Manuscript version: Author’s Accepted Manuscript
The version presented in WRAP is the author’s accepted manuscript and may differ from the published version or Version of Record.

Persistent WRAP URL:
http://wrap.warwick.ac.uk/134773

How to cite:
Please refer to published version for the most recent bibliographic citation information. If a published version is known of, the repository item page linked to above, will contain details on accessing it.

Copyright and reuse:
The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Publisher’s statement:
Please refer to the repository item page, publisher’s statement section, for further information.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk.
Abstract—A spatial modulation aided non-orthogonal multiple access with full-duplex relay (SM-NOMA-FDR) scheme is proposed for the coordinated direct and relay transmission in this paper. Specifically, the signal of the near user is mapped to an $M$-ary modulated symbol and the signal of the far user is mapped to an SM symbol. The base station first transmits signals to the near user and relay via SM-NOMA, and then the relay decodes and retransmits the signal of the far user. An SM-assisted FDR is used in this scheme to improve the spectral efficiency while reducing energy consumption and making full use of the antenna resources at the relay, since SM only activates one antenna in each transmission. We derive the ergodic capacity and bit error rate of the proposed scheme over independent Rayleigh fading channels. Numerical results validate the accuracy of the theoretical analysis and show the superior performance of the proposed SM-NOMA-FDR scheme.

Index Terms—Full-duplex relay, non-orthogonal multiple access, performance analysis, spatial modulation.

I. INTRODUCTION

Non-orthogonal multiple access (NOMA) is an emerging technique to achieve high spectral efficiency (SE) [1], which allows multiple users to share radio resources using same time, frequency and code. Specifically, the base station (BS) mixes the signals from multiple users with different power levels, and the users detect their signals via successive interference cancellation (SIC) [2]. Recently, relaying has been used to enhance the reliability of communications [3]–[5]. A cooperative NOMA scheme was proposed in [3], where the user with strong channel condition serves as a relay to improve the reliability of the poor-channel user. Chen et al. presented a secondary NOMA relay in [4], where users communicate with the BS via a NOMA relay, and a selected secondary user acts as a relay for the primary user.

Spatial modulation (SM) has been regarded as a promising multiple-antenna technique [6]. In SM, only a single antenna is activated in each time slot, which reduces system complexity and energy consumption [7]. Consequently, SM is inherently suitable for full-duplex relay (FDR) systems, where one antenna of the relay is activated to send signals, while the remaining antennas can be used to receive the signal from the BS [8]. On the other hand, the combination of NOMA and SM has also attracted growing attention. In particular, Wang et al. proposed an SM aided NOMA (SM-NOMA) system in [9], where the BS transmits the superimposed SM symbols of users. In [10], a two-user SM-NOMA scheme was introduced, where the signal of one user is conveyed in the spatial domain and the signal of the other user is conveyed in the signal domain, and thus it completely cancels the intra-cluster interference. In [11], SM-NOMA was invoked for the vehicle-to-vehicle (V2V) system, where Rician channels are considered for V2V links. Li et al. proposed a three-node cooperative system using SM-NOMA in [12], where a BS serves two users, and the near user acts as a relay for the far user.

Coordinated direct and relay transmission (CDRT) is an effective strategy to extend the cell coverage and achieve high SE in 5G networks. However, the main issue of realizing CDRT is that the side information is required for interference cancellation. To address this problem, the authors in [13] introduced NOMA into CDRT with the half-duplex (HD) relay. In order to further improve the SE of [13], an FDR cooperative NOMA scheme with imperfect self-interference (SI) removal was developed for CDRT in [14]. However, for the concept of green communications, low energy consumption is also an important goal except the high SE. Motivated by this, in this paper, we focus on introducing SM into NOMA-based CDRT to improve the SE without increasing energy consumption. The main contributions of this paper are summarized as follows:

- We propose a two user SM-NOMA with FDR (SM-NOMA-FDR) scheme for CDRT, where the BS directly communicates with the near user and communicates with the far user through an FDR. Both the BS and the FDR adopt the SM technique.
- Taking into account the impact of residual SI at FDR, the performance of SM-NOMA-FDR in terms of ergodic capacity and bit error rate (BER) for two users are derived with closed-form expressions.
- Numerical results validate the correctness of the theoretical analysis. Meanwhile, we compare the proposed SM-NOMA-FDR with the the state of the arts and demonstrate superior performance of the proposed scheme.

Copyright (c) 2015 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org. The work of N. Zhao was supported by the National Natural Science Foundation of China (NSFC) under Grant 61871065. (Corresponding author: Minglu Jin.)

Q. Si, M. Jin, and N. Zhao are with the Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian 116024, China (e-mail: sqty@mail.dlut.edu.cn; mljin@dlut.edu.cn; zhaonan@dlut.edu.cn).

Y. Chen is with the School of Engineering, University of Warwick, Coventry CV4 7AL, U.K. (e-mail: Yunfei.Chen@warwick.ac.uk).

X. Wang is with the Department of Electrical and Computer Engineering, University of Western Ontario, London, ON N6A 5B9, Canada (e-mail: xianbin.wang@uwo.ca).
antennas of the relay can be given as thus $U_1$ can perfectly remove the interference from the relay. and

ture amplitude modulation (QAM) symbol of $U_1$ ($U_2$) with

the column of $\mathbf{F}$ from the BS simultaneously.

inactive antennas of FDR can be used for receiving signals from the BS simultaneously.

Hereafter, subscripts $B$, 1, 2 and $R$ denote the BS, $U_1$, $U_2$, and the relay, respectively. Then, the channels from the BS to $U_1$, from the BS to the relay and from the relay to $U_2$ can be denoted by $\mathbf{H} \in \mathbb{C}^{M \times N}$, $\mathbf{G} \in \mathbb{C}^{N \times N}$ and $\mathbf{F} \in \mathbb{C}^{M \times N}$, where the entries of these channels are complex Gaussian random variables with zero mean and variance $\lambda_h$, $\lambda_g$ and $\lambda_f$, respectively. We assume that the distance between the relay and the BS is larger than the distance between the BS and $U_1$, hence we have $\lambda_h > \lambda_g$. In the proposed scheme, the information of $U_1$ and $U_2$ are mapped to a conventional modulated symbol and an SM symbol, respectively. According to SM, the first $\log_2 N$ bits of $U_2$ are used for selecting one active antenna and the remaining bits of $U_2$ are mapped to an $M_2$-ary modulated symbol.

In the first time slot, the BS activates the $k$-th antenna to transmit the signal $\sqrt{a_1 \gamma_m} + \sqrt{a_2 x_n}$ based on NOMA, where $a_1$ and $a_2$ are power allocation coefficients satisfying $a_1 + a_2 = 1$, and $\gamma_m (x_n)$ is the $M_1$-ary ($M_2$-ary) quadrature amplitude modulation (QAM) symbol of U1 (U2) with $E[|\gamma_m|^2] = 1 \,(E[|x_n|^2] = 1)$. Thus, we have $b_1 = \log_2 M_1$ and $b_2 = \log_2 N + \log_2 M_2$. SIC is implemented at $U_1$, and thus $U_1$ can perfectly remove the interference from the relay. Accordingly, the received signals at $U_1$ and $N - 1$ inactive antennas of the relay can be given as

$$y_1 = \sqrt{P_B} h_k (\sqrt{a_1 \gamma_m} + \sqrt{a_2 x_n}) + n_1, \quad (1)$$

$$y_R = \sqrt{P_B} g_{pk} (\sqrt{a_1 \gamma_m} + \sqrt{a_2 x_n}) + \sqrt{I} x_R + n_R, \quad (2)$$

where: i) $P_B$ is the transmit power of the BS; ii) $h_k$ is the $k$-th column of $\mathbf{H}$; iii) $y_R = [y_1, \ldots, y_{p-1}, \ldots, y_N]^T$; iv) $g_{pk}$ denotes the channel vector between the $k$-th transmit antenna at the BS and $\{1, 2, \ldots, N\}\setminus p$ receive antennas at the relay; v) $x_R$ denotes a vector with entries of the relay’s signal $x_R$ and $I$ is the power of SI; vi) $n_1$ and $n_R$ are additive white Gaussian noise (AWGN) vectors with elements of zero mean and variance $\sigma_I^2$ and $\sigma_R^2$, respectively.

The signal of $U_2$ is first detected at $U_1$ by using the maximum likelihood (ML) method, which can be given by

$$\hat{k}, x_n = \arg \min_{k, n} \| y_1 - \sqrt{P_B} h_k \sqrt{a_2 y_n} \|^2. \quad (3)$$

After removing the interference of $U_2$, $U_1$ performs another ML detection to detect its own signal, which is expressed as

$$\gamma_m = \arg \min_m \| \tilde{y}_1 - \sqrt{P_B} h_k \sqrt{a_1 \gamma_m} \|^2, \quad (4)$$

where $\tilde{y}_1 = y_1 - \sqrt{P_B} h_k \sqrt{a_2 x_n}$.

At the relay side, it demodulates the data of $U_2$ by treating the signal of $U_1$ as noise, which can be given as

$$\hat{k}, x_n = \arg \min_{k, n} \| y_R - \sqrt{P_B} g_{pk} \sqrt{a_2 x_n} \|^2. \quad (5)$$

In the second time slot, the BS transmits a new signal to the relay and $U_1$. Meanwhile, the relay activates the $k$-th antenna to send the estimated symbol to $U_2$. Let $p = \hat{k}$ and $x_R = x_n$, then the observation at $U_2$ can be formulated as

$$y_2 = \sqrt{P_R} f_p x_R + n_2, \quad (6)$$

where $P_R$ is the transmit power of the relay and $f_p$ is the $p$-th column of $\mathbf{F}$. Finally, $U_2$ can decode the signal via the ML method.

III. PERFORMANCE ANALYSIS

A. Capacity Analysis

In this subsection, we will derive the instantaneous and ergodic capacities of $U_1$ and $U_2$ under Gaussian inputs.

1) Capacity of $U_1$: Recall that $U_2$ can perfectly eliminate the interference from the relay since the signal of $U_2$ is already known at $U_1$. Then, by assuming that all transmit antennas of the BS are activated with an equal probability of $1/N$, the instantaneous capacity of $U_1$ after SIC can be given by [6]

$$C_1 = \frac{1}{N} \sum_{k=1}^{N} \log_2 \left( 1 + \frac{P_B \| h_k \|^2 a_1}{\sigma_I^2} \right). \quad (7)$$

where $\| h_k \|^2$ obeys the central chi-square distribution with $2M$ degrees of freedom whose probability density function (PDF) is given by [15]

$$f_{\| h_k \|^2}(x) = \frac{1}{(N-1)!} x^{N-1} e^{-x}. \quad (8)$$

Let $\Theta = P_B \| h_k \|^2 a_1 / \sigma_I^2$, the ergodic capacity of $U_1$ can be expressed as [16]

$$C_1 = E_{\mathbf{H}} \log_2(1 + \Theta) = \frac{1}{\ln 2} \int_{0}^{\infty} \frac{1 - F_{\Theta}(\theta)}{1 + \theta} d\theta, \quad (9)$$

where $F_{\Theta}(\theta)$ is the cumulative distribution function (CDF) of $\Theta$, which can be given by [15]

$$F_{\Theta}(\theta) = p_r \left( \frac{P_B \| h_k \|^2 a_1}{\sigma_I^2} < \theta \right) = p_r \left( \| h_k \|^2 < \frac{\theta \sigma_I^2}{P_B a_1} \right) \quad (10)$$

$$= 1 - e^{-\frac{\theta \sigma_I^2}{P_B a_1 \lambda_h}} \sum_{k=0}^{M-1} \frac{\left( \frac{\theta \sigma_I^2}{P_B a_1 \lambda_h} \right)^k}{k!} \quad (10)$$
Then, by substituting (10) into (9), we can obtain the ergodic capacity of U1 as in (13), shown at the top of the next page, where \( Ei(x) = \int_{-\infty}^{x} e^{t}/t \, dt \) is the exponential integral function [17].

2) Capacity of U2: For the proposed scheme, the capacity of U2 is determined by the weakest link, hence we have

\[
C_2 = \min(C_{2,R}, C_{2,2}),
\]

(14)

where \( C_{2,R} \) and \( C_{2,2} \) are the capacity of U2 over the BS-to-relay and relay-to-U2 links, respectively. Since the data of U2 function \([17]\), we use an upper bound to facilitate the theoretical solution. When we assuming that the SI is extremely small, we can obtain

\[
C_{2,R}^{sp} = C_{2,2}^{sp} = C^{sp} \approx \log_2 N
\]

(22)
at the high signal-to-noise ratio (SNR) region (Please refer to [18] for the proof).

Then, we derive the ergodic capacity of U2. Based on (22), the ergodic capacity of U2 can be calculated as

\[
\bar{C}_2 = E_{G,F} \left[ \min(C_{2,R}^{sp} + C^{sp}, C_{2,2}^{sp} + C^{sp}) \right]
\]

(23)

where

\[
\min(\Xi_1, \Xi_2) = \min(\frac{P_B \|\mathbf{g}_{pk}\|^2 a_2}{P_B \|\mathbf{g}_{pk}\|^2 a_1 + I + \sigma_R^2}, \frac{P_R \|\mathbf{f}_k\|^2}{\sigma_2^2}).
\]

(24)

Similarly, let \( \Xi = \min(\Xi_1, \Xi_2) \), \( \bar{C}_2 \) can be rewritten as

\[
\bar{C}_2 = \frac{1}{\ln 2} \int_{0}^{+\infty} \frac{1 - F_{\Xi}(\xi) - C^{sp}}{1 + \xi} \, d\xi + C^{sp},
\]

(25)

where \( F_{\Xi}(\xi) \) denotes the CDF of the random variable \( \Xi \) as

\[
F_{\Xi}(\xi) = p_{r} (\min(\Xi_1, \Xi_2) < \xi)
\]

\[
= 1 - p_{r} (\min(\Xi_1, \Xi_2) > \xi)
\]

(26)

In (26), \( F_{\Xi_1}(\xi) \) and \( F_{\Xi_2}(\xi) \) can be derived by

\[
F_{\Xi_1}(\xi) = \begin{cases} 
1 - e^{-c_1 \sum_{n=0}^{N-2} \frac{1}{n!} (c_1)^n}, & \text{for } \xi < \frac{a_2}{a_1}, \\
1, & \text{otherwise}
\end{cases}
\]

(27)

where \( c_1 = \xi(I + \sigma_R^2)/(P_B \lambda_g a_2 - P_B \lambda_g a_1), \ c_2 = \xi \sigma_2^2 / P_R \lambda_f \).

Therefore, the ergodic capacity of U2 can be obtained as

\[
\bar{C}_2 = \frac{1}{\ln 2} \int_{0}^{+\infty} \frac{1 - F_{\Xi_1}(\xi) - C^{sp}}{1 + \xi} \, d\xi + C^{sp}
\]

(28)

\[
= \frac{1}{\ln 2} \int_{0}^{+\infty} \frac{1}{1 + \xi} \sum_{n=0}^{N-2} \frac{1}{n!} (c_1)^n \times \sum_{k=0}^{M-1} \frac{1}{k!} \left( \frac{c_2}{k} \right)^k \, d\xi + C^{sp}.
\]

(29)

Notice that it is difficult to derive the closed-form expression of (29). Thus, numerical integration is used to evaluate the integral.
B. BER Performance Analysis

In this subsection, we will derive the BER of U1 and U2 in closed-form expressions by using the union bound theory [15].

1) BER of U1: The BER of U1 depends on whether the signal of U2 is successfully eliminated. Thus, it can be divided into two parts as

\[ BER_1 = (1 - p_{SIC}) p_{c,1} + p_{SIC} p_{e,1} \]  \hspace{1cm} (30)

where \( p_{SIC} \) is the error probability that U1 detects the signal of U2, \( p_{c,1} \) is the BER of U1 under the condition that the signal of U2 is correctly detected, and \( p_{e,1} \) denotes the BER of U1 when errors occur in SIC.

Let us denote \( v = \| e_k x_n - e_k x_n \|^2 \), where \( e_k \) is the \( k \)-th row of an \( N \)-dimensional unit matrix. Consequently, the formulation of \( p_{SIC} \) can be given by [19]

\[ p_{SIC} \approx \frac{1}{NM_2} \sum_{k=1}^{N} \sum_{n=1}^{N} \sum_{m=k}^{M_2} \sum_{\hat{n} \neq n} f \left( \frac{\lambda_k P_{B} a_{0} v}{4(\sigma_1^2 + a_1)}, M \right), \]  \hspace{1cm} (31)

\[ f(x,y) = \Upsilon(y) \sum_{t=0}^{y-1} \left( y - 1 + t \right) \left[ 1 - \Upsilon(x) \right]^t, \]  \hspace{1cm} (32)

where \( \Upsilon(x) = 0.5(1 - \sqrt{x/1}) \).

Recall that the signal of U1 is only transmitted in the signal domain. Thus the expression of \( p_{c,1} \) can be given as

\[ p_{c,1} \approx \frac{1}{b_1 ^{2}} \sum_{m=1}^{M_1} \sum_{\gamma_{m} \neq \gamma_{\hat{m}}} D f \left( \frac{\lambda_f P_{B} a_{1} u}{4\sigma_1^2}, M \right), \]  \hspace{1cm} (33)

where \( u = |\gamma_m - \gamma_{\hat{m}}|^2 \) and \( D \) denotes the Hamming distance between the information bits of \( \gamma_m \) and \( \gamma_{\hat{m}} \).

On the other hand, if the interference due to failing SIC exists, \( p_{c,1} \) can be approximated by \( p_{c,1} \approx 1 - n_{M_1}/\log_2 M_1 \), where \( n_{M_1} \) is the average correct bits of the detection of \( M_1 \)-ary modulated symbols. \( n_{M_1} \) can be expressed as [12]

\[ n_{M_1} = \frac{1}{M_1 - 1} \sum_{t=1}^{\log_2 M_1} t \left( \log_2 M_1 \right). \]  \hspace{1cm} (34)

2) BER of U2: Similarly, the BER of U2 can be given by

\[ BER_2 = (1 - p_R) p_{c,2} + p_R p_{e,2}, \]  \hspace{1cm} (35)

where \( p_R \) is the error probability that the relay detects the signal of U2. \( p_{c,2} \) is the BER of U2 when the signal of U2 is correctly detected at the relay and \( p_{e,2} \) denotes the BER of U2 when the relay forwards an erroneous signal.

According to (2) and (6), both in the BS-to-relay and relay-to-U2 links, the SM symbol is used to carry the information of U2. In addition, \( N - 1 \) antennas are used to receive the information at the relay. Therefore, \( p_R \) can be given by

\[ p_R \approx A \sum_{k=1}^{N} \sum_{n=1}^{N} \sum_{m=k}^{M_2} \sum_{\hat{n} \neq n} f \left( \frac{p_R \lambda_f a_{0} v}{4(\sigma_1^2 + a_1 + t)}, N - 1 \right), \]  \hspace{1cm} (36)

where \( A = 1/NM_2 \). The value of \( p_{c,2} \) can be determined from \( p_{c,2} \approx ABEP_{SM} (\lambda_f P_R/\sigma_2^2) \), where \( ABEP_{SM} (\lambda_f P_R/\sigma_2^2) \) represents the average BER of SM over the Rayleigh fading channel with SNR of \( \lambda_f P_R/\sigma_2^2 \) at the receiver [19].

Finally, we assume that U2 is less likely to correctly detect its own signal when errors occur at the relay. Thus we consider the worst case, i.e., \( p_{c,2} = 1 \).

IV. SIMULATION RESULTS

In this section, the capacity and BER of the proposed scheme are presented and compared with the multi-input multi-output aided NOMA with FDR (MIMO-NOMA-FDR) and SM-aided orthogonal multiple access with half-duplex relay (SM-OMA-HDR) systems. The transmit SNR is defined as \( SNR = P/\sigma^2 \), where \( P \) is the transmit power and \( \sigma^2 \) is the variance of noise, respectively. In the simulation, we assume \( \lambda_h = \lambda_f = 1, \lambda_g = 0.8, \sigma_1^2 = \sigma_2^2 = \sigma_R^2 = \sigma^2 \) and \( P_B = P_R = R = 1 \).

In MIMO-NOMA-FDR, for a fair comparison, we consider that the BS has one transmit antenna, and the relay has one transmit antenna and \( N - 1 \) receive antennas. In SM-OMA-HDR, the BS transmits a modulated symbol to U1 during the \((t-1)\)-th time slot. Meanwhile, the relay transmits an SM symbol of U2 that received at the last time slot to U2. Then, during the \( t \)-th time slot, the BS transmits a new SM symbol of U2 to the relay.

In Fig. 2 and Fig. 3, the simulation results for the achievable capacity and the ergodic capacity of SM-NOMA-FDR are provided and compared with those of MIMO-NOMA-FDR and SM-OMA-HDR systems. The results verify the accuracy of ergodic capacities obtained from (13) and (29). We can observe that SM-NOMA-FDR provides a significant improvement in the sum capacity compared to MIMO-NOMA-FDR and SM-OMA-HDR. Fig. 2 and Fig. 3 show that U1 achieves the same performance in SM-NOMA-FDR and MIMO-NOMA-FDR. On the other hand, we can find that the capacity of U2 for SM-NOMA-FDR and MIMO-NOMA-FDR becomes a constant at the high SNR regime, which means that the relay suffers from the interference caused by U1 in this case. Moreover, in SM-OMA-HDR, U2 achieves a better capacity performance than...
In this paper, we have proposed an SM-NOMA-FDR scheme for the CDRT network, where the BS can directly communicate with the near user and communicates with the far user by using an FDR. The SI of the FDR has also been considered. Theoretical analyses of the ergodic capacity and the BER have been derived. Finally, the simulation results have verified the theoretical analyses and revealed that the proposed system achieves superior performances compared with MIMO-NOMA-FDR and SM-OMA-HDR systems.

V. Conclusion

In this paper, we have proposed an SM-NOMA-FDR scheme for the CDRT network, where the BS can directly communicate with the near user and communicates with the far user by using an FDR. The SI of the FDR has also been considered. Theoretical analyses of the ergodic capacity and the BER have been derived. Finally, the simulation results have verified the theoretical analyses and revealed that the proposed system achieves superior performances compared with MIMO-NOMA-FDR and SM-OMA-HDR systems.

REFERENCES


