Joint Transceiver Design With Antenna Selection for Large-Scale MU-MIMO mmWave Systems

Xiongfei Zhai, *Student Member, IEEE*, Yunlong Cai, *Senior Member, IEEE*, Qingjiang Shi, *Member, IEEE*, Minjian Zhao, *Member, IEEE*, Geoffrey Ye Li, *Fellow, IEEE*, and Benoit Champagne, *Senior Member, IEEE*

***Abstract*— This paper considers the uplink of large-scale multiple-user multiple-input multiple-output millimeter wave systems, where several mobile stations (MSs) communicate with a single base station (BS) equipped with a large-scale antenna array, for application to fifth generation wireless networks. Within this context, the use of hybrid transceivers along with antenna selection can significantly reduce the implementation cost and energy consumption of analog phase shifters and low-noise amplifiers. We aim to jointly design the MS beamforming vectors, the hybrid receiving matrices (baseband and analog), and the antenna selection matrix at the BS in order to maximize the achievable system sum-rate under a set of constraints. The corresponding optimization problem is nonconvex and difficult to solve, mainly due to the receive antenna selection and constant modulus constraints on the analog receiving matrix. By exploiting the special structure of the problem and linear relaxation, we first convert this problem into three subproblems, which are solved via an alternating optimization method. The latter iteratively updates the antenna selection matrix, the transmit beamforming vectors, and the hybrid receiving matrices by sequentially addressing each subproblem while keeping the other variables fixed. Specifically, the antenna selection matrix is optimized via the concave–convex procedure; the weighted mean-square error minimization approach is used to find the solution for the transmit beamformer; and the hybrid receiver is obtained via manifold optimization. The convergence of the proposed algorithm is analyzed and its effectiveness is verified by simulation.**

***Index Terms*— Hybrid transceiver, millimeter-wave, antenna selection, CCCP, manifold optimization.**

I. Introduction

IN ORDER to mitigate spectrum shortage and increase transmission rates, millimeter wave (mmWave) communications are now considered as a key enabling technology for the fifth generation (5G) wireless systems and beyond [1]–[4]. Recently, mmWave systems (with operating frequency in the range 30-300GHz) combined with large-scale multiple-input multiple-output (MIMO) techniques have drawn significant attention within the research community [5]–[8] as a means to tremendously improve system capacity.

Nevertheless, mmWave-based large-scale MIMO will increase fabrication cost and energy consumption of the radio frequency (RF) chains as well as of the analog-to-digital (A/D) converters. To address these issues, the use of a hybrid transceiver structure consisting of a baseband digital and RF analog units has been considered as an attractive solution, since it allows to employ a number of RF chains that is much smaller than the number of antenna elements. Two basic kinds of RF electronic modules have been considered for hybrid transceiver implementation, namely: analog phase shifters and analog switches. If analog phase shifters are used [9]–[12], a constant modulus constraint on the elements of the analog transceiver matrices will be imposed. If the simpler and less expensive analog switches are used in the hybrid transceiver [13]–[15], the latter cannot achieve full diversity gain in correlated channels. A typical algorithm to design hybrid transceivers is based on orthogonal matching pursuit (OMP) [10], [12], [15], where the columns of the analog transceiver matrix can be selected from certain candidate vectors, such as e.g., the array response vectors of the channel matrices. To solve for the analog transceivers under the constant modulus constraint, an alternating algorithm based on manifold optimization (MO) has been proposed in [16].

In large-scale MIMO mmWave systems with hybrid transceivers, the numbers of analog phase shifters and low-noise amplifiers (LNA) can be quite large and their energy consumption is considerable. By employing antenna selection, we can turn off some antennas, phase shifters and LNAs, making the system more energy efficient. Therefore, it is of interest to consider transceiver design with antenna selection. Optimal antenna selection requires an exhaustive search whose complexity grows exponentially with the number of antennas, which is computationally prohibitive for massive MIMO systems. Reducing the computational complexity of antenna selection in these systems is therefore of great practical and theoretical interest, as further discussed in [17] and [18]. For the traditional MIMO spatial multiplexing systems, Gorokhov *et al* [19] have provided a comprehensive study of antenna selection algorithms and developed an efficient decremental algorithm to this end. Specifically, it has been proved in [19] that the sum-rate of the system with receive antenna selection is statistically lower bounded by that of a set of parallel

Manuscript received November 4, 2016; revised March 15, 2017 and April 16, 2017; accepted April 25, 2017. Date of publication June 27, 2017; date of current version August 18, 2017. The work of Y. Cai was supported in part by the National Natural Science Foundation of China under Grant 61471319 and in part by the Fundamental Research Funds for the Central Universities. The work of Q. Shi was supported by the National Nature Science Foundation of China under Grant 61671411 and Grant 61631020. *(Corresponding authors: Qingjiang Shi; Yunlong Cai.)*

X. Zhai, Y. Cai, and M. Zhao are with the College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China (e-mail: feifei4006@zju.edu.cn; ylcai@zju.edu.cn; mjzhao@zju.edu.cn).

Q. Shi is with the College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China (e-mail: qing.j.shi@gmail.com).

G. Y. Li is with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA (e-mail: liye@ece.gatech.edu).

B. Champagne is with the Department of Electrical and Computer Engineering, McGill University, Montreal, QC H3A 0E9, Canada (e-mail: benoit.champagne@mcgill.ca).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSAC.2017.2720197

independent single-input multiple-output (SIMO) channels. Receive antenna selection based on a maximum-norm criterion for independent and identically distributed (i.i.d) Rayleigh fading channels has been proposed in [20]. From the optimization perspective, linear relaxation has been employed in [21]–[23] to simplify the convex optimization problem into a form that is solvable in polynomial time. For mmWave systems, a new solution to minimize the side lobe levels for antenna selection is developed in [24] based on compressed sensing techniques and convex optimization relaxation. Joint design of transceiver and antenna selection can significantly enhance the utility of communication systems, which has been studied in [25]–[27]. However, this type of optimization problems are generally nonconvex and NP-hard, even in the single-antenna case [28], while in the case of large-scale MU-MIMO, the need for hybrid transceiver design renders the corresponding problem even more challenging.

In this paper, the transceiver design for the uplink of large-scale multiple-user MIMO (MU-MIMO) mmWave systems is investigated. In particular, to exploit the spatial diversity and reduce the energy consumption at the BS, we study the joint design of the hybrid transceiver matrices and receive antenna selection scheme, where the goal is to maximize the system sum-rate under constraints on the transmit power, analog phase-shifter modulus, and receive antenna selection matrix. We develop an efficient iterative algorithm based on alternating optimization (AO) to address the resulting nonconvex problem. By fixing the transceiver design, we optimize the receive antenna selection matrix based on the concave-convex procedure (CCCP) [29]. Then by fixing the hybrid receiver and the receive antenna selection matrix, we use the weighted minimum mean square error (WMMSE) approach [30] to optimize the transmit beamforming matrices. Finally, MO [16] is adopted to update the hybrid receiver while fixing the remaining variables. The convergence and sum-rate performance of the proposed algorithm is also investigated.

The contributions of this paper are summarized as follows:

1) We investigate hybrid transceiver design with antenna selection for the uplink of large-scale MU-MIMO mmWave systems, where the underlying objective is to maximize the total sum-rate. To the best of our knowledge, there is currently no work that considers the sum-rate maximization problem for this system model.

2) We propose a CCCP-based method to optimize the antenna selection matrix, which admits a binary solution instead of a fractional solution following the linear relaxation.

3) We employ MO instead of the conventional relaxation based on Frobenius norm [10], [16] to solve the original sum-rate maximization problem. By using MO, the space of feasible solutions for analog receiving matrix is not restricted and the solution is guaranteed to converge to a stationary point.

4) We develop a low-complexity iterative algorithm that approaches the performance of an exhaustive antenna selection scheme, and the convergence of the proposed algorithm is analysed.

This paper is organized as follows. In Section II, the system model is introduced and the related optimization problem is formulated. Then we describe the CCCP-based method for antenna selection in Section III while the WMMSE and the MO approaches for hybrid transceiver design are presented in Section IV and Section V, respectively. The convergence of the proposed algorithm is analyzed in Section VI. Simulation results are presented in Section VII to demonstrate the performance of the proposed algorithm. Finally, Section VIII draws the conclusion.

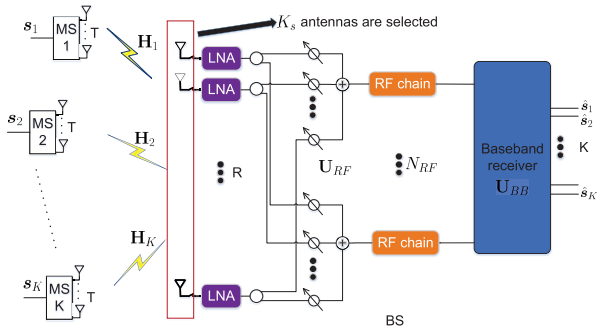

Fig. 1. A mmWave uplink system.

Throughout this paper, we use bold upper-case letters for matrices while keeping the bold lower-case for vectors and small normal face for scalars. For a matrix $\mathbf{A}$, $[\mathbf{A}]_{ij}$ is the entry on the i^{th} row and j^{th} column. The superscript $\mathbf{A}^{\dagger}$, $\mathbf{A}^{*}$, $\mathbf{A}^{T}$ and $\mathbf{A}^{H}$ denote the Moore-Penrose pseudo inverse, conjugate, transpose, and Hermitian transpose of $\mathbf{A}$, respectively. Furthermore, $\mathbf{I}$ is the identity matrix whose dimension will be clear from the context, and $\mathbb{C}^{m\times n}$ denotes the m by n dimensional complex space. The notations $E(\cdot)$, $\mathrm{Tr}(\cdot)$, $\det(\cdot)$, $vec(\cdot)$ and $\mathbb{R}(\cdot)$ represent the expectation, trace, determinant, vectorization and real part, respectively. $\circ$ is the Hadamard product between two matrices. The complex circular-normal distribution is denoted by $\mathcal{CN}(\cdot,\cdot)$.

II. System Model and Problem Formulation

In this section, we first introduce the mmWave uplink system model and then formulate the problem under study.

A. System Model

We consider the uplink of a large-scale MU-MIMO mmWave system as shown in Fig. 1, where K mobile stations (MSs) indexed with $k \in \mathcal{K} \triangleq \{1, 2, \ldots, K\}$ and each equipped with T antennas, simultaneously communicate with a common base station (BS) through a mmWave channel. The BS is equipped with R receive antenna elements that are individually fed to LNAs. In the envisaged 5G application, the MS is equipped with a small number of antennas while this number for the BS is much larger, that is, we assume $R \gg T$ and $R \geq TK$. Hence, a fully digital transmit beamformer is assumed at the MS, while a hybrid structure with $N_{RF} \ll R$ chains along with antenna selection is used at the BS.

Referring to Fig. 1, the transmitted signal by the k^{th} MS is given by

$$\boldsymbol{x}_k = \boldsymbol{v}_k s_k, \tag{1}$$

where $s_k \in \mathbb{C}$ is the transmitted symbol and $\boldsymbol{v}_k \in \mathbb{C}^{T\times 1}$ denotes the corresponding transmit beamforming vector.

For convenience, we assume that $E(|s_k|^2) = \frac{1}{K}$, i.e., the original symbol power in the system is constant, although this condition is not essential and can be relaxed.[1] A constraint on the total transmit power for each MS is enforced by setting $\mathrm{Tr}(\boldsymbol{v}_k \boldsymbol{v}_k^H) \leq 1$, for $k \in \mathcal{K}$.

In this work, as further discussed below, we consider a narrowband frequency-flat channel model. Accordingly, the received signal vector at the BS is given by

$$\boldsymbol{y} = \sqrt{\rho} \sum_{k=1}^{K} \mathbf{H}_k \boldsymbol{v}_k s_k + \boldsymbol{n}, \tag{2}$$

where $\mathbf{H}_k \in \mathbb{C}^{R \times T}$ is a normalized random channel matrix with $E[\mathrm{Tr}(\mathbf{H}_k \mathbf{H}_k^H)] = TR$, ρ represents the maximum average received power and $\boldsymbol{n}$ denotes the additive white Gaussian noise (AWGN) vector at the BS with distribution $\mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I})$. If we denote

$$\begin{aligned} \boldsymbol{s} &\triangleq [s_1, s_2, \ldots, s_K]^T, \\ \mathbf{V} &\triangleq [\mathbf{H}_1 \boldsymbol{v}_1, \mathbf{H}_2 \boldsymbol{v}_2, \ldots, \mathbf{H}_K \boldsymbol{v}_K], \end{aligned}$$

then (2) can be compactly written as

$$\boldsymbol{y} = \sqrt{\rho} \mathbf{V} \boldsymbol{s} + \boldsymbol{n}. \tag{3}$$

In order to reduce the energy consumption and the computational complexity in optimizing the beamformers while still exploiting spatial diversity for the BS with large-scale receive antenna arrays, antenna selection is invoked. Specifically, we select K_s out of R antennas at the receiver side, where we assume $N_{RF} \leq K_s \ll R$. This is conveniently expressed by means of a diagonal antenna selection matrix $\mathbf{\Delta} \in \mathbb{C}^{R \times R}$, with entries $[\mathbf{\Delta}]_{ii} = 1$ if the i^{th} antenna is selected and $[\mathbf{\Delta}]_{ii} = 0$ otherwise, where $\sum_{i=1}^{R} [\mathbf{\Delta}]_{ii} = K_s$. Then the corresponding received signal vector after antenna selection will be

$$\boldsymbol{y}_{\mathbf{\Delta}} \triangleq \mathbf{\Delta} \boldsymbol{y} = \sqrt{\rho} \mathbf{\Delta} \mathbf{V} \boldsymbol{s} + \boldsymbol{n}_{\mathbf{\Delta}}, \tag{4}$$

where $\boldsymbol{n}_{\mathbf{\Delta}} = \mathbf{\Delta} \boldsymbol{n}$.

To reduce the hardware complexity at the BS, only N_{RF} RF chains are used with analog phase shifters, where we assume that $K \leq N_{RF} \ll R$. We denote by $\mathbf{U}_{RF} \in \mathbb{C}^{R \times N_{RF}}$ the corresponding analog receiving matrix, whose entries have constant modulus, i.e. $|[\mathbf{U}_{RF}]_{ij}| = 1$, for $1 \leq i \leq R$, $1 \leq j \leq N_{RF}$. The hybrid beamformer output signal after analog phase shifting and baseband processing can be expressed as

$$\hat{\mathbf{s}} = \mathbf{U}_{BB}^H \mathbf{U}_{RF}^H \boldsymbol{y}_{\mathbf{\Delta}}, \tag{5}$$

where $\mathbf{U}_{BB} \in \mathbb{C}^{N_{RF} \times K}$ denotes the baseband processing matrix.

Under the above processing assumptions, the sum-rate of the hybrid beamforming system in Fig. 1 will be given by Eq. (6), as shown at the bottom of this page, [10].

B. Channel Model

Within the context of narrow-band bandpass transmission, there is only limited spatial selectivity [10] due to the high path loss at mmWave frequencies and the high antenna correlation of tightly packed arrays. Based on the Saleh Valenzuela model [31]–[33], the channel matrix can be expressed as[2]

$$\mathbf{H} = \sqrt{\frac{TR}{L}} \sum_{l=1}^{L} \alpha_l \Lambda^r(\phi_l^r) \Lambda^t(\phi_l^t) \boldsymbol{a}_r(\phi_l^r) \boldsymbol{a}_t(\phi_l^t)^H, \tag{7}$$

where L is the number of rays, $\alpha_l \sim \mathcal{CN}(0, 1)$ is the complex gain of the l^{th} ray, and $\phi_l^r(\phi_l^t)$ represents its azimuth angles of arrival (departure) (AoA and AoD). The functions $\Lambda^r(\phi_l^r)$ and $\Lambda^t(\phi_l^t)$ represent the directivity gains for the receive and transmit antenna elements at the corresponding AOA and AoD. The receive and transmit array response vectors are denoted by $\boldsymbol{a}_r(\phi_l^r)$ and $\boldsymbol{a}_t(\phi_l^t)$, respectively, whose mathematical expressions depend on the particular antenna array geometries.

C. Problem Formulation

In this work, we aim to select the receive antennas at the BS, $\mathbf{\Delta}$, the transmit beamforming vectors at the MSs, $\boldsymbol{v}_k$, for $k \in \mathcal{K}$, and the hybrid receiving matrices, $\mathbf{U}_{RF}$ and $\mathbf{U}_{BB}$, in order to maximize the uplink system sum-rate in (6) under the constant modulus constraints on $\mathbf{U}_{RF}$ and the transmit power constraints on $\boldsymbol{v}_k$. Mathematically, the problem can be formulated as

$$\begin{aligned} \max_{\mathbf{\Delta}, \boldsymbol{v}_k, \mathbf{U}_{RF}, \mathbf{U}_{BB}} \quad & \mathcal{R} \\ \text{s.t. } & \mathrm{Tr}(\boldsymbol{v}_k \boldsymbol{v}_k^H) \leq 1, \quad k \in \mathcal{K}, \\ & |[\mathbf{U}_{RF}]_{ij}| = 1, \quad 1 \leq i \leq R, \ 1 \leq j \leq N_{RF}, \\ & [\mathbf{\Delta}]_{ii} \in \{0, 1\}, \quad i = 1, 2, \ldots, R, \\ & \sum_{i=1}^{R} [\mathbf{\Delta}]_{ii} = K_s. \end{aligned} \tag{8}$$

Note that problem (8) is nonconvex and difficult to solve due to the constant modulus and the antenna selection constraints. Even if we fix $\mathbf{U}_{BB}$, $\mathbf{U}_{RF}$, and $\boldsymbol{v}_k$, there is still no simple approach to obtain the optimal antenna selection matrix $\mathbf{\Delta}$ unless an exhaustive search is performed. To address the optimization problem in (8), we exploit the special structure of the problem and convert it into three subproblems. We then resort to an AO method that iteratively updates the antenna

[1] Here we remark that our proposed approach can be extended to the case where the different users have their own power constraints. For example, by assuming $E(|s_k|^2) = p_k$ and defining $\boldsymbol{s} \triangleq [s_1, s_2, \ldots, s_K]^T$, we have $E(\mathrm{Tr}(\boldsymbol{s}\boldsymbol{s}^H)) = \sum_{k=1}^{K} p_k$. Further, if we define $\rho' \triangleq \rho / \sum_{k=1}^{K} p_k$, then Eq. (3) can be rewritten as $\boldsymbol{y} = \sqrt{\rho'} \mathbf{V} \boldsymbol{s} + \boldsymbol{n}$. As a result, the corresponding problem has the same form as that of problem (8) and thus it can be solved by using our proposed algorithm.

[2] For simplicity, a planar propagation geometry is assumed, but extension to three-dimensional space is possible.

$$\mathcal{R} \triangleq \log\det(\mathbf{I} + \frac{\rho}{K\sigma^2} \mathbf{U}_{BB}^H \mathbf{U}_{RF}^H \mathbf{\Delta} \mathbf{V} \mathbf{V}^H \mathbf{\Delta}^H \mathbf{U}_{RF} \mathbf{U}_{BB} (\mathbf{U}_{BB}^H \mathbf{U}_{RF}^H \mathbf{U}_{RF} \mathbf{U}_{BB})^{-1}) \tag{6}$$

selection matrix, the transmit beamformers, and the hybrid receiving matrices by sequentially addressing each subproblem while keeping the other variables fixed. In the following sections, the antenna selection matrix is first derived via the CCCP-based method. Then, a non-iterative WMMSE approach is proposed to find the transmit beamformers. Finally, the hybrid receiver is obtained by using MO.

III. Antenna Selection

In this section, we optimize the antenna selection matrix $\mathbf{\Delta}$ via the CCCP-based method while fixing the transmit beamformers and hybrid receiver. To this end, we first reformulate problem (8) into a more tractable form with respect to $\mathbf{\Delta}$.

Specifically, we abstract the BS hybrid receiver operation by assuming ideal multi-user MMSE decoding, and focus on optimizing the antenna selection matrix. As a result, the original problem in (8) can be simplified into the following approximate problem [10]:

$$\begin{aligned}&\max_{\mathbf{\Delta}} \ \log\det(\mathbf{I}+\frac{\rho}{K\sigma^2}\mathbf{\Delta}\mathbf{V}\mathbf{V}^H\mathbf{\Delta}^H)\\ &\text{s.t. } [\mathbf{\Delta}]_{ii}\in\{0,1\},\quad i=1,2,\ldots,R,\\ &\quad\sum_{i=1}^{R}[\mathbf{\Delta}]_{ii}=K_s.\end{aligned}\tag{9}$$

According to the Sylvester's determinant theorem in Eq. (24) of [35], we obtain

$$\log\det(\mathbf{I}+\frac{\rho}{K\sigma^2}\mathbf{\Delta}\mathbf{V}\mathbf{V}^H\mathbf{\Delta}^H) = \log\det(\mathbf{I}+\frac{\rho}{K\sigma^2}\mathbf{V}^H\mathbf{\Delta}^H\mathbf{\Delta}\mathbf{V}).$$

Then problem (9) can be rewritten as

$$\begin{aligned}&\max_{\mathbf{\Delta}} \ \log\det(\mathbf{I}+\frac{\rho}{K\sigma^2}\mathbf{V}^H\mathbf{\Delta}^H\mathbf{\Delta}\mathbf{V})\\ &\text{s.t. } [\mathbf{\Delta}]_{ii}\in\{0,1\},\quad i=1,2,\ldots,R,\\ &\quad\sum_{i=1}^{R}[\mathbf{\Delta}]_{ii}=K_s.\end{aligned}\tag{10}$$

Furthermore, since each $[\mathbf{\Delta}]_{ii}$ is either 0 or 1 and $\mathbf{\Delta}$ is a diagonal matrix, we have $[\mathbf{\Delta}]_{ii}^2=[\mathbf{\Delta}]_{ii}$, implying $\mathbf{\Delta}^H\mathbf{\Delta}=\mathbf{\Delta}$ and $\sum_{i=1}^{R}[\mathbf{\Delta}]_{ii}^2 = K_s$ (due to the second equality constraint). Therefore, problem (10) can be converted to

$$\begin{aligned}&\max_{\mathbf{\Delta}} \ \log\det(\mathbf{I}+\frac{\rho}{K\sigma^2}\mathbf{V}^H\mathbf{\Delta}\mathbf{V})+\beta(\sum_{i=1}^{R}[\mathbf{\Delta}]_{ii}^2-K_s)\\ &\text{s.t. } [\mathbf{\Delta}]_{ii}\in\{0,1\},\quad i=1,2,\ldots,R,\\ &\quad\sum_{i=1}^{R}[\mathbf{\Delta}]_{ii}=K_s,\end{aligned}\tag{11}$$

where β is a positive constant. Note that the second term in the objective function of the above problem is strongly convex while the Hessian of the first term in the objective function is bounded over the compact constraint set. Thus, it can be shown that there exists a sufficiently large β such that the objective of problem (11) is convex. Since maximizing a convex function over a simplex must yields solution at the vertex of the simplex, problem (11) is equivalent to

$$\begin{aligned}&\max_{\mathbf{\Delta}} \ \log\det(\mathbf{I}+\frac{\rho}{K\sigma^2}\mathbf{V}^H\mathbf{\Delta}\mathbf{V})+\beta(\sum_{i=1}^{R}[\mathbf{\Delta}]_{ii}^2-K_s)\\ &\text{s.t. } 0\le[\mathbf{\Delta}]_{ii}\le 1,\quad i=1,2,\ldots,R,\\ &\quad\sum_{i=1}^{R}[\mathbf{\Delta}]_{ii}=K_s,\end{aligned}\tag{12}$$

which is obtained by relaxing the binary constraints, i.e., linear relaxation [21], [34].

From the above discussion, it is important to determine β, which is explained in the following Theorem 1, whose proof is presented in Appendix A.

Theorem 1: Let a be the second-order derivative of $\sum_{i=1}^{R}[\mathbf{\Delta}]_{ii}^2-K_s$ with respect to $[\mathbf{\Delta}]_{ii}$ and let b represent an upper bound on the absolute value of the second-order derivative of $\log\det(\mathbf{I}+\frac{\rho}{K\sigma^2}\mathbf{V}^H\mathbf{\Delta}\mathbf{V})$ with respect to $[\mathbf{\Delta}]_{ii}$. Then, problem (12) admits a binary solution for $[\mathbf{\Delta}]_{ii}$ when $\beta=b/a$.

Now we are ready to solve the convex problem (12). Since the objective of the problem is to maximize a convex function with a special constraint (which is a simplex), we can apply the CCCP-based method [29] to it. To illustrate this approach, let us define

$$f(\mathbf{\Delta})\triangleq\log\det(\mathbf{I}+\frac{\rho}{K\sigma^2}\mathbf{V}^H\mathbf{\Delta}\mathbf{V})+\beta(\sum_{i=1}^{R}[\mathbf{\Delta}]_{ii}^2-K_s),\tag{13}$$

and

$$\mathcal{S}\triangleq\{\mathbf{\Delta}\mid 0\le[\mathbf{\Delta}]_{ii}\le 1, i\in\{1,2,\ldots,R\},\sum_{i=1}^{R}[\mathbf{\Delta}]_{ii}=K_s\}.\tag{14}$$

Since $f(\mathbf{\Delta})$ is convex, according to Appendix A, the following inequality always holds at the given point $\mathbf{\Delta}_{\text{ini}}$,

$$f(\mathbf{\Delta})\ge f(\mathbf{\Delta}_{\text{ini}})+\sum_{i=1}^{R}g_i([\mathbf{\Delta}]_{ii}-[\mathbf{\Delta}_{\text{ini}}]_{ii}),\ \ \mathbf{\Delta}\in\mathcal{S},\tag{15}$$

where

$$\begin{aligned}g_i&\triangleq\left.\frac{\partial f(\mathbf{\Delta})}{\partial[\mathbf{\Delta}]_{ii}}\right|_{\mathbf{\Delta}=\mathbf{\Delta}_{\text{ini}}}\\ &=[\frac{\rho}{K\sigma^2}\mathbf{V}(\mathbf{I}+\frac{\rho}{K\sigma^2}\mathbf{V}^H\mathbf{\Delta}_{\text{ini}}\mathbf{V})^{-1}\mathbf{V}^H]_{ii}+2\beta[\mathbf{\Delta}_{\text{ini}}]_{ii}.\end{aligned}\tag{16}$$

Hence we can maximize the tight lower bound of $f(\mathbf{\Delta})$, i.e. the right side of (15), leading to the following problem[3]

$$\mathbf{\Delta}_{\text{CCCP}}=\arg\max_{\mathbf{\Delta}\in\mathcal{S}}\ \sum_{i=1}^{R}g_i[\mathbf{\Delta}]_{ii}.\tag{17}$$

[3]It is worth mentioning that problem (17) can be also viewed as a subproblem of the Frank-Wolfe method [36, pp. 262–266] as applied to (12). However, since we are maximizing a convex function, it is not necessary to further update the initial value of $\mathbf{\Delta}$ used for the next iteration as in the Frank-Wolfe method. Hence, although the CCCP method shares the same subproblem with the Frank-Wolfe method, these two methods are essentially different when applied to problem (12).

TABLE I
PSEUDO CODE OF THE CCCP-BASED METHOD FOR OPTIMIZING $\mathbf{\Delta}$

1 Initialize $\mathbf{\Delta}_{\text{ini}}$ such that $\mathbf{\Delta}_{\text{ini}} \in \mathbf{S}$;
2 Determine β according to Theorem 1;
3 Solve problem (17);
4 Return $\mathbf{\Delta}_{\text{CCCP}}$.

Since problem (17) is a linear program with constraint set (14), its solution has a special structure. Specifically, for the indices of the K_s largest values of g_i, the corresponding entries of the optimal $\mathbf{\Delta}$ are 1 while the remaining entries are 0. This implies that problem (17) automatically admits a closed-form binary solution. Thus the CCCP-based method applied to (12) can be efficiently implemented; the detailed procedure is summarized in Table I.

IV. OPTIMIZING BEAMFORMER

In this section, we optimize the transmit beamformers $\boldsymbol{v}_k$, given the antenna selection matrix and the hybrid beamformer at the BS receiver. As mentioned earlier, it is difficult to optimize the $\boldsymbol{v}_k$ on the basis of the complete sum-rate objective in (8). Here, we use the WMMSE approach to transform the sum-rate maximization problem into a weighted MSE minimization problem, which is easier to solve.

To obtain sub-optimal beamformers $\boldsymbol{v}_k$, we focus on the following optimization problem,

$$\begin{aligned} \max_{\boldsymbol{v}_k} \ & \log\det(\mathbf{I} + \frac{\rho}{K\sigma^2}\mathbf{\Delta}\mathbf{V}\mathbf{V}^H\mathbf{\Delta}^H) \\ \text{s.t. } & \text{Tr}(\boldsymbol{v}_k\boldsymbol{v}_k^H) \le 1, \quad k \in \mathcal{K}. \end{aligned} \tag{18}$$

The WMMSE approach can obtain a stationary solution to this problem [30] by using an iterative block coordinate descent (BCD) method. However, in our proposed approach, we do not really need to find a stationary point of (18). Instead, to reduce the computational complexity, we propose to use a simplified, non-iterative form of weighted MSE reduction that leads to improved beamforming vectors, as explained below.

We first reformulate the sum-rate maximization problem (18) into an equivalent problem involving weighted MSE minimization. The MSE matrix of the received signal vector at the BS can be expressed as

$$\begin{aligned} \mathbf{E}(\mathbf{U},\mathbf{V}) &\triangleq E[(\tilde{\boldsymbol{s}}-\boldsymbol{s})(\tilde{\boldsymbol{s}}-\boldsymbol{s})^H] \\ &= \frac{1}{K}(\mathbf{I}-\sqrt{\rho}\mathbf{U}^H\mathbf{\Delta}\mathbf{V})(\mathbf{I}-\sqrt{\rho}\mathbf{U}^H\mathbf{\Delta}\mathbf{V})^H \\ &\quad + \sigma^2\mathbf{U}^H\mathbf{U}, \end{aligned} \tag{19}$$

where $\tilde{\boldsymbol{s}} = \mathbf{U}^H\boldsymbol{y}_{\mathbf{\Delta}}$ and $\mathbf{U}$ is an abstract fully digital receiver. At this time, $\mathbf{U}$ only serves as an intermediate variable in the update of the beamformers $\boldsymbol{v}_k$ but it will be further decomposed into an hybrid product $\mathbf{U}_{RF}\mathbf{U}_{BB}$ later in Section V. The corresponding MSE minimization problem can then be written as

$$\begin{aligned} \min_{\boldsymbol{v}_k,\mathbf{U}} \ & \text{Tr}(\mathbf{E}(\mathbf{U},\mathbf{V})) \\ \text{s.t. } & \text{Tr}(\boldsymbol{v}_k\boldsymbol{v}_k^H) \le 1, \ k \in \mathcal{K}. \end{aligned} \tag{20}$$

By fixing $\boldsymbol{v}_k$, solution of the above problem leads to the well-known MMSE receiver

$$\mathbf{U}^{\text{mmse}} = \frac{\sqrt{\rho}}{K}(\frac{\rho}{K}\mathbf{\Delta}\mathbf{V}\mathbf{V}^H\mathbf{\Delta}^H + \sigma^2\mathbf{I})^{-1}\mathbf{\Delta}\mathbf{V}. \tag{21}$$

Substituting (21) into (19), the corresponding MSE matrix reduces to

$$\mathbf{E}^{\text{mmse}} = \frac{1}{K}(\mathbf{I} - \frac{\rho}{K}\mathbf{V}^H\mathbf{\Delta}^H(\frac{\rho}{K}\mathbf{\Delta}\mathbf{V}\mathbf{V}^H\mathbf{\Delta}^H + \sigma^2\mathbf{I})^{-1}\mathbf{\Delta}\mathbf{V}). \tag{22}$$

Directly following the proof in [30], we have the following theorem which establishes the equivalence between the sum-rate maximization problem and the weighted MSE minimization problem.

Theorem 2: Let $\mathbf{W} \succeq 0$ be a weighting matrix. The problem

$$\begin{aligned} \min_{\boldsymbol{v}_k,\mathbf{U},\mathbf{W}} \ & \text{Tr}(\mathbf{W}\mathbf{E}) - \log\det(\mathbf{W}) \\ \text{s.t. } & \text{Tr}(\boldsymbol{v}_k\boldsymbol{v}_k^H) \le 1, \quad k \in \mathcal{K}, \end{aligned} \tag{23}$$

is equivalent to problem (18), in the sense that they share the same KKT solution set.

It can be observed that the weighted MSE cost function in (23) is convex in each block of the optimization variables, that is $\mathbf{U}$, $\mathbf{W}$, and $\boldsymbol{v}_k$. Moreover, the constraints are separable across the block variables. Therefore, the BCD method [36, pp. 323–330] applies to problem (23). Specifically, we minimize the weighted-MSE cost function with respect to one block of variables while fixing the other blocks, hence leading to the following three steps.

i) Update $\mathbf{U}$ while fixing $\boldsymbol{v}_k$ and $\mathbf{W}$: this step yields the MMSE receiver $\mathbf{U}^{\text{mmse}}$ given by (21).
ii) Update $\mathbf{W}$ while fixing $\mathbf{U}$ and the $\boldsymbol{v}_k$: this step yields a closed-form solution as follows

$$\mathbf{W}^{opt} = (\mathbf{E}^{\text{mmse}})^{-1}. \tag{24}$$

iii) Update $\boldsymbol{v}_k$ while fixing $\mathbf{W}$ and $\mathbf{U}$: in this step, we solve the following problem for $k = 1, 2, \ldots, K$,

$$\begin{aligned} \min_{\boldsymbol{v}_k} \ & \text{Tr}[\frac{1}{K}\mathbf{W}(\mathbf{I}-\sqrt{\rho}\mathbf{U}^H\mathbf{\Delta}\mathbf{V})(\mathbf{I}-\sqrt{\rho}\mathbf{U}^H\mathbf{\Delta}\mathbf{V})^H] \\ \text{s.t. } & \text{Tr}(\boldsymbol{v}_k\boldsymbol{v}_k^H) \le 1, \ k \in \mathcal{K}. \end{aligned} \tag{25}$$

The above is a convex quadratic optimization problem with K separable constraints; accordingly, it can be solved by Lagrange multipliers.

We do these three steps only once, which ensures that $\boldsymbol{v}_k$ are improved (in the sense of improving the objective).

V. OPTIMIZING HYBRID RECEIVER

In this section, inspired by [16], we use MO to search the solution for $\mathbf{U}_{BB}$ and $\mathbf{U}_{RF}$ by fixing $\mathbf{\Delta}$ and $\boldsymbol{v}_k$. However, our approach considers the original sum-rate maximization problem (8), which is different from the Frobenius norm based approach in [16].

TABLE II

PSEUDO CODE OF MO ALGORITHM FOR $\mathbf{U}_{RF}$ (SEE ALSO APPENDIX B)

1 Initialize $\mathbf{U}_{RF,0}$ such that $\boldsymbol{x}_0 = vec(\mathbf{U}_{RF,0}) \in \mathcal{M}^m$;
2 $\mathbf{d}_0 = -\text{grad}p(\boldsymbol{x}_0)$ and $r = 0$;
3 **repeat**
4 Choose the Armijo backtracking line search step size α_r;
5 Find the next point $\boldsymbol{x}_{r+1}$ and the corresponding $\mathbf{U}_{RF,r+1}$using retraction in (37);
6 Determine Riemannian gradient $\mathbf{g}_{r+1} = \text{grad}p(\boldsymbol{x}_{r+1})$ according to (35) and (36);
7 Calculate vector transports $\mathbf{g}_r^+$ and $\mathbf{d}_r^+$ of the gradient $\mathbf{g}_r$ and conjugate direction $\mathbf{d}_r$ from $\boldsymbol{x}_r$ to $\boldsymbol{x}_{r+1}$;
8 Choose the Polak-Ribiere parameter γ_{r+1};
9 Compute conjugate direction $\mathbf{d}_{r+1} = -\mathbf{g}_{r+1} + \gamma_{r+1}\mathbf{d}_r^+$;
10 $r \leftarrow r + 1$;
11 **until** a stopping criterion is satisfied.

TABLE III

PSEUDO CODE OF THE CCCP-WMMSE-MO ALGORITHM

1 Initialize $\boldsymbol{\Delta}$, $\boldsymbol{v}_k$, $\mathbf{U}_{RF}$ and $\mathbf{U}_{BB}$, such that they meet all the constraints;
2 **repeat**
3 use CCCP-based method to find $\boldsymbol{\Delta}$;
4 $\mathbf{U} \leftarrow \frac{\sqrt{\rho}}{K}(\frac{\rho}{K}\boldsymbol{\Delta}\mathbf{V}\mathbf{V}^H\boldsymbol{\Delta}^H + \sigma^2\mathbf{I})^{-1}\boldsymbol{\Delta}\mathbf{V}$;
5 $\mathbf{W} \leftarrow (\frac{1}{K}(\mathbf{I} - \sqrt{\rho}\mathbf{V}^H\boldsymbol{\Delta}^H\mathbf{U}))^{-1}$;
6 for $k = 1 : K$
7 use Lagrange multiplier to solve (25) and find $\boldsymbol{v}_k$;
8 end
9 **until** a stopping criterion is satisfied;
10 **repeat**
11 $\mathbf{U}_{BB} \leftarrow \mathbf{U}_{RF}^{\dagger}\mathbf{U}$;
12 use MO algorithm to find $\mathbf{U}_{RF}$;
13 **until** a stopping criterion is satisfied.

A. Baseband Receiver Design

First, we design the baseband receiver $\mathbf{U}_{BB}$ while fixing the analog receiver $\mathbf{U}_{RF}$. Since we already have fully digital MMSE receiver $\mathbf{U}^{\text{mmse}}$ from the beamformer optimization, we simply need to exploit the decomposition of $\mathbf{U}^{mmse}$ into the product of $\mathbf{U}_{RF}$ and $\mathbf{U}_{BB}$. As a result, we can obtain the well-known least squares solution for the baseband beamformer, as given by

$$\mathbf{U}_{BB} = \mathbf{U}_{RF}^{\dagger}\mathbf{U}^{\text{mmse}}. \tag{26}$$

Clearly, this is a globally optimal solution for the baseband receiver design with a fixed analog receiver.

B. Analog Receiver Design

We here focus on optimizing the analog receiver while fixing $\boldsymbol{\Delta}$, $\boldsymbol{v}_k$ and $\mathbf{U}_{BB}$, i.e., update $\mathbf{U}_{RF}$ by finding the solution of the following optimization problem

$$\begin{aligned} &\max_{\mathbf{U}_{RF}} \ \mathcal{R} \\ &\text{s.t. } |[\mathbf{U}_{RF}]_{ij}| = 1, \quad 1 \le i \le R, \ 1 \le j \le N_{RF}. \end{aligned} \tag{27}$$

This problem is complex mainly due to the constant modulus constraints, which is intrinsically nonconvex. Here we use MO to address the issue, which has the following two advantages: 1) the rich geometry of Riemannian manifolds makes it possible to define gradients of cost functions in terms of vector fields; 2) optimization over a Riemannian manifold is locally analogous to that over a Euclidean space with smooth constraints. As a result, we can resort to the well developed conjugate gradient algorithm in Euclidean spaces to find its counterpart in Riemannian manifolds, as further explained in Appendix B. In [16], the hybrid receiver is designed by minimizing its Frobenius distance to the optimal fully digital receiver. Here, by exploiting the idea of MO, we directly solve the sum-rate maximization problem (27), which results in better performance.

Based on Riemannian geometry, we can develop an iterative algorithm to optimize the analog receiver matrix $\mathbf{U}_{RF}$, which is summarized in Table II. In this algorithm, the use of the well-known Armijo backtracking line search step size in Step 1 and of the Polak-Ribiere parameter in Step 8 ensures that the objective function is non-decreasing at each iteration according to [36, p. 129]. In Step 7, a transport is a specific mapping between two tangent vectors in different tangent spaces, which can be expressed as

$$\text{Transp}_{\boldsymbol{x}_r \to \boldsymbol{x}_{r+1}}(\mathbf{d}) \triangleq \mathbf{d} - \mathbb{R}(\mathbf{d} \circ \boldsymbol{x}_{r+1}^*) \circ \boldsymbol{x}_{r+1}. \tag{28}$$

According to [16] and [37, pp. 63–65], MO for problem (27) converges to a stationary point, i.e., the point where the gradient of the objective function is zero.

C. Complete Algorithm and Computational Complexity Analysis

According to the aforementioned results, we summarize the proposed joint antenna selection and transceiver design algorithm in Table III, which is referred to as CCCP-WMMSE-MO.

Since R is a large number, the dominant computational complexity of optimizing the transmit beamformers in the transceiver design without antenna selection is given by $O(R^3 + R^2K + RK^2 + RTK^2 + RT^2K + RK^3 + RTK)$. With antenna selection the computational complexity of this optimization reduces to $O(R^3+RKK_s+RK_s+RT^2K+RK^2+RTK+RK^2K_s)$. Compared to the optimization of the transmit beamformers, the computational complexity of the MO-based hybrid receiver design is relatively low since it only requires the nested loops of a line search process and the Kronecker products of two matrices (see in Appendix B). Finally, the computational complexity of the proposed antenna selection method is given by $O(R^2K + RK^2)$. Hence, the overall complexity of the proposed transceiver design algorithm is similar to that of the transceiver design using all antennas. However, with the aid of the proposed algorithm with antenna selection, the energy consumption at the BS can be significantly reduced, as will be demonstrated by simulation results in Section VII.

VI. Convergence Analysis

The proposed iterative transceiver design algorithm, shown in Table III, is a two-phase algorithm. In this section, we analyze the convergence of our proposed iterative transceiver design algorithm. The convergence property for the first phase (i.e. Steps 2-10 in Table III) is summarized in Theorem 3.

Theorem 3: Any limit point of the sequence generated by the first phase of the algorithm in Table III is a stationary point of the following optimization problem:

$$\begin{aligned} &\max_{\mathbf{\Delta}, \boldsymbol{v}_k} \log\det(\mathbf{I} + \frac{\rho}{K\sigma^2}\mathbf{\Delta}\mathbf{V}\mathbf{V}^H\mathbf{\Delta}^H) + \beta(\sum_{i=1}^{R}[\mathbf{\Delta}]_{ii}^2 - K_s) \\ &\text{s.t. } \mathrm{Tr}(\boldsymbol{v}_k\boldsymbol{v}_k^H) \le 1, \quad k \in \mathcal{K}, \\ &\quad \mathbf{\Delta} \in \mathbf{S}, \end{aligned} \tag{29}$$

which is an approximative problem assuming a specific choice of receiver, i.e. MMSE decoder.

The detailed proof is presented in Appendix C, where we first establish the equivalence among the objective functions in (9), (12), (18), (23) and (29) by using Theorem 1 and Theorem 2. Note that step 3 in Table III is equivalent to globally solving problem (9). We then demonstrate that the non-iterative WMMSE is equivalent to minimizing a locally tight upper bound of (18). Hence, the first phase is in essence the block successive upper-bound minimization (BSUM) algorithm [38] applied to the optimization problem (29). Finally, based on the convergence properties of the BSUM algorithm [38], it follows that repeating Steps 2-10 can finally reach a stationary solution of problem (29)

For the second phase (Steps 11-14) in Table III, directly following the proof in [37, pp. 63–65], we can conclude that:

Theorem 4: Any limit point of the sequence generated by the second phase of the algorithm in Table III is a stationary point of (27).

From Theorem 3 and Theorem 4, the objective function values of (27) and (29) generated by the CCCP-WMMSE-MO algorithm increase monotonically. In practice, both loops will end after only a few steps, which will be shown in Section VII through numerical experiments. According to our simulation results, a near-optimal solution is often obtained by applying the CCCP-WMMSE-MO algorithm to the optimization problem (8).

VII. Simulation Results

In this section, we evaluate the sum-rate performance of the proposed transceiver design. Without loss of generality, we define $SNR \triangleq \frac{\rho}{\sigma^2}$ as in [10] and [16]. For the mmWave channel, the number of rays is set to $L = 20$. The directivity gain can be expressed as

$$\Lambda(\phi_l) = \begin{cases} 1, & \forall \phi_l \in [\phi_{\min}, \phi_{\max}], \\ 0, & \text{otherwise.} \end{cases} \tag{30}$$

We implement the Saleh Valenzuela model assuming planar geometry and uniform linear arrays for convenience. The azimuth AoDs and AoAs are generated randomly with uniform distribution over $[0, 2\pi)$. The antenna elements in the uniform linear array (ULA) are separated by a half wavelength and all

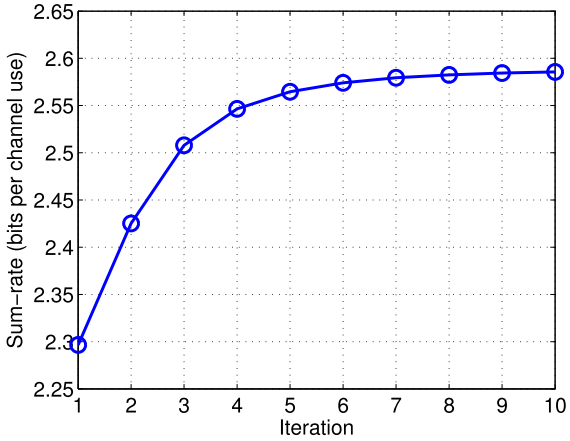

Fig. 2. The convergence performance of the first loop in the CCCP-WMMSE-MO algorithm with $R = 64$, $K_s = 32$, $K = 4$, $N_{RF} = 8$ and SNR $= 0$ dB.

simulation results are averaged over 500 channel realizations. Moreover, the initial phases of the analog receiver, $\mathbf{U}_{RF}$, follow a uniform distribution over $[0, 2\pi)$. Among all the analyzed techniques in this paper, we consider the following:

- CCCP-WMMSE-MO: the proposed algorithm described in Table III;
- CCCP-OMP-MO: the CCCP-based method and the algorithm in [16] used jointly;
- EX-WMMSE-MO: all possible combinations of BS antennas are tried and the best one is selected as the output before resorting to the WMMSE approach for $\boldsymbol{v}_k$ and MO for the hybrid receiver;
- FULL-WMMSE-MO: the CCCP-WMMSE-MO algorithm without antenna selection ($K_s = R$);
- FULL-OMP-MO: the CCCP-OMP-MO algorithm without antenna selection;
- CCCP-WMMSE-FD: the hybrid receiver is replaced with the fully digital MMSE receiver after the antenna selection matrix is obtained;
- EX-WMMSE-FD: assuming that the BS uses a fully digital MMSE receiver, we verify all possible combinations of selected antennas using exhaustive search. Then, for each fixed antenna selection, we optimize the precoding matrices using the WMMSE approach. Finally, we select the best solution, which provides an upper bound for the performance of CCCP-WMMSE-MO.

A. Convergence Performance

Fig. 2 and Fig. 3 illustrate the convergence behavior for the first and second loops of the proposed algorithm (respectively lines 2-10 and lines 11-14 in Table III) in a typical proposed algorithm in a typical situation. From these figures and other similar results, we conclude that both loops in the CCCP-WMMSE-MO algorithm converge monotonically in a few iterations.

B. Performance Versus the Number of Selected Antennas

Fig. 4 and Fig. 5 show the system sum-rate versus the number of selected antennas K_s under SNR $= 0$ dB.

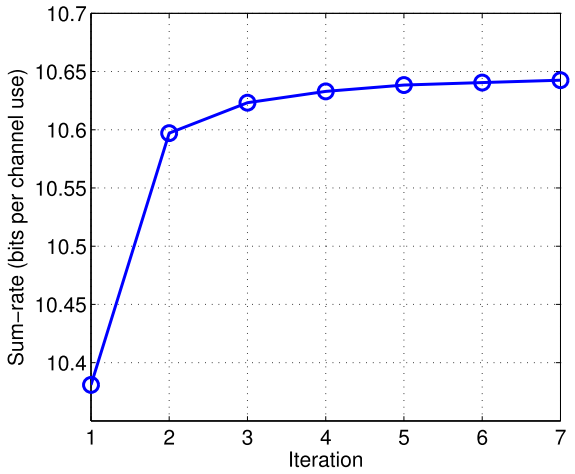

Fig. 3. The convergence performance of the second loop in the CCCP-WMMSE-MO algorithm with $R = 64$, $K_s = 32$, $K = 4$, $N_{RF} = 8$ and SNR $= 0$ dB.

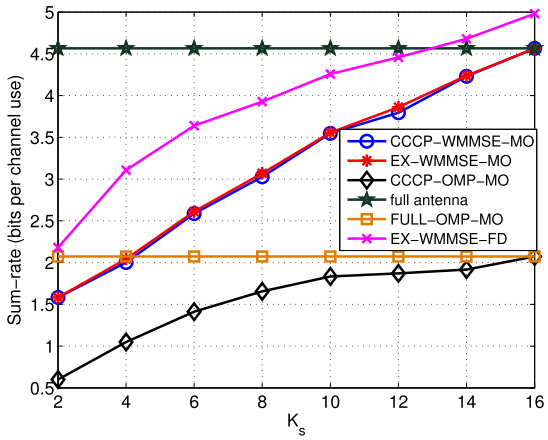

Fig. 4. The system sum-rate versus K_s when $R = 16$, $T = 2$, $K = 2$, $N_{RF} = 2$ and SNR $=$ 0dB.

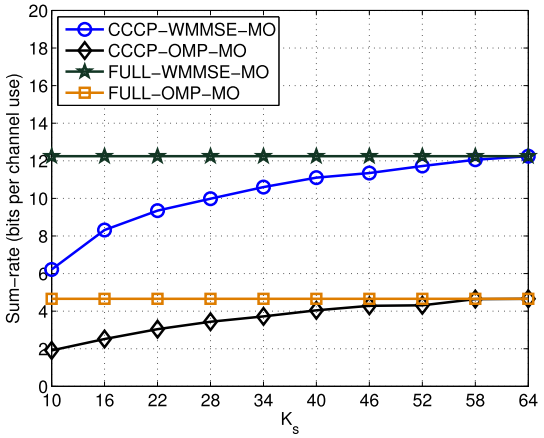

Fig. 5. The system sum-rate versus K_s when $R = 64$, $T = 2$, $K = 4$, $N_{RF} = 8$ and SNR $= 0$ dB.

To illustrate the performance of the exhaustive antenna selection method as a benchmark, we here only consider moderate-scale systems (with tens of antennas) in Fig. 4 when testing the EX-WMMSE-MO algorithm and the EX-WMMSE-FD algorithm, although it can be applied to large-scale MU-MIMO systems (possibly with hundreds of antennas). Consequently, we do not show the performance of the EX-WMMSE-MO algorithm in Fig. 5 where $R = 64$.

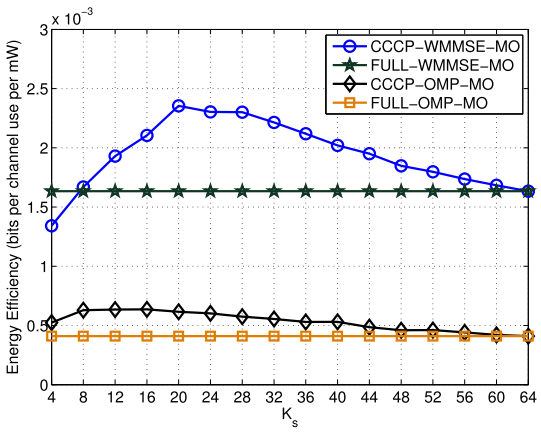

Fig. 6. The energy efficiency versus K_s when $R = 64$, $T = 2$, $K = 4$, $N_{RF} = 4$ and SNR $= 0$ dB.

From these figures, we can see that if a performance loss is accepted, the proposed algorithm can select much fewer antennas than the FULL-WMMSE-MO and therefore reduce energy consumption and implementation costs at the BS. The sum-rate performance of the proposed algorithm is very close to that of the exhaustive search method, which is the optimal antenna selection method. Hence, the CCCP-based antenna selection algorithm can perform as well as the exhaustive search method but with much lower complexity. It can also be observed that the CCCP-WMMSE-MO algorithm always significantly outperforms the CCCP-OMP-MO algorithm. Furthermore, if we select all the antennas at the BS, one can see that the sum-rate performance provided by the proposed FULL-WMMSE-MO algorithm is much higher than that of the FULL-OMP-MO algorithm, which verifies the effectiveness of the proposed transceiver design approach. In addition, the performance of the EX-WMMSE-FD algorithm is presented in Fig. 4, which serves as an upper bound for the CCCP-WMMSE-MO algorithm. As we can see, the gap between the CCCP-WMMSE-MO algorithm and the EX-WMMSE-FD algorithm is small, implying a high-quality solution provided by the CCCP-WMMSE-MO algorithm to problem (8).

Next, we consider the energy efficiency at the BS and compare the proposed CCCP-WMMSE-MO, the FULL-WMMSE-MO and the conventional CCCP-OMP-MO algorithms. The energy efficiency is defined as $\frac{\mathcal{R}}{P_r}$, where $\mathcal{R}$ is the system sum-rate and P_r is the energy consumption at the BS. In particular, $P_r \triangleq P_{r,1} + P_{r,2} + P_{r,3}$, where $P_{r,1}$ denotes the total energy consumption in the baseband receiver, $P_{r,2} = N_{RF}\tilde{P}_{r,2}$ denotes the total energy consumption of the RF chains, with $\tilde{P}_{r,2}$ being the energy consumption of each RF chain, and $P_{r,3} = K_s(\tilde{P}_{r,3} + N_{RF}\bar{P}_{r,3})$ denotes the total energy consumption of the analog phase shifters and LNAs, with $\tilde{P}_{r,3}$ and $\bar{P}_{r,3}$ denoting the individual LNA and phase shifter energy consumptions, respectively. According to [39]–[42], we have $P_{r,1} = 200mW$, $\tilde{P}_{r,2} = 120mW$ and $\tilde{P}_{r,3} = \bar{P}_{r,3} = 20mW$ in a small cell scenario. Fig. 6 shows the energy efficiency versus K_s under SNR $= 0$ dB. From the figure, we can see that the CCCP-WMMSE-MO algorithm outperforms the CCCP-OMP-MO, the FULL-WMMSE-MO, and the FULL-OMP-MO algorithms when $K_s \geq 8$. If the antenna selection strategy is not used at the BS, one can see

TABLE IV
THE AVERAGE CPU TIME OF DIFFERENT ALGORITHMS TO SOLVE PROBLEM (8) WHEN $T = 2$, $K = 2$ AND SNR $= 0$ dB

	CCCP-WMMSE-MO	EX-WMMSE-MO	CCCP-OMP-MO
$R = 16, K_s = 2, N_{RF} = 2$	6.38×10^{-2}s	10.14×10^{-2}s	2.78×10^{-2}s
$R = 20, K_s = 2, N_{RF} = 2$	6.99×10^{-2}s	16.10×10^{-2}s	2.89×10^{-2}s
$R = 16, K_s = 4, N_{RF} = 2$	18.82×10^{-2}s	88.26×10^{-2}s	8.62×10^{-2}s
$R = 16, K_s = 2, N_{RF} = 4$	6.20×10^{-2}s	12.56×10^{-2}s	8.83×10^{-2}s

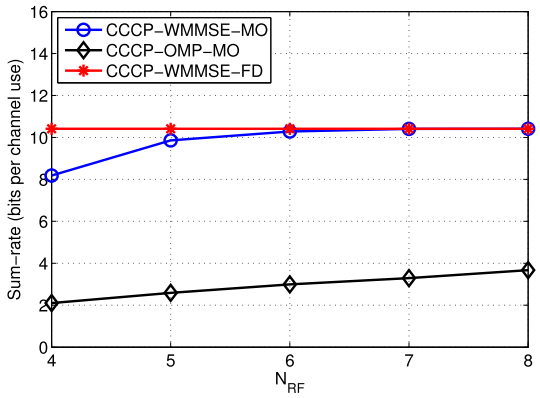

Fig. 7. The system sum-rate versus N_{RF} for $R = 64$, $T = 2$, $K = 4$ and $K_s = 32$ when SNR $= 0$ dB.

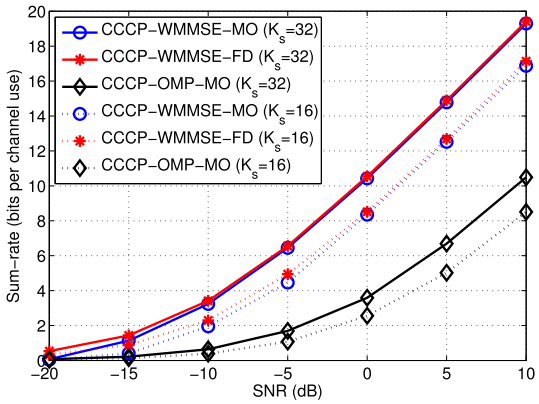

Fig. 8. The system sum-rate versus SNR for $R = 64$, $T = 2$, $K = 4$, $N_{RF} = 8$.

that the FULL-WMMSE-MO algorithm is still more energy efficient than the FULL-OMP-MO algorithm. Besides, the results show that with an increase in K_s, the BS energy efficiency of the CCCP-WMMSE-MO algorithm increases first and then decreases, and the maximum value of the energy efficiency occurs when K_s is around 20. It is because the power consumption of the antennas becomes significant when their number is large enough.

C. Performance Versus the Number of RF Chains

Fig. 7 shows the system sum-rate performance versus the number of RF chains N_{RF} when SNR $=$ 0 dB. From the figure, the proposed algorithm outperforms the CCCP-OMP-MO algorithm with a 4 to 5 bits performance gap. Furthermore, as N_{RF} increases, the performance of the CCCP-WMMSE-MO algorithm is enhanced monotonically and coincides with the CCCP-WMMSE-FD algorithm when $N_{RF} = 8$. This means that the proposed algorithm can achieve optimal sum-rate with much less RF chains, which in turns reduces the hardware complexity and implementation cost.

D. Performance Versus SNR

Fig. 8 compares the performance of different algorithms versus SNR for $K_s = 32$ and $K_s = 16$. From the figure, the proposed algorithm outperforms the conventional CCCP-OMP-MO algorithm with a 10 dB gap for both $K_s = 32$ and 16 cases, and its performance comes very close to that of the CCCP-WMMSE-FD algorithms. For such system configurations, the proposed CCCP-WMMSE-MO algorithm serves as an excellent candidate for joint transceiver design with antenna selection, achieving both good performance and low complexity.

E. Complexity Evaluation

We present the average CPU time of different algorithms in Table IV. From these results, we note that the CCCP-WMMSE-MO algorithm takes less time to solve problem (8) than the EX-WMMSE-MO algorithms for the given choices of parameters. While the CPU times of the proposed algorithm exceed that of the CCCP-OMP-MO algorithm, it can achieve much higher sum-rate as demonstrated earlier. Besides, the CCCP-WMMSE-MO and the EX-WMMSE-MO algorithms are sensitive to K_s, while it can be shown that the average CPU time of the CCCP-OMP-MO algorithm increases significant with K_s and N_{RF}.

VIII. CONCLUSIONS

In this paper, we have investigated joint transceiver design with antenna selection at the BS to maximize the sum-rate for massive MU-MIMO uplink systems operating at mmWave. The problem was separated into three subproblems and a novel alternating algorithm named CCCP-WMMSE-MO was proposed for their iterative solution. By exploiting the CCCP-based method, we solved the antenna selection problem and directly obtained a binary antenna selection matrix. Then the WMMSE approach was used to optimize the transmit beamformers while we resorted to the MO method to update the hybrid receiver. The computational complexity and convergence behavior of the proposed CCCP-WMMSE-MO algorithm have been analyzed. The numerical results have shown that the proposed algorithm can achieve an outstanding performance, which is close to that of the algorithms using fully digital transceiver and exhaustive search method. In terms of energy efficiency at the BS, the CCCP-WMMSE-MO algorithm outperforms the FULL-WMMSE-MO algorithm. Our results and analysis indicate that the proposed CCCP-WMMSE-MO algorithm can serve as an excellent candidate for joint MU-MIMO transceiver design with antenna selection in mmWave uplink applications, achieving good spectrum efficiency, as well as low energy consumption and implementation complexity.

APPENDIX A
PROOF OF THEOREM 1

Proof: Considering the idea behind CCCP, since $\sum_{i=1}^{R}[\mathbf{\Delta}]_{ii}^2 - K_s$ is strongly convex with second-order derivative with respect to $[\mathbf{\Delta}]_{ii}$, we can choose β according to Theorem 1 such that the objective function in (12) is convex. Then problem (12) is to maximize a convex function subject to the linear constraints. Therefore the optimal solution for $[\mathbf{\Delta}]_{ii}$ must be one of the bounds of the interval, i.e. 0 or 1. ■

APPENDIX B
GRADIENT SEARCH ON THE RIEMANNIAN MANIFOLDS

The background theory on manifolds and manifold optimization can be found in [37, pp. 17–53], [43, pp. 1–36], and [44, pp. 1–31]. We define the Euclidean metric of the complex plane $\mathbb{C}$ as

$$\langle x_1, x_2\rangle = \mathbb{R}(x_1^* x_2). \tag{31}$$

Therefore, the complex circle can be denoted as

$$\mathcal{M}_{cc} = \{x \in \mathbb{C} : \langle x, x\rangle = 1\}. \tag{32}$$

Letting x be an arbitrary point on the manifold $\mathcal{M}_{cc}$, we characterize the directions along which it can move as the tangent vectors. Accordingly, the tangent space at the point x can be represented by

$$\mathcal{T}_x\mathcal{M}_{cc} = \{z \in \mathbb{C} : z^* x + x^* z = 2\langle x, z\rangle = 0\}. \tag{33}$$

With reference to problem (27), we denote $\boldsymbol{x} = vec(\mathbf{U}_{RF})$. This vector forms a complex circle manifold $\mathcal{M}_{cc}^m = \{\boldsymbol{x} \in \mathbb{C}^{m\times 1} : |[\boldsymbol{x}]_1| = |[\boldsymbol{x}]_2| = \cdots = |[\boldsymbol{x}]_m| = 1\}$, where $m = RN_{RF}$. Then we can search the solution of problem (27) over a Cartesian product of m circles in the complex plane, which is a Riemannian submanifold of $\mathbb{C}^{m\times 1}$. Hence, the tangent space at a given point $\boldsymbol{x}$ can be expressed as

$$\mathcal{T}_x\mathcal{M}_{cc}^m = \{\boldsymbol{z} \in \mathbb{C}^{m\times 1} : \mathbb{R}(\boldsymbol{z} \circ \boldsymbol{x}^*) = \mathbf{0}\}. \tag{34}$$

Similar to the Euclidean space, we can find the direction of greatest increase of a function $p(\boldsymbol{x})$ defined on $\mathcal{M}_{cc}^m$ among all the tangent vectors. Considering the relationship between the tangent vector $\mathrm{grad}p(\boldsymbol{x})$ on $\mathcal{M}_{cc}^m$ and the Euclidean gradient $\triangledown p(\boldsymbol{x})$, $\mathrm{grad}p(\boldsymbol{x})$ can be defined as

$$\mathrm{grad}p(\boldsymbol{x}) \triangleq \triangledown p(\boldsymbol{x}) - \mathbb{R}[\triangledown p(\boldsymbol{x}) \circ \boldsymbol{x}^*] \circ \boldsymbol{x}. \tag{35}$$

The Euclidean gradient of the cost function $p(\boldsymbol{x}) = -\mathcal{R}(\boldsymbol{x})$ in (27) is obtained as Eq. (36), as shown at the bottom of this page.

Once a tangent vector is obtained, we can do retraction to map a vector from the tangent space onto the manifold itself. Following this key step of manifold optimization, the destination on the manifold when moving along a tangent vector is determined. The retraction of a tangent vector $\alpha\mathbf{d}$ at $\boldsymbol{x}$ can be expressed as

$$\mathrm{Retr}_{\mathbf{x}}(\alpha\mathbf{d}) \triangleq vec\left[\frac{[\boldsymbol{x} + \alpha\mathbf{d}]_i}{|[\boldsymbol{x} + \alpha\mathbf{d}]_i|}\right]. \tag{37}$$

Exploiting the concepts of tangent space, Riemannian gradient and retraction on the complex manifold $\mathcal{M}_{cc}^m$, we can develop a line search based conjugate gradient algorithm to maximize the sum-rate as shown in Table II, which is a classical search technique in Euclidean space.

APPENDIX C
PROOF OF THEOREM 3

Proof: Following Theorem 1, we have the following identities:

$$\log\det(\mathbf{I} + \frac{\rho}{K\sigma^2}\mathbf{\Delta}\mathbf{V}\mathbf{V}^H\mathbf{\Delta}^H) = \log\det(\mathbf{I} + \tfrac{\rho}{K\sigma^2}\mathbf{V}^H\mathbf{\Delta}\mathbf{V}) \tag{38}$$

Thus, Step 3 in Table III is equivalent to applying CCCP to solve

$$\min_{\mathbf{\Delta}\in\mathbf{S}} \; -\log\det(\mathbf{I} + \frac{\rho}{K\sigma^2}\mathbf{\Delta}\mathbf{V}\mathbf{V}^H\mathbf{\Delta}^H) - \beta(\sum_{i=1}^{R}[\mathbf{\Delta}]_{ii}^2 - K_s). \tag{39}$$

Note that the CCCP-based method is equivalent to successively minimizing a sequence of upper bounds of the objective of the above problem.

Furthermore, given the current $\mathbf{V}$, denoted by $\tilde{\mathbf{V}}$, we can easily obtain that [45]

$$\log\det(\mathbf{I} + \frac{\rho}{K\sigma^2}\mathbf{\Delta}\mathbf{V}\mathbf{V}^H\mathbf{\Delta}^H) \geq l_\tau(\mathbf{V}, \tilde{\mathbf{V}}), \quad \forall \boldsymbol{v}_k, \tag{40}$$

where

$$l_\tau(\mathbf{V}, \tilde{\mathbf{V}}) \triangleq \log\det(\mathbf{W}^{\mathrm{mmse}}) - \mathrm{Tr}(\mathbf{W}^{\mathrm{mmse}}\mathbf{E}(\mathbf{U}^{\mathrm{mmse}}, \mathbf{V})) + \tau, \tag{41}$$

$$\mathbf{W}^{\mathrm{mmse}} = \frac{1}{K}(\mathbf{I} - \sqrt{\rho}\mathbf{V}^H\mathbf{\Delta}^H\mathbf{U}))^{-1}, \tag{42}$$

$$\mathbf{U}^{\mathrm{mmse}} = \frac{\sqrt{\rho}}{K}(\frac{\rho}{K}\mathbf{\Delta}\mathbf{V}\mathbf{V}^H\mathbf{\Delta}^H + \sigma^2\mathbf{I})^{-1}\mathbf{\Delta}\mathbf{V}, \tag{43}$$

and τ is a small positive number.

Moreover, since $\mathbf{W}^{\mathrm{mmse}} = (\mathbf{E}(\mathbf{U}^{\mathrm{mmse}}, \tilde{\mathbf{V}}))^{-1}$, we have

$$\log\det(\mathbf{I} + \frac{\rho}{K\sigma^2}\mathbf{\Delta}\tilde{\mathbf{V}}\tilde{\mathbf{V}}^H\mathbf{\Delta}^H) = l_\tau(\tilde{\mathbf{V}}, \tilde{\mathbf{V}}), \forall\tilde{\mathbf{V}}. \tag{44}$$

$$\begin{aligned}\triangledown p(\boldsymbol{x}) = vec[&-\frac{2\rho}{K\sigma^2}\mathbf{\Delta}\mathbf{V}\mathbf{V}^H\mathbf{\Delta}^H\mathbf{U}_{RF}\mathbf{U}_{BB}(\mathbf{U}_{BB}^H\mathbf{U}_{RF}^H\mathbf{U}_{RF}\mathbf{U}_{BB})^{-H}\\ &(\mathbf{I} + \frac{\rho}{K\sigma^2}\mathbf{U}_{BB}^H\mathbf{U}_{RF}^H\mathbf{\Delta}\mathbf{V}\mathbf{V}^H\mathbf{\Delta}^H\mathbf{U}_{RF}\mathbf{U}_{BB}(\mathbf{U}_{BB}^H\mathbf{U}_{RF}^H\mathbf{U}_{RF}\mathbf{U}_{BB})^{-1})^{-H}\mathbf{U}_{BB}^H\\ &+\frac{2\rho}{K\sigma^2}\mathbf{U}_{RF}\mathbf{U}_{BB}(\mathbf{U}_{BB}^H\mathbf{U}_{RF}^H\mathbf{U}_{RF}\mathbf{U}_{BB})^{-H}\mathbf{U}_{BB}^H\mathbf{U}_{RF}^H\mathbf{\Delta}\mathbf{V}\mathbf{V}^H\mathbf{\Delta}^H\mathbf{U}_{RF}\mathbf{U}_{BB}\\ &(\mathbf{I} + \frac{\rho}{K\sigma^2}\mathbf{U}_{BB}^H\mathbf{U}_{RF}^H\mathbf{\Delta}\mathbf{V}\mathbf{V}^H\mathbf{\Delta}^H\mathbf{U}_{RF}\mathbf{U}_{BB}(\mathbf{U}_{BB}^H\mathbf{U}_{RF}^H\mathbf{U}_{RF}\mathbf{U}_{BB})^{-H}(\mathbf{U}_{BB}^H\mathbf{U}_{RF}^H\mathbf{U}_{RF}\mathbf{U}_{BB})^{-H}\mathbf{U}_{BB}^H]\end{aligned} \tag{36}$$

Hence, it can be verified that

$$\nabla_{\mathbf{V}} \log\det(\mathbf{I} + \frac{\rho}{K\sigma^2}\mathbf{\Delta}\mathbf{V}\mathbf{V}^H\mathbf{\Delta}^H) = \nabla_{\mathbf{V}} l_\tau(\mathbf{V}, \tilde{\mathbf{V}}). \quad (45)$$

As a result, combining (40), (44) and (45), we infer that l_τ is a locally tight lower bound [38] of $\log\det(\mathbf{I} + \frac{\rho}{K\sigma^2}\mathbf{\Delta}\mathbf{V}\mathbf{V}^H\mathbf{\Delta}^H)$. Hence, steps 4-9 are equivalent to minimizing a locally tight upper bound of $-\log\det(\mathbf{I} + \frac{\rho}{K\sigma^2}\mathbf{\Delta}\mathbf{V}\mathbf{V}^H\mathbf{\Delta}^H)$ subject to the power constraints. In sum, the first phase in Table III is in essence the BSUM algorithm [38] applied to the optimization problem (29). Based on the convergence properties of the BSUM algorithm [38], it follows that repeated iteration of steps 2-10 can reach a stationary solution of problem (29). ■

REFERENCES

[1] 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.

[2] S. K. Yong and C.-C. Chong, "An overview of multigigabit wireless through millimeter wave technology: Potentials and technical challenges," *EURASIP J. Wireless Commun. Netw.*, vol. 2007, no. 1, p. 078907, 2006.

[3] 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.

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

[5] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742–758, Oct. 2014.

[6] F. Rusek *et al.*, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[7] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: Potentials and challenges," *Proc. IEEE*, vol. 102, no. 3, pp. 366–385, Mar. 2014.

[8] M. R. Akdeniz *et al.*, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1164–1179, Jun. 2014.

[9] V. Venkateswaran and A.-J. van der Veen, "Analog beamforming in MIMO communications with phase shift networks and online channel estimation," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4131–4143, Aug. 2010.

[10] O. El Ayach, R. W. Heath, Jr., S. Abu-Surra, S. Rajagopal, and Z. Pi, "Low complexity precoding for large millimeter wave MIMO systems," in *Proc. ICC*, Jun. 2012, pp. 3724–3729.

[11] S. Hur, T. Kim, D. J. Love, J. V. Krogmeier, T. A. Thomas, and A. Ghosh, "Millimeter wave beamforming for wireless backhaul and access in small cell networks," *IEEE Trans. Commun.*, vol. 61, no. 10, pp. 4391–4403, Oct. 2013.

[12] 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.

[13] A. F. Molisch, M. Z. Win, Y.-S. Choi, and J. H. Winters, "Capacity of MIMO systems with antenna selection," *IEEE Trans. Wireless Commun.*, vol. 4, no. 4, pp. 1759–1772, Jul. 2005.

[14] 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.

[15] R. Méndez-Rial, C. Rusu, N. González-Prelcic, A. Alkhateeb, and R. W. Heath, Jr., "Hybrid MIMO architectures for millimeter wave communications: Phase shifters or switches?" *IEEE Access*, vol. 4, pp. 247–267, Jan. 2016.

[16] X. Yu, J.-C. Shen, J. Zhang, and K. B. Letaief, "Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 485–500, Apr. 2016.

[17] A. F. Molisch and M. Z. Win, "MIMO systems with antenna selection," *IEEE Microw. Mag.*, vol. 5, no. 1, pp. 46–56, Mar. 2004.

[18] S. Sanayei and A. Nosratinia, "Antenna selection in MIMO systems," *IEEE Commun. Mag.*, vol. 42, no. 10, pp. 68–73, Oct. 2004.

[19] A. Gorokhov, D. A. Gore, and A. J. Paulraj, "Receive antenna selection for MIMO spatial multiplexing: Theory and algorithms," *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2796–2807, Nov. 2003.

[20] Q. Ma and C. Tepedelenlioglu, "Antenna selection for unitary space-time modulation," *IEEE Trans. Inf. Theory*, vol. 51, no. 10, pp. 3620–3631, Oct. 2005.

[21] A. Dua, K. Medepalli, and A. Paulraj, "Receive antenna selection in MIMO systems using convex optimization," *IEEE Trans. Wireless Commun.*, vol. 5, no. 9, pp. 2353–2357, Sep. 2006.

[22] S. Mahboob, R. Ruby, and V. C. M. Leung, "Transmit antenna selection for downlink transmission in a massively distributed antenna system using convex optimization," in *Proc. Conf. Broadband, Wireless Comput., Commun. Appl.*, Nov. 2012, pp. 228–233.

[23] X. Gao, O. Edfors, J. Liu, and F. Tufvesson, "Antenna selection in measured massive MIMO channels using convex optimization," in *Proc. Globecom Workshops*, Dec. 2013, pp. 129–134.

[24] C. Rusu, N. González-Prelcic, and R. W. Heath, Jr., "Array thinning for antenna selection in millimeter wave MIMO systems," in *Proc. ICASSP*, Mar. 2016, pp. 3416–3420.

[25] J. Park, J. Chun, and H. Park, "Generalised singular value decompositionbased algorithm for multi-user multiple-input multiple-output linear precoding and antenna selection," *IET Commun.*, vol. 4, no. 16, pp. 1899–1907, Nov. 2010.

[26] S.-C. Huang, W.-H. Fang, H.-S. Chen, and Y.-T. Chen, "Hybrid genetic algorithm for joint precoding and transmit antenna selection in multiuser MIMO systems with limited feedback," in *Proc. VTC-Spring*, May 2010, pp. 1–5.

[27] J. H. Lee and D. Park, "Antenna selection and unitary precoding for interference alignment with ML receiver," *IEEE Commun. Lett.*, vol. 16, no. 8, pp. 1216–1219, Aug. 2012.

[28] Z.-Q. Luo and S. Zhang, "Dynamic spectrum management: Complexity and duality," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 57–73, Feb. 2008.

[29] A. L. Yuille and A. Rangarajan, "The concave-convex procedure," *Neural Comput.*, vol. 15, pp. 915–936, Apr. 2003.

[30] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4331–4340, Sep. 2011.

[31] P. F. M. Smulders and L. M. Correia, "Characterisation of propagation in 60 GHz radio channels," *Electron. Commun. Eng. J.*, vol. 9, no. 2, pp. 73–80, Apr. 1997.

[32] H. Xu, V. Kukshya, and T. S. Rappaport, "Spatial and temporal characteristics of 60-GHz indoor channels," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 3, pp. 620–630, Apr. 2002.

[33] (Sep. 2011). *IEEE 802.15 WPAN Millimeter Wave Alternative PHY Task Group 3c*. [Online]. Available: http://www.ieee802.org/15/pub/TG3c.html

[34] *Linear Programming Relaxation*, accessed on Jan. 2014. [Online]. Available: https://en.wikipedia.org/wiki/Linear_programming_relaxation

[35] K. B. Petersern and M. S. Pedersern. (Nov. 2008). *The Matrix Cookbook*. [Online]. Available: http://matrixcookbook.com

[36] D. P. Bertsekas, *Nonlinear Programming*, 3nd ed. Belmont, MA, USA: Athena Scientific, 2016.

[37] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*. Princeton, NJ, USA: Princeton Univ. Press, 2009.

[38] M. Razaviyayn, M. Hong, and Z.-Q. Luo, "A unified convergence analysis of block successive minimization methods for nonsmooth optimization," *SIAM J. Optim.*, vol. 23, no. 2, pp. 1126–1153, 2013.

[39] K. Okada *et al.*, "Full four-channel 6.3-Gb/s 60-GHz CMOS transceiver with low-power analog and digital baseband circuitry," *IEEE J. Solid-State Circuits*, vol. 48, no. 1, pp. 46–64, Jan. 2013.

[40] W. Shin, B.-H. Ku, O. Inac, Y.-C. Ou, and G. M. Rebeiz, "A 108–114 GHz 4×4 wafer-scale phased array transmitter with high-efficiency on-chip antennas," *IEEE J. Solid-State Circuits*, vol. 48, no. 9, pp. 2041–2055, Sep. 2013.

[41] N. Deferm and P. Reynaert, *CMOS Front Ends for Millimeter Wave Wireless Communication Systems*. New York, NY, USA: Springer, 2015.

[42] L. A. Belov, S. M. Smolskiy, and V. N. Knochemasov, *Handbook of RF, Microwave, and Millimeter-Wave Components*. Norwood, MA, USA: Artech House, 2012.

[43] Y. Ma and Y. Fu, *Manifold Learning Theory and Applications*. Boca Raton, FL, USA: CRC Press, 2012.

[44] J. M. Lee, *Introduction to Smooth Manifolds*. New York, NY, USA: Springer, 2012.

[45] Q. Shi, M. Hong, X. Gao, E. Song, Y. Cai, and W. Xu, "Joint source-relay design for full-duplex MIMO AF relay systems," *IEEE Trans. Signal Process.*, vol. 64, no. 23, pp. 6118–6131, Dec. 2016.

Xiongfei Zhai received the B.S. degree in information and communication engineering from Zhejiang University, Hangzhou, China, in 2009, where he is currently pursuing the Ph.D. degree with the College of Information Science and Electronic Engineering. His research interests include massive MIMO communications, millimeter-wave communications, and application of optimization algorithm in communications.

Yunlong Cai (S'07–M'10–SM'16) received the B.S. degree in computer science from Beijing Jiaotong University, Beijing, China, in 2004, the M.Sc. degree in electronic engineering from the University of Surrey, Guildford, U.K., in 2006, and the Ph.D. degree in electronic engineering from the University of York, York, U.K., in 2010. From 2010 to 2011, he was a Post-Doctoral Fellow with the Electronics and Communications Laboratory of the Conservatoire National des Arts et Metiers, Paris, France. From 2016 to 2017, he was a Visiting Scholar with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, USA. Since 2011, he has been with the College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou, China, where he is currently an Associate Professor. His research interests include transceiver design for multiple-antenna systems, sensor array processing, adaptive filtering, full-duplex communications, cooperative and relay communications, and wireless information and energy transfer.

Qingjiang Shi received the Ph.D. degree in communication engineering from Shanghai Jiao Tong University, Shanghai, China, in 2011. From 2009 to 2010, he visited Prof. Z.-Q. (Tom) Luo's Research Group, University of Minnesota, Twin Cities. In 2011, he was a Research Scientist with the Research and Innovation Center (Bell Labs China), Alcatel-Lucent, Shanghai. He is currently a Professor with the College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, China. His current research interests lie in algorithm design for signal processing in advanced MIMO, cooperative communication, physical layer security, energy-efficient communication, wireless information, and power transfer.

He received the National Excellent Doctoral Dissertation Nomination Award in 2013, the Shanghai Excellent Doctoral Dissertation Award in 2012, and the Best Paper Award from the IEEE PIMRC'09 conference.

Minjian Zhao (M'10) received the M.Sc. and Ph.D. degrees in communication and information systems from Zhejiang University, Hangzhou, China, in 2000 and 2003, respectively. He is currently a Professor with the Department of Information Science and Electronic Engineering, Zhejiang University. His research interests include modulation theory, channel estimation and equalization, and signal processing for wireless communications.

Geoffrey Ye Li (S'93–M'95–SM'97–F'06) received the B.S.E. and M.S.E. degrees from the Department of Wireless Engineering, Nanjing Institute of Technology, Nanjing, China, in 1983 and 1986, respectively, and the Ph.D. degree from the Department of Electrical Engineering, Auburn University, Alabama, in 1994. He was a Teaching Assistant and then a Lecturer with Southeast University, Nanjing, from 1986 to 1991, a Research and Teaching Assistant with Auburn University, Alabama, from 1991 to 1994, and a Post-Doctoral Research Associate with the University of Maryland at College Park, Maryland, from 1994 to 1996. He was with AT&T Labs—Research at Red Bank, New Jersey, as a Senior Technical Staff Member and then a Principal Technical Staff Member from 1996 to 2000. Since 2000, he has been with the School of Electrical and Computer Engineering, Georgia Institute of Technology as an Associate Professor and then a Full Professor. He has been holding a Cheung Kong Scholar title at the University of Electronic Science and Technology of China since 2006. His general research interests include statistical signal processing and communications, with emphasis on cross-layer optimization for spectral- and energy-efficient networks, cognitive radios and opportunistic spectrum access, and practical issues in LTE systems. In these areas, he has published around 200 journal papers in addition to around 40 granted patents and numerous conference papers. His publications have been cited around 28 000 times and he has been recognized as the World's Most Influential Scientific Mind, also known as a Highly Cited Researcher, by Thomson Reuters. He was awarded the IEEE Fellow for his contributions to signal processing for wireless communications in 2005. He received the 2010 Stephen O. Rice Prize Paper Award, the 2013 WTC Wireless Recognition Award, and the 2017 Award for Advances in Communication from the IEEE Communications Society, and the 2013 James Evans Avant Garde Award and the 2014 Jack Neubauer Memorial Award from the IEEE Vehicular Technology Society. He also received the 2015 Distinguished Faculty Achievement Award from the School of Electrical and Computer Engineering, Georgia Tech. He has been involved in editorial activities for over 20 technical journals for the IEEE, including the Founding Editor-in-Chief of the IEEE 5G Tech Focus. He has organized and chaired many international conferences, including a Technical Program Vice Chair of the IEEE ICC'03, a Technical Program Co-Chair of the IEEE SPAWC'11, a General Chair of the IEEE GlobalSIP'14, and a Technical Program Co-Chair of the IEEE VTC'16 (Spring).

Benoit Champagne (S'87–M'89–SM'03) received the B.Ing. degree in engineering physics from the École Polytechnique de Montréal in 1983, the M.Sc. degree in physics from the Université de Montréal in 1985, and the Ph.D. degree in electrical engineering from the University of Toronto in 1990. From 1990 to 1999, he was an Assistant and then an Associate Professor with the INRS-Telecommunications, Université du Quebec, Montréal. In 1999, he joined McGill University, Montreal, where he is currently a Full Professor with the Department of Electrical and Computer Engineering; he also served as an Associate Chairman of Graduate Studies with the Department of Electrical and Computer Engineering from 2004 to 2007. His research focuses on the study of advanced algorithms for the processing of communication signals by digital means. His interests span many areas of statistical signal processing, including detection and estimation, sensor array processing, adaptive filtering, and applications thereof to broadband communications and audio processing, where he has co-authored nearly 250 referred publications. His research has been funded by the Natural Sciences and Engineering Research Council of Canada, the "Fonds de Recherche sur la Nature et les Technologies" from the Government of Quebec, and some major industrial sponsors, including Nortel Networks, Bell Canada, InterDigital, and Microsemi.

He was an Associate Editor of the *EURASIP Journal on Applied Signal Processing* from 2005 to 2007, the IEEE SIGNAL PROCESSING LETTERS from 2006 to 2008, and the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 2010 to 2012, and a Guest Editor for two special issues of the *EURASIP Journal on Applied Signal Processing* published in 2007 and 2014, respectively. He has also served on the Technical Committees of several international conferences in the fields of communications and signal processing. In particular, he was the Registration Chair of the IEEE ICASSP 2004, a Co-Chair of the Antenna and Propagation Track and the IEEE VTC-Fall 2004, a Co-Chair of Wide Area Cellular Communications Track and the IEEE PIMRC 2011, a Co-Chair of Workshop on D2D Communications and the IEEE ICC 2015, and the Publicity Chair of the IEEE VTC-Fall 2016.