Received 19 April 2022; accepted 27 April 2022. Date of publication 4 May 2022; date of current version 12 May 2022.

Digital Object Identifier 10.1109/OJCOMS.2022.3172621

User Selection for Massive MIMO under Line-of-Sight Propagation

RAFAEL S. CHAVES [1,2] (Member, IEEE), MARKUS V. S. LIMA [2] (Member, IEEE), EDIZ CETIN [1] (Member, IEEE), AND WALLACE A. MARTINS [3] (Senior Member, IEEE)

[1]School Engineering, Macquarie University, Sydney, NSW 2109, Australia

[2]Electrical Engineering Program (PEE/Coppe), Universidade Federal do Rio de Janeiro, Rio de Janeiro 21941-909, Brazil

[3]Interdisciplinary Centre for Security Reliability and Trust, University of Luxembourg, 4365 Esch-sur-Alzette, Luxembourg

CORRESPONDING AUTHOR: R. S. CHAVES (e-mail: rafael.chaves@smt.ufrj.br)

This work was supported in part the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Brazil—Finance Code 001; in part by the International Macquarie University Research Excellence Scholarship (iMQRES); and in part by CNPq and FAPERJ, Brazilian funding agencies.

ABSTRACT This paper provides a review of user selection algorithms for massive multiple-input multiple-output (MIMO) systems under the line-of-sight (LoS) propagation model. Although the LoS propagation is extremely important to some promising technologies, like in millimeter-wave communications, massive MIMO systems are rarely studied under this propagation model. This paper fills this gap by providing a comprehensive study encompassing several user selection algorithms, different linear precoders and simulation setups, and also considers the effect of partial channel state information (CSI). One important result is the existence of practical cases in which the LoS propagation model may lead to significant levels of interference among users within a cell; these cases are not satisfactorily addressed by the existing user selection algorithms. Motivated by this issue, a new user selection algorithm based on inter-channel interference (ICI) called ICI-based selection (ICIBS) is proposed. Unlike other techniques, the ICIBS accounts for the ICI in a global manner, thus yielding better results, especially in cases where there are many users interfering with each other. In such scenarios, simulation results show that when compared to the competing algorithms, the proposed approach provided an improvement of at least 10.9% in the maximum throughput and 7.7% in the 95%-probability throughput when half of the users were selected.

INDEX TERMS Massive MIMO, favorable propagation, user selection, line-of-sight channel, inter-channel interference.

I. INTRODUCTION

MASSIVE *multiple-input multiple-output* (MIMO) systems employ a large number of antennas at the *base station* (BS), serving users within the same time-frequency radio resource [1]. Current MIMO systems commonly employ up to eight antennas at the BS, whereas massive MIMO proposals consider using hundreds of antennas [2]. By using a large number of antennas, the small-scale fading effect can be eliminated through the so-called channel hardening, resulting in a deterministic scalar channel model [3]. Furthermore, under *favorable propagation* (FP) condition, massive MIMO systems can achieve very high throughput by using simple linear processing [2]. These advantages make massive MIMO very attractive for wireless communication systems.

One of the most critical aspects for the realization of massive MIMO systems is meeting the FP condition, which requires the channel vectors to be mutually orthogonal. However, this condition is rarely met in real propagation environments. Fortunately, the *asymptotically favorable propagation* (AFP) condition is sufficient to guarantee high throughput by using simple linear processing [2]. This condition is achieved by *non-line-of-sight* (NLoS) environments with rich scattering, which can be modeled by an i.i.d. Rayleigh fading channel [4]. The i.i.d. Rayleigh channel has the advantage of yielding tractable analytical expressions

for the achievable rate of massive MIMO systems [5]. However, in practical applications, the propagation channel is composed of NLoS and *line-of-sight* (LoS) components. Studying the extreme cases, i.e., NLoS and LoS, is important to have performance bounds since, in practice, we will have a case lying in-between the LoS and NLoS cases. Therefore, although massive MIMO systems are frequently studied under the assumption of i.i.d. Rayleigh channel models, the LoS channels are also important [6]–[10]. In fact, the massive MIMO technology is a potential candidate for use in *millimeter-wave* (mmWave) bands, where LoS is the main form of propagation due to the high attenuation and low penetration of the electromagnetic waves at these frequencies [11]–[14].

In the LoS propagation environment, the array geometry and the *angle of arrival* (AoA) corresponding to each user determines the channel [15]. This is an example of real propagation environment in which the FP condition may be met, albeit in several cases this condition does not hold [16]. For instance, both the FP and AFP conditions are violated when users are close to each other, i.e., when they have similar AoAs [6], [7], [17]–[19]. The probability of having close users tends to increase as the number of users served by a given BS grows. This proximity of users has major impact on the *spectral efficiency* (SE) of massive MIMO systems, degrading it severely when users have similar AoAs [6]. This scenario is recurrent in sports and cultural events, where the number of users in a given region is very high. Dealing with such crowded scenarios is also important for the *Internet of Things* (IoT) since a large number of devices may be connected to the wireless network [20]. Additionally, although high, the finite number of antennas at the BS limits the number of transmit beams that can be formed, thus constraining the number of users that can be served simultaneously. Therefore, performing user scheduling, which includes *user selection* in its core, is of paramount importance for the proper operation of massive MIMO systems.

A. USER SELECTION IN MASSIVE MIMO SYSTEMS

The SE of massive MIMO systems is directly related to the *signal-to-interference-plus-noise ratio* (SINR) associated with each user. Intuitively, users that cause strong interference among each other should not be transmitting simultaneously in order to maximize the overall SINR during a time slot. However, selecting users whose transmissions have the highest SINRs is computationally complex, requiring an exhaustive search across many possible combinations of users. The goal of this subsection is to provide a literature review of user selection algorithms suitable to massive MIMO systems. To this end, not only the user selection algorithms proposed for massive MIMO systems are discussed, but also other algorithms that work in the conventional *multi-user* MIMO (MU-MIMO) setup.

For conventional MU-MIMO systems, the most popular user selection algorithms are the *greedy zero-forcing dirty-paper algorithm* [21], *greedy scheduler with equal power allocation* [22], and *semi-orthogonal selection* (SOS) [23]. While the former ones perform user selection by maximizing directly the *downlink* (DL) SE of the system, the latter selects users in order to yield a subset comprised of the "most orthogonal" channel vectors. The greedy zero-forcing dirty-paper and greedy scheduler with equal power allocation algorithms are based on the *dirty paper coding* (DPC) [24], which achieves the capacity of DL MIMO systems with Gaussian channels [25]. However, the implementation of DPC is problematic in practical situations due to the high computational burden, particularly for a large number of antennas and users [23].

The SOS algorithm, on the other hand, does not suffer from the aforementioned problems and, therefore, has been widely employed for user selection in massive MIMO systems. In [26], the SOS algorithm is used along with an antenna selection algorithm to solve a joint antenna and user selection problem. In [27], the SOS algorithm is used for user scheduling in a scenario where the number of antennas is slightly larger than the number of users. Moreover, a new version of the SOS algorithm, called *massive MIMO pair-wise SOS* algorithm is proposed, in which the SOS is initialized with the full set of users instead of an empty one, which may lead to reduced complexity when the system has more antennas than users. In [28], the SOS algorithm is used to select users in a crowded environment, simulating an IoT application. Additionally, a new version of the SOS algorithm, called *modified SOS* (SOS-M) is proposed in which the SOS operation is no longer limited by the number of antennas, but by a given number of users to be selected.

Unlike the SOS algorithm, there exist several user selection solutions aimed directly at the massive MIMO case, such as the *random selection* (RS) [29], [30], the *delete the minimum lambda* (DML) [31], and the *correlation-based selection* (CBS) [7], [32], [33] approaches. The RS approach is a naive way to perform user selection without optimizing any performance criterion. The DML is a decremental user selection algorithm based on *zero-forcing* (ZF) precoding and, consequently, it can only be used in cases where the number of users is smaller than the number of antennas, not being suitable for applications involving crowded scenarios. The CBS algorithm, on the other hand, selects users based on the correlation between pairs of users' channels, iteratively removing those users that strongly interfere with one another. Hence, the CBS aims to maximize the SINR gain of a particular user and does not guarantee the best achievable overall SINR gain for the whole system. Other classes of algorithms that can be used to perform user selection in massive MIMO are the *user grouping* algorithms [8], [34], [35] and recently proposed *machine learning-based selection* algorithms [36]. These algorithms separate users into clusters, serving a reduced number of users per cluster in the same time-frequency resource in order to decrease the interference between users within the same cluster. However, user grouping algorithms are more focused on the precoder design than on the selection of the users inside the clusters.

In summary, most of the aforementioned user selection algorithms depend on particular precoding and/or power allocation schemes, not being suitable to a wide range of applications involving massive MIMO systems. Therefore, in this paper, we focus on the SOS and CBS algorithms as they require only *channel state information* (CSI) and are not constrained to a single precoding scheme.

B. CONTRIBUTIONS

The main goal of this paper is to provide a review of user selection algorithms for massive MIMO systems under the LoS channel, a propagation environment of great interest, albeit poorly studied, with very particular challenges that can deteriorate the performance of massive MIMO systems, as explained in this article. Additionally, a new user selection algorithm based on *inter-channel interference* (ICI), namely *ICI-based selection* (ICIBS), is proposed.[1] Unlike the CBS, the ICIBS accounts for the ICI in a global manner, thus yielding better results especially when there are many users interfering with one another. Therefore, this paper essentially plays the role of a review paper, addressing some existing gaps for the first time.

The specific contributions of this paper are:

1) *Literature review:* This paper provides a comprehensive literature review of user selection algorithms for massive MIMO focusing on two important techniques, namely, the SOS and CBS algorithms. These algorithms work quite well in the massive MIMO scenario and require only CSI, thus they are not constrained to some specific precoding and power allocation schemes. These two techniques are covered in detail and a thorough analysis of their computational complexity is provided.
2) *New algorithm:* This paper proposes a new user selection algorithm, called ICIBS, which has the same flexibility of SOS and CBS algorithms. However, unlike SOS and CBS, the ICIBS is designed to address those situations in which there are many users interfering with each other, which yields higher interference levels. In these cases, a significant advantage of ICIBS over the other algorithms can be observed. On the other hand, when dealing with lower interference levels, e.g., when the users were spread over the cell, the ICIBS performed as good as the other algorithms. Further, the computational complexity of ICIBS is much lower than that of SOS, and only slightly higher than that of CBS.
3) *Study of massive MIMO under LoS propagation:* This paper also provides a comprehensive study of massive MIMO systems under LoS propagation, covering some of the most fundamental aspects of massive MIMO systems, such as the FP and the SE. Additionally, the paper shows the main differences between Rayleigh and LoS propagation models in terms of FP and extends previous results from the literature. Furthermore, an extensive set of simulation results considering several user selection algorithms, different linear precoders, various setups covering low and high interference levels, and also the effect of partial CSI are provided. To the best of our knowledge, there is no other work providing such a detailed review and study of massive MIMO systems under LoS propagation model.

1. A preliminary version of the ICIBS algorithm has been presented in [37]. Here this preliminary work has been significantly extended to include a fast implementation, evaluation of its computational complexity, and an extensive set of analyses and results.

C. ORGANIZATION

Section II describes the general DL transmission model for massive MIMO systems, presents the LoS channel propagation model, defines the favorable propagation, and presents the SE for a massive MIMO system under LoS propagation. Section III presents the main characteristics of the *maximum ratio transmitter* MRT and ZF precoders, whereas Section IV details the SOS, CBS, and ICIBS user selection algorithms. In Section V, the complexity of the user selection algorithms is discussed. In Section VI, simulation results considering many different setups are presented to compare the performance of the aforementioned user selection algorithms. The concluding remarks are drawn in Section VII.

D. NOTATION

Vectors and matrices are represented by boldface lowercase and uppercase letters, respectively. The notation $\mathbf{X}^{\mathrm{T}}$, $\mathbf{X}^{*}$, $\mathbf{X}^{\mathrm{H}}$, and $\mathbf{X}^{-1}$ stand for transpose, complex conjugate, Hermitian (conjugate transpose), and inverse operations on $\mathbf{X}$, respectively. $\mathsf{Diag}(\mathbf{x})$ is a diagonal matrix with $\mathbf{x}$ on its main diagonal, and $\mathsf{span}\{\mathbf{x}_1, \ldots, \mathbf{x}_K\}$ is the set of all linear combinations of the vectors $\mathbf{x}_1, \ldots, \mathbf{x}_K$. The symbols $\mathbb{C}$, $\mathbb{R}$, $\mathbb{R}_+$, $\mathbb{Z}$, and $\mathbb{N}$ denote the sets of complex, real, non-negative real, integer, and natural numbers, respectively. The set $\mathbb{C}^{M\times K}$ denotes all $M \times K$ matrices comprised of complex-valued entries. The symbol $\mathbf{I}_M$ denotes an $M\times M$ identity matrix and $\mathbf{0}_{M\times K}$ denotes an $M \times K$ zero matrix. The symbol $\otimes$ stands for the Kronecker product. The notation $\mathsf{abs}(\mathbf{X}) \in \mathbb{C}^{M\times K}$ represents the element-wise absolute value of the entries of $\mathbf{X} \in \mathbb{C}^{M\times K}$, whereas $|\mathcal{X}|$ stands for the number of elements (cardinality) of the set $\mathcal{X}$. The symbols $\mathcal{CN}(\mathbf{m}, \mathbf{C})$ and $\mathcal{U}(a, b)$ respectively denote a circularly symmetric Gaussian distribution with mean $\mathbf{m}$ and covariance matrix $\mathbf{C}$ and a uniform distribution between a and b. The notations $\mathsf{E}\{x\}$ and $\mathsf{Var}\{x\}$ stand for the expected value and variance of x, and $x \xrightarrow{p} y$ means that x converges to y in probability.

II. SYSTEM MODEL

A. DOWNLINK TRANSMISSION

Consider a single-cell massive MIMO system equipped with an M-antenna base station that serves K single-antenna users as illustrated in Fig. 1. For a DL transmission operating in

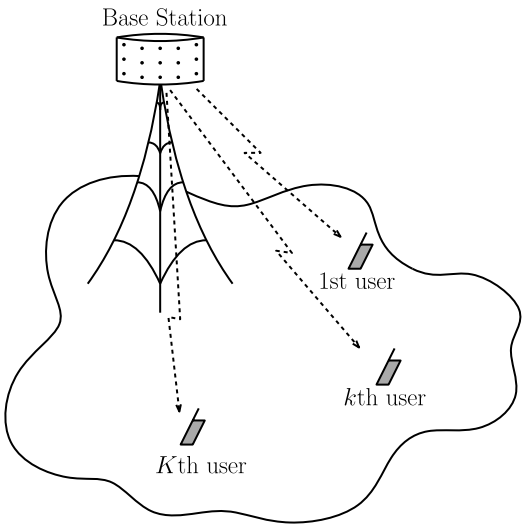


FIGURE 1. Massive MIMO system with an *M*-antenna base station serving *K* single-antenna users.

time-division duplex (TDD) mode, the received signal by the users can be expressed as

$$\mathbf{y} = \sqrt{\rho}\,\mathbf{G}^{\mathrm{T}}\mathbf{x} + \mathbf{n}, \tag{1}$$

where vector $\mathbf{y} \in \mathbb{C}^{K\times 1}$ contains the signals received by the K users, $\rho \in \mathbb{R}_+$ is the DL *signal-to-noise ratio* (SNR), $\mathbf{G} \in \mathbb{C}^{M\times K}$ is the massive MIMO channel matrix, $\mathbf{x} \in \mathbb{C}^{M\times 1}$ is the precoded signal transmitted by the BS, and $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}_{K\times 1},\ \mathbf{I}_K)$ is the additive noise. Due to power limitations, the precoded signal is constrained by

$$\mathsf{E}\{\mathbf{x}^{\mathrm{H}}\mathbf{x}\} \leq 1. \tag{2}$$

The transmitted signal is digitally precoded (i.e., the number of radio-frequency chains is equal to the number of antennas) before the transmission in order to mitigate the effects of the channel. In massive MIMO, linear precoding is asymptotically optimal in terms of achievable spectral efficiency. The linear-precoded signal can be written as

$$\mathbf{x} = \mathbf{W}\mathrm{Diag}(\boldsymbol{\eta})^{1/2}\mathbf{s}, \tag{3}$$

where $\mathbf{W} \in \mathbb{C}^{M\times K}$ is the precoding matrix, whose columns $\mathbf{w}_k$ are normalized (i.e., $\|\mathbf{w}_k\|_2 = 1$), $\boldsymbol{\eta} \in \mathbb{R}_+^{K\times 1}$ is the power allocation vector, and $\mathbf{s} \sim \mathcal{CN}(\mathbf{0}_{K\times 1},\ \mathbf{I}_K)$ is the vector of transmitted symbols.

B. MASSIVE MIMO CHANNEL

The massive MIMO channel can be written as [1]

$$\mathbf{G} = \mathbf{H}\,\mathrm{Diag}(\boldsymbol{\beta})^{1/2}, \tag{4}$$

where $\mathbf{H} \in \mathbb{C}^{M\times K}$ is the small-scale fading matrix and $\boldsymbol{\beta} \in \mathbb{R}_+^{K\times 1}$ is the large-scale fading vector. For massive MIMO systems, the most common models for the small-scale fading are the i.i.d. Rayleigh fading and the *uniformly random line-of-sight* (UR-LoS). These two types of channels model extreme propagation scenarios; for example, the UR-LoS models an LoS scenario whereas the i.i.d. Rayleigh models a NLoS scenario. The former is more suitable for short-range communications, and the latter is more suited for long-range communications. In practice, the actual propagation channel is likely to lie in-between these two extreme cases.

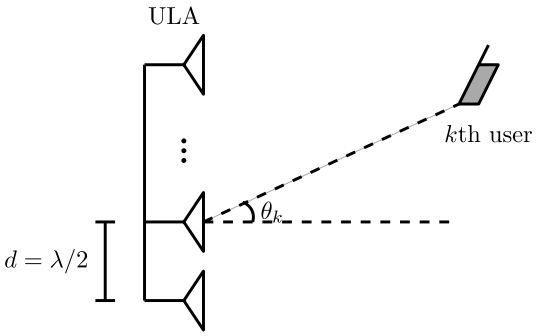


FIGURE 2. Line-of-sight channel between the *k*th user and the *M*-antenna ULA.

The UR-LoS channel vector with the M-antenna standard *uniform linear array* (ULA) illustrated in Fig. 2 is given by[2]

$$\mathbf{g}_k = \sqrt{\beta_k}\mathbf{h}_k,\ \ \forall\, k \in \mathcal{K}, \tag{5}$$

where $\mathbf{h}_k \in \mathbb{C}^{M\times 1}$ is the small-scale fading vector denoted by [38]

$$\mathbf{h}_k = \mathrm{e}^{\mathrm{j}\phi_k}\left[1\ \mathrm{e}^{-\mathrm{j}\pi\sin(\theta_k)}\ \cdots\ \mathrm{e}^{-\mathrm{j}(M-1)\pi\sin(\theta_k)}\right]^{\mathrm{T}}, \tag{6}$$

$\phi_k \sim \mathcal{U}(-\pi,\ \pi)$ is the phase shift associated with the array and the kth user, $\theta_k \sim \mathcal{U}(-\pi,\ \pi)$ is the AoA for the kth user, and $\mathcal{K} = \{1, 2, \ldots, K\}$ is the set of all user indices.

Another popular array used in massive MIMO system is the *uniform rectangular array* (URA). This type of array, in addition to the azimuth angle, can also steer a beam in the elevation angle. Further, the use of URA enables the placement of the same number of antennas of ULA in a smaller area, allowing its use in space-limited scenarios. The small-scale fading vector for this array is given by [7]

$$\mathbf{h}_k = \mathrm{e}^{\mathrm{j}\phi_k}\mathbf{h}_k^{(1)} \otimes \mathbf{h}_k^{(2)} \in \mathbb{C}^{M_1M_2\times 1}, \tag{7}$$

where

$$\mathbf{h}_k^{(1)} = \left[1\ \mathrm{e}^{-\mathrm{j}\pi\sqrt{1-\nu_k^2}\cos(\phi_k)}\ \cdots\ \mathrm{e}^{-\mathrm{j}(M_1-1)\pi\sqrt{1-\nu_k^2}\cos(\phi_k)}\right]^{\mathrm{T}},$$
$$\mathbf{h}_k^{(2)} = \left[1\ \mathrm{e}^{-\mathrm{j}\pi\sqrt{1-\nu_k^2}\sin(\phi_k)}\ \cdots\ \mathrm{e}^{-\mathrm{j}(M_2-1)\pi\sqrt{1-\nu_k^2}\sin(\phi_k)}\right]^{\mathrm{T}},$$

$M_1 \in \mathbb{N}$ and $M_2 \in \mathbb{N}$ are the number of antennas in x- and y-axes, $\phi_k \sim \mathcal{U}(-\pi,\ \pi)$ is the elevation angle for the kth user, and $\nu_k = \sin(\theta_k)$.

This paper adopts the ULA as standard for the derivations due to its mathematical tractability. In practice, however, the use of ULAs may be limited for some frequency ranges due to the physical space required to deploy a ULA with so many antennas. On the other hand, the ULA yields better SE than

2. A standard ULA is an array in which the distance between adjacent antennas d is equal to half of the signal wavelength λ [38].

other arrays such as the URA and the *uniform circular array* (UCA) in practical scenarios, with and without LoS [16], [39]–[41]. Moreover, the ULA simplifies the *inter-channel interference* model, which plays a fundamental role in user selection algorithms.

Remark 1: One of the main differences between the i.i.d. Rayleigh channel and the UR-LoS channel is the1 ℓ_2-norm of their small-scale fading vectors. For the UR-LoS channel, $\|\mathbf{h}_k\|_2^2 = M$, for any value of M. However, for the i.i.d. Rayleigh channel, $\|\mathbf{h}_k\|_2^2 \xrightarrow{p} M$, as $M \longrightarrow \infty$, which is the so-called *channel hardening* property [3].

C. FAVORABLE PROPAGATION

Intuitively, to maximize the performance of a wireless communication system from an information-theoretic perspective, the massive MIMO channel vectors should be as different as possible, according to some appropriate metric. This appropriate metric is the so-called *favorable propagation* offered by the channel [6], [42], [43], defined as

$$\mathbf{h}_k^{\mathrm{H}}\mathbf{h}_{k'} = 0, \quad \forall k, k' \in \mathcal{K}, \text{ with } k \neq k', \tag{8}$$

meaning that FP requires the massive MIMO channel vectors corresponding to different users to be pair-wise orthogonal. Since there are many cases that violate the FP condition, as it will be discussed later, it is interesting to measure how close two channel vectors are from being orthogonal. To do so, we define

$$r_{kk'} = \frac{|\mathbf{h}_k^{\mathrm{H}}\mathbf{h}_{k'}|}{\|\mathbf{h}_k\|_2\|\mathbf{h}_{k'}\|_2}, \tag{9}$$

i.e., $r_{kk'}$ is the magnitude of the channel correlation coefficient. For the UR-LoS channel, $r_{kk'}$ is given in *Theorem* 1.

Theorem 1: The magnitude of the channel correlation coefficient of a UR-LoS massive MIMO channel is given by [6]

$$r_{kk'} = \frac{1}{M}\left|\frac{\sin\left(M\frac{\pi}{2}(\sin\theta_{k'} - \sin\theta_k)\right)}{\sin\left(\frac{\pi}{2}(\sin\theta_{k'} - \sin\theta_k)\right)}\right|. \tag{10}$$

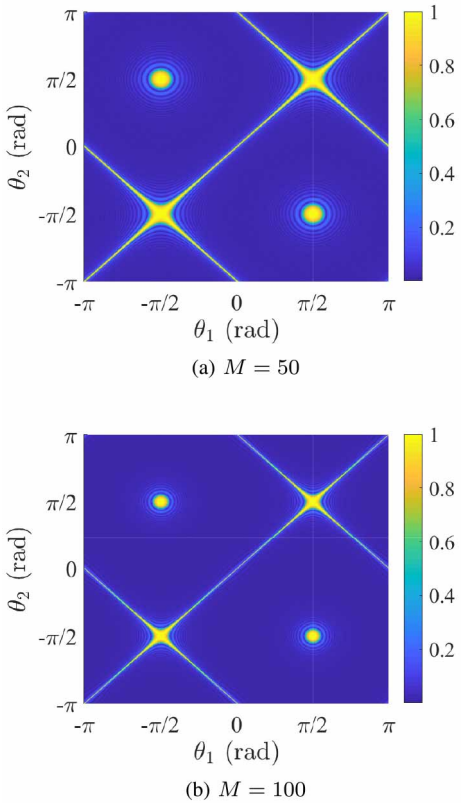
From Theorem 1, the UR-LoS channel offers FP when (10) is equal to zero for all $k, k' \in \mathcal{K}$, with $k \neq k'$. Thus, the numerator of (10) must be equal to zero and the denominator of (10) must be nonzero. The condition for which FP holds in an LoS channel is given in the following corollary of *Theorem* 1.

Corollary 1: The UR-LoS channel offers favorable propagation iff for all $k' \in \mathcal{K}\backslash\{k\}$, $\exists\, n_{k'} \in \{\pm1, \pm2, \ldots, \pm(M-1)\}$ such that

$$\theta_{k'} = \arcsin\left(\sin\theta_k + \frac{2n_{k'}}{M}\right), \tag{11}$$

and all of the following three conditions are satisfied:

1) $\theta_{k'} \neq \theta_k$, $\forall k, k' \in \mathcal{K}$, with $k \neq k'$;
2) $\theta_{k'} \neq \pi - \theta_k$, $\forall k, k' \in \mathcal{K}$, with $k \neq k'$;
3) If $\theta_k = \pm\pi/2$ for some $k \in \mathcal{K}$, then $\theta_{k'} \neq \mp\pi/2$ $\forall k' \in \mathcal{K} \setminus \{k\}$.



(a) $M = 50$

(b) $M = 100$

FIGURE 3. Magnitude of the correlation coefficient for $K = 2$ and $M \in \{50,100\}$.

Proof: The proof is given in Appendix A. ■

Corollary 1 is a more general version of the result provided in [6], in that we state the exact user positions that lead to favorable propagation.

Fig. 3 depicts the magnitude of the correlation coefficient with varying positions for two users, $K = 2$, for BS with $M = 50$ and 100 antennas. As can be observed, the massive MIMO channel offers favorable propagation or a moderately favorable propagation for different values of θ_1 and θ_2. However, due to the geometry of the array the users have high correlation when they have the same AoA, or when $\theta_1 = \pm\pi/2$ with $\theta_2 = \mp\pi/2$, or when $\theta_2 = -\theta_1 \pm \pi$. In these regions, the users have high mutual interference, degrading the SE performance of the massive MIMO system.

Remark 2: In order to validate *Corollary* 1, consider two users, k_1 and k_2, whose channels are orthogonal to the kth user, then

$$\theta_{k_1} = \arcsin\left(\sin\theta_k + \frac{2n_{k_1}}{M}\right), \tag{12}$$

$$\theta_{k_2} = \arcsin\left(\sin\theta_k + \frac{2n_{k_2}}{M}\right), \tag{13}$$

for some $n_{k_1}, n_{k_2} \in \{\pm1, \ldots, \pm(M-1)\}$. The magnitude of the correlation coefficient between these two users can be

written as

$$r_{k_1k_2} = \frac{1}{M}\left|\frac{\sin\left(M\frac{\pi}{2}\left(\sin\theta_{k_2}-\sin\theta_{k_1}\right)\right)}{\sin\left(\frac{\pi}{2}\left(\sin\theta_{k_2}-\sin\theta_{k_1}\right)\right)}\right|, \tag{14}$$

as long as the denominator is nonzero, which boils down to the condition $\sin\theta_{k_2}-\sin\theta_{k_1}\notin\{-2,0,2\}$. That means $\theta_{k_2}\neq\theta_{k_1}$, $\theta_{k_2}\neq\pi-\theta_{k_1}$, and if $\theta_{k_1}=\pm\pi/2$, then $\theta_{k_2}\neq\mp\pi/2$. Note that, based on (12) and (13), the sine function in the numerator of the correlation coefficient in (14) can be rewritten as $\sin(\pi(n_{k_2}-n_{k_1}))$, which is always zero. This result shows that if two different users are orthogonal to a third one, then they will themselves be orthogonal to one another if they satisfy *Corollary* 1.

It is worth highlighting that although the set of possible values for $n_{k'}$ is comprised of $2(M-1)$ integers, as stated in *Corollary* 1, only $M-1$ of these values generate a valid solution. By a valid solution, we mean choices of $n_{k'}$ that, simultaneously, yield $\sin\theta_k+2n_{k'}/M$ lying in $[-1,1]$ and $\theta_{k'}$ satisfying the conditions stated in *Corollary* 1.

Remark 3: Another important difference between the i.i.d. Rayleigh and UR-LoS channels lies in the asymptotically favorable propagation property. The i.i.d. Rayleigh channel does not offer favorable propagation, only asymptotically favorable propagation, which is defined as [44]

$$\frac{1}{M}\mathbf{h}_k^{\mathrm{H}}\mathbf{h}_{k'} \xrightarrow{p} 0, \quad \forall k,k'\in\mathcal{K}, \text{ with } k\neq k', \tag{15}$$

and $K\ll M\longrightarrow\infty$. The asymptotically favorable propagation is one of the key factors to enable massive MIMO systems with linear precoding [44]. In general, UR-LoS channel can provide more favorable propagation condition than the i.i.d. Rayleigh channel. However, the UR-LoS channel model can also provide non favorable channels with a non negligible probability, especially when users are very close to each other in the cell. In these cases, employing a user selection algorithm is of paramount importance for the SE of the massive MIMO system.

Remark 4: For the the URAs, there is an extra degree of freedom in addition to the azimuth angle, which is the elevation angle. This means that URAs can distinguish users (through their channel vectors exhibiting low correlation coefficient $r_{kk'}$) in both azimuth and elevation, whereas ULAs have only the azimuth, thus cannot distinguish users that have the same azimuth but different elevations. For a ULA and a URA with the same number of antennas, the ULA presents higher resolution in the azimuth coordinate and no resolution in the elevation coordinate. In comparison with the ULA, one can think that the URA trades resolution in the azimuth for resolution in the elevation coordinate/angle.

D. SPECTRAL EFFICIENCY

Having CSI in the BS is critical for the proper operation of massive MIMO systems. The BS estimates the channel through an uplink pilot transmission by receiving K signals related to orthogonal pilot signals of length (in samples) $\tau_{\mathrm{p}}\in\mathbb{N}$ sent by the users. Mathematically, the kth user transmits a pilot signal $\boldsymbol{\varphi}_k\in\mathbb{C}^{\tau_{\mathrm{p}}\times 1}$, which is the kth column of a matrix $\boldsymbol{\Phi}\in\mathbb{C}^{\tau_{\mathrm{p}}\times K}$ that satisfies $\boldsymbol{\Phi}^{\mathrm{H}}\boldsymbol{\Phi}=\mathbf{I}_K$. Then, the signal transmitted by the kth user is given by

$$\mathbf{x}_{\mathrm{p},k} = \sqrt{\tau_{\mathrm{p}}}\boldsymbol{\varphi}_k^{\mathrm{H}}, \tag{16}$$

where $\tau_{\mathrm{c}}\geq\tau_{\mathrm{p}}\geq K$, with $\tau_{\mathrm{c}}\in\mathbb{N}$ being the coherence time in samples. Therefore, the signal received by the BS is given by

$$\begin{aligned}\mathbf{Y}_{\mathrm{p}} &= \sqrt{\rho_{\mathrm{p}}}\,\mathbf{G}\mathbf{X}_{\mathrm{p}}+\mathbf{N}_{\mathrm{p}}\\ &= \sqrt{\rho_{\mathrm{p}}\tau_{\mathrm{p}}}\,\mathbf{H}\mathrm{Diag}(\boldsymbol{\beta})^{1/2}\boldsymbol{\Phi}^{\mathrm{H}}+\mathbf{N}_{\mathrm{p}},\end{aligned} \tag{17}$$

where $\mathbf{X}_{\mathrm{p}}=\sqrt{\tau_{\mathrm{p}}}\boldsymbol{\Phi}^{\mathrm{H}}$, $\rho_{\mathrm{p}}\in\mathbb{R}_+$ is the pilot transmission SNR, and the entries of $\mathbf{N}_{\mathrm{p}}\in\mathbb{C}^{M\times\tau_{\mathrm{p}}}$ are i.i.d. circularly symmetric Gaussian distributed with zero mean and unitary variance. In order to estimate the channel of the kth user, the BS processes the received signal $\mathbf{Y}_{\mathrm{p}}$ as follows:

$$\begin{aligned}\mathbf{y}'_{\mathrm{p},k} &= \mathbf{Y}_{\mathrm{p}}\boldsymbol{\varphi}_k\\ &= \sqrt{\rho_{\mathrm{p}}\tau_{\mathrm{p}}}\,\mathbf{H}\mathrm{Diag}(\boldsymbol{\beta})^{1/2}\boldsymbol{\Phi}^{\mathrm{H}}\boldsymbol{\varphi}_k+\mathbf{N}_{\mathrm{p}}\boldsymbol{\varphi}_k\\ &= \sqrt{\rho_{\mathrm{p}}\tau_{\mathrm{p}}\beta_k}\,\mathbf{h}_k+\mathbf{n}'_{\mathrm{p},k},\end{aligned} \tag{18}$$

where $\mathbf{n}'_{\mathrm{p},k}\sim\mathcal{CN}(\mathbf{0}_{M\times 1},\mathbf{I}_M)$ since $\|\boldsymbol{\varphi}_k\|_2=1$. Finally, assuming the knowledge of the large-scale fading, the channel estimate for the kth user is obtained from $\mathbf{y}'_{\mathrm{p},k}$ and is given by

$$\begin{aligned}\hat{\mathbf{h}}_k &= \mathbf{h}_k+\frac{\mathbf{n}'_{\mathrm{p},k}}{\sqrt{\rho_{\mathrm{p}}\tau_{\mathrm{p}}\beta_k}}\\ &= \mathbf{h}_k+\boldsymbol{\varepsilon}_k,\ \forall k\in\mathcal{K},\end{aligned} \tag{19}$$

where $\boldsymbol{\varepsilon}_k\sim\mathcal{CN}(\mathbf{0}_{M\times 1},\ \sigma_\varepsilon^2\mathbf{I}_M)$ is the channel estimation error and $\sigma_\varepsilon^2=1/\rho_{\mathrm{p}}\tau_{\mathrm{p}}\beta_k$ represents the "quality" of the channel estimation. Moreover, this estimate is normalized to guarantee that $\left\|\hat{\mathbf{h}}_k\right\|_2=\sqrt{M}$.

Under an LoS propagation scenario, the AoA is constant for few coherence time intervals. Thus, one can assume that the effective channel is known by the users. The SE for a massive MIMO UR-LoS channel is given in the following theorem.

Theorem 2: The spectral efficiency of a massive MIMO system under LoS propagation with linear precoding and channel estimate as (19) is lower bounded by

$$R_{\mathrm{sum}}\geq\frac{1}{2}\left(1-\frac{\tau_{\mathrm{p}}}{\tau_{\mathrm{c}}}\right)\sum_{k\in\mathcal{K}}\log_2(1+\gamma_k), \tag{20}$$

where $\gamma_k\in\mathbb{R}_+$ is the SINR in (21) on the bottom of the next page. Additionally, the equality holds when $\sigma_\varepsilon^2=0$.

Proof: The proof is given in Appendix B. ■

Remark 5: The LoS propagation is expected to occur within the context of short-range communications, while inter-cell interference is expected to be i.i.d. Rayleigh distributed [7], [14]. Thus, in a multi-cell model, the interference from different cells, i.e., pilot contamination, can be incorporated in the channel estimation error since the multi-cell interference is i.i.d. Rayleigh distributed. However,

the single-cell scenario is frequently adopted since it can easily isolate the effect of the LoS channel [7], [8], [10], [45].

Remark 6: Although in this section we used uplink pilot transmission to estimate MK complex-valued coefficients, it is possible to use different estimation approaches since we are dealing with LoS propagation. Under LoS propagation, it is possible to estimate only the real-valued AoAs and the complex gain for each user. This approach is usually less computationally expensive and demands less pilot signals, which leads to a better usage of the communication resources in general. However, it is important to highlight that such approach would provide a different model to (19), (20), and (21).

III. PRECODING SCHEMES

Precoding is a crucial step for the DL operation of massive MIMO systems that is employed to mitigate multi-user interference. Linear precoding schemes are particularly useful for massive MIMO systems, achieving the same performance of the dirty paper code in the asymptotic case [2]. By definition, the precoding vectors have $\|\mathbf{w}_k\|_2^2 = 1$, and hence, every linear precoding vector can be written as

$$\mathbf{w}_k = \frac{\mathbf{v}_k}{\|\mathbf{v}_k\|_2}. \tag{22}$$

A. MAXIMUM RATIO TRANSMITTER

The maximum ratio transmitter aims to amplify the signal of interest for each user, disregarding the impact of this amplification on other users. In the *multiple-input single-output* (MISO) case, the MRT is the precoding scheme that yields the highest capacity for this type of system by maximizing the SNR of the desired user. The MRT precoding vector is given by [42]

$$\mathbf{w}_k^{\text{MRT}} = \frac{\hat{\mathbf{h}}_k^*}{\sqrt{M}}, \ \forall\ k \in \mathcal{K}. \tag{23}$$

The DL SE of a massive MIMO UR-LoS channel with MRT can be easily derived from (20), by replacing $\mathbf{w}_k$ with $\mathbf{w}_k^{\text{MRT}}$. Therefore, the SE is bounded by

$$R_{\text{sum}}^{\text{MRT}} \geq \frac{1}{2}\left(1 - \frac{\tau_{\text{p}}}{\tau_{\text{c}}}\right) \sum_{k\in\mathcal{K}} \log_2\left(1 + \gamma_k^{\text{MRT}}\right), \tag{24}$$

where $\gamma_k^{\text{MRT}} \in \mathbb{R}_+$ is denoted by

$$\gamma_k^{\text{MRT}} = \frac{\rho M \beta_k \eta_k}{1 + \rho\beta_k M \sum_{k'\in\mathcal{K}\setminus\{k\}} \eta_{k'} r_{kk'}^2 + \rho\beta_k\sigma_\varepsilon^2 \sum_{k'\in\mathcal{K}} \eta_{k'}}. \tag{25}$$

It is worth highlighting the dependence of the small-scale fading in (24). The magnitude of the channel correlation can be used as a measure of favorable propagation for the massive MIMO channel. In this case, a non-favorable channel can have a significant impact on the system performance. For the i.i.d. Rayleigh channel, the favorable propagation condition is basically impacted by the number of antennas M. However, for the UR-LoS channel, even with a very large number of antennas, the system is still susceptible to have a non-favorable channel as explained in Section II-C.

B. ZERO-FORCING

In the zero-forcing precoder, the precoding vectors $\mathbf{w}_k$ are selected in order to satisfy the zero-interference condition $\mathbf{h}_k^{\text{T}}\mathbf{w}_{k'} = 0\ \forall k' \in \mathcal{K}\setminus\{k\}$. The zero-interference condition can be achieved by using the Moore-Penrose inverse of the channel matrix as the precoding matrix. In the case of full CSI knowledge, the ZF precoding can perfectly eliminate the interference among the users. The ZF precoding vector is given by [42]

$$\mathbf{w}_k^{\text{ZF}} = \frac{\hat{\mathbf{H}}^*\mathbf{r}_k}{\left\|\hat{\mathbf{H}}^*\mathbf{r}_k\right\|_2}, \ \forall\ k \in \mathcal{K}, \tag{26}$$

where $\mathbf{r}_k$ is the kth column of $\mathbf{R} = (\hat{\mathbf{H}}^{\text{T}}\hat{\mathbf{H}}^*)^{-1}$. Note that the Moore-Penrose inverse computation is only possible if $M \geq K$ and $\hat{\mathbf{H}}^{\text{T}}$ is a full-row rank matrix.

Corollary 2: The DL SE of a massive MIMO UR-LoS channel with ZF precoding is lower bounded by

$$R_{\text{sum}}^{\text{ZF-1}} \geq \frac{1}{2}\left(1 - \frac{\tau_{\text{p}}}{\tau_{\text{c}}}\right) \sum_{k\in\mathcal{K}} \log_2\left(1 + \gamma_k^{\text{ZF-1}}\right), \tag{27}$$

where $\gamma_k^{\text{ZF-1}} \in \mathbb{R}_+$ is denoted by

$$\gamma_k^{\text{ZF-1}} = \frac{\rho\beta_k\eta_k}{[(\hat{\mathbf{H}}^{\text{T}}\hat{\mathbf{H}}^*)^{-1}]_{kk}\left(1 + \rho\beta_k\sigma_\varepsilon^2 \sum_{k'\in\mathcal{K}} \eta_{k'}\right)}. \tag{28}$$

Proof: The proof is given in Appendix C. ■

Like in (24), (27) also depends on the small-scale fading, albeit this dependence is less evident in comparison with the MRT case. If $\hat{\mathbf{H}}^{\text{T}}$ is a full-row rank matrix, then all users are geographically separated, i.e., have different AoAs. However, even in this case, if they have similar AoAs, then $\hat{\mathbf{H}}^{\text{T}}$ becomes ill-conditioned leading to performance degradation.

Remark 7: Instead of normalizing the precoding vector $\mathbf{w}$, one can normalize the precoded vector $\mathbf{x}$ as explained below. Let the ZF-precoded signal $\mathbf{x}$ be

$$\mathbf{x} = \sqrt{c}\,\hat{\mathbf{H}}^*\left(\hat{\mathbf{H}}^{\text{T}}\hat{\mathbf{H}}^*\right)^{-1}\text{Diag}(\boldsymbol{\eta})^{1/2}\mathbf{s}, \tag{29}$$

$$\gamma_k = \frac{\rho\beta_k\eta_k\left|\hat{\mathbf{h}}_k^{\text{T}}\mathbf{w}_k\right|^2}{1 + \rho\beta_k \sum_{k'\in\mathcal{K}\setminus\{k\}} \eta_{k'}\left|\hat{\mathbf{h}}_k^{\text{T}}\hat{\mathbf{w}}_{k'}\right|^2 + \rho\beta_k\sigma_\varepsilon^2 \sum_{k'\in\mathcal{K}} \eta_{k'}}, \ \forall k \in \mathcal{K} \tag{21}$$

where $c \in \mathbb{R}_+$ is the normalization constant. The power of $\mathbf{x}$ is given by

$$\mathsf{E}\{\mathbf{x}^{\mathrm{H}}\mathbf{x}\} = c\,\mathsf{E}\left\{\mathbf{s}^{\mathrm{H}}\mathsf{Diag}(\boldsymbol{\eta})^{1/2}(\hat{\mathbf{H}}^{\mathrm{H}}\hat{\mathbf{H}})^{-1}\mathsf{Diag}(\boldsymbol{\eta})^{1/2}\mathbf{s}\right\}. \tag{30}$$

Using the equality in (2) and (30), one gets [7]

$$c = \frac{1}{\sum_{k\in\mathcal{K}}[(\hat{\mathbf{H}}^{\mathrm{H}}\hat{\mathbf{H}})^{-1}]_{kk}\eta_k}. \tag{31}$$

Therefore, the new SE expression is bounded by

$$R_{\mathrm{sum}}^{\mathrm{ZF\text{-}2}} \geq \frac{1}{2}\left(1-\frac{\tau_{\mathrm{p}}}{\tau_{\mathrm{c}}}\right)\sum_{k\in\mathcal{K}}\log_2\left(1+\gamma_k^{\mathrm{ZF\text{-}2}}\right), \tag{32}$$

where $\gamma_k^{\mathrm{ZF\text{-}2}} \in \mathbb{R}_+$ is denoted by

$$\gamma_k^{\mathrm{ZF\text{-}2}} = \frac{\rho\beta_k\eta_k}{\left(1+\rho\beta_k\sigma_\varepsilon^2\sum_{k'\in\mathcal{K}}\eta_{k'}\right)\sum_{k'\in\mathcal{K}}\left[\left(\hat{\mathbf{H}}^{\mathrm{H}}\hat{\mathbf{H}}\right)^{-1}\right]_{k'k'}\eta_{k'}}. \tag{33}$$

IV. USER SELECTION

As discussed in Sections II and III, the small-scale fading plays an important role in the SE of massive MIMO systems under UR-LoS channels, degrading the performance in non-favorable propagation conditions. Additionally, the number of users can also be a critical aspect when the ZF precoder is used and $K > M$. One way to cope with these issues is by performing user selection.

The user selection problem has not received much attention in the massive MIMO research for several reasons: Firstly, massive MIMO systems rely on the favorable propagation and channel hardening properties, which consider that the number of antennas at the BS is much larger than the number of users ($M \gg K$). Secondly, i.i.d. Rayleigh channel model is assumed, which is not severely affected by the position of the users [6]. However, these assumptions may not fully hold in practical scenarios; for example, there may be situations where the number of users is very close to the number of BS antennas or even larger. The MRT precoding algorithm used in massive MIMO does not have a restriction on the number of users, but the ZF precoding does and requires the number of users to be smaller or equal to the number of BS antennas. One critical issue here is how to select the best set of users in order to yield the highest spectral efficiency to the network and the remaining users.

Consider a case where K is large enough to degrade the massive MIMO system performance, yielding a non-favorable propagation environment or simply a situation where $K > M$. In this scenario, the BS should choose L out of K users to transmit and receive data in a given time slot in order to guarantee $M \geq K$ and provide favorable propagation. Mathematically, let $\mathcal{S} \subset \mathcal{K}$ be the set of selected users in a given time slot. The user selection problem can be written as an optimization problem, which consists of finding $\mathcal{S}$, with $|\mathcal{S}| = L$, that maximizes the SE,[3] i.e.,

$$\begin{aligned} &\underset{\mathcal{S}}{\text{maximize}} \quad \frac{1}{2}\left(1-\frac{\tau_{\mathrm{p}}}{\tau_{\mathrm{c}}}\right)\sum_{k\in\mathcal{S}}\log_2(1+\gamma_k) \\ &\text{subject to} \quad |\mathcal{S}| = L. \end{aligned} \tag{34}$$

The user selection problem in (34) is not only non-convex, but also involves combinatorial optimization. In general, the solution can only be found through an exhaustive search, which is impractical due to the high-dimensional search space in massive MIMO systems. Some suboptimal approaches for the user selection problem are the semi-orthogonal selection [23], correlation-based selection [7], and *inter-channel interference-based selection* (ICIBS), which is a new proposed approach described in Section IV-C.

Remark 8: Despite not providing performance improvements to massive MIMO systems under i.i.d. Rayleigh fading channels, user selection algorithms can bring advantages under Rician fading channels. In [46], the authors show that in the asymptotic case ($M \longrightarrow \infty$) the SE of a massive MIMO system under Rician fading is maximized when the channel is dominated by LoS propagation, and the LoS channel components corresponding to different users are orthogonal to one another. The orthogonality among these LoS components can be guaranteed by some user selection algorithms. Furthermore, the user selection may help the statistical beamforming approach proposed in [46] in dealing with some of the limitations due to the finite number of antennas and the scheduling process. Also, the user selection algorithms could be adapted to take advantage of the statistical CSI when dealing with Rician channels like in [46]. That is, under the statistical CSI hypothesis, the BS knows not only the LoS component but also other parameters of the Rician κ-factor model [47]; therefore, when dealing with these channel models, it would be interesting to add this additional piece of information into the user selection algorithms. In the Rician fading scenario, using only the LoS components is not enough to perform a good user selection since the Rician κ-factor is also relevant to maximize the SE of massive MIMO systems under Rician fading channels.

A. SEMI-ORTHOGONAL SELECTION

The semi-orthogonal selection is a user selection algorithm that is asymptotically optimum when $K \longrightarrow \infty$ in a MIMO system under i.i.d. Rayleigh fading channel [23]. This algorithm aims to select the set of users that yields the most orthogonal effective small-scale fading channel matrix. This selection is made in an iterative manner by choosing the user with the most orthogonal small-scale fading channel with respect to the previously selected users. At the lth iteration, the algorithm first calculates the orthogonal complement of $\mathbf{h}_k$ to the subspace $\mathsf{span}\{\mathbf{q}_1, \ldots, \mathbf{q}_{l-1}\}$ for all the $K-l+1$

3. An alternative criterion is used in [7], but the maximization of the SE is the most widely used [21]–[23], [26]–[31], and [37].

Algorithm 1 SOS Algorithm

Input: L and $\mathbf{H}$
Initialization: $\mathcal{T}_0 = \mathcal{K}$ and $\mathcal{S}_0 = \emptyset$
for $l = 1$ to L **do**
  if $l = 1$ **then**
    $\mathbf{h}_k^{\perp} = \mathbf{h}_k, \ \forall\, k \in \mathcal{T}_{l-1}$
  else
    $\mathbf{P}_{l-1} = \mathbf{I}_M - \mathbf{Q}_{l-1}\mathbf{Q}_{l-1}^{\mathrm{H}}$
    $\mathbf{h}_k^{\perp} = \mathbf{P}_{l-1}\mathbf{h}_k, \ \forall\, k \in \mathcal{T}_{l-1}$
  end if
  $k^{\star} = \underset{k \in \mathcal{T}_{l-1}}{\operatorname{argmax}} \left\| \mathbf{h}_k^{\perp} \right\|_2$
  $\mathbf{q}_l = \mathbf{h}_{k^{\star}}^{\perp} / \left\| \mathbf{h}_k^{\perp} \right\|_2$
  $\mathcal{T}_l = \mathcal{T}_{l-1} \setminus \{k^{\star}\}$
  $\mathcal{S}_l = \mathcal{S}_{l-1} \cup \{k^{\star}\}$
  $\mathbf{Q}_l = \left[\mathbf{q}_1 \ \mathbf{q}_2 \cdots \mathbf{q}_l\right]$
end for
return $\mathcal{S} = \mathcal{S}_L$

remaining users in $\mathcal{T}_{l-1}$. The orthogonal component $\mathbf{h}_k^{\perp}$ is given by

$$\mathbf{h}_k^{\perp} = \mathbf{P}_{l-1}\mathbf{h}_k, \ \forall\, k \in \mathcal{T}_{l-1}, \tag{35}$$

where $\mathbf{P}_{l-1} \in \mathbb{C}^{M \times M}$ is the orthogonal projection matrix denoted by

$$\mathbf{P}_{l-1} = \mathbf{I}_M - \mathbf{Q}_{l-1}\mathbf{Q}_{l-1}^{\mathrm{H}}, \tag{36}$$

and $\mathbf{Q}_{l-1} \in \mathbb{C}^{M \times l}$ is denoted by

$$\mathbf{Q}_{l-1} = \left[\mathbf{q}_1 \ \mathbf{q}_2 \cdots \mathbf{q}_{l-1}\right]. \tag{37}$$

The users are then selected based on the orthogonal complement with the largest ℓ_2-norm as follows:

$$k^{\star} = \underset{k \in \mathcal{T}_{l-1}}{\operatorname{argmax}} \left\| \mathbf{h}_k^{\perp} \right\|_2. \tag{38}$$

Finally, vector $\mathbf{q}_l$, sets $\mathcal{T}_l$, and $\mathcal{S}_l$ are updated as

$$\mathbf{q}_l = \frac{\mathbf{h}_{k^{\star}}^{\perp}}{\left\| \mathbf{h}_k^{\perp} \right\|_2}, \tag{39}$$

$$\mathcal{T}_l = \mathcal{T}_{l-1} \setminus \left\{k^{\star}\right\}, \tag{40}$$

$$\mathcal{S}_l = \mathcal{S}_{l-1} \cup \left\{k^{\star}\right\}. \tag{41}$$

This process is repeated until $l = L$. The SOS algorithm is summarized in Algorithm 1. Note that the first selected user is randomly chosen since for the UR-LoS fading channel, $\|\mathbf{h}_k\|_2 = \sqrt{M}, \ \forall k \in \mathcal{K}$.

Remark 9: The classical version of the SOS has a last step that forces the semi-orthogonality of the non-selected users at a given iteration. In the classical algorithm, $\mathcal{T}_l$ is updated as follows:

$$\mathcal{T}_l = \left\{ k \in \mathcal{T}_{l-1}, k \neq k^{\star} \middle| \frac{|\mathbf{h}_k^{\mathrm{H}}\mathbf{q}_l|}{\|\mathbf{h}_k\|_2 \|\mathbf{q}_l\|_2} < \delta \right\}, \tag{42}$$

where $\delta \in \mathbb{R}_+$ is the level of semi-orthogonality. This step speeds up the selection by eliminating the users with poor levels of orthogonality. This last step is useful for MIMO systems since the user selection is only performed in cases where $K > M$ and terminated when $L = M$. However, this is not necessarily the case for massive MIMO systems.

Algorithm 2 S-SOS Algorithm

Input: L and $\mathbf{H}$
Initialization: $\mathbf{P}_0 = \mathbf{I}_M$, $\mathcal{T}_0 = \mathcal{K}$, and $\mathcal{S}_0 = \emptyset$
for $l = 1$ to L **do**
  $\mathbf{h}_k^{\perp} = \mathbf{P}_{l-1}\mathbf{h}_k, \ \forall\, k \in \mathcal{T}_{l-1}$
  $k^{\star} = \underset{k \in \mathcal{T}_{l-1}}{\operatorname{argmax}} \left\| \mathbf{h}_k^{\perp} \right\|_2$
  $\mathbf{q}_l = \mathbf{h}_{k^{\star}}^{\perp} / \left\| \mathbf{h}_k^{\perp} \right\|_2$
  $\mathbf{P}_l = \mathbf{P}_{l-1} - \mathbf{q}_l\mathbf{q}_l^{\mathrm{H}}$
  $\mathcal{T}_l = \mathcal{T}_{l-1} \setminus \{k^{\star}\}$
  $\mathcal{S}_l = \mathcal{S}_{l-1} \cup \{k^{\star}\}$
end for
return $\mathcal{S} = \mathcal{S}_L$

The Algorithm 1 can be optimized if a recursive approach is used to compute the orthogonal projection matrix [48]. Equation (36) can be re-written as

$$\begin{aligned} \mathbf{P}_l &= \mathbf{I}_M - \mathbf{Q}_l\mathbf{Q}_l^{\mathrm{H}} \\ &= \mathbf{I}_M - \left[\mathbf{Q}_{l-1} \ \mathbf{q}_l\right]\begin{bmatrix}\mathbf{Q}_{l-1}^{\mathrm{H}} \\ \mathbf{q}_l^{\mathrm{H}}\end{bmatrix} \\ &= \mathbf{I}_M - \mathbf{Q}_{l-1}\mathbf{Q}_{l-1}^{\mathrm{H}} - \mathbf{q}_l\mathbf{q}_l^{\mathrm{H}} \\ &= \mathbf{P}_{l-1} - \mathbf{q}_l\mathbf{q}_l^{\mathrm{H}}. \end{aligned} \tag{43}$$

By using (43), we have the new SOS algorithm, named *simplified-SOS* (S-SOS), summarized in Algorithm 2.

B. CORRELATION-BASED SELECTION

The correlation-based selection algorithm is an alternative to the SOS algorithm. The CBS algorithm is a greedy method that searches for a pair of users (k, k') with the highest $r_{kk'}$ and removes the one with the highest magnitude of the correlation coefficient with the remaining users. From the SINR perspective, the CBS aims to maximize the SINR of one specific user, disregarding the SINR of the remaining users. It must be noted, however, that in some cases removing a user might inadvertently result in increased SINR of the remaining users. For example, the user with the highest correlation coefficient may have a high correlation with only one user and a small correlation with the others, and another user may have a moderate correlation with all other users. In this case, dropping the second user may lead to better improvement in SE than dropping the first one. The CBS algorithm is summarized in Algorithm 3.

C. INTER-CHANNEL INTERFERENCE-BASED SELECTION

The proposed inter-channel interference-based selection algorithm, ICIBS, can be seen as a generalization of the CBS algorithm [37]. Unlike CBS, which takes into account local (pair-wise) interference information, ICIBS considers

Algorithm 3 CBS Algorithm

Input: L and $\mathbf{G}$
Initialization: $\mathcal{S}_0 = \mathcal{K}$
for $l = 1$ to $K - L$ **do**
  $r_{kk'} = \frac{|\mathbf{h}_k^{\mathrm{H}}\mathbf{h}_{k'}|}{\|\mathbf{h}_k\|_2\|\mathbf{h}_{k'}\|_2}\ \forall\,(k,k') \in \mathcal{S}_{l-1} \times \mathcal{S}_{l-1}\,|\,k \neq k'$
  $(i,j) = \underset{(k,k')\in\mathcal{S}_{l-1}\times\mathcal{S}_{l-1}|k\neq k'}{\operatorname{argmax}} r_{kk'}$
  if $\max_{k'\in\mathcal{S}_{l-1}\setminus\{i,j\}} r_{ik'} > \max_{k'\in\mathcal{S}_{l-1}\setminus\{i,j\}} r_{jk'}$ **then**
    $k^\star = i$
  else if $\max_{k'\in\mathcal{S}_{l-1}\setminus\{i,j\}} r_{ik'} < \max_{k'\in\mathcal{S}_{l-1}\setminus\{i,j\}} r_{jk'}$ **then**
    $k^\star = j$
  end if
  $\mathcal{S}_l = \mathcal{S}_{l-1} \setminus \{k^\star\}$
end for
return $\mathcal{S} = \mathcal{S}_{K-L}$

Algorithm 4 ICIBS Algorithm

Input: L and $\mathbf{H}$
Initialization: $\mathcal{S}_0 = \mathcal{K}$
for $l = 1$ to $K - L$ **do**
  $\psi_k^{(l)} = \frac{1}{|\mathcal{S}_{l-1}|-1}\sum_{k'\in\mathcal{S}_{l-1}\setminus\{k\}} r_{kk'},\ \forall\, k \in \mathcal{S}_{l-1}$
  $k^\star = \underset{k\in\mathcal{S}_{l-1}}{\operatorname{argmax}}\ \psi_k^{(l)}$
  $\mathcal{S}_l = \mathcal{S}_{l-1} \setminus \{k^\star\}$
end for
return $\mathcal{S} = \mathcal{S}_{K-L}$

the global interference.[4] That is, at each iteration it finds the user whose removal will lead to the highest overall SINR gain for the remaining/selected users. Thus, starting with set $\mathcal{S}_0 = \mathcal{K}$, it iteratively generates $\mathcal{S}_l \subset \mathcal{S}_{l-1}$, $l \in \mathbb{N}$, by removing the user that maximizes the ICI defined as

$$\psi_k^{(l)} = \frac{1}{|\mathcal{S}_{l-1}| - 1} \sum_{k'\in\mathcal{S}_{l-1}\setminus\{k\}} r_{kk'},\ \forall\, k \in \mathcal{S}_{l-1}. \tag{44}$$

The algorithm stops at iteration $K - L$ since $|\mathcal{S}_{K-L}| = L$. It is worth highlighting the intuitive connection between the ICI and the SE. For example, by dropping the user with the highest ICI, we indirectly reduce the interference on the remaining users, increasing the SINR and consequently increasing the SE. Additionally, when the MRT precoder is used, the ICI has a more straightforward connection with the SE since the SE when MRT is used is upper and lower bounded by functions of the ICI [37]. The ICIBS method is summarized in Algorithm 4.

ICIBS can be improved by computing step 2 in an efficient manner through the following recursion

$$\psi_k^{(l)} = \frac{(|\mathcal{S}_{l-2}| - 1)\psi_k^{(l-1)} - r_{kk^\star}}{|\mathcal{S}_{l-1}| - 1}\ \forall\, k \in \mathcal{S}_{l-1}, \tag{45}$$

4. The term interference is being used loosely since the transmission powers are not taken into account. In fact, it is a *potential interference*. Further, we perform power allocation after user selection as in [7], [23], [26]–[28], [30], [31], [49].

Algorithm 5 S-ICIBS Algorithm

Input: L and $\mathbf{G}$
Initialization: $\mathcal{S}_0 = \mathcal{K}$
for $l = 1$ to $K - L$ **do**
  if $l = 1$ **then**
    $\psi_k^{(1)} = \frac{1}{|\mathcal{S}_0|-1}\sum_{k'\in\mathcal{S}_0\setminus\{k\}} r_{kk'},\ \forall\, k \in \mathcal{S}_0$
  else
    $\psi_k^{(l)} = \frac{(|\mathcal{S}_{l-2}|-1)\psi_k^{(l-1)} - r_{kk^*}}{|\mathcal{S}_{l-1}|-1}$
  end if
  $k^\star = \underset{k\in\mathcal{S}_{l-1}}{\operatorname{argmax}}\ \psi_k^{(l)}$
  $\mathcal{S}_l = \mathcal{S}_{l-1} \setminus \{k^\star\}$
end for
return $\mathcal{S} = \mathcal{S}_{K-L}$

where k^* is the user index that is removed from iteration $l-1$ to l. Hence, the summation in step 2 is computed just once at the first iteration, to obtain $\psi_k^{(1)}$. The new version of the ICIBS, aptly named *simplified-ICIBS* (S-ICIBS), is summarized in Algorithm 5.

V. COMPUTATIONAL COMPLEXITY

This section discusses and summarizes the computational complexities of the user selection approaches previously detailed in Section IV. Table 1 details the number of additions, multiplications, divisions, and square root operations that are required by each algorithm. In the following, we discuss the computational complexities of the SOS, S-SOS, CBS, ICIBS and S-ICIBS respectively.

A. SEMI-ORTHOGONAL SELECTION

The complexity of SOS presented in Algorithm 1 is measured in terms of the number of operations used to select the best set of users. The most expensive operations are in (35), (36), (38), and (39). At the lth iteration, Algorithm 1 requires $2(M^3 + M^2 - M)(K - l + 1)$ additions and $4M^3(K - l + 1)$ multiplications to compute (35). Moreover, it requires $4M^2(l - 1)$ additions and $4M^2(l - 1)$ multiplications to compute (36). Additionally, Algorithm 1 requires $2M$ divisions, $(K-l+1)$ ℓ_2-norm calculations, and one linear search in a $(K - l + 1)$- dimensional space to compute (38) and (39). However, in order to compute $\|\mathbf{x}\|_2$, where $\mathbf{x} \in \mathbb{C}^{M\times 1}$, $2M - 1$ additions, $2M$ multiplications, and 1 square root operation are needed. Therefore, the number of additions A_l^{SOS}, multiplications M_l^{SOS}, divisions D_l^{SOS}, and square root S_l^{SOS} computations required by Algorithm 1 per iteration are given by

$$\begin{aligned}
A_l^{\mathrm{SOS}} &= 2(M^3 + M^2 - M)(K - l + 1) \\
&\quad + (2M - 1)(K - l + 1)4M^2(l - 1), \\
M_l^{\mathrm{SOS}} &= 4M^3(K - l + 1) + 4M^2(l - 1) \\
&\quad + 2M(K - l + 1), \\
D_l^{\mathrm{SOS}} &= 2M, \\
S_l^{\mathrm{SOS}} &= K - l + 1,
\end{aligned}$$

TABLE 1. Number of operations and complexity of the user selection algorithms.

	Addition	Multiplication	Division	Square Root
SOS	$(M^3+M^2-M)(L-1)(2K-L)$ $+(2M-1)(KL-L(L-1)/2)$ $+2M^2L(L-1)$	$2M^2(L-1)(2MK-ML+L)$ $+M(2KL-L(L-1))$	$2ML$	$KL-L(L-1)/2$
S-SOS	$(M^3+M^2-M)(L-1)(2K-L)$ $+(2M-1)(KL-L(L-1)/2)$ $+4M^2(L-1)$	$2M^2(L-1)(2MK-ML+2)$ $+M(2KL-L(L-1))$	$2ML$	$KL-L(L-1)/2$
CBS	$4K^2M+K(K-1)/2$	$4K^2M+K^2-K$	$2MK$	$K(K-1)/2$
ICIBS	$4K^2M+K(K-1)(2K+5)/6$ $-L(L+1)(L-1)/3$	$4K^2M+K^2-K$	$2MK+K(K+1)/2$ $-L(L+1)/2$	$K(K-1)/2$
S-ICIBS	$4K^2M+K^2-L(L+1)/2$	$4K^2M+K(3K-1)/2$ $-L(L+1)/2$	$2MK+K(K+1)/2$ $-L(L+1)/2$	$K(K-1)/2$

respectively. By summing these quantities across L iterations the total number of a given operation can be obtained as summarized in Table 1. It is worth highlighting that the additions and multiplications are only computed from the second iteration onward.

One characteristic of the SOS algorithm is its incremental approach for selecting users, i.e., it starts with an empty set of selected users. This behavior can be seen in Table 1 with the number of operations necessary to compute the algorithm increasing with L. This incremental characteristic leads to low computational complexity when the number of selected users is very small.

B. SIMPLIFIED SEMI-ORTHOGONAL SELECTION

For the S-SOS the most expensive operations are equations (35), (38), (39), and (43). The only difference between S-SOS and SOS is in (36) that is replaced with (43), which leads to the following required number of additions $A_l^{\text{S-SOS}}$, multiplications $M_l^{\text{S-SOS}}$, divisions $D_l^{\text{S-SOS}}$, and square root $S_l^{\text{S-SOS}}$ computations per iteration are given by

$$\begin{aligned} A_l^{\text{S-SOS}} &= 2(M^3+M^2-M)(K-l+1)+4M^2 \\ &\quad + (2M-1)(K-l+1), \\ M_l^{\text{S-SOS}} &= 4M^3(K-l+1)+4M^2+2M(K-l+1), \\ D_l^{\text{S-SOS}} &= 2M, \\ S_l^{\text{S-SOS}} &= K-l+1, \end{aligned}$$

respectively. One can observe that the second term of $A_l^{\text{S-SOS}}$ and $M_l^{\text{S-SOS}}$ is not linear with index l, differently from A_l^{SOS} and M_l^{SOS}. This difference yields a reduced computational burden for S-SOS. The total number of operations required for S-SOS is summarized in Table 1.

C. CORRELATION-BASED SELECTION

The complexity of the CBS algorithm is mostly concentrated in the computation of the magnitude of the correlation coefficient $r_{kk'}$. In order to help in the calculation of $r_{kk'}$, we define the matrix $\boldsymbol{\Upsilon} \in \mathbb{R}_+^{K\times K}$ as

$$\boldsymbol{\Upsilon} = \begin{bmatrix} 0 & r_{12} & \cdots & r_{1K} \\ r_{21} & 0 & \cdots & r_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ r_{K1} & r_{K2} & \cdots & 0 \end{bmatrix}. \tag{46}$$

Note that matrix $\boldsymbol{\Upsilon}$ is not exactly a correlation matrix because it is a hollow matrix and does not have the r_{kk} information. Another definition for $\boldsymbol{\Upsilon}$ is given by

$$\boldsymbol{\Upsilon} = \text{abs}\big(\mathbf{I}_K - \bar{\mathbf{H}}^{\text{H}}\bar{\mathbf{H}}\big), \tag{47}$$

where $\bar{\mathbf{H}} \in \mathbb{C}^{M\times K}$ is the normalized small-scale fading channel matrix of the users. In order to build $\boldsymbol{\Upsilon}$ it is necessary to carry out $4K^2M$ additions, $4K^2M$ multiplications, and $K(K-1)/2$ absolute value operations, which requires one addition, two multiplications, and one square root operation. Note that it is only necessary to do $K(K-1)/2$ absolute value operations because $\boldsymbol{\Upsilon}$ is a symmetric hollow matrix. Another important step is the ℓ_2-norm calculations and divisions to build the matrix $\bar{\mathbf{H}}$. However, since $\|\mathbf{h}_k\|_2 = \sqrt{M}\ \forall k \in \mathcal{K}$, it is just necessary to undertake $2KM$ divisions. The total number of operations required by CBS is summarized in Table 1.

After calculating $\boldsymbol{\Upsilon}$, three linear searches are carried out to find the user to be dropped at that iteration. Therefore, at the lth iteration, the CBS first searches in $(K-l+1)(K-l)/2$ elements in order to find the pair (k,k') with the highest $r_{kk'}$. Then, it performs two searches in $(K-l-1)$-dimensional space to find which user has the second highest magnitude of the correlation coefficient. These three searches have computational complexities of $\mathcal{O}((K-l+1)(K-l)/2)$, $\mathcal{O}(K-l-1)$, and $\mathcal{O}(K-l-1)$, respectively.

There are no arithmetic calculations inside the loop of CBS, which is a big advantage compared to SOS and S-SOS. However, it requires 3 searches that depend on K, which can be a real bottleneck depending on the size of the massive MIMO system.

D. INTER-CHANNEL INTERFERENCE-BASED SELECTION

The computational burden of ICIBS is very similar to that of CBS, with one of the main differences being the additional step required by ICIBS to compute the ICI. Equation (44) can be rewritten as matrix equation given by

$$\boldsymbol{\psi}^{(l)} = \frac{1}{|\mathcal{S}_{l-1}| - 1} \boldsymbol{\Upsilon}_{\mathcal{S}_{l-1}} \mathbf{1}_{|\mathcal{S}_{l-1}| \times 1}. \tag{48}$$

This additional step requires $(K-l+1)(K-l)$ additions and $K-l+1$ divisions per iteration since $|\mathcal{S}_{l-1}| = K-l+1$. The total number of operations required by ICIBS is summarized in Table 1. Additionally, ICIBS performs one linear search in a $(K-l-1)$-dimensional space in order to find the user with the highest ICI.

Compared to CBS, ICIBS performs additional arithmetic operations inside its loop, which increases the computational burden. However, ICIBS also performs fewer searches in order to find the user to be dropped from the transmission, which is an advantage compared to CBS. Besides that, ICIBS, like CBS, is a decremental user selection approach, which, differently from SOS and S-SOS, leads to reduced computational burden when the number of selected users L is close to the actual number of users K.

E. SIMPLIFIED INTER-CHANNEL INTERFERENCE-BASED SELECTION

The complexity of the S-ICIBS is concentrated in the computation of (45), which requires $(K-l)$ additions, $(K-l)$ multiplications, and $(K-l)$ divisions. The total number of operations is summarized in Table 1. At first glance, the S-ICIBS does not seem to yield an improvement because it performs extra multiplications compared to ICIBS. However, the computational advantages of (45) become more evident as K increases.

VI. SIMULATION RESULTS

A. OVERVIEW

In this section, the proposed ICIBS algorithm is compared to SOS [23] and CBS [7] algorithms. These algorithms were combined with MRT, ZF, and the *minimum mean squared error* (MMSE) [15] precoders. Equal power allocation was used since the aim is to evaluate the impact of user selection. This power allocation scheme is considered the most basic one and, in fact, it can be considered as if no power allocation was implemented [42]. The performance is assessed via numerical simulations by analyzing the effect of the number of selected users on the system's throughput and computational complexity.[5] All the code used in this work are available on GitHub [50]. The algorithms were evaluated under three different scenarios:

- *Perfect CSI:* to evaluate the improvement afforded by the user selection algorithms when complete and perfect knowledge of the channel is available.
- *Partial CSI:* to evaluate the performance of the user selection algorithms when only partial knowledge of the channel is available.
- *Ultra Clustered-crowded environment:* to evaluate the improvement provided by the user selection algorithms when all the users are clustered in a small section of the cell. This scenario is really challenging for user selection algorithms due to the proximity of the users, i.e., the AoAs corresponding to these users are very similar, thus increasing the inter-channel interference level.

TABLE 2. Simulation parameters.

	Parameters
Number of Antennas	$M \in \{50, 100\}$
Number of Users	$K \in \{10, 25, 50, 100, 150\}$
Cell Radius	$R = 500$ m
Bandwidth	$B = 20$ MHz
Large-scale Fading	$\beta_k = -148 - 37.6 \log_{10}(d_k \times 10^{-3})$ dB
BS Power	10 W
BS Antenna Gain	0 dBi
User Antenna Gain	0 dBi
Monte-Carlo Ensemble	$5,000$
Precoding Algorithms	MRT, ZF, and MMSE
User Selection Algorithms	SOS, CBS, and ICIBS

B. SYSTEM PARAMETERS

The simulation set-up was strongly inspired by that of [7]. For the simulations, a 500-m radius hexagonal single-cell massive MIMO system with $M \in \{50, 100\}$ and $K \in \{10, 50, 75, 100, 150\}$ was used. The UR-LoS channel with a ULA, carrier frequency of 2 GHz, and bandwidth $B = 20$ MHz were used in the simulations. The large-scale coefficient was known by the BS and is defined as [51]

$$\beta_k = -148 - 37.6 \log_{10}\left(\frac{d_k}{1 \text{ km}}\right) \text{ [dB]}, \tag{50}$$

where $d_k \in \mathbb{R}_+$ is the distance between the kth user and the BS. Further, we considered the worst possible scenario in the cell, where all users were at the cell edge, yielding $\beta = -137$ dB for each user. Although this assumption seems restrictive at first glance, it makes it possible to highlight the impact of small-scale fading on the system performance and how it can be compensated through user selection. The radiated power at the BS was 10 W, the BS and user antenna gain was 0 dBi, and the noise figure for the users was 9 dB. Hence, the DL SNR was 132 dB, yielding an effective SNR $\bar{\rho}$ of -5 dB. Moreover, the throughput was calculated using $2,000$ realizations of the UR-LoS fading channel. The simulations parameters are summarized in Table 2.

C. PERFECT CSI

In this subsection, the throughput of massive MIMO systems was analyzed considering perfect CSI at the BS. This setup

5. The throughput is given by

$$\mu = BR_{\text{sum}} \text{ [bps]}, \tag{49}$$

where $B \in \mathbb{R}_+$ is the bandwidth in Hz.

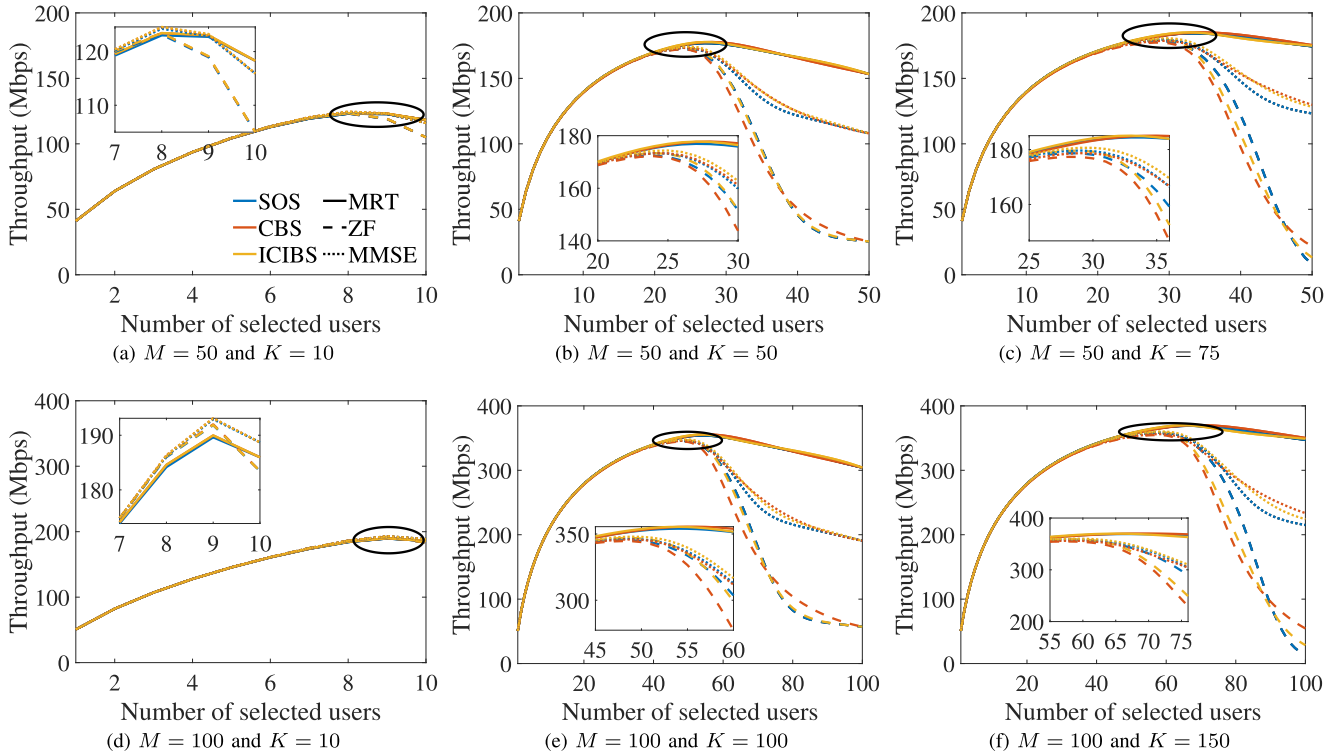


FIGURE 4. Average throughput *versus* the number of selected users L for $M \in \{50,100\}$, considering different number of users K. The line style (solid, dashed or dotted line) determines the precoder, whereas the colors specify the user selection algorithm. The yellow solid line, e.g., represents the results achieved by the ICIBS scheme considering an MRT precoder.

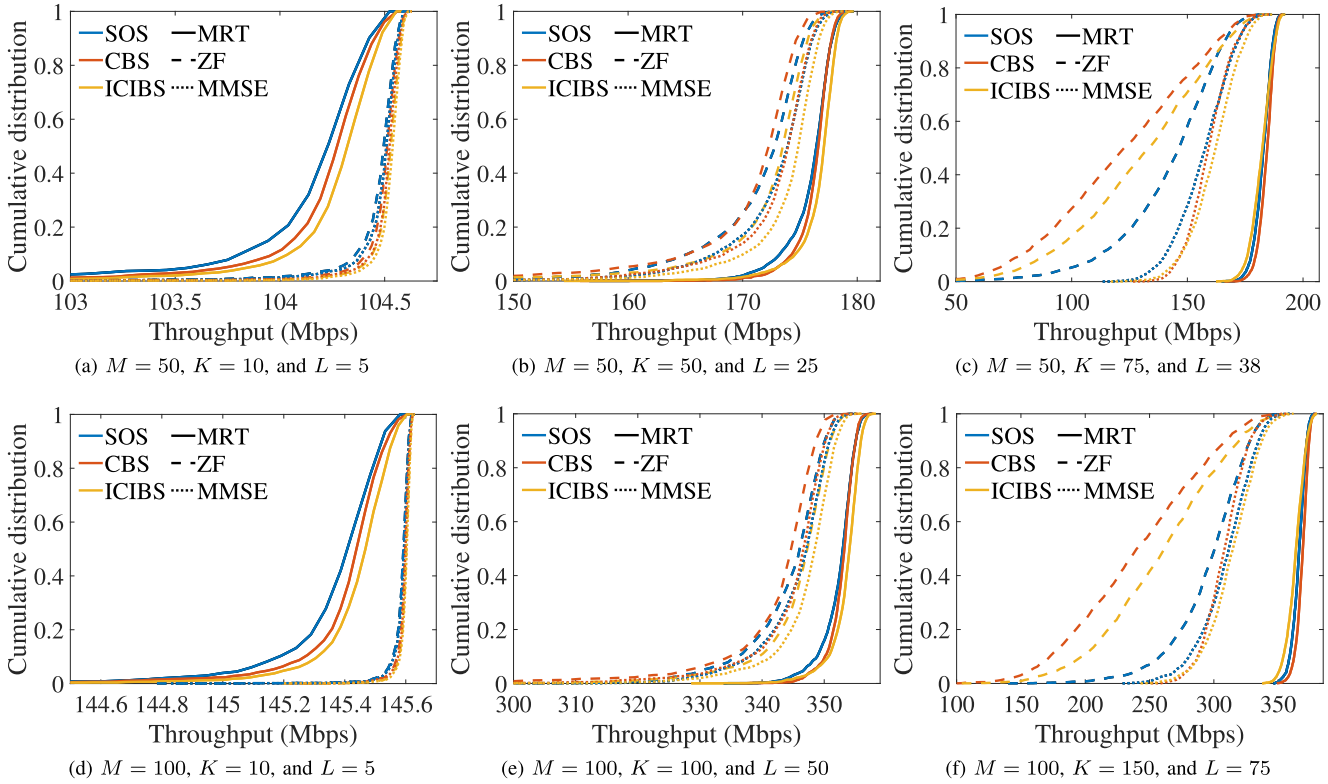
deliberately aimed to avoid the effect of errors in the precoders in order to highlight the improvement brought about by user selection to the throughput. For this case, two different scenarios were simulated: first one with $M = 50$ and $K \in \{10, 50, 75\}$, and the second one with $M = 100$ and $K \in \{10, 100, 150\}$.

Fig. 4 depicts the average throughput *versus* the number of selected users L for $M \in \{50, 100\}$ and different number of users $K \in \{10, 50, 75, 100, 150\}$, wherein $L = K$ corresponds to keeping all users (i.e., *no user selection*). The figure shows that there was a number of selected users that maximizes the throughput and this number varied depending on the particular user selection algorithm. The concavity of the curves shows that the user selection benefited the system throughput, even when $M \gg K$. In the worst case, which happened when the user selection algorithms were used with the ZF precoder, the user selection improved the throughput by at least 16.88% and 4.52% for $M = 50$ and $M = 100$, respectively, considering the maximum achieved throughput. Figs. 4(a)–4(c) illustrate user selection for $M = 50$ and $K \in \{10, 50, 75\}$. In this case, all the algorithms achieved a very close maximum performance. Indeed, in the best case, ICIBS was only 0.56% and 1.29% better than SOS and CBS, respectively. For $M = 100$, the achieved maximum throughput performances were still very similar with ICIBS being 0.47% and 1.23% better than SOS and CBS, as depicted in Figs. 4(d)–4(f). It is worth highlighting that in Fig. 4(c) and 4(f), the user selection enabled the use of ZF and MMSE precoders since they can only be used when $K \leq M$, which is already an advantage by itself.

Fig. 5 shows the *cumulative distribution function* (CDF) of the throughput for $M \in \{50, 100\}$, $L = \lceil K/2 \rceil$, and varying the number of users K. As can be observed in this figure, the performance of the user selection algorithms varied with K and the precoder used. For $M = 50$, the ICIBS algorithm outperformed the others for all values of K, except for $K = 75$, as shown in Figs. 5(a)–5(c). For $K = 75$, the SOS provided the best performance with ZF, ICIBS with MMSE, and CBS performing slightly better with MRT. Moreover, with this number of selected users, the precoding algorithms had a significant impact on the performance; the ZF and MMSE performing better with small K and MRT yielding a better performance for large K. In Figs. 5(d)–5(f), for $M = 100$, the same pattern of curves was observed as in Figs. 5(a)–5(c), with the only difference being the achieved throughput, which was higher due to the higher multiplexing gain provided by the larger array.

D. COMPUTATIONAL COMPLEXITY

This subsection analyzes the computational burden behavior of the user selection algorithms with the number of selected users L and the number of antennas M. For this analysis, we used a scenario with $K = 100$ and $M \in \{50, 100\}$. The computational burden of the user selection algorithms was compared in terms of the flops count. We consider that a



(a) $M = 50$, $K = 10$, and $L = 5$ (b) $M = 50$, $K = 50$, and $L = 25$ (c) $M = 50$, $K = 75$, and $L = 38$

(d) $M = 100$, $K = 10$, and $L = 5$ (e) $M = 100$, $K = 100$, and $L = 50$ (f) $M = 100$, $K = 150$, and $L = 75$

FIGURE 5. ***Cumulative distribution function*** **of the throughput for $L = \lceil K/2 \rceil$ and $M \in \{50,100\}$, considering different number of users K. The line style (solid, dashed or dotted line) determines the precoder, whereas the colors specify the user selection algorithm. The yellow solid line, e.g., represents the results achieved by the ICIBS scheme considering an MRT precoder.**

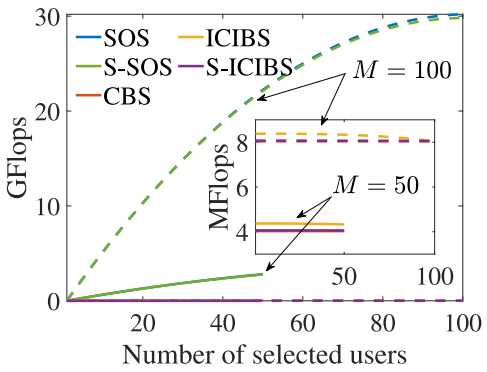


FIGURE 6. **Giga flops count *versus* the number of selected users for $K = 100$ and $M \in \{50,100\}$.**

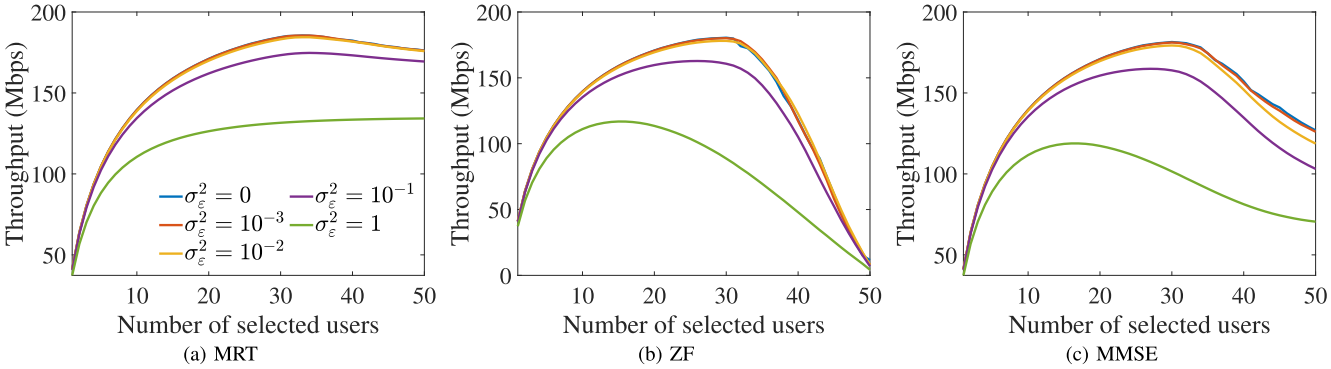
division and a square root operation are equivalent to one flop, similar to that of an addition and a multiplication.

Fig. 6 depicts the flops count *versus* the number of selected users L. As can be observed in this figure, the computational burden of SOS and S-SOS grows rapidly with L due to the incremental nature of this type of algorithm. Additionally, the simplified version does present a significant decrease of the computational burden in this case. Moreover, the computational burden of SOS and S-SOS is significantly higher than that of CBS, ICIBS, and S-ICIBS, requiring at least 10^9 more operations. Furthermore, S-ICIBS reduces the computational burden compared to ICIBS, achieving a similar one to CBS without requiring expensive linear searches like CBS. In conclusion, SOS and S-SOS can be prohibitive for massive MIMO systems, depending on the system parameters.

E. PARTIAL CSI

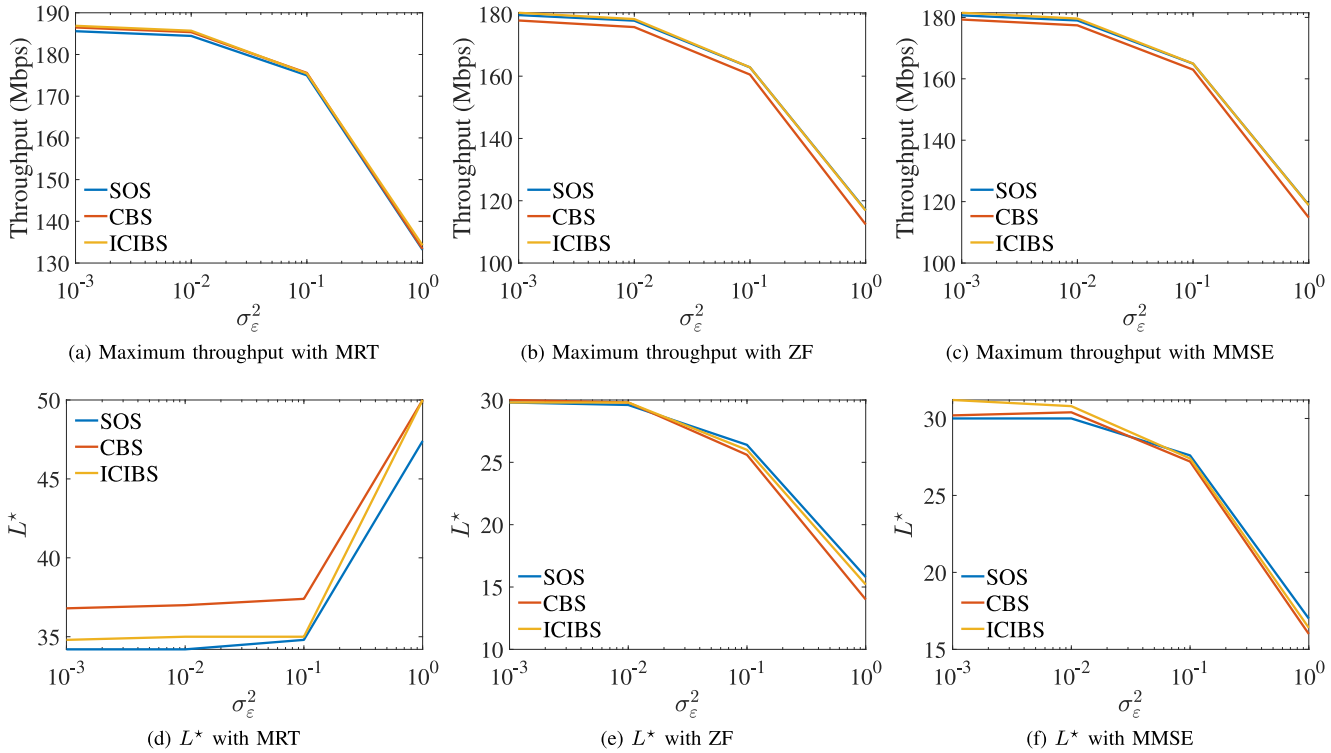
This subsection analyzes the robustness of the user selection algorithms to a possible error, which leads to the partial CSI scenario. This scenario is closer to what happens in practical massive MIMO systems, wherein channel estimation errors stem from several sources, such as inherent inaccuracies of the channel estimation algorithm, and/or from the pilot contamination from different cells in a multi-cell scenario, as explained in Section II-D. For this analysis, we considered the channel estimate as in (19) with $\sigma_\varepsilon^2 \in \{0, 10^{-3}, 10^{-2}, 10^{-1}, 10^{0}\}$.

Fig. 7 shows the average throughput *versus* the number of selected users L with partial CSI knowledge for the ICIBS algorithm, considering $M = 50$, $K = 75$, different precoding algorithms, and different σ_ε^2. For the sake of simplicity, we only show the results for the ICIBS algorithms since the other user selection algorithms have similar performances. As it can be observed from this figure, the performance of ICIBS has the same pattern as in Fig. 4 with variations in the throughput due to the errors in the channel estimation. For small σ_ε^2, the performance of ICIBS is close to the case when perfect CSI is available for all precoders. For large σ_ε^2, the ICIBS has a significant performance gap



(a) MRT (b) ZF (c) MMSE

FIGURE 7. Average throughput *versus* the number of selected users L with partial CSI knowledge for the ICIBS algorithm, considering $M = 50$, $K = 75$, different precoding algorithms, and different σ_ε^2.



(a) Maximum throughput with MRT (b) Maximum throughput with ZF (c) Maximum throughput with MMSE

(d) $L^\star$ with MRT (e) $L^\star$ with ZF (f) $L^\star$ with MMSE

FIGURE 8. Maximum achieved throughput and optimum number of selected users $L^\star$ *versus* the variance of the channel estimation error σ_ε^2 for $M \in \{50,100\}$, considering different user selection and precoding algorithms.

when compared to the case with perfect CSI knowledge. Furthermore, it is worth highlighting that, in the case with $\sigma_\varepsilon^2 = 1$, the throughput as a function of the number of selected users when the MRT precoder is used is monotonically increasing, unlike in the other cases. This phenomenon happens due to the fact that the channel estimate for this case with high channel estimation error is more like an i.i.d. Rayleigh than an LoS channel, which explains the monotonic behavior.

Fig. 8 depicts the performance comparison between ICIBS and the other user selection algorithms in terms of the maximum achieved throughput and the optimum number of selected user $L^\star$, i.e., the number of selected users that leads to the maximum achieved throughput. As it can be observed in Figs. 8(a)–8(c), the maximum achieved throughput is very similar for all of the user selection algorithms considering all precoding algorithms. On the other hand, the optimum number of selected users varies for different user selection algorithms operating under different levels of channel estimation error as can be observed in Figs. 8(d)–8(f). It is desirable to serve as many users as possible. The major difference in the amount of users served by the algorithms happens when the MRT precoder is used in a scenario with $\sigma_\varepsilon^2 = 10^{-1}$. In this case, the CBS is able to serve two more

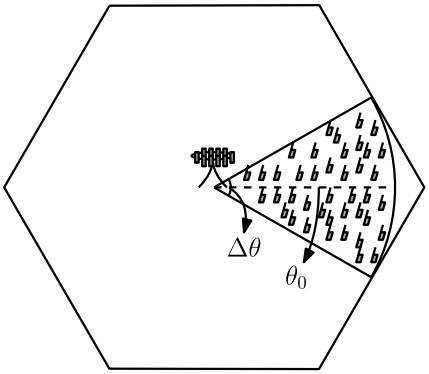


FIGURE 9. Example of an ultra clustered-crowded scenario used in the simulations, where θ_0 is the direction of the sector of the cell and $\Delta\theta$ is the aperture angle.

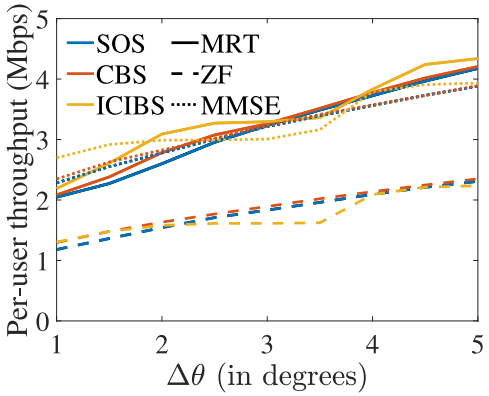
users than SOS and ICIBS. In general, ICIBS and CBS serve more users than SOS in scenarios with low levels of channel estimation errors ($\sigma_\varepsilon^2 \leq 10^{-2}$), whereas for high levels of channel estimation error, the optimum number of selected users depends on the precoder besides the user selection algorithm.

F. ULTRA CLUSTERED-CROWDED SCENARIO

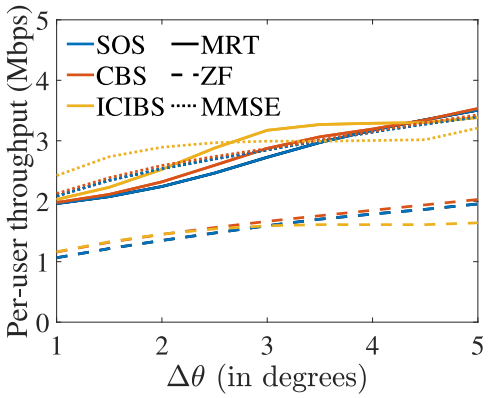
In this section, we analyze the performance of the user selection algorithms in an *ultra clustered-crowded scenario*, meaning that $K > M$ and all the users are clustered in the same section of the cell, as depicted in Fig. 9. For the simulations in this case, we considered $M = 50$, $K = 75$, $\theta_0 \in \{0°, 45°\}$, and $\Delta\theta \in \{1°, 1.5°, \ldots, 4.5°, 5°\}$.

Fig. 10 illustrates the average per-user throughput *versus* the size of the sector where the users are distributed. As can be observed in Fig. 10(a), ICIBS performs better than SOS and CBS for some values of $\Delta\theta$, yielding slightly increased throughput. This improvement is more evident for MRT and MMSE precoders since the performance of all user selection algorithms can be considered equal for the ZF precoder. It is worth highlighting the drop in the ICIBS performance around $\Delta\theta = 3.6°$, which seems consistent for all precoding algorithms. This drop in the performance can be explained by the fact that for a ULA with $M = 50$ antennas, the directivity of the array is π/M, thus for $\Delta\theta$ around π/M, it is already possible to have small correlation between some users, which may be a problem for ICIBS since it takes advantage from situations where the correlation among the users is very high and its global usage of the interference among the users leverages the selection. In Fig. 10(b), we can observe the same pattern of Fig. 10(a), with the performance drop around $\Delta\theta = 5°$, which may be due to the fact that the array is steered to 45°.

Fig. 11 depicts the average throughput *versus* the number of selected users L for $\Delta\theta = 1.5°$. Like in the case presented in Section VI-C, the user selection also improved the system throughput, achieving maximum throughput for all algorithms when $L^\star = 2$ users are served with both ZF and MMSE, and $L^\star = 1$ user is served with MRT. In this



(a) $\theta_0 = 0°$



(b) $\theta_0 = 45°$

FIGURE 10. Average per-user throughput *versus* $\Delta\theta$ for $M = 50$ and $K = 75$. The line style (solid, dashed or dotted line) determines the precoder, whereas the colors specify the user selection algorithm. The yellow solid line, e.g., represents the results achieved by the ICIBS scheme considering an MRT precoder.

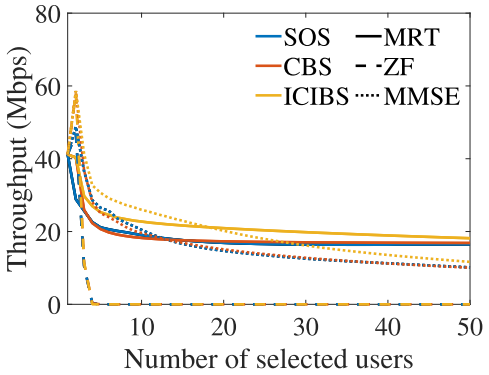


FIGURE 11. Average throughput *versus* the number of selected users L for $M = 50$, $K = 75$, $\Delta\theta = 1.5°$, and $\theta_0 = 0°$. The line style (solid, dashed or dotted line) determines the precoder, whereas the colors specify the user selection algorithm. The yellow solid line, e.g., represents the results achieved by the ICIBS scheme considering an MRT precoder.

case, both CBS and ICIBS achieved the same maximum performance, and, at their peaks, they were 19.98% and 23.03% better than SOS for ZF and MMSE, respectively. Moreover, the performance of SOS and CBS decreased faster with L than that of ICIBS, which is an advantage

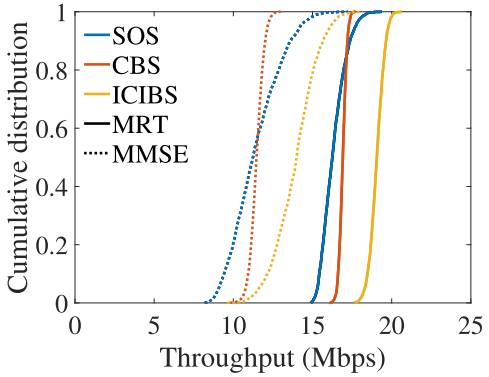


FIGURE 12. *Cumulative distribution function* of the throughput for $M = 50$, $K = 75$, $L = 38$, $\Delta\theta = 1.5°$, and $\theta_0 = 0°$. The line style (solid, dashed or dotted line) determines the precoder, whereas the colors specify the user selection algorithm. The yellow solid line, e.g., represents the results achieved by the ICIBS scheme considering an MRT precoder.

if it is necessary to work with $L > L^{\star}$. For example, for $L = 38$, ICIBS achieved a throughput of 19.09 Mbps and 14.02 Mbps with MRT and MMSE, which were still 10.94% and 21.91% better than SOS and CBS with the same precoders. The performance of the user selection algorithms with ZF precoder was impaired in this scenario due to the proximity of the users, which lead to an ill-conditioned channel matrix, and even the user selection could not help in this case.

Fig. 12 shows the CDF of the throughput for $\Delta\theta = 1.5°$ and $L = 38$. The benefit that ICIBS provided in this type of scenario is clearly evident from this figure. For both MRT and MMSE precoders, ICIBS yielded a significant gain in the *95%-probability throughput* (95PT) over the other user selection algorithms. ICIBS achieved a 95PT of 18.27 Mbps and 11.57 Mbps with MRT and MMSE precoders, resulting in improved performance of at least 7.66% over the other user selection algorithms. Although the ICIBS did not yield significant improvements in terms of system throughput in the results presented in Section VI-C, the results due to ICIBS presented in this section were significantly superior in comparison to the competing user selection algorithms. The fundamental difference between the scenario simulated in this section from the scenario in Section VI-C lies in the inter-channel interference level. Indeed, in this section, since the users are concentrated in a small portion of the cell, there are more users interfering with each other simultaneously, thus the inter-channel interference is more severe. This is why the ICIBS outperformed other competing approaches.

G. PRACTICAL GUIDELINES

This subsection discusses the suitability of the user selection algorithms in practical scenarios, summarizing the main findings of massive MIMO systems under LoS propagation, and explaining how to use the obtained results for aiding the design of massive MIMO systems. Firstly, it is worth highlighting that the user selection would not imply an additional cost to the BS since all the user selection algorithms are implemented digitally and can share the BS's *digital signal processor* (DSP) units. Therefore, user selection algorithms can be easily used in massive MIMO systems.

As it can be observed from the results presented throughout this section, the performance of all of the user selection algorithms are very sensitive to the parameters of the system. Thus, the most suitable algorithm depends on the application and the scenario. For example, if the aim is to maximize the achieved throughput without concerning with the number of served users, the most suitable algorithm is the ICIBS with ZF and MMSE precoders, followed by CBS with MRT precoder. However, if the system has a requirement to serve specific minimum number of users L, then the most suitable algorithm is the one that yields the highest throughput for that particular L. In general, for the case where $K \leq M$ and $L \leq \lceil K/2 \rceil$, the ICIBS outperforms the other user selection algorithms by a small amount for all precoder types. On the other hand, when $K \leq M$ and $L = \lceil K/2 \rceil$, CBS yields the highest throughput using the MRT precoder, SOS yields the highest throughput using the ZF precoder, and ICIBS yields the highest throughput using the MMSE precoder. Moreover, SOS generally tends to perform better than CBS and ICIBS when ZF is used for $K/2 < L \leq M$. However, the performance with ZF precoder is so degraded in LoS propagation for that range of L such that it is advisable to use another precoder in this case.

Another important factor is the computational complexity, which can be a hindrance in practical applications. The computational burden of the SOS rapidly grows with the number of antennas M and the number of selected users L, which is not desirable since massive MIMO systems use very large M. The SOS should be considered only in cases where one can have a significant gain in the throughput. Therefore, the high computational burden of the SOS is an impediment for its use in massive MIMO systems.

The final point to consider in the design of massive MIMO systems with user selection is the cell usage. In the simulations, we considered two possible scenarios where the whole cell and only a section of the cell was used. The previous comments are based on the case when the whole cell was used. However, when only a small section of the cell is used, the ICIBS outperforms all of the other algorithms. Further, the ultra clustered-crowded scenario used in the performance analysis represents some real scenarios that are increasingly common nowadays, such as large scale social, cultural, and sporting events, where a dedicated and efficient communication system is necessary.

VII. CONCLUDING REMARKS

In this paper a comprehensive review of the user selection algorithms for massive MIMO systems has been provided, and their performance when combined with different linear precoders for various system configurations considering perfect and partial CSI under the LoS propagation channel has been evaluated. In addition, a thorough analysis of the computational complexity of the user selection algorithms is also

provided. Findings highlight the fact that unlike the well-studied case of i.i.d. Rayleigh channel, for LoS channels both the favorable propagation and asymptotically favorable propagation conditions can be violated irrespective of the number of users and antennas in the massive MIMO system, degrading the system performance. Further, it is shown that practical cases exist in which the LoS propagation model may lead to significant levels of interference among users within a cell and such cases are not satisfactorily addressed by the existing user selection algorithms. To this end, a new user selection algorithm based on ICI, called ICIBS, is proposed and its performance evaluated. Unlike other techniques, the ICIBS accounts for the ICI in a global manner, thus yielding better results than the other algorithms especially in cases where there are many users interfering with one another and similar results in scenarios having low-interference levels. Although the paper has focused on LoS propagation, all algorithms can also be applied to different channel models, such as the Rician fading model. From our experience, however, the benefits in SE due to the use of user selection algorithms are more prominent as the Rician fading model tends to the LoS model, whereas they become negligible for i.i.d. Rayleigh fading.

APPENDIX A
PROOF OF COROLLARY 1

An LoS environment offers FP when (10) is equal to zero for all $k, k' \in \mathcal{K}$, with $k \neq k'$. First, we show the condition that guarantees orthogonality between two users. Then, we expand this result to guarantee orthogonality for all users. The kth user and the k'th user are orthogonal when (10) is equal to zero, iff

$$\sin\Big(M\frac{\pi}{2}(\sin\theta_{k'} - \sin\theta_k)\Big) = 0, \tag{51}$$

and

$$\sin\Big(\frac{\pi}{2}(\sin\theta_{k'} - \sin\theta_k)\Big) \neq 0. \tag{52}$$

Equation (51) is zero if there exists an integer $n_{k'}$ such that

$$\sin\theta_{k'} - \sin\theta_k = \frac{2n_{k'}}{M}, \tag{53}$$

where $n_{k'} \in \{0, \pm 1, \ldots, \pm M\}$, as the left-hand side of the previous equation lies in the interval $[-2, 2]$. Solving this equation for $\theta_{k'}$, yields

$$\theta_{k'} = \arcsin\left(\sin\theta_k + \frac{2n_{k'}}{M}\right), \tag{54}$$

for some $n_{k'} \in \{0, \pm 1, \ldots, \pm M\}$.

Moreover, to guarantee (52), we must satisfy

$$\sin\theta_{k'} - \sin\theta_k \notin \{-2, 0, 2\}, \tag{55}$$

which results in $\theta_{k'} \neq \theta_k$ and $\theta_{k'} \neq \pi - \theta_k$ for the case $\sin\theta_{k'} - \sin\theta_k \neq 0$, and if $\theta_k = \pm\pi/2$, then $\theta_{k'} \neq \mp\pi/2$ for the cases $\sin\theta_{k'} - \sin\theta_k \neq \pm 2$, respectively. To satisfy (53), (55) and (56) simultaneously. Therefore,

$$r_{kk'} = 0 \Longleftrightarrow \theta_{k'} = \arcsin\left(\sin\theta_k + \frac{2n_{k'}}{M}\right), \tag{56}$$

for some $n_{k'} \in \{\pm 1, \ldots, \pm(M-1)\}$.

Given that (54) does not depend on a specific pair of users, to have $r_{kk'} = 0$ for all $k, k' \in \mathcal{K}$, with $k \neq k'$, the users need to satisfy (56) for all $k, k' \in \mathcal{K}$, with $k \neq k'$. ■

APPENDIX B
PROOF OF THEOREM 2

The signal received by the kth user y_k in (1) can be expanded in

$$y_k = \sqrt{\rho\beta_k\eta_k}\hat{\mathbf{h}}_k^{\mathrm{T}}\mathbf{w}_k s_k + \sum_{k'\in\mathcal{K}\setminus\{k\}} \sqrt{\rho\beta_k\eta_{k'}}\hat{\mathbf{h}}_k^{\mathrm{T}}\mathbf{w}_{k'}s_{k'} - \sum_{k'\in\mathcal{K}} \sqrt{\rho\beta_k\eta_{k'}}\boldsymbol{\epsilon}_k^{\mathrm{T}}\mathbf{w}_{k'}s_{k'} + n_k, \tag{57}$$

where $\mathbf{h}_k$ and $\mathbf{w}_k$ are the kth columns of matrices $\mathbf{H}$ and $\mathbf{W}$, respectively. The first term in the RHS of (57) corresponds to the signal of interest, the second one to the interference caused from the other users, the third one to the interference caused by the uncertainty in the channel estimation, and the last one represents the additive noise.

Equation (57) is a point-to-point channel with deterministic channel $\hat{\mathbf{h}}_k^{\mathrm{T}}\mathbf{w}_k$ and additive non-Gaussian noise. Therefore, assuming that the kth user knows the equivalent channel $\mathbf{h}_k^{\mathrm{T}}\mathbf{w}_k$ for decoding the received signal, the DL SE is given by [42], [52]

$$R_{\mathrm{sum}} \geq \frac{1}{2}\left(1 - \frac{\tau_{\mathrm{p}}}{\tau_{\mathrm{c}}}\right)\sum_{k\in\mathcal{K}}\log_2(1+\gamma_k), \tag{58}$$

where $\tau_{\mathrm{p}} \in \mathbb{N}$ is the pilot time in samples, $\tau_{\mathrm{c}} \in \mathbb{N}$ is the coherence time in samples, and $\gamma_k \in \mathbb{R}_+$ is the DL SINR related to the kth user. In this case, due to the independence of the random variables, the SINR for the kth user is given by

$$\gamma_k = \frac{p_{\mathrm{s}}}{p_{\mathrm{n}} + p_{\mathrm{i}} + p_{\mathrm{u}}}, \tag{59}$$

where

$$p_{\mathrm{s}} = \mathsf{Var}\left\{\sqrt{\rho\beta_k\eta_k}\hat{\mathbf{h}}_k^{\mathrm{T}}\mathbf{w}_k s_k\right\}, \tag{60}$$

$$p_{\mathrm{n}} = \mathsf{Var}\{n_k\} = 1, \tag{61}$$

$$p_{\mathrm{i}} = \mathsf{Var}\left\{\sum_{k'\in\mathcal{K}\setminus\{k\}} \sqrt{\rho\beta_k\eta_{k'}}\hat{\mathbf{h}}_k^{\mathrm{T}}\mathbf{w}_{k'}s_{k'}\right\}, \tag{62}$$

$$p_{\mathrm{u}} = \mathsf{Var}\left\{\sum_{k'\in\mathcal{K}} \sqrt{\rho\beta_k\eta_{k'}}\boldsymbol{\epsilon}_k^{\mathrm{T}}\mathbf{w}_{k'}s_{k'}\right\}. \tag{63}$$

Then, yielding (21).

When $\sigma_\varepsilon^2 = 0$, i.e., the BS has perfect CSI knowledge, (57) is reduced to

$$y_k = \sqrt{\rho\beta_k\eta_k}\hat{\mathbf{h}}_k^{\mathrm{T}}\mathbf{w}_k s_k + \sum_{k'\in\mathcal{K}\setminus\{k\}} \sqrt{\rho\beta_k\eta_{k'}}\hat{\mathbf{h}}_k^{\mathrm{T}}\mathbf{w}_{k'}s_{k'} + n_k, \quad (64)$$

which corresponds to a point-to-point channel with deterministic channel $\hat{\mathbf{h}}_k^{\mathrm{T}}\mathbf{w}_k$ and additive Gaussian noise. Therefore, assuming that the kth user knows the equivalent channel $\mathbf{h}_k^{\mathrm{T}}\mathbf{w}_k$ for decoding the received signal, the DL SE is given by [42], [52]

$$R_{\mathrm{sum}} = \frac{1}{2}\left(1 - \frac{\tau_{\mathrm{p}}}{\tau_{\mathrm{c}}}\right)\sum_{k\in\mathcal{K}} \log_2(1+\gamma_k), \quad (65)$$

where γ_k is given by (21) with $\sigma_\varepsilon^2 = 0$. ■

APPENDIX C
PROOF OF COROLLARY 2

Using (26) in (21), we have

$$\gamma_k^{\text{ZF-1}} = \frac{\rho\beta_k\eta_k\left|\frac{\hat{\mathbf{h}}_k^{\mathrm{T}}\hat{\mathbf{H}}^*\mathbf{r}_k}{\left\|\hat{\mathbf{H}}^*\mathbf{r}_k\right\|_2}\right|^2}{1+\rho\beta_k\sum_{k'\in\mathcal{K}\setminus\{k\}}\eta_{k'}\left|\frac{\hat{\mathbf{h}}_k^{\mathrm{T}}\hat{\mathbf{H}}^*\mathbf{r}_{k'}}{\left\|\hat{\mathbf{H}}^*\mathbf{r}_{k'}\right\|_2}\right|^2 + \rho\beta_k\sigma_\varepsilon^2\sum_{k'\in\mathcal{K}}\eta_{k'}}. \quad (66)$$

The vector $\mathbf{r}_k$ can be written as

$$\mathbf{r}_k = \left(\hat{\mathbf{H}}^{\mathrm{T}}\hat{\mathbf{H}}^*\right)^{-1}\mathbf{e}_k, \quad (67)$$

where $\mathbf{e}_k$ is the kth column of $\mathbf{I}_K$. Using (67) in (66), $\gamma_k^{\text{ZF-1}}$ can be rewritten as

$$\begin{aligned}\gamma_k^{\text{ZF-1}} &= \rho\beta_k\eta_k\frac{|\mathbf{e}_k^{\mathrm{T}}\mathbf{e}_k|^2}{\left\|\hat{\mathbf{H}}^*\mathbf{r}_k\right\|_2^2} \\ &\quad\times\left(1+\rho\beta_k\sum_{k'\in\mathcal{K}\setminus\{k\}}\eta_{k'}\frac{|\mathbf{e}_k^{\mathrm{T}}\mathbf{e}_{k'}|^2}{\left\|\hat{\mathbf{H}}^*\mathbf{r}_{k'}\right\|_2^2} + \rho\beta_k\sigma_\varepsilon^2\sum_{k'\in\mathcal{K}}\eta_{k'}\right)^{-1} \\ &= \frac{\rho\beta_k\eta_k}{\left\|\hat{\mathbf{H}}^*\mathbf{r}_k\right\|_2^2\left(1+\rho\beta_k\sigma_\varepsilon^2\sum_{k'\in\mathcal{K}}\eta_{k'}\right)} \\ &= \frac{\rho\beta_k\eta_k}{\left[\left(\hat{\mathbf{H}}^{\mathrm{T}}\hat{\mathbf{H}}^*\right)^{-1}\right]_{kk}\left(1+\rho\beta_k\sigma_\varepsilon^2\sum_{k'\in\mathcal{K}}\eta_{k'}\right)}. \end{aligned} \quad (68)$$

■

REFERENCES

[1] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[2] F. Rusek *et al.*, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[3] B. M. Hochwald, T. L. Marzetta, and V. Tarokh, "Multiple-antenna channel hardening and its implications for rate feedback and scheduling," *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 1893–1909, Sep. 2004.

[4] E. Björnson, E. G. Larsson, and T. L. Marzetta, "Massive MIMO: Ten myths and one critical question," *IEEE Commun. Mag.*, vol. 54, no. 2, pp. 114–123, Feb. 2016.

[5] T. V. Chien and E. Björnson, "Massive MIMO communications," in *5G Mobile Communications*. Cham, Switzerland: Springer, 2017, pp. 77–116.

[6] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Aspects of Favorable propagation in massive MIMO," in *Proc. 22nd Eur. Signal Process. Conf.*, Lisbon, Portugal, Sep. 2014, pp. 76–80.

[7] H. Yang and T. L. Marzetta, "Massive MIMO with max-min power control in line-of-sight propagation environment," *IEEE Trans. Commun.*, vol. 65, no. 11, pp. 4685–4693, Nov. 2017.

[8] H. Yang, "User scheduling in massive MIMO," in *Proc. IEEE 19th Int. Workshop Signal Process. Adv. Wireless Commun.*, Kalamata, Greece, Jun. 2018, pp. 1–5.

[9] H. Yang, H. Q. Ngo, and E. G. Larsson, "Multi-cell massive MIMO in LoS," in *Proc. IEEE Global Commun. Conf.*, Dec. 2018, pp. 1–6.

[10] R. S. Chaves, E. Cetin, M. V. S. Lima, and W. A. Martins, "On the convergence of max-min fairness power allocation in massive MIMO systems," *IEEE Commun. Lett.*, vol. 24, no. 12, pp. 2873–2877, Dec. 2020.

[11] W. Roh *et al.*, "Millimeter-wave beamforming as an enabling technology for 5G cellular communications: Theoretical feasibility and prototype results," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 106–113, Feb. 2014.

[12] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 831–846, Oct. 2014.

[13] Z. Zhou, J. Fang, L. Yang, H. Li, Z. Chen, and S. Li, "Channel estimation for millimeter-wave multiuser MIMO systems via PARAFAC decomposition," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7501–7516, Nov. 2016.

[14] E. G. Larsson, T. L. Marzetta, H. Q. Ngo, and H. Yang, "Antenna count for massive MIMO: 1.9 GHz vs. 60 GHz," *IEEE Commun. Mag.*, vol. 56, no. 9, pp. 132–137, Sep. 2018.

[15] D. Tse and P. Viswanath, *Capacity of Wireless Channels*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[16] X. Gao, O. Edfors, F. Rusek, and F. Tufvesson, "Massive MIMO performance evaluation based on measured propagation data," *IEEE Trans. Wireless Commun.*, vol. 14, no. 7, pp. 3899–3911, Jul. 2015.

[17] A. Osseiran *et al.*, "Scenarios for 5G mobile and wireless communications: The vision of the METIS project," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 26–35, May 2014.

[18] F. Challita, M.-T. Martinez-Ingles, M. Lienard, J.-M. Molina-Garcia-Pardo, and D. P. Gaillot, "Line-of-sight massive MIMO channel characteristics in an indoor scenario at 94 GHz," *IEEE Access*, vol. 6, pp. 62361–62370, 2018.

[19] J. Flordelis, F. Rusek, X. Gao, G. Dahman, O. Edfors, and F. Tufvesson, "Spatial separation of closely-located users in measured massive MIMO channels," *IEEE Access*, vol. 6, pp. 40253–40266, 2018.

[20] A. Ijaz *et al.*, "Enabling massive IoT in 5G and beyond systems: PHY radio frame design considerations," *IEEE Access*, vol. 4, pp. 3322–3339, Jun. 2016.

[21] G. Dimić and N. D. Sidiropoulos, "On downlink beamforming with greedy user selection: Performance analysis and a simple new algorithm," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3857–3868, Oct. 2005.

[22] J. Jiang, R. M. Buehrer, and W. H. Tranter, "Greedy scheduling performance for a zero-forcing dirty-paper coded system," *IEEE Trans. Commun.*, vol. 54, no. 5, pp. 789–793, May 2006.

[23] T. Yoo and A. Goldsmith, "On the optimality of Multiantenna broadcast scheduling using zero-forcing beamforming," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 528–541, Mar. 2006.

[24] M. Costa, "Writing on dirty paper," *IEEE Trans. Inf. Theory*, vol. 29, no. 3, pp. 439–441, May 1983.

[25] H. Weingarten, Y. Steinberg, and S. S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3936–3964, Sep. 2006.

[26] M. Benmimoune, E. Driouch, W. Ajib, and D. Massicotte, "Joint transmit antenna selection and user scheduling for massive MIMO systems," in *Proc. IEEE Wireless Commun. Netw. Conf.*, New Orleans, LA, USA, Mar. 2015, pp. 381–386.
[27] S. Dierks and N. Juenger, "Scheduling for massive MIMO with few excess antennas," in *Proc. 20th Int. ITG Workshop Smart Antennas*, Munich, Germany, Mar. 2016, pp. 40–44.
[28] M. Kuerbis, N. M. Balasubramanya, L. Lampe, and A. Lampe, "User scheduling in massive MIMO systems with a large number of devices," in *Proc. IEEE 28th Annu. Int. Symp. Pers. Indoor Mobile Radio Commun.*, Montreal, QC, Canada, Oct. 2017, pp. 1–6.
[29] H. Liu, H. Gao, S. Yang, and T. Lv, "Low-complexity downlink user selection for massive MIMO systems," *IEEE Syst. J.*, vol. 11, no. 2, pp. 1072–1083, Jun. 2017.
[30] T. A. Sheikh, J. Bora, and M. A. Hussain, "Sum-rate performance of massive MIMO systems in highly scattering channel with semi-orthogonal and random user selection," *Radioelectron. Commun. Syst.*, vol. 61, no. 12, pp. 547–555, Dec. 2018.
[31] S. Huang, H. Yin, H. Li, and V. C. M. Leung, "Decremental user selection for large-scale multi-user MIMO downlink with zero-forcing Beamforming," *IEEE Wireless Commun. Lett.*, vol. 1, no. 5, pp. 480–483, Oct. 2012.
[32] A. Farsaei, A. Alvarado, F. M. J. Willems, and U. Gustavsson, "An improved dropping algorithm for line-of-sight massive MIMO with max-min power control," *IEEE Commun. Lett.*, vol. 23, no. 6, pp. 1109–1112, Jun. 2019.
[33] A. Farsaei, A. Alvarado, F. M. J. Willems, and U. Gustavsson, "An improved dropping algorithm for line-of-sight massive MIMO with Tomlinson–Harashima Precoding," *IEEE Commun. Lett.*, vol. 23, no. 11, pp. 2099–2103, Nov. 2019.
[34] J. Nam, A. Adhikary, J.-Y. Ahn, and G. Caire, "Joint spatial division and multiplexing: Opportunistic beamforming, user grouping and simplified downlink scheduling," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 876–890, Oct. 2014.
[35] G. Lee and Y. Sung, "A new approach to user scheduling in massive multi-user MIMO broadcast channels," *IEEE Trans. Commun.*, vol. 66, no. 4, pp. 1481–1495, Apr. 2018.
[36] A. Farsaei, A. Sheikh, U. Gustavsson, A. Alvarado, and F. M. J. Willems, "DropNet: An improved dropping algorithm based on neural networks for line-of-sight massive MIMO," *IEEE Access*, vol. 9, pp. 29441–29448, 2021.
[37] R. S. Chaves, E. Cetin, M. V. S. Lima, and W. A. Martins, "User selection based on inter-channel interference for massive MIMO under line-of-sight propagation," in *Proc. XXXIVth General Assem. Scientific Symp. Int. Union Radio Sci.*, Rome, Aug. 2021, pp. 1–4.
[38] H. L. Van Trees, *Optimum Array Processing*, 1st ed. New York, NY, USA: Wiley, 2002.
[39] X. Gao, O. Edfors, J. Liu, and F. Tufvesson, "Antenna selection in measured massive MIMO channels using convex optimization," in *Proc. IEEE Globecom Workshops*, Atlanta, GA, USA, Dec. 2013, pp. 129–134.
[40] X. Gao, O. Edfors, F. Tufvesson, and E. G. Larsson, "Multi-switch for antenna selection in massive MIMO," in *Proc. IEEE Global Commun. Conf.*, San Diego, CA, USA, Dec. 2014, pp. 1–6.
[41] X. Gao, O. Edfors, F. Tufvesson, and E. G. Larsson, "Massive MIMO in real propagation environments: Do all antennas contribute equally?" *IEEE Trans. Commun.*, vol. 63, no. 11, pp. 3917–3928, Nov. 2015.
[42] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of Massive MIMO*, 1st ed. Cambridge, U.K.: Cambridge Univ. Press, 2016.
[43] X. Wu, N. C. Beaulieu, and D. Liu, "On favorable propagation in massive MIMO systems and different antenna configurations," *IEEE Access*, vol. 3, pp. 5578–5593, Apr. 2017.
[44] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742–758, Oct. 2014.
[45] H. Yang and T. L. Marzetta, "Max-min SINR dependence on channel correlation in line-of-sight massive MIMO," in *Proc. IEEE Global Commun. Conf.*, Singapore, Dec. 2017, pp. 1–6.
[46] X. Li, S. Jin, H. A. Suraweera, J. Hou, and X. Gao, "Statistical 3-D beamforming for large-scale MIMO downlink systems over Rician fading channels," *IEEE Trans. Commun.*, vol. 64, no. 4, pp. 1529–1543, Apr. 2016.
[47] O. Ozdogan, E. Björnson, and E. G. Larsson, "Massive MIMO with spatially correlated Rician fading channels," *IEEE Trans. Commun.*, vol. 67, no. 5, pp. 3234–3250, May 2019.
[48] J. Mao, J. Gao, Y. Liu, and G. Xie, "Simplified semi-orthogonal user selection for MU-MIMO systems with ZFBF," *IEEE Wireless Commun. Lett.*, vol. 1, no. 1, pp. 42–45, Feb. 2012.
[49] J. Liu, X. She, and L. Chen, "A low complexity capacity-greedy user selection scheme for zero-forcing Beamforming," in *Proc. IEEE 69th Veh. Technol. Conf.*, Barcelona, Spain, Apr. 2009, pp. 1–5.
[50] R. S. Chaves, "User Selection for Massive MIMO Under LoS Propagation." 2022. [Online]. Available: https://github.com/rafaelschaves/user-selection-for-massive-mimo-under-los-propagation
[51] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO networks: Spectral, energy, and hardware efficiency," *Found. Trends® Signal Process.*, vol. 11, nos. 3–4, pp. 154–655, 2017.
[52] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Hoboken, NJ, USA: Wiley, 2006.

RAFAEL S. CHAVES (Member, IEEE) was born in Rio de Janeiro, Brazil, in 1992. He received the Electronics and Computing Engineering degree from the Polytechnic School, Federal University of Rio de Janeiro (UFRJ) in 2016 when he was awarded as the best undergraduate student of that year, and the received the M.Sc degree in electrical engineering from COPPE/UFRJ in 2018. He is currently pursuing the cotutelle Ph.D. degree with COPPE/UFRJ and School of Engineering, Macquarie University, Sydney. His research interests are in the fields of digital signal processing, adaptive signal processing, tensor signal processing, wireless communications, massive MIMO systems, underwater acoustic communications, and machine learning. He is a member of the Brazilian Telecommunications Society (SBrT).

MARKUS V. S. LIMA (Member, IEEE) received the B.Sc. degree in electronics engineering from the Federal University of Rio de Janeiro (UFRJ) in 2008, and the M.Sc. and D.Sc. degrees in electrical engineering from COPPE/UFRJ in 2009 and 2013, respectively. He is currently an Associate Professor of the Program of Electrical Engineering (PEE/COPPE) and the Department of Electrical Engineering (DEE/Poli), UFRJ. His research interests include adaptive filtering and machine learning, digital signal processing, sound source localization and enhancement through microphone arrays, and wireless communications systems, in particular the physical layer techniques employed in 4G and 5G systems. He also collaborates with the Institute of Electrical and Electronics Engineers (IEEE) in a volunteer basis, having integrated the Technical Program Committee of the International Symposium on Industrial Electronics in 2015, the board of the IEEE Section in Rio de Janeiro as a Secretary from 2016 to 2017, and he has been the Chair of the IEEE Signal Processing Chapter in Rio de Janeiro since 2018. He is a member of the Brazilian Telecommunications Society (SBrT).

EDIZ CETIN (Member, IEEE) received the B.Eng. degree (Hons.) in control and computer engineering and the Ph.D. degree in unsupervised adaptive signal processing techniques for wireless receivers from the University of Westminster, London, U.K., in 1996 and 2002, respectively.

From 2002 to 2011, he was with the University of Westminster, initially as a Postdoctoral Research Fellow and subsequently, from 2006 to 2011, as a Senior Lecturer. From 2011 to 2017, he was a Senior Research Associate with the Australian Centre for Space Engineering Research, University of New South Wales, Sydney, NSW, Australia. He is currently a Senior Lecturer with the School of Engineering, Macquarie University, Sydney. To date, he has authored or coauthored more than 70 technical publications, a book chapter, and holds two patents in the areas of communications and GNSS receivers. His research interests include radio frequency interference detection and localization, fault-tolerant reconfigurable circuits for space applications, adaptive techniques for RF impairment mitigation for communications and global navigation satellite system (GNSS) receivers, and design and low-power implementation of digital circuits.

Dr. Cetin served as the Chair of the Institute of Electrical and Electronics Engineers New South Wales (NSW) Circuits and Systems/Solid-State Circuits/Photonics/Electron Devices joint chapter. He is a member of the Institution of Engineering and Technology and served as the Chair of the IET NSW Local Network.

WALLACE A. MARTINS (Senior Member, IEEE) received the Electronics Engineering degree and the M.Sc. and D.Sc. degrees in electrical engineering from the Federal University of Rio de Janeiro (UFRJ), Rio de Janeiro, Brazil, in 2007, 2009, and 2011, respectively. He was a Research Visitor with the University of Notre Dame, USA, in 2008, Université de Lille 1, France, in 2016, and Universidad de Alcalá, Spain, in 2018. He was an Associate Professor with the Department of Electronics and Computer Engineering (DEL/Poli) and the Electrical Engineering Program (PEE/COPPE), UFRJ from 2013 to 2022. He was an Academic Coordinator and the Deputy Department Chairman (DEL/Poli), UFRJ from 2016 to 2017. He is currently a Research Scientist working with the Interdisciplinary Centre for Security, Reliability and Trust (SnT), University of Luxembourg. His research interests include digital signal processing and telecommunications, with focus on equalization and beamforming/precoding for terrestrial and non-terrestrial (satellite) wireless communications. He was the recipient of the Best Student Paper Award from EURASIP at EUSIPCO-2009, Glasgow, Scotland, the 2011 Best Brazilian D.Sc. Dissertation Award from Capes, and the Best Paper Award at SBrT-2020, Florianópolis, Brazil. He is a member (Associate Editor) of the Editorial Boards of the IEEE SIGNAL PROCESSING LETTERS and the *Journal on Advances in Signal Processing* (EURASIP).