

ESPRIT-based Channel Estimation for Frequency-Selective Millimeter Wave Massive MIMO System

Wenyan Ma and Chenhao Qi
School of Information Science and Engineering
Southeast University, Nanjing 210096, China
Email: qch@seu.edu.cn

Abstract—Channel estimation for frequency-selective millimeter wave (mmWave) massive MIMO system is investigated. To overcome the frequency-selective fading, orthogonal frequency division multiplexing (OFDM) is employed. First, the channel structure of frequency-selective channel is analyzed to show that all the OFDM subcarriers share the same angles of arrival (AoA) and angles of departure (AoD). Then a two dimensional ESPRIT (TDE)-based channel estimation scheme is proposed, where the super-resolution estimation of AoA and AoD can be obtained by utilizing the rotation invariance of the channel steering vectors. Finally the AoA and AoD are paired to reconstruct the channel. Simulation results show that the proposed TDE-based channel estimation scheme outperforms the existing schemes at high SNR region.

Index Terms—Millimeter wave (mmWave) communications, channel estimation, hybrid precoding, massive MIMO.

I. INTRODUCTION

Millimeter wave (mmWave) communication is a promising technology for next generation wireless communication owing to its abundant frequency spectrum resource [1]. Due to the large number of antennas at both transmitter and receiver sides, the size of channel matrix is very large, therefore the channel estimation is rather time consuming. Moreover, recently it is found that the mmWave channel is frequency-selective [2], which brings more challenges for channel estimation.

Some approaches for channel estimation in frequency-selective mmWave channels have been proposed. An orthogonal matching pursuit (OMP)-based channel estimation scheme is proposed in [3], which explores the sparsity of beamspace channel by incorporating sparse signal processing techniques. Moreover, a simultaneous weighted orthogonal matching pursuit (SWOMP)-based channel estimation scheme is proposed in [2], where the spatial noise components are whitened to improve the estimation accuracy using the OMP method. However, considering the limited beamspace resolution, the sparsity of beamspace channel may be impaired by power leakage, indicating that the beamspace channel is not ideally sparse and there are many small nonzero entries [4]. To solve this problem, a distributed grid matching pursuit (DGMP)-based channel estimation scheme is proposed to iteratively detect and adjust the channel support [5]. However, it only estimates a single path, leading to large estimation error for

multi-path channels.

To improve the channel estimation accuracy, some super-resolution channel estimation schemes have been proposed. However, these schemes assume that the channels are with flat-fading. An ESPRIT-based and a MUSIC-based channel estimation schemes are proposed in [6] and [7], respectively, where the ESPRIT and MUSIC methods are directly applied for mmWave channel estimation with frequency-flat channels. However, in order to directly use the ESPRIT and MUSIC methods, these two schemes have to turn off approximately half of the antennas so that the number of powered antennas is equal to that of time slots for channel estimation, which will reduce the total transmission power and signal coverage. Moreover, they do not consider the frequency-selective mmWave channels.

In this paper, we investigate the channel estimation for frequency-selective mmWave massive MIMO system. To overcome the frequency-selective fading, orthogonal frequency division multiplexing (OFDM) is employed. First, we analyze the channel structure of frequency-selective channel to show that all the OFDM subcarriers share the same angles of arrival (AoA) and angles of departure (AoD). Then we propose a two dimensional ESPRIT (TDE)-based channel estimation scheme, where the super-resolution estimation of the AoA and AoD are obtained by utilizing the rotation invariance of the channel steering vectors. Finally the AoA and AoD are paired to reconstruct the channel. Note that unlike the existing schemes, the TDE-based scheme only needs to turn off one antenna, which has little impact on the total transmission power.

The notations are defined as follows. Symbols for matrices (upper case) and vectors (lower case) are in boldface. $(\cdot)^T$, $(\cdot)^H$, $(\cdot)^*$, $(\cdot)^{-1}$, $(\cdot)^\dagger$, \mathbf{I}_L , $\mathbf{1}_L$, $\mathbb{C}^{M \times N}$, \otimes , \circ , $\text{vec}(\cdot)$, $\mathbb{E}\{\cdot\}$, $\mathcal{O}(\cdot)$, $\mathbf{0}^M$, $\mathbf{A}[p, q]$, \mathbb{Z} and \mathcal{CN} denote the transpose, conjugate transpose (Hermitian), conjugate, inverse, pseudo inverse, identity matrix of size L , vector of size L with all entries being 1, the set of $M \times N$ complex-valued matrices, Kronecker product, Khatri-Rao product, vectorization, the square diagonal matrix with the elements of vector \mathbf{a} on the main diagonal, expectation, order of complexity, zero vector of size M , entry of \mathbf{A} at the p th row and q th column, set of integer and complex Gaussian distribution, respectively.

II. SYSTEM MODEL AND PROBLEM FORMULATION

A. System Model

We consider an uplink multi-user mmWave massive MIMO system comprising a base station (BS) and U users. OFDM modulation with K subcarriers is employed to deal with the frequency-selective fading channels [2]. Both the BS and users are equipped with uniform linear arrays (ULAs). Let N_A , M_A , N_R and M_R denote the number of antennas at the BS, number of antennas at each user, number of RF chains at the BS and number of RF chains at each user, respectively. For uplink transmission, each user performs analog precoding in RF and digital precoding in the baseband, while the BS performs analog combining in RF and digital combining in the baseband. The received signal vector at the BS for the $k(k = 0, 1, \dots, K-1)$ th OFDM subcarrier is represented as

$$\mathbf{y}^k = \mathbf{W}_B^k \mathbf{W}_R \sum_{u=1}^U \mathbf{H}_u^k \mathbf{F}_{R,u} \mathbf{F}_{B,u}^k \mathbf{s}_u^k + \mathbf{W}_B^k \mathbf{W}_R \mathbf{n} \quad (1)$$

where $\mathbf{F}_{B,u}^k \in \mathbb{C}^{M_R \times M_R}$, $\mathbf{F}_{R,u} \in \mathbb{C}^{M_A \times M_R}$, $\mathbf{W}_B^k \in \mathbb{C}^{N_R \times N_R}$, and $\mathbf{W}_R \in \mathbb{C}^{N_R \times N_A}$ are the digital precoding matrix, analog precoding matrix, digital combining matrix, and analog combining matrix for the $u(u = 1, 2, \dots, U)$ th user, respectively. Note that the analog precoding and combining matrices are frequency-flat, while the digital precoding and combining matrices are different for every subcarrier [2]. $\mathbf{s}_u^k \in \mathbb{C}^{M_R}$ denotes the signal vector satisfying $\mathbb{E}\{\mathbf{s}_u^k \mathbf{s}_u^{kH}\} = \mathbf{I}_{M_R}$. $\mathbf{n} \in \mathbb{C}^{N_A}$ denotes additive white Gaussian noise (AWGN) vector satisfying $\mathbf{n} \sim \mathcal{CN}(0, \sigma^2 \mathbf{I}_{N_A})$. $\mathbf{H}_u^k \in \mathbb{C}^{N_A \times M_A}$ denotes the channel matrix at the k th subcarrier between the BS and the u th user and can be expressed as

$$\mathbf{H}_u^k = \sum_{d=0}^{D-1} \mathbf{H}_{u,d} e^{-j \frac{2\pi k d}{K}} \quad (2)$$

where D denotes the number of delay taps of the channel. According to the widely used Saleh-Valenzuela channel model [1], the channel matrix at the $d(d = 0, 1, \dots, D-1)$ th delay tap is defined as

$$\mathbf{H}_{u,d} = \gamma \sum_{i=1}^{L_u} g_{u,i} p_{rc,u}(dT_s - \tau_{u,i}) \boldsymbol{\alpha}(N_A, \theta_{u,i}) \boldsymbol{\alpha}^H(M_A, \phi_{u,i}) \quad (3)$$

where L_u , $g_{u,i}$, $p_{rc,u}(t)$, T_s and $\tau_{u,i}$ denote the total number of resolvable paths, the channel gain, the pulse shaping, the sampling interval and the delay of the $i(i = 1, 2, \dots, L_u)$ th path for the u th user, respectively. In practice, $\tau_{u,i}$ obeys the uniform distribution $[0, (K-1)T_s]$. Define $\gamma \triangleq \sqrt{N_A M_A / L_u}$. Further define the steering vector $\boldsymbol{\alpha}(N, \theta)$ as

$$\boldsymbol{\alpha}(N, \theta) = \frac{1}{\sqrt{N}} [1, e^{j\pi\theta}, \dots, e^{j\pi\theta(N-1)}]^T. \quad (4)$$

Define the AoA and AoD of the i th path of the u th user as $\vartheta_{u,i}$ and $\varphi_{u,i}$, respectively. Further define $\theta_{u,i} \triangleq 2d_{BS} \sin \vartheta_{u,i} / \lambda$ and $\phi_{u,i} \triangleq 2d_{UE} \sin \varphi_{u,i} / \lambda$, where d_{BS} and d_{UE} denote the antenna interval of the BS and users, respectively. We

usually set $d_{BS} = d_{UE} = \lambda/2$, where λ is the wavelength of mmWave signal. In practice, both $\vartheta_{u,i}$ and $\varphi_{u,i}$ obey the uniform distribution $[-\pi, \pi]$.

B. Problem Formulation

Note that \mathbf{y}^k in (1) is a combination of signal from different users. We use T_1 different digital precoding matrices and analog precoding matrices, denoted as $\mathbf{F}_{B,u}^{k,t_1} \in \mathbb{C}^{M_R \times M_R}$ and $\mathbf{F}_{R,u}^{t_1} \in \mathbb{C}^{M_A \times M_R}$, respectively, $t_1 = 1, 2, \dots, T_1$, at the $u(u = 1, 2, \dots, U)$ th user. We use T_2 different digital combining matrices and analog combining matrices, denoted as $\mathbf{W}_B^{k,t_2} \in \mathbb{C}^{N_R \times N_R}$ and $\mathbf{W}_R^{t_2} \in \mathbb{C}^{N_R \times N_A}$, respectively, $t_2 = 1, 2, \dots, T_2$, at the BS. To distinguish different user signal at the BS, each user repeatedly transmits an orthogonal pilot sequence $\mathbf{p}_u^k \in \mathbb{C}^U$ for $T_1 T_2$ times. For simplicity, we suppose each user transmits the same pilot sequence for all M_R RF chains, where the pilot matrix for the u th user can be defined as $\mathbf{P}_u^k \triangleq [\mathbf{p}_u^k, \mathbf{p}_u^k, \dots, \mathbf{p}_u^k]^H = \mathbf{1}_{M_R} \mathbf{p}_u^{kH} \in \mathbb{C}^{M_R \times U}$. The channel keeps constant during $T \triangleq T_1 T_2 U$ time slots [4]. During the T_1 repetitive transmission of pilot sequence from the $((t_2-1)T_1+1)$ th transmission to $(t_2 T_1)$ th transmission, we use T_1 different $\mathbf{F}_{B,u}^{k,t_1}$ and $\mathbf{F}_{R,u}^{t_1}$ for hybrid precoding while use the same \mathbf{W}_B^{k,t_2} and $\mathbf{W}_R^{t_2}$ for hybrid combining, where the received pilot matrix $\mathbf{Y}^{k,t_1,t_2} \in \mathbb{C}^{N_R \times U}$ is denoted as

$$\mathbf{Y}^{k,t_1,t_2} = \mathbf{W}_B^{k,t_2} \mathbf{W}_R^{t_2} \sum_{u=1}^U \mathbf{H}_u^k \mathbf{F}_{R,u}^{t_1} \mathbf{F}_{B,u}^{k,t_1} \mathbf{P}_u^k + \widetilde{\mathbf{N}}^{t_1,t_2} \quad (5)$$

with $\widetilde{\mathbf{N}}^{t_1,t_2} \triangleq \mathbf{W}_B^{k,t_2} \mathbf{W}_R^{t_2} \mathbf{N}^{t_1,t_2} \in \mathbb{C}^{N_R \times U}$. Each entry of the AWGN matrix $\mathbf{N}^{t_1,t_2} \in \mathbb{C}^{N_A \times U}$ independently obeys complex Gaussian distribution with zero mean and variance of σ^2 . Due to the orthogonality of \mathbf{p}_u^k , i.e., $\mathbf{p}_u^{kH} \mathbf{p}_u^k = 1$ and $\mathbf{p}_u^{kH} \mathbf{p}_i^k = 0, \forall u, i \in \{1, 2, \dots, U\}, i \neq u$ [8], we can obtain the measurement vector $\mathbf{r}_u^{k,t_1,t_2} \in \mathbb{C}^{N_R}$ for the u th user by multiplying \mathbf{Y}^{k,t_1,t_2} with \mathbf{p}_u^k as

$$\mathbf{r}_u^{k,t_1,t_2} = \mathbf{Y}^{k,t_1,t_2} \mathbf{p}_u^k = \mathbf{W}_B^{k,t_2} \mathbf{H}_u^k \mathbf{f}_u^{k,t_1} + \widetilde{\mathbf{n}}^{t_1,t_2} \quad (6)$$

where

$$\mathbf{W}_B^{k,t_2} \triangleq \mathbf{W}_B^{k,t_2} \mathbf{W}_R^{t_2}, \quad \mathbf{f}_u^{k,t_1} \triangleq \mathbf{F}_{R,u}^{t_1} \mathbf{F}_{B,u}^{k,t_1} \mathbf{1}_{M_R}, \quad (7)$$

$$\widetilde{\mathbf{n}}^{t_1,t_2} \triangleq \widetilde{\mathbf{N}}^{t_1,t_2} \mathbf{p}_u^k.$$

Define $T_3 \triangleq T_2 N_R$. We stack the T_2 received pilot sequences together and have

$$\mathbf{r}_u^{k,t_1} = \mathbf{W}^k \mathbf{H}_u^k \mathbf{f}_u^{k,t_1} + \widetilde{\mathbf{n}}^{t_1} \quad (8)$$

where

$$\mathbf{r}_u^{k,t_1} \triangleq [(\mathbf{r}_u^{k,t_1,1})^T, (\mathbf{r}_u^{k,t_1,2})^T, \dots, (\mathbf{r}_u^{k,t_1,T_2})^T]^T \in \mathbb{C}^{T_3},$$

$$\mathbf{W}^k \triangleq [(\mathbf{W}^{k,1})^T, (\mathbf{W}^{k,2})^T, \dots, (\mathbf{W}^{k,T_2})^T]^T \in \mathbb{C}^{T_3 \times N_A},$$

$$\widetilde{\mathbf{n}}^{t_1} \triangleq [(\widetilde{\mathbf{n}}^{t_1,1})^T, (\widetilde{\mathbf{n}}^{t_1,2})^T, \dots, (\widetilde{\mathbf{n}}^{t_1,T_2})^T]^T \in \mathbb{C}^{T_3}. \quad (9)$$

Further define

$$\mathbf{R}_u^k \triangleq [\mathbf{r}_u^{k,1}, \mathbf{r}_u^{k,2}, \dots, \mathbf{r}_u^{k,T_1}] \in \mathbb{C}^{T_3 \times T_1},$$

$$\mathbf{F}_u^k \triangleq [\mathbf{f}_u^{k,1}, \mathbf{f}_u^{k,2}, \dots, \mathbf{f}_u^{k,T_1}] \in \mathbb{C}^{M_A \times T_1}, \quad (10)$$

$$\widetilde{\mathbf{n}} \triangleq [\widetilde{\mathbf{n}}^1, \widetilde{\mathbf{n}}^2, \dots, \widetilde{\mathbf{n}}^{T_1}] \in \mathbb{C}^{T_3 \times T_1}.$$

We have

$$\mathbf{R}_u^k = \mathbf{W}^k \mathbf{H}_u^k \mathbf{F}_u^k + \tilde{\mathbf{n}}. \quad (11)$$

We need to estimate \mathbf{H}_u^k based on \mathbf{R}_u^k , \mathbf{W}^k and \mathbf{F}_u^k , which will be discussed in the following section.

III. ESPRIT-BASED CHANNEL ESTIMATION

In this section, we first analyze the channel structure of frequency-selective channel to show that all the subcarriers share the same AoA and AoD. Then we propose a TDE-based channel estimation scheme, where the AoA and AoD are estimated with super-resolution utilizing the rotation invariance of the channel steering vectors. Finally the AoA and AoD are paired to reconstruct the channel.

A. Analysis of Frequency-Selective Channel

Define $\tilde{g}_{u,d,i} \triangleq g_{u,i} p_{rc,u}(dT_s - \tau_{u,i})$, $i = 1, 2, \dots, L_u$ to ease the notation. The channel matrix at the d th delay tap in (3) can be represented in a more compact way as

$$\mathbf{H}_{u,d} = \gamma \mathbf{A}_{u,R} \mathbf{\Delta}_{u,d} \mathbf{A}_{u,T}^H \quad (12)$$

where $\mathbf{A}_{u,R} \in \mathbb{C}^{N_A \times L_u}$, $\mathbf{A}_{u,T} \in \mathbb{C}^{M_A \times L_u}$ and $\mathbf{\Delta}_{u,d} \in \mathbb{C}^{L_u \times L_u}$ are denoted as

$$\begin{aligned} \mathbf{A}_{u,R} &\triangleq [\boldsymbol{\alpha}(N_A, \theta_{u,1}), \boldsymbol{\alpha}(N_A, \theta_{u,2}), \dots, \boldsymbol{\alpha}(N_A, \theta_{u,L_u})], \\ \mathbf{A}_{u,T} &\triangleq [\boldsymbol{\alpha}(M_A, \phi_{u,1}), \boldsymbol{\alpha}(M_A, \phi_{u,2}), \dots, \boldsymbol{\alpha}(M_A, \phi_{u,L_u})], \\ \mathbf{\Delta}_{u,d} &\triangleq \text{diag}([\tilde{g}_{u,d,1}, \tilde{g}_{u,d,2}, \dots, \tilde{g}_{u,d,L_u}]). \end{aligned} \quad (13)$$

Then (2) can be further rewritten as

$$\mathbf{H}_u^k = \gamma \mathbf{A}_{u,R} \mathbf{\Lambda}_u^k \mathbf{A}_{u,T}^H \quad (14)$$

where $\mathbf{\Lambda}_u^k$ is defined as $\mathbf{\Lambda}_u^k \triangleq \sum_{d=0}^{D-1} \mathbf{\Delta}_{u,d} e^{-j\frac{2\pi kd}{K}} \in \mathbb{C}^{L_u \times L_u}$. $\mathbf{\Lambda}_u^k$ is also a diagonal matrix with L_u nonzero entries, where the i ($i = 1, 2, \dots, L_u$)th diagonal entry $\mathbf{\Lambda}_u^k[i, i]$ is the equivalent channel gain for the i th path of the k th subcarrier. It is seen from (14) that all K subcarriers share the same AoA and AoD, indicating that we can estimate AoA and AoD utilizing only one subcarrier. Therefore we use the subcarrier $k = 0$ to estimate AoA and AoD. Then the channel gain is estimated and the channel matrix is reconstructed for each subcarrier $k = 0, 1, \dots, K - 1$ successively.

B. TDE-based Channel Estimation Scheme

Now we propose the TDE-based channel estimation scheme to estimate the AoA and AoD with super-resolution in three stages. Then the AoA and AoD are paired to reconstruct the channel.

1) *AoA Estimation*: In the first stage, the BS and U users turn off the N_A th and M_A th antenna of the total N_A and M_A antennas, respectively. We use all K subcarriers to transmit pilot sequences for T time slots. To ease the notation, denote \mathbf{W}^k in (9) and \mathbf{F}_u^k in (10) with $k = 0$ in the first stage as $\tilde{\mathbf{W}}^{(1)}$ and $\tilde{\mathbf{F}}_u^{(1)}$, respectively, which can be represented as

$$\tilde{\mathbf{W}}^{(1)} = [\tilde{\mathbf{W}}, \mathbf{0}^{T_3}], \quad \tilde{\mathbf{F}}_u^{(1)} = [(\tilde{\mathbf{F}}_u)^T, (\mathbf{0}^{T_1})^T]^T \quad (15)$$

where $\tilde{\mathbf{W}} \in \mathbb{C}^{T_3 \times (N_A-1)}$ and $\tilde{\mathbf{F}}_u \in \mathbb{C}^{(M_A-1) \times T_1}$ denote the hybrid combining and precoding matrix connected to the powered N_A-1 and M_A-1 antennas, respectively. Combining (11), (14) and (15), at the subcarrier $k = 0$, we have

$$\begin{aligned} \tilde{\mathbf{R}}_u^{(1)} &= \tilde{\mathbf{W}}^{(1)} \mathbf{H}_u^0 \tilde{\mathbf{F}}_u^{(1)} + \tilde{\mathbf{n}} = \gamma \tilde{\mathbf{W}}^{(1)} \mathbf{A}_{u,R} \mathbf{\Lambda}_u^0 \mathbf{A}_{u,T}^H \tilde{\mathbf{F}}_u^{(1)} + \tilde{\mathbf{n}} \\ &= \gamma \tilde{\mathbf{W}} \mathbf{A}_{u,R}^1 \mathbf{\Lambda}_u^0 (\mathbf{A}_{u,T}^1)^H \tilde{\mathbf{F}}_u + \tilde{\mathbf{n}} \end{aligned} \quad (16)$$

where $\mathbf{A}_{u,R}^1 \in \mathbb{C}^{(N_A-1) \times L_u}$ and $\mathbf{A}_{u,T}^1 \in \mathbb{C}^{(M_A-1) \times L_u}$ are consisted of the first $N_A - 1$ and $M_A - 1$ rows of $\mathbf{A}_{u,R}$ and $\mathbf{A}_{u,T}$, respectively. Entry of $\mathbf{A}_{u,R}^1$ at the n ($n = 1, 2, \dots, N_A - 1$)th row and i ($i = 1, 2, \dots, L_u$)th column and that of $\mathbf{A}_{u,T}^1$ at the m ($m = 1, 2, \dots, M_A - 1$)th row and i th column are denoted as

$$\begin{aligned} \mathbf{A}_{u,R}^1[n, i] &= e^{j\pi\theta_{u,i}(n-1)} / \sqrt{N_A}, \\ \mathbf{A}_{u,T}^1[m, i] &= e^{j\pi\phi_{u,i}(m-1)} / \sqrt{M_A}. \end{aligned} \quad (17)$$

In the second stage, the BS and U users turn off the 1st and M_A th antenna of the total N_A and M_A antennas, respectively. However, in this stage, we only use the subcarrier $k = 0$ to transmit pilot sequences for T time slots. $\tilde{\mathbf{F}}_u^{(1)}$ is set to be the same as that in (15), while \mathbf{W}^k with $k = 0$ in the second stage is denoted as

$$\tilde{\mathbf{W}}^{(2)} = [\mathbf{0}^{T_3}, \tilde{\mathbf{W}}] \quad (18)$$

where $\tilde{\mathbf{W}}$ is set to be the same as that in (15). Therefore (11) can be represented as

$$\begin{aligned} \tilde{\mathbf{R}}_u^{(2)} &= \gamma \tilde{\mathbf{W}}^{(2)} \mathbf{A}_{u,R} \mathbf{\Lambda}_u^0 \mathbf{A}_{u,T}^H \tilde{\mathbf{F}}_u^{(1)} + \tilde{\mathbf{n}} \\ &= \gamma \tilde{\mathbf{W}} \mathbf{A}_{u,R}^2 \mathbf{\Lambda}_u^0 (\mathbf{A}_{u,T}^1)^H \tilde{\mathbf{F}}_u + \tilde{\mathbf{n}} \end{aligned} \quad (19)$$

where $\mathbf{A}_{u,R}^2 \in \mathbb{C}^{(N_A-1) \times L_u}$ is consisted of the last $N_A - 1$ rows of $\mathbf{A}_{u,R}$ as

$$\mathbf{A}_{u,R}^2[n, i] = e^{j\pi\theta_{u,i}n} / \sqrt{N_A}. \quad (20)$$

Combining (17) and (20) and utilizing the rotation invariance property of the channel steering vectors, we have

$$\mathbf{A}_{u,R}^2 = \mathbf{A}_{u,R}^1 \mathbf{\Theta} \quad (21)$$

where $\mathbf{\Theta} \in \mathbb{C}^{L_u \times L_u}$ is the diagonal matrix denoted as

$$\mathbf{\Theta} = \text{diag}([e^{j\pi\theta_{u,1}}, e^{j\pi\theta_{u,2}}, \dots, e^{j\pi\theta_{u,L_u}}]). \quad (22)$$

As shown in Algorithm 1, we use $\tilde{\mathbf{R}}_u^{(1)}$ and $\tilde{\mathbf{R}}_u^{(2)}$ to obtain an estimate of $\theta_{u,i}$, $i = 1, 2, \dots, L_u$. Define $\tilde{\mathbf{R}}_u \in \mathbb{C}^{2T_3 \times T_1}$. By stacking $\tilde{\mathbf{R}}_u^{(1)}$ and $\tilde{\mathbf{R}}_u^{(2)}$ at step 2, we have

$$\tilde{\mathbf{R}}_u = \begin{bmatrix} \tilde{\mathbf{R}}_u^{(1)} \\ \tilde{\mathbf{R}}_u^{(2)} \end{bmatrix} \stackrel{(a)}{\approx} \mathbf{B}_R \mathbf{B}_T \quad (23)$$

where

$$\begin{aligned} \mathbf{B}_R &\triangleq \begin{bmatrix} \gamma \tilde{\mathbf{W}} \mathbf{A}_{u,R}^1 \\ \gamma \tilde{\mathbf{W}} \mathbf{A}_{u,R}^1 \mathbf{\Theta} \end{bmatrix} \in \mathbb{C}^{2T_3 \times L_u}, \\ \mathbf{B}_T &\triangleq \mathbf{\Lambda}_u^0 (\mathbf{A}_{u,T}^1)^H \tilde{\mathbf{F}}_u \in \mathbb{C}^{L_u \times T_1}. \end{aligned} \quad (24)$$

Algorithm 1 TDE-based Channel Estimation for AoA

- 1: *Input:* $\tilde{\mathbf{R}}_u^{(1)}, \tilde{\mathbf{R}}_u^{(2)}$.
 - 2: Obtain $\hat{\mathbf{R}}_u$ via (23).
 - 3: Obtain \mathbf{B} via (25).
 - 4: Obtain \mathbf{U} via (26).
 - 5: Obtain \mathbf{U}_1 and \mathbf{U}_2 via (27).
 - 6: Obtain Ψ via (29).
 - 7: Obtain L_u eigenvalues of Ψ as $\{\lambda_i, i = 1, 2, \dots, L_u\}$.
 - 8: Obtain $\hat{\theta}_{u,i}$ via (30), $i = 1, 2, \dots, L_u$.
 - 9: *Output:* $\hat{\theta}_{u,i}, i = 1, 2, \dots, L_u$.
-

The equality marked by (a) holds by neglecting the term from the additive noise. Further define $\mathbf{B} \in \mathbb{C}^{2T_3 \times 2T_3}$ at step 3 as

$$\mathbf{B} \triangleq \tilde{\mathbf{R}}_u (\tilde{\mathbf{R}}_u)^H = \mathbf{B}_R (\mathbf{B}_T \mathbf{B}_T^H) \mathbf{B}_R^H. \quad (25)$$

At step 4, the singular value decomposition (SVD) of the positive semi-definite matrix \mathbf{B} can be represented as

$$\mathbf{B} = \mathbf{U} \Sigma \mathbf{U}^H \quad (26)$$

where $\mathbf{U} \in \mathbb{C}^{2T_3 \times L_u}$ is consisted of the first L_u columns of an unitary matrix and $\Sigma \in \mathbb{C}^{L_u \times L_u}$ is a real diagonal matrix. It shows that \mathbf{B}_R and \mathbf{U} share the same basis of their respective L_u column vectors. Therefore there exists an invertible matrix $\mathbf{T}_R \in \mathbb{C}^{L_u \times L_u}$ satisfying $\mathbf{U} = \mathbf{B}_R \mathbf{T}_R$. Dividing \mathbf{U} into two submatrices $\mathbf{U}_1 \in \mathbb{C}^{T_3 \times L_u}$ and $\mathbf{U}_2 \in \mathbb{C}^{T_3 \times L_u}$, we have

$$\mathbf{U} = \begin{bmatrix} \mathbf{U}_1 \\ \mathbf{U}_2 \end{bmatrix} = \begin{bmatrix} \gamma \tilde{\mathbf{W}} \mathbf{A}_{u,R}^1 \mathbf{T}_R \\ \gamma \tilde{\mathbf{W}} \mathbf{A}_{u,R}^1 \Theta \mathbf{T}_R \end{bmatrix}. \quad (27)$$

Considering \mathbf{T}_R is an invertible matrix, we obtain

$$\mathbf{U}_2 = \mathbf{U}_1 \mathbf{T}_R^{-1} \Theta \mathbf{T}_R. \quad (28)$$

Define $\Psi \triangleq \mathbf{T}_R^{-1} \Theta \mathbf{T}_R \in \mathbb{C}^{L_u \times L_u}$, i.e., the eigenvalue decomposition of Ψ , where \mathbf{T}_R is consisted of L_u eigenvectors and Θ is consisted of L_u eigenvalues on the diagonal. We obtain Ψ at step 6 as

$$\Psi = (\mathbf{U}_1)^\dagger \mathbf{U}_2. \quad (29)$$

Define the L_u eigenvalues of Ψ as $\{\lambda_i, i = 1, 2, \dots, L_u\}$. Then the estimated AoA of the i th path of the u th user can be expressed at step 8 as

$$\hat{\theta}_{u,i} = \arg(\lambda_i) / \pi, \quad i = 1, 2, \dots, L_u \quad (30)$$

where $\arg(\lambda)$ denotes the phase angle of the complex number λ . Finally we output the estimated AoA of the i th path of the u th user at step 9.

2) *AoD Estimation:* In the third stage, the BS and U users turn off the N_A th and 1st antenna of the total N_A and M_A antennas, respectively. Similarly, in this stage, we only use the subcarrier $k = 0$ to transmit pilot sequences for T time slots. $\tilde{\mathbf{W}}^{(1)}$ is set to be the same as that in (15), while \mathbf{F}_u^k with $k = 0$ in the third stage is denoted as

$$\tilde{\mathbf{F}}_u^{(2)} = [(\mathbf{0}^{T_1})^T, (\tilde{\mathbf{F}}_u)^T]^T \quad (31)$$

where $\tilde{\mathbf{F}}_u$ is set to be the same as that in (15). Therefore (11) can be represented as

$$\begin{aligned} \tilde{\mathbf{R}}_u^{(3)} &= \gamma \tilde{\mathbf{W}}^{(1)} \mathbf{A}_{u,R} \Lambda_u^0 \mathbf{A}_{u,T}^H \tilde{\mathbf{F}}_u^{(2)} + \tilde{\mathbf{n}} \\ &= \gamma \tilde{\mathbf{W}} \mathbf{A}_{u,R}^1 \Lambda_u^0 (\mathbf{A}_{u,T}^2)^H \tilde{\mathbf{F}}_u + \tilde{\mathbf{n}} \end{aligned} \quad (32)$$

where $\mathbf{A}_{u,T}^2 \in \mathbb{C}^{(M_A-1) \times L_u}$ is consisted of the last $M_A - 1$ rows of $\mathbf{A}_{u,T}$ as

$$\mathbf{A}_{u,T}^2[m, i] = e^{j\pi\phi_{u,i}m} / \sqrt{M_A}. \quad (33)$$

Combining (17) and (33), we have

$$\mathbf{A}_{u,T}^2 = \mathbf{A}_{u,T}^1 \Phi \quad (34)$$

where $\Phi \in \mathbb{C}^{L_u \times L_u}$ is the diagonal matrix denoted as

$$\Phi = \text{diag}([e^{j\pi\phi_{u,1}}, e^{j\pi\phi_{u,2}}, \dots, e^{j\pi\phi_{u,L_u}}]). \quad (35)$$

It is seen that (34) has the same structure as (21). Therefore we can directly run Algorithm 1 to obtain the estimated AoD of the i th path of the u th user $\hat{\phi}_{u,i}$, $i = 1, 2, \dots, L_u$ by replacing

$$\begin{aligned} &[\tilde{\mathbf{R}}_u^{(1)}, \tilde{\mathbf{R}}_u^{(2)}, \tilde{\mathbf{W}}, \mathbf{A}_{u,R}^1, \Theta, \\ &\Lambda_u^0, (\mathbf{A}_{u,T}^1)^H, \tilde{\mathbf{F}}_u, T_3, T_1, \hat{\theta}_{u,i}] \end{aligned}$$

with

$$\begin{aligned} &[(\tilde{\mathbf{R}}_u^{(1)})^H, (\tilde{\mathbf{R}}_u^{(3)})^H, (\tilde{\mathbf{F}}_u)^H, \mathbf{A}_{u,T}^1, \Phi, \\ &(\Lambda_u^0)^H, (\mathbf{A}_{u,R}^1)^H, (\tilde{\mathbf{W}})^H, T_1, T_3, \hat{\phi}_{u,i}] \end{aligned}$$

3) AoA and AoD Pairing and Channel Reconstruction:

After $\theta_{u,i}$ and $\phi_{u,i}$, $i = 1, 2, \dots, L_u$ are estimated, we can obtain the estimation of $\hat{\mathbf{A}}_{u,R}$ and $\hat{\mathbf{A}}_{u,T}$ as

$$\begin{aligned} \hat{\mathbf{A}}_{u,R} &= [\alpha(N_A, \hat{\theta}_{u,1}), \alpha(N_A, \hat{\theta}_{u,2}), \dots, \alpha(N_A, \hat{\theta}_{u,L_u})], \\ \hat{\mathbf{A}}_{u,T} &= [\alpha(M_A, \hat{\phi}_{u,1}), \alpha(M_A, \hat{\phi}_{u,2}), \dots, \alpha(M_A, \hat{\phi}_{u,L_u})]. \end{aligned} \quad (36)$$

Define $\delta_u^k \triangleq [\Lambda_u^k[1, 1], \Lambda_u^k[2, 2], \dots, \Lambda_u^k[L_u, L_u]]^T \in \mathbb{C}^{L_u}$. Then \mathbf{R}_u^k with $k = 0, 1, \dots, K - 1$ in (11) obtained in the first stage can be represented in vector form as

$$\begin{aligned} \text{vec}(\mathbf{R}_u^k) &= \text{vec}(\mathbf{W}^k \mathbf{H}_u^k \mathbf{F}_u^k) + \text{vec}(\tilde{\mathbf{n}}) \\ &\stackrel{(a)}{=} ((\mathbf{F}_u^k)^T \otimes \mathbf{W}^k) \text{vec}(\mathbf{H}_u^k) + \text{vec}(\tilde{\mathbf{n}}) \\ &\stackrel{(b)}{=} \gamma ((\mathbf{F}_u^k)^T \otimes \mathbf{W}^k) (\mathbf{A}_{u,T}^* \circ \mathbf{A}_{u,R}) \delta_u^k + \text{vec}(\tilde{\mathbf{n}}) \end{aligned} \quad (37)$$

where the equality marked by (a) holds due to the fact that $\text{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A})\text{vec}(\mathbf{B})$, the equality marked by (b) follows from the channel model in (14) and the properties of the Khatri-Rao product. Then δ_u^k can be estimated as

$$\hat{\delta}_u^k = \left(((\mathbf{F}_u^k)^T \otimes \mathbf{W}^k) (\hat{\mathbf{A}}_{u,T}^* \circ \hat{\mathbf{A}}_{u,R}) \right)^\dagger \text{vec}(\mathbf{R}_u^k) / \gamma \quad (38)$$

However, in this case, it requires that the L_u columns of $\hat{\mathbf{A}}_{u,R}$ and $\hat{\mathbf{A}}_{u,T}$ are paired, i.e., $\hat{\theta}_{u,i}$ and $\hat{\phi}_{u,i}$, $i = 1, 2, \dots, L_u$ are paired. Otherwise $(\hat{\mathbf{A}}_{u,T}^* \circ \hat{\mathbf{A}}_{u,R}) \neq (\mathbf{A}_{u,T}^* \circ \mathbf{A}_{u,R})$, resulting estimation error of $\hat{\delta}_u^k$. Therefore we should pair $\hat{\theta}_{u,i}$

and $\hat{\phi}_{u,i}$ before estimation of δ_u^k , which can be realized by estimating Λ_u^0 .

Λ_u^0 can be estimated as

$$\hat{\Lambda}_u^0 = (\widetilde{\mathbf{W}}^{(1)} \hat{\mathbf{A}}_{u,R})^\dagger \widetilde{\mathbf{R}}_u^{(1)} (\hat{\mathbf{A}}_{u,T}^H \widetilde{\mathbf{F}}_u^{(1)})^\dagger / \gamma. \quad (39)$$

Λ_u^0 is a diagonal matrix with L_u nonzero entries. However, since $\hat{\theta}_{u,i}$ and $\hat{\phi}_{u,i}$ are not paired, $\hat{\Lambda}_u^0$ may not be a diagonal matrix. Then we analyze the structure of $\hat{\Lambda}_u^0$. Define

$$\begin{aligned} \boldsymbol{\theta}_u &\triangleq [\theta_{u,1}, \theta_{u,2}, \dots, \theta_{u,L_u}], \boldsymbol{\phi}_u \triangleq [\phi_{u,1}, \phi_{u,2}, \dots, \phi_{u,L_u}], \\ \hat{\boldsymbol{\theta}}_u &\triangleq [\hat{\theta}_{u,1}, \hat{\theta}_{u,2}, \dots, \hat{\theta}_{u,L_u}], \hat{\boldsymbol{\phi}}_u \triangleq [\hat{\phi}_{u,1}, \hat{\phi}_{u,2}, \dots, \hat{\phi}_{u,L_u}]. \end{aligned}$$

$\hat{\boldsymbol{\theta}}_u$ and $\hat{\boldsymbol{\phi}}_u$ are two permutations of $\boldsymbol{\theta}_u$ and $\boldsymbol{\phi}_u$, respectively, if neglecting the additive noise. Define $\mathbf{p} \in \mathbb{Z}^{L_u}$ and $\mathbf{q} \in \mathbb{Z}^{L_u}$ as two permutations of $\{1, 2, \dots, L_u\}$. Then we have

$$\hat{\theta}_{u,i} = \theta_{u,\mathbf{p}[i]}, \quad \hat{\phi}_{u,i} = \phi_{u,\mathbf{q}[i]}. \quad (40)$$

Denote $\mathbf{C}_{\mathbf{p}}$ to be a $L_u \times L_u$ square matrix where only $\mathbf{C}_{\mathbf{p}}[\mathbf{p}[i], i] = 1, i = 1, 2, \dots, L_u$ and all the other entries are zero. Then (40) can be converted in matrix form as

$$\hat{\boldsymbol{\theta}}_u = \boldsymbol{\theta}_u \mathbf{C}_{\mathbf{p}}, \quad \hat{\boldsymbol{\phi}}_u = \boldsymbol{\phi}_u \mathbf{C}_{\mathbf{q}}. \quad (41)$$

Similarly, (36) can be further expressed as

$$\hat{\mathbf{A}}_{u,R} = \mathbf{A}_{u,R} \mathbf{C}_{\mathbf{p}}, \quad \hat{\mathbf{A}}_{u,T} = \mathbf{A}_{u,T} \mathbf{C}_{\mathbf{q}}. \quad (42)$$

Denote $\mathbf{Z}_{\mathbf{p},\mathbf{q}}$ to be a $L_u \times L_u$ square matrix where only $\mathbf{Z}_{\mathbf{p},\mathbf{q}}[\mathbf{p}[i], \mathbf{q}[i]] = 1$ and all the other entries are zero. Combining (39) and (42), we have

$$\begin{aligned} \hat{\Lambda}_u^0 &= (\mathbf{C}_{\mathbf{p}})^{-1} \Lambda_u^0 ((\mathbf{C}_{\mathbf{q}})^H)^{-1} \\ &\stackrel{(a)}{=} (\mathbf{C}_{\mathbf{p}})^H \Lambda_u^0 \mathbf{C}_{\mathbf{q}} = \sum_{i=1}^{L_u} \Lambda_u^0[i, i] \mathbf{Z}_{\mathbf{p}[i], \mathbf{q}[i]} \end{aligned} \quad (43)$$

where the equality marked by (a) holds due to the fact that $(\mathbf{C}_{\mathbf{p}})^{-1} = (\mathbf{C}_{\mathbf{p}})^H$, and $\mathbf{p} \in \mathbb{Z}^{L_u}$ is a permutation of $\{1, 2, \dots, L_u\}$ satisfying $(\mathbf{C}_{\mathbf{p}})^H = \mathbf{C}_{\mathbf{p}}$. It is seen that $\hat{\Lambda}_u^0$ is a permutation of Λ_u^0 , where $(\mathbf{p}[i], \mathbf{q}[i])$ represent the coordinate of the i th diagonal entry of Λ_u^0 . Since \mathbf{p} and \mathbf{q} are two permutations of $\{1, 2, \dots, L_u\}$, there are no repeated numbers in \mathbf{p} and \mathbf{q} respectively, indicating that there is one and only one nonzero entry in each row and column of $\hat{\Lambda}_u^0$. For example, if $L_u = 3$, $\mathbf{p} = \{1, 3, 2\}^T$ and $\mathbf{q} = \{3, 2, 1\}^T$, $\hat{\Lambda}_u^0$ can be represented as

$$\hat{\Lambda}_u^0 = \begin{bmatrix} 0 & 0 & \Lambda_u^0[1, 1] \\ \Lambda_u^0[3, 3] & 0 & 0 \\ 0 & \Lambda_u^0[2, 2] & 0 \end{bmatrix}. \quad (44)$$

By finding the coordinate of these L_u nonzero entries, we can pair \mathbf{p} and \mathbf{q} with the same indexes, i.e., $(\mathbf{p}[i], \mathbf{q}[i]), i = 1, 2, \dots, L_u$. Define $\hat{\mathbf{p}}$ and $\hat{\mathbf{q}}$ as the estimate of \mathbf{p} and \mathbf{q} , respectively. Then we can obtain the paired AoA and AoD of the L_u paths of the u th user as

$$\tilde{\boldsymbol{\theta}}_u = \hat{\boldsymbol{\theta}}_u (\mathbf{C}_{\mathbf{p}})^{-1} = \hat{\boldsymbol{\theta}}_u \mathbf{C}_{\mathbf{p}}, \quad \tilde{\boldsymbol{\phi}}_u = \hat{\boldsymbol{\phi}}_u (\mathbf{C}_{\mathbf{q}})^{-1}. \quad (45)$$

Algorithm 2 AoA and AoD Pairing for TDE-based Channel Estimation

- 1: *Input:* $\hat{\boldsymbol{\theta}}_u, \hat{\boldsymbol{\phi}}_u, \hat{\Lambda}_u^0$.
 - 2: Initialization: $\hat{\mathbf{p}} \leftarrow \mathbf{0}^{L_u}, \hat{\mathbf{q}} \leftarrow \mathbf{0}^{L_u}, \Gamma_1 \leftarrow \{1, 2, \dots, L_u\}$.
 - 3: **for** $i = 1, 2, \dots, L_u$ **do**
 - 4: Obtain q_i via (46).
 - 5: Obtain the i th entry of $\hat{\mathbf{p}}$ and $\hat{\mathbf{q}}$ via (47).
 - 6: Update Γ_{i+1} via (48).
 - 7: **end for**
 - 8: Obtain $\tilde{\boldsymbol{\theta}}_u$ and $\tilde{\boldsymbol{\phi}}_u$ via (45).
 - 9: *Output:* $\boldsymbol{\theta}_u, \boldsymbol{\phi}_u$.
-

Now we propose Algorithm 2 to pair the AoA and AoD of the L_u paths of the u th user, considering the structure of $\hat{\Lambda}_u^0$. There are L_u iterations for Algorithm 2 from step 3 to step 7. Define Γ_i as a set of column indexes with $L_u - i + 1$ entries. We initialize Γ_1 to be $\{1, 2, \dots, L_u\}$ at step 2. For the i th iteration, we compare the amplitude of entries in the i th row of $\hat{\Lambda}_u^0$ with the column indexes denoted by Γ_i to find the largest one at step 4, which is expressed as

$$q_i = \arg \max_{d \in \Gamma_i} |\hat{\Lambda}_u^0[i, d]|. \quad (46)$$

(i, q_i) is paired to represent the coordinate of the nonzero entry of $\hat{\Lambda}_u^0$ in the i th row. At step 5, we add i and q_i to $\hat{\mathbf{p}}$ and $\hat{\mathbf{q}}$ respectively as

$$\hat{\mathbf{p}}[i] = i, \quad \hat{\mathbf{q}}[i] = q_i. \quad (47)$$

Since there is one and only one nonzero entry in each column of $\hat{\Lambda}_u^0$, i.e., $\hat{\Lambda}_u^0[n, q_i] = 0, n \in \{1, 2, \dots, L_u\}, n \neq i$, we delete q_i from Γ_i to obtain Γ_{i+1} at step 6 as

$$\Gamma_{i+1} = \Gamma_i \setminus q_i. \quad (48)$$

We repeat this procedure for L_u times to obtain the paired \mathbf{p} and \mathbf{q} . Then the paired AoA and AoD of the L_u paths of the u th user can be obtained from (45) at step 8. Finally we output the paired AoA and AoD of the L_u paths of the u th user at step 9.

Then we can obtain the column-paired $\tilde{\mathbf{A}}_{u,R}$ and $\tilde{\mathbf{A}}_{u,T}$ from (36) by replacing $\hat{\theta}_{u,i}, \hat{\phi}_{u,i}$ with $\tilde{\theta}_{u,i}, \tilde{\phi}_{u,i}$, respectively. Then the estimation of δ_u^k can be obtained from (38) by replacing $\hat{\mathbf{A}}_{u,R}, \hat{\mathbf{A}}_{u,T}$ with $\tilde{\mathbf{A}}_{u,R}, \tilde{\mathbf{A}}_{u,T}$, respectively. Finally the estimated channel matrix at the k th subcarrier between the BS and the u th user can be reconstructed as

$$\hat{\mathbf{H}}_u^k = \gamma \tilde{\mathbf{A}}_{u,R} \text{diag}(\hat{\boldsymbol{\delta}}_u^k) \tilde{\mathbf{A}}_{u,T}^H. \quad (49)$$

As shown in (30), (38) and (49), the channel estimation error equals to zero without noise for the TDE-based scheme, which shows the superiority of the super-resolution TDE-based scheme in its small estimation error.

IV. SIMULATION RESULTS

Now we evaluate the performance of the proposed TDE-based scheme. We consider uplink transmission of a multi-user mmWave massive MIMO system. The BS serving $U = 4$

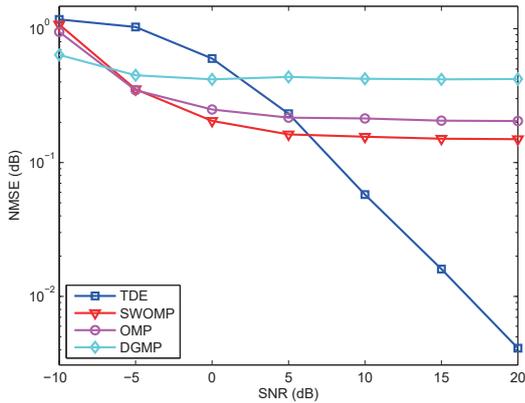


Fig. 1. Comparisons of NMSE for different SNR.

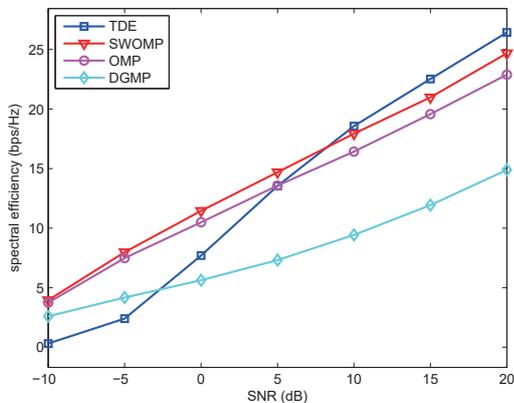


Fig. 2. Comparisons of spectral efficiency for different SNR.

users has $N_A = 64$ antennas and $N_R = 4$ RF chains, while each user has $M_A = 16$ antennas and $M_R = 1$ RF chain. The number of resolvable paths in mmWave channel is set to be $L_u = 3$, while $g_{u,i} \sim \mathcal{CN}(0, 1)$ for $i = 1, 2, \dots, L_u$. We use $K = 16$ subcarriers. We set $T_1 = 8$ and $T_2 = 8$. In order to make fair comparisons, we set the total time slots of the DGMP-based, SWOMP-based and OMP-based schemes the same as the proposed scheme. We do not compare with the ESPRIT-based and the MUSIC-based channel estimation schemes because they consider the frequency-flat channel.

As shown in Fig. 1, we compare the channel estimation performance in terms of normalized mean-squared error (NMSE) for the proposed TDE-based, the SWOMP-based [2], the OMP-based [3] and the DGMP-based schemes [5] with different SNR. It is observed from Fig. 1 that the TDE-based scheme outperforms the SWOMP-based, OMP-based and DGMP-based schemes at high SNR region. At SNR of 15dB, the TDE-based scheme has 89.4%, 92.2% and 96.2% performance improvement compared with the SWOMP-based,

OMP-based and DGMP-based schemes, respectively. The reason for the unsatisfactory performance of the SWOMP-based and OMP-based schemes is that both the SWOMP-based and OMP-based schemes do not consider the power leakage due to the limited beamspace resolution, while the reason for the unsatisfactory performance of the DGMP-based scheme is that it only estimates a single path.

As shown in Fig. 2, we compare the spectral efficiency for the proposed TDE-based scheme, the SWOMP-based scheme, the OMP-based scheme and the DGMP-based scheme. It is seen that the proposed scheme achieves better performance than the other schemes at high SNR region. At SNR of 15dB, the TDE-based scheme has 9.4%, 15.1% and 88.7% performance improvement compared with the SWOMP-based, OMP-based and DGMP-based schemes, respectively.

V. CONCLUSIONS

In this paper, we have proposed a TDE-based channel estimation scheme, which includes the AoA and AoD estimation as well as AoA and AoD pairing and channel reconstruction. Simulation results have verified the effectiveness of our work and have shown that the proposed scheme outperforms the existing schemes. Future work will focus on the interference mitigation for multiuser mmWave MIMO transmission.

ACKNOWLEDGMENT

This work is supported in part by the National Natural Science Foundation of China (NSFC) under Grant 61871119 and by the Natural Science Foundation of Jiangsu Province under Grant BK20161428.

REFERENCES

- [1] 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.
- [2] J. Rodriguez-Fernandez, N. Gonzalez-Prelcic, K. Venugopal, and R. W. Heath, "Frequency-domain compressive channel estimation for frequency-selective hybrid mmWave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 2946–2960, May 2018.
- [3] K. Venugopal, A. Alkhateeb, R. W. Heath, and N. Gonzalez-Prelcic, "Time-domain channel estimation for wideband millimeter wave systems with hybrid architecture," in *Proc. IEEE ICASSP*, New Orleans, USA, Mar. 2017, pp. 6493–6497.
- [4] F. Shu, Y. Qin, T. Liu *et al.*, "Low-complexity and high-resolution DOA estimation for hybrid analog and digital massive MIMO receive array," *IEEE Trans. Commun.*, vol. 66, no. 6, pp. 2487–2501, Jun. 2018.
- [5] Z. Gao, C. Hu, L. Dai, and Z. Wang, "Channel estimation for millimeter-wave massive MIMO with hybrid precoding over frequency-selective fading channels," *IEEE Commun. Lett.*, vol. 20, no. 6, pp. 1259–1262, Jun. 2016.
- [6] A. Liao, Z. Gao, Y. Wu, H. Wang, and M.-S. Alouini, "2D unitary ESPRIT based super-resolution channel estimation for millimeter-wave massive MIMO with hybrid precoding," *IEEE Access*, vol. 5, pp. 24 747–24 757, Nov. 2017.
- [7] Z. Guo, X. Wang, and W. Heng, "Millimeter-Wave channel estimation based on 2-D beamspace MUSIC method," *IEEE Trans. Wireless Commun.*, vol. 16, no. 8, pp. 5384–5394, Aug. 2017.
- [8] 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.