User-Directed Analog Beamforming for Multiuser Millimeter-Wave Hybrid Array Systems

J. Andrew Zhang[1], Hang Li[1], Xiaojing Huang[1], Y. Jay Guo[1] and Antonio Cantoni[2]
[1]University of Technology Sydney, Global Big Data Technologies Centre (GBDTC), Australia
[2]University of Western Australia, Perth, Australia
{Andrew.Zhang; Hang.Li; Xiaojing.Huang; Jay.Guo}@uts.edu.au; antonio.cantoni@uwa.edu.au.

Abstract—Beamforming design for millimeter-Wave hybrid array with the subarray structure is very challenging. There is neither known optimal solution that maximizes the sum rate capacity nor near-optimal solution. This paper proposes some low-complexity user-directed analog radio-frequency (RF) beamforming design schemes. The basic idea is to iteratively allocate different subarrays to different users such that users' channel correlation can be efficiently reduced via RF beamforming. Several new but less efficient schemes are also presented to shed light on RF beamforming design, and to serve as comparisons for the user-directed schemes. Simulation results are provided for these proposed schemes, existing ones in the literature and an upper-bound for hybird array with a fully-connected structure. The user-directed schemes demonstrate significantly better sum-rate and BER performance over other schemes, although the gap to the upperbound is still large.

I. INTRODUCTION

Millimeter wave (mmWave) hybrid antenna array [1], [2] is an attractive solution that balances performance, system complexity and cost for 5G cellular communications. There are two basic hybrid array structures [2]: a *fully-connected structure* where each antenna is connected to multiple phase shifters and multiple radio-frequency (RF) frontend chains; and a *subarray structure* where the array is divided into subarrays, each antenna is only connected to one phase shifter and each subarray has only one RF chain [1]. A lot of research has been focusing on the fully-connected structure as mathematically it is more tractable. It provides full flexibility and can achieve better performance, however, it requires many more phase shifters and lacks scalability as phase arrays are typically integrated with antenna elements in a chip today. For these reasons, subarray structure seems to be a better option for practical deployment.

Hybrid precoding/combining, named as beamforming (BF) for simplicity in this paper, for mm-wave hybrid array has been a topic of substantial current research. It includes both analog RF and digital BF that can be developed either independently or jointly [1]. Since RF BF vectors form part of the *equivalent digital channels*, RF BF has more significant impact on the final system capacity and performance. For the fully-connected structure, closed-form optimal solution is known and sub-optimal approximations have been reported in the literature.

But neither optimal solution maximizing multi-user sum-capacity nor near-optimal solution for subarray structure is known due to its special BF matrix of block-diagonal structure [3].

Most of existing work on BF design for mmWave hybrid array with subarray structure considers simple RF BF and focuses more on optimizing digital BF, e.g., [4]–[6]. In [4], point-to-point MIMO capacity is studied by finding phase vectors closest to the optimal eigen-vectors in terms of Frobenius norm. For multiuser MIMO, [5] proposes a multi-stage hybrid diversity maximization precoding, which combines channel diagonalization to minimize multiuser interference and sum-mean-square-error optimization to minimize mean symbol error. In [6], hybrid block diagonalization is proposed by combining phase-only equal-gain transmission (EGT) for subarray precoding and digital block diagonalization in baseband. The achieved sum-rate is shown to approach to that achieved by block diagonalization in a full digital array. The authors in [3] consider a different approach to optimize the RF BF directly. They propose an iterative approach to optimize the achievable capacity of each subarray from the first to the last, utilizing the idea of successive interference cancelation (SIC). Each BF vector in this case is found to be the singular-vector of an updated matrix in each iteration. Although claimed as a near-optimal solution, it is actually sub-optimal and achieves capacity far from a benchmark as is discussed in Section IV-B.

In this paper, we propose some user-directed BF schemes and provide comprehensive comparison with known and other new but less efficient schemes. The basic idea of our user-directed BF schemes is to allocate different subarrays to different users such that users' channel correlation can be reduced while optimizing other metrics such as signal power. This turns out to be an effective approach. We also present several schemes that one would intuitively expect to work but simulation results prove otherwise. These methods include generating BF vectors leading to orthogonal equivalent channel matrix, least square approximation to the eigen-beamforming for the fully-connected structure, and the SIC version of our system to [3]. Simulation results show that the proposed user-directed schemes achieve much better performance than the others, although there still exist a large gap between them and the upper-bound for the fully-connected structure.

Notation: We use superscript $(\cdot)^T$ for transpose, $(\cdot)^H$ for conjugate transpose, and $(\cdot)^*$ for conjugate.

II. PROBLEM FORMULATION

We consider a multiuser MIMO system where a base-station (BS) with a one-dimensional localized hybrid array [1] communicates with a group of K users each with a single omnidirectional antenna. The hybrid array contains M linearly positioned analog subarrays. Each subarray is a uniform linear array (ULA) and has N antenna elements. The distance between neighbouring elements is $\lambda/2$, where λ is the wavelength, and the distance between subarrays (between the first elements of two neighbouring subarrays) is assumed to be $d = N\lambda/2$. Assume that each antenna element in the hybrid array has an omni-directional radiation pattern, i.e., spatial response 1 over all angles. For simplicity, we only consider the case of $M = K$ and a narrow-band beamforming model in this paper. We consider the multiuser access uplink channel, and the proposed schemes can be easily adapted to the downlink channel. The results in this paper shall be extended to subarrays of other structures such as square array.

The purpose of this paper is to do some general investigation for different BF schems, and hence we do not constraint the values of the RF BF vector. Let $\mathbf{w}_m, |\mathbf{w}_m| = 1$ be the BF vector for the m-th subarray, where $|\cdot|$ is used to denote the absolute value of a single element, the norm of a vector or the determinant of a matrix. Let the $NM \times K$ channel matrix for the whole array be $\mathbf{H} = (\mathbf{H}_1^T, \cdots, \mathbf{H}_M^T)^T$, where $\mathbf{H}_m = (\mathbf{h}_{m,1}, \cdots, \mathbf{h}_{m,K})$ and $\mathbf{h}_{m,k}$ denotes the $N \times 1$ channel vector between the N antenna elements in the m-th subarray and user k. The equivalent digital baseband channel between the BS and users can be represented as

$$\mathbf{Q} = ((\mathbf{w}_1^T \mathbf{H}_1)^T, \cdots, (\mathbf{w}_M^T \mathbf{H}_M)^T)^T. \tag{1}$$

Consider a narrow-band spatially sparse multipath channel model for each user, where the line-of-sight (LOS) path is dominating. Assume a plane wave propagation where each multipath signal arrives at all antenna elements in parallel. Then $\mathbf{h}_{m,k}$ can be expressed as

$$\mathbf{h}_{m,k} = \sum_{\ell=1}^{L_k} a_{m,k,\ell} \mathbf{v}_{k,\ell} = \mathbf{V}_k \mathbf{a}_{m,k}, \tag{2}$$

where $\mathbf{V}_k = (\mathbf{v}_{k,1}, \cdots, \mathbf{v}_{k,L_k})$, $\mathbf{v}_{k,\ell} = (1, e^{j\pi \sin(\theta_{k,\ell})}, \cdots, e^{j\pi(N-1)\sin(\theta_{k,\ell})})^T$ is the subarray response to the ℓ-th spatial multipath signal with angle of arrival (AoA) $\theta_{k,\ell}$, $\mathbf{a}_{m,k} = (a_{m,k,1}, \cdots, a_{m,k,L_k})^T$ and $a_{m,k,\ell} = b_{k,\ell} e^{j\pi m N \sin(\theta_{k,\ell})}$ is the complex channel amplitude including the magnitude of the ℓ-th multipath $b_{k,\ell}$, and phase difference between subarrays. Constant phase differences between subarrays and users that will not affect the results are ignored. Let η denote the *path-power-ratio* between the power of the dominating LOS path and the mean power of other paths.

In this paper, we refer the BF design to maximizing the sum-rate capacity for the uplink multiple access channel (MAC) that is achievable by dirty paper coding [7]. Without the uplink channel knowledge, each user's best option is to transmit at its maximum power, which is assumed to be the same $p_k =$

p. We use the following upper-bound equation for the sum-rate capacity for MAC [7] as the optimality condition for BF optimization

$$C \triangleq \log \left| \mathbf{I} + \frac{p}{\varrho^2} \mathbf{Q}\mathbf{Q}^H \right|, \tag{3}$$

where $\mathbf{I}$ is an identity matrix, ϱ^2 is the combined noise variance in each subarray and is assumed to be the same for all subarrays. For simplicity, we will just call C sum-rate.

There are no known optimal solutions for $\mathbf{w}_m$s that maximize (3) yet. The sum-rate capacity of a SDMA system is collective effects of received signal power and the correlation between different channels. Next, we investigate a set of solutions from each of these two directions.

III. USER-DIRECTED ITERATIVE BEAMFORMING DESIGN

The basic idea here is to determine BF vectors for subarrays one-by-one, using an iterative cancellation approach. When one subarray's BF vector is determined, its corresponding channel or signal will be removed from the rest for determining the BF of other subarrays. This process attempts to reduce channel correlation while allowing optimization of a chosen metric. We consider the details of two approaches, namely user based and signal base.

A. User-based Approach

This approach determines the allocation of subarrays to users. The actual implementation can be optimization in either one-dimension or two-dimension. In one-dimension optimization, we can allocate a subarray to one user through applying one of the following rules: (1) Allocate a subarray to user k if the k-th column vector in $\mathbf{H}_m$ has the maximum power, and (2) Allocate subarray m' to user k if $\mathbf{h}_{m',k}$ has the maximum power among all $\mathbf{h}_{m,k}, m = 1, \cdots, M$. The two-dimension algorithm is a joint optimization process across both column and row dimensions.

1) 1-D Optimization: In 1-D optimization, we actually want to find a subarray for the right user, or find a user for the right subarray. Define the set of users as $\mathcal{U}$. The iterative process of the one-dimension algorithm using the first rule is exemplified as follows.

S1: Initialization: $\mathcal{U} = [1, \cdots, K], m = 1$;
S2: In the m-th iteration, for $\mathbf{H}_m$, among the columns with indexes in $\mathcal{U}$, find the one with largest power. Denote this column as $\mathbf{h}_{m,k}$;
S3: Generate the BF vector $\mathbf{w}_m$ for subarray m based on $\mathbf{h}_{m,k}$;
S4: Update $\mathcal{U}$ by removing k from it. Let $m = m + 1$ and repeat from S2, until all BF vectors are determined.

In S3, $\mathbf{w}_m$ can be generated from $\mathbf{h}_{m,k}$ using either equal gain combining (EGC) or maximal ratio combining (MRC). For EGC, $\mathbf{w}_m = \exp(-j\angle(\mathbf{h}_{m,k}))/\sqrt{N}$ where the $\angle(\cdot)$ operation is element-wise; for MRC, $\mathbf{w}_m = \mathbf{h}_{m,k}^*/|\mathbf{h}_{m,k}|$.

2) 2-D Optimization: For 2-D optimization, in each iteration we try to find the right subarray for the right user. Define the user set $\mathcal{U}_t$ and subarray set $\mathcal{S}_t$ at the t-th iteration. Both sets will be updated in each iteration once an allocation is done. The algorithm is similar to that in 1-D optimization, but iteration can start from and continue to any subarray and user. The processing in the t-th iteration can be described as follows: Find

$$m_k = \arg \max_{m \in \mathcal{S}_t} |\mathbf{h}_{m,k}|^2, \ \epsilon_k = |\mathbf{h}_{m_k,k}|^2, \ \text{for all } k \in \mathcal{U}_t,$$

and determine

$$\hat{k} = \arg \max_{k \in \mathcal{U}_t} \epsilon_k. \tag{4}$$

Then the $m_{\hat{k}}$-th subarray is allocated to the $\hat{k}$-th user, and $\mathcal{U}_t$ and $\mathcal{S}_t$ are updated by removing $\hat{k}$-th user and $m_{\hat{k}}$-th subarray, respectively.

B. Signal-based Approach

Instead of allocating subarrays to users directly based on signal power, we compute the eigen-beamforming of an updated channel matrix by removing a user's signal in each iteration. In this way, each subarray will be assigned to a different user implicitly in each iteration. The process is as follows.

S1: Initialization: $\bar{\mathbf{H}}_1 = \mathbf{H}_1$, $m = 1$.
S2: Compute the singular value decomposition (SVD) of $\bar{\mathbf{H}}_m \bar{\mathbf{H}}_m^H = \mathbf{U}_h \mathbf{D}_h \mathbf{U}_h^H$, where the diagonal elements in $\mathbf{D}$ are ordered descendingly. Determine $\mathbf{w}_m$ as the conjugate of the first column of $\mathbf{U}_h$;
S3: Decide the allocation of subarray m to a user implicitly: Find the element with maximal power in $\mathbf{w}_m^T \bar{\mathbf{H}}_m$, and allocate subarray m to the user corresponding to this element (through its original index in $\mathbf{H}_m$.);
S4: Update: Let $m = m + 1$. Construct $\bar{\mathbf{H}}_m$ by removing the columns corresponding to all users that have been allocated in the previous iterations. Repeat from S2.

Using eigen-beamforming in each iteration can lead to a local optimal BF solution that would maximize the rate of a single specific subarray if it is the only subarray in the array. By implicitly allocating the eigen-beamforming to a user and removing the user's signal from the following calculations, we can make sure that each new eigen-beamforming could be optimized with respect to other users. Hence this achieves low correlation between all users' equivalent channels. The complexity of this approach is much higher than the user-based approach due to the SVD operations.

IV. OTHER APPROACHES FOR COMPARISON

In this section, we provide a sum-rate upper-bound for the fully-connected structure and use it as a benchmark for our proposed schemes. We also present several other schemes that we have tried. Intuitively, these schemes would work as they were developed from one or two directions of achieving high sum-rate, by, e.g., maximizing the total received power or making the equivalent channel vectors orthogonal. However, simulation results demonstrate that they do not perform as expected. They are provided here as comparisons to the user-directed schemes, as well as shedding some lights on the RF BF design.

A. Upper-bound of Fully Connected Hybrid Array and Its Least Square (LS) Approximation

We establish an upper-bound using eigen-beamforming for a fully-connected hybrid array. Define an $M \times MN$ BF matrix $\mathbf{W}$, where each element can take any value and each row vector has a norm of 1. Let the SVD of $\mathbf{H}\mathbf{H}^H$ be $\mathbf{U}_h \mathbf{D}_h \mathbf{U}_h^H$. The optimal $\mathbf{W}$ that maximizes the sum-rate is given by

$$\mathbf{W}_{\text{opt}} = \mathbf{U}_{h,s}^H, \tag{5}$$

where $\mathbf{U}_{h,s}$ is the $MN \times K$ submatrix of $\mathbf{U}_h$, corresponding to the K largest singular value of $\mathbf{H}\mathbf{H}^H$. This optimal $\mathbf{W}_{\text{opt}}$ achieves maximal sum-rate for a fully-connected array, which also serves as an upper bound for any array with the subarray structure.

Finding a sub-optimal solution by approximating the optimal vector with one in the sense of least square (LS) or minimum mean square is a common approach. Here we want to find such a sub-optimal solution for the subarray structure using sub-matrix LS approximation to $\mathbf{W}_{\text{opt}}$. That is, calculating the suboptimal solutions based on

$$\begin{aligned} \mathbf{w}_{m,\text{LS}} &= \arg \min_{\mathbf{w}_m} |\mathbf{w}_m^T \mathbf{H}_m - \mathbf{w}_{m,\text{opt}} \mathbf{H}|^2 \\ &= \mathbf{H}_m (\mathbf{H}_m^H \mathbf{H}_m)^{-1} (\mathbf{w}_{m,\text{opt}} \mathbf{H})^T, \end{aligned} \tag{6}$$

where $\mathbf{w}_{m,\text{opt}}$ is the m-th row in $\mathbf{W}_{\text{opt}}$. Normalize $\mathbf{w}_{m,\text{LS}}$ so that its power becomes 1.

Simulation results, however, do not support the effectiveness of this approach. The reason is being investigated.

B. Successive Interference Cancellation (SIC)

We derive a SIC solution following that in [3] for our system model here. The *matrix determinant lemma* states that if a matrix $\mathbf{A}$ is invertible, then

$$|\mathbf{A} + \mathbf{u}\mathbf{v}^H| = (1 + \mathbf{v}^H \mathbf{u}) |\mathbf{A}|, \tag{7}$$

where $\mathbf{u}$ and $\mathbf{v}$ are column vectors. Using this lemma, the sum-rate in (3) can be represented as

$$\begin{aligned} \mathbf{C} &= \log \left| \mathbf{I} + \frac{p}{\varrho^2} \sum_{m=1}^{M} (\mathbf{w}_m^T \mathbf{H}_m)^H (\mathbf{w}_m^T \mathbf{H}_m) \right| \\ &= \sum_{m=1}^{M} \log \left| 1 + \frac{p}{\varrho^2} (\mathbf{w}_m^T \mathbf{H}_m \mathbf{P}_{m-1}^{-1} \mathbf{H}_m^H \mathbf{w}_m^*) \right|, \end{aligned} \tag{8}$$

where

$$\mathbf{P}_m = \mathbf{I} + \frac{p}{\varrho^2} \sum_{i=1}^{m} (\mathbf{w}_i^T \mathbf{H}_i)^H (\mathbf{w}_i^T \mathbf{H}_i), \ m \in [1, M], \tag{9}$$

$$\mathbf{P}_0 = \mathbf{I}. \tag{10}$$

By applying the principle of SIC, a *sub-optimal solution* can be found for (8) by iteratively "switching on" subarrays one-by-one and optimize $\mathbf{w}_m$ starting from subarray

1. Mathematically, this sub-optimal solution maximizes each of $\log\left|1 + \frac{p}{\varrho^2}(\mathbf{w}_m^T\mathbf{H}_m\mathbf{P}_{m-1}^{-1}\mathbf{H}_m^H\mathbf{w}_m^*)\right|, m \in [1, M]$ ordered from $m = 1$ to M. This is equivalent to seeking optimal $\mathbf{w}_m$s without caring about their impact on the rest $M-m$ subarrays, but only considering the first m subarrays. Hence the solution is not really optimal as claimed in the paper [3], as validated by our simulation results.

The solution $\mathbf{w}_m$ is given by the conjugate of the singular vector corresponding to the maximal singular value in the SVD of $\mathbf{H}_m\mathbf{P}_{m-1}^{-1}\mathbf{H}_m^H$. Note that the computation complexity of this approach is quite high, and it needs to be updated for different SNR values.

C. Constructing Orthogonal Equivalent Channel Matrix

This approach tries to construct $\mathbf{Q}$ with orthogonal rows to minimize the channel correlation. For the first BF vector $\mathbf{w}_1$, we can use several different methods to get it. For example, let it be the eigen-BF vector for the first subarray, or using any of the previous approaches. Let $\mathbf{q}_m = \mathbf{w}_m^T\mathbf{H}_m$ be the m-th row of $\mathbf{Q}$, and define $\mathbf{Q}_m = (\mathbf{q}_1^T, \cdots, \mathbf{q}_m^T)^T$. Starting from $m = 1$, we want to determine $\mathbf{w}_{m+1}$ such that

$$\mathbf{Q}_m\mathbf{q}_{m+1}^H = \mathbf{Q}_m\mathbf{H}_{m+1}^H\mathbf{w}_{m+1}^* = \mathbf{0}. \tag{11}$$

Equation (11) suggests that $\mathbf{w}_{m+1}$ can be obtained from the null space of $\mathbf{Q}_m\mathbf{H}_{m+1}^H$. Let $\mathbf{g}_i$ be the i-th column vector in the null space matrix. We get $\mathbf{w}_{m+1}$ as

$$\mathbf{w}_{m+1}^* = \arg\max_{\mathbf{g}_i} |\mathbf{g}_i^H\mathbf{H}_{m+1}|^2, \tag{12}$$

so that the power of $\mathbf{q}_{m+1}$ can be maximized. The following steps describe the whole process.

S1: Initialization: given $\mathbf{w}_1$, $\mathbf{q}_1 = \mathbf{w}_1^T\mathbf{H}_1$, $\mathbf{Q}_1 = \mathbf{q}_1$, $m = 1$;
S2: Compute the null space matrix of $\mathbf{Q}_m\mathbf{H}_{m+1}^H$, and obtain $\mathbf{w}_{m+1}$ from (12);
S3: Normalize $\mathbf{w}_{m+1}$ so that its power is 1. Update $\mathbf{Q}_{m+1} = (\mathbf{Q}_m^T, \mathbf{H}_{m+1}^T\mathbf{w}_{m+1})^T$. Let $m = m + 1$ and repeat from S2.

Although this approach generates an orthogonal equivalent channel, the sum-rate and system performance is not satisfactory due to the power of the resulted channel matrix is much lower than those of other approaches, as is seen from the simulation results in Section V. This seems to suggest that orthogonalization is not a good solution given the natural signal correlation between different subarrays, although a strict theoretical analysis is not available yet.

D. Maximizing Total Received Power

If we only want to maximize received signal power from all users, we are equivalently maximizing the sum power of the diagonal elements in $\mathbf{Q}\mathbf{Q}^H$. Doing this basically disregards the channel correlation. Consider a more general problem by adding per-user rate constraints. This problem can be formulated as

$$\max_{\mathbf{W}}\ \rho(\mathbf{W}) \triangleq \sum_{m=1}^M \mathbf{w}_m^T\mathbf{H}_m\mathbf{H}_m^H\mathbf{w}_m^* = \sum_{k=1}^K\sum_{m=1}^M |\mathbf{w}_m^T\mathbf{h}_{m,k}|^2,$$

s.t.

$$\sum_{m=1}^M |\mathbf{w}_m^T\mathbf{h}_{m,k}|^2 \geq r_k \text{ (per-user rate constraints)},$$

$$|\mathbf{w}_m|^2 = 1, \text{ (Normalization constraints)}, \tag{13}$$

where r_k is the given per-user rate constraint for user k.

Without constraints on $\mathbf{w}_m$, (13) becomes a quadratic constrained quadratic programming (QCQP) problem. We can use the Lagrange multiplier to show that the optimal solution satisfies

$$\begin{cases} \left(\mathbf{H}_m\mathbf{H}_m^H - \sum_{k\in\mathcal{K}} u_k\mathbf{h}_{m,k}\mathbf{h}_{m,k}^H\right)\mathbf{w}_m = \lambda_m\mathbf{w}_m, \\ \qquad m = 1, \cdots, M, \\ \sum_{m=1}^M \mathbf{w}_m^T\mathbf{h}_{m,k}\mathbf{h}_{m,k}^H\mathbf{w}_m^* = r_k, \ k \in \mathcal{K}, \end{cases} \tag{14}$$

where $\mathcal{K}$ is a subset of $\{1, \cdots, K\}$ or empty, depending on how the per-user rate constraint is met. It can be seen that the optimal solution will be the eigenvector corresponding to the maximal eigenvalue of the matrix $\left(\mathbf{H}_m\mathbf{H}_m^H - \sum_{k\in\mathcal{K}} u_k\mathbf{h}_{m,k}\mathbf{h}_{m,k}^H\right)$. However, a closed-form expression for the solution cannot be obtained as u_ks are not explicitly known. Instead, this problem can be solved by using iterative convex optimization algorithms, e.g., the interior-point method [8].

Without any rate constrains, the optimal solution would be the eigen-beamforming, and $\mathbf{w}_m$ would be the conjugate of the singular-vector corresponding to the maximal singular value of $\mathbf{H}_m\mathbf{H}_m^H$.

Although the signal power can be maximized in this approach, there is no consideration of reducing channel correlation for different users. Hence the achieved sum-rate is very low in most of the time according to simulation results to be presented.

V. SIMULATION RESULTS

A hybrid array system with $M = 4$, $N = 8$ and $K = 4$ are simulated. Each user's multipath signals arrive in cluster, and their number L_k in each cluster is set as 6, with one LOS-path dominating in power. The multipath signals are generated following a complex zero-mean Gaussian distribution. Different variances are used for LOS and non-LOS paths with their ratio equal to a given *path-power-ratio* η. The AoAs of multipath in one cluster are uniformly distributed in the range of $[\theta_k - \delta, \theta_k + \delta]$, where θ_k is the AoA of the LOS path for user k. The default values used in simulation are $\delta = 10$ degrees, SNR= 15 dB, $\eta = 10$ dB and $\theta_k, k = 1, \cdots, K$ are generated using a uniform distribution over $[-80, 80]$ degrees, unless specified otherwise. Both sum-rate C and averaged bit-error-rate (BER) for all users under AWGN and zero-forcing (ZF) equalization are simulated. The sum rate C is calculated from (3) and the sum-rate values of different schemes are normalized to the upper-bound of the fully connected array.
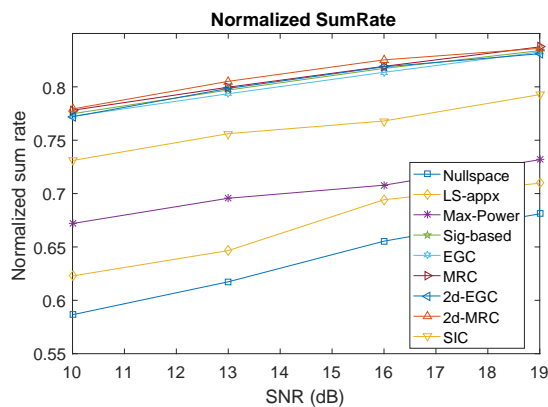
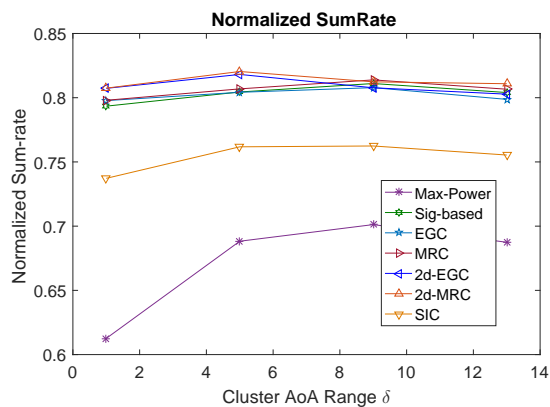Fig. 1. Normalized sum rate C versus SNR for all the presented schemes.
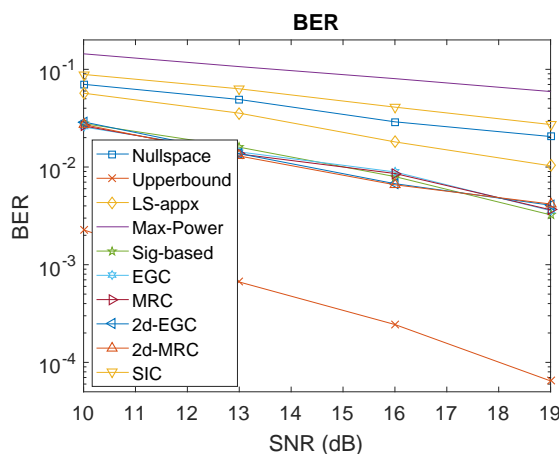

Fig. 2. BER versus SNR for all the presented schemes.

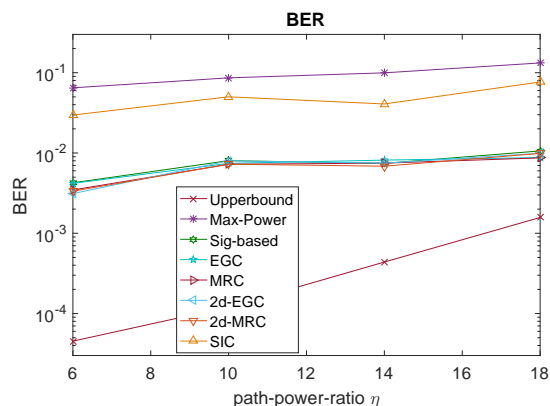Fig. 3. Sum-rate versus cluster AoA range δ.

Fig. 4. BER versus path-power-ratio η.

In Figs. 1 and 2, we show how the normalized sum rate and BER vary with SNR, respectively. The proposed user-directed schemes achieve close performance in terms of both sum-rate and BER. Although performing best in all the presented schemes for the subarray structure, they still have a large performance gap to the upper bound for the fully-connected structure. Looking into the details, the ranking from high to low in sum rate is consistent for different SNRs and is in the order of *2d-MRC, MRC, 2d-EGC, sig-based*, and *EGC* schemes. The *SIC* scheme is inferior to the user-directed schemes in sum-rate and better than the other three ones; however, it becomes the second worst in BER, probably because the ZF equalization is inappropriate for it. The *nullspace, LS-approximation* and *maximizing power* schemes all have bad performance, for the possible reasons discussed in respective sections. Hence to show other results clearly, we will not present the null space and LS approximation results in the subsequent plots.

Fig. 3 demonstrates how sum-rate C is affected by different cluster AoA range δ. The user-directed schemes show good robustness against different δ values, and are notably better than the *SIC* and *maximizing-power* methods.

Fig. 4 demonstrates the impact of different path-power-ratios on BER. It can be seen that BER values generally increase with η increasing for all curves, probably because the correlation between the propagation channels increases in this case.

VI. CONCLUSIONS

We presented several new schemes for designing RF beam-forming for mmWave hybrid array with the subarray structure. The user-directed approach can efficiently reduce channel correlation between users and enable good signal separation in multiuser MIMO systems. Schemes based on this approach generally have low complexity and demonstrate good numerical performance with respect to sum rate and BER. Hence they are promising for practical implementation. We also presented some schemes that would intuitively work but show inferior performance in simulation, including generating orthogonal equivalent digital channels, least square approximation to the optimal solution for the fully-connected structure, and maximizing the total power. These results strike the importance of achieving balanced channel correlation and signal power, and provide insights on BF design for mmWave hybrid array.

REFERENCES

[1] J. Zhang, X. Huang, V. Dyadyuk, and Y. Guo, "Massive hybrid antenna array for millimeter-wave cellular communications," *Wireless Communications, IEEE*, vol. 22, no. 1, pp. 79–87, 2015.

[2] R. Heath, N. Gonzalez-Prelcic, S. Rangan, W. Roh, and A. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE Journal on Selected Topics in Signal Processing*, 2016.

[3] 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 *2015 IEEE International Conference on Communications (ICC)*, June 2015, pp. 1334–1339.

[4] W. Ni, X. Dong, and W. Lu, "Near-optimal hybrid processing for massive MIMO systems via matrix decomposition," *CoRR*, vol. abs/1504.03777, 2015. [Online]. Available: http://arxiv.org/abs/1504.03777

[5] F. Khalid and J. Speidel, "Robust hybrid precoding for multiuser MIMO wireless communication systems," *IEEE Transactions on Wireless Communications*, vol. 13, no. 6, pp. 3353–3363, June 2014.

[6] W. Ni and X. Dong, "Hybrid block diagonalization for massive multiuser MIMO systems," *IEEE Transactions on Communications*, vol. 64, no. 1, pp. 201–211, Jan 2016.

[7] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of mimo channels," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 5, pp. 684–702, June 2003.

[8] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.