Signal Processing 167 (2020) 107321

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Spatial overlapping index based joint beam selection for millimeter-wave multiuser MIMO systems

Anzhong Hu[a,*], Shaoshi Yang[b]

[a] School of Communication Engineering, Hangzhou Dianzi University, Hangzhou 310018, China
[b] School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Key Laboratory of Universal Wireless Communications, Ministry of Education, Beijing 100876, China

ARTICLE INFO

Article history:
Received 6 May 2019
Revised 17 September 2019
Accepted 25 September 2019
Available online 25 September 2019

Keywords:
MIMO systems
Beamforming
Millimeter-wave communication
Interference
5G

ABSTRACT

In this paper, beam selection for millimeter-wave multiuser multi-input multi-output (MIMO) systems that employ zero-forcing based analog-digital hybrid processing is investigated. The existing approaches have tackled neither the joint beam selection at both sides of the communication system nor the strong interference imposed by the mobile stations (MSs) which share beams with some other MSs. Hence, by first assuming full channel state information, the problem of joint beam selection is formulated and the concept of spatial overlapping index (SOI) is proposed to deal with the strong interference imposed by the MSs which share beams with some other MSs. Furthermore, by considering the realistic scenario of having channel estimation, the joint beam selection is divided and integrated into the channel estimation process, and the SOI is redefined based on the channel estimates. Finally, our analysis and simulations show that the upper bound of the sum rate increases with the number of MSs and the number of base station antennas, the proposed approach outperforms the existing approaches, and the computational complexity of the proposed approach is close to or lower than that of the existing approaches.

1. Introduction

Millimeter-wave (mm-wave) communication is capable of substantially increasing the capacity per geographical area for mobile communication systems [1]. The large frequency band in mm-wave communication is the dominant factor, and the employment of large antenna array for high array gain further facilitates mm-wave communication.

In mm-wave multi-input multi-output (MIMO) communication systems, the number of radio frequency (RF) chains is usually constrained by appropriate power consumption and cost. Consequently, hybrid precoding-and-combining that jointly processes signals in both the analog and the digital circuits are preferred [2–5]. The number of beams generated by the analog processing is limited by the number of RF chains. Therefore, the set of beams that can be generated is usually predefined, and the specific beam invoked by each transmission is selected from this set.

There have been several schemes proposed for beam selection in mm-wave multiuser MIMO systems. In [6], the beamspace channel sparsity is utilized to generate the beam sparsity mask, which is used to select beams that can achieve the highest beamspace channel gain. In [7], the authors employ the maximum signal-to-interference-plus-noise ratio (SINR) as the selection criterion. They also investigate a criterion that maximizes the capacity, which is achieved by deleting one beam that results in the lowest decrease of the capacity each time. In [8], the authors divide the users into the interference group and the non-interference group, and then select beams for the two groups separately. The authors of Chen et al. [9] suggest to avoid severe interference between mobile stations (MSs) that have correlated channels by reselecting beams for these MSs. The authors of Guo and Li [10] propose to use machine learning to select beams in a limited beam range to reduce the computational complexity. However, in these schemes, the beam selection strategy is based on the sequential greedy method, which may not find the globally optimal solution to the multi-beam selection problem subject to mutually interfering MSs. In other words, they cannot properly deal with the inter-user interference associated with multiple simultaneously served co-channel MSs. In [11], the beams are selected by solving a sparse recovery problem with the aid of out-of-band channels, but the inter-user interference is not tackled either. The author of Choi [12] employs compressive sensing to select the beams without relying on explicit channel estimation, but the beam selection requires four training phases.

* Corresponding author.
E-mail addresses: huaz@hdu.edu.cn (A. Hu), shaoshi.yang@ieee.org (S. Yang).

https://doi.org/10.1016/j.sigpro.2019.107321
0165-1684/

Against the above background, in this paper, a spatial overlapping index (SOI) based beam selection approach is proposed for mm-wave multiuser MIMO systems that use zero-forcing (ZF) based analog-digital hybrid processing. In the proposed approach, the joint beam selection at both the MS and the base station (BS) is carried out. By analysing the system's sum rate, we show that the beams that are used by each MS can be indicated with the direction-of-arrival (DOA) information. Furthermore, our sum-rate analysis demonstrates that the main interference comes from MSs that share beams with another MS. Then, the concept of SOI is proposed to indicate these interfering MSs. Based on the SOI, a joint beam selection approach is proposed to deal with the above-mentioned main interference. The major contributions of this paper are summarized as follows.

(1) A joint beam selection approach based on SOI while assuming full channel state information (CSI) is proposed. The SOI can be used to deal with the main interference imposed by the MSs which share beams with another MS. Notably, the MS-side beams and the BS-side beams are jointly selected. Thus, inter-user interference can be more efficiently mitigated.

(2) Since the assumption of full CSI is challenged by mm-wave channel characteristics and the transceiver structures, we further propose a beam selection approach based on SOI and the assumption of using channel estimation. The MS-side beams are selected in the downlink training stage, and then the uplink beamspace channel is estimated by using these beams. The SOI is redefined based on the channel estimates and is also used to deal with the strong interference imposed by MSs that share beams with another MS.

(3) The upper bound of the sum rate achieved by the proposed approach is derived, which demonstrates that the sum rate increases with the number of BS antennas and the number of MSs. Moreover, the computational complexity of the proposed approach is analyzed and compared with other approaches. The comparison shows that the proposed approach has similar or lower computational complexity in typical system configurations.

The rest of this paper is organized as follows. In Section 2 we present the system model. In Section 3, the joint beam selection approach relying on the SOI concept and the full CSI assumption is described. In Section 4 we describe the joint beam selection approach relying on channel estimation and the redefined SOI. We analyse the upper bound of the sum rate and the computational complexity of our approach in Section 5, and provide the simulation results in Section 6. Finally, our conclusions are drawn in Section 7.

Notations: Lower-case (upper-case) boldface symbols denote vectors (matrices); $(\cdot)^H$ and $\mathbb{E}\{\cdot\}$ denote the conjugate transpose and the expectation, respectively; $[\cdot]_j$ is the jth element of a vector; $[\cdot]_{:,j}$, $[\cdot]_{j,:}$, $[\cdot]_{j,k}$, $[\cdot]_{j_1:j_2,k_1:k_2}$, tr$(\cdot)$, and det$(\cdot)$ represent the jth column, the jth row, the element in the jth row and kth column, the submatrix composed of the j_1th to the j_2th row and the k_1th to the k_2th column, the trace, and the determinant of a matrix, respectively; $a \sim b$ means that a is proportional to b; $\mathbf{0}_{a,b}$ is a zero matrix with a rows and b columns; $\mathbf{I}_a$ is an identity matrix with a rows; $||\cdot||$ is the Euclidean norm of a vector; and i is the imaginary unit.

2. System model

Consider an mm-wave multiuser MIMO system, in which K MSs simultaneously transmit signals to a BS. The BS is equipped with a uniform linear array (ULA) that is composed of N_B antennas and we have $N_\mathrm{B} > K$. Each MS is equipped with a ULA of N_M antennas. The signal vector received at the BS can be expressed as

$$\mathbf{x} = \sum_{k=1}^{K} \mathbf{H}_k \mathbf{w}_k s_k + \mathbf{n} \in \mathbb{C}^{N_\mathrm{B}\times 1}, \tag{1}$$

where $\mathbf{H}_k \in \mathbb{C}^{N_\mathrm{B}\times N_\mathrm{M}}$ is the channel matrix from the kth MS to the BS; $\mathbf{w}_k \in \mathbb{C}^{N_\mathrm{M}\times 1}$ is the precoder for the kth MS; s_k is the symbol transmitted by the kth MS and is a random variable with zero mean and variance σ^2; $\mathbf{n} \in \mathbb{C}^{N_\mathrm{B}\times 1}$ is the received noise vector at the BS whose elements are independent and identically distributed (i.i.d.) complex Gaussian random variables with zero mean and variance one.

Since the number of RF chains is limited for the sake of low-cost digital signal processing in mm-wave MIMO systems, it is assumed that the BS has K RF chains and each MS has one RF chain. Additionally, each RF chain is fully connected to all the antennas. Consequently, analog-digital hybrid receiver should be employed at the BS and analog precoding should be employed at each MS. In this paper, we consider hybrid ZF receiver at the BS. More specifically, the received signal is processed as

$$\tilde{\mathbf{x}} = \mathbf{F}_\mathrm{D}\mathbf{F}_\mathrm{A}^H \mathbf{x}, \tag{2}$$

where $\mathbf{F}_\mathrm{A} \in \mathbb{C}^{N_\mathrm{B}\times K}$ is the analog receiving matrix that satisfies

$$\mathbf{F}_\mathrm{A}^H \mathbf{F}_\mathrm{A} = \mathbf{I}_K, \tag{3}$$

and $\mathbf{F}_\mathrm{D} \in \mathbb{C}^{K\times K}$ is the digital receiving matrix.

According to the ZF criterion, we have

$$\mathbf{F}_\mathrm{D} = \tilde{\mathbf{H}}^{-1}, \tag{4}$$

where $\tilde{\mathbf{H}} \in \mathbb{C}^{K\times K}$ should be of full rank and

$$[\tilde{\mathbf{H}}]_{:,k} = \mathbf{F}_\mathrm{A}^H \mathbf{H}_k \mathbf{w}_k. \tag{5}$$

Substituting (1) into (2) yields

$$\begin{aligned}\tilde{\mathbf{x}} &= \sum_{k=1}^{K} \mathbf{F}_\mathrm{D}\mathbf{F}_\mathrm{A}^H \mathbf{H}_k \mathbf{w}_k s_k + \mathbf{F}_\mathrm{D}\mathbf{F}_\mathrm{A}^H \mathbf{n} \\ &= \sum_{k=1}^{K} \tilde{\mathbf{H}}^{-1}[\tilde{\mathbf{H}}]_{:,k} s_k + \tilde{\mathbf{H}}^{-1}\mathbf{F}_\mathrm{A}^H \mathbf{n} \\ &= [s_1, s_2, \cdots, s_K]^T + \tilde{\mathbf{H}}^{-1}\mathbf{F}_\mathrm{A}^H \mathbf{n},\end{aligned} \tag{6}$$

where the second equation is derived with (4) and (5).

According to (6), the uplink sum rate of the system is

$$\begin{aligned}C &= \sum_{k=1}^{K} \log_2\left(1 + \frac{\sigma^2}{||[\tilde{\mathbf{H}}^{-1}]_{k,:}\mathbf{F}_\mathrm{A}^H||^2}\right) \\ &= \sum_{k=1}^{K} \log_2\left(1 + \frac{\sigma^2}{[\tilde{\mathbf{H}}^{-1}]_{k,:}\mathbf{F}_\mathrm{A}^H\mathbf{F}_\mathrm{A}[\tilde{\mathbf{H}}^{-1}]_{k,:}^H}\right) \\ &= \sum_{k=1}^{K} \log_2\left(1 + \frac{\sigma^2}{||[\tilde{\mathbf{H}}^{-1}]_{k,:}||^2}\right),\end{aligned} \tag{7}$$

where the last equation is derived with (3).

2.1. Channel model

As in [13,14], the clustered channel model with N_cl clusters is adopted, and each cluster is composed of N_ray propagation paths. The channel matrix is written as

$$\mathbf{H}_k = \sqrt{\frac{N_\mathrm{M}N_\mathrm{B}}{N_\mathrm{cl}N_\mathrm{ray}}} \sum_{l_\mathrm{cl}=1}^{N_\mathrm{cl}} \sum_{l_\mathrm{ray}=1}^{N_\mathrm{ray}} \alpha_{kl_\mathrm{cl}l_\mathrm{ray}} \mathbf{a}_\mathrm{B}\left(\theta^\mathrm{B}_{kl_\mathrm{cl}l_\mathrm{ray}}\right)\mathbf{a}_\mathrm{M}^H\left(\theta^\mathrm{M}_{kl_\mathrm{cl}l_\mathrm{ray}}\right), \tag{8}$$

where $\alpha_{kl_\mathrm{cl}l_\mathrm{ray}}$ is the channel gain of the l_rayth path in the l_clth cluster from the kth MS to the BS, and it is complex Gaussian distributed with zero mean and variance one; $\theta^\mathrm{M}_{kl_\mathrm{cl}l_\mathrm{ray}}$ is the direction-of-departure (DOD) of this path at the kth MS, $\theta^\mathrm{B}_{kl_\mathrm{cl}l_\mathrm{ray}}$ is the DOA

of this path at the BS. Additionally, these angles are i.i.d. random variables. The mean DOD of the l_{cl}th cluster at the kth MS is denoted as $\bar{\theta}^{\text{M}}_{kl_{\text{cl}}}$, and the mean DOA of the l_{cl}th cluster from the kth MS to the BS is denoted as $\bar{\theta}^{\text{B}}_{kl_{\text{cl}}}$, and they distribute uniformly in the range $[-\pi/2, \pi/2]$. The DODs at the kth MS in the l_{cl}th cluster have a Laplace distribution with mean $\bar{\theta}^{\text{M}}_{kl_{\text{cl}}}$ and standard deviation $\sigma^{\text{M}}_{kl_{\text{cl}}}$. The DOAs from the kth MS to the BS in the l_{cl}th cluster have a Laplace distribution with mean $\bar{\theta}^{\text{B}}_{kl_{\text{cl}}}$ and standard deviation $\sigma^{\text{B}}_{kl_{\text{cl}}}$.

In addition, $\mathbf{a}_{\text{M}}(\theta^{\text{M}}_{kl_{\text{cl}}l_{\text{ray}}}) \in \mathbb{C}^{N_{\text{M}}\times 1}$ and $\mathbf{a}_{\text{B}}(\theta^{\text{B}}_{kl_{\text{cl}}l_{\text{ray}}}) \in \mathbb{C}^{N_{\text{B}}\times 1}$ are ULA steering vectors and are defined as

$$\left[\mathbf{a}_{\text{M}}\left(\theta^{\text{M}}_{kl_{\text{cl}}l_{\text{ray}}}\right)\right]_n = \frac{1}{\sqrt{N_{\text{M}}}} e^{i2\pi d/\lambda (n-1)\sin\left(\theta^{\text{M}}_{kl_{\text{cl}}l_{\text{ray}}}\right)}, \tag{9}$$

$$\left[\mathbf{a}_{\text{B}}\left(\theta^{\text{B}}_{kl_{\text{cl}}l_{\text{ray}}}\right)\right]_m = \frac{1}{\sqrt{N_{\text{B}}}} e^{i2\pi d/\lambda (m-1)\sin\left(\theta^{\text{B}}_{kl_{\text{cl}}l_{\text{ray}}}\right)}, \tag{10}$$

where $n = 1, 2, \cdots, N_{\text{M}}$, $m = 1, 2, \cdots, N_{\text{B}}$, d is the distance between adjacent antennas and λ is the wavelength.

2.2. Problem formulation

Similar to Alkhateeb and Heath [4], Alkhateeb et al. [5], it is assumed that predefined quantization codebooks are employed for analog processing, and the equally spaced ULA steering vectors constitute the codebooks. More specifically, we have analog processing vectors $\mathbf{w}_k = \mathbf{a}_{\text{M}}(\varphi^{\text{M}}_k)$ and $[\mathbf{F}_{\text{A}}]_{:,k} = \mathbf{a}_{\text{B}}(\varphi^{\text{B}}_k)$, where

$$\varphi^{\text{M}}_k \in \mathcal{G}_{\text{M}}, \tag{11}$$

$$\mathcal{G}_{\text{M}} = \left\{\phi^{\text{M}}_n = -\frac{\pi}{2} + \frac{\pi}{2N_{\text{M}}} + \frac{(n-1)\pi}{N_{\text{M}}} | n = 1, 2, \cdots, N_{\text{M}}\right\}, \tag{12}$$

$$\varphi^{\text{B}}_k \in \mathcal{G}_{\text{B}}, \tag{13}$$

$$\mathcal{G}_{\text{B}} = \left\{\phi^{\text{B}}_m = -\frac{\pi}{2} + \frac{\pi}{2N_{\text{B}}} + \frac{(m-1)\pi}{N_{\text{B}}} | m = 1, 2, \cdots, N_{\text{B}}\right\}. \tag{14}$$

Additionally, the analog processing vectors, which are also termed as beams, should not be reused at the BS. Thus, we have the following constraint

$$\varphi^{\text{B}}_k \neq \varphi^{\text{B}}_{k'}, \forall k \neq k'. \tag{15}$$

Note that the choice of these beams will affect the system performance. According to (5), the choice of these beams in (11), (13), and (15) will determine $\tilde{\mathbf{H}}$, which will determine the sum rate. Thus, we have the beam selection problem as

$$\max_{\{\varphi^{\text{B}}_k, \varphi^{\text{M}}_k | k=1,2,\cdots,K\}} C$$
$$\text{s.t.} \quad (11), (13), (15). \tag{16}$$

Remark 1. The beam selection is performed at the BS and the MS-side beam selection results are sent from the BS to the MSs, i.e., $\mathbf{w}_k$ is sent to the kth MS. As there are multiple clusters for each MS, the investigated problem is to choose one cluster for each MS and avoids to choose two clusters that are of close angles, which corresponds to restriction (15). Moreover, the investigated problem can be easily modified for the scenario that multiple clusters should be used for each MS. For example, each MS is equipped with N RF chains, and the BS is equipped KN RF chains, Eq. (1) can be changed into

$$\mathbf{x} = \sum_{k=1}^{K} \mathbf{H}_k \sum_{n=1}^{N} \mathbf{w}_{kn} s_{kn} + \mathbf{n} \in \mathbb{C}^{N_{\text{B}}\times 1}, \tag{17}$$

s_{kn} is the symbol transmitted on the nth RF chain of the kth user, $\mathbf{w}_{kn}$ is the corresponding analog beamforming vector. $\mathbf{F}_{\text{A}} \in \mathbb{C}^{N_{\text{B}}\times KN}$, $\tilde{\mathbf{H}} \in \mathbb{C}^{KN\times KN}$, Eq. (5) can be changed into

$$[\tilde{\mathbf{H}}]_{:,(k-1)N+n} = \mathbf{F}^H_{\text{A}} \mathbf{H}_k \mathbf{w}_{kn}. \tag{18}$$

Eq. (6) can be changed into

$$\tilde{\mathbf{x}} = [s_1, s_2, \cdots, s_{KN}]^T + \tilde{\mathbf{H}}^{-1} \mathbf{F}^H_{\text{A}} \mathbf{n}, \tag{19}$$

Eq. (7) can be changed into

$$C = \sum_{k=1}^{K}\sum_{n=1}^{N} \log_2\left(1 + \frac{\sigma^2}{||[\tilde{\mathbf{H}}^{-1}]_{(k-1)N+n,:}||^2}\right), \tag{20}$$

and the problem (16) can be changed into

$$\max_{\{\varphi^{\text{B}}_{kn}, \varphi^{\text{M}}_{kn} | k=1,2,\cdots,K, n=1,2,\cdots,N\}} C$$
$$\text{s.t.} \quad (11), (13), (15). \tag{21}$$

In the following, we will select the beams by assuming full CSI. The more practical system configuration with channel estimates at the BS will be discussed in Section 4.

3. Spatial overlapping index based beam selection with full CSI

As discussed in Section 1, the existing beam selection approaches [6–8] cannot effectively tackle the inter-user interference. In this section, we first analyze the sum rate. Then, we define the concept of SOI to indicate the sharing of one beam by two MSs, and correspondingly propose a joint beam selection approach to deal with the interference imposed by the MSs that share beams.

3.1. Beam restriction

In this subsection, the beams are further restricted in addition to (11) and (13) for maximizing the sum rate.

According to the results in [15] and [16], we know that $|[\tilde{\mathbf{H}}]_{k,k}|$ achieves a large value only when $|\varphi^{\text{B}}_k - \theta^{\text{B}}_{kl_{\text{cl}}l_{\text{ray}}}| \sim 1/N_{\text{B}}$ and $|\varphi^{\text{M}}_k - \theta^{\text{M}}_{kl_{\text{cl}}l_{\text{ray}}}| \sim 1/N_{\text{M}}$. Meanwhile, with the sum rate expression in (7), we know that in an interference-free environment, i.e., $\tilde{\mathbf{H}}$ is a diagonal matrix, the sum rate is maximized only when $|[\tilde{\mathbf{H}}]_{k,k}|$ is maximized. Thus, it is beneficial to make $|\varphi^{\text{B}}_k - \theta^{\text{B}}_{kl_{\text{cl}}l_{\text{ray}}}| \sim 1/N_{\text{B}}$ and $|\varphi^{\text{M}}_k - \theta^{\text{M}}_{kl_{\text{cl}}l_{\text{ray}}}| \sim 1/N_{\text{M}}$ in an interference-free environment. Hence, based on the fact that the angles in each cluster distribute close to the mean angle, we propose the following beam selection criterion:

$$\begin{aligned}\left(\varphi^{\text{M}}_k, \varphi^{\text{B}}_k\right) \in \Big\{&\left(\varphi^{\text{M}}, \varphi^{\text{B}}\right) | \varphi^{\text{M}} \in \mathcal{G}_{\text{M}}, \varphi^{\text{B}} \in \mathcal{G}_{\text{B}}, |\varphi^{\text{M}} - \bar{\theta}^{\text{M}}_{kl_{\text{cl}}}| \\ &\leq \frac{\pi}{2N_{\text{M}}}, |\varphi^{\text{B}} - \bar{\theta}^{\text{B}}_{kl_{\text{cl}}}| \leq \frac{\pi}{2N_{\text{B}}}, \\ &l_{\text{cl}} \in \{1, 2, \cdots, N_{\text{cl}}\}\Big\}.\end{aligned} \tag{22}$$

Note that the constraints (11) and (13) in (16) are included in this criterion. Moreover, the two variables $\varphi^{\text{B}}_k, \varphi^{\text{M}}_k$ are paired, which means only one of them needs to be optimized. Then, the optimization problem (16) is reformulated as the following joint beam selection problem:

$$\max_{\{\varphi^{\text{B}}_k | k=1,2,\cdots,K\}} C$$
$$\text{s.t.} \quad (15), (22). \tag{23}$$

3.2. Spatial overlapping index with full CSI

The problem (23) is not easy to solve. The brute-force search can take into account all interference, but will cost N^K_{cl} computation steps. The approaches in [6–8] consider neither the interference imposed by the MSs that share beams nor the joint beam

selection at both the MS and the BS. Here, we propose to analyze the sum rate expression (7) to find the main interference imposed by the MSs sharing beams.

According to the results in [15] and [16], with $|\theta^{\mathrm{B}}_{kl_{\mathrm{cl}}l_{\mathrm{ray}}} - \varphi^{\mathrm{B}}_{k'}| > \pi/2/N_{\mathrm{B}}, \forall l_{\mathrm{cl}}, l_{\mathrm{ray}}, \forall k \neq k'$, $\tilde{\mathbf{H}}$ can be approximated to be diagonal, which means $||[\tilde{\mathbf{H}}^{-1}]_{k,:}||^2 \approx [\tilde{\mathbf{H}}]^{-2}_{k,k}$. According to the sum rate expression (7), this means that there is almost no inter-user interference. Thus, we can infer there is almost no interference when $|\theta^{\mathrm{B}}_{kl_{\mathrm{cl}}l_{\mathrm{ray}}} - \varphi^{\mathrm{B}}_{k'}| > \pi/2/N_{\mathrm{B}}, \forall l_{\mathrm{cl}}, l_{\mathrm{ray}}, \forall k \neq k'$; and vice versa. Since the multipath DOAs in each cluster distribute close to the mean DOA of this cluster, we can further simplify the almost no interference condition into

$$|\bar{\theta}^{\mathrm{B}}_{kl_{\mathrm{cl}}} - \varphi^{\mathrm{B}}_{k'}| > \frac{\pi}{2N_{\mathrm{B}}}, \forall l_{\mathrm{cl}}, \forall k \neq k'. \tag{24}$$

Thus, we propose to use SOI for MSs

$$s_{k,l_{\mathrm{cl}},m} = \begin{cases} 1, & |\bar{\theta}^{\mathrm{B}}_{kl_{\mathrm{cl}}} - \phi^{\mathrm{B}}_{m}| < \frac{\pi}{2N_{\mathrm{B}}}, \\ 0, & \text{otherwise}, \end{cases} \tag{25}$$

to measure the interference, where $\phi^{\mathrm{B}}_m \in \mathcal{G}_{\mathrm{B}}$, cf. (14). If the beam of the kth MS is temporarily selected as $\varphi^{\mathrm{B}}_k = \phi^{\mathrm{B}}_m$, and $s_{k',l'_{\mathrm{cl}},m} = 1$ is satisfied, for $k \neq k'$ and specific values of l'_{cl}, then (24) cannot be satisfied, which means there is strong interference between the kth MS and the k'th MS. Consequently, the interference from the k'th MS should be taken into consideration in the selection of ϕ^{B}_m for the kth MS. Thus, we propose to sequentially select the beams for the MSs in a greedy way and take such strong interference into consideration at the same time.

The corresponding SOI based beam selection algorithm for (23) is shown in Algorithm 1. In the first step, we calculate the SOI defined in (25) based on the full CSI. Then, we sequentially select the beams for the MSs in a greedy way, i.e., we select the beam for the kth MS in the kth iteration, as shown in steps 2–16. For the beam selection of the kth MS, all the N_{cl} clusters are tried, i.e., the temporal sum rate corresponding to the selection of the l_{cl}th cluster is calculated, which is denoted as $C_{k,l_{\mathrm{cl}}}$, as shown in steps 3–13. Moreover, the cluster corresponding the maximal $C_{k,l_{\mathrm{cl}}}$ is selected and denoted as $l^{\star}_k$, and the corresponding beamforming angles are selected, as shown in steps 14–15.

Algorithm 1 SOI Based beam selection algorithm with full CSI.

Input: $\mathbf{H}_k, k = 1, 2, \cdots, K, \bar{\theta}^{\mathrm{B}}_{kl_{\mathrm{cl}}}, \bar{\theta}^{\mathrm{M}}_{kl_{\mathrm{cl}}}, k = 1, 2, \cdots, K, l_{\mathrm{cl}} = 1, 2, \cdots, N_{\mathrm{cl}}$.
Output: $\varphi^{\mathrm{M}}_k, \varphi^{\mathrm{B}}_k, k = 1, 2, \cdots, K$.
1: calculate $s_{k,l_{\mathrm{cl}},m}, \forall k, l_{\mathrm{cl}}, m$ with (25)
2: **for** $k = 1 \to K$
3: **for** $l_{\mathrm{cl}} = 1 \to N_{\mathrm{cl}}$
4: set $\varphi^{\mathrm{M}}_k, \varphi^{\mathrm{B}}_k$ as (22) with l_{cl}
5: calculate m with (14), $k_1 \leftarrow 0$
6: **for** $k' = k+1 \to K$
7: **if** $\sum_{l'_{\mathrm{cl}}=1}^{N_{\mathrm{cl}}} s_{k',l'_{\mathrm{cl}},m} > 0$
8: $k_1 \leftarrow k_1 + 1$, $f_{k,l_{\mathrm{cl}},k_1} \leftarrow k'$
9: set $\varphi^{\mathrm{B}}_{f_{k,l_{\mathrm{cl}},k_1}}, \varphi^{\mathrm{M}}_{f_{k,l_{\mathrm{cl}},k_1}}$ as (22) with $\varphi^{\mathrm{B}}_{f_{k,l_{\mathrm{cl}},k_1}} \neq \varphi^{\mathrm{B}}_k$
10: **end if**
11: **end for**
12: calculate $C_{k,l_{\mathrm{cl}}}$ with (26)
13: **end for**
14: $l^{\star}_k \leftarrow \arg\max_{l_{\mathrm{cl}}} C_{k,l_{\mathrm{cl}}}$
15: set $\varphi^{\mathrm{M}}_k, \varphi^{\mathrm{B}}_k$ as (22) with $l_{\mathrm{cl}} = l^{\star}_k$
16: **end for**

For the calculation of the temporal sum rate $C_{k,l_{\mathrm{cl}}}$, the beamforming angles corresponding to the l_{cl}th cluster of the kth MS should be calculated, and the corresponding beam index should also be calculated, as shown in steps 4–5. Additionally, the MSs with indices of 1, 2, $\cdots$, k should be taken into consideration, and the MSs with nonzero SOI should also be taken into consideration, i.e., k'th MS should be taken into consideration if $\sum_{l'_{\mathrm{cl}}=1}^{N_{\mathrm{cl}}} s_{k',l'_{\mathrm{cl}},m} > 0$, where $k' > k$ and m is the beam index for the l_{cl}th cluster of the kth MS. For these MSs with indices larger than k, the beams for them have not been selected yet, thus we temporally select one beam for them, and the selected should obey the basic rule in (22) and be not the same as the selected beam for the kth MS. These calculations correspond to steps 6–11.

If there are $S_{k,l_{\mathrm{cl}}} - k$ MSs taken into consideration based on the SOI, i.e., $S_{k,l_{\mathrm{cl}}} - k$ MSs with indices in $\{k+1, k+2, \cdots, K\}$ and satisfy $\sum_{l'_{\mathrm{cl}}=1}^{N_{\mathrm{cl}}} s_{k',l'_{\mathrm{cl}},m} > 0$, there will be $S_{k,l_{\mathrm{cl}}}$ MSs taken into consideration. Then, the temporal sum rate can be calculated as

$$C_{k,l_{\mathrm{cl}}} = \sum_{\tilde{k}=1}^{k} \log_2\left(1 + \frac{\sigma^2}{||[\check{\mathbf{H}}^{-1}_{k,l_{\mathrm{cl}}}]_{\tilde{k},:}||^2}\right), \tag{26}$$

where $\check{\mathbf{H}}_{k,l_{\mathrm{cl}}} \in \mathbb{C}^{S_{k,l_{\mathrm{cl}}} \times S_{k,l_{\mathrm{cl}}}}$ is the beamspace channel matrxi with $[\check{\mathbf{H}}_{k,l_{\mathrm{cl}}}]_{k_1,k_2} = \mathbf{a}^H_{\mathrm{B}}(\varphi^{\mathrm{B}}_{f_{k,l_{\mathrm{cl}},k_1}})\mathbf{H}_k\mathbf{a}_{\mathrm{M}}(\varphi^{\mathrm{M}}_{f_{k,l_{\mathrm{cl}},k_2}})$, where $f_{k,l_{\mathrm{cl}},k_1}, k_1 = 1, 2, \cdots, S_{k,l_{\mathrm{cl}}}$ are the indices of the MSs considered in this stage and $f_{k,l_{\mathrm{cl}},k_1} = k_1, \forall k_1 \leq k$. This calculation corresponds to step 12.

Remark 2. In step 9, there are N_{cl} choices of the combination $\varphi^{\mathrm{B}}_{f_{k,l_{\mathrm{cl}},k_1}}, \varphi^{\mathrm{M}}_{f_{k,l_{\mathrm{cl}},k_1}}$, and we randomly choose one such that $\varphi^{\mathrm{B}}_{f_{k,l_{\mathrm{cl}},k_1}} \neq \varphi^{\mathrm{B}}_k$.

4. Spatial overlapping index based beam selection with channel estimation

In practical mm-wave systems, the number of clusters KN_{cl} may be large, and there are N_{ray} multipaths in each cluster. Moreover, there are multiple antennas at both the BS and the MSs while the number of RF chains is much less than the number of antennas at each transceiver. These facts make it difficult to acquire full CSI at the BS. In addition, the BS should transmit the beam selection results to the MSs for their beamforming. For the proposed SOI based algorithm, full CSI $\mathbf{H}_k, k = 1, 2, \cdots, K$ and $\bar{\theta}^{\mathrm{B}}_{kl_{\mathrm{cl}}}, \bar{\theta}^{\mathrm{M}}_{kl_{\mathrm{cl}}}, k = 1, 2, \cdots, K, l_{\mathrm{cl}} = 1, 2, \cdots, N_{\mathrm{cl}}$ are necessary, which makes the proposed approach infeasible for mm-wave systems with channel estimation that can only acquire the partial CSI. In this section, we propose to select beams and estimate channels in the channel training stages. First, we propose one beam selection approach for the MSs in the downlink training stage. Then, we propose another SOI based beam selection algorithm which is based on the channel estimation for the BS in the uplink training stage.

4.1. MS-Side beam selection

In this training stage, the BS transmits signals to the MSs and the received signal vector at the kth MS is denoted as

$$\mathbf{y}_k = \sqrt{\frac{\rho_{\mathrm{B}}}{K}}\mathbf{H}^H_k \sum_{k'=1}^{K} [\tilde{\mathbf{F}}_{\mathrm{A}}]_{:,k} + \mathbf{z}_k \in \mathbb{C}^{N_{\mathrm{M}} \times 1}, \tag{27}$$

where ρ_{B} is the transmission power at the BS, $\tilde{\mathbf{F}}_{\mathrm{A}} \in \mathbb{C}^{N_{\mathrm{B}} \times K}$ is the analog beamforming matrix at the BS, and $\mathbf{z}_k \in \mathbb{C}^{N_{\mathrm{M}} \times 1}$ is the received noise vector at the kth MS whose elements are i.i.d. complex Gaussian random variables with zero mean and variance one. Moreover, the beamforming vectors are in the form of the steering vectors with uniform angles, i.e.,

$$[\tilde{\mathbf{F}}_{\mathrm{A}}]_{:,k} = \mathbf{a}_{\mathrm{B}}(\tilde{\phi}^{\mathrm{B}}_k), \tag{28}$$

where

$$\tilde{\phi}_k^{\mathrm{B}} = -\frac{\pi}{2} + \frac{\pi}{2N_{\mathrm{B}}} + \frac{\lfloor \frac{k-1}{K} N_{\mathrm{B}} \rfloor \pi}{N_{\mathrm{B}}}. \tag{29}$$

The kth MS sequentially employs $\tilde{\mathbf{w}}_n = \mathbf{a}_{\mathrm{M}}(\phi_n^{\mathrm{M}}) \in \mathbb{C}^{N_{\mathrm{M}}\times 1}$, $n = 1, 2, \cdots, N_{\mathrm{M}}$, cf. (12), as the receiving vector and yields

$$y_{kn} = \tilde{\mathbf{w}}_n^H \mathbf{y}_k. \tag{30}$$

Then, the MS-side beam is selected to maximize the received signal power, i.e.,

$$\mathbf{w}_k = \mathbf{a}_{\mathrm{M}}(\phi_{n_k^\star}^{\mathrm{M}}), \tag{31}$$

where $n_k^\star = \arg\max_n |y_{kn}|^2$. As can be seen, the downlink training and MS-side beam selection cost N_{M} downlink training slots.

4.2. Spatial overlapping index with channel estimation

In this training stage, the K MSs sequentially transmits to the BS. For the kth MS, the received signal vector at the BS is denoted as

$$\mathbf{r}_k = \sqrt{\rho_{\mathrm{M}}}\mathbf{H}_k\mathbf{w}_k + \mathbf{v}_k \in \mathbb{C}^{N_{\mathrm{B}}\times 1}, \tag{32}$$

where ρ_{M} is the transmission power at the MS, $\mathbf{v}_k \in \mathbb{C}^{N_{\mathrm{B}}\times 1}$ is the received noise vector at the BS whose elements are i.i.d. complex Gaussian random variables with zero mean and variance one.

The BS employs $\hat{\mathbf{F}}_{\mathrm{A}} \in \mathbb{C}^{N_{\mathrm{B}}\times N_{\mathrm{P}}}$ as the receiving matrix, where

$$[\hat{\mathbf{F}}_{\mathrm{A}}]_{:,p} = \mathbf{a}_{\mathrm{B}}(\hat{\phi}_p^{\mathrm{B}}), \tag{33}$$

and

$$\hat{\phi}_p^{\mathrm{B}} = -\frac{\pi}{2} + \frac{\pi}{2N_{\mathrm{B}}} + \frac{\lfloor \frac{p-1}{N_{\mathrm{P}}} N_{\mathrm{B}} \rfloor \pi}{N_{\mathrm{B}}}, \tag{34}$$

$p = 1, 2 \cdots, N_{\mathrm{P}}$. Note that the BS can only employ K beamforming vectors in one time slot, the employment of N_{P} beamforming vectors means that we need $\lceil N_{\mathrm{P}}/K \rceil$ pilot slots for each MS, and the total number of pilot slots for uplink training and beam selection is thus $\lceil N_{\mathrm{P}}/K \rceil K$. Then, the received signal is processed into

$$\begin{aligned}\hat{\mathbf{r}}_k &= \hat{\mathbf{F}}_{\mathrm{A}}^H \mathbf{r}_k \\ &= \sqrt{\rho_{\mathrm{M}}}\hat{\mathbf{F}}_{\mathrm{A}}^H\mathbf{H}_k\mathbf{w}_k + \hat{\mathbf{F}}_{\mathrm{A}}^H\mathbf{v}_k \in \mathbb{C}^{K\times 1},\end{aligned} \tag{35}$$

where the second equation is derived by substituting (32) into the first equation.

4.2.1. Definition of SOI

Substituting (8) into (35) results into

$$\begin{aligned}\hat{\mathbf{r}}_k &= \sqrt{\frac{N_{\mathrm{M}}N_{\mathrm{B}}\rho_{\mathrm{M}}}{N_{\mathrm{cl}}N_{\mathrm{ray}}}} \sum_{l_{\mathrm{cl}}=1}^{N_{\mathrm{cl}}} \sum_{l_{\mathrm{ray}}=1}^{N_{\mathrm{ray}}} \alpha_{kl_{\mathrm{cl}}l_{\mathrm{ray}}} \mathbf{a}_{\mathrm{M}}^H\left(\theta_{kl_{\mathrm{cl}}l_{\mathrm{ray}}}^{\mathrm{M}}\right)\mathbf{w}_k\hat{\mathbf{F}}_{\mathrm{A}}^H\mathbf{a}_{\mathrm{B}}\left(\theta_{kl_{\mathrm{cl}}l_{\mathrm{ray}}}^{\mathrm{B}}\right) + \hat{\mathbf{F}}_{\mathrm{A}}^H\mathbf{v}_k \\ &\approx \sum_{l_{\mathrm{cl}}=1}^{N_{\mathrm{cl}}} \beta_{k,l_{\mathrm{cl}}}\hat{\mathbf{F}}_{\mathrm{A}}^H\mathbf{a}_{\mathrm{B}}\left(\bar{\theta}_{kl_{\mathrm{cl}}}^{\mathrm{B}}\right) + \hat{\mathbf{F}}_{\mathrm{A}}^H\mathbf{v}_k,\end{aligned} \tag{36}$$

where the approximation is based on the fact that the DOAs of the multipaths in each cluster are close to the mean DOA of each cluster, and

$$\beta_{k,l_{\mathrm{cl}}} = \sqrt{\frac{N_{\mathrm{M}}N_{\mathrm{B}}\rho_{\mathrm{M}}}{N_{\mathrm{cl}}N_{\mathrm{ray}}}} \sum_{l_{\mathrm{ray}}=1}^{N_{\mathrm{ray}}} \alpha_{kl_{\mathrm{cl}}l_{\mathrm{ray}}} \mathbf{a}_{\mathrm{M}}^H\left(\theta_{kl_{\mathrm{cl}}l_{\mathrm{ray}}}^{\mathrm{M}}\right)\mathbf{w}_k \tag{37}$$

is the effective channel gain. Based on the results in [15] and [16], we have

$$[\hat{\mathbf{F}}_{\mathrm{A}}]_{:,p}^H\mathbf{a}_{\mathrm{B}}\left(\bar{\theta}_{kl_{\mathrm{cl}}}^{\mathrm{B}}\right) \approx 0, \forall |\hat{\phi}_p^{\mathrm{B}} - \bar{\theta}_{kl_{\mathrm{cl}}}^{\mathrm{B}}| > \frac{\pi}{2N_{\mathrm{B}}}, \tag{38}$$

and the approximation error is negligible for a wider angle range when N_{B} increases. Thus, with a large number of antennas at the BS and naturally separated clusters, the received signal can be further approximated as

$$\begin{aligned}[\hat{\mathbf{r}}_k]_p &\approx \beta_{k,\acute{l}_{k,p}}[\hat{\mathbf{F}}_{\mathrm{A}}]_{:,p}^H\mathbf{a}_{\mathrm{B}}\left(\bar{\theta}_{k\acute{l}_{k,p}}^{\mathrm{B}}\right) + [\hat{\mathbf{F}}_{\mathrm{A}}]_{:,p}^H\mathbf{v}_k \\ &= \beta_{k,\acute{l}_{k,p}}\mathbf{a}_{\mathrm{B}}^H\left(\hat{\phi}_p^{\mathrm{B}}\right)\mathbf{a}_{\mathrm{B}}\left(\bar{\theta}_{k\acute{l}_{k,p}}^{\mathrm{B}}\right) + \mathbf{a}_{\mathrm{B}}^H\left(\hat{\phi}_p^{\mathrm{B}}\right)\mathbf{v}_k,\end{aligned} \tag{39}$$

where $\acute{l}_{k,p} \in \mathcal{L}_{k,p} = \{1, 2, \cdots, N_{\mathrm{cl}}\} \cap \{l_{\mathrm{cl}} | |\hat{\phi}_p^{\mathrm{B}} - \bar{\theta}_{kl_{\mathrm{cl}}}^{\mathrm{B}}| \le \pi/2/N_{\mathrm{B}}\}$, and the second equation is derived by substituting (28) into the first equation. Note that (39) corresponds to the case of $\mathcal{L}_{k,p} \neq \emptyset$. When $\mathcal{L}_{k,p} = \emptyset$, the part $\beta_{k,\acute{l}_{k,p}}\mathbf{a}_{\mathrm{B}}^H(\hat{\phi}_p^{\mathrm{B}})\mathbf{a}_{\mathrm{B}}(\bar{\theta}_{k\acute{l}_{k,p}}^{\mathrm{B}})$ in (39) does not exist, and the received signal can be approximated as

$$[\hat{\mathbf{r}}_k]_p \approx \mathbf{a}_{\mathrm{B}}^H\left(\hat{\phi}_p^{\mathrm{B}}\right)\mathbf{v}_k. \tag{40}$$

For the case of $\mathcal{L}_{k,p} \neq \emptyset$, it means that there are strong clusters in the direction of the pth beam, cf. (34). For the case of $\mathcal{L}_{k,p} = \emptyset$, it means that there is no strong cluster in the direction of the pth beam. Thus, we propose to define the SOI as

$$\acute{s}_{k,p} = \begin{cases} 1, & |[\hat{\mathbf{r}}_k]_p| > \eta, \\ 0, & |[\hat{\mathbf{r}}_k]_p| \le \eta, \end{cases} \tag{41}$$

where

$$\eta = \frac{1}{2}\mathbb{E}\{|\beta_{k,l_{\mathrm{cl}}}|\} \tag{42}$$

is the proposed threshold. Note that the SOI here is defined for indicating whether $\mathcal{L}_{k,p} \neq \emptyset$ or $\mathcal{L}_{k,p} = \emptyset$. Comparing (39) and (40), we can see that the signal part $\beta_{k,\acute{l}_{k,p}}\mathbf{a}_{\mathrm{B}}^H(\hat{\phi}_p^{\mathrm{B}})\mathbf{a}_{\mathrm{B}}(\bar{\theta}_{k\acute{l}_{k,p}}^{\mathrm{B}})$ is in (39) but not in (40). Thus, the detection of the existence or nonexistence of the signal part (i.e., $\mathcal{L}_{k,p} \neq \emptyset$ or $\mathcal{L}_{k,p} = \emptyset$) is to compare as (41). Because $\mathbf{a}_{\mathrm{B}}^H(\hat{\phi}_p^{\mathrm{B}})\mathbf{v}_k$ is Gaussian distributed with zero mean, the optimal detection threshold is $|\beta_{k,\acute{l}_{k,p}}\mathbf{a}_{\mathrm{B}}^H(\hat{\phi}_p^{\mathrm{B}})\mathbf{a}_{\mathrm{B}}(\bar{\theta}_{k\acute{l}_{k,p}}^{\mathrm{B}})|/2$. Since the DOA $\bar{\theta}_{k\acute{l}_{k,p}}^{\mathrm{B}}$ and the effective channel gain $\beta_{k,\acute{l}_{k,p}}$ are random, we propose to replace the optimal detection threshold with an approximated value. More specifically, we approximate $|\mathbf{a}_{\mathrm{B}}^H(\hat{\phi}_p^{\mathrm{B}})\mathbf{a}_{\mathrm{B}}(\bar{\theta}_{k\acute{l}_{k,p}}^{\mathrm{B}})|$ as one, and approximate $|\beta_{k,\acute{l}_{k,p}}|$ as its expectation $\mathbb{E}\{|\beta_{k,l_{\mathrm{cl}}}|\}$. Then, we get the proposed detection threshold in (42).

Moreover, (42) can be further simplified as

$$\eta = \sqrt{\frac{N_{\mathrm{M}}N_{\mathrm{B}}\rho_{\mathrm{M}}}{4N_{\mathrm{cl}}N_{\mathrm{ray}}}}\mathbb{E}\left\{\left|\sum_{l_{\mathrm{ray}}=1}^{N_{\mathrm{ray}}} \alpha_{kl_{\mathrm{cl}}l_{\mathrm{ray}}}\mathbf{a}_{\mathrm{M}}^H(\theta_{kl_{\mathrm{cl}}l_{\mathrm{ray}}}^{\mathrm{M}})\mathbf{w}_k\right|\right\} \tag{43}$$

$$\approx \sqrt{\frac{N_{\mathrm{M}}N_{\mathrm{B}}\rho_{\mathrm{M}}}{4N_{\mathrm{cl}}N_{\mathrm{ray}}}}\mathbb{E}\left\{\left|\sum_{l_{\mathrm{ray}}=1}^{N_{\mathrm{ray}}} \alpha_{kl_{\mathrm{cl}}l_{\mathrm{ray}}}\right|\right\} \tag{44}$$

$$= \sqrt{\frac{N_{\mathrm{M}}N_{\mathrm{B}}\rho_{\mathrm{M}}}{4N_{\mathrm{cl}}}}\Gamma(1.5), \tag{45}$$

where (43) is derived by substituting (37), the approximation (44) is based on $\mathbf{a}_{\mathrm{M}}^H(\theta_{kl_{\mathrm{cl}}l_{\mathrm{ray}}}^{\mathrm{M}})\mathbf{w}_k \approx 1$, (45) is based on the fact that

$$\left|\sum_{l_{\mathrm{ray}}=1}^{N_{\mathrm{ray}}} \alpha_{kl_{\mathrm{cl}}l_{\mathrm{ray}}}\right|^2 = \left|\sum_{l_{\mathrm{ray}}=1}^{N_{\mathrm{ray}}} \alpha_{kl_{\mathrm{cl}}l_{\mathrm{ray}}}^{\mathrm{re}}\right|^2 + \left|\sum_{l_{\mathrm{ray}}=1}^{N_{\mathrm{ray}}} \alpha_{kl_{\mathrm{cl}}l_{\mathrm{ray}}}^{\mathrm{im}}\right|^2, \tag{46}$$

where $\alpha_{kl_{\mathrm{cl}}l_{\mathrm{ray}}}^{\mathrm{re}}$ and $\alpha_{kl_{\mathrm{cl}}l_{\mathrm{ray}}}^{\mathrm{im}}$ are the real and the imaginary parts of $\alpha_{kl_{\mathrm{cl}}l_{\mathrm{ray}}}$ and are i.i.d. Gaussian variables with zero mean and variance 0.5, cf. the definition of $\alpha_{kl_{\mathrm{cl}}l_{\mathrm{ray}}}$ below (8). Thus, $\sum_{l_{\mathrm{ray}}=1}^{N_{\mathrm{ray}}} \alpha_{kl_{\mathrm{cl}}l_{\mathrm{ray}}}^{\mathrm{re}}$ and $\sum_{l_{\mathrm{ray}}=1}^{N_{\mathrm{ray}}} \alpha_{kl_{\mathrm{cl}}l_{\mathrm{ray}}}^{\mathrm{im}}$ are i.i.d. Gaussian variables with zero mean and variance $0.5N_{\mathrm{ray}}$, which means $2/N_{\mathrm{ray}}|\sum_{l_{\mathrm{ray}}=1}^{N_{\mathrm{ray}}} \alpha_{kl_{\mathrm{cl}}l_{\mathrm{ray}}}|^2$ is distributed according to the Chi-square distribution with degrees of

freedom being 2. As can be seen, $\sqrt{2/N_{\text{ray}}}|\sum_{l_{\text{ray}}=1}^{N_{\text{ray}}}\alpha_{kl_{\text{cl}}l_{\text{ray}}}|$ is distributed according to the Chi distribution with degrees of freedom being 2, and correspondingly we have [17]

$$\mathbb{E}\left\{\sqrt{\frac{2}{N_{\text{ray}}}}|\sum_{l_{\text{ray}}=1}^{N_{\text{ray}}}\alpha_{kl_{\text{cl}}l_{\text{ray}}}|\right\}=\sqrt{2}\Gamma(1.5). \tag{47}$$

4.2.2. Channel estimation

Furthermore, as $\acute{\mathbf{F}}_{\text{A}}$ is employed for channel estimation, the beams for data transmission should be selected from it, i.e.,

$$\mathbf{F}_{\text{A}}=\acute{\mathbf{F}}_{\text{A}}\mathbf{P}, \tag{48}$$

where $\mathbf{P}\in\mathbb{R}^{N_{\text{P}}\times K}$ is a binary matrix with $||[\mathbf{P}]_{:,k}||=1,\forall k$ and $[\mathbf{P}]_{p,k}\in\{0,1\}$, $\forall k, p$. Additionally, $\mathbf{P}$ depends on the beam selection. With beam selection, the received pilot in (35) can be processed into

$$\frac{1}{\sqrt{\rho_{\text{M}}}}\mathbf{P}^H\acute{\mathbf{r}}_k=\mathbf{P}^H\acute{\mathbf{F}}_{\text{A}}^H\mathbf{H}_k\mathbf{w}_k+\frac{1}{\sqrt{\rho_{\text{M}}}}\mathbf{P}^H\acute{\mathbf{F}}_{\text{A}}^H\mathbf{v}_k \tag{49}$$

$$=\mathbf{F}_{\text{A}}^H\mathbf{H}_k\mathbf{w}_k+\frac{1}{\sqrt{\rho_{\text{M}}}}\mathbf{F}_{\text{A}}^H\mathbf{v}_k \tag{50}$$

$$=[\tilde{\mathbf{H}}]_{:,k}+\frac{1}{\sqrt{\rho_{\text{M}}}}\mathbf{F}_{\text{A}}^H\mathbf{v}_k, \tag{51}$$

where the second equation is derived by substituting (48) into the first equation, and the last equation is derived by substituting (5) into the second equation. Thus, with beam selection, we propose to estimate the beamspace channel of the kth MS $[\tilde{\mathbf{H}}]_{:,k}$ as

$$[\hat{\tilde{\mathbf{H}}}]_{:,k}=\frac{1}{\sqrt{\rho_{\text{M}}}}\mathbf{P}^H\acute{\mathbf{r}}_k, \tag{52}$$

where $\hat{\tilde{\mathbf{H}}}\in\mathbb{C}^{K\times K}$ is the estimate of $\tilde{\mathbf{H}}$.

4.2.3. Proposed beam selection with SOI and channel estimation

As can be seen, the channel estimation relies on the beam selection result $\mathbf{P}$, cf. (52), and the beam selection relies on the channel estimation. Here, we propose to try all beam choices, i.e., all realizations of $\mathbf{P}$, with the SOI used to take into consideration of interference, and estimate the corresponding beamspace channels as (52); additionally, the sum rates corresponding to these choices are compared to find the best beam. The sketch of the SOI based beam selection algorithm is as Algorithm 2. In the first step, we calculate the SOI according to (41). Then, we sequentially select beams for the MSs in a greedy manner, i.e., we select a beam for the kth MS in the kth iteration, as steps 2–18. For the kth MS, the temporal sum rate corresponding to the selection of the pth beam is calculated, which is denoted as $\acute{C}_{k,p}$ here, and the beam corresponding to the maximal $\acute{C}_{k,p}$ is selected, as steps 3–17. For the pth beam, we check whether it has been selected by other MSs, with $||[\mathbf{P}]_{p,:}||<1$ meaning that it has not been selected, and $||[\mathbf{P}]_{p,:}||\geq 1$ meaning that it has been selected, as step 4. If the pth beam has been selected by other MSs, $\acute{C}_{k,p}$ is set as zero, as in the initialization; otherwise, $\acute{C}_{k,p}$ is calculated, as steps 5–14. With the pth beam being temporally selected, we set $[\mathbf{P}]_{p,k}=1$, as step 5. Then, we search for MSs with SOI being one for the pth beam, i.e., we search for the k'th MS with $k'\in\{k+1,k+2,\cdots,K\}$ and $\acute{s}_{k',p}>0$, as step 7. For the k'th MS, we temporally select one beam which has not been selected, i.e., the p'th beam with $\sum_{k''=1}^{k'-1}[\mathbf{P}]_{p',k''}<1$, as step 9. Suppose there are $\acute{S}_{k,p}-k$ MSs taken into consideration based on the SOI, i.e., MSs with $\acute{s}_{k',p}=1, k'>k$, then the number of MSs taken into consideration is $\acute{S}_{k,p}$. We denote the indices of these MSs as $\acute{f}_{k,p,k_1}, k_1=1,2,\cdots,\acute{S}_{k,p}$, where $\acute{f}_{k,p,k_1}=k_1,\forall k_1\leq k$. For $k_1>k$, the index k_1 is used to denote the considered MS, i.e., k_1 denotes the MS with k' that satisfies $\acute{s}_{k',p}>0$, and we have $\acute{f}_{k,p,k_1}=k'$; additionally, k_1 increases continuously, as steps 8–9. Correspondingly, we construct the beamspace channel matrix $\acute{\mathbf{H}}_{k,p}\in\mathbb{C}^{\acute{S}_{k,p}\times\acute{S}_{k,p}}$ as

$$[\acute{\mathbf{H}}_{k,p}]_{k_1,k_2}=[\hat{\tilde{\mathbf{H}}}]_{\acute{f}_{k,p,k_1},\acute{f}_{k,p,k_2}}, \tag{53}$$

as step 12. Note that $[\hat{\tilde{\mathbf{H}}}]_{\acute{f}_{k,p,k_1},\acute{f}_{k,p,k_2}}$ is calculated with (52), which means the channel estimation is based on the beam selection. Then, the temporal sum rate is calculated as

$$\acute{C}_{k,p}=\sum_{\tilde{k}=1}^{k}\log_2\left(1+\frac{\sigma^2}{||[\acute{\mathbf{H}}_{k,p}^{-1}]_{\tilde{k},:}||^2}\right), \tag{54}$$

as step 13. Then, all the temporal beam selections are removed, as step 14. Finally, the beam that results into the maximum sum rate is selected for the kth MS, as step 17, which means the beam selection relies on the channel estimation.

Algorithm 2 SOI Based beam selection algorithm with channel estimation.

Input: $\acute{\mathbf{r}}_k, k=1,2,\cdots,K$.
Output: $\mathbf{P}$.
Initialization: $\mathbf{P}=\mathbf{0}_{N_{\text{P}},K}$, $\acute{C}_{k,p}=0, k=1,2,\cdots,K, p=1,2,\cdots,N_{\text{P}}$.
1: calculate $\acute{s}_{k,p}, k=1,2\cdots,K, p=1,2,\cdots,N_{\text{P}}$ with (41)
2: **for** $k=1\rightarrow K$
3: **for** $p=1\rightarrow N_{\text{P}}$
4: **if** $|||[\mathbf{P}]_{p,:}||<1$
5: $[\mathbf{P}]_{p,k}\leftarrow 1, k_1\leftarrow 0$
6: **for** $k'=k+1\rightarrow K$
7: **if** $\acute{s}_{k',p}>0$
8: $k_1\leftarrow k_1+1$
9: $\acute{f}_{k,p,k_1}\leftarrow k'$, $[\mathbf{P}]_{p',k'}\leftarrow 1$with $\sum_{k''=1}^{k'-1}[\mathbf{P}]_{p',k''}<1$
10: **end if**
11: **end for**
12: calculate $\acute{\mathbf{H}}_{k,p}$ with (53)
13: calculate $\acute{C}_{k,p}$ with (54)
14: $[\mathbf{P}]_{:,k:K}\leftarrow\mathbf{0}_{N_{\text{P}},K-k+1}$
15: **end if**
16: **end for**
17: $p_k^\star\leftarrow\arg\max_p\acute{C}_{k,p}$, $[\mathbf{P}]_{p_k^\star,k}\leftarrow 1$
18: **end for**

Remark 3. In order to improve channel estimation precision, more uplink training slots are needed, and it is expected that the beam selection performance will benefit from the added training slots.

5. Performance and complexity analysis

As the relation between the sum rate expression and the beams is not clear, in this section, the upper bound of the ergodic sum rate of the proposed approach and its asymptotic performance will be analyzed. Additionally, the computational complexity of the proposed approach will be analyzed and compared with that of other approaches.

5.1. Sum rate upper bound analysis

Lemma 1. *There exists an upper bound of the sum rate*

$$\check{C}=\sum_{k=1}^{K}\log_2\left(1+\sigma^2|[\mathbf{R}]_{k,k}|^2\right), \tag{55}$$

where $\mathbf{R} \in \mathbb{C}^{K\times K}$ *is an upper triangular matrix that results from the QR decomposition of* $\tilde{\mathbf{H}}$ *as*

$$\tilde{\mathbf{H}} = \mathbf{QR}, \tag{56}$$

and $\mathbf{Q} \in \mathbb{C}^{K\times K}$ *is a unitary matrix.*

Proof. Refer to Appendix A □

According to (55), we have the following upper bound of the ergodic sum rate

$$\mathbb{E}\{\check{C}\} \le \tilde{C} = \sum_{k=1}^{K} \log_2\left(1+\sigma^2\mathbb{E}\left\{|[\mathbf{R}]_{k,k}|^2\right\}\right), \tag{57}$$

which is derived by applying the Jensen's inequality on $\check{C}$ in (55).

With the upper bound of the ergodic sum rate of the proposed approach derived, we will continue to derive two properties of the upper bound.

Proposition 1. *With high SNR, i.e.,* $\sigma^2\mathbb{E}\{|[\mathbf{R}]_{k,k}|^2\} \gg 1$*, the ergodic sum rate upper bound increases approximately as*

$$\lim_{N_\mathrm{B}\to\infty} \Delta_{\tilde{C}} \sim \log_2 M_{N_\mathrm{B}}, \tag{58}$$

where $\Delta_{\tilde{C}}$ *and* M_{N_B} *are the change of the sum rate and the multiple of the number of the BS antennas.*

Proof. Refer to Appendix B □

Proposition 2. *When* $N_\mathrm{B}\to\infty$ *and* $\sigma^{\mathrm{B}}_{kl_{\mathrm{cl}}} \to 0, \forall k, l_{\mathrm{cl}}$*, the ergodic sum rate upper bound with the constraint in (22) satisfies*

$$\lim_{N_\mathrm{B}\to\infty} \lim_{\sigma^{\mathrm{B}}_{kl_{\mathrm{cl}}}\to 0,\forall k,l_{\mathrm{cl}}} \tilde{C} \sim K. \tag{59}$$

Proof. Refer to Appendix C. □

5.2. Complexity analysis

We use the big O notation $O(n)$ to indicate that the number of complex multiplications is on the order of n and assume that $N_\mathrm{B} \gg K$. For the beam selection approach in [6], the BS-side channel matrix $\mathbf{H} \in \mathbb{C}^{N_\mathrm{B}\times K}$ with

$$[\mathbf{H}]_{:,k} = \sum_{l_{\mathrm{cl}}=1}^{N_{\mathrm{cl}}} \sum_{l_{\mathrm{ray}}=1}^{N_{\mathrm{ray}}} \alpha_{kl_{\mathrm{cl}}l_{\mathrm{ray}}} \mathbf{a}_\mathrm{B}(\theta^{\mathrm{B}}_{kl_{\mathrm{cl}}l_{\mathrm{ray}}}), \tag{60}$$

is transformed into beamspace as $\mathbf{H}_\mathrm{b} = \mathbf{U}^H\mathbf{H} \in \mathbb{C}^{N_\mathrm{B}\times K}$, where the columns of $\mathbf{U} \in \mathbb{C}^{N_\mathrm{B}\times N_\mathrm{B}}$ are in the form of beam steering vector (10) with angles given in (14). Then, the K beams corresponding to the largest K values in $\sqrt{\sum_{k=1}^{K}|[\mathbf{H}_\mathrm{b}]_{m,k}|^2/||[\mathbf{H}_\mathrm{b}]_{:,k}||^2}$, $m = 1, 2, \cdots, N_\mathrm{B}$ are selected. As can be seen, this approach is of computational complexity $O(N_\mathrm{B}^2K)$.

For the maximization of SINR approach in [7], in the first step, each $[\mathbf{H}_\mathrm{b}]_{m,:}$ is deleted from $\mathbf{H}_\mathrm{b}$ and results into $\mathbf{H}_\mathrm{b}^{(1,m)} \in \mathbb{C}^{(N_\mathrm{B}-1)\times K}$, and the $m_1^\star$th row is deleted to form $\mathbf{H}_\mathrm{b}^{(1)} \in \mathbb{C}^{(N_\mathrm{B}-1)\times K}$, where $m_1^\star = \arg\min_m \mathrm{tr}((\mathbf{H}_\mathrm{b}^{(1,m)^H}\mathbf{H}_\mathrm{b}^{(1,m)})^{-1})$. In the second step, one row of $\mathbf{H}_\mathrm{b}^{(1)}$ should be deleted in a similar manner, and this process continues to the $N_\mathrm{B}-K$th step, results into $\mathbf{H}_\mathrm{b}^{(N_\mathrm{B}-K)} \in \mathbb{C}^{K\times K}$. Then, the corresponding beams are selected. As can be seen, this approach is of computational complexity $O(N_\mathrm{B}^3K)$.

The maximization of capacity approach in [7] is similar to the maximization of SINR approach in [7], the difference only lies on the criterion of deleting the rows. For example, in the first step, $m_1^\star = \arg\max_m \sum_{k=1}^{K}\log_2(1+\sigma^2/[(\mathbf{H}_\mathrm{b}^{(1,m)^H}\mathbf{H}_\mathrm{b}^{(1,m)})^{-1}]_{k,k})$, which is derived in a similar manner as the sum rate expression (7). Apparently, this approach is also of computational complexity $O(N_\mathrm{B}^3K)$.

For the grouping approach in [8], the elements of $[\mathbf{H}_\mathrm{b}]_{:,k}$ are sorted by the absolute value and the beam corresponding to the largest one is denoted as the strongest beam of the kth MS. Then, if the strongest beam of the kth MS is different from that of any other MS, this MS is grouped into the non-interference group, and others are grouped into the interference group. For the non-interference group, the strongest beam is selected. For the interference group, the beam with the maximum contribution to the sum rate is selected. For example, in the first step, the $m_1^\star$th row of $\mathbf{H}_\mathrm{b}$ is selected, where $m_1^\star = \arg\max_m \sum_{k=1}^{K}\log_2(1+\sigma^2/[(\mathbf{A}^H\mathbf{A}+\varepsilon\mathbf{I}_K+[\mathbf{H}_\mathrm{b}]_{m,:}^H[\mathbf{H}_\mathrm{b}]_{m,:})^{-1}]_{k,k})$, and $\mathbf{A} \in \mathbb{C}^{r\times K}$ corresponds to the rows selected for the non-interference group and r is the number of rows. In addition, $\varepsilon\mathbf{I}_K$ is used to make sure the matrix inverse exists and is set as $\varepsilon = 10^{-3}$. It can be seen that the computational complexity is $O(N_\mathrm{B}^2K + N_\mathrm{B}K^4)$.

For the proposed approach, it can be seen that the main computation lies on the calculation of $C_{k,l_{\mathrm{cl}}}$ in step 12 of Algorithm 1, which is of computational complexity $O(N_\mathrm{B}N_\mathrm{M}K)$. Thus, the proposed approach is of computational complexity $O(N_\mathrm{B}N_\mathrm{M}K^2N_{\mathrm{cl}})$. When $N_\mathrm{B}^2 > N_\mathrm{M}KN_{\mathrm{cl}}$ and $N_\mathrm{B}+K^3 > N_\mathrm{M}KN_{\mathrm{cl}}$, the proposed approach is of lower computational complexity than the other approaches. For example, with the simulation parameters given in Table 1, the proposed approach is of computational complexity $O(10^6)$, while approaches [6]-[7] are of computational complexity $O(10^8)$, and the approach [8] is of computational complexity $O(7\times 10^5)$. As can be seen, the proposed approach is of similar or lower computational complexity.

Table 1
Simulation parameters.

Parameter	Value	Parameter	Value
N_B	256	N_M	16
N_R	6	K	6
N_{cl}	8	N_{ray}	10
$\sigma^{\mathrm{B}}_{kl_{\mathrm{cl}}}$	5°	$\sigma^{\mathrm{M}}_{kl_{\mathrm{cl}}}$	5°
σ^2	100	ρ_B	100
ρ_M	10^4	d/λ	0.5

6. Numerical results

In this section, simulations are carried out to evaluate the sum rate. The simulation parameters are given in Table 1 and are partially adopted from Sun et al. [18]. Note that as the existing approaches only select beams at the BS, the beams at the MSs are selected based on the maximum gain criterion in the simulations. The MS-side beam selection for the approaches in [6–8] are as follows. The MS-side channel matrix at the kth MS $\tilde{\mathbf{h}}_k = \sum_{l_{\mathrm{cl}}=1}^{N_{\mathrm{cl}}}\sum_{l_{\mathrm{ray}}=1}^{N_{\mathrm{ray}}}\alpha_{kl_{\mathrm{cl}}l_{\mathrm{ray}}}\mathbf{a}_\mathrm{M}(\theta^{\mathrm{M}}_{kl_{\mathrm{cl}}l_{\mathrm{ray}}}) \in \mathbb{C}^{N_\mathrm{M}\times 1}$ is transformed into $\tilde{\mathbf{h}}_k^\mathrm{b} = \tilde{\mathbf{U}}^H\tilde{\mathbf{h}}_k \in \mathbb{C}^{N_\mathrm{M}\times 1}$, where the columns of $\tilde{\mathbf{U}} \in \mathbb{C}^{N_\mathrm{M}\times N_\mathrm{M}}$ are in the form of beam steering vector (9) with angles given in (12). Then, the MS-side beam corresponding to $\max_n[|\tilde{\mathbf{h}}_k^\mathrm{b}|]_n$ is selected at the kth MS. The case with fully digital processing at the BS is also simulated, which means there are N_B RF chains at the BS, and ZF receiving is employed. In the simulations, the beam selection approach in [6] is denoted as "Beam mask [6]", the maximization of SINR approach in [7] is denoted as "M-SINR [7]", the maximization of capacity approach in [7] is denoted as "M-capacity [7]", the grouping approach in [8] is denoted as "Group [8]", the proposed SOI approach is denoted as "SOI", the upper bound in (55) is denoted as "Upper bound", and the fully digital processing is denoted as "Fully-digital".

In Fig. 1, the sum rate versus the SNR is depicted. The proposed SOI based beam selection approach is of much higher sum rate than other approaches. Moreover, the sum rate of the SOI approach is close to the derived upper bound, and increases in the same rate as the fully digital receiving scenario. These results verify that the

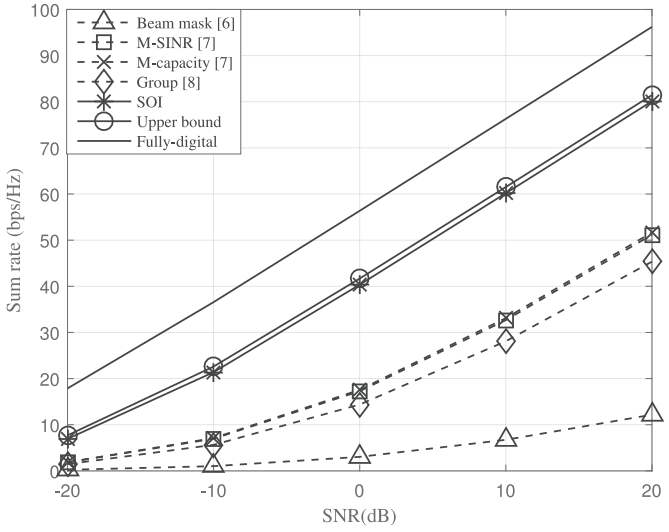

Fig. 1. The sum rate versus the SNR.

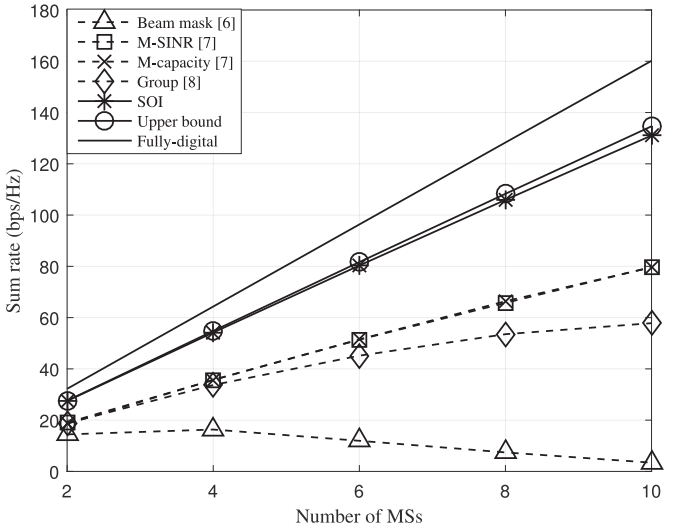

Fig. 3. The sum rate versus the number of MSs.

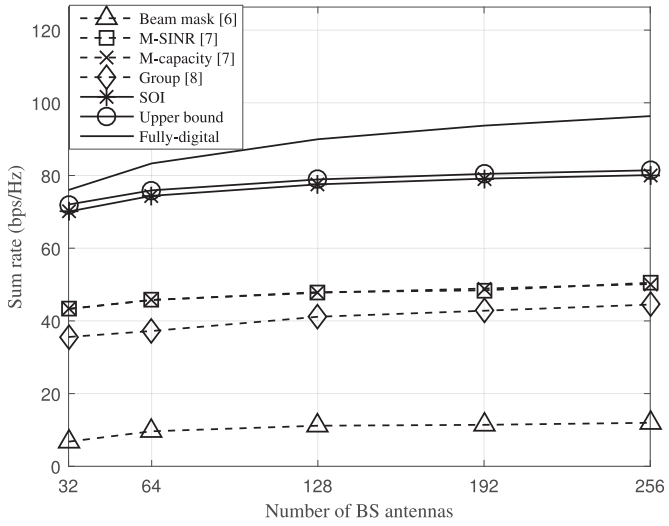

Fig. 2. The sum rate versus the number of BS antennas.

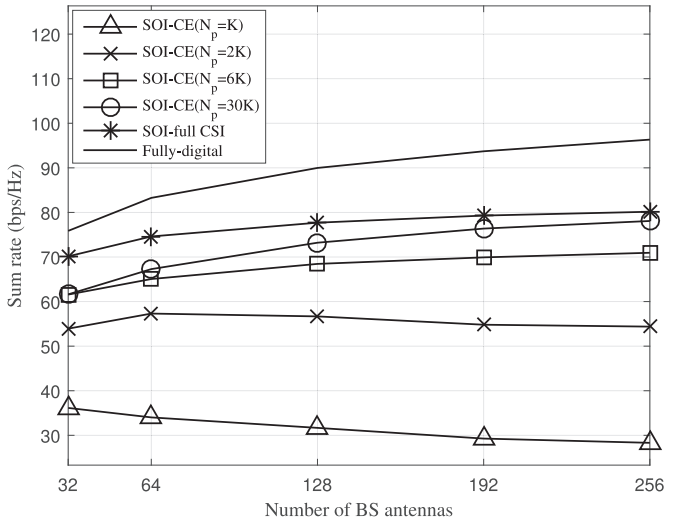

Fig. 4. The sum rate comparison of the two proposed SOI based beam selection approaches.

proposed approach is superior in eliminating interference and the derived upper bound is tight.

In Fig. 2, the sum rate versus the number of BS antennas is shown. The sum rate of the proposed SOI approach increases with the increase of N_B and is always higher than that of other approaches. In addition, the gap between the upper bounds and the sum rate is almost invariant with the increase of N_B. Moreover, with N_B multiplied by 2 and changes sequentially as 32, 64, 128, 256, the corresponding sum rate changes are 3.9, 3, and 2.5, which are close to each other. This result verifies the approximate effectiveness of the derived upper bounds in Proposition 1.

In Fig. 3, the sum rate versus the number of MSs is shown. The proposed approach still performs better than other approaches and is close to the upper bound. Moreover, the upper bound increases linearly with the increase of the number of MSs. This result verifies the effectiveness of the upper bound in Proposition 2.

In Fig. 4, the sum rate with channel estimation and SOI based beam selection proposed in Section 4 is compared with the sum rate of the proposed SOI approach with full CSI in Section 3 and the case with fully digital processing. Here, the SOI approach with channel estimation in Section 4 is denoted as "SOI-CE($N_P = \cdot$)", and the $\cdot$ represents different number of uplink pilot slots, the proposed SOI with full CSI in Section 3 is denoted as "SOI-full CSI". As can be seen, the sum rate of the SOI with channel estimation increases with the increase of the number of uplink pilot symbols. Moreover, with $N_P = 30K$, the sum rate of the SOI with channel estimation gets close to the sum rate of the SOI with full CSI when the number of BS antennas increases. These results verify that the SOI with channel estimation can also have good performance with enough pilot symbols and can even have similar performance with the SOI with full CSI.

In order to show the superiority of the proposed method, the more practical 3GPP spatial channel model in [19] will also be used here. More specifically, the carrier frequency is $f_c = 30$ GHz, the bandwidth is 100 MHz, the noise figure is 9 dB, the BS height is 10 m, the MSs distribute uniformly in the 120° sector in front of the BS and in the distance range [10, 100] m. The MS height is 1.5 m. The path loss is modeled as $32.4 + 21\log_{10} d + 20\log_{10} f_c$, where d is the distance between the MS and the BS. The number of clusters and the number of rays are changed into $N_{cl} = 12$, $N_{ray} = 20$. In Fig. 5, the sum rate versus the SNR is shown. In this figure, we can see that the proposed approach still performs better than other approaches and is close to the upper bound. In Fig. 6, the transmission power is 60 dBm, the sum rate with channel estimation

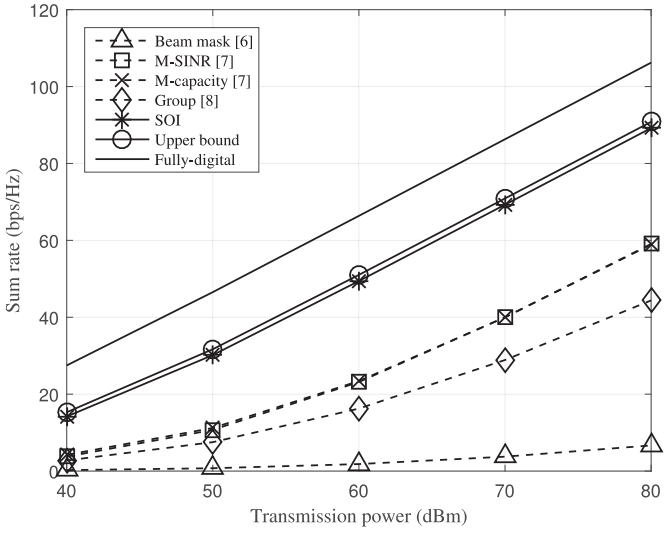

Fig. 5. The sum rate versus the SNR with the 3GPP channel model.

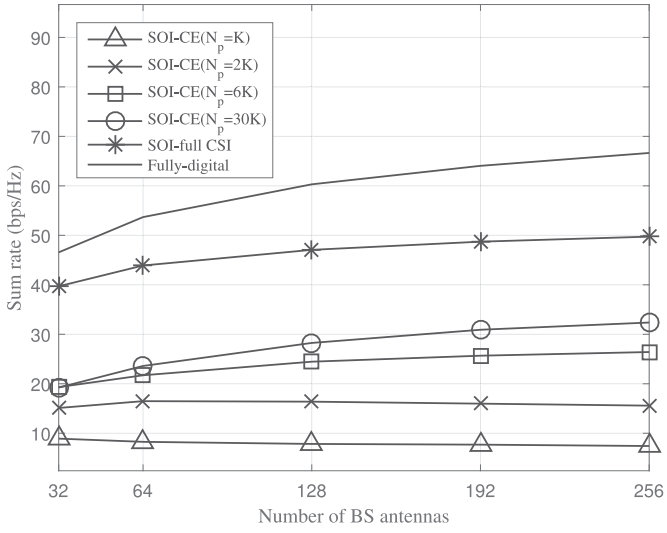

Fig. 6. The sum rate comparison of the two proposed SOI based beam selection approaches with the 3GPP channel model.

and SOI based approaches as well as the fully digital processing are compared. It can be observed that the sum rate achieved with the SOI based approach increases with the increase of the number of pilots and can even gets close to that with full CSI.

7. Conclusions

In this paper, the uplink sum rate of mm-wave multiuser MIMO systems that employ ZF based hybrid precoding is analyzed. By introducing the SOI, the most strong interference can be identified. Then, by resorting to the SOI, the joint beam selection with beam constraints is proposed. As full channel information may be infeasible, a simple channel estimation approach is proposed. With channel estimation, the SOI is defined in another way, and the corresponding SOI based beam selection is also proposed. Additionally, the upper bound of the sum rate is analyzed. Simulation results show that the proposed SOI approach performs better than other beam selection approaches and is close to the upper bound. Moreover, the SOI approach with channel estimation can also perform similarly as the SOI approach with full CSI.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This research was supported by Project 61601152 and Project U1609216 supported by National Natural Science Foundation of China.

Appendix A. Proof of Lemma 1

According to the QR decomposition, we have

$$||[\tilde{\mathbf{H}}^{-1}]_{k,:}||^2 = [\mathbf{R}^{-1}]_{k,:}\mathbf{Q}^H\mathbf{Q}[\mathbf{R}^{-1}]^H_{k,:} \tag{61}$$

$$= ||[\mathbf{R}^{-1}]_{k,:}||^2 \tag{62}$$

$$= \sum_{k'=1}^{K} \frac{|M_{k',k}|^2}{|\det(\mathbf{R})|^2}, \tag{63}$$

where $M_{k',k}$ is the determinant of the matrix that results from deleting the k'th row and the kth column of $\mathbf{R}$. Since $\mathbf{R}$ is an upper triangular matrix, we have $|\det(\mathbf{R})| = \prod_{j=1}^{K} |[\mathbf{R}]_{j,j}|$, $|M_{k,k}| = \prod_{j=1,j\neq k}^{K} |[\mathbf{R}]_{j,j}|$. Accordingly, we have

$$|M_{k',k}| = 0, \forall k' < k, \tag{64}$$

$$|M_{k',k}| \geq 0, \forall k' > k. \tag{65}$$

Then, we have

$$||[\tilde{\mathbf{H}}^{-1}]_{k,:}||^2 \geq \frac{|M_{k,k}|^2}{|\det(\mathbf{R})|^2} = \frac{1}{|[\mathbf{R}]_{k,k}|^2}. \tag{66}$$

Substituting (66) into (7) yields

$$C \leq \sum_{k=1}^{K} \log_2\left(1 + \sigma^2 |[\mathbf{R}]_{k,k}|^2\right). \tag{67}$$

Appendix B. Proof of Proposition 1

According to the structure of $\mathbf{R}$, we have $[\tilde{\mathbf{H}}]_{:,k} = \sum_{j=1}^{k} [\mathbf{R}]_{j,k}[\mathbf{Q}]_{:,j}$. Since $\mathbf{Q}$ is unitary, we have $||[\tilde{\mathbf{H}}]_{:,k}||^2 = \sum_{j=1}^{k} |[\mathbf{R}]_{j,k}|^2$. Accordingly, we have

$$|[\mathbf{R}]_{k,k}|^2 = ||[\tilde{\mathbf{H}}]_{:,k}||^2 - \sum_{j=1}^{k-1} |[\mathbf{R}]_{j,k}|^2 \tag{68}$$

$$= ||[\tilde{\mathbf{H}}]_{:,k}||^2 - \sum_{j=1}^{k-1} |[\mathbf{Q}]^H_{:,j}[\tilde{\mathbf{H}}]_{:,k}|^2 \tag{69}$$

$$\sim ||[\tilde{\mathbf{H}}]_{:,k}||^2. \tag{70}$$

Thus, we have

$$\mathbb{E}\left\{|[\mathbf{R}]_{k,k}|^2\right\} \sim \mathbb{E}\left\{||[\tilde{\mathbf{H}}]_{:,k}||^2\right\}. \tag{71}$$

According to (5), we have

$$\mathbb{E}\left\{||[\tilde{\mathbf{H}}]_{:,k}||^2\right\} = \frac{N_{\mathrm{M}}N_{\mathrm{B}}}{N_{\mathrm{cl}}N_{\mathrm{ray}}} \sum_{l_{\mathrm{cl}}=1}^{N_{\mathrm{cl}}} \sum_{l_{\mathrm{ray}}=1}^{N_{\mathrm{ray}}} \mathbb{E}\left\{|\mathbf{a}^H_{\mathrm{M}}\left(\theta^{\mathrm{M}}_{kl_{\mathrm{cl}}l_{\mathrm{ray}}}\right)\mathbf{a}_{\mathrm{M}}\left(\varphi^{\mathrm{M}}_{k}\right)|^2\right\} \times \sum_{k'=1}^{K} \mathbb{E}\left\{|\mathbf{a}^H_{\mathrm{B}}\left(\varphi^{\mathrm{B}}_{k'}\right)\mathbf{a}_{\mathrm{B}}\left(\theta^{\mathrm{B}}_{kl_{\mathrm{cl}}l_{\mathrm{ray}}}\right)|^2\right\} \tag{72}$$

$$\sim N_{\mathrm{B}}. \tag{73}$$

Substituting (72) into (71) yields

$$\mathbb{E}\left\{|[\mathbf{R}]_{k,k}|^2\right\} \sim N_{\mathrm{B}}. \tag{74}$$

On the other hand, with $\sigma^2\mathbb{E}\{|[\mathbf{R}]_{k,k}|^2\} \gg 1$, the upper bound in (55) can be approximated as

$$\begin{aligned}\check{C} &\approx \sum_{k=1}^{K}\log_2\left(\sigma^2|[\mathbf{R}]_{k,k}|^2\right)\\ &= K\log_2(\sigma^2)+\sum_{k=1}^{K}\log_2\left(|[\mathbf{R}]_{k,k}|^2\right).\end{aligned} \tag{75}$$

Substituting (74) into (75) results into

$$\check{C} - K\log_2(\sigma^2) \sim \log_2 N_{\mathrm{B}}. \tag{76}$$

Thus, when N_{B} is multiplied into $N_{\mathrm{B}}M_{N_{\mathrm{B}}}$, the corresponding change of sum rate satisfies

$$\check{C} + \Delta_{\check{C}} - K\log_2(\sigma^2) \sim \log_2(N_{\mathrm{B}}M_{N_{\mathrm{B}}}). \tag{77}$$

Subtracting (77) from (76) results into (58).

Appendix C. Proof of Proposition 2

According to (8), we have

$$\lim_{\sigma^{\mathrm{B}}_{kl_{\mathrm{cl}}}\to 0,\forall l_{\mathrm{cl}}} \mathbf{H}_k \sim \sum_{l_{\mathrm{cl}}=1}^{N_{\mathrm{cl}}}\sum_{l_{\mathrm{ray}}=1}^{N_{\mathrm{ray}}} \alpha_{kl_{\mathrm{cl}}l_{\mathrm{ray}}}\mathbf{a}_{\mathrm{B}}\left(\bar{\theta}^{\mathrm{B}}_{kl_{\mathrm{cl}}}\right)\mathbf{a}^H_{\mathrm{M}}\left(\theta^{\mathrm{M}}_{kl_{\mathrm{cl}}l_{\mathrm{ray}}}\right). \tag{78}$$

Substituting (78) into (5) yields

$$\lim_{\sigma^{\mathrm{B}}_{kl_{\mathrm{cl}}}\to 0,\forall l_{\mathrm{cl}}} [\tilde{\mathbf{H}}]_{:,k} \sim \mathbf{F}^H_{\mathrm{A}}\sum_{l_{\mathrm{cl}}=1}^{N_{\mathrm{cl}}}\sum_{l_{\mathrm{ray}}=1}^{N_{\mathrm{ray}}} \alpha_{kl_{\mathrm{cl}}l_{\mathrm{ray}}}\mathbf{a}_{\mathrm{B}}\left(\bar{\theta}^{\mathrm{B}}_{kl_{\mathrm{cl}}}\right)\mathbf{a}^H_{\mathrm{M}}\left(\theta^{\mathrm{M}}_{kl_{\mathrm{cl}}l_{\mathrm{ray}}}\right)\mathbf{w}_k. \tag{79}$$

Additionally, as $N_{\mathrm{B}}\to\infty$, the probability $P(|\bar{\theta}^{\mathrm{B}}_{k'l'_{\mathrm{cl}}} - \bar{\theta}^{\mathrm{B}}_{kl_{\mathrm{cl}}}| \le \pi/\sqrt{N_{\mathrm{B}}}) \to 0, \forall k'\neq k, \forall l_{\mathrm{cl}}, l'_{\mathrm{cl}}$. According to (22), we have

$$P\left(|\varphi^{\mathrm{B}}_{k'} - \bar{\theta}^{\mathrm{B}}_{kl_{\mathrm{cl}}}| \le \frac{\pi}{2\sqrt{N_{\mathrm{B}}}}\right) = 0, \forall k'\neq k, \forall l_{\mathrm{cl}}. \tag{80}$$

Applying the correlation results in [16] to (80) results into

$$\lim_{N_{\mathrm{B}}\to\infty} \mathbf{a}^H_{\mathrm{B}}\left(\varphi^{\mathrm{B}}_{k'}\right)\mathbf{a}_{\mathrm{B}}\left(\bar{\theta}^{\mathrm{B}}_{kl_{\mathrm{cl}}}\right) = 0, \forall k'\neq k, \forall l_{\mathrm{cl}}. \tag{81}$$

With (13), we have

$$\lim_{N_{\mathrm{B}}\to\infty} [\mathbf{F}_{\mathrm{A}}]^H_{:,k'}\mathbf{a}_{\mathrm{B}}(\bar{\theta}^{\mathrm{B}}_{kl_{\mathrm{cl}}}) = 0, \forall k'\neq k, \forall l_{\mathrm{cl}}. \tag{82}$$

Substituting (82) into (79) results into

$$\lim_{N_{\mathrm{B}}\to\infty}\lim_{\sigma^{\mathrm{B}}_{kl_{\mathrm{cl}}}\to 0,\forall l_{\mathrm{cl}}} \frac{[\tilde{\mathbf{H}}]_{k',k}}{[\tilde{\mathbf{H}}]_{k,k}} = 0, \forall k'\neq k. \tag{83}$$

Thus, we have

$$\lim_{N_{\mathrm{B}}\to\infty}\lim_{\sigma^{\mathrm{B}}_{kl_{\mathrm{cl}}}\to 0,\forall l_{\mathrm{cl}}} |[\mathbf{R}]_{k,k}|^2 = |||[\tilde{\mathbf{H}}]_{:,k}||^2. \tag{84}$$

Substituting (84) into (57) yields

$$\lim_{N_{\mathrm{B}}\to\infty}\lim_{\sigma^{\mathrm{B}}_{kl_{\mathrm{cl}}}\to 0,\forall k,l_{\mathrm{cl}}} \check{C} = \sum_{k=1}^{K}\log_2\left(1+\sigma^2\mathbb{E}\left\{|||[\tilde{\mathbf{H}}]_{:,k}||^2\right\}\right) \sim K. \tag{85}$$

References

[1] Z. Pi, F. Khan, An introduction to millimeter-wave mobile broadband systems, IEEE Commun. Mag. 49 (6) (2011) 101–107.
[2] J. Li, L. Xiao, X. Xu, et al., Robust and low complexity hybrid beamforming for uplink multiuser MmWave MIMO systems, IEEE Commun. Lett. 20 (6) (2016) 1140–1143.
[3] F. Sohrabi, W. Yu, Hybrid digital and analog beamforming design for large-scale antenna arrays, IEEE J. Sel. Top. Signal Process. 10 (3) (2016) 501–513.
[4] A. Alkhateeb, R.W. Heath Jr., Frequency selective hybrid precoding for limited feedback millimeter wave systems, IEEE Trans. Commun. 64 (5) (2016) 1801–1818.
[5] A. Alkhateeb, O.E. Ayach, G. Leus, et al., Channel estimation and hybrid precoding for millimeter wave cellular systems, IEEE J. Sel. Top. Signal Process 8 (5) (2014) 831–846.
[6] J. Hogan, A. Sayeed, Beam selection for performance-complexity optimization in high-dimensional MIMO systems, in: Proc. of 2016 Annu. Conf. Inform. Sci. and Syst., 2016, pp. 337–342.
[7] P.V. Amadori, C. Masouros, Low RF-complexity millimeter-wave beamspace-MIMO systems by beam selection, IEEE Trans. Commun. 63 (6) (2015) 2212–2223.
[8] X. Gao, L. Dai, Z. Chen, et al., Near-optimal beam selection for beamspace MmWave massive MIMO systems, IEEE Commun. Lett. 20 (5) (2016) 1054–1057.
[9] J. Chen, P. Zhao, Z. Wang, et al., Enhanced beam selection for multi-user mm-wave massive MIMO systems, Electron. Lett. 52 (14) (2016) 1268–1270.
[10] Z. Guo, J. Li, Time-variant beam selection for lens-based millimetre-wave massive MIMO systems, Electron. Lett. 53 (17) (2017) 1226–1228.
[11] A. Ali, N. Gon'alez-Prelcic, R.W. Heath Jr., Millimeter wave beam-selection using out-of-band spatial information, IEEE Trans. Wirel. Commun. 17 (2) (2018) 1038–1052.
[12] J. Choi, Beam selection in mm-wave multiuser MIMO systems using compressive sensing, IEEE Trans. Commun. 63 (8) (2015) 2936–2947.
[13] O.E. Ayach, S. Rajagopal, S. Abu-Surra, et al., Spatially sparse precoding in millimeter wave MIMO systems, IEEE Trans. Wirel. Commun. 13 (3) (2014) 1499–1513.
[14] A. Adhikary, E.A. Safadi, M.K. Samimi, et al., Joint spatial division and multiplexing for MmWave channels, IEEE J. Sel. Areas Commun. 32 (6) (2014) 1239–1255.
[15] Z. Chen, C. Yang, Pilot decontamination in wideband massive MIMO systems by exploiting channel sparsity, IEEE Trans. Wirel. Commun. 15 (7) (2016) 5087–5100.
[16] A. Hu, Antenna tilt design for millimeter-wave beamspace MIMO systems, Wirel. Pers. Commun. 94 (3) (2017) 1701–1713.
[17] E.W. Weisstein, Chi distribution, 2019. [Online]. Available: http://mathworld.wolfram.com/ChiDistribution.html.
[18] S. Sun, T.S. Rappaport, M. Shafi, et al., Propagation models and performance evaluation for 5g millimeter-wave bands, IEEE Trans. Veh. Technol. 67 (9) (2018) 8422–8439.
[19] 3GPP, Study on Channel Model for Frequencies from 0.5 to 100 GHz. 3rd Gener. Partnership Project, 2017.