Hybrid NOMA/OMA with Buffer-Aided Relaying for Cooperative Uplink System

Jianping Quan[1], Peng Xu[1(✉)], Yunwu Wang[1], and Zheng Yang[2]

[1] Chongqing Key Laboratory of Mobile Communications Technology, School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China
quanjianp@163.com, xupeng@cqupt.edu.cn, wangyunw@163.com

[2] Fujian Provincial Engineering Technology Research Center of Photoelectric Sensing Application, Key Laboratory of OptoElectronic Science and Technology for Medicine of Ministry of Education, Fujian Normal University, Fuzhou 350007, China
zyfjnu@163.com

Abstract. In this paper, we consider a cooperative uplink network consisting of two users, a half-duplex decode-and-forward (DF) relay and a base station (BS). In the relaying network, the two users transmit packets to the buffer-aided relay using non-orthogonal multiple access (NOMA) or orthogonal multiple access (OMA) technology. We proposed a hybrid NOMA/OMA based mode selection (MS) scheme, which adaptively switches between the NOMA and OMA transmission modes according to the instantaneous strength of wireless links and the buffer state. Then, the state transmission matrix probabilities of the corresponding Markov chain is analyzed, and the performance in terms of sum throughput, outage probability, average packet delay and diversity gain are evaluated with closed form expressions. Numerical results are provided to demonstrate that hybrid NOMA/OMA achieves significant performance gains compared to conventional NOMA and OMA in most scenarios.

Keywords: Hybrid NOMA/OMA · Buffer-aided relaying · Cooperative uplink system

1 Introduction

Non-orthogonal multiple access (NOMA) technology, is recognized to be a promising mobile communication technology in future communication systems

The work of J. Quan and P. Xu was supported by the National Natural Science Foundation of China under Grant 61701066, in part by Chongqing Natural Science Foundation Project under Grant cstc2019jcyj-msxm1354, in part by Chongqing College Students' Innovative Entrepreneurial Training Plan Program under Grant S201910617032, and in part by Special Funded Undergraduate Research Training Program under Grant A2019-39. The work of Z. Yang was supported by National Natural Science Foundation of China under Grant 61701118.

X. Chu et al. (Eds.): QShine 2019, LNICST 300, pp. 169–183, 2020.
https://doi.org/10.1007/978-3-030-38819-5_11

since it can significantly enhance spectrum efficiency [2]. Different to orthogonal multiple access (OMA) technology which serves multiple users in orthogonal time/frequze/code domain, NOMA allows multiple users to share communication resources in the same time/frequze/code domain with different power levels, and achieves better performance than conventional OMA [4,13].

Cooperative communication enables efficient utilization of communication resources, which has many advantages in improving system efficiency and reliability, and expanding the coverage of wireless communication networks. The work in [3] first exploited cooperation between the users, i.e., the stronger users help the other weaker users by using the decode-and-forward (DF) scheme, such that the performance of NOMA can be enhanced and the optimal diversity gain can be achieved. On the other hand, the work in [6] investigated cooperative NOMA with a single relay, where multiple users were helped by a dedicated relay node using the DF and amplify-and-forward (AF) schemes. In addition, relay selection scheme has also been proposed for cooperative NOMA networks with multiple relays in some existing works (e.g., [15,16]).

For the Buffer-aided relaying system, the data buffer enables the relay to transmit when the source-to-relay link is in outage and to receive when the relay-to-destination link is in outage, which can provide additional freedom for wireless cooperative communication networks and overcome the bottleneck effect of conventional cooperative communication technologies [17,19]. Exiting works related to buffer-aided cooperative communication mainly considered the design of adaptive link or mode selection (MS) schemes for single-relay systems (e.g., [5,20]), and the design of relay selection (RS) schemes for multiple-relay systems [7,9,10,14]. In addition, buffer-aided cooperative NOMA for downlink transmission has also been investigated in recent existing works [8,11,17]. The works in [8] and [17] considered cooperative NOMA with a single buffer-aided relay. In [8], adaptive and fixed rates were assumed for the source-to-relay and relay-to-user transmissions, respectively, and the sum throughput was maximized based on the optimal MS scheme. In [17], fixed rate was assumed for both the source-to-relay and relay-to-user transmissions, and a relay decision scheme was proposed to enhance outage performance. Recently, the work in [11] proposed a hybrid buffer-aided NOMA/OMA RS scheme, which is shown to significantly outperform NOMA and OMA RS schemes, but its performance is hard to be analyzed with closed-form expressions.

In previous works [8,11,17], downlink buffer-aided relay systems has been widely considered. In this paper, we focus on an uplink buffer-aided relay system with two users, a DF relay, and a base station (BS). For the considered system, we propose a buffer-aided hybrid NOMA/OMA based MS scheme, which adaptively switches between the NOMA and OMA transmission modes according to the instantaneous channel state information (CSI) and buffer state. The basic idea of the proposed hybrid NOMA/OMA MS scheme is to give priority to the NOMA transmission mode, i.e., NOMA will be adopted to transmit the two users' messages simultaneously. However, NOMA transmission mode might not be successful, especially for weak channel conditions; in this case, the transmission mode will switch to OMA.

The Markov chain (MC) of the proposed hybrid NOMA/OMA based MS scheme is formulated, and the corresponding state transmission matrix probabilities are analyzed. Then, we evaluate the performance of the proposed scheme. In particular, closed-form expressions for sum throughput, outage probability, and average delay are obtained, and it is demonstrated that the proposed scheme can achieve a diversity gain of 2 as long as the buffer size is not smaller than 3. It is worth noting that performance analysis of the proposed hybrid NOMA/OMA scheme is non-trivial since NOMA and OMA have different requirements of the channel state and buffer state. Numerical results are provided to demonstrate that hybrid NOMA/OMA can significantly outperform conventional NOMA and OMA in most scenarios, especially for sum throughput and outage probability.

2 System Model and Preliminaries

2.1 System Model

Consider a buffer-aided uplink DF relaying system which consists of two users, a relay, and a BS, as shown in Fig. 1. We assume that the direct links between the two users and the are sufficiently weak to be ignored, since they are blocked due to long-distance path loss or obstacles [7,16,17]. It is assumed that the time duration is partitioned into slots with equal length and each transmitted packet spans one time slot. In each time slot, the users or the relay may be selected to transmit packets. When each user is selected, it assembles an information symbol intended for the BS into a packet with r_0 bits, where r_0 denotes the target transmission rate, i.e., the same target rate is assumed for each user to guarantee fairness [17]. The relay is equipped with two buffers, i.e., B_1 and B_2. Each buffer consists of $L \geq 2$ storage units and each storage unit can store a data packet received from any user. A storage unit at buffer B_u is used to store an information symbol transmitted by user u, $u = 1, 2$. If the relay is selected, it retrieves information symbols from the buffers and transmits them to the BS. Assume that each user always has information symbols to transmit. The channel gain from user u to the relay is denoted as h_u. The channel gain from the relay to the BS is denoted as h_R. These channels are assumed to be independent flat Rayleigh block fading channels which remain constant during one time slot and change randomly from one time slot to anther. Denote $H_u \triangleq |h_u|^2$ and $H_R \triangleq |h_R|^2$ for the sake of brevity, which follow exponential distributions and their expectations are denoted by

$$\mathbb{E}[H_1] = \frac{1}{\Omega_1},\ \mathbb{E}[H_2] = \frac{1}{\Omega_2} \text{ and } \mathbb{E}[H_R] = \frac{1}{\Omega_R}. \tag{1}$$

In addition, it is assumed that each transmitter is constrained by the maximum transmit power P, and each receiver has the same noise power σ^2.

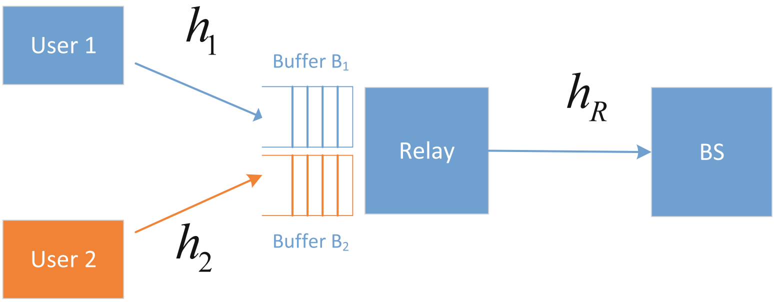

Fig. 1. System Model of an uplink buffer-aided relaying system.

Table 1. Necessary requirements for each transmission mode (T: Transmit; R: Receive; S: Silent)

Mode	User 1	User 2	Relay	BS	**CSI requirement**	**Buffer requirement**
$\mathcal{M}_1$	T	S	R	S	$\mathcal{R}_1 \triangleq \{H_1 \geq \epsilon_0\}$	$l_1 < L$
$\mathcal{M}_2$	S	T	R	S	$\mathcal{R}_2 \triangleq \{H_2 \geq \epsilon_0\}$	$l_2 < L$
$\mathcal{M}_3$	T	T	R	S	$\mathcal{R}_3 \triangleq \left\{ \begin{array}{l} H_{\pi_1} \geq \epsilon_0 2^{r_0} \\ H_{\pi_2} \geq \epsilon_0 \end{array} \right\}$	$\max\{l_1, l_2\} < L$
$\mathcal{M}_4$	S	S	T	R	$\mathcal{R}_4 \triangleq \{H_R \geq \epsilon_R\}$	$\min\{l_1, l_2\} > 0$
$\mathcal{M}_5$	S	S	S	S	$\forall H_1, H_2, H_R$	$\forall l_1, l_2$

2.2 Transmission Modes and CSI Requirements

For the proposed system, we consider five possible transmission modes, denoted by $\mathcal{M}_1, \cdots, \mathcal{M}_5$. Specifically, $\mathcal{M}_1$, $\mathcal{M}_2$, and $\mathcal{M}_3$ denote the user-to-relay modes, where the opportunistic hybrid NOMA/OMA is utilized: $\mathcal{M}_1$ and $\mathcal{M}_2$ utilize OMA for which only one of the users is selected to transmit a packet to the relay, and $\mathcal{M}_3$ utilizes NOMA for which the two users transmit packets to the relay simultaneously. $\mathcal{M}_4$ denotes the relay-to-BS mode, where the relay selects a packet from each user's buffer and blends the two packets into a mixed packet with $2r_0$ bits and then transmits it to the BS[1]; and $\mathcal{M}_5$ denotes the silent mode.

The instantaneous CSI requirement for each mode is summarized in Table 1, where the CSI region of mode $\mathcal{M}_k$ is defined as $\mathcal{R}_k$. Specifically, for the mode $\mathcal{M}_u$, $u = 1, 2$, it requires $H_u \geq \epsilon_0$ so that the relay can decode packets correctly, where $\epsilon_0 \triangleq \frac{2^{r_0}-1}{\rho}$ and $\rho \triangleq \frac{P}{\sigma^2}$ is the transmit signal-to-noise-ratio (SNR). For the mode $\mathcal{M}_4$, the transmission rate from the relay to the BS should be $2r_0$ bits per channel use (BPCU), so $H_R \geq \epsilon_R$ is required, where $\epsilon_R \triangleq \frac{2^{2r_0}-1}{\rho}$. For

[1] Note that, different to the user-to-relay transmission, we consider only one mode for the relay-to-BS transmission (i.e. $\mathcal{M}_4$), where the relay transmits both the two users' messages simultaneously, such that the two users' messages can reach the BS at the same time slot and short-term user fairness can be guaranteed [13].

the NOMA mode $\mathcal{M}_3$, the relay receives the following signal:

$$y_R = h_1\sqrt{P_1}x_1 + h_2\sqrt{P_2}x_2 + n_r, \tag{2}$$

where $P_u \leq P$ is the transmit power at user u, $u = 1, 2$, and n_r is the additive Gaussian noise at the relay with zero mean and variance σ^2. In addition, some existing modulation classification algorithms [18] can be more piratical in existing complex networks, which is out the scope of this paper.

The relay uses successive interference cancellation (SIC)[2] to decode the two users' messages. Specifically, assume that the users are sorted according to their channel qualities, i.e., $|h_{\pi_1}| \geq |h_{\pi_2}|$, where $\pi_1, \pi_2 \in \{1, 2\}$. The relay first decodes the message of the stronger user π_1 by treating the other user's signal as pure noise, which requires $\frac{H_{\pi_1}P_{\pi_1}}{\sigma^2+H_{\pi_2}P_{\pi_2}} \geq \rho\epsilon_0$; then, it cancels the signal of the stronger user π_1 from the observed signal, and decodes the message of the weaker user π_2, which requires $H_{\pi_2}P_{\pi_2} \geq P\epsilon_0$.

Remark 1. *For user π_1, we set $P_{\pi_1} \triangleq P$; for user π_2, we set the minimum required transmit power, i.e., $P_{\pi_2} \triangleq \frac{P\epsilon_0}{H_{\pi_2}}$ if $\frac{\epsilon_0}{H_{\pi_2}} \leq 1$, and $P_{\pi_2} \triangleq 0$, otherwise, in order to control the inter-user interference power when decoding user π_1's messages. Using this power setting, the required CSI region $\mathcal{R}_3$ can be easily obtained as shown in Table 1. Such a power setting requires the weaker user π_2 to know the perfect CSI of h_{π_2} at the beginning of each time slot. Hybrid NOMA with imperfect CSI would be an interesting future topic.*

2.3 The Buffer Requirements

Let l_u denote the number of packets in buffer B_u at the end of each time slot, $l_u \in \{0, 1, ...L\}$. The buffer requirement for each mode is also summarized in Table 1, where mode $\mathcal{M}_u$ requires that buffer B_u is not full, i.e., $l_u < L$, $u = 1, 2$; mode $\mathcal{M}_3$ requires that both the two buffers are not full, i.e., $\max\{l_1, l_2\} < L$; and mode $\mathcal{M}_4$ requires that both the two buffers are not empty, i.e., $\min\{l_1, l_2\} > 0$.

3 Hybrid NOMA/OMA Mode Selection

The design of optimal buffer-aided relaying schemes for delay-constrained networks is still a challenging issue, which has not been solved even for the single-user case [19]. Alternatively, a heuristic but efficient delay-constrained buffer-aided MS scheme will be proposed in this section.

The basic idea is to allocate each mode $\mathcal{M}_k$ a weight, denoted by W_k, to determine the priority of each mode, which is given in Table 2, where $0 < \delta < 1/2$ is used to differentiate two weights with the same integer part. In addition, α_1, α_2 and α_3 denote three layers of the corresponding transmission modes, where $\alpha_1 \ll \alpha_2 \ll \alpha_3$. In particular, when $\min\{l_1, l_2\} \geq 2$, $\mathcal{M}_4$ lies in layer α_3, which

[2] Compared to "joint decoding" [1], SIC enjoys much lower decoding complexity, and hence this paper adopts the SIC detection at the relay.

enjoys the highest priority. The motivation of the threshold of 2 is to achieve the tradeoff between outage probability minimization and average packet delay minimization [9,17]. In this case, each buffer will be prone to remain at the size of 1 or 2 especially at high SNRs, which means that each buffer is neither full nor empty in most time slots as long as $L \geq 3$. When $\min\{l_1, l_2\} = 1$, $\mathcal{M}_4$ falls down to layer α_2, the same layer with $\mathcal{M}_3$. Moreover, the OMA modes $\mathcal{M}_1$ and $\mathcal{M}_2$ lies in layer α_1, and the silent mode $\mathcal{M}_5$ will be selected only if the weight of any other mode is smaller than 2δ or its CSI requirement is not satisfied.

Table 2. Weight for Each Mode, where $0 < \delta < 1/2$, $0 \ll \alpha_1 \ll \alpha_2 \ll \alpha_3$

Mode	Weight for each mode, i.e., W_k
$\mathcal{M}_1$	$\alpha_1(L - l_1)$
$\mathcal{M}_2$	$\alpha_1(L - l_2) + \delta$
$\mathcal{M}_3$	$\alpha_2(L - \max\{l_1, l_2\}) + \delta$
$\mathcal{M}_4$	$\alpha_3\,(\min\{l_1, l_2\} - 1) + \alpha_2$
$\mathcal{M}_5$	2δ

With the allocated weights, the hybrid NOMA/OMA based MS scheme can be mathematically expressed as follows. In particular, mode $\mathcal{M}_{k^*}$ is selected in each time slot, where

$$k^* = \arg \max_{k \in \mathcal{R}_k, W_k \geq 2\delta} W_k. \tag{3}$$

Remark 2. *Hybrid NOMA/OMA will reduce to NOMA, if we disable the transmission mode $\mathcal{M}_1$ and $\mathcal{M}_2$ by setting $W_1 = W_2 = 0$, i.e., only modes $\mathcal{M}_3$ and $\mathcal{M}_4$ are used to receive and transmit messages at the relay, respectively. In addition, if we disable mode $\mathcal{M}_3$ by setting $W_3 = 0$, hybrid NOMA/OMA will reduce to OMA where the users transmit messages to the relay in orthogonal time slots.*

4 Performance Analysis

In this section, the performance of the proposed hybrid NOMA/OMA based MS scheme will be analyzed by formulating a MC as well as its transition matrix to model the evolution of the relay buffers.

4.1 State Transmission Matrix

Let $s_n = (l_1, l_2)$, $n \in \{1, 2, ..., (L+1)^2\}$, denote the states in the MC, which describes the queues of the two buffers at the relay. Let $\mathbf{A}$ denote the $(L+1) \times (L+1)$ state transition matrix, whose entry $\mathbf{A}_{i,j} = p(s_j \rightarrow s_i) = \mathbb{P}\{(l_1(t+1), l_2(t+1)) = s_i | (l_1(t), l_2(t)) = s_j\}$ is the transition probability to move from state s_j at time t to s_i at time $t+1$. The transition probabilities for the proposed scheme can be summarized in the following proposition.

Proposition 1. *The transition probabilities of the states of the MC for the proposed hybrid NOMA/OMA based MS scheme are given in (5)–(9), shown in the next page, where $P_{(l_1,l_2)}^{(l_1,l_2+1)}$ is identical to $P_{(l_1,l_2)}^{(l_1+1,l_2)}$, after switching "1" and "2" in (7). Note that, for the sake of brevity, $\phi(\Omega_1,\Omega_2)$ in (7)–(9) is defined as follows:*

$$\phi(\Omega_1,\Omega_2) \triangleq e^{-\epsilon_0(\Omega_1 2^{r_0}+\Omega_2)} + e^{-\epsilon_0(\Omega_2 2^{r_0}+\Omega_1)} - e^{-(\Omega_1+\Omega_2)\epsilon_0 2^{r_0}}. \tag{4}$$

Proof. Please refer to Appendix A.

$$P_{(l_1,l_2)}^{(l_1',l_2')} = 0,$$
$$\text{if } |l_1'-l_1| \geq 2 \vee |l_2'-l_2| \geq 2 \vee \{l_1' = l_1-1 \wedge l_2' \neq l_2-1\} \vee \{l_1' \neq l_1-1 \wedge l_2' = l_2-1\}. \tag{5}$$

$$P_{(l_1,l_2)}^{(l_1,l_2)} = \begin{cases} (1-e^{-\Omega_1\epsilon_0})(1-e^{-\Omega_2\epsilon_0})(1-e^{-\Omega_R\epsilon_R}) & \text{if } \max\{l_1,l_2\} < L \wedge \min\{l_1,l_2\} > 0, \\ (1-e^{-\Omega_1\epsilon_0})(1-e^{-\Omega_2\epsilon_0}) & \text{if } \max\{l_1,l_2\} < L \wedge \min\{l_1,l_2\} = 0, \\ (1-e^{-\Omega_2\epsilon_0})(1-e^{-\Omega_R\epsilon_R}) & \text{if } l_1 = L \wedge 0 < l_2 < L, \\ (1-e^{-\Omega_1\epsilon_0})(1-e^{-\Omega_R\epsilon_R}) & \text{if } l_2 = L \wedge 0 < l_1 < L, \\ 1-e^{-\Omega_1\epsilon_0} & \text{if } l_1 = 0 \wedge l_2 = L, \\ 1-e^{-\Omega_2\epsilon_0} & \text{if } l_2 = 0 \wedge l_1 = L, \\ 1-e^{-\Omega_R\epsilon_R} & \text{if } l_1 = l_2 = L, \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

$$P_{(l_1,l_2)}^{(l_1+1,l_2)} = \begin{cases} e^{-\Omega_1\epsilon_0}(1-e^{-\Omega_2\epsilon_0})(1-e^{-\Omega_R\epsilon_R}) & \text{if } 0 < l_2 \leq l_1 < L, \\ \left[e^{-\Omega_1\epsilon_0} - \phi(\Omega_1,\Omega_2)\right](1-e^{-\Omega_R\epsilon_R}) & \text{if } 0 < l_1 < l_2 < L, \\ e^{-\Omega_1\epsilon_0}(1-e^{-\Omega_R\epsilon_R}) & \text{if } 0 < l_1 < l_2 = L, \\ e^{-\Omega_1\epsilon_0}(1-e^{-\Omega_2\epsilon_0}) & \text{if } 0 = l_2 \leq l_1 < L, \\ e^{-\Omega_1\epsilon_0} - \phi(\Omega_1,\Omega_2) & \text{if } 0 = l_1 < l_2 < L, \\ e^{-\Omega_1\epsilon_0} & \text{if } l_1 = 0, l_2 = L, \\ 0 & \text{otherwise.} \end{cases} \tag{7}$$

$$P_{(l_1,l_2)}^{(l_1+1,l_2+1)} = \begin{cases} (1-e^{-\Omega_R\epsilon_R})\phi(\Omega_1,\Omega_2) & \text{if } \max\{l_1,l_2\} < L \wedge \min\{l_1,l_2\} \geq 2, \\ \phi(\Omega_1,\Omega_2) & \text{if } \max\{l_1,l_2\} < L \wedge \min\{l_1,l_2\} < 2, \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

$$P_{(l_1,l_2)}^{(l_1-1,l_2-1)} = \begin{cases} e^{-\Omega_R\epsilon_R}(1-\phi(\Omega_1,\Omega_2)) & \text{if } \max\{l_1,l_2\} < L \wedge \min\{l_1,l_2\} = 1, \\ e^{-\Omega_R\epsilon_R} & \text{if } \min\{l_1,l_2\} \geq 2 \vee \{\min\{l_1,l_2\} = 1 \wedge \max\{l_1,l_2\} = L\}, \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

4.2 Performance of the Proposed Scheme

One can verify that the transition matrix $\mathbf{A}$ is column stochastic and irreducible[3], so the stationary state probability vector can be obtained as follows [7]:

$$\boldsymbol{\pi} = (\mathbf{A} - \mathbf{I} + \mathbf{B})^{-1}\mathbf{b}, \tag{10}$$

[3] Column stochastic means that all entries in any column sum up to one; irreducible means that it is possible to move from any state to any state [12].

where $\boldsymbol{\pi} = [\pi_{s_1}, \cdots, \pi_{s_{(L+1)^2}}]^T$, $\mathbf{b} = [1, 1, \cdots, 1]^T$ and $\mathbf{B}_{i,j} = 1$, $\forall i, j$. In the next, we will use $\pi_{(l_1,l_2)}$ to denote the stationary state probability of the buffer state $s_n = (l_1, l_2)$ for simplicity.

In the following, the performance of the sum throughput, the outage probability, and the average delay will be analyzed.

Throughput. Over a long time of period, obviously the sum receive and the transmit throughputs at the relay will be the same, and the sum throughput of the system can be expressed as follows:

$$\begin{aligned} \bar{R}_{\text{sum}} &= r_0 \sum_{l_1=0}^{L-1} \sum_{l_2=0}^{L-1} (P_{(l_1,l_2)}^{(l_1+1,l_2)} + P_{(l_1,l_2)}^{(l_1,l_2+1)} + 2P_{(l_1,l_2)}^{(l_1+1,l_2+1)}) \pi_{(l_1,l_2)} \\ &= 2r_0 \sum_{l_1=1}^{L} \sum_{l_2=1}^{L} P_{(l_1,l_2)}^{(l_1-1,l_2-1)} \pi_{(l_1,l_2)}. \end{aligned} \tag{11}$$

Outage Probability. Since the target sum rate is r_0 BPCU ($r_0/2$ BPCU for each user), the outage probability of the system can be expressed as follows:

$$P_{\text{sys}}^{\text{out}} = 1 - \bar{R}_{\text{sum}}/r_0. \tag{12}$$

Proposition 2. *The diversity gain of* 2 *can be achieved by the proposed hybrid NOMA/OMA MS scheme, i.e.,* $-\lim_{\text{SNR}\to\infty} \frac{\log P_{\text{sys}}^{\text{out}}}{\log \text{SNR}} = 2$, *as long as* $L \geq 3$.

Proof. Please refer to Appendix B.

Average Delay. Denote P_k as the probability that mode $\mathcal{M}_k$ is selected. Over a long period of time, based on (11), we obtain

$$P_1 + P_2 + 2P_3 = 2P_4 = 1 - P_{\text{sys}}^{\text{out}}. \tag{13}$$

Moreover, denote η_U and η_R as the transmit sum throughputs (in number of packets) of the users and the relay, respectively, which can be expressed as

$$\eta_U = \eta_R = \bar{R}_{\text{sum}}/r_0 = 1 - P_{\text{sys}}^{\text{out}}. \tag{14}$$

Since in each time slot, at most two packets are transmitted from the two users, the average sum queuing length (in number of time slots) at two users can be obtained as

$$Q_U = 2 - (P_1 + P_2 + 2P_3) = 1 + P_{\text{sys}}^{\text{out}}. \tag{15}$$

Thus, the average delay at the two users is

$$D_U = \frac{Q_U}{\eta_U} = \frac{1 + P_{\text{sys}}^{\text{out}}}{1 - P_{\text{sys}}^{\text{out}}}. \tag{16}$$

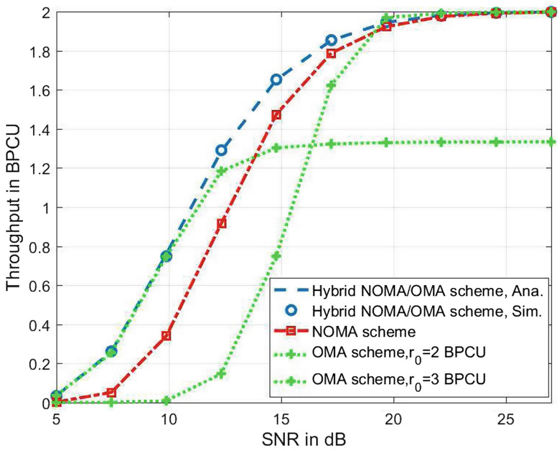

Fig. 2. Sum throughput vs. transmit SNR, where $L = 5$.

In addition, the average delay at the relay is $D_R = \bar{Q}_R/\eta_R$, where $\bar{Q}_R$ is the average sum queuing length of the two buffers, which can be expressed as

$$\bar{Q}_R = \sum_{l_1=0}^{L}\sum_{l_2=0}^{L}(l_1 + l_2)\pi_{(l_1,l_2)}. \tag{17}$$

In summary, the total average packet delay of the system is $D_U + D_R$.

5 Numerical Results

In this section, we evaluate the performance of the proposed hybrid NOMA/OMA based MS scheme by using computer simulations, in terms of sum throughput, outage probability and average delay. NOMA and OMA mentioned in Remark 2 are taken as the comparative ones. Each channel is modeled as $h_i = d_i^{-\beta/2} g_i$, where the small scale fading gain is Rayleigh distributed, i.e., $g_i \sim \mathcal{CN}(0, 1)$, $i \in \{1, 2, R\}$. Furthermore, asymmetric distances are considered, which are set as $d_1 = 1$, $d_2 = 2$ and $d_R = 1$, and the path loss exponent is chosen as 2 to reflect a favorable propagation condition. This means that $\Omega_1 = 1$, $\Omega_2 = 4$, and $\Omega_R = 1$. In addition, the target rate is set as $r_0 = 2$ BPCU unless stated otherwise. In Fig. 2, sum throughput comparison is presented for hybrid NOMA/OMA, NOMA and OMA schemes, where the buffer size is set as $L = 5$. One can observe that, when $r_0 = 2$, hybrid NOMA/OMA and NOMA achieve the maximum sum throughput of 2 BPCU at high SNRs, whereas OMA can only achieve about 1.3 BPCU in this case. This is because only one packet can be transmitted from the users to the relay in one time slot for OMA. If we set $r_0 = 3$, OMA can achieve the sum throughput of 2 BPCU at high SNRs, but has a very poor performance at low or moderate SNRs. In addition, one can also

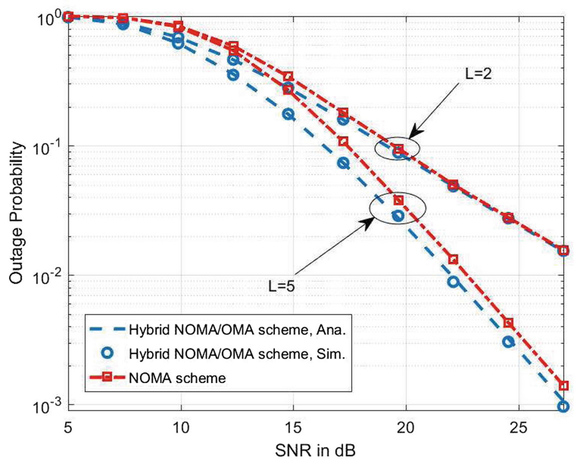

Fig. 3. System outage probability vs. transmit SNR, where $r_0 = 2$ BPCU.

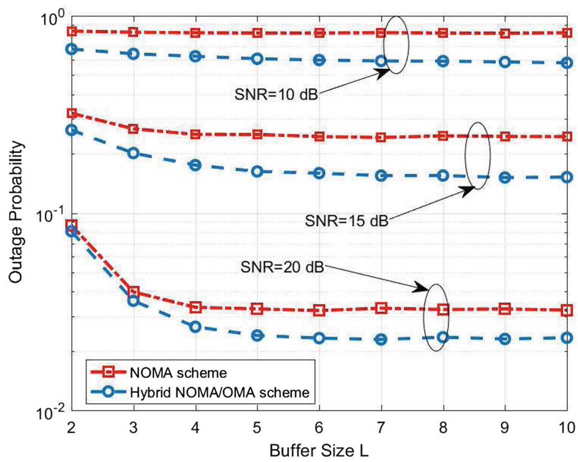

Fig. 4. System outage probability vs. buffer size L, where $r_0 = 2$ BPCU.

observe that hybrid NOMA/OMA outperforms NOMA significantly especially at low or moderate SNRs. For example, when SNR=10 dB, hybrid NOMA/OMA and NOMA achieve the sum throughputs of 0.75 and 0.35 BPCU, respectively, i.e., there is a improvement of more than 100%.

The outage probability performance of hybrid NOMA/OMA and NOMA schemes are presented in Figs. 3 and 4 versus transmit SNR and buffer size L, respectively. In Fig. 3, one can observe that the gap between the two schemes is slight when $L = 2$, especially at high SNRs, but significant performance gap exists when $L = 5$. In Fig. 4, one can observe that hybrid NOMA/OMA achieves lower outage probability compared to NOMA for different buffer sizes and SNRs.

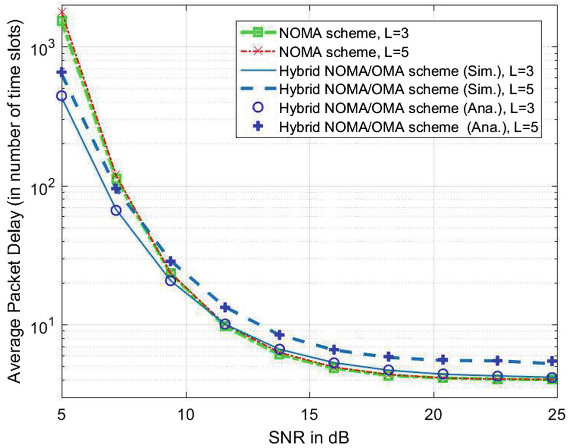

Fig. 5. Average packet delay vs. transmit SNR, where $r_0 = 2$ BPCU.

In particular, hybrid NOMA/OMA can benefit from enlarging L significantly, whereas the outage probability of NOMA almost does not decrease when $L \geq 5$.

In Fig. 5, we present average packet delay comparison of hybrid NOMA/OMA and NOMA schemes when $L = 3$ and $L = 5$. One can observe that, at low SNRs, hybrid NOMA/OMA achieves a much shorter average delay. This is because the average delay at the users is the dominant factor when the outage probability is high at low SNRs (shown in Sect. 4.2). At high SNRs, hybrid NOMA/OMA suffers from a longer average delay especially when $L = 5$. This is because the average delay at the relay is the dominant factor at high SNRs. For hybrid NOMA/OMA, a single packet is transmitted when an OMA transmission mode ($\mathcal{M}_1$ or $\mathcal{M}_2$) is selected, which may obstruct the following received packets in the same buffer. However, it can be seen that the average delay of hybrid NOMA/OMA is just slightly longer than NOMA at high SNRs.

6 Conclusion

This paper has investigated a cooperative uplink system with two users, a buffer-aided relay, and a BS. A hybrid NOMA/OMA based MS scheme has been proposed, which combines NOMA and OMA, and all possible transmission modes are allocated layered weights according to the buffer states in order to determine their priorities. Then, we have also analyzed the state transmission matrix probabilities of the corresponding MC, and derived closed form expressions for sum throughput, outage probability, and average delay. A diversity gain of 2 can be achieved when the buffer size is not smaller than 3. Numerical results have shown that hybrid NOMA/OMA significantly outperforms conventional NOMA and OMA in most scenarios.

Appendix A

Proof of Proposition 1

To prove this proposition, we first analyze the probability of the required CSI region for each mode $\mathcal{M}_k$ (shown in Table 1), denoted by $P_{\mathcal{R}_k}$, $k = [1:4]$, which is given as follows:

$$P_{\mathcal{R}_1} = e^{-\Omega_1 \epsilon_0},\ P_{\mathcal{R}_2} = e^{-\Omega_2 \epsilon_0} \tag{18}$$

$$P_{\mathcal{R}_3} = \phi(\Omega_1, \Omega_2),\ P_{\mathcal{R}_4} = e^{-\Omega_R \epsilon_R}. \tag{19}$$

We then consider the following cases:

1. Since each buffer at most receives or transmits only one packet in one time slot, $P_{(l_1,l_2)}^{(l_1',l_2')} = 0$ if $|l_u' - l_u| \geq 2$, $u = 1, 2$. Moreover, the two buffers transmit at the same time slot in the proposed scheme, and hence (5) can be easily obtained.
2. $P_{(l_1,l_2)}^{(l_1,l_2)}$ corresponds to the case that m ode $\mathcal{M}_5$ is selected. Since weight W_5 has the smallest value when $\max\{l_1, l_2\} < L \wedge \min\{l_1, l_2\} > 0$ compared to the other modes' weights, mode $\mathcal{M}_5$ can only be selected if all channels are so weak that the other modes' CSI requirements (shown in Table 1) cannot be satisfied. In this subcase, $P_{(l_1,l_2)}^{(l_1,l_2)} = (1 - P_{\mathcal{R}_1})(1 - P_{\mathcal{R}_2})(1 - P_{\mathcal{R}_4})$. The values of $P_{(l_1,l_2)}^{(l_1,l_2)}$ in the other subcases can be obtained similarly shown in (6).
3. $P_{(l_1,l_2)}^{(l_1+1,l_2)}$ corresponds to the case that mode $\mathcal{M}_1$ is selected. Take the subcase $0 < l_2 \leq l_1 < L$ for example. In this subcase, $W_5 < W_1 < \min\{W_2, W_3, W_4\}$, so mode $\mathcal{M}_1$ can be selected only if the CSI requirement of $\mathcal{M}_1$ can be satisfied but the CSI requirement of $\mathcal{M}_i$ cannot be satisfied, $i = 2, 3, 4$, and thus $P_{(l_1,l_2)}^{(l_1+1,l_2)} = P_{\mathcal{R}_1}(1 - P_{\mathcal{R}_2})(1 - P_{\mathcal{R}_4})$. $P_{(l_1,l_2)}^{(l_1+1,l_2)}$ can be calculated for the other subcases shown in (7).
4. $P_{(l_1,l_2)}^{(l_1+1,l_2+1)}$ corresponds to the case that mode $\mathcal{M}_3$ is selected. If $\max\{l_1, l_2\} < L \wedge \min\{l_1, l_2\} = 2$, $W_3 > W_i$, $i = 1, 2, 5$, and $W_3 < W_4$, so mode $\mathcal{M}_3$ can be selected only if the CSI requirement of $\mathcal{M}_3$ can be satisfied but the CSI requirement of $\mathcal{M}_4$ cannot be satisfied, and thus $P_{(l_1,l_2)}^{(l_1+1,l_2+1)} = P_{\mathcal{R}_3}(1 - P_{\mathcal{R}_4})$. If $\max\{l_1, l_2\} < L \wedge \min\{l_1, l_2\} < 2$, W_3 has the largest value, and hence $P_{(l_1,l_2)}^{(l_1+1,l_2+1)} = P_{\mathcal{R}_3}$.
5. $P_{(l_1,l_2)}^{(l_1-1,l_2-1)}$ corresponds to the case that mode $\mathcal{M}_4$ is selected, and (9) can be easily obtained, following similar derivation steps for the previous case.

Appendix B

Proof of Proposition 2

The transition matrix $\mathbf{A}$ is too complicated (shown in Proposition 1) to obtain an explicit approximation of the outage probability $P_{\text{sys}}^{\text{out}}$ in (12) at high SNRs.

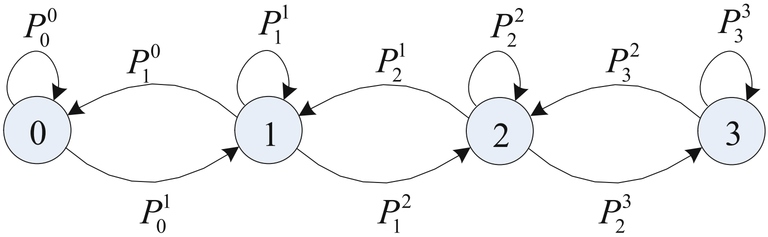

Fig. 6. Diagram of the MC of the simplified NOMA scheme with $L = 3$.

Alternatively, we wish to derive an upper bound on $P_{\mathrm{sys}}^{\mathrm{out}}$ in order to obtain an achievable diversity gain of the proposed scheme. In particular, it should be noted that the throughput achieved by NOMA (mentioned in Remark 2) is just a lower bound of hybrid NOMA/OMA. This is because the relay can still receive messages by using modes $\mathcal{M}_1$ and $\mathcal{M}_2$ for hybrid NOMA/OMA, even if the CSI requirements of $\mathcal{M}_3$ and $\mathcal{M}_4$ cannot be satisfied. Thus, the outage probability of NOMA, denoted by $P_{\mathrm{NOMA}}^{\mathrm{out}}$, is an upper bound of $P_{\mathrm{sys}}^{\mathrm{out}}$.

Using NOMA, there exists only three modes ($\mathcal{M}_k$, $k = 3, 4, 5$) and $(L + 1)$ states since the two buffers have the same size in each time slot. The MC of the simplified NOMA scheme for the case $L = 3$ is presented in Fig. 6, where each transition probability from state i to state j, denoted by P_i^j, can be easily approximated as

$$P_0^0 \approx \epsilon_0(\Omega_1 + \Omega_2), \; P_1^1 \approx \epsilon_0\epsilon_R\Omega_R(\Omega_1 + \Omega_2), \tag{20}$$

$$P_3^3 \approx \epsilon_R\Omega_R, \; P_0^1 \approx 1 - \epsilon_0(\Omega_1 + \Omega_2), \tag{21}$$

$$P_2^3 \approx \epsilon_R\Omega_R, \; P_3^2 \approx 1 - \epsilon_R\Omega_R, \tag{22}$$

$$P_1^0 \approx \epsilon_0(\Omega_1 + \Omega_2), \tag{23}$$

at high SNRs. Based on the above transition probabilities, the stationary state probabilities of the MC can be obtained, which are approximately given by

$$\pi_0^{\mathrm{NOMA}} \approx \frac{1}{2}\epsilon_0(\Omega_1 + \Omega_2), \pi_1^{\mathrm{NOMA}} \approx \frac{1}{2}, \tag{24}$$

$$\pi_2^{\mathrm{NOMA}} \approx \frac{1}{2}, \pi_3^{\mathrm{NOMA}} \approx \frac{1}{2}\epsilon_R\Omega_R, \tag{25}$$

at high SNRs. Thus, the outage probability of NOMA can be obtained as follows:

$$P_{\mathrm{sys}}^{\mathrm{NOMA}} \approx \frac{1}{2}[\epsilon_0(\Omega_1 + \Omega_2) + \epsilon_R\Omega_R]^2. \tag{26}$$

Furthermore, it is easy to prove that the diversity gain regarding to $P_{\mathrm{sys}}^{\mathrm{NOMA}}$ is 2. On the other hand, increasing L obviously benefits to decrease the outage probability, and hence hybrid NOMA/OMA achieves the diversity gain of 2 as long as $L \geq 3$.

References

1. Cover, T.M., Thomas, J.A.: Elements of Information Theory, 2nd edn. Wiley-Interscience, New York (2006)
2. Ding, Z., Lei, X., Karagiannidis, G.K., Schober, R., Yuan, J., Bhargava, V.K.: A survey on non-orthogonal multiple access for 5G networks: research challenges and future trends. IEEE J. Sel. Areas Commun. **35**(10), 2181–2195 (2017)
3. Ding, Z., Peng, M., Poor, H.: Cooperative non-orthogonal multiple access in 5G systems. IEEE Commun. Lett. **19**(8), 1462–1465 (2015)
4. Ding, Z., Yang, Z., Fan, P., Poor, H.V.: On the performance of non-orthogonal multiple access in 5G systems with randomly deployed users. IEEE Signal Process. Lett. **21**(12), 1501–1505 (2014)
5. Jamali, V., Zlatanov, N., Schober, R.: Bidirectional buffer-aided relay networks with fixed rate transmission-part II: delay-constrained case. IEEE Trans. Wireless Commun. **14**(3), 1339–1355 (2015)
6. Kim, J.B., Lee, I.H.: Non-orthogonal multiple access in coordinated direct and relay transmission. IEEE Commun. Lett. **19**(11), 2037–2040 (2015)
7. Krikidis, I., Charalambous, T., Thompson, J.S.: Buffer-aided relay selection for cooperative diversity systems without delay constraints. IEEE Trans. Wireless Commun. **11**(5), 1957–1967 (2012)
8. Luo, S., Teh, K.C.: Adaptive transmission for cooperative NOMA system with buffer-aided relaying. IEEE Commun. Lett. **21**(4), 937–940 (2017)
9. Luo, S., Teh, K.C.: Buffer state based relay selection for buffer-aided cooperative relaying systems. IEEE Trans. Wireless Commun. **14**(10), 5430–5439 (2015)
10. Nomikos, N., Charalambous, T., Krikidis, I., Skoutas, D.N., Vouyioukas, D., Johansson, M.: A buffer-aided successive opportunistic relay selection scheme with power adaptation and inter-relay interference cancellation for cooperative diversity systems. IEEE Trans. Commun. **63**(5), 1623–1634 (2015)
11. Nomikos, N., Charalambous, T., Vouyioukas, D., Karagiannidis, G.K., Wichman, R.: Hybrid NOMA/OMA with buffer-aided relay selection in cooperative networks. IEEE J. Sel. Top. Signal Process. **13**(3), 524–537 (2019)
12. Norris, J.R.: Markov Chains. Cambridge University Press, Cambridge (1997). https://doi.org/10.1017/CBO9780511810633
13. Xu, P., Cumanan, K.: Optimal power allocation scheme for non-orthogonal multiple access with alpha-fairness. IEEE J. Sel. Areas Commun. **35**(10), 2357–2369 (2017)
14. Xu, P., Ding, Z., Krikidis, I., Dai, X.: Achieving optimal diversity gain in buffer-aided relay networks with small buffer size. IEEE Trans. Veh. Technol. **65**(10), 8788–8794 (2016)
15. Xu, P., Yang, Z., Ding, Z., Zhang, Z.: Optimal relay selection schemes for cooperative NOMA. IEEE Trans. Veh. Technol. **67**(8), 7851–7855 (2018)
16. Yang, Z., Ding, Z., Wu, Y., Fan, P.: Novel relay selection strategies for cooperative NOMA. IEEE Trans. Veh. Technol. **66**(11), 10114–10123 (2017)
17. Zhang, Q., Liang, Z., Li, Q., Qin, J.: Buffer-aided non-orthogonal multiple access relaying systems in Rayleigh fading channels. IEEE Trans. Commun. **65**(1), 95–106 (2017)
18. Zhang, Z., Wang, C., Gan, C., Sun, S., Wang, M.: Automatic modulation classification using convolutional neural network with features fusion of SPWVD and BJD. IEEE Trans. Signal Inf. Process. Netw. **5**(3), 469–478 (2019)

19. Zlatanov, N., Ikhlef, A., Islam, T., Schober, R.: Buffer-aided cooperative communications: opportunities and challenges. IEEE Commun. Mag. **52**(4), 146–153 (2014)
20. Zlatanov, N., Schober, R., Popovski, P.: Buffer-aided relaying with adaptive link selection. IEEE J. Sel. Areas Commun. **31**(8), 1530–1542 (2013)