Channel Estimation for 3-D Lens Millimeter Wave Massive MIMO System

Wenyan Ma, *Student Member, IEEE*, and Chenhao Qi, *Senior Member, IEEE*

***Abstract*— Channel estimation for 3-D lens millimeter wave massive MIMO system is investigated. First, the structure of beamspace channel is analyzed to show that the dominant entries of beamspace channel matrix form a dual crossing (DC) shape. Then, a DC-based channel estimation algorithm is proposed to iteratively refine the selection of dominant entries until the stop condition is met. Simulation results show that the DC-based algorithm outperforms the existing algorithms, including orthogonal matching pursuit and adaptive support detection (ASD). At signal-to-noise ratio of 15 dB, the DC-based algorithm has 57.7% improvement in terms of normalized mean squared error compared with ASD while the computational complexity is only half of ASD due to fast computation of least squares estimation.**

***Index Terms*— Millimeter wave communications, channel estimation, massive MIMO, lens antenna array.**

I. INTRODUCTION

MILLIMETER wave (mmWave) communications, which range from 30GHz to 300GHz in frequency, have recently attracted significant attention due to large available bandwidth and high spectrum efficiency [1]. To compensate the path loss, mmWave system is usually equipped with massive MIMO antenna arrays to form analog beamforming for directional transmission. Recently, lens-based mmWave massive MIMO which can simplify the design of analog beamforming has been proposed [2]. Lens can work as discrete fourier transform (DFT) matrix and transform the spatial channel into beamspace channel [3]. Since the number of effective channel paths is limited in mmWave system, a beam selector is employed to select dominant beams from beamspace channel.

Channel estimation for 2D lens mmWave massive MIMO system is studied in [4] by detecting the dominant entries of beamspace channel matrix, which is essentially the 2D power leakage problem of DFT [5], [6]. In [7], an adaptive support detection (ASD) algorithm for beamspace channel estimation is proposed for 3D lens mmWave massive MIMO System, where the width and the length of a rectangle shaped by the channel dominant entries is adaptively modified. To our best knowledge, [7] is the only literature to address the 3D beamspace channel estimation.

In this letter, we consider the beamspace channel estimation for 3D lens mmWave massive MIMO system. First, we analyze the structure of beamspace channel and show that the dominant entries of channel matrix form a dual crossing (DC) shape. Then we propose a DC-based channel estimation algorithm which can iteratively refine the selection of dominant entries until the stop condition is met.

The notations are defined as follows. Symbols for matrices (upper case) and vectors (lower case) are in boldface. $(\cdot)^T$, $(\cdot)^H$, $(\cdot)^{-1}$, $\boldsymbol{I}_L$, $\mathbb{C}^{M\times N}$, $\otimes$, vec$(\cdot)$, E$\{\cdot\}$, $\boldsymbol{0}^M$, $\|\cdot\|_0$, $\|\cdot\|_2$, $\boldsymbol{A}[p,q]$ and $\mathcal{CN}$, denote the transpose, conjugate transpose (Hermitian), inversion, identity matrix of size L, set of $M\times N$ complex-valued matrices, kronecker product, vectorization, expectation, zero vector of size M, l_0-norm, l_2-norm, entry of $\boldsymbol{A}$ at the pth row and qth column and complex Gaussian distribution, respectively.

II. SYSTEM MODEL

We consider a downlink multi-user mmWave massive MIMO system comprising a BS and U single-antenna users. The BS is equipped with an uniform planar array (UPA) which has M_v antennas in vertical and M_h antennas in horizontal. The total number of antennas at BS is $M \triangleq M_v M_h$. We adopt the widely used 3D Saleh-Valenzuela channel model [1], where the channel vector $\boldsymbol{h}_u \in \mathbb{C}^M$ between the BS and the $u(u = 1, 2, \ldots, U)$th user can be expressed as

$$\boldsymbol{h}_u = \sqrt{\frac{M}{L_u}} \sum_{i=1}^{L_u} \boldsymbol{h}_{u,i} = \sqrt{\frac{M}{L_u}} \sum_{i=1}^{L_u} g_{u,i} \boldsymbol{\alpha}(M_h, \theta_{u,i}) \otimes \boldsymbol{\alpha}(M_v, \varphi_{u,i}) \tag{1}$$

where L_u, $\boldsymbol{h}_{u,i}$ and $g_{u,i}$ denote the total number of resolvable paths, channel vector and channel fading coefficient of the ith path, respectively. The steering vector $\boldsymbol{\alpha}(M,\theta)$ is defined as $\boldsymbol{\alpha}(M,\theta) = \frac{1}{M}\left[1, e^{-j2\pi\theta}, ..., e^{-j2\pi\theta(M-1)}\right]^T$. Define physical azimuth and physical elevation of the ith path of the uth user as $\Theta_{u,i}$ and $\Phi_{u,i}$, respectively. Further define $\theta_{u,i} = \frac{d_h}{\lambda}\sin\Theta_{u,i}$ and $\varphi_{u,i} = \frac{d_v}{\lambda}\sin\Phi_{u,i}$, where d_h and d_v denote the antenna interval in horizontal and in vertical, respectively. We usually set $d_v = d_h = \lambda/2$, where λ is the wavelength of mmWave signal. Both $\theta_{u,i}$ and $\varphi_{u,i}$ obey uniform distribution $[-0.5, 0.5]$.

We place 3D lens behind the UPA to concentrate signal power on specified antennas. The 3D lens $\boldsymbol{G} \in \mathbb{C}^{M\times M}$ can be mathematically represented as the kronecker product of two DFT matrix $\boldsymbol{G} = \boldsymbol{D}(M_h) \otimes \boldsymbol{D}(M_v)$, where the DFT matrix of size M can be defined as $\boldsymbol{D}(M) = [\boldsymbol{\alpha}(M,0), \boldsymbol{\alpha}(M,1/M), ..., \boldsymbol{\alpha}(M,(M-1)/M)]^H$. The received signal of all U users, denoted as $\boldsymbol{y} \in \mathbb{C}^U$ can be expressed as

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

where $\boldsymbol{F} \in \mathbb{C}^{U\times U}$ is a digital beamforming matrix, $\boldsymbol{s} \in \mathbb{C}^U$ is a signal vector satisfying $\text{E}\{\boldsymbol{s}\boldsymbol{s}^H\} = \boldsymbol{I}_U$, and $\boldsymbol{n} \in \mathbb{C}^U$ is

Manuscript received April 10, 2017; revised May 12, 2017; accepted May 24, 2017. Date of publication June 5, 2017; date of current version September 8, 2017. This work was supported in part by National Natural Science Foundation of China under Grant 61302097, Natural Science Foundation of Jiangsu Province under Grant BK20161428, and National Science and Technology Major Project of China under Grant 2016ZX03001016-003. The associate editor coordinating the review of this letter and approving it for publication was F. Gao. *(Corresponding author: Chenhao Qi.)*

The authors are with the School of Information Science and Engineering, Southeast University, Nanjing 210096, China (e-mail: qch@seu.edu.cn).

Digital Object Identifier 10.1109/LCOMM.2017.2712132

an additive white Gaussian noise (AWGN) vector satisfying $\boldsymbol{n} \sim \mathcal{CN}(0, \sigma^2 \boldsymbol{I}_U)$. We define

$$\boldsymbol{H}^b \triangleq \boldsymbol{H}\boldsymbol{G} \tag{3}$$

as a beamspace channel matrix where $\boldsymbol{H} \triangleq [\boldsymbol{h}_1, ..., \boldsymbol{h}_U]^T \in \mathbb{C}^{U\times M}$ is the mmWave massive MIMO channel matrix between the BS and all users. Due to the limited scattering of mmWave channel, only dominant beams need to be selected and used for signal transmission. We denote the beam selector as $\boldsymbol{B} \in \mathbb{C}^{M\times U}$.

Note that the design of $\boldsymbol{B}$ and $\boldsymbol{F}$ relies on $\boldsymbol{H}^b$. To acquire an estimate of $\boldsymbol{H}^b$, uplink channel estimation is usually employed so that an estimate of downlink channel can be obtained based on channel reciprocity. Suppose each user repeatedly sends an orthogonal pilot sequence with length of U for K times and the channel keeps constant during $V \triangleq KU$ time slots. The U orthogonal pilot sequences sent from U users make up a pilot matrix $\boldsymbol{P} \in \mathbb{C}^{U\times U}$. For the $k(k = 1, 2, \ldots, K)$th sending, the downlink beam selector works as an uplink combiner $\boldsymbol{C}_k^T \in \mathbb{C}^{U\times M}$, so the pilot sequences received by BS are denoted as

$$\boldsymbol{Y}_k = \boldsymbol{C}_k^T (\boldsymbol{H}^b)^T \boldsymbol{P} + \boldsymbol{C}_k^T \boldsymbol{N}_k \tag{4}$$

where $\boldsymbol{N}_k$ is an AWGN matrix. Each entry of $\boldsymbol{N}_k$ independently obeys complex Gaussian distribution with zero mean and variance of σ^2. Due to the orthogonality of $\boldsymbol{P}$, i.e., $\boldsymbol{P}\boldsymbol{P}^H = \boldsymbol{I}_U$, we can multiply $\boldsymbol{Y}_k$ with $\boldsymbol{P}^H$ as

$$\boldsymbol{R}_k = \boldsymbol{Y}_k \boldsymbol{P}^H = \boldsymbol{C}_k^T (\boldsymbol{H}^b)^T + \widetilde{\boldsymbol{N}}_k \tag{5}$$

where $\widetilde{\boldsymbol{N}}_k \triangleq \boldsymbol{C}_k^T \boldsymbol{N}_k \boldsymbol{P}^H$. After all users repeatedly send pilot sequences for K times, we can stack $\boldsymbol{R}_k, k = 1, 2, \ldots, K$ into a matrix $\boldsymbol{R}$ as

$$\boldsymbol{R} = [\boldsymbol{R}_1^T, \ldots, \boldsymbol{R}_K^T]^T = \boldsymbol{C}^T (\boldsymbol{H}^b)^T + \widetilde{\boldsymbol{N}} \tag{6}$$

where $\boldsymbol{C} \triangleq [\boldsymbol{C}_1, \ldots, \boldsymbol{C}_K]^T$ and $\widetilde{\boldsymbol{N}} \triangleq [\widetilde{\boldsymbol{N}}_1^T, \ldots, \widetilde{\boldsymbol{N}}_K^T]^T$. The $u(u = 1, 2, \ldots, U)$th column of $\boldsymbol{R}$, denoted as $\boldsymbol{r}_u$, can be represented as

$$\boldsymbol{r}_u = \boldsymbol{C}^T \boldsymbol{h}_u^b + \widetilde{\boldsymbol{n}}_u \tag{7}$$

where $\boldsymbol{h}_u^b$ is the uth column of $(\boldsymbol{H}^b)^T$ and $\widetilde{\boldsymbol{n}}_u$ is the uth column of $\widetilde{\boldsymbol{N}}$. In fact, $\boldsymbol{h}_u^b = \boldsymbol{G}^T \boldsymbol{h}_u$ is the beamspace channel vector from the uth user to the BS. Due to the sparse property of $\boldsymbol{h}_u^b$, (7) is essentially a sparse recovery problem. However, considering the limited beamspace resolution of $\boldsymbol{G}$, the sparsity of $\boldsymbol{h}_u^b$ may be impaired by power leakage [3], indicating that $\boldsymbol{h}_u^b$ is not ideally sparse and there are many small nonzero entries of $\boldsymbol{h}_u^b$. Therefore, it brings challenge for the sparse recovery.

III. Analysis of Beamspace Channel

We first analyze the structure of beamspace channel. Based on our analysis, we will then propose a DC-based beamspace channel estimation algorithm for 3D lens mmWave massive MIMO system.

Define the beamspace channel vector of the $i(i = 1, 2, \ldots, L_u)$th path from the uth user to the BS as $\boldsymbol{h}_{u,i}^b$.

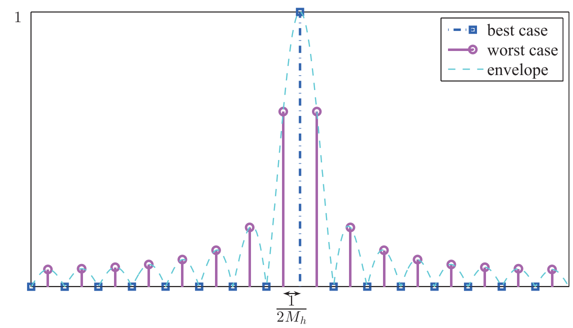

Fig. 1. Envelop of $\boldsymbol{\alpha}^b(M_h, \theta_{u,i})$.

According to (1), $\boldsymbol{h}_{u,i}^b$ can be represented as

$$\begin{aligned}\boldsymbol{h}_{u,i}^b &= \boldsymbol{G}^T \boldsymbol{h}_{u,i} \\ &= \boldsymbol{D}(M_h)^T \otimes \boldsymbol{D}(M_v)^T g_{u,i} \boldsymbol{\alpha}(M_h, \theta_{u,i}) \otimes \boldsymbol{\alpha}(M_v, \varphi_{u,i}) \\ &= g_{u,i} \left(\boldsymbol{D}(M_h)^T \boldsymbol{\alpha}(M_h, \theta_{u,i})\right) \otimes \left(\boldsymbol{D}(M_v)^T \boldsymbol{\alpha}(M_v, \varphi_{u,i})\right) \\ &= g_{u,i} \boldsymbol{\alpha}^b(M_h, \theta_{u,i}) \otimes \boldsymbol{\alpha}^b(M_v, \varphi_{u,i}) \\ &= g_{u,i} \mathrm{vec}\left(\boldsymbol{\alpha}^b(M_v, \varphi_{u,i}) \left(\boldsymbol{\alpha}^b(M_h, \theta_{u,i})\right)^T\right)\end{aligned} \tag{8}$$

where

$$\boldsymbol{\alpha}^b(M_h, \theta_{u,i}) \triangleq \boldsymbol{D}(M_h)^T \boldsymbol{\alpha}(M_h, \theta_{u,i}), \tag{9}$$

$$\boldsymbol{\alpha}^b(M_v, \varphi_{u,i}) \triangleq \boldsymbol{D}(M_v)^T \boldsymbol{\alpha}(M_v, \varphi_{u,i}), \tag{10}$$

is the DFT of channel steering vector. Since $\theta_{u,i}$ and $\varphi_{u,i}$ can be any value from $-1/2$ to $1/2$ while the DFT resolution is fixed to be $1/M_h$ and $1/M_v$, the envelop of $\boldsymbol{\alpha}^b(M_h, \theta_{u,i})$ will appear as single peak only if $\theta_{u,i} \in \boldsymbol{\Lambda} \triangleq \{j/M_h - 1/2, j = 0, 1, ..., M_h - 1\}$, which is the best case without power leakage. But in most cases $\theta_{u,i} \notin \boldsymbol{\Lambda}$, the envelop of $\boldsymbol{\alpha}^b(M_h, \theta_{u,i})$ will appear as two high peaks in main lobe and many low peaks in side lobes, which is caused by the power leakage. In the worst case as shown in Fig. 1, the envelop of $\boldsymbol{\alpha}^b(M_h, \theta_{u,i})$ will appear as two equally high peaks in main lobe and the highest peak in each side lobe, where the power leakage is the largest. The envelop of $\boldsymbol{\alpha}^b(M_v, \varphi_{u,i})$ is similar as $\boldsymbol{\alpha}^b(M_h, \theta_{u,i})$.

To effectively detect the dominant entries of $\boldsymbol{h}_{u,i}^b$, we reshape $\boldsymbol{h}_{u,i}^b$ into a two-dimensional channel matrix $\boldsymbol{H}_{u,i}^b$ with M_v rows and M_h columns as

$$\boldsymbol{H}_{u,i}^b \triangleq g_{u,i} \boldsymbol{\alpha}^b(M_v, \varphi_{u,i}) \left(\boldsymbol{\alpha}^b(M_h, \theta_{u,i})\right)^T. \tag{11}$$

In fact, the entry at the $p(p = 1, 2, \ldots, M_v)$th row and the $q(q = 1, 2, \ldots, M_h)$th column of $\boldsymbol{H}_{u,i}^b$, denoted as $\boldsymbol{H}_{u,i}^b[p, q]$ can be written as

$$\boldsymbol{H}_{u,i}^b[p, q] = \boldsymbol{h}_{u,i}^b\left[(q-1)M_v + p\right]. \tag{12}$$

It is rare event that $\boldsymbol{H}_{u,i}^b$ is ideally sparse only when both $\boldsymbol{\alpha}^b(M_v, \varphi_{u,i})$ and $\boldsymbol{\alpha}^b(M_h, \theta_{u,i})$ are the best cases. In most cases, $\boldsymbol{H}_{u,i}^b$ is not sparse since there are many nonzero entries due to the power leakage. The four peaks in main lobe are significantly higher than the other peaks in side lobes, meaning that the two rows and two columns of $\boldsymbol{H}_{u,i}^b$ including the four peaks in main lobe occupy a large portion of channel power. For example, if $M_v = M_h = 32$ and $\theta_{u,i} = \varphi_{u,i} = 1/64$, in the worst case, the largest 35 entries of $\boldsymbol{H}_{u,i}^b$ occupying 90% power of $\boldsymbol{H}_{u,i}^b$ are all in the $\{M_v/2, M_v/2+1\}$th rows

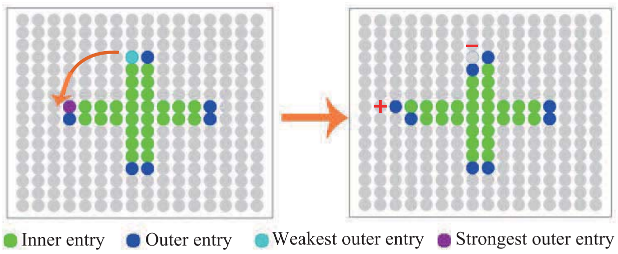

Fig. 2. Adjustment of dominant entries while keeping the DC shape.

and the $\{M_h/2, M_h/2+1\}$th columns, where the positions of these 35 entries form a DC shape as shown in Fig. 2.

IV. DC-Based Beamspace Channel Estimation

Based on the analysis of the beamspace channel, now we propose a DC-based beamspace channel estimation algorithm for each user. First we have to determine the number of dominant entries, denoted as J, for each channel vector. Define the power ratio of dominant channel entries over total channel entries as $\eta, \eta \in (0, 1]$. Since in the worst case the power leakage of the channel is most serious, J will be largest to satisfy η, compared to that in other cases. Given M_h and M_v, we denote the channel vector in the worse case as

$$\boldsymbol{h}_w^b = \boldsymbol{\alpha}^b(M_h, 1/(2M_h)) \otimes \boldsymbol{\alpha}^b(M_v, 1/(2M_v)). \tag{13}$$

Given η, we can obtain J via

$$J = \arg \min_{J=\|\boldsymbol{\Gamma}\|_0} \left\{ \sum_{i\in\boldsymbol{\Gamma}} \left| \boldsymbol{h}_w^b[i] \right|^2 > \eta \right\} \tag{14}$$

where $\boldsymbol{\Gamma}$ denoting a set of positions of dominant channel entries, is a subset of $\boldsymbol{\Omega} \triangleq \{1, 2, \ldots, M\}$, i.e., $\boldsymbol{\Gamma} \subset \boldsymbol{\Omega}$. In fact, we can sort $\{|\boldsymbol{h}_w^b[i]|, i \in \boldsymbol{\Omega}\}$ in descending order and then select the first J entries satisfying (14).

As shown in **Algorithm 1**, we use $\boldsymbol{r}_{u,i}$ to obtain an estimate of $\boldsymbol{h}_{u,i}^b$, denoted as $\hat{\boldsymbol{h}}_{u,i}^b$, $i = 1, 2, \ldots, \hat{L}_u$. Note that in practice L_u is unknown and we use an estimate of it as $\hat{L}_u$ for the algorithm. At step 2, we initialize $\boldsymbol{r}_{u,1}$ to be $\boldsymbol{r}_u$. At step 13, we obtain $\boldsymbol{r}_{u,i+1}$ based on $\boldsymbol{r}_{u,i}$ via

$$\boldsymbol{r}_{u,i+1} \leftarrow \boldsymbol{r}_{u,i} - \boldsymbol{C}^T \hat{\boldsymbol{h}}_{u,i}^b. \tag{15}$$

From step 4 to step 14, we iteratively estimate the positions of dominant entries for the ith channel path, denoted as $\boldsymbol{\Gamma}_i, i = 1, 2, \ldots, \hat{L}_u$. At each iteration, we first find two adjacent columns indexed by $\{s_q, s_q+1\}$ with the largest channel power from $\boldsymbol{H}_{u,i}^b$ via

$$s_q = \arg \max_{q=1,2,\ldots,M_h-1} \left\| (\boldsymbol{C}_q^H \boldsymbol{C}_q)^{-1} \boldsymbol{C}_q^H \boldsymbol{r}_{u,i} \right\|_2 \tag{16}$$

where $\boldsymbol{C}_q$ represents a submatrix consisted of $2M_v$ consecutive columns, with column indices denoted as $(q-1)M_v+1, (q-1)M_v+2, \ldots, (q-1)M_v+M_v, qM_v+1, qM_v+2, \ldots, qM_v+M_v$ from $\boldsymbol{C}^T$. Similarly, we find two adjacent rows indexed by $\{s_p, s_p+1\}$ with the largest channel power from $\boldsymbol{H}_{u,i}^b$ via

$$s_p = \arg \max_{p=1,2,\ldots,M_v-1} \left\| (\boldsymbol{C}_p^H \boldsymbol{C}_p)^{-1} \boldsymbol{C}_p^H \boldsymbol{r}_{u,i} \right\|_2 \tag{17}$$

Algorithm 1 DC-Based Beamspace Channel Estimation

1: *Input:* $\boldsymbol{r}_u$, $\boldsymbol{C}$, $\hat{L}_u$, M_v, M_h, η.
2: Initialization: $\boldsymbol{r}_{u,1} \leftarrow \boldsymbol{r}_u$. $\hat{\boldsymbol{h}}_u^b \leftarrow \boldsymbol{0}^M$.
3: Obtain J via (14).
4: **for** $i = 1, 2, \ldots, \hat{L}_u$ **do**
5: Obtain s_q and s_p via (16) and (17), respectively.
6: Obtain J dominant entries of $\boldsymbol{H}_{u,i}^b$.
7: **repeat**
8: Obtain $\boldsymbol{\Gamma}_i$ based on J dominant entries.
9: Compute $\hat{\boldsymbol{h}}_{u,i}^b[\boldsymbol{\Gamma}_i]$ via (18).
10: Update $\hat{\boldsymbol{H}}_{u,i}^b$ based on $\hat{\boldsymbol{h}}_{u,i}^b[\boldsymbol{\Gamma}_i]$ according to (12).
11: Adjust the selection of J dominant entries.
12: **until** *Stop Condition* is met.
13: Obtain $\boldsymbol{r}_{u,i+1}$ via (15).
14: **end for**
15: $\boldsymbol{\Gamma}_u = \bigcup_{i=1,2,\ldots,\hat{L}_u} \boldsymbol{\Gamma}_i$.
16: Obtain $\hat{\boldsymbol{h}}_u^b[\boldsymbol{\Gamma}_u]$ via (18).
17: *Output:* $\hat{\boldsymbol{h}}_u^b$.

where $\boldsymbol{C}_p$ represents a submatrix consisted of $2M_h$ inconsecutive columns, with colum indices denoted as $p, p+1, p+M_v, p+1+M_v, p+2M_v, p+1+2M_v, \ldots, p+(M_h-1)M_v, p+1+(M_h-1)M_v$ from $\boldsymbol{C}^T$. Now the selected two adjacent rows and two adjacent columns form a DC shape, where the crossing four peaks in main lobe can be determined. Based on these four peaks, we start to select the other $J-4$ dominant entries from $\boldsymbol{H}_{u,i}^b$ at step 6. Initially we select $Q \triangleq \lfloor J/4 \rfloor - 1$ entries on the top, bottom, left and right of the four peaks, respectively, while the left $J-4Q-4$ entries can be uniformly placed on any corners of DC, as shown in Fig. 2.

From step 7 to step 12, we repeat the procedures to refine the selection of J dominant entries while keeping the DC shape. Note that in 2D lens mmWave massive MIMO channel estimation, the dominant entries of channel vector can be directly calculated without refinement, since the dominant entries are symmetrically distributed on the two sides of the largest entry. At step 8, the positions of entries of $\boldsymbol{h}_{u,i}^b$ corresponding to the initially selected J entries, denoted as $\boldsymbol{\Gamma}_i$, can be obtained according to the relationship described in (12). We estimate the J entries of $\hat{\boldsymbol{h}}_{u,i}^b$ at step 9 which is essentially least squares (LS) estimation as

$$\hat{\boldsymbol{h}}_{u,i}^b[\boldsymbol{\Gamma}_i] = (\boldsymbol{C}_{\Gamma_i}^H \boldsymbol{C}_{\Gamma_i})^{-1} \boldsymbol{C}_{\Gamma_i}^H \boldsymbol{r}_{u,i} \tag{18}$$

where $\boldsymbol{C}_{\Gamma_i}$ represents a submatrix consisted of J columns from $\boldsymbol{C}^T$ with column indices $\boldsymbol{\Gamma}_i$. Then we update the entries of $\hat{\boldsymbol{H}}_{u,i}^b$ corresponding to the entries of $\hat{\boldsymbol{h}}_{u,i}^b[\boldsymbol{\Gamma}_i]$ according to (12) at step 10. As shown in Fig. 2, there are totally eight outer entries at four DC corners. At step 11, we adjust the selection of J dominant entries by deleting the weakest outer entry and then adding a new entry outside of the strongest outer entry along DC.

Definition 1 (Stop Condition): A new entry is to be added to the position where an old entry has been deleted, or a newly added entry is to be deleted.

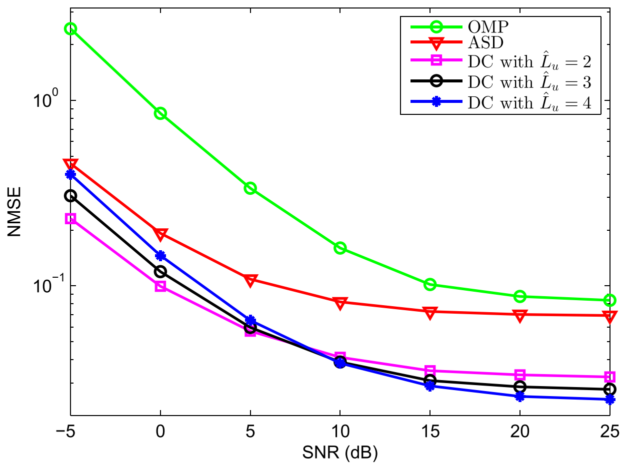

Fig. 3. Comparisons of NMSE for different channel estimation algorithms.

The *Stop Condition* implies that the incremental property or decremental property, marked as "+" or "-" in Fig. 2, can be determined by the operations of adding a new entry or deleting an old entry, respectively. The property of eight line ends of four DC corners should be determined. Once the property is determined, it can not be changed. If the property is to be violated, we stop the iterations. It indicates that the line end can not be extended anymore if an entry on the line end is deleted. In this manner, we can keep on extending the line end that has larger dominant entry outside of the current outer entry until convergence.

After the positions of J dominant entries are obtained for each channel path, we make an union of them at step 15. We obtain $\hat{\boldsymbol{h}}_u^b[\boldsymbol{\Gamma}_u]$ via (18) by replacing $\boldsymbol{\Gamma}_i$, $\hat{\boldsymbol{h}}_{u,i}^b$ and $\boldsymbol{r}_{u,i}$ with $\boldsymbol{\Gamma}_u$, $\hat{\boldsymbol{h}}_u^b$ and $\boldsymbol{r}_u$, respectively.

Note that we only change two entries for each adjustment of J dominant entries at step 11. Therefore at step 9 we can fast compute LS obtaining coefficients of newly selected J entries based on the coefficients of previously selected J entries, instead of making a new LS estimation in (18), which leads to substantial reduction in computational complexity from $\mathcal{O}(2VJ^2 + 4J^3)$ to $\mathcal{O}(4VJ + 10J^2 + 2J)$ [8].

V. Simulation Results

We consider an mmWave massive MIMO system including a BS equipped with $M_h = M_v = 32$ antennas and $U = 16$ single-antenna users. The number of resolvable paths in mmWave channel is set to be $L_u = 3$, while $g_{u,1} \sim \mathcal{CN}(0, 1)$ and $g_{u,i} \sim \mathcal{CN}(0, 0.01)$ for $i = 2, 3$. We use $V = 256$ time slots to transmit pilot sequences for uplink channel estimation. Uplink combiner $\boldsymbol{C}^T$ is assumed to be a Bernoulli random matrix with each entry independently obeying bivariate uniform distribution $\frac{1}{V}\{-1, +1\}$. We set $J = 64$ to be the same as [7] so that we can fairly compare DC-based algorithm with ASD-based algorithm proposed in [7].

As shown in Fig. 3, we make comparisons of normalized mean squared error (NMSE) for different channel estimation algorithms, including DC-based algorithm, ASD-based algorithm and orthogonal matching pursuit (OMP) sparse recovery algorithm. Since L_u is unknown in practice, we set $\hat{L}_u = 2, 3, 4$, respectively. It is seen that the performance of DC-based algorithm with $\hat{L}_u = 2, 3, 4$ is similar, while DC-based algorithm and ASD-based algorithm outperform OMP sparse recovery especially in low signal-to-noise ratio (SNR) region. In high SNR region, DC-based algorithm performs better than ASD-based algorithm, which has 57.7% improvement at SNR = 15dB.

TABLE I
Comparisons of Computational Complexity for DC-Based Algorithm and ASD-Based Algorithm

SNR(dB)	-5	0	5	10	15	20	25
DC	3.219	3.331	3.968	5.592	7.449	8.544	9.076
ASD	2.408	2.374	2.383	2.393	2.392	2.391	2.385
Ratio	0.446	0.454	0.462	0.483	0.510	0.525	0.534

As shown in Table I, we make comparisons of computational complexity for DC-based algorithm and ASD-based algorithm. The second row and the third row of Table I show the average number of repeating the procedures during the refinement of selection of J dominant entries for DC-based algorithm and ASD-based algorithm, respectively. The last row of Table I shows the ratio of computational complexity of DC-based algorithm over ASD-based algorithm regarding the fast computation of LS addressed at the end of Section IV. In fact, the ratio is $((2VJ^2 + 4J^3) + (t_1 - 1)(4VJ + 10J^2 + 2J))/(2VJ^2t_2 + 4J^3t_2)$, where t_1 and t_2 denote the corresponding entry in the second row and in the third row, respectively. It is shown that at SNR = 15dB, the computational complexity of DC-based algorithm is only half of that of ASD-based algorithm.

VI. Conclusion

In this letter, we have analyzed the beamspace channel to show that the dominant entries of channel matrix form a DC shape. Then a DC-based channel estimation algorithm is proposed to iteratively refine the selection of dominant entries until the stop condition is met.

References

[1] R. W. Heath, N. González-Prelcic, S. Rangan, W. Roh, and A. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 436–453, Apr. 2016.

[2] 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, Apr. 2016.

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

[4] L. Dai, X. Gao, S. Han, I. Chih-Lin, and X. Wang, "Beamspace channel estimation for millimeter-wave massive MIMO systems with lens antenna array," *IEEE Trans. Wireless Commun.*, to be published.

[5] H. Xie, F. Gao, S. Zhang, and S. Jin, "A unified transmission strategy for TDD/FDD massive MIMO systems with spatial basis expansion model," *IEEE Trans. Veh. Technol.*, vol. 66, no. 4, pp. 3170–3184, Apr. 2017.

[6] H. Xie, F. Gao, and S. Jin, "An overview of low-rank channel estimation for massive MIMO systems," *IEEE Access*, vol. 4, pp. 7313–7321, 2016.

[7] X. Gao, L. Dai, S. Han, I. Chih-Lin, and F. Adachi, "Beamspace channel estimation for 3D lens-based millimeter-wave massive MIMO systems," in *Proc. IEEE WCSP*, Yangzhou, China, Oct. 2016, pp. 1–5.

[8] A. Jennings and J. J. McKeown, *Matrix Computation*. Hoboken, NJ, USA: Wiley, 1992.