Low-complexity Subarray-based RF Precoding for Wideband Multiuser Millimeter Wave Systems

Zhitong Ni, *Student Member, IEEE*, J. Andrew Zhang, *Senior Member, IEEE*,
Kai Yang, *Member, IEEE*, Fei Gao, and Jianping An, *Member, IEEE*

Abstract—This correspondence paper proposes a novel low-complexity radio-frequency (RF) precoding and combining scheme for wideband multiuser millimeter wave hybrid-array systems, targeting at maximizing the system energy efficiency. We first derive a nearly-optimal fully-connected RF precoder, via minimizing the correlation across different users and subcarriers. We then extend the optimization solution to subarray-based architectures by exploiting the unitary matrix feature of subarrays. With the obtained phase values of the precoder, we optimize the power allocation on each subcarrier of the baseband precoder. Simulation results are provided and validate the effectiveness of our proposed hybrid precoding scheme.

Index Terms—Millimeter wave communication, hybrid precoding, wideband, subarrays

I. Introduction

Hybrid precoding, as well as the combining at receivers, has been a research focus for millimeter wave (mmWave) hybrid systems in the last several years [1]–[5]. Hybrid precoding consists of a radio-frequency (RF) precoder with phase shifters in the RF domain and a digital precoder in the baseband domain. Balancing cost and performance, as well as beamforming (BF) gain and multiplexing gain, the hybrid array can be effectively used in both base station (BS) and user equipment (UE).

Hybrid precoding in wideband multiuser systems is one of the main challenges at present when being extended from a narrowband beamforming model [6]–[9]. Typically, a narrowband precoding/beamforming structure is considered for RF precoding, even in wideband systems [10], [11], to make the cost of the system affordable. Hence, a common RF precoder is applied to the whole frequency band of the signal. Optimal RF precoding design in wideband systems generally requires high-complexity processing, such as singular value decomposition (SVD) of the channel [10]. Alternative designs can be based on a pre-designed codebook, which reduces the overhead of channel feedback at the cost of degraded performance [8], [11], [12]. In addition, the authors in [9]

exploited directional precoding structures to reduce the complexity associated with obtaining the hybrid precoder.

Energy efficiency (EE) is another major concern in mmWave hybrid array [13]–[16]. The authors in [14] proposed energy-efficient hybrid precoding, taking into consideration the number of RF chains and phase shifters. In [15], the authors further optimized the number of RF chains, which is related to the number of receivers connected to the transmitter. In [16], the authors optimized EE using the Dinkelbach's algorithm [17], via optimal power allocation for different users. These works are all based on the fully-connected RF precoder, which has a mathematical form of a matrix with the modulus of all entries being equal.

Subarray-based precoding, where only a limited number of elements in the precoding matrix is non-zero, is a promising way to further reduce the power consumption at the cost of lower spectral efficiency, compared to the fully-connected one [18], [19]. There are two common types of subarrays, fixed ones and dynamic ones [20], depending on whether the locations of non-zero elements in the precoding matrix are fixed or dynamically designed. In [18], the authors investigated data rate maximization based on the water-filling method for fixed subarrays. In [19], the authors studied the minimization of the mean squared error of angle-of-arrival (AoA) and angle-of-departure (AoD) estimates for subarray-based RF precoding in wideband systems. In [20], the authors developed dynamic subarrays for single-user wideband systems, where connections are dynamically allowed between RF chains and antennas with a given number of phase shifters. Intuitively, using fixed subarrays can achieve better EE than using fully-connected arrays and dynamic subarrays due to the less number of phase shifters and simpler circuit design. However, there is little work reported for EE optimization for subarrays.

In this correspondence paper, we propose a novel hybrid precoding scheme that maximizes EE in wideband multiuser multi-input multi-output orthogonal-frequency-division-multiplexing (MIMO-OFDM) systems. The main innovation of this paper is a low-complexity RF precoding algorithm. The EE maximization problem is divided into two parts, i.e., the optimization of precoding matrices and the power allocation. The baseband precoder is applied in the frequency domain and is subcarrier-dependent, while the RF precoder is common for all subcarriers. Our main contributions are summarized as follows.

- We propose a low-complexity fully-connected RF precoding algorithm via path-wise implementation. Using an optimal baseband precoder, only the left singular matrix

of the RF precoder influences the system performance.

- We extend the precoding algorithm from fully-connected RF precoder to subarray-based RF precoder, by exploiting the unitary matrix feature of the subarrays.
- We optimize the power allocation on each subcarrier by using Jensen's inequality. We illustrate that the subarrays can achieve higher EE than the fully-connected one, despite slightly lower sum rates.

Notations: $\mathbf{a}$ denotes a vector, $\mathbf{A}$ denotes a matrix, italic English letters like N and lower-case Greek letters α are a scalar, $\angle a$ is the phase angle of complex value a. $|\mathbf{A}|$, $\mathbf{A}^T$, $\mathbf{A}^H$, $\mathbf{A}^\dagger$ represent determinant value, transpose, conjugate transpose, pseudo inverse respectively. We denote Frobenius norm of a matrix as $\|\mathbf{A}\|_F$. $[\mathbf{A}]_{m,n}$ is the (m, n)th entry of a matrix.

II. SYSTEM AND CHANNEL MODELS

We consider a multiuser MIMO-OFDM system model with a BS and U UEs. The BS is equipped with N_T antennas and U RF chains. Each UE is equipped with an antenna array that has N_R phase shifters. The BS communicates with all UEs, with one spatial stream for each UE.

At the BS, a baseband precoder, $\mathbf{F}_{BB}[k] = [\mathbf{f}_{BB1}[k], \cdots, \mathbf{f}_{BBU}[k]]$, of the dimension $U \times U$, is applied on each subcarrier k, and $\mathbf{f}_{BBu}[k]$ is the uth column of $\mathbf{F}_{BB}[k]$, denoting the baseband precoding vector for the uth UE. The baseband precoder processes the signal in the frequency domain before transforming the signal into the time domain by using K-point inverse fast Fourier transform's (IFFT's), where K is the number of subcarriers. The cyclic prefix of length D is added to avoid inter-carrier interference. The time-domain signals are then precoded by an RF precoder of the dimension $N_T \times N_P$, denoted as $\mathbf{F}_{RF}$, such that $[\mathbf{F}_{RF}]_{m,n} = e^{j\psi_{m,n}}$, where $\psi_{m,n}$ is a quantized phase value. We note that all entries of $\mathbf{F}_{RF}$ have equal modulus and quantized phase shifts, and is subcarrier-independent, while $\mathbf{F}_{BB}[k]$ varies with k. Let the original symbol vector on the subcarrier k be $\mathbf{s}[k]$, which is a $U \times 1$ data vector with the uth entry being the data symbol transmitted from BS to the uth UE. The transmitted signal is written as

$$\mathbf{x}[k] = \mathbf{F}_{RF}\mathbf{F}_{BB}[k]\mathbf{s}[k]. \tag{1}$$

Considering the power consumption constraints, we normalize the hybrid precoder, such that $\sum_{k=1}^{K} \|\mathbf{F}_{RF}\mathbf{F}_{BB}[k]\|_F^2 = P \leq P_{ow}$, where P is total transmitted power and P_{ow} is the power constraint. Both fully-connected RF precoder and subarrays are considered in this paper. For subarrays, they are partially implemented by N_T phase shifters and have a form of block diagonal matrix,

$$\mathbf{F}_{RF} = \mathrm{diag}(\mathbf{f}_{RF_1}, \cdots, \mathbf{f}_{RF_{N_P}}), \tag{2}$$

where $\mathbf{f}_{RF_u}$ is a subarray for the uth UE with N_T/N_P antenna elements.

At the uth UE, the received signals from N_R antennas are combined using a combining vector $\mathbf{w}_u$ that satisfies $[\mathbf{w}_u]_i = e^{j\psi_i'}$ with ψ_i' being the quantized phase values. The combined signal is transformed into the frequency/digital domain using K-point FFT's, after removing the cyclic prefix of length D.

The combined baseband signal on the subcarrier k is written as

$$\mathbf{y}_u[k] = \mathbf{w}_u^H \left(\mathbf{H}_u[k]\mathbf{F}_{RF}\mathbf{F}_{BB}[k]\mathbf{s}[k] + \mathbf{n}_u[k] \right), \tag{3}$$

where $\mathbf{H}_u[k]$ represents the wideband mmWave channel matrix on the subcarrier k between the uth UE and BS, and $\mathbf{n}_u[k] \sim \mathcal{CN}(0, \sigma^2\mathbf{I})$ represents a complex Gaussian noise vector.

The channel matrix can be obtained from the FFT of a time-domain channel matrix [21]. The adopted mmWave channel has L paths between the BS and each UE. For notational simplicity, we drop the subscript u and write the delay-d channel matrix as

$$\mathbf{H}'[d] = \rho \sum_{l=1}^{L} p(dT_s - \tau_l)\alpha_l \mathbf{a}_R(\theta_l^r, \varphi_l^r)\mathbf{a}_T^H(\theta_l^t, \varphi_l^t), \tag{4}$$

where $\rho = \sqrt{\frac{N_T N_R}{L}}$, α_l, τ_l, and $p(\tau)$ represent the normalization factor, complex path loss, time delay, and the pulse shaping function, respectively, T_s is the sampling interval, θ^r, φ^r, θ^t, and φ^t represent elevation AoAs, azimuth AoAs, elevation AoDs, and azimuth AoDs, respectively, $\mathbf{a}_R(\cdot)$ and $\mathbf{a}_T(\cdot)$ denote the array response vectors at the UE and the BS, respectively. The frequency selective channel matrix on the subcarrier k is given by

$$\mathbf{H}[k] = \sum_{d=0}^{D-1} \mathbf{H}'[d]e^{-j\frac{2\pi k}{K}d}. \tag{5}$$

III. PROBLEM STATEMENT

The proposed scheme aims to maximize the EE for multiuser systems in three stages. In Stage 1, all UEs send the channel state information (CSI) through an uplink feedback channel. In Stage 2, the BS optimizes both the combiners and the hybrid precoder using the CSI obtained from UEs. In Stage 3, the BS sends the optimized combiners to each UE through a downlink feedback channel. We assume that the CSI is already available, and focus on the optimization in Stage 2 in this paper.

The optimization goal is to maximize the EE as defined by [14], [16], i.e.,

$$\eta = \frac{R}{\xi P + N_P P_{RF} + N_{PS} P_{PS} + P_{UE}}, \tag{6}$$

where R is the sum rate, ξ is the efficiency of the power amplifier, P_{RF} and P_{PS} are the energy consumed by RF chains and phase shifters, respectively, P_{UE} is the energy consumed by UEs, and N_{PS} is the number of phase shifters. The sum rate is

$$R = \sum_{\substack{k=1 \\ u=1}}^{K,U} \log_2 \left(1 + \frac{\left| \mathbf{w}_u^H \mathbf{H}_u[k]\mathbf{F}_{RF}\mathbf{f}_{BB_u}[k] \right|^2}{\sum_{n \neq u} \left| \mathbf{w}_u^H \mathbf{H}_u[k]\mathbf{F}_{RF}\mathbf{f}_{BB_n}[k] \right|^2 + \sigma^2} \right). \tag{7}$$

For a given system with fixed number of RF chains and phase shifters, the optimization problem can be represented as

$$\{\mathbf{F}_{\mathrm{RF}}^{\star}, \mathbf{F}_{\mathrm{BB}}^{\star}[k], \mathbf{w}_u^{\star}, P_k^{\star}\} = \arg\max_{\substack{\mathbf{F}_{\mathrm{RF}}, \mathbf{F}_{\mathrm{BB}}[k] \\ \mathbf{w}_u, P}} \eta,$$

$$\text{s.t.} \quad \sum_{k=1}^{K} P_k \le P_{\mathrm{ow}}, [\mathbf{F}_{\mathrm{RF}}]_{m,n} = e^{j\psi_{m,n}}, [\mathbf{w}_u]_i = e^{j\psi_i'}, \quad (8)$$

where $P_k = \|\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}[k]\|_F^2$. Note that the entries of both the RF precoder and the combiners have equal modulus and quantized phase shifts. The problem above becomes non-convex and difficult to be addressed.

As in [11], [20], we let $\mathbf{F}_{\mathrm{BB}}[k] = \left(\mathbf{F}_{\mathrm{RF}}^H\mathbf{F}_{\mathrm{RF}}\right)^{-\frac{1}{2}}\mathbf{F}[k]$ with $\mathbf{F}[k]$ being a matrix to be determined, and rewrite the sum rate as

$$R = \sum_{\substack{k=1 \\ u=1}}^{K,U} \log_2\left(1 + \frac{|\mathbf{w}_u^H\mathbf{H}_u^{\mathrm{RF}}[k]\mathbf{f}_u[k]|^2}{\sum_{n \ne u}|\mathbf{w}_u^H\mathbf{H}_u^{\mathrm{RF}}[k]\mathbf{f}_n[k]|^2 + \sigma^2}\right), \quad (9)$$

where $\mathbf{f}_u[k]$ is the uth column of $\mathbf{F}[k]$, and $\mathbf{H}_{\mathrm{u}}^{\mathrm{RF}}[k] = \mathbf{H}_u[k]\mathbf{F}_{\mathrm{RF}}\left(\mathbf{F}_{\mathrm{RF}}^H\mathbf{F}_{\mathrm{RF}}\right)^{-\frac{1}{2}} = \mathbf{H}_u[k]\mathbf{U}_{\mathrm{RF}}\mathbf{V}_{\mathrm{RF}}^H$ denotes the channel before baseband precoding, where $\mathbf{U}_{\mathrm{RF}}$ and $\mathbf{V}_{\mathrm{RF}}$ are the left and right singular matrices of $\mathbf{F}_{\mathrm{RF}}$ with the dimension of $N_{\mathrm{T}} \times N_{\mathrm{P}}$ and $N_{\mathrm{P}} \times U$, respectively.

We further define an equivalent channel for all UEs as

$$\overline{\mathbf{H}}[k] \triangleq \left[\mathbf{H}_1^H[k]\mathbf{w}_1, \cdots, \mathbf{H}_U^H[k]\mathbf{w}_U\right]^H, \quad (10)$$

where $\overline{\mathbf{H}}[k]$ is the equivalent channel. Based on $\overline{\mathbf{H}}[k]$, we obtain the optimal $\mathbf{F}_{\mathrm{BB}}[k]$ that minimizes the multiuser interference (MUI) as

$$\mathbf{F}_{\mathrm{BB}}^{\star}[k] = \mu_k(\mathbf{F}_{\mathrm{RF}}^H\mathbf{F}_{\mathrm{RF}})^{-\frac{1}{2}}\left(\overline{\mathbf{H}}[k]\mathbf{F}_{\mathrm{RF}}(\mathbf{F}_{\mathrm{RF}}^H\mathbf{F}_{\mathrm{RF}})^{-\frac{1}{2}}\right)^{\dagger}\mathbf{\Gamma}[k], \quad (11)$$

where μ_k is a coefficient for meeting the power constraint, i.e., $\|\mathbf{F}_{\mathrm{RF}}\mathbf{F}_{\mathrm{BB}}[k]\|_F^2 = P_k$, and $\mathbf{\Gamma}[k]$ is a diagonal matrix that allocates the power to each UE. The actual values of the power allocation matrix $\mathbf{\Gamma}[k]$ does not affect our RF precoding design. It can be optimized using the water-filling method. For simplicity, we consider the unitary constraint as in [11] and assume $\mathbf{\Gamma}[k] = \mathbf{I}$ here.

Substituting $\mathbf{F}_{\mathrm{BB}}^{\star}[k]$ into the power constraint, μ_k^2 is obtained as

$$\mu_k^2 = \frac{P_k}{\left\|\mathbf{F}_{\mathrm{RF}}\left(\mathbf{F}_{\mathrm{RF}}^H\mathbf{F}_{\mathrm{RF}}\right)^{-\frac{1}{2}}\left(\overline{\mathbf{H}}[k]\mathbf{F}_{\mathrm{RF}}(\mathbf{F}_{\mathrm{RF}}^H\mathbf{F}_{\mathrm{RF}})^{-\frac{1}{2}}\right)^{\dagger}\right\|_F^2}$$

$$= \frac{P_k}{\left\|\mathbf{U}_{\mathrm{RF}}\mathbf{V}_{\mathrm{RF}}\left(\overline{\mathbf{H}}[k]\mathbf{U}_{\mathrm{RF}}\mathbf{V}_{\mathrm{RF}}^H\right)^{\dagger}\right\|_F^2} = \frac{P_k}{\left\|\left(\overline{\mathbf{H}}[k]\mathbf{U}_{\mathrm{RF}}\right)^{\dagger}\right\|_F^2}. \quad (12)$$

The third equation of (12) is obtained based on the fact that the unitary matrix, $\mathbf{V}_{\mathrm{RF}}$, has no influence on the Frobenius norm. Substituting (11) and (12) into (6), we can rewrite the EE as

$$\eta = \frac{\sum_{k=1}^{K}\sum_{u=1}^{U}\log_2\left(1 + \frac{1}{\sigma^2}|\mathbf{w}_u^H\mathbf{H}_u^{\mathrm{RF}}[k]\mathbf{f}_u[k]|^2\right)}{\xi P + N_{\mathrm{P}}P_{\mathrm{RF}} + N_{\mathrm{PS}}P_{\mathrm{PS}} + P_{\mathrm{UE}}}$$

$$= \frac{\sum_{k=1}^{K}\sum_{u=1}^{U}\log_2\left(1 + \frac{\mu_k^2}{\sigma^2}\right)}{\xi P + N_{\mathrm{P}}P_{\mathrm{RF}} + N_{\mathrm{PS}}P_{\mathrm{PS}} + P_{\mathrm{UE}}}$$

$$= \frac{\sum_{k=1}^{K}U\log_2\left(1 + \frac{P_k}{\sigma^2\left\|\left(\overline{\mathbf{H}}[k]\mathbf{U}_{\mathrm{RF}}\right)^{\dagger}\right\|_F^2}\right)}{\xi P + N_{\mathrm{P}}P_{\mathrm{RF}} + N_{\mathrm{PS}}P_{\mathrm{PS}} + P_{\mathrm{UE}}}. \quad (13)$$

To this end, we obtain the optimal baseband precoder, with the variables of $\mathbf{U}_{\mathrm{RF}}$, $\mathbf{w}_u$, and $\{P_k\}_{k=1}^{K}$ to be optimized. The original optimization problem in (8) can be recast as

$$\{\mathbf{U}_{\mathrm{RF}}^{\star}, \mathbf{w}_u^{\star}, P_k^{\star}\} = \arg\max_{\mathbf{U}_{\mathrm{RF}}, \mathbf{w}_u, P_k} \eta,$$

$$\text{s.t.} \quad \sum_{k=1}^{K} P_k \le P_{\mathrm{ow}}, [\mathbf{F}_{\mathrm{RF}}]_{m,n} = e^{j\psi_{m,n}}[\mathbf{w}_u]_i = e^{j\psi_i'}. \quad (14)$$

It is noted that the channel is subcarrier-dependent, while the RF precoder and the combiner are subcarrier-independent, which makes the problem still hard to solve.

IV. PROPOSED RF PRECODING SCHEME

In this section, we first derive a nearly-optimal fully-connected RF precoder. Then, we transform the optimization problem into the subarray-based one and optimize the subarrays using its unitary matrix feature. Finally, we allocate the power on each subcarrier using Jensen's inequality.

One main property of $\mathbf{U}_{\mathrm{RF}}$ and $\mathbf{w}_u$ is that they are normalized to unity. Additionally, they have influence on the sum rate only, whereas the power constraints $\{P_k\}_{k=1}^{K}$ have influence on both sum rate and power consumption. Based on these two properties, we can divide the optimization problem of (14) into two sub-problems that are easier to solve. We first determine $\mathbf{U}_{\mathrm{RF}}$ and $\mathbf{w}_u$ by maximizing the sum rate, and then optimize the value of P_k's.

A. Optimizing RF Precoder and Combiner

The channel matrix, $\mathbf{H}_u[k]$, is the key to optimizing both $\mathbf{U}_{\mathrm{RF}}$ and $\mathbf{w}_u$. It is nearly impossible to derive globally optimal $\mathbf{U}_{\mathrm{RF}}$ and $\mathbf{w}_u$ that maximize the sum rate on each subcarrier, since $\mathbf{H}_u[k]$ is subcarrier dependent. We note that AoAs and AoDs of the channel keep stable for all subcarriers, i.e.,

$$\mathbf{H}_u[k] = \mathbf{A}_u\mathbf{\Lambda}_u[k]\mathbf{D}_u^H, \quad (15)$$

where $\mathbf{A}_u = [\mathbf{a}_{R,u}(\theta_1^r, \varphi_1^r), \cdots, \mathbf{a}_{R,u}(\theta_L^r, \varphi_L^r)]$, $\mathbf{D}_u = [\mathbf{a}_{T,u}(\theta_1^t, \varphi_1^t), \cdots, \mathbf{a}_{T,u}(\theta_L^t, \varphi_L^t)]$, and $\mathbf{\Lambda}_u[k]$ is a diagonal matrix that is related to the time delay and the path loss of the channel. Our following proposed solution uses $\mathbf{A}_u$ and $\mathbf{D}_u$ to obtain the RF precoder and the combiners by selecting optimal columns from $\mathbf{A}_u$ and $\mathbf{D}_u$. Noting that (5) and (15) are equivalent, the lth entry of $\mathbf{\Lambda}_u[k]$ is given by

$$\lambda_l[k] = \rho\sum_{d=0}^{D-1} p(dT_s - \tau_l)\alpha_l e^{-j2\pi\frac{kd}{K}}, l \in \{1, \cdots, L\}. \quad (16)$$

Substituting (15) into (13), we define a scalar in the denominator of sum rate as

$$\Delta[k] = \left\| \left(\overline{\mathbf{H}}[k]\mathbf{U}_{\mathrm{RF}} \right)^{\dagger} \right\|_F^2$$

$$= \left\| \left(\begin{bmatrix} \mathbf{w}_1^H \mathbf{A}_1 \mathbf{\Lambda}_1[k] \mathbf{D}_1^H \\ \vdots \\ \mathbf{w}_U^H \mathbf{A}_U \mathbf{\Lambda}_U[k] \mathbf{D}_U^H \end{bmatrix} \mathbf{U}_{\mathrm{RF}} \right)^{\dagger} \right\|_F^2. \quad (17)$$

We note that the sum rate is only influenced by $\Delta[k]$ that needs to be minimized.

Thanks to the sparsity of mmWave channels, the columns of both $\mathbf{A}_u$ and $\mathbf{D}_u$ have low correlation, especially for $\mathbf{D}_u$ of which columns are approximately orthogonal to each other. By exploiting such low correlation, the RF precoder and combiners can be optimized separately, where $\mathbf{w}_u$ only depends on $\mathbf{A}_u$ and $\mathbf{U}_{\mathrm{RF}}$ only depends on $\mathbf{D}_u$.

To magnify the channel gain for each UE and suppress the MUI, we attempt to select one column from each $\mathbf{D}_u$ and each $\mathbf{A}_u$, respectively, and use the selected columns to generate $\mathbf{U}_{\mathrm{RF}}$ and $\mathbf{w}_u$. The selected columns should make the non-diagonal entries of $\overline{\mathbf{H}}[k]\mathbf{U}_{\mathrm{RF}}$ become as small as possible. Firstly, the MUI has a dominant impact on the sum rate, which is also the reason why we design $\mathbf{F}[k]$ to minimize the denominator of R instead of maximizing the nominator of R. Secondly, when $\mathbf{U}_{\mathrm{RF}}$ is made up by columns from each $\mathbf{D}_u$ and $\mathbf{w}_u$ is made up by one column of $\mathbf{A}_u$, the channel gain in the diagonal entries of $\overline{\mathbf{H}}[k]\mathbf{U}_{\mathrm{RF}}$ can be maintained, and hence we can guarantee a satisfactory spectral efficiency.

Since the channel gain varies on different subcarriers, it is impossible to obtain a fixed set of columns from $\mathbf{D}_u$ and $\mathbf{A}_u$, respectively. For obtaining subcarrier-independent RF precoder and combiners, we attempt to make the expectation of MUI be minimized,

$$\{\mathbf{U}_{\mathrm{RF}}^{\star}, \mathbf{w}_u^{\star}\} = \arg \min_j \mathbb{E}(\|\overline{\mathbf{H}}[k]\mathbf{U}_{\mathrm{RF}}\|_F^2 - \|\mathrm{diag}(\overline{\mathbf{H}}[k]\mathbf{U}_{\mathrm{RF}})\|_F^2)$$

$$\text{s.t.} \quad \mathbf{U}_{\mathrm{RF}} = [\mathbf{d}_{1_j}, \cdots, \mathbf{d}_{U_j}], \mathbf{w}_u = \mathbf{a}_{u_j}, \quad (18)$$

where $\mathbf{d}_{u_j}$ is a column of $\mathbf{D}_u$, $\mathbf{a}_{u_j}$ is a column of $\mathbf{A}_u$ that has the same selected index with $\mathbf{d}_{u_j}$, and $\mathbb{E}(\cdot)$ takes the expectation value from $k = 1$ to $k = K$. The objective function of (18) is minimized by selecting the optimal combination of the array response vectors.

The optimized fully-connected RF precoder is given by

$$\mathbf{F}_{\mathrm{RFc}} = \exp\left(j\angle \mathbf{U}_{\mathrm{RF}}^{\star} \mathbf{E}_{\mathrm{RF}} \mathbf{V}_{\mathrm{RF}}^H\right) \triangleq \mathbf{D} \quad (19)$$

where $\mathbf{D} = [\mathbf{d}_{1_j}, \cdots, \mathbf{d}_{U_j}]$ with each column being an array response vector. Since the phase shifters have limited number of shifts, the optimal RF precoder and combiners can be obtained as an element-wise solution of $\mathbf{D}$ and $\mathbf{a}_{u_j}$, respectively.

By substituting $\mathbf{U}_{\mathrm{RF}}$ into $\Delta[k]$, we obtain

$$\Delta^{\star}[k] \approx \left\| \left(\left[w_1 \lambda_1[k]\mathbf{d}_{1_j}, \cdots, w_U \lambda_U[k]\mathbf{d}_{U_j} \right]^H \mathbf{D} \right)^{\dagger} \right\|_F^2$$

$$= \left\| \left(\overline{\mathbf{\Lambda}}[k]\mathbf{D}^H\mathbf{D} \right)^{\dagger} \right\|_F^2, \quad (20)$$

where $\lambda_u[k]$ is the u_jth entry of $\mathbf{\Lambda}_u[k]$, $w_u = \mathbf{a}_{u_j}^H \mathbf{w}_u$ is the combined gain at the UE, and $\overline{\mathbf{\Lambda}}[k] = \mathrm{diag}(w_1 \lambda_1, \cdots, w_U \lambda_U)$. The result of (20) approaches to its minimal value because the vectors in $\mathbf{D}$ are approximately orthogonal, i.e., $\mathbf{D}^H \mathbf{D} \approx \mathbf{I}$.

B. Extension to Subarrays

As for subarrays that have a limited number of non-zero elements, we need to redesign $\mathbf{F}_{\mathrm{RF}}$ accordingly. Note that the left singular matrix of subarrays is equivalent to $\mathbf{F}_{\mathrm{RF}}$ itself, due to the unitary property from (2). The unitary property makes it easy to determine the subarrays.

Similar to (18), we aim to minimize the gain of the non-diagonal entries in $\overline{\mathbf{H}}_{\mathrm{RF}}\mathbf{F}_{\mathrm{RF}}$. Different from the fully-connected RF precoder, the dimension of selected columns from each $\mathbf{D}_u$ needs to be reduced due to the reduced number of phase shifters in subarrays. Hence, the optimization problem for subarrays is given by

$$\{\mathbf{f}_{\mathrm{RF}_u}, \mathbf{w}_u\} = \arg \min_j \mathbb{E}(\|\overline{\mathbf{H}}[k]\mathbf{F}_{\mathrm{RF}}\|_F^2 - \|\mathrm{diag}(\overline{\mathbf{H}}[k]\mathbf{F}_{\mathrm{RF}})\|_F^2)$$

$$\text{s.t.} \quad \mathbf{f}_{\mathrm{RF}_u} = \tilde{\mathbf{d}}_{u_j}, \mathbf{w}_u = \mathbf{a}_{u_j}, \quad (21)$$

where $\tilde{\mathbf{d}}_{u_j}$ is the array response vector used for designing $\mathbf{f}_{\mathrm{RF}_u}$, i.e., $\tilde{\mathbf{d}}_{u_j} = [\mathbf{D}_u]_{1+(u-1)*N_{\mathrm{T}}/U:N_{\mathrm{T}}/U+(u-1)*N_{\mathrm{T}}/U, j}$. Note that there are UL candidates to be tested for determining $\mathbf{F}_{\mathrm{RF}}$, with each candidate being a combination of $\tilde{\mathbf{d}}_{u_j}$. We test all candidates exhaustively and select the one that minimizes the objective function of (21).

C. Power Allocation

After determining the entries of RF precoder and combiners, the EE is a function of P_k only, i.e.,

$$\eta = \frac{\sum\limits_{k=1}^{K} U\log_2\left(1 + \frac{P_k}{\sigma^2 \Delta^{\star}[k]}\right)}{\xi P + N_{\mathrm{P}}P_{\mathrm{RF}} + N_{\mathrm{PS}}P_{\mathrm{PS}} + P_{\mathrm{UE}}}. \quad (22)$$

It is noted that η is a non-convex function due to the fractional form. Additionally, there are K individual variables of P_k that need to be determined, which could increase the computational complexity greatly. We use Jensen's inequality to further reduce the complexity, i.e.,

$$\eta = \frac{U \sum\limits_{k=1}^{K} \log_2\left(1 + \frac{P_k}{\sigma^2 \Delta^{\star}[k]}\right)}{\xi P + N_{\mathrm{P}}P_{\mathrm{RF}} + N_{\mathrm{PS}}P_{\mathrm{PS}} + P_{\mathrm{UE}}}$$

$$\leq \frac{KU\log_2\left(1 + \frac{1}{K}\sum\limits_{k=1}^{K} \frac{P_k}{\sigma^2 \Delta^{\star}[k]}\right)}{\xi P + N_{\mathrm{P}}P_{\mathrm{RF}} + N_{\mathrm{PS}}P_{\mathrm{PS}} + P_{\mathrm{UE}}}. \quad (23)$$

It is clear that η is maximized when it equals its upper bound, i.e., $\frac{P_k}{\Delta^{\star}[k]} = \sum\limits_{k=1}^{K} \frac{P_k}{K\Delta[k]}$. By noting that $\sum\limits_{k=1}^{K} P_k = P$, the optimal P_k is obtained as

$$P_k^{\star} = \frac{P\Delta^{\star}[k]}{\sum\limits_{k=1}^{K} \Delta^{\star}[k]}. \quad (24)$$

TABLE I
COMPUTATIONAL COMPLEXITY

The proposed RF precoding algorithm	
Inputs: $\mathbf{D}_u$ and $\mathbf{A}_u$, $\forall u \le U$.	
Operation	Complexity
Select one common column from $\mathbf{D}_u$ and $\mathbf{A}_u$	$\mathcal{O}(L)$
Generate all candidates for $\mathbf{U}_{\mathrm{RF}}$ and $\{\mathbf{w}_u\}_{u=1}^U$	$\mathcal{O}(UL)$
For each candidate: obtain $\overline{\mathbf{H}}[k]$ in (10)	$\mathcal{O}(UN_{\mathrm{R}}N_{\mathrm{T}})$
For each candidate: calculate $\overline{\mathbf{H}}[k]\mathbf{U}_{\mathrm{RF}}$	$\mathcal{O}(U^2N_{\mathrm{T}})$
For each candidate: calculate $\|\overline{\mathbf{H}}[k]\mathbf{U}_{\mathrm{RF}}\|_F^2$	$\mathcal{O}(U^2)$
Take expectation from $k=1$ to $k=K$	
Overall (fully-connected)	$\mathcal{O}(U^2LN_{\mathrm{R}}N_{\mathrm{T}})$
Overall (subarray-based)	$\mathcal{O}(ULN_{\mathrm{R}}N_{\mathrm{T}})$

With P_k optimized, P is the final parameter to be determined at BS. Noticing that the power constraint of $P \le P_{\mathrm{ow}}$, we can obtain its optimal value via letting $\frac{\partial \eta}{\partial P} = 0$. If the optimal P is larger than P_{ow}, the optimal value of P should be P_{ow}.

D. Complexity Analysis

As summarized in Table I, the overall computational complexity for obtaining the fully-connected RF precoder is $\mathcal{O}(U^2LN_{\mathrm{R}}N_{\mathrm{T}})$. Note that the selected column for each combiner should have the same index with that of each column for $\mathbf{U}_{\mathrm{RF}}$. Hence, there are UL candidates of $\mathbf{U}_{\mathrm{RF}}$ to be tested in (18). For each candidate, the complexity for solving the objective function in (18) is $\mathcal{O}(UN_{\mathrm{R}}N_{\mathrm{T}})$, because the numbers of multiplication for obtaining $\overline{\mathbf{H}}[k]$ and for calculating $\overline{\mathbf{H}}[k]\mathbf{U}_{\mathrm{RF}}$ are $\mathcal{O}(UN_{\mathrm{R}}N_{\mathrm{T}})$ and $\mathcal{O}(U^2N_{\mathrm{T}})$, respectively, and $\mathcal{O}(U^2N_{\mathrm{T}})$ is smaller than $\mathcal{O}(UN_{\mathrm{R}}N_{\mathrm{T}})$. For subarrays, the overall complexity is reduced to $O(ULN_{\mathrm{T}}N_{\mathrm{R}})$ due to the reduced number of phase shifters.

V. SIMULATION RESULTS

In this section, we present simulation results to evaluate the performance of our proposed subarray-based RF precoder, compared with the proposed fully-connected one and other typical schemes in [9], [10]. We simulate a wideband channel with $L = 3$ paths for each UE with the number of UEs being $U = 4$. The delay of each path is uniformly distributed from 0 to $DT_s = 1.28$ ms, where T_s is 0.01 ms. The path gain, α_l, has zero mean and variance of 1. The AoAs and the AoDs are uniformly distributed in $[0, 2\pi] \times [0, \pi]$. The BS has a 16×16 hybrid uniform planar array (UPA) with the inter-element spacing being half a wavelength. The system has $K = 512$ subcarriers and the number of RF chains at BS is $N_{\mathrm{P}} = 4$. We adopt a raised-cosine filter with the roll-off factor being 1 for the pulse-shaping filter as in [11]. For subarrays, there are four subarrays and each one has a 8×8 UPA. Each UE has a 2×2 UPA.

Fig. 1 illustrates how the sum rate increases with SNR, i.e., P/σ^2. To make a fair comparison, we adopt the fully-connected RF precoder for all schemes except our proposed subarrays. Assuming that the perfect CSI is available, we quantize the phase values of RF precoders and combiners with 4 bits. For the scheme in [9], the thresholds of the gaps of AoDs and magnitudes are set as $\pi/12$ and 1, respectively. The two-step SVD in [10] adopts the SVD of the channel to design
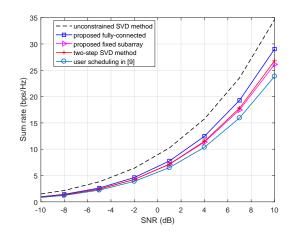

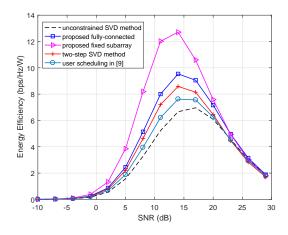Fig. 1. System sum rates versus SNR with using different schemes.

Fig. 2. System EE versus SNR with using different schemes.

the RF precoder. The figure shows that our proposed fully-connected RF precoding algorithm outperforms the schemes in [9] and [10], and approaches to the unconstrained SVD method. The proposed subarray-based RF precoding achieves a slightly reduced sum rate compared with our proposed fully-connected one due to the reduced number of phase shifters. The two-step SVD method has a high complexity due to the SVD operation. The user scheduling method in [9] needs to quantize the modulus of combiners, which results in an overall minor performance gain.

Fig. 2 illustrates the EE achieved by different schemes versus SNR. We use the practical values for $\xi = 0.8$, $P_{\mathrm{RF}} = 250$ mW, $P_{\mathrm{PS}} = 2$ mW, and $P_{\mathrm{UE}} = 50$ mW. The power constraint, P_{ow}, is sufficiently large. The system setup is the same as that in Fig. 1. We see that the EE of unconstrained SVD method remains the lowest since it employs N_{T} RF chains, which result in huge power consumption. Both of our proposed fully-connected RF precoding and subarray-based RF precoding achieve better EE than the states of the art [9], [10]. We also see that the EE shows an upward trend with SNR increasing from -10 dB to nearly 13 dB, followed by a significant decrease with SNR increasing from 15 to 30 dB. This can be explained by the fact that, with the SNR
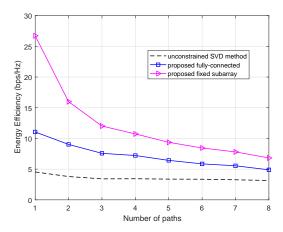
Fig. 3. System EE versus the number of paths.

increasing, the nominator of η is a logarithm form of P and the denominator of η is a linear function of P. Hence, η approaches to zero when P approaches to both zero and infinity. It is worth pointing out that if the SNR constraint, P_{ow}/σ^2, is smaller than 13 dB, the optimal value of P should be equal to P_{ow}.

Fig. 3 illustrates the EE achieved by different hybrid precoders versus the number of paths. The SNR is fixed at 10 dB. The system setup is the same as that in Fig. 1. We see that the EE of unconstrained SVD method remains the lowest and keeps stable with the number of paths increasing since the unconstrained SVD method does not exploit the sparsity of mmWave channels. Both of our proposed fully-connected RF precoding and subarray-based RF precoding achieve better EE than the unconstrained SVD method. We also see that the EE drops with L increasing. For fully-connected RF precoder, the EE drops linearly from $L = 1$ to $L = 8$. As for subarrays, there is a sharp decrease of EE when L increasing from 1 to 3. The possible reason for the sharp decrease is that, when L increases, the achievable sum rate of fixed subarrays tends to drop rapidly due to the less sparse channel matrix.

VI. Conclusion

In this paper, we proposed a nearly-optimal low-complexity hybrid precoding scheme, in particular, RF precoding and combining algorithms for wideband multiuser mmWave systems. The scheme is designed to maximize EE. We first obtained the RF precoder for fully-connected arrays and then extended it to subarrays by utilizing the unitary property of the subarrays. We optimized the power allocation on different subcarriers with the application of Jensen's inequality. The simulation results show that our proposed subarray-based RF precoding achieves higher EE than the proposed fully-connected RF precoding, the unconstrained SVD, and other typical hybrid precoding schemes.

References

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

[2] J. Wang, Z. Lan, C. Pyo, T. Baykas, Chin-sean Sum, M. A. Rahman, J. Gao, R. Funada, F. Kojima, H. Harada, and S. Kato, "Beam codebook based beamforming protocol for multi-Gbps millimeter-wave WPAN systems," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 8, pp. 1390–1399, Oct. 2009.

[3] T. Nitsche, C. Cordeiro, A. B. Flores, E. W. Knightly, E. Perahia, and J. C. Widmer, "IEEE 802.11ad: directional 60 GHz communication for multi-Gigabit-per-second Wi-Fi [invited paper]," *IEEE Commun. Mag.*, vol. 52, no. 12, pp. 132–141, Dec. 2014.

[4] J. An, K. Yang, J. Wu, N. Ye, S. Guo, and Z. Liao, "Achieving sustainable ultra-dense heterogeneous networks for 5G," *IEEE Commun. Mag.*, vol. 55, no. 12, pp. 84–90, Dec. 2017.

[5] W. Roh, J. Seol, J. Park, B. Lee, J. Lee, Y. Kim, J. Cho, K. Cheun, and F. Aryanfar, "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] C. Huang, L. Liu, C. Yuen, and S. Sun, "Iterative channel estimation using LSE and sparse message passing for mmwave MIMO systems," *IEEE Trans. Signal Process.*, vol. 67, no. 1, pp. 245–259, Jan. 2019.

[7] X. Lin, S. Wu, C. Jiang, L. Kuang, J. Yan, and L. Hanzo, "Estimation of broadband multiuser millimeter wave massive MIMO-OFDM channels by exploiting their sparse structure," *IEEE Trans. Wireless Commun.*, vol. 17, no. 6, pp. 3959–3973, Jun. 2018.

[8] M. R. Castellanos, V. Raghavan, J. H. Ryu, O. H. Koymen, J. Li, D. J. Love, and B. Peleato, "Channel-reconstruction-based hybrid precoding for millimeter-wave multi-user MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 2, pp. 383–398, May 2018.

[9] V. Raghavan, S. Subramanian, J. Cezanne, A. Sampath, O. H. Koymen, and J. Li, "Single-user versus multi-user precoding for millimeter wave MIMO systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 6, pp. 1387–1401, Jun. 2017.

[10] H. Yuan, J. An, N. Yang, K. Yang, and T. Q. Duong, "Low complexity hybrid precoding for multiuser millimeter wave systems over frequency selective channels," *IEEE Trans. Veh. Technol.*, vol. 68, no. 1, pp. 983–987, Jan. 2019.

[11] A. Alkhateeb and R. W. Heath, "Frequency selective hybrid precoding for limited feedback millimeter wave systems," *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 1801–1818, May 2016.

[12] V. Raghavan, J. Cezanne, S. Subramanian, A. Sampath, and O. Koymen, "Beamforming tradeoffs for initial UE discovery in millimeter-wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 543–559, Apr. 2016.

[13] Q. Wu, G. Y. Li, W. Chen, D. W. K. Ng, and R. Schober, "An overview of sustainable green 5G networks," *IEEE Wireless Commun.*, vol. 24, no. 4, pp. 72–80, Aug. 2017.

[14] X. Gao, L. Dai, S. Han, C. I, and R. W. Heath, "Energy-efficient hybrid analog and digital precoding for mmwave MIMO systems with large antenna arrays," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 998–1009, Apr. 2016.

[15] C. Lin and G. Y. Li, "Energy-efficient design of indoor mmwave and sub-thz systems with antenna arrays," *IEEE Trans. Wireless Commun.*, vol. 15, no. 7, pp. 4660–4672, Jul. 2016.

[16] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, "Reconfigurable intelligent surfaces for energy efficiency in wireless communication," *IEEE Trans. Wireless Commun.*, vol. 18, no. 8, pp. 4157–4170, Aug. 2019.

[17] S. Schaible, "Fractional programming. II, on Dinkelbach' s algorithm," *Manag. Sci.*, vol. 22, no. 8, pp. 868–873, 1976, [Online]. Available: https://doi.org/10.1287/mnsc.22.8.868.

[18] O. El Ayach, R. W. Heath, S. Rajagopal, and Z. Pi, "Multimode precoding in millimeter wave MIMO transmitters with multiple antenna sub-arrays," in *2013 IEEE Global Communications Conference (GLOBECOM)*, Dec. 2013, pp. 3476–3480.

[19] X. Huang and Y. J. Guo, "Frequency-domain AoA estimation and beamforming with wideband hybrid arrays," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2543–2553, Aug. 2011.

[20] S. Park, A. Alkhateeb, and R. W. Heath, "Dynamic subarrays for hybrid precoding in wideband mmwave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 2907–2920, May 2017.

[21] P. Schniter and A. Sayeed, "Channel estimation and precoder design for millimeter-wave communications: The sparse way," in *2014 48th Asilomar Conference on Signals, Systems and Computers*, Nov. 2014, pp. 273–277.