Off-grid Compressive Sensing Based Channel Estimation with Non-uniform Grid in Millimeter Wave MIMO System

You You*, Li Zhang†,
*Purple Mountain Laboratories, Nanjing, China, Email: youyou@pmlabs.com.cn
†School of Electronic and Electrical Engineering, University of Leeds, Leeeds, UK, Email: l.x.zhang@leeds.ac.uk

***Abstract*—Channel estimation is challenging for millimeter-wave (mmWave) communications because of the use of hybrid architecture and massive multiple-input multiple-output (MIMO) technology. By utilizing the sparsity in the angular domain, conventional on-grid compressive sensing methods can efficiently recover the channel state information (CSI). However, the channel estimation accuracy is severely affected by the off-grid errors and the selection of grid angles. The off-grid compressive sensing methods and the non-uniform grid angles can improve the channel estimation accuracy. In this paper, we investigate the impact of the non-uniform grid angles for the off-grid compressive sensing methods and the on-grid compressive sensing methods. We propose to employ the orthogonal matching pursuit (OMP) algorithm with interior point (IP) method based off-grid error mitigation to implement the channel estimation using the selected angle design. The simulation results demonstrate the advantages of the proposed off-grid compressive sensing method and show the impact of the non-uniform grid angles.**

***Index Terms*—channel estimation, compressive sensing, optimization methods, off-grid errors.**

I. INTRODUCTION

Millimeter-wave (mmWave) communication is considered a promising technology for the future wireless systems because of the large amount of available spectrum [1]. In order to overcome the huge propagation loss, massive multiple input multiple output (MIMO) can be used to provide desirable beamforming gains. To reduce the power consumption and the hardware cost caused by the fully digital architecture, a hybrid MIMO architecture consisting of an analog beamformer in the radio frequency (RF) domain cascaded with a digital MIMO processor in baseband has been proposed for the mmWave communication [2]. However, the hybrid architecture and the large number of antennas at both transmitter and receiver make it challenging to obtain accurate channel state information (CSI) which is essential for hybrid precoding.

MmWave channel has been proved to have sparsity in the angular domain [1]. Instead of estimating all the entries in the channel matrix, only the angle-of-departures (AoDs), angle-of-arrivals (AoAs) and corresponding path gains of the dominant paths are estimated. By using the virtual channel representation [3] and the compressive sensing algorithms [4], CSI acquisition can be efficient. Several channel estimation schemes for mmWave Massive MIMO systems have been proposed recently. Specifically, [5] is a close-loop codebook-based beam training scheme. The authors of [5] propose a multistage process that the transmitter emits the pilot beams with wide beams based on the designed code-book that cover all of the angles of interest at the first stage. According to the feed back from receiver, transmitter is able to select beam patterns wisely in the following stages. However, the performance for close-loop methods are limited by the design of codebooks. On the other hand, without the feedback from receiver, several open-loop channel estimation schemes are developed [6]. By exploiting the sparsity in angular domain, CSI can be efficiently estimated by the orthogonal matching pursuit (OMP) algorithm [6]. However, such solutions in [6] are all on-grid methods which assume that the AoDs /AoAs lie on discrete grid angles (called on-grid AoAs/AoDs). This assumption results in power leakage problem caused by angle quantification, because the actual AoDs/AoAs are continuous angles (called off-grid AoAs/AoDs) as shown in Fig. 1. Several off-grid methods are proposed to mitigate the power leakage problem such as [7] for close-loop solutions and [8] for open-loop solutions. But, these off-grid methods are all based on the uniformly distributed grid angles. The selection of grid angles also effects the channel estimation performance significantly for both on-grid methods and off-grid methods. Paper [6] shows, compared with the uniformly distributed grid angles, using non-uniform virtual gird angles achieves better channel estimation performance. In this paper, we propose an OMP algorithm with optimization method and non-uniform grid angles to improve the channel estimation performance and investigate the impact of different virtual grid angles on both on-grid methods and off-grid methods.

Specifically, we adopt non-uniform grid angles for our proposed method. First, conventional OMP algorithm is used to estimate the 'rough' AoAs/AoDs. Then we optimize the 'rough' AoAs/AoDs iteratively using the interior point (IP) method to increase the coherence between the received signal and the sensing matrix. Comparing with the conventional OMP algorithm, the proposed off-grid mitigation method with the non-uniform grid angles significantly improve the channel estimation accuracy. In addition, comparing with the off-grid mitigation method with uniform grid angles, employing non-uniform virtual grid angles achieves super-resolution channel estimation at high signal-to-noise ratio (SNR).

The organization of the paper is as follows. Section II intro-

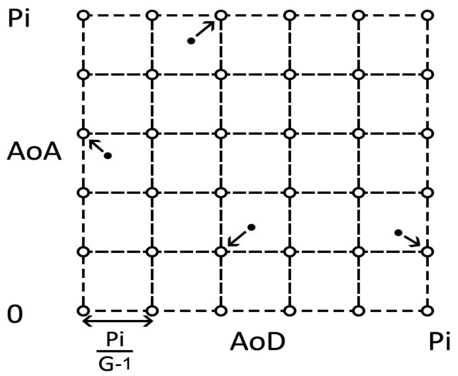

Fig. 1. An illustration of angle grid and the off grid angles [8].

duces the system model and CS based problem formulation. Section III presents the proposed off-grid mitigation method and the non-uniform grid angles. Simulation results illustrating the performance of the proposed algorithm are presented in Section IV. Finally, the conclusion is drawn in Section V.

II. System model

We consider a single user hybrid massive MIMO mmWave system, where the transmitter equipped with N_T antennas and N_T^{RF} RF chains communicating with a receiver equipped with N_R antennas and N_R^{RF} RF chains ($N_T^{RF} \leq N_T$, $N_R^{RF} \leq N_R$). In the hybrid system, RF chains are much fewer in number than the total adopted antenna numbers. Each RF chain is connected to all antennas via phase shifters. The precoding/combining processing is divided between the analog and digital domains. The discrete-time model for the received signal $\mathbf{r} \in \mathbb{C}^{\mathbf{N_R^{RF}} \times \mathbf{1}}$ of a single symbol period $\mathbf{s} \in \mathbb{C}^{\mathbf{N_T^{RF}} \times \mathbf{1}}$ can be formulated as

$$\mathbf{r} = \mathbf{C}^H \mathbf{H} \mathbf{F} \mathbf{s} + \mathbf{C}^H \mathbf{n}, \tag{1}$$

where $\mathbf{n} \in \mathbb{C}^{N_R \times 1}$ is the noise vector with $\mathcal{CN}(0, \sigma_n^2)$ entries. $\mathbf{C} \in \mathbb{C}^{N_R \times N_R^{RF}}$ is the hybrid combining matrix. $\mathbf{F} \in \mathbb{C}^{N_T \times N_T^{RF}}$ is the hybrid precoding matrix. $\mathbf{H} \in \mathbb{C}^{N_R \times N_T}$ is the channel matrix. $\mathbf{s} \in \mathbb{C}^{N_T^{RF} \times 1}$ is the pilot signal in one symbol period.

In the channel estimation stage, $\mathbf{x} = \mathbf{F}\mathbf{s} \in \mathbb{C}^{N_T \times 1}$ is the pilot sequence, where the n-th element of $\mathbf{x}$ is transmitted by the n-th antenna at transmitter. We assume that the transmitter sends $N_T^{Beam} (N_T^{Beam} < N_T)$ different pilot sequences $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{N_T^{Beam}}$ and the receiver obtains N_R^{Beam} dimension of received pilot sequence. Because the receiver only has N_R^{RF} RF chains, M time slots are required to to obtain an N_R^{Beam} dimension received pilot sequence, where $M = N_R^{Beam}/N_R^{RF}$. We assume N_R^{Beam} is multiples of N_R^{RF} so that the training time is MN_T^{Beam}. Considering the m-th time slot for the p-th pilot sequence, received vector $\mathbf{y}_{p,m} \in \mathbb{C}^{N_R^{RF} \times 1}$ is given by

$$\mathbf{y}_{p,m} = \mathbf{W}_m^H \mathbf{H} \mathbf{x}_p + \mathbf{W}_m^H \mathbf{n}_{p,m}, \tag{2}$$

where $\mathbf{W}_m \in \mathbb{C}^{N_R \times N_R^{RF}}$ is the combining matrix, $\mathbf{n}_{p,m} \in \mathbb{C}^{N_R \times 1}$ is the noise vector. Collecting $\mathbf{y}_{p,m}$ for $m \in \{1, 2, \ldots, M\}$, we get the received pilots in M time slots as

$$\mathbf{y}_p = \mathbf{W}^H \mathbf{H} \mathbf{x}_p + \mathbf{W}^H \mathbf{n}_p, \tag{3}$$

where $\mathbf{n}_p \in \mathbb{C}^{N_R \times 1}$ is the noise vector and

$$\begin{aligned} \mathbf{y}_p &= [\mathbf{y}_{p,1}^T, \mathbf{y}_{p,2}^T, \ldots, \mathbf{y}_{p,M}^T]^T \in \mathbb{C}^{N_R^{Beam} \times 1}, \\ \mathbf{W} &= [\mathbf{W}_1, \mathbf{W}_2, \ldots, \mathbf{W}_M] \in \mathbb{C}^{N_R \times N_R^{Beam}}. \end{aligned} \tag{4}$$

Collecting $\mathbf{y}_p$ for $p \in \{1, 2, \ldots, N_T^{Beam}\}$, we get the received pilots for all N_T^{Beam} transmitted pilot sequences as

$$\mathbf{Y} = \mathbf{W}^H \mathbf{H} \mathbf{X} + \mathbf{N}, \tag{5}$$

where

$$\begin{aligned} \mathbf{Y} &= [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{N_T^{Beam}}] \in \mathbb{C}^{N_R^{Beam} \times N_T^{Beam}}, \\ \mathbf{X} &= [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{N_T^{Beam}}] \in \mathbb{C}^{N_T \times N_T^{Beam}}, \\ \mathbf{N} &= [\mathbf{n}_1, \mathbf{n}_2, \ldots, \mathbf{n}_{N_T^{Beam}}] \in \mathbb{C}^{N_R^{Beam} \times N_T^{Beam}}. \end{aligned} \tag{6}$$

A. Channel model

A geometric channel model [2] is widely adopted to approximate the mmWave narrowband channel as

$$\mathbf{H} = \sqrt{\frac{N_T N_R}{L}} \sum_{\ell=1}^{L} \alpha_\ell \mathbf{a}_R(\theta_{R,\ell}) \mathbf{a}_T^H(\theta_{T,\ell}), \tag{7}$$

where L is the number of scatterers, α_ℓ is the complex gain, $\theta_{R,\ell}$ and $\theta_{T,\ell}$ are the AoA and AoD of the l-th path, respectively. The mmWave channel is regarded unchanged within the channel coherence time for channel estimation and each scatterer only contributes one propagation path. $\mathbf{a}_R(\theta_{R,\ell})$ and $\mathbf{a}_T^H(\theta_{T,\ell})$ are the steering vector at receiver and transmitter respectively. The steering vectors depend on the array geometry. Ignoring the subscripts without loss of generality, we consider an N-element uniform linear arrays (ULA) geometry so that the steering vectors are denoted as

$$\mathbf{a}(\theta) = [1, e^{-j2\pi \frac{d}{\lambda} \cos\theta}, e^{-j4\pi \frac{d}{\lambda} \cos\theta}, ..., e^{-j2\pi(N-1) \frac{d}{\lambda} \cos\theta}]^T, \tag{8}$$

where d is the antenna spacing, λ denotes the wavelength of operation. The channel model in (7) can be also written in matrix form as

$$\mathbf{H} = \mathbf{A}_R \mathbf{H}_a \mathbf{A}_T^H, \tag{9}$$

where

$$\begin{aligned} \mathbf{H}_a &= \text{diag}(\alpha_1, \alpha_2, \ldots, \alpha_L), \\ \mathbf{A}_R &= [\mathbf{a}_R(\theta_{R,1}), \mathbf{a}_R(\theta_{R,2}), \ldots, \mathbf{a}_R(\theta_{R,L})], \\ \mathbf{A}_T &= [\mathbf{a}_T(\theta_{T,1}), \mathbf{a}_T(\theta_{T,2}), \ldots, \mathbf{a}_T(\theta_{T,L})]. \end{aligned} \tag{10}$$

To exploit the sparsity of the mmWave channel, virtual channel representation and on-grid compressive sensing algorithms are widely exploited in mmWave CE. Specifically, the continuous true AoDs/AoAs are approximated as discrete angles based on pre-defined virtual angles called 'grid angles', defined as

$$\{\varphi_g : \varphi_g \in [0, \pi], g = 1, \ldots, G\}, \tag{11}$$

where $\varphi_1 = 0, \varphi_G = \pi$. Uniform grid is widely applied in mmWave CE as $\varphi_g \in \{0, \frac{\pi}{G-1}, \ldots, \frac{\pi(G-1)}{G-1}\})$, with $G \gg L$ to achieve the desired resolution [5], [6]. In this paper, we use non-uniform grid angles for our proposed algorithm. The non-uniform grid is determined so that $\{\cos(\varphi_g)\}$ are uniformly distributed in $(-1, 1]$: specifically, φ_g satisfies

$$\cos(\varphi_g) = \frac{2}{G}(g-1) - 1, \tag{12}$$

for $g \in \{1, 2, \ldots, G\}$. Note that, this grid is distributed non-uniformly in angular interval $[0, \pi]$. When $d = \frac{\lambda}{2}$, the rows of these array response matrices are orthogonal as proved in [6] that $\bar{\mathbf{A}}_T \bar{\mathbf{A}}_T^H = \frac{G}{N_T}\mathbf{I}_{N_T}$ and $\bar{\mathbf{A}}_R \bar{\mathbf{A}}_R^H = \frac{G}{N_R}\mathbf{I}_{N_R}$. In [6], the orthogonality property has been proved to be able to reduce the coherence of the proposed CS formulation.

Collecting all the steering vectors with angles $\{\varphi_g\}$, we assume that the AoD and AoA use the same grid and named as $\varphi_{T,g} = \varphi_g$ and $\varphi_{R,g} = \varphi_g$. $\bar{\mathbf{A}}_T = [\mathbf{a}_T(\varphi_{T,1}), \mathbf{a}_T(\varphi_{T,2}), \ldots, \mathbf{a}_T(\varphi_{T,G})] \in \mathbb{C}^{N_T \times G}$ and $\bar{\mathbf{A}}_R = [\mathbf{a}_R(\varphi_{R,1}), \mathbf{a}_R(\varphi_{R,2}), \ldots, \mathbf{a}_R(\varphi_{R,G})] \in \mathbb{C}^{N_R \times G}$ are the array response matrices of non-uniform grid angles. $\mathbf{H}$ can be approximated in terms of a L-sparse matrix $\mathbf{H}_b \in \mathbb{C}^{G \times G}$, with L non zero elements in the positions on the non-uniform grid AoDs/AoAs as

$$\mathbf{H} = \bar{\mathbf{A}}_R \mathbf{H}_b \bar{\mathbf{A}}_T^H + \mathbf{E}. \tag{13}$$

$\mathbf{E}$ is the off-grid errors and caused by quantizing the AoDs/AoAs by the grid angles. Increasing the grid size G can improve the angle resolution and reduce the off-grid errors, but the grid size G can not be too large to break the Restricted Isometry Property (RIP) for compressive sensing algorithms. Too large grid size leads to even worse estimation performance with exponentially increasing complexity [8]. In this paper, we propose to employ interior point (IP) method to mitigate the impact of the off-grid errors for on-grid compressive sensing method (OMP) with different virtual grids.

B. Formulation of channel estimation

Considering the system model in (5), we can formulate the mmWave CE problem as a signal recovery problem by vectorizing the received signal matrix $\mathbf{Y}$ in (5) as

$$\begin{aligned} \mathbf{y}_v &= (\mathbf{X}^T \otimes \mathbf{W}^H) \cdot \mathrm{vec}(\mathbf{H}) + \mathrm{vec}(\mathbf{N}) \\ &= \mathbf{Q} \cdot \mathrm{vec}(\mathbf{H}) + \mathbf{n}_Q. \end{aligned} \tag{14}$$

Using the property of the Khatri-Rao product

$$\mathrm{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A}) \cdot \mathrm{vec}(\mathbf{B}). \tag{15}$$

The matrices $\mathbf{W}^H$, $\mathbf{H}$ and $\mathbf{X}$ in (14) are regarded as $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ in (15) respectively. $\mathbf{n}_Q$ is the vectorized noise. Let $\mathbf{Q} = (\mathbf{X}^T \otimes \mathbf{W}^H) \in \mathbb{C}^{N_T^{Beam} N_R^{Beam} \times N_T N_R}$. Least square (LS) is a widely used method to estimate $\mathrm{vec}(\mathbf{H})$ as $(\mathbf{Q}^H\mathbf{Q})^{-1}\mathbf{Q}^H\mathbf{y}_v$. However, the LS solution requires $N_T^{Beam} N_R^{Beam} \geq N_T N_R$ so that $\mathbf{Q}^H\mathbf{Q}$ has full rank. Considering the large number of antennas in the mmWave MIMO system, training overhead with LS method is huge. This difficulty can be overcome by using CS methods, because the number of entries to be estimated is proportional to the sparsity level which is much less than $N_T N_R$.

Using the virtual channel representation in (13) but neglecting the grid error $\mathbf{E}$, (14) can be rewritten as a sparse recovery problem as

$$\begin{aligned} \mathbf{y}_v &= \mathbf{Q} \cdot \mathrm{vec}(\bar{\mathbf{A}}_R \mathbf{H}_b \bar{\mathbf{A}}_T^H) + \mathbf{n}_Q \\ &= \mathbf{Q}\mathbf{A}_D \cdot \mathbf{h}_b + \mathbf{n}_Q \\ &= \bar{\mathbf{Q}} \cdot \mathbf{h}_b + \mathbf{n}_Q. \end{aligned} \tag{16}$$

$\mathbf{A}_D = \bar{\mathbf{A}}_T^* \otimes \bar{\mathbf{A}}_R$ is an $N_T N_R \times G^2$ dictionary matrix that consists of the G^2 column vectors of the form $\mathbf{a}_T^H(\varphi_{T,g}) \otimes \mathbf{a}_R(\varphi_{R,g})$, where $\{g = 1, \ldots, G\}$. $\bar{\mathbf{Q}} = \mathbf{Q}\mathbf{A}_D$ is a $N_T^{Beam} N_R^{beam} \times G^2$ sensing matrix. $\mathbf{h}_b$ is a $G^2 \times 1$ sparse vector that has only L non-zero elements and $L \ll G^2$. Given the formulation in (16), CS algorithms such as OMP can be employed to solve this sparse signal recovery problem.

III. Proposed off-grid channel estimation method

In this section, OMP method is used in conjunction with the IP method with non-uniform grid angles to effectively estimate the AoDs/AoAs and path gains. The proposed algorithm is named as IP-OMP-D. Compared with the method proposed in [8], we apply non-uniform grid angles and different optimization process.

The proposed IP-OMP-D algorithm for recovering the sparse channel vector $\mathbf{h}_b$ is summarized in Algorithm 1. Algorithm 1 requires known sensing matrix $\bar{\mathbf{Q}}$, measurement vector $\mathbf{y}_v$, sparsity K and grid size G. In the initial stage, the iteration counter is set as $t = 1$ and the residual is set as $\mathbf{r}_0 = \mathbf{y}_v$. When $t \leq K$, this algorithm chooses the column j of $\bar{\mathbf{Q}}$ which is the most strongly correlated with the residual $\mathbf{r}_{t-1}$ in step 3. The chosen column number j is stored in set Ω_t in step 4. In step 5, based on the column number j and predetermined dictionary matrix $\mathbf{A}_D$, a rough AoD/AoA estimated value can be obtained. Specifically, $\mathbf{A}_D = \bar{\mathbf{A}}_T^* \otimes \bar{\mathbf{A}}_R$ is an $N_T N_R \times G^2$ dictionary matrix that consists of the G^2 column vectors corresponding to G^2 possible discrete AoD/AoA pairs according to the grid $\{\varphi_{T,g}\}$ and $\{\varphi_{R,g}\}$. For the t-th iteration, the rough AoD/AoA pair is estimated as the m-th and n-th virtual angle in the grid as $\varphi_{T,m} = \arccos(\frac{2}{G}(m-1) - 1)$ and $\varphi_{R,n} = \arccos(\frac{2}{G}(n-1) - 1)$, where $m = \mathrm{ceil}(\frac{j}{G})$ and $n = \mathrm{mod}(j, G)$. Note that, if $\mathrm{mod}(j, G) = 0$, $n = G$. The main problem for rough AoD/AoA pair is the off grid angles caused by quantization. Specifically, the maximum value in $|\bar{\mathbf{Q}}(i)^H \mathbf{r}_{t-1}|$ can be larger the results in the step 3, if we can modify the discrete grid angles to decrease to the off-grid errors. In step 6, a more accurate AoD/AoA pair is estimated by maximizing $|\bar{\mathbf{Q}}(i)^H \mathbf{r}_{t-1}|$.

Algorithm 1 IP-OMP-D method for mmWave channel estimation

Require: sensing matrix $\bar{\mathbf{Q}}$, measurement vector $\mathbf{y}_v$, sparsity K and grid G

1: Ω_{t-1} =empty set, residual $\mathbf{r}_0 = \mathbf{y}_v$, set the iteration counter $t = 1$

2: **while** $t \le K$ **do**

3: $j = \arg \max_{i=1,\ldots,G^2} |\bar{\mathbf{Q}}(i)^H \mathbf{r}_{t-1}|$

4: $\Omega_t = \Omega_{t-1} \cup \{j\}$

5: $m = \text{ceil}(\frac{j}{G})$, $n = \begin{cases} \text{mod}(j, G) & (m \ne 0) \\ G & (m = 0) \end{cases}$

$AoD_t = \arccos(\frac{2}{G}(m-1) - 1)$

$AoA_t = \arccos(\frac{2}{G}(n-1) - 1)$

$x_t = (AoD_t, AoA_t)$

6: $\max_{AoD'_t, AoA'_t} f(AoD'_t, AoA'_t)$

$x'_t = (AoD'_t, AoA'_t)$

7: $\mathbf{p} = (\mathbf{X}^T \otimes \mathbf{W}^H)(\mathbf{a}^*(AoD'_t) \otimes \mathbf{a}(AoA'_t))$

8. $\bar{\mathbf{Q}}_j = \mathbf{p}$

9: $\mathbf{h}_t = \arg\min_{\mathbf{h}} \|\mathbf{y}_v - \bar{\mathbf{Q}}_{\Omega_t} \mathbf{h}\|_2$

10: $\mathbf{r}_t = \mathbf{y}_v - \bar{\mathbf{Q}}_{\Omega_t} \mathbf{h}_t$

11: $t = t + 1$

12: **end while**

13: $\mathbf{h}_b(i) = \mathbf{h}_{t-1}$ for $i \in \Omega_{t-1}$ and $\mathbf{h}_b(i) = 0$ otherwise

14: **return** $\mathbf{h}_b$

IP method is employed and $x_t = (AoD_t, AoA_t)$ is set as the original point corresponding to the jth column in $\bar{\mathbf{Q}}$. Objective function is defined as the correlation between the sensing column and the residual as $f(AoD'_t, AoA'_t) = |(\mathbf{X}^T \otimes \mathbf{W}^H)(\mathbf{a}^*_T(AoD'_t) \otimes \mathbf{a}_R(AoA'_t)^H)\mathbf{r}_{t-1}|$. Through maximizing the objective function between the adjacent grid angles, a more accurate angle pair $x'_t = (AoD'_t, AoA'_t)$ can be obtained. This optimization process in step 6 is formulated as

$$\max_{AoD'_t, AoA'_t} f(AoD'_t, AoA'_t)$$

$$s.t. \begin{cases} AoD_t - \frac{\varphi_{T,m} - \varphi_{T,m-1}}{2} < AoD'_t < AoD_t + \frac{\varphi_{T,m+1} - \varphi_{T,m}}{2} \\ AoA_t - \frac{\varphi_{R,n} - \varphi_{R,n-1}}{2} < AoA'_t < AoA_t + \frac{\varphi_{R,n+1} - \varphi_{R,n}}{2}. \end{cases}$$

After obtaining x'_t by optimization, the new most correlated column is calculated as $\mathbf{p} = (\mathbf{X}^T \otimes \mathbf{W}^H)(\mathbf{a}^*(AoD'_t) \otimes \mathbf{a}(AoA'_t))$ in step 7. In step 8, the column j in sensing matrix $\bar{\mathbf{Q}}$ is replaced by $\mathbf{p}$, so that the AoD'_t and AoA'_t become the new $\varphi_{T,m}$ and $\varphi_{R,n}$ in the grid. In this way, the estimated AoDs/AoAs are moved towards the optimal solutions. In step 9, the channel gains associated with the estimated AoDs/AoAs are obtained by evaluating the LS solution of $\mathbf{y}_v = \bar{\mathbf{Q}}_{\Omega_t}\mathbf{h}$, where $\bar{\mathbf{Q}}_{\Omega_t} \in \mathbb{C}^{N_t^{Beam} N_r^{Beam} \times t}$ is the sub-matrix of $\bar{\mathbf{Q}}$ that only contains the columns whose indices are included in Ω_t and $\mathbf{h} \in \mathbb{C}^{t \times 1}$ is a vector with varying size. In step 10, the contributions of the column j to $\mathbf{y}_v$ are subtracted to update the residual. After obtaining the new residual, in step 11, $t = t + 1$. This procedure is repeated from step 3 to step 11 until $t = K$. In step 13, the sparse channel vector $\mathbf{h}_b \in \mathbb{C}^{G^2 \times 1}$ is made by recording the K estimated channel gains into the corresponding position according to elements in Ω_t, so that $\mathbf{h}_b(i) = \mathbf{h}_{t-1}$ for $i \in \Omega_{t-1}$ and $\mathbf{h}_b(i) = 0$, otherwise. $\mathbf{h}_b$ is the channel matrix as in (16).

IV. Simulation Results

In this section, performance of the proposed method is examined by simulations. We consider a mmWave massive MIMO system with $N_T = N_R = 64$ antennas (ULA) and $N_R^{RF} = N_T^{RF} = 4$ RF chains at both transmitter and receiver. $N_T^{Beam} = N_R^{Beam} = 32$ training beams are used in the simulation. Note that $N_T^{Beam} N_R^{Beam} < N_T N_R$. The path gains are assumed i.i.d. random variables with distribution $\mathcal{CN}(0, \sigma_\alpha^2)$, where $\sigma_\alpha^2 = 1$. The AoAs and AoDs are modelled by the Laplacian distribution whose mean is uniformly distributed over $[0, \pi)$, and angular standard deviation is $\sigma_{AS} = 1$. Each element of the transmitted pilots $\mathbf{X}$ satisfies $x_{i,j} = \sqrt{\frac{P}{N_T}} e^{jw_{i,j}}$, where $P = 1$ is the transmitted power, $w_{i,j}$ is the random phase uniformly distributed in $[0, 2\pi)$. The SNR is defined by SNR=$\frac{P\sigma_\alpha^2}{\sigma_n^2}$, where σ_n^2 is the noise variance. All the simulation results are averaged over 100 channel realizations with a carrier frequency of 60GHz. At each channel realization, we assume that the number of scatterers L is determined by $L = \max\{P_{10}, 1\}$, where P_{10} is the outcome of the Poisson random variable with mean 10. OMP algorithm[6] and IP-OMP [8] are adopted for performance comparison and show the impact of the non-uniform grid angles on both on-grid algorithms and off-grid algorithms.

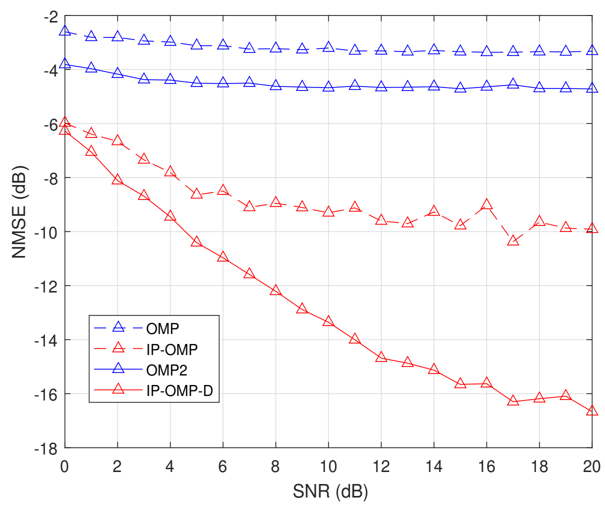

Fig. 2. NMSEs of IP-OMP-D with G=64 at different SNRs (dB)

Fig. 2 compares the normalized mean square error (NMSE) defined as $10\log_{10}\left(\mathbb{E}(\|\mathbf{H} - \mathbf{H}^{CS}\|_F^2 / \|\mathbf{H}\|_F^2)\right)$, where $\mathbf{H}^{CS}$ is the estimated channel matrix. OMP and IP-OMP with uniform grid angles (uniformly distributed over $[0, \pi)$) are used for comparison. OMP with designed non-uniform grid angles (defined as (12)) is also adopted for comparison and named as OMP2. The grid size G is 64 for all the algorithms. As shown in Fig. 2, as expected, the performances of OMP2 and

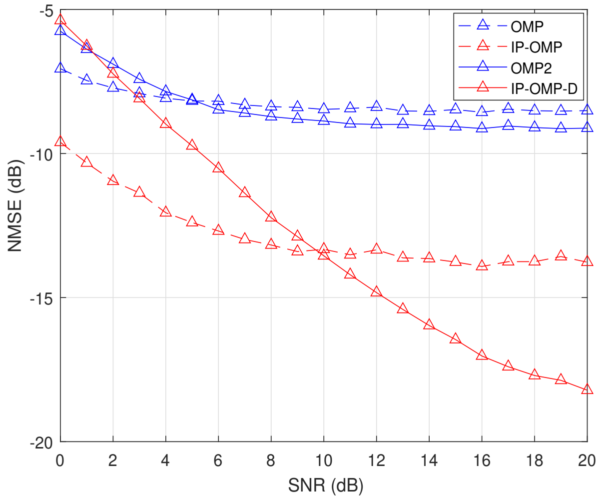

Fig. 3. NMSEs of IP-OMP-D with G=128 at different SNRs (dB)

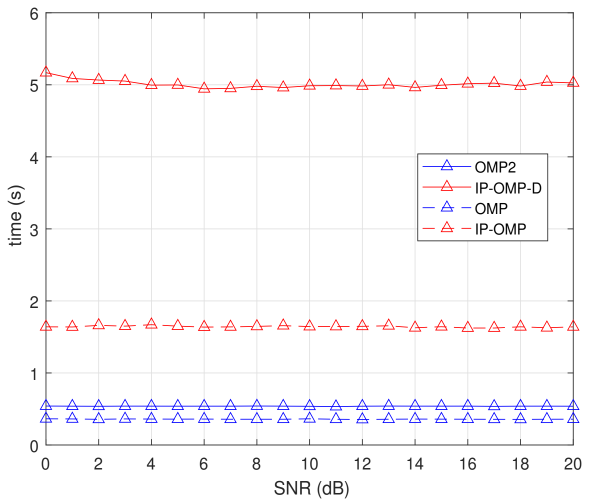

Fig. 4. Runtime of IP-OMP-D at different SNRs (dB)

IP-OMP-D are better than OMP and IP-OMP at all SNRs due to non-uniform grid angles when the grid size is small. The non-uniform grid angles are able to help reduce the coherence of the sensing matrix and improve the CE accuracy.

Fig. 3 compares the NMSE of IP-OMP-D at different SNRs but $G = 128$. As expected, OMP and IP-OMP are able to perform better with larger G at cost of higher computational complexity. OMP2 and IP-OMP-D also achieve better CE accuracy at high SNRs. However, OMP2 and IP-OMP-D deteriorate at low SNRs and are even worse than OMP and IP-OMP. With large noise and large grid size, OMP are more likely to obtain wrong 'rough' estimated AoDs/AoAs, so that the optimization process in IP-OMP-D can not perform correctly.

Fig. 4 shows the runtime of IP-OMP-D at different SNRs with $G = 128$. Considering the optimization process in IP-OMP and IP-OMP-D, we use runtime in MATLAB to compare the computational complexity. Compared with OMP, the results show that the non-uniform grid angles leads to slightly higher complexity for OMP2. But IP-OMP-D has almost 3 times complexity compared with IP-OMP because of the non-uniform range of optimization.

In summary, with small grid size, adopting non-uniform grid angles is able to improve the CE accuracy at cost of slightly increased complexity. But with large grid size, non-uniform grid angles only perform better at high SNRs and large noise deteriorate the CE accuracy significantly.

V. Conclusion

In this paper, we propose to employ non-uniform grid angles with an off-grid compressive sensing method for mmWave CE. The simulation results demonstrate that the proposed off-grid compressive sensing method with non-uniform grid angles are able to achieve better performance at cost of affordable complexity, especially at high SNRs. But, with large grid size, employing non-uniform grid angles is difficult to perform superior CE accuracy at low SNRs.

References

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

[2] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, Mar. 2014.

[3] A. M. Sayeed, "Deconstructing multiantenna fading channels," *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2563–2579, Oct. 2002.

[4] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[5] 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.

[6] J. Lee, G.-T. Gil, and Y. H. Lee, "Channel estimation via orthogonal matching pursuit for hybrid MIMO systems in millimeter wave communications," *IEEE Trans. Commun.*, vol. 64, no. 6, pp. 2370–2386, Jun. 2016.

[7] S. Sun and T. S. Rappaport, "Millimeter wave MIMO channel estimation based on adaptive compressed sensing," in *Proc. IEEE Int. Conf. Commun.Workshops (ICC Workshops)*, May 2017, pp. 47–53.

[8] Y. You, L. Zhang, and M. Liu, "IP aided OMP based channel estimation for millimeter wave massive MIMO communication," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2019, pp. 1–6.