Joint Design of Beam Selection and Precoding for mmWave MU-MIMO Systems with Lens Antenna Array

Rongbin Guo*, Yunlong Cai*, Qingjiang Shi†, Minjian Zhao* and Benoit Champagne‡
*College of ISEE, Zhejiang University, Hangzhou, China, 310027
†School of Info. Sci & Tech., Zhejiang Sci-Tech University, Hangzhou, China, 310018
‡Department of Electrical and Computer Engineering, McGill University, Montreal, QC, Canada, H3A 0E9
Emails: {rbguoisee, ylcai, mjzhao}@zju.edu.cn, qing.j.shi@gmail.com and benoit.champagne@mcgill.ca

***Abstract*—Wireless transmission with lens antenna arrays is becoming more and more attractive for millimeter wave (mmWave) multiple-input multiple-output (MIMO) systems with limited radio frequency (RF) chains due to their energy-focusing capability. In this paper, we consider the joint design of beam selection and precoding to maximize the sum rate of a downlink single-sided lens MU-MIMO mmWave system under transmit power constraints. We first formulate the optimization problem into a tractable form using the popular weighted minimum mean squared error (WMMSE) approach. To solve this problem, we then propose an efficient joint beam selection and precoding algorithm based on the innovative penalty dual decomposition (PDD) method. Simulation results demonstrate that our proposed algorithm can achieve near-optimal performance when compared to the fully digital precoding scheme and thus outperform the competing methods.[1]**

***Index Terms*—mmWave MU-MIMO, lens antenna array, beam selection, WMMSE, PDD method.**

I. Introduction

Communications over millimeter wave (mmWave) frequencies will be a key feature of the fifth generation (5G) cellular networks by supporting unprecedented data rates in the 30-300GHz band [1], [2]. Indeed, the significantly reduced wavelength makes it possible to realize massive MIMO systems that use a large number of antennas within a small physical size, and yet achieve high array gain for directional communications by exploiting precoding techniques [3].

However, in the case of massive MIMO, the conventional fully digital precoding techniques lead to unaffordable costs in terms of radio frequency (RF) chains and power consumption. To deal with this limitation, a number of studies have proposed the concept of *beamspace MIMO* based on the discrete lens array (DLA) [4]. In effect, this approach can transform the traditional MIMO spatial channels into beamspace channels with angle-dependent energy-focusing capabilities [5]. In practice, only a small number of beams are needed due to the sparse nature of beamspace channels. Since each beam corresponds to a single RF chain, this effectively reduces the cost of RF chains in mmWave massive MIMO systems. Within this context, a critical problem with mmWave lens array systems is the design of the beam selection and digital precoding schemes.

[1]This work was supported by the Fundamental Research Funds for the Central Universities and the National Science Foundation of China (NSFC) under Grant 91538103

Currently, studies on the beam selection problem with DLA concentrate on choosing beams with maximum magnitude (expressed as "MM-BS" in the following) to obtain as much power from each user as possible [5]–[7]. Subsequently, [8] considers the potential multiuser interference and proposes an interference-aware beam selection (IA-BS) strategy which achieves better results than MM-BS schemes. However, all the aforementioned schemes are based on fixed digital precoding methods – such as zero forcing (ZF), maximum ratio combining (MRC), etc.– which might suffer from significant performance degradation since the beam selection and precoding are designed separately.

In this paper, we consider the joint design of beam selection and precoding to maximize the sum rate of a downlink single-sided lens mmWave MU-MIMO system with limited RF chains at the base station (BS). Firstly we formulate the optimization problem into a tractable form by using the weighted minimum mean squared error (WMMSE) approach [9]. Then in order to solve the resulting problem, we propose an efficient joint beam selection and precoding algorithm based on the innovative penalty dual decomposition (PDD) method [10]. Simulation results demonstrate that our proposed algorithm can achieve near-optimal performance when compared to the fully digital precoding scheme and thus outperform the competing methods.

Notation: In this paper, lower-case and upper-case boldface letters $\boldsymbol{a}$ and $\boldsymbol{A}$ denote a vector and a matrix, respectively. Lower-case letter a_{ij} means the (i,j)-th element of matrix $\boldsymbol{A}$ and $|a|$ denotes the amplitude of scalar a. $\boldsymbol{A}^H$, $\boldsymbol{A}^{-1}$, $\boldsymbol{A}^{\dagger}$, Tr$(\boldsymbol{A})$ and $||\boldsymbol{A}||_2$ denote the conjugate transpose, inverse, pseudo-inverse, trace and frobenius norm of matrix $\boldsymbol{A}$, respectively; $\boldsymbol{I}_K$ is the $K \times K$ identity matrix and $E\{\cdot\}$ represents the expectation operation.

II. System model and problem formulation

We consider a downlink single-sided lens mmWave MU-MIMO system, where the BS employs M_s antennas and N_{RF} RF chains to simultaneously serve K spatially distributed users, each equipped with single-antenna receiver. Note that the number of RF chains should satisfy $N_{RF} \geq K$ to guarantee spatial multiplexing gain for the K users. Without loss of generality, we choose $N_{RF} = K$ in this paper.

A. System Model

As shown in Fig 1, the BS transmits K data streams carrying independent messages, each one intended to

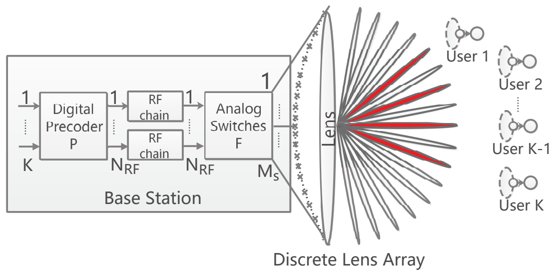

Fig. 1: A mmWave MU-MIMO system with single-sided lens antenna array.

a specific user. The precoded baseband data vector at the BS can be expressed as

$$\boldsymbol{x} = \boldsymbol{P}\boldsymbol{s} = \sum_{k=1}^{K} \boldsymbol{p}_k s_k, \tag{1}$$

where $\boldsymbol{s} = [s_1, s_2, \ldots, s_K]^T$, s_k is the data symbol intended for user k with zero mean and normalized power $\mathbb{E}\{|s_k|^2\} = 1$, and $\boldsymbol{P} = [\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_K] \in \mathbb{C}^{N_{RF} \times K}$ is the precoding matrix and $\boldsymbol{p}_k$ is the digital precoding vector for user k. The $K \times 1$ received signal vector $\boldsymbol{y}$ of all K users can be expressed as

$$\boldsymbol{y} = \boldsymbol{H}^H \boldsymbol{F}\boldsymbol{P}\boldsymbol{s} + \boldsymbol{n}, \tag{2}$$

in which $\boldsymbol{H} \in \mathbb{C}^{M_s \times K}$ is the beamspace channel matrix, $\boldsymbol{F} \in \mathbb{C}^{M_s \times N_{RF}}$ is the beam selection matrix whose entries are either 0 or 1, and $\boldsymbol{n} \sim \mathcal{CN}(0, \sigma^2 \boldsymbol{I}_K)$ is the $K \times 1$ additive white Gaussian noise (AWGN) vector. The beamspace channel is developed below.

B. Beamspace Channel Model

The beamspace channel matrix $\boldsymbol{H}$ is obtained from the spatial MIMO channels by Fourier transformation:

$$\boldsymbol{H} = [\boldsymbol{h}_1, \boldsymbol{h}_2, \cdots, \boldsymbol{h}_K] = [\boldsymbol{U}\boldsymbol{g}_1, \boldsymbol{U}\boldsymbol{g}_2, \ldots, \boldsymbol{U}\boldsymbol{g}_K], \tag{3}$$

where $\boldsymbol{U} \in \mathbb{C}^{M_s \times M_s}$ is a spatial discrete Fourier transformation (DFT) matrix corresponding to a carefully designed DLA [6] and $\boldsymbol{g}_k \in \mathbb{C}^{M_s \times 1}$ is the spatial channel for user k. The DFT matrix $\boldsymbol{U}$ consists of the array steering vectors of M_s orthogonal directions (beams) over the entire space, i.e.:

$$\boldsymbol{U} = [\boldsymbol{a}(\varphi_1), \boldsymbol{a}(\varphi_2), \ldots, \boldsymbol{a}(\varphi_{M_s})]^H, \tag{4}$$

in which $\varphi_n = \frac{1}{M_s}(n - \frac{M_s+1}{2})$ with $n = 1, 2, \ldots, M_s$ are the normalized spatial directions (beams) [8] and $\boldsymbol{a}(\varphi_n) = \frac{1}{\sqrt{M_s}}[e^{-j2\pi\varphi_n i}]_{i \in \mathcal{I}}$ is the $M_s \times 1$ array steering vector, where $\mathcal{I} = \{j - (M_s - 1)/2 | j = 0, 1, \cdots, M_s - 1\}$ is an index set of array elements. The spatial direction is defined as $\varphi \triangleq \frac{d}{\lambda} sin\theta$ where θ is the physical direction, λ is the signal wavelength, and $d = \lambda/2$ is the antenna spacing[2].

In this work, we employ the widely used Saleh-Valenzuela channel model for the considered mmWave system [5]–[8]:

$$\boldsymbol{g}_k = \beta_k^{(0)} \boldsymbol{a}(\phi_k^{(0)}) + \sum_{l=1}^{L} \beta_k^{(l)} \boldsymbol{a}(\phi_k^{(l)}), \tag{5}$$

where $\beta_k^{(0)} \boldsymbol{a}(\phi_k^{(0)})$ and $\beta_k^{(l)} \boldsymbol{a}(\phi_k^{(l)})$ represent the line-of-sight (LOS) and the l-th non-line-of-sight (NLOS) channel between the BS and user k, respectively. Meanwhile, β_k represent the complex gains and ϕ_k are the spatial directions of the LOS and NLOS channels.

As we know, the number of dominant scatters in a mmWave channel is quite limited [1]. Thus, the number of NLOS components L in (5) is typically much less than M_s, which means that the beamspace channel matrix $\boldsymbol{H}$ has a sparse structure.

C. Problem Formulation

Our work concentrates on the joint design of the digital precoding matrix $\boldsymbol{P}$ and the beam selection matrix $\boldsymbol{F}$ in order to maximize the downlink sum rate. The signal-to-interference-plus-noise ratio (SINR) of user k can be expressed as γ_k:

$$\gamma_k = \frac{|\boldsymbol{h}_k^H \boldsymbol{F}\boldsymbol{p}_k|^2}{\sum_{i \neq k}^{K} |\boldsymbol{h}_k^H \boldsymbol{F}\boldsymbol{p}_i|^2 + \sigma^2}. \tag{6}$$

Then the sum rate maximization problem can be mathematically formulated as

$$\begin{aligned} &\max_{\boldsymbol{F},\boldsymbol{P}} \sum_{k=1}^{K} \log(1 + \gamma_k) \\ &\text{s.t. } \text{Tr}(\boldsymbol{P}^H \boldsymbol{F}^H \boldsymbol{F}\boldsymbol{P}) \leq P_s, \\ &\sum_{i=1}^{M_s} f_{ij} = 1, \sum_{j=1}^{N_{RF}} f_{ij} \leq 1, \\ &f_{ij} \in \{0, 1\}, \forall (i, j) \in \mathcal{S}, \end{aligned} \tag{7}$$

where P_s is the transmit power upper bound of the BS, f_{ij} is the (i, j)-th element of $\boldsymbol{F}$, and $\mathcal{S} \triangleq \{(i, j) | i = 1, 2, \ldots, M_s, j = 1, 2, \ldots, N_{RF}\}$. The constraints $\sum_{i=1}^{M_s} f_{ij} = 1$ ensure that each RF chain is connected with one beam, while $\sum_{j=1}^{N_{RF}} f_{ij} \leq 1$ guarantee that each beam is selected for at most one RF chain. These constraints ensure that N_{RF} beams are selected to serve all users accurately.

Note that problem (7) is nonconvex and very challenging due to the beam selection constraints. In the next section, we propose an efficient joint beam selection and precoding algorithm to solve (7).

III. The Proposed Joint Design Algorithm

We first use the WMMSE approach to transform problem (7) into a tractable form. Then we propose an efficient joint beam selection and precoding algorithm to solve the equivalently converted problem based on the PDD method.

A. Reformulation of Problem (7)

Define the mean square error (MSE) of user k as

$$\begin{aligned} e_k &\triangleq E\{|\hat{s}_k - s_k|^2\} \\ &= |u_k \boldsymbol{h}_k^H \boldsymbol{F}\boldsymbol{p}_k|^2 - 2\Re e(u_k \boldsymbol{h}_k^H \boldsymbol{F}\boldsymbol{p}_k) \\ &\quad + 1 + \sigma^2 |u_k|^2 + \sum_{i \neq k}^{K} |u_k \boldsymbol{h}_k^H \boldsymbol{F}\boldsymbol{p}_i|^2, \end{aligned} \tag{8}$$

where $\hat{s}_k = u_k y_k$, u_k and y_k are the receiver gain and received signal of user k, respectively.

Let $w_k > 0$ be a weight factor, the problem

$$\begin{aligned} &\min_{\boldsymbol{F},\boldsymbol{P},w_k,u_k} \sum_{k=1}^{K} w_k e_k - \log w_k \\ &\text{s.t. } \text{Tr}(\boldsymbol{P}^H \boldsymbol{F}^H \boldsymbol{F}\boldsymbol{P}) \leq P_s, \sum_{i=1}^{M_s} f_{ij} = 1, \\ &\sum_{j=1}^{N_{RF}} f_{ij} \leq 1, f_{ij} \in \{0, 1\}, \forall (i, j) \in \mathcal{S}, \end{aligned} \tag{9}$$

[2]We consider a 2D formulation but extension to 3D is not difficult, and the channel is known by the BS [8].

is equivalent to problem (7), in the sense that the global optimal solution for the two problems are identical [9]. Note that u_k is only applied in (8) to satisfy the inverse MSE/rate relation as explained in [9].

Hence, the sum rate maximization problem (7) can be solved via the weighted MSE minimization (9). The latter problem, which is defined over the parameter space $\{\boldsymbol{F}, \boldsymbol{P}, w_k, u_k\}$, is easier to handle. In particular, optimizing each variable separately while holding the others fixed amounts to a convex problem that can be solved easily (e.g., closed-form).

Before proceeding to the derivation of the proposed algorithm, let us reformulate problem (9) into an equivalent, yet more tractable form. To this end, we introduce auxiliary variables $\{\hat{f}_{ij}\}$ and $\boldsymbol{V}$ subject to constraints $f_{ij} = \hat{f}_{ij}$, $f_{ij}(1-\hat{f}_{ij}) = 0$, $0 \le \hat{f}_{ij} \le 1$, $\boldsymbol{F}\boldsymbol{P} = \boldsymbol{V}$, $\text{Tr}(\boldsymbol{V}\boldsymbol{V}^H) \le P_s$. We also rewrite the constraint $\sum_{j=1}^{N_{RF}} f_{ij} \le 1$ as $\boldsymbol{f}_i^T \boldsymbol{1} \le 1$ with $\boldsymbol{f}_i^T \in \mathbb{C}^{1\times M_s}$ being the i-th row of $\boldsymbol{F}$ and $\boldsymbol{1} \in \mathbb{C}^{M_s \times 1}$ denoting a vector whose all elements are 1. Similarly, f_{ij} can be expressed as $\boldsymbol{f}_i^T \boldsymbol{b}_j$ with $\boldsymbol{b}_j \in \mathbb{C}^{N_{RF}\times 1}$ denoting the j-th column of $\boldsymbol{I}_{N_{RF}}$. With these notations, the problem (9) can be equivalently written as

$$\min_{\boldsymbol{V},\boldsymbol{P},w_k,u_k,f_{ij},\hat{f}_{ij}} \sum_{k=1}^{K} w_k e_k - \log w_k \tag{10a}$$

$$\text{s.t. } \boldsymbol{F}\boldsymbol{P} = \boldsymbol{V}, \boldsymbol{f}_i^T \boldsymbol{b}_j (1-\hat{f}_{ij}) = 0, \tag{10b}$$

$$\boldsymbol{f}_i^T \boldsymbol{b}_j = \hat{f}_{ij}, \sum_{i=1}^{M_s} \boldsymbol{f}_i^T \boldsymbol{b}_j = 1, \tag{10c}$$

$$\text{Tr}(\boldsymbol{V}\boldsymbol{V}^H) \le P_s, \sum_{j=1}^{N_{RF}} \boldsymbol{f}_i^T \boldsymbol{b}_j \le 1, \tag{10d}$$

$$0 \le \hat{f}_{ij} \le 1, \forall (i,j) \in \mathcal{S}. \tag{10e}$$

In the following, we develop the proposed algorithm based on the PDD method to solve problem (10).

B. The Proposed PDD-based Algorithm

In this subsection, we propose a PDD based algorithm, which is characterized by embedded double loop where the inner loop serves to approximately solve the augmented Lagrangian (AL) subproblem while the outer loop aims to update the dual variables or penalty parameter based on the constraint violation. We first write the AL problem of (10) as follows,

$$\begin{aligned}
&\min_{\boldsymbol{V},\boldsymbol{P},f_{ij},\hat{f}_{ij},w_k,u_k} \sum_{k=1}^{K} w_k e_k - \log w_k \\
&+ \frac{1}{2\rho} \sum_{i=1}^{M_s} \sum_{j=1}^{N_{RF}} (\boldsymbol{f}_i^T \boldsymbol{b}_j (1-\hat{f}_{ij}) + \rho \lambda_{ij})^2, \\
&+ \frac{1}{2\rho} \sum_{i=1}^{M_s} \sum_{j=1}^{N_{RF}} (\boldsymbol{f}_i^T \boldsymbol{b}_j - \hat{f}_{ij} + \rho \hat{\lambda}_{ij})^2, \\
&+ \frac{1}{2\rho} \sum_{j=1}^{N_{RF}} (\sum_{i=1}^{M_s} \boldsymbol{f}_i^T \boldsymbol{b}_j - 1 + \rho \mu_j)^2, \\
&+ \frac{1}{2\rho} ||\boldsymbol{F}\boldsymbol{P} - \boldsymbol{V} + \rho \boldsymbol{\xi}||_2^2, \\
&\text{s.t. } \text{Tr}(\boldsymbol{V}\boldsymbol{V}^H) \le P_s, \boldsymbol{f}_i^T \boldsymbol{1} \le 1, 0 \le \hat{f}_{ij} \le 1, \forall (i,j) \in \mathcal{S},
\end{aligned} \tag{11}$$

where $\{\lambda_{ij}\}$, $\{\hat{\lambda}_{ij}\}$, $\{\mu_j\}$, $\boldsymbol{\xi} \in \mathbb{C}^{M_s \times K}$ and ρ denote the dual variables associated with the equality constraints (10b), (10c) and a penalty factor, respectively. We address the AL problem (11) in the inner loop with the block successive upper-bound minimization (BSUM) method [11] and divide the variables into five blocks: $\{u_k\}$, $\{w_k\}$, $\{\hat{f}_{ij}, \boldsymbol{P}\}$, $\{\boldsymbol{V}\}$ and $\{\boldsymbol{f}_i^T\}$. In the following, we explain the methods to solve the subproblems for each one of these blocks.

Step 1. We optimize $\{u_k\}$ by fixing the remaining variables. In this case, (11) simplifies to the unconstrained problem:

$$\min_{u_k} \sum_{k=1}^{K} w_k e_k. \tag{12}$$

By examining the first order optimality condition of above problem, we obtain the closed-form solution [9]:

$$u_k^{opt} = u_k^{mmse} = J_k^{-1} \boldsymbol{v}_k^H \boldsymbol{h}_k, \tag{13}$$

where $J_k = \sum_{i=1}^{K} \boldsymbol{h}_k^H \boldsymbol{v}_i \boldsymbol{v}_i^H \boldsymbol{h}_k + \sigma^2$ and $\boldsymbol{v}_i$ is the i-th column of $\boldsymbol{V}$.

Step 2. We optimize $\{w_k\}$ by fixing the remaining variables. Following the approach in **Step 1**, we obtain the closed-form solution [9]:

$$w_k^{opt} = (e_k^{mmse})^{-1} = (1 - \boldsymbol{v}_k^H \boldsymbol{h}_k J_k^{-1} \boldsymbol{h}_k^H \boldsymbol{v}_k)^{-1}. \tag{14}$$

Step 3. We optimize $\{\hat{f}_{ij}, \boldsymbol{P}\}$ by fixing the remaining variables. The subproblem of optimizing $\{\hat{f}_{ij}\}$ can be expressed as

$$\begin{aligned}
\min_{\hat{f}_{ij}} &\frac{1}{2\rho} (\boldsymbol{f}_i^T \boldsymbol{b}_j (1-\hat{f}_{ij}) + \rho \lambda_{ij})^2 \\
&+ \frac{1}{2\rho} (\boldsymbol{f}_i^T \boldsymbol{b}_j - \hat{f}_{ij} + \rho \hat{\lambda}_{ij})^2 \\
\text{s.t. } & 0 \le \hat{f}_{ij} \le 1.
\end{aligned} \tag{15}$$

Problem (15) features a scalar quadratic objective function of $\hat{f}_{ij}$. We can obtain the solution without constraint by enforcing the first order optimality condition of the objective function:

$$\hat{f}_{ij}^{oc} = \frac{f_{ij}^2 + f_{ij}\rho\lambda_{ij} + f_{ij} + \rho\hat{\lambda}_{ij}}{1 + f_{ij}^2}. \tag{16}$$

Recalling that $\hat{f}_{ij}$ satisfies $0 \le \hat{f}_{ij} \le 1$, we can easily obtain the solution for $\hat{f}_{ij}$:

$$\hat{f}_{ij}^{opt} = \begin{cases} 1, & 1 \le \hat{f}_{ij}^{oc}, \\ \hat{f}_{ij}^{oc}, & 0 < \hat{f}_{ij}^{oc} < 1, \\ 0, & \hat{f}_{ij}^{oc} \le 0. \end{cases} \tag{17}$$

Similarly, the subproblem of optimizing $\{\boldsymbol{P}\}$ can be stated as

$$\min_{\boldsymbol{P}} \frac{1}{2\rho} ||\boldsymbol{F}\boldsymbol{P} - \boldsymbol{V} + \rho \boldsymbol{\xi}||_2^2. \tag{18}$$

Since this is a quadratic programming problem without constraint, we can obtain the following closed-form solution:

$$\boldsymbol{P}^{opt} = (\boldsymbol{F}^H \boldsymbol{F})^{\dagger} \boldsymbol{F}^H (\boldsymbol{V} - \rho \boldsymbol{\xi}). \tag{19}$$

As the variables $\boldsymbol{P}$ and $\hat{f}_{ij}$ are uncoupled in the block, we should update them simultaneously.

Step 4. We optimize $\{\boldsymbol{V}\}$ by fixing the remaining variables. The subproblem of optimizing $\{\boldsymbol{V}\}$ can be expressed as

$$\min_{\boldsymbol{V}} P_1(\boldsymbol{V}) \tag{20a}$$

$$\text{s.t. } \text{Tr}(\boldsymbol{V}\boldsymbol{V}^H) \le P_s, \tag{20b}$$

where

$$P_1(\boldsymbol{V}) \triangleq \sum_{k=1}^{K} w_k \big(|u_k \boldsymbol{h}_k^H \boldsymbol{V} \boldsymbol{b}_k|^2 - 2\Re e(u_k \boldsymbol{h}_k^H \boldsymbol{V} \boldsymbol{b}_k) + 1 + \sigma^2 |u_k|^2 + \sum_{i\neq k}^{K} |u_k \boldsymbol{h}_k^H \boldsymbol{V} \boldsymbol{b}_i|^2\big) + \frac{1}{2\rho} ||\boldsymbol{F}\boldsymbol{P} - \boldsymbol{V} + \rho\boldsymbol{\xi}||_2^2. \tag{21}$$

This is a convex quadratic optimization subproblem with a quadratic constraint. By introducing the Lagrange multiplier λ for the constraint (20b), we define the Lagrangian function associated with problem (20) as follows

$$\mathcal{L}(\boldsymbol{V}, \lambda) \triangleq P_1(\boldsymbol{V}) + \lambda(Tr(\boldsymbol{V}\boldsymbol{V}^H) - P_s). \tag{22}$$

The first order optimality condition of (22) with respect to $\boldsymbol{V}$ when $\lambda \geq 0$ yields:

$$\boldsymbol{V}^{opt} = (\sum_{k=1}^{K} w_k \boldsymbol{h}_k u_k^* u_k \boldsymbol{h}_k^H + \frac{1}{2\rho}\boldsymbol{I} + \lambda \boldsymbol{I})^{-1} \cdot \big(\frac{1}{2\rho}(\boldsymbol{F}\boldsymbol{P} + \rho\boldsymbol{\xi}) + \sum_{k=1}^{K} w_k u_k^* \boldsymbol{h}_k \boldsymbol{b}_k^H\big). \tag{23}$$

If the solution $\boldsymbol{V}$ with $\lambda = 0$ satisfies the constraint (20b), the optimal λ is zero. Otherwise we can obtain solution of λ through the slackness condition:

$$\text{Tr}(\boldsymbol{V}\boldsymbol{V}^H) - P_s = 0. \tag{24}$$

We can find λ by solving the problem (24) with the bisection method [12].

Step 5. We optimize $\{\boldsymbol{f}_i^T\}$ by fixing the remaining variables. The subproblem of optimizing $\{\boldsymbol{f}_i^T\}$ can be given by

$$\begin{aligned} &\min_{\boldsymbol{f}_i^T} P_2(\boldsymbol{f}_i^T) \\ &\text{s.t. } \boldsymbol{f}_i^T \boldsymbol{1} \leq 1, i = 1, 2, \ldots, M_s, \end{aligned} \tag{25}$$

where

$$\begin{aligned} P_2(\boldsymbol{f}_i^T) \triangleq &\frac{1}{2\rho} \sum_{i=1}^{M_s} \sum_{j=1}^{N_{RF}} (\boldsymbol{f}_i^T \boldsymbol{b}_j (1 - \hat{f}_{ij}) + \rho\lambda_{ij})^2 \\ &+ \frac{1}{2\rho} \sum_{i=1}^{M_s} \sum_{j=1}^{N_{RF}} (\boldsymbol{f}_i^T \boldsymbol{b}_j - \hat{f}_{ij} + \rho\hat{\lambda}_{ij})^2 \\ &+ \frac{1}{2\rho} \sum_{j=1}^{N_{RF}} (\sum_{i=1}^{M_s} \boldsymbol{f}_i^T \boldsymbol{b}_j - 1 + \rho\mu_j)^2 \\ &+ \frac{1}{2\rho} ||\boldsymbol{F}\boldsymbol{P} - \boldsymbol{V} + \rho\boldsymbol{\xi}||_2^2. \end{aligned} \tag{26}$$

We find that the subproblems for each $\boldsymbol{f}_i^T$ are independent. Each subproblem is convex with an affine constraint. Hence, similar to (20), the first order optimality condition of the Lagrangian function of (25) with respect to $\boldsymbol{f}_i^T$ when $\theta_i \geq 0$ yields:

$$\begin{aligned} \boldsymbol{f}_i^{opt} = &\big(3\boldsymbol{I} + \sum_{j=1}^{N_{RF}} (\hat{f}_{ij}^2 - 2\hat{f}_{ij}) \boldsymbol{b}_j \boldsymbol{b}_j^T + \Re e\{\boldsymbol{P}\boldsymbol{P}^H\}\big)^{-1} \\ &\cdot \big(- \Re e\{\boldsymbol{c}_i\} - \rho\theta_i \boldsymbol{1} - \sum_{j=1}^{N_{RF}} \big(\rho\lambda_{ij}(1 - \hat{f}_{ij}) \\ &- (\rho\hat{\lambda}_{ij} - \hat{f}_{ij}) - (\sum_{k\neq i}^{M_s} \boldsymbol{f}_k^T \boldsymbol{b}_j + \rho\mu_j - 1)\big)\boldsymbol{b}_j\big). \end{aligned} \tag{27}$$

where $\boldsymbol{f}_i = \{\boldsymbol{f}_i^T\}^T$, $\theta_i \geq 0$ is the Lagrangian multiplier, $\boldsymbol{C} \triangleq \boldsymbol{P}(\rho\boldsymbol{\xi} - \boldsymbol{V})^H$ and $\boldsymbol{c}_i$ denotes the i-th column of $\boldsymbol{C}$. If the solution $\boldsymbol{f}_i^T$ with $\theta_i = 0$ satisfies the constraint in (25), the optimal θ_i is zero. Otherwise we can obtain θ_i from the slackness condition:

$$\boldsymbol{f}_i^T \boldsymbol{1} = 1. \tag{28}$$

We can find θ_i by using the bisection method.

Besides, the dual variables $\{\lambda_{ij}, \hat{\lambda}_{ij}, \mu_j, \boldsymbol{\xi}\}^m$ can be updated by the following expressions:

$$\lambda_{ij}^{m+1} = \lambda_{ij}^m + \frac{1}{\rho^m}(\boldsymbol{f}_i^T \boldsymbol{b}_j (1 - \hat{f}_{ij})), \forall (i,j) \in \mathcal{S}, \tag{29a}$$

$$\hat{\lambda}_{ij}^{m+1} = \hat{\lambda}_{ij}^m + \frac{1}{\rho^m}(\boldsymbol{f}_i^T \boldsymbol{b}_j - \hat{f}_{ij}), \forall (i,j) \in \mathcal{S}, \tag{29b}$$

$$\mu_j^{m+1} = \mu_j^m + \frac{1}{\rho^m}(\sum_{i=1}^{M_s}(\boldsymbol{f}_i^T \boldsymbol{b}_j - 1)), \forall (i,j) \in \mathcal{S}, \tag{29c}$$

$$\boldsymbol{\xi}^{m+1} = \boldsymbol{\xi}^m + \frac{1}{\rho^m}(\boldsymbol{F}\boldsymbol{P} - \boldsymbol{V}). \tag{29d}$$

where m denotes the outer iteration number. We define the constraint violation h as:

$$h = \max_{\forall (i,j)\in\mathcal{S}} \left\{ \begin{array}{l} |\boldsymbol{f}_i^T \boldsymbol{b}_j (1 - \hat{f}_{ij})|, |\boldsymbol{f}_i^T \boldsymbol{b}_j - \hat{f}_{ij}|, \\ |\sum_{i=1}^{M_s} (\boldsymbol{f}_i^T \boldsymbol{b}_j - 1)|, ||\boldsymbol{F}\boldsymbol{P} - \boldsymbol{V}||_2. \end{array} \right\}. \tag{30}$$

The proposed joint design algorithm based on PDD method for problem (10) is summarized in Table I. Following [10], it can be shown that every limit point of the sequence generated by our algorithm is a stationary point of problem (10).

TABLE I: The proposed PDD based algorithm

1 Initialize dual variables $\{\lambda_{ij}, \hat{\lambda}_{ij}, \mu_j, \boldsymbol{\xi}\}^0$, primal variables $\{\boldsymbol{V}, \boldsymbol{P}, \boldsymbol{f}_i^T, \hat{f}_{ij}, w_k, u_k\}^{0,0}$, ϵ, N^{max}, $\rho^0 > 0$, τ_0, c and* $\eta^0 = \tau_0^{1/6}$. Set $m = 0$, $n = 0$.
2 **Repeat**
  2.1 **Repeat**
    2.1.1 Update $\{u_k\}^{m,n+1}$ by (13) with fixed $\{\boldsymbol{V}, \boldsymbol{P}, \boldsymbol{f}_i^T, \hat{f}_{ij}, w_k\}^{m,n}$.
    2.1.2 Update $\{w_k\}^{m,n+1}$ by (14) with fixed $\{\boldsymbol{V}, \boldsymbol{P}, \boldsymbol{f}_i^T, \hat{f}_{ij}\}^{m,n}$ and $\{u_k\}^{m,n+1}$.
    2.1.3 Update $\{\hat{f}_{ij}, \boldsymbol{P}\}^{m,n+1}$ by (17), (19) with fixed $\{\boldsymbol{V}, \boldsymbol{f}_i^T\}^{m,n}$ and $\{u_k, w_k\}^{m,n+1}$.
    2.1.4 Update $\{\boldsymbol{V}\}^{m,n+1}$ by (23) with fixed $\{\boldsymbol{f}_i^T\}^{m,n}$, and $\{\boldsymbol{P}, \hat{f}_{ij}, u_k, w_k\}^{m,n+1}$.
    2.1.5 Update $\{\boldsymbol{f}_i^T\}^{m,n+1}$ by (27) with fixed $\{\boldsymbol{V}, \boldsymbol{P}, \hat{f}_{ij}, u_k, w_k\}^{m,n+1}$ and set $n = n + 1$.
  2.2 **Until** $n > N^{max}$
  2.3 Assign the solution of the primal variables from (m, n) to $(m + 1, n)$.
  2.4 Calculate the constraint violation h by (30).
  2.5 **if** $h \leq \eta^m$ **then** update dual variables by (29),
  2.6 **else** set $\rho^{m+1} = c\rho^m$ **end if**.
  2.7 Set* $\tau_{m+1} = 0.6\tau_m$, $\eta^{m+1} = \tau_{m+1}^{1/6}$ and $m = m+1$.
3 **Until** $h < \epsilon$.

* Here we set N^{max}, ρ, c, η, and τ empirically.

IV. Simulation Results

This section presents simulation results of the proposed PDD based joint beam selection and precoding algorithm where the BS is equipped with a DLA consisting of $M_s = 128$ antennas and $N_{RF} = 16$ RF chains, in order to serve $K = 16$ users. The channel model parameters of user k are set according to [6]: 1) one LOS link with $L = 2$ NLOS links; 2) $\phi_k^{(0)}$ and $\phi_k^{(l)}$ obey the uniform independent identical distribution (i.i.d.) within $[-\frac{1}{2}, \frac{1}{2}]$. 3) $\beta_k^{(0)} \sim \mathcal{CN}(0, 1)$, $\beta_k^{(l)} \sim \mathcal{CN}(0, 10^{-1})$ with $l = 1, 2$. For the PDD based algorithm, we set the initial penalty factor $\rho^0 = 10^{-2}$ and the control parameter $c = 0.1$. In addition, we set $\tau^0 = 1$ and the maximum inner iteration number $N^{max} = 100$ which is large enough to observe convergence.

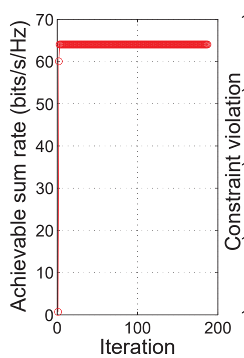

Fig. 2: Convergence properties of the PDD based algorithm (SNR = 25dB).

We choose SNR = 25dB to present the convergence properties of the PDD based algorithm without loss of generality in Fig. 2. The achievable sum rate converges rapidly in less than 5 iterations, and the constraint violation h reduces to a threshold $\epsilon = 10^{-7}$ in less than 200 iterations, which means that the solution has met equality constraints of problem (10).

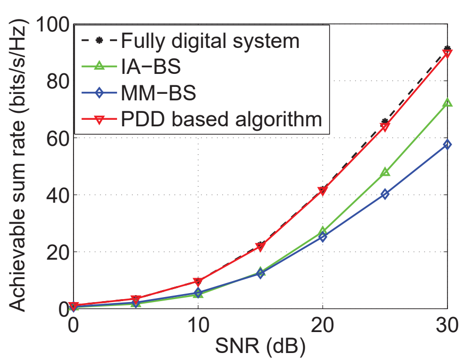

Fig. 3: Achievable sum rate comparison.

Fig. 3 compares the sum rate of the proposed joint design algorithm with the IA-BS [8], and MM-BS [5]–[7] schemes averaged over 100 channel realizations. Meanwhile we simulate the fully digital precoding as a benchmark. We observe that the proposed algorithm achieves a near-optimal performance when compared to the fully digital precoding scheme and thus outperform the competing methods. This demonstrates the merits of the proposed joint design algorithm for the beam selection and digital precoding.

V. Conclusion

In this paper, we have considered the joint design of beam selection and precoding for the downlink of a single-sided lens mmWave MU-MIMO system. We reformulated the sum rate maximization problem into a tractable form by using the WMMSE approach. We then proposed an efficient joint beam selection and precoding algorithm based on the PDD method. Simulation results demonstrate that our proposed algorithm can achieve near-optimal performance when compared to the fully digital precoding scheme and thus outperform the competing methods. The proposed algorithm exploited specific structure present in lens antenna array systems and could be used for any joint beam selection and precoding rate maximization problem.

References

[1] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi and F. Gutierrez, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, pp. 335-349, May 2013.

[2] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong and J. C. Zhang, "What will 5G be ?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065-1082, Jun. 2014.

[3] W. Roh, J. Y. Seol, J. Park, B. Lee, J. Lee, Y. Kim, J. Cho, K. Cheun and F. Aryanfar, "Millimeter-wave beamforming as an enabling technology for 5G cellular communications: Theoretical feasibility and prototype results," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 106-113, Feb. 2014.

[4] J. Brady, N. Behdad, and A. Sayeed, "Beamspace MIMO for millimeterwave communications: System architecture, modeling, analysis, and measurements," *IEEE Trans. Antennas Propag.*, vol. 61, no. 7, pp. 3814-3827, Jul. 2013.

[5] Y. Zeng and R. Zhang, "Millimeter wave MIMO With lens Antenna Array: A new path division multiplexing paradigm," *IEEE Trans. Commun.*, vol. 64, no. 4, pp. 1557-1571, April 2016.

[6] A. Sayeed and J. Brady, "Beamspace MIMO for high-dimensional multiuser communication at millimeter-wave frequencies," in *Proc. IEEE GLOBECOM*, Dec. 2013, pp. 3679-3684.

[7] P. Amadori and C. Masouros, "Low RF-complexity millimeter-wave beamspace-MIMO systems by beam selection," *IEEE Trans. Commun.*, vol. 63, no. 6, pp. 2212-2222, Jun. 2015.

[8] X. Gao, L. Dai, Z. Chen, Z. Wang, and Z. Zhang, "Near-optimal beam selection for beamspace mmWave massive MIMO systems," *IEEE Commun. Lett.*, vol. 20, no. 5, pp. 1054-1057, May 2016.

[9] 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.

[10] Q. Shi and M. Hong, "Penalty dual decomposition method with application in signal processing," accepted in *2017 IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Mar. 2017.

[11] M. Hong, T.-H. Chang, X. Wang, M. Razaviyayn, S. Ma, and Z.-Q. Luo. "A block successive upper bound minimization method of multipliers for linearly constrained convex optimization," 2013. Preprint, [Online]. Available: arXiv:1401.7079.

[12] Q. Shi, W. Xu, J. Wu, E. Song, and Y. Wang, "Secure beamforming for MIMO broadcasting with wireless information and power transfer," *IEEE Trans. Wireless Commun.*, vol. 14, no. 5, pp. 2841-2853, May 2015.