Comparative Analysis of NOMA and OMA Schemes: GSVD-based NOMA Systems and the Role of Mobile Edge Computing

Yunus Dursun

University of Manchester, Manchester, United Kingdom

https://doi.org/10.26636/jtit.2023.3.1354

Abstract — This paper presents a comprehensive study that examines the fundamental concept of the non-orthogonal multiple access (NOMA) scheme and provides its detailed comparison with the orthogonal multiple access (OMA) technique. Furthermore, the paper explores the application of the generalized singular value decomposition (GSVD) method in conjunction with NOMA, accompanied by a detailed review of GSVD-based NOMA systems. This study also introduces the concept of mobile edge computing (MEC) and extensively discusses its key parameters. Furthermore, a comprehensive analysis of NOMA MEC is presented, shedding light on its potential advantages and challenges. The aims of this study are to provide a comprehensive understanding of the aforementioned topics and contribute to the advancement of MIMO-NOMA systems.

Keywords — generalized singular value decomposition, MIMO, mobile edge computing, non-orthogonal multiple-access

1. Non-orthogonal Multiple Access

Non-orthogonal multiple access (NOMA) is a multiple access technique that allows multiple users to simultaneously share the same frequency, time, or code resources to communicate with a base station or an access point. Several NOMA schemes are distinguished in the literature [1]. Some of their well-known varieties are summarized below. The sparse code multiple access (SCMA) technique enables multiple users to communicate with a base station simultaneously by utilizing separate sparse codes assigned to each user based on a multidimensional codebook [2]. In the pattern division multiple access (PDMA) approach, the available bandwidth is divided into multiple non-overlapping frequency patterns or slots. Thus, the users rely on a unique pattern or slot to modulate their signals [3]. Resource spread multiple access (RSMA) employs a distinct spreading sequence for users to disperse their data over the frequency band. The receiver then reverses the spreading process by applying the identical spreading sequence to recover the user's information [4]. Multi-user shared access (MUSA) is based on code domain multiplexing, where symbols are multiplexed using the same spreading code. These symbols are transmitted over an orthogonal channel, such as a sub-carrier, as in OFDMA. At the receiver end, SIC decodes the received symbols [5]. Interleave-grid multiple access (IGMA) is another technique in which the user data is segmented and interleaved based on a specific pattern, creating a grid-like structure that helps minimize user interference and improves the overall spectral efficiency of the system [6]. Rate-splitting multiple access is also evoking significant interest within the research community. In rate-splitting multiple access, users partition their data into shared and exclusive components [7]. The shared components of each user are aggregated and modulated jointly, while the unique components of each user are modulated separately. This results in a transmitted signal containing shared and unique components for all users. Both users initially decode the shared component at the receiver, treating any interference from the unique signals as noise. Both users use SIC to decode their exclusive signals in the subsequent stage.

In 3GPP Release 13, the standardization of power domain NOMA (PD-NOMA), known as multi-user superposition transmission (MUST), has been introduced for a broadcast channel. In PD-NOMA, multiple users use different power levels to share the same time and frequency resources. At the transmitter, superposition coding is employed to multiplex the users, while the receiver uses successive interference cancellation to decode the superposed signals [8]. PD-NOMA is considered a promising technique for 5G and beyond wireless communication systems, as it is capable of significantly increasing spectral efficiency and supporting multiple users with diverse communication requirements. PD-NOMA can also improve user fairness and energy efficiency, enabling users with weaker channel conditions to share the same resources with stronger users, without sacrificing their quality of service [9].

1.1. Overview of Power-domain NOMA

This section introduces the basic concepts of PD-NOMA for downlink and uplink networks. Additionally, we analyze and compare the sum rate and signal-to-interference-plus-noise ratio (SINR) of NOMA and OMA.

1.2. Downlink NOMA Network

Figure 1 illustrates a downlink NOMA scheme consisting of a base station or an access point and K receivers, where the BS broadcasts a superposed signal to all the receivers. The BS combines complex-valued symbols with superpo-

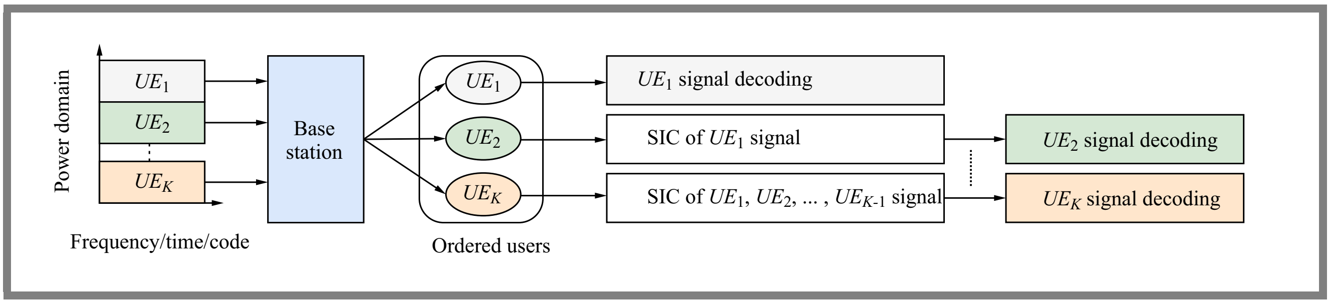

Fig. 1. Basic concept of a downlink NOMA [9].

sition coding (SC), and the receivers employ the successive interference cancellation (SIC) technique to decode their respective signals. Each receiver, except for the weakest one or receivers without SIC capabilities, performs the SIC process at the receiver. Even though there is a tendency for the SIC process as follows: the users first extract the strongest signal from the combined signal and then subtract it to eliminate interference from the remaining signals, and the SIC process is repeated until the receiver's signal is decoded. This strategy may not be optimal. Ding *et al.* showed that dynamic decoding orders relying on users' QoS and CSI-based SIC orders might improve the system's performance [10]. To simplify the analysis, we consider a downlink NOMA system with a base station and two users to derive the SINRs and sum rates. Additionally, we assume that the base station and users are equipped with a single antenna and the system bandwidth B is one. The information-bearing signals, x_N for the near user UE_1 and x_F for the far user UE_2, are superimposed at the transmitter as follows:

$$x = \sqrt{P_N} x_N + \sqrt{P_F} x_F \ , \tag{1}$$

where P_N and P_F denote the transmission power allocation coefficients for the near and far users, respectively. P_{tot} represents the total transmit power which equals the sum of P_N and P_F. The received signal at the receivers are:

$$y_i = h_i x + n_i, \ i \in \{N, F\} \ , \tag{2}$$

where h_i denotes the channel coefficient between the BS and user UE_i and n_i represents the additive white Gaussian noise (AWGN) with zero mean and σ_i^2 variance for UE_i. Let's assume the users are ordered using the CSI-based method at the receiver and the near user has a strong signal than of the far user, i.e., $\frac{|h_N|^2}{\sigma_N^2} \geqslant \frac{|h_F|^2}{\sigma_F^2}$. Therefore, the SINR expression of the near user and far user are given by:

$$SNR_N = \frac{P_N |h_N|^2}{\sigma_N^2} \ , \tag{3}$$

$$SINR_F = \frac{P_F |h_F|^2}{P_N |h_N|^2 + \sigma_F^2} \ . \tag{4}$$

Accordingly, the data rate for the near user and far user can be written as follows:

$$R_N = \log_2 \left(1 + \frac{P_N |h_N|^2}{\sigma_N^2} \right) \ , \tag{5}$$

$$R_F = \log_2 \left(1 + \frac{P_F |h_F|^2}{P_N |h_N|^2 + \sigma_F^2} \right) \ . \tag{6}$$

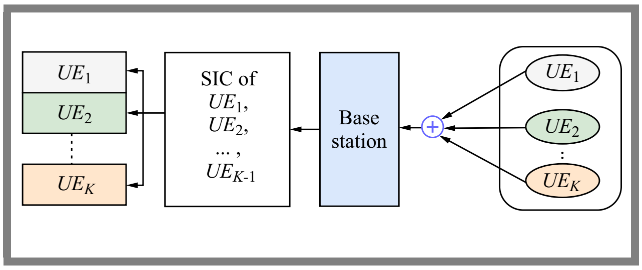

Fig. 2. Basic concept of uplink NOMA [9].

1.3. Uplink NOMA Network

As illustrated in Fig. 2, an uplink NOMA system allows K users to simultaneously transmit their data to the BS using the same spectrum. The base station employs SIC to decode the signals from different users. We again assume that the near user has better channel gain than the far user, i.e. $\frac{|h_N|^2}{\sigma_N^2} \geqslant \frac{|h_F|^2}{\sigma_F^2}$. The received signal at the receiver is:

$$y = h_N x_N + h_F x_F + n_B, \tag{7}$$

where n_B is an AWGN with zero mean and σ_N^2 variance at the receiver. If the BS decodes the received signal in descending order, the data rate for the near and far users are:

$$R_N = \log_2 \left(1 + \frac{P_N |h_N|^2}{P_F |h_F|^2 + \sigma_B^2} \right) \ , \tag{8}$$

$$R_F = \log_2 \left(1 + \frac{P_F |h_F|^2}{\sigma_B^2} \right) \ . \tag{9}$$

On the other hand, if the BS decodes the received signal in ascending order, the data rate for the near and far users becomes:

$$R_N = \log_2 \left(1 + \frac{P_N |h_N|^2}{\sigma_B^2} \right) \ , \tag{10}$$

$$R_F = \log_2 \left(1 + \frac{P_F |h_F|^2}{P_N |h_N|^2 + \sigma_B^2} \right) \ . \tag{11}$$

It is worth mentioning that in each case, the sum rate for the users is the same as given in:

$$R_N + R_F = \log_2 \left(\frac{P_N |h_N|^2 + P_F |h_F|^2 + \sigma_B^2}{\sigma_B^2} \right) \ . \tag{12}$$

In other words, the sum rate in the uplink NOMA does not depend on the order of SIC, assuming no error propagation occurs in the SIC process. However, according to Benjebbour [8], it is more practical to perform SIC in the descending order of channel quality levels.

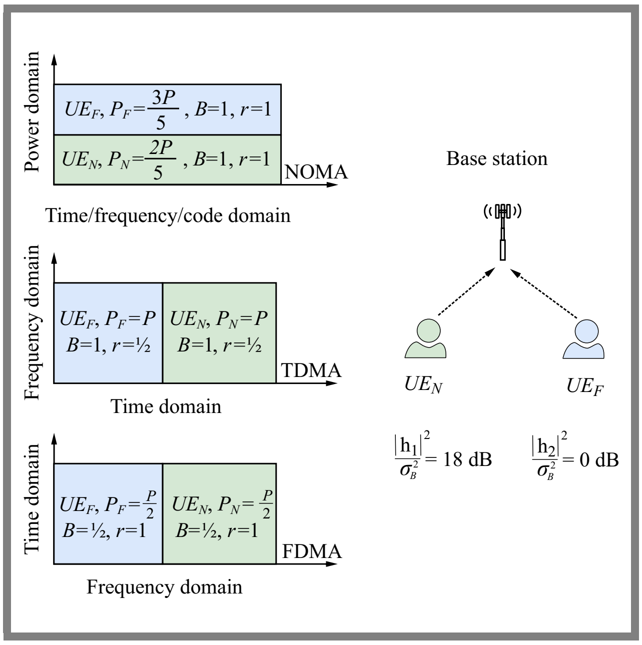

Fig. 3. Sum rate comparison between NOMA, TDMA, and FDMA networks.

1.4. Sum Rate Comparison Between NOMA, TDMA, and FDMA Networks

This section compares NOMA with TDMA and FDMA schemes for an uplink scenario, as illustrated in Fig. 3. The data rates of the near and far users in the NOMA, FDMA, and TDMA systems are given as follows [11]:

$$NOMA = \begin{cases} R_N^{NOMA} = B\log_2(1+\frac{P_N|h_N|^2}{B\sigma_B^2}) \\ R_F^{NOMA} = B\log_2(1+\frac{P_F|h_F|^2}{P_N|h_N|^2+B\sigma_B^2}) \\ 0 \leqslant P_N,\ P_F \leqslant P \end{cases}$$
$$TDMA = \begin{cases} R_N^{TDMA} = B(1-\tau)\log_2(1+\frac{P_N|h_N|^2}{B\sigma_B^2}) \\ R_F^{TDMA} = B\tau\log_2(1+\frac{P_F|h_F|^2}{B\sigma_B^2}) \\ 0 \leqslant \tau \leqslant 1 \end{cases}$$
$$FDMA = \begin{cases} R_N^{FDMA} = B(1-\omega)\log_2(1+\frac{P_N|h_N|^2}{B(1-\omega)\sigma_B^2}) \\ R_F^{FDMA} = B\omega\log_2(1+\frac{P_F|h_F|^2}{B\omega\sigma_B^2}) \\ 0 \leqslant \omega \leqslant 1. \end{cases} \tag{13}$$

Assuming two uplink users UE_N and UE_F with channel gains of $\frac{h_N}{\sigma_B^2} = 18$ dB and $\frac{h_F}{\sigma_B^2} = 0$ dB, respectively, the total power is the same in all schemes such that $P_N + P_F = P$, where P is the maximum transmit power [8]. In the TDMA scheme, the users are allocated equal time slots, i.e., $\tau = 0.5$. The data rates for the near and far users are $R_N^{TDMA} = 3.0011$ bps and $R_F^{TDMA} = 0.5$ bps, respectively. In the FDMA scheme, the bandwidth is split equally between the users, i.e., $\omega = 0.5$, and the resulting data rates are $R_N^{FDMA} = 3.0011$ bps and $R_F^{FDMA} = 0.5$ bps for the near and far users, respectively. In the NOMA case, the power is split between the users by δ, with two out of five for the near user and three out of five for the far user. Thus, the data rates for the near and far users are $R_N^{NOMA} = 4.0682$ bps and $R_F^{NOMA} = 0.6781$ bps. The total sum rates achieved by the TDMA, FDMA, and NOMA schemes are 3.5011 bps, 3.5011 bps, and 4.7463 bps, respectively. Based on this example, it can be concluded that NOMA offers a significant advantage over OMA schemes in terms of spectral efficiency, resulting in 35.57% higher sum data rates. Figure 4 emphasizes the influence of the power allocation coefficient on the sum rate performance. In Fig. 5, a comparison is presented between the sum rates of NOMA, TDMA, and FDMA, with the power allocation coefficient δ, the bandwidth allocation coefficient ω, and the time allocation coefficient τ set to 0.3. As the channel gain difference increases, it is observed that NOMA outperforms OMA schemes, exhibiting greater improvement in performance.

2. Enhancing MIMO-NOMA Systems Through GSVD

Generalized singular value decomposition (GSVD) is a powerful matrix factorization technique that extends the standard singular value decomposition (SVD) to accommodate rectangular matrices of potentially different dimensions. This technique encompasses two primary types: real-valued GSVD and complex-valued GSVD. Real GSVD is utilized for real-valued matrices, while complex GSVD is tailored for complex-valued matrices. Various algorithms, such as Van Loan's, which was first introduced in 1976, as well as Paige and Sounders' algorithms [12], facilitate the computation of GSVD. By decomposing matrices into their singular components, GSVD finds a widespread application across diverse domains. It is employed in signal processing tasks, such as adaptive filtering, blind source separation, and channel estimation, as well as in analyzing biological data in bioinformatics [13], [14]. In wireless communication, GSVD decomposes MIMO channels into orthogonal SISO channels. In other words, the main use case of the GSVD is beamforming design. Table 1 lists

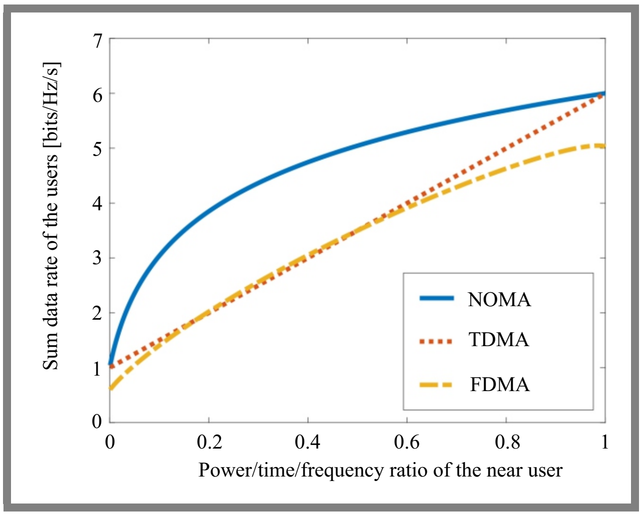

Fig. 4. Sum data rate comparison for different resource ratios for the near user.

some studies that employ GSVD to solve some problems in wireless communication.

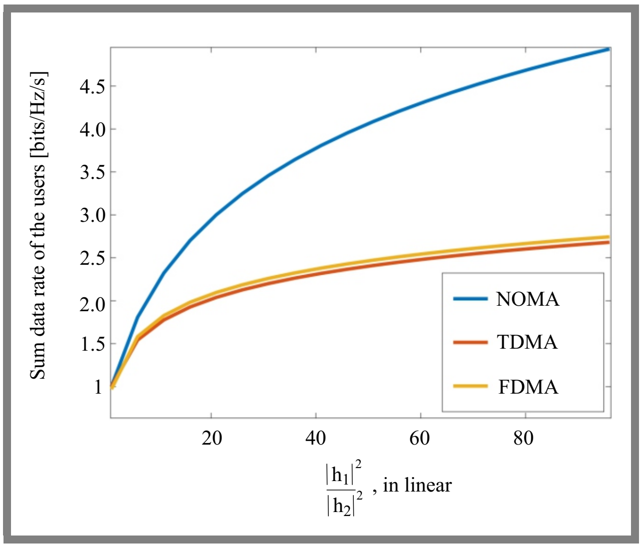

Fig. 5. Sum data rate comparison under different SNR rates.

2.1. Definition of GSVD

Let us consider two matrices, $\mathbf{H_1} \in \mathbf{C}^{m\times n}$ and $\mathbf{H_2} \in \mathbf{C}^{m\times n}$. By applying the GSVD method, we can decompose these matrices into three components: a unitary matrix, a non-singular matrix, and a non-negative singular matrix. The decomposition can be expressed as follows [15]:

$$\mathbf{\Sigma_1} = \mathbf{U}\mathbf{H_1}\mathbf{Q} \quad \text{and} \quad \mathbf{\Sigma_2} = \mathbf{V}\mathbf{H_2}\mathbf{Q}, \tag{14}$$

where the matrices $\mathbf{U} \in \mathbf{C}^{m\times m}$ and $\mathbf{V} \in \mathbf{C}^{m\times m}$ are unitary, while $\mathbf{Q} \in \mathbf{C}^{n\times n}$ is non-singular, and $\mathbf{\Sigma_1} \in \mathbf{C}^{m\times n}$ and $\mathbf{\Sigma_2} \in \mathbf{C}^{m\times n}$ are diagonal matrices with non-negative elements. The dimension of the matrices form $\mathbf{\Sigma_1}$ and $\mathbf{\Sigma_2}$ as follows:

– If $m \geqslant n$, then $\mathbf{\Sigma_1} = \begin{pmatrix} \mathbf{0}_{(m-n)\times n} \\ \mathbf{S}_1 \end{pmatrix}$ and $\mathbf{\Sigma_2} = \begin{pmatrix} \mathbf{S}_2 \\ \mathbf{0}_{(m-n)\times n} \end{pmatrix}$.

– If $m \leqslant n \leqslant 2m$, $r = n - m$ and $q = 2m - n$, then $\mathbf{\Sigma_1} = \begin{pmatrix} \mathbf{I}_r & \mathbf{O}_{r\times q} & \mathbf{O}_{r\times r} \\ \mathbf{O}_{q\times r} & \mathbf{S}_1 & \mathbf{O}_{q\times r} \end{pmatrix}$ and $\mathbf{\Sigma_2} = \begin{pmatrix} \mathbf{O}_{q\times r} & \mathbf{S}_2 & \mathbf{O}_{q\times r} \\ \mathbf{O}_{r\times r} & \mathbf{O}_{r\times q} & \mathbf{I}_r \end{pmatrix}$.

– If $2m \geqslant n$, then $\mathbf{\Sigma_1} = \begin{pmatrix} \mathbf{I}_m & \mathbf{O}_{m\times(n-m)} \end{pmatrix}$ and $\mathbf{\Sigma_2} = \begin{pmatrix} \mathbf{O}_{m\times(n-m)} & \mathbf{I}_m \end{pmatrix}$,

where $\mathbf{O}$ and $\mathbf{I}$ represent the zero and identity matrices, respectively. Moreover, $\mathbf{S}_1$ and $\mathbf{S}_2$ are non-negative diagonal matrices, with elements between zero and one. Notably, the elements of $\mathbf{S}_1$ are sorted in descending order, while those of $\mathbf{S}_2$ are sorted in ascending order.

2.2. Application of GSVD in MIMO-NOMA

We consider a base station (BS) with n antennas communicating with two downlink users, each equipped with m antennas. The channels between the BS and the users can be represented by $\mathbf{G}_i = \frac{\mathbf{H}_i}{\sqrt{d_i^\tau}}$, where $\mathbf{H}_i$ denotes the small-scale fading coefficients. The near user, denoted as UE_N, and the far user, denoted as UE_F, are sorted based on their large-scale fading element $\sqrt{d_i^\tau}$. Here, d represents the distance of the i-th user, and τ represents their path loss component. The received signals at the receivers are:

$$\mathbf{y}_N = \frac{\mathbf{H}_N\mathbf{x}}{\sqrt{d_N^\tau}} + \mathbf{n}_N\ , \tag{15}$$

$$\mathbf{y}_F = \frac{\mathbf{H}_F\mathbf{x}}{\sqrt{d_F^\tau}} + \mathbf{n}_F\ . \tag{16}$$

The noise at the i-th receiver, $\mathbf{n}_i,\ i \in N, F$, modeled by the additive white Gaussian noise, is given by mutually independent and identically distributed elements with zero mean and variance σ_i. The transmitted signal, denoted as $\mathbf{x} \in \mathbf{C}^{n\times 1}$, is subject to interference mitigation techniques using precoding and decoding matrices $\mathbf{P}_b \in \mathbf{C}^{n\times n}$ and $\mathbf{D}_i \in \mathbf{C}^{m\times m}$, respectively. The decomposition of channels, as described in Eq. (14), leads to the selection of detection matrices $\mathbf{D}_i$ as $\mathbf{U}$ and $\mathbf{V}$ for the near and far users. Additionally, the precoding matrix $\mathbf{P}_b$ is modified to $\mathbf{Q}\sqrt{P}/t$, where P represents the maximum transmission power and t is a power normalization coefficient [15]. Consequently, the MIMO receivers obtain the transmitted signal as follows:

$$\begin{aligned}\mathbf{U}\mathbf{y}_N &= \mathbf{U}\mathbf{H}_N\mathbf{P}_b\mathbf{x} + \mathbf{U}\mathbf{n_N} = \frac{P}{t\sqrt{d_N^\tau}}\mathbf{\Sigma}_N + \mathbf{U}\mathbf{n}_N \\ \mathbf{V}\mathbf{y}_F &= \mathbf{V}\mathbf{H}_F\mathbf{P}_b\mathbf{x} + \mathbf{V}\mathbf{n_F} = \frac{P}{t\sqrt{d_F^\tau}}\mathbf{\Sigma}_F + \mathbf{V}\mathbf{n}_F.\end{aligned} \tag{17}$$

Please note that $\mathbf{U}$ and $\mathbf{V}$ are the unitary matrices. Therefore, the unitary matrices $\mathbf{U}$ and $\mathbf{V}$ preserve the variance of noise after multiplication with therewith.

Example: In this example, we analyze a basic setup comprising a BS with four transmitter antennas. The near user is equipped with three receiver antennas, while the distant users have two receiver antennas. The near, which small-scale channel coefficient denoted as $\mathbf{H}_N \in \mathbf{C}^{3\times 4}$, is located $d_N = 40$ meters away from the base station. On the other hand, the far user, represented by $\mathbf{H}_F \in \mathbf{C}^{2\times 4}$, is located $d_F = 75$ meters away from the base station. The value of the path loss component, denoted as α, is 3.2.

$$\mathbf{H}_N = \begin{bmatrix} 0.629 + 0.735i & 0.066 + 0.931i & 0.193 + 0.616i & 0.924 + 0.556i \\ 0.210 + 0.772i & 0.260 + 0.013i & 0.639 + 0.949i & 0.263 + 0.915i \\ 0.752 + 0.907i & 0.804 + 0.234i & 0.524 + 0.950i & 0.065 + 0.641i \end{bmatrix}.$$

$$\mathbf{H}_F = \begin{bmatrix} 0.390+0.173i & 0.604+0.135i & 0.926+0.021i & 0.394+0.827i \\ 0.485+0.126i & 0.549+0.505i & 0.918+0.947i & 0.963+0.015i \end{bmatrix}.$$

Decoding matrices $\mathbf{U} \in \mathbf{C}^{3\times 3}$ and $\mathbf{V} \in \mathbf{C}^{2\times 2}$ can be found as follows:

$$\mathbf{U} = \begin{bmatrix} 0.5886+0.0000i & 0.0000+0.0000i & 0.8084+0.0000i \\ 0.2571-0.6199i & 0.4855+0.2742i & -0.1872+0.4513i \\ -0.2400+0.3816i & 0.6442+0.5235i & 0.1747-0.2778i \end{bmatrix}.$$

$$\mathbf{V} = \begin{bmatrix} -0.5294+0.7844i & 0.3201+0.0453i \\ -0.2029-0.2517i & 0.4024-0.8565i \end{bmatrix},$$

and the precoding matrix $\mathbf{Q}$ becomes:

$$\mathbf{Q} = \begin{bmatrix} -0.2004+0.3013i & 0.2468-0.5664i & 1.2740-0.5080i & 0.6976-0.7226i \\ -0.4527+0.5098i & -0.0139-0.7738i & 0.7713+0.3349i & 0.0869-0.8978i \\ -0.8984+0.6990i & -0.1623-1.2709i & 1.4061-0.6229i & 0.2928-0.3442i \\ 0.2405+0.5082i & 0.6040-1.2088i & 0.7572-0.7511i & 0.9449-0.2902i \end{bmatrix}.$$

Let's rearrange the super-positioned transmitted signal $\mathbf{x} \in \mathbf{C}^{4\times 1}$ with power allocation such that $\mathbf{x} = \mathbf{P}\mathbf{s}$, where $\mathbf{P} \in \mathbf{C}^{4\times 4}$ is the diagonal non-negative power allocation matrix and $\mathbf{s} \in \mathbf{C}^{4\times 1}$ contains the coded signals for both users as follows:

$$\mathbf{x} = \underbrace{\begin{bmatrix} p_{1,1} & 0 & 0 & 0 \\ 0 & p_{2,2} & 0 & 0 \\ 0 & 0 & p_{3,3} & 0 \\ 0 & 0 & 0 & p_{4,4} \end{bmatrix}}_{\mathbf{P}} \times \underbrace{\begin{bmatrix} l_1 s_{1,1} + \sqrt{(1-l_1^2)} s_{1,2} \\ l_2 s_{2,1} + \sqrt{(1-l_2^2)} s_{2,2} \\ l_3 s_{3,1} + \sqrt{(1-l_3^2)} s_{3,2} \\ l_4 s_{4,1} + \sqrt{(1-l_4^2)} s_{4,2} \end{bmatrix}}_{\mathbf{s}}$$

$$= \begin{bmatrix} p_{1,1}(l_1 s_{1,1} + \sqrt{(1-l_1^2)} s_{1,2}) \\ p_{2,2}(l_2 s_{2,1} + \sqrt{(1-l_2^2)} s_{2,2}) \\ p_{3,3}(l_3 s_{3,1} + \sqrt{(1-l_3^2)} s_{3,2}) \\ p_{4,4}(l_4 s_{4,1} + \sqrt{(1-l_4^2)} s_{4,2}) \end{bmatrix},$$

where $s_{i,1}$ and $s_{i,2}$ represent the corresponding message, and $l_{i,1}$ and $l_{i,2}$ are the power allocation coefficients for the near and far users, respectively. Also, we assume that s encoded with unit power i.e, $\|\mathbf{s}_i\|^2 = 1,\ i \in \{1,2,3,4\}$. After the GSVD is applied to the downlink channels, the channels become:

$$\mathbf{\Sigma}_N = \begin{bmatrix} 0 & 0.4526 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

and

$$\mathbf{\Sigma}_F = \begin{bmatrix} 1.0000 & 0 & 0 & 0 \\ 0 & 0.8917 & 0 & 0 \end{bmatrix}.$$

Observations at the near user are equal to $\mathbf{y}_N = \frac{\mathbf{\Sigma}_N \mathbf{x}}{\sqrt{d_N^{\alpha}}} + \mathbf{n}_N$ that can be written as follows:

$$\mathbf{y}_N = \begin{bmatrix} 0 & 0.4526 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} p_{1,1}(l_1 s_{1,1} + \sqrt{(1-l_1^2)} s_{1,2}) \\ p_{2,2}(l_2 s_{2,1} + \sqrt{(1-l_2^2)} s_{2,2}) \\ p_{3,3}(l_3 s_{3,1} + \sqrt{(1-l_3^2)} s_{3,2}) \\ p_{4,4}(l_4 s_{4,1} + \sqrt{(1-l_4^2)} s_{4,2}) \end{bmatrix} \times \frac{1}{\sqrt{d_N^{\alpha}}} + \begin{bmatrix} n_{N,1} \\ n_{N,2} \\ n_{N,3} \end{bmatrix}$$

$$= \begin{bmatrix} 0.4526 \times (l_2 s_{2,1} + \sqrt{(1-l_2^2)} s_{2,2}) \sqrt{d_N^{-\alpha}} + n_{N,1} \\ 1 \times (l_3 s_{3,1} + \sqrt{(1-l_3^2)} s_{3,2}) \sqrt{d_N^{-\alpha}} + n_{N,2} \\ 1 \times (l_4 s_{4,1} + \sqrt{(1-l_4^2)} s_{4,2}) \sqrt{d_N^{-\alpha}} + n_{N,3} \end{bmatrix}.$$

Similarly, observations at the far user $\mathbf{y}_F$ are equal to $\mathbf{y}_F = \frac{\mathbf{\Sigma}_F \mathbf{x}}{\sqrt{d_F^{\alpha}}} + \mathbf{n}_F$ and can be given as follows:

$$\mathbf{y}_F = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.8917 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} p_{1,1}(l_1 s_{1,1} + \sqrt{(1-l_1^2)} s_{1,2}) \\ p_{2,2}(l_2 s_{2,1} + \sqrt{(1-l_2^2)} s_{2,2}) \\ p_{3,3}(l_3 s_{3,1} + \sqrt{(1-l_3^2)} s_{3,2}) \\ p_{4,4}(l_4 s_{4,1} + \sqrt{(1-l_4^2)} s_{4,2}) \end{bmatrix} \times \frac{1}{\sqrt{d_F^{\alpha}}} + \begin{bmatrix} n_{F,1} \\ n_{F,2} \end{bmatrix}$$

$$= \begin{bmatrix} 1 \times (l_1 s_{1,1} + \sqrt{(1-l_1^2)} s_{1,2}) \sqrt{d_F^{-\alpha}} + n_{F,1} \\ 0.8917 \times (l_2 s_{2,1} + \sqrt{(1-l_2^2)} s_{2,2}) \sqrt{d_F^{-\alpha}} + n_{F,2} \end{bmatrix}.$$

We can consider each sub-channel an individual SISO channel. Therefore, they may require different SIC ordering regard-

ing their effective channel gains. Furthermore, by examining the expressions for $\mathbf{y}_N$ and $\mathbf{y}_F$, we can deduce that $\mathbf{s}_1$ corresponds to a private stream intended for the far user. On the other hand, $\mathbf{s}_3$ and $\mathbf{s}_4$ represent private streams dedicated to the near user, while $\mathbf{s}_2$ serves as the common stream shared by both users. The OMA transmission strategy can be employed for private streams. For example, the first steam may equal $\mathbf{s}_1 = s_{1,2}$.
Now, let us have a look at the decoding of the common stream. It can be calculated that the near user has a stronger channel gain than the far user, given by $\frac{0.4525}{\sqrt{40^{3.2}}} \geqslant \frac{0.8917}{\sqrt{75^{3.2}}}$. Therefore, the near and far users decode their signals as follows:

$$R_{N,2} = \log_2\left(1 + \frac{p_{2,2} \times l_2^2 \frac{0.4525}{40^{3.2}}}{\mathrm{Var}(n_{N,2})}\right),$$

$$R_{F,2} = \log_2\left(1 + \frac{p_{2,2} \times (1 - l_2^2) \times \frac{0.8917}{75^{3.2}}}{p_{2,2} \times l_2^2 \frac{0.4525}{40^{3.2}} + \mathrm{Var}(n_{F,2})}\right).$$

Likewise, data rate expressions can be derived for private channels by eliminating inter-user interference. Additionally, the transmitted power from each antenna, denoted as $\mathbf{P}$, holds a specific physical interpretation, such as maximum transmit power from each antenna element. Hence, the opportunity for further exploration emerges from optimizing the elements within $\mathbf{P}$, encompassing enhancing secrecy, improving data rate, ensuring fairness, and selecting optimal antennas.

3. Mobile Edge Computing

Mobile edge computing, also known as multi-access edge computing, brings the processing of traffic and services from centralized cloud servers to the edge of the network, closer to the end-users, as illustrated in Fig. 6. Instead of transmitting all the data to the cloud for analysis, MEC devices are responsible for the processing, storage, and analysis of the data [24]. This approach minimizes latency, thus improving performance of high-bandwidth applications in real-time [25]. The combination of NOMA and MEC holds immense potential, as it not only enhances the spectral efficiency of MEC users, but also empowers IoT devices at the edge to handle computationally intensive tasks. Combining NOMA and MEC requires an optimal approach to resource and power allocation and time management. In order to minimize the time spent offloading tasks, UE needs to determine the optimal task partition coefficient β and power allocation p_{off}. The offloading time (T_{off}) can be defined as follows:

$$T_{off} = \frac{\beta N}{R} \text{ [s]}, \tag{18}$$

Here, N represents the data size of the task, and R is the data rate of the UE.
The energy consumed during the offloading time T_{off} can be calculated as:

$$E_{off} = T_{off} \times p_{off} \text{ [J]}, \tag{19}$$

In Eq. (19), p_{off} denotes the transmit power of the UE. Once the data is offloaded to the MEC server, the duration for the

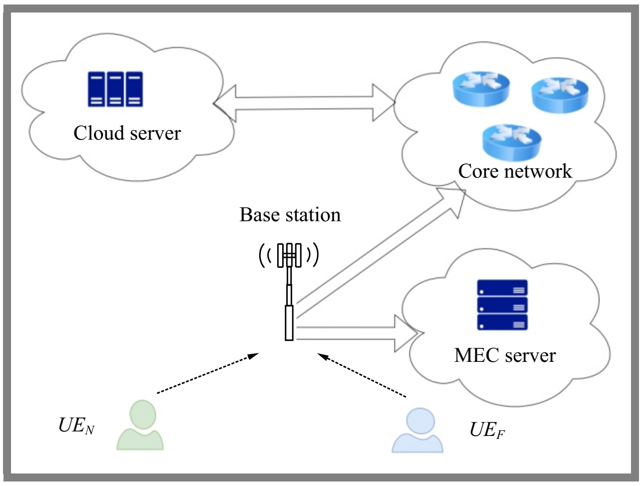

Fig. 6. NOMA-assisted MEC model.

mobile execution time T_{mec} can be determined using:

$$T_{mec} = \frac{\beta N C_m}{f_m} \text{ [s]}, \tag{20}$$

In this equation, C_m represents the required CPU cycles to execute a bit, and f_m is the CPU frequency of the MEC server. The energy consumption during T_{mec} can be calculated as:

$$E_{mec} = \xi \beta N C_m f_m^2 \text{ [J]}, \tag{21}$$

where ξ denotes the energy consumption coefficient for the MEC. Tables 2–3 summarize existing works on combining NOMA and MEC. It provides valuable insights into different research papers that have explored this subject, with a particular emphasis placed on potential optimization benefits.

4. Discussions and Future Works

In this paper, we have presented a comprehensive analysis of NOMA and compared it with OMA schemes. The study also explored the integration of NOMA with MIMO technologies using the GSVD method. Additionally, we discussed the key parameters of the MEC technology and provided a literature review on NOMA with MEC.
Future research directions in this field involve addressing the limitations that come with assuming availability of perfect channel state information (CSI) at the transmitters and receivers. It is recommended to investigate the implications and benefits of incorporating imperfect channel estimation with random error matrices, as demonstrated in [36]. By considering such realistic scenarios with imperfect CSI, the findings of this study can be extended to real-world applications.
Furthermore, the constraints of the GSVD technique, which currently allow the combination of only two users simultaneously, pose another limitation. However, recent advancements have shown that the GSVD technique can be extended to accommodate more than two users, as highlighted in [37] and [38]. The results of these studies indicate the potential of assigning more than two users to a resource block, using the proposed methods.
An intriguing path for future research involves integrating the

Tab. 1. GSVD-based MIMO applications.

Ref.	Objective	Technology	System analysis	Optimization variable	Constraints	UL/DL	Result
[16]	Minimize outage probability while improving physical layer security	NOMA-MIMO	Performance analysis and optimization	Power allocation coefficients	N/A	DL	Compared with GSVD-OMA based transmission, NOMA has superior outage performance
[17]	Design low complexity and highly efficient GSVD-based beamforming to maximize secrecy capacity	OMA-MIMO	Optimization	Power allocation coefficients	Average power consumption	DL	GSVD-MIMO achieves nearly identical performance with secure dirty paper coding (S-DPC)
[18]	Maximize secrecy rate	OMA-MIMO	Optimization	Sub-channel and power allocation	Quality of service	DL	GSVD-based precoding outperforms a TDMA-based system
[19]	Obtain the expressions of the average data rate and outage in a MIMO-NOMA relaying	NOMA-MIMO	Performance analysis	N/A	Finite number of users	DL	GSVD-NOMA achieves a higher sum rate than GSVD-OMA
[20]	Maximize minimum data rate	NOMA-MIMO	Optimization	Power allocation coefficients	Imperfect channel estimation	DL	The SINR balancing problem was solved using error bounds. The proposed solution has better performance than non-robust or OMA-based solutions
[21]	Minimize offloading delay	H-NOMA-MIMO	Optimization	Power allocation coefficients	Total power	UL	Hybrid NOMA-MIMO based solution has better delay performance compared with OMA-based solution
[22]	Minimize energy consumption	NOMA-MIMO	Optimization	Task assignment and power allocation coefficients	Total power, offloading time, and RF chains energy consumption	UL	NOMA-MIMO performs better than OMA, especially when the data is high and time is stringent
[23]	Maximize secrecy sum rate	NOMA-MIMO	Optimization	Power allocation coefficients	Total power and QoS	UL	NOMA has better SSR performance than OMA

Tab. 2. Summary of some existing works on NOMA-MEC.

Ref.	Objective	Method	Technology	Constraints	Optimization parameters	Offloading policy
[26]	Minimize task offloading and computing delay	Linear problem using auxiliary variable	Massive MIMO	Transmit power, MEC computing capacity	Power allocation, computation frequency allocation	Partial
[24]	Minimize delay	Bisection search	SISO-MEC	Energy and offloading power, computation time at MEC	Task partition coefficient, power allocation	Partial
[22]	Minimize total energy during local task offloading and MEC computing	AO-SCA	MIMO-MEC	Total power, time, energy consumption on RF chains	Task partition coefficient, power allocation	Partial
[23]	Minimize delay	Dinkelbach transform-SCA	MIMO-MEC	Total power	Power allocations	Full
[27]	Minimize system energy consumption	SCA	SISO-MEC	Time, transmission power, decoding power	Power allocation	Full
[28]	Minimize delay	Alternating optimization (AO)	UAV assisted SISO-MEC	Energy and QoS	Trajectory of UAV, power allocation, user scheduling	Full

proposed system models with emerging technologies, such as intelligent reflecting surfaces and unmanned aerial vehicles. The integration of these technologies holds significant promise and provides opportunities to enhance the signal-to-interference-plus-noise ratio (SINR) of MIMO-NOMA systems. Overall, this research contributes to the understanding of NOMA, its comparison with OMA schemes, the application of GSVD in NOMA systems, and the exploration of key parameters associated with the MEC technology. The identified research directions open up avenues for further advancements in wireless communication systems, enabling more efficient and reliable transmissions in diverse scenarios.

References

[1] B. Makki *et al.*, "A Survey of NOMA: Current Status and Open Research Challenges", *IEEE Open Journal of the Communications Society*, vol. 1, pp. 179–189, 2020 (https://doi.org/10.1109/OJCOMS.2020.2969899).

[2] M. Taherzadeh, H. Nikopour, A. Bayesteh, and H. Baligh, "SCMA Codebook Design", in: *2014 IEEE 80th Vehicular Technology Conference (VTC2014-Fall)*, Vancouver, Canada, pp. 1–5, 2014 (https://doi.org/10.1109/VTCFall.2014.6966170).

[3] J. Zeng *et al.*, "Pattern Division Multiple Access (PDMA) for Cellular Future Radio Access", in: *2015 International Conference on Wireless Communications & Signal Processing (WCSP)*, Nanjing, China, pp. 1–5, 2015 (https://doi.org/10.1109/WCSP.2015.7341229).

[4] Y. Cao, H. Sun, J. Soriaga, and T. Ji, "Resource Spread Multiple Access – A Novel Transmission Scheme for 5G Uplink", in: *2017 IEEE 86th Vehicular Technology Conference (VTC-Fall)*, Toronto, Canada, 2017 (https://doi.org/10.1109/VTCFall.2017.8288412).

[5] Z. Yuan *et al.*, "Multi-User Shared Access for Internet of Things", in: *2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)*, Nanjing, China, 2016 (https://doi.org/10.1109/VTCSpring.2016.7504361).

[6] S. Hu *et al.*, "Nonorthogonal Interleave-Grid Multiple Access Scheme for Industrial Internet of Things in 5G Network", *IEEE Transactions on Industrial Informatics*, vol. 14, no. 12, pp. 5436–5446, 2018 (https://doi.org/10.1109/TII.2018.2858142).

[7] B. Clerckx *et al.*, "Rate Splitting for MIMO Wireless Networks: a Promising Physical Layer Strategy for LTE Evolution", *IEEE Communications Magazine*, vol. 54, no. 5, pp. 98–105, 2016 (https://doi.org/10.1109/MCOM.2016.7470942).

[8] A. Benjebbour, "An Overview of Non-Orthogonal Multiple Access", *ZTE Communications*, vol. 15, no. 1, pp. 28–34, 2017 (https://doi.org/10.3969/j.issn.1673-5188.2017.01.005).

[9] M. Aldababsa *et al.*, "A Tutorial on Non-orthogonal Multiple Access for 5G and Beyond", *Wireless Communications and Mobile Computing*, pp. 1–24, 2018 (https://doi.org/10.1155/2018/9713450).

[10] Z. Ding, R. Schober, and H.V. Poor, "Unveiling the Importance of SIC in NOMA Systems – Part 1: State of the Art and Recent Findings", *IEEE Communications Letters*, vol. 24, no. 11, pp. 2373–2377, 2020 (https://doi.org/10.1109/LCOMM.2020.3012604).

[11] Z. Yang *et al.*, "Sum-Rate Maximization of Uplink Rate Splitting Multiple Access (RSMA) Communication", in: *2019 IEEE Global Communications Conference (GLOBECOM)*, Waikoloa, USA, 2019 (https://doi.org/10.1109/GLOBECOM38437.2019.9013344).

[12] C.C. Paige, "Computing the Generalized Singular Value Decomposition", *SIAM Journal on Scientific and Statistical Computing*, vol. 7, no. 4, pp. 1126–1146,1986 (https://doi.org/10.1137/0907077).

Tab. 3. Summary of some existing works on NOMA-MEC – continued.

Ref.	Objective	Method	Technology	Constraints	Optimization parameters	Offloading policy
[29]	Maximize computation capacity	AO, DC programming	Backscatter-assisted SISO_NO-MA	Energy and QoS	Energy harvesting time coefficient, BackCom time coefficient, transmission time, computing resource allocation	Partial
[30]	Maximize computation capacity	AO, concave-convex procedure and SDR	IRS and UAV assisted SISO-NOMA	Total transmit power	Phase shift of IRS, transmit power, computational resource allocation, the trajectory of UAV	Binary
[31]	Minimize system energy	AO, matching algorithm, SCA	SISO-MEC	Latency	Power allocation, time, sub-channel allocation	Binary
[32]	Maximize computation capacity	Deep reinforcement learning	SISO-multi-MEC	Delay, limited sub-channel	Task scheduling, power allocation	Binary
[25]	Minimize latency	AO	WPT, IRS	Transmit power	Power allocation, phase shift of IRS	Partial
[33]	Minimize total energy consumption	TD3	SISO-MEC	Transmit power, latency	Task partition, power allocation	Binary
[34]	Maximize computation probability	Meta-heuristic-based algorithms, PSO, GA	SISO-MEC	Transmit power, computational resource	Task partition, power allocation	Binary, partial, full
[35]	Minimize power consumption	AO, Riemann gradient descent	IRS-MEC	Delay, computational resource	Bandwidth allocation, computational resource allocation, power allocation, the phase shift of IRS	Binary

[13] S. Doclo and M. Moonen, "GSVD-based Optimal Filtering for Single and Multi-microphone Speech Enhancement", *IEEE Transactions on Signal Processing*, vol. 50, no. 9, pp. 2230–2244, 2002 (https://doi.org/10.1109/TSP.2002.801937).

[14] O. Alter, P.O. Brown, and D. Botstein, "Generalized Singular Value Decomposition for Comparative Analysis of Genome-scale Expression Data Sets of Two Different Organisms", *Proceedings of the National Academy of Sciences*, vol. 100, no. 6, pp. 3351–3356, 2003 (https://doi.org/10.1073/pnas.0530258100).

[15] 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 (https://doi.org/10.1109/TWC.2018.2883102).

[16] 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 (https://doi.org/10.1109/LWC.2021.3071365).

[17] S. Ali., A. Fakoorian, and A.L. Swindlehurst, "Dirty Paper Coding versus Linear GSVD-Based Precoding in MIMO Broadcast Channel with Confidential Messages", in: *2011 IEEE Global Telecommunications Conference – GLOBECOM 2011*, Houston, USA, 2011 (https://doi.org/10.1109/GLOCOM.2011.6134129).

[18] W. Mei, Z. Chen, and J. Fang, "GSVD-based Precoding in MIMO Systems with Integrated Services", *IEEE Signal Processing Letters*, vol. 23, no. 11, pp. 1528–1532, 2016 (https://doi.org/10.1109/LSP.2016.2606349).

[19] C. Rao, Z. Ding, and X. Dai, "Application of GSVD-based Precoding in MIMO-NOMA Relaying Systems", *IET Communications*, vol. 14, no. 21, pp. 3802–3812, 2020 (https://doi.org/10.1049/iet-com.2020.0555).

[20] 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 (https://doi.org/10.1109/LWC.2019.2926277).

[21] Y. Dursun, F. Fang, and Z. Ding, "Hybrid NOMA based MIMO Offloading for Mobile Edge Computing in 6G Networks", *China Communications*, vol. 19, no. 10, pp. 12–20, 2022 (https://doi.org/10.23919/JCC.2022.00.024).

[22] Y. Dursun, M.B. Goktas, and Z. Ding, "Green NOMA based MU-

MIMO Transmission for MEC in 6G Networks", *Computer Networks*, vol. 228, art. no. 109749, 2023 (https://doi.org/10.1016/j.comnet.2023.109749).

[23] Y. Dursun, K. Wang, and Z. Ding, "Secrecy sum rate maximization for a MIMO-NOMA uplink transmission in 6G networks", *Physical Communication*, vol. 53, art. no. 101675, 2022 (https://doi.org/10.1016/j.phycom.2022.101675).

[24] F. Fang *et al.*, "Optimal Resource Allocation for Delay Minimization in NOMA-MEC Networks", *IEEE Transactions on Communications*, vol. 68, no. 12, pp. 7867–7881, 2020 (https://doi.org/10.1109/TCOMM.2020.3020068).

[25] G. Li *et al.*, "Latency Minimization for IRS-Aided NOMA MEC Systems with WPT-enabled IoT Devices", *IEEE Internet of Things Journal*, vol. 10, no. 14, pp. 12156–12168, 2023 (https://doi.org/10.1109/JIOT.2023.3240395).

[26] S.S. Yılmaz and B. Ozbek, "Massive MIMO-NOMA Based MEC in Task Offloading for Delay Minimization", *IEEE Access*, vol. 11, pp. 162–170, 2022 (https://doi.org/10.1109/ACCESS.2022.3232731).

[27] Y. Zhang, Z. Na, Y. Wang, and C. Ji, "Joint power allocation and deployment optimization for HAP-assisted NOMA-MEC System", *Wireless Networks*, pp. 1–13, 2022 (https://doi.org/10.1007/s11276-022-03201-8).

[28] F. Guo *et al.*, "Joint Trajectory and Computation Offloading Optimization for UAV-assisted MEC with NOMA", in: *IEEE INFOCOM 2019 – IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, Paris, France, 2019, (https://doi.org/10.1109/INFOCOMWKSHPS47286.2019.9093764).

[29] C. Zheng and W. Zhou, "Computation Bits Maximization in Backscatter-assisted Wireless-powered NOMA-MEC Networks", *EURASIP J. on Wireless Communications and Networking*, art. no. 23, 2022 (https://doi.org/10.1186/s13638-022-02097-4).

[30] Y. Xu, T. Zhang, Y. Zou, and Y. Liu, "Reconfigurable Intelligence Surface Aided UAV-MEC Systems with NOMA", *IEEE Communications Letters*, vol. 26, no. 9, pp. 2121–2125, 2022 (https://doi.org/10.1109/LCOMM.2022.3183285).

[31] I. Altin and M. Akar, "A Joint Resource Allocation Method for Hybrid NOMA MEC Offloading", *Physical Communication*, vol. 54, art. no. 101809, 2022 (https://doi.org/10.1016/j.phycom.2022.101809).

[32] L. Lin, W. Zhou, Z. Yang, and J. Liu, "Deep Reinforcement Learning-based Task Scheduling and Resource Allocation for NOMA-MEC in Industrial Internet of Things", *Peer-to-Peer Networking and Applications*, vol. 16, no. 1, pp. 170–188, 2023 (https://doi.org/10.1007/s12083-022-01348-x).

[33] C. Li, H. Wang, and R. Song, "Mobility-Aware Offloading and Resource Allocation in NOMA-MEC Systems via DC", *IEEE Communications Letters*, vol. 26, no. 5, pp. 1091–1095, 2022 (https://doi.org/10.1109/LCOMM.2022.3154434).

[34] T.V. Truong and A. Nayyar, "System Performance and Optimization in NOMA Mobile Edge Computing Surveillance Network using GA and PSO", *Computer Networks*, vol. 223, art. no. 109575, 2023 (https://doi.org/10.1016/j.comnet.2023.109575).

[35] M. Chen, Y. Wan, M. Wen, and T. Zhou, "Fairness Optimization in IRS-assisted MEC Computational Offloading", *Physical Communication*, vol. 54, art. no. 101855, 2022 (https://doi.org/10.1016/j.phycom.2022.101855).

[36] C-L. Wang, Y.-C. Ding, Y.-C. Wang, "A Low-Complexity Power Allocation Scheme for MIMO-NOMA Systems with Imperfect Channel Estimation", in: *2022 IEEE 33rd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Kyoto, Japan, pp. 234–239, 2022 (https://doi.org/10.1109/PIMRC54779.2022.9977487).

[37] L. Khamidullina, A.L.F. de Almeida, and M. Haardt, "Multilinear Generalized Singular Value Decomposition (ML-GSVD) and its Application to Multiuser MIMO Systems", *IEEE Transactions on Signal Processing*, vol. 70, pp. 2783–2797, 2022 (https://doi.org/10.1109/TSP.2022.3178902).

[38] L. Khamidullina, A.L.F. de Almeida, and M. Haardt, "Rate Splitting and Precoding Strategies for Multi-User MIMO Broadcast Channels with Common and Private Streams", in: *ICASSP 2023 – 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Rhodes Island, Greece, 2023 (https://doi.org/10.1109/ICASSP49357.2023.10095138).

Yunus Dursun, Ph.D.
School of Electrical and Electronic Engineering
https://orcid.org/0000-0003-3405-8369
E-mail: yunus.dursun@manchester.ac.uk
University of Manchester, Manchester, United Kingdom
https://www.manchester.ac.uk