Model based Beamspace Channel Estimation for Millimeter Wave Massive MIMO System

Zhikang Xia, Chenhao Qi and Tengxiang Zhang
School of Information Science and Engineering
Southeast University, Nanjing 210096, China
Email: qch@seu.edu.cn

***Abstract*—Channel estimation concerning channel power leakage is investigated for millimeter wave (mmWave) massive MIMO system where the base station (BS) is equipped with uniform planar array (UPA). To model the channel leakage, the leakage effect of two dimensional beamspace channel is firstly analyzed. A theorem is proved to show that a four-point star (FS)-based model is better than the existing dual-crossing (DC)-based model. Then a FS-based beamspace channel estimation algorithm is proposed, where the initial selection of dominant entries is computed offline and then the selection is iteratively refined until the stop condition is satisfied. Simulation results verify the effectiveness of our work and show that the FS-based channel estimation algorithm outperforms the DC-based channel estimation algorithm.**

***Index Terms*—Millimeter wave (mmWave) communications, massive MIMO, channel estimation, channel power leakage.**

I. Introduction

Millimeter wave (mmWave) communication is a promising technology for next-generation wireless communication due to large available bandwidth and high spectrum efficiency [1], [2]. However, there is large propagation loss caused by high carrier frequency [3]. In order to compensate the propagation loss, massive multi-input multi-ouput (MIMO) antenna arrays is usually equipped in mmWave system to form analog beamforming for directional transmission [4]. However, in order to apply mmWave massive MIMO communication, there are a lot of challenges to be overcome, one of which is to make efficient channel estimation [5], [6], [7], [8]. Note that there exists considerable channel power leakage in practice when performing channel estimation [9]. In [10], the dominant entries of beamspace channel matrix are required for the channel estimation of 2D lens mmWave massive MIMO system. In order to obtain the beamspace channel estimation, an adaptive support detection (ASD) scheme is proposed in [11], which uses an adaptively updated rectangle to determine the selection of dominant entries and reduce the channel power leakage. In [12], a dual crossing (DC)-based beamspace channel estimation scheme which uses two dual crossings and iteratively makes optimizations to select the dominant entries, is proposed to improve the channel estimation accuracy.

In this paper, we investigate the channel estimation concerning channel power leakage for mmWave massive MIMO system where the base station (BS) is equipped with uniform planar array (UPA). First we analyze the leakage effect of two dimensional beamspace channel. We provide a theorem to show that a four-point star (FS)-based model is better than the existing DC-based model. Then we propose a FS-based beamspace channel estimation algorithm, where the initial selection of dominant entries is computed offline and then the selection is iteratively refined until the stop condition is satisfied.

Notations: symbols in boldface represent matrices (upper case) and vectors (lower case). $(\cdot)^T$, $(\cdot)^H$ and $(\cdot)^{-1}$ denote the transpose, conjugate transpose and inversion, respectively. $\boldsymbol{I}_L$ is a matrix of size L. $\mathbb{C}^{M\times N}$, vec$(\cdot)$, $\mathcal{CN}$ represent the set of $M\times N$ complex-valued matrices, vectorization and complex Gaussian distribution, respectively. $\mathbf{0}^M$ means a zero vector of size M. $\boldsymbol{A}[p,q]$, $\otimes$ and $\lfloor\cdot\rfloor$ represent the entry of $\boldsymbol{A}$ at the pth row and qth column, kronecker product and round to infinite, respectively. $\|\cdot\|_0$ and $\|\cdot\|_2$ represent l_0-norm and l_2-norm.

II. Problem Formulation

In this paper, we consider a downlink multi-user mmWave massive MIMO system in which U single-antenna users are served by a BS. The BS transmits signals via a UPA in vertical plane, which has $N_{\rm v}$ vertical antennas and $N_{\rm h}$ horizontal antennas, or equivalently $N \triangleq N_{\rm v}N_{\rm h}$ antennas in total. Note that both $N_{\rm h}$ and $N_{\rm v}$ are even in practice. Let $\boldsymbol{h}_u \in \mathbb{C}^N$ denote the channels between the BS and the $u(u = 1, 2, \ldots, U)$th user. According to the popular 3D Saleh-Valenzuela channel model [4], $\boldsymbol{h}_u$ can be expressed as

$$\boldsymbol{h}_u = \sqrt{\frac{N}{L_u}}\sum_{i=1}^{L_u}\boldsymbol{h}_{u,i} = \sqrt{\frac{N}{L_u}}\sum_{i=1}^{L_u} g_{u,i}\boldsymbol{\alpha}(N_{\rm h},\theta_{u,i})\otimes\boldsymbol{\alpha}(N_{\rm v},\varphi_{u,i}) \tag{1}$$

where L_u and $g_{u,i}$ denote the total number of resolvable paths and channel gain for the ith path, respectively. Let $\Theta_{u,i}$ and $\Phi_{u,i}$ denote the physical azimuth and physical elevation of the ith path, respectively. Then $\theta_{u,i} = \frac{d_{\rm h}}{\lambda}\sin\Theta_{u,i}$ and $\varphi_{u,i} = \frac{d_{\rm v}}{\lambda}\sin\Phi_{u,i}$ [12], where $d_{\rm h}$ and $d_{\rm v}$ denote the interval between the antennas in horizontal and in vertical, respectively. Normally we set $d_{\rm h}$ and $d_{\rm v}$ as half of the wavelength of mmWave signal. The channel steering vector $\boldsymbol{\alpha}(N,\theta)$ in (1) is denoted as

$$\boldsymbol{\alpha}(N,\theta) = \frac{1}{\sqrt{N}}\left[1, e^{-j2\pi\theta}, ..., e^{-j2\pi\theta(N-1)}\right]^T \tag{2}$$

Suppose the BS has U RF chains, where each RF chain independently serves an user. The BS sends mutually orthogonal pilot sequences with length of U to different users, where each pilot sequence is repetitively sent for K times during the downlink channel estimation. As an example, we may set $K = 1$. With the commonly used blocking fading channel model,

978-1-5386-6119-2/18/$31.00 ©2018 IEEE

the channel parameters are assumed to be constant during $V \triangleq KU$ time slots. We define the pilot matrix as $\boldsymbol{P} \in \mathbb{C}^{U\times U}$, where the uth row of $\boldsymbol{P}$ denotes the pilot sequence sent to the uth user. For the $k(k = 1, 2, \ldots, K)$th transmission of pilot sequences, the BS adopts an analog precoder $\boldsymbol{F}_k \in \mathbb{C}^{U\times N}$ formed by phase shifter networks, where each entry of $\boldsymbol{F}_k$ has constant envelop. Different with data transmission, we do not consider the digital precoder during the channel estimation, where the digital precoder can be simply set to be an identity matrix. Then the $k(k = 1, 2, \ldots, K)$th pilot sequence received by the $u(u = 1, 2, \ldots, U)$th user is expressed as

$$\boldsymbol{y}_{u,k}^T = \boldsymbol{h}_u^T \boldsymbol{F}_k^T \boldsymbol{P} + \boldsymbol{n}_{u,k}^T, \tag{3}$$

where $\boldsymbol{n}_{u,k} \in \mathbb{C}^U$ denote an additive white Gaussian noise (AWGN) vector which satisfying $\boldsymbol{n}_{u,k} \sim \mathcal{CN}(\boldsymbol{0}, \sigma^2 \boldsymbol{I}_U)$. Because of the orthogonality of $\boldsymbol{P}$, i.e., $\boldsymbol{P}\boldsymbol{P}^{\boldsymbol{H}} = \boldsymbol{I}_U$, we multiply $\boldsymbol{y}_{u,k}^T$ with $\boldsymbol{P}^H$, obtaining

$$\boldsymbol{r}_{u,k} \triangleq (\boldsymbol{y}_{u,k}^T \boldsymbol{P}^H)^T = \boldsymbol{F}_k \boldsymbol{h}_u + \widetilde{\boldsymbol{n}}_{u,k} \tag{4}$$

where $\widetilde{\boldsymbol{n}}_{u,k} \triangleq (\boldsymbol{n}_{u,k}^T \boldsymbol{P}^H)^T$. Since the mmWave MIMO channel is typically sparse in beamspace [13], we consider the beamspace representation of $\boldsymbol{h}_u$. Define

$$\boldsymbol{G} = \boldsymbol{D}(N_{\mathrm{h}}) \otimes \boldsymbol{D}(N_{\mathrm{v}}) \tag{5}$$

where

$$\boldsymbol{D}(M) \triangleq [\boldsymbol{\alpha}(M, 0), \boldsymbol{\alpha}(M, 1/M), ..., \boldsymbol{\alpha}(M, (M-1)/M)]^H \tag{6}$$

denotes the standard DFT matrix of size M. Therefore, the beamspace channel for the uth user based on (1) can be represented as

$$\boldsymbol{h}_u^{\mathrm{b}} \triangleq \boldsymbol{G}\boldsymbol{h}_u \tag{7}$$

where the superscript "b" means the beamspace. Because of the orthogonality of $\boldsymbol{G}$, i.e., $\boldsymbol{G}\boldsymbol{G}^{\boldsymbol{H}} = \boldsymbol{I}_N$, we have

$$\boldsymbol{h}_u = \boldsymbol{G}^H \boldsymbol{h}_u^{\mathrm{b}}. \tag{8}$$

Then substituting (8) into (4) leads to

$$\boldsymbol{r}_{u,k} = \boldsymbol{C}_k \boldsymbol{h}_u^{\mathrm{b}} + \widetilde{\boldsymbol{n}}_{u,k} \tag{9}$$

where $\boldsymbol{C}_k = \boldsymbol{F}_k \boldsymbol{G}^H$. After the BS repetitively sends all pilot sequences for K times, we obtain $\boldsymbol{r}_{u,1}, \boldsymbol{r}_{u,2}, \ldots, \boldsymbol{r}_{u,K}$, which can be stacked together into a column vector as $\boldsymbol{r}_u \triangleq [\boldsymbol{r}_{u,1}^T, \boldsymbol{r}_{u,2}^T, \ldots, \boldsymbol{r}_{u,K}^T]^T$. Note that the channel vector $\boldsymbol{h}_u^{\mathrm{b}}$ keeps constant during the K times repetitive pilot transmission. We have

$$\boldsymbol{r}_u = \boldsymbol{C}\boldsymbol{h}_u^{\mathrm{b}} + \widetilde{\boldsymbol{n}}_u \tag{10}$$

where

$$\boldsymbol{C} \triangleq [\boldsymbol{C}_1^T, \boldsymbol{C}_2^T, \ldots, \boldsymbol{C}_K^T]^T, \tag{11}$$

$$\widetilde{\boldsymbol{n}}_u \triangleq [\widetilde{\boldsymbol{n}}_{u,1}^T, \widetilde{\boldsymbol{n}}_{u,2}^T, \ldots, \widetilde{\boldsymbol{n}}_{u,K}^T]^T. \tag{12}$$

Since the mmWave MIMO beamspace channel is sparse, i.e., $\boldsymbol{h}_u^{\mathrm{b}}$ is a sparse vector, (10) is essentially a problem of sparse recovery. However, regarding the fact that the number of antennas is finite, meaning that the beamspace resolution of $\boldsymbol{G}$ is finite, there is more or less channel power leakage effect. In fact, large number of antennas leads to less power leakage and vice versa. In this context, $\boldsymbol{h}_u^{\mathrm{b}}$ is not ideally sparse, i.e., $\boldsymbol{h}_u^{\mathrm{b}}$ has many small nonzero entries, which makes the sparse recovery problem more complicated as a result.

III. Modeling of Channel Power Leakage

To model the channel power leakage, we first analyze the leakage effect of beamspace channel. We define $\boldsymbol{\alpha}^b(N_{\mathrm{h}}, \theta_{u,i}) \triangleq \boldsymbol{D}(N_{\mathrm{h}})^T \boldsymbol{\alpha}(N_{\mathrm{h}}, \theta_{u,i})$ and $\boldsymbol{\alpha}^b(N_{\mathrm{v}}, \varphi_{u,i}) \triangleq \boldsymbol{D}(N_{\mathrm{v}})^T \boldsymbol{\alpha}(N_{\mathrm{v}}, \varphi_{u,i})$. According to (1), the beamspace representation of the $i(i = 1, 2, \ldots, L)$th path channel vector between the BS and the uth user, denoted as $\boldsymbol{h}_{u,i}^{\mathrm{b}}$ can be expressed as

$$\begin{aligned}\boldsymbol{h}_{u,i}^{\mathrm{b}} &= \boldsymbol{G}^T \boldsymbol{h}_{u,i} \\ &= \boldsymbol{D}(N_{\mathrm{h}})^T \otimes \boldsymbol{D}(N_{\mathrm{v}})^T g_{u,i} \boldsymbol{\alpha}(N_{\mathrm{h}}, \theta_{u,i}) \otimes \boldsymbol{\alpha}(N_{\mathrm{v}}, \varphi_{u,i}) \\ &= g_{u,i} \mathrm{vec}\left(\boldsymbol{\alpha}^b(N_{\mathrm{v}}, \varphi_{u,i}) \left(\boldsymbol{\alpha}^b(N_{\mathrm{h}}, \theta_{u,i})\right)^T\right).\end{aligned} \tag{13}$$

To better analyze the characteristic of $\boldsymbol{h}_{u,i}^{\mathrm{b}}$, we reshape $\boldsymbol{h}_{u,i}^{\mathrm{b}}$ into a matrix with N_{v} rows and N_{h} columns, which is denoted as $\boldsymbol{H}_{u,i}^{\mathrm{b}}$. $\boldsymbol{H}_{u,i}^{\mathrm{b}}$ can be expressed as

$$\boldsymbol{H}_{u,i}^{\mathrm{b}}[p, q] = \boldsymbol{h}_{u,i}^{\mathrm{b}}\left[(q-1)N_{\mathrm{v}} + p\right]. \tag{14}$$

where $\boldsymbol{H}_{u,i}^{\mathrm{b}}[p, q]$ denotes the entry at the $p(p = 1, 2, \ldots, N_{\mathrm{v}})$th row and the $q(q = 1, 2, \ldots, N_{\mathrm{h}})$th column of $\boldsymbol{H}_{u,i}^{\mathrm{b}}$.

According to [12], generally there will be four peaks being significantly higher in the main lobe than other peaks in side lobes. As shown in Fig. 1, the four strongest entries corresponding to these four peaks are marked in red. The entries within the selection of two adjacent rows and columns of $\boldsymbol{H}_{u,i}^{\mathrm{b}}[p, q]$, illustrated in a shape of DC, occupying a large proportion of the total channel power. As shown in Fig. 1, the DC includes the entries marked in blue and red. In fact, some entries around the four strongest entries are even stronger than the entries within the DC. For example, the entry marked in a green cross may be larger in amplitude than the entry marked in a blue cross. To be in detail, we substitute (13) into (14), having

$$\boldsymbol{H}_{u,i}^{\mathrm{b}}[p, q] = g_{u,i} \boldsymbol{\alpha}_p^b(N_{\mathrm{v}}, \varphi_{u,i}) \boldsymbol{\alpha}_q^b(N_{\mathrm{h}}, \theta_{u,i}) \tag{15}$$

where $\boldsymbol{\alpha}_i^b(M, \theta)$ representing the ith entry of the vector $\boldsymbol{\alpha}^b(M, \theta)$, can be expressed as

$$\boldsymbol{\alpha}_i^b(M, \theta) = \frac{1}{M} \sum_{m=0}^{M-1} e^{-j2\pi m\left(\frac{i-1}{M} + \theta\right)}. \tag{16}$$

If $\theta = (M+1)/(2M)$, we have

$$\left|\boldsymbol{\alpha}_{\frac{M}{2}}^b\left(M, \frac{M+1}{2M}\right)\right| = \left|\boldsymbol{\alpha}_{\frac{M}{2}+1}^b\left(M, \frac{M+1}{2M}\right)\right|. \tag{17}$$

Once $\varphi_{u,i} = (N_{\mathrm{v}} + 1)/(2N_{\mathrm{v}})$ and $\theta_{u,i} = (N_{\mathrm{h}} + 1)/(2N_{\mathrm{h}})$, the channel power leakage is the largest, which has the largest channel impairment for the transmitted signal. In this context, we denote the two-dimensional channel matrix as $\boldsymbol{H}_{\mathrm{w}}^{\mathrm{b}}$, where the postscript "w" means the worst case. According to (17), we have

$$\begin{aligned}&\left|\boldsymbol{H}_{\mathrm{w}}^{\mathrm{b}}[N_{\mathrm{h}}/2 + 1, N_{\mathrm{v}}/2 + 1]\right| = \left|\boldsymbol{H}_{\mathrm{w}}^{\mathrm{b}}[N_{\mathrm{h}}/2, N_{\mathrm{v}}/2]\right| \\ &= \left|\boldsymbol{H}_{\mathrm{w}}^{\mathrm{b}}[N_{\mathrm{h}}/2, N_{\mathrm{v}}/2 + 1]\right| = \left|\boldsymbol{H}_{\mathrm{w}}^{\mathrm{b}}[N_{\mathrm{h}}/2 + 1, N_{\mathrm{v}}/2]\right|\end{aligned} \tag{18}$$

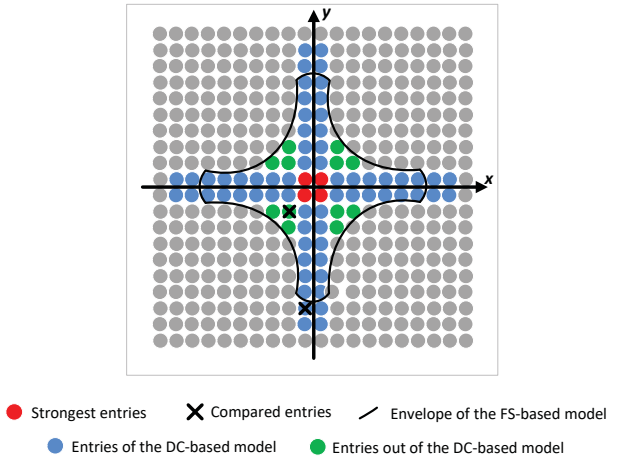

Fig. 1. Modeling of channel power leakage.

which indicating the four strongest entries marked in red in Fig. 1 are equally large in amplitude. The entry marked in the green cross is denoted as $\left|\boldsymbol{H}_{\mathrm{w}}^{\mathrm{b}}[N_{\mathrm{h}}/2-1, N_{\mathrm{v}}/2-1]\right|$. We denote the entry marked in the blue cross is denoted as $\left|\boldsymbol{H}_{\mathrm{w}}^{\mathrm{b}}[N_{\mathrm{h}}/2, z]\right|$, where z is a positive integer. In the following theorem, we will show that if z satisfies some conditions, the entry marked in the green cross is larger in amplitude than the entry marked in the blue cross, i.e., $\left|\boldsymbol{H}_{\mathrm{w}}^{\mathrm{b}}[N_{\mathrm{h}}/2-1, N_{\mathrm{v}}/2-1]\right| > \left|\boldsymbol{H}_{\mathrm{w}}^{\mathrm{b}}[N_{\mathrm{h}}/2, z]\right|$.

Theorem 1: If

$$z \leq \left\lfloor \frac{1}{2} + \frac{N_{\mathrm{h}}}{2\pi} \arccos\left(\frac{(1-\cos\frac{3\pi}{N_{\mathrm{h}}})(1-\cos\frac{3\pi}{N_{\mathrm{v}}})}{1-\cos\frac{\pi}{N_{\mathrm{h}}}} - 1 \right) \right\rfloor, \tag{19}$$

we have

$$\left|\boldsymbol{H}_{\mathrm{w}}^{\mathrm{b}}[N_{\mathrm{h}}/2-1, N_{\mathrm{v}}/2-1]\right| > \left|\boldsymbol{H}_{\mathrm{w}}^{\mathrm{b}}[N_{\mathrm{h}}/2, z]\right|. \tag{20}$$

Proof: If (19) is satisfied, we have

$$(1-\cos\frac{3\pi}{N_{\mathrm{h}}})(1-\cos\frac{3\pi}{N_{\mathrm{v}}}) < (1-\cos\frac{\pi}{N_{\mathrm{h}}})(1+\cos\frac{(2z-1)\,\pi}{N_{\mathrm{v}}}). \tag{21}$$

Note that we always have $|1-e^{j\theta}|^2 = 2-2\cos\theta$. Then (21) can be written as

$$\left|1-e^{j\frac{3\pi}{N_{\mathrm{h}}}}\right|\left|1-e^{j\frac{3\pi}{N_{\mathrm{v}}}}\right| < \left|1-e^{j\frac{\pi}{N_{\mathrm{h}}}}\right|\left|1-e^{-j\pi\frac{N_{\mathrm{v}}+2z-1}{N_{\mathrm{v}}}}\right|. \tag{22}$$

Since N_{v} is even, $N_{\mathrm{v}}+2z-1$ is odd, which means $e^{-j\pi(N_{\mathrm{v}}+2z-1)} = -1$. We take the reciprocal for both sides of (22), having

$$\left|\frac{1-e^{j3\pi}}{1-e^{j\frac{3\pi}{N_{\mathrm{h}}}}}\right|\left|\frac{1-e^{j3\pi}}{1-e^{j\frac{3\pi}{N_{\mathrm{v}}}}}\right| > \left|\frac{1-e^{j\pi}}{1-e^{j\frac{\pi}{N_{\mathrm{h}}}}}\right|\left|\frac{1-e^{-j\pi(N_{\mathrm{v}}+2z-1)}}{1-e^{-j\pi\frac{N_{\mathrm{v}}+2z-1}{N_{\mathrm{v}}}}}\right|. \tag{23}$$

Note that we have $\frac{1-\epsilon^M}{1-\epsilon} = \sum_{m=0}^{M-1}\epsilon^m$ for any $\epsilon \neq 1$. Then (23) can be rewritten as

$$\left|\sum_{m=0}^{N_{\mathrm{h}}-1} e^{j\frac{3\pi m}{N_{\mathrm{h}}}}\right|\left|\sum_{m=0}^{N_{\mathrm{v}}-1} e^{j\frac{3\pi m}{N_{\mathrm{v}}}}\right| > \left|\sum_{m=0}^{N_{\mathrm{h}}-1} e^{j\frac{\pi m}{N_{\mathrm{h}}}}\right|\left|\sum_{m=0}^{N_{\mathrm{v}}-1} e^{-j\pi m\frac{N_{\mathrm{v}}+2z-1}{N_{\mathrm{v}}}}\right|. \tag{24}$$

Based on (16), we have

$$\left|\boldsymbol{\alpha}_{\frac{N_{\mathrm{h}}}{2}-1}^{b}\left(N_{\mathrm{h}}, \frac{N_{\mathrm{h}}+1}{2N_{\mathrm{h}}}\right)\right|\left|\boldsymbol{\alpha}_{\frac{N_{\mathrm{v}}}{2}-1}^{b}\left(N_{\mathrm{v}}, \frac{N_{\mathrm{v}}+1}{2N_{\mathrm{v}}}\right)\right| > \left|\boldsymbol{\alpha}_{\frac{N_{\mathrm{h}}}{2}}^{b}\left(N_{\mathrm{h}}, \frac{N_{\mathrm{h}}+1}{2N_{\mathrm{h}}}\right)\right|\left|\boldsymbol{\alpha}_{z}^{b}\left(N_{\mathrm{v}}, \frac{N_{\mathrm{v}}+1}{2N_{\mathrm{v}}}\right)\right|. \tag{25}$$

According to (15) and (25), we can get (20). ■

For example, suppose $N_{\mathrm{h}} = N_{\mathrm{v}} = 32$. If $z \leq 11$, we have $|\boldsymbol{H}_{\mathrm{w}}^{\mathrm{b}}[15,15]| > |\boldsymbol{H}_{\mathrm{w}}^{\mathrm{b}}[16,z]|$.

Remark 1: For the entry selection, it is better to include the entry marked in the green cross than that marked in the blue cross, since the former is larger in amplitude than the latter. In fact, the entries marked in green without the cross are also large. However, in [12] which proposes a DC-based beamspace channel estimation method, these green entries are neglected. Therefore, in this work, we propose a FS-based beamspace channel estimation method, where the envelop of the FS-based model is marked in black in Fig. 1.

IV. FS-based Beamspace Channnel Estimation

As shown in Fig. 1, the envelop of the FS-based model is similar to the hyperbola, we use the following expression to represent the envelop of the FS-based model as

$$|y| = \frac{c}{|x|^a} \tag{26}$$

where c and a are real positive parameters to be determined in terms of N_{h}, N_{v} and η. $\eta \in (0,1]$ is defined as the power ratio of the entries within the envelop over all the entries. The x-axis and y-axis are illustrated in Fig. 1. In fact, a is a parameter related to the number of antennas in horizontal and vertical. If $N_{\mathrm{h}} > N_{\mathrm{v}}$, the figure is a flat rectangle instead of a square in Fig. 1, which implies we should set $a > 1$. If $N_{\mathrm{h}} < N_{\mathrm{v}}$, the figure is a tall rectangle, where we should set $0 < a < 1$. If $N_{\mathrm{h}} = N_{\mathrm{v}}$, we may simply set $a = 1$. Given N_{h}, N_{v} and a, larger η leads to larger c, which indicates that the envelop of the FS-based model is larger and more entries need to be selected.

Given N_{h}, N_{v} and η, we can search a and c to find appropriate parameters for the FS-based model. We denote the starting position, the stop position and the search interval for the parameter a as a_{D}, a_{U}, Δ_a, respectively. Similarly, we denote the starting position, the stop position and the search interval for the parameter c as c_{D}, c_{U}, Δ_c, respectively. Given a pair of a and c, we obtain a set of entries within the envelop of the FS-based model as

$$\boldsymbol{\Omega} = \Big\{(p,q)\Big|\left|p-\frac{N_{\mathrm{v}}+1}{2}\right|\left|q-\frac{N_{\mathrm{h}}+1}{2}\right|^a \leq c \quad p=1,2,\ldots,N_{\mathrm{v}},\ q=1,2,\ldots,N_{\mathrm{h}}\Big\}. \tag{27}$$

Then we compute the power ratio of the entries in $\boldsymbol{\Omega}$ over all the entries as

$$\tilde{\eta} = \sum_{(p,q)\in\boldsymbol{\Omega}} \left|\boldsymbol{H}_{\mathrm{w}}^{\mathrm{b}}[p,q]\right|^2 \Big/ \sum_{p=1}^{N_{\mathrm{v}}}\sum_{q=1}^{N_{\mathrm{h}}} \left|\boldsymbol{H}_{\mathrm{w}}^{\mathrm{b}}[p,q]\right|^2. \tag{28}$$

We iteratively search for a set of entries with $\tilde{\eta} \geq \eta$ while the number of entries in the set is minimized. After running steps

Algorithm 1 FS-based Beamspace Channel Estimation

1: *Input:* $\boldsymbol{r}_u$, $\hat{L}_u$, N_v, N_h, η, $\boldsymbol{C}$, a_D, a_U, Δ_a, c_D, c_U, Δ_c.
2: Initialization: $J \leftarrow N$. $\boldsymbol{\Gamma} \leftarrow \emptyset$.
3: **for** $a = a_\mathrm{D} : \Delta_a : a_\mathrm{U}$ **do**
4: **for** $c = c_\mathrm{D} : \Delta_c : c_\mathrm{U}$ **do**
5: Obtain $\boldsymbol{\Omega}$ via (27).
6: Obtain $\tilde{\eta}$ via (28).
7: **if** $\tilde{\eta} \geq \eta$ and $\|\boldsymbol{\Omega}\|_0 < J$ **then**
8: $\boldsymbol{\Gamma} \leftarrow \boldsymbol{\Omega}$. $J \leftarrow \|\boldsymbol{\Omega}\|_0$.
9: **end if**
10: **end for**
11: **end for**
12: $\boldsymbol{r}_u^{(1)} \leftarrow \boldsymbol{r}_u$. $\hat{\boldsymbol{h}}_u^{\mathrm{b}} \leftarrow \boldsymbol{0}^N$.
13: **for** $i = 1, 2, \ldots, \hat{L}_u$ **do**
14: Obtain $\zeta_p^{(i)}$ and $\zeta_q^{(i)}$ via (30) and (31), respectively.
15: Obtain the dominant entries $\boldsymbol{\Psi}_i$ via (32).
16: **repeat**
17: $\hat{\boldsymbol{h}}_{u,i}^{\mathrm{b}} \leftarrow \boldsymbol{0}^N$. Obtain $\hat{\boldsymbol{h}}_{u,i}^{\mathrm{b}}[\boldsymbol{\Psi}_i]$ via (33).
18: Update $\hat{\boldsymbol{H}}_{u,i}^{b}$ based on $\hat{\boldsymbol{h}}_{u,i}^{\mathrm{b}}[\boldsymbol{\Psi}_i]$ according to (14).
19: Adjust the dominant entries $\boldsymbol{\Psi}_i$.
20: **until** *Stop Condition* is met.
21: Calculate $\boldsymbol{r}_u^{(i+1)}$ via (29).
22: **end for**
23: $\boldsymbol{\Upsilon}_u = \bigcup_{i=1,2,\ldots,\hat{L}_u} \boldsymbol{\Psi}_i$.
24: Obtain $\hat{\boldsymbol{h}}_u^{\mathrm{b}}[\boldsymbol{\Upsilon}_u]$ via (33).
25: *Output:* $\hat{\boldsymbol{h}}_u^{\mathrm{b}}$.

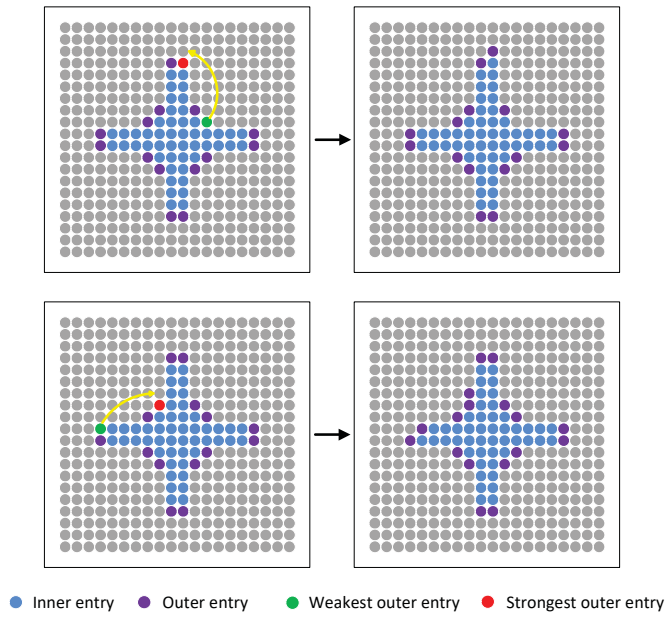

Fig. 2. Adjustment of the FS-based algorithm.

from step 3 to step 11, we obtain a best entry set $\boldsymbol{\Gamma}$, which is initialized to be an empty set in step 2.

From step 12 to step 24, we use $\boldsymbol{r}_u^{(i)}$ to get an estimation of $\boldsymbol{h}_{u,i}^{\mathrm{b}}$, which is denoted as $\hat{\boldsymbol{h}}_{u,i}^{\mathrm{b}}$, $i = 1, 2, \ldots, \hat{L}_u$. We initialize $\boldsymbol{r}_u^{(1)}$ to be $\boldsymbol{r}_u$, as indicated by step 12. Since the genuine L_u is unknown in practice, we can only use an estimate of it as $\hat{L}_u$.

At step 17, we compute $\boldsymbol{r}_u^{(i+1)}$ based on $\boldsymbol{r}_u^{(i)}$ via

$$\boldsymbol{r}_u^{(i+1)} \leftarrow \boldsymbol{r}_u^{(i)} - \boldsymbol{C}\hat{\boldsymbol{h}}_{u,i}^{\mathrm{b}}[\boldsymbol{\Psi}_i]. \tag{29}$$

where $\boldsymbol{\Psi}_i$ is a set including the position information of dominant entries for the ith channel path. We obtain $\boldsymbol{\Psi}_i$ as follows. First, we find out two adjacent rows numbered as $\{\zeta_p^{(i)}, \zeta_p^{(i)} + 1\}$, whose channel power is the largest in $\boldsymbol{H}_{u,i}^{\mathrm{b}}$ via

$$\zeta_p^{(i)} = \arg \max_{p=1,2,\ldots,N_v-1} \left\| (\boldsymbol{C}_{(p)}^H \boldsymbol{C}_{(p)})^{-1} \boldsymbol{C}_{(p)}^H \boldsymbol{r}_u^{(i)} \right\|_2 \tag{30}$$

and two adjacent columns numbered as $\{\zeta_q^{(i)}, \zeta_q^{(i)} + 1\}$, whose channel power is the largest in $\boldsymbol{H}_{u,i}^{\mathrm{b}}$ via

$$\zeta_q^{(i)} = \arg \max_{q=1,2,\ldots,N_h-1} \left\| (\boldsymbol{C}_{(q)}^H \boldsymbol{C}_{(q)})^{-1} \boldsymbol{C}_{(q)}^H \boldsymbol{r}_u^{(i)} \right\|_2 \tag{31}$$

where $\boldsymbol{C}_{(q)}$ and $\boldsymbol{C}_{(p)}$ represents a submatrix of $\boldsymbol{C}$ consisted of $2N_\mathrm{v}$ consecutive columns and $2N_\mathrm{h}$ inconsecutive columns, respectively, with the column numbers respectively denoted as $(q-1)N_\mathrm{v}+1, (q-1)N_\mathrm{v}+2, \ldots, (q-1)N_\mathrm{v}+N_\mathrm{v}, qN_\mathrm{v}+1, qN_\mathrm{v}+2, \ldots, qN_\mathrm{v}+N_\mathrm{v}$ and $p, p+1, p+N_\mathrm{v}, p+1+N_\mathrm{v}, p+2N_\mathrm{v}, p+1+2N_\mathrm{v}, \ldots, p+(N_\mathrm{h}-1)N_\mathrm{v}, p+1+(N_\mathrm{h}-1)N_\mathrm{v}$. Then we can obtain $\boldsymbol{\Psi}_i$ by

$$\boldsymbol{\Psi}_i = \left\{ \left(q - \frac{N_\mathrm{h}}{2} + \zeta_q^{(i)} - 1\right) N_\mathrm{v} + p - \frac{N_\mathrm{v}}{2} + \zeta_p^{(i)}, (p,q) \in \boldsymbol{\Gamma} \right\}. \tag{32}$$

From step 16 to step 20, we iteratively execute the procedures to optimize $\boldsymbol{\Psi}_i$ while ensuring the FS shape at the same time. We estimate the dominant entries of $\hat{\boldsymbol{h}}_{u,i}^{\mathrm{b}}$ by the least squares (LS) method as

$$\hat{\boldsymbol{h}}_{u,i}^{\mathrm{b}}[\boldsymbol{\Psi}_i] = (\boldsymbol{C}_{\Psi_i}^H \boldsymbol{C}_{\Psi_i})^{-1} \boldsymbol{C}_{\Psi_i}^H \boldsymbol{r}_u^{(i)} \tag{33}$$

where $\boldsymbol{C}_{\Psi_i}$ represents a submatrix which is made up of columns with column numbers $\boldsymbol{\Psi_i}$ from $\boldsymbol{C}$. Based on the entries of $\hat{\boldsymbol{h}}_{u,i}^{\mathrm{b}}[\boldsymbol{\Psi}_i]$, we can update the entries of $\hat{\boldsymbol{H}}_{u,i}^{b}$ according to (14) at step 18. As shown in Fig 2, we consider the entries at the eight corners, which are at the top, bottom, left, right, top right, top left, bottom right and bottom left directions respectively, as the outer entries. At step 19, we first delete the outer entry with the weakest power. Then we add a new entry outside of the strongest outer entry. If the strongest outer entry is at the top, bottom, right or left directions, we add a new entry along the row or the column of the strongest outer entry. If the strongest outer entry is at the rest four directions, there are two candidate entries either on the row direction or column direction, where we select the stronger entry among these two candidate entries.

Definition 1 (Stop Condition): The entry to be added or the entry to be deleted is the entry just deleted or added, respectively.

In fact, the adjustment of the dominant entries ensures that the total power of the dominant entries is keeping increasing. When the *Stop Condition* is satisfied, the total power of the dominant entries is the largest, which best approaches the real channel.

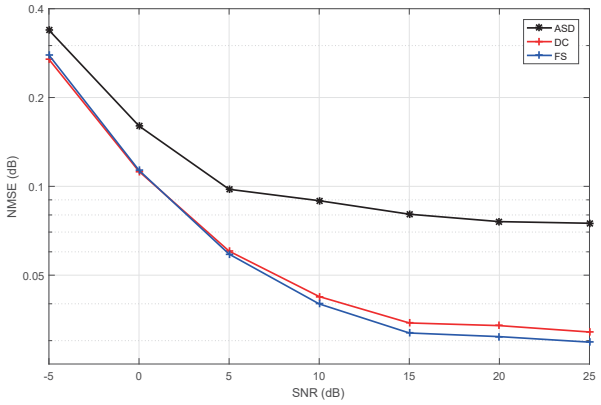

Fig. 3. Comparisons of different channel estimation algorithms ($N_{\rm h} = 32, N_{\rm v} = 32$).

After the positions of the dominant entries have been calculated for all of $\hat{L}_u$ channel paths, we combine all of them at step 23. We obtain $\hat{\boldsymbol{h}}_u^{\rm b}[\boldsymbol{\Upsilon}_u]$ via (33) by replacing $\boldsymbol{\Psi}_i$, $\hat{\boldsymbol{h}}_{u,i}^{\rm b}$ and $\boldsymbol{r}_u^{(i)}$ with $\boldsymbol{\Upsilon}_u$, $\hat{\boldsymbol{h}}_u^{\rm b}$ and $\boldsymbol{r}_u$, respectively. Finally, we output the channel estimation $\hat{\boldsymbol{h}}_u^{\rm b}$ for the uth user.

It should be mentioned that we don't have to consider the computational complexity of step 3 to step 11 because they can be operated offline.

V. Simulation Results

We set $N_{\rm h} = N_{\rm v} = 32$, $U = 16$, $V = 256$, and $\hat{L}_u = 3$, where $g_{u,1} \sim \mathcal{CN}(0,1)$ and $g_{u,i} \sim \mathcal{CN}(0, 0.01)$ for $i = 2, 3$. We set $\eta = 0.93$ and correspondingly $J = 56$. We set $a_{\rm D} = 0.5$, $a_{\rm U} = 2$, $\Delta_a = 0.01$, $c_{\rm D} = 0$, $c_{\rm U} = 100$ and $\Delta_c = 0.01$ so that we can ensure the selected $\boldsymbol{\Gamma}$ approaches the best choice. Downlink combiner matrix $\boldsymbol{C}$ is supposed to be a matrix obeying Bernoulli random distribution, in which all of its entries satisfying bivariate uniform distribution $\frac{1}{V}\{-1, +1\}$ independently.

As shown in Fig. 3, we make comparisons of different channel estimation algorithms for their normalized mean squared error (NMSE), including the FS-based channel estimation algorithm, DC-based channel estimation algorithm [12] and ASD-based channel estimation algorithm [11]. We can find that in low signal-to-noise ratio (SNR) region, the result of the FS-based channel estimation algorithm is similar as that of the DC-based channel estimation algorithm. In high SNR region, the FS-based channel estimation algorithm enjoys a better performance than that of DC-based channel estimation algorithm. At SNR=20dB, the result of FS-based channel estimation algorithm has more than 8% improvement than that of the DC-based channel estimation algorithm.

The comparisons of computational complexity for these three algorithms are shown in Table I. The average number of iterations to run the three algorithms for different SNR is listed from the first row to the third row of Table I, respectively. The fourth row is the ratio of average iterations of the FS-based algorithm over the DC-based algorithm, which essentially equals $((2VJ^2+4J^3)+(t_1-1)(4VJ+10J^2+2J))/((2VJ^2+4J^3)+(t_2-1)(4VJ+10J^2+2J))$, where t_1 and t_2 are the entries in the first row and the second row, respectively. The fifth row is the ratio of average iterations of the FS-based algorithm over the ASD-based algorithm, which essentially equals $((2VJ^2+4J^3)+(t_1-1)(4VJ+10J^2+2J))/(2VJ^2t_3+4J^3t_3)$, where t_1 and t_3 are the entries in the first row and the third row, respectively. According to Table I, the computational complexity of the FS-based algorithm is much lower than the ASD-based algorithm. Although the computational complexity of the FS-based algorithm is a bit higher than that of the DC-based algorithm, the channel estimation performance of the former is better than the latter, especially in the high SNR region.

TABLE I
Comparisons of computational complexity for different algorithms.

SNR(dB)	-5	0	5	10	15	20	25
FS	8.376	9.135	10.486	13.068	14.439	14.446	14.961
DC	3.138	3.357	3.969	5.276	7.257	8.549	9.113
ASD	2.398	2.384	2.383	2.395	2.387	2.390	2.385
FS over DC	118.6%	120.4%	122.5%	125.7%	122.2%	117.6%	117.1%
FS over ASD	53.5%	55.1%	57.3%	61.1%	63.6%	63.5%	64.5%

TABLE II
Ratio of the FS-based algorithm over the DC-based algorithm in terms of NMSE for different $N_{\rm h}$ and $N_{\rm v}$.

SNR(dB)	-5	0	5	10	15	20	25
64×64	103.4%	98.4%	95.6%	95.1%	93.7%	94.1%	92.6%
64×32	104.7%	103.2%	98.0%	95.5%	93.7%	93.2%	92.5%
64×16	105.6%	103.3%	101.9%	98.1%	96.6%	94.4%	95.1%
32×32	103.4%	100.9%	97.7%	94.4%	92.5%	91.8%	92.4%
32×16	106.3%	108.3%	103.4%	99.9%	98.1%	96.6%	94.8%

We also consider mmWave massive MIMO systems with some other BS antenna configuration, including $N_{\rm h} = 64$ and $N_{\rm v} = 64$, $N_{\rm h} = 64$ and $N_{\rm v} = 32$, $N_{\rm h} = 64$ and $N_{\rm v} = 16$, $N_{\rm h} = 32$ and $N_{\rm v} = 32$, $N_{\rm h} = 32$ and $N_{\rm v} = 16$. As shown in Table II, we provide the ratio of the FS-based algorithm over the DC-based algorithm in terms of NMSE for different $N_{\rm h}$ and $N_{\rm v}$ and different SNR. It is seen that as the SNR increases, the channel estimation performance of the FS-based algorithm gets better than that of the DS-based algorithm. The reason is that the FS-based model can better approach the real channel with power leakage than the DC-based model, especially in the high SNR region. We also observe that the FS-based model performs well with the square BS antenna configuration, e.g., $N_{\rm h} = 64$ and $N_{\rm v} = 64$, $N_{\rm h} = 32$ and $N_{\rm v} = 32$, since the parameters of the FS-based model expressed in (26) in square configuration can be better figured out than the other configuration.

VI. Conclusions

In this paper, we have investigated the channel estimation concerning channel power leakage for mmWave massive MIMO system. We have analyze the leakage effect of two dimensional beamspace channel and have provided a theorem to show that the FS-based model is better than the

existing DC-based model. We have proposed the FS-based beamspace channel estimation algorithm, where the initial selection of dominant entries is computed offline and then the selection is iteratively refined until the stop condition is satisfied. Simulation results have verified the effectiveness of our work and have shown that the FS-based channel estimation algorithm outperforms the DC-based channel estimation algorithm.Future work will focus on other methods to reduce the channel power leakage.

ACKNOWLEDGMENT

This work is supported in part by the National Natural Science Foundation of China under Grants 61871119 and 61302097, and in part by the Natural Science Foundation of Jiangsu Province under Grant BK20161428.

REFERENCES

[1] Y. Zeng and R. Zhang, "Cost-effective millimeter-wave communications with lens antenna array," *IEEE Commun. Lett.*, vol. 24, no. 4, pp. 81–87, 2017.

[2] V. W.S.Wong, R. Schober, D. W. K. Ng, and L.-C. Wang, *Key Technologies for 5G Wireless Systems*. Cambridge University Press, 2017.

[3] T. S. Rappaport, Y. Xing, G. R. Maccartney, A. F. Molisch, E. Mellios, and J. Zhang, "Overview of millimeter wave communications for fifth-generation (5G) wireless networks-with a focus on propagation models," *IEEE Trans. Antennas Propag.*, vol. 65, no. 12, pp. 6213–6230, 2017.

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

[5] Y. Peng, Y. Li, and P. Wang, "An enhanced channel estimation method for millimeter wave systems with massive antenna arrays," *IEEE Commun. Lett.*, vol. 19, no. 9, pp. 1592–1595, Sep. 2015.

[6] I. Bahceci, M. Hasan, T. M. Duman, and B. A. Cetiner, "Efficient channel estimation for reconfigurable MIMO antennas: Training techniques and performance analysis," *IEEE Trans. Antennas Propag.*, vol. 16, no. 1, pp. 565–580, 2017.

[7] L. Zhao, G. Geraci, T. Yang, D. W. K. Ng, and J. Yuan, "A tone-based aoa estimation and multiuser precoding for millimeter wave massive mimo," *IEEE Transactions on Communications*, vol. 65, no. 12, pp. 5209–5225, 2017.

[8] W. Ma and C. Qi, "Beamspace channel estimation for millimeter wave massive mimo system with hybrid precoding and combining," *IEEE Transactions on Signal Processing*, vol. 66, no. 18, pp. 4839–4853, 2018.

[9] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. Sel. Top. Signal. Process.*, vol. 8, no. 5, pp. 831–846, Oct. 2014.

[10] 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," in *Proc. IEEE/CIC Chengdu, China*, 2016, pp. 1–6.

[11] X. Gao, L. Dai, S. Han, C.-L. I, 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.

[12] W. Ma and C. Qi, "Channel estimation for 3-D lens millimeter wave massive MIMO system," *IEEE Commun. Lett.*, vol. PP, no. 99, pp. 1–1, 2017.

[13] J. Brady, N. Behdad, and A. 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.