu} s_1 \quad \text{(21a)} \\ 0 & \text{otherwise} \end{cases}
\]

\[
P_2^*(s) = \begin{cases} \frac{\mu}{\gamma} & \text{if } s_1 < \frac{\gamma}{\mu} s_2 \cap s_2 > \frac{\gamma \ln 2}{\mu} \quad \text{(21b)} \\ 0 & \text{otherwise} \end{cases}
\]

A similar observation was also made in [12] and [30] for the multiple-access channel. We note that the sum rate maximization problem with power allocation was also considered in [16].

Remark 8: Similar to the previous subsection, after having obtained the optimal \( \mu^* \) and \( \gamma^* \) from Proposition 2 and the optimal \( \mathbf{q}^* \) and \( \mathbf{P}^* \) from Theorem 2, we can calculate \( \mathbf{C}^* \) from (11). The optimal rate pair \( (R_{12}^*, R_{21}^*) \) can still be obtained from (12).

C. Achievability of the Outer Bounds

In this subsection, we provide the conditions as well as the optimal transmission strategies, i.e., the optimal mode selection, rate selection, and/or power allocation policies, which achieve the outer bound \( R_{\text{out}}^* \). Before formally stating the optimal transmission strategies, we introduce some variables that we require for the optimal mode selection policy. First, for the multiple-access mode, we employ successive decoding where the decoding order is determined by a coin flip. To this end, we define \( T(s) \in \{0,1\} \) as the outcome of a coin flip in the \( i \)-th time slot where the probabilities of the possible outcomes of the coin flip are \( \Pr{T(s) = 1} = p_t \) and \( \Pr{T(s) = 0} = 1 - p_t \). Specifically, if \( T(s) = 0 \), the relay first decodes the codeword received from user 1 and treats the codeword from user 2 as noise. Then, the relay subtracts the contribution of user 1’s codeword from the received codeword and decodes the codeword received from user 2. Similarly, if \( T(s) = 1 \), the relay first decodes the codeword received from user 2 and treats the codeword of user 1 as noise, and then the relay decodes the codeword received from user 1. Second, for the optimal rate selection policy, nodes transmit with constant fractions of the instantaneous capacities. These constant fractions are collected in vector \( \mathbf{\rho} = [\rho_1, \rho_2, \rho_{r1}, \rho_{r2}] \). The value of \( \mathbf{\rho} \) depends on the channel statistics, the powers (budgets) of the nodes, and the value of \( \eta \). For given channel statistics, the powers of the nodes, and \( \eta \), the optimal value of \( \mathbf{\rho} \) is unique. The following lemma reveals how the outer bound \( R_{\text{out}}^* \) can be achieved and specifies the corresponding optimal rate selection policy.

Lemma 2: The outer bound \( R_{\text{out}}^* \) is achievable provided that the relay nodes are equipped with infinite-size buffers, i.e., \( Q_j^{\text{max}} \to \infty, j = 1, 2 \), the optimal values of \( \mathbf{q}^* \) and/or \( \mathbf{P}^* \) are given by Theorem 1 or Theorem 2 depending on the adopted power constraint, and the following rate selection policy is used

\[
R_{1r}(s) = \begin{cases} \rho_1 C_{1r}(s), & \text{if } T(s) = 1 \\ \rho_1 (C_{r} - C_{2r}(s)), & \text{otherwise} \end{cases} \quad (22a)
\]

\[
R_{2r}(s) = \begin{cases} \rho_2 C_{2r}(s), & \text{if } T(s) = 0 \\ \rho_2 (C_{r} - C_{1r}(s)), & \text{otherwise} \end{cases} \quad (22b)
\]

\[
R_{r1}^* = \min\{\rho_1 C_{r1}(s), Q_2(i-1)\} \quad (22c)
\]

\[
R_{r2}^* = \min\{\rho_2 C_{r2}(s), Q_1(i-1)\} \quad (22d)
\]

where constants \( \rho_{1r}, \rho_{2r}, \rho_{r1}, \) and \( \rho_{r2} \) are given by

\[
\rho_{1r} = \frac{R_{12}^*}{p_1(C_{r1}^* + (1 - p_1)(C_{r} - C_{2r}^*))} \quad (23a)
\]

\[
\rho_{2r} = \frac{R_{21}^*}{p_2(C_{r}^* - C_{r2}^*) + (1 - p_2)C_{2r}^*} \quad (23b)
\]

\[
\rho_{r1} = \frac{R_{r1}^*}{C_{r1}^*} \quad (23c)
\]

\[
\rho_{r2} = \frac{R_{r2}^*}{C_{r2}^*} \quad (23d)
\]

and \( C^* \) and \( R^* \) are obtained from (11) and (12), respectively. Moreover, the decoding order probability is given by

\[
p_t = \begin{cases} p_t^{\text{min}}, & \text{if } \eta < \frac{1}{2} \\ [p_t^{\text{min}}, p_t^{\text{max}}], & \text{if } \eta = \frac{1}{2} \\ p_t^{\text{max}}, & \text{if } \eta > \frac{1}{2} \end{cases} \quad (24)
\]

where

\[
p_t^{\text{min}} = \left[ \frac{R_{12}^* + C_{r2}^* - C_{r1}^*}{C_{r1}^* + C_{r2}^* - C_{r}^*} \right]_0 \quad (25a)
\]

\[
p_t^{\text{max}} = \left[ \frac{C_{r}^* - R_{12}^*}{C_{r1}^* + C_{r2}^* - C_{r}^*} \right]_0 \quad (25b)
\]

Note that if \( \eta = \frac{1}{2} \) holds, any value of \( p_t \) in the interval \([p_t^{\text{min}}, p_t^{\text{max}}]\) is optimal.

Proof: Please refer to Appendix E.

Remark 9: The outer bound \( R_{\text{out}}^* \) is achievable by the proposed protocol only if the effect of the queues at the relay becomes negligible, i.e., in each time slot, the relay node has space available in its buffers for information reception and sufficient information for transmission. Assuming that \( Q_j^{\text{max}} \to \infty, j = 1, 2 \), the buffers have always enough space available to store information. Hence, Lemma 2 effectively indicates that, if we use the optimal variables for the outer bound, i.e., \( \mathbf{q}^* \) and/or \( \mathbf{P}^* \), the proposed rate selection policy operates the queues of the buffers such that the effect of
Remark 10: Constants $\rho_{jj'}$ in Lemma 2 are needed to stabilize the queues especially for the case of asymmetric channels. For instance, if $\Omega_1 \gg \Omega_2$ holds, the capacity of the channel between user 2 and the relay is the bottleneck capacity which limits the achievable long-term rates, e.g., $\bar{R}_{12}^* = \bar{C}_{12}^* < \bar{C}_{1r}^*$ holds. Hence, in time slots where the multiple-access mode is adopted, if user 1 always transmitted with full capacity rate, i.e., $C_r(s)$ for $T(s) = 1$ and $C_r(s) - C_{2r}(s)$ for $T(s) = 0$, some information bits would be trapped in buffer $B_1$ and would never be forwarded to user 2. In order to avoid this information loss, user 1 transmits only with a fraction of the capacity rates such that the relay is able to forward this information to user 2 at some later time slots. Therefore, the proposed protocol chooses $\rho$ according to (23) such that rate pair $(\hat{R}_{12}^*, \hat{R}_{21}^*)$ is achievable.

Remark 11: In Lemma 2, we employ a deterministic approach to stabilize the queues, i.e., the nodes transmit with a fixed fraction of the channel capacity in all time slots in which they are chosen to transmit. It is also possible to employ a probabilistic approach to stabilize the queues and obtain the same average long-term rate region [16]–[18]. For instance, the protocol in [17] employs the four possible point-to-point modes in the network as additional, separate transmission modes, i.e., the user 1-to-relay, user 2-to-relay, relay-to-user 1, and relay-to-user 2 transmission modes. Moreover, when the multiple-access and broadcast modes are used in [16], [17], the nodes transmit with a transmission rate pair on the boundary surface of the instantaneous capacity region, see Fig. 2. Thereby, using coin flips, the protocol in [17] chooses probabilistically between the multiple-access mode and the user 1-to-relay and user 2-to relay modes. Similarly, it chooses probabilistically between the broadcast mode and the relay-to-user 1 and relay-to-user 2 modes. As an example for this approach, if $\bar{R}_{12}^* = \bar{C}_{2r}^* < \bar{C}_{1r}^*$ holds, the protocols in [16], [17] chooses probabilistically between the multiple-access mode and the user 2-to-relay mode to reduce the average amount of information transmitted from user 1 to the relay until the the queue in buffer $B_1$ becomes stable, i.e., $\bar{R}_{1r}^* = \bar{C}_{2r}^*$.

IV. DISCUSSION AND PRACTICAL CHALLENGES

In this section, we discuss some of the challenges arising in the implementation of the proposed protocols and propose possible strategies to address these challenges.

A. Availability of the Channel Statistics

In order to calculate constants $\mu$ and/or $\gamma$ in Algorithms 1 and 2, and $\rho$ and $p_t$ in Lemma 2, perfect knowledge of the distribution of the fading gains, i.e., $f_j(s_j)$, $j = 1, 2$, is required. In practice, this information can be obtained based on a sufficiently large number of channel measurements. Nevertheless, obtaining an accurate estimate of the fading distribution involves many practical challenges. Hence, we propose an adaptive algorithm which updates the values of $\mu$, $\gamma$, $\rho$, and $p_t$ in each time slot based on a limited history of fading realizations. Thus, this algorithm does not require a priori knowledge of the channel distribution to compute estimates of $\mu$, $\gamma$, $\rho$, and $p_t$ and is robust to changing channel statistics. In the following, we provide the adaptive algorithm for the case of fixed node transmit powers. The extension of the algorithm to the case of long-term per-node power constraint is straightforward. We denote the estimates of $\mu^*$, $\rho^*$, and $p_t$ in the $i$-th time slot by $\hat{\mu}(i) = [\hat{\mu}_1(i), \hat{\mu}_2(i), \hat{\mu}_r(i)]$, $\hat{\rho}(i) = [\hat{\rho}_{1r}(i), \hat{\rho}_{2r}(i), \hat{\rho}_{12}(i), \hat{\rho}_{21}(i), \hat{\rho}_{2r}(i)]$, and $\hat{p}_t(i)$, respectively.\footnote{Note that if the channel statistics are available, in an initial phase before transmission starts, Algorithms 1 or 2 are employed to find $\mu^*$ and/or $\gamma^*$. However, if the channel statistics are unknown, the values of $\hat{\mu}$ and/or $\hat{\gamma}$ are updated directly in each time slot and there is no initial phase. This is the reason why we employ the notation $\hat{\mu}(m)$ and/or $\hat{\gamma}(m)$ in Algorithms 1 and 2 where $m$ is the iteration index and $\hat{\mu}(i)$ in Algorithm 3 where $i$ is the time slot number.}

Fig. 5. The optimal transmit powers of the nodes in (17) for given fading states $(s_1, s_2)$: a) four possible regions for the powers of user 1 and user 2 for the multiple-access transmission mode, and b) two possible regions for the optimal power of the relay for the broadcast transmission mode.
Moreover, we introduce the historical averages of the capacity terms within a time window of length $T_c$. $C(i) = \{C_{1r}(i), C_{2r}(i), C_{1r}(i), C_{2r}(i), C_{r}(i)\}$ where $C_{1r}(i) = \frac{1}{T_c} \sum_{i' = \mu}^i q_{MAC}(i') C_{1r}(i')$, $C_{2r}(i) = \frac{1}{T_c} \sum_{i' = \mu}^i q_{MAC}(i') C_{2r}(i')$, $C_{1r}(i) = \frac{1}{T_c} \sum_{i' = \mu}^i q_{BC}(i') C_{1r}(i')$, $C_{2r}(i) = \frac{1}{T_c} \sum_{i' = \mu}^i q_{BC}(i') C_{2r}(i')$, and $C_{r}(i) = \frac{1}{T_c} \sum_{i' = \mu}^i q_{MAC}(i') C_{r}(i')$ where $i' = \max\{1, i - T_c + 1\}$.

Using these notations, $\tilde{\rho}$ and $\tilde{\mu}$ are updated in the $i$-th time slot based on (23) and (24), respectively, by employing the historical average capacity terms instead of $C^*$ given in (11). Moreover, $\tilde{\mu}$ is updated in the each time slot using the following update rules

\[
\tilde{\mu}_1(i + 1) = \left[\tilde{\mu}_1(i) + \lambda_1(i) \left(\tilde{C}_{r2}(i) - \tilde{C}_{1r}(i)\right)\right]_0 \tag{25a}
\]

\[
\tilde{\mu}_2(i + 1) = \left[\tilde{\mu}_2(i) + \lambda_2(i) \left(\tilde{C}_{r1}(i) - \tilde{C}_{2r}(i)\right)\right]_0 \tag{25b}
\]

\[
\tilde{\mu}_r(i + 1) = \left[\tilde{\mu}_r(i) + \lambda_r(i) \left(\tilde{C}_{r1}(i) + \tilde{C}_{r2}(i) - \tilde{C}_{r}(i)\right)\right]_{\min\{\eta, 1 - \eta\}} \tag{25c}
\]

where $\lambda_j(i), j \in \{1, 2, r\}$, is the step size in the $i$-th time slot.

We formally present this online algorithm for updating $\tilde{\mu}$ in Algorithm 3. If the channel gains are stationary processes, i.e., the channel statistics are fixed and do not change, the value of $\tilde{\mu}^*$ is an accurate estimate of $\mu^*$ after a sufficiently large number of time slots. For this case, adaptive step sizes, e.g., the square summable but not absolutely summable step size and the nonsummable diminishing step size, are preferable as they can lead to a faster convergence compared to a constant step size. However, if the channel gains are not stationary processes, e.g., the channel statistics vary due to e.g. a changing mobility of the nodes, Algorithm 3 should update $\tilde{\mu}(i)$ on a regular basis in order to track the possible changes of the channel statistics [23]. For this case, a constant step size is preferable over adaptive step sizes. The convergence behavior of Algorithm 3 depends mainly on the choice of the step size $\lambda_j(i), j \in \{1, 2, r\}$, and some examples are given in Section V.

**B. Delay-Constrained Transmission**

The assumption of infinite-size buffers at the relay, i.e., $Q^\text{max}_j \to \infty, j = 1, 2$, is necessary to achieve the outer bound $R^\text{out}$, cf. Lemma 2. However, the transmission strategies provided in Lemma 2 for infinite-size buffers do not lead to a bounded average delay. Hence, the protocol in Lemma 2 is delay-unconstrained and provides a performance upper bound for delay-constrained transmission. More precisely, the advantages of buffering and adaptive mode selection come at the expense of an increased end-to-end delay. However, our numerical results in Section VII show that with the simple heuristic modifications proposed in this subsection, the protocol for delay-unconstrained transmission can be also employed for delay-constrained transmission at the expense of a small performance degradation due to the delay constraint.

Specifically, we modify the protocol in Lemma 2 for limited-size buffers at the relay and by choosing $Q^\text{max}_j$ and $Q^\text{max}_2$ appropriately, the average delay can be limited to any desired value.

Let $T_j(i), j = 1, 2$, denote the waiting time that a bit transmitted from user $j$ in the $i$-th time slot stays in buffer $B_j$ before it is transmitted to the respective user. Then, according to Little’s Law [31], the average waiting time/delay is given by

\[
E\{T_j\} = \frac{E\{Q_j\}}{E\{R_j\}}, \quad j = 1, 2. \tag{26}
\]

For infinite-size buffers, the maximum average arrival rates, i.e., $E\{R_j\}, j = 1, 2$, are bounded. However, although, with the rate selection policy in Lemma 2, the queues are stable, i.e., $R_1 = R_2$ and $R_2 = R_1$, we cannot guarantee that $E\{Q_j\}, j = 1, 2$, is bounded, and consequently, we cannot guarantee a delay-constrained transmission.

On the other hand, for a buffer of size $Q^\text{max}_j$, the average size of the queue is bounded, i.e., $E\{Q_j\} < Q^\text{max}_j$. However, the protocols in Theorems 1 and 2 do not consider the effect of the states of the queues for transmission mode selection which leads to a performance degradation for limited-size buffers.

For instance, if $Q^\text{max}_1 - Q_1(i - 1) < R_1(s)$ and/or $Q^\text{max}_2 - Q_2(i - 1) < R_2(s)$ occur, the multiple-access mode users cannot transmit to the relay with rates $R_1(s)$ and $R_2(s)$. On the other hand, if $Q_1(i - 1) < R_2(s)$ and/or $Q_2(i - 1) < R_1(s)$ occur for the broadcast mode, the relay cannot transmit to the users with rates $R_1(s)$ and $R_2(s)$. Although these effects are negligible for infinite-size buffers, they can considerably reduce the achievable rates for the case of finite-size buffers. Motivated by these observations, we propose the following modified protocol for fixed node transmit powers which guarantees a delay-constrained transmission. A similar delay-constrained protocol can be also developed for per-node long-term power constraints.

**Delay-Constrained Protocol:** For finite-size buffers at the relay, a delay-constrained protocol for the half-duplex bidirectional relay channel with AWGN, block fading, and a fixed transmit power for each node is obtained by replacing the capacity terms $C(s) = \{C_{1r}(s), C_{2r}(s), C_{r1}(s), C_{r2}(s), C_r(s)\}$ in Theorem 1, Corollary 1, and Lemma 2 by the following virtual capacities

\[
\tilde{C}_{1r}(s) = \min\{C_{1r}(s), Q^\text{max}_1 - Q_1(i - 1)\} \tag{27a}
\]

\[
\tilde{C}_{2r}(s) = \min\{C_{2r}(s), Q^\text{max}_2 - Q_2(i - 1)\} \tag{27b}
\]

\[
\tilde{C}_{r1}(s) = \min\{C_{r1}(s), Q_2(i - 1)\} \tag{27c}
\]

\[
\tilde{C}_{r2}(s) = \min\{C_{r2}(s), Q_1(i - 1)\} \tag{27d}
\]

\[
\tilde{C}_r(s) = \min\{C_r(s), Q^\text{max}_1 + Q^\text{max}_2\}
\]

\[
\tilde{C}_{1r}(s) = \min\{C_{1r}(s), Q^\text{max}_1 - Q_1(i - 1)\} \tag{27a}
\]

\[
\tilde{C}_{2r}(s) = \min\{C_{2r}(s), Q^\text{max}_2 - Q_2(i - 1)\} \tag{27b}
\]

\[
\tilde{C}_{r1}(s) = \min\{C_{r1}(s), Q_2(i - 1)\} \tag{27c}
\]

\[
\tilde{C}_{r2}(s) = \min\{C_{r2}(s), Q_1(i - 1)\} \tag{27d}
\]

\[
\tilde{C}_r(s) = \min\{C_r(s), Q^\text{max}_1 + Q^\text{max}_2\}
\]
As stated before, by appropriately choosing $Q^\text{max}_1$ and $Q^\text{max}_2$, the average delays of the transmissions from user 1-to-user 2 and user 2-to-user 1 can be limited to any desired value larger than one time slot, respectively. In particular, if the target delays are sufficiently large, we obtain $(\bar{R}_{1r}, \bar{R}_{2r}) \rightarrow (\bar{R}^*_1, \bar{R}^*_2)$, where $(\bar{R}^*_1, \bar{R}^*_2)$ is the achievable long-term rate pair for delay-unconstrained transmission. Moreover, since $E\{Q_j\} \leq Q^\text{max}_j$, $j = 1, 2$ holds, the average delays are upper bounded by $E\{T_1\} \leq Q^\text{max}_1/\bar{R}_{1r} \approx Q^\text{max}_1/\bar{R}^*_1$ and $E\{T_2\} \leq Q^\text{max}_2/\bar{R}_{2r} \approx Q^\text{max}_2/\bar{R}^*_2$. Therefore, if we choose $Q^\text{max}_1 = I\bar{R}^*_1$, and $Q^\text{max}_2 = I\bar{R}^*_2$, the average delays are upper bounded by 1 time slot.

**Remark 12:** We note that for the proposed delay-constrained protocol, Algorithm 3 is more effective for computing constant $\mu$ than Algorithm 1. The reason is that the value of $\mu$ which is obtained from Algorithm 1 is optimal for the infinite-size buffers. In particular, the effect of the finite-size buffers is not included in Algorithm 1 while the effect of the finite buffer sizes can be incorporated in Algorithm 3 by calculating the historical average capacities based on the virtual instantaneous capacities in (27). Hence, we employ Algorithm 3 to find constant $\mu$ in the proposed delay-constrained protocol.

**Remark 13:** Although the above modifications are heuristic, the numerical results in Section V confirm their effectiveness for the considered set of parameters. Nevertheless, finding the optimal protocol, i.e., optimal mode selection, rate selection, and/or power allocation policies, for a given average delay requirement is an interesting problem for future work.

## V. Numerical Results

In this section, we evaluate the achievable rate region obtained with the proposed protocols for the bidirectional relay channel with block fading. We assume Rayleigh fading, i.e., the channel gains $s_1$ and $s_2$ follow exponential distributions with means $\Omega_1$ and $\Omega_2$, respectively, and change independently from one time slot to the next. In the following, we first introduce benchmark schemes which are used to evaluate the performance of the proposed protocols. Subsequently, we provide numerical results for both proposed protocols and the benchmark schemes.

### A. Benchmark Schemes

As benchmark schemes, we consider protocols which use a fixed schedule for using the possible transmission modes. This is in contrast to the proposed protocols, cf. Lemma 2 and Theorems 1 and 2, which optimally employ adaptive mode selection in each time slot. We also refer to the benchmark schemes as conventional protocols in the following. We consider two types of conventional protocols as benchmark schemes: 1) Conventional protocols with delay-constrained transmission, and 2) conventional protocols with delay-unconstrained transmission. In particular, in the protocols considered in [2]–[5], [7], [9], [32], the users transmit to the relay for a fraction of each time slot and the relay forwards the information to the respective users in the remaining fraction of the time slot. These protocols lead to delay-constrained transmissions since the information stays in the buffers at the relay for less than one time slot. In order to make the comparison fair with respect to the delay, we also consider benchmark schemes that exploit the buffering capability. In this case, although the benchmark protocols have a fixed and predetermined schedule of using the transmission modes, the users are allowed to transmit to the relay in multiple time slots as long as the buffers at the relay are not full, and the relay forwards the information to the respective users using again multiple time slots. As a performance bound for this protocol, we assume $Q^\text{max}_j \rightarrow \infty$, $j = 1, 2$, i.e., infinite-size buffers at the relay. These protocols can be interpreted as extensions of the protocol proposed in [14] for one-way relaying.

For the conventional protocols with delay-constrained transmission, the transmission is accomplished in one time slot. In particular, based on the CSI of the involved links, a fraction of the time slot is allocated to each transmission mode in each time slot. Let us define $\Delta = [\Delta_{\text{MAC}}(s), \Delta_{\text{BC}}(s)]$, $\forall s$, where $\Delta_{\text{MAC}}(s)$ and $\Delta_{\text{BC}}(s)$ are the fractions of the time slot which are allocated to the multiple-access and broadcast modes, respectively. Moreover, we define $R = [R_{12}(s), R_{21}(s)]$, $\forall s$, where $R_{12}(s)$ and $R_{21}(s)$ denote the transmission rates for the user 1-to-user 2 and user 2-to-user 1 information flows, respectively. Furthermore, the achievable long-term rates for the user 1-to-user 2 and user 2-to-user 1 information flows are given by $\bar{R}_{12} = E\{R_{12}(s)\}$ and $\bar{R}_{21} = E\{R_{21}(s)\}$, respectively. The boundary surface of the achievable rate region of the considered conventional protocols with delay-constrained transmission is obtained from the following optimization problem

\[
\begin{align*}
\text{maximize} & \quad \eta \bar{R}_{12} + (1 - \eta) \bar{R}_{21} \\
\text{subject to} & \quad C_1 : \ R_{12}(s) \leq \min\{\Delta_{\text{MAC}}(s)C_1(s), \Delta_{\text{BC}}(s)C_2(s)\} \\
& \quad C_2 : \ R_{21}(s) \leq \min\{\Delta_{\text{MAC}}(s)C_2(s), \Delta_{\text{BC}}(s)C_1(s)\} \\
& \quad C_3 : \ R_{12}(s) + R_{21}(s) \leq \Delta_{\text{MAC}}(s)C_P(s), \Delta_{\text{BC}}(s)C_P(s), (28)
\end{align*}
\]

where $D$ is the feasible set for $\Delta$ which imposes constraints $0 \leq \Delta_k(s) \leq 1$, $k \in \{\text{MAC, BC}\}$ and $\Delta_{\text{MAC}}(s) + \Delta_{\text{BC}}(s) = 1$, $\forall s$.

On the other hand, in the conventional protocols with delay-unconstrained transmission, the buffers at the relay are filled by using the multiple-access mode and then the relay forwards the information to the users using the broadcast mode. We define $\Delta = [\Delta_{\text{MAC}}, \Delta_{\text{BC}}]$ with $\Delta_k = \lim_{N \to \infty} \frac{N_k}{N}$, $k \in \{\text{MAC, BC}\}$ and $N_{\text{MAC}}$ and $N_{\text{BC}}$ are the number of time slots allocated to the multiple-access and broadcast modes, respectively. Assuming $Q^\text{max}_j \to \infty$, $j = 1, 2$, for these protocols, the value of $\Delta_k$ depends only on the long-term statistics of the channel, the transmit powers of the nodes, and the value of $\eta$. Furthermore, we define $\bar{C}_1 = \Delta_{\text{MAC}}E\{C_1(s)\}$, $\bar{C}_2 = \Delta_{\text{MAC}}E\{C_2(s)\}$, $\bar{C}_1 = \Delta_{\text{BC}}E\{C_1(s)\}$, $\bar{C}_2 = \Delta_{\text{BC}}E\{C_2(s)\}$, and $\bar{C}_j = \Delta_{\text{MAC}}E\{C_j(s)\}$. Thereby, the boundary surface of the achievable rate region of the conventional protocols with delay-unconstrained transmission is
obtained from the following optimization problem

\[
\begin{align*}
\text{maximize} & \quad \eta \tilde{R}_{12} + (1 - \eta) \tilde{R}_{21} \\
\text{subject to} \quad & C1: \quad \tilde{R}_{12} \leq \min\{C_{1r}, \tilde{C}_{r2}\} \\
& C2: \quad \tilde{R}_{21} \leq \min\{C_{2r}, \tilde{C}_{r1}\} \\
& C3: \quad \tilde{R}_{12} + \tilde{R}_{21} \leq \tilde{C}_r,
\end{align*}
\]

where \( D \) is the feasible set for \( \Delta \) which imposes constraints

\[\Delta \in D, \ P \in P, R \geq 0.\]

where \( \Delta \) is the feasible set for \( \Delta \) which imposes constraints

\[0 \leq \Delta_k \leq 1, \ k \in \{MAC, BC\} \text{ and } \Delta_{MAC} + \Delta_{BC} = 1.\]

B. Evaluation of the Proposed Protocols and the Benchmark Schemes

We first evaluate the performance of the proposed protocols assuming infinite-size buffers, cf. Lemma 2, Theorem 1, and Theorem 2. In order to numerically evaluate the proposed protocols, we first obtain constants \( \mu \) and/or \( \gamma \) from Propositions 1 or 2 by using the optimal mode selection and/or power allocation policies in Theorems 1 and 2 for given channel statistics, node power/power budgets, and \( \eta \). In Propositions 1 or 2, we employ step sizes \( \lambda_j[m] = \frac{0.01}{\sqrt{m}}, \ j \in \{1, 2, r\} \) and \( \kappa_j[m] = \frac{0.001}{\sqrt{m}}, \ j \in \{1, 2, r\} \). Then, equipped with \( \mu^* \) and \( \gamma^* \), we are able to numerically compute \( \mathbf{C}^* \) from (11).

The optimal rate pair \( (\tilde{R}_{12}^*, \tilde{R}_{21}^*) \) can be obtained from (12) as a function of \( \mathbf{C}^* \). To evaluate the proposed protocols by simulation, we adopt the proposed transmission strategies, i.e., the optimal mode selection and power allocation policies in Theorems 1 and 2, and the optimal rate selection policy in Lemma 2, and simulate transmission for \( N = 10^5 \) time slots. Thereby, the effect of the queues in (2) and (3) is also taken into account. Finally, the rate pair \( (\tilde{R}_{12}^*, \tilde{R}_{21}^*) \) is calculated as the average amount of information which is received at the destinations. For the benchmark scheme, we consider the conventional protocols with delay-unconstrained transmission. The boundary surface of the achievable long-term rate region for the conventional protocols are obtained by solving the optimization problem in (29) using the CVX solver [33].

Before presenting the performance results, in Fig. 6, we show the convergence behavior of the iterative algorithm proposed in Proposition 1. We plot the value of dual variable \( \mu_r \) versus the iteration number \( m \) for different step sizes. We assume average channel gains of \( \Omega_1 = \Omega_2 = 1 \), transmit powers \( P_1 = P_2 = P_r = 10 \text{ dB} \), and \( \eta = \frac{1}{2} \), i.e., the sum rate point. We adopt both square summable but not absolutely summable step sizes, i.e., \( \lambda[m] = \frac{0.10}{\sqrt{m}}, \frac{0.05}{\sqrt{m}}, \frac{0.001}{\sqrt{m}} \), and nonsummable diminishing step sizes, i.e., \( \lambda[m] = \frac{0.10}{\sqrt{m}}, \frac{0.05}{\sqrt{m}}, \frac{0.001}{\sqrt{m}} \), cf. [27]. We observe that for the considered choices of the step sizes, the dual variable obtained from Proposition 1 converges to \( \mu_r^* = 0.3 \). Moreover, from Fig. 6, we can conclude that the convergence behavior of the adaptive algorithm in Proposition 1 depends on the choice of the step sizes. For example, smaller step sizes lead to a slower convergence; compare e.g., the curve for \( \lambda[m] = \frac{0.05}{\sqrt{m}} \) with the curve for \( \lambda[m] = \frac{0.001}{\sqrt{m}} \).

However, increasing the value of the step size may lead to oscillation and even divergence, see the curve for \( \lambda[m] = \frac{0.10}{\sqrt{m}} \). Fig. 7 shows the achievable long-term rate regions of the proposed protocols and the conventional protocols for both types of power constraints considered in this paper. We assume transmit powers \( P_1 = P_2 = P_r = 0.10 \text{ dB} \) for the case of fixed node transmit powers, power budgets \( P_{\text{max}} = P_{\text{max}} = P_{\text{r}} = 0 \text{ dB} \) for the long-term per-node power constraint, and symmetric channels with average gains \( \Omega_1 = \Omega_2 = 1 \). We observe that the rate regions of the conventional protocols are in the interior of the rate regions of the proposed protocols for both power constraints. Moreover, a considerable gain is obtained by adaptive mode selection compared to using the modes according to a fixed schedule. From Fig. 7, we can conclude that the gain of adaptive mode selection is higher for \( \eta \to 0, 1 \) compared to \( \eta = \frac{1}{2} \). Furthermore, by comparing the performance gaps between the proposed protocols and the conventional protocols, we observe that applying joint adaptive mode selection and power allocation is significantly more effective than applying only adaptive mode selection or power allocation especially for the sum rate point. Additionally, we observe that power allocation leads to larger performance gains in the low signal-to-noise (SNR) regime compared to the high SNR regime. For example, let \( \tilde{R}_{\text{sum}}^{\text{fix}} \) and \( \tilde{R}_{\text{sum}}^{\text{adapt}} \) denote the maximum sum rate of the proposed protocols for fixed node transmit powers and the per-node long-term power constraint, respectively. Then, we obtain \( \tilde{R}_{\text{sum}}^{\text{adapt}} / \tilde{R}_{\text{sum}}^{\text{fix}} = 1.362 / 2.492 = 1.34 \) for \( P_j = P_{\text{max}} = 10 \text{ dB} \) and \( \tilde{R}_{\text{sum}}^{\text{adapt}} / \tilde{R}_{\text{sum}}^{\text{fix}} = 1.459 / 0.911 = 1.55 \) for \( P_j = P_{\text{max}} = 0 \text{ dB} \). Finally, we observe a perfect match between numerical and simulation results.

In many practical scenarios, the nodes have different capabilities as far as the transmit powers are concerned. Hence, in Fig. 8, we plot the achievable long-term rate regions for symmetric channels with average gains \( \Omega_1 = \Omega_2 = 1 \) and the following fixed node transmit power scenarios: 1) \( P_1 = P_2 = P_r = 10 \text{ dB} \) and \( P_r = 0 \text{ dB} \) for applications where the relay can afford less transmit power than the users, e.g., satellite communication where the satellite and the earth stations represent the relay and the users, respectively, and 2)
Due to the symmetry of the information flows, we obtain achievable long-term rate region corresponding to two extreme (corner) points on the boundary surface of the infinite-size buffers. This is indicated by a perfect match between numerical and simulation results.

In these cases, the proposed protocol with adaptive mode selection considerably outperforms the conventional protocols and the match between numerical and simulation results is perfect. Moreover, let us define \( (\hat{R}_{12}^{\text{max}}, \hat{R}_{21}^{\text{min}}) \) and \( (\hat{R}_{12}^{\text{min}}, \hat{R}_{21}^{\text{max}}) \) as the two extreme (corner) points on the boundary surface of the achievable long-term rate region corresponding to \( \eta = 0, 1 \). Due to the symmetry of the information flows, we obtain \( \hat{R}_{12}^{\text{max}} = \hat{R}_{21}^{\text{max}} \) and \( \hat{R}_{12}^{\text{min}} = \hat{R}_{21}^{\text{min}} \). Moreover, the achievable long-term rates for both information flows, \( \hat{R}_{12}^{\text{max}} \) and \( \hat{R}_{21}^{\text{max}} \), are the same for both transmit power scenarios. However, the highest achievable rate that can be provided for one information flow while achieving the maximum transmission rate for the other information flow, i.e., \( \hat{R}_{12}^{\text{min}} \) and \( \hat{R}_{21}^{\text{min}} \), is higher for transmit power scenario 2 than for transmit power scenario 1.

Next, we investigate the case when the channel statistics are not known and employ Algorithm 3 to find \( \mu \). In particular, in Fig. 9, we plot the stochastic estimate of \( \mu \) versus the
We consider both symmetric channels with average gains $\Omega_1 = \Omega_2 = 1$ and asymmetric channels with average gains $\Omega_1 = \frac{3}{2}$ and $\Omega_2 = \frac{1}{2}$. We observe that with average delays of five and ten time slots around 93% and 97% of the rate with infinite delay are achieved, respectively. Therefore, the proposed heuristic protocol for delay-constrained transmission is indeed efficient for the considered set of parameters. Even for an average delay of only 5 time slots, a considerable gain is obtained by the proposed protocol compared to the conventional protocol which causes a delay of one time slot. For the case of asymmetric channels, we observe that the rate region is asymmetric in favor of the rate of the user 1-to-user 2 information flow, i.e., $R^{\text{min}}_{21} < R^{\text{min}}_{12}$. The main reason for this behavior is that for the corner point $(R^{\text{max}}_{12}, R^{\text{max}}_{21})$, we obtain $p_t = 1$, i.e., the relay always first decodes the codeword received from user 2 while treating the codeword received from the user 1 as noise. While for the corner point $(R^{\text{min}}_{12}, R^{\text{min}}_{21})$, the relay always first decodes the codeword received from the user 1 while treating the codeword received from the user 2 as noise. Since the user 1-to-relay link is statistically stronger than the user 2-to-relay link, treating the codeword received from user 1 as noise reduces the average rate $R^{\text{min}}_{21}$ considerably compared to treating the codeword received from user 2 as noise for the average rate $R^{\text{min}}_{12}$.

We also note that the maximum achievable rate for one-way relaying in one direction is the same as that in the other direction for both symmetric and asymmetric channels, i.e., $R^{\text{max}}_{12} = R^{\text{max}}_{21}$. Note that for the corner point $(R^{\text{max}}_{12}, R^{\text{max}}_{21})$, the rate of the user 1-to-user 2 information flow is limited by the capacity rates $C_1(s)$ and $C_2(s)$, if $q_{\text{MAC}}(s) = 1$ and $q_{\text{BC}}(s) = 1$ hold, respectively. In contrast, for the corner point $(R^{\text{min}}_{12}, R^{\text{min}}_{21})$, the rate of the user 2-to-user 1 information flow is limited by the capacity rates $C_2(s)$ and $C_1(s)$, if $q_{\text{MAC}}(s) = 1$ and $q_{\text{BC}}(s) = 1$ hold, respectively. Since the transmit powers of the nodes are identical, the rates of both information flows are limited by the same values of the capacities, i.e., $C_1(s) = C_{12}(s)$ and $C_2(s) = C_{21}(s)$, which leads to $R^{\text{max}}_{12} = R^{\text{max}}_{21}$.

VI. CONCLUSION

In this paper, we considered the half-duplex bidirectional buffer-aided relay channel with block fading and proposed protocols which are not restricted to adhere to a predefined schedule for using the available transmission modes. In particular, the proposed protocols specify the optimal transmission strategy, i.e., the optimal transmission mode, transmission rates, and/or transmit powers of the nodes, in each time slot based on the instantaneous CSI of the involved links and their long-term statistics. The optimal transmission strategy was obtained such that the weighted long-term sum rate is maximized where each weight corresponds to a point on the boundary surface of the achievable long-term rate region of the proposed protocol. To enable adaptive mode selection, the relay has to be equipped with two buffers for storage of the information received from both users. Moreover, we considered both fixed node transmit powers as well as per-node long-term power constraints. For the cases when the channel statistics are not known a priori, we proposed an...
adaptive algorithm to calculate the long-term variables of the proposed protocols. This algorithm only employs instantaneous CSI and does not require knowledge of the channel statistics. Furthermore, as data buffering increases the end-to-end delay, we also developed a delay-constrained protocol by taking into account the effect of finite-size buffers in the mode selection policy. Simulation results confirmed that the proposed protocols outperform protocols with a fixed schedule of transmission for all points of the rate region.

**APPENDIX A**

**PROOF OF THEOREM 1**

In this appendix, our aim is to find the optimal mode selection policy as a solution of the relaxed version of the optimization problem given in (7) assuming a fixed transmit power for each node. Since the cost function and the constraints in (7) are affine in the optimization variables \( q \in \mathcal{Q} \) and \( \mathbf{R} \) and the feasible set is nonempty, Slater’s condition is satisfied, hence, the duality gap is zero [26]. Therefore, the solution of the primal problem in (7) can be found from the solution of the dual problem of (7) [26]. Denoting the Lagrange multipliers corresponding to constraints \( \bar{R}_{12} \leq C_{1r} \), \( \bar{R}_{12} \leq C_{2r} \), \( R_{21} \leq C_{1r} \), \( R_{21} \leq C_{2r} \), \( \bar{R}_{12} + R_{21} \leq C_{r} \), by \( \mu_{1r} \), \( \mu_{2r} \), \( \mu_{r1} \), \( \mu_{r2} \), \( \mu_r \), respectively, the Lagrangian function corresponding to the optimization problem in (7) is obtained as

\[
\mathcal{L}(q, \mathbf{R}, \hat{\mu}) = \eta \bar{R}_{12} + (1 - \eta) \bar{R}_{21} + \mu_{1r}(C_{1r} - \bar{R}_{12}) + \mu_{2r}(C_{2r} - \bar{R}_{21}) + \mu_{r1}(C_{r} - \bar{R}_{12} - \bar{R}_{21}) + \mu_{r2}(C_{r} - R_{21} - \bar{R}_{12})
\]

where \( \hat{\mu} = [\mu_{1r}, \mu_{2r}, \mu_r, \mu_{r1}, \mu_{r2}] \). The dual function is then given by

\[
\mathcal{D}(\hat{\mu}) = \max_{q \in Q, R \geq 0} \mathcal{L}(q, \mathbf{R}, \hat{\mu})
\]

and the dual problem is given by

\[
\min_{\hat{\mu} \geq 0} \mathcal{D}(\hat{\mu}). \tag{32}
\]

To solve (7) using the dual problem in (32), we first obtain the primal variables \( q \) and \( \mathbf{R} \) for a given dual variable \( \mu \) from (31). Then, the optimal \( \hat{\mu} \) is obtained in Appendix B by solving the dual problem in (32). The optimal rate pair, \( (\bar{R}_{12}, \bar{R}_{21}) \), and the optimal mode selection policy, \( q^*_s(s) \), \( \forall s, k \), are either the boundary points of the feasible sets, i.e., \( q \in \mathcal{Q} \) and \( \mathbf{R} \geq 0 \), or the stationary points which can be obtained by setting the derivatives of the Lagrangian function in (30) with respect to \( \bar{R}_{12}, \bar{R}_{21}, \) and \( q^*_s(s) \) to zero. The derivatives of the Lagrangian function in (30) are given by

\[
\frac{\partial \mathcal{L}}{\partial \bar{R}_{12}} = \eta - \mu_{1r} - \mu_{2r} - \mu_r \tag{33a}
\]

\[
\frac{\partial \mathcal{L}}{\partial \bar{R}_{21}} = 1 - \eta - \mu_{2r} - \mu_{r1} - \mu_r \tag{33b}
\]

\[
\frac{\partial \mathcal{L}}{\partial q^*_s(s)} = \Pr\{s\} [\mu_{1r} C_{1r}(s) + \mu_{2r} C_{2r}(s) + \mu_r C_r(s)] \tag{33c}
\]

\[
\frac{\partial \mathcal{L}}{\partial q^*_r(s)} = \Pr\{s\} [\mu_{1r} C_{1r}(s) + \mu_{2r} C_{2r}(s)]. \tag{33d}
\]

If the derivatives \( \frac{\partial \mathcal{L}}{\partial \bar{R}_{12}} \) and \( \frac{\partial \mathcal{L}}{\partial \bar{R}_{21}} \) are non-zero, the optimal values of \( \mathbf{R}^* \) are at the boundary of the feasible set for \( \mathbf{R} \), i.e., \( \bar{R}_{12}, \bar{R}_{21} \rightarrow \infty \) or \( \bar{R}_{12}, \bar{R}_{21} \rightarrow 0 \), respectively, which cannot be the optimal solution. Therefore, the derivatives \( \frac{\partial \mathcal{L}}{\partial \bar{R}_{12}} \) and \( \frac{\partial \mathcal{L}}{\partial \bar{R}_{21}} \) in (33a) and (33b) have to be zero which leads to

\[
\mu_{1r} = \eta - \mu_{2r} - \mu_r \triangleq \mu_1 \quad \text{and} \quad \mu_{2r} = 1 - \eta - \mu_{r1} - \mu_r \triangleq \mu_2.
\]

In order to be able calculate the derivative of \( \mathcal{L} \) with respect to \( q^*_s(s) \), we implicitly assume that the fading gains can take discrete values from set \( s \in S_1 \times S_2 \) where \( \Pr\{s\} \) is the probability of fading state \( s \). Note that the continuous distribution can be represented by \( |S_1|, |S_2| \rightarrow \infty \).
a given channel state, the derivative \( \frac{\partial C}{\partial \mu(s)} \) is always positive. However, since \( q_{\text{MAC}}(s) + q_{\text{BC}}(s) = 1 \), \( \forall s \), we obtain

\[
q_{\text{MAC}}^*(s) = \begin{cases} 
1, & \text{if } \Lambda_{\text{MAC}}(s) > \Lambda_{\text{BC}}(s) \\
0, & \text{if } \Lambda_{\text{MAC}}(s) < \Lambda_{\text{BC}}(s),
\end{cases}
\]  

(34)

where \( \Lambda_{\text{MAC}}(s) = \mu_1 C_1(s_1) + \mu_2 C_2(s_2) + \mu_\gamma C_\gamma(s) \) and \( \Lambda_{\text{BC}}(s) = \mu_1 C_1(s_1) + \mu_2 C_2(s_2) \) are referred to as selection metrics. Note that for ergodic fading with continuous probability density function, the probability that \( \Lambda_{\text{MAC}}(s) = \Lambda_{\text{BC}}(s) \) holds is zero except for the following two cases: 1) \( \mu_1 = \gamma \neq 0 \) and \( P_1 = P_2 \), and 2) \( \mu_2 = \mu_\gamma = 0 \). The former case corresponds to channels where \( \Omega_1 \gg \Omega_2 \) and the latter case to channels where \( \Omega_1 \ll \Omega_2 \). For these cases, the optimal selection policy is to select the multiple-access and broadcast modes with equal probability. This leads to the optimal mode selection policy in (8). Moreover, since \( \mu_1 = \gamma - \mu_2 \neq \mu_\gamma \neq \mu_1 \), \( \mu_2 = 1 - \gamma - \mu_1 \neq \mu_\gamma \neq \mu_2 \), we give the optimal selection policy in (8) in terms of \( \mu = [\mu_1, \mu_2, \mu_\gamma] \) instead of in terms of \( \mu_\gamma \). This completes the proof.

**APPENDIX B**

**Proof of Proposition 1**

The optimal values of \( \mu_1 \), \( \mu_2 \), and \( \mu_\gamma \) are obtained by solving the dual problem in (32). In particular, substituting, the optimal values of \( q \) in (34) for a given \( \mu \) into (30), we obtain the dual function in (31) as

\[
D(\mu) = \eta C_r^*(s) + (1 - \eta) \tilde{R}^*(s) + \mu_1 (C_1^*(s) - \tilde{C}_r^*(s)) + \mu_2 (\tilde{C}_1^* - \tilde{C}_r^*(s)) + \mu_\gamma (\tilde{C}_r^*(s) - \tilde{C}_1^*(s) - \tilde{C}_2^*(s)).
\] 

(35)

In order to solve the dual problem in (32), we use the well-known sub-gradient method [27], [28]. The sub-gradient method reduces to a gradient-based search if the functions in the optimization problem are differentiable. The idea is to minimize \( D(\mu) \) by updating all the components of \( \mu \) at the same time along the subgradient search directions. The updates are performed as specified in Algorithm 1 with the updates given in (10). Note that, \( \mu_1 = \gamma - \mu_2 \neq \mu_\gamma \neq \mu_1 \) and \( \mu_2 = 1 - \gamma - \mu_1 \neq \mu_\gamma \neq \mu_2 \). The subgradient method is guaranteed to converge to the optimal dual variable \( \mu^* \) provided that the step sizes \( \lambda_j[m] \), \( \forall j, m \) are chosen sufficiently small [27], [28]. This completes the proof.

**APPENDIX C**

**Proof of Theorem 2**

In this appendix, our aim is to obtain the optimal mode selection and power allocation policies by solving the relaxed version of the optimization problem in (7) under the per-node long-term power constraint in (4). Since this optimization problem is convex in the optimization variables \( q, P, \) and \( \bar{R}, \) cf. Lemma 1, and the feasible set is nonempty, Slater’s condition is satisfied. Therefore, the duality gap is zero [26] which allows us to solve the dual problem of (7) instead of directly solving the primal problem. Denoting the Lagrange multipliers corresponding to the power constraint in (4) by \( \gamma_j \), \( j = \{1, 2, r\} \), the Lagrangian function corresponding to the optimization problem in (7) is obtained as

\[
\mathcal{L}(q, P, \bar{R}, \bar{\mu}, \gamma) = \eta \tilde{R}_1 + (1 - \eta) \tilde{R}_2 + \mu_1 (\tilde{C}_1 - \tilde{R}_1) + \mu_2 (\tilde{C}_2 - \tilde{R}_2) + \mu_\gamma (\tilde{C}_\gamma - \tilde{R}_\gamma) + \gamma_1 (P_1^\mu - \tilde{P}_1) + \gamma_2 (P_2^\mu - \tilde{P}_2) + \gamma_3 (P_r^\mu - \tilde{P}_r),
\]

(36)

where \( \gamma = [\gamma_1, \gamma_2, \gamma_3] \). The dual function is then given by

\[
D(\bar{\mu}, \gamma) = \max_{q \in \mathbb{Q}, \bar{R} \geq 0} \min_{P \geq 0} \mathcal{L}(q, P, \bar{R}, \bar{\mu}, \gamma)
\]

(37)

and the dual problem is

\[
\min_{\bar{\mu} \geq 0, \gamma \geq 0} D(\bar{\mu}, \gamma).
\]

(38)

To solve (7) using the dual problem in (38), we first obtain the primal variables \( q, P, \bar{R} \) for given dual variables \( \bar{\mu} \) and \( \gamma \) from (37). Then, we solve the dual problem in (37) in Appendix D to find the optimal values of \( \bar{\mu} \) and \( \gamma \). The optimal values of \( \tilde{R}_1, \tilde{R}_2, q_j, q_k, q_l, \forall j, k, l \) are either boundary points of the corresponding feasible set or stationary points which can be obtained by setting the derivatives of the Lagrangian function in (36) with respect to \( \tilde{R}_1, \tilde{R}_2, q_j \), and \( P_j \) to zero. The derivatives with respect to \( \tilde{R}_1, \tilde{R}_2, q_j \) lead to (33a) and (33b), respectively. Hence, similar to the case considered in Appendix A, we obtain \( \mu_1 = \gamma - \mu_2 - \mu_\gamma \neq \mu_1 \) and \( \mu_2 = 1 - \gamma - \mu_1 - \mu_\gamma \neq \mu_2 \). The derivatives of the Lagrangian function in (36) with respect to \( q_{\text{MAC}}(s) \) and \( q_{\text{BC}}(s) \) lead to

\[
\frac{\partial \mathcal{L}}{\partial q_{\text{MAC}}(s)} = \Pr(s) \left[ \mu_1 \log_2 \left( 1 + \frac{P_1(s) s_1}{q_{\text{MAC}}(s)} \right) \right]
\]

\[
\frac{\partial \mathcal{L}}{\partial q_{\text{BC}}(s)} = \Pr(s) \left[ \mu_2 \log_2 \left( 1 + \frac{P_2(s) s_2}{q_{\text{MC}}(s)} \right) \right]
\]

(39a)

(39b)

respectively. The derivatives \( \frac{\partial \mathcal{L}}{\partial q_{\text{MAC}}(s)} \) and \( \frac{\partial \mathcal{L}}{\partial q_{\text{BC}}(s)} \) are non-
negative\(^9\), however, since \(q_{\text{MAC}}(s) + q_{\text{BC}}(s) = 1\), \(\forall s\), we obtain
\[
q^*_{\text{MAC}}(s) = \begin{cases} 
1, & \text{if } \Lambda_{\text{MAC}}(s) > \Lambda_{\text{BC}}(s) \\
0, & \text{if } \Lambda_{\text{MAC}}(s) < \Lambda_{\text{BC}}(s)
\end{cases} 
\] (40)
where
\[
\Lambda_{\text{MAC}}(s) = \mu_{1r} \log_2 \left(1 + \frac{P_1(s)s_1}{q_{\text{MAC}}(s)}\right) + \mu_{2r} \log_2 \left(1 + \frac{P_2(s)s_2}{q_{\text{MAC}}(s)}\right) + \mu_r \log_2 \left(1 + \frac{P_r(s)s_1 + P_2(s)s_2}{q_{\text{MAC}}(s)}\right)
\]
\[
\Lambda_{\text{BC}}(s) = \mu_{1r} \log_2 \left(1 + \frac{P_1(s)s_1}{q_{\text{BC}}(s)}\right) + \mu_{2r} \log_2 \left(1 + \frac{P_2(s)s_2}{q_{\text{BC}}(s)}\right) + \mu_r \log_2 \left(1 + \frac{P_r(s)s_1 + P_2(s)s_2}{q_{\text{BC}}(s)}\right)
\]
Since we assumed that the fading gains \(s_1\) and \(s_2\) have continuous probability density functions, the probability that \(\Lambda_{\text{MAC}}(s) = \Lambda_{\text{BC}}(s)\) holds is zero. Hence, without loss of generality, we set \(q^*_{\text{MAC}}(s) = 1\) if \(\Lambda_{\text{MAC}}(s) = \Lambda_{\text{BC}}(s)\) occurs. This leads to the optimal mode selection policy in (15).

Since \(q^*_s(s)\) is either zero or one, we only have to obtain the optimal powers of the users if \(q^*_{\text{MAC}}(s) = 1\) holds and the optimal power of the relay if \(q^*_{\text{BC}}(s) = 1\). If a node is not selected, its optimal power is zero. Therefore, we calculate the derivatives of the Lagrangian function in (36) with respect to \(P_j(s), j \in \{1, 2\}\) if \(q^*_{\text{MAC}}(s) = 1\) and with respect to \(P_r(s)\) if \(q^*_{\text{BC}}(s) = 1\). This leads to
\[
\frac{\partial \mathcal{L}}{\partial P_1(s)} = \Pr[s]\left[\frac{1}{\ln 2} \cdot \frac{\mu_{1r}s_1 + \mu_r s_1}{1 + P_1(s)s_1} - \gamma_1\right] 
\]
(42a)
\[
\frac{\partial \mathcal{L}}{\partial P_2(s)} = \Pr[s]\left[\frac{1}{\ln 2} \cdot \frac{\mu_{2r}s_2 + \mu_r s_2}{1 + P_2(s)s_2} - \gamma_2\right] 
\]
(42b)
\[
\frac{\partial \mathcal{L}}{\partial P_r(s)} = \Pr[s]\left[\frac{1}{\ln 2} \cdot \frac{\mu_{1r}s_1 + \mu_r s_1}{1 + P_r(s)s_1} - \gamma_r\right]. 
\]
(42c)

\(^9\)We can conclude that \(\frac{\partial \mathcal{L}}{\partial q_{\text{MAC}}(s)}\) and \(\frac{\partial \mathcal{L}}{\partial q_{\text{BC}}(s)}\) are non-negative based on the same argument as employed in the proof of Lemma 1. In particular, we can rewrite \(\frac{\partial \mathcal{L}}{\partial q_{\text{MAC}}(s)}\) and \(\frac{\partial \mathcal{L}}{\partial q_{\text{BC}}(s)}\) as sum of functions of the form \(\frac{\partial f(x,y)}{\partial x}\), where \(f(x,y) = x \log (1 + \frac{x}{y})\) and \(x\) and \(y\) represent the mode selection and power variables, respectively. Moreover, from Lemma 1, \(f(x,y)\) is an increasing function in \(x\) for given \(y > 0\) [24, Proposition 1] which leads to \(\frac{\partial \mathcal{L}}{\partial q_{\text{MAC}}(s)} \geq 0\) and \(\frac{\partial \mathcal{L}}{\partial q_{\text{BC}}(s)} \geq 0\).

Setting the above equation to zero, we obtain
\[
\frac{\mu_{1r}s_1}{1 + P_1(s)s_1} + \frac{\mu_r s_1}{1 + P_1(s)s_1} = \gamma_1 \ln 2 
\]
(43a)
\[
\frac{\mu_{2r}s_2}{1 + P_2(s)s_2} + \frac{\mu_r s_2}{1 + P_2(s)s_2} = \gamma_2 \ln 2 
\]
(43b)
\[
\frac{\mu_{1r}s_1}{1 + P_r(s)s_1} + \frac{\mu_r s_1}{1 + P_r(s)s_1} = \gamma_r \ln 2. 
\]
(43c)

Considering that the left hand sides of (43a) and (43b) are monotonically decreasing in both \(P_1(s)\) and \(P_2(s)\), we distinguish the following four mutually exclusive cases to find \(P^*_1(s)\) and \(P^*_2(s)\).

**Case 1:** If \(s_1 < \frac{\gamma_1 \ln 2}{\mu_{1r} + \mu_r}\) and \(s_2 < \frac{\gamma_2 \ln 2}{\mu_{2r} + \mu_r}\) hold, from (42a) and (42b), we obtain \(\frac{\partial \mathcal{L}}{\partial P_1(s)} < 0\) and \(\frac{\partial \mathcal{L}}{\partial P_2(s)} < 0\) for \(P_1(s) \geq 0\) and \(P_2(s) \geq 0\). In other words, solving (43a) and (43b) yield negative values for \(P_1(s)\) and \(P_2(s)\). Hence, the optimal values of \(P_1(s)\) and \(P_2(s)\) have to be at the lower bounds of the feasible set \(P \geq 0\) which leads to \(P^*_1(s) = 0\) and \(P^*_2(s) = 0\).

**Case 2:** If \(s_1 > \frac{\gamma_1 \ln 2}{\mu_{1r} + \mu_r}\) and \(s_2 > \frac{\gamma_2 \ln 2}{\mu_{2r} + \mu_r}\), solving (43a) and (43b) does not yield positive values for both \(P_1(s)\) and \(P_2(s)\). In other words, one or both of the derivatives \(\frac{\partial \mathcal{L}}{\partial P_1(s)}\) and \(\frac{\partial \mathcal{L}}{\partial P_2(s)}\) have to be negative which means that the transmit powers of one or both users have to be zero. If we assume \(P^*_1(s) = 0\) and obtain \(P^*_2(s)\) from (43b), the derivative \(\frac{\partial \mathcal{L}}{\partial P_2(s)}\) in (42a) will be positive. This is not possible since if \(\frac{\partial \mathcal{L}}{\partial P_2(s)} > 0\), the optimal value of \(P_1(s)\) has to be at the upper bound of the feasible set, i.e., \(P^*_1(s) \rightarrow \infty\), which contradicts the earlier assumption of \(P^*_1(s) = 0\). However, if we assume \(P^*_2(s) = 0\) and calculate \(P^*_1(s)\) from (43a) as
\[
P^*_1(s) = \frac{\mu_{1r} + \mu_r}{\gamma_1 \ln 2} - \frac{1}{s_1}, 
\]
(44)
we obtain \(\frac{\partial \mathcal{L}}{\partial P^*_2(s)} < 0\) from (42b). Therefore, \(P_2(s)\) has to be at the lower boundary of the feasible set, i.e., \(P^*_2(s) = 0\), which is consistent with the earlier assumption of \(P^*_2(s) = 0\).

**Case 3:** If \(s_1 < \frac{\gamma_1 \ln 2}{\mu_{1r} + \mu_r}\) and \(s_2 > \frac{\gamma_2 \ln 2}{\mu_{2r} + \mu_r}\) hold, with similar arguments as those given for Case 2, we obtain \(P^*_1(s) = 0\) and
\[
P^*_2(s) = \frac{\mu_{2r} + \mu_r}{\gamma_2 \ln 2} - \frac{1}{s_2}, 
\]
(45)
where \(d\) is chosen to satisfy
\[
\frac{\mu_{1r}s_1}{\gamma_1 \ln 2 - s_1 d} + \frac{\mu_{2r}s_2}{\gamma_2 \ln 2 - s_2 d} = \frac{\mu_r}{d} + 1. 
\]
(47)

Equation (43c) is a quadratic equation and has two solutions for \(P^*_r(s)\) in general. However, since \(P^*_r(s) \geq 0\) has to hold, the left hand side of (43c) is monotonically decreasing in \(P^*_r(s)\). Thus, the maximum value of the left hand side of (43c) occurs when \(P^*_r(s) = 0\) which leads to the necessary condition
$\mu_{r_1}s_1 + \mu_{r_2}s_2 > \gamma_r \ln 2$ for obtaining a unique positive solution for $P_1(s)$. In particular, if $\mu_{r_1}s_1 + \mu_{r_2}s_2 < \gamma_r \ln 2$ holds, both of the two roots of (43c) are negative, and if $\mu_{r_1}s_1 + \mu_{r_2}s_2 < \gamma_r \ln 2$ holds, one of the roots of (43c) is negative and the other one is positive. Since the relay power has to be positive, we select the maximum of the roots of (43c) if it is positive. Thus, we obtain
\[
P^*_{r_1}(s) = \left[ \frac{-b + \sqrt{b^2 - 4ac}}{2a} \right]^+, \tag{48}
\]
where $a = \gamma_r \ln 2s_1s_2$, $b = \gamma_r \ln 2(s_1 + s_2) - (\mu_{r_1} + \mu_{r_2})s_1s_2$, and $c = \gamma_r \ln 2 - \mu_{r_1}s_1 - \mu_{r_2}s_2$.

Since $\mu_{r_1} = \eta - \mu_{r_2} - \mu_r \triangleq \mu_1$ and $\mu_{r_2} = 1 - \eta - \mu_{r_1} - \mu_r \triangleq \mu_2$ hold, we give the optimal mode selection and power allocation policies in Theorem 2 in terms of $\mu = [\mu_1, \mu_2, \mu_r]$ instead of in terms of $\mu$. This completes the proof.

**APPENDIX D**

**PROOF OF PROPOSITION 2**

The optimal values of $\mu$ and $\gamma$ are obtained by solving the dual problem in (38). In particular, we substitute the optimal values of $q$ and $P$ in (40) and $R$ in (12) for a given $\mu$ and $\gamma$ into (36) to obtain the dual function in (37). Similar to the proof of Proposition 1, we employ the sub-gradient method to solve the dual problem in (38). The updates are performed as specified in Algorithm 2 with the updates for $\mu$ in (10) and the updates for $\gamma$ in (19). Note that $\mu_{r_1} = \eta - \mu_{r_2} - \mu_r \triangleq \mu_1$ and $\mu_{r_2} = 1 - \eta - \mu_{r_1} - \mu_r \triangleq \mu_2$ have to hold from (33a) and (33b), respectively. Moreover, the dual feasibility condition, $\mu \geq 0$ has to hold. Hence, we obtain $0 \leq \mu_1 \leq \eta$, $0 \leq \mu_2 \leq 1 - \eta$, and $0 \leq \mu_r \leq \min\{\eta, 1 - \eta\}$. In the following, we also obtain a feasible interval for $\gamma$. First, due to the dual feasibility condition, $\gamma_j \geq 0$, $j \in \{1, 2, r\}$, has to hold. In order to find an upper bound, we remove some of the terms in the denominators of the left hand side of (43a) which leads to
\[
\gamma_1 \ln 2 < \frac{\mu_1s_1}{P_1(s)s_1} + \frac{\mu_2s_2}{P_2(s)s_1} = \frac{\mu_1 + \mu_r}{P_1(s)}.
\]
Now, if constraint $P_1 \leq P^\text{max}$ is active, we obtain $E\{P_1(s)\} = P^\text{max} < \frac{\mu_1 + \mu_r}{\gamma_1 \ln 2}$ which leads to $\gamma_1 < \frac{\mu_1 + \mu_r}{P^\text{max} \ln 2}$. On the other hand, if constraint $P_1 \leq P^\text{max}$ is inactive, Lagrange multiplier $\gamma_1$ has to be zero due to the complementary slackness condition [26]. Nevertheless, $\frac{\mu_1 + \mu_r}{P^\text{max} \ln 2}$ is an upper bound for $\gamma_1$. With a similar procedure, we obtain $\gamma_2 < \frac{\mu_2 + \mu_r}{P^\text{max} \ln 2}$ and $\gamma_r < \frac{1 - \mu_1 - \mu_2 - 2\mu_r}{P^\text{max} \ln 2}$. This completes the proof.

**APPENDIX E**

**PROOF OF LEMMA 2**

The main idea of proof of Lemma 2 is to show that, for $Q^\text{max}_j \rightarrow \infty$, $j = 1, 2$, if we use the optimized variables for outer bound $R^\text{out}$, i.e., optimal $q^*$ and $P^*$ from (7), there exists a rate selection policy that makes the effect of the queues at the relay nodes negligible as $N \rightarrow \infty$. To this end, let $(R^*_{12}, R^*_{21})$ be a point on the boundary surface of the rate region of the outer bound $R^\text{out}$. Moreover, for the rest of the proof, we assume that $q^*$ and/or $P^*$ are employed to calculate $R_{1r}, R_{2r}, R_{r1},$ and $R_{r2}$.

For infinite-size buffers at the relay nodes, i.e., $Q^\text{max}_j \rightarrow \infty$, $j = 1, 2$, the state of the queues does not influence the average arrival transmission rates, i.e., the effect of min function in (2) is removed. However, the average departure transmission rates can still be limited by the amount of information available at the buffers, i.e., the effect of the min function in (3) cannot be neglected even if $Q^\text{max}_j \rightarrow \infty$, $j = 1, 2$. Mathematically, the bounds on average rates $R_{1r}$ and $R_{2r}$ can be expressed independently from the dynamics of the queues as
\[
\begin{align*}
R_{1r} &\leq \begin{cases} C_{s_1}^*, & \text{if } R_{2r} > C_{s_1}^* \\ R_{2r}, & \text{if } R_{2r} < C_{s_1}^* \end{cases} \tag{50a} \\
R_{2r} &\leq \begin{cases} C_{s_2}^*, & \text{if } R_{1r} > C_{s_2}^* \\ R_{1r}, & \text{if } R_{1r} < C_{s_2}^* \end{cases} \tag{50b}
\end{align*}
\]
respectively. In particular, if $R_{2r} > C_{s_1}^*$ holds, i.e., the average rate flowing into the buffer is larger than the average capacity of the respective departure channel, then, as $N \rightarrow \infty$, buffer $B_2$ always has enough information to supply because the amount of information in the queues increases over time and we obtain bound $R_{1r} \leq C_{s_1}^*$. On the other hand, if $R_{2r} < C_{s_1}^*$ holds, i.e., the average information flowing into the buffer is less than the average capacity of the respective departure channel, then by the law of conservation of flow, we obtain $R_{r1} \leq R_{2r}$. If $R_{2r} > C_{s_2}^*$ holds, we obtain $R_{r1} \leq R_{2r} = C_{s_1}^*$. For a detailed proof of this relation, we refer to the proof of [15, Theorem 2]. Similar results hold for $R_{1r}, R_{1r},$ and $C_{s_2}^*$.

In the following, using the bounds in (50) and depending on which set $A_k$, $k = 1, \ldots, 7$, the optimal $(C_{s_2}^*, C_{s_1}^*)$ belongs to, see Fig. 4, we propose a rate selection policy which achieves the optimal rate pair $(R^*_{12}, R^*_{21})$ on the boundary surface of the rate region of the outer bound $R^\text{out}$.

**Case 1:** If $(C_{s_2}^*, C_{s_1}^*) \in A_1$ holds, from constraints C1, C2, and C3 in (7), the active constraints are $R_{12}^* \leq C_{s_1}^*, R_{21}^* \leq C_{s_2}^*$, and $R_{12}^* + R_{21}^* = C_{s_1}^*$. Thereby, in the time slots for which $q^*_\text{MAC}(s) = 1$ holds, we let the users transmit with capacity rates, i.e., $\rho_{1r} = 1$ and $\rho_{2r} = 1$, where the decoding order probability is given by $p_1 = \frac{R_{12}^* + C_{s_1}^* - C_{s_2}^*}{R_{12}^* - C_{s_2}^*}$. With this strategy, we obtain $R_{1r} = R_{12}^* \leq R_{2r}^*$ and $R_{r1} = R_{12}^* \leq R_{2r}$. Moreover, since for this case, $R_{1r} = R_{12}^* < C_{s_1}^*$ and $R_{2r} = R_{21}^* < C_{s_1}^*$ hold, we obtain the bounds $R_{1r} \leq R_{2r}$, and $R_{r2} \leq R_{r1}$, from (50).

Thereby, we set $\rho_{1r} = \frac{R_{12}^*}{C_{s_1}^*}$ and $\rho_{2r} = \frac{R_{21}^*}{C_{s_1}^*}$ which leads to $R_{1r} = R_{21}^*$ and $R_{2r} = R_{12}^*$. Hence, we obtain $R_{1r} = R_{r2} = R_{12}^* = R_{r1} = R_{21}^*$.

**Case 2:** If $(C_{s_2}^*, C_{s_1}^*) \in A_2$ holds, we obtain the active constraints in (7) as $R_{12}^* \leq C_{s_2}^*, R_{21}^* \leq C_{s_1}^*$, and $R_{12}^* + R_{21}^* = C_{s_1}^*$. The same rate selection policy as for $(C_{s_2}^*, C_{s_1}^*) \in A_1$ leads to $R_{1r} = R_{r2} = R_{12}^* \leq R_{2r} = R_{r1} = R_{21}^*$.

**Case 3:** If $(C_{s_2}^*, C_{s_1}^*) \in A_3$ holds, we obtain the active constraints in (7) as $R_{12}^* = C_{s_2}^*$ and $R_{21}^* = C_{s_1}^*$. Thereby, for the time slots with $q^*_\text{MAC}(s) = 1$, if we assume that
the users transmit with their capacity rates, i.e., $\rho_{1r} = 1$ and $\rho_{2r} = 1$, where the decoding order probability is given by $p_t \in [p_{t \min}^{\rho_{1r}}, p_{t \max}^{\rho_{2r}}]$ with $p_{t \min}^{\rho_{1r}} = \frac{C_r^* - C_{r_{12}}^* - C_{r_{21}}^* - C_{r_{12}}^*}{C_{r_{12}}^* + C_{r_{21}}^* - C_{r_{12}}^*}$ and $p_{t \max}^{\rho_{2r}} = \frac{C_{r_{12}}^* - C_{r_{12}}^*}{C_{r_{12}}^* + C_{r_{21}}^* - C_{r_{12}}^*}$. Consequently, we obtain the bounds $R_{1r} = C_{r_{12}}^*$ and $R_{2r} = C_{r_{21}}^*$. Case 4: If $(C_{r_{12}}^*, C_{r_{12}}^*) \in A_4$, we obtain the active constraints in (7) as $R_{12} = C_{r_{12}}^*$ and $R_{21} = C_{r_{21}}^*$. With a similar discussion as for the previous cases, we propose the following parameters for this case: $p_{t} = 1$, $\rho_{1r} = \frac{R_{12}^*}{C_{r_{12}}^* - C_{r_{21}}^*}$, $\rho_{2r} = 1$, and $\rho_{r1} = \frac{R_{12}^*}{C_{r_{12}}^* - C_{r_{21}}^*}$. Hence, we obtain $R_{1r} = C_{r_{12}}^*$ and $R_{2r} = C_{r_{21}}^*$, which, based on (50), leads to $R_{1r} = R_{2r} = R_{12}$ and $R_{2r} = R_{t1} = R_{21}$. Case 6: If $(C_{r_{21}}^*, C_{r_{12}}^*) \in A_6$, we obtain the active constraints in (7) as $R_{12} = C_{r_{21}}^*$ and $R_{21} = C_{r_{12}}^*$. The same rate selection policy for $(C_{r_{21}}^*, C_{r_{12}}^*) \in A_1$ leads to $R_{1r} = R_{2r} = R_{12}$ and $R_{2r} = R_{t1} = R_{21}$. The transmission rate selection is similar to the one given for case $(C_{r_{12}}^*, C_{r_{12}}^*) \in A_4$ after switching the roles of user 1 and user 2.

To summarize, provided that $Q_{j \max} \to \infty$, $j = 1, 2$ holds and $q^*$ and $P^*$ are employed, we proposed a rate selection policy for each of the seven mutually exclusive cases $(C_{r_{12}}^*, C_{r_{12}}^*) \in A_k$, $k = 1, \ldots, 7$, which achieves rate pair $(R_{12}, R_{21})$ on the boundary surface of the outer bound. In other words, the proposed rate selection policy operates the queues of the buffers such that the effect of the number of events that the buffers do not have enough information to supply, due to the ’niin’ functions in (22c) and (22d), is negligible as $N \to \infty$. Note that the values of constants $p_{t}$, $\rho_{1r}$, $\rho_{2r}$, $\rho_{r1}$, and $\rho_{r2}$ given in Lemma 2 are identical to those obtained in this proof but are presented in a more compact form. This completes the proof.

REFERENCES

[1] T. M. Cover and J. A. Thomas, Elements of Information Theory. Wiley, John and Sons, Incorporated, 1991.
[2] S. J. Kim, P. Mitran, and V. Tarokh, “Performance Bounds for Bidirectional Coded Cooperation Protocols,” IEEE Trans. Inf. Theory, vol. 54, no. 11, pp. 5235–5241, Nov. 2008.
[3] T. Oechtering, C. Schnurr, I. Bjelakovic, and H. Boche, “Broadcast Capacity Region of Two-Phase Bidirectional Relaying,” IEEE Trans. Inf. Theory, vol. 54, no. 1, pp. 454–458, Jan. 2008.
[4] Y. Wu, P. A. Chou, and S.-Y. Kung, “Information Exchange in Wireless Networks with Network Coding and Physical-Layer Broadcast,” in Proc. 39th Ann. Conf. Inf. Sci. Syst., Mar. 2005.
[5] B. Rankov and A. Wittneben, “Spectral Efficient Signaling for Half-duplex Relay Channels,” in Proc. Asilomar Conf. Signals, Syst., Comput., Oct. 2005, pp. 1066–1071.
[6] T. J. Oechtering and H. Boche, “Optimal Resource Allocation for a Bidirectional Regenerative Half-duplex Relaying,” in Proc. Int. Symp. on Inf. Theory and its App., Oct. 2006.
[7] L. Gerdes, M. Riemensberger, and W. Utschick, “On Achievable Rate Regions for Half-Duplex Gaussian MIMO Relay Channels: A Decomposition Approach,” IEEE J. Sel. Areas Commun., vol. 30, no. 8, pp. 1319–1330, Sep. 2012.
[8] ——, “Bounds on the Capacity Regions of Half-Duplex Gaussian MIMO Relay Channels,” EURASIP J. Adv. Sig. Proc., pp. 1–17, March 2013.
[9] R. F. Wyrembelski, T. J. Oechtering, and H. Boche, “Decode-and-Forward Strategies for Bidirectional Relaying,” in Proc. IEEE PIMRC, Sep. 2008, pp. 1 –6.
[10] P. Liu and I. Kim, “Performance Analysis of Bidirectional Communication Protocols Based on Decode-and-Forward Relaying,” IEEE Trans. Commun., vol. 58, no. 10, pp. 2683–2690, Oct. 2010.
[11] A. Goldsmith and P. Varaiya, “Capacity of Fading Channels with Channel Side Information,” IEEE Trans. Inf. Theory, vol. 43, no. 6, pp. 1986–1992, Nov 1997.
[12] D. Tse and S. Hanly, “Multiaccess Fading Channels. Part I: Polymatroid Structure, Optimal Resource Allocation and Throughput Capacities,” IEEE Trans. Inf. Theory, vol. 44, no. 7, pp. 2796–2815, Nov 1998.
[13] A. Host-Madsen and J. Zhang, “Capacity Bounds and Power Allocation for Wireless Relay Channels,” IEEE Trans. Inf. Theory, vol. 51, no. 6, pp. 2020–2040, 2005.
[14] B. Xia, Y. Fan, J. Thompson, and H. Poor, “Buffering in a Three-Node Relay Network,” IEEE Trans. Wireless Commun., vol. 7, no. 11, pp. 4492–4496, Nov. 2008.
[15] N. Zlatanov, R. Schober, and P. Popovski, “Buffer-Aided Relaying with Adaptive Link Selection,” IEEE J. Select. Areas Commun., vol. 31, no. 8, pp. 1530–1542, Aug. 2013.
[16] V. Jamali, N. Zlatanov, A. Ihklef, and R. Schober, “Adaptive Mode Selection and Power Allocation in Bidirectional Buffer-aided Relay Networks,” in Proc. IEEE Globecom, Dec. 2013.
[17] ——, “Adaptive Mode Selection in Bidirectional Buffer-aided Relay Networks with Fixed Transmit Powers,” in Proc. EUSIPCO, Sep. 2013.
[18] H. Liu, P. Popovski, E. de Carvalho, and Y. Zhao, “Sum-Rate Optimization in a Two-Way Relay Network with Buffering,” IEEE Commun. Lett., vol. 17, no. 1, pp. 95–98, Jan. 2013.
[19] L.-L. Yang, D. C., and L. Hanzo, “Multihop Diversity - A Precious Source of Fading Mitigation in Multihop Wireless Networks,” in Proc. IEEE Globecom, Dec. 2011, pp. 1–5.
[20] A. Ihklef, D. Michalopoulos, and R. Schober, “Max-Max Relay Selection for Relays with Buffers,” IEEE Trans. Wireless Commun., vol. 11, no. 3, pp. 1124–1135, Mar. 2012.
[21] I. Krikidis, T. Charalambous, and J. Thompson, “Buffer-Aided Relay Selection for Cooperative Diversity Systems without Delay Constraints,” IEEE Trans. Wireless Commun., vol. 11, no. 5, pp. 1957–1967, May 2012.
[22] M. J. Neely, Stochastic Network Optimization with Application to Communication and Queuing Systems. Morgan & Claypool.
[23] X. Wang and N. Gao, “Stochastic Resource Allocation Over Fading Multiple Access and Broadcast Channels,” IEEE Trans. Inf. Theory, vol. 56, no. 5, pp. 2382–2391, May 2010.
[24] X. Wang and G. Giannakis, “Resource Allocation for Wireless Multisiter OFDM Networks,” IEEE Trans. Inf. Theory, vol. 57, no. 7, pp. 4359–4372, Jul. 2011.
[25] A. E. Gamal and Y. Kim, Network Information Theory. Cambridge University Press, 2011.
[26] S. Boyd and L. Vandenberghe, Convex Optimization. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[27] S. Boyd, EE392o Course Notes Stanford Univ. Stanford, CA, 2004. [Online]. Available: http://www.stanford.edu/class/ee392o
[28] W. Yu and R. Lui, “Dual Methods for Nonconvex Spectrum Optimization of Multicarrier Systems,” IEEE Trans. Commun., vol. 54, no. 7, Jul. 2006.
[29] H. Mu and J. Tugnait, “Joint Soft-Decision Cooperative Spectrum Sensing and Power Control in Multiband Cognitive Radios,” IEEE Trans. Sig. Proc., vol. 60, no. 10, pp. 5334–5346, Oct. 2012.
Vahid Jamali (S12) was born in Fasa, Iran, in 1988. He received his B.S. and M.S. degrees in electrical engineering from K. N. Toosi University of Technology (KNTU), in 2010 and 2012, respectively. Currently, he is working toward his Ph.D. degree at the Friedrich-Alexander University (FAU), Erlangen, Germany. His research interests include multiuser information theory, wireless communications, cognitive radio network, LDPC codes, and optimization theory.

Nikola Zlatanov (S06) received his Dipl.Ing. and M.S. degrees in electrical engineering from SS. Cyril and Methodius University, Skopje, Macedonia, in 2007 and 2010, respectively. Currently, he is working toward his Ph.D. degree at the University of British Columbia (UBC), Vancouver, Canada. His current research interests include the wireless communications and information theory. He received several awards for his work including UBCs Four-Year Doctoral Fellowship in 2010, UBCs Killam Doctoral Scholarship and Macedonias Young Scientist of the Year Award in 2011, Vanier Canada Graduate Scholarship in 2012, DAAD Research Grant in 2013, and best paper award from the German Information Technology Society (ITG) in 2014.

Aissa Ikhlef (M’09) was born in Constantine, Algeria. He received the B.S. degree in Electrical Engineering from the University of Constantine, Constantine, Algeria, in 2001 and the M.Sc. and Ph.D. degrees in Electrical Engineering from the University of Rennes 1, Rennes, France, in 2004 and 2008, respectively. From 2004 to 2008, he was with Supélec, France, pursuing his Ph.D. degree. From 2007 to 2008, he was a Lecturer at the University of Rennes 1. From 2008 to 2010 he was a Postdoctoral Fellow with the Communication and Remote Sensing Laboratory, Catholic University of Louvain, Louvain La Neuve, Belgium. He was a visiting Postdoctoral Fellow at the University of British Columbia, Vancouver, BC, Canada, from August to November 2009. From 2010 to 2013 he was with the data communications group, University of British Columbia, Vancouver, Canada, as a Postdoctoral Fellow. From June 2013 to August 2014 he was with Toshiba Research Europe Limited, Bristol, UK, as a Senior Research Engineer. Since September 2014 he has been with the School of Electrical and Electronic Engineering, Newcastle University, Newcastle, UK, as a Lecturer (Assistant Professor). His current research interests include cooperative and energy harvesting communications.

Robert Schober (S’98, M’01, SM’08, F’10) was born in Neuendettelsau, Germany, in 1971. He received the Diplom (Univ.) and the Ph.D. degrees in electrical engineering from the University of Erlangen-Nuernberg in 1997 and 2000, respectively. From May 2001 to April 2002 he was a Postdoctoral Fellow at the University of Toronto, Canada, sponsored by the German Academic Exchange Service (DAAD). Since May 2002 he has been with the University of British Columbia (UBC), Vancouver, Canada, where he is now a Full Professor. Since January 2012 he is an Alexander von Humboldt Professor and the Chair for Digital Communication at the Friedrich Alexander University (FAU), Erlangen, Germany. His research interests fall into the broad areas of Communication Theory, Wireless Communications, and Statistical Signal Processing.

Dr. Schober received several awards for his work including the 2002 Heinz Maier-Leibnitz Award of the German Science Foundation (DFG), the 2004 Innovations Award of the Vodafone Foundation for Research in Mobile Communications, the 2006 UBC Killam Research Prize, the 2007 Wilhelm Friedrich Bessel Research Award of the Alexander von Humboldt Foundation, the 2008 Charles McDowell Award for Excellence in Research from UBC, a 2011 Alexander von Humboldt Professorship, and a 2012 NSERC E.W.R. Steacie Fellowship. In addition, he received best paper awards from the German Information Technology Society (ITG), the European Association for Signal, Speech and Image Processing (EURASIP), IEEE WCNC 2012, IEEE Globecom 2011, IEEE ICUWB 2006, the International Zurich Seminar on Broadband Communications, and European Wireless 2000. Dr. Schober is a Fellow of the Canadian Academy of Engineering and a Fellow of the Engineering Institute of Canada. He is currently the Editor-in-Chief of the IEEE TRANSACTIONS ON COMMUNICATIONS.