Joint Antenna Selection and Transceiver Design for MU-MIMO mmWave Systems

Xiongfei Zhai[1], Yunlong Cai[1], Qingjiang Shi[2], Minjian Zhao[1], Geoffrey Ye Li[3] and Benoit Champagne[4]
[1]College of ISEE, Zhejiang University, Hangzhou 310027, China
[2]School of Info. Sci & Tech., Zhejiang Sci-Tech University, Hangzhou 310018, China
[3]School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, USA
[4]Department of Electrical and Computer Engineering, McGill University, Montreal, QC, Canada, H3A 0E9
{feifei4006, ylcai, mjzhao}@zju.edu.cn, qing.j.shi@gmail.com, liye@ece.gatech.edu, benoit.champagne@mcgill.ca

***Abstract*—This paper considers the uplink of large-scale multiple-user multiple-input multiple-output (MU-MIMO) millimeter wave (mmWave) systems, where a number of mobile stations (MSs) communicate with a single base station (BS) equipped with a large-scale antenna array, for application to fifth generation (5G) 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 (LNA). 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. 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 (AO) method. Specifically, the antenna selection matrix is optimized via the concave-convex procedure (CCCP); the weighted mean-square error minimization (WMMSE) approach is used to find the solution for the transmit beamformer; and the hybrid receiver is obtained via manifold optimization (MO). The convergence of the proposed algorithm is analysed 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 (with operating frequency in the range 30-300GHz) combined with large-scale multiple-input multiple-output (MIMO) techniques is now considered as a key enabling technology for the fifth generation (5G) wireless systems and beyond, which have drawn a great deal of attention within the research community recently [1], [2].

Nevertheless, mmWave-based based large-scale MIMO will increase fabrication cost and power consumption of the radio frequency (RF) chain as well as of the analog-to-digital (A/D) converters. To address these issues, the use of hybrid transceiver structure consisting of baseband digital and RF analog units has been considered as an attractive solution. Two basic kinds of elements have been considered for hybrid transceiver implementation, namely: analog phase shifters [3], [4] and analog switches [5], [6]. A typical algorithm to design hybrid transceivers is based on orthogonal matching pursuit (OMP) [4], where the columns of the analog transceiver matrices can be selected from certain candidate vectors. To solve for the analog transceiver under the constant modulus constraint, an alternating algorithm based on manifold optimization (MO) has been proposed in [7].

In large-scale MIMO mmWave systems with hybrid transceiver, the numbers of analog phase shifters and low-noise amplifiers (LNA) can be quite large and their energy consumption is considerable. An efficient method to reduce the system energy consumption and the implementation cost is through antenna selection. Optimal antenna selection requires an exhaustive search, whose complexity grows exponentially with number of the antennas. Reducing the computational complexity of antenna selection in large-scale MIMO systems is therefore of great practical and theoretical interest, as further discussed in [8], [9]. From the optimization perspective, linear relaxation has been employed in [10] to simplify the convex optimization problem into a form that is solvable in polynomial time.

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. 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) [11]. Then by fixing the hybrid receiver and the receive antenna selection matrix, we use the weighted mean-square minimization (WMMSE) approach [12] to optimize the transmit beamforming matrices. Finally, MO [7] is adopted to update the hybrid receiver. The convergence and sum-rate performance of the proposed algorithm is also investigated.

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. 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 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 a matrix, respectively. 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 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.

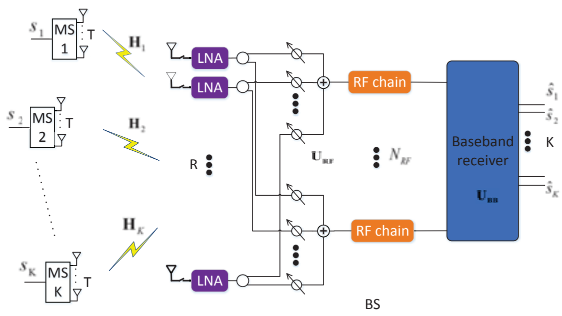

Fig. 1. A mmWave uplink system.

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 $R \gg T$. Hence, a fully digital transmit beamformer is assumed at the MS, while a hybrid structure with N_{RF} 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 with $E(|s_k|^2) = \frac{1}{K}$ and $\boldsymbol{v}_k \in \mathbb{C}^{T\times 1}$ denotes the corresponding transmit beamforming vector. A constraint on the total transmit power for each MS is enforced by setting $\mathrm{Tr}(\boldsymbol{v}_k \boldsymbol{v}_k^H) \le 1$, 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 $\mathbf{H}_k \in \mathbb{C}^{T\times R}$ 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

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

then Eq. (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 $K_s \ll R$. This is conveniently expressed by means of a diagonal antenna selection matrix $\boldsymbol{\Delta} \in \mathbb{C}^{R\times R}$, with entries $[\boldsymbol{\Delta}]_{ii} = 1$ if the i^{th} antenna is selected and $[\boldsymbol{\Delta}]_{ii} = 0$ otherwise, where $\sum_{i=1}^{R} [\boldsymbol{\Delta}]_{ii} = K_s$. Then the corresponding received signal vector after antenna selection will be

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

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

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

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

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

With the above discussed prossing, the sum-rate of the system in Fig. 1 will be expressed in (6) [4].

B. Problem formulation

In this work, we aim to select the receive antennas at the BS, $\boldsymbol{\Delta}$, 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

$$\begin{aligned} &\max_{\boldsymbol{\Delta}, \boldsymbol{v}_k, \mathbf{U}_{RF}, \mathbf{U}_{BB}} \mathcal{R} \\ &\text{s.t. } \mathrm{Tr}(\boldsymbol{v}_k \boldsymbol{v}_k^H) \le 1, k \in \mathcal{K}, \\ &\quad |[\mathbf{U}_{RF}]_{ij}| = 1, 1 \le i \le R, 1 \le j \le N_{RF}, \\ &\quad [\boldsymbol{\Delta}]_{ii} \in \{0, 1\}, i = 1, 2, \ldots, R, \\ &\quad \sum_{i=1}^{R} [\boldsymbol{\Delta}]_{ii} = K_s. \end{aligned} \tag{7}$$

Note that problem (7) 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 $\boldsymbol{\Delta}$ unless an exhaustive search is performed. To address the optimization problem in (7), 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 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 via the CCCP-based method while fixing the transmit beamformers and hybrid receiver. In this case, we reformulate problem (7) into a more tractable form with respect to the antenna selection matrix, $\boldsymbol{\Delta}$.

If the hybrid receiver is designed based on the MMSE criterion, the original problem in (7) can be decomposed into

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

the following approximate problem:

$$\begin{aligned} &\max_{\mathbf{\Delta}} \log\det(\mathbf{I} + \frac{\rho}{K\sigma^2}\mathbf{\Delta}\mathbf{V}\mathbf{V}^H\mathbf{\Delta}^H) \\ &\text{s.t. } [\mathbf{\Delta}]_{ii} \in \{0,1\}, i = 1,2,\ldots,R, \\ &\quad \sum_{i=1}^{R}[\mathbf{\Delta}]_{ii} \;=\; K_s. \end{aligned} \tag{8}$$

Since the antenna selection matrix is diagonal and furthermore with only 0 or 1 on the diagonal,

$$\mathbf{\Delta}^H\mathbf{\Delta} \;=\; \mathbf{\Delta}. \tag{9}$$

Based on (9), the Sylvester's determinant theorem in [13], and linear relaxation, problem (8) can be rewritten as

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

If, the popular software packages CVX [14] can be used to solve the above optimization problem, the solution of $[\mathbf{\Delta}]_{ii}$ may be fractional, rather than binary as required. In order to obtain the final antenna selection matrix, we must round the fractional elements to the nearest integers, 0 or 1, which will incur performance degradation.

In order to address the issue, we can exploit the special constraint structure in problem (10) and use the CCCP-based method to directly obtain a binary solution for this problem. According to Theorem 1 in [11], we can reformulate problem (10) as

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

where the auxiliary variable β is positive. Considering the constraint $\sum_{i=1}^{R}[\mathbf{\Delta}]_{ii} = K_s$, if $[\mathbf{\Delta}]_{ii}$ is 0 or 1, then $\beta(\sum_{i=1}^{R}[\mathbf{\Delta}]_{ii}^2 - K_s) = 0$, which means that the objective function value of (11) is the same as that of (10). The following theorem can ensure Problem (11) to have a binary solution.

***Theorem** 1:* Problem (11) admits a binary solution for $\mathbf{\Delta}$ when $\beta = b/a$, where a is the second-order derivative of $\sum_{i=1}^{R}[\mathbf{\Delta}]_{ii}^2 - K_s$ and the absolute value of the second-order derivative of $\log\det(\mathbf{I} + \frac{\rho}{K\sigma^2}\mathbf{V}^H\mathbf{\Delta}\mathbf{V})$ is upper bounded with b.

Proof: Considering the idea behind CCCP, since $\sum_{i=1}^{R}[\mathbf{\Delta}]_{ii}^2 - K_s$ is strong convex with second-order derivative, we can choose β according to Theorem 1 such that the objective function in (11) is convex. Then problem (11) is to maximize a convex function subject to the linear constraints. Therefore the optimal solution for $[\mathbf{\Delta}]_{ii}$ must be the bound of the interval, i.e. 0 or 1. ■

We can obtain the solution of problem (11) using the Frank-Wolfe method [15]. Denote

$$f(\mathbf{\Delta}) \triangleq \log\det(\mathbf{I} + \frac{\rho}{K\sigma^2}\mathbf{V}^H\mathbf{\Delta}\mathbf{V}) + \beta(\sum_{i=1}^{R}[\mathbf{\Delta}]_{ii}^2 - K_s), \tag{12}$$

and

$$\mathbf{S} \triangleq \left\{\mathbf{\Delta} \mid 0 \le [\mathbf{\Delta}]_{ii} \le 1, \forall i, \sum_{i=1}^{R}[\mathbf{\Delta}]_{ii} = K_s\right\}. \tag{13}$$

Given the initial point $\mathbf{\Delta}_{ini}$, we solve the following problem

$$\mathbf{\Delta}_{CCCP} = \arg\max_{\mathbf{\Delta}\in\mathbf{S}} \;\; \sum_{i=1}^{R} g_i[\mathbf{\Delta} - \mathbf{\Delta}_{ini}]_{ii} \tag{14}$$

where $g_i \triangleq \left.\frac{\partial f(\mathbf{\Delta})}{\partial [\mathbf{\Delta}]_{ii}}\right|_{\mathbf{\Delta}=\mathbf{\Delta}_{ini}}$.

For the indexes of the K_s greatest g_i, we set the corresponding entries in $\mathbf{\Delta}$ to be 1 while the others are 0. Since problem (14) is a special linear programming and admits a binary closed-form solution, the CCCP-based method applying to (11) can be efficiently implemented.

IV. Optimizing Beamformer

In this section, we optimize the transmit beamformers for given the antenna selection matrix and the hybrid receiver. As mentioned earlier, it is pretty hard to optimize beamformer. Here, we use the WMMSE approach to transform the sum-rate maximization problem to the weighted MSE minimization problem, which is easier to solve.

To optimize beamformer, we need to solve the following optimization problem,

$$\begin{aligned} &\max_{\boldsymbol{v}_k} \;\; \log\det(\mathbf{I} + \frac{\rho}{K\sigma^2}\mathbf{\Delta}\mathbf{V}\mathbf{V}^H\mathbf{\Delta}^H) \\ &\text{s.t. } \mathrm{Tr}(\boldsymbol{v}_k\boldsymbol{v}_k^H) \le 1, k \in \mathcal{K}. \end{aligned} \tag{15}$$

The WMMSE approach applies to the above problem and it can solve the problem to a stationary solution [12]. However, we do not need to solve (15) to a stationary point. Instead, to reduce the computational complexity, we perform only non-iterative WMMSE algorithm to get better beamformers, which is described as follows.

We first reformulate the sum-rate maximization problem to an equivalent problem involving weighted MSE minimization.

The MSE matrix of the received signal vector at the BS can be expressed as

$$\begin{aligned} &\mathbf{E}(\mathbf{U},\mathbf{V}) \triangleq E[(\tilde{\boldsymbol{s}} - \boldsymbol{s})(\tilde{\boldsymbol{s}} - \boldsymbol{s})^H] \\ &= \frac{1}{K}(\mathbf{I} - \sqrt{\rho}\mathbf{U}^H\mathbf{\Delta}\mathbf{V})(\mathbf{I} - \sqrt{\rho}\mathbf{U}^H\mathbf{\Delta}\mathbf{V})^H + \sigma^2\mathbf{U}^H\mathbf{U}, \end{aligned} \tag{16}$$

where $\tilde{\boldsymbol{s}}$ is given by $\tilde{\boldsymbol{s}} = \mathbf{U}^H\boldsymbol{y}_{\mathbf{\Delta}}$ and $\mathbf{U}$ is an adaptive fully digital receiver. Note that $\mathbf{U}$ is only an intermediate variable in the update of beamformers and it will be decomposed into

$\mathbf{U}_{RF}$ and $\mathbf{U}_{BB}$ later in Section V. Then the corresponding MSE minimization problem can be written as

$$\begin{aligned} &\min_{\boldsymbol{v}_k,\mathbf{U}} \ \mathrm{Tr}(\mathbf{E}(\mathbf{U},\mathbf{V})) \\ &\text{s.t. } \mathrm{Tr}(\boldsymbol{v}_k\boldsymbol{v}_k^H) \le 1, k \in \mathcal{K}. \end{aligned} \tag{17}$$

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

$$\mathbf{U}^{mmse} = \frac{\sqrt{\rho}}{K}(\frac{\rho}{K}\mathbf{\Delta}\mathbf{V}\mathbf{V}^H\mathbf{\Delta}^H + \sigma^2\mathbf{I})^{-1}\mathbf{\Delta}\mathbf{V}. \tag{18}$$

Plugging (18) into (16), the corresponding MSE matrix reduces to

$$\mathbf{E}^{mmse} = \frac{1}{K}(\mathbf{I} - \frac{\rho}{K}\mathbf{V}^H\mathbf{\Delta}^H(\frac{\rho}{K}\mathbf{\Delta}\mathbf{V}\mathbf{V}^H\mathbf{\Delta}^H + \sigma^2\mathbf{I})^{-1}\mathbf{\Delta}\mathbf{V}). \tag{19}$$

Directly following the proof in [12], we have the following theorem that establishes the equivalence between the sum-rate maximization problem and the weighted MSE minimization problem.

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

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

is equivalent to problem (15), in the sense that they share the same KKT solution set.

It can be observed that the cost function in (20) is convex in each block of the optimization variables, $\mathbf{U}$, $\mathbf{W}$, and $\boldsymbol{v}_k$. Moreover, the constraints are separable across the block variables. Therefore, the block coordinate descent (BCD) method in [15] applies to problem (20). Specifically, we minimize the weighted-MSE cost function with respect to one block variable while fixing the other blocks, consisting of the following three steps.

i) Update $\mathbf{U}$ while fixing $\boldsymbol{v}_k$: this step yields the MMSE receiver given by (18).

ii) Update $\mathbf{W}$ while fixing $\mathbf{U}$ as $\mathbf{U}^{mmse}$: this step is to update $\mathbf{W}$, which admits a closed-form solution as follows

$$\mathbf{W}^{opt} = \mathbf{E}^{mmse-1}. \tag{21}$$

iii) Update $\boldsymbol{v}_k$ while fixing $\mathbf{W}$ and $\mathbf{U}$: after obtaining $\mathbf{U}$ and $\mathbf{W}$ in the previous two steps, we update $\boldsymbol{v}_k$. That is, we solve the following problem in the order of $k = 1, 2, \ldots, K$,

$$\begin{aligned} &\min_{\boldsymbol{v}_k} \ \mathrm{Tr}[\frac{1}{K}\mathbf{W}(\mathbf{I} - \sqrt{\rho}\mathbf{U}^H\mathbf{\Delta}\mathbf{V})(\mathbf{I} - \sqrt{\rho}\mathbf{U}^H\mathbf{\Delta}\mathbf{V})^H] \\ &\text{s.t. } \mathrm{Tr}(\boldsymbol{v}_k\boldsymbol{v}_k^H) \le 1, k \in \mathcal{K}. \end{aligned} \tag{22}$$

The above problem is a convex quadratic optimization problem with K separable constraints. Thus, Lagrange multiplier can solve it.

V. Optimizing Hybrid receiver

Inspired by [7], we use MO to search the solution for $\mathbf{U}_{BB}$ and $\mathbf{U}_{RF}$ by fixing $\mathbf{\Delta}$ and $\boldsymbol{v}_k$ in this section, which is different from the Frobenius norm based approach in [7]. The original sum-rate maximization problem is considered instead.

A. Baseband receiver design

First, we design the baseband receiver $\mathbf{U}_{BB}$ while fixing analog receiver $\mathbf{U}_{RF}$. Since we have got the fully digital MMSE receiver $\mathbf{U}^{mmse}$, what we need is to decompose $\mathbf{U}^{mmse}$ into $\mathbf{U}_{RF}$ and $\mathbf{U}_{BB}$. As a result, we can obtain a well-known least squares solution given by

$$\mathbf{U}_{BB} = \mathbf{U}_{RF}^{\dagger}\mathbf{U}^{mmse}. \tag{23}$$

Obviously, it 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 $\mathbf{\Delta}$, $\boldsymbol{v}_k$ and $\mathbf{U}_{BB}$, i.e., update $\mathbf{U}_{RF}$ by finding the solution of the following optimization problem

$$\begin{aligned} &\max_{\mathbf{U}_{RF}} \ \mathcal{R} \\ &\text{s.t. } |[\mathbf{U}_{RF}]_{ij}| = 1, 1 \le i \le R, 1 \le j \le N_{RF}. \end{aligned} \tag{24}$$

This optimization problem is complicated 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; 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 a well developed conjugate gradient algorithm in Euclidean spaces to find its counterpart on the Riemannian manifolds. In [7], the hybrid receiver is designed by minimizing its Frobenius distance to the optimal fully digital receiver. By exploiting the idea of MO in this paper, we directly solve the sum-rate maximization problem (24), which results in better performance.

Based on Riemannian geometry, we can develop an iterative algorithm to optimize analog receiver, which is summarized in Table I. In Table I, the gradient, retraction, and vector transport are defined as

$$\begin{aligned} &\mathrm{grad}\mathcal{R}(\boldsymbol{x}) \triangleq \nabla\mathcal{R}(\boldsymbol{x}) - \mathbb{R}\{\nabla\mathcal{R}(\boldsymbol{x}) \circ \boldsymbol{x}^*\} \circ \boldsymbol{x}^*, \\ &\mathrm{Retr}_{\mathbf{x}}(\alpha\mathbf{d}) \triangleq vec\left[\frac{[\boldsymbol{x}+\alpha\mathbf{d}]_i}{|[\boldsymbol{x}+\alpha\mathbf{d}]_i|}\right], \\ &\mathrm{Transp}_{\boldsymbol{x}_r \to \boldsymbol{x}_{r+1}}(\mathbf{d}) \triangleq \mathbf{d} - \mathbb{R}(\mathbf{d} \circ \boldsymbol{x}_{r+1}^*) \circ \boldsymbol{x}_{r+1}. \end{aligned}$$

From Table I, the well-known Armijo backtracking line search step and Polak-Ribiere parameter ensure the objective function to be non-decreasing in each iteration [15]. According to [7], MO for problem (24) converges to a stationary point, i.e., the point where the gradient of the objective function is zero.

VI. Analysis of the proposed algorithm

In this section, we first summarize the proposed joint antenna selection and transceiver design algorithm in Table II, which is referred to as CCCP-WMMSE-MO. Then the computational complexity and convergence of our algorithm are analyzed.

TABLE I
PSEUDO CODE OF MO ALGORITHM FOR $\mathbf{U}_{RF}$

Initialize $\mathbf{U}_{RF,0}$ such that $\boldsymbol{x}_0 = vec(\mathbf{U}_{RF,0}) \in \mathcal{M}^m$;
$\mathbf{d}_0 = \text{grad}\mathcal{R}(\boldsymbol{x}_0)$ and $r = 0$;
repeat
  Choose Armijo backtracking line search step size α_r;
  Find the next point $\boldsymbol{x}_{r+1}$ and the corresponding $\mathbf{U}_{RF,r+1}$using retraction ;
  Determine Riemannian gradient $\mathbf{g}_{r+1} = -\text{grad}\mathcal{R}(\boldsymbol{x}_{r+1})$;
  Calculate the vector transports $\mathbf{g}_r^+$ and $\mathbf{d}_r^+$ of gradient $\mathbf{g}_r$ and conjugate direction $\mathbf{d}_r$ from $\boldsymbol{x}_r$ to $\boldsymbol{x}_{r+1}$;
  Choose Polak-Ribiere parameter γ_{r+1};
  Compute conjugate direction $\mathbf{d}_{r+1} = -\mathbf{g}_{r+1} + \gamma_{r+1}\mathbf{d}_r^+$;
  $r \leftarrow r + 1$;
until a stopping criterion triggers.

TABLE II
PSEUDO CODE OF THE CCCP-WMMSE-MO ALGORITHM

1 Initialize $\mathbf{\Delta}$, $\boldsymbol{v}_k$, $\mathbf{U}_{RF}$ and $\mathbf{U}_{BB}$, such that they meet all the constraints;
2 **repeat**
3   use CCCP-based method to find $\mathbf{\Delta}$;
4   use non-iterative WMMSE approach to update $\boldsymbol{v}_k$;
5 **until** a stopping criterion triggers;
6 use MO algorithm to find $\mathbf{U}_{BB}$ and $\mathbf{U}_{RF}$.

A. Complexity analysis

Considering the derivation of the proposed algorithm, since R is a large number, the dominated computational complexity of optimizing transmit beamformers in the transceiver design without antenna selection is given by $\mathcal{O}(R^3 + R^2K + RK^2 + RTK^2 + RT^2K + RK^3 + RTK)$. With antenna selection the computational complexity of optimizing beamformers reduces to $\mathcal{O}(R^3 + RKK_s + RK_s + RT^2K + RK^2 + RTK + RK^2K_s)$. Compared to the optimization of 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. Moreover, note that the computational complexity of the proposed antenna selection method is given by $\mathcal{O}(R^2K + RK^2)$, therefore the overall complexity of the proposed algorithm is similar to that of the transceiver design using all antennas. However, with the aid of the proposed transceiver algorithm with antenna selection, the energy consumption at the BS can be significantly reduced. We will show simulation results in Section VII to validate the proposed algorithm.

B. Convergence analysis

First, for the first loop in Table II (steps 2-5), the convergence is summarized in Theorem 3.

Theorem 3: Any limit point of the sequence generated by the first loop in Table II is a stationary point of (25), which is the corresponding problem with fixed hybrid receiver.

$$\begin{aligned} &\max_{\mathbf{\Delta}, \boldsymbol{v}_k} \log\det(\mathbf{I} + \frac{\rho}{K\sigma^2}\mathbf{\Delta}\mathbf{V}\mathbf{V}^H\mathbf{\Delta}^H) \\ &\text{s.t. } \operatorname{Tr}(\boldsymbol{v}_k\boldsymbol{v}_k^H) \le 1, k \in \mathcal{K}, \\ &\qquad \mathbf{\Delta} \in \mathbf{S}. \end{aligned} \tag{25}$$

Proof: Due to the space limitation, we only outline the main steps of the proof here. We first establish the equivalence among the objective functions in (8), (11), (15), (20) and (25) by using Theorem 1 and Theorem 2. Note that step 3 in Table II is equivalent to globally solving problem (8). We then demonstrate that the non-iterative WMMSE is equivalent to minimizing a locally tight upper bound of (15). Hence, the first loop is in essence the block successive upper-bound minimization (BSUM) algorithm [16] applied to the optimization problem (25). Finally, in terms of the convergence result of the BSUM algorithm [16], steps 2-5 can reach a stationary solution of problem (25) ■

For the MO algorithm, directly following the discussion in [7], we can conclude that:

Theorem 4: Any limit point of the sequence generated by the MO algorithm is a stationary point of (24).

From Theorem 3 and Theorem 4, we can obtain a high quality solution by applying the CCCP-WMMSE-MO algorithm to optimization problem (7).

VII. SIMULATION RESULTS

In this section, we evaluate the sum-rate performance of the proposed algorithm. Without loss of generality, the SNR is defined as $\frac{\rho}{\sigma^2}$. We use Saleh-Valenzuela model [17] for mmWave channel and the number of rays is set to 20. The antenna elements in the uniform linear array (ULA) are separated by a half wavelength and all simulation results are averaged over 500 channel realizations. In the simulation results, the term "CCCP-OMP-MO" resorts to the CCCP-based method and the algorithm in [7] jointly, while "FULL-WMMSE-MO" denotes the CCCP-WMMSE-MO algorithm with full antenna and "FULL-OMP-MO" is the CCCP-OMP-MO algorithm without antenna selection. Besides, "CCCP-WMMSE-FD" means that we replace the hybrid receiver with the fully digital MMSE receiver after the antenna selection matrix is obtained.

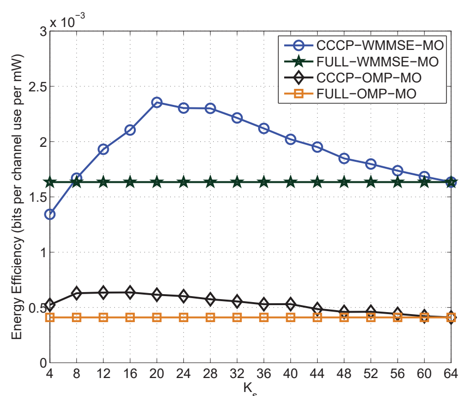

Fig. 2. The energy efficiency versus K_s when $R = 64$, $T = 2$, $K = 4$, $N_{RF} = 8$ and $SNR = 0dB$.

We consider the energy efficiency at the BS to compare the proposed CCCP-WMMSE-MO algorithm, the FULL-WMMSE-MO algorithm and the conventional CCCP-OMP-MO algorithm. 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, $P_{r,2} = N_{RF} \times \tilde{P}_{r,2}$ denotes the total energy consumption of RF chains, with $\tilde{P}_{r,2}$ being the energy consumption of each RF chain, and $P_{r,3} = K_s \times (\tilde{P}_{r,3} + N_{RF} \times \bar{P}_{r,3})$ denotes the total energy consumption of the analog phase shifters and LNAs, with $\tilde{P}_{r,3}$ and $\bar{P}_{r,3}$ being each LNA energy consumption and each phase shifter energy consumption, respectively. According to [18]–[21], 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. For the FULL-WMMSE-MO algorithm, we have $K_s = R$. Fig. 2 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 algorithm, the FULL-WMMSE-MO algorithm, and the FULL-OMP-MO algorithm when $K_s \geq 8$. If antenna selection strategy is not used at the BS, one can see that the CCCP-WMMSE-MO FULL-WMMSE-MO algorithm is still more energy efficient than the FULL-OMP-MO algorithm. Besides, the results show that with the increasing of K_s, the BS energy efficiency of the proposed algorithm increases first and then decreases, and the maximum value of the energy efficiency occurs when K_s is around 20.

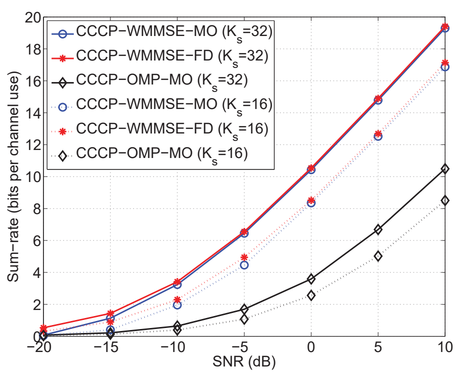

Fig. 3. The system sum-rate versus SNR for $R = 64$, $T = 2$, $K = 4$, $N_{RF} = 8$.

Fig. 3 compares the performance of different algorithms versus SNR for $K_s = 32$ and $K_s = 16$, respectively. 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. The proposed algorithm reaches the performance of the fully digital receiver. Under this system setting, the CCCP-WMMSE-MO algorithm serves as an excellent candidate for joint transceiver design with antenna selection, achieving both good performance and low complexity.

VIII. CONCLUSIONS

In this paper, we have investigated joint transceiver design with antenna selection at the BS to maximize system sum-rate for MU-MIMO mmWave uplink systems. The problem is separated into three subproblems and a novel alternating algorithm named the CCCP-WMMSE-MO algorithm is proposed. By exploiting the CCCP-based method, we have solved the antenna selection problem and directly obtained a binary antenna selection matrix. Then the WMMSE approach is used to optimize the transmit beamformers while the MO method is resorted to update the hybrid receiver. The computational complexity of the CCCP-WMMSE-MO algorithm has been analyzed. Besides, the convergence of the proposed algorithm has been analytically analysed. The numerical results have suggested that the proposed algorithm achieves outstanding performance, which is close to that of the algorithm using fully digital transceiver. In terms of energy efficiency at the BS, the CCCP-WMMSE-MO algorithm outperforms the FULL-WMMSE-MO algorithm.

IX. ACKNOWLEDGEMENT

The work of Q. Shi is supported by the National Nature Science Foundation of China under grant 61302076, 61374020. The work of Y. Cai and M. Zhao is supported in part by the National Natural Science Foundation of China under Grant 61471319, Zhejiang Provincial Natural Science Foundation of China under Grant LY14F010013, the Fundamental Research Funds for the Central Universities, the National Science Foundation of China under Grant 91538103 and the National High Technology Research and Development Program (863 Program) of China under Grant 2014AA01A707.

REFERENCES

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

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

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

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

[6] 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.

[7] X. Yu, J. 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.

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

[9] S. Sanayei and A. Nosratinia, "Antenna selection in MIMO systems," *IEEE Commun. Mag.,* vol. 42, pp. 68-73, Oct. 2004.

[10] 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, Sept. 2006.

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

[12] 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, Sept. 2011.

[13] K. B. Petersern and M. S. Pedersern, *The Matrix Cookbook*. http://matrixcookbook.com, Nov. 2008.

[14] CVX Research, Inc. CVX: Matlab software for disciplined convex programming, version 2.0 beta. http://cvxr.com/cvx, Sept. 2012.

[15] D. Bertsekas, *Nonlinear Programming*, 2nd ed. Belmont, MA: Athena Scientific, 1999.

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

[17] IEEE 802.15 WPAN Millimeter Wave Alternative PHY Task Group 3c. [online]. www. ieee802.org/15/pub/TG3c.html, Sept. 2011.

[18] 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.

[19] W. Shin, B.-H. Ku, O. Inac, Y.-C. Ou, and G. M. Rebeiz, "108-114 GHz 44 wafer-scale phased array transmitter with high-efficiency on-chip antennas," *IEEE J. Solid-State Circuits*, vol. 48, no.9 pp. 2041-2055, May 2013.

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

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