Received August 20, 2019, accepted September 1, 2019, date of publication September 12, 2019,
date of current version September 25, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2940719

An Efficient Network-Coded ARQ Scheme for Two-Way Wireless Communication With Full-Duplex Relaying

HONGYI ZHU[1], (Student Member, IEEE), BESMA SMIDA[2], (Senior Member, IEEE), AND DAVID J. LOVE[1], (Fellow, IEEE)

[1]School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907, USA
[2]Department of Electrical and Computer Engineering, The University of Illinois at Chicago, Chicago, IL 60607, USA

Corresponding author: Hongyi Zhu (zhyd1010@hotmail.com)

This work was supported by the NSF CCF-CIF under Grant 1620794.

ABSTRACT There has been renewed interest in using relaying systems as a low-cost solution in next generation wireless broadband networks. In this paper, we take a fresh look at relaying and propose a new retransmission protocol for a two-way wireless network with one base-station, one full-duplex relay and multiple users exchanging messages over fading channels. In this scenario, the relay and each user may hear the message from the base-station and other users. To improve the throughput, we apply network-coded ARQ with reverse-link assistance (NC-RLA) at the relay to exploit the overheard side-information acquired from both the downlink and uplink. We analyze the throughput performance enhancements of NC-RLA and validate our mathematical derivations with Monte Carlo simulation. The simulation results show that significant throughput improvements are obtained with the NC-RLA scheme for low/moderate SNRs. We also prove that full-duplex relay increases the throughput if it is located at the optimal position. Otherwise, it is not beneficial to add a relay, especially with a large number of users using network coding.

INDEX TERMS Automatic repeat request (ARQ), full-duplex relay, network coding, outage probability, two-way communication, throughput analysis.

I. INTRODUCTION

The demand for high data-rate wireless access continues to rise greatly due to the growth of wireless subscribers. A recent study showed that the number of global mobile devices and connections is projected to increase from 8.6 billion in 2017 to 12.3 billion in 2022 [1]. Unfortunately, the available spectrum for wireless communication is a limited resource, making it crucial to exploit the spectrum efficiently. Therefore, great effort has been made to develop technologies that enhance spectral efficiency to deliver higher data rates over the same amount of spectrum [2].

A. RELATED WORK

One recent and promising area of research is the exploitation of a full-duplex relay. Originally known as an on-frequency repeater, the full-duplex relay was proposed in the 1980s [3]. In [4], the author used antenna isolation measurements to implement an on-frequency repeater. Multiple-input-multiple-output on-frequency repeaters were studied in [5], [6]. Full-duplex relays perform simultaneous transmission and reception on the same spectrum, which holds the potential to double the spectral efficiency over that of a half-duplex relay [7]. In practice, however, the performance of full-duplex relaying is limited due to the presence of self-interference. For this reason, some works have proposed switching the relay between the two operation modes (i.e., hybrid relay) [8]–[10] or imitating full-duplex operation with multiple connected half-duplex relays [11], [12]. Research into hybrid relaying has shown the trade-offs associated with both half-duplex and full-duplex relays. As proved in [8], full-duplex relaying can achieve a higher spectral efficiency when self-interference is suppressed to an acceptable level. Many recent research studies on self-interference cancellation techniques have shown that self-interference can be drastically reduced by combining

The associate editor coordinating the review of this manuscript and approving it for publication was Daniel Benevides Da Costa.

passive suppression [13], active cancellation [14] and advanced antenna design [15]. These achievements have increased the practicality of employing full-duplex relays. With the application of resource allocation [9], [16] and relay selection [10], [17], the benefits of full-duplex relays have been explored for various types of wireless networks.

Most research has aimed to increase spectral efficiency with a focus on one-way communications, either in the downlink or uplink. However, this abstraction ignores the fact that wireless systems are nearly always *two-way*. The broadcast nature of wireless transmission means that users can often overhear downlink and/or uplink transmissions corresponding to other users. When we consider a two-way system, the overheard side-information can be utilized to help the retransmissions.

A combination of network coding (NC) [18] and automatic repeat request (ARQ) can effectively utilize the side-information generated from simultaneous retransmissions to multiple users. Network coding was initially proposed for wired systems [19], but was soon extended to improve noisy wireless systems [20]. As proved in some recent network coding work [21]–[27], network coding can be strengthened by using an ARQ approach. Network coding strategies with ARQ (i.e., NCed-ARQ) were thoroughly studied in prior work for multi-cast channels (e.g., [21], [22]) and broadcast channels (e.g., [23], [28]). The throughput of regular ARQ can be improved greatly by making the receivers combine a packet with its retransmitted copies, which is called Hybrid-ARQ (HARQ). A NCed HARQ scheme, an extension to NCed-ARQ, was proposed and analyzed in [24], [25]. Carrier frequency offset problem was addressed for OFDM modulated network coding used with a two-way relay channel [29]. Using network coded cooperative communication, massive video multicasting can be realized in cellular networks as shown in [30]. A buffer-aided physical layer network coding scheme was proposed for cooperative networks [31].

B. CONTRIBUTIONS

In this paper, we consider two-way wireless communication when a full-duplex relay is used between the base-station (BS) and users as shown in Fig. 2 (Fig. 1 illustrates the case without relay). To achieve a complete coverage of wireless access, people have considered to use relays as a low-cost method to extend the wireless signal. This wireless system depicts the current and future cellular networks. The detailed descriptions for the system models will be given in Section II. In this scenario, the users can hear from and transmit to the base-station and relay, which means the direct links between the base-station and users are considered in the system. Since the relay deploys full duplexing, during the downlink, the users will receive the packets from the base-station and relay at the same time. We also assume that each user can overhear other users' transmissions during both the downlink and uplink. In this system, the base-station or the relay can work as the central node for NCed-ARQ scheme. To achieve an efficient usage of spectrum and simpler channel

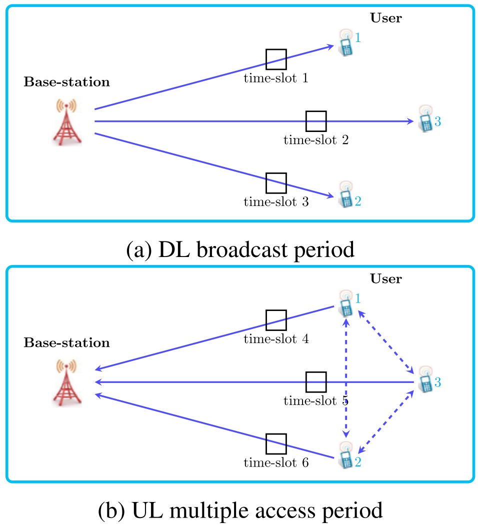


FIGURE 1. MABC models.

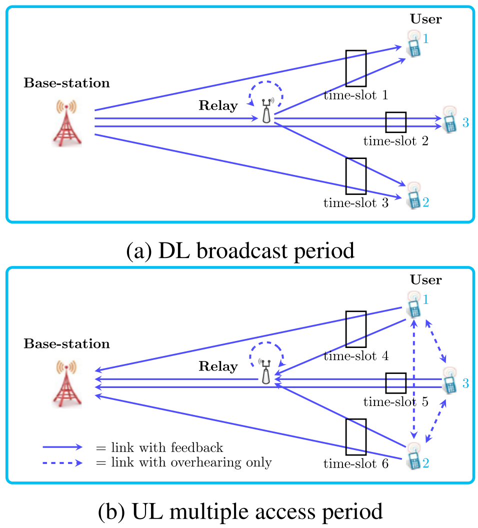


FIGURE 2. MABRC models.

equalization, many communication protocols adopt time division duplexing (TDD). We also assume TDMA and TDD for the system models used in this paper.

For this two-way wireless relaying system, we design a new retransmission protocol inspired by NCed-ARQ scheme, and derive the outage probability and throughput. During retransmissions in downlink and uplink, each user can overhear other users' packets as side-information. To utilize the side-information and improve the throughput performance, we propose a network coding scheme with reverse-link assistance (NC-RLA) that allows users to send network-coded packets (NCPs) on the uplink to help the downlink, which has been partially discussed in our preliminary work work [32]. Compared to the scheme proposed in [24], our strategy utilizes the side-information overheard during both downlink and uplink to improve the downlink throughput performance without hurting the uplink transmissions. In a relaying system with direct links between the base-station and users, the base-station and relay can be the center of network coding. Since the direct links are weaker than the link between the base-station and relay, most packets will be decoded by the relay. Hence, we prove that it is more efficient to apply NC-RLA at the relay and optimize the position of the relay.

With full-duplex relaying, each user receives the packets from the base-station and relay at the same time. It forms a multiple access channel at each user. To take advantage of the two simultaneous transmissions, we consider the achievable rate region for the users. Users are assumed to use interference cancellation to jointly decode the two packets received from the base-station and relay. Using ARQ protocol, both the base-station and relay will retransmit the packets that are not decoded by the receivers. In this setting, packets may be decoded by the relay in advance. The base-station will stop transmitting the packets decoded by the relay, and the relay will retransmit the packets to their target users. The whole transmission of each packet ends when the two transmitters finish their transmissions. Since both the base-station and relay have constraints on the transmit power, the transmit powers of the base-station and relay are optimized. To fairly compare the system with and without a relay, we assume a fixed total transmit power.

We derive closed-form outage probabilities and achievable rate expressions for the two-way wireless systems with and without relaying. We adopt Monte Carlo simulation to validate our mathematical derivations. The simulation and numerical results show that our proposed network coding scheme greatly improves the downlink throughput, especially for low and moderate signal-to-noise-ratios (SNRs). We compare the performance for the different network settings. We also investigate whether it is beneficial to adopt a relay under different situations when assuming an equivalent energy cost. With the numerical results, we study and optimize the location of the relay, if we want to set up a relay for a target area with many users.

The remainder of this paper is organized as follows. Section II provides an overview of the channel models considered. In Section III, we describe the proposed protocols for the two-way communication without a full-duplex relay, which works as the basis and comparison for Section IV. In Section IV, we analyze the two-way system with a full-duplex relay and derive expressions of its throughput. Simulation results and performance comparisons for different settings are provided in Section V. Finally, conclusions are given in Section VI.

II. PRELIMINARIES

We consider two physical channel models: 1) an M-user multiple-access-broadcast channel (MABC) and 2) an M-user multiple-access-broadcast channel with one relay (MABRC), as depicted for $M = 3$ in Fig. 1 and Fig. 2, respectively. MABC model is much simpler and used as comparison for MARBC model. In these two system models, two-way communication must occur between multiple users and a base-station. Without loss of generality, we refer to the users and base-station as nodes. These nodes are all assumed to be half-duplex, which means that the node cannot transmit and receive on a specific frequency at the same time. To improve the overall throughput, we study the scenario when a full-duplex relay, which can transmit and receive simultaneously on a given frequency, is also placed within the communication network.

To ensure reliable communications, we use an ARQ protocol for both the MABC and MABRC model. Since the rate region for a multiple-access channel with HARQ is not known, we only consider basic ARQ in this paper. The receiver sends an ACK/NACK to indicate the success/failure of transmission. The transmitter retransmits the same message if it receives a NACK or moves on to the next message if it receives an ACK. Note that the user-to-user channel has no feedback. With ARQ, each packet will be retransmitted until it gets decoded by the receiver. In fact, retransmissions can be regarded as the transmissions from correlated sources whose outage analysis was discussed in [33]. A reduction in the outage probability can be achieved from rate allocation for each retransmission. In this paper, we consider the case of a fixed data rate and only require perfect channel state information at the receivers. To make our analysis tractable, we also assume *free and perfect* ARQ feedback (i.e., ARQ feedback is sent over individual control information channel without being lost), identical packet length, capacity-achieving codes, and no packet overhead. These ideal assumptions are applied to both the classical and proposed network coding for fair comparison.

We use the notation $t_x \to r_x$ to represent the channel links, where $t_x, r_x \in \{b, r, u\}$ refers to base-station b, relay r and users u, respectively. Although the proposed protocol does not require a symmetric system, to simply analyze the average throughput performance, we assume all channels are symmetric across each user. We also assume the forward and reverse links between any two users have the same distribution for overhearing. The proposed NC-RLA scheme helps downlink transmissions without affecting the uplink. The analysis of uplink transmissions is still necessary due to the dependency of the downlink and uplink in network coding protocol. The performance of the uplink transmission can be derived by an analogy with the downlink.

A. MULTIPLE-ACCESS-BROADCAST CHANNEL

This scenario models two-way communication between a base-station and multiple users. Since time division is assumed, the transmission time is divided into two periods: the uplink multiple access period and the downlink broadcast period. The downlink period is shown in Fig. 1(a). The base-station transmits to the user nodes sequentially. The uplink period is shown in Fig. 1(b). Similarly to the downlink, the user nodes take turns and transmit sequentially to the base-station. During both periods, each user is allowed to overhear and keep the unintended packets for network coding.

During the downlink period of the MABC models, we only consider channel $b \rightarrow u$ (with channel-gain h_{bu}) for each user u. We define x_b as the symbol, with the constant transmit power per symbol fixed at P_b, transmitted by the base-station and define y_u as the symbol received by user u. We also assume a block Rayleigh fading channel and additive white Gaussian noise (AWGN). The input-output relationship under this model is given by

$$y_u = h_{bu}x_b + w_u, \tag{1}$$

where the AWGN term w_u at each user is distributed as $\mathcal{CN}(0, N_0)$. Under the assumption of perfect channel state information at the receivers, the achievable rate (in nats/Hz/s) is $I_{bu} = \ln(1 + \rho_{bu})$, where the signal-to-noise ratio (SNR) is $\rho_{bu} = |h_{bu}|^2 P_b/N_0$. We assume the path loss for channel $b \rightarrow u$ is $L_{bu} = Kd_{bu}^{\gamma}$, where γ is the path loss exponent, K is a constant for the effect of propagation and d_{bu} is the distance of channel $b \rightarrow u$. Considering fading, the overall channel gain is $h_{bu} = h/\sqrt{L_{bu}}$, where h is a zero-mean complex Gaussian random variable and $|h|^2$ is exponentially distributed with mean one. Here, $Exp(\cdot)$ denotes the exponential distribution. The receive SNR over channel $b \rightarrow u$ under Rayleigh fading is distributed as $\rho_{bu} = \frac{|h|^2 P_b}{L_{bu}N_0} \sim Exp(\lambda_{bu})$, where $\lambda_{bu} \triangleq \frac{L_{bu}N_0}{P_b}$.

B. MULTIPLE-ACCESS-BROADCAST CHANNEL WITH ONE RELAY

This scenario models two-way communication between a base-station and multiple user nodes when using a relay. Practically, the relay could be thought of as a femtocell or repeater depending on its capabilities. During the broadcast period, the base-station and the relay transmit to each user in a sequence of time-slots as shown in Fig. 2(a). Due to the full-duplex relaying, the transmissions from the base-station and relay to one user share the same time-slot. During the multiple-access period, similarly, the users and the relay transmit to the base-station sequentially, as shown in Fig. 2(b). We assume the message is first decoded at the relay and then transmitted to the target user (i.e., a decode-and-forward relay operation). Hence, it forms a multiple-access channel at each user and generates self-interference at the relay.

During the downlink period of the MABRC models, we need to consider channel $b \rightarrow u$ with channel gain h_{bu}, channel $r \rightarrow u$ with h_{ru} for each user u, and channel $b \rightarrow r$ with h_{br}. Because the base-station and relay transmit in the same time-slot, channel $b \rightarrow u$ and $r \rightarrow u$ form a multiple-access channel at the user u. In addition, since the relay operates in full-duplex mode, some self-interference is generated at the relay. We define x_r as the symbol, with the transmit power per symbol fixed at P_r, transmitted by the relay, y_r as the symbol received by the relay, and i_r as the self-interference term at the relay. Under this model, the input-output relationships are given by

$$y_r = h_{br}x_b + i_r + w_r, \; y_u = h_{bu}x_b + h_{ru}x_r + w_u, \tag{2}$$

where the AWGN terms w_r and w_u are $\mathcal{CN}(0, N_0)$. Assuming that each user has perfect channel state information knowledge, the achievable rate over channel $b \rightarrow r$ is given by $I_{br} = \ln(1 + \rho_{br})$, where the receive signal-to-interference-and-noise ratio (SINR) is $\rho_{br} = |h_{br}|^2 P_b/(N_{SI} + N_0)$. N_{SI} is defined as the self-interference power at the relay. We define R_b and R_r as the data rate of the base-station and relay. The rate region of the Gaussian multiple-access channel [34] is

$$\begin{aligned} R_b &\le \ln(1 + \rho_{bu}), \quad R_r \le \ln(1 + \rho_{ru}), \\ R_b + R_r &\le \ln(1 + \rho_{bu} + \rho_{ru}). \end{aligned} \tag{3}$$

where the receive SNRs are $\rho_{bu} = |h_{bu}|^2 P_b/N_0$ and $\rho_{ru} = |h_{ru}|^2 \, P_r/N_0$. The path losses are $L_{bu} = Kd_{bu}^{\gamma}$, $L_{br} = Kd_{br}^{\gamma}$ and $L_{ru} = Kd_{ru}^{\gamma}$, respectively. The variables d_{bu}, d_{br} and d_{ru} are the distances of channel $b \rightarrow u$, $b \rightarrow r$, and $r \rightarrow u$. The power of self-interference of the relay can be modeled as $N_{SI} = P_r/\psi$ [35], where ψ is a self-interference cancellation factor due to both antenna separation and active self-interference cancellation techniques. Therefore, the receive SINR over channel $b \rightarrow r$ and SNR over channel $r \rightarrow u$ are distributed as $\rho_{br} \sim Exp(\lambda_{br})$ and $\rho_{ru} \sim Exp(\lambda_{ru})$, respectively, where $\lambda_{br} \triangleq \frac{L_{br}(N_{SI}+N_0)}{P_b}$ and $\lambda_{ru} \triangleq \frac{L_{ru}N_0}{P_r}$.

III. PROPOSED PROTOCOL – MABC MODELS

In this section, we start with the MABC model – a two-way wireless system consisting of one base-station and multiple user nodes (i.e., with no relay). During each downlink period, the base-station transmits to the users one after another. Since we use an ARQ protocol, each user replies with an ACK or NACK during the subsequent uplink period to indicate whether the user successfully decoded the last downlink packet or not. By utilizing the ARQ feedback and side-information overheard during transmissions, we apply the NC-RLA scheme to improve the downlink throughput. We analyze the MABC model first, so we can focus on the NC-RLA protocol.

A. PROTOCOL: NETWORK CODING WITH REVERSE-LINK-ASSISTANCE

Network coding, as considered in this paper, utilizes regular packets (RPs) and network coded packets (NCPs). As shown in the flowchart, Fig. 3, each RP is always sent individually at first. If an RP is decoded at the user it is intended for, the packet has been successfully delivered. If the RP is not decoded by the intended user but is decoded by an unintended user, it will be saved as side-information. In both cases, the base-station will transmit a new RP. If no one decodes the RP, it will be retransmitted. Note that any successfully transmitted RPs will not be included in NCPs. Based on the ARQ feedback, the base-station can group the downlink RPs for the users who overheard all the packets for the other users in the same group but failed to decode their own packets. In this case, these RPs are ready for network coding transmissions. The base-station can send the NCP that consists of the RPs. Users can extract their own packets from the

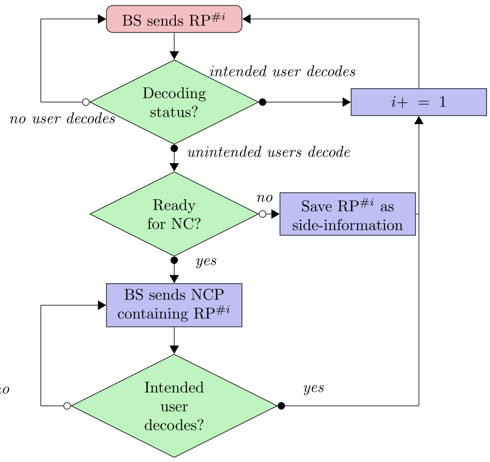


FIGURE 3. **The flowcharts for the MABC models.**

NCP due to the overheard side-information. Therefore, sending one NCP is equivalent to sending multiple RPs simultaneously, which greatly increases the throughput.

We define $\omega_{B,u}$ as the downlink RP from the base-station to user u and $\omega_{u,B}$ as the uplink RP from user u to the base-station for $u \in \{1, \ldots, M\}$, where M is the total number of users.

1) EXAMPLE ($M = 2$ CASE)

Fig. 4 illustrates the NC-RLA scheme for the case of $M = 2$ users. In the figure, a solid line indicates an intended transmission for the target user and the dashed line indicates an unintended overhearing of one packet. Additionally, the line with a '×' mark represents a decoding failure at the receiver. In the first transmission, we assume that the two users successfully overhear each other user's downlink packet but fail to decode their own packets (i.e., user 1 overheard $\omega_{B,2}$ and user 2 overheard $\omega_{B,1}$). During the uplink transmissions, the base-station fails to decode the two uplink packets. We also assume that user 1 overhears $\omega_{2,B}$, but user 2 does not overhear $\omega_{1,B}$. In this case, the base-station sends the NCP $\omega_{B,1} \oplus \omega_{B,2}$ where $\oplus$ indicates the network coding operation of packets. With the side-information, each user is able to extract its own packet from the NCP. For example, user 1 can get $\omega_{B,1}$ by using the network coding operation $(\omega_{B,1} \oplus \omega_{B,2}) \oplus \omega_{B,2}$. One NCP is intended for both users. They are able to decode their intended downlink packets at the same time. The NCed-ARQ approach increases the downlink throughput because it will take the base-station fewer transmissions of the NCP to send the packets to their intended users.

In our proposed protocol, we also allow users to send NCPs, which we call RLA-NCPs to distinguish them from NCPs sent by the base-station. Intuitively, these RLA-NCPs benefit downlink performance even though the are transmitted on the uplink. Each user sends the NCP that consists of its own uplink packet combined with the downlink packet of the other user (i.e., user 1 sends $\omega_{1,B} \oplus \omega_{B,2}$ and user 2 sends $\omega_{2,B} \oplus \omega_{B,1}$). Since the base-station knows the downlink packets $\omega_{B,1}$ and $\omega_{B,2}$, the uplink transmissions will not be affected. Similarly, each user can decodes the RLA-NCP using the overheard uplink packet to recover its own packet. Based on the side-information collected by the users, each RLA-NCP can be either useful or useless. In this example, user 2 cannot decode the RLA-NCP from user 1 because it failed to overhear $\omega_{1,B}$ in Round I. Since user 1 overheard $\omega_{2,B}$, it gets a ''useful'' transmission of RLA-NCP from user 2.

2) GENERAL CASE ($M > 2$)

We apply this idea to the case with $M > 2$ users. At first, all nodes transmit RPs only. The packets received by unintended users will be saved as side-information. The base-station uses the acknowledgements to group the downlink packets. Assuming that m users , with $2 \leq m \leq M$, have overheard $m - 1$ other users' packets but fail to decode their own packets (e.g., user 2 decodes $\omega_{B,1}, \omega_{B,3}, \ldots, \omega_{B,m}$), the base-station groups the downlink packets for the m users for network coding transmissions. Note that there may be multiple groups of packets to apply the network coding. For the example with three users, we have four possible groups of packets. If user 1 and user 2 failed to decode their own packets and overheard the other user's packet, the packets for user 1 and user 2 can be grouped in a NCP $(\omega_{B,1} \oplus \omega_{B,2})$. Similarly, we have two other possible groups, those intended for user 1 and user 3 $(\omega_{B,1} \oplus \omega_{B,3})$ and those intended for user 2 and user 3 $(\omega_{B,2} \oplus \omega_{B,3})$. In the case that all three users successfully overheard the others' packets but failed to decode their own packets, the base-station will group all three packets $(\omega_{B,1} \oplus \omega_{B,2} \oplus \omega_{B,3})$.

For a group with m users, the base-station can send NCP $\omega_{B,1} \oplus \omega_{B,2} \oplus \ldots \oplus \omega_{B,m}$. In the uplink, each user sends a network-coded combination of that user's uplink packet and the $m - 1$ downlink packets intended for the other users. For example, user 1 can send an RLA-NCP, $\omega_{1,B} \oplus \omega_{B,2} \oplus \ldots \oplus \omega_{B,m}$. User 2 can decode the RLA-NCP and recover its downlink packet if it has decoded $\omega_{1,B}$. Therefore, our proposed scheme works well with $M > 2$ users. As the number of users increases, the downlink transmissions will achieve more gains.

B. DERIVATION OF RETRANSMISSION NUMBER

This section is dedicated to the derivation of the average number of retransmissions for the downlink packets. We define R_d as the average data rate for a downlink packet per user. Note that we assume the same rate for the base-station and relay. As a result, we have $R_d = R_b = R_r$ for the following analysis. For the downlink transmission from one base-station to M users, the average number of transmission required for successful decoding of one packet in the downlink is denoted by $S_{d,M}$. Assuming that we have no limit on maximum number of retransmissions, the downlink throughput with M users is defined as

$$T_{d,M} \triangleq \frac{R_d}{S_{d,M}}. \tag{4}$$

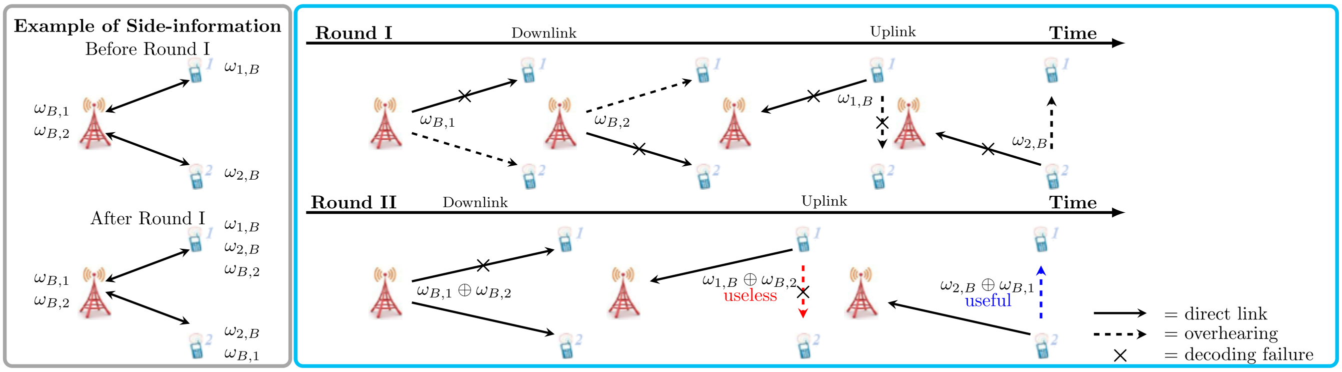


FIGURE 4. NC-RLA for a 2-user case.

After the k^{th} downlink packet transmission, we define $I_{bu,k}$ as mutual information over channel $b \to u$ for user u, and also define $I_{RLA,k}$ as mutual information of the RLA-NCPs sent by other users during the uplink transmissions. The achievable mutual information (in nats/Hz/s) is $I_{bu,k} = \ln(1 + \rho_{bu,k})$, where we define the SNR random variable at the k^{th} transmission of each downlink packet as $\rho_{bu,k} \sim Exp(\lambda_{bu})$ which is identically and independently distributed (i.i.d.) across k. We also define events $A_{bu,k} = \{I_{bu,k} > R_d\}$ and $A_{RLA,k} = \{I_{RLA,k} > R_d\}$. The probability that the sequences of mutual information $I_{bu,1}, I_{bu,2}, \ldots, I_{bu,k}$ and $I_{RLA,1}, I_{RLA,2}, \ldots, I_{RLA,k}$ at each user do not achieve rate R_d (i.e., decoding is unsuccessful) is respectively given by

$$\begin{aligned} p_b(k) &= \Pr\{\bar{A}_{bu,1}, \bar{A}_{bu,2}, \ldots, \bar{A}_{bu,k}\}, \\ p_{RLA}(k) &= \Pr\{\bar{A}_{RLA,1}, \bar{A}_{RLA,2}, \ldots, \bar{A}_{RLA,k}\}, \end{aligned} \tag{5}$$

where $\Pr\{A, B\}$ means $\Pr\{A \cap B\}$.

To evaluate the average number of retransmissions required for delivering a packet, we analyze the downlink using two different phases.

1) PHASE I OF THE DOWNLINK

In this phase, the base-station sends RPs until at least one user successfully decodes the downlink packet. If the packet is successfully decoded by the intended user, the transmission of this packet is completed. If not, the packet is saved for Phase II.

After all downlink packets have been sent once, the users take turns transmitting their uplink packets to the base-station. In this downlink phase, users also transmit RPs only during uplink and we have no RLA in Phase I. Hence, $p_{RLA}(k) = 1$ for all k and can be omitted in this phase. Considering that the decoding results of users are i.i.d., the probability that all M users fail to decode one downlink packet after k^{th} transmission from the base-station is $p_b(k)^M$. After Phase I, the average number of transmissions for one downlink packet to be successfully decoded by at least one user is

$$S_{bu,1,M} = \sum_{k=1}^{\infty} k(p_b(k-1)^M - p_b(k)^M) = \sum_{k=0}^{\infty} p_b(k)^M. \tag{6}$$

2) PHASE II OF THE DOWNLINK

After receiving ACKs, the base-station sends NCPs during Phase II. In this downlink phase, users are allowed to transmit RLA-NCPs during the uplink periods. We define a specific event $\Omega(m, j, t)$ in which exactly m unintended users decode one identical downlink packet in Phase I at the j^{th} transmission and the intended user of the packet needs additional t transmissions to decode the packet, ($m \in \{1, \ldots, M-1\}$, $j \in \{1, \ldots, \infty\}$ and $t \in \{1, \ldots, \infty\}$). In this case, this packet can be included into a network coding group with $m + 1$ packets. In Phase II, the base-station transmits the downlink NCP that consists of these packets. The average number of retransmissions per packet needed in the second phase is

$$S_{bu,2,M} = \sum_{j=1}^{\infty} \sum_{t=1}^{\infty} \sum_{m=1}^{M-1} \frac{t}{m+1} \Pr\{\Omega(m, j, t)\}. \tag{7}$$

Because the NCP contains $m + 1$ packets, the scale factor of $m + 1$ is included. Event $\Omega(m, j, t)$, which is worth mentioning, also indicates that Phase II and RLA start after the j^{th} downlink transmission. Hence, $p_{RLA}(k) = 1$ for $k = 1, 2, \ldots, j$.

The probability $\Pr\{\Omega(m, j, t)\}$ can be derived through the following independent events.

(a) All the $M - 1$ unintended users failed to decode the packet before the j^{th} transmission with probability $P_{(a)}(j) = \Pr\{\bar{A}_{bu,1}, \bar{A}_{bu,2}, \ldots \bar{A}_{bu,j-1}\}^M = p_b(j-1)^{M-1}$.

(b) Exactly m unintended users decode the packet at the j^{th} transmission during Phase I with probability $P_{(b)}(m, j) = \binom{M-1}{m} \Pr\{A_{bu,j} | I_{bu,j-1}\}^m \Pr\{\bar{A}_{bu,j} | I_{bu,j-1}\}^{M-1-m}$.

(c) The packet is successfully decoded by its intended user at exactly the $(t + j)^{th}$ transmission with probability $P_{(c)}(t, j) = p_b(t+j-1)p_{RLA}(t+j-1) - p_b(t+j)p_{RLA}(t+j)$.

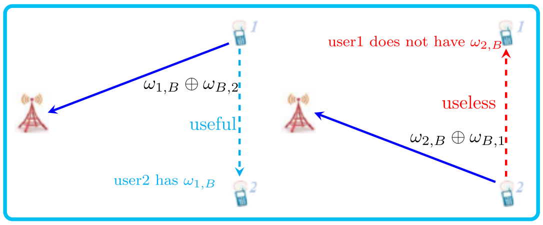


FIGURE 5. Transmissions of RLA-NCPs for a 2-user case.

As a result, $\Pr\{\Omega(m,j,t)\} = P_{(a)}(j)P_{(b)}(m,j)P_{(c)}(t,j)$, and we have

$$S_{bu,2,M} = \sum_{m=1}^{M-1} \frac{1}{m+1} \sum_{j=1}^{\infty} P_{(a)}(j)P_{(b)}(m,j) \sum_{t=1}^{\infty} tP_{(c)}(t,j). \quad (8)$$

C. DERIVATION OF OUTAGE PROBABILITY

In this section, we derive the outage probabilities needed to evaluate the average number of retransmissions. As defined earlier, the achievable mutual information at the k^{th} transmission is $I_{bu,k} = \ln(1+\rho_{bu,k})$, where $\rho_{bu,k} \overset{i.i.d.}{\sim} Exp(\lambda_{bu})$. The probability that the user fails to decode its downlink packet after one retransmission from the base-station (i.e., the outage probability for one transmission over channel $b \to u$) is given by

$$q_{bu} = \Pr\{\bar{A}_{bu,k}\} = \Pr\{\rho_{bu,k} \le e^{R_d} - 1\} = 1 - e^{-\lambda_{bu}\Theta_d}, \quad (9)$$

where $\Theta_d = e^{R_d} - 1$. With our ARQ protocol, the receiver discards the packets that it cannot decode. Hence, the retransmissions of each packet are independent, and we have

$$p_b(k) = \Pr\{\bar{A}_{bu,1}, \bar{A}_{bu,2}, \ldots, \bar{A}_{bu,k}\} = \Pr\{\bar{A}_{bu,1}\}^k = q_{bu}^k. \quad (10)$$

Similarly, we define q_{ub} as the outage probability for one uplink transmission over channel $u \to b$. We assume the channels between any two different users ($u \to u'$) are symmetric and have the same average receive SNR. We define $q_{uu'}$ as the probability that mutual information of an uplink packet (RP or RLA-NCP) collected in one retransmission does not achieve the rate at another user. For uplink transmissions, we define P_u as the transmit power for the symbols sent by the users and define L_{ub} and $L_{uu'}$ as the path loss over these two channels. Obviously, $L_{ub} = L_{bu}$ and $L_{uu'} = Kd_{uu'}^{\gamma}$, where $d_{uu'}$ is the average distance between any two users. The SNRs over these two channels are distributed as $\rho_{ub,k} \overset{i.i.d.}{\sim} Exp(\lambda_{ub})$ and $\rho_{uu',k} \overset{i.i.d.}{\sim} Exp(\lambda_{uu'})$, respectively, where we define $\lambda_{ub} = L_{ub}N_0/P_u$ and $\lambda_{uu'} = L_{uu'}N_0/P_u$ and have

$$q_{ub} = 1 - e^{-\lambda_{ub}\Theta_u}, \quad q_{uu'} = 1 - e^{-\lambda_{uu'}\Theta_u}, \quad (11)$$

where $\Theta_u = e^{R_u} - 1$ and R_u is the uplink data rate in units of nats/Hz/s.

In Phase I of the downlink, we assume that $m+1$ users overheard the others' downlink packets (i.e., these $m+1$ users will start receiving downlink NCPs in Phase II). Each of these $m+1$ users is also allowed to send RLA-NCPs that consist of their own uplink RP and all other m users' downlink RPs. However, since we have no feedback for channel $u \to u'$, a RLA-NCP is considered "useful" for a specific user if and only if that user has already overheard the uplink RP needed to extract the downlink packet from RLA-NCP (as illustrated in Fig. 5). Indeed, each new uplink RP has to be sent at least once before being combined in an RLA-NCP. We assume that a user will send a new uplink RP once an old one is successfully decoded by the base-station. Each user will either send uplink RP or RLA-NCP in each uplink transmission.

Since Phase I and Phase II of the downlink end at the j^{th} and k^{th} downlink transmissions, respectively, all $m+1$ users are allowed to transmit RLA-NCPs during the $k-j$ transmissions. To evaluate the average number of RLA-NCPs sent by users during a specific number of uplink transmissions, we define the probability $\Phi(k-j, n)$ that, a user sends n different RPs, $n \in \{0, \ldots, k-j\}$ in total $k-j$ uplink transmissions during Phase II of the downlink (i.e., the user sends $k-j-n$ RLA-NCPs during $k-j$ transmissions). Considering i.i.d. uplink transmissions, n is binomially distributed as

$$\Phi(k-j, n) = \binom{k-j}{n} q_{ub}^{k-j-n}(1-q_{ub})^n. \quad (12)$$

Due to the symmetric network, the number of RLA-NCPs that one user transmits to another user equals the number of RLA-NCPs received from the same user. Hence, each user receives $k-j-n$ RLA-NCPs from every other user during the $k-j$ uplink transmissions. The total number of RLA-NCPs one user receives from other users is $m(k-j-n)$, in which the number of useful RLA-NCPs for the user is

$$\xi(m, k-j, n) = m(1-q_{uu'})(k-j-n). \quad (13)$$

In (13), we set the probability for each user to successfully decode a RP or RLA-NCP (whether it is "useful" or not) to $1-q_{uu'}$, which is identical for all users because we assumed a symmetric network. Then we have

$$\begin{aligned} p_{RLA}(k) &= \sum_{n=0}^{k-j} q_{uu'}^{\xi(m,k-j,n)} \Phi(k-j, n) \\ &= (1 - q_{ub} + q_{uu'}^{m(1-q_{uu'})} q_{ub})^{k-j} = q_{RLA}^{k-j}, \quad (14) \end{aligned}$$

where $q_{RLA} \triangleq 1 - q_{ub} + q_{uu'}^{m(1-q_{uu'})} q_{ub}$. Here, the variable m, k and j are the same variables in (8), whose probabilities are considered in $\Pr\{\Omega(m,j,t)\}$ when performing throughput analysis. When two users are close to each other, the channel between the two users is often higher capacity than the channel between either of the users and the base-station. It is easier for a user to overhear uplink RPs or RLA-NCPs successfully from the close users. The channels between users can be regarded as lossless ones and $q_{uu'} \to 0$. Hence, we have $q_{RLA} = 1 - q_{ub}$ as a best case for RLA.

D. DERIVATION OF DOWNLINK THROUGHPUT IN MABC

To derive the average number of retransmissions for one downlink packet in both of the phases, we plug the outage probabilities provided in (10) and (14) into (6) and (7), respectively. In the first downlink phase, the base-station and all users only transmit RPs, and we can derive

$$S_{bu,1,M} = \sum_{k=0}^{\infty} p_b(k)^M = \sum_{k=0}^{\infty} q_{bu}^{kM} = \frac{1}{1-q_{bu}^M}. \quad (15)$$

To simplify $S_{bu,2,M}$, we eliminate the two summations with respect to j and t. In the second downlink phase, the average number of retransmissions of one packet is derived as

$$\begin{aligned} S_{bu,2,M} &= \sum_{m=1}^{M-1} \frac{1}{m+1} \sum_{j=1}^{\infty} P_{(a)}(j) P_{(b)}(m,j) \sum_{t=1}^{\infty} t P_{(c)}(t,j) \\ &= \sum_{m=1}^{M-1} \frac{1}{m+1} \binom{M-1}{m} \sum_{j=1}^{\infty} q_{bu}^{(j-1)(M-1)} \\ &\quad \times (1-q_{bu})^m q_{bu}^{M-1-m} \sum_{t=0}^{\infty} q_{bu}^{t+j} q_{RLA}^t \\ &= \sum_{m=1}^{M-1} \binom{M}{m+1} \frac{q_{bu}^{M-m}(1-q_{bu})^m}{M(1-q_{bu}^M)(1-q_{bu}q_{RLA})}. \quad (16) \end{aligned}$$

With NC-RLA and M users, the average number of transmissions per downlink packet is

$$\begin{aligned} S_{d,M} = S_{d,1,M} + S_{d,2,M} &= \frac{1}{1-q_{bu}^M} \\ &+ \sum_{m=1}^{M-1} \binom{M}{m+1} \frac{q_{bu}^{M-m}(1-q_{bu})^m}{M(1-q_{bu}^M)(1-q_{bu}q_{RLA})}, \quad (17) \end{aligned}$$

and the downlink throughput is

$$T_{d,M} = \frac{R_d}{1/(1-q_{bu}^M) + \sum_{m=1}^{M-1} \binom{M}{m+1} \frac{q_{bu}^{M-m}(1-q_{bu})^m}{M(1-q_{bu}^M)(1-q_{bu}q_{RLA})}}. \quad (18)$$

Note that, if we set $M = 1$, then $T_{d,1} = R_d(1-q_{bu})$ and we get the downlink throughput without network-coding. If we only have two users (i.e., $M = 2$), the summation in (18) can be removed and the throughput becomes

$$T_{d,2} = \frac{R_d}{1/(1-q_{bu}^2) + q_{bu}/(1-q_{bu}q_{RLA})}. \quad (19)$$

IV. PROPOSED PROTOCOL – MABRC MODELS

In this section, we consider two-way communication with one base-station, one relay, and M users. As introduced in Section II, a full-duplex relay transmits and receives simultaneously over the same frequency-band. This full-duplex operation, however, generates self-interference at the relay. To combat this, we let both the relay and the users send ARQ feedback during the uplink period. As assumed earlier, the ARQ feedback is perfect. The retransmissions by the base-station depend on the ACK/NACKs from both the relay and the users, and the relay's retransmissions depend on the ACKs from the users only. Each downlink packet is transmitted by the base-station until it is successfully decoded by the relay or the intended user. If the relay decodes the downlink packet but the target user fails to decode it, the retransmission of this packet is forwarded to the relay. To improve the downlink performance, we apply NC-RLA at the relay.

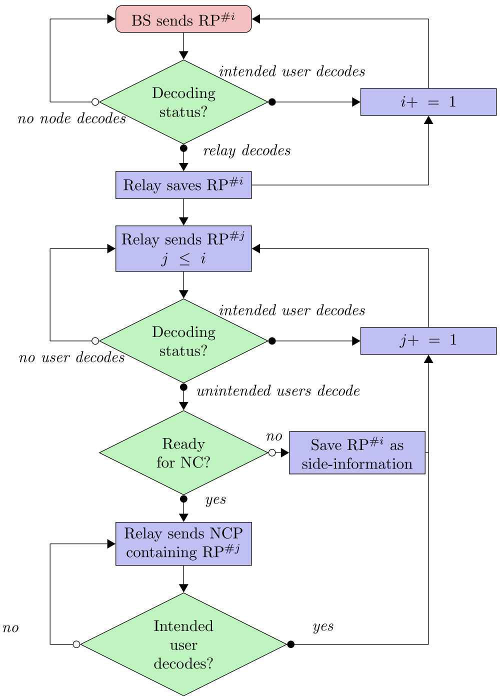
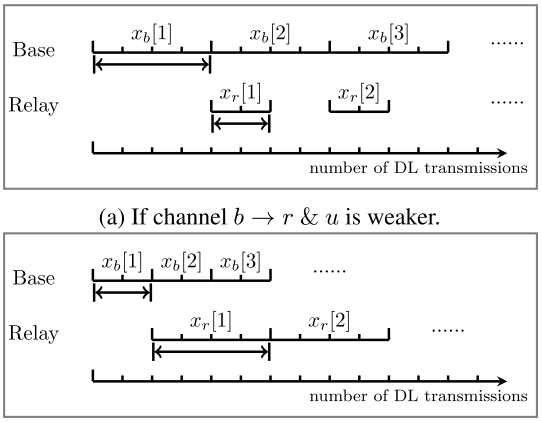


FIGURE 6. The flowcharts for the MABRC models.

A. PROTOCOL: NETWORK CODING AT RELAY AND POWER ALLOCATION

In MABRC systems, either the base-station or the relay can send downlink NCPs. If we apply network coding at the base-station level, the network-coding gain from the downlink NCPs sent by the base-station can be achieved when multiple users overhear all the others' downlink RPs but fail to decode their own RPs. The probability that unintended users overhear the downlink packets in a MABRC system is much smaller than that of downlink packets being decoded by and forwarded to the relay. Indeed, most retransmissions to the users are done at the relay. Therefore, as shown in the flowchart, Fig. 6, we apply network coding at the relay instead of the base-station to improve the overall downlink performance. Similar to the protocol introduced in Section III, the relay works as a central node that sends downlink NCPs. We also divide the transmissions of each downlink packet by the relay into two phases, where the relay transmits downlink RPs in Phase I and downlink NCPs in Phase II. The relay starts to transmit downlink NCPs when m users successfully decode all the others' RPs from the relay but not their own. In addition, all m users are allowed to transmit RLA-NCPs during the uplink periods.

Due to differences in transmit power, distance, and interference, the channels $b \to r$ & u and $r \to u$ are usually unbalanced. As a result, it takes more retransmissions for either



(a) If channel $b \rightarrow r$ & u is weaker.

(b) If channel $r \rightarrow u$ is weaker.

FIGURE 7. The unbalance between channels.

the base-station or the relay to transmit a downlink packet successfully. In Fig. 7, $x_b[i]$ and $x_r[i]$ indicate the i^{th} (for $i \in \{1, 2, \ldots\}$) downlink packets sent by the base-station and the relay, respectively. Obviously, it takes more transmissions on average to transmit one packet over the weaker link. The whole downlink communication is completed when both the base-station and the relay transmit all their downlink packets successfully. As a result, the overall downlink throughput is always dominated by the weaker link.

The downlink throughput reaches its maximum when the channels $b \rightarrow r$ & u and $r \rightarrow u$ are balanced (i.e., it takes almost the same number of retransmissions for the base-station and relay to transmit a downlink packet successfully). To achieve this optimal performance, we allow the base-station and relay to allocate their transmit power. As defined in Section II, P_b and P_r are the transmit powers of the base-station and relay, respectively. We assume a fixed sum of transmit power at the base-station and the relay (i.e., $P_b + P_r$ is fixed). We start to increase the relay's transmit power from zero and decrease the base-station's transmit power. At first, the throughput is dominated by channel $r \rightarrow u$. As the relay's transmit power increases, channel $r \rightarrow u$ becomes stronger and the throughput increases. The decrease in the base-station's transmit power and increase in interference causes channel $b \rightarrow r$ & u to weaken. The throughput will be dominated by channel $b \rightarrow r$ & u once becomes the weaker link. The throughput decreases as the relay's transmit power increases. Hence, we aim to balance the two channels with power allocation to achieve the best performance.

B. DERIVATION OF RETRANSMISSION NUMBER

This section is dedicated to derive the average number of retransmissions for MABRC models using the same assumptions listed in Section II. As defined earlier, R_b and R_r are the downlink data rates of the base-station and relay. Generally, we set $R_b = R_r = R_d$, where R_d is the downlink rate. We define S_b as the average number of transmissions from the base-station for a downlink packet to be successfully decoded by either the relay or the target user. Let $S_{r,M}$ denote the average number of transmissions from the relay for a downlink packet assuming network coding at the relay and M users. To adapt the transmit power of the base-station and the relay, we also define β as the power allocation parameter, where $\beta P_b = (1-\beta)P_r$. With power allocation, we then define

$$S_{d,M}^{pa} \triangleq \min_{\beta}\{\max\{S_b, S_{r,M}\}\}, \tag{20}$$

as the average number of transmissions required for a user to successfully decode a downlink packet minimized over all possible power allocation.

After the k^{th} transmission of one downlink packet, we define $I_{bu,k}$, $I_{br,k}$, and $I_{ru,k}$ as mutual information over channel $b \rightarrow u$, $b \rightarrow r$, and $r \rightarrow u$, respectively. As discussed in Section II, $I_{br,k} = \ln(1+\rho_{br,k})$ (in nats/Hz/s), where we define the random variable $\rho_{br,k} \overset{i.i.d.}{\sim} Exp(\lambda_{br})$. Due to multiple-access transmissions at users, $I_{bu,k}$ and $I_{ru,k}$ can be derived through the capacity region in (3). We also define events $A_{bu,k} = \{I_{bu,k} > R_b := R_d\}$, $A_{br,k} = \{I_{br,k} > R_b := R_d\}$ and $A_{ru,k} = \{I_{ru,k} > R_r := R_d\}$. The probability that the sequences of mutual information $I_{bu,1}, I_{bu,2}, \ldots, I_{bu,k}, I_{br,1}, I_{br,2}, \ldots, I_{br,k}$ fail to achieve the rate R_b at the receiver is given by $p_b(k) = Pr\{\bar{A}_{bu,1}, \bar{A}_{bu,2}, \ldots, \bar{A}_{bu,k}, \bar{A}_{br,1}, \bar{A}_{br,2}, \ldots, \bar{A}_{br,k}\}$. The probability that the sequences of mutual information $I_{ru,1}, I_{ru,2}, \ldots, I_{ru,k}$ at the receiver do not achieve rate R_r is given by $p_r(k) = Pr\{\bar{A}_{ru,1}, \bar{A}_{ru,2}, \ldots, \bar{A}_{ru,k}\}$.

The average number of transmissions per packet from the base-station is given by

$$S_b = \sum_{k=1}^{\infty} k(p_b(k-1) - p_b(k)) = \sum_{k=0}^{\infty} p_b(k). \tag{21}$$

To evaluate the average number of transmissions per packet from the relay assuming M users, we define $P_{rem,r}$ as the probability that the relay successfully decodes one downlink packet before the target user (i.e., the downlink packet is forwarded to the relay for future retransmissions). The notation $S_{rem,r,M}$ denotes the average number of transmissions per packet from the relay to be decoded successfully by the intended user if the packet is forwarded to the relay. Then we have

$$\begin{aligned} P_{rem,r} = & \sum_{k=1}^{\infty} \Pr\{\bar{A}_{bu,1}, \bar{A}_{bu,2}, \ldots, \bar{A}_{bu,k}\} \\ & \times (\Pr\{\bar{A}_{br,1}, \bar{A}_{br,2}, \ldots, \bar{A}_{br,k-1}\} \\ & - \Pr\{\bar{A}_{br,1}, \bar{A}_{br,2}, \ldots, \bar{A}_{br,k}\}). \end{aligned} \tag{22}$$

Note that $S_{rem,r,M}$ can be derived using (17) because the relay acts as a central node. The average number of transmissions per packet from the relay is given by

$$S_{r,M} = P_{rem,r} S_{rem,r,M}. \tag{23}$$

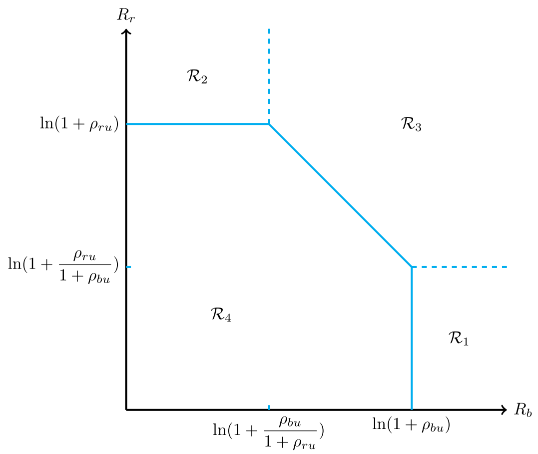


FIGURE 8. Rate region for multiple access channel.

C. DERIVATION OF OUTAGE PROBABILITIES FOR MULTIPLE ACCESS CHANNEL

Similar to the case in MABC models, with ARQ protocol, the outage probabilities over channels $b \to u$, $b \to r$, and $r \to u$ are i.i.d. over each retransmission. We define the outage probabilities for each transmission over these three channels as $q_{bu} = \Pr\{\bar{A}_{bu,k}\}$, $q_{br} = \Pr\{\bar{A}_{br,k}\}$, and $q_{ru} = \Pr\{\bar{A}_{ru,k}\}$, respectively. These outage probabilities are derived as follows:

1) CHANNEL $b \to r$

As defined earlier, $I_{br,k} = \ln(1+\rho_{br,k})$. The outage probability for each transmission over this channel is

$$q_{br} = \Pr\{\bar{A}_{br,1}\} = \Pr\{I_{br,1} \le R_d\} = 1 - e^{-\lambda_{br}\Theta_d}. \quad (24)$$

2) CHANNEL $b \to u$ AND $r \to u$

Due to the multiple-access transmissions, according to the relations between mutual information and the target rates of channel $b \to u$ and $r \to u$, we have the following rate regions (as shown in Fig. 8 and the variables ρ_{bu} and ρ_{ru} given in Section II):

$$\mathcal{R}_1 = \{\mathbf{R} : R_b > \ln(1+\rho_{bu}), R_r \le \ln(1+\frac{\rho_{ru}}{1+\rho_{bu}})\}, \quad (25)$$

$$\mathcal{R}_2 = \{\mathbf{R} : R_r > \ln(1+\rho_{ru}), R_b \le \ln(1+\frac{\rho_{bu}}{1+\rho_{ru}})\}, \quad (26)$$

$$\mathcal{R}_3 = \Omega - \mathcal{R}_1 \cup \mathcal{R}_2 \cup \mathcal{R}_4, \quad (27)$$

$$\begin{aligned}\mathcal{R}_4 = \{\mathbf{R} : R_b \le \ln(1+\rho_{bu}), R_r \le \ln(1+\rho_{ru}),\\ R_b + R_r \le \ln(1+\rho_{bu}+\rho_{ru})\},\end{aligned} \quad (28)$$

where $R_b = R_r = R_d$, as assumed before. We can derive the probability of each region as

$$\begin{aligned}\Pr\{\mathcal{R}_1\} &= \Pr\{\rho_{bu} < \Theta_d, \rho_{ru} \ge (1+\rho_{bu})\Theta_d\}\\ &= \int_0^{\Theta_d} [1 - F_{\rho_{ru}}((1+x)\Theta_d)] f_{\rho_{bu}}(x)dx\\ &= \frac{\lambda_{bu}e^{-\lambda_{ru}\Theta_d}}{\lambda_{bu}+\lambda_{ru}\Theta_d}\left(1 - e^{-(\lambda_{bu}+\lambda_{ru}\Theta_d)\Theta_d}\right),\end{aligned} \quad (29)$$

$$\begin{aligned}\Pr\{\mathcal{R}_2\} &= \Pr\{\rho_{ru} < \Theta_d, \rho_{bu} \ge (1+\rho_{ru})\Theta_d\}\\ &= \int_0^{\Theta_d} [1 - F_{\rho_{bu}}((1+x)\Theta_d)] f_{\rho_{ru}}(x)dx\\ &= \frac{\lambda_{ru}e^{-\lambda_{bu}\Theta_d}}{\lambda_{ru}+\lambda_{bu}\Theta_d}\left(1 - e^{-(\lambda_{ru}+\lambda_{bu}\Theta_d)\Theta_d}\right),\end{aligned} \quad (30)$$

$$\begin{aligned}\Pr\{\mathcal{R}_4\} &= \Pr\{\rho_{bu} \ge \Theta_d, \rho_{ru} \ge \Theta_d\}\\ &\quad - \Pr\{\rho_{bu} \ge \Theta_d, \rho_{ru} \ge \Theta_d, \rho_{bu}+\rho_{ru} < \Theta_d^2 + 2\Theta_d\}\\ &= [1 - F_{\rho_{bu}}(\Theta_d)][1 - F_{\rho_{ru}}(\Theta_d)]\\ &\quad - \int_{\Theta_d}^{\Theta_d^2+\Theta_d} [F_{\rho_{ru}}(\Theta_d^2+2\Theta_d - x) - F_{\rho_{ru}}(\Theta_d)] f_{\rho_{bu}}(x)dx\\ &= \frac{\lambda_{bu}e^{-\lambda_{ru}\Theta_d^2} - \lambda_{ru}e^{-\lambda_{bu}\Theta_d^2}}{\lambda_{bu}-\lambda_{ru}} e^{-(\lambda_{bu}+\lambda_{ru})\Theta_d},\end{aligned} \quad (31)$$

where $F_{\rho_{bu}}, f_{\rho_{bu}}$ and $F_{\rho_{ru}}, f_{\rho_{ru}}$ are the cumulative and probability distribution functions of ρ_{bu} and ρ_{ru}, respectively. The outage probabilities for channel $b \to u$ and channel $r \to u$ are

$$q_{bu} = 1 - \Pr\{\mathcal{R}_2\} - \Pr\{\mathcal{R}_4\}, \quad q_{ru} = 1 - \Pr\{\mathcal{R}_1\} - \Pr\{\mathcal{R}_4\}. \quad (32)$$

Due to the i.i.d. retransmissions in our ARQ scheme, with these outage probabilities, we have

$$\begin{aligned}p_b(k) &= \Pr\{\bar{A}_{bu,1}\}^k \Pr\{\bar{A}_{br,1}\}^k = q_{bu}^k q_{br}^k,\\ p_r(k) &= \Pr\{\bar{A}_{ru,1}\}^k = q_{ru}^k.\end{aligned} \quad (33)$$

D. DERIVATION OF DOWNLINK THROUGHPUT IN MABRC

To derive the average number of transmission, we plug $p_b(k)$ into (21) to get

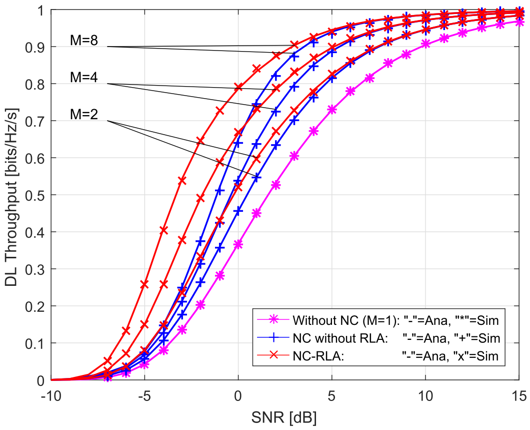
$$S_b = \sum_{k=0}^{\infty} p_b(k) = \sum_{k=0}^{\infty} q_{bu}^k q_{br}^k = \frac{1}{1 - q_{bu}q_{br}}. \quad (34)$$

Due to the i.i.d. retransmissions, we can simplify (22) as $P_{rem,r} = \sum_{k=1}^{\infty} q_{bu}^k (q_{br}^{k-1} - q_{br}^k) = \frac{q_{bu}(1-q_{br})}{1-q_{bu}q_{br}}$. Since the protocol used at the base-station in the MABC model is employed at the relay in the MABRC model, we replace q_{bu} with q_{ru} in (17) to get

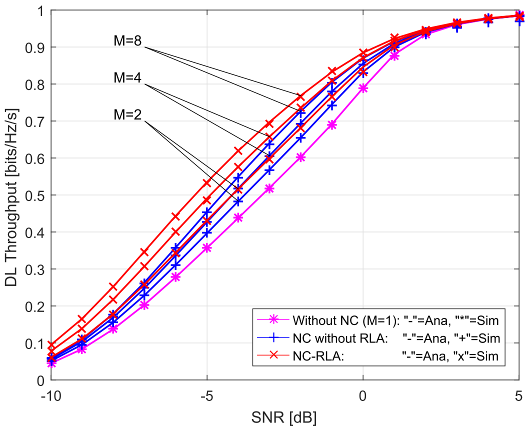
$$\begin{aligned}S_{rem,r,M} &= \frac{1}{1-q_{ru}^M}\\ &\quad + \sum_{m=1}^{M-1} \binom{M}{m+1} \frac{q_{ru}^{M-m}(1-q_{ru})^m}{M(1-q_{ru}^M)(1-q_{ru}q_{RLA})},\end{aligned} \quad (35)$$

The average number of transmissions per packet from the relay is given by

$$\begin{aligned}S_{r,M} &= P_{rem,r}S_{rem,r,M} = \frac{q_{bu}(1-q_{br})}{1-q_{bu}q_{br}}\left(\frac{1}{1-q_{ru}^M}\right.\\ &\quad \left. + \sum_{m=1}^{M-1} \binom{M}{m+1} \frac{q_{ru}^{M-m}(1-q_{ru})^m}{M(1-q_{ru}^M)(1-q_{ru}q_{RLA})}\right).\end{aligned} \quad (36)$$

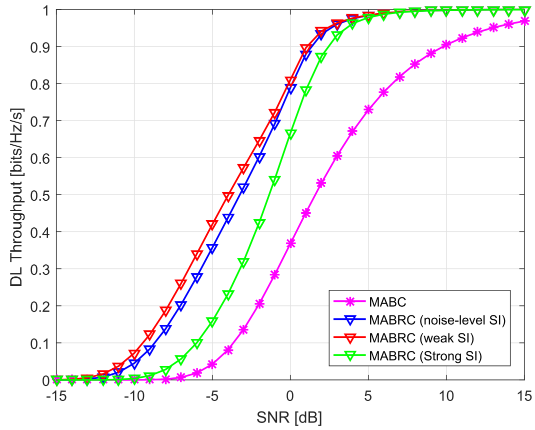


(a) MABC model

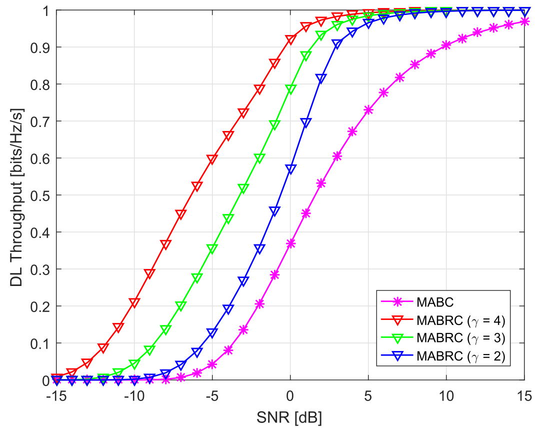


(b) MABRC model

FIGURE 9. Analytical and simulated DL throughput for MABC/MABRC.



(a) $\alpha = 0.5, \gamma = 3$



(b) $\alpha = 0.5$, noise-level SI

FIGURE 10. DL throughput for MABC/MABRC without NC.

Hence, in the MABRC models with power allocation, the downlink throughput is

$$T_{d,M} = \frac{R_d}{S^{pa}_{d,M}} = \frac{R_d}{\min\limits_{\beta}\{\max\{S_b, S_{r,M}\}\}}. \tag{37}$$

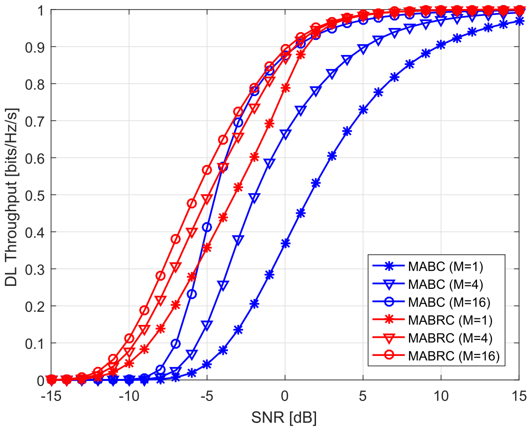
V. PERFORMANCE RESULTS

In this section, we validate our mathematical derivation with the Monte Carlo simulation. We numerically evaluate and compare the throughput for both the MABC and MABRC models using different settings. The associated parameters are as follows. The data-rates $R_d, R_u \in \{1, 2\}$ bits/Hz/s. The transmit power $P_b = P_0$ in the MABC models, and $P_b = (1-\beta)P_0$ and $P_r = \beta P_0$ in the MABRC models. $P_u = P_0$ in both system models. Users are assumed to be close to each other and distant from the base-station. The relay is set between the base-station and users. We set the distance as $d_{bu} = d_0, d_{br} = \alpha d_0, d_{ru} = (1-\alpha)d_0$ and $d_{uu'} = \alpha' d_0$, where α and α' are the distance ratios. Assuming close users, we set $\alpha' = 0.1$. The path loss exponent $\gamma \in \{2, 3, 4\}$. The self-interference level $\psi = \frac{P_0}{N_0}\psi'$, and $\psi' = 10$ (weak), 1 (noise-level) or 0.1 (strong). The mean SNR for channel $b \to u$ is defined as $\lambda_0^{-1} = \frac{P_0}{N_0} \in [-15, 15]$ dB.

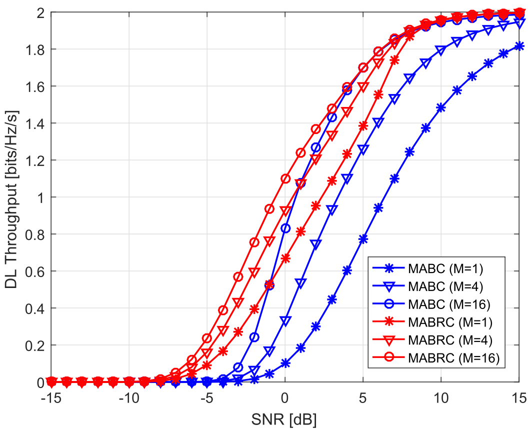
For comparison, the mean inverse SNR/SINRs for all channels are given by $\lambda_{bu} = \frac{Kd_0^2N_0}{P_0} = \lambda_0 = \lambda_{ub}$ and $\lambda_{uu'} = \frac{K(\alpha' d_0)^2N_0}{P_0} = \alpha'^2\lambda_0$ for MABC models, and $\lambda_{bu} = \frac{Kd_0^2N_0}{((1-\beta)P_0)} = \frac{1}{1-\beta}\lambda_0$, $\lambda_{br} = \frac{K(\alpha d_0)^2(\beta N_0/\psi' + N_0)}{((1-\beta)P_0)} = \frac{\alpha^2(\beta+\psi')}{\psi'(1-\beta)}\lambda_0$ and $\lambda_{ru} = \frac{K((1-\alpha)d_0)^2N_0}{(\beta P_0)} = \frac{(1-\alpha)^2}{\beta}\lambda_0$ for MABRC models.

1) VALIDATION WITH MONTE CARLO SIMULATION

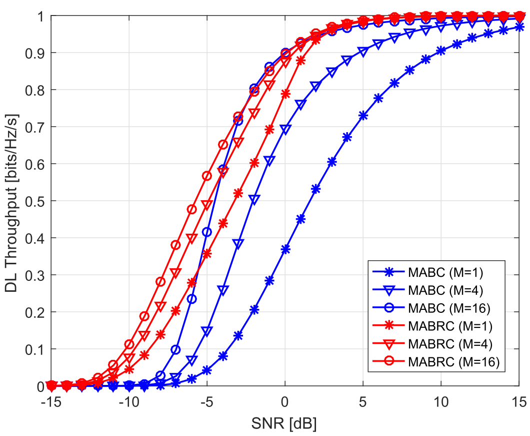
Fig. 9 shows the downlink throughput for both the MABC and MABRC models with $M = 1, 2, 4, 8$ users. The lines show the analytical results, and the markers show the Monte Carlo simulation results. The comparison shows that the analytical and simulation results match well, which validates our derivations. Note that our proposed NC-RLA scheme greatly outperforms the classical NCed-ARQ scheme for both the MABC (Fig. 7(a)) and MABRC (Fig. 7(b)) models. The greatest gain is achieved when the SNR is between -10 and 5 dB. kThe downlink throughput performance significantly improves as the number of users M increases. This means that the proposed NC-RLA scheme offers the potential for a sizable throughput improvement, especially for a large number of users in low/moderate SNRs.



(a) $R_d = R_u = 1$

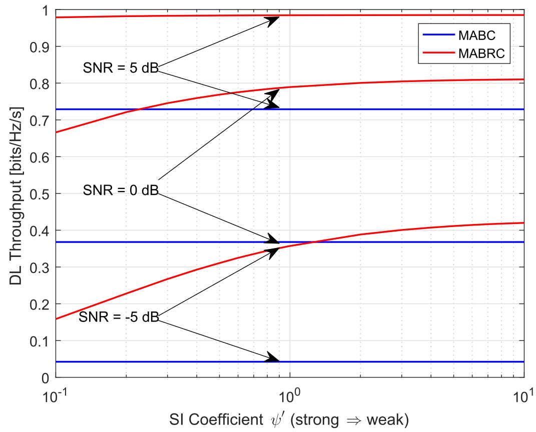


(b) $R_d = 2, R_u = 1$

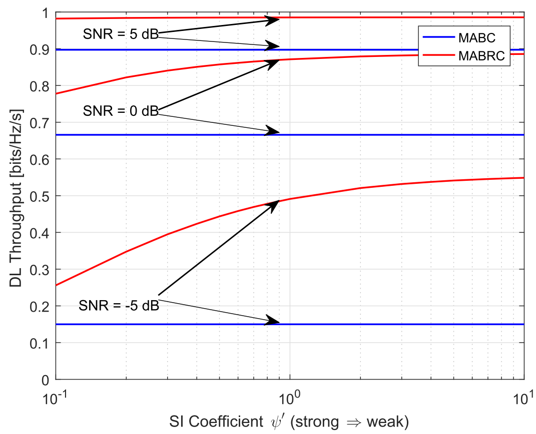


(c) $R_d = 1, R_u = 2$

FIGURE 11. DL throughput for NC-RLA in MABC/MABRC.



(a) $M = 1$ (no NC-RLA)



(b) $M = 4$

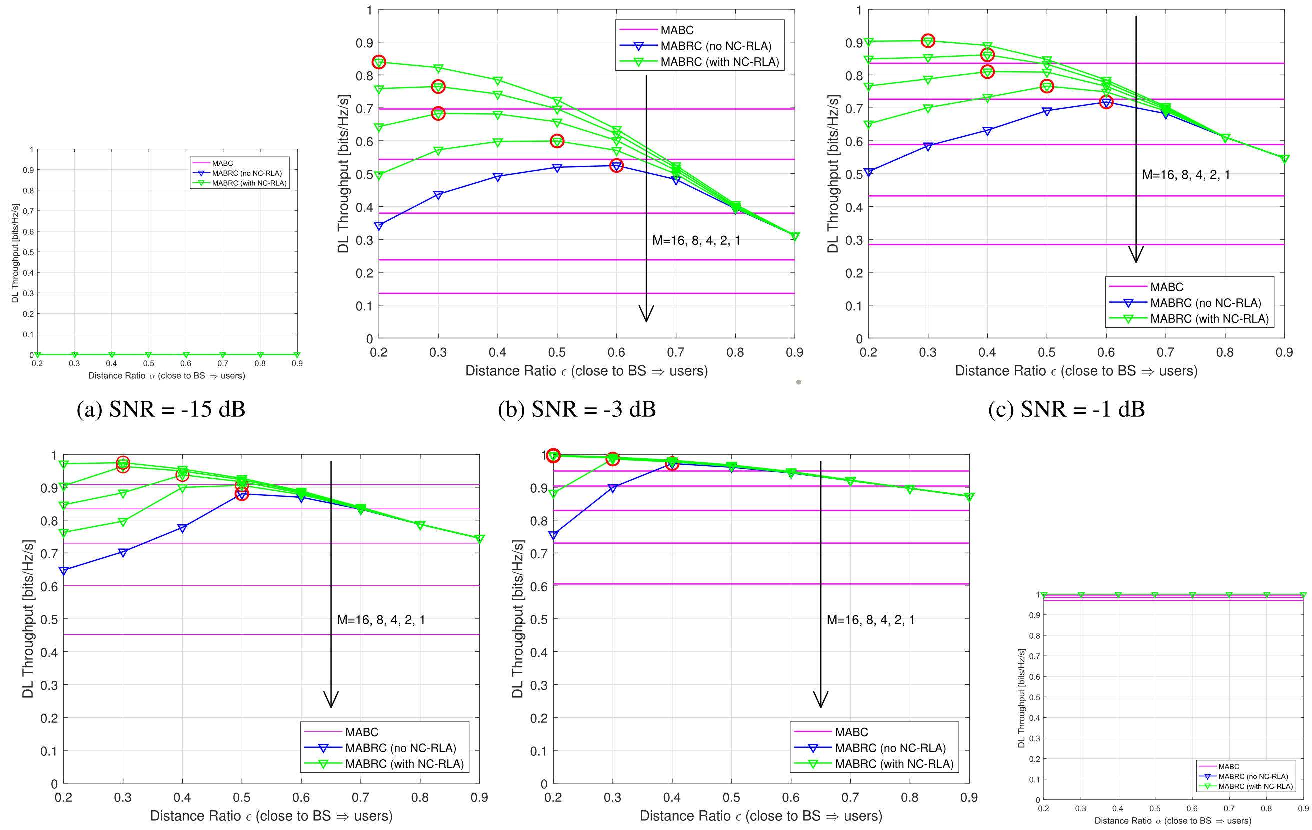
FIGURE 12. DL throughput as function of self-interference coefficient.

2) DOWNLINK THROUGHPUT PERFORMANCE WITHOUT NETWORK CODING

For comparison, we evaluate the performance of the MABC and MABRC without NC-RLA. The achievable downlink throughput is shown in Fig. 10. For the MABRC, we consider a different self-interference level (weak $\psi' = 10$, noise-level $\psi' = 1$, and strong $\psi' = 0.1$) and a different path loss exponent. The average self-interference at the relay equals $N_{SI} = P_r/\psi = (P_r/P_0)N_0/\psi' = \beta N_0/\psi'$; therefore, the self-interference decreases with ψ'. The relay is set at the midpoint between the base-station and users ($\alpha = 0.5$). In Fig. 10(a), we observe that at any level of self-interference, the MABRC outperforms MABC due to power and location optimization. Even with strong self-interference at the relay, the MABRC can still achieve a better performance. Fig. 10(b) shows the performance of the MABRC with a different path loss exponent. Note that the comparison is based on an identical average SNR from the base-station to the users. Obviously, with a stronger path loss, adding a full-duplex relay provides a greater improvement of the MABC.

3) DOWNLINK THROUGHPUT PERFORMANCE WITH NC-RLA

Next, we address the effect of NC-RLA on the throughput performance in both MABC and MABRC. In Fig. 11, we plot the values of achievable throughput with $M = 1, 4, 16$ users and different rate. For the MABRC, we assume that the relay suffers a noise-level self-interference ($\psi' = 1$). We also set $\gamma = 3$ and $\alpha = 0.5$. Note that this is not the best position for the relay, especially for a large number of users. The simulation results demonstrate that using



(a) SNR = -15 dB (b) SNR = -3 dB (c) SNR = -1 dB

(d) SNR = 1 dB (e) SNR = 3 dB (f) SNR = 15 dB

FIGURE 13. DL throughput as a function of distance ratio.

NC-RLA improves downlink throughput, and this improvement increases with M. Similarly, NC-RLA is particularly efficient under low/moderate SNRs for both systems. In Fig. 11(b) and 11(c), we observe that NC-RLA also works well for different downlink/uplink rates.

In Fig. 12, the downlink throughput is plotted as a function of the self-interference coefficient ψ' for the MABC and MABRC with (a) $M = 1$ (i.e., no NC-RLA) and (b) $M = 4$. We observe that a stronger self-interference level leads to a much lower downlink throughput for the MABRC. The gap between the MABC and MABRC decreases as M increases.

4) NUMERICAL OPTIMIZATION FOR THE POSITION OF RELAY

In Fig. 13, we plot the downlink throughput as a function of the distance ratio α in MABC and MABRC with noise-level self-interference. We consider a range of $\alpha \in [0.2, 0.9]$. The lines from bottom to top indicate performances with $M = 1, 2, 4, 8, 16$ users for both systems. As shown in Fig. 13(a) and 13(f), the performance of the MABRC is unaffected by the position of the relay, at both low SNR (= −15 dB) or high SNR (= 15 dB). Hence, in these two cases, there is no need to optimize the position for the relay.

In Fig. 13(b), (c), (d) and (e), the optimal positions of the relay are circled. We see that

- Without NC-RLA, the optimal performance for the MABRC is achieved at $\alpha \in (0.5, 0.6)$ which indicates that the relay should be set at the midpoint between the base-station and the users, but a little closer to the users. If we ignore the direct link $b \rightarrow u$ in the MABRC, the optimal position for the relay is in the middle to balance the resulting SNR over $b \rightarrow r$ and $r \rightarrow u$. When taking $b \rightarrow u$ into consideration, the transmission from the base-station becomes better, but the retransmission from the relay becomes worse due to the competition of multiple access at the users. Therefore, the relay should be set closer to the users to achieve the highest downlink throughput in the MABRC without NC-RLA.
- With NC-RLA, the optimal position for the relay is closer to the base-station as the number of users increases. Indeed, NC-RLA greatly improves the transmissions from the relay. Hence, the position of the relay that balances the two links is set farther from the users.
- It is also important to point out that the MABC can outperform the MABRC if the position of the relay is not optimized with a large number of users. For example in Fig. 13(b), if we set $\alpha \geq 0.6$, with 16 users, the MABC can achieve a higher throughput than the MABRC.

VI. SUMMARY AND CONCLUSION

We proposed a new retransmission protocol for the MABRC model. In this setting, one base-station, one full-duplex relay, and multiple users exchange messages over two-way wireless channels. In this scenario, the relay and each user may hear from the base-station and other users. To improve the performance, we optimized the transmit powers of the base-station and relay in the MABRC. We derived closed-form expressions for the achievable downlink throughput. Our simulation results showed that the new scheme significantly improves the performances of both the MABC and MABRC, particularly for low and moderate SNRs. A combination of both NC-RLA and full-duplex relay provides the best performance; however, if the position is not optimized, adding a relay will not be beneficial when M is large and the self-interference level is high.

ACKNOWLEDGMENT

This article was presented in part at the IEEE International Conference on Communication (ICC'2016), Kuala Lumpur, Malaysia, May 2016.

REFERENCES

[1] "Cisco visual networking index: Global mobile data traffic forecast update, 2017–2022," Cisco, San Jose, CA, USA, White Paper, 2018.
[2] P. Rysavy, "Challenges and considerations in defining spectrum efficiency," *Proc. IEEE*, vol. 102, no. 3, pp. 386–392, Mar. 2014.
[3] V. S. Kaunismaa, "On frequency radio (OFR) repeater as fade filler," in *Proc. IEEE 39th Veh. Technol. Conf.*, May 1989, pp. 528–531.
[4] W. T. Slingsby and J. P. McGeehan, "Antenna isolation measurements for on-frequency radio repeaters," in *Proc. 9th Int. Conf. Antennas Propag. (ICAP)*, vol. 1, Apr. 1995, pp. 239–243.
[5] L. Weng and R. D. Murch, "Multi-user MIMO relay system with self-interference cancellation," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, 2007, pp. 958–962.
[6] P. Larsson and M. Prytz, "MIMO on-frequency repeater with self-interference cancellation and mitigation," in *Proc. IEEE 69th Veh. Technol. Conf. (VTC)*, Apr. 2009, pp. 1–5.
[7] M. Heino, D. Korpi, T. Huusari, E. Antonio-Rodriguez, S. Venkatasubramanian, T. Riihonen, L. Anttila, C. Icheln, K. Haneda, R. Wichman, and M. Valkama, "Recent advances in antenna design and interference cancellation algorithms for in-band full duplex relays," *IEEE Commun. Mag.*, vol. 53, no. 5, pp. 91–101, May 2015.
[8] T. Riihonen, S. Werner, and R. Wichman, "Hybrid full-duplex/half-duplex relaying with transmit power adaptation," *IEEE Trans. Wireless Commun.*, vol. 10, no. 9, pp. 3074–3085, Sep. 2011.
[9] D. W. K. Ng, E. S. Lo, and R. Schober, "Dynamic resource allocation in MIMO-OFDMA systems with full-duplex and hybrid relaying," *IEEE Trans. Commun.*, vol. 60, no. 5, pp. 1291–1304, May 2012.
[10] I. Krikidis, H. A. Suraweera, P. J. Smith, and C. Yuen, "Full-duplex relay selection for amplify-and-forward cooperative networks," *IEEE Trans. Wireless Commun.*, vol. 11, no. 12, pp. 4381–4393, Dec. 2012.
[11] A. Ikhlef, J. Kim, and R. Schober, "Mimicking full-duplex relaying using half-duplex relays with buffers," *IEEE Trans. Veh. Technol.*, vol. 61, no. 7, pp. 3025–3037, Sep. 2012.
[12] H. Jin and V. C. M. Leung, "Performance analysis of full-duplex relaying employing fiber-connected distributed antennas," *IEEE Trans. Veh. Technol.*, vol. 63, no. 1, pp. 146–160, Jan. 2014.
[13] D. Choi and D. Park, "Effective self interference cancellation in full duplex relay systems," *Electron. Lett.*, vol. 48, no. 2, pp. 129–130, Jan. 2012.
[14] R. López-Valcarce, E. Antonio-Rodríguez, C. Mosquera, and F. Pérez-González, "An adaptive feedback canceller for full-duplex relays based on spectrum shaping," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 8, pp. 1566–1577, Sep. 2012.
[15] B. Chun and H. Park, "A spatial-domain joint-nulling method of self-interference in full-duplex relays," *IEEE Commun. Lett.*, vol. 16, no. 4, pp. 436–438, Apr. 2012.
[16] H. Kim, S. Lim, H. Wang, and D. Hong, "Optimal power allocation and outage analysis for cognitive full duplex relay systems," *IEEE Trans. Wireless Commun.*, vol. 11, no. 10, pp. 3754–3765, Oct. 2012.
[17] H. Cui, M. Ma, L. Song, and B. Jiao, "Relay selection for two-way full duplex relay networks with amplify-and-forward protocol," *IEEE Trans. Wireless Commun.*, vol. 13, no. 7, pp. 3768–3777, Jul. 2014.
[18] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1204–1216, Jul. 2000.
[19] D. S. Lun, M. Médard, R. Koetter, and M. Effros, "On coding for reliable communication over packet networks," *Phys. Commun.*, vol. 1, no. 1, pp. 3–20, 2008.
[20] P. Larsson, N. Johansson, and K.-E. Sunell, "Coded bi-directional relaying," in *Proc. IEEE 63rd Veh. Technol. Conf.*, vol. 2, May 2006, pp. 851–855.
[21] M. Ghaderi, D. Towsley, and J. Kurose, "Network coding performance for reliable multicast," in *Proc. IEEE Mil. Commun. Conf. (MILCOM)*, Oct. 2007, pp. 1–7.
[22] P. Larsson, "Multicast multiuser ARQ," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar./Apr. 2008, pp. 1985–1990.
[23] F. Chiti, R. Fantacci, and A. Tassi, "An efficient network coding scheme with symbol combining: Performance evaluation, optimization, and comparisons," in *Proc. IEEE Trans. Veh. Technol.*, vol. 62, no. 3, pp. 1267–1275, Mar. 2013.
[24] P. Larsson, B. Smida, T. Koike-Akino, and V. Tarokh, "Analysis of network coded HARQ for multiple unicast flows," *IEEE Trans. Commun.*, vol. 61, no. 2, pp. 722–732, Feb. 2013.
[25] S. Pu, Z. He, X. Lin, and W. Wu, "Performance analysis of joint Chase combining and network coding in wireless broadcast retransmission," in *Proc. 4th Int. Conf. Wireless Commun., Netw. Mobile Comput.*, Oct. 2008, pp. 1–4.
[26] F. Gabriel, G. T. Nguyen, R.-S. Schmoll, J. A. Cabrera, M. Muehleisen, and F. H. Fitzek, "Practical deployment of network coding for real-time applications in 5G networks," in *Proc. IEEE 15th Annu. Consum. Commun. Netw. Conf. (CCNC)*, Jan. 2018, pp. 1–2.
[27] S. S. Ullah, G. Liva, and S. C. Liew, "Short packet physical-layer network coding with mismatched channel state information," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2018, pp. 1–5.
[28] A. Osseiran, M. Xiao, S. B. Slimane, M. Skoglund, and J. Manssour, "Advances in wireless network coding for IMT-Advanced & beyond," in *Proc. 2nd Int. Conf. Wireless Commun., Veh. Technol., Inf. Theory Aerosp. Electron. Syst. Technol. (VITAE)*, Feb./Mar. 2011, pp. 1–6.
[29] L. F. Xie, I. W.-H. Ho, Z. H. Situ, L. Lu, and W. Lu, "Channel-coded physical-layer network coding with OFDM modulation," *IEEE Access*, vol. 6, pp. 22267–22280, 2018.
[30] S. Pandi, R. T. Arranz, G. T. Nguyen, and F. H. P. Fitzek, "Massive video multicasting in cellular networks using network coded cooperative communication," in *Proc. IEEE 15th Annu. Consum. Commun. Netw. Conf. (CCNC)*, Jan. 2018, pp. 1–2.
[31] J. Gu, R. C. de Lamare, and M. Huemer, "Buffer-aided physical-layer network coding with optimal linear code designs for cooperative networks," *IEEE Trans. Commun.*, vol. 66, no. 6, pp. 2560–2575, Jun. 2018.
[32] H. Zhu, B. Smida, and D. Love, "An efficient network coding scheme for two-way communication with ARQ feedback," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2016, pp. 1–6.
[33] M. Cheng, K. Anwar, and T. Matsumoto, "Outage analysis of correlated source transmission in block Rayleigh fading channels," in *Proc. IEEE Veh. Technol. Conf. (VTC)*, Sep. 2012, pp. 1–5.
[34] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2011.
[35] N. Shende, O. Gurbuz, and E. Erkip, "Half-duplex or full-duplex relaying: A capacity analysis under self-interference," in *Proc. 47th Annu. Conf. Inf. Sci. Syst. (CISS)*, Mar. 2013, pp. 1–6.

HONGYI ZHU received the B.S. degree in communication engineering from Fudan University, Shanghai, China, in 2013, and the Ph.D. degree in electrical engineering from Purdue University, in 2018. His research interests include information theory and wireless systems for 5G and beyond. His previous works were presented at ICC 2016 and ICC 2017.

BESMA SMIDA received the Ph.D. and M.Sc. degrees from the University of Quebec (INRS), Montreal, Canada. After completing her appointment as a Postdoc and later a Lecturer at Harvard University, she became an Assistant Professor of electrical and computer engineering at Purdue University Calumet. Additionally, she was a Research Engineer with the Technology Evolution and Standards Group of Microcell, Inc. (now Rogers Wireless), Montreal, Canada. She took part in wireless normalization committees (3GPP, T1P1). She is currently an Associate Professor of electrical and computer engineering with The University of Illinois at Chicago. Her research interests include wireless communication theory, full-duplex communication, radar, and backscatter modulation. Dr. Smida is a member of the IEEE ComSoc Emerging Technologies Committee and an Associate Editor of the IEEE Communications Letters. She was awarded the INSIGHT Into Diversity Magazine's 2015 ''100 Inspiring Women in STEM''. She received the Academic Gold Medal of the Governor General of Canada, in 2007, and the NSF CAREER Award, in 2015.

DAVID J. LOVE (S'98–M'05–SM'09–F'15) received the B.S. (Hons.), M.S.E., and Ph.D. degrees in electrical engineering from The University of Texas at Austin, in 2000, 2002, and 2004, respectively. Since 2004, he has been with the School of Electrical and Computer Engineering, Purdue University, where he is currently the Nick Trbovich Professor of electrical and computer engineering and leads the College of Engineering Preeminent Team on Efficient Spectrum Usage. His research interests include in the design and analysis of broadband wireless communication systems, 5G wireless systems, multiple-input multiple-output (MIMO) communications, millimeter wave wireless, software defined radios and wireless networks, coding theory, and MIMO array processing. He is a member of the Executive Committee for the National Spectrum Consortium. He currently serves as a Senior Editor for the *IEEE Signal Processing Magazine* and previously served as an Editor for the IEEE Transactions on Communications, an Associate Editor for the IEEE Transactions on Signal Processing, and a Guest Editor for special issues of the IEEE Journal on Selected Areas in Communications and the *EURASIP Journal on Wireless Communications and Networking*.