Multi-User Angle-Domain MIMO-NOMA System for MmWave Communications

Israa Khaled, Charlotte Langlais, Ammar El Falou, Bachar Elhassan, Michel Jezequel

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3112225, IEEE Access

Digital Object Identifier

Multi-User Angle-Domain MIMO-NOMA System for MmWave Communications

ISRAA KHALED[1,2], CHARLOTTE LANGLAIS[1], AMMAR EL FALOU[2,3], BACHAR ELHASSAN[2], and MICHEL JEZEQUEL[1]

[1]IMT Atlantique, Lab-STICC, F-29238 Brest, France
[2]Lebanese University, Faculty of Engineering, Tripoli, Lebanon
[3]ESYCOM, Gustave Eiffel University, EISEE Paris, Noisy-le-Grand, France

Corresponding author: Israa Khaled (e-mail: israa.khaled@imt-atlantique.fr).

This work was supported in part by Institut Mines-Télécom (IMT) Atlantique, and in part by the Lebanese University research fund program and the AZM association.

ABSTRACT Thanks to the high directionality of millimeter-wave (mmWave) channels, angle-domain beamforming is an appealing technique for multi-user multiple-input multiple-output (MU-MIMO) in terms of sum-throughput performance and limited feedback. By utilizing only the angular information of users at the transmitter, we propose an angle-domain non-orthogonal multiple access (NOMA) scheme to enhance the sum-throughput of the mmWave MU-MIMO system, especially in congested cells. We first derive a set of angular-based performance metrics, such as the inter-user spatial interference, the user channel quality, and the sum-throughput, by exploiting the specific features of the mmWave propagation. Then, a multi-user clustering algorithm is developed based on the spatial interference metric, and a new user ordering strategy is proposed using the angular-based channel quality metric. Additionally, we design a power allocation method that maximizes the angular-based sum-throughput. Extensive numerical results show that the proposed scheme significantly improves the performance of the mmWave MU-MIMO system by achieving up to 39% increase in the spectral efficiency when the number of users is closed to the number of antennas. Moreover, we find that the proposed user ordering strategy outperforms other limited feedback strategies, and the angular-based power allocation allows for efficient successive interference cancellation.

INDEX TERMS Millimeter-wave networks, multiple-input multiple-output, beamforming, non-orthogonal multiple access.

I. INTRODUCTION

A. MOTIVATION AND RELATED WORK

Millimeter-wave (mmWave) communication has been recognized as an emerging solution to meet the high data rate requirements in future cellular networks [1]. Indeed, mmWave systems can provide large blocks of contiguous spectrum. Additionally, the short mmWave wavelengths facilitate the practical deployment of the multiple-input multiple-output (MIMO) technology, which subsequently overcomes the severe propagation loss. To further enhance the network capacity, the non-orthogonal multiple access (NOMA) technique can be adopted. Using superposition coding (SC) at the base station (BS) and successive interference cancellation (SIC) at the receiver side, NOMA efficiently improves the connection density by multiplexing users in the power domain [2]. Moreover, NOMA achieves a significant gain over the conventional orthogonal multiple access (OMA) in terms of sum-rate and user fairness for 2-user and multi-user scenarios [3]–[5].

NOMA-based multi-user MIMO systems have been extensively studied in the literature for ergodic channels. Most of the existing works on MIMO-NOMA design the beamforming weight vector, user clustering and power allocation techniques based on the assumption of perfect and imperfect full instantaneous channel state information (CSI) at the transmitter [6]–[11]. However, acquiring full CSI (FCSI) is challenging due to user mobility, and on the other hand, because of the overall system overhead. One possible solution is to use statistical CSI (SCSI), such as the spatial correlation or the channel mean, which can be easily and accurately obtained at the transmitter [12]. For instance, the authors in [3] formulate an ergodic capacity maximization problem under SCSI to design the power allocation for a Rayleigh fading 2-user MIMO-NOMA system. In [13], an optimal statistical precoder for a MIMO-NOMA system is designed to maximize the ergodic capacity. In [14], a joint rate and

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3112225, IEEE Access

power allocation scheme is proposed for downlink NOMA to minimize the total transmission power under SCSI.

However, in the context of mmWave transmissions, the FCSI challenge can be mitigated thanks to the specific properties of mmWave channels [15], [16]: i) the highly directional propagation of mmWave wireless links and ii) the sensitivity of the mmWave signals to blockages. Thus, the non-line-of-sight (NLoS) components are severely attenuated compared to the line-of-sight (LoS) link. As a result, the spatial location of the user, i.e., the angle and the distance with respect to (w.r.t..) the BS, can efficiently characterize the mmWave channels [17]. Motivated by these properties, several works design NOMA systems with limited feedback for the mmWave networks. To reduce the system overhead, the authors in [18] propose a NOMA scheme with random analog beamforming, where only the users' distance information is fed back to the BS instead of FCSI. Specifically, the BS randomly generates a single beam, and users falling in this beam are served using NOMA with a fixed power allocation strategy. Meanwhile, the users are ordered according to their distance information instead of their effective channel gain. In [19], a joint user selection and analog beamforming scheme based on the user position is proposed for mmWave NOMA systems. First, the authors define the users' pairing angle based on the angle difference between the two users. Thereby, they develop a user selection method that assigns a near user with a far one w.r.t.. the BS, which are expected to be within the specified beamwidth and have the minimum pairing angle. In [17], a location-based low-complexity user pairing algorithm is proposed for the mmWave MIMO-NOMA transmission by applying the maximum ration transmission technique. The authors consider the Rician fading channels and derive the switching boundary between NOMA and spatial division multiple access in the angle-distance plane for the 2-user scenario. They found that NOMA is preferable when the angle difference between the two users is smaller than a threshold, or their distance difference is larger than a threshold. By applying these results in the multi-user scenario, they design a low-complexity location-based user pairing algorithm. In [20], the authors propose a user pairing strategy based on both angle and distance. The BS pairs the nearest user with another having the minimum relative angle distance, such that the latter is located within a distance threshold from the BS.

B. CONTRIBUTION

Consequently, both user angle and distance information w.r.t.. the BS has significant potential in the field of mmWave MIMO-NOMA systems with limited feedback. In the light of these articles, we consider a downlink MIMO-NOMA scheme with limited feedback that exploits only the angle-domain information (ADI), i.e., the user's spatial angle[1] that are collected at the BS, to provide: 1) user clustering, 2) user ordering, 3) power allocation, 4) fully digital beamforming performing simultaneous addressing of users, named angle-domain digital beamforming (AD-DBF). In the rest of the paper, ADI is assumed to be perfectly estimated and known at the transmitter. Figure 1 illustrates our proposed MIMO-NOMA scheme: thanks to the definition of an inter-user spatial interference metric, users are grouped within clusters, themselves addressed simultaneously by the beams of the digital beamformer. NOMA is used in each cluster to manage multiple access interference. This paper extends the results of our work [23][2] to the more general multi-user scenario instead of 2 users. Moreover, we consider a multi-path propagation environment. Coupling digital beamforming and NOMA helps to manage high user density in a congested cell by offering spatial and non-orthogonal multiple access degrees of freedom.

The main contributions of this paper can be summarized as follows

- The sum-throughput expression of the proposed MIMO-NOMA system with ADI feedback is derived for the general multi-user scenario. It is built based on a newly defined user channel quality metric that considers the inter-cluster spatial interference and only uses ADI. Thanks to the defined metric, we propose a new user ordering strategy for the NOMA protocol, with only the ADI feedback.
- A multi-user clustering algorithm is developed so that users with high spatial interference are grouped within the same cluster themselves served using NOMA. For that, we derive an inter-user spatial interference metric based on ADI by assuming the predominance of the LOS link. Unlike prior works on angular-based NOMA systems that only use the spatial angular distance between users (a.k.a. azimuth angle difference), this metric includes the beam-width information of each main lobe and does not require additional feedback. The beam-width is very dependent on the antennasâĂŹ number and the steering angle of the beam.
- A power optimization problem that maximizes the system sum-throughput is formulated with either FCSI or ADI feedbacks and under total power and efficient SIC constraints. The obtained problem is simplified to a convex problem, thanks to a fixed inter-cluster power allocation strategy. Then, this simplification allows for the derivation of the closed-form intra-cluster power

[1]The user's spatial angle w.r.t. the BS represents the angle-of-departure (AoD) of the LoS path and is a slowly time-varying CSI compared to the other CSI components. This information can be estimated and tracked with standard AoDs estimation algorithms such as [21], [22] at the BS or the receiver side. If the users estimate this information, then they send it back to the BS.

[2]In our previous work [23], we investigate the angle-domain (AD) MIMO-NOMA scheme with only ***two users*** within the NOMA cluster in a ***mono-path*** environment. A ***2-user clustering algorithm*** is developed, and a ***fixed power allocation*** method is applied for the two NOMA users.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3112225, IEEE Access

allocation solution, obtained via Karush-Kuhn-Tucker (KKT) conditions.
- Extensive simulations are conducted in mmWave multi-path rural environments to verify the performance of the proposed schemes, using the stochastic mmWave channel model, called NYUSIM [24]. We compare the performance of the proposed ordering strategy with other limited feedback-based ordering strategies. We show the effectiveness of SIC when applying the angular-based power allocation. Moreover, we compare the proposed multi-user AD MIMO-NOMA schemes with the conventional AD-DBF (C-AD-DBF) [25], the 2-user AD MIMO-NOMA [23] and the OMA scheme.

C. ORGANIZATION

The rest of this paper is organized as follows: Section II presents the model of the proposed multi-user AD MIMO-NOMA scheme that integrates AD-DBF with the NOMA transmission. In Section III, we derive the ADI-based performance metrics: the user channel quality, the sum-throughput and the inter-user spatial interference. Section IV discusses the multi-user clustering algorithm. In Section V, we formulate the power allocation problem, which will be solved using KKT conditions, thanks to a fixed inter-cluster power allocation. Section VI numerically evaluates the performance of the proposed schemes, and Section VII concludes the paper.

D. NOTATIONS

Throughout this paper, $\mathbf{A}$, $\mathbf{a}$ and a denote matrix, vector and scalar, respectively. $(.)^T$, $(.)^H$ and $\mathrm{Tr}(\cdot)$ represent the transpose, the Hermitian transpose and the trace, respectively. $\mathcal{N}(\nu, \sigma^2)$ is a Gaussian random variable with mean ν and variance σ^2.

II. DESCRIPTION OF MULTI-USER ANGLE-DOMAIN MIMO-NOMA SCHEME

We consider a mmWave AD MIMO-NOMA system in the downlink, as depicted in Fig. 1, with one BS and K single-antenna user equipments (UEs). At the BS, a uniform linear array (ULA) with M antenna elements are stacked horizontally along the x-axis. All UEs are represented by the ID set $\mathcal{K} = \{1, \cdots, K\}$. According to the user clustering algorithm described in Section IV, the K UEs are grouped into $C \leq K$ spatial clusters with K_c UEs in the c-th cluster, $c \in \mathcal{C} = \{1, \cdots, C\}$. NOMA is adopted within the c-th cluster where multiple UEs exist. By utilizing digital phase shifters and the knowledge of ADI, i.e., the user's spatial direction, AD-DBF [25] forms beams towards each cluster.

Consequently, two different types of clusters exist, namely single-user (SU) and NOMA clusters, as shown in Fig. 1. In the former, only one UE is served thanks to a beam generated by AD-DBF. However, in the latter, multiple UEs share the same spatial and time-frequency resource blocks using NOMA. $\mathcal{C}_n = \{1, \cdots, C_n\}$ and $\mathcal{C}_s = \{C_n+1, \cdots, C\}$ are the ID sets of NOMA and SU clusters, respectively, with C_n and C_s being respectively the number of NOMA and SU clusters. Besides, $\mathcal{K}_c = \{1, \cdots, K_c\}$ is the newly defined ID set of UEs in the c-th cluster, with K_c being the number of UEs in the c-th cluster. Hence, $K_c = 1$, $\forall c \in \mathcal{C}_s$ and $\mathcal{C} = \mathcal{C}_n \cup \mathcal{C}_s$. In this work, we extend the study of [23] to allow for multi-user communication per NOMA cluster, i.e., $K_c \geq 2$ [3], $\forall c \in \mathcal{C}_n$, and we consider a multi-path environment.

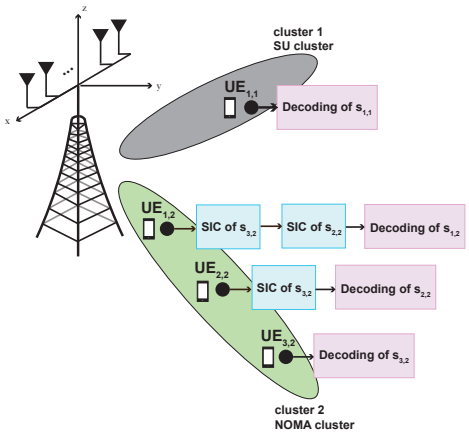


FIGURE 1. Multi-user MIMO-NOMA system.

A. CHANNEL MODEL

Using ULA along the x-axis, the array steering vector $\mathbf{a}(\theta)$ corresponding to the spatial angle θ is defined as

$$\mathbf{a}(\theta) = \left[1, e^{j2\pi \frac{d}{\lambda}\cos(\theta)}, \cdots, e^{j2\pi(M-1)\frac{d}{\lambda}\cos(\theta)}\right]^T, \quad (1)$$

where λ is the wavelength and $d = \frac{\lambda}{2}$ is the inter-antenna spacing.

The mmWave channel vector $\mathbf{h}_{l,c} \in \mathbb{C}^{1\times M}$ between the BS and the lth UE in the c-th cluster, denoted as $\mathrm{UE}_{l,c}$, is

$$\mathbf{h}_{l,c} = \sum_{q=1}^{N_{l,c}} \alpha_{q,l,c}\mathbf{a}^H(\theta_{q,l,c}), \quad (2)$$

where $N_{l,c}$ is the number of multi-path components in $\mathrm{UE}_{l,c}$'s channel, $\alpha_{q,l,c}$ and $\theta_{q,l,c}$ are respectively the complex gain and the spatial angle associated with the q-th path in $\mathrm{UE}_{l,c}$'s channel, and $\mathbf{a}_{q,l,c} = \mathbf{a}(\theta_{q,l,c}) \in \mathbb{C}^{M\times 1}$ is the array steering vector corresponding to $\theta_{q,l,c}$. In this work, we assume that the LoS path exists in the channel of each user. And, the case where the LoS path does not exist for some users is left as a future work. Denote $q = 1$ as the index of the LoS path, i.e., the strongest path. Hence, $\theta_{1,l,c}$ represents the spatial direction of $\mathrm{UE}_{l,c}$.

[3] In this work, we do not consider the complexity of the SIC decoding, and hence the NOMA cluster size is not defined. This consideration is left for future work.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3112225, IEEE Access

B. DATA TRANSMISSION

The $\text{UE}_{l,c}$ receives the following signal

$$y_{l,c} = \sqrt{\eta}\mathbf{h}_{l,c}\mathbf{w}_c\mathbf{s}_c + \sum_{b\in\mathcal{C},b\neq c} \sqrt{\eta}\mathbf{h}_{l,c}\mathbf{w}_b\mathbf{s}_b + n_{l,c}, \quad (3)$$

where η is the normalization factor that eliminates the beamforming effect on the transmission power, $\mathbf{w}_c \in \mathbb{C}^{M\times 1}$ is the beamforming weight vector corresponding to the c-th cluster, $\mathbf{s}_c$ is the superposed signal of the c-th cluster and $n_{l,c} \sim \mathcal{N}(0, \sigma_n^2)$ is the additive white Gaussian noise at $\text{UE}_{l,c}$.

In (3), the first term represents the received superposed signal corresponding to the c-th cluster, while the second term is the inter-cluster interference. Denote by $\mathbf{W} = [\mathbf{w}_1 \cdots \mathbf{w}_C] \in \mathbb{C}^{M\times C}$ the transmit beamforming matrix, then $\eta = \frac{1}{\text{Tr}(\mathbf{W}^H\mathbf{W})}$. We apply an AD-DBF technique that uses digital phase shifters to generate and steer a spatial beam toward each cluster, as shown in Fig. 1. Thereby, $\mathbf{w}_c = \mathbf{a}_c$, where $\mathbf{a}_c = \mathbf{a}(\theta_c)$ is the array steering vector corresponding to the spatial direction θ_c of the c-th cluster, and then $\eta = \frac{1}{MC}$. The angle θ_c is calculated based on the spatial direction $\theta_{1,l,c}$ of UEs within the c-th cluster as follows

$$\theta_c = \begin{cases} \theta_{1,1,c} & \text{if } K_c = 1, \\ \frac{\min\limits_{l\in\mathcal{K}_c}\{\theta_{1,l,c}\} + \max\limits_{l\in\mathcal{K}_c}\{\theta_{1,l,c}\}}{2} & \text{if } K_c \geq 2. \end{cases} \quad (4)$$

The set $\boldsymbol{\Theta} = \{\theta_{1,l,c},\ l \in \mathcal{K}_c,\ c \in \mathcal{C}\}$ is called ADI. The main objective of this work is to design a MIMO-NOMA scheme with ADI feedback. This implies that the only channel information to be estimated is the set of spatial directions of UEs. This information can be estimated and tracked with standard AoDs estimation algorithms such as [21], [22] at the BS or the receiver side. In this paper, we assume a perfect estimation and feedback if needed. And, the impact of the imperfection of ADI on the proposed scheme is left for future work.

Using SC at the BS, NOMA superposes in power-domain the signals of UEs within the same cluster. The superposed signal $\mathbf{s}_c$ of the c-th cluster is thereby given by

$$\mathbf{s}_c = \sum_{l\in\mathcal{K}_c} \sqrt{\gamma_{l,c}p_c}s_{l,c}, \quad (5)$$

where $s_{l,c}$ is the modulated signal relative to $\text{UE}_{l,c}$, p_c is the power allocated to the c-th cluster, and $\gamma_{l,c}$ is the power allocation coefficient for $s_{l,c}$ such that $\sum_{l\in\mathcal{K}_c}\gamma_{l,c} = 1$.

C. AD MIMO-NOMA WITH SIC

Without loss of generality, let UEs in the c-th NOMA cluster be sorted in the descending order of their channel qualities, i.e., $\text{UE}_{l,c}$ is stronger than $\text{UE}_{t,c}$ $(\forall t > l \in \mathcal{K}_c)$, where the user ordering strategy is detailed later in this subsection. Then, according to the NOMA principle, the PA coefficients assigned to UEs in the c-th cluster are sorted in the ascending order of their channel qualities (to be defined later), i.e., $\gamma_{1,c} \leq \cdots \leq \gamma_{K_c,c}$. Thus, using SIC decoding at the receiver side, every $\text{UE}_{l,c}$ can remove the interference coming from the weaker UEs before decoding its own signal.

In the case of single-antenna NOMA, the channel quality of $\text{UE}_{l,c}$ is the corresponding channel gain $|\mathbf{h}_{l,c}|$. Then, $|\mathbf{h}_{1,c}| \geq \cdots \geq |\mathbf{h}_{K_c,c}|$. And, SIC works properly if the PA coefficients satisfy the following condition [26]

$$\frac{\gamma_{l,c}p_c|\mathbf{h}_{l-1,c}|^2}{\sigma_n^2} - \sum_{m=1}^{l-1}\frac{\gamma_{m,c}p_c|\mathbf{h}_{l-1,c}|^2}{\sigma_n^2} \geq P_{\min},\ l = 2,\cdots K_c, \quad (6)$$

where $P_{\min}$ is the minimum power difference between $\text{UE}_{l,c}$ and UEs with lower channel quality. Thus, SIC at $\text{UE}_{l,c}$ can decode first the signals of weak UEs with high PA, i.e., $\text{UE}_{t,c}$ $(\forall t > l)$, and then subtract them from the superposed signal to decode its own signal.

In the case of multi-antenna NOMA with a single beam, e.g., using analog beamforming, the users are ordered according to their effective channel gain $|\mathbf{h}_{l,c}\mathbf{a}_c|^2$ or their distance w.r.t. the BS, both are considered in [18]. However, in the AD MIMO-NOMA system, UE also suffers from the inter-cluster interference as seen in (3), unlike the single-antenna NOMA or the multi-antenna NOMA with a single beam. Therefore, the ordering strategy of the NOMA users and the constraint of efficient SIC in (6) must consider this interference. To this end, we generalize the channel quality $\zeta_{l,c}$ metric as the superposed-signal to other-clusters interference plus noise ratio at $\text{UE}_{l,c}$, as follows

$$\zeta_{l,c} = \frac{\eta p_c|\mathbf{h}_{l,c}\mathbf{a}_c|^2}{\sum\limits_{b\in\mathcal{C},b\neq c}\eta p_b|\mathbf{h}_{l,c}\mathbf{a}_b|^2 + \sigma_n^2}. \quad (7)$$

And, we order UEs such that $\zeta_{1,c} \geq \zeta_{2,c} \geq \cdots \geq \zeta_{K_c,c}$. Besides, the constraint of efficient SIC in the c-th cluster for MIMO-NOMA is given by

$$\gamma_{l,c}\zeta_{l-1,c} - \sum_{m=1}^{l-1}\gamma_{m,c}\zeta_{l-1,c} \geq P_{\min}, \quad l = 2,\cdots K_c. \quad (8)$$

This user ordering strategy thereby requires FCSI at the transmitter, which involves a high link overhead. For that, in Section III, we propose another ordering strategy with only ADI feedback.

Assuming a perfect SIC decoding, the received signal $y_{l,c}$ at $\text{UE}_{l,c}$ in (3) can be rewritten as

$$y_{l,c} = \underbrace{\sqrt{\eta}\sqrt{\gamma_{l,c}p_c}\mathbf{h}_{l,c}\mathbf{a}_c s_{l,c}}_{\text{desired signal}} + \underbrace{\sum_{m=1}^{l-1}\sqrt{\eta}\sqrt{\gamma_{m,c}p_c}\mathbf{h}_{l,c}\mathbf{a}_c s_{m,c}}_{\text{intra-cluster interference}} + \underbrace{\sum_{b\in\mathcal{C},b\neq c}\sqrt{\eta}\sqrt{p_b}\mathbf{h}_{l,c}\mathbf{a}_b\mathbf{s}_b}_{\text{inter-cluster interference}} + n_{l,c}. \quad (9)$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3112225, IEEE Access

Accordingly, $\text{UE}_{l,c}$ in the c-th NOMA cluster has the following signal-to-interference-plus-noise ratio (SINR), while decoding its own message

$$\begin{aligned}\text{SINR}_{l,c}^{\text{NOMA}} &= \frac{\eta\gamma_{l,c}p_c|\mathbf{h}_{l,c}\mathbf{a}_c|^2}{\sum_{\substack{b\in\mathcal{C}\\ b\neq c}}\eta p_b|\mathbf{h}_{l,c}\mathbf{a}_b|^2+\sum_{m=1}^{l-1}\eta\gamma_{m,c}p_c|\mathbf{h}_{l,c}\mathbf{a}_c|^2+\sigma_n^2},\\ &= \frac{\gamma_{l,c}\zeta_{l,c}}{\sum_{m=1}^{l-1}\gamma_{m,c}\zeta_{l,c}+1}.\end{aligned} \tag{10}$$

In (10), the first term in the denominator is the inter-cluster interference, and the second one is the residual intra-cluster interference after SIC.

For the UE in the c-th SU cluster, $\text{SINR}_{1,c}^{\text{SU}}$ is given by

$$\text{SINR}_{1,c}^{\text{SU}} = \frac{\eta p_c|\mathbf{h}_{1,c}\mathbf{a}_c|^2}{\sum_{b\in\mathcal{C},b\neq c}\eta p_b|\mathbf{h}_{1,c}\mathbf{a}_b|^2+\sigma_n^2} = \zeta_{1,c}. \tag{11}$$

Considering a successful decoding at each UE and no propagation error within the total bandwidth B, $\text{UE}_{l,c}$ achieves the following throughput according to Shannon formula

$$R_{l,c} = B\log_2(1+\text{SINR}_{l,c}). \tag{12}$$

Therefore, the system sum-throughput R can be expressed as

$$R = \sum_{c\in\mathcal{C}}\sum_{l\in\mathcal{K}_c}R_{l,c}, \tag{13a}$$

$$= \underbrace{\sum_{c\in\mathcal{C}_s}B\log_2(1+\zeta_{1,c})}_{\text{SU clusters}} + \underbrace{\sum_{c\in\mathcal{C}_n}\sum_{l\in\mathcal{K}_c}B\log_2\left(1+\frac{\gamma_{l,c}\zeta_{l,c}}{\sum_{m=1}^{l-1}\gamma_{m,c}\zeta_{l,c}+1}\right)}_{\text{NOMA clusters}}. \tag{13b}$$

As seen in (13b), R depends on the PA coefficients $\boldsymbol{\gamma} = \{\gamma_{l,c},\ c\in\mathcal{C}_n,\ l\in\mathcal{K}_c\}$ and on the UEs' channel qualities $\boldsymbol{\zeta} = \{\zeta_{l,c},\ c\in\mathcal{C},\ l\in\mathcal{K}_c\}$ built using FCSI $\mathbf{h}_{l,c}$ as shown in (7). Therefore, R can be improved by carefully designing a user clustering (UC) algorithm and a power allocation (PA) technique, including inter-cluster PA $\boldsymbol{p} = \{p_c,\ c\in\mathcal{C}\}$ and intra-cluster PA $\boldsymbol{\gamma}$ strategies. However, since we seek an AD MIMO-NOMA system only based on ADI, in the next section, we derive a channel quality metric (for user ordering strategy), a sum-throughput expression (for PA), and a spatial interference metric (for UC), all only based on $\boldsymbol{\Theta}$.

III. ADI-BASED PERFORMANCE METRICS

In this section, we derive ADI-based performance metrics: the user channel quality, the sum-throughput, and the inter-user spatial interference, only based on ADI. These metrics will be used in the user ordering strategy in Section III-A and the design of UC and PA methods in Section IV and V, respectively.

A. ADI-BASED CHANNEL QUALITY METRIC

The FCSI-based channel quality $\zeta_{l,c}$ in (7) is based on the channel vector $\mathbf{h}_{l,c}$, including both LoS and NLoS paths. However, AD-DBF doesn't exploit the NLoS paths and generates a single beam in the LOS path direction. To maintain low channel feedback, we consider a simplifying assumption that the NLOS path gains are negligible compared to those of the LOS path, i.e., $\frac{\alpha_{q,l,c}}{\alpha_{1,l,c}} \to 0,\ \forall 1 < q \leq N_{l,c}$. It is reasonable in mmWave channels since the transmission is highly directional at mmWave frequencies [16]. In addition, the authors in [27] show that the amplitudes of the NLoS paths are 20 dB weaker than the LoS path. Thereby, $\mathbf{h}_{l,c}\mathbf{a}_c$ can be simplified as

$$\mathbf{h}_{l,c}\mathbf{a}_c \overset{\frac{\alpha_{q,l,c}}{\alpha_{1,l,c}}\to 0}{\longrightarrow} \alpha_{1,l,c}\mathbf{a}_{1,l,c}^H\mathbf{a}_c. \tag{14}$$

(14) includes the gain $\alpha_{1,l,c}$ and the spatial angle $\theta_{1,l,c}$ of the LoS path. By substituting (14) in (7), we obtain the following channel quality $\breve{\zeta}_{l,c}$ of $\text{UE}_{l,c}$

$$\zeta_{l,c} \overset{\frac{\alpha_{q,l,c}}{\alpha_{1,l,c}}\to 0}{\longrightarrow} \breve{\zeta}_{l,c} = \frac{\eta p_c|\alpha_{1,l,c}\mathbf{a}_{1,l,c}^H\mathbf{a}_c|^2}{\sum_{b\in\mathcal{C},b\neq c}\eta p_b|\alpha_{1,l,c}\mathbf{a}_{1,l,c}^H\mathbf{a}_b|^2+\sigma_n^2}. \tag{15}$$

In high signal-to-noise ratio (SNR) regime, the channel quality $\breve{\zeta}_{l,c}$ can be approximated by

$$\breve{\zeta}_{l,c} \approx \frac{p_c|\mathbf{a}_{1,l,c}^H\mathbf{a}_c|^2}{\sum_{b\in\mathcal{C},b\neq c}p_b|\mathbf{a}_{1,l,c}^H\mathbf{a}_b|^2}. \tag{16}$$

As seen in (16), $\breve{\zeta}_{l,c}$ depends on the inter-cluster PA strategy $\boldsymbol{p}$, the spatial direction of the considered UE $\theta_{1,l,c}$ and the angular information of each cluster θ_b $(\forall b\in\mathcal{C})$.

The prior works on MIMO-NOMA with limited feedback use the angular distance [28], the distance from the BS [18], [29], the large-scale fading coefficients [30] or the effective channel gains $|\mathbf{h}_{l,c}\mathbf{a}_c|^2$ [23], [31] as the user channel quality, to determine the order of SIC decoding. Different from these works, we leverage the highly directionality of the transmission properties and we propose a sub-optimal user ordering strategy that orders the users in the c-th NOMA cluster according to the following criterion

$$\breve{\zeta}_{1,c} \geq \breve{\zeta}_{2,c} \geq \cdots \geq \breve{\zeta}_{K_c,c}. \tag{17}$$

Interestingly, unlike prior works and our previous work [23], this strategy considers the inter-cluster spatial interference, represented by the term in the denominator of (16). However, the distance or the path gain is not needed anymore. And, we will show in Section VI that it is a winning strategy in the context of congested cells where the system is in an interference-limited regime.

B. ADI-BASED SUM-THROUGHPUT METRIC

The objective of this work is to design a PA method that maximizes the system sum-throughput with ADI feedback. However, the objective function R in (13) requires FCSI, and

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3112225, IEEE Access

hence, we obtain an FCSI-based PA. To develop an ADI-based PA, we replace (16) with (7) in (13) such that the ADI-based sum-throughput is defined as follows

$$\breve{R} = \sum_{c\in\mathcal{C}_s} B\log_2(1+\breve{\zeta}_{1,c}) + \sum_{c\in\mathcal{C}_n}\sum_{l\in\mathcal{K}_c} B\log_2\left(1+\frac{\gamma_{l,c}\breve{\zeta}_{l,c}}{\sum_{m=1}^{l-1}\gamma_{m,c}\breve{\zeta}_{l,c}+1}\right). \tag{18}$$

We note that $\breve{R}$ is obtained by considering the simplifying assumption $\frac{\alpha_{q,l,c}}{\alpha_{1,l,c}} \to 0$ in a high SNR regime. These assumptions will be verified by simulations in Section VI. $\breve{R}$ depends on ADI $\boldsymbol{\Theta}$ and the PA technique, including inter-cluster PA $\boldsymbol{p}$ and intra-cluster PA $\boldsymbol{\gamma}$ strategies, which will be described in Section V.

C. ADI-BASED INTER-USER SPATIAL INTERFERENCE METRIC

In this subsection, we focus on the system before clustering, and we use the subscript k ($\forall k \in \mathcal{K}$) as UE ID instead of (l,c) used after clustering. The objective is to define an ADI-based inter-user spatial interference metric needed for UC. Most of the existing works on AD MIMO-NOMA utilize the angular distance for UC. However, this metric does not describe the inter-user interference properly when digital beamforming is adopted. Indeed, the beamwidth is very dependent on the antennas' number and the steering angle of the beam. Interestingly, we propose an inter-user spatial interference metric that includes the angular distance in addition to the main-beam width information.

In a mono-path environment [23], we defined an inter-user spatial interference metric as follows

$$\beta_{k,u} = \frac{1}{M}|\mathbf{a}_{1,k}^H\mathbf{a}_{1,u}|,\ k,u\in\mathcal{K}, \tag{19}$$

where $\mathbf{a}_{1,u}$ is the array steering vector of the beam generated toward UE u spatial direction, i.e., $\theta_{1,u}$ and $\mathbf{a}_{1,k}^H$ is the array steering vector of the LoS path in UE k channel. Thus, $\beta_{k,u}$ is the normalized spatial interference of the beam generated toward UE u with the LoS path in UE k channel. By applying the simplifying assumption $\frac{\alpha_{q,l,c}}{\alpha_{1,l,c}} \to 0$ in the multi-path environment, we can use the same definition given by (19) as in [23].

Using (1), the spatial interference $\beta_{k,u}$ in (19) can be rewritten as

$$\beta_{k,u} = \frac{1}{M}\left|\sum_{m=1}^{M} e^{j2\pi(m-1)\frac{d}{\lambda}(\cos(\theta_{1,k})-\cos(\theta_{1,u}))}\right| \tag{20a}$$

$$= \left|\frac{\sin\left(\frac{M\pi d}{\lambda}(\cos(\theta_{1,k})-\cos(\theta_{1,u}))\right)}{M\sin\left(\frac{\pi d}{\lambda}(\cos(\theta_{1,k})-\cos(\theta_{1,u}))\right)}\right|. \tag{20b}$$

$\beta_{k,u}$ only depends on the spatial directions $\theta_{1,k}$ and $\theta_{1,u}$. According to (20b), $\beta_{k,u} = \beta_{u,k} \in (0,1)$, and $\beta_{k,u}$ represents the value of the normalized array factor $AF_{(\theta_{1,u})}(\theta)$ of the beam pointed at the spatial direction $\theta_{1,u}$ for $\theta = \theta_{1,k}$, such that [23]

$$\beta_{k,u} = |AF_{(\theta_{1,u})}(\theta_{1,k})|. \tag{21}$$

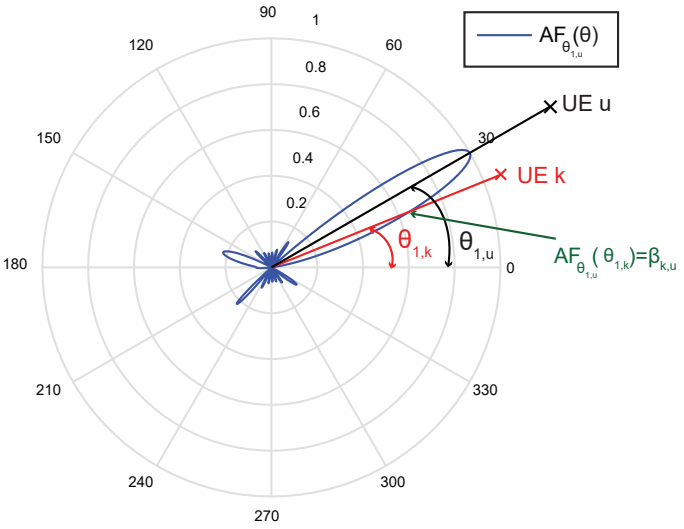


FIGURE 2. Illustration of the normalized array factor $AF_{(\theta_{1,u})}(\theta)$ of the beam pointed toward UE u and the value $AF_{(\theta_{1,u})}(\theta_{1,k})$ for the spatial direction $\theta_{1,k}$ corresponding to $\beta_{k,u}$ of UE k and UE u.

To illustrate (21), we depict the normalized array factor $AF_{(\theta_{1,u})}(\theta)$ (blue curve) of the beam pointed toward UE u and the corresponding value of $\beta_{k,u}$ in Fig. 2.

Remark 1. *Based on (20), we can see that if the angular distance $\Delta\theta = |\theta_{1,k}-\theta_{1,u}| \to 0$, then $\beta_{k,u} \to 1$. Thus, the smaller $\Delta\theta$ is, i.e., UE k and UE u are closely in the same direction, the larger $\beta_{k,u}$ is.*

Remark 2. *$\Omega_u^{3dB} = |\theta_{1,u}-\theta|$ with $AF_{(\theta_{1,u})}(\theta) = \sqrt{\frac{1}{2}}$ gives a measure of the 3dB-width of the beam generated at UE u. This 3dB-beamwidth Ω_u^{3dB} depends on the beam direction $\theta_{1,u}$ and the number of antennas M.*

Thus, β is an important metric to determine the level of the inter-user spatial interference using only the user spatial directions $\theta_{1,k}$ ($\forall k \in \mathcal{K}$), as depicted in Fig. 2. Moreover, we define the interference threshold β_0, such that $\beta_{k,u} \geq \beta_0$ means that the LoS path of UE k lies in the UE u beam. Particularly, the β_0-beamwidth $\Omega_u^{\beta_0}$ defines the angular distance $|\theta_{1,u}-\theta_0|$ satisfying $|AF_{(\theta_{1,u})}(\theta_0)| = |AF_{(\theta_{1,u})}(\theta_{1,u}\pm\Omega_u^{\beta_0})| = \beta_0$. Using a fixed β_0, $\Omega_u^{\beta_0}$ is varying w.r.t. the number of antennas and the spatial direction of UE u. Therefore, with the defined spatial interference metric β, we are able to define a unique threshold β_0 whatever the number of antennas or the users' spatial directions.

Lemma 1. *UE k, having $\beta_{k,u} \geq \beta_0$, is located in the β_0-width of the main lobe of the UE u beam, only if $\beta_0 > 0.217$.*

Proof. With $d = \frac{\lambda}{2}$, a single main lobe exists in $(0,\pi)$, surrounded by a number of side lobes [32]. $\beta_{k,u} \geq \beta_0$ ensures that UE k belongs to $\Omega_u^{\beta_0}$. However, $\Omega_u^{\beta_0}$ can be associated to the main lobe or one of the side lobes. To select only the users covered by the main lobe, β_0 must be

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3112225, IEEE Access

lower bounded by the first-side lobe (FSL) level β^{FSL}, i.e., $\beta_0 > \beta^{\text{FSL}}$. This value is given by $\beta^{\text{FSL}} = 0.217$ for any values of the antennas number M and the steering angle $\theta_{1,u}$ [32], when $d = \frac{\lambda}{2}$, as illustrated by Fig. 3. Here, Lemma 1 is proved. □

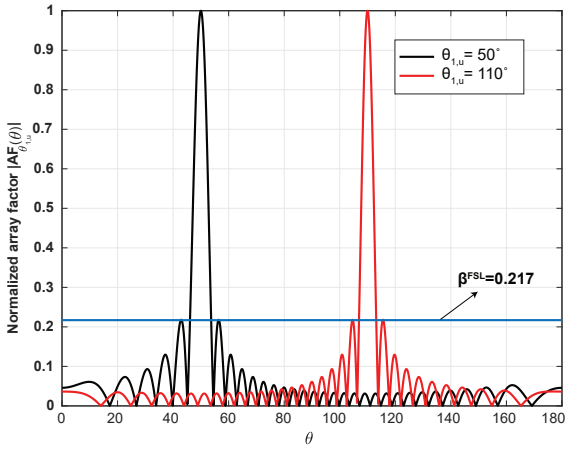


FIGURE 3. Illustration of the normalized array factor $|AF^{n}_{(\theta_{1,u})}(\theta)|$ of the beam generated toward UE u for $\theta_{1,u} \in \{50°, 110°\}$, with $M = 32$.

We will use this metric in the design of the UC algorithm described in Section IV. The value 0.217 in Lemma 1 will be observed by simulations in Section VI.

IV. USER CLUSTERING ALGORITHM

The UC optimization problem that maximizes the sum-throughput R in (13) is a non-convex problem, which can only be solved using an exhaustive search. However, that would result in an exponential complexity, brought by calculating the sum-throughput among all the possible sets of clusters. Indeed, the NOMA cluster size is not restricted in this paper, i.e., $1 \leq K_c \leq K$, hence the number of all the possible combinations increases exponentially with the number of users, K. The previous UC methods using FCSI or partial CSI [33]–[35] [36], such as the large-scale fading coefficients or the channel correlation, are based on the K-means or agglomerative algorithms, which require the target number of clusters as an input parameter. However, for practical applications, this parameter is a priori unknown at the BS and an arbitrary value may lead to suboptimal solutions. To reduce the exponential complexity of the exhaustive search method on the one hand, and on the other hand the CSI overhead and the number of input parameters, we leverage the potentiality of the angular information in the highly directional mmWave transmission. Specifically, we take advantage of the ADI-based inter-user spatial interference metric to group the users in SU and NOMA clusters so that the inter-user interference is reduced, and thereby, the sum-throughput is improved.

To the best of our knowledge, it is used for the first time in the context of UC. Let $\mathcal{G}_c$ be the set of UE IDs in $\mathcal{K}$ belonging to the c-th cluster. Accordingly, if $\mathcal{G}_c\{l\} = k$, then UE $k \in \mathcal{K}$ is the l-th UE in the c-th cluster. We assume that each user belongs to only one cluster, thereby $\mathcal{G}_c \cap \mathcal{G}_{c'} = \emptyset$, $\forall c \neq c' \in \mathcal{C}$ and $\cup_{c=1}^{C} \mathcal{G}_c = \mathcal{K}$. Let β_0 be the inter-user spatial interference threshold. Our proposed UC algorithm, denoted as $\beta-$UC, aims to find the set $\mathcal{G}_c$ $\forall c \in \mathcal{C}$, so that all UEs having interference with each others are grouped in the same cluster. To do so, we first find the NOMA clusters by using (22) given at the top of the next page, while starting by the pair of UEs (o, r) having the largest spatial interference such that $\beta_{o,r} \geq \beta_0$. Once all the NOMA clusters are obtained, we assign to each one of the remaining UEs a single cluster as mentioned in (23), which is given at the top of the next page. The beam angle associated to each cluster is calculated as in (4). More details of the proposed $\beta-$UC algorithm are given in Algorithm 1.

The computational complexity of the proposed UC algorithm is dominated by phase 1: clustering users with high inter-user interference into NOMA clusters. This phase is a nested loop, where the outer loop finds the C_n NOMA clusters, and the inner loop searches the users belonging to the current c-th NOMA cluster with $\lfloor \frac{K_c}{2} \rfloor$ iterations, where $\lfloor \cdot \rfloor$ stands for the integer part. Complexity is essentially present in the inner loop at step 5 that carries out $\mathcal{O}(|\mathcal{F}|)$ searches to find the highest inter-user spatial interference $\beta_{o,r}$ in $\mathcal{F}$. Thereby, the complexity order of the UC algorithm is given by

$$\mathcal{O}\left(\sum_{c=1}^{C_n}\sum_{p=1}^{\lfloor \frac{K_c}{2} \rfloor} \mathbf{card}(\mathcal{F}_{p,c})\right), \tag{24}$$

where $\mathcal{F}_{p,c}$ denotes the set $\mathcal{F}$ corresponding to the c-th and p-th iterations of the outer and the inner loop, respectively.

Lemma 2. *Once $\beta-$UC is applied with $\beta_0 > \beta^{FSL}$, UEs in the c-th NOMA cluster are located within the β_0-width of the main lobe corresponding to the c-th NOMA cluster, i.e., $\beta_{g,c} \overset{def}{=} |AF_{(\theta_c)}(\theta_{1,g})| \geq \beta_0$ $\forall g \in \mathcal{G}_c$.*

Proof. The set $\mathcal{G}_c$ of UE IDs within the c-th NOMA cluster satisfies (22), thereby $\beta_{g,d} \geq \beta_0$ $\forall d \neq g \in \mathcal{G}_c$. Let $\min_{g \in \mathcal{G}_c}\{\theta_{1,g}\} = \theta_{1,v}$ and $\max_{g \in \mathcal{G}_c}\{\theta_{1,g}\} = \theta_{1,w}$, then the spatial angle θ_c corresponding to the c-th NOMA cluster is equal to $\theta_c = \frac{\theta_{1,v}+\theta_{1,w}}{2}$, as defined in (4).

We define $\beta_{v,c} \overset{def}{=} |AF_{(\theta_c)}(\theta_{1,v})|$, which can be rewritten by $\beta_{v,c} = |AF_{(\theta_{1,v})}(\theta_c)|$, with $|AF_{(\theta_{1,v})}(\theta)|$ is the normalized array factor of the beam steered toward UE v. Now, we aim to demonstrate that $|AF_{(\theta_{1,v})}(\theta_c)| \geq \beta_0$. Since $\beta_{v,w} \geq \beta_0$ and $\beta_0 > \beta^{\text{FSL}}$, then UE w is covered by the β_0-width of the main lobe of the UE v beam according to Lemma 1. Therefore, and by exploiting the property of the function in (20b), $|AF_{(\theta_{1,v})}(\theta)|$ is a decreasing function of θ in the angle interval $I = [\theta_{1,v}, \theta_{1,w}]$. Accordingly, $|AF_{(\theta_{1,v})}(\theta_c)| > |AF_{(\theta_{1,v})}(\theta_{1,w})| \geq \beta_0$.

Similarly, we can demonstrate that $|AF_{(\theta_{1,w})}(\theta_c)| \geq \beta_0$.

However, $|AF_{(\theta_c)}(\theta_{1,v})| = |AF_{(\theta_{1,v})}(\theta_c)| \geq \beta_0$ and $|AF_{(\theta_c)}(\theta_{1,w})| = |AF_{(\theta_{1,w})}(\theta_c)| \geq \beta_0$. Generally, $\theta_{1,v} \leq \theta_{1,g} \leq \theta_{1,w}$ $\forall g \in \mathcal{G}_c$, and $|AF^{n}_{(\theta_c)}(\theta)|$ has one maximum at

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3112225, IEEE Access

$$\mathcal{G}_c = \begin{cases} \{g : g \in \mathcal{K}, g \notin \mathcal{G}_j\ (\forall j < c), \beta_{g,d} \geq \beta_0\ \forall d \neq g \in \mathcal{G}_c\}, & \forall c \in \mathcal{C}_n; \quad (22) \\ \{g : g \in \mathcal{K}, g \notin \mathcal{G}_j\ (\forall j < c), |\mathcal{G}_c| = 1\}, & \forall c \in \mathcal{C}_s. \quad (23) \end{cases}$$

$\theta = \theta_c$ in the interval I. This implies that $|AF_{(\theta_c)}(\theta_{1,g})| \geq \min\{|AF_{(\theta_c)}(\theta_{1,v})|, |AF_{(\theta_c)}(\theta_{1,w})|\} \geq \beta_0\ \forall g \in \mathcal{G}_c$. In other words, UE g is located within the β_0-width of the main beam corresponding to the c-th NOMA cluster, according to Lemma 1. Here, Lemma 2 is proved. □

If the condition in Lemma 1 is met, i.e., $\beta_{k,u} \geq \beta_0$ with $\beta_0 > 0.217$, Lemma 2 ensures that the choice of the cluster angle in (4) guarantees that the users are covered by the main lobe corresponding to the c-th NOMA cluster within the β_0-width. This in turn ensures that the beam misalignment is not severe.

V. MULTI-USER POWER ALLOCATION TECHNIQUE

In this section, we consider FCSI- and ADI-based PA problems that maximize the FCSI- and the ADI-based sum-throughput, respectively. Throughout this section, we propose a fixed inter-cluster PA strategy to make a convex optimization problem. Then, the closed-form intra-cluster PA can be given using the KKT conditions. To the best of our knowledge, this is the first work that utilizes the ADI-based sum-throughput to design a PA policy with ADI feedback.

A. PROBLEM FORMULATION

Given the set $\mathcal{G}_c$ $(\forall c \in \mathcal{C})$, we aim to design a PA technique that maximizes the system sum-throughput of the multi-user AD MIMO-NOMA scheme. The PA technique consists to allocate the power between the clusters $\boldsymbol{p}$ and between UEs in each NOMA cluster $\boldsymbol{\gamma}$. By defining P_c as the emitted power toward the c-th cluster as follows

$$P_c = \eta \|\mathbf{w}_c\|^2 \sum_{l=1}^{K_c} \gamma_{l,1} p_c = \eta \|\mathbf{w}_c\|^2 p_c, \quad (25)$$

the optimization problem can be formulated as follows

$$(\mathcal{P}_1): \quad \{\overset{\star}{\boldsymbol{\gamma}}, \overset{\star}{\boldsymbol{p}}\}^{\text{type}} = \max_{\boldsymbol{\gamma}, \boldsymbol{p}} R^{\text{type}}(\boldsymbol{\gamma}, \boldsymbol{p}), \quad (26)$$

$$\text{s.t.} \quad \sum_{l \in \mathcal{K}_c} \gamma_{l,c} = 1,\ \forall c \in \mathcal{C}, \quad (26a)$$

$$\sum_{c \in \mathcal{C}} P_c \leq P_e, \quad (26b)$$

$$\delta(\boldsymbol{\gamma}, \zeta^{\text{type}}_{l-1,c}) \geq P_{\min},\ 2 \leq l \leq K_c,\ \forall c \in \mathcal{C}_n. \quad (26c)$$

where type $= \{\text{FCSI, ADI}\}$, $R^{\text{FCSI}} = R$ in (13), $R^{\text{ADI}} = \breve{R}$ in (18), $\zeta^{\text{FCSI}}_{l,c} = \zeta_{l,c}$ in (7), $\zeta^{\text{ADI}}_{l,c} = \breve{\zeta}_{l,c}$ in (16), P_e is the total transmit power and $\delta(\boldsymbol{\gamma}, \zeta^{\text{type}}_{l-1,c}) = \gamma_{l,c}\zeta^{\text{type}}_{l-1,c} - \sum_{m=1}^{l-1} \gamma_{m,c}\zeta^{\text{type}}_{l-1,c}$. (26a) and (26b) are respectively the total intra-cluster and inter-cluster PA constraints, while (26c) is the SIC constraint given by (8). $(\mathcal{P}_1)$ maximizes the system sum-throughput by jointly optimizing intra- and inter-cluster PA. However, the PA factors from different UEs are very coupled in (26). To obtain closed-form solutions, we apply a fixed inter-cluster PA[4], i.e., the c-th cluster is assigned with a predefined fixed power p_c, as detailed in Section V-B. Hence, $(\mathcal{P}_1)$ is reformulated as an intra-cluster PA maximization problem subject to the intra-cluster PA and efficient SIC constraints as follows

$$(\mathcal{P}_2): \quad \{\overset{\star}{\boldsymbol{\gamma}}\}^{\text{type}} = \max_{\boldsymbol{\gamma}} R^{\text{type}}(\boldsymbol{\gamma}), \quad (27)$$

$$\text{s.t.} \quad \sum_{l \in \mathcal{K}_c} \gamma_{l,c} = 1,\ \forall c \in \mathcal{C} \quad (27a)$$

$$\delta(\boldsymbol{\gamma}, \zeta^{\text{type}}_{l-1,c}) \geq P_{\min},\ 2 \leq l \leq K_c,\ \forall c \in \mathcal{C}_n. \quad (27b)$$

Denote R^{type}_c as the sum-throughput of UEs within the c-th cluster. For SU cluster, $R^{\text{type}}_c = B \log_2(1 + \zeta^{\text{type}}_{1,c}) =$ constant $\forall \boldsymbol{\gamma}$. For NOMA cluster, $R^{\text{type}}_c = \sum_{l \in \mathcal{K}_c} B \log_2 \left(1 + \frac{\gamma_{l,c}\zeta^{\text{type}}_{l,c}}{\sum_{m=1}^{l-1} \gamma_{m,c}\zeta^{\text{type}}_{l,c} + 1}\right)$ depends only on the PA coefficients $\boldsymbol{\gamma}_c = \{\gamma_{l,c},\ l \in \mathcal{K}_c\}$ of UEs in the c-th cluster. And, $\delta(\boldsymbol{\gamma}_c, \zeta^{\text{type}}_{l-1,c}) = \delta(\boldsymbol{\gamma}, \zeta^{\text{type}}_{l-1,c})$. Consequently, solving $(\mathcal{P}_2)$ is equivalent to solve C_n problems, independently. Therefore, for each NOMA cluster, the sum-throughput optimization problem defining the intra-cluster PA strategy $\{\overset{\star}{\boldsymbol{\gamma}}_c\}^{\text{type}}\ \forall c \in \mathcal{C}_n$ can be modeled as

$$(\mathcal{P}_3): \{\overset{\star}{\boldsymbol{\gamma}}_c\}^{\text{type}} = \max_{\boldsymbol{\gamma}_c} \sum_{l \in \mathcal{K}_c} B \log_2 \left(1 + \frac{\gamma_{l,c}\zeta^{\text{type}}_{l,c}}{\sum_{m=1}^{l-1} \gamma_{m,c}\zeta^{\text{type}}_{l,c} + 1}\right), \quad (28)$$

$$\text{s.t.} \quad \sum_{l \in \mathcal{K}_c} \gamma_{l,c} = 1, \quad (28a)$$

$$\delta(\boldsymbol{\gamma}_c, \zeta^{\text{type}}_{l-1,c}) \geq P_{\min},\ 2 \leq l \leq K_c. \quad (28b)$$

After designing a fixed inter-cluster PA strategy in Section V-B, we aim in Section V-C to design an intra-cluster PA strategy by solving $(\mathcal{P}_3)$ using the KKT conditions.

B. FIXED INTER-CLUSTER POWER ALLOCATION

In our previous work with two users per NOMA clusters [23], we applied a uniform PA per cluster, i.e., $p_c =$ constant $(\forall c)$. Moreover, the inter-cluster power was allocated such that $\sum_{c=1}^{C} P_c = P_e$. This implied that the uniform PA per cluster satisfies $P_c = \frac{P_e}{C}$, using (25) and $\eta \|\mathbf{w}_c\|^2 = \frac{1}{C}$. In multi-user AD MIMO-NOMA scheme, the number of UEs per cluster is variable. So, we allocate a power p_c to the c-th cluster proportional to the number K_c of UEs served in the c-th cluster. Accordingly, based on (25), P_c is also proportional

[4] An iterative sub-optimal inter-cluster PA solution is left for future work.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3112225, IEEE Access

Algorithm 1 β-based User Clustering Algorithm

Input: $\theta_{1,k}$, $\beta_{k,u}$ $\forall k < u \in \mathcal{K}$, β_0.
Output: C_n, C, $\mathcal{G}_c$ $\forall c \in \mathcal{C}$.
Initialization : $c = 0$
Select the set of groups of 2-UEs having a spatial interference greater than β_0 s.t. $\mathcal{D} = \{(k,u),\ \beta_{k,u} \geq \beta_0, k < u \in \mathcal{K}\}$.
Phase 1: Clustering of users with high inter-user spatial interference in NOMA clusters
repeat
  $c = c + 1$, $\mathcal{F} = \mathcal{D}$, $\mathcal{G}_c = \{\}$;
  repeat
    Locate in $\mathcal{F}$ the group of 2-UEs having the largest spatial interference: $\beta_{o,r} = \max_{(k,u)\in\mathcal{F}} \{\beta_{k,u}\}$, $\mathcal{G}_c = \{\mathcal{G}_c, o, r\}$.
    Remove from $\mathcal{D}$ the couple (o,r) and their coupling with any other UEs to prevent their presence in another cluster:
    $\mathcal{D} \leftarrow \mathcal{D} - \{(o,r)\} - \{(o,w)\} - \{(v,o)\} - \{(r,m)\} - \{(t,r)\}$ $(\forall w > o, v < o, m > r, t < r \in \mathcal{K})$
    Select the set of UEs in interference with UE o:
    $\mathcal{V} = \{v : (o,v) \in \mathcal{F}, o < v$ or $(v,o) \in \mathcal{F}, o > v\}$.
    Select the set of UEs in interference with UE r:
    $\mathcal{W} = \{w : (r,w) \in \mathcal{F}, r < w$ or $(w,r) \in \mathcal{F}, r > w\}$
    Select UEs in interference with both UE$_o$ and UE$_r$:
    $\mathcal{Y} = \mathcal{V} \cap \mathcal{W}$
    if $|\mathcal{Y}| == 1$ **then**
      // One user exists in $\mathcal{F}$
      Update $\mathcal{F}$ and $\mathcal{G}_c$:
      $\mathcal{F} = \emptyset$, $y = \mathcal{Y}$, $\mathcal{G}_c = [\mathcal{G}_c, y]$,
      Remove from $\mathcal{D}$ the coupling of UE y with any other UEs to prevent their presence in another cluster:
      $\mathcal{D} \leftarrow \mathcal{D} - \{(y,w)\} - \{(v,y)\}$ $(\forall w > y, v < y \in \mathcal{K})$.
    else
      // Multiple users exist in $\mathcal{F}$
      Update $\mathcal{F}$ and select the set of groups of 2-UEs in $\mathcal{Y}$ who have a spatial interference greater than β_0:
      $\mathcal{F} = \{(k,u),\ \beta_{k,u} \geq \beta_0, k < u \in \mathcal{Y}\} - \{(o,r)\}$.
    end if
  until $\mathcal{F} = \emptyset$
until $\mathcal{D} = \emptyset$
$C_n = c$.
Phase 2: Clustering of the remaining users in SU clusters
for $k = 1 : K$ **do**
  if $k \notin \cup_{l=1}^{c} \mathcal{G}_l$ **then**
    $c = c + 1$; $\mathcal{G}_c = \{k\}$.
  end if
end for
$C = c$.

to K_c. This is supposed to guarantee a certain power fairness among UEs. Therefore, we calculate P_c as follows

$$P_c = K_c \frac{P_e}{K}. \tag{29}$$

From (25) and (29) we obtain that

$$p_c = K_c C \frac{P_e}{K}. \tag{30}$$

Thus, unlike [23], the power is non uniformly allocated to the clusters.

C. FCSI- AND ADI-BASED INTRA-CLUSTER POWER ALLOCATION

We apply a similar methodology as in [26][5] to find the closed-form solution of $(\mathcal{P}_3)$, since it is convex under the constraints (28a) and (28b). Thus, the corresponding Lagrange function $\mathcal{L}^{\text{type}}$ is defined as follows

$$\mathcal{L}^{\text{type}}(\boldsymbol{\gamma}_c, \rho, \boldsymbol{\chi}) = B\sum_{l=1}^{K_c} \log_2\left(1 + \frac{\gamma_{l,c}\zeta_{l,c}^{\text{type}}}{\sum_{m=1}^{l-1}\gamma_{m,c}\zeta_{l,c}^{\text{type}} + 1}\right) + \rho\left(1 - \sum_{l=1}^{K_c}\gamma_{l,c}\right) + \sum_{l=2}^{K_c}\chi_l\left\{\gamma_{l,c}\zeta_{l-1,c}^{\text{type}} - \sum_{m=1}^{l-1}\gamma_{m,c}\zeta_{l-1,c}^{\text{type}} - P_{\min}\right\} \tag{31}$$

where ρ and $\boldsymbol{\chi} = \{\chi_l,\ l = 2, \cdots K_c\}$ are the Lagrange multipliers corresponding to the constraints (28a) and (28b), respectively, such that $\rho > 0$ and $\chi_l \geq 0$ $(i = 2, \cdots K_c)$. The optimal PA solution satisfies the following KKT conditions

$$\frac{\partial\mathcal{L}^{\text{type}}}{\partial\gamma_{1,c}} = \frac{B\zeta_{1,c}^{\text{type}}}{\gamma_{1,c}\zeta_{1,c}^{\text{type}} + 1} - \sum_{m=2}^{K_c}\frac{B\gamma_{m,c}(\zeta_{m,c}^{\text{type}})^2}{(\sum_{q=1}^{m}\gamma_{q,c}\zeta_{m,c}^{\text{type}} + 1)(\sum_{q'=1}^{m-1}\gamma_{q',c}\zeta_{m,c}^{\text{type}} + 1)} - \rho - \sum_{m=2}^{K_c}\chi_m\zeta_{m-1,c}^{\text{type}} = 0, \text{ if } \gamma_{1,c} \geq 0, \tag{32}$$

$$\frac{\partial\mathcal{L}^{\text{type}}}{\partial\gamma_{l,c}} = \frac{B\zeta_{l,c}^{\text{type}}}{\sum_{m=1}^{l}\gamma_{m,c}\zeta_{l,c}^{\text{type}} + 1} - \rho + \chi_l(\zeta_{l-1,c}^{\text{type}}) - \sum_{m=l+1}^{K_c}\frac{B\gamma_{m,c}(\zeta_{m,c}^{\text{type}})^2}{(\sum_{q=1}^{m}\gamma_{q,c}\zeta_{m,c}^{\text{type}} + 1)(\sum_{q'=1}^{m-1}\gamma_{q',c}\zeta_{m,c}^{\text{type}} + 1)} - \sum_{m=l+1}^{K_c}\chi_m(\zeta_{m-1,c}^{\text{type}}) = 0, \text{ if } \gamma_{l,c} \geq 0, \ \forall l = 2, \cdots, K_c, \tag{33}$$

[5]In [26], the authors consider a single-antenna NOMA system in both uplink and downlink. Given the set of NOMA clusters, they derive a PA policy that maximizes the sum-throughput per NOMA cluster under three constraints: (1) minimum data rate requirements, (2) transmission power and (3) efficient SIC constraints. The closed-form solutions for optimal PA are derived using KKT conditions for any cluster size.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3112225, IEEE Access

$$\frac{\partial \mathcal{L}^{\text{type}}}{\partial \rho} = f(\boldsymbol{\gamma}) = 1 - \sum_{l=1}^{K_c} \gamma_{l,c} = 0, \text{ if } \rho > 0, \tag{34}$$

$$\frac{\partial \mathcal{L}^{\text{type}}}{\partial \chi_l} = g_l(\boldsymbol{\gamma}) = \gamma_{l,c}\zeta_{l-1,c}^{\text{type}} - \sum_{m=1}^{l-1} \gamma_{m,c}\zeta_{l-1,c}^{\text{type}} - P_{\min} \geq 0, \text{ if } \chi_l \geq 0, \ \forall l = 2, \cdots, K_c, \tag{35}$$

$$\chi_l g_l(\boldsymbol{\gamma}) = 0, \ \forall l = 2, \cdots K_c. \tag{36}$$

As seen in $(\mathcal{P}_3)$, we have $K_c - 1$ inequality constraints, leading to 2^{K_c-1} combinations of Lagrange multipliers as stated in (36). Note that $\gamma_{l,c} > 0 \ \forall l \in \mathcal{K}_c$ in our problem, then obtaining optimal feasible solutions for K_c decision variables requires exactly K_c equations [37]. This implies that only one combination can be checked to find a solution for the problem $(\mathcal{P}_3)$, i.e., $\chi_l > 0$ and $g_l(\boldsymbol{\gamma}) = 0 \, \forall l \in \{2, \cdots K_c\}$. This is one of the several combinations in [26], where only the equations of the total power and SIC constraints are considered to find the optimal solutions. In other words, $f(\boldsymbol{\gamma}) = 0$ and $g_l(\boldsymbol{\gamma}) = 0 \ (\forall l)$, provide the optimal solution, while the other constraints provide the necessary conditions for the obtained optimal solution. Thereby, the closed-form optimal solutions can be calculated similarly to [26] by

$$\overset{\star}{\gamma}{}_{l,c}^{\text{type}} = \begin{cases} \dfrac{1}{2^{Kc-1}} - \displaystyle\sum_{m=1}^{K_c-1} \frac{P_{\min}}{2^m \zeta_{m,c}^{\text{type}}}, & \text{if } l = 1, \\ \dfrac{1}{2^{Kc-l+1}} + \dfrac{P_{\min}}{\zeta_{l-1,c}^{\text{type}}} - \displaystyle\sum_{m=l}^{K_c} \frac{P_{\min}}{2^{m-l+1}\zeta_{m-1,c}^{\text{type}}}, & \text{if } l \neq 1 \in \mathcal{K}_c. \end{cases} \tag{37}$$

For the dynamic intra-cluster PA solutions in [26], the authors verify that the corresponding set of Lagrange multipliers satisfies the KKT conditions for the different combinations. Therefore, since (37) is one of those in [26], the corresponding set of Lagrange multipliers $\{\rho, \boldsymbol{\chi}\}$ also satisfies the KKT conditions. Unlike our previous work [23] applying a fixed intra-cluster PA for the two NOMA users, the proposed PA policy dynamically maximizes the sum-throughput per each NOMA cluster under the SIC constraints. The closed-form solution in (37) for the problem $(\mathcal{P}_3)$ is obtained with a linear complexity w.r.t. the number of all NOMA users, i.e., $\mathcal{O}(\sum_{c=1}^{C_n} K_c)$.

D. PERFORMANCE COMPARISON: FCSI-BASED VERSUS ADI-BASED SCHEMES

In this subsection, we compare the realistically achievable sum-throughput R_Σ^{FCSI} and R_Σ^{ADI} of the multi-user AD MIMO-NOMA schemes using FCSI or ADI feedbacks, respectively. They are expressed as follow

$$\begin{cases} R_\Sigma^{\text{FCSI}} = R(\{\overset{\star}{\boldsymbol{\gamma}}\}^{\text{FCSI}}), & (38) \\ R_\Sigma^{\text{ADI}} = R(\{\overset{\star}{\boldsymbol{\gamma}}\}^{\text{ADI}}), & (39) \end{cases}$$

where $R(\boldsymbol{\gamma})$ is the realistic sum-throughput given by (13), $\{\overset{\star}{\boldsymbol{\gamma}}\}^{\text{FCSI}}$ and $\{\overset{\star}{\boldsymbol{\gamma}}\}^{\text{ADI}}$ are calculated using (37).

UC and inter-cluster PA are identical for both sum-throughput. The difference is in the user ordering strategy and the computation of intra-cluster PA within each NOMA cluster. Likewise, the FCSI- and ADI-based sum-throughput of UEs within the c-th cluster are given by $R_c^{\text{FCSI}} = R_c(\{\overset{\star}{\boldsymbol{\gamma}}_c\}^{\text{FCSI}})$ and $R_c^{\text{ADI}} = R_c(\{\overset{\star}{\boldsymbol{\gamma}}_c\}^{\text{ADI}})$, respectively, with $R_c(\boldsymbol{\gamma}_c) = \sum_{l \in \mathcal{K}_c} B \log_2 \left(1 + \frac{\gamma_{l,c}\zeta_{l,c}^{\text{FCSI}}}{\sum_{m=1}^{l-1} \gamma_{m,c}\zeta_{l,c}^{\text{FCSI}} + 1}\right)$. Obviously, $R_c(\{\overset{\star}{\boldsymbol{\gamma}}_c\}^{\text{FCSI}}) \geq R_c(\{\overset{\star}{\boldsymbol{\gamma}}_c\}^{\text{ADI}}) \ \forall c \in \mathcal{C}_n$. The sum-throughput gain ΔR_Σ defined as $\Delta R_\Sigma = R_\Sigma^{\text{FCSI}} - R_\Sigma^{\text{ADI}}$ can be written as

$$\Delta R_\Sigma = \sum_{c \in \mathcal{C}_n} \left(R_c(\{\overset{\star}{\boldsymbol{\gamma}}_c\}^{\text{FCSI}}) - R_c(\{\overset{\star}{\boldsymbol{\gamma}}_c\}^{\text{ADI}}) \right) \geq 0. \tag{40}$$

The gain ΔR_Σ only results from the sum-throughput change within the NOMA clusters due to different intra-cluster PA and user ordering strategies. Consequently, ΔR_Σ rises as the number of NOMA clusters grows.

In this paper, we apply the proposed schemes in a mmWave rural environment, using NYUSIM. Thus, a single LoS path exists solely or with another NLoS path [24]. Thereby, we analyze ΔR_Σ for the two following cases: (i) mono-path environment, i.e., $N_{l,c} = 1$ and (ii) multi-path environment with $N_{l,c} = 2$.

1) Mono-path environment

In mono-path environments, $\mathbf{h}_{l,c}\mathbf{a}_c$ is given by (14). And, in a high SNR regime, we obtain that $\boldsymbol{\zeta}^{\text{ADI}} \approx \boldsymbol{\zeta}^{\text{FCSI}}$. Thereby, the same user ordering strategy is applied using FCSI or ADI feedbacks, and we obtain $\overset{\star}{\boldsymbol{\gamma}}{}^{\text{ADI}} \approx \overset{\star}{\boldsymbol{\gamma}}{}^{\text{FCSI}}$ according to (37) and $\Delta R_\Sigma \approx 0$. Consequently, the proposed schemes have similar performance in a mono-path environment, using the channel vector $\mathbf{h}_{l,c}$ information or only ADI $\boldsymbol{\Theta}$.

2) Multi-path environment

In multi-path environments, we applied the simplifying assumption $\frac{\alpha_{2,l,c}}{\alpha_{1,l,c}} \to 0$. In practice, $\frac{\alpha_{2,l,c}}{\alpha_{1,l,c}} = \epsilon_0$, with $0 < |\epsilon_0| < 1$, which in turn implies that $\boldsymbol{\zeta}^{\text{ADI}} \neq \boldsymbol{\zeta}^{\text{FCSI}}$, and $\overset{\star}{\boldsymbol{\gamma}}{}^{\text{ADI}} \neq \overset{\star}{\boldsymbol{\gamma}}{}^{\text{FCSI}}$. In a high SNR regime, ΔR only depends on ϵ_0. If $\epsilon_0 \to 0$, then $\Delta R_\Sigma \to 0$. Therefore, in multi-path environments, the performance of ADI-based AD MIMO-NOMA scheme w.r.t. FCSI-based AD MIMO-NOMA depends on the complex gain of the NLoS path, if it exists, against that of the LoS path.

VI. SIMULATION RESULTS

In this section, extensive simulations are conducted to verify the performance of our proposed multi-user AD MIMO-NOMA schemes. The system is simulated in a rural environment over 5000 trials. The channel impulse responses in both time and space are generated using the mmWave channel model called NYUSIM [24], which adopts the time-cluster spatial-lobe approach. The NYUSIM model is built on channel statistics extracted from extensive measurements conducted in various outdoor urban and rural environments at

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3112225, IEEE Access

mmWave frequencies. According to statistical distributions, NYUSIM generates the distance and angles of each user. FCSI and ADI are assumed to be perfectly known at the transmitter. The channel and system parameters are listed in Table 1. We assume that the users belong to the same side of BS, and their angles range from 10° to 165° to prevent the endfire beamforming. The noise power in Table 1 represents the thermal noise power at a temperature of $25^\circ C$.

TABLE 1. Simulation parameters

Number of transmit antennas, M	32
Carrier frequency	28 GHz
Channel bandwidth, B	20 MHz
Cell edge radius	100 m
Transmission power, P_e	30 dBm
Noise power, σ_n^2	-101 dBm
Minimum power difference, $P_{\min}$	1 mW
Number of paths per time cluster in rural environment	$\{1, 2\}$

The proposed multi-user AD MIMO-NOMA schemes with either FCSI or ADI NOMA involve the same ADI-based DBF and UC, with single and multi-user clusters according to Section IV. The main difference is made by the NOMA protocol inside multi-user clusters. For FCSI, user ordering and intra-cluster PA are built based on FCSI, whereas for ADI, only the spatial directions include. Here, the 2-user AD MIMO-NOMA scheme uses the β-UC algorithm developed in [23] and the PA technique designed in Section V with $K_c = 2$ $(\forall c \in \mathcal{C}_n)$.

A. CHANNEL ESTIMATION OVERHEAD

To assess the channel estimation overhead of the simulated schemes, we provide the number of real coefficients needed per channel estimation for each user in Table 2. It is clear that ADI-based schemes require $2M$ fewer coefficients than those based on FCSI since only the user spatial direction is necessary.

	Number of real coefficients per channel estimation for each user
C-AD-DBF	1 $[\theta_{1,k}]$
AD MIMO-OMA	1 $[\theta_{1,k}]$
2-user AD MIMO-NOMA with FCSI NOMA	$2\times M+1$ $[\mathbf{h}_k \in \mathbb{C}^{1\times M}, \theta_{1,k}]$
2-user AD MIMO-NOMA with ADI NOMA	1 [$\theta_{1,k}$]
multi-user AD MIMO-NOMA with FCSI NOMA	$2\times M+1$ $[\mathbf{h}_k \in \mathbb{C}^{1\times M}, \theta_{1,k}]$
multi-user AD MIMO-NOMA with ADI NOMA	1 [$\theta_{1,k}$]

TABLE 2. Channel Estimation Overhead

B. USER ORDERING STRATEGY

In Fig. 4, the impact of various user ordering strategies using FCSI or limited feedback on the 2-user AD MIMO-NOMA spectral efficiency (sum-throughput per bandwidth) is presented. The optimal strategy orders the users w.r.t. their realistic channel qualities ζ, which is given in (7), such that $\zeta_{1,c} \geq \cdots \geq \zeta_{K_c,c}$. To relieve the burden of feeding back FCSI information, distance- and effective channel gain $|\mathbf{h}_{l,c}\mathbf{a}_c|^2$- based ordering strategies are considered in the literature. In this paper, we propose a new ADI-based ordering strategy that uses ADI-based channel quality $\breve{\zeta}$ in (16) and orders the users according to (17). For comparison, we consider the 2-user AD MIMO-NOMA scheme with fixed PA (FPA) [38], where a coefficient $0 \leq \psi \leq \frac{1}{2}$ controls the amount of power associated to each one of the pairing users such as $\gamma_{1,c} = \psi$ and $\gamma_{2,c} = 1 - \psi$.

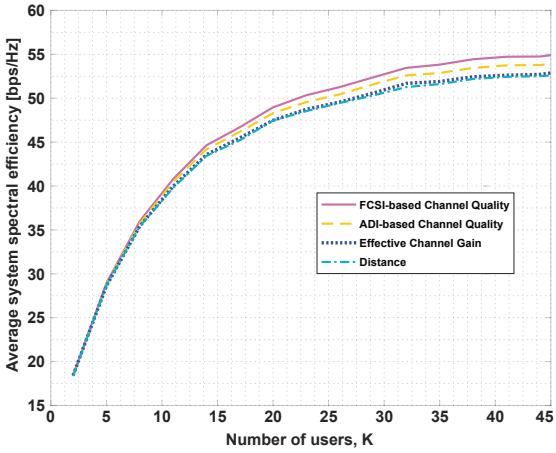


FIGURE 4. Average system spectral efficiency of 2-user AD MIMO-NOMA for different ordering strategies using distance, effective channel gain $|\mathbf{h}_{l,c}\mathbf{a}_c|^2$, ADI-based or FCSI-based channel quality with FPA $(\psi = 0.4)$.

As seen in Fig. 4, and not surprisingly, the distance- and the effective channel gain $|\mathbf{h}_{l,c}\mathbf{a}_c|^2$-based ordering strategies have very similar performance even in a congested cell. This is because $|\mathbf{h}_{l,c}\mathbf{a}_c|^2$ is inversely proportional to the distance, and thus the users will be sorted similarly from strong to weak users with the ascending order of distances or the descending order of effective channel gains. Interestingly, the proposed ordering strategy offers slightly better spectral efficiency than these two strategies. In fact, unlike the distance- or the effective channel gain-based ordering strategies, the proposed one takes into account the spatial inter-cluster interference. In addition, the performance gap between the proposed ADI- and the optimal FCSI-based ordering strategies is fairly small, around 2% loss when 45 users are served simultaneously. With only the angular information feedback, i.e., with $2\times M = 64$ fewer coefficients per channel estimation as shown in Table 2, the ADI-based ordering strategy provides a near-optimal sum-rate performance w.r.t. the optimal FCSI-based ordering strategy. Consequently, it is safe to conclude that the proposed ADI-based ordering strategy can be a viable solution in a directional mmWave environment.

C. POWER ALLOCATION POLICY

In this subsection, we aim to assess the intra-cluster PA proposed in Section V-C, denoted by KKT-PA, for the 2-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3112225, IEEE Access

user AD MIMO-NOMA scheme by comparing it with FPA and fractional transmit PA (FTPA) [38]. In addition, through numerical simulations, we answer the question: does KKT-PA with ADI NOMA make an efficient SIC possible?

FTPA is similar to the power control used in the LTE uplink. FTPA is a channel dependent PA technique that dynamically allocates powers by taking into account the channel gains of users [38] as given by (41) and (42) for 2-user AD MIMO-NOMA with FCSI or ADI NOMA.

$$\gamma_{1,c}^{\text{type}} = \frac{\left(\zeta_{1,c}^{\text{type}}\right)^{-\tau}}{\left(\zeta_{1,c}^{\text{type}}\right)^{-\tau} + \left(\zeta_{2,c}^{\text{type}}\right)^{-\tau}}, \tag{41}$$

$$\gamma_{2,c}^{\text{type}} = \frac{\left(\zeta_{2,c}^{\text{type}}\right)^{-\tau}}{\left(\zeta_{1,c}^{\text{type}}\right)^{-\tau} + \left(\zeta_{2,c}^{\text{type}}\right)^{-\tau}}, \tag{42}$$

The decay factor $0 \leq \tau \leq 1$ controls the power associated to users in NOMA cluster. ψ and τ are kept constant over all the NOMA clusters.

Fig. 5 and Fig. 6 show the average system spectral efficiency of 2-user AD MIMO-NOMA when KKT-PA, FPA with $\psi \in \{0.1, 0.3, 0.5\}$ and FTPA with $\tau \in \{0, 1\}$ are applied, with FCSI and ADI NOMA, respectively.

As seen in Fig. 5 and Fig. 6, the average system spectral efficiency of the 2-user AD MIMO-NOMA scheme using FPA grows as ψ increases with either FCSI or ADI NOMA. This result is consistent with the analysis in Appendix A, which indicates that the sum-throughput is an increasing function of $\gamma_{1,c}$. In other words, increasing ψ increases the power allocated to the strong user, which in turn enhances the sum-throughput at the expense of decreasing the power allocated to the weak user. Using FTPA, the average system spectral efficiency decreases as τ increases. This is because more power is allocated to the weak user. When $\tau = 0$ and $\psi = 0.5$, FTPA and FPA are equivalent to equal transmit PA among users, that is the optimal solution in the absence of the efficient SIC constraint. However, KKT-PA involves the optimal solution $\overset{\star}{\gamma}{}_{1,c}^{\text{type}} = \frac{1}{2} - \frac{P_{\min}}{2\zeta_{1,c}^{\text{type}}}$ under the SIC constraints. From Fig. 5 and Fig. 6, it can be seen that the gap between KKT-PA and FPA with $\psi = 0.5$ in terms of spectral efficiency is negligible. More importantly, the dynamic KKT-PA maximizes the sum-throughput under the efficient SIC constraints and can be easily applied to any cluster size, unlike FPA and FTPA.

Using KKT conditions, the closed-form solutions $\overset{\star}{\gamma}{}_c^{\text{FCSI}}$ and $\overset{\star}{\gamma}{}_c^{\text{ADI}}$ of $(\mathcal{P}_3)$ always satisfy that $\delta(\overset{\star}{\gamma}{}_c^{\text{FCSI}}, \zeta_{l-1,c}^{\text{FCSI}}) \geq P_{\min}$ and $\delta(\overset{\star}{\gamma}{}_c^{\text{ADI}}, \zeta_{l-1,c}^{\text{ADI}}) \geq P_{\min}$, respectively, as seen in (28b). However, we need to examine if the intra-cluster PA $\overset{\star}{\gamma}{}_c^{\text{ADI}}$ derived from the ADI-based SIC constraint $\delta(\overset{\star}{\gamma}{}_c^{\text{ADI}}, \zeta_{l-1,c}^{\text{ADI}}) \geq P_{\min}$ performs an efficient SIC in reality, i.e., if $\overset{\star}{\gamma}{}_c^{\text{ADI}}$ satisfies that $\delta(\overset{\star}{\gamma}{}_c^{\text{ADI}}, \zeta_{l-1,c}^{\text{FCSI}}) \geq P_{\min}$. To this end, we calculate the probability Υ of efficient SIC among all the $C_n(K_c - 1)$ users applying SIC when the KKT-PA with ADI NOMA is applied, such that $\Upsilon = P(\cup_{c=1}^{C_n} \cup_{l=2}^{K_c} \delta(\overset{\star}{\gamma}{}_c^{\text{ADI}}, \zeta_{l-1,c}^{\text{FCSI}}) \geq P_{\min}) = \sum_{c=1}^{C_n} \sum_{l=2}^{K_c} P(\delta(\overset{\star}{\gamma}{}_c^{\text{ADI}}, \zeta_{l-1,c}^{\text{FCSI}}) \geq P_{\min})$. We evaluate Υ in % for both 2-user and multi-user AD MIMO-NOMA with ADI NOMA when a different number of users, $K \in \{10, 20, 30, 40\}$, is served. $\Upsilon\%$ is averaged over 1000 trials. The numerical results provide that the average $\Upsilon\%$ is equal to 100% for both schemes and $\forall K \in \{10, 20, 30, 40\}$, which means that all users carefully apply efficient SIC. Consequently, the simplifying assumption used to design the angular-based NOMA scheme does not affect the effectiveness of SIC in highly directional channels.

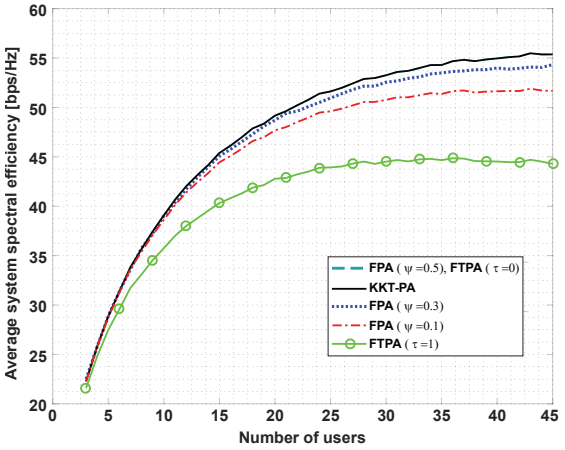


FIGURE 5. Average spectral efficiency of 2-user AD MIMO-NOMA when KKT-PA, FPA and FTPA are used with FCSI NOMA.

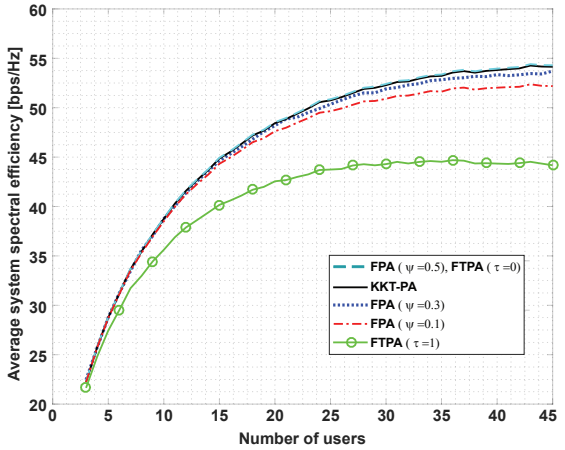


FIGURE 6. Average spectral efficiency of 2-user AD MIMO-NOMA when KKT-PA, FPA and FTPA are used with ADI NOMA.

D. AD MIMO-NOMA PERFORMANCE: AVERAGE SPECTRAL EFFICIENCY VERSUS TOTAL NUMBER OF USERS

To illustrate the advantages of the proposed multi-user and 2-user AD MIMO-NOMA schemes with either FCSI or ADI NOMA, we adopt two baseline schemes in our simulations. For baseline 1, C-AD-DBF scheme in [25] is considered, where BS generates K directive beams, each one is steered in the spatial direction of the intended user. Note that C-AD-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3112225, IEEE Access

DBF can not support more than M users. For baseline 2, labeled as AD MIMO-OMA, ADI-based digital beamforming is associated with OMA scheme. For a fair comparison, the users are grouped into two categories of clusters, namely SU and 2-user clusters according to the same β-UC algorithm as in 2-user AD MIMO-NOMA [23]. In 2-user OMA clusters, each user is served in its dedicated fraction of the orthogonal resource, e.g., time or frequency. Assume an equal resource allocation among users in the OMA cluster. Compared to C-AD-DBF, this scheme prevents users with a high level of spatial interference from getting a rate close to zero by grouping them in the same cluster and assigning them with orthogonal resources.

In Fig. 7, we investigate the average system spectral efficiency of the aforementioned schemes versus the total number of users for a spatial interference threshold of $\beta_0 = 0.5$. Performance sensibility w.r.t. the threshold is addressed in Section VI-F.

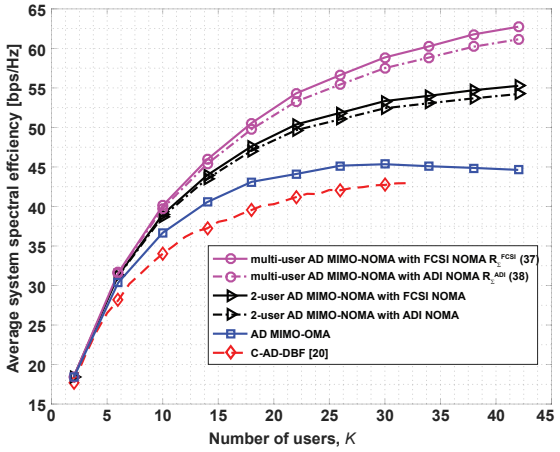


FIGURE 7. Average system spectral efficiency versus the total number of users, K, for $\beta_0 = 0.5$.

As observed from Fig. 7, attractive results on the performance of the proposed AD MIMO-NOMA with ADI NOMA are achieved, compared to FCSI NOMA. We see that the highest spectral efficiency gain $\frac{\Delta R_\Sigma}{B}$ (ΔR_Σ given by (40)), obtained for $K = 40$, of the proposed 2-user and multi-user AD MIMO-NOMA schemes using FCSI over that using ADI, is only 1.8% and 2.5%, respectively. Thus, the performance degradation due to the simplifying assumption, i.e., the ignorance of the NLOS paths and the high SNR regime, is very low. This observation verifies the interest of the proposed limited feedback NOMA protocol, including user ordering and intra-cluster power allocation, based only on ADI.

Moreover, Fig. 7 shows that ΔR_Σ increases with the number of users. However, since the number of NOMA clusters increases with the number of users, the simulation results in Fig. 7 are consistent with the analysis in Section V-D that ΔR_Σ rises with the number of NOMA clusters. Fig. 8 will analyze more precisely the variation of the number of users per cluster as K grows.

In Fig. 7, we observe that multiple access schemes such as NOMA and OMA improve the C-AD-DBF performance in terms of spectral efficiency. Indeed, all curves are above the C-AD-DBF one. This is due to the ability of the proposed β-UC algorithm to group users with high spatial interference and the potential of OMA and NOMA schemes to reduce this interference. Multi-user AD MIMO-NOMA with ADI NOMA achieves a spectral efficiency gain up to 36% w.r.t.. C-AD-DBF, when K is closed to M. More importantly, C-AD-DBF can not support simultaneously more than $M = 32$ users. Accordingly, NOMA is a promising technique to significantly boost the spectral efficiency of the C-AD-DBF system, particularly in a congested cell, at the expense of the additional complexity brought by the resource allocation (UC and PA) and the SIC decoding, at the transmitter and the receiver sides, respectively. A more detailed discussion of the complexity of our proposed scheme is given in Section VI-F.

When comparing OMA and NOMA protocols, we note that the average spectral efficiencies for the NOMA schemes increase as K grows, while it remains stable from $K = 25$ for OMA. Thus, NOMA schemes extensively outperform OMA for $K \geq 25$. Indeed, the NOMA schemes leverage the additional degrees of freedom brought by the power domain for multiple access, while OMA serves only one user in the same spectral or time resource block. This confirms the potential of AD MIMO-NOMA to exploit the multi-user diversity in a congested cell, unlike AD MIMO-OMA. However, using the same UC algorithm, the spectral efficiency gain of NOMA over OMA is achieved at the cost of greater decoding complexity at the receivers given by SIC and PA (in the case of equal repartition of PA for OMA clusters).

Finally, as expected, the proposed multi-user schemes outperform those with 2-user, especially in a congested cell. Indeed, as the number of users grows, the number of users per NOMA cluster also increases in the multi-user scheme to manage the inter-user interference. In contrast, for the 2-user scheme, the number of users is fixed to two, and more beams, potentially in interference, are needed to address the growing number of users. This result will be discussed in the following Section VI-E.

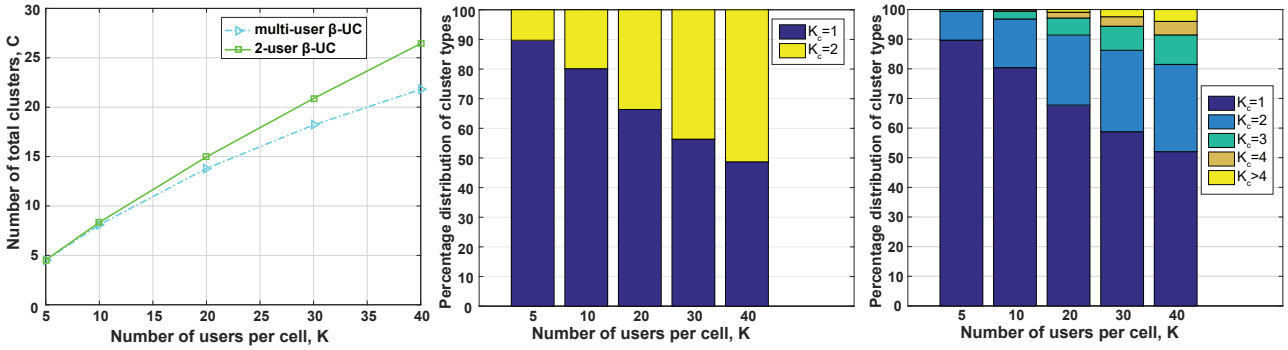
In the rest of the paper, we will only consider AD MIMO-NOMA with partial CSI, i.e., with ADI NOMA.

E. AD MIMO-NOMA PERFORMANCE: NUMBER OF USERS PER NOMA CLUSTER

Fig. 8 shows the impact of the proposed 2-user and multi-user β-UC algorithms on the scheduled clusters; their number, their types (SU or NOMA cluster) and the number of users in NOMA cluster, for a spatial interference threshold of $\beta_0 = 0.5$. Fig. 8a depicts the average number of clusters, using either 2-user or multi-user β-UC. Fig. 8b and Fig. 8c depict the cluster's percentage distribution per type for 2-user and multi-user β-UC, respectively, for $K \in \{5, 10, 20, 30, 40\}$.

As seen in Fig. 8a, the total number of clusters using 2-user β-UC is larger than when using multi-user β-UC, since multiple users $K_c \geq 2$ can be scheduled within the same

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3112225, IEEE Access



(a) Average number of clusters for 2-user and multi-user β-UC

(b) Cluster's percentage distribution per type for 2-user β-UC.

(c) Cluster's percentage distribution per type for multi-user β-UC.

FIGURE 8. Average number of clusters (a), cluster's percentage distribution per type for 2-user (b) and multi-user AD MIMO-NOMA (c) versus the total number of users, K, for $\beta_0 = 0.5$.

cluster in the latter scheme. In Fig. 8b and Fig. 8c, we observe that, as the number of users increases, the number of 2- or multi-user clusters also grows to deal with the higher level of spatial interference. Moreover, even if for $K = 5$, we get practically the same type of clusters, i.e., only 2-user, for both schemes, when K increases, multi-user clusters appear, with K_c equal to 3, 4 or greater than 4. Thanks to this additional degree of freedom, the inter-cluster interference is reduced by passing from 2-user to multi-user scheme. Consequently, the multi-user scheme achieves a higher spectral efficiency w.r.t.. the 2-user scheme, as shown in Fig. 7, especially when $K \geq 20$.

F. AD MIMO-NOMA COMPLEXITY

As already stated, the computational complexity of the proposed MIMO-NOMA schemes is brought by the resource allocation (UC and PA) at the BS and the SIC decoding at the receiver side. We discuss each of these independently for the 2-user and the multi-user schemes.

In our previous work [23], the proposed 2-user β-UC algorithm searches the largest β as the largest element in a matrix of size $(K-1) \times K$. Hence, it costs $\mathcal{O}((K-1)K)$ to find this element for each iteration. Moreover, $K_c = 2$ $(\forall c \in \mathcal{C}_n)$, then there is no inner loop $(\lfloor \frac{K_c}{2} \rfloor = 1)$, and the complexity order is given by $\mathcal{O}(C_n(K-1)K)$. The complexity order of the multi-user β-UC algorithm is given by (24). Thanks to the set $\mathcal{F}$ with a reduced size after each iteration, the multi-user β-UC has a lower complexity than 2-user UC [23], which uses the same matrix over all the iterations. In this paper, to give a fair comparison, we apply the same trick to the 2-user β-UC algorithm. Thus, the complexity order of the 2-user β-UC algorithm is also given by (24), with $K_c = 2$ $(\forall c \in \mathcal{C}_n)$ and $\lfloor \frac{K_c}{2} \rfloor = 1$. Therefore, it can be rewritten as $\mathcal{O}\left(\sum_{c=1}^{C_n} \mathbf{card}(\mathcal{F}_{1,c})\right)$, with $\mathbf{card}(\mathcal{F}_{1,c}) < (K-1)K \ \forall c \in \mathcal{C}_n$. Now, we compare the complexity order of UC for both schemes by calculating the corresponding average of $\sum_{c=1}^{C_n} \sum_{p=1}^{\lfloor \frac{K_c}{2} \rfloor} \mathbf{card}(\mathcal{F}_{p,c})$ in Table 3. In addition, we evaluate in Table 3 the number of NOMA clusters C_n^{2-UE} and C_n^{m-UE} for 2-user and multi-user schemes, respectively. As seen in the Table, the multi-user β-UC algorithm has a lower complexity than the 2-user one, and the gap in the complexity between multi-user and 2-user schemes increases as K grows. This is because more NOMA clusters are scheduled in the 2-user scheme, as shown in Table 3. And on the other hand, because $\mathbf{card}(\mathcal{F}_{p,c})$ of the multi-user β-UC is decreasing faster than $\mathbf{card}(\mathcal{F}_{1,c})$ of the 2-user β-UC after each iteration due to the ability of multi-user β-UC to group multiple users in the same NOMA cluster.

	2-user $\sum_{c=1}^{C_n} \mathbf{card}(\mathcal{F}_{1,c})$	multi-user $\sum_{c=1}^{C_n} \sum_{p=1}^{\lfloor \frac{K_c}{2} \rfloor} \mathbf{card}(\mathcal{F}_{p,c})$	C_n^{2-UE}	C_n^{m-UE}
$K = 10$	5.7	5.7	1.6	1.6
$K = 20$	35.6	33	4.9	4.4
$K = 30$	112.3	95.7	9.5	7.6
$K = 40$	258	201	13.6	10.2

TABLE 3. The complexity of 2-user and multi-user β-UC algorithms and the number of NOMA clusters of both schemes, when the BS serves $K \in \{10, 20, 30, 40\}$ users.

The complexity order of the SIC decoding is linear with the number of SIC users[6], i.e., $\mathcal{O}(\sum_{c=1}^{C_n}(K_c - 1))$. In addition, the complexity order of the PA method is linear with the number of all NOMA users, i.e., $\mathcal{O}(\sum_{c=1}^{C_n} K_c)$. Note that these expressions are the same for the multi-user and the 2-user schemes, but with different numbers of NOMA clusters, C_n, and different NOMA cluster sizes, K_c. Thus, in Table 4, to compare PA complexity orders (resp. SIC complexity orders), we calculate the relative percentage of NOMA users (resp. SIC users) of the multi-user scheme, as compared to the 2-user scheme. We observe that the relative percentage of NOMA users is low, even in a congested cell (e.g., 6.7% for $K = 40$). Indeed, the number of SU clusters is globally the same for both schemes and the remaining users are NOMA users. When looking at SIC complexity, however, with more

[6]SIC users are the NOMA users applying SIC.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2021.3112225, IEEE Access

	Relative percentage of NOMA users	Relative percentage of SIC users
$K = 20$	4.7%	19.3%
$K = 30$	5.9%	23.7%
$K = 40$	6.7%	25.8%

TABLE 4. Relative percentage of NOMA users and SIC users for the multi-user MIMO-NOMA scheme, as compared to the 2-user scheme, when the BS serves $K \in \{20, 30, 40\}$ users.

users in each NOMA cluster (as shown in Fig. 8(c)), the relative percentage of SIC users is quite high, especially in a congested cell (e.g. 25.8% for $K = 40$). In other words, the gap in terms of SIC complexity increases as the number of users grows, unlike that of the PA method.

Finally, we conclude that SIC decoding highly contributes to the difference in complexity between the 2-user and the multi-user schemes.

G. AD MIMO-NOMA PERFORMANCE: AVERAGE SPECTRAL EFFICIENCY VERSUS SPATIAL INTERFERENCE THRESHOLD

Fig. 9 investigates the impact of the spatial interference threshold β_0 on the system performance in terms of the average system spectral efficiency when $K \in \{10, 20, 40\}$ users are served simultaneously.

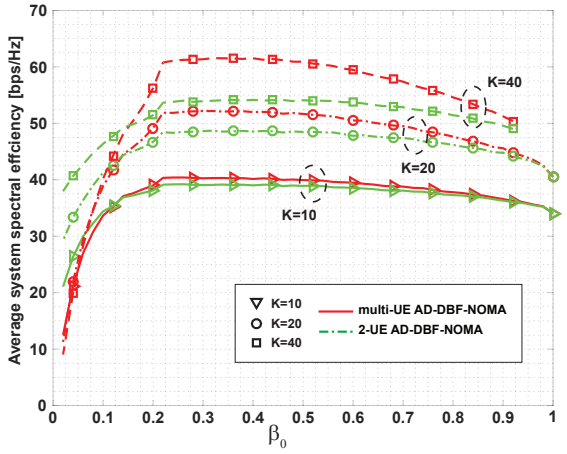


FIGURE 9. Average system spectral efficiency of AD MIMO-NOMA with ADI NOMA, for 2-user and multi-user clusters, when $K \in \{10, 20, 40\}$ users exist.

Firstly, we highlight that AD MIMO-NOMA with $\beta_0 = 1$ is equivalent to C-AD-DBF for $K \leq M$. The spectral efficiency curves of the proposed schemes change the slope around $\beta_0 = 0.21$. This observation is coherent with the value $\beta^{\text{FSL}} = 0.217$ given in Lemma 1. We recall that we consider the users as interferes, and we group them in the same NOMA cluster, only if the main lobe covers them within the β_0-width with $\beta_0 \geq 0.217$. Instead, when $\beta_0 < 0.217$, users can interfere with secondary lobes. Based on this Lemma, the useful range is $0.217 \leq \beta_0 < 1$. In this range, we find that the two proposed schemes outperform C-AD-DBF (or AD MIMO-NOMA with $\beta_0 = 1$), as previously seen on Fig. 7.

In addition, the proposed multi-user scheme outperforms the 2-user one $\forall \beta_0 > 0.217$. In fact, the performance gap between these two schemes depends on two factors: (i) the total number of users, K and (ii) the spatial interference threshold, β_0. This gap increases (resp. decreases) with growing K (resp. decreasing β_0). The increasing gap with K has already been observed in Fig. 7 and explained in section VI-E by the additional degrees of freedom brought by the multi-user clusters. Varying β_0 captures the minimum level of the inter-user interference between users in the same cluster. Then, when choosing lower β_0 satisfying Lemma 1, more users can be located in the same cluster. Thanks to the fewer clusters, the inter-cluster interference is reduced, and the performance of the multi-user scheme is improved w.r.t. the 2-user one. Therefore, the lowest β_0 in the useful range produces the highest system spectral efficiency.

When comparing the multi-user AD MIMO-NOMA scheme with $\beta_0 = 0.217$ and $\beta_0 = 0.5$, the performance gap is negligible for $K = \{10, 20\}$. For K equal to 40, it is slightly in favor of $\beta_0 = 0.217$. Ultimately, as previously obtained, $\beta_0 = 0.217$ is the better choice.

VII. CONCLUSION

In this paper, we have proposed a multi-user MIMO-NOMA transmission scheme, denoted as AD MIMO-NOMA, at mmWave frequencies using only the angular information to address the problems of FCSI estimation and feedback. This was done by considering the key features of mmWave channels, i.e., the high directionality and the potential blockage. We first derived a set of angular-based performance metrics, which participate in the design of user clustering, user ordering, and power allocation techniques. Specifically, the user clustering algorithm groups the users with high spatial interference to perform NOMA, while the power allocation technique maximizes the system throughput. Numerical results verified that the proposed AD MIMO-NOMA scheme effectively improves the multi-user MIMO system throughput and guarantees the efficient SIC. Besides, the results show the potential of NOMA against OMA to exploit the multi-user diversity in a congested cell. As future works, we will consider the complexity and the imperfection of SIC decoding in the proposed multi-user AD MIMO-NOMA scheme and evaluate the system performance in terms of bit error rate. Moreover, it would be interesting to develop an error model of angular information and investigate the impact of imperfect ADI on the performance of the proposed scheme. Since the mmWave bands are attractive due to the large bandwidth available, the wideband implementation with the orthogonal frequency division modulation is an interesting topic to investigate in the context of the mmWave AD MIMO-NOMA system.

.

APPENDIX A

For 2-user AD MIMO-NOMA with FCSI or ADI NOMA, the sum-throughput R_c^{type} of UEs within the c-th cluster can be expressed as

$$R_c^{\text{type}} = B\log_2(1+\gamma_{1,c}\zeta_{1,c}^{\text{type}})+B\log_2\left(1+\frac{\gamma_{2,c}\zeta_{2,c}^{\text{type}}}{\gamma_{1,c}\zeta_{2,c}^{\text{type}}+1}\right). \tag{43}$$

The first-order derivative of R_c^{type} w.r.t.. $\gamma_{1,c} = 1-\gamma_{2,c}$ is given by

$$\frac{\mathrm{d}R_c^{\text{type}}}{\mathrm{d}\gamma_{1,c}} = B\frac{\zeta_{1,c}^{\text{type}}-\zeta_{2,c}^{\text{type}}}{(1+\zeta_{1,c}^{\text{type}}\gamma_{1,c})(1+\zeta_{2,c}^{\text{type}}\gamma_{1,c})} \tag{44}$$

Since $\zeta_{1,c}^{\text{type}} > \zeta_{2,c}^{\text{type}}$, then $\frac{\mathrm{d}R_c^{\text{type}}}{\mathrm{d}\gamma_{1,c}} > 0$ and the objective function R_c^{type} in $(\mathcal{P}_3)$ is a monotonically increasing function of $\gamma_{1,c}$. However, the constraints (28a) and (28b) in $(\mathcal{P}_3)$ implies that $0 \le \gamma_{1,c} \le \frac{1}{2}-\frac{P_{\min}}{2\zeta_{1,c}^{\text{type}}}$. Therefore, $\overset{\star}{\gamma}{}_{1,c}^{\text{type}} = \frac{1}{2}-\frac{P_{\min}}{2\zeta_{1,c}^{\text{type}}}$ maximizes the sum-throughput within the 2-user NOMA cluster. This result is also obtained using KKT as in (37).

REFERENCES

[1] R. W. Heath, N. Gonzalez-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 436–453, 2016.

[2] L. Dai, B. Wang, Y. Yuan, S. Han, I. Chih-Lin, and Z. Wang, "Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends," *IEEE Commun. Mag.*, vol. 53, no. 9, pp. 74–81, 2015.

[3] Q. Sun, S. Han, I. Chin-Lin, and Z. Pan, "On the ergodic capacity of MIMO NOMA systems," *IEEE Wireless Commun. Lett.*, vol. 4, no. 4, pp. 405–408, 2015.

[4] Y. Liu, G. Pan, H. Zhang, and M. Song, "On the capacity comparison between MIMO-NOMA and MIMO-OMA," *IEEE Access*, vol. 4, pp. 2123–2129, 2016.

[5] Z. Wei, L. Yang, D. W. K. Ng, J. Yuan, and L. Hanzo, "On the performance gain of NOMA over OMA in uplink communication systems," *IEEE Transactions on Communications*, vol. 68, no. 1, pp. 536–568, 2019.

[6] Z. Ding, F. Adachi, and H. V. Poor, "The application of MIMO to non-orthogonal multiple access," *IEEE Trans. Wireless Commun.*, vol. 15, no. 1, pp. 537–552, 2015.

[7] T.-V. Nguyen, V.-D. Nguyen, D. B. da Costa, and B. An, "Hybrid user pairing for spectral and energy efficiencies in multiuser MISO-NOMA networks with SWIPT," *IEEE Trans. Commun.*, vol. 68, no. 8, pp. 4874–4890, 2020.

[8] A. S. de Sena, F. R. M. Lima, D. B. da Costa, Z. Ding, P. H. Nardelli, U. S. Dias, and C. B. Papadias, "Massive MIMO-NOMA networks with imperfect SIC: Design and fairness enhancement," *IEEE Trans. Wireless Commun.*, vol. 19, no. 9, pp. 6100–6115, 2020.

[9] J. Choi, "Minimum power multicast beamforming with superposition coding for multiresolution broadcast and application to NOMA systems," *IEEE Trans. Commun.*, vol. 63, no. 3, pp. 791–800, 2015.

[10] Z. Xiao, L. Zhu, J. Choi, P. Xia, and X.-G. Xia, "Joint power allocation and beamforming for non-orthogonal multiple access (NOMA) in 5G millimeter wave communications," *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 2961–2974, 2018.

[11] B. Wang, L. Dai, Z. Wang, N. Ge, and S. Zhou, "Spectrum and energy-efficient beamspace MIMO-NOMA for millimeter-wave communications using lens antenna array," *IEEE J. Sel. Areas in Commun.*, vol. 35, no. 10, pp. 2370–2382, 2017.

[12] J. Wang, S. Jin, X. Gao, K.-K. Wong, and E. Au, "Statistical eigenmode-based SDMA for two-user downlink," *IEEE Trans. Signal Processing*, vol. 60, no. 10, pp. 5371–5383, 2012.

[13] Y. Sun, J. Zhou, Q. Cao, and S. Li, "Precoder design in statistical CSI aided non-orthogonal multiple access," *IEEE Access*, vol. 6, pp. 16 484–16 492, 2018.

[14] J. Choi, "Joint rate and power allocation for NOMA with statistical CSI," *IEEE Trans. Commun.*, vol. 65, no. 10, pp. 4519–4528, 2017.

[15] T. S. Rappaport, Y. Xing, G. R. MacCartney, A. F. Molisch, E. Mellios, and J. Zhang, "Overview of millimeter wave communications for fifth-generation (5G) wireless networksâĂŤwith a focus on propagation models," *IEEE Trans. Antennas and Propagation*, vol. 65, no. 12, pp. 6213–6230, 2017.

[16] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE J. Sel. Areas in Commun.*, vol. 32, no. 6, pp. 1164–1179, 2014.

[17] J. Wang, Y. Li, C. Ji, Q. Sun, S. Jin, and T. Q. Quek, "Location-based MIMO-NOMA: Multiple access regions and low-complexity user pairing," *IEEE Trans. Commun.*, vol. 68, no. 4, pp. 2293–2307, 2020.

[18] Z. Ding, P. Fan, and H. V. Poor, "Random beamforming in millimeter-wave NOMA networks," *IEEE access*, vol. 5, pp. 7667–7681, 2017.

[19] E. M. Mohamed, "Joint users selection and beamforming in downlink millimetre-wave noma based on users positioning," *IET Communications*, vol. 14, no. 8, pp. 1234–1240, 2020.

[20] X. Lu, Y. Zhou, and V. W. Wong, "A joint angle and distance based user pairing strategy for millimeter wave NOMA networks," in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, 2020.

[21] Z. Zheng, W.-Q. Wang, H. Meng, H. C. So, and H. Zhang, "Efficient beamspace-based algorithm for two-dimensional DOA estimation of incoherently distributed sources in massive MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 67, no. 12, pp. 11 776–11 789, 2018.

[22] T. Wang, B. Ai, R. He, and Z. Zhong, "Two-dimension direction-of-arrival estimation for massive MIMO systems," *IEEE Access*, vol. 3, pp. 2122–2128, 2015.

[23] I. Khaled, C. Langlais, A. El Falou, M. Jezequel, and B. ElHasssan, "Joint SDMA and power-domain NOMA system for multi-user mm-Wave communications," in *Proc. IEEE Int. Wireless Commun. and Mobile Computing (IWCMC)*, 2020, pp. 1112–1117.

[24] M. K. Samimi and T. S. Rappaport, "3-D millimeter-wave statistical channel model for 5G wireless system design," *IEEE Trans. Microw. Theory Tech.*, vol. 64, no. 7, pp. 2207–2225, 2016.

[25] I. Khaled, A. El Falou, C. Langlais, B. ElHassan, and M. Jezequel, "Multi-user digital beamforming based on path angle information for mm-Wave MIMO systems," in *Proc. Int. ITG Workshop on Smart Antennas (WSA)*. VDE, 2020.

[26] M. S. Ali, H. Tabassum, and E. Hossain, "Dynamic user clustering and power allocation for uplink and downlink non-orthogonal multiple access (NOMA) systems," *IEEE access*, vol. 4, pp. 6325–6343, 2016.

[27] T. S. Rappaport, E. Ben-Dor, J. N. Murdock, and Y. Qiao, "38 GHz and 60 GHz angle-dependent propagation for cellular & peer-to-peer wireless communications," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2012, pp. 4568–4573.

[28] N. Rupasinghe, Y. Yapıcı, I. Güvenç, M. Ghosh, and Y. Kakishima, "Angle feedback for NOMA transmission in mmwave drone networks," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 3, pp. 628–643, 2019.

[29] Q. Yang, H.-M. Wang, D. W. K. Ng, and M. H. Lee, "NOMA in downlink SDMA with limited feedback: Performance analysis and optimization," *IEEE J. Sel. Areas in Commun.*, vol. 35, no. 10, pp. 2281–2294, 2017.

[30] M. Morales-Céspedes, O. A. Dobre, and A. García-Armada, "Semi-blind interference aligned NOMA for downlink MU-MISO systems," *IEEE Trans. Commun.*, vol. 68, no. 3, pp. 1852–1865, 2019.

[31] X. Hu, C. Zhong, Y. Han, X. Chen, J. Zhao, and Z. Zhang, "Angle-domain mmwave MIMO NOMA systems: analysis and design," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2019.

[32] H. J. Visser, *Array and phased array antenna basics*. Wiley Online Library, 2005.

[33] Q. N. Le, V.-D. Nguyen, N.-P. Nguyen, S. Chatzinotas, O. A. Dobre, and R. Zhao, "Learning-assisted user clustering in cell-free massive MIMO-NOMA networks," *arXiv preprint arXiv:2011.07549*, 2020.

[34] S. R. Islam, M. Zeng, O. A. Dobre, and K.-S. Kwak, "Resource allocation for downlink NOMA systems: Key techniques and open issues," *IEEE Wireless Communications*, vol. 25, no. 2, pp. 40–47, 2018.

[35] J. Cui, Z. Ding, P. Fan, and N. Al-Dhahir, "Unsupervised machine learning-based user clustering in millimeter-wave-NOMA systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 11, pp. 7425–7440, 2018.

[36] L. Zhu, J. Zhang, Z. Xiao, X. Cao, D. O. Wu, and X.-G. Xia, "Millimeter-wave NOMA with user grouping, power allocation and hybrid beamforming," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5065–5079, 2019.
[37] E. K. Chong and S. H. Zak, *An introduction to optimization*. John Wiley & Sons, 2004.
[38] N. Otao, Y. Kishiyama, and K. Higuchi, "Performance of non-orthogonal multiple access with SIC in cellular downlink using proportional fair-based resource allocation," *IEICE Trans. Commun.*, vol. 98, no. 2, pp. 344–351, 2015.

ISRAA KHALED received the computer and communications engineering degree from Lebanese University, in 2018. She is currently pursuing the Ph.D. degree in information and communication engineering with IMT Atlantique and Lebanese University. Her current research interests millimeter wave, multiple-input multiple-output, beamforming and non-orthogonal multiple access.

CHARLOTTE LANGLAIS was born in Le Mans, France, in 1976. She received the M.E degree in Telecommunications, the M.S. degree in electrical engineering, and the Ph.D. degree in electrical engineering from the Institut National des Sciences AppliquÃl'es (INSA), Rennes, France, in 1999 and 2002, respectively. She received the accreditation to supervise research from the University of Southern Brittany, France, in 2021. In 2002, she joined the Mathematical and Electrical Engineering Department, IMT Atlantique, France, as an Associate Professor. Since 2020, she has been Study Director with the Mathematical and Electrical Engineering Department, IMT Atlantique. She is a member of the CNRS Lab-STICC Laboratory. She is the author of 2 book chapters and more than 70 articles. Her research interests include wireless communications, MIMO systems, multiuser and cooperative communications.

AMMAR EL FALOU (SâĂŹ09, MâĂŹ14) received his M.Sc. degree in Digital Telecommunication Systems from TELECOM Paris and UPMC (Univ. Paris VI), Paris, France and his M.E. in Telecommunications Engineering from the Lebanese University, Faculty of Engineering, Lebanon, both in 2009. In 2013, he received his Ph.D. in Communication and Information Science from IMT Atlantique (ex. TELECOM Bretagne), Brest, France. In 2014, he worked as a postdoctoral at Orange Labs, Sophia Antipolis, France. From 2015 to 2020, he was a part-timer assistant professor at Lebanese University (LU), Beirut Arab University (BAU) and Lebanese International University (LIU) in Lebanon. Currently, he is an assistant professor at ESIEE Paris, Gustave Eiffel University, Noisy-le-Grand, France.

BACHAR A. ELHASSAN received the Diploma degree in engineering from the Faculty of Engineering, Lebanese University, Tripoli, Lebanon, in 1991, the M.S. degree in signal and image processing from the National Polytechnic Institute (INPG), Grenoble, France, in 1992, and the Ph.D. degree in electronics from INPG, France, in 1995, in collaboration with France Telecom. Since 1996, he has been with the Faculty of Engineering, Lebanese University, Tripoli, Lebanon, where he is currently a Full Professor. He served as a Chairman for the Electrical Engineering Department for more than six years. He is also the Founder and the Director of the laboratory of Telecommunication, Networking and Microwaves (LTRM), Doctoral School of Sciences and Technologies(EDST), Lebanese University. His research interests are as follows: image processing, digital communications, intelligence in wireless sensor networks, and Internet of Things. He was the Chair of the IEEE ComSoc Lebanon chapter from January 2013 to December 2016 and the Vice Chair of the IEEE Lebanon Section form January 2017 to December 2018. He has been the Chair of the IEEE Lebanon Section Elected since January 2019. He founded and was the Director of the Development and Training Center, the Order of Engineers and Architects Tripoli, Lebanon, from July 2009 to December 2013. He was the Head of the Employed Engineers Branch,Order of Engineers and Architects, Tripoli, Lebanon, from April 2016 to April 2019. He has also been the member of the ICT Arab Committee since March 2017.

MICHEL JEZEQUEL received the M.E degree in electronics from the ÃL'cole Nationale SupÃl'rieure de l'ÃL'lectronique et de ses Applications, Paris, France in 1982, and the accreditation to supervise research from the University of Southern Brittany, France, in 2008. He first worked at CIT ALCATEL in Lannion, France as a design engineer (1983-1986) and after an experience in a small company, he followed a one-year course on software design. In 1988, he joined IMT Atlantique where he is currently Full Professor. He served as a Head of the Electronics Department for more than 25 years. He is a member of the CNRS Lab-STICC Laboratory, and a member of Pracom (Advanced Research Centre for Communications). His research interests include algorithm and silicon interaction. More precisely he works on error correction codes, turbo codes and iterative processing, iterative detection, and circuit design for communication systems.