Paper

An Approximate Evaluation of BER Performance for Downlink GSVD-NOMA with Joint Maximum-likelihood Detector

Ngo Thanh Hai and Dang Le Khoa

Department of Telecommunications and Networks, University of Science, VNU-HCM, District 5, Ho Chi Minh City, Vietnam

https://doi.org/10.26636/jtit.2022.160922

Abstract — Generalized Singular Value Decomposition (GSVD) is the enabling linear precoding scheme for multiple-input multiple-output (MIMO) non-orthogonal multiple access (NOMA) systems. In this paper, we extend research concerning downlink MIMO-NOMA systems with GSVD to cover bit error rate (BER) performance and to derive an approximate evaluation of the average BER performance. Specifically, we deploy, at the base station, the well-known technique of joint-modulation to generate NOMA symbols and joint maximum-likelihood (ML) to recover the transmitted data at end user locations. Consequently, the joint ML detector offers almost the same performance, in terms of average BER as ideal successive interference cancellation. Next, we also investigate BER performance of other precoding schemes, such as zero-forcing, block diagonalization, and simultaneous triangularization, comparing them with GSVD. Furthermore, BER performance is verified in different configurations in relation to the number of antennas. In cases where the number of transmit antennas is greater than twice the number of receive antennas, average BER performance is superior.

Keywords — generalized singular value decomposition (GSVD), joint maximum-likelihood, joint modulation, MIMO, non-orthogonal multiple access (NOMA)

1. Introduction

Non-orthogonal multiple access (NOMA) has emerged as a promising technology for the next generation of wireless networks (5G and beyond). This is due to the fact that NOMA is capable of improving spectrum efficiency, providing better fairness, as well as reducing latency in serving users all those factors are necessary for intelligent and dynamic next generation wireless networks [1], [2]. In conventional orthogonal multiple access (OMA), multiple users are assigned to different radio resources, such as frequency and time, meaning that the number of users severed is limited. However, NOMA can provide massive connections by simultaneously serving multiple users using the same spectrum resources [3], but this is done at the expense of increased intra-cell interference. To mitigate intra-cell interference, NOMA exploits successive interference cancellation (SIC) at receivers to the detect desired signals [4]. Therefore, the key principle of NOMA is based on superposition coding (SC) at the transmitter and SIC at the receiver.

Recently, the combination of multiple-input multiple-output with NOMA (MIMO-NOMA) has received a lot of attention in wireless communication due to its high spectral efficiency. In [5], ergodic capacity maximization was studied for the Rayleigh fading channel in MIMO-NOMA with statistical channel state information at the transmitter. The authors of [6] have investigated problems affecting the downlink MIMO-NOMA system with regards to clustering, beamforming, and power allocation. Many works have shown that the performance of MIMO-NOMA is superior to that of MIMO-OMA. However, MIMO-NOMA with a precoder scheme was realized and offered potential performance gains [7]. The authors of [7] have proposed a signal alignment based framework with precoding that is not only general and applicable to both uplink and downlink MIMO-NOMA systems, but also achieves a significant performance gain compared to MIMO-NOMA without precoding.

Precoding schemes are usually classified into two categories: nonlinear precoding and linear precoding. Nonlinear precoding is commonly known as dirty paper precoding (DPC) [8], [9], which can reach the maximum capacity region of MIMO channels if the transmitters perfectly estimate channel state information. However, DPC is difficult to implement due to computational complexity of the detection process. In order to reduce decoding complexity on the user side, linear precoding is necessary. The key principle of precoding consist in transforming channel matrices into diagonal matrices in the process of zero-forcing (ZF)-based precoding, block diagonalization (BD)-based precoding [10], and generalized singular value decomposition (GSVD) [11]. All of the above methods are referred to as simultaneous diagonalization (SD). In addition, the channel matrices, after being detected at the users', may have the form of triangular matrices when simultaneous triangularization (ST)-based precoding [12] is applied by relying on QR decomposition. The authors of [12] have revealed that the performance of ST precoding is close to that of the upper bounds of DPC and outperforms SD precoding as GSVD in terms of total system capacity. As far as

antenna configurations are concerned, ZF precoding and BD precoding are valid only when the total number of receive antennas of all users is lower than that of transmit antennas at the base station. Furthermore, ST precoding is capable of achieving better total system capacity if the number of transmit antennas is greater than that of receive antennas of each user. GSVD precoding, meanwhile, may apply to all antenna configurations.

As mentioned above, GSVD is a simple tool for linear precoding schemes for MIMO-NOMA implementations. In essence, GSVD can be extended to a point-to-point MIMO channel, where singular value decomposition (SVD) is applied during the conversion process. In [13], the authors proposed a transmission protocol combining GSVD and NOMA and evaluated the system's performance based on the expected data rates. Here, the scheme was considered in the asymptotic regime and the number of transmit antennas and receive antennas approached the infinite value. Moreover, the authors came up with limiting the distribution of the squared generalized singular value of the two users' channel matrices. The authors of [13] continued to make important contributions regarding GSVD-NOMA by achieving some new results on the distribution of the squared generalized singular value, as shown in [14]. In this paper, we take advantage of the joint density probability function of GSVD singular values in [14] to derive the average BER performance. In [15], the GSVD-NOMA scheme has been considered with a channel estimation error. This research has proposed three models of uncertainty and realized power allocation to balance signal-to-interference-plus-noise ratio (SINR).

Distribution of the squared generalized singular value function presented in [14] is only applicable for average results computations. However, in some research schemes concerned with secure transmission analysis and channel power allocation, the marginal probability density function (PDF) is necessary. Hence, the authors of [16] have obtained the distribution characteristics of the ordered GSVD singular values. The theoretical analysis of GSVD-based security transmission has first been presented in [14], where performance of a GSVD-based MIMO-OMA system was investigated for secrecy outage probability. Focusing on security of transmission in GSVD-based MIMO-NOMA schemes, the authors of [17], [18] analyzed theoretical secrecy outage probability. The results they obtained revealed the superiority of GSVD-NOMA in terms of efficiency and security, compared to GSVD-OMA.

As far as BER performance of NOMA is concerned, a relatively small number of studies has been carried out. In [19], the exact closed-form BER expression of the QPSK constellation for an uplink NOMA system was expressed over an additive white Gaussian noise (AWGN) channel. In [20], an exact closed-form BER expression under SIC error for downlink NOMA over Rayleigh fading channels was derived. Besides, the authors have also derived one-degree integral form exact expression and closed-form approximate BER expression for uplink NOMA. Moreover, over the Nakagami-m flat fading channel, the exact BER of downlink NOMA systems with SIC was derived for two and three user systems [21]. However, the performance of MIMO-NOMA, has been only studied in terms of overall system capacity and outage probability [5], [6]. The BER performance of the system has not been studied extensively. Recently, in [22], BER performance of an uplink NOMA was investigated with the use of the joint maximum-likelihood detector, where the base station was assumed to be equipped with N antennas. Apart from the SIC technique at the receivers, the authors in [23] came up with a technique to detect desired signals at the receivers, known as log-likelihood ratios (LLRs). For a downlink NOMA, the LLRs are characterized by almost the same error probability performance as ideal SIC probability.

In this paper, we consider a MIMO-NOMA system with GSVD, consisting of two users communicating with a base station (BS). The BS modulates the data of the two users using quadrature phase shift keying (QPSK) and superposes the said data by joint-modulation or multi-user superposition transmission case 2 (MUST-2) [24] to generate their respective NOMA symbols. For each user, we use the joint maximum-likelihood to recover data on each parallel GSVD-MIMO channel. The main contribution of this paper is that we derive the approximate expression of the average BER performance for the near user and the far user in downlink MIMO-NOMA systems with GSVD, as well as verify the correctness of the approximate expression obtained in the course of the Monte Carlo simulation. By relying on the approximate expression and simulation results, precoding schemes are compared with each other in order to choose the suitable precoding method for each antenna configuration. Moreover, by evaluating BER performance of GSVD, applicable antenna configurations are determined that may by designed.

The paper is organized as follows. Section 2 presents the system's model, the fundamental theory of GSVD, signal processing (MUST-2) at BS, and the joint maximum-likelihood (ML) detector to decode signals at the end users. In Section 3, we analyze numerical average BER performance, as well as derive its approximate closed-form expressions. In Section 4, works related to other precoding schemes, such as ZF, BD, and ST, are presented. In Section 5, numerical results are obtained to verify the precision of the analysis performed, the approximate expressions, and the simulation results. This section also shows the comparison with detection techniques and different precoding schemes. Finally, conclusion are presented in Section 6. Lemma and Theorem proofs are given in the Appendix.

2. System Model and Signal Processing

2.1. System Model

In this paper, we consider a MIMO-NOMA downlink system with one BS and two users: near user (NU) and far user (FU). The BS is equipped with N antennas and M antennas for each user (Fig. 1): $\boldsymbol{H}_n$ and $\boldsymbol{H}_f$ are $M \times N$ channel matrices from BS to NU and FU, respectively. Each element of the channel matrices is a mutually independent and identically distributed

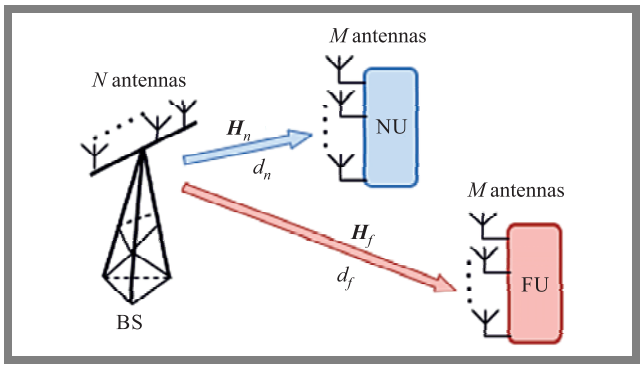

Fig. 1. A two-users MIMO downlink system model.

(i.i.d.) complex Gaussian random variable with zero mean and unit variance $\mathcal{CN}(0,1)$. The user channel is assumed to be constant in terms of the transmission duration of one codeword and changes independently from one codeword to the next. As such, it is viewed as a quasi-static channel. Moreover, to apply GSVD to the linear precoding scheme, we assume that channel state information (CSI) is known fully at both the base station and the users. d_n and d_f denote distances between the base station and NU and FU, respectively. α is the path loss exponent.

Let $\boldsymbol{S} \in \mathbb{C}^{L\times 1}$ be the transmit signal vector with the length L. The transmit signal is precoded with the linear precoder matrix $\boldsymbol{V} \in \mathbb{C}^{N\times L}$. The precoded signal vector is used to transmit the result as:

$$\boldsymbol{S}_p = \frac{1}{t}\boldsymbol{V}\boldsymbol{S}, \tag{1}$$

where t denotes the power normalization factor. Assuming that the average transmit power at BS is P, the value of t is chosen that need, to satisfy the following condition:

$$P = \frac{1}{t^2}E\left[\text{trace}\left(\boldsymbol{V}\boldsymbol{S}\boldsymbol{S}^{\text{H}}\boldsymbol{V}^{\text{H}}\right)\right]. \tag{2}$$

At the near user and the far user, the received signal is presented, respectively, as:

$$\begin{aligned}\tilde{\boldsymbol{Y}}_n &= \frac{d_n^{-\frac{\alpha}{2}}}{t}\boldsymbol{H}_n\boldsymbol{V}\boldsymbol{S} + \boldsymbol{N}_n,\\ \tilde{\boldsymbol{Y}}_f &= \frac{d_f^{-\frac{\alpha}{2}}}{t}\boldsymbol{H}_f\boldsymbol{V}\boldsymbol{S} + \boldsymbol{N}_f,\end{aligned} \tag{3}$$

where $\boldsymbol{N}_j \sim \mathcal{CN}(0, N_0\cdot\boldsymbol{I}_M)$, $j \in \{n,f\}$ is the additive white Gaussian noise (AWGN) vector and $\boldsymbol{I}_M$ denotes the identity matrix of size M. Moreover, at each user signals $\tilde{\boldsymbol{Y}}_j$ are detected with the linear matrices $\boldsymbol{U}_j^{\text{H}} \in \mathbb{C}^{K\times M}$, leading to:

$$\begin{aligned}\boldsymbol{Y}_n &= \frac{d_n^{-\frac{\alpha}{2}}}{t}\boldsymbol{U}_n^{\text{H}}\boldsymbol{H}_n\boldsymbol{V}\boldsymbol{S} + \tilde{\boldsymbol{N}}_n,\\ \boldsymbol{Y}_f &= \frac{d_f^{-\frac{\alpha}{2}}}{t}\boldsymbol{U}_f^{\text{H}}\boldsymbol{H}_f\boldsymbol{V}\boldsymbol{S} + \tilde{\boldsymbol{N}}_f,\end{aligned} \tag{4}$$

where $\tilde{\boldsymbol{N}}_j$ denotes AWGN after the detection process. The choice of $\boldsymbol{U}_j^{\text{H}}$ and $\boldsymbol{V}$ needs to satisfy diagonalization or triangularization conditions. In this paper, we apply GSVD to diagonalization. Then the product of three matrices $\boldsymbol{U}_j^{\text{H}}, \boldsymbol{H}_j$ and $\boldsymbol{V}$ is the diagonal matrix $\boldsymbol{D}_j$.

2.2. GSVD and the Joint PDF of Squared Generalized Singular Values

GSVD is found in [25] under the assumption of the same number of columns in two channel matrices and is presented in more detail in [13], [16]. By applying GSVD, $\boldsymbol{H}_n, \boldsymbol{H}_f$ are decomposed as follows:

$$\boldsymbol{H}_n = \boldsymbol{U}_n\boldsymbol{D}_n\boldsymbol{V}^{-1} \text{ and } \boldsymbol{H}_f = \boldsymbol{U}_f\boldsymbol{D}_f\boldsymbol{V}^{-1}, \tag{5}$$

where $\boldsymbol{U}_n, \boldsymbol{U}_f \in \mathbb{C}^{M\times M}$ are two unitary matrices, $\boldsymbol{V} \in \mathbb{C}^{N\times N}$ is an invertible matrix. $\boldsymbol{D}_n, \boldsymbol{D}_f \in \mathbb{C}^{M\times N}$ are two non-negative diagonal matrices whose structure depends on the choices of M and N.

a) The case when $M \geqslant N$**.**

$\boldsymbol{D}_n, \boldsymbol{D}_f$ are given by:

$$\boldsymbol{D}_n = \begin{bmatrix}\boldsymbol{S}_1\\ \boldsymbol{O}_{(M-N)\times N}\end{bmatrix} \text{ and } \boldsymbol{D}_f = \begin{bmatrix}\boldsymbol{O}_{(M-N)\times N}\\ \boldsymbol{S}_2\end{bmatrix}, \tag{6}$$

where $O_{(M-N)\times N}$ denotes the zero matrix of size $(M-N)\times N$, $\boldsymbol{S}_1 = \text{diag}\,(\alpha_1,\ldots,\alpha_N)$, $\boldsymbol{S}_2 = \text{diag}\,(\beta_1,\ldots,\beta_N)$ satisfying $1 \geqslant \alpha_1 \geqslant \ldots \geqslant \alpha_N \geqslant 0$, $1 \geqslant \beta_N \geqslant \ldots \geqslant \beta_1 \geqslant 0$ and $\alpha_i^2 + \beta_i^2 = 1$, $i = 1,\ldots,N$. The generalized singular values are defined as α_i. Based on proposition 1.2 from [26], the unordered generalized singular values, squared, of the pair $\boldsymbol{H}_n, \boldsymbol{H}_f$ $(X_i = \alpha_i^2)$ follow the law of the beta-Jacobi ensemble. Moreover, by combining them with the introduction from [27], we achieve the following joint probability density function of $X_i \in [0,1]$:

$$\begin{aligned}f_{X_1,\ldots,X_N}(x_1,\ldots,x_N) &= c_{J1}\prod_{1\leqslant i<j\leqslant N}(x_i - x_j)^2\\ &\times\prod_{i=1}^{N}x_i^{M-N}(1-x_i)^{M-N},\end{aligned} \tag{7}$$

where $c_{J1} = \prod_{j=1}^{N}\frac{\Gamma(2M-N+j)}{\Gamma(1+j)[\Gamma(M-N+j)]^2}$. Let $Y_i = \beta_i^2$, due to $\beta_i^2 = 1-\alpha_i^2$, so the joint probability density function of $Y_i \in [0,1]$ can be:

$$\begin{aligned}f_{Y_1,\ldots,Y_N}(y_1,\ldots,y_N) &= c_{J1}\prod_{1\leqslant i<j\leqslant N}(y_i - y_j)^2\\ &\times\prod_{i=1}^{N}y_i^{M-N}(1-y_i)^{M-N}.\end{aligned} \tag{8}$$

b) The case when $M < N < 2M$**.**

Put $q = 2M - N$ and $r = N - M$, $\boldsymbol{D}_n$ and $\boldsymbol{D}_f$ are written as follows:

$$\begin{aligned}\boldsymbol{D}_n &= \begin{bmatrix}\boldsymbol{I}_r & \boldsymbol{O}_{r\times q} & \boldsymbol{O}_{r\times r}\\ \boldsymbol{O}_{q\times r} & \boldsymbol{S}_1 & \boldsymbol{O}_{q\times r}\end{bmatrix},\\ \boldsymbol{D}_f &= \begin{bmatrix}\boldsymbol{O}_{q\times r} & \boldsymbol{S}_2 & \boldsymbol{O}_{q\times r}\\ \boldsymbol{O}_{r\times r} & \boldsymbol{O}_{r\times q} & \boldsymbol{I}_r\end{bmatrix},\end{aligned} \tag{9}$$

where $\boldsymbol{S}_1 = \text{diag}\,(\alpha_1,\ldots,\alpha_q)$, and $\boldsymbol{S}_2 = \text{diag}\,(\beta_1,\ldots,\beta_q)$, satisfying $1 \geqslant \alpha_1 \geqslant \ldots \geqslant \alpha_q \geqslant 0$, $1 \geqslant \beta_q \geqslant \ldots \geqslant \beta_1 \geqslant 0$ and $\alpha_i^2 + \beta_i^2 = 1$, $i = 1,\ldots,q$.

Lemma 1. When $M < N < 2M$, the joint probability density function of the unordered generalized singular values, squared, of the pair $\boldsymbol{H}_n, \boldsymbol{H}_f \in \mathbb{C}^{M\times N}$ $\left(X_i = \alpha_i^2\right)$ is given by:

$$f_{X_1,\ldots,X_q}(x_1,\ldots,x_q) = c_{J2} \prod_{1\leqslant i<j\leqslant q} (x_i - x_j)^2 \times \prod_{i=1}^{q} x_i^r (1-x_i)^r, \tag{10}$$

with $c_{J2} = \frac{1}{q!}\prod_{i=1}^{q} \frac{\Gamma(2M-i+1)}{\Gamma(q-i+1)[\Gamma(M-i+1)]^2}$.

Proof: see Appendix A.

The joint probability density function of $Y_i \in [0,1]$ can be easily concluded as:

$$f_{Y_1,\ldots,Y_q}(y_1,\ldots,y_q) = c_{J2} \prod_{1\leqslant i<j\leqslant q} (y_i - y_j)^2 \times \prod_{i=1}^{q} y_i^r (1-y_i)^r, \tag{11}$$

c) The case when $N \geqslant 2M$**.**

$\boldsymbol{D}_n$ and $\boldsymbol{D}_f$ are expressed as follows:

$$\begin{aligned} \boldsymbol{D}_n &= \left[\boldsymbol{I}_M \;\; \boldsymbol{O}_{M\times(N-M)}\right], \\ \boldsymbol{D}_f &= \left[\boldsymbol{O}_{M\times(N-M)} \;\; \boldsymbol{I}_M\right]. \end{aligned} \tag{12}$$

The structure of $\boldsymbol{D}_n$ and $\boldsymbol{D}_f$ in Eq. (12) is completely independent of small-scale fading properties.

2.3. Modulation MUST-2 at BS

Clearly, from the GSVD diagonalization for two channel matrices, we get the length of the transmit signal vector $\boldsymbol{S} \in \mathbb{C}^{N\times 1}$. The precoding matrix and detection matrices, respectively, are $\boldsymbol{V} \in \mathbb{C}^{N\times N}$ and $\boldsymbol{U}_j^{\mathrm{H}} \in \mathbb{C}^{M\times M}$. The received vectors $\boldsymbol{Y}_n$ and $\boldsymbol{Y}_f$ at NU and FU are expressed as:

$$\begin{aligned} \text{NU: } \boldsymbol{Y}_n &= \frac{d_n^{-\frac{\alpha}{2}}}{t}\boldsymbol{D}_n\boldsymbol{S} + \tilde{\boldsymbol{N}}_n, \\ \text{FU: } \boldsymbol{Y}_f &= \frac{d_f^{-\frac{\alpha}{2}}}{t}\boldsymbol{D}_f\boldsymbol{S} + \tilde{\boldsymbol{N}}_f, \end{aligned} \tag{13}$$

where $\tilde{\boldsymbol{N}}_j = \boldsymbol{U}_j^{\mathrm{H}}\boldsymbol{N}_j$. Due to the fact that $\boldsymbol{U}_j$ is a unitary matrix, $\tilde{\boldsymbol{N}}_j \sim \mathcal{CN}(0, N_0 \cdot \boldsymbol{I}_M)$.

At the BS, we consider three types of symbols. The first type is the QPSK symbol of NU's signal denoted as s_i^n, $E\left(|s_i^n|^2\right) = 1$. Next, s_i^f is the QPSK symbol of FU's signal, $E\left(\left|s_i^f\right|^2\right) = 1$. Finally, the NOMA symbol for the two users is denoted as s_i. The NOMA symbol has a generic form of:

$$s_i = \sqrt{\phi P}s_i^n + \sqrt{\theta P}s_i^f, \tag{14}$$

where ϕ, θ are the power allocation coefficients satisfying $\phi + \theta = 1$, $\phi < \theta$, for efficient SIC at NU. The modulation in Eq. (14) is referred to as multi-user superposition transmission case 1 (MUST-1) [24]. Due to independent modulation in conventional NOMA, the constellation of s_i does not follow the Gray mapping rule. Therefore, we modulate NOMA symbols using MUST-2 or joint-modulation, which means that bits from different users are mapped to one symbol taking into account the allocated power and the number of bits of each user. In this paper, we use a 16-QAM Gray-mapped constellation for joint mapping since 2 bits are assigned for NU and 2 bits are assigned for FU. The allocated power for NU and FU are respectively ϕP and θP, we have (Fig. 2):

$$d_1 = \frac{\sqrt{\theta P} - \sqrt{\phi P}}{\sqrt{2}} \quad \text{and} \quad d_2 = \frac{\sqrt{\theta P} + \sqrt{\phi P}}{\sqrt{2}},$$

It turns out that the constellation of MUST-2 is generated by permuting the position of points in the MUST-1 constellation satisfying the Gray mapping rule. Therefore, if users have different modulation orders according to their constellation's IQ, MUST-2 is valid for modulation at BS. Consider, for instance, a joint symbol at BS that has a single bit for FU and two bits for NU. Then, their constellation IQ will be the 8-QAM mapped Gray rule, with the positions of points arranged based on the power allocated to each user.

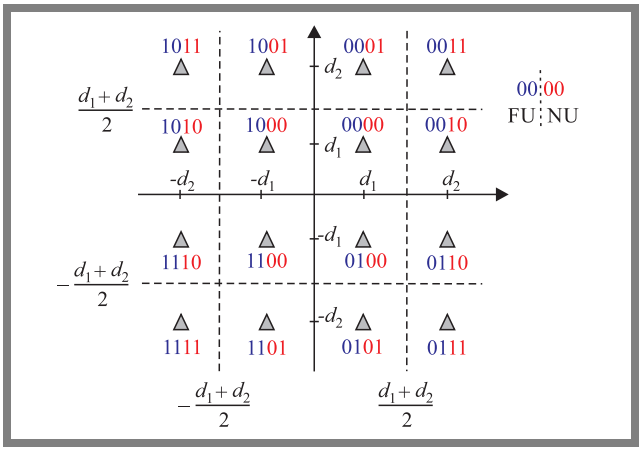

Fig. 2. Constellation of MUST-2 with 2 bits for NU and 2 bits for FU.

Based on the structure of $\boldsymbol{D}_n, \boldsymbol{D}_f$ and the received vector in Eq. (13), we formulate the forms of the transmitted vector $\boldsymbol{S}$ at BS as:

a) The case when $M \geqslant N$**.**

The structure of $\boldsymbol{S}$ is expressed as:

$$\boldsymbol{S} = (s_1, s_2, \ldots, s_N)^{\mathrm{T}}, \tag{15}$$

where $\boldsymbol{S}$ comprises N NOMA symbols s_i, $i = 1,\ldots,N$. The received symbol decomposed into parallel channels is written as:

$$\begin{aligned} y_i^n &= \frac{d_n^{-\frac{\alpha}{2}}}{t_1}\alpha_i s_i + \tilde{n}_i^n, \\ y_i^f &= \frac{d_f^{-\frac{\alpha}{2}}}{t_1}\beta_i s_i + \tilde{n}_i^f. \end{aligned} \tag{16}$$

The power normalization factor is given by:
$t = t_1 = \sqrt{\frac{N}{2M-N}}$ [13] and $\tilde{n}_i^n, \tilde{n}_i^f \sim \mathcal{CN}(0, N_0)$.

b) The case when $M < N < 2M$**.**

Assume that BS transmits the symbols vector $\boldsymbol{S}$ with N NOMA symbols. At the receiver, FU receives only the first M symbols, whereas, FU only gets the last M symbols.

So, we formulate a structure of symbols $\boldsymbol{S}$ transmitted at BS as:

$$\boldsymbol{S} = \Bigg(\underbrace{\sqrt{P}s_1^n, \ldots, \sqrt{P}s_r^n}_{r \text{ NU's symbols}}, \underbrace{s_1, \ldots, s_q}_{q \text{ NOMA symbols}} \underbrace{\sqrt{P}s_1^f, \ldots, \sqrt{P}s_r^f}_{r \text{ FU's symbols}} \Bigg)^{\mathrm{T}}. \tag{17}$$

The signals received at NU and FU are represented as:

$$\begin{aligned} y_i^n &= \frac{d_n^{-\frac{\alpha}{2}}}{t_2}\sqrt{P}s_i^n + \tilde{n}_i^n, \;\; i = 1, \ldots, r, \\ y_i^n &= \frac{d_n^{-\frac{\alpha}{2}}}{t_2}\alpha_i s_i + \tilde{n}_i^n, \;\; i = r+1, \ldots, M, \\ y_i^f &= \frac{d_f^{-\frac{\alpha}{2}}}{t_2}\beta_i s_i + \tilde{n}_i^f, \;\; i = 1, \ldots, q, \\ y_i^f &= \frac{d_f^{-\frac{\alpha}{2}}}{t_2}\sqrt{P}s_i^f + \tilde{n}_i^f, \;\; i = q+1, \ldots, M, \end{aligned} \tag{18}$$

where: $t_2 = t_1$ [13],
$\alpha_{i|i=r+1,\ldots,M} = \alpha_{i|i=1,\ldots,q}$,
$s_{i|i=r+1,\ldots,M} = s_{i|i=1,\ldots,q}$,
$s^f_{i|i=q+1,\ldots,M} = s^f_{i|i=1,\ldots,r}$.

c) The case when $N > 2M$**.**

The form of the symbol's vector at transmitted at BS is represented by:

$$\boldsymbol{S} = \Bigg(\underbrace{\sqrt{P}s_1^n, \ldots, \sqrt{P}s_M^n}_{M \text{ NU's symbols}}, \underbrace{0, \ldots, 0}_{N-2M \text{ symbols } 0} \underbrace{\sqrt{P}s_1^f, \ldots, \sqrt{P}s_M^f}_{M \text{ FU's symbols}} \Bigg)^{\mathrm{T}}. \tag{19}$$

NU and FU obtain the received symbols as:

$$\begin{aligned} y_i^n &= \frac{d_n^{-\frac{\alpha}{2}}}{t_3}\sqrt{P}s_i^n + \tilde{n}_i^n, \\ y_i^f &= \frac{d_f^{-\frac{\alpha}{2}}}{t_3}\sqrt{P}s_i^f + \tilde{n}_i^f, \; i = 1, \ldots, M, \end{aligned} \tag{20}$$

where $t_3 = \sqrt{\frac{2M}{N-2M}}$ [13].

2.4. Joint Maximum-likelihood Detector at NU and FU

As far as the QPSK symbols are concerned, users demodulate them easily by means of the maximum likelihood decision [28] on the parallel channels for $M < N < 2M$ and $2M < N$ scenarios. However, with NOMA symbols, we apply the joint maximum-likelihood (ML) detector to the estimation of NU signals and FU signals. This approach is mentioned in [22] to analyze BER performance of the uplink NOMA system with multiple receive antennas over the Rayleigh fading channel. This means that each user estimates firstly the joint symbols (NOMA symbols) on the 16-QAM constellation and, after that, based on their correct-order bits, obtains their own symbols. The detection of joint symbols is:

$$r_i^* = \underset{r_i \in \boldsymbol{X}}{\arg\min} \left| z_i^j - r_i \right|^2, \tag{21}$$

where:

$$z_i^n = \frac{t_1}{\alpha_i d_n^{-\frac{\alpha}{2}}} y_i^n, \;\; z_i^f = \frac{t_1}{\beta_i d_f^{-\frac{\alpha}{2}}} y_i^f$$

and $j \in \{n, f\}$ on the NOMA symbol channels. $\boldsymbol{X}$ is a set of the constellation point coordinates. The users use r_i^* for bit mapping and obtain decoded bits for NU and FU. Here, the first two bits of a joint symbol correspond to FU, and the remaining two bits are represented as two bits of NU.

3. BER Performance Analysis

In this section, we derive the approximate expression of the average BER performance of NU and FU.

Let us define the generic form of the constellation point as $\overline{b_1 b_2 b_3 b_4}$, where b_1, b_2 are two bits of FU corresponding to the blue bits in Fig. 2 and b_3, b_4 represent two red bits shown in this figure, being the two bits of NU. First, we investigate BER performance in one codeword. After that, the average BER performance is calculated for the overall fading domain.

3.1. BER of NU

a) The case when $M \geqslant N$**.**

NU receives signals on N parallel NOMA symbol channels, so the average BER is:

$$P_{n1} = \frac{1}{N}\sum_{i=1}^{N} P_i^{n1}, \tag{22}$$

where P_i^{n1} is the average BER in the i-th parallel channel. P_i^{n1} is represented by the error probability for bit b_3 as P_{b3} and the error probability for bit b_4 as P_{b4} in the form of $\overline{b_1 b_2 b_3 b_4}$ as follows:

$$P_i^{n1} = \frac{1}{2}\left(P_{b3} + P_{b4}\right). \tag{23}$$

Bit $b_3 = 1$ when the real part of the transmitted symbol s_{iI} equals either $-d_2$ or d_2 and $b_3 = 0$ implies that $s_{iI} = -d_1$ or $s_{iI} = d_1$. Then, P_{b3} is:

$$\begin{aligned} P_{b3} = \frac{1}{4}\big(&P_{b3|s_{iI}=-d_2} + P_{b3|s_{iI}=-d_1} + P_{b3|s_{iI}=d_1} \\ &+ P_{b3|s_{iI}=d_2}\big), \end{aligned} \tag{24}$$

where $P_{b3|s_{iI}=x}$ is the error probability for bit b_3 when the real part of the transmitted symbol assumes the value of x. From Eq. (21), z_i^n is written as: $z_i^n = s_i + w_i^n$ and $w_i^n = \frac{t_1}{\alpha_i d_n^{-\frac{\alpha}{2}}}\tilde{n}_i^n$. By investigating the constellation in Fig. 2, $P_{b3|s_{iI}=-d_2}$ can be defined by:

$$\begin{aligned} P_{b3|s_{iI}=-d_2} &= \Pr\left(-\frac{d_1+d_2}{2} < -d_2 + w_{iI}^n < \frac{d_1+d_2}{2}\right) \\ &= \Pr\left(-\frac{d_1-d_2}{2} < w_{iI}^n < \frac{d_1+3d_2}{2}\right), \end{aligned} \tag{25}$$

where w_{iI}^n is the real part of w_i^n and $w_{iI}^n \sim \mathcal{N}\left(0, \frac{t_1^2 N_0}{2\alpha_i^2 d_n^{-\alpha}}\right)$. Putting $\rho = \frac{P}{N_0}$, $a_1 = d_n^{-\alpha}\phi$, $a_2 = d_n^{-\alpha}\left(2\sqrt{\theta} - \sqrt{\phi}\right)^2$ and $a_3 = d_n^{-\alpha}\left(2\sqrt{\theta} + \sqrt{\phi}\right)^2$. By integrating the probability density function of w_{iI}^n over the value domain in Eq. (25), we obtain $P_{b3|s_{iI}=-d_2}$ as:

$$P_{b3|s_{iI}=-d_2} = Q\left(\sqrt{\frac{a_1\rho}{t_1^2}\alpha_i^2}\right) - Q\left(\sqrt{\frac{a_3\rho}{t_1^2}\alpha_i^2}\right). \quad (26)$$

Similarly as in $P_{b3|s_{iI}=-d_2}$, $P_{b3|s_{iI}=-d_1}$ is:

$$\begin{aligned} P_{b3|s_{iI}=-d_1} &= \Pr\left(w_{iI}^n < \frac{d_1 - d_2}{2}\right) \\ &\quad + \Pr\left(w_{iI}^n > \frac{3d_1 + d_2}{2}\right) \\ &= Q\left(\sqrt{\frac{a_1\rho}{t_1^2}\alpha_i^2}\right) + Q\left(\sqrt{\frac{a_2\rho}{t_1^2}\alpha_i^2}\right). \end{aligned} \quad (27)$$

Additionally, we also show that $P_{b3|s_{iI}=d_1} = P_{b3|s_{iI}=-d_1}$ and $P_{b3|s_{iI}=d_2} = P_{b3|s_{iI}=-d_2}$. Due to the symmetrical property of the constellation in Fig. 2, we obtain $P_{b3} = P_{b4}$. Therefore, the average BER for one codeword on the i-th parallel channel is:

$$\begin{aligned} P_i^{n1} &= \frac{1}{2}\left[2Q\left(\sqrt{\frac{a_1\rho}{t_1^2}\alpha_i^2}\right) + Q\left(\sqrt{\frac{a_2\rho}{t_1^2}\alpha_i^2}\right)\right. \\ &\quad \left. - Q\left(\sqrt{\frac{a_3\rho}{t_1^2}\alpha_i^2}\right)\right]. \end{aligned} \quad (28)$$

Next, in the overall fading domain we evaluate the average BER of NU:

$$\begin{aligned} \overline{P}_{n1} &= \int_0^1 \cdots \int_0^1 \frac{1}{N}\sum_{i=1}^{N} P_i^{n1}(x_i) \\ &\quad \times f_{X_1,\ldots,X_N}(x_1,\ldots,x_N)\mathrm{d}x_1 \ldots \mathrm{d}x_N. \end{aligned} \quad (29)$$

Theorem 1. The average BER of NU in the overall fading domain can be approximated as:

$$\begin{aligned} \overline{P}_{n1} \simeq \frac{c_{J1}}{2N}\Bigg[&\frac{1}{2}\sum_{\sigma\in \boldsymbol{S}_N}\sum_{j=1}^{N} B(p_{j1}+1, q_1+1)G(t_1, p_{j1}, q_1) \\ &\times \prod_{\substack{i=1\\ i\neq j}}^{N} B(p_{i1}+1, q_1+1) + \sum_{\sigma_1,\sigma_2\in \boldsymbol{S}_N} \mathrm{sgn}(\sigma_1)\mathrm{sgn}(\sigma_2) \\ &\times \sum_{j=1}^{N} B(p_{j2}+1, q_1+1)G(t_1, p_{j2}, q_1) \\ &\times \prod_{\substack{i=1\\ i\neq j}}^{N} B(p_{i2}+1, q_1+1)\Bigg], \end{aligned} \quad (30)$$

where $\boldsymbol{S}_N$ is the set of the permutations of $\{1, 2, \ldots, N\}$ and sgn (σ) denotes the sign of the permutations σ, $q_1 = M - N$, $p_{k1} = M - N + 2\sigma(k) - 2$, $p_{k2} = M - N + \sigma_1(k) + \sigma_2(k) - 2$, $k \in \{i, j\}$. $B(x, y)$ is the Beta function defined in [29]. $G(t, x, y)$ is:

$$\begin{aligned} G(t,x,y) =& \frac{1}{3}F\left(\frac{a_1}{2}\right) + \frac{1}{6}F\left(\frac{a_2}{2}\right) - \frac{1}{6}F\left(\frac{a_3}{2}\right) \\ &+ F\left(\frac{2a_1}{3}\right) + \frac{1}{2}F\left(\frac{2a_2}{3}\right) - \frac{1}{2}F\left(\frac{2a_3}{3}\right), \end{aligned}$$

where $F(u) = {}_1F_1(x+1; x+y+2; -\frac{u\rho}{t^2})$ and ${}_1F_1(a; b; z)$ is the generalized hypergeometric function [30].

Proof: See Appendix B.

b) The case when $M < N < 2M$**.**

NU receives symbols on $N - M$ QPSK symbol channels and $2M - N$ NOMA symbol channels, so the average BER is:

$$P_{n2} = \frac{1}{M}\left(rP_n^{\mathrm{QPSK}} + \sum_{i=1}^{q} P_i^{n2}\right), \quad (31)$$

where P_n^{QPSK} is the average BER performance on the QPSK symbol channel [28]:

$$P_n^{\mathrm{QPSK}} \approx Q\left(\sqrt{\frac{Pd_n^{-\alpha}}{t_2^2 N_0}}\right). \quad (32)$$

Considering the NOMA symbol channels, only the number of channels differs between $M \geqslant N$ and $M < N < 2M$ cases, so P_i^{n2} can be expressed similarly as P_i^{n1}:

$$\begin{aligned} P_i^{n2} &= \frac{1}{2}\left[2Q\left(\sqrt{\frac{a_1\rho}{t_1^2}\alpha_i^2}\right) + Q\left(\sqrt{\frac{a_2\rho}{t_1^2}\alpha_i^2}\right)\right. \\ &\quad \left. - Q\left(\sqrt{\frac{a_3\rho}{t_1^2}\alpha_i^2}\right)\right]. \end{aligned} \quad (33)$$

The average BER of NU is evaluated in the overall fading domain:

$$\overline{P}_{n2} = \frac{1}{M}\left(rP_n^{\mathrm{QPSK}} + \sum_{i=1}^{q} \overline{P}_i^{n2}\right). \quad (34)$$

By applying **Lemma 1** and the same argument as in **Theorem 1**, we obtain the approximate expression of $T_n = \sum_{i=1}^{q} \overline{P}_i^{n2}$ as:

$$\begin{aligned} T_n \simeq \frac{c_{J2}}{2}\Bigg[&\frac{1}{2}\sum_{\sigma\in \boldsymbol{S}_q}\sum_{j=1}^{q} \mathrm{B}(p'_{j1}+1, q_2+1)\mathrm{G}(t_2, p'_{j1}, q_2) \\ &\times \prod_{\substack{i=1\\ i\neq j}}^{q} \mathrm{B}(p'_{i1}+1, q_2+1) + \sum_{\sigma_1,\sigma_2\in \boldsymbol{S}_q} \mathrm{sgn}(\sigma_1)\mathrm{sgn}(\sigma_2) \\ &\times \sum_{j=1}^{q} \mathrm{B}(p'_{j2}+1, q_2+1)\mathrm{G}(t_2, p'_{j2}, q_2) \\ &\times \prod_{\substack{i=1\\ i\neq j}}^{q} \mathrm{B}(p'_{i2}+1, q_2+1)\Bigg], \end{aligned} \quad (35)$$

where $\boldsymbol{S}_q$ is the set of the permutations of $\{1, 2, \ldots, q\}$ and $p'_{k1} = N - M + 2\sigma(k) - 2$, $p'_{k2} = N - M + \sigma_1(k) + \sigma_2(k) - 2$, $k \in \{i, j\}$ and $q_2 = N - M$.

Substituting Eqs. (32) and (35) into Eq. (34), we achieve the approximate average BER for NU.

c) The case when $N > 2M$**.**

The user's channel is decomposed into M complex Gaussian channels. From Eq. (20), the average BER for NU is:

$$\overline{P}_{n3} \approx Q\left(\sqrt{\frac{Pd_n^{-\alpha}}{t_3^2 N_0}}\right). \tag{36}$$

3.2. BER of FU

a) The case when $M \geqslant N$**.**

In this case, FU also receives signals on N parallel NOMA symbol channels, so the average BER is expressed as:

$$P_{f1} = \frac{1}{N}\sum_{i=1}^{N} P_i^{f1}, \tag{37}$$

where P_i^{f1} is the average BER on the i-th parallel channel. FU's signals can be identified by the first two bits b_1, b_2 of the transmitted symbol. This generates the result of $P_i^{f1} = \frac{1}{2}(P_{b1} + P_{b2})$. P_{bj} is the error probability of j-th bit $j = 1, 2$. Along similar lines, in the NU case, we also get $P_{b1} = P_{b2}$ and:

$$P_i^{f1} = \frac{1}{2}\left[Q\left(\sqrt{\frac{c_1\rho}{t_1^2}\beta_i^2}\right) + Q\left(\sqrt{\frac{c_2\rho}{t_1^2}\beta_i^2}\right)\right], \tag{38}$$

where:

$$c_1 = d_f^{-\alpha}\left(\sqrt{\theta} - \sqrt{\phi}\right)^2, \quad c_2 = d_f^{-\alpha}\left(\sqrt{\theta} + \sqrt{\phi}\right)^2.$$

Due to the similarity of Joint-PDF of X_i and Y_i in Eqs. (7) and (8), it can be shown that:

$$\begin{aligned}\overline{P}_{f1} \simeq \frac{c_{J1}}{4N}\Bigg[&\frac{1}{2}\sum_{\sigma\in\boldsymbol{S}_N}\sum_{j=1}^{N}\mathrm{B}(p_{j1}+1, q_1+1)\mathrm{H}(t_1, p_{j1}, q_1)\\ &\times\prod_{\substack{i=1\\ i\neq j}}^{N}\mathrm{B}(p_{i1}+1, q_1+1) + \sum_{\sigma_1,\sigma_2\in\boldsymbol{S}_N}\mathrm{sgn}(\sigma_1)\mathrm{sgn}(\sigma_2)\\ &\times\sum_{j=1}^{N}\mathrm{B}(p_{j2}+1, q_1+1)\mathrm{H}(t_1, p_{j2}, q_1)\\ &\times\prod_{\substack{i=1\\ i\neq j}}^{N}\mathrm{B}(p_{i2}+1, q_1+1)\Bigg],\end{aligned} \tag{39}$$

where $p_{i1}, p_{i2}, p_{j1}, p_{j2}$ and q_1 are defined in **Theorem 1**. $\mathrm{H}(t, x, y)$ is expressed through $\mathrm{F}(u)$ as:

$$\mathrm{H}(t, x, y) = \frac{1}{3}\mathrm{F}\left(\frac{c_1}{2}\right) + \frac{1}{3}\mathrm{F}\left(\frac{c_2}{2}\right) + \mathrm{F}\left(\frac{2c_1}{3}\right) + \mathrm{F}\left(\frac{2c_2}{3}\right).$$

b) The case when $M < N < 2M$**.**

By the same argument as in the NU case, we get the average BER of FU on the overall fading domain:

$$\overline{P}_{f2} = \frac{1}{M}\left(rP_f^{\mathrm{QPSK}} + \sum_{i=1}^{q}\overline{P}_i^{f2}\right). \tag{40}$$

Similarly as T_n, let $T_f = \sum_{i=1}^{q}\overline{P}_i^{f2}$. We can prove that

$$\begin{aligned}T_f \simeq \frac{c_{J2}}{4}\Bigg[&\frac{1}{2}\sum_{\sigma\in\boldsymbol{S}_q}\sum_{j=1}^{q}\mathrm{B}(p'_{j1}+1, q_2+1)\mathrm{H}(t_2, p'_{j1}, q_2)\\ &\times\prod_{\substack{i=1\\ i\neq j}}^{q}\mathrm{B}(p'_{i1}+1, q_2+1) + \sum_{\sigma_1,\sigma_2\in\boldsymbol{S}_q}\mathrm{sgn}(\sigma_1)\mathrm{sgn}(\sigma_2)\\ &\times\sum_{j=1}^{q}\mathrm{B}(p'_{j2}+1, q_2+1)\mathrm{H}(t_2, p'_{j2}, q_2)\\ &\times\prod_{\substack{i=1\\ i\neq j}}^{q}\mathrm{B}(p'_{i2}+1, q_2+1)\Bigg],\end{aligned} \tag{41}$$

where $p'_{i1}, p'_{i2}, p'_{j1}, p'_{j2}$ and q_2 are defined in Eq. (35). P_f^{QPSK} is the average BER in the QPSK symbol channel given by:

$$P_f^{\mathrm{QPSK}} \approx Q\left(\sqrt{\frac{Pd_f^{-\alpha}}{t_2^2 N_0}}\right). \tag{42}$$

By substituting Eqs. (41) and (42) into (40), we derive the approximate expression for the average BER of FU.

c) The case when $N > 2M$**.**

The average BER for FU is evaluated as:

$$\overline{P}_{f3} \approx Q\left(\sqrt{\frac{Pd_f^{-\alpha}}{t_3^2 N_0}}\right). \tag{43}$$

The summary theoretical analysis BER performance is shown in Table 1.

4. Works Relating to ZF, BD, and ST

In this section, we briefly mention the precoding techniques as ZF, BD, and ST. This serves as a basis for comparing them with GSVD in terms of the BER performance.

4.1. ZF Based Precoding

Zero-forcing based precoding [10] is valid only when $2M \leqslant N$. In this case, the detection matrices at the users are $\boldsymbol{U}_n = \boldsymbol{I}_M$, and $\boldsymbol{U}_f = \boldsymbol{I}_M$. The precoding matrix is:

$$\boldsymbol{V} = \boldsymbol{H}^{\mathrm{H}}\left(\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}\right)^{-1}, \tag{44}$$

where $\boldsymbol{H} \in \mathbb{C}^{2M\times N}$ is denoted as $\boldsymbol{H} = \left[\boldsymbol{H}_n^{\mathrm{H}} \;\; \boldsymbol{H}_f^{\mathrm{H}}\right]^{\mathrm{H}}$. The precoding and detection processes are presented through the following equations:

$$\begin{aligned}\boldsymbol{U}_n^{\mathrm{H}}\boldsymbol{H}_n\boldsymbol{V} &= \left[\boldsymbol{I}_M \;\; \boldsymbol{O}\right],\\ \boldsymbol{U}_f^{\mathrm{H}}\boldsymbol{H}_f\boldsymbol{V} &= \left[\boldsymbol{O} \;\; \boldsymbol{I}_M\right].\end{aligned} \tag{45}$$

Using $E\left(\boldsymbol{S}\boldsymbol{S}^{\mathrm{H}}\right) = \boldsymbol{I}_N$, from Eq. (2), we easily obtain the power normalization factor given by:

$$t_{ZF} = \sqrt{\frac{2M}{N-2M}}. \tag{46}$$

Tab. 1. Analysis of BER performance of NU and FU for GSVD based precoding.

BER	Antenna configurations		
	$M \geqslant N$	$M < N < 2M$	$N > 2M$
NU	$\overline{P}_{n1} = E_{X_1,\ldots,X_N}\left(\frac{1}{N}\sum_{i=1}^{N} P_i^{n1}\right)$ $\overline{P}_{n1} \simeq$ Eq. (30)	$\overline{P}_{n2} = \frac{1}{M}\left[rQ\left(\sqrt{\frac{Pd_n^{-\alpha}}{t_2^2 N_0}}\right) + E_{X_1,\ldots,X_q}\left(\sum_{i=1}^{q} P_i^{n2}\right)\right]$ $\overline{P}_{n2} \simeq \frac{1}{M}\left[rQ\left(\sqrt{\frac{Pd_n^{-\alpha}}{t_2^2 N_0}}\right) + T_n\right], T_n \simeq$ Eq. (35)	$\overline{P}_{n3} \approx Q\left(\sqrt{\frac{Pd_n^{-\alpha}}{t_3^2 N_0}}\right)$
FU	$\overline{P}_{f1} = E_{Y_1,\ldots,Y_N}\left(\frac{1}{N}\sum_{i=1}^{N} P_i^{f1}\right)$ $\overline{P}_{f1} \simeq$ Eq. (39)	$\overline{P}_{f2} = \frac{1}{M}\left[rQ\left(\sqrt{\frac{Pd_f^{-\alpha}}{t_2^2 N_0}}\right) + E_{Y_1,\ldots,Y_q}\left(\sum_{i=1}^{q} P_i^{f2}\right)\right]$ $\overline{P}_{f2} \simeq \frac{1}{M}\left[rQ\left(\sqrt{\frac{Pd_f^{-\alpha}}{t_2^2 N_0}}\right) + T_f\right], T_f \simeq$ Eq. (41)	$\overline{P}_{f3} \approx Q\left(\sqrt{\frac{Pd_f^{-\alpha}}{t_3^2 N_0}}\right)$

4.2. BD Based Precoding

Similarly to ZF based precoding [10], BD based precoding is valid only when $2M \leqslant N$. Carrying out SVD of $\boldsymbol{H}_n$ and $\boldsymbol{H}_f$, can be obtained as:

$$\begin{aligned} \boldsymbol{H}_n &= \tilde{\boldsymbol{U}}_n \begin{bmatrix} \tilde{\boldsymbol{D}}_n^{(1)} & \boldsymbol{O} \end{bmatrix} \begin{bmatrix} \tilde{\boldsymbol{V}}_n^{(1)} & \tilde{\boldsymbol{V}}_n^{(0)} \end{bmatrix}^{\mathrm{H}}, \\ \boldsymbol{H}_f &= \tilde{\boldsymbol{U}}_f \begin{bmatrix} \boldsymbol{O} & \tilde{\boldsymbol{D}}_f^{(1)} \end{bmatrix} \begin{bmatrix} \tilde{\boldsymbol{V}}_f^{(0)} & \tilde{\boldsymbol{V}}_f^{(1)} \end{bmatrix}^{\mathrm{H}}, \end{aligned} \tag{47}$$

where $\tilde{\boldsymbol{D}}_n^{(1)}, \tilde{\boldsymbol{D}}_f^{(1)} \in \mathbb{C}^{M\times M}$ and $\tilde{\boldsymbol{V}}_n^{(0)}, \tilde{\boldsymbol{V}}_f^{(0)} \in \mathbb{C}^{N\times(N-M)}$. Next, using SVD to $\boldsymbol{H}_n\tilde{\boldsymbol{V}}_f^{(0)}$ and $\boldsymbol{H}_f\tilde{\boldsymbol{V}}_n^{(0)}$, the results of the analyses are presented as:

$$\begin{aligned} \boldsymbol{H}_n\tilde{\boldsymbol{V}}_f^{(0)} &= \boldsymbol{U}_n \begin{bmatrix} \boldsymbol{D}_n^{(1)} & \boldsymbol{O} \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_n^{(1)} & \boldsymbol{V}_n^{(0)} \end{bmatrix}^{\mathrm{H}}, \\ \boldsymbol{H}_f\tilde{\boldsymbol{V}}_n^{(0)} &= \boldsymbol{U}_f \begin{bmatrix} \boldsymbol{O} & \boldsymbol{D}_f^{(1)} \end{bmatrix} \begin{bmatrix} \boldsymbol{V}_f^{(0)} & \boldsymbol{V}_f^{(1)} \end{bmatrix}^{\mathrm{H}}, \end{aligned} \tag{48}$$

with $\boldsymbol{D}_n^{(1)}, \boldsymbol{D}_f^{(1)} \in \mathbb{C}^{M\times M}$ and $\boldsymbol{V}_n^{(1)}, \boldsymbol{V}_f^{(1)} \in \mathbb{C}^{(N-M)\times M}$. The precoding matrix $\boldsymbol{V}$ can be obtained by concatenation of the precoding matrices as:

$$\boldsymbol{V} = \begin{bmatrix} \tilde{\boldsymbol{V}}_f^{(0)}\boldsymbol{V}_n^{(1)} & \tilde{\boldsymbol{V}}_n^{(0)}\boldsymbol{V}_f^{(1)} \end{bmatrix}. \tag{49}$$

The strategies of precoding and detection at BS and the users respectively can be performed by:

$$\begin{aligned} \boldsymbol{U}_n^{\mathrm{H}}\boldsymbol{H}_n\boldsymbol{V} &= \begin{bmatrix} \boldsymbol{D}_n^{(1)} & \boldsymbol{O} \end{bmatrix}, \\ \boldsymbol{U}_f^{\mathrm{H}}\boldsymbol{H}_f\boldsymbol{V} &= \begin{bmatrix} \boldsymbol{O} & \boldsymbol{D}_f^{(1)} \end{bmatrix}. \end{aligned} \tag{50}$$

By using $E\left(\boldsymbol{S}\boldsymbol{S}^{\mathrm{H}}\right) = \boldsymbol{I}_N$, from Eq. (2) the power normalization factor can be:

$$t_{BD} = \sqrt{2M}. \tag{51}$$

4.3. ST Based Precoding

ST based precoding is mentioned in [12] and is valid when $M \leqslant N$.

a) The case when $M \leqslant N < 2M$**.**

From (47), we concatenate $\tilde{\boldsymbol{V}}_n^{(0)}$ and $\tilde{\boldsymbol{V}}_f^{(0)}$, the matrix $\boldsymbol{H}$ is:

$$\boldsymbol{H} = \begin{bmatrix} \tilde{\boldsymbol{V}}_n^{(0)} & \tilde{\boldsymbol{V}}_f^{(0)} \end{bmatrix}^{\mathrm{H}}. \tag{52}$$

Next, realizing SVD decomposition $\boldsymbol{H}$, we obtain:

$$\boldsymbol{H} = \tilde{\boldsymbol{U}}\tilde{\boldsymbol{D}} \begin{bmatrix} \tilde{\boldsymbol{V}}^{(1)} & \tilde{\boldsymbol{V}}^{(0)} \end{bmatrix}^{\mathrm{H}}, \tag{53}$$

where $\tilde{\boldsymbol{V}}^{(0)} \in \mathbb{C}^{N\times(2M-N)}$. Let, QR decomposition be:

$$\begin{aligned} \boldsymbol{Q}_n\boldsymbol{R}_n &= \boldsymbol{H}_n \begin{bmatrix} \tilde{\boldsymbol{V}}^{(0)} & \tilde{\boldsymbol{V}}_f^{(0)} \end{bmatrix}, \\ \boldsymbol{Q}_f\boldsymbol{R}_f &= \boldsymbol{H}_f \begin{bmatrix} \tilde{\boldsymbol{V}}^{(0)} & \tilde{\boldsymbol{V}}_n^{(0)} \end{bmatrix}, \end{aligned} \tag{54}$$

where $\boldsymbol{Q}_n, \boldsymbol{Q}_f \in \mathbb{C}^{M\times M}$. By setting the precoding matrix $\boldsymbol{V} = \begin{bmatrix} \tilde{\boldsymbol{V}}^{(0)} & \tilde{\boldsymbol{V}}_f^{(0)} & \tilde{\boldsymbol{V}}_n^{(0)} \end{bmatrix}$ and choosing $\boldsymbol{U}_n = \boldsymbol{Q}_n, \boldsymbol{U}_f = \boldsymbol{Q}_f$, the simultaneous triangularization of $\boldsymbol{H}_n$ and $\boldsymbol{H}_f$ is:

$$\begin{aligned} \boldsymbol{U}_n^{\mathrm{H}}\boldsymbol{H}_n\boldsymbol{V} &= \begin{bmatrix} \boldsymbol{R}_n & \boldsymbol{O} \end{bmatrix}, \\ \boldsymbol{U}_f^{\mathrm{H}}\boldsymbol{H}_f\boldsymbol{V} &= \begin{bmatrix} \boldsymbol{R}'_f & \boldsymbol{O} & \boldsymbol{R}''_f \end{bmatrix}, \end{aligned} \tag{55}$$

where $\boldsymbol{R}_n \in \mathbb{C}^{M\times M}$, $\boldsymbol{R}'_f \in \mathbb{C}^{M\times(2M-N)}$ and $\boldsymbol{R}''_f \in \mathbb{C}^{M\times(N-M)}$. Moreover, $\boldsymbol{R}_n$ and $\boldsymbol{R}_f = \begin{bmatrix} \boldsymbol{R}'_f & \boldsymbol{R}''_f \end{bmatrix}$ are upper-triangular matrices with real-valued entries on their main diagonals.

From Eq. (2), the power normalization factor in ST based precoding case is given as:

$$t_{ST} = \sqrt{N}. \tag{56}$$

b) The case when $N \geqslant 2M$**.**

From Eq. (47), we get the precoding matrix such as:

$$\boldsymbol{V} = \begin{bmatrix} \tilde{\boldsymbol{V}}_f^{(0)} & \tilde{\boldsymbol{V}}_n^{(0)} \end{bmatrix}. \tag{57}$$

Realizing triangularization $\boldsymbol{H}_n\tilde{\boldsymbol{V}}_f^{(0)}$ and $\boldsymbol{H}_f\tilde{\boldsymbol{V}}_n^{(0)}$ by QR decomposition:

$$\begin{aligned} \boldsymbol{Q}_n\boldsymbol{R}_n &= \boldsymbol{H}_n\tilde{\boldsymbol{V}}_f^{(0)}, \\ \boldsymbol{Q}_f\boldsymbol{R}_f &= \boldsymbol{H}_f\tilde{\boldsymbol{V}}_n^{(0)}, \end{aligned} \tag{58}$$

with $\boldsymbol{R}_n, \boldsymbol{R}_f \in \mathbb{C}^{M\times(N-M)}$. Letting $\boldsymbol{U}_n = \boldsymbol{Q}_n$ and $\boldsymbol{U}_f = \boldsymbol{Q}_f$ the simultaneous triangularization of $\boldsymbol{H}_n$ and $\boldsymbol{H}_f$ is:

$$\begin{aligned} \boldsymbol{U}_n^{\mathrm{H}}\boldsymbol{H}_n\boldsymbol{V} &= \begin{bmatrix} \boldsymbol{R}_n^{(1)} & \boldsymbol{R}_n^{(0)} & \boldsymbol{O} \end{bmatrix}; \\ \boldsymbol{U}_f^{\mathrm{H}}\boldsymbol{H}_f\boldsymbol{V} &= \begin{bmatrix} \boldsymbol{O} & \boldsymbol{R}_f^{(1)} & \boldsymbol{R}_f^{(0)} \end{bmatrix}, \end{aligned} \tag{59}$$

with $\boldsymbol{R}_n^{(0)}$, $\boldsymbol{R}_f^{(0)} \in \mathbb{C}^{M\times(N-2M)}$. The two upper-triangular matrices are $\boldsymbol{R}_n^{(1)}$, $\boldsymbol{R}_f^{(1)} \in \mathbb{C}^{M\times M}$ with real-valued entries on their main diagonals.

From Eq. (2), we can easily obtain the power normalization factor as:

$$t_{ST} = \sqrt{2(N-M)}. \tag{60}$$

5. Numerical Results

For numerical simulations, we carry out the Monte Carlo simulation over 10^5 independent trials to verify the correctness of derived theoretical and approximate expressions of BER performance for NU and FU in GSVD-NOMA.

Suppose that the path-loss factor $\alpha = 2$, the distances $d_n = 1$, $d_f = 3$, the range of the power allocation ratio $\theta = 0.55 : 0.05 : 0.95$, and SNR $= \frac{P}{N_0} = 0 : 2.5 : 35$ dB. The simulation parameters are given in Table 2. We analyze the BER performance as a function of the transmission's SNR and θ using precoding schemes and different antenna configurations scenarios.

Tab. 2. Simulation parameters.

Path-loss factor	$\alpha = 2$
Distances	$d_n = 1, d_f = 3$
Power allocation ratio	$\theta = 0.55 : 0.05 : 0.95$
Number of transmit antennas	$N = 2, 3, 5, 7, 9$
Number of receive antennas	$N = 2, 4, 7$
Signal-to-noise ratio [dB]	SNR $= 0 : 2.5 : 35$
Number of trials	10^5

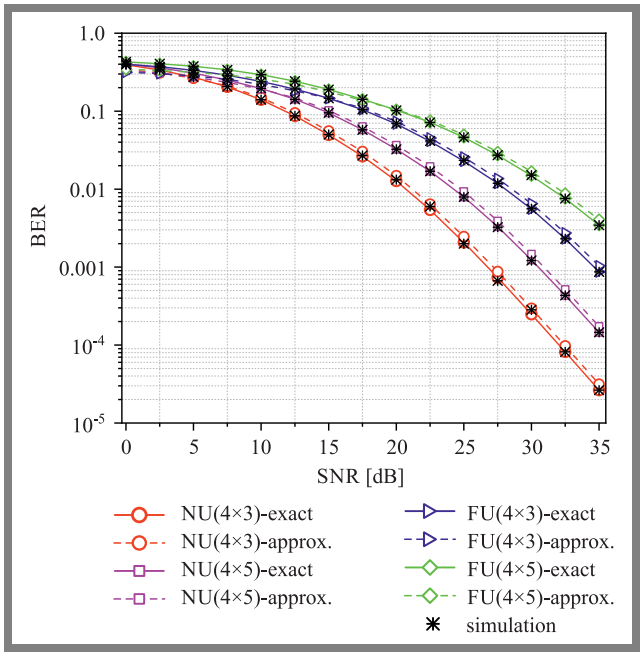

Fig. 3. BER performance of the near user and the far user, system performance curves are shown for two scenarios: $M \geqslant N$ (4×3) and $2M > N > M$ (4×5).

In Fig. 3, we calculate numerically the average BER performance of NU and FU, using Eqs. (30), (34), (39), and (40). Then, we validate the derived results by means of simulations under scenarios with the transmit and receive antenna configurations of 4×3 and 4×5. The examples under consideration correspond with the $N < 2M$ scenario. As a result, the relatively high number of trials makes the simulation results more precise and closer to the theoretical results. Moreover, the approximate derivations agree quite well with the actual analysis.

Figure 4 shows the comparison of ST and GSVD precoding for $M = 4$ and $N = 5$, in the $M < N < 2M$ scenario. One may clearly observe that BER performance of ST is superior to that of GSVD for NU and FU. For example, the loss in BER performance of two users for GSVD, when compared to ST precoding, equals approx. 5 dB for BER of 0.016. This is due to the fact that the values of fading channel coefficients α and β in the decomposition process performed by GSVD are lower than or equal to 1, as mentioned in Section 2.2, whereas the entries of ST diagonal matrices do not apply to all conditions. Therefore, if an antenna configuration is chosen that belongs to the $M < N < 2M$ case, ST precoding should be taken into consideration. Furthermore, this choice is completely relevant due to the outperformance of ST precoding in terms of the ergodic rate region compared to GSVD [12].

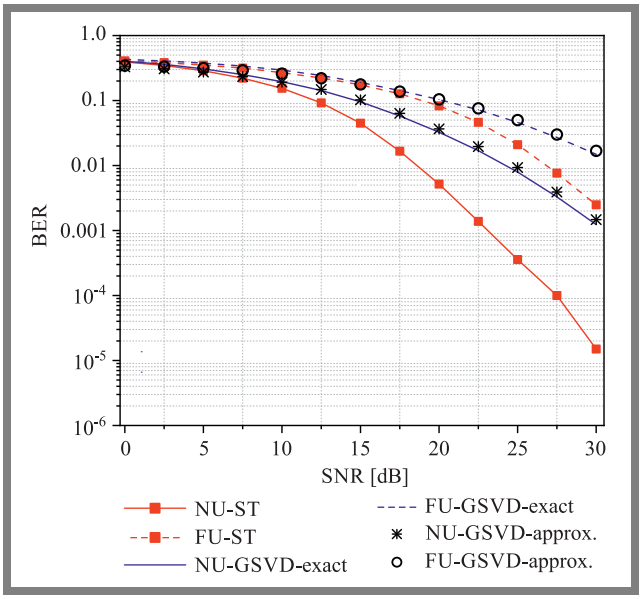

Fig. 4. Comparison of BER performance for ST based precoding and GSVD based precoding in the case of $2M > N > M$ for the (4×5) antenna configuration.

BER performance gain continues to be investigated in the case $N \geqslant 2M$ for ZF, BD, ST, and GSVD based precoding. Specifically, the number of transmit antennas is $N = 9$ and the users are equipped with $M = 4$ antennas, as shown in Fig. 5. We observed that BER performance of ST is dropped significantly compared to ZF, BD, and GSVD. This problem can be interpreted in such a way that based on triangular channel matrices in ST, the detection at the users' is undertaken in reverse order of the transmit signal vector. Moreover, for each subsequent

symbol, self-interference caused by the previously detected symbols needs to be eliminated. The self-interference cancellation process is usually imperfect, meaning that the system's performance is negatively impacted. As the above analysis shows, the power normalization factor is equal for GSVD and ZF. Moreover, after decomposition, the channel matrices have diagonal entries equal to 1. As a result, we can observe that BER performance gain of ZF is the same as in GSVD.

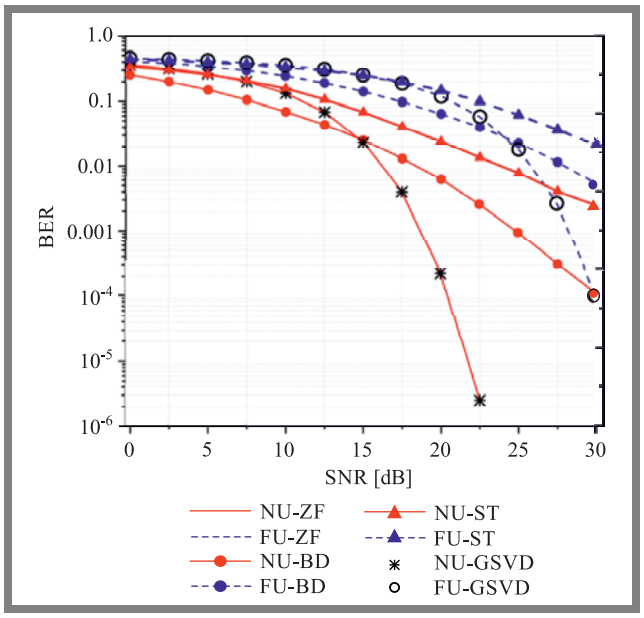

Fig. 5. BER performance comparison of ZF, BD, ST and GSVD based precoding schemes in the case of $N \geqslant 2M$ for antenna configurations.

In the low SNR regime, BD precoding performs better than ZF and GSVD. However, when transmit SNR is in the higher regime, ZF and GSVD precoding dramatically outperform BD precoding in terms of BER. In this antenna configuration, the parallel SISO channel in BD is dependent on small-scale fading elements, whereas in the case of ZF and GSVD the MIMO-NOMA channel is decomposed completely into the parallel AWGN channel. Therefore, when the average transmit power increases at the transmitters, BER performance of ZF and GSVD is considerably more superior. In scenarios in which the number of transmit antennas is greater than twice the number of receive antennas, GSVD and ZF based precoding schemes are chosen to improve the system's BER performance.

Figure 6 shows the result of a comparison of two detection techniques applied to NU, namely joint ML and symbol-level SIC (SL-SIC) with the ideal SIC. SL-SIC is the technique studied in [23]. In SL-SIC, NU demodulates the FU's signals and a hard decision is made, with channel coding not being performed. After that, NU regenerates the signals of FU and uses SIC to cancel them. With the ideal SIC, we assume that signals from FU are completely cancelled by NU. It is observed that joint ML significantly outperforms SL-SIC and offers almost the same performance as ideal SIC in terms of average BER. Performance of SL-SIC depends on power allocation coefficients θ and significantly degrades

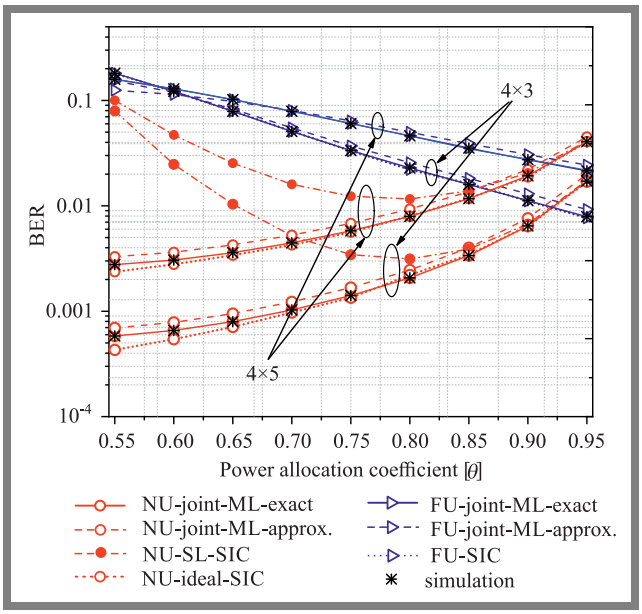

Fig. 6. BER performance comparison of the detector schemes: joint maximum-likelihood (ML), symbol-level SIC (SL-SIC) and ideal SIC under effect of power allocation coefficient θ.

BER for NU at small θ. A decrease in power allocated to FU symbols causes FU symbols to be detected erroneously by implementing SIC at NU. When θ is high, interference from NU may exert a weak impact on FU performance. However, at high θ, detection at NU is problematic due to low power allocation to NU. If the SL-SIC detector is applied to NU, power allocation should be considered.

In Fig. 7, we plot the average BER performance achieved by the joint ML detector versus transmit SNR when the number of transmit antennas changes, i.e. $N = 2, 3, 5, 7, 9$ and $M = 4$ for receive antennas. Based on the results shown, an increase in N decays BER performance of NU and FU be-

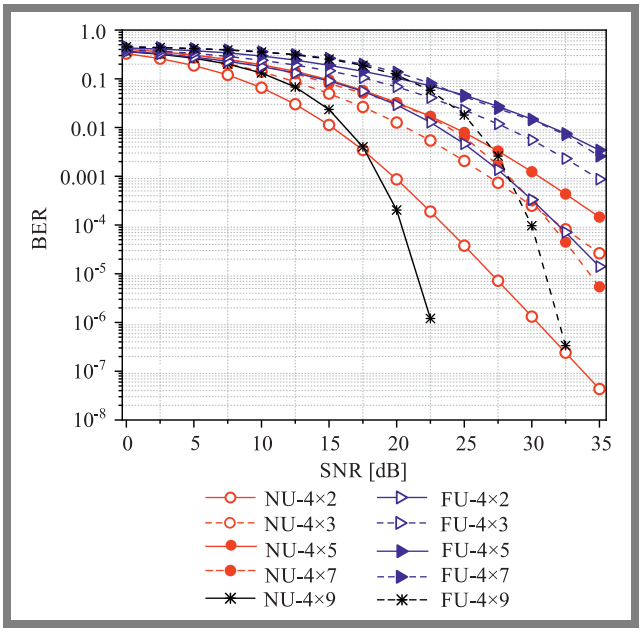

Fig. 7. BER performance of NU and FU for a varying number of transmit antennas N.

cause of the trade-off between diversity gain and multiplexing gain. More specifically, in the 4×2 and 4×3 scenarios, due to an increased number of parallel channels in the structure of GSVD, a decrease in performance is observed. Additionally, with an increase in the number of BS antennas in the $2M > N > M$ scenario, the number of parallel channels is constant (M), meaning that BER performance remains almost unchanged for NU and FU. Considering $N \geqslant 2M$, GSVD-MIMO channels decomposed into a number M of parallel Gaussian channels. It is shown in Eqs. (36) and (43), that the number of antennas does not affect the system's performance. In this case, BER outperforms other scenarios in the high SNR regime. Therefore, allowing a fixed number of receive antennas and an adjustable number of transmit antennas, in the low SNR regime, the transmitter should be equipped with a small number of aerials. However, assuming that the transmitted power at BS can be allocated at high levels permissively, BER performance is better when the number of transmit antennas satisfies the $N = 2M + 1$ condition.

Figure 8 shows the average BER performance of NU and FU when the number of receive antennas M increases. Here, the number of transmit antennas is modeled as $N = 5$. With an increase in the number of users' antennas M, it delivers better results in terms of BER performance. However, in the special case of $N \geqslant 2M (2 \times 5)$, the performance achieved is superior to all other solutions. So, if the number of transmit antennas is fixed, the number of receive antennas should satisfy the $N \geqslant 2M$ condition.

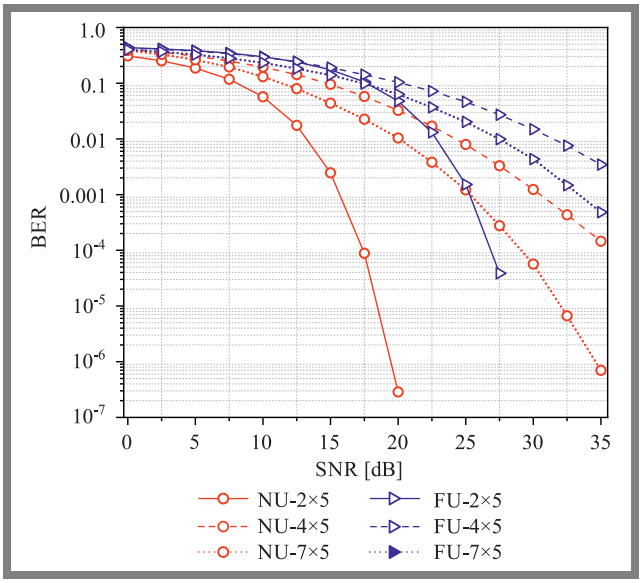

Fig. 8. BER performance of NU and FU for a varying number of receive antennas M.

6. Conclusions

In this paper, we consider the downlink MIMO-NOMA system with one base station and two users: near user and far user. The generalized singular value decomposition (GSVD) is applied to linear precoding and detection schemes at the BS and the end users respectively. Through mathematical analyses, we obtain the approximate expression BER performance for NU and FU with the joint-modulation at BS and the joint maximum-likelihood detector at each user. It was shown that the exact results, approximate expressions, and simulation results are completely consistent with each other. Furthermore, the joint maximum-likelihood detector is almost similar to the ideal SIC and significantly outperforms symbol-level SIC in terms of average BER performance. In comparison with other precoding schemes, GSVD offers the same performance as zero-forcing precoding, outperforming block diagonalization and simultaneous triangularization in terms of BER performance when the antenna configuration satisfies the $N \geqslant 2M$ condition. In $M \leqslant N < 2M$ scenarios, simultaneous triangularization precoding should be considered due to its superior performance not only in terms of BER, but also in terms of the ergodic achievable rate region [12]. However, fact that it may be applied to all antenna configurations in a significant advantage of GSVD. We also investigated average BER performance of GSVD with a varying number of antennas. It has been observed that the system's performance is superior when the number of antennas satisfies the $N \geqslant 2M$ condition. The analysis performed may be extended to downlink and uplink MIMO-NOMA with more than two users, where each user uses higher-order modulation level (M-ary modulation), to conduct further studies. Moreover, it is better to consider multiple performance metrics to arise a trade-off.

Acknowledgements

This research has been funded by University of Science, Vietnam National University, Ho Chi Minh City (VNU-HCM) under grant number DT-VT 2022-05.

Appendix A: proof of Lemma 1

The joint probability density function of unordered $W_i = \frac{\alpha_i^2}{\beta_i^2}$ in [14] is represented as:

$$f_{W_1,\ldots,W_q}(w_1,\ldots,w_q) = c_{J2} \prod_{i=1}^{q} \frac{w_i^{M-q}}{(1+w_i)^{N+q}} \times \prod_{i<j}^{q} (w_i - w_j)^2, \quad (61)$$

where c_{J2} is defined in Lemma 1. We can rewrite $W_i = \frac{X_i}{1-X_i}$, and the joint probability distribution of X_i will be expressed as:

$$\begin{aligned} F_{X_1,\ldots,X_q}(x_1,\ldots,x_q) &= \Pr(X_1 \leqslant x_1,\ldots,X_q \leqslant x_q) \\ &= \Pr\left(W_1 \leqslant \frac{x_1}{1-x_1},\ldots,W_q \leqslant \frac{x_q}{1-x_q}\right) \quad (62) \\ &= \int_0^{\frac{x_1}{1-x_1}} \cdots \int_0^{\frac{x_q}{1-x_q}} f_{W_1,\ldots,W_q}(w_1,\ldots w_q)\,\mathrm{d}w_1 \ldots \mathrm{d}w_q. \end{aligned}$$

Using the Leibniz integral rule, we obtain the joint probability density function of unordered X_i:

$$f_{X_1,\ldots,X_q}(x_1,\ldots,x_q) = \frac{\partial F_{X_1,\ldots,X_q}(x_1,\ldots,x_q)}{\partial x_1 \ldots \partial x_q} \tag{63}$$
$$= c_{J2}\prod_{i<j}^{q}\left(\frac{x_i}{1-x_i}-\frac{x_j}{1-x_j}\right)^2\prod_{i=1}^{q}x_i^{M-q}(1-x_i)^{N-M+2q-2}$$
$$= c_{J2}\prod_{i<j}^{q}(x_i-x_j)^2\prod_{i=1}^{q}x_i^r(1-x_i)^r.$$

The proof is completed.

Appendix B: proof of Theorem 1

Using exponential bound for Q-function in [31]:

$$Q(x) \simeq \frac{1}{12}\mathrm{e}^{-\frac{1}{2}x^2}+\frac{1}{4}\mathrm{e}^{-\frac{2}{3}x^2}. \tag{64}$$

Then, the error probability on the i-th parallel channel in Eq. (28) is:

$$P_i^{n1} \simeq \frac{1}{4}\left[\frac{1}{3}\mathrm{e}^{-\frac{a_1\rho\alpha_i^2}{2t_1^2}}+\frac{1}{6}\mathrm{e}^{-\frac{a_2\rho\alpha_i^2}{2t_1^2}}-\frac{1}{6}\mathrm{e}^{-\frac{a_3\rho\alpha_i^2}{2t_1^2}}\right. \tag{65}$$
$$\left.+\,\mathrm{e}^{-\frac{2a_1\rho\alpha_i^2}{3t_1^2}}+\frac{1}{2}\mathrm{e}^{-\frac{2a_2\rho\alpha_i^2}{3t_1^2}}-\frac{1}{2}\mathrm{e}^{-\frac{2a_3\rho\alpha_i^2}{3t_1^2}}\right]=\frac{1}{4}g(\alpha_i^2).$$

From joint-PDF function in Eq. (7), conducting the average BER in fading channel:

$$\overline{P}_{n1} \simeq \frac{c_{J1}}{4N}\int_0^1\ldots\int_0^1\left[\sum_{i=1}^{N}g(x_i)\right]\left[\prod_{i=1}^{N}x_i^{M-N}(1-x_i)^{M-N}\right.$$
$$\left.\times\prod_{i<j}^{N}(x_i-x_j)^2\right]\mathrm{d}x_1\ldots\mathrm{d}x_N. \tag{66}$$

Consider term:

$$\prod_{i<j}^{N}(x_j-x_i) = \det\begin{pmatrix}1 & 1 & \ldots & 1\\ x_1 & x_2 & \ldots & x_N\\ x_1^2 & x_2^2 & \ldots & x_N^2\\ \vdots & \vdots & \ddots & \vdots\\ x_1^{N-1} & x_2^{N-1} & \ldots & x_N^{N-1}\end{pmatrix}$$
$$= \det\boldsymbol{V}, \tag{67}$$

where $\boldsymbol{V}$ is the Vandermonde matrix. Applying the Leibniz formula for the determinant of $\boldsymbol{V}$:

$$\prod_{i<j}^{N}(x_j-x_i)=\sum_{\sigma\in\boldsymbol{S}_N}\mathrm{sgn}(\sigma)\prod_{i=1}^{N}x_i^{\sigma(i)-1}, \tag{68}$$

$\boldsymbol{S}_N$ is the set of the permutations of $\{1,2,\ldots,N\}$. Carrying out algebraic manipulation, we get:

$$\prod_{i<j}^{N}(x_j-x_i)^2 = \sum_{\sigma\in\boldsymbol{S}_N}\prod_{i=1}^{N}x_i^{2\sigma(i)-2}+2\sum_{\sigma_1,\sigma_2\in\boldsymbol{S}_N}\mathrm{sgn}(\sigma_1)$$
$$\times\,\mathrm{sgn}(\sigma_2)\prod_{i=1}^{N}x_i^{\sigma_1(i)+\sigma_2(i)-2}. \tag{69}$$

Substituting Eq. (69) into Eq. (66):

$$\overline{P}_{n1} \simeq \frac{c_{J1}}{4N}\int_{[0,1]^N}\left[\sum_{i=1}^{N}g(x_i)\right]\left[\prod_{i=1}^{N}x_i^{M-N}(1-x_i)^{M-N}\right]$$
$$\times\left[\sum_{\sigma\in\boldsymbol{S}_N}\prod_{i=1}^{N}x_i^{2\sigma(i)-2}\right]\mathrm{d}x_1\ldots\mathrm{d}x_N \tag{70}$$
$$+\frac{c_{J1}}{2N}\int_{[0,1]^N}\left[\sum_{i=1}^{N}g(x_i)\right]\left[\prod_{i=1}^{N}x_i^{M-N}(1-x_i)^{M-N}\right]$$
$$\times\left[\sum_{\sigma_1,\sigma_2\in\boldsymbol{S}_N}\mathrm{sgn}(\sigma_1)\mathrm{sgn}(\sigma_2)\times\prod_{i=1}^{N}x_i^{\sigma_1(i)+\sigma_2(i)-2}\right]\mathrm{d}x_1\ldots\mathrm{d}x_N,$$

where $\int_{[0,1]^N}=\int_0^1\ldots\int_0^1$ for brevity. By performing algebraic operations, we obtain the approximate expression of $\overline{P}_{n1}$ as:

$$\overline{P}_{n1} \simeq \frac{c_{J1}}{4N}\sum_{\sigma\in\boldsymbol{S}_N}\sum_{j=1}^{N}\int_0^1 g(x_j)x_j^{p_{j1}}(1-x_j)^{q_1}\,\mathrm{d}x_j$$
$$\times\prod_{\substack{i=1\\ i\neq j}}^{N}\int_0^1 x_i^{p_{i1}}(1-x_i)^{q_1}\,\mathrm{d}x_i+\frac{c_{J1}}{2N}\sum_{\sigma_1,\sigma_2\in\boldsymbol{S}_N}\mathrm{sgn}(\sigma_1)$$
$$\times\,\mathrm{sgn}(\sigma_2)\sum_{j=1}^{N}\int_0^1 g(x_j)x_j^{p_{j2}}(1-x_j)^{q_1}\,\mathrm{d}x_j$$
$$\times\prod_{\substack{i=1\\ i\neq j}}^{N}\int_0^1 x_i^{p_{i2}}(1-x_i)^{q_1}\,\mathrm{d}x_i. \tag{71}$$

Considering the following integral, $I_1=\int_0^1 t^p(1-t)^q\,\mathrm{d}t$ and $I_2=\int_0^1\mathrm{e}^{-at}t^p(1-t)^q\,\mathrm{d}t$ and applying two equations Eqs. (8.380) and (8.384) from [29], we can rewrite I_1 as:

$$I_1=\mathrm{B}(p+1,q+1), \tag{72}$$

$\mathrm{B}(x,y)$ is the beta function. Using Eq. (3.383) from [29] I_2 is:

$$I_2=\mathrm{B}(p+1,q+1)_1F_1(p+1;p+q+2;-a), \tag{73}$$

where ${}_1F_1(a;b;z)$ is the generalized hypergeometric function. Applying Eqs. (72) and (73) to (71), the proof is completed.

References

[1] L. Dai, *et al.*, "Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends", *IEEE Communications Magazine*, vol. 53, no. 9, pp. 74–81, 2015 (DOI: 10.1109/MCOM.2015.7263349).

[2] Y. Liu, Z. Qin, M. Elkashlan, Z. Ding, A. Nallanathan, and L. Hanzo, "Nonorthogonal Multiple Access for 5G and Beyond", *Proceedings of the IEEE*, vol. 105, no. 12, pp. 2347–2381, 2017 (DOI: 10.1109/JPROC.2017.2768666).

[3] Y. Saito, *et al.*, "Non-Orthogonal Multiple Access (NOMA) for Cellular Future Radio Access", in *2013 IEEE 77th Vehicular Technology Conference (VTC Spring)*, pp. 1–5, 2013 (DOI: 10.1109/VTCSpring.2013.6692652).

[4] K. Higuchi and A. Benjebbour, "Non-orthogonal Multiple Access (NOMA) with Successive Interference Cancellation for Future Radio Access", *IEICE Transactions on Communications*, vol. E98.B, pp. 403–414, 2015 (DOI: 10.1587/transcom.E98.B.403).

[5] Q. Sun, S. Han, I. C-L, and Z. Pan, "On the Ergodic Capacity of MIMO NOMA Systems", *IEEE Wireless Communications Letters*, vol. 4, no. 4, pp. 405–408, 2015 (DOI: 10.1109/LWC.2015.2426709).

[6] S. Ali, E. Hossain, and D.I. Kim, "Non-Orthogonal Multiple Access (NOMA) for Downlink Multiuser MIMO Systems: User Clustering, Beamforming, and Power Allocation", *IEEE Access*, vol. 5, pp. 565–577, 2017 (DOI: 10.1109/ACCESS.2016.2646183).

[7] Z. Ding, R. Schober, and H.V. Poor, "A General MIMO Framework for NOMA Downlink and Uplink Transmission Based on Signal Alignment", *IEEE Transactions on Wireless Communications*, vol. 15, no. 6, pp. 4438–4454, 2016 (DOI: 10.1109/TWC.2016.2542066).

[8] H. Weingarten, Y. Steinberg, and S.S. Shamai, "The Capacity Region of the Gaussian Multiple-Input Multiple-Output Broadcast Channel", *IEEE Transactions on Information Theory*, vol. 52, no. 9, pp. 3936–3964, 2006 (DOI: 10.1109/TIT.2006.880064).

[9] Z. Chen and X. Dai, "MED Precoding for Multiuser MIMO-NOMA Downlink Transmission", *IEEE Transactions on Vehicular Technology*, vol. 66, no. 6, pp. 5501–5505, 2017 (DOI: 10.1109/TVT.2016.2627218).

[10] A. Krishnamoorthy, Z. Ding, and R. Schober, "Precoder Design and Statistical Power Allocation for MIMO-NOMA via User-Assisted Simultaneous Diagonalization", *IEEE Transactions on Communications*, vol. 69, no. 2, pp. 929–945, 2021 (DOI: 10.1109/TCOMM.2020.3036453).

[11] D. Senaratne and C. Tellambura, "GSVD Beamforming for Two-User MIMO Downlink Channel", *IEEE Transactions on Vehicular Technology*, vol. 62, no. 6, pp. 2596–2606, 2013 (DOI: 10.1109/TVT.2013.2241091).

[12] A. Krishnamoorthy, M. Huang, and R. Schober, "Precoder Design and Power Allocation for Downlink MIMO-NOMA via Simultaneous Triangularization", in *2021 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–6, 2021 (DOI: 10.1109/WCNC49053.2021.9417424).

[13] Z. Chen, Z. Ding, X. Dai, and R. Schober, "Asymptotic Performance Analysis of GSVD-NOMA Systems with a Large-Scale Antenna Array", *IEEE Transactions on Wireless Communications*, vol. 18, no. 1, pp. 575–590, 2019 (DOI: 10.1109/TWC.2018.2883102).

[14] Z. Chen, Z. Ding, and X. Dai, "On the Distribution of the Squared Generalized Singular Values and Its Applications", *IEEE Transactions on Vehicular Technology*, vol. 68, no. 1, pp. 1030–1034, 2019 (DOI: 10.1109/TVT.2018.2885122).

[15] M.F. Hanif and Z. Ding, "Robust Power Allocation in MIMO-NOMA Systems", *IEEE Wireless Communications Letters*, vol. 8, no. 6, pp. 1541–1545, 2019 (DOI: 10.1109/LWC.2019.2926277).

[16] C. Rao, Z. Ding, and X. Dai, "The Distribution Characteristics of Ordered GSVD Singular Values and its Applications in MIMO-NOMA", *IEEE Communications Letters*, vol. 24, no. 12, pp. 2719–2722, 2020 (DOI: 10.1109/LCOMM.2020.3017796).

[17] C. Rao, Z. Ding, and X. Dai, "GSVD-Based MIMO-NOMA Security Transmission", *IEEE Wireless Communications Letters*, vol. 10, no. 7, pp. 1484–1487, 2021 (DOI: 10.1109/LWC.2021.3071365).

[18] Y. Qi and M. Vaezi, "Secure Transmission in MIMO-NOMA Networks", *IEEE Communications Letters*, vol. 24, no. 12, pp. 2696–2700, 2020 (DOI: 10.1109/LCOMM.2020.3016999).

[19] X. Wang, F. Labeau, and L. Mei, "Closed-Form BER Expressions of QPSK Constellation for Uplink Non-Orthogonal Multiple Access", *IEEE Communications Letters*, vol. 21, no. 10, pp. 2242–2245, 2017 (DOI: 10.1109/LCOMM.2017.2720583).

[20] F. Kara and H. Kaya, "BER Performances of Downlink and Uplink NOMA in the Presence of SIC Errors over Fading Channels", *IET Communications*, vol. 12, no. 15, pp. 1834–1844, 2018 (DOI: 10.1049/iet-com.2018.5278).

[21] T. Assaf, A. Al-Dweik, M.E. Moursi, and H. Zeineldin, "Exact BER Performance Analysis for Downlink NOMA Systems Over Nakagamim Fading Channels", *IEEE Access*, vol. 7, pp. 134539–134555, 2019 (DOI: 10.1109/ACCESS.2019.2942113).

[22] J.S. Yeom, H.S. Jang, K.S. Ko, and B.C. Jung, "BER Performance of Uplink NOMA With Joint Maximum-Likelihood Detector", *IEEE Transactions on Vehicular Technology*, vol. 68, no. 10, pp. 10295–10300, 2019 (DOI: 10.1109/TVT.2019.2933253).

[23] C. Yan, *et al.*, "Receiver Design for Downlink Non-Orthogonal Multiple Access (NOMA)", in *2015 IEEE 81st Vehicular Technology Conference (VTC Spring)*, pp. 1–6, 2015 (DOI: 10.1109/VTCSpring.2015.7146043).

[24] —, "Wireless Technology Evolution Towards 5G: 3GPP release 13 to release 15 and beyond", 2017, (https://www.5gamericas.org/wireless-technology-evolution-towards-5g-3gpp-release-13-to-release-15-and-beyond/).

[25] C.F. Van Loan, "A General Matrix Eigenvalue Algorithm", *SIAM Journal on Numerical Analysis*, vol. 12, no. 6, pp. 819–834, 1975 (https://www.jstor.org/stable/2156413).

[26] A. Edelman and B.D. Sutton, "The Beta-Jacobi Matrix Model, the CS Decomposition, and Generalized Singular Value Problems", *Foundations of Computational Mathematics*, vol. 8, no. 2, pp. 259–285, 2008 (DOI: 10.1007/s10208-006-0215-9).

[27] J. Tiefeng, "Limit theorems for beta-Jacobi ensembles", *Bernoulli*, vol. 19, no. 3, pp. 1028–1046, 2013 (DOI: 10.3150/12-BEJ495, https://projecteuclid.org/journalArticle/Download?urlId=10.3150%2F12-BEJ495).

[28] A. Goldsmith, "Wireless Communications", *Cambridge University Press*, 2005 (DOI: 10.1017/CBO9780511841224).

[29] D. Zwillinger and A. Jeffrey, *Table of integrals, series, and products*, 7th ed. Elsevier, 2007 (ISBN 978-0-12-373637-6).

[30] F.W.L. Oliver, D.W. Lozier, R.F. Boisvert, and C.W Clark, *NIST Handbook of Mathematical Functions*, 2010 (https://assets.cambridge.org/97805211/92255/copyright/9780521192255_copyright_info.pdf).

[31] M. Chiani, D. Dardari, and M.K. Simon, "New exponential bounds and approximations for the computation of error probability in fading channels", *IEEE Transactions on Wireless Communications*, vol. 2, no. 4, pp. 840–845, 2003 (DOI: 10.1109/TWC.2003.814350).

Ngo Thanh Hai received his B.Sc. from the University of Science, Vietnam National University, Ho Chi Minh City (VNU-HCM) and M.Sc. in Electronics and Telecommunications from Posts and Telecommunications Institute of Technology. Since 2019, he has been working as a researcher at the Faculty of Electronics and Telecommunications, University of Science, VNU-HCM. His research activities focus on non-orthogonal multiple access (NOMA), multiple-input multiple-output (MIMO) and covert wireless communications.
E-mail: ngohai@hcmus.edu.vn
Department of Telecommunications and Networks, University of Science, VNU-HCM, District 5, Ho Chi Minh City, Vietnam

Dang Le Khoa received his B.E. and Ph.D. degrees in Radio Physics and Electronics from the University of Science, Vietnam National University, Ho Chi Minh City (VNU-HCM). He is the head of the Telecommunications and Networks Department, University of Science, VNU-HCM. His current research interests are in the areas of wireless communications and digital signal processing for telecommunication.
E-mail: dlkhoa@hcmus.edu.vn
Department of Telecommunications and Networks, University of Science, VNU-HCM, District 5, Ho Chi Minh City, Vietnam