IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received September 1, 2018, accepted September 23, 2018, date of publication September 28, 2018,
date of current version October 19, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2872516

Beam Grouping Based User Scheduling in Multi-Cell Millimeter-Wave MIMO Systems

ANZHONG HU, (Member, IEEE)
School of Communication Engineering, Hangzhou Dianzi University, Hangzhou 310018, China
e-mail: huaz@hdu.edu.cn

This work was supported by the National Natural Science Foundation of China under Project 61601152.

ABSTRACT This paper investigates the user scheduling in multi-cell millimeter-wave multiple-input multiple-output systems. The asymptotic analysis reveals that scheduling of users with larger effective channel gains is beneficial. Based on the relation between the inter-cell interference and the channel correlation, the users are grouped according to the beams they belong to. In accordance, scheduling of users in the same group will cause high inter-cell interference. Then, a user scheduling criterion which minimizes the inter-cell interference and maximizes the effective channel gain is proposed. Finally, the computational complexity of the proposed approach is analyzed and is compared with that of other approaches. The simulations verify that the proposed approach can mitigate more inter-cell interference than other approaches in favorable situations.

INDEX TERMS MIMO systems, beams, millimeter-wave communication, scheduling.

I. INTRODUCTION

The future wireless systems are expected to provide much higher spectral efficiency and sum rate [1], [2]. In the emerging technologies, the millimeter-wave (mm-wave) multiple-input multiple-output (MIMO) [3] is of great potential. In mm-wave MIMO systems, wide bandwidth can be used to increase sum rate. Moreover, many antennas can be equipped in mm-wave MIMO systems, which can be used to achieve diversity gain and multiplexing gain. Meanwhile, the large number of antennas constitute the massive MIMO system [4], [5], which is also one candidate for future wireless systems. As can be seen, the spectral efficiency and the sum rate can be greatly improved in mm-wave MIMO systems.

User scheduling is critical for mitigating interference in mm-wave MIMO systems [3]. In the mm-wave frequency, the free-space path loss is much higher than that in microwave frequency, which means the scattering in mm-wave transmission is limited [6]. Thus, the channels from the mobile stations (MSs) to the base station (BS) may be highly correlated for common scatterers, and cause severe interference. User scheduling is a direct way for mitigating inter-cell interference, and has been extensively studied in wireless systems.

Most of the user scheduling approaches try to maximize the sum rate of the system. The optimal method is the exhaustive search, which lists all the subsets of MSs and is computationally prohibitive when the number of MSs is large. By terminating when an acceptable sum rate is reached, the complexity of exhaustive search can be decreased [7]. In order to further reduce the complexity, several suboptimal approaches have been proposed, in which the greedy algorithm is most popular. The basic greedy scheduling method selects the MS that can provide the maximal increase of the sum rate in each iteration [8]. As the sum rate is closely related to the channel norm and the capacity bound, the corresponding metrics are maximized in the greedy methods in [8]–[12], and the complexity can be decreased. In order to further decrease the complexity of the greedy method, semi-orthogonal user selection is proposed to pre-select the MSs of less interference [13], [14]. Another suboptimal approach is opportunistic scheduling, in which the MS with the best channel condition or the maximum sum rate is selected to transmit in each resource slot [15], as those in [16]–[18].

Some of the user scheduling approaches emphasize fairness. The simplest method is round-robin (RR) scheduling, which schedules each MS with equal resource blocks and in circular order [19], [20]. The most common fairness criterion is the max-min criterion [21]. The max-min criterion is known to be employed in resource allocation, as in [22]–[24]. When the max-min criterion is employed in user scheduling, the MS with the minimum average throughput should be scheduled. Another method that emphasizes fairness is the proportional fair (PF) method [21], [25]. In PF scheduling, the MS with the maximum ratio of the current data rate to the average throughput is scheduled, which means that the MS with the current data rate near its peak is likely to be scheduled [16]. The PF scheduling has been employed in various kinds of systems, such as those in [26] and [27].

In multi-cell mm-wave systems, the interference channels may be highly correlated with the transmission channel, which means scheduling of MSs according to the channel correlation is necessary. However, this topic has not been addressed in these existing articles. In this paper, user scheduling for maximizing the sum rate of multi-cell mm-wave MIMO systems is investigated. The asymptotic signal-to-interference-plus-noise ratio (SINR) and the inter-cell interference is analyzed. Based on the analysis, the MSs are grouped according to the beams of the channels and the scheduling criterion is formed. The main contributions of this paper are three-fold.

1) The asymptotic signal-to-interference-plus-noise ratio (SINR) of the system is analyzed. The analysis shows that scheduling MSs with higher effective channels gain is beneficial for the sum rate.
2) The relation between the inter-cell interference and the channel correlation is analyzed. As the analysis demonstrates that larger distances between the inter-cell interference channel and the transmission channel in the beamspace results into lower inter-cell interference, the idea of grouping the MSs according to the beams of the channels is proposed. The grouping makes MSs that cause low inter-cell interference of higher priority in scheduling.
3) The computational complexity of the proposed approach is analyzed. The analysis demonstrates that the computational complexity of the proposed approach is comparable to or lower than that of the existing approaches.

This paper is organized as follows. In Section II, the system model and the assumptions are given. Section III presents the analysis of the SINR, the inter-cell interference, and the beam grouping approach is proposed in this section. In Section IV, the computational complexities of the approaches are analyzed and compared. Section V gives the simulation parameters and the numerical results. Finally, conclusions are drawn in Section VI.

Notations: Lower-case (upper-case) boldface symbols denote vectors (matrices); $\mathbf{I}_K$ represents the $K \times K$ identity matrix; $(\cdot)^H$ and $\mathbb{E}\{\cdot\}$ denote the conjugate transpose and the expectation, respectively; $[\cdot]_j$ is the j-th element of a vector or the j-th column of a matrix; $||\mathbf{A}||_\mathrm{F}$ is the Frobenius norm of $\mathbf{A}$; $[\cdot]_{j,k}$ is the element in the j-th row and k-th column of a matrix; $\mathcal{A} \cup \mathcal{B}$ is the union of $\mathcal{A}$ and $\mathcal{B}$; $(j, k) \neq (l, k')$ means $j \neq l$ or $k \neq k'$; and i is the imaginary unit.

II. SYSTEM MODEL

Suppose there are L cells in the mm-wave MIMO system, each cell consists of one BS and K MSs. The BS is equipped with one uniform rectangular array (URA) of N_B antennas and each MS has one URA of N_M antennas. Each BS has N_R radio frequency (RF) chains, and each MS has only one RF chain. As can be seen, the BS can only transmit to N_R MSs simultaneously. Moreover, the full-connected hybrid structure is employed, which means each RF chain is connected to all the antennas through analog phase shifters. The MSs distribute randomly in front of the array. In this paper, we focus on the downlink transmission, and the uplink transmission can be tackled in a similar way.

A. SIGNAL AND CHANNEL

The received signal of the (j, k)-th MS, which means the k-th MS in the j-th cell, is denoted as $\mathbf{y}_{j,k} \in \mathbb{C}^{N_\mathrm{M}\times 1}$ and can be written as

$$\mathbf{y}_{j,k} = \sum_{l=1}^{L} \mathbf{H}_{jl,k}\mathbf{x}_l + \mathbf{n}_{j,k}, \tag{1}$$

where $\mathbf{H}_{jl,k} \in \mathbb{C}^{N_\mathrm{M}\times N_\mathrm{B}}$ is the channel matrix from the MSs in the j-th cell to the l-th BS, $\mathbf{x}_l \in \mathbb{C}^{N_\mathrm{B}\times 1}$ is the transmitted signal from the l-th BS, $\mathbf{n}_{j,k} \in \mathbb{C}^{N_\mathrm{M}\times 1}$ is the noise received at the k-th MS in the j-th cell, and is composed of independent and identically distributed (i.i.d.) complex random variables with zero mean and variance σ^2.

In mm-wave MIMO systems, the channel from the (j, k)-th MS to the l-th BS $\mathbf{H}_{jl,k} \in \mathbb{C}^{N_\mathrm{M}\times N_\mathrm{B}}$ can be expressed as [28]–[30]

$$\mathbf{H}_{jl,k} = \sum_{p=1}^{P} \alpha_{jl,pk}\mathbf{a}_\mathrm{M}(\theta^\mathrm{M}_{jl,pk}, \phi^\mathrm{M}_{jl,pk})\mathbf{a}^H_\mathrm{B}(\theta^\mathrm{B}_{jl,pk}, \phi^\mathrm{B}_{jl,pk}), \tag{2}$$

where P is the number of paths, $\alpha_{jl,pk}$ is the complex path gain of the p-th path from (j, k)-th MS to the l-th BS; $\alpha_{jl,pk} \geq \alpha_{jl,\tilde{p}k}$ for $p > \tilde{p}$; $\theta^\mathrm{B}_{jl,pk}$ and $\phi^\mathrm{B}_{jl,pk}$ are the corresponding azimuth direction-of-departure (DOD) and elevation DOD at the l-th BS, respectively; $\theta^\mathrm{M}_{jl,pk}$ and $\phi^\mathrm{M}_{jl,pk}$ are the corresponding azimuth direction-of-arrival (DOA) and elevation DOA at the (j, k)-th MS, respectively. Note that one of the P paths is the line-of-sight (LOS) path, and other paths are multipaths. When the LOS path undergoes blockage, the corresponding path gain is zero.

In addition, $\mathbf{a}_\mathrm{B}(\theta^\mathrm{B}_{jl,pk}, \phi^\mathrm{B}_{jl,pk}) \in \mathbb{C}^{N_\mathrm{B}\times 1}$ is the array manifold defined as

$$\begin{aligned}[\mathbf{a}_\mathrm{B}(\theta^\mathrm{B}_{jl,pk}, \phi^\mathrm{B}_{jl,pk})]_n &= \frac{1}{\sqrt{N_\mathrm{B}}} e^{i\pi(n_\mathrm{Bx}-0.5(N_\mathrm{Bx}-1))\cos\phi^\mathrm{B}_{jl,pk}\sin\theta^\mathrm{B}_{jl,pk}} \\ &\quad \times e^{i\pi(n_\mathrm{By}-0.5(N_\mathrm{By}-1))\sin\phi^\mathrm{B}_{jl,pk}},\end{aligned} \tag{3}$$

where N_Bx and N_By are the numbers of antennas in the horizontal and vertical directions of the URA at the BS, and satisfy $N_\mathrm{B} = N_\mathrm{Bx}N_\mathrm{By}$; $n = n_\mathrm{By}N_\mathrm{Bx} + n_\mathrm{Bx} + 1$, $n_\mathrm{Bx} = 0, 1, \cdots, N_\mathrm{Bx} - 1, n_\mathrm{By} = 0, 1, \cdots, N_\mathrm{By} - 1$. Similarly, $\mathbf{a}_\mathrm{M}(\theta^\mathrm{M}_{jl,pk}, \phi^\mathrm{M}_{jl,pk}) \in \mathbb{C}^{N_\mathrm{M}\times 1}$ is the array manifold defined as

$$[\mathbf{a}_{\mathrm{M}}(\theta^{\mathrm{M}}_{jl,pk}, \phi^{\mathrm{M}}_{jl,pk})]_n = \frac{1}{\sqrt{N_{\mathrm{M}}}} e^{i\pi(n_{\mathrm{Mx}}-0.5(N_{\mathrm{Mx}}-1))\cos\phi^{\mathrm{M}}_{jl,pk}\sin\theta^{\mathrm{M}}_{jl,pk}} \times e^{i\pi(n_{\mathrm{My}}-0.5(N_{\mathrm{My}}-1))\sin\phi^{\mathrm{M}}_{jl,pk}}, \tag{4}$$

where N_{Mx} and N_{My} are the numbers of antennas in the horizontal and vertical directions of the URA at the MS, and satisfy $N_{\mathrm{M}} = N_{\mathrm{Mx}}N_{\mathrm{My}}$; $n = n_{\mathrm{My}}N_{\mathrm{Mx}} + n_{\mathrm{Mx}} + 1$, $n_{\mathrm{Mx}} = 0, 1, \cdots, N_{\mathrm{Mx}} - 1$, $n_{\mathrm{My}} = 0, 1, \cdots, N_{\mathrm{My}} - 1$.

B. TRANSCEIVER PROCESSING

By employing hybrid precoding, the transmitted signal from the l-th BS can be expressed as

$$\mathbf{x}_l = \mathbf{U}_l\mathbf{F}_l\mathbf{s}_l, \tag{5}$$

where $\mathbf{U}_l \in \mathbb{C}^{N_{\mathrm{B}}\times N_{\mathrm{R}}}$ is the analog precoding matrix of the l-th BS, $\mathbf{F}_l \in \mathbb{C}^{N_{\mathrm{R}}\times N_{\mathrm{R}}}$ is the digital precoding matrix of the l-th BS, $\mathbf{s}_l \in \mathbb{C}^{N_{\mathrm{R}}\times 1}$ is composed of the transmitted data symbols for the scheduled MSs in the l-th cell, and each element of $\mathbf{s}_l$ is of zero mean and variance one.

At the (j, k)-th MS, the received signal in (1) is processed with analog combining and is changed into

$$y_{j,k} = \mathbf{w}^H_{j,k}\mathbf{y}_{j,k} = \sum_{l=1}^{L} \mathbf{w}^H_{j,k}\mathbf{H}_{jl,k}\mathbf{x}_l + \mathbf{w}^H_{j,k}\mathbf{n}_{j,k}, \tag{6}$$

where $\mathbf{w}_{j,k} \in \mathbb{C}^{N_{\mathrm{M}}\times 1}$ is the analog combiner used at the MS. Substituting (5) into (6) yields

$$y_{j,k} = \sum_{l=1}^{L} \mathbf{w}^H_{j,k}\mathbf{H}_{jl,k}\mathbf{U}_l\mathbf{F}_l\mathbf{s}_l + \mathbf{w}^H_{j,k}\mathbf{n}_{j,k}. \tag{7}$$

Note that the employment of hybrid precoding is the result of the limited number of RF chains. As there are N_{R} RF chains at the BS, only N_{R} digital signals can be simultaneously converted to analog signals, which are further converted to N_{B} analog signals. The digital and analog signals should be processed separately and this kind of precoding is hybrid precoding. Furthermore, the employment of hybrid precoding will affect the sum rate of the system. Consequently, the scheduling with hybrid precoding will also be different from that with fully digital precoding.

C. ASSUMPTIONS AND PROBLEM FORMULATION

The main assumptions used in this paper are as follows.

1) The j-th BS has knowledge of the DODs of all the MSs inside the j-th cell, i.e., $\theta^{\mathrm{B}}_{jj,pk}$, $\phi^{\mathrm{B}}_{jj,pk}$, $\forall p, k$, and the LOS DODs of MSs in other cells, which are denoted as $\bar{\theta}^{\mathrm{B}}_{jl,k}$, $\bar{\phi}^{\mathrm{B}}_{jl,k}$, $\forall l \neq j, k$. The (j, k)-th MS knows the DOAs of the strongest path from the j-th BS, i.e., $\theta^{\mathrm{M}}_{jj,1k}$ and $\phi^{\mathrm{M}}_{jj,1k}$. The channel is invariant in one channel coherence interval of T slots. Note that the existing angle estimation algorithms like those in [31] and [32] can be employed for angle information.

2) The analog precoder vector, i.e., each column of $\mathbf{U}_l$, can only be chosen from the codebook $\mathcal{F}_{\mathcal{B}} = \{\mathbf{a}_{\mathrm{B}}(\theta, \phi)|\cos\phi\sin\theta = -1 + 2m/N_{\mathrm{Bx}}, \sin\phi = -1 + 2n/N_{\mathrm{By}}, m = 0, 1, \cdots, N_{\mathrm{Bx}} - 1, n = 0, 1, \cdots, N_{\mathrm{By}} - 1\}$. The analog combining vector, i.e., $\mathbf{w}_{j,k}$, can only be chosen from the codebook $\mathcal{F}_{\mathcal{M}} = \{\mathbf{a}_{\mathrm{M}}(\theta, \phi)|\cos\phi\sin\theta = -1 + 2m/N_{\mathrm{Mx}}, \sin\phi = -1 + 2n/N_{\mathrm{My}}, m = 0, 1, \cdots, N_{\mathrm{Mx}} - 1, n = 0, 1, \cdots, N_{\mathrm{My}} - 1\}$.

3) The number of RF chains at the BS is less than the number of MSs in each cell, i.e., $N_{\mathrm{R}} < K$. This means that user scheduling is necessary and the number of MSs that can be scheduled is N_{R}. Additionally, the set of the indices of the scheduled MSs in the j-th cell is denoted as $\mathcal{S}_j = \{c_{j,n}|n = 1, 2, \cdots, N_{\mathrm{R}}\}$, in which $c_{j,n}$ is the index of the n-th scheduled MS in the j-th cell and we have $|\mathcal{S}_j| = N_{\mathrm{R}}$.

4) The (j, k)-th MS employs directional combining, i.e.,

$$\mathbf{w}_{j,k} = \arg\min_{\mathbf{a}_{\mathrm{M}}(\theta,\phi)\in\mathcal{F}_{\mathcal{M}}} |\cos\phi\sin\theta - \cos\phi^{\mathrm{M}}_{jj,1k}\sin\theta^{\mathrm{M}}_{jj,1k}| + |\sin\phi - \sin\phi^{\mathrm{M}}_{jj,1k}|.$$

5) The l-th BS employs directional beamforming, i.e.,

$$[\mathbf{U}_l]_n = \mathbf{a}_{\mathrm{B}}(\theta_{l,c_{l,n}}, \phi_{l,c_{l,n}}) \tag{8}$$

where

$$\mathbf{a}_{\mathrm{B}}(\theta_{l,k}, \phi_{l,k}) \triangleq \arg\min_{\mathbf{a}_{\mathrm{B}}(\theta,\phi)\in\mathcal{F}_{\mathcal{B}}} |\cos\phi\sin\theta - \cos\phi^{\mathrm{B}}_{ll,1k}\sin\theta^{\mathrm{B}}_{ll,1k}| + |\sin\phi - \sin\phi^{\mathrm{B}}_{ll,1k}|. \tag{9}$$

Note that the analog beamforming vectors can also be designed by maximizing the inner product of the analog beamforming vector and the channel vector, as in [33] and [34].

6) The l-th BS has knowledge of the effective channel gain

$$h_{l,k,k'} \triangleq \mathbf{w}^H_{l,k}\mathbf{H}_{ll,k}\mathbf{a}_{\mathrm{B}}(\theta_{l,k'}, \phi_{l,k'}), \quad \forall k, k' = 1, 2, \cdots, K \tag{10}$$

and employs ZF precoding, i.e.,

$$\mathbf{F}_l = \sqrt{\eta_l}\mathbf{Z}^{-1}_{ll} \in \mathbb{C}^{N_{\mathrm{R}}\times N_{\mathrm{R}}}, \tag{11}$$

where η_l is the power amplification factor, $\mathbf{Z}_{jl} \in \mathbb{C}^{N_{\mathrm{R}}\times N_{\mathrm{R}}}$ is defined as

$$[\mathbf{Z}_{jl}]_{n,:} = \mathbf{w}^H_{j,c_{j,n}}\mathbf{H}_{jl,c_{j,n}}\mathbf{U}_l. \tag{12}$$

Thus, we have

$$[\mathbf{Z}_{jj}]_{n,n'} = h_{j,c_{j,n},c_{j,n'}}. \tag{13}$$

7) The transmission power of each BS is fixed to σ^2_{x}, i.e., $\mathbb{E}\{||\mathbf{x}_l||^2_{\mathrm{F}}\} = \sigma^2_{\mathrm{x}}$. Substituting (5) and (11) into this equation yields

$$\eta_l = \frac{\sigma^2_{\mathrm{x}}}{||\mathbf{Z}^{-1}_{ll}||^2_{\mathrm{F}}}. \tag{14}$$

As a result, the n-th scheduled MS in the j-th cell at the t-th slot, cf. (7), can be written as

$$y_{j,c_{j,n}} = \sqrt{\eta_j}[\mathbf{s}_j]_n + \sum_{l=1,l\neq j}^{L} [\mathbf{Z}_{jl}]_{n,:}\mathbf{F}_l\mathbf{s}_l + \mathbf{w}_{j,c_{j,n}}^H \mathbf{n}_{j,c_{j,n}}. \quad (15)$$

Then, the SINR is

$$\gamma_{j,c_{j,n}} = \eta_j \left(\sum_{l=1,l\neq j}^{L} ||[\mathbf{Z}_{jl}]_{n,:}\mathbf{F}_l||^2 + \sigma^2 \right)^{-1}. \quad (16)$$

Note the SINR of the unscheduled MS is denoted as zero, i.e., $\gamma_{j,k} = 0, \forall k \notin \mathcal{S}_j$. Correspondingly, the accumulate sum rate of the system in one channel coherence interval is

$$R = T\sum_{j=1}^{L}\sum_{k=1}^{K} \log_2(1+\gamma_{j,k}). \quad (17)$$

Then, the user scheduling problem is how to maximize the sum rate, which is

$$\max_{\mathcal{S}_j,\forall j} R, \text{ s.t. } |\mathcal{S}_j| = N_{\mathrm{R}}, \quad \mathcal{S}_j \subset \{1, 2, \cdots, K\}. \quad (18)$$

Note that the limited number of RF chains means only N_{R} out of K users can be scheduled in each cell each time. Though the object means that the "best" N_{R} users will be scheduled in each cell all the time, it can be easily modified for ensuring that all the K MSs can be scheduled. For example, the users scheduled in the past can be temporarily abandoned for scheduling and the duration of the abandon depends on the requirements on sum rate and fairness.

With the assumptions and the user scheduling problem, a beam grouping based approach will be investigated in the following section.

III. BEAM GROUPING BASED USER SCHEDULING

As shown earlier, user scheduling for multi-cell mm-wave MIMO systems is important but is rarely investigated. The existing user scheduling approaches cannot tackle the inter-cell interference. Since the sum rate R is not a convex function of $\mathcal{S}_j$, the optimization problem (18) cannot be directly solved by using interior-point methods. In order to tackle the problem, in this section, we first maximize the signal power by resorting to the asymptotic analysis, and then restrict the inter-cell interference with beam grouping, and finally propose a scheduling criterion that maximizes the sum rate, with which we form a beam grouping based user scheduling algorithm.

A. MAXIMIZATION OF THE RECEIVING POWER

From the sum rate expression (17), it can be seen that the SINR $\gamma_{j,k}$ determines that sum rate. Substituting (11) and (14) into (16) yields

$$\gamma_{j,c_{j,n}}^{-1} = ||\mathbf{Z}_{jj}^{-1}||_{\mathrm{F}}^2 \left(\sum_{l=1,l\neq j}^{L} \frac{||[\mathbf{Z}_{jl}]_{n,:}\mathbf{Z}_{ll}^{-1}||^2}{||\mathbf{Z}_{ll}^{-1}||_{\mathrm{F}}^2} + \frac{\sigma^2}{\sigma_{\mathrm{x}}^2} \right). \quad (19)$$

Before we continue deriving, the asymptotic result of $\mathbf{Z}_{jj}$ should be given first.

Lemma 1: When the scheduled MSs in the j-th cell do not share the same path, i.e., $\theta_{jj,pk}^{\mathrm{B}} \neq \theta_{jj,\tilde{p}\tilde{k}}^{\mathrm{B}}$, $\phi_{jj,pk}^{\mathrm{B}} \neq \phi_{jj,\tilde{p}\tilde{k}}^{\mathrm{B}}$, $\forall p, \tilde{p}, k \neq \tilde{k}$, $\mathbf{Z}_{jj}$ tends to be a diagonal matrix as N_{Bx}, $N_{\mathrm{By}} \to \infty$.

Proof: According to the definition of $\mathbf{Z}_{jl}$ in (12), we have

$$[\mathbf{Z}_{jj}]_{n,n'} = \mathbf{w}_{j,c_{j,n}}^H \mathbf{H}_{jj,c_{j,n}} [\mathbf{U}_j]_{n'}. \quad (20)$$

Substituting (2) and (8) into (23) results into

$$[\mathbf{Z}_{jj}]_{n,n'} = \sum_{p=1}^{P} \alpha_{jj,pc_{j,n}} \mathbf{w}_{j,c_{j,n}}^H \mathbf{a}_{\mathrm{M}}(\theta_{jj,pc_{j,n}}^{\mathrm{M}}, \phi_{jj,pc_{j,n}}^{\mathrm{M}}) \times \mathbf{a}_{\mathrm{B}}^H(\theta_{jj,pc_{j,n}}^{\mathrm{B}}, \phi_{jj,pc_{j,n}}^{\mathrm{B}}) \mathbf{a}_{\mathrm{B}}(\theta_{j,c_{j,n'}}, \phi_{j,c_{j,n'}}).$$

As $N_{\mathrm{Bx}}, N_{\mathrm{By}} \to \infty$, we have $\theta_{jj,pc_{j,n'}}^{\mathrm{B}} \to \theta_{j,c_{j,n'}}$, $\phi_{jj,pc_{j,n'}}^{\mathrm{B}} \to \phi_{j,c_{j,n'}}$. Since $\theta_{jj,pk}^{\mathrm{B}} \neq \theta_{jj,\tilde{p}\tilde{k}}^{\mathrm{B}}$, $\phi_{jj,pk}^{\mathrm{B}} \neq \phi_{jj,\tilde{p}\tilde{k}}^{\mathrm{B}}$, $\forall p, \tilde{p}, k \neq \tilde{k}$, we have $|\theta_{jj,pc_{j,n}}^{\mathrm{B}} - \theta_{j,c_{j,n'}}| \nrightarrow 0$, $|\phi_{jj,pc_{j,n}}^{\mathrm{B}} - \phi_{j,c_{j,n'}}| \nrightarrow 0$, $\forall n \neq n'$. As a result, $\mathbf{a}_{\mathrm{B}}^H(\theta_{jj,pc_{j,n}}^{\mathrm{B}}, \phi_{jj,pc_{j,n}}^{\mathrm{B}})\mathbf{a}_{\mathrm{B}}(\theta_{j,c_{j,n'}}, \phi_{j,c_{j,n'}}) \to 0, \forall n \neq n'$. Hence, $\mathbf{Z}_{jj}$ tends to be a diagonal matrix. ■

Then, we have the following corollary.

Corollary 1: When the scheduled MSs in the j-th cell do not share the same path and the interference $\mathbf{Z}_{jl} = \mathbf{0}_{N_{\mathrm{R}}\times N_{\mathrm{R}}}, \forall l \neq j$, $\lim_{N_{\mathrm{Bx}},N_{\mathrm{By}}\to\infty} \gamma_{j,c_{j,n}}$ increases with the increase of $|h_{j,c_{j,n},c_{j,n}}|, \forall n$.

Proof: With $\mathbf{Z}_{jl}, \forall l \neq j$ being a zero matrix, we have $\gamma_{j,c_{j,n}}^{-1} = ||\mathbf{Z}_{jj}^{-1}||_{\mathrm{F}}^2\sigma^2/\sigma_{\mathrm{x}}^2$. According to Lemma 1, we have $\lim_{N_{\mathrm{Bx}},N_{\mathrm{By}}\to\infty} \gamma_{j,c_{j,n}} = \sigma_{\mathrm{x}}^2/\sigma^2/\sum_{n=1}^{N_{\mathrm{R}}} |[\mathbf{Z}_{jj}^{-1}]_{n,n}|^2$. According to (13), we have $[\mathbf{Z}_{jj}^{-1}]_{n,n} = 1/h_{j,c_{j,n},c_{j,n}}$. Then, we have $\lim_{N_{\mathrm{Bx}},N_{\mathrm{By}}\to\infty} \gamma_{j,c_{j,n}} = \sigma_{\mathrm{x}}^2/\sigma^2/\sum_{n=1}^{N_{\mathrm{R}}} |h_{j,c_{j,n},c_{j,n}}|^{-2}$. As can be seen, when $|h_{j,c_{j,n},c_{j,n}}|, \forall n$ increases, $\lim_{N_{\mathrm{Bx}},N_{\mathrm{By}}\to\infty} \gamma_{j,c_{j,n}}$ increases. ■

According to the asymptotic result in Corollary 1, we know that MSs with larger $|h_{j,c_{j,n},c_{j,n}}|$ should be scheduled in the j-th cell for maximizing the receiving power, i.e.,

$$\max_{\mathcal{S}_j,\forall j} |h_{j,c_{j,n},c_{j,n}}|, \quad \forall n. \quad (21)$$

B. RESTRICTION OF THE INTER-CELL INTERFERENCE

In order to maximize $\gamma_{j,c_{j,n}}, \forall n$, we also need to minimize $||\mathbf{Z}_{jl}||_{\mathrm{F}}, \forall l \neq j$, cf. (19), i.e.,

$$\min_{\mathcal{S}_j} ||\mathbf{Z}_{jl}||_{\mathrm{F}}, \quad \forall j, \ \forall l \neq j. \quad (22)$$

According to the definition of $\mathbf{Z}_{jl}$ in (12), one apparent way to achieve (22) is to

$$\min_{\mathcal{S}_j} |[\mathbf{H}_{jl,c_{j,n}}]_{m,:}[\mathbf{U}_l]_{m'}|, \quad \forall j, \ \forall l \neq j, \ \forall n, m, m'. \quad (23)$$

Substituting (2) and (8) into (23) results into

$$\min_{\mathcal{S}_j}\left|\sum_{p=1}^{P}\alpha_{jl,pc_{j,n}}\left[\mathbf{a}_{\mathrm{M}}(\theta^{\mathrm{M}}_{jl,pc_{j,n}},\phi^{\mathrm{M}}_{jl,pc_{j,n}})\right]_m \mathbf{a}^H_{\mathrm{B}}(\theta^{\mathrm{B}}_{jl,pc_{j,n}},\phi^{\mathrm{B}}_{jl,pc_{j,n}})\mathbf{a}_{\mathrm{B}}(\theta_{l,m'},\phi_{l,m'})\right|, \quad \forall j,\ \forall l\neq j,\ \forall n,m,m'. \tag{24}$$

Since we only have knowledge of the LOS DODs, $\bar{\theta}^{\mathrm{B}}_{jl,k}$, $\bar{\phi}^{\mathrm{B}}_{jl,k}$, we can only resort to

$$\min_{\mathcal{S}_j}\left|\mathbf{a}^H_{\mathrm{B}}(\bar{\theta}^{\mathrm{B}}_{jl,k},\bar{\phi}^{\mathrm{B}}_{jl,k})\mathbf{a}_{\mathrm{B}}(\theta_{l,m'},\phi_{l,m'})\right|, \quad \forall j,\ \forall l\neq j,\ \forall n,m,m'. \tag{25}$$

The correlation $|\mathbf{a}^H_{\mathrm{B}}(\bar{\theta}^{\mathrm{B}}_{jl,k},\bar{\phi}^{\mathrm{B}}_{jl,k})\mathbf{a}_{\mathrm{B}}(\theta_{l,m'},\phi_{l,m'})|$ is a function of $|\cos(\bar{\phi}^{\mathrm{B}}_{jl,k})\sin(\bar{\theta}^{\mathrm{B}}_{jl,k})-\cos(\phi_{l,n})\sin(\theta_{l,n})|$ and $|\sin(\bar{\phi}^{\mathrm{B}}_{jl,k})-\sin(\phi_{l,n})|$. Moreover, the function curve is similar to that of the sinc function, which generally tends to zero, i.e., be minimized, with the increase of the variable.

Thus, we seek to restrict the differences and the restrictions are

$$|\cos(\bar{\phi}^{\mathrm{B}}_{jl,k})\sin(\bar{\theta}^{\mathrm{B}}_{jl,k})-\cos(\phi_{l,n})\sin(\theta_{l,n})|>\frac{D_{\mathrm{Bx}}}{N_{\mathrm{Bx}}}, \quad \forall j,\quad \forall l\neq j,\ \forall n,m,m', \tag{26}$$

$$|\sin(\bar{\phi}^{\mathrm{B}}_{jl,k})-\sin(\phi_{l,n})|>\frac{D_{\mathrm{By}}}{N_{\mathrm{By}}}, \quad \forall j,\quad \forall l\neq j,\ \forall n,m,m', \tag{27}$$

where D_{Bx} and D_{By} are parameters that can be designed. The two dimensional space with axes being $\cos(\phi)\sin(\theta)$ and $\sin(\phi)$ is denoted as beamspace. With the restrictions in (26) and (27), the beamspace can be divided into N_{B} subspaces. Additionally, the set of the axes of points in the $mN_{\mathrm{By}}+n$-th subspace is denoted as $\mathcal{C}_{mN_{\mathrm{By}}+n}$, which is defined as

$$\begin{aligned}\mathcal{C}_{mN_{\mathrm{By}}+n} \triangleq \{(x,y)\,|\,&|x-(-1+2m/N_{\mathrm{Bx}})|\leq D_{\mathrm{Bx}}/N_{\mathrm{Bx}},\\ &|y-(-1+2n/N_{\mathrm{By}})|\leq D_{\mathrm{By}}/N_{\mathrm{By}}\},\\ &m=0,1,\cdots,N_{\mathrm{Bx}}-1,\ n=0,1,\cdots,N_{\mathrm{By}}-1.\end{aligned}$$

The union of all the sets should cover the whole beamspace, which means $D_{\mathrm{Bx}}\geq 1, D_{\mathrm{By}}\geq 1$. Then, we can group MSs according to the sets which the beamspace points of their paths belong to. More specifically, for the (l,k')-th MS, its group is named the (l,k')-th group, which is defined as

$$\begin{aligned}\mathcal{G}_{l,k'} = \{(j,k)|\{&(\cos(\bar{\phi}^{\mathrm{B}}_{jl,k})\sin(\bar{\theta}^{\mathrm{B}}_{jl,k}),\sin(\bar{\phi}^{\mathrm{B}}_{jl,k})),\\ &(\cos(\phi^{\mathrm{B}}_{ll,1k'})\sin(\theta^{\mathrm{B}}_{ll,1k'}),\sin(\phi^{\mathrm{B}}_{ll,1k'}))\}\in\mathcal{C}_{mN_{\mathrm{By}}+n},\\ &\forall(j,k)\neq(l,k'),\forall m,n\},\end{aligned} \tag{28}$$

which is a set of the indices of the MSs with beamspace points of paths to the l-th BS in the same set as the beamspace point of the first path from the (l,k')-th MS to the l-th BS. In order to achieve the goal in (26) and (27), MSs in $\mathcal{G}_{l,k'}$ should not be scheduled when the (l,k')-th MS is scheduled, i.e.,

$$k\notin\mathcal{S}_j,(j,k)\in\mathcal{G}_{l,k'},\quad \forall k'\in\mathcal{S}_l,\ l\neq j. \tag{29}$$

According to the scheduling restriction above, it can be seen that the parameters D_{Bx} and D_{By} are crucial to the scheduling performance. When they grow large, the correlation in (24) tends to zero, which means the inter-cell interference tends to vanish. However, as the two parameters grow large, more MSs are incorporated into one group, which means more MSs should not be scheduled when the (l,k')-th MS is scheduled. Thus, the number of MSs that can be scheduled decreases, which may deteriorate the sum rate. These two parameters can be designed by taking simulations, and the numerical results demonstrate that the simplest yet effective way is to set $D_{\mathrm{Bx}}=D_{\mathrm{By}}=4$ so as to avoid the strongest inter-cell interference.

Note that the idea of grouping users has been investigated in some existing articles. For example, [35] group users according to the channel spaces, [36] proposes several grouping criterions and a dynamic scheduling approach, [37] adopts a user grouping scheme similar to that in [35] and proposes a channel estimation approach. However, user grouping here is based on the beamspace distances for the LOS paths, and is different from these grouping approaches.

C. SCHEDULING CRITERION AND ALGORITHM

In the last two subsections, we have separated the object in (18) into two sub-objects, which are (21) and (29). However, it is possible that the two sub-objects contradict with each other. More specifically, MSs that achieve the goal in (21) may violate the restriction in (29). In order to mitigate inter-cell interference, we propose to give priority to the restriction in (29).

As a result, we have the following beam grouping based user scheduling criterion. When MSs in the l-th cell have been scheduled, MSs in the j-th cell are partitioned into two groups, which are named low-interference group and high-interference group. In addition, the (j,k)-th MS is in the low-interference group if this MS accords with (29), and is in the high-interference group otherwise. Then, MSs in the low-interference group have priority in scheduling. Moreover, in each group, one MS has priority in scheduling if this MS corresponds to a higher value of $|h_{j,c_{j,n},c_{j,n}}|$. Correspondingly, the K MSs in the j-th cell are ordered in the scheduling process. Moreover, the index of the a_j-th MS in the order is denoted as g_{a_j}. For the case that MSs in the j-th cell are firstly scheduled, the MSs only need to be ordered according to $|h_{j,c_{j,n},c_{j,n}}|$.

Additionally, the scheduling of the MSs should make $\mathbf{Z}_{jj}$ invertible. More specifically, the scheduling is executed in a successive manner. In the a_j-th step, whether the g_{a_j}-th MS should be scheduled is judged. This MS cannot be scheduled if $\tilde{\mathbf{Z}}_{jj}(a_j)\in\mathbb{C}^{|\tilde{\mathcal{S}}_j(a_j)|\times|\tilde{\mathcal{S}}_j(a_j)|}$ is ill-conditioned, where

$$[\tilde{\mathbf{Z}}_{jj}(a_j)]_{n,n'}=h_{j,\tilde{c}_{j,n},\tilde{c}_{j,n'}}. \tag{30}$$

Algorithm 1 Beam Grouping Based User Scheduling Algorithm

Initialize: $a_j = 0$, $\mathcal{S}_j = \emptyset$, $\forall j$, $\delta = 1000$
1: **for** $j = 1 \to L$ **do**
2: sort MSs according to the scheduling criterion
3: **while** $|\mathcal{S}_j| < N_{\mathrm{R}}$ and $a_j < K$ **do**
4: choose the (j, g_{a_j})-th MS, $a_j \leftarrow a_j + 1$
5: **if** the condition number of $\tilde{\mathbf{Z}}_{jj}(a_j)$ is larger
6: than δ
7: **continue**
8: **else**
9: $\mathcal{S}_j = \mathcal{S}_j \cup (j, g_{a_j})$
10: **end if**
11: **end while**
12: **end for**

Moreover, $\tilde{\mathcal{S}}_j(a_j) = \{\bar{\mathcal{S}}_j(a_j), g_{a_j}\}$, where $\bar{\mathcal{S}}_j(a_j)$ is the set of the indices of the scheduled MSs, and the elements of $\tilde{\mathcal{S}}_j(a_j)$ are denoted as $\tilde{c}_{j,n}$, $n = 1, 2, \cdots, |\tilde{\mathcal{S}}_j(a_j)|$.

With the scheduling criterion, we can also form a scheduling algorithm based on beam grouping, which is given as follows.

With the beam grouping based user scheduling algorithm presented, we will continue analyzing the computational complexity and compare the proposed approach with other traditional approaches in the next section.

IV. COMPUTATIONAL COMPLEXITY ANALYSIS AND COMPARISON

In this section, the computational complexities of the user scheduling approaches are analyzed and compared, which are denoted with the big O notation.

The greedy approach is the most popular user scheduling approach for maximizing the sum rate. Since we only have knowledge of the LOS DODs, $\bar{\theta}^{\mathrm{B}}_{jl,k}$, $\bar{\phi}^{\mathrm{B}}_{jl,k}$, the inter-cell interference cannot be tackled in the greedy approach. More specifically, each time this approach adds one MS to the already scheduled MSs, and then calculates $||\bar{\mathbf{Z}}_{jj}(n)^{-1}||^2_{\mathrm{F}}$ in the process of choosing the n-th best MS, where $\bar{\mathbf{Z}}_{jj}(n) \in \mathbb{C}^{n\times n}$. In each cell, when we try to add the n-th MS, there are $K-n+1$ trials each of which results in $O(n^3)$. Thus, the computational complexity of the greedy approach is $O(\sum_{n=1}^{N_{\mathrm{R}}} (K-n+1)n^3 L)$.

The max-min scheduling approach favors the MS with the lower accumulate sum rate, i.e., it only emphasizes the scheduling fairness. Then, the MSs are chosen in an increasing order of the accumulate sum rate of each MS. However, those MSs with poor transmission conditions will also be scheduled with many time resources, and this will decrease the system capacity. More specifically, it chooses the N_{R} MSs with the lowest $\gamma_{j,k}$ in cell j for scheduling unless this makes $\mathbf{Z}_{jj}$ rank-deficient, in which case the MS that causes the rank-deficiency is replaced with the next MS in the order. However, this scheduling should be executed for each time slot. Thus, the computational complexity of the max-min approach is T times as large as that of the proposed approach.

The PF scheduling favors the MS with the largest ratio of the current rate to the accumulate rate. As the current rate depends on the scheduling decision, it can be approximated as the rate with only this MS scheduled. Then, in each cell, the MSs are chosen in an increasing order of this ratio unless $\mathbf{Z}_{jj}$ is rank-deficient, and this MS will be replaced with the next MS in the order. However, this scheduling should be executed for each time slot. Thus, the computational complexity of the PF approach is T times as large as that of the proposed approach.

For the proposed approach, the main computations are in step 5. For step 5, the condition number of $\tilde{\mathbf{Z}}_{jj}(a_j)$ should be calculated, the computational complexity of which is $O(|\tilde{\mathcal{S}}_j(a_j)|^3)$. Since $|\tilde{\mathcal{S}}_j(a_j)| < N_{\mathrm{R}}$, the computational complexity of the proposed approach is $O(N_{\mathrm{R}}^3 KL)$.

From the above analysis, it can be seen that the proposed approach is of similar computational complexity as the greedy approach, and is only lower computational complexity than the max-min and the PF approaches. With the analysis and comparison, we will continue to present numerical results to verify the better performance of the proposed approach.

V. NUMERICAL RESULTS

In this section, numerical results are provided to compare the performance of the proposed approach and other existing approaches. Without other statements, the simulation parameters are as Table 1. Each cell is of a rhombus shape. The BS height is 10 meters and the angle between the horizontal plane and the BS array surface normal is 10.3°. The distance between the MS and the BS in the horizontal plane is in the range of 90 meters to 100 meters. The horizontal angles between the line connecting the MS and the BS and the reference line is in the range of −60° to 60°. The probability of blockage of the LOS path is 0.1. The reflection loss of the multipath is in the range of 15 dB to 20 dB. More specifically, the complex path loss is given as [29]

$$\alpha_{jl,pk} = \sqrt{A(\theta^{\mathrm{B}}_{jl,pk}, \phi^{\mathrm{B}}_{jl,pk})}\frac{\lambda\beta_{jl,pk}e^{i\phi_{jl,pk}}}{4\pi d_{jl,k}},$$

TABLE 1. Simulation parameters.

Parameter	Value	Parameter	Value
N_{Bx}	12	N_{By}	100
N_{Mx}	2	N_{My}	2
N_{R}	4	K	20
P	4	L	3
T	20	$\sigma_{\mathrm{x}}^2/\sigma^2$	10^{18}
D_{Bx}	4	D_{By}	4

where

$$10\log_{10} A(\theta^{\mathrm{B}}_{jl,pk}, \phi^{\mathrm{B}}_{jl,pk}) = -\min\left(\min\left(\frac{12\theta^{\mathrm{B}\,2}_{jl,pk}}{(7\pi/18)^2}, 25\right) + \min\left(\frac{12\phi^{\mathrm{B}\,2}_{jl,pk}}{(7\pi/180)^2}, 20\right), 20\right)$$

is the directional antenna gain at the BS; $\lambda = 3.75$ millimeters is the wavelength; $\beta_{jl,pk}$ represents the blockage coefficient or the reflection loss; $\beta_{jl,pk}$ is zero for the former case and $20\log_{10}\beta_{jl,pk}$ is a random variable between -15 dB and -20 dB for the latter case; $\phi_{jl,pk}$ is a random phase. For the non-LOS path from one MS in the l-th cell to the j-th BS, where $l \neq j$, the reflector is in the l-th cell. In addition, the Jain's index can be used as a fairness index, and the definition is [38]

$$J = \frac{R^2}{KL\sum_{j=1}^{L}\sum_{k=1}^{K} R_{j,k}^2}, \tag{31}$$

where $R_{j,k} = \log_2(1+\gamma_{j,k})$.

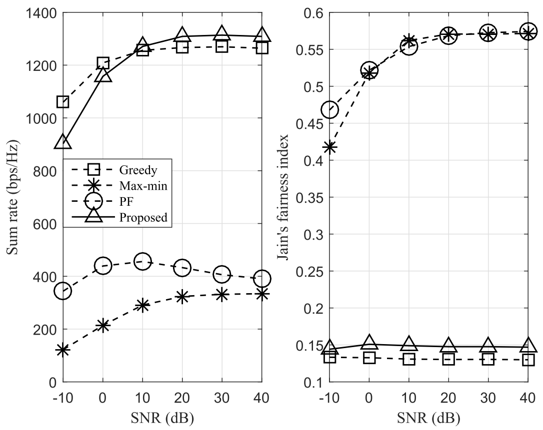

FIGURE 1. **The left figure demonstrates the sum rate versus the SNR. The right figure demonstrates the Jain's fairness index versus the SNR.**

In Fig. 1, the sum rate and the Jain's fairness index versus the SNR is demonstrated. Note that the SNR in the figure is given as $10\log_{10}\sigma_x^2/\sigma^2 - 140$, where 140 dB is the typical fading value in the transmission scenario. The figure shows that the proposed approach surpasses the greedy approach in regards to the sum rate as the SNR is larger than 10 dB. Since the inter-cell interference constrains the increase of the sum rate when the SNR is higher than 10 dB, the results mean that the proposed approach can mitigate more inter-cell interference than the greedy approach. Meanwhile, the Jain's fairness index of the proposed approach is always larger than that of the greedy approach. This means that the proposed approach can provide more fairness.

In Fig. 2, the sum rate and the Jain's fairness index versus the number of the MSs K is demonstrated. It can be seen that the sum rate of the proposed approach increases faster than that of other approaches as K increases. In addition, the proposed approach surpasses the greedy approach when $K \geq 20$. As the number of choices increases when K increases, this result verifies that the proposed approach is more efficient in mitigating inter-cell interference when there are more choices. It can also be seen that the Jain's fairness index of the proposed approach is larger than that of the greedy approach.

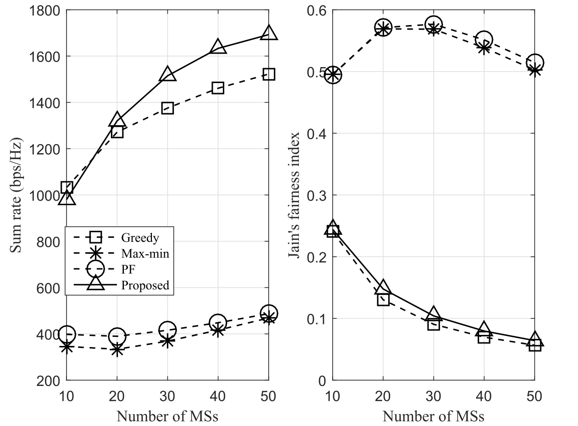

FIGURE 2. **The left figure demonstrates the sum rate versus the number of the MSs K. The right figure demonstrates the Jain's fairness index versus the number of the MSs K.**

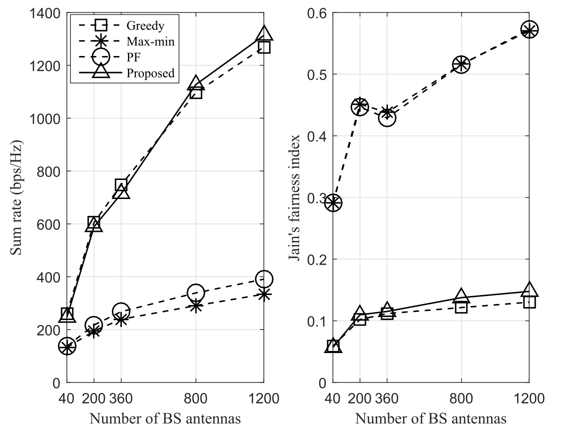

FIGURE 3. **The left figure demonstrates the sum rate versus the number of the BS antennas N_B. The right figure demonstrates the Jain's fairness index versus the number of the BS antennas N_B.**

In Fig. 3, the sum rate and the Jain's fairness index versus the number of the BS antennas N_B is demonstrated. Here, the utilized combinations of N_{Bx} and N_{By} are $(2, 20)$, $(5, 40)$, $(6, 60)$, $(10, 80)$, $(12, 100)$. As can be seen, the sum rate of the proposed approach is still higher than that of other approaches when $N_B \geq 800$. This result is in accordance with Corollary 1, which shows that in specific conditions the scheduling of MSs with larger $|h_{j,c_{j,n},c_{j,n}}|$ is beneficial as $N_{Bx}, N_{By} \to \infty$.

In Fig. 4, the sum rate and the Jain's fairness index versus the minimum distance between the MS and the BS in the horizontal plane is demonstrated. As the distance increases, the MSs distribute more close to the cell edge, which causes higher inter-cell interference. It can be seen that the proposed approach tends to perform better than the greedy approach as the distance increases, which also verifies that the proposed approach can mitigate more inter-cell interference.

In Fig. 5, the sum rate and the Jain's fairness index versus the probability of the blockage of the LOS path is demonstrated. As the probability increases, the sum rate of

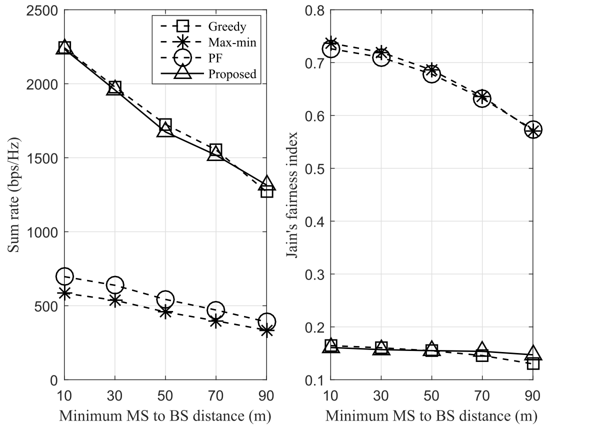

FIGURE 4. The left figure demonstrates the sum rate versus the minimum distance between the MS and the BS in the horizontal plane. The right figure demonstrates the Jain's fairness index versus the minimum distance between the MS and the BS in the horizontal plane.

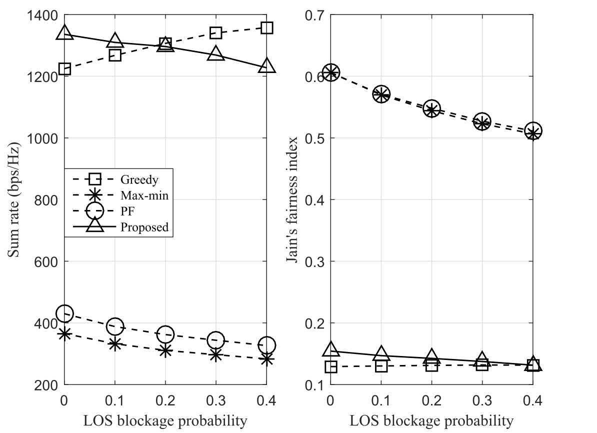

FIGURE 5. The left figure demonstrates the sum rate versus the probability of the blockage of the LOS path. The right figure demonstrates the Jain's fairness index versus the probability of the blockage of the LOS path.

the proposed approach decreases and is of lower sum rate than the greedy approach when the probability is larger than 0.2. This is also in accordance with the derivation of the approach. Since the proposed approach utilizes the LOS DODs to measure the inter-cell interference, the measure accuracy deteriorates as the probability of the blockage of the LOS path increases. Thus, the less inter-cell interference can be mitigated as the probability increases.

VI. CONCLUSIONS

In this paper, the user scheduling in multi-cell mm-wave MIMO systems is investigated. The asymptotic SINR of the system is first derived, which demonstrates that MSs with larger effective channel gains should be scheduled first. Based on the correlation property of the array manifold, the MSs are grouped according to the beams that the DODs belong to. Correspondingly, MSs not in the groups of scheduled MSs should be scheduled first. Then, the scheduling criterion which emphasizes both the maximization of the effective channel gain and the minimization of the inter-cell interference is proposed. Correspondingly, a user scheduling algorithm is proposed. After that, the computational complexity of the proposed approach is analyzed and compared with that of other approaches. Finally, simulations are carried out and the results demonstrate the better performance of the proposed approach.

REFERENCES

[1] Z. Pi and F. Khan, "An introduction to millimeter-wave mobile broadband systems," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 101–107, Jun. 2011.

[2] T. S. Rappaport *et al.*, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, 2013.

[3] A. L. Swindlehurst, E. Ayanoglu, P. Heydari, and F. Capolino, "Millimeter-wave massive MIMO: The next wireless revolution?" *IEEE Commun. Mag.*, vol. 52, no. 9, pp. 56–62, Sep. 2014.

[4] 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.

[5] 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.

[6] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, Jr., "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, Mar. 2014.

[7] M. Egan, I. B. Collings, W. Ni, and C. K. Sung, "User scheduling for the broadcast channel using a sum-rate threshold," in *Proc. IEEE Int. Conf. Commun.*, Kyoto, Japan, Jun. 2011, pp. 1–6.

[8] Z. Shen, R. Chen, J. G. Andrews, R. W. Heath, Jr., and B. L. Evans, "Low complexity user selection algorithms for multiuser MIMO systems with block diagonalization," *IEEE Trans. Signal Process.*, vol. 54, no. 9, pp. 3658–3663, Sep. 2006.

[9] L.-N. Tran, M. Bengtsson, and B. Ottersten, "Iterative precoder design and user scheduling for block-diagonalized systems," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3726–3739, Jul. 2012.

[10] Z. Tu and R. S. Blum, "Multiuser diversity for a dirty paper approach," *IEEE Commun. Lett.*, vol. 7, no. 8, pp. 370–372, Aug. 2003.

[11] X. Zhang and J. Lee, "Low complexity MIMO scheduling with channel decomposition using capacity upperbound," *IEEE Trans. Commun.*, vol. 56, no. 6, pp. 871–876, Jun. 2008.

[12] A. Razi, D. J. Ryan, I. B. Collings, and J. Yuan, "Sum rates, rate allocation, and user scheduling for multi-user MIMO vector perturbation precoding," *IEEE Trans. Wireless Commun.*, vol. 9, no. 1, pp. 356–365, Jan. 2010.

[13] T. Yoo, N. Jindal, and A. Goldsmith, "Multi-antenna downlink channels with limited feedback and user selection," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 7, pp. 1478–1491, Sep. 2007.

[14] 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.

[15] A. Asadi and V. Mancuso, "A survey on opportunistic scheduling in wireless communications," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 4, pp. 1671–1688, Jan. 2013.

[16] P. Viswanath, D. N. C. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1277–1294, Jun. 2002.

[17] D. Gesbert, S. G. Kiani, A. Gjendemsjo, and G. E. Oien, "Adaptation, coordination, and distributed resource allocation in interference-limited wireless networks," *Proc. IEEE*, vol. 95, no. 12, pp. 2393–2409, Dec. 2007.

[18] M. Sharif and B. Hassibi, "On the capacity of MIMO broadcast channels with partial side information," *IEEE Trans. Inf. Theory*, vol. 51, no. 2, pp. 506–522, Feb. 2005.

[19] A. Silberschatz, P. B. Galvin, and G. Gagne, *Operating System Concepts*. Hoboken, NJ, USA: Wiley, 2013.

[20] C. Simon and G. Leus, "Round-robin scheduling for orthogonal beamforming with limited feedback," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2486–2496, Jun. 2011.

[21] F. Kelly, "Charging and rate control for elastic traffic," *Eur. Trans. Telecommun.*, vol. 8, no. 1, pp. 33–37, Jan./Feb. 1997.

[22] S. Bai, W. Zhang, Y. Liu, and C. Wang, "Max-min fair scheduling in OFDMA-based multi-hop WiMAX mesh networks," in *Proc. IEEE Int. Conf. Commun.*, Tokyo, Japan, Jun. 2011, pp. 1–5.

[23] J. Tang, G. Xue, and W. Zhang, ''Maximum throughput and fair bandwidth allocation in multi-channel wireless mesh networks,'' in *Proc. IEEE Int. Conf. Comput. Commun.*, Barcelona, Spain, Apr. 2006, pp. 1–10.

[24] B. Song, Y.-H. Lin, and R. L. Cruz, ''Weighted max-min fair beamforming, power control, and scheduling for a MISO downlink,'' *IEEE Trans. Wireless Commun.*, vol. 7, no. 2, pp. 464–469, Feb. 2008.

[25] A. Jalali, R. Padovani, and R. Pankaj, ''Data throughput of CDMA-HDR a high efficiency-high data rate personal communication wireless system,'' in *Proc. IEEE 51st Veh. Technol. Conf.*, Tokyo, Japan, May 2000, pp. 1854–1858.

[26] X. Chen, H. Hu, H. Wang, H.-H. Chen, and M. Guizani, ''Double proportional fair user pairing algorithm for uplink virtual MIMO systems,'' *IEEE Trans. Wireless Commun.*, vol. 7, no. 7, pp. 2425–2429, Jul. 2008.

[27] R. Fritzsche, P. Rost, and G. P. Fettweis, ''Robust rate adaptation and proportional fair scheduling with imperfect CSI,'' *IEEE Trans. Wireless Commun.*, vol. 14, no. 8, pp. 4417–4427, Aug. 2015.

[28] J. Brady, N. Behdad, and A. M. Sayeed, ''Beamspace MIMO for millimeter-wave communications: System architecture, modeling, analysis, and measurements,'' *IEEE Trans. Antennas Propag.*, vol. 61, no. 7, pp. 3814–3827, Jul. 2013.

[29] J. Brady and A. Sayeed, ''Beamspace MU-MIMO for high-density gigabit small cell access at millimeter-wave frequencies,'' in *Proc. IEEE 15th Int. Workshop Signal Process. Adv. Wireless Commun.*, Toronto, ON, Canada, Jun. 2014, pp. 80–84.

[30] T. S. Rappaport, E. Ben-Dor, J. N. Murdock, and Y. Qiao, ''38 GHz and 60 GHz angle-dependent propagation for cellular & peer-to-peer wireless communications,'' in *Proc. IEEE Int. Conf. Commun.*, Ottawa, ON, Canada, Jun. 2012, pp. 4568–4573.

[31] A. Hu, T. Lv, H. Gao, Z. Zhang, and S. Yang, ''An ESPRIT-based approach for 2-D localization of incoherently distributed sources in massive MIMO systems,'' *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 996–1011, Oct. 2014.

[32] A. Liao, Z. Gao, Y. Wu, H. Wang, and M.-S. Alouini, ''2D unitary ESPRIT based super-resolution channel estimation for millimeter-wave massive MIMO with hybrid precoding,'' *IEEE Access*, vol. 5, pp. 24747–24757, Nov. 2017.

[33] J. Mao, Z. Gao, Y. Wu, and M.-S. Alouini, ''Over-sampling codebook-based hybrid minimum sum-mean-square-error precoding for millimeter-wave 3D-MIMO,'' *IEEE Wireless Commun. Lett.*, to be published.

[34] L. Liang, W. Xu, and X. Dong, ''Low-complexity hybrid precoding in massive multiuser MIMO systems,'' *IEEE Wireless Commun. Lett.*, vol. 3, no. 6, pp. 653–656, Dec. 2014.

[35] A. Adhikary, J. Nam, J. Y. Ahn, and G. Caire, ''Joint spatial division and multiplexing—The large-scale array regime,'' *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6441–6463, Oct. 2013.

[36] Y. Xu, G. Yue, and S. Mao, ''User grouping for massive MIMO in FDD systems: New design methods and analysis,'' *IEEE Access*, vol. 2, pp. 947–959, 2014.

[37] W. Huang, Y. Huang, W. Xu, and L. Yang, ''Beam-blocked channel estimation for FDD massive MIMO with compressed feedback,'' *IEEE Access*, vol. 5, pp. 11791–11804, 2017.

[38] R. Jain, D. Chiu, and W. Hawe, ''A quantitative measure of fairness and discrimination for resource allocation in shared systems,'' Digit. Equip. Corp., Maynard, MA, USA, Tech. Rep. DEC-TR-301, 1984.

ANZHONG HU (S'13–M'17) received the Ph.D. degree from the Beijing University of Posts and Telecommunications, Beijing, China, in 2014. He is currently a Lecturer with the School of Communication Engineering, Hangzhou Dianzi University, Hangzhou, China. His current research interests include channel estimation and array processing.