IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 2, 2017, accepted March 20, 2017, date of publication March 24, 2017, date of current version May 17, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2687078

User Selection Algorithms for Block Diagonalization Aided Multiuser Downlink mm-Wave Communication

RAKSHITH RAJASHEKAR AND LAJOS HANZO, (Fellow, IEEE)
School of ECS, University of Southampton, Southampton SO17 1BJ, U.K.

Corresponding author: L. Hanzo (lh@ecs.soton.ac.uk)

This work was supported in part by the EPSRC under Grant EP/Noo4558/1 and Grant EP/L018659/1, in part by the European Research Council's Advanced Fellow Grant under the Beam-Me-Up Project and of the Royal Society's Wolfson Research Merit Award is gratefully acknowledged.

ABSTRACT In order to meet the ever increasing data rate demands, the next generation of wireless communication systems is being designed for exploiting the large amounts of unused spectrum in the millimeter-wave (mm-wave) band. Since operating at mm-wave frequencies imposes several challenges, such as high-path loss, as well as both spatial and temporal channel sparsity, there is a significant research interest focused on designing feasible solutions for establishing reliable and high-throughput links at mm-wave frequencies. In this paper, we consider a cellular system relying on hybrid beamforming aided base station (BS) as well as user equipment and study the user selection problem, which has not been hitherto studied in the literature. More specifically, we study the problem of selecting K'-users by the BS for communication out of K-users whilst ensuring that the sum rate is maximized. Specifically, we propose a user selection algorithm, which relies on the knowledge of both the channel gains and the angle of departure (AoD) of the channel paths spanning to the various users, which is termed the AoD aided user selection (AoD-US). Furthermore, we devise a pair of subspace metrics based on: 1) the angle between the subspaces spanned by the BS array response vectors; 2) the ratio of interference and signal space dimensions of various users, in order to reduce the user search space in AoD-US. This modified user selection algorithm is termed the AoD aided user selection with user set pruning (AoD-US-P). Furthermore, we study the attainable sum-rate performance of the block-diagonalization aided downlink and show that the proposed selection algorithms guarantee both multiuser diversity and multiplexing gains. Additionally, the proposed algorithms are studied in the round-robin scheduling scenario, where all the K-users are scheduled for achieving fairness. Our simulation results revealed that the AoD-US-P achieves nearly the same performance as that achieved by the AoD-US despite having a small user set, while both are observed to outperform the channel power-based selection scheme.

INDEX TERMS User selection, scheduling, block diagonalization, mm-wave communication, precoding and combining.

I. INTRODUCTION

The large unused spectral resources available in the millimeter wave (mm-wave) frequencies are expected to be utilized for meeting the throughput demands of the next generation wireless communication systems [1]–[5]. While attaining higher throughputs is the main goal, as a benefit of having higher user bandwidths, operating at mm-wave frequencies comes with its own challenges [3]. A major challenge imposed by the mm-wave channel is the high signal attenuation and the channel sparsity in the spatial and temporal domains [6], [7]. Beamforming (BF) is a popular solution conceived for mitigating the pathloss imposed by the channel, where both the transmitter and the receiver are *aligned* to the dominant singular vectors of the channel matrix in order to improve the received signal quality. Furthermore, the performance of the *hybrid* BF (HBF) architecture has been extensively studied in the literature [8]–[20], where several analog phase shifters are employed in conjunction with the digital processing implemented with the aid of a reduced number of RF chains compared to the number of antennas available

at the transmitter/receiver. Specifically, the basis pursuit based HBF was studied in [11]–[13], a Convex Quadratic Programming based HBF method was proposed in [14], while in [15] and [16] the authors studied the HBF relying on dominant beam selection and MMSE based successive interference cancellation methods, respectively. In [17], the authors proposed a low-complexity directional BF approach, while in [18] the HBF problem was studied under the finite alphabet constraint. Rajashekar and Hanzo [19] proposed an iterative matrix decomposition based HBF (IMD-HBF), which gives an accurate representation of the unconstrained BF matrices.

While the above HBF methods were proposed for single user scenarios, there are several methods have also been conceived for multiuser mm-wave communication in the literature [19]–[25]. Specifically, Liang *et al.* [21] presented a phased-zero forcing (PZF) method, where a conventional zero forcing (ZF) approach was employed in the baseband domain. As a further development, Stirling-Gallacher and Rahman [22] presented beam selection based analog BF followed by minimum mean square error (MMSE) based digital precoding. Choi [23] proposed an orthogonal matching pursuit (OMP) [27] based beam selection method, while Alkhateeb *et al.* [24] proposed a two-stage approach to the multiuser BF problem, where the analog beams were selected in the first stage followed by digital precoding based on the conventional ZF approach. Ni and Dong [25] proposed a block diagonalization (BD) [28], [29] based digital precoding and equal gain transmission (EGT) [26] based analog BF arrangement. The schemes in [21]–[25] rely on the idealized simplifying assumption of having full channel state information (CSI) for designing BD precoder. In [19], we proposed a BD scheme relying only on the knowledge of the angle of departure (AoD) of the signals at the base station (BS). Recently, Cao and Thompson [20] considered a low complexity multiuser transmission scheme for mm-wave communication, where the cell is divided into *orthogonal* beam sectors and then an SDMA-TDMA based transmission scheme is employed. Furthermore, the authors consider a stochastic approach in order to anlyse the coverage probability of achievable rate and SINR, in contrast to the physical layer perspective taken in this paper.

Against this background, the following are the new contributions of this paper.

1) In the conventional block diagonalization schemes [28], [29], the number of users that can be supported by the BS is limited by the number of transmit RF chains. Specifically, if K is the number of users in the cell and M_t is the number of transmit RF chains at the BS, then the number of users that can be supported by the BS with the aid of N_s streams per user is $K' = M_t/N_s$. As a result, only a fixed number of K' users can be scheduled for communication having a given channel use. In this scenario, two important questions arise. Namely
   a) How to choose K' users out of the K users for a given channel use ?
   b) How to schedule the K users in the cell to achieve fairness among users ?

   The above issues are indeed quite well investigated in the literature [30]–[33], but in the context of microwave communication, where the channel is usually assumed to be Rayleigh or Rician [34]. In this paper, we revisit the above problems in the context of mm-wave communication, which has not been hitherto disseminated in the literature. Note that the existing literature [19]–[25] mainly deal with the multiuser communication in the mm-wave frequencies, but does not address the user selection problem which is the main focus of this paper. Unlike the existing user selection schemes [30]–[33], which rely on full CSI at the BS, we propose a low-complexity user selection algorithm that requires only partial CSI at the BS. Specifically, our solution relies on the knowledge of the channel gains and the AoD of the channel paths between the BS and each user, while being oblivious to the angle of arrival (AoA) of the channel paths impinging on each user. This gives a significant benefit, since the BD can be achieved once uplink training has been carried out without suffering from any additional delay due to the independent downlink training required by each user. Our proposed solution is termed as the AoD aided user selection (AoD-US).
2) Furthermore, we propose a modified AoD-US algorithm termed as AoD aided user selection with user set pruning (AoD-US-P). Specifically, we propose a simple method for reducing the size of the user set in AoD-US, which offers a reduced user search space. Two subspace measures are investigated in order to achieve this, which are based on (i) the angle between the subspaces spanned by the BS array response vectors, (ii) normalized dimension of the interference and the signal spaces of various users. Furthermore, we study the performance of these algorithms in the round-robin (RR) scheduling scenario and quantify the sum rate achieved by the BS employing both IMD-HBF [19] and EGT-HBF [25].

The remainder of the paper is organized as follows. Section II outlines the system model that describes the mm-wave channel and the BD based downlink communication employed in IMD-HBF [19] and EGT-HBF [25]. In Section III, we present our proposed AoD-US, AoD-US-P and their counterparts in the context of RR scheduling. Finally, our simulation results and discussions are presented in Section IV, while Section V concludes the paper.

Notations: $\mathbb{C}$ and $\mathbb{R}$ represent the field of complex and real numbers, respectively. Lowercase and uppercase boldface letters represent matrices and vectors, respectively. The two-norm of a vector or the Frobenius norm of a matrix is represented by $\|\cdot\|$. The spectral norm of a matrix $\mathbf{X}$ is represented by $\|\mathbf{X}\|_\sigma$. If $\mathbf{A} \in \mathbb{C}^{m\times n}$ such that $a_{i,j}e^{jb_{i,j}}$ is the polar representation of the $(i,j)^{\text{th}}$ element $\mathbf{A}_{(i,j)}$, then $\angle\mathbf{A}$ represents a matrix whose $(i,j)^{\text{th}}$ element $\angle\mathbf{A}_{(i,j)} = e^{jb_{i,j}}$.

The $|\mathbf{A}|$ represents the determinant of the matrix $\mathbf{A}$ and $|\cdot|$ represents the cardinality of a given set. The span($\mathbf{A}$) represents the space spanned by the columns of $\mathbf{A}$, while $\mathbf{P}(\mathbf{A})$ represents the projection matrix onto the space spanned by the columns of $\mathbf{A}$. The $\mathtt{rank}(\mathbf{A})$ represents the rank of the matrix $\mathbf{A}$. Furthermore, the notations of $(\cdot)^*$, $(\cdot)^H$, and $(\cdot)^T$ indicate the complex conjugate, Hermitian transpose, and transpose of a vector/matrix, respectively. Furthermore, $\mathbf{A}([q:r],:)$ defines a matrix with rows $q, q+1, \ldots, r-1, r$ of $\mathbf{A}$ and $\mathbf{A}(:,[p \ : \ q])$ is a matrix with columns $p, p+1, \ldots, q-1, q$ of $\mathbf{A}$. The expected value of a random variable $\mathbf{X}$ is represented by $\mathbb{E}[\mathbf{X}]$. Furthermore, $\mathcal{CN}(\mu, \sigma^2)$ denotes the distribution of a complex Gaussian random variable with variance σ^2 and mean μ.

II. SYSTEM MODEL

A. CHANNEL MODEL

The mm-wave channel between the BS and each user is modeled by the geometric model of [11]–[14].[1] That is, the channel between the j^{th} user and the BS is given by

$$\mathbf{H}_j = \sqrt{N_t N_r} \sum_{i=1}^{L_j} \gamma_i^{(j)} \mathbf{e}_r(\psi_i^{(j)}) \mathbf{e}_t^H(\phi_i^{(j)}), \quad 1 \le j \le K, \quad (1)$$

where N_t and N_r represent the number of antennas at the BS and each user, respectively, L_j is the number of channel paths between the j^{th} user and the BS, $(\psi_i^{(j)}, \phi_i^{(j)})$ represent the AoA and AoD of the i^{th} path of the j^{th} user, $\gamma_i^{(j)} \sim \mathcal{CN}(0,1)$ is the gain of the i^{th} path of the j^{th} user's channel, $\mathbf{e}_t$ and $\mathbf{e}_r$ represent the spatial transmit and receive signatures of a uniform linear array (ULA),[2] respectively, which are given by

$$\mathbf{e}_t(\phi) = \frac{1}{\sqrt{N_t}} \left[1, e^{j\frac{2\pi}{\lambda} d \cos\phi}, \ldots, e^{j\frac{2\pi}{\lambda} d(N_t-1)\cos\phi}\right]^T,$$
$$\mathbf{e}_r(\psi) = \frac{1}{\sqrt{N_r}} \left[1, e^{j\frac{2\pi}{\lambda} d \cos\psi}, \ldots, e^{j\frac{2\pi}{\lambda} d(N_r-1)\cos\psi}\right]^T,$$

where λ is the wavelength of the carrier and d is the separation between the antenna elements. We assume that $d = \lambda/2$, $L_j = L$ [24], $(\psi_i^{(j)}, \phi_i^{(j)}) \in [0, 2\pi]$ for $1 \le j \le K$. Equation (1) can be expressed in a compact form as

$$\mathbf{H}_j = \mathbf{E}_r^{(j)} \mathbf{D}^{(j)} \mathbf{E}_t^{(j)H}, \quad (2)$$

where

$$\mathbf{E}_r^{(j)} = \left[\mathbf{e}_r(\psi_1^{(j)}), \mathbf{e}_r(\psi_2^{(j)}), \ldots, \mathbf{e}_r(\psi_L^{(j)})\right], \quad (3)$$
$$\mathbf{E}_t^{(j)} = \left[\mathbf{e}_t(\phi_1^{(j)}), \mathbf{e}_t(\phi_2^{(j)}), \ldots, \mathbf{e}_t(\phi_L^{(j)})\right], \quad (4)$$

[1]Note that the mm-wave channel model based on practical measurements as proposed in [6] and [7] considers a set of clusters, each having a number of paths associated with a low angular spread. Since paths within a cluster are highly correlated, often a simplified model is adopted in the literature [11]–[14], where each scattering cluster is assumed to contribute a single propagation path. For the ease of exposition, we have adopted the simplified channel model in this paper.

[2]Note that the results presented in this paper can be readily applied to other antenna array structures as well, such as uniform planar arrays or uniform circular arrays.

and $\mathbf{D}^{(j)}$ is a diagonal matrix whose k^{th} diagonal entry is given by $\mathbf{D}^{(j)}_{(k,k)} = \sqrt{N_t N_r}\gamma_k^{(j)}$.

B. HYBRID BEAMFORMING (HBF) SYSTEM

In order to attain high array gains, N_t and N_r are taken to be large, which are of the order of tens or hundreds of antennas [1]–[5]. However, it becomes impractical to employ the same number of RF chains, which has led to a new *hybrid* architecture where analog phase shifters are employed in conjunction with fewer RF chains at both the transmitter and the receiver. Let us assume that the BS has M_t RF chains and supports N_s streams per user, where each user is equipped with $M_r = N_s$ RF chains to receive the N_s streams. The downlink (DL) signal received at the j^{th} user is given by

$$\mathbf{y}_j = \mathbf{G}_j^H \Phi_j^H \mathbf{H}_j \Theta \mathbf{C} \mathbf{x} + \mathbf{G}_j^H \Phi_j^H \mathbf{n}_j \in \mathbb{C}^{N_s}, \quad (5)$$

where $\mathbf{x} = \left[\mathbf{x}_1^T, \mathbf{x}_2^T, \ldots, \mathbf{x}_K^T\right]^T \in \mathbb{C}^{KN_s}$ is the transmit vector in which the j^{th} user's symbols are represented by $\mathbf{x}_j \in \mathbb{C}^{N_s}$ such that $\mathbb{E}[\mathbf{x}\mathbf{x}^H] = \frac{P}{KN_s}\mathbf{I}_{KN_s}$, where P is the total transmit power at the BS and $\mathbf{n}_j \in \mathbb{C}^{N_r}$ is the additive white Gaussian noise vector, whose elements are from $\mathcal{CN}(0,1)$. Furthermore, $\mathbf{G}_j \in \mathbb{C}^{N_s \times N_s}$ and $\Phi_j \in \mathbb{C}^{N_r \times N_s}$ are the digital and analog combining matrices of the j^{th} user, respectively, $\mathbf{C} \in \mathbb{C}^{KN_s \times KN_s}$ and $\Theta \in \mathbb{C}^{N_t \times KN_s}$ are the digital and analog transmit precoding (TPC) matrices, respectively. The digital and analog precoding matrices are normalized to satisfy $\|\Theta\mathbf{C}\|^2 = KN_s$, while the digital and analog combining matrices of each user are normalized to satisfy $\|\Phi_j\mathbf{G}_j\|^2 = N_s$. Figure 1 gives a pictorial portrayal of the multiuser mm-wave system employing HBF both at the BS and at the user equipment.

C. BLOCK DIAGONALIZATION IN HBF SYSTEMS

The composite user channel can be written as

$$\mathbf{H}_{comp} = \left[\mathbf{H}_1^T, \mathbf{H}_2^T, \cdots, \mathbf{H}_K^T\right]^T \in \mathbb{C}^{KN_r \times N_t}, \quad (6)$$

while the DL TPC matrix at the BS can be formulated as

$$\mathbf{F} = \Theta\mathbf{C} = [\mathbf{F}_1, \mathbf{F}_2, \cdots, \mathbf{F}_K] \in \mathbb{C}^{N_t \times KN_s}, \quad (7)$$

where $\mathbf{F}_i \in \mathbb{C}^{N_t \times N_s}$ is the TPC matrix associated with the i^{th} user.

Definition 1: A TPC matrix $\mathbf{F}$ is said to block-diagonalize the composite user channel $\mathbf{H}_{comp}$, if we have $\mathbf{H}_i\mathbf{F}_j = \mathbf{O}_{N_r \times N_s}$ for $1 \le i \ne j \le K$.

The existing BD schemes conceived for mm-wave communication [19], [24], [25] essentially design analog and digital precoders that satisfy Definition 1. However, it is not essential to have full CSI at the BS in order to design TPCs that achieve BD. The knowledge of AoD is sufficient for achieving BD, as stated by the following proposition.

Proposition 1 ([29] Spencer et. al.): Given a composite user channel $\mathbf{H}_{comp}$, the knowledge of the AoDs of various users given by $\left\{\mathbf{E}_t^{(1)}, \mathbf{E}_t^{(2)}, \cdots, \mathbf{E}_t^{(K)}\right\}$ is sufficient for obtaining a block-diagonalizing TPC $\mathbf{F}$.

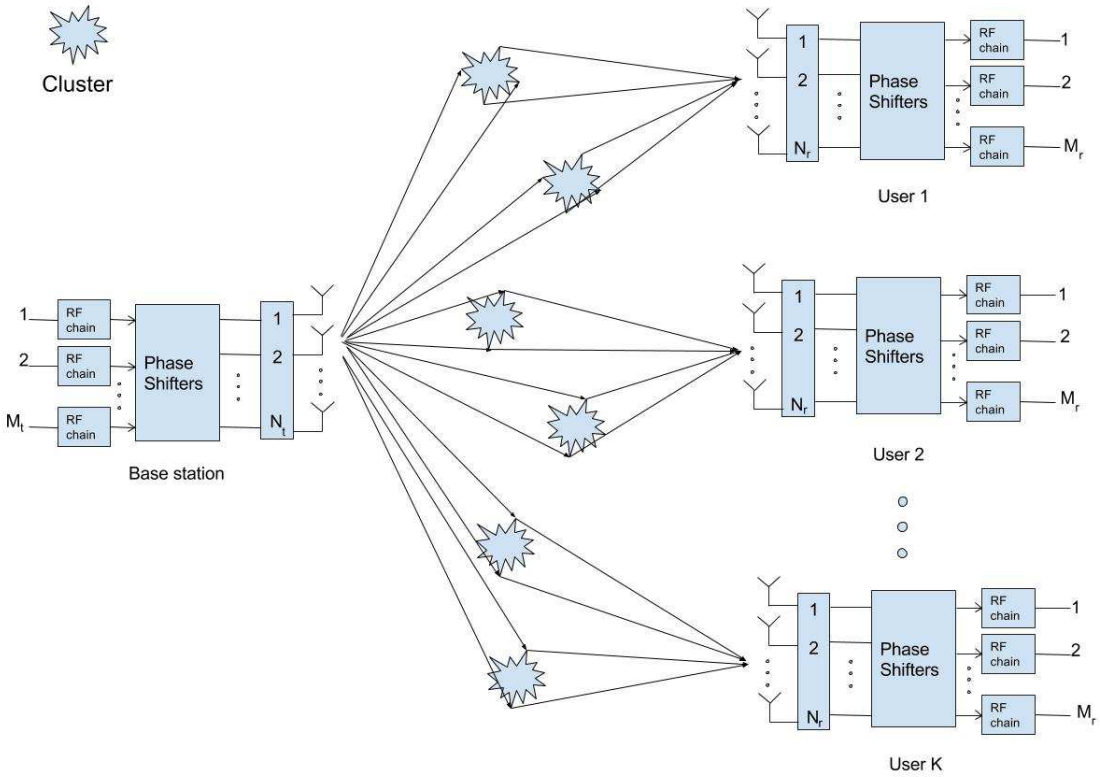


FIGURE 1. The multiuser MIMO system with a two-path channel between each user and the BS, where each user is assumed to have N_r antennas as well as M_r RF chains and the BS is assumed to have N_t antennas and M_t RF chains.

The sum rate achievable by the BD scheme in our HBF system can be expressed as

$$R_{BD} = \sum_{i=1}^{K} \log_2 \left(\left| \mathbf{I}_{N_s} + \frac{P}{KN_s} (\mathbf{G}_i^H \Phi_i^H \Phi_i \mathbf{G}_i)^{-1} \mathbf{G}_i^H \Phi_i^H \mathbf{H}_i \mathbf{F}_i \mathbf{F}_i^H \mathbf{H}_i^H \Phi_i \mathbf{G}_i \right| \right), \quad (8)$$

which is measured in terms of bits per channel use (bpcu).[3]

The TPCs proposed in [19] exploit the sufficiency of partial CSI for achieving BD for designing analog and digital TPCs for the BS, while the hybrid TPCs proposed in [21]–[25] rely on full CSI. In this paper, we consider the IMD-HBF [19] and the EGT-HBF [25] as candidate BD schemes designed for studying the sum rate performance in our proposed user selection algorithms. For further details on these BD schemes, the reader is referred to [19] and [25] as well as to the references there in. In the next section, we present the user selection problem formulated for DL mm-wave communication and present our proposed algorithms.

III. PROPOSED AoD-US AND AoD-US-P ALGORITHMS FOR mm-WAVE COMMUNICATION

Recall that there are K users in the cell and only M_t/N_s users can be scheduled for a given channel use. Let the number of users supported by the BS during a given channel use be $K' = M_t/N_s$ and the indexed set of users be represented by $\mathcal{S} = \{1, 2, \ldots, K\}$. The problem of selecting K' users that maximize the sum rate (8) is given by

$$\mathcal{S}' = \arg \max_{\mathcal{S}_s \subset \mathcal{S}} R_{BD}(\mathcal{S}_s), \quad (9)$$

where

$$R_{BD}(\mathcal{S}_s) = \sum_{i \in \mathcal{S}_s} \log_2 \left(\left| \mathbf{I}_{N_s} + \frac{P}{K'N_s} (\mathbf{G}_i^H \Phi_i^H \Phi_i \mathbf{G}_i)^{-1} \mathbf{G}_i^H \Phi_i^H \mathbf{H}_i \mathbf{F}_i \mathbf{F}_i^H \mathbf{H}_i^H \Phi_i \mathbf{G}_i \right| \right), \quad (10)$$

so that $|\mathcal{S}_s| = K'$. It is evident from (9) that the user selection problem[4] is combinatorial and imposes a high computational complexity, especially when K is large. The analogue counterpart of (9) in microwave communication has been addressed by employing greedy user-selection algorithms [30]–[33]. *In this section, we propose greedy user selection algorithms analogous to that in [31], however, in contrast to the solution proposed in [31], our solution only relies on partial CSI at the BS, while supporting multiple streams to each user.*

[3] In (8), it is implicitly assumed that the baseband user-channel diagonalization matrices are a part of $\mathbf{G}_i$ and $\mathbf{F}_i$.

[4] Note that other combined metrics such as weighted sum rate, and proportional fair scheduling [31] with power allocation for optimizing the worst-case user performance are not considered in this paper, and they are relegated for our future work.

A. PROPOSED AoD-US AND AoD-US-P ALGORITHMS

Let $\bar{\mathbf{E}}_t^{(j)} = \mathbf{E}_t^{(j)}\mathbf{D}^{*(j)}$ for $1 \le j \le K$ represent the partial CSI available at the BS, which captures both the channel gains and the AoD of the channel paths spanning to each user. Algorithm 1 provides the proposed AoD-US procedure, whose details are highlighted as follows. The first user from the set $\mathcal{S}$ is selected based on the maximum channel gain criterion. Upon selecting the first user, the next user is selected based on the criterion that maximizes the channel energy in the space orthogonal to the first user's channel. This process is repeated iteratively until K' users have been selected.

Algorithm 1 Proposed AoD-US Algorithm

Require:
$i = \arg\max_j \|\bar{\mathbf{E}}_t^{(j)}\|^2$,
$\mathcal{S} = \{1, 2, \ldots, K\}$, $\mathcal{S}_s = \{\cdot\}$,
$\mathcal{S} := \mathcal{S}\backslash i$,
$\mathcal{S}_s := \mathcal{S}_s \cup i$,
$r = 1$.
while $r \le (K' - 1)$ **do**
STEP1.
$\mathbf{P}(\mathcal{S}_s)$: Projection matrix onto $\mathtt{span}\{\bigcup_{i\in\mathcal{S}_s} \bar{\mathbf{E}}_t^{(i)}\}$
$\mathbf{P}^{\perp}(\mathcal{S}_s) = \mathbf{I} - \mathbf{P}(\mathcal{S}_s)$
STEP2.
$i = \arg\max_{j\in\mathcal{S}} \|\mathbf{P}^{\perp}(\mathcal{S}_s)\bar{\mathbf{E}}_t^{(j)}\|^2$,
$\mathcal{S} := \mathcal{S}\backslash i$,
$\mathcal{S}_s := \mathcal{S}_s \cup i$,
$r := r + 1$
end while

In order to gain insight into the interference between various users, let us consider the block-QR factorization of the CSI matrix as follows. If the ordered set of users selected from Algorithm 1 is given by $\{i_1, i_2, \ldots, i_{K'}\}$, we have

$$\left[\bar{\mathbf{E}}_t^{(i_1)}, \bar{\mathbf{E}}_t^{(i_2)}, \ldots, \bar{\mathbf{E}}_t^{(i_{K'})}\right] = [\mathbf{G}_1, \mathbf{G}_2, \cdots, \mathbf{G}_{K'}] \times \begin{bmatrix} \mathbf{I}_L & \Xi_{1,2} & \cdots & \Xi_{1,K'} \\ \mathbf{O} & \mathbf{I}_L & \cdots & \Xi_{2,K'} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{O} & \mathbf{O} & \cdots & \mathbf{I}_L \end{bmatrix}, \tag{11}$$

where

$$\mathbf{G}_1 = \bar{\mathbf{E}}_t^{(i_1)}, \quad \mathbf{G}_m = \bar{\mathbf{E}}_t^{(i_m)} - \sum_{i=1}^{m-1} \mathbf{G}_i \Xi_{i,m}, \tag{12}$$

so that

$$\Xi_{i,m} = (\mathbf{G}_i^H \mathbf{G}_i)^{-1} \mathbf{G}_i^H \bar{\mathbf{E}}_t^{(i_m)}, \tag{13}$$

for $2 \le m \le K' \le K$. If the array response vectors of various users are nearly orthogonal to each other, then the matrices $\Xi_{i,m}$ are nearly zero. Thus, $\Xi_{i,m}$ gives a measure of the amount of interference between various users. Below, we introduce two metrics, which are used for pruning the user set $\mathcal{S}$ in *Step 2* of Algorithm 1.

Lemma 1: If $\bar{\mathbf{E}}_t^{(i_k)}$ and $\mathbf{G}_k$ for $1 \le k \le K' \le K$ are defined as in (11), then

$$\mathbf{P}(\bar{\mathbf{E}}_t^{(i_k)}) = \mathbf{G}_k(\bar{\mathbf{E}}_t^{(i_k)H}\bar{\mathbf{E}}_t^{(i_k)})^{-1}\bar{\mathbf{E}}_t^{(i_k)H} + \sum_{i=1}^{k-1} \mathbf{P}(\mathbf{G}_i)\mathbf{P}(\bar{\mathbf{E}}_t^{(i_k)}). \tag{14}$$

Proof: The proof is provided in Appendix A. ■

The physical interpretation of Lemma 1 is that the projection matrix onto each user's AoD space can be decomposed into projection over the *interference space* and the *residual signal space*. In (14), each of the elements $\mathbf{P}(\mathbf{G}_i)\mathbf{P}(\bar{\mathbf{E}}_t^{(i_k)})$ in the summation corresponds to the interference space between the k^{th} and the i^{th} user's AoD space, while the term $\mathbf{G}_k(\bar{\mathbf{E}}_t^{(i_k)H}\bar{\mathbf{E}}_t^{(i_k)})^{-1}\bar{\mathbf{E}}_t^{(i_k)H}$ corresponds to the k^{th} user's residual signal space, which is free from the interference due to other users' signal. This is plausible due to the projection onto $\mathtt{span}(\mathbf{G}_k) \subseteq \mathtt{span}(\bar{\mathbf{E}}_t^{(i_k)})$ owing to (12). Thus, it is meaningful to consider $\mathbf{P}(\mathbf{G}_i)\mathbf{P}(\bar{\mathbf{E}}_t^{(i_k)})$ in order to study the amount of interference between users.

The following proposition quantifies the proportion of the interference space dimension and the residual signal space dimension in a given user's AoD space.

Proposition 2: If $\bar{\mathbf{E}}_t^{(i_k)}$ and $\mathbf{G}_k$ for $1 \le k \le K' \le K$ are defined as in (11), then

$$\|\mathbf{G}_k(\bar{\mathbf{E}}_t^{(i_k)H}\bar{\mathbf{E}}_t^{(i_k)})^{-1}\bar{\mathbf{E}}_t^{(i_k)H}\|^2 + \sum_{i=1}^{k-1} \|\mathbf{P}(\mathbf{G}_i)\mathbf{P}(\bar{\mathbf{E}}_t^{(i_k)})\|^2 = L. \tag{15}$$

Proof: The proof is provided in Appendix B. ■

From (15), we have

$$\underbrace{\frac{\|\mathbf{G}_k(\bar{\mathbf{E}}_t^{(i_k)H}\bar{\mathbf{E}}_t^{(i_k)})^{-1}\bar{\mathbf{E}}_t^{(i_k)H}\|^2}{L}}_{\substack{\text{Normalized residual} \\ \text{signal space dimension}}} + \underbrace{\frac{\sum_{i=1}^{k-1}\|\mathbf{P}(\mathbf{G}_i)\mathbf{P}(\bar{\mathbf{E}}_t^{(i_k)})\|^2}{L}}_{\substack{\text{Normalized interference space} \\ \text{dimension}}} = 1. \tag{16}$$

It can be observed from (16) that the term $\|\mathbf{P}(\mathbf{G}_i)\mathbf{P}(\bar{\mathbf{E}}_t^{(i_k)})\|^2/L$ appears in all of $\|\mathbf{P}(\bar{\mathbf{E}}_t^{(i_k)})\|^2$ for $1 \le i \le k$ and also in the corresponding term of any user selected subsequently. Thus, we can use $\|\mathbf{P}(\mathbf{G}_i)\mathbf{P}(\bar{\mathbf{E}}_t^{(i_k)})\|^2/L$ as a metric for measuring the amount of interference and in turn prune the user set in Algorithm 1. In terms of the notation used in Algorithm 1, we have $\mathbf{P}(\mathbf{G}_i)\mathbf{P}(\bar{\mathbf{E}}_t^{(i_k)}) \equiv \mathbf{P}\left(\mathbf{P}^{\perp}(\mathcal{S}_s)\bar{\mathbf{E}}_t^{(i)}\right)\mathbf{P}(\bar{\mathbf{E}}_t^{(i_k)})$.

Below, we introduce another metric for quantifying the interference between the users, which is also used for pruning the user set in Algorithm 1. Before proceeding further, let us introduce the following definition.

Definition 2 [35], [36]: Given two subspaces $\mathcal{L}$ and $\mathcal{M}$, the *minimal angle* between the two subspaces is defined as

the number in $[0, \pi/2]$ whose cosine is defined by

$$c_0(\mathcal{L}, \mathcal{M}) = \sup\left\{|\mathbf{y}^H\mathbf{z}| : \mathbf{y} \in \mathcal{L}, \|\mathbf{y}\| \leq 1, \mathbf{z} \in \mathcal{M}, \|\mathbf{z}\| \leq 1\right\}. \tag{17}$$

Proposition 3 [36]: Given two subspaces $\mathcal{L} = \texttt{span}(\mathbf{P})$ and $\mathcal{M} = \texttt{span}(\mathbf{Q})$, where $\mathbf{P}$ and $\mathbf{Q}$ are projection matrices, the minimal angle $c_0(\mathcal{L}, \mathcal{M})$ is equal to $\|\mathbf{PQ}\|_\sigma$.

Motivated by Proposition 3, we propose another metric for quantifying the interference between various users, which is given by $\|\mathbf{P}(\mathbf{G}_i)\mathbf{P}(\bar{\mathbf{E}}_t^{(i_k)})\|_\sigma$. This metric measures the minimum angle between the subspaces spanned by $\mathbf{G}_i$ and $\bar{\mathbf{E}}_t^{(i_k)}$. Note that this metric is a natural extension of the metric used in [31] to a higher dimensional space.[5] In terms of the notation used in Algorithm 1, we have $\|\mathbf{P}(\mathbf{G}_i)\mathbf{P}(\bar{\mathbf{E}}_t^{(i_k)})\|_\sigma \equiv \|\mathbf{P}\left(\mathbf{P}^\perp(\mathcal{S}_s)\bar{\mathbf{E}}_t^{(i)}\right)\mathbf{P}(\bar{\mathbf{E}}_t^{(i_k)})\|_\sigma$.

Let the proposed pair of metrics be represented by

$$m_{i,k}^{(1)} = \|\mathbf{P}\left(\mathbf{P}^\perp(\mathcal{S}_s)\bar{\mathbf{E}}_t^{(i)}\right)\mathbf{P}(\bar{\mathbf{E}}_t^{(k)})\|_\sigma, \tag{18}$$

$$m_{i,k}^{(2)} = \frac{\|\mathbf{P}\left(\mathbf{P}^\perp(\mathcal{S}_s)\bar{\mathbf{E}}_t^{(i)}\right)\mathbf{P}(\bar{\mathbf{E}}_t^{(k)})\|^2}{L}. \tag{19}$$

The proposed AoD-US algorithm relying on pruning (AoD-US-P), which relies on the metrics given in (19) and (18) is presented in Algorithm 2.

Algorithm 2 Proposed AoD-US-P Algorithm

Require:
$\mathcal{S} = \{1, 2, \ldots, K\}$, $\mathcal{S}_s = \{\cdot\}$,
$i = \arg\max_{j\in\mathcal{S}} \|\bar{\mathbf{E}}_t^{(j)}\|^2$,
$\mathcal{S} := \mathcal{S}\backslash i$,
$\mathcal{S}_s := \mathcal{S}_s \cup i$,
$0 \leq \alpha < 1$, $r = 1$.
while $r \leq (K' - 1)$ **do**
  STEP1.
  $\mathbf{P}(\mathcal{S}_s)$: Projection matrix onto $\texttt{span}\{\bigcup_{i\in\mathcal{S}_s} \bar{\mathbf{E}}_t^{(i)}\}$
  $\mathbf{P}^\perp(\mathcal{S}_s) = \mathbf{I} - \mathbf{P}(\mathcal{S}_s)$
  STEP2.
  $i = \arg\max_{j\in\mathcal{S}} \|\mathbf{P}^\perp(\mathcal{S}_s)\bar{\mathbf{E}}_t^{(j)}\|^2$,
  $\mathcal{S} := \mathcal{S}\backslash i$,
  $\mathcal{S}_s := \mathcal{S}_s \cup i$,
  Compute $m_{i,k}^{(p)}$ given by (19) or (18) $(p = 1, 2)$ $\forall k \in \mathcal{S}$
  for all $k \in \mathcal{S}$ **do**
    if $m_{i,k}^{(p)} > \alpha$ **then**
      if $|\mathcal{S}| \geq K' - |\mathcal{S}_s|$ **then**
        $\mathcal{S} := \mathcal{S}\backslash k$
      end if
    end if
  end for
  $r := r + 1$
end while

In Step 2 of Algorithm 2, satisfying the condition $|\mathcal{S}| \geq K' - |\mathcal{S}_s|$ ensures that we do not have less than K' number of users selected at the end of Algorithm 2. Note that in Step 2 of Algorithm 2, we choose the user before pruning the set, so that the performance of Algorithm 2 is as close to that of Algorithm 1 as possible.

In Section IV we show using our simulation results that the proposed algorithms give both the benefits of multiuser diversity and multiplexing gain. Specifically, we show that the achievable sum rate of the BD scheme operating with the aid of the proposed algorithms scales at high SNR as

$$R_{asy} = K'N_s \log\left(\frac{P}{K'N_s}(N_tN_r)\eta\log(K)\right) \tag{20}$$

w.r.t. the number of users K, where N_tN_r corresponds to the BF gain, $0 \leq \eta \leq 1$ corresponds to the BF efficiency, $\log(K)$ corresponds to the multiuser diversity gain [31], and the multiplication factor $K'N_s$ corresponds to the multiplexing gain.[6] In the next part of the paper, we discuss the application of the proposed user selection algorithms in the RR scheduling.

B. ROUND-ROBIN (RR) SCHEDULING

So far we have dealt with the problem of choosing K' users out of K users for DL communication. In order for the remaining $(K - K')$ users to be able to communicate, they have to be scheduled as well. There are several algorithms in the literature that deal with this problem [31], [37]–[39]. In this paper we study the achievable sum rate of the RR scheduling, whose details are given as follows.

We iteratively apply the proposed user selection algorithms, until all the K users have been scheduled. For the ease of exposition, we assume $K = cK'$, where the scheduler is invoked c times in order to schedule all the K users.

In the RR scheduling scenario, the users are scheduled in the groups given by $\mathcal{S}_s^{(1)}, \mathcal{S}_s^{(2)}, \ldots, \mathcal{S}_s^{(c)}$, where $c = K/K'$ channel uses are required for scheduling all the users. The average sum rate achievable by the BD scheme in the RR scheduling scenario is given by

$$R_{RR-BD}(\mathcal{S}) = \frac{1}{c}\sum_{i=1}^{c} R_{BD}(\mathcal{S}_s^{(i)}) \text{ bpcu}, \tag{21}$$

where $R_{BD}(\mathcal{S}_s^{(i)})$ is computed based on (10). Note that in each iteration of the RR scheduling in Algorithm 3, the users are chosen optimally, while ignoring the impact of residual users on the sum rate performance in the remaining RR slots. As a result, the chances of sum rate degradation would increase, as the RR iterations proceed, which is one of the limitations of the proposed RR scheduling.

IV. SIMULATION RESULTS AND DISCUSSIONS

In this section, we study the performance of the proposed algorithms w.r.t. various parameters, such as the signal-to-noise ratio (SNR), the number of users K in the system, the pruning parameter α, etc. We consider the IMD-HBF [19]

[5] In [31], the authors use a similar metric, which is applicable to only one dimensional subspaces $\mathcal{L}$ and $\mathcal{M}$.

[6] The BF efficiency factor η will be close to one when the values of N_t and N_r are large, which results in near-orthogonal array response vectors [34].

Algorithm 3 Proposed RR Scheduling Algorithm

Require: $\mathcal{S}' = \{1, 2, \ldots, K\}$, $l = 1$,
$0 \leq \alpha < 1$, $c = K/K'$.
while $l \leq c$ **do**
  $\mathcal{S}_s^{(l)} = \{\cdot\}$,
  $\mathcal{S} = \mathcal{S}'$,
  $i = \arg\max_{j \in \mathcal{S}} \|\bar{\mathbf{E}}_t^{(j)}\|^2$,
  $\mathcal{S}_s^{(l)} := \mathcal{S}_s^{(l)} \cup i$, $r = 1$.
  while $r \leq (K' - 1)$ **do**
    STEP1.
    $\mathbf{P}(\mathcal{S}_s^{(l)})$: Projection matrix onto $\mathtt{span}\{\bigcup_{i \in \mathcal{S}_s^{(l)}} \bar{\mathbf{E}}_t^{(i)}\}$
    $\mathbf{P}^{\perp}(\mathcal{S}_s^{(l)}) = \mathbf{I} - \mathbf{P}(\mathcal{S}_s^{(l)})$
    STEP2.
    $i = \arg\max_{j \in \mathcal{S}} \|\mathbf{P}^{\perp}(\mathcal{S}_s^{(l)})\bar{\mathbf{E}}_t^{(j)}\|^2$,
    $\mathcal{S} := \mathcal{S} \backslash i$,
    $\mathcal{S}_s^{(l)} := \mathcal{S}_s^{(l)} \cup i$
    Compute $m_{i,k}^{(p)}$ given by (19) or (18) ($p = 1, 2$) $\forall k \in \mathcal{S}$
    for all $k \in \mathcal{S}$ **do**
      if $m_{i,k}^{(p)} > \alpha$ **then**
        if $|\mathcal{S}| \geq K' - |\mathcal{S}_s^{(l)}|$ **then**
          $\mathcal{S} := \mathcal{S} \backslash k$
        end if
      end if
    end for
    $r := r + 1$
  end while
  $\mathcal{S}' := \mathcal{S}' \backslash \mathcal{S}_s^{(l)}$,
  $l := l + 1$
end while

and EGT-HBF [25] schemes as candidate BD schemes for studying the performance of the user selection algorithms.

A. SIMULATION SCENARIO

In all our simulations, we assume that the BS is equipped with $N_t = 32$ antennas, $M_t = KN_s$ RF chains, where we assume $N_s = M_r$, and each user is equipped with $N_r = 16$ antennas and $M_r = 2$ RF chains. Both the user equipment and the BS are assumed to have their antenna elements arranged in ULA with an inter-element spacing of $d = \lambda/2$. Furthermore, we assume the geometric channel model of Section II-A in conjunction with $L = 4$. Note that the SNR values depicted in various figures correspond to the pre-beamforming SNR values.

1) SUM RATE AS A FUNCTION OF α

In this section, we study the achievable sum rate performance of the IMD-HBF and EGT-HBF schemes employing Algorithm 2 as a function of α by considering $K' = 4$ and $K = \{6, 18, 30\}$. In addition to studying the performance of the proposed algorithms, we consider channel power based user selection, which naively selects the K' users having the maximum channel gains. This selection scheme is referred to as the channel power based selection (CPS). This scheme is considered in order to highlight the importance of AoD information in user selection in addition to the channel gains.

a: ALGORITHM 2 WITH $m_{i,k}^{(1)}$ (18)

Figure 2 compares the achievable sum rate in IMD-HBF and EGT-HBF based BD schemes as a function of α, where both the schemes are assumed to be operating at an SNR of -10dB and use the metric given by (18), while employing Algorithm 2. It is evident from Fig. 2 that the sum rate performance of both the schemes employing Algorithm 2 approaches that of Algorithm 1 as α increases. This is expected, since a small value of α corresponds to *near-orthogonal* set of users and hence results in a small user set. As α increases, the user set will also grow and in turn gives better sum rate performance. Also, it becomes evident from Fig. 2 that the loss in the sum rate performance is insignificant when $\alpha \geq 0.5$ and K is sufficiently large. *Note that the CPS counterpart suffers from a significant performance loss compared to the user selection schemes of Algorithm 1 and Algorithm 2, indicating that the choice of user set based on largest channel gains is inferior compared to that, which considers both the channel gains and the AoDs of various users.*

Fig. 3 portrays the average number of users pruned in Algorithm 2 as a function of α. It is clear from Fig. 3 that a significant number of users are pruned even when α is as large as 0.5. In the subsequent part of the paper, we fix $\alpha = 0.5$ for studying the performance of Algorithm 2 that employs $m_{i,k}^{(1)}$ of (18).

b: ALGORITHM 2 WITH $m_{i,k}^{(2)}$ (19)

Figure 4 compares the achievable sum rate of the IMD-HBF and EGT-HBF based BD schemes as a function of α, where both the schemes are assumed to be operating at an SNR of -10dB and use the metric given by (19), while employing Algorithm 2. In contrast to Algorithm 2 employing $m_{i,k}^{(1)}$ of (18), the sum rate performance converges to that of Algorithm 1 for smaller values of α. While this is an important benefit, the limitation of using the metric $m_{i,k}^{(2)}$ of (19) is that of the reduced number of pruned users, as portrayed in Fig. 5. In contrast to Fig. 3, in Fig. 5 we see a sudden drop in the number of users pruned. This is not surprising, since the chances of finding the users with marginal overlap in their AoD subspaces is very low. In the subsequent part of the paper, we fix $\alpha = 0.2$ while studying the performance of Algorithm 2 that employs $m_{i,k}^{(2)}$ of (19).[7] Furthermore, it can be seen from both Fig. 2 and Fig. 4 that the CPS scheme suffers from a significant performance degradation compared to the proposed algorithms, indicating that the knowledge of the AoD in addition to channel gains is crucial for attaining a good sum rate performance. It can be observed from Fig. 2(a)

[7]Note that the values of $\alpha = 0.5$ and $\alpha = 0.2$ for $m_{i,k}^{(1)}$ and $m_{i,k}^{(2)}$, respectively, are sufficiently good for other channel conditions where $L \in \{2, 3, 4, 5\}$. Owing to the limited space, the plots corresponding to these values of L are not included.

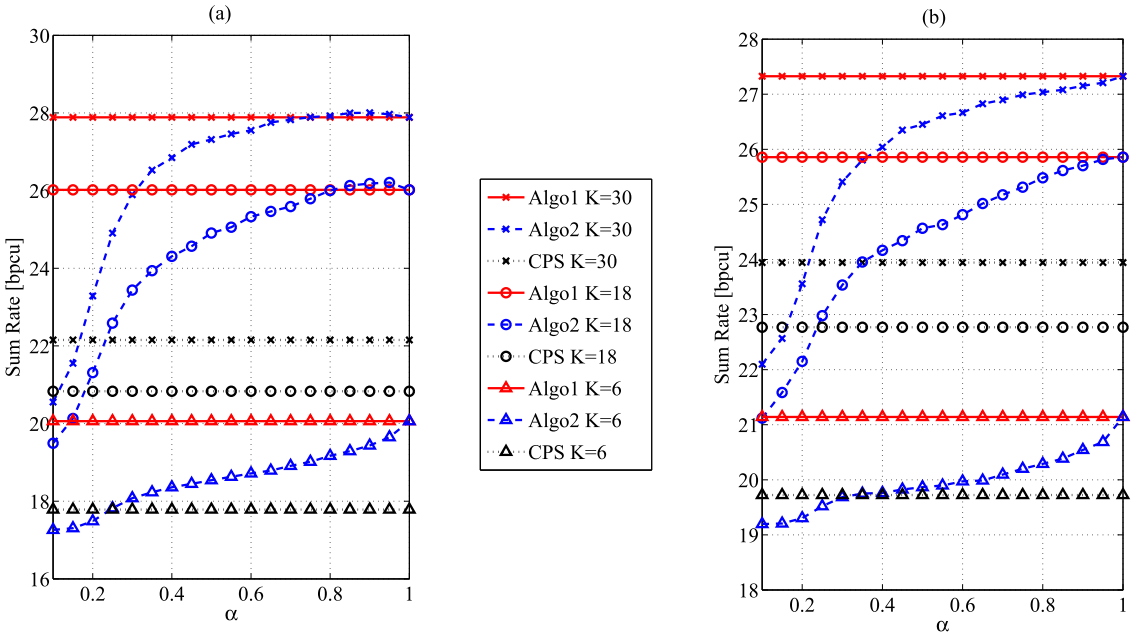


FIGURE 2. Comparison of the achievable sum rate as a function of α in the IMD-HBF and EGT-HBF based BD schemes employing Algorithm 1 and Algorithm 2. Specifically, plot (a) corresponds to the IMD-HBF scheme and plot (b) corresponds to the EGT-HBF scheme, where both the schemes are assumed to be operating at an SNR of -10dB and use the metric $m_{i,k}^{(1)}$ (18) while employing Algorithm 2.

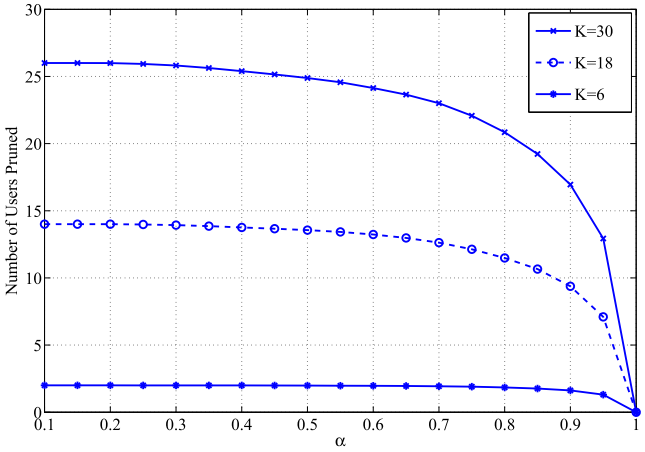


FIGURE 3. Average number of users pruned in Algorithm 2 as a function of α. Algorithm 2 is assumed to use the metric $m_{i,k}^{(1)}$ (18) for user set pruning. The plots correspond to three different values of $K \in \{6, 18, 30\}$.

and Fig. 4(a) that Algorithm 2 performs marginally better than Algorithm 1 for certain values of α, when K is large. This is not surprising, since Algorithm 2 is not optimal and hence higher sum rates are indeed achievable. Note that although the user set of Algorithm 2 is a subset of that in Algorithm 1, the former can have a different set of users due to pruning, which in turn may result in higher rates.

2) SUM RATE AS A FUNCTION OF K

In this section, we study the achievable sum rate of the IMD-HBF and EGT-HBF schemes employing Algorithm 2 as a function of K by considering $K' = 4$, $\alpha = 0.5$ when using $m_{i,k}^{(1)}$ and $\alpha = 0.2$ when using $m_{i,k}^{(2)}$ of (19). Again, we consider CPS in order to highlight the importance of the AoD information in user selection in addition to the channel gains. Furthermore, we use (20) in order to study the achievable multiplexing gains and the multiuser diversity gains offered by the proposed algorithms.

a: ALGORITHM 2 WITH $m_{i,k}^{(1)}$ (18)

Figure 6 compares the achievable sum rate in the IMD-HBF and EGT-HBF schemes as a function of K for various values of SNR. It is evident from Figure 6 that the performance of Algorithm 2 is nearly the same as that of Algorithm 1, especially for larger values of K. Specifically, when we have $K = 20$ and SNR = 0dB, there is about 1 bpcu degradation in case of IMD-HBF and about 2 bpcu degradation in case of EGT-HBF. Furthermore, it is evident that the sum rate attained at high SNR values scales according to (20) in both the IMD-HBF and EGT-HBF schemes. Specifically, the BF efficiency observed in case of IMD-HBF is about $\eta = 98\%$ and that in case of EGT-HBF is about $\eta = 85\%$. Thus, we conclude that the proposed algorithms guarantee achieving both multiplexing gains and multiuser diversity gains. Additionally, the achievable sum rate based on exhaustive search is also provided for comparison. The computational complexity of the exhaustive search grows exponentially with K, hence we have restricted our study to $K = 18$. It can be observed from Fig. 6(a) and Fig. 6(b) that the sum rate performance gap between the proposed algorithms and the exhaustive search is marginal at low SNRs and only grows slowly as the SNR

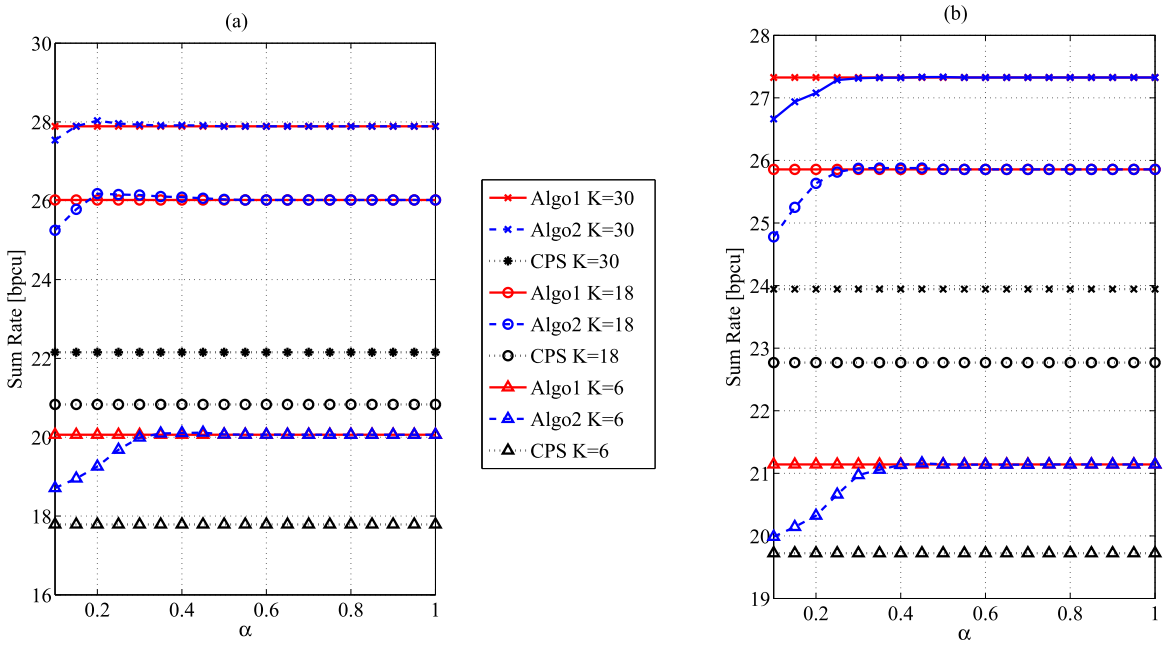


FIGURE 4. Comparison of the achievable sum rate as a function of α in the IMD-HBF and EGT-HBF based BD schemes employing Algorithm 1 and Algorithm 2. Specifically, plot (a) corresponds to the IMD-HBF scheme and plot (b) corresponds to the EGT-HBF scheme, where both the schemes are assumed to be operating at an SNR of -10dB and use the metric $m_{i,k}^{(2)}$ (19) while employing Algorithm 2.

increases. Specifically, at an SNR of -10dB and $K = 14$ the performance of Algorithm 1 is about 3 bpcu lower than that of the exhaustive search in case of IMD-HBF and that in case of EGT-HBF is about 2.65 bpcu. For the same values of SNR and K, the performance of Algorithm 2 is about 4.4 bpcu lower than that of the exhaustive search in case of IMD-HBF and that in case of EGT-HBF is about 4.2 bpcu.

b: ALGORITHM 2 WITH $m_{i,k}^{(2)}$ (19)

Figure 7 compares the achievable sum rate of the IMD-HBF and EGT-HBF schemes as a function of K for various values of the SNR. It is clear from Figure 6 that the performance of Algorithm 2 is the same as that of Algorithm 1, especially for larger values of K, unlike the case of Algorithm 2 using $m_{i,k}^{(1)}$, where a marginal degradation was observed. Furthermore, it is evident that the sum rate achieved at high SNR values scales according to (20) in both the IMD-HBF and EGT-HBF schemes, as in the case of Algorithm 2 using $m_{i,k}^{(1)}$. It can be seen from both Fig. 6 and Fig. 7 that the CPS scheme suffers from a significant performance degradation compared to the proposed algorithms, indicating that the knowledge of the AoD in addition to the channel gains is crucial for attaining a good sum rate performance. It can be observed from Fig. 7(a) and Fig. 7(b) that the sum rate performance gap between the proposed algorithms and the exhaustive search is marginal at low SNRs but increases with the SNR, as observed in the earlier case. Specifically, at an SNR of -10dB and $K = 14$ the performance of Algorithm 1 is about 3.1 bpcu lower than that of the exhaustive search in case of IMD-HBF and that in case of EGT-HBF is about 3.2 bpcu. Furthermore, it can be observed from Fig. 7(a) and Fig. 7(b) that the performance of Algorithm 2 is nearly the same as that of Algorithm 1.

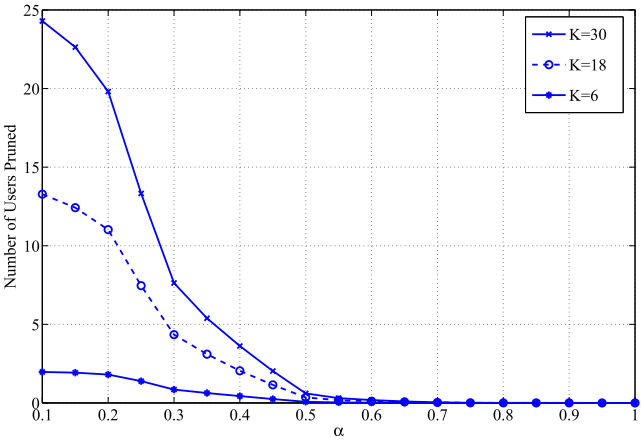


FIGURE 5. Average number of users pruned in Algorithm 2 as a function of α. Algorithm 2 is assumed to use the metric $m_{i,k}^{(2)}$ (19) for user set pruning. The plots correspond to three different values of $K \in \{6, 18, 30\}$.

3) SUM RATE IN RR SCHEDULING

In this section, we study the achievable sum rate as a function of the SNR in the IMD-HBF and EGT-HBF schemes employing Algorithm 3 by considering $K' = 4$, $K = 8$, $\alpha = 0.5$ when using $m_{i,k}^{(1)}$ of (18) and $\alpha = 0.2$ when using $m_{i,k}^{(2)}$ of (19). Figure 8 compares the sum rate of the IMD-HBF

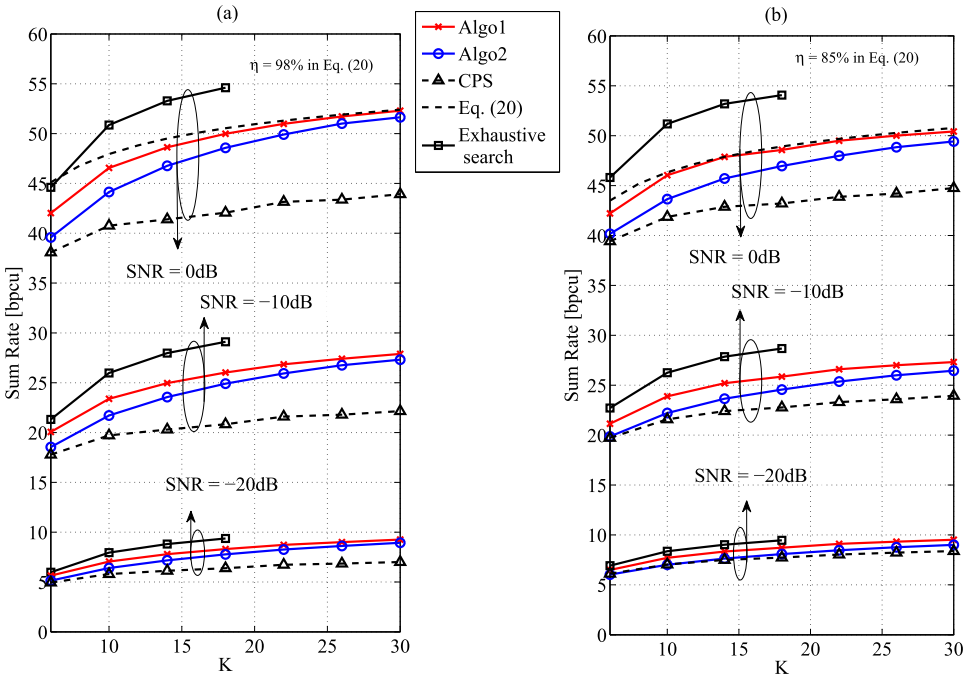


FIGURE 6. Comparison of the achievable sum rate as a function of K in the IMD-HBF and EGT-HBF based BD schemes employing Algorithm 1 and Algorithm 2. Specifically, plot (a) corresponds to the IMD-HBF scheme and plot (b) corresponds to the EGT-HBF scheme, where both the schemes are assumed to use the metric $m_{i,k}^{(1)}$ (18) while employing Algorithm 2. The high SNR approximation of sum rate in (20) with appropriately chosen η is also provided for reference.

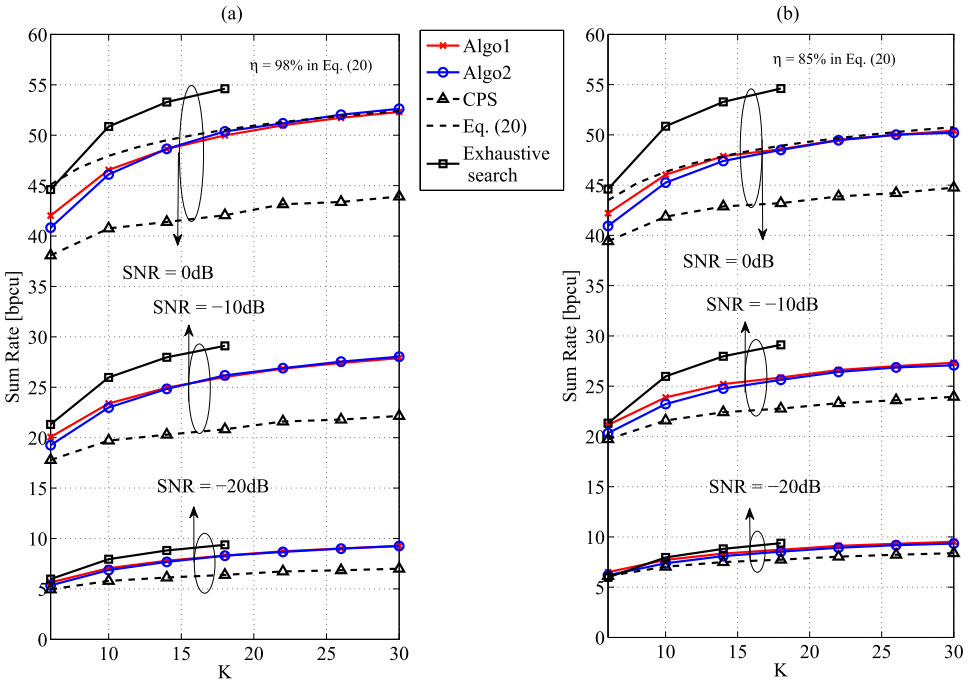


FIGURE 7. Comparison of the achievable sum rate as a function of K in the IMD-HBF and EGT-HBF based BD schemes employing Algorithm 1 and Algorithm 2. Specifically, plot (a) corresponds to the IMD-HBF scheme and plot (b) corresponds to the EGT-HBF scheme, where both the schemes are assumed to use the metric $m_{i,k}^{(2)}$ (19) while employing Algorithm 2. The high SNR approximation of sum rate in (20) with appropriately chosen η is also provided for reference.

and EGT-HBF schemes in the RR scheduling scenario, where the scheduler is assumed to use $m_{i,k}^{(1)}$ of (18). It can be seen from the inset provided in Fig. 8(a) that the CPS suffers from about 1.2 bpcu degradation compared to the case operating without pruning. However, in case of the EGT-HBF scheme, it is evident that at an SNR of -10dB the CPS scheme suffers

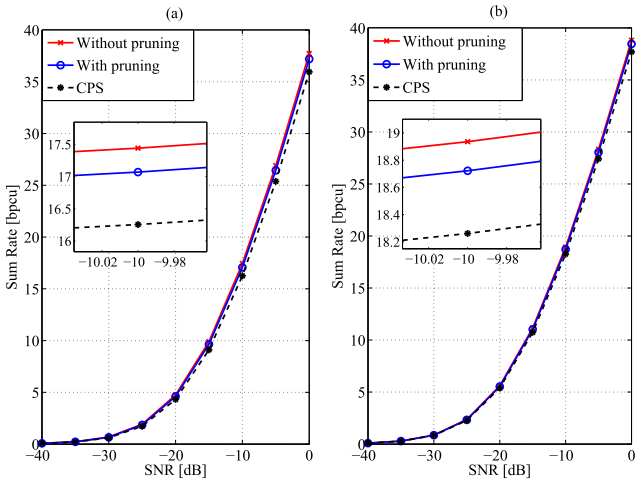


FIGURE 8. Comparison of the achievable sum rate in IMD-HBF and EGT-HBF schemes operating in the RR scheduling scenario. The RR scheduling algorithm is assumed to employ $m_{i,k}^{(1)}$ (18). The achievable rate when employing CPS algorithm is also provided for reference.

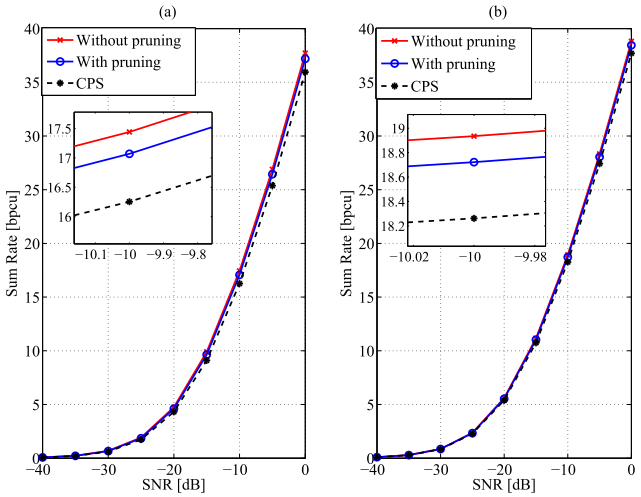


FIGURE 9. Comparison of the achievable sum rate in IMD-HBF and EGT-HBF schemes operating in the RR scheduling scenario. The RR scheduling algorithm is assumed to employ $m_{i,k}^{(2)}$ (19). The achievable rate when employing CPS algorithm is also provided for reference.

only about 0.7 bpcu compared to the case operating without pruning. Similar observations can be made in case of the RR scheduling, which employs $m_{i,k}^{(2)}$ of (19), whose sum rate is provided in Fig. 9.

4) FUTURE WORK

Some of the potential open problems related to the solutions presented are discussed as follows. In this work, we have studied the user selection problem by considering two different metrics that are related to the angle between the interference as well as the signal subspaces and the dimension of the signal and interference spaces. Note that there are several other distance measures between subspaces, such as the *Chordal*, *Procrustus*, *Fubini* and other distances [40], which are to be explored for their suitability to optimal and efficient user selection schemes. The solution to the problem of user association in cellular communication mainly relies on the signal strength. In case of cell edge users, this is not a suitable metric, since the signal power gleaned from all the surrounding BSs will be nearly same. In case of mm-wave communication, the AoA/AoD play an important role in addition to the signal strength. Thus, the problem of user association for the cell edge users in case of mm-wave communication requires a thorough investigation, where the proposed metrics would play an important role. Note that for the case, where the users have different SNR values as for the cell edge users, the sum rate metric and the RR scheduling algorithm have to be suitably modified in order to account for different user SNR values. These limitations have to be overcome in our future work. Additionally, as stated in Section III-B, in each iteration of the RR scheduling algorithm the users are chosen optimally, while not taking into account the impact of the residual users on the sum rate performance in the remaining RR slots. Further improvements might be possible, if the impact of the residual users is taken into consideration.

V. CONCLUSIONS

We have studied the user selection problem in the context of mm-wave communication, which has hitherto not been explored in the literature. Specifically, user selection algorithms have been proposed, which rely only on the knowledge of channel gains and AoD of the channel paths of various users at the BS. Furthermore, two novel metrics have been formulated for eliminating the interfering set of users and for reducing the user set space. Both the proposed AoD-US and AoD-US-P algorithms have been shown to achieve both multiuser diversity and multiplexing gains, when employed in the IMD-HBF and EGT-HBF schemes. Furthermore, both the proposed algorithms have been studied in the RR scheduling scenario, where both the algorithms attained nearly the same performance, while outperforming the channel power based selection scheme.

APPENDIX A

Proof: The proof is straightforward and is provided for the sake of completeness. From (12), we have

$$\begin{aligned}
\bar{\mathbf{E}}_t^{(i_1)} &= \mathbf{G}_1 \\
\bar{\mathbf{E}}_t^{(i_2)} &= \mathbf{G}_2 + \mathbf{G}_1 \Xi_{1,2} \\
\bar{\mathbf{E}}_t^{(i_3)} &= \mathbf{G}_3 + \mathbf{G}_1 \Xi_{1,3} + \mathbf{G}_2 \Xi_{2,3} \\
&\vdots \\
\bar{\mathbf{E}}_t^{(i_k)} &= \mathbf{G}_k + \mathbf{G}_1 \Xi_{1,k} + \cdots + \mathbf{G}_{k-1} \Xi_{k-1,k}.
\end{aligned}$$

Multiplying both sides of the l^{th} equation by $(\bar{\mathbf{E}}_t^{(i_l)H} \bar{\mathbf{E}}_t^{(i_l)})^{-1} \bar{\mathbf{E}}_t^{(i_l)H}$, we have (22)-(24) as shown at the top of the next page. Since, we have

$$\mathbf{G}_{i-1} \Xi_{i-1,k} (\bar{\mathbf{E}}_t^{(i_k)H} \bar{\mathbf{E}}_t^{(i_k)})^{-1} \bar{\mathbf{E}}_t^{(i_k)H} = \mathbf{P}(\mathbf{G}_i) \mathbf{P}(\bar{\mathbf{E}}_t^{(i_k)}), \quad (25)$$

(22)-(24) as shown at the top of this page, reduce to (24)-(28) as shown at the top of the next page. This concludes the proof. ■

$$\bar{\mathbf{E}}_t^{(i_1)}(\bar{\mathbf{E}}_t^{(i_1)H}\bar{\mathbf{E}}_t^{(i_1)})^{-1}\bar{\mathbf{E}}_t^{(i_1)H} = \mathbf{G}_1(\bar{\mathbf{E}}_t^{(i_1)H}\bar{\mathbf{E}}_t^{(i_1)})^{-1}\bar{\mathbf{E}}_t^{(i_1)H} \tag{22}$$

$$\bar{\mathbf{E}}_t^{(i_2)}(\bar{\mathbf{E}}_t^{(i_2)H}\bar{\mathbf{E}}_t^{(i_2)})^{-1}\bar{\mathbf{E}}_t^{(i_2)H} = \mathbf{G}_2(\bar{\mathbf{E}}_t^{(i_2)H}\bar{\mathbf{E}}_t^{(i_2)})^{-1}\bar{\mathbf{E}}_t^{(i_2)H} + \mathbf{G}_1\Xi_{1,2}(\bar{\mathbf{E}}_t^{(i_2)H}\bar{\mathbf{E}}_t^{(i_2)})^{-1}\bar{\mathbf{E}}_t^{(i_2)H} \tag{23}$$

$$\vdots$$

$$\bar{\mathbf{E}}_t^{(i_k)}(\bar{\mathbf{E}}_t^{(i_k)H}\bar{\mathbf{E}}_t^{(i_k)})^{-1}\bar{\mathbf{E}}_t^{(i_k)H} = \mathbf{G}_k(\bar{\mathbf{E}}_t^{(i_k)H}\bar{\mathbf{E}}_t^{(i_k)})^{-1}\bar{\mathbf{E}}_t^{(i_k)H} + \mathbf{G}_1\Xi_{1,k}(\bar{\mathbf{E}}_t^{(i_k)H}\bar{\mathbf{E}}_t^{(i_k)})^{-1}\bar{\mathbf{E}}_t^{(i_k)H} + \cdots + \mathbf{G}_{k-1}\Xi_{k-1,k}(\bar{\mathbf{E}}_t^{(i_k)H}\bar{\mathbf{E}}_t^{(i_k)})^{-1}\bar{\mathbf{E}}_t^{(i_k)H} \tag{24}$$

$$\mathbf{P}(\bar{\mathbf{E}}_t^{(i_1)}) = \mathbf{G}_1(\bar{\mathbf{E}}_t^{(i_1)H}\bar{\mathbf{E}}_t^{(i_1)})^{-1}\bar{\mathbf{E}}_t^{(i_1)H} = \mathbf{P}(\mathbf{G}_1) \tag{26}$$

$$\mathbf{P}(\bar{\mathbf{E}}_t^{(i_2)}) = \mathbf{G}_2(\bar{\mathbf{E}}_t^{(i_2)H}\bar{\mathbf{E}}_t^{(i_2)})^{-1}\bar{\mathbf{E}}_t^{(i_2)H} + \mathbf{P}(\mathbf{G}_1)\mathbf{P}(\bar{\mathbf{E}}_t^{(i_2)}) \tag{27}$$

$$\vdots$$

$$\mathbf{P}(\bar{\mathbf{E}}_t^{(i_k)}) = \mathbf{G}_k(\bar{\mathbf{E}}_t^{(i_k)H}\bar{\mathbf{E}}_t^{(i_k)})^{-1}\bar{\mathbf{E}}_t^{(i_k)H} + \sum_{i=1}^{k-1}\mathbf{P}(\mathbf{G}_i)\mathbf{P}(\bar{\mathbf{E}}_t^{(i_k)}). \tag{28}$$

APPENDIX B

Proof: Considering $\|\cdot\|^2$ of (14), we have

$$\|\mathbf{P}(\bar{\mathbf{E}}_t^{(i_k)})\|^2 = \|\mathbf{G}_k(\bar{\mathbf{E}}_t^{(i_k)H}\bar{\mathbf{E}}_t^{(i_k)})^{-1}\bar{\mathbf{E}}_t^{(i_k)H} + \sum_{i=1}^{k-1}\mathbf{P}(\mathbf{G}_i)\mathbf{P}(\bar{\mathbf{E}}_t^{(i_k)})\|^2. \tag{29}$$

We have $\|\mathbf{P}(\bar{\mathbf{E}}_t^{(i_k)})\|^2 = \sum_i^L \lambda_i$, where λ_i denotes the i^{th} eigenvalue of $\mathbf{P}(\bar{\mathbf{E}}_t^{(i_k)})$. Since $\mathbf{P}(\bar{\mathbf{E}}_t^{(i_k)})$ is a projection matrix, we have $\lambda_i = 1$ for $1 \le i \le \mathtt{rank}(\bar{\mathbf{E}}_t^{(i_k)})$. We have $\mathtt{rank}(\bar{\mathbf{E}}_t^{(i_k)}) = L$ with probability one, since the AoDs of channel paths are assumed to be from continuous uniform distribution over $[0, 2\pi]$. Let $\Upsilon_{i,k}$ represent $\mathbf{P}(\mathbf{G}_i)\mathbf{P}(\bar{\mathbf{E}}_t^{(i_k)})$ and Γ_k represent $\mathbf{G}_k(\bar{\mathbf{E}}_t^{(i_k)H}\bar{\mathbf{E}}_t^{(i_k)})^{-1}\bar{\mathbf{E}}_t^{(i_k)H}$. From (29), we have

$$\|\Gamma_k + \sum_{i=1}^{k-1}\Upsilon_{i,k}\|^2 = \mathtt{Tr}\left\{\left(\Gamma_k + \sum_{i=1}^{k-1}\Upsilon_{i,k}\right)^H\left(\Gamma_k + \sum_{i=1}^{k-1}\Upsilon_{i,k}\right)\right\}. \tag{30}$$

Furthermore, we have

$$\Gamma_k^H\Upsilon_{i,k} = \bar{\mathbf{E}}_t^{(i_k)}(\bar{\mathbf{E}}_t^{(i_k)H}\bar{\mathbf{E}}_t^{(i_k)})^{-1}\mathbf{G}_k^H\mathbf{P}(\mathbf{G}_i)\mathbf{P}(\bar{\mathbf{E}}_t^{(i_k)}) = \mathbf{O}, \tag{31}$$

since the columns of $\mathbf{G}_k$ and $\mathbf{G}_i$ are orthogonal to each other. Similarly, it can be shown that $\Upsilon_{i,k}^H\Upsilon_{j,k} = \mathbf{O}$ for $i \ne j$. Thus, we have

$$\|\Gamma_k + \sum_{i=1}^{k-1}\Upsilon_{i,k}\|^2 = \mathtt{Tr}\left(\Gamma_k^H\Gamma_k + \sum_{i=1}^{k-1}\Upsilon_{i,k}^H\Upsilon_{i,k}\right), \tag{32}$$

$$= \mathtt{Tr}\left(\Gamma_k^H\Gamma_k\right) + \sum_{i=1}^{k-1}\mathtt{Tr}\left(\Upsilon_{i,k}^H\Upsilon_{i,k}\right). \tag{33}$$

Thus, we have

$$L = \mathtt{Tr}\left(\Gamma_k^H\Gamma_k\right) + \sum_{i=1}^{k-1}\mathtt{Tr}\left(\Upsilon_{i,k}^H\Upsilon_{i,k}\right), \tag{34}$$

$$= \|\mathbf{G}_k(\bar{\mathbf{E}}_t^{(i_k)H}\bar{\mathbf{E}}_t^{(i_k)})^{-1}\bar{\mathbf{E}}_t^{(i_k)H}\|^2 + \sum_{i=1}^{k-1}\|\mathbf{P}(\mathbf{G}_i)\mathbf{P}(\bar{\mathbf{E}}_t^{(i_k)})\|^2. \tag{35}$$

This concludes the proof. ■

ACKNOWLEDGMENTS

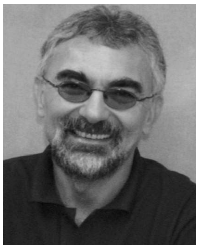
The authors would like to thank the Editor as well as the anonymous reviewers for their valuable suggestions, which significantly improved the paper. The authors also acknowledge the use of the IRIDIS High Performance Computing Facility, and associated support services at the University of Southampton, in the completion of this work. All data supporting this study are openly available from the University of Southampton repository under http://doi.org/10.5258/SOTON/D0038.

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, May 2013.
[3] J. G. Andrews *et al.*, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.
[4] F. Boccardi, R. W. Heath, Jr., A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.
[5] W. Roh *et al.*, "Millimeter-wave beamforming as an enabling technology for 5G cellular communications: Theoretical feasibility and prototype results," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 106–113, Feb. 2014.
[6] Q. C. Li, G. Wu, and T. S. Rappaport, "Channel model for millimeter-wave communications based on geometry statistics," in *Proc. Globecom Workshops (GC Wkshps)*, Dec. 2014, pp. 427–432.
[7] M. K. Samimi and T. S. Rappaport, "3-D statistical channel model for millimeter-wave outdoor mobile broadband communications," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2015, pp. 2430–2436.
[8] X. Zhang, A. F. Molisch, and S.-Y. Kung, "Variable-phase-shift-based RF-baseband codesign for MIMO antenna selection," *IEEE Trans. Signal Process.*, vol. 53, no. 11, pp. 4091–4103, Nov. 2005.
[9] P. Sudarshan, N. B. Mehta, A. F. Molisch, and J. Zhang, "Channel statistics-based RF pre-processing with antenna selection," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 3501–3511, Dec. 2006.
[10] S. Han, I. Chih-Lin, Z. Xu, and C. Rowell, "Large-scale antenna systems with hybrid analog and digital beamforming for millimeter wave 5G," *IEEE Commun. Mag.*, vol. 53, no. 1, pp. 186–194, Jan. 2015.

[11] O. E. Ayach, R. W. Heath, Jr., S. Abu-Surra, S. Rajagopal, and Z. Pi, "Low complexity precoding for large millimeter wave MIMO systems," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2012, pp. 3724–3729.
[12] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, Jr., "Hybrid precoding for millimeter wave cellular systems with partial channel knowledge," in *Proc. Inf. Theory Appl. Workshop (ITA)*, San Diego, CA, USA, Feb. 2013, pp. 1–5.
[13] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, Jr., "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 831–846, Oct. 2014.
[14] W. Ni, X. Dong, and W.-S. Lu. (2015). "Near-optimal hybrid processing for massive MIMO systems via matrix decomposition." [Online]. Available: https://arxiv.org/abs/1504.03777
[15] J. Singh and S. Ramakrishna, "On the feasibility of codebook-based beamforming in millimeter wave systems with multiple antenna arrays," *IEEE Trans. Wireless Commun.*, vol. 14, no. 5, pp. 2670–2683, May 2015.
[16] L. Dai, X. Gao, J. Quan, S. Han, and C.-L. I, "Near-optimal hybrid analog and digital precoding for downlink mmWave massive MIMO systems," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2015, pp. 1334–1339.
[17] V. Raghavan, S. Subramanian, J. Cezanne, and A. Sampath, "Directional beamforming for millimeter-wave MIMO systems," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2015, pp. 1–7.
[18] R. Rajashekar and L. Hanzo, "Hybrid beamforming in mm-wave MIMO systems having a finite input alphabet," *IEEE Trans. Commun.*, vol. 64, no. 8, pp. 3337–3349, Aug. 2016.
[19] R. Rajashekar and L. Hanzo, "Iterative matrix decomposition aided block diagonalization for mm-wave multiuser MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1372–1384, Mar. 2017.
[20] P. Cao and J. S. Thompson, "Practical multi-user transmission design in millimeter wave cellular networks: Is the joint SDMA-TDMA technique the answer?" in *Proc. IEEE 17th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jul. 2016, pp. 1–5.
[21] 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.
[22] R. A. Stirling-Gallacher and M. S. Rahman, "Multi-user MIMO strategies for a millimeter wave communication system using hybrid beam-forming," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2015, pp. 2437–2443.
[23] J. Choi, "Beam selection in mm-Wave multiuser MIMO systems using compressive sensing," *IEEE Trans. Commun.*, vol. 63, no. 8, pp. 2936–2947, Aug. 2015.
[24] A. Alkhateeb, G. Leus, and R. W. Heath, Jr., "Limited feedback hybrid precoding for multi-user millimeter wave systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 11, pp. 6481–6494, Nov. 2015.
[25] W. Ni and X. Dong, "Hybrid block diagonalization for massive multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 64, no. 1, pp. 201–211, Jan. 2016.
[26] D. J. Love and R. W. Heath, Jr., "Equal gain transmission in multiple-input multiple-output wireless systems," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, vol. 2. Nov. 2002, pp. 1124–1128.
[27] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.
[28] L.-U. Choi and R. D. Murch, "A transmit preprocessing technique for multiuser MIMO systems using a decomposition approach," *IEEE Trans. Wireless Commun.*, vol. 3, no. 1, pp. 20–24, Jan. 2004.
[29] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 461–471, Feb. 2004.
[30] G. Dimic and N. D. Sidiropoulos, "On downlink beamforming with greedy user selection: Performance analysis and a simple new algorithm," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3857–3868, Oct. 2005.
[31] 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.
[32] Z. Shen, R. Chen, J. G. Andrews, R. W. Heath, 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.
[33] M. Sharif and B. Hassibi, "A comparison of time-sharing, DPC, and beamforming for MIMO broadcast channels with many users," *IEEE Trans. Commun.*, vol. 55, no. 1, pp. 11–15, Jan. 2007.
[34] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
[35] F. Deutsch, "The angle between subspaces of a Hilbert space," in *Approximation Theory, Wavelets and Applications*, S. P. Singh, Ed. Dordrecht, The Netherlands: Kluwer, 1995, pp. 107–130.
[36] O. M. Baksalary and G. Trenkler, "On angles and distances between subspaces," *Linear Algebra Appl.*, vol. 431, no. 11, pp. 2243–2260, Nov. 2009.
[37] 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.
[38] C. Swannack, E. Uysal-Biyikoglu, and G. W. Wornell, "Low complexity multiuser scheduling for maximizing throughput in the MIMO broadcast channel," in *Proc. Allerton Conf. Commun., Control, Comput.*, Allerton, IL, USA, Oct. 2004, pp. 440–449.
[39] H. Viswanathan and K. Kumaran, "Rate scheduling in multiple antenna downlink wireless systems," *IEEE Trans. Commun.*, vol. 53, no. 4, pp. 645-655, Apr. 2005.
[40] K. Ye and L.-H. Lim. (2014). "Schubert varieties and distances between subspaces of different dimensions." [Online]. Available: https://arxiv.org/abs/1407.0900

RAKSHITH RAJASHEKAR received the B.E. degree in electrical communication engineering from Visvesvaraya Technological University, Karnataka, India, in 2007, and the Ph.D. degree from the Department of Electrical Communication Engineering, Indian Institute of Science, India, in 2014. He is currently a Research Fellow with the University of Southampton (UoS), U.K. Before joining UoS, he was with Accord Software & Systems, Bangalore, India, as a Systems Engineer, from 2007 to 2009, and as a Senior Scientist with Broadcom Communications, Bangalore, from 2014 to 2015. His current research interests include millimeter wave communication, energy efficient differential communication, with a focus on space-time signal processing and coding.



LAJOS HANZO (M'91–SM'92–F'04) received the degree in electronics in 1976 and the Ph.D. degree in 1983. During his 38-year career in telecommunications he has held various research and academic posts in Hungary, Germany, and U.K. Since 1986, he has been with the School of Electronics and Computer Science, University of Southampton, U.K. He has successfully supervised about 100 Ph.D. students. He is currently directing a 100-strong academic research team, working on a range of research projects in the field of wireless multimedia communications sponsored by industry, the Engineering and Physical Sciences Research Council, U.K., the European Research Council's Advanced Fellow Grant, and the Royal Society's Wolfson Research Merit Award. He is an enthusiastic supporter of industrial and academic liaison and he offers a range of industrial courses. He has co-authored 20 John Wiley/IEEE Press books on mobile radio communications totaling in excess of 10 000 pages, published over 1500 research entries at the IEEE Xplore. He is a fellow of the Royal Academy of Engineering, of the Institution of Engineering and Technology, and of the European Association for Signal Processing. In 2009, he received the honorary doctorate Doctor Honoris Causa by the Technical University of Budapest. He served as the TPC Chair and the General Chair of the IEEE conferences, presented keynote lectures and has been received a number of distinctions. He is also a Governor of the IEEE VTS. From 2008 to 2012, he was the Editor-in-Chief of the IEEE Press and a Chaired Professor at Tsinghua University, Beijing. He is the Chair in telecommunications with the University of Southampton. He has 22,000+ citations.