High-Performance Channel Estimation for mmWave Wideband Systems with Hybrid Structures

Pengcheng Zhu, *Member, IEEE,* Huixin Lin, Jiamin Li, *Member, IEEE,* Dongming Wang, *Member, IEEE,* and Xiaohu You, *Fellow, IEEE*

***Abstract*—In this paper, a channel estimation problem for millimeter-wave (mmWave) multiple-input multiple-output (MIMO) orthogonal frequency division multiplexing (OFDM) systems with hybrid structures is studied. Firstly, a beamspace multiple signal classification (MUSIC) algorithm for mmWave wideband channels is proposed to simultaneously estimate the angles of arrival (AOAs), angles of departure (AODs) and transmission delays. Since the traditional spectral peak search method has high complexity, a multi-spectral peak search method is skillfully designed to search for multiple spectral peaks on the MUSIC spatial spectrum more quickly and accurately. Then, the proposed channel estimator is extended to more actual systems equipped with uniform planar arrays (UPAs). Finally, the Cramér–Rao bound (CRB) results of these channel parameters are derived for evaluating the performance of the proposed channel parameter estimator. Simulation results demonstrate that the proposed channel estimator has greatly high channel estimation accuracy.**

***Index Terms*—Millimeter-Wave, MIMO-OFDM, MUSIC, UPA, Parameter Estimation.**

I. Introduction

MILLIMETER-wave (mmWave) communication is one of the potential technologies for future wireless communication systems [1]–[3]. Compared with microwave systems, mmWave systems have the advantages of narrow beam, large communication capacity, short wave length and good security. However, due to these characteristics, the transmission loss of mmWave systems is large. Since antenna arrays can make signals more directional, large antenna arrays can be equipped at both base stations (BS) and mobile stations (MS) to compensate for significant transmission loss, thereby ensuring sufficient beamforming gains for mmWave communications [4].

Hybrid analog/digital beamforming structures have been employed in mmWave multiple-input multiple-output (MIMO) systems to process antenna arrays [5]–[7], which consists of an analog radio frequency (RF) beamformer and a digital baseband processor. An analog RF beamformer can be effectively implemented by phased arrays or lens arrays [8], and the number of RF chains can be significantly less than the number of antennas due to the sparsity of mmWave channels, which means that hybrid analog/digital structures can reduce hardware overhead [9].

To maintain the high beamforming gains of mmWave hybrid MIMO systems, it is necessary to accurately obtain the channel state information (CSI) by channel estimation. Since the responses of mmWave channels are dominated by only a small number of scattering paths, the traditional MIMO channel model with rich scattering cannot characterize the spatial sparsity of mmWave channels. Alternatively, a mmWave channel can be modeled in parametric form, which means that the mmWave channel estimation problem is no longer a channel matrix estimation problem, but a channel parameter estimation problem [10].

Recently, many studies have focused on addressing the channel parameter estimation problem in mmWave hybrid MIMO systems. In [11], an open-loop channel estimator for mmWave hybrid MIMO systems was proposed, which adopted a grid-based orthogonal matching pursuit (OMP) algorithm to estimate the angle of arrival (AOA), angle of departure (AOD) and corresponding gain of each resolvable path. However, OMP algorithm had angle quantization error, which can affect the performance of this channel estimator. To improve channel estimation accuracy, two super-resolution channel estimation methods respectively based on the iterative reweight (IR) and multiple signal classification (MUSIC) algorithm were proposed in [12] and [10]. However, the performance of the former was very sensitive to the selection of initial values and its convergence speed became slow when approaching the minimum value. Fortunately, the latter did not have these problems and can distinguish spatially close signals for accurate direction detection.

Most of the previous studies have focused on narrowband channels, but seldom consider the wideband channels with frequency selectivity [10]–[12]. Compared with narrowband channels, a wideband channel has a non-negligible multipath delay extension [13], which means that each resolvable path of a mmWave wideband channel is usually characterized by path gains, AOAs, AODs and transmission delays. Besides, orthogonal frequency division multiplexing (OFDM) has been commonly employed in mmWave wideband communication systems due to its robustness to channel frequency selectivity [14]–[16]. Thus, it is necessary to study the wideband channel estimation problem in mmWave MIMO-OFDM systems with hybrid structures.

Similar to the mmWave narrowband channel parameter estimation problem, a mmWave wideband channel parameter estimation problem can also be equally expressed as a

This work was supported by the National Key R&D Program of China under Grant 2021YFB2900300, and the National Natural Science Foundation of China under Grants 62171126 and 61971127. (*Corresponding author: Jiamin Li.*)

The authors are with the National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China, and also with the Purple Mountain Laboratories, Nanjing 211111, China. (Email: {p.zhu, huixin, jiaminli, wangdm, xhyu}@seu.edu.cn).

sparse signal recovery problem by exploiting the sparse scattering nature of mmWave channels, which can be solved by compressed sensing (CS) method [17]–[19]. Contrastingly, the channel estimation method that exploited tensor for signal processing had better channel estimation accuracy [20]. In [21], a tensor-based channel estimation method was proposed, which first employed an alternating least squares (ALS) procedure to obtain the estimates of factor matrices, and then employed a simple correlation-based method to estimate the channel parameters. However, the alternating results of ALS heavily depended on the initial state which may converge to a biased local optimal solution. To solve this problem, an improved channel parameter estimation scheme based on the structured CANDECOMP/PARAFAC decomposition (SCPD) algorithm was proposed in [22] by exploiting the Vandermonde nature of factor matrices. However, the specially designed beamformers of this scheme were proved to worsen the performance of channel estimation.

Different from the studies mentioned above, a channel parameter estimation scheme based on the spatial spectrum estimation (SSE) method [23] is creatively proposed for mmWave hybrid MIMO-OFDM systems. Specially, the beamspace MUSIC algorithm is employed to estimate the transmission delays, AOAs and AODs of mmWave channels, and the least-squares (LS) method is employed to estimate the path gains of mmWave channels. Beamspace MUSIC is one of the classical SSE methods [24], which has been applied to mmWave narrowband channels [25]–[27], but has not been applied to mmWave wideband channels to the best of our knowledge. Compared with tensor-based channel parameter estimator [20]–[22], [28], the proposed SSE-based channel parameter estimator has the following advantages. Firstly, its channel estimation performance is not affected by the initial state. Secondly, it is a super-resolution channel parameter estimator with high accuracy. Thirdly, its beamspace is designed by the discrete Fourier transform (DFT) beamspace framework which has been proved to improve the performance of channel estimation [10]. The main contributions of this paper are summarized as follows.

- To accurately and simultaneously acquire the estimates of transmission delays, AOAs and AODs of mmWave wideband channels, the received signal is reconstructed, and then a beamspace MUSIC algorithm for mmWave hybrid MIMO-OFDM systems is specially designed, whose spatial spectrum is three-dimensional.
- The traditional MUSIC algorithm exploits a grid search method to find spectral peaks on the spatial spectrum, which has high accuracy but large complexity, because the spatial spectrum is three-dimensional. Therefore, a chicken swarm optimization (CSO) algorithm is adopted to optimize the spectral peak search of the proposed beamspace MUSIC algorithm.
- To estimate the parameters of multiple paths, it is necessary to find multiple spectral peaks on the spatial spectrum of MUSIC algorithm. However, CSO algorithm can output only one global optimal solution at a time. Hence, a multi-spectral peak search method is designed by skillfully combining the proposed beamspace MUSIC algorithm with CSO algorithm according to the principle of dichotomy.
- Compared with uniform linear arrays (ULAs), uniform planar arrays (UPAs) have higher space efficiency [29] and can avoid beam-squint effect [30]. Therefore, the proposed channel estimator is extend to mmWave hybrid MIMO-OFDM systems equipped with UPAs, which is rarely studied in mmWave wideband channel estimation.
- To provide a benchmark for evaluating the performance of the proposed channel parameter estimator, the Cramér–Rao bound (CRB) results of these channel parameters are derived, which can represent the optimal asymptotic achievable performance of solving the channel parameter estimation problem studied in this paper.

The remainder of this paper is organized as follows. In Section II, system model is presented, and the channel estimation strategy for mmWave hybrid MIMO-OFDM systems is proposed in Section III. CRB results are derived in Section IV, and simulation results are presented in Section V, followed by concluding remarks in Section VI.

Notations: $\mathcal{A}$, $\mathbf{A}$, $\mathbf{a}$ and a represent the set, matrix, vector and scalar respectively. $\mathbf{A}^T$, $\mathbf{A}^*$, $\mathbf{A}^H$ and $\mathbf{A}^{-1}$ respectively denote the matrix transpose, conjugate, conjugate transpose and inverse. $\mathrm{diag}\left(\mathbf{a}\right)$ is a diagonal matrix with the elements of $\mathbf{a}$ on the diagonal. $\mathbf{A}\otimes\mathbf{B}$ and $\mathbf{A}\odot\mathbf{B}$ are the Kronecker product and Khatri-Rao product of $\mathbf{A}$ and $\mathbf{B}$, respectively. $\mathrm{vec}\left(\mathbf{A}\right)$ stacks columns of $\mathbf{A}$ into a vector. $\left[\mathbf{A}\right]_{l,l}$ represents the (l,l)-th element of $\mathbf{A}$. $\delta\left(\cdot\right)$ is the delta function.

II. System Model

A. Channel Model with ULAs

A mmWave MIMO-OFDM system with one base station (BS) and one mobile station (MS) is shown in Fig.1, where the BS and MS are respectively equipped with Q_a-antenna and U_a-antenna uniform linear arrays (ULAs). To reduce hardware overhead, hybrid analog/digital beamforming structures are employed in this system, where the antenna array of BS is connected to an analog RF combiner $\mathbf{W}\in\mathbb{C}^{Q_a\times Q_r}$ with $Q_r\left(<Q_a\right)$ RF chains, and the antenna array of MS is connected to an analog RF precoder $\mathbf{F}\in\mathbb{C}^{U_a\times U_r}$ with $U_r\left(<U_a\right)$ RF chains. The digital processors of BS and MS are responsible for channel estimation and digital precoding. The total number of OFDM subcarriers is M_0, where the first M subcarriers are selected for channel estimation [31].

Due to the sparse nature of mmWave channels, a ray-tracing channel model with L resolvable scatter paths between BS and MS is adopted, where L is generally small, because the signals in mmWave systems suffer from severe path loss and tend to reflect less to surrounding environment [32], [33]. To facilitate channel estimation, a time frame is divided into N time blocks. Therefore, the delay-domain time-varying channel matrix $\mathbf{H}\left(n,\tau\right)\in\mathbb{C}^{Q_a\times U_a}$ at the n-th $(n=1,\cdots,N)$ block can be modeled as [32]

$$\mathbf{H}\left(n,\tau\right)=\sum_{l=1}^{L}\alpha_l\left(n\right)\mathbf{a}_B\left(\theta_l\right)\mathbf{a}_U^T\left(\phi_l\right)\delta\left(\tau-\tau_l\right),\qquad(1)$$

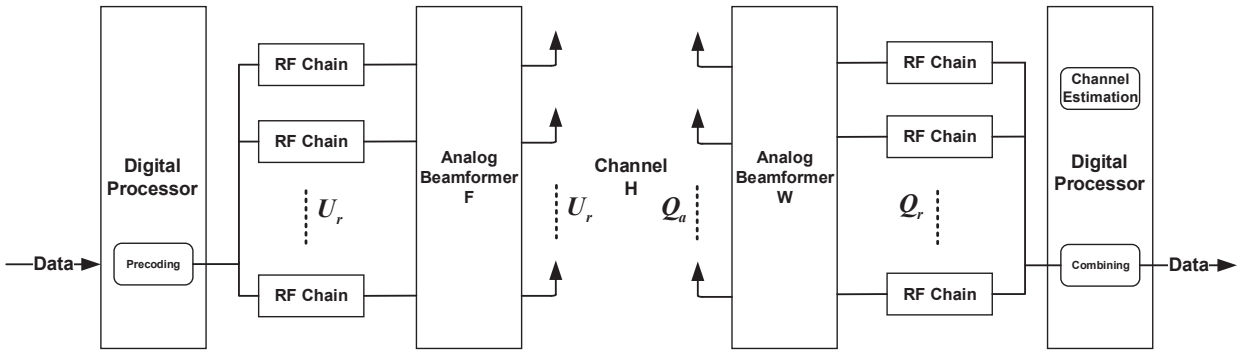


Fig. 1. A mmWave MIMO-OFDM system with hybrid structures

where τ_l is the transmission delay of the l-th $(l = 1, \cdots, L)$ path, $\alpha_l(n)$ is the path gain of the l-th path at the n-th block, $\mathbf{a}_B(\cdot) \in \mathbb{C}^{Q_a}$ and $\mathbf{a}_U(\cdot) \in \mathbb{C}^{U_a}$ are respectively the spatial steering vectors of BS and MS antenna arrays, which are respectively expressed as [31]

$$\begin{aligned} \mathbf{a}_B(\theta_l) &= \frac{1}{\sqrt{Q_a}}\left[1, e^{j2\pi\theta_l}, \cdots, e^{j2\pi(Q_a-1)\theta_l}\right]^T, \\ \mathbf{a}_U(\phi_l) &= \frac{1}{\sqrt{U_a}}\left[1, e^{j2\pi\phi_l}, \cdots, e^{j2\pi(U_a-1)\phi_l}\right]^T. \end{aligned} \tag{2}$$

Here, $\phi_l = \frac{d}{\lambda}\sin(\varphi_l)$ and $\theta_l = \frac{d}{\lambda}\sin(\vartheta_l)$ are respectively denoted as the AOD and AOA of the l-th path, where λ is the carrier wavelength, d is the adjacent antenna spacing assumed to be $\frac{\lambda}{2}$, φ_l and ϑ_l are respectively the physical AOD and physical AOA of the l-th path.

By performing the Fourier transform of (1) over τ, the frequency-domain time-varying channel matrix $\mathbf{H}_m(n) \in \mathbb{C}^{Q_a \times U_a}$ at the m-th $(m = 1, \cdots, M)$ subcarrier and n-th block is

$$\mathbf{H}_m(n) = \sum_{l=1}^{L} \alpha_l(n) \underbrace{e^{-j2\pi m\Delta_f \tau_l}}_{\triangleq e_m(\tau_l)} \mathbf{a}_B(\theta_l)\mathbf{a}_U^T(\phi_l), \tag{3}$$

where $\Delta_f = \frac{f_s}{M_0}$ is the subcarrier spacing with sampling rate f_s, and $e_m(\tau_l)$ is the m-th element of frequency-domain steering vector $\mathbf{e}(\tau_l) \in \mathbb{C}^M$ denoted as

$$\begin{aligned} \mathbf{e}(\tau_l) &= [e_1(\tau_l), e_2(\tau_l), \cdots, e_M(\tau_l)]^T \\ &= \left[e^{-j2\pi\tau_l\Delta_f}, e^{-j2\pi\tau_l 2\Delta_f}, \cdots, e^{-j2\pi\tau_l M\Delta_f}\right]^T. \end{aligned} \tag{4}$$

Remark 1: It is assumed that the variation of mmWave channels is only caused by the path gains, while the transmission delays, AOAs and AODs remain unchanged during a time frame. This assumption is based on [26], [32], [34], [35], which indicate that the transmission delays, AOAs and AODs belong to large-scale fading, while the path gains belong to small-scale fading.

B. Channel Model with UPA

Different from the above channel model, the BS is equipped with Q_a-antenna uniform planar array (UPA) as shown in Fig.2 and connected to an analog RF combiner $\mathbf{W} \in \mathbb{C}^{Q_a \times Q_r}$ with $Q_r\ (< Q_a)$ RF chains. Here, $Q_a = Q_a^x Q_a^z$ and $Q_r = Q_r^x Q_r^z$ in which $Q_a^x\,(Q_r^x)$ and $Q_a^z\,(Q_r^z)$ are respectively the number of antennas (RF chains) in the $\mathcal{X}$ direction and $\mathcal{Z}$ direction. The angle of elevation (e-angle) and angle of azimuth (a-angle) of the l-th path are respectively denoted as μ_l and ν_l. Therefore, the delay-domain time-varying channel matrix $\mathbf{H}(n,\tau) \in \mathbb{C}^{Q_a \times U_a}$ at the n-th block can be expressed as [30]

$$\mathbf{H}(n,\tau) = \sum_{l=1}^{L} \alpha_l(n)\mathbf{a}_B(\mu_l,\nu_l)\mathbf{a}_U^T(\phi_l)\delta(\tau-\tau_l), \tag{5}$$

where

$$\begin{aligned} \mathbf{a}_B(\mu_l,\nu_l) = &\frac{1}{\sqrt{Q_a}}[1, \cdots, \\ &e^{j\frac{2\pi d}{\lambda}((Q_a^x-1)\cos\mu_l\cos\nu_l + (Q_a^z-1)\sin\mu_l)}]^T. \end{aligned} \tag{6}$$

It can be found that $\mathbf{a}_B(\mu_l,\nu_l) = \mathbf{a}_B^x(\mu_l,\nu_l) \otimes \mathbf{a}_B^z(\mu_l)$, where $\mathbf{a}_B^x(\cdot) \in \mathbb{C}^{Q_a^x}$ and $\mathbf{a}_B^z(\cdot) \in \mathbb{C}^{Q_a^z}$ are respectively the spatial steering vectors of BS antenna array in the $\mathcal{X}$ direction and $\mathcal{Z}$ direction, which can be respectively expressed as

$$\begin{aligned} \mathbf{a}_B^x(\mu_l,\nu_l) &= \frac{1}{\sqrt{Q_a^x}}\left[1, \cdots, e^{j\frac{2\pi d}{\lambda}(Q_a^x-1)\cos\mu_l\cos\nu_l}\right]^T, \\ \mathbf{a}_B^z(\mu_l) &= \frac{1}{\sqrt{Q_a^z}}\left[1, \cdots, e^{j\frac{2\pi d}{\lambda}(Q_a^z-1)\sin\mu_l}\right]^T. \end{aligned} \tag{7}$$

Denoting $\theta_l^x = \frac{d}{\lambda}\cos\mu_l\cos\nu_l$ and $\theta_l^z = \frac{d}{\lambda}\sin\mu_l$, (7) can be rewritten as

$$\begin{aligned} \mathbf{a}_B^x(\theta_l^x) &= \frac{1}{\sqrt{Q_a^x}}\left[1, \cdots, e^{j2\pi(Q_a^x-1)\theta_l^x}\right]^T, \\ \mathbf{a}_B^z(\theta_l^z) &= \frac{1}{\sqrt{Q_a^z}}\left[1, \cdots, e^{j2\pi(Q_a^z-1)\theta_l^z}\right]^T, \end{aligned} \tag{8}$$

and $\mathbf{a}_B(\theta_l^x,\theta_l^z) = \mathbf{a}_B^x(\theta_l^x) \otimes \mathbf{a}_B^z(\theta_l^z)$. Then, (5) can be rewritten as

$$\begin{aligned} \mathbf{H}(n,\tau) &= \sum_{l=1}^{L} \alpha_l(n)\mathbf{a}_B(\theta_l^x,\theta_l^z)\mathbf{a}_U^T(\phi_l)\delta(\tau-\tau_l) \\ &= \sum_{l=1}^{L} \alpha_l(n)\mathbf{a}_B^x(\theta_l^x) \otimes \mathbf{a}_B^z(\theta_l^z)\mathbf{a}_U^T(\phi_l)\delta(\tau-\tau_l). \end{aligned} \tag{9}$$

By performing the Fourier transform of (9) over τ, the frequency-domain time-varying channel matrix $\mathbf{H}_m(n) \in \mathbb{C}^{Q_a \times U_a}$ at the m-th subcarrier and n-th block is

$$\mathbf{H}_m(n) = \sum_{l=1}^{L} \alpha_l(n) \underbrace{e^{-j2\pi m \Delta_f \tau_l}}_{\triangleq e_m(\tau_l)} \mathbf{a}_B^x(\theta_l^x) \otimes \mathbf{a}_B^z(\theta_l^z) \mathbf{a}_U^T(\phi_l). \tag{10}$$

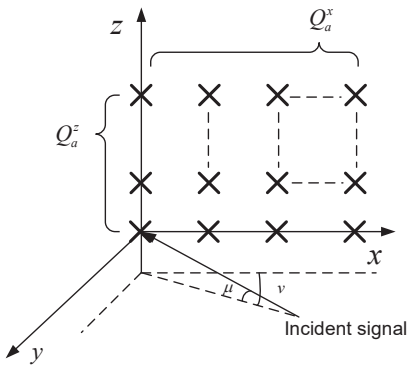


Fig. 2. Structures of UPA equipped at the BS

C. Uplink Transmission Model

As shown in Fig.3, each time block contains T_0 time slots, where the first T_p time slots are utilized for channel estimation, and the remaining time slots are utilized for data transmission.

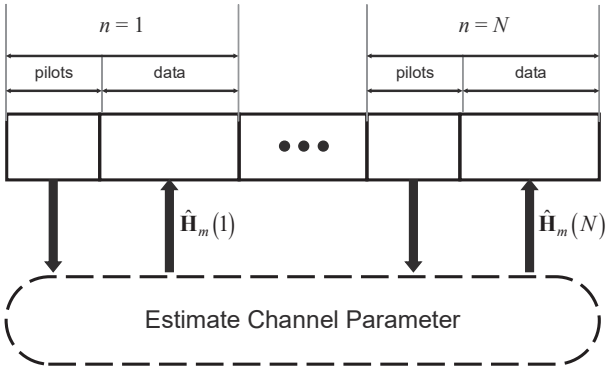


Fig. 3. Time frame for channel estimation

Assuming that the pilot signals transmitted by all RF chains at the MS are identical for all selected subcarriers and the same during each time block, the pilot signal transmitted by the i-th $(i = 1, \cdots, U_r)$ RF chain of MS during each time block is $\mathbf{s}_i = [s_i(1), \cdots, s_i(T_p)] \in \mathbb{C}^{1 \times T_p}$, and then the pilot signal transmitted by all RF chains of MS is $\mathbf{S} = [\mathbf{s}_1^T, \cdots, \mathbf{s}_{U_r}^T]^T \in \mathbb{C}^{U_r \times T_p}$ which satisfies orthogonality that the pilot signals transmitted by different RF chains of MS are orthogonal to each other, i.e.,

$$\begin{aligned} \mathbf{s}_i \mathbf{s}_j^H &= \sum_{t=1}^{T_p} s_i(t) s_j^*(t) \\ &= \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \quad i, j = 1, \cdots, U_r, \end{aligned} \tag{11}$$

where $T_p > U_r$ to ensure orthogonality. Thus, the output signal $\mathbf{Y}_m(n) \in \mathbb{C}^{Q_r \times T_p}$ of the receiving beamformer at the m-th subcarrier and n-th block is

$$\mathbf{Y}_m(n) = \mathbf{W}^H \mathbf{H}_m(n) \mathbf{F}^* \mathbf{S} + \mathbf{W}^H \mathbf{N}_m(n), \tag{12}$$

where $\mathbf{N}_m(n) \in \mathbb{C}^{Q_a \times T_p}$ is the received noise matrix, each element of which satisfies the i.i.d. complex Gaussian distribution with $\mathcal{CN}(0, \sigma_0^2)$, and the beamformers connected with ULAs at the MS and BS are designed by the DFT beamspace framework [10], i.e.,

$$\begin{aligned} \mathbf{F} &= \left[\mathbf{a}_U(\tilde{\phi}_1) \quad \mathbf{a}_U(\tilde{\phi}_2) \quad \cdots \quad \mathbf{a}_U(\tilde{\phi}_{U_r})\right], \\ \mathbf{W} &= \left[\mathbf{a}_B(\tilde{\theta}_1) \quad \mathbf{a}_B(\tilde{\theta}_2) \quad \cdots \quad \mathbf{a}_B(\tilde{\theta}_{Q_r})\right], \end{aligned} \tag{13}$$

where $\tilde{\phi}_i = (i-1)/U_a$, $i = 1, \cdots, U_r$ and $\tilde{\theta}_j = (j-1)/Q_a$, $j = 1, \cdots, Q_r$. Furthermore, the beamformer connected with UPA at the BS satisfies $\mathbf{W} = \mathbf{W}_x \otimes \mathbf{W}_z$, where $\mathbf{W}_x \in \mathbb{C}^{Q_a^x \times Q_r^x}$ and $\mathbf{W}_z \in \mathbb{C}^{Q_a^z \times Q_r^z}$ are respectively the analog RF combiners employed at the BS in the $\mathcal{X}$ direction and $\mathcal{Z}$ direction [36], both of which are designed by the DFT beamspace framework as (13).

According to the orthogonality of pilot signals in (11), the signal $\mathbf{y}_{m,i}(n) \in \mathbb{C}^{Q_r}$ from the i-th RF chain of MS to all RF chains of BS at the m-th subcarrier and n-th block is

$$\begin{aligned} \mathbf{y}_{m,i}(n) &= \mathbf{Y}_m(n) \mathbf{s}_i^H \\ &= \mathbf{W}^H \mathbf{H}_m(n) \mathbf{F}^* \mathbf{S} \mathbf{s}_i^H + \mathbf{W}^H \mathbf{N}_m(n) \mathbf{s}_i^H \\ &= \mathbf{W}^H \mathbf{H}_m(n) \mathbf{f}_i^* + \mathbf{n}_{m,i}(n), \end{aligned} \tag{14}$$

where $\mathbf{f}_i \in \mathbb{C}^{U_a}$ is the i-th column of matrix $\mathbf{F}$ and $\mathbf{n}_{m,i}(n) = \mathbf{W}^H \mathbf{N}_m(n) \mathbf{s}_i^H \sim \mathcal{CN}\left(0, \sigma^2 \mathbf{I}_{Q_r}\right)$.

III. Channel Estimation Strategy

The four parameters of mmWave wideband channels with ULAs are first estimated, where the transmission delays, AODs and AOAs are simultaneously estimated by the proposed beamspace MUSIC algorithm, and the path gains are estimated by the LS method. Afterwards, the proposed channel estimator is extend to mmWave MIMO-OFDM systems with UPAs.

A. Beamspace MUSIC for Wideband Channels

To simultaneously estimate the transmission delays, AOAs and AODs of mmWave wideband channels, the signal $\mathbf{y}_{m,i}(n)$ in (14) should be first reconstructed. Take (3) into (14), i.e.,

$$\begin{aligned} \mathbf{y}_{m,i}(n) &= \sum_{l=1}^{L} \alpha_l(n) e_m(\tau_l) \left(\mathbf{f}_i^H \mathbf{a}_U(\phi_l)\right) \left(\mathbf{W}^H \mathbf{a}_B(\theta_l)\right) + \mathbf{n}_{m,i}(n) \\ &= \Big[e_m(\tau_1) \left(\mathbf{f}_i^H \mathbf{a}_U(\phi_1)\right) \left(\mathbf{W}^H \mathbf{a}_B(\theta_1)\right) \cdots \\ &\quad e_m(\tau_L) \left(\mathbf{f}_i^H \mathbf{a}_U(\phi_L)\right) \left(\mathbf{W}^H \mathbf{a}_B(\theta_L)\right)\Big] \boldsymbol{\alpha}(n) + \mathbf{n}_{m,i}(n) \\ &= \mathbf{e}_m \odot \left(\mathbf{f}_i^H \mathbf{A}_U\right) \odot \left(\mathbf{W}^H \mathbf{A}_B\right) \boldsymbol{\alpha}(n) + \mathbf{n}_{m,i}(n), \end{aligned} \tag{15}$$

where

$$\begin{aligned} \mathbf{e}_m &= [e_m(\tau_1), \cdots, e_m(\tau_L)] \in \mathbb{C}^{1 \times L}, \\ \mathbf{A}_U &= [\mathbf{a}_U(\phi_1), \cdots, \mathbf{a}_U(\phi_L)] \in \mathbb{C}^{U_a \times L}, \\ \mathbf{A}_B &= [\mathbf{a}_B(\theta_1), \cdots, \mathbf{a}_B(\theta_L)] \in \mathbb{C}^{Q_a \times L}, \\ \boldsymbol{\alpha}(n) &= [\alpha_1(n), \cdots, \alpha_L(n)]^T \in \mathbb{C}^{L \times 1}. \end{aligned}$$

Stacking U_r vectors $\mathbf{y}_{m,i}(n)$ in a column, an U_rQ_r vector is obtained as

$$\begin{aligned}\mathbf{y}_m(n) &= \left[\mathbf{y}_{m,1}^T(n),\cdots,\mathbf{y}_{m,U_r}^T(n)\right]^T \\ &= \mathbf{e}_m \odot \begin{bmatrix}\mathbf{f}_1^H\mathbf{A}_U \\ \vdots \\ \mathbf{f}_{U_r}^H\mathbf{A}_U\end{bmatrix} \odot \left(\mathbf{W}^H\mathbf{A}_B\right)\boldsymbol{\alpha}(n) + \mathbf{n}_m(n) \\ &= \mathbf{e}_m \odot \left(\mathbf{F}^H\mathbf{A}_U\right) \odot \left(\mathbf{W}^H\mathbf{A}_B\right)\boldsymbol{\alpha}(n) + \mathbf{n}_m(n)\end{aligned} \tag{16}$$

where

$$\mathbf{n}_m(n) = \left[\mathbf{n}_{m,1}^T(n),\cdots,\mathbf{n}_{m,U_r}^T(n)\right]^T \sim \mathcal{CN}\left(0,\sigma^2\mathbf{I}_{U_rQ_r}\right).$$

To accurately estimate the transmission delays of mmWave channels, it is necessary to acquire the characteristics of transmission delays from the signals on all selected subcarriers. Then, (16) can be reconstructed and an MU_rQ_r vector is obtained as

$$\begin{aligned}\mathbf{y}(n) &= \left[\mathbf{y}_1^T(n),\cdots,\mathbf{y}_M^T(n)\right]^T \\ &= \begin{bmatrix}\mathbf{e}_1 \\ \vdots \\ \mathbf{e}_M\end{bmatrix} \odot \left(\mathbf{F}^H\mathbf{A}_U\right) \odot \left(\mathbf{W}^H\mathbf{A}_B\right)\boldsymbol{\alpha}(n) + \mathbf{n}(n) \\ &= \left[\mathbf{e}(\tau_1)\cdots\mathbf{e}(\tau_L)\right] \odot \left(\mathbf{F}^H\mathbf{A}_U\right) \odot \left(\mathbf{W}^H\mathbf{A}_B\right)\boldsymbol{\alpha}(n) + \mathbf{n}(n) \\ &= \underbrace{\mathbf{E} \odot \left(\mathbf{F}^H\mathbf{A}_U\right) \odot \left(\mathbf{W}^H\mathbf{A}_B\right)}_{\triangleq\mathbf{B}}\boldsymbol{\alpha}(n) + \mathbf{n}(n),\end{aligned} \tag{17}$$

where

$$\mathbf{E} = \left[\mathbf{e}(\tau_1),\cdots,\mathbf{e}(\tau_L)\right] \in \mathbb{C}^{M\times L},$$
$$\mathbf{n}(n) = \left[\mathbf{n}_1^T(n),\cdots,\mathbf{n}_M^T(n)\right]^T \sim \mathcal{CN}\left(0,\sigma^2\mathbf{I}_{MU_rQ_r}\right).$$

and $\mathbf{B} \in \mathbb{C}^{MU_rQ_r\times L}$ is the spatial matrix, whose l-th column is expressed as

$$\mathbf{b}(\tau_l,\phi_l,\theta_l) = \mathbf{e}(\tau_l) \otimes \left(\mathbf{F}^H\mathbf{a}_U(\phi_l)\right) \otimes \left(\mathbf{W}^H\mathbf{a}_B(\theta_l)\right). \tag{18}$$

The covariance matrix of $\mathbf{y}(n)$ can be calculated by

$$\mathbf{R}_y \triangleq \mathbb{E}\left[\mathbf{y}(n)\mathbf{y}(n)^H\right] = \frac{1}{N}\sum_{n=1}^{N}\mathbf{y}(n)\mathbf{y}(n)^H. \tag{19}$$

Take (17) into (19), i.e.,

$$\begin{aligned}\mathbf{R}_y &= \mathbf{B}\mathbb{E}\left[\boldsymbol{\alpha}(n)\boldsymbol{\alpha}(n)^H\right]\mathbf{B}^H + \mathbb{E}\left[\mathbf{n}(n)\mathbf{n}(n)^H\right] \\ &= \mathbf{B}\mathbf{R}_\alpha\mathbf{B}^H + \sigma^2\mathbf{I}_{MU_rQ_r},\end{aligned} \tag{20}$$

where $\mathbf{R}_\alpha$ is the covariance matrix of $\boldsymbol{\alpha}(n)$. The eigenvalue decomposition of $\mathbf{R}_y$ is

$$\mathbf{R}_y = \mathbf{U}\mathrm{diag}\left(\bar{\eta}_1,\cdots,\bar{\eta}_{MU_rQ_r}\right)\mathbf{U}^H, \tag{21}$$

where $\mathbf{U}$ is the eigenvector matrix and $\{\bar{\eta}_j\}_{j=1}^{MU_rQ_r}$ are the eigenvalues which satisfy

$$\bar{\eta}_1 > \bar{\eta}_2 > \cdots > \bar{\eta}_L > \bar{\eta}_{L+1} = \cdots = \bar{\eta}_{MU_rQ_r} = \sigma^2. \tag{22}$$

Define the diagonal matrices of large eigenvalues and small eigenvalues respectively as

$$\begin{aligned}\boldsymbol{\Sigma}_s &= \mathrm{diag}\left(\bar{\eta}_1,\cdots,\bar{\eta}_L\right), \\ \boldsymbol{\Sigma}_n &= \mathrm{diag}\left(\bar{\eta}_{L+1},\cdots,\bar{\eta}_{MU_rQ_r}\right) = \sigma^2\mathbf{I}_{MU_rQ_r-L}.\end{aligned} \tag{23}$$

Then, (21) can be rewritten as

$$\begin{aligned}\mathbf{R}_y &= \mathbf{U}\begin{bmatrix}\boldsymbol{\Sigma}_s & \\ & \boldsymbol{\Sigma}_n\end{bmatrix}\mathbf{U}^H \\ &= \begin{bmatrix}\mathbf{U}_s & \mathbf{U}_n\end{bmatrix}\begin{bmatrix}\boldsymbol{\Sigma}_s & \\ & \boldsymbol{\Sigma}_n\end{bmatrix}\begin{bmatrix}\mathbf{U}_s & \mathbf{U}_n\end{bmatrix}^H \\ &= \mathbf{U}_s\boldsymbol{\Sigma}_s\mathbf{U}_s^H + \mathbf{U}_n\boldsymbol{\Sigma}_n\mathbf{U}_n^H,\end{aligned} \tag{24}$$

where $\mathbf{U}_s \in \mathbb{C}^{MU_rQ_r\times L}$ is the signal subspace corresponding to the large eigenvalues, and $\mathbf{U}_n \in \mathbb{C}^{MU_rQ_r\times(MU_rQ_r-L)}$ is the noise subspace corresponding to the small eigenvalues. Ideally, signal subspace and noise subspace are orthogonal to each other, so the steering vectors in signal subspace are also orthogonal to noise subspace [24], i.e.,

$$\left|\mathbf{B}^H\mathbf{U}_n\right| = 0 \quad and \quad \left|\mathbf{b}(\tau_l,\phi_l,\theta_l)^H\mathbf{U}_n\right| = 0. \tag{25}$$

Therefore, the three-dimensional spatial spectrum of the proposed beamspace MUSIC algorithm is calculated by

$$P(\tau,\phi,\theta) = \frac{1}{\left|\mathbf{b}(\tau,\phi,\theta)^H\mathbf{U}_n\right|}, \tag{26}$$

where $\theta \in [\tilde{\theta}_1,\tilde{\theta}_{Q_r}]$, $\phi \in \left[\tilde{\phi}_1,\tilde{\phi}_{U_r}\right]$ and $\tau \in [0,\tau_{max}]$ with maximum transmission delay τ_{max}.

According to (25), $\mathbf{b}(\tau_l,\phi_l,\theta_l)$ and $\mathbf{U}_n$ are theoretically orthogonal to each other, i.e., the denominator of $P(\tau,\phi,\theta)$ is zero when $(\tau,\phi,\theta) = (\tau_l,\phi_l,\theta_l)$. However, due to the existence of noise, the denominator of $P(\tau,\phi,\theta)$ is not zero but a very small value, i.e., there is a spectrum peak on the spatial spectrum. Therefore, the estimates of AOAs, AODs and transmission delays of L paths can be obtained by searching for L spectrum peaks on the spatial spectrum, which can be described as

$$\begin{aligned}\left(\hat{\tau}_l,\hat{\phi}_l,\hat{\theta}_l\right) =& \arg\max_{\tau,\phi,\theta} P(\tau,\phi,\theta) \quad l = 1,\cdots,L \\ =& \arg\min_{\tau,\phi,\theta} \left|\mathbf{b}(\tau,\phi,\theta)^H\mathbf{U}_n\right| \\ \text{s.t.} \quad & \theta \in [\tilde{\theta}_1,\tilde{\theta}_{Q_r}], \phi \in \left[\tilde{\phi}_1,\tilde{\phi}_{U_r}\right], \tau \in [0,\tau_{max}],\end{aligned} \tag{27}$$

where $\hat{\theta}_l$, $\hat{\phi}_l$ and $\hat{\tau}_l$ are respectively the estimates of the AOA, AOD and transmission delay of the l-th path. To solve the problem described in (27), the traditional MUSIC algorithm exploits a grid search method which has high accuracy but large complexity, because the spatial spectrum in (26) is three-dimensional, whose three dimensions are respectively the index of subcarriers, the index of RF chains equipped at the MS and the index of RF chains equipped at the BS. For this reason, a chicken swarm optimization (CSO) algorithm is adopted to optimize the spectral peak search of the proposed beamspace MUSIC algorithm.

Remark 2: In this paper, the number of paths is assumed to be known, which can be generally estimated by the AIC and MDL methods [37].

B. Multi-Spectral Peak Search Method

Chicken swarm optimization (CSO) algorithm which describes the optimization problem as the process of chickens

searching for food in a multi-dimensional space is a swarm intelligent stochastic optimization algorithm [38]. Compared with other optimization algorithms, CSO has the advantages of fast convergence speed and strong searching ability. The following rules are used to idealize this algorithm [39].

- The whole chicken swarm is divided into several groups according to the fitness values of chickens, and the number of groups are equal to that of roosters. Moreover, hens randomly select groups to determine the mate relationship with roosters, and chicks randomly select hens to follow to establish the mother-child relationship.
- In each group, a rooster is dominant with the strongest searching ability, so its fitness value is the lowest; Hens have slightly worse search ability and follow the rooster closely to search for food, so their fitness values are slightly higher; Chicks have the worst search ability and only search for food around hens, which can realize local search function, so their fitness values are the largest.
- The grouping and relationship of chickens can be changed when the number of iterations is the multiple of the update interval of chicken swarm. Besides, the positions of chickens are updated according to their respective position updating strategies when the fitness values of new positions are smaller than those of original positions. This process continues until the maximum number of iterations is reached or the global optimal solution is found.

Here, the position updating strategy of roosters is

$$\begin{aligned} x_{i,j}^{t+1} &= x_{i,j}^{t}\left[1+\mathrm{randn}\left(0,\sigma^2\right)\right], \\ \sigma^2 &= \begin{cases} 1, & f_i \le f_k \\ \exp\left(\frac{f_k-f_i}{|f_i|+\varepsilon}\right), & else \end{cases}, \end{aligned} \tag{28}$$

where $x_{i,j}^{t}$ represents the position of the i-th rooster at the j-th dimension and the t-th moment, $\mathrm{randn}\left(0,\sigma^2\right)$ represents the Gaussian distribution function with mean 0 and variance σ^2, $\varepsilon>0$ represents the infinite decimal, f_i represents the fitness value of the i-th rooster, and f_k represents the fitness value of the k-th $(k\neq i)$ rooster which is randomly selected from the chicken swarm. The position updating strategy of hens is

$$\begin{aligned} x_{i,j}^{t+1} &= x_{i,j}^{t} + k_1\times\mathrm{rand}\times\left(x_{r_1,j}^{t}-x_{i,j}^{t}\right) \\ &\quad + k_2\times\mathrm{rand}\times\left(x_{r_2,j}^{t}-x_{i,j}^{t}\right), \\ k_1 &= \exp\left(\frac{f_i-f_{r_1}}{|f_i|+\varepsilon}\right), \quad k_2=\exp\left(f_{r_2}-f_i\right), \end{aligned} \tag{29}$$

where $x_{i,j}^{t}$ represents the position of the i-th hen at the j-th dimension and the t-th moment, rand represents the uniform distribution function with mean 0 and variance 1, r_1 represents the index of the companion rooster of the i-th hen, and $r_2\,(r_1\neq r_2)$ represents the index of the rooster randomly selected from the chicken swarm. The chick position updating strategy is

$$x_{i,j}^{t+1} = x_{i,j}^{t} + F\left(x_{m,j}^{t}-x_{i,j}^{t}\right), \quad F\in[0,2], \tag{30}$$

where $x_{i,j}^{t}$ represents the position of the i-th chick at the j-th dimension and the t-th moment, and m represents the index of the hen followed by the i-th chick.

From the above analysis, it can be seen that the spectral peak search of the proposed beamspace MUSIC algorithm can be optimized by the CSO algorithm. In this case, the objective optimization problem is that described in (27), and the global optimal solution is the coordinate corresponding to the spectral peak on the spatial spectrum. However, CSO algorithm output only one global optimal solution at a time. To estimate the parameters of multiple paths, it is necessary to find multiple spectral peaks on the spatial spectrum of MUSIC algorithm. Thus, a multi-spectral peak search method is designed by skillfully combining the proposed beamspace MUSIC algorithm with the CSO algorithm according to the principle of dichotomy as follows.

1) *Initialization of Parameters:* The maximum number of iterations is $L-1$, the initial set of global optimal solutions is $\mathcal{I}=\{(0,\tilde{\phi}_1,\tilde{\theta}_1),(\tau_{\max},\tilde{\phi}_{U_r},\tilde{\theta}_{Q_r})\}$, the initial search scope is $\{(\tau,\phi,\theta)\,|(0,\tilde{\phi}_1,\tilde{\theta}_1)\le(\tau,\phi,\theta)\le(\tau_{\max},\tilde{\phi}_{U_r},\tilde{\theta}_{Q_r})\}$, the fitness values of coordinates $(0,\tilde{\phi}_1,\tilde{\theta}_1)$ and $(\tau_{\max},\tilde{\phi}_{U_r},\tilde{\theta}_{Q_r})$ are set to $+\infty$.
2) *Calculation of Global Optimal Solution:* Execute the CSO algorithm in the initial search scope, and then output a global optimal solution (τ_0,ϕ_0,θ_0).
3) *Update of set $\mathcal{I}$:* Add (τ_0,ϕ_0,θ_0) to set $\mathcal{I}$ and sort $\mathcal{I}$ in ascending order. The element $(\tau_{\min},\phi_{\min},\theta_{\min})$ with the smallest fitness value in set $\mathcal{I}$ is the coordinate corresponding to the first spectral peak on the spatial spectrum.
4) *Starting Iterations:* Set the fitness value of coordinate $(\tau_{\min},\phi_{\min},\theta_{\min})$ to $+\infty$.
5) *Update of Search Scopes:* Update the search scopes to $\{(\tau,\phi,\theta)\,|\mathcal{I}_{i-1}\le(\tau,\phi,\theta)\le\mathcal{I}_i\}$ and $\{(\tau,\phi,\theta)\,|\mathcal{I}_i\le(\tau,\phi,\theta)\le\mathcal{I}_{i+1}\}$, where $\mathcal{I}_i$ represents the i-th element of set $\mathcal{I}$, whose value is $(\tau_{\min},\phi_{\min},\theta_{\min})$.
6) *Calculation of Global Optimal Solutions:* Execute the CSO algorithm in these two search scopes respectively to obtain two global optimal solutions $(\tau_{i1},\phi_{i1},\theta_{i1})$ and $(\tau_{i2},\phi_{i2},\theta_{i2})$.
7) *Update of set $\mathcal{I}$:* Add $(\tau_{i1},\phi_{i1},\theta_{i1})$ and $(\tau_{i2},\phi_{i2},\theta_{i2})$ to set $\mathcal{I}$ and sort $\mathcal{I}$ in ascending order. Then, the element $(\tau_{\min},\phi_{\min},\theta_{\min})$ with the smallest fitness value in set $\mathcal{I}$ is the coordinate corresponding to another spectral peak on the spatial spectrum.
8) *Stopping Iterations:* (4)-(7) continues until the maximum number of iterations is reached, i.e., the coordinates corresponding to L spectral peaks on the spatial spectrum are obtained.

C. LS Method for Path Gain Estimation

After obtaining the estimates of transmission delays, AOAs and AODs, the estimated spatial matrix $\hat{\mathbf{B}}$ can be calculated. Then, the path gains can be estimated by the LS method, i.e.,

$$\hat{\boldsymbol{\alpha}}(n)=\left(\hat{\mathbf{B}}^H\hat{\mathbf{B}}\right)^{-1}\hat{\mathbf{B}}^H\mathbf{y}(n), \quad n=1,\cdots,N. \tag{31}$$

Note that the l-th element of $\hat{\boldsymbol{\alpha}}(n)$ is related to the l-th column of $\hat{\mathbf{B}}$. Then, the channel matrix at the m-th subcarrier and n-th

block can be estimated by

$$\hat{\mathbf{H}}_m(n) = \sum_{l=1}^{L} \hat{\alpha}_l(n)\, e_m(\hat{\tau}_l)\, \mathbf{a}_B\left(\hat{\theta}_l\right) \mathbf{a}_U^T\left(\hat{\phi}_l\right). \tag{32}$$

The procedure of the proposed channel parameter estimator for mmWave wideband channels with ULAs is summarized in Algorithm 1.

Algorithm 1 Channel Estimator for mmWave Wideband Channels with ULAs

Input: M, N, L, $\mathbf{y}(n)$
Output: $\hat{\tau}_l$, $\hat{\phi}_l$, $\hat{\theta}_l$, $\hat{\mathbf{H}}_m(n)$

1: Calculate the covariance matrix $\mathbf{R}_y$ of $\mathbf{y}(n)$ by (19).
2: Perform the eigenvalue decomposition of $\mathbf{R}_y$ and obtain the noise subspace $\mathbf{U}_n$ by (21).
3: Calculate the spatial spectrum $P(\tau,\phi,\theta)$ by (26).
4: Initialization: set $\mathcal{I} = \{(0,\tilde{\phi}_1,\tilde{\theta}_1),(\tau_{\max},\tilde{\phi}_{U_r},\tilde{\theta}_{Q_\mathrm{t}})\}$, search scope is $\{(\tau,\phi,\theta)\,|(0,\tilde{\phi}_1,\tilde{\theta}_1) \le (\tau,\phi,\theta) \le (\tau_{\max},\tilde{\phi}_{U_r},\tilde{\theta}_{Q_r})\}$, the fitness values of $(0,\tilde{\phi}_1,\tilde{\theta}_1)$ and $(\tau_{\max},\tilde{\phi}_{U_r},\tilde{\theta}_{Q_r})$ are set to $+\infty$.
5: Execute CSO algorithm in the initial search scope and obtain the global optimal solution $(\tau_{\min},\phi_{\min},\theta_{\min})$, i.e., the coordinate of the first spectral peak.
6: Add $(\tau_{\min},\phi_{\min},\theta_{\min})$ to set $\mathcal{I}$ and sort $\mathcal{I}$ in ascending order.
7: **for** $l = 2$ to L **do**
8: Set the fitness value of $\mathcal{I}_i = (\tau_{\min},\phi_{\min},\theta_{\min})$ to $+\infty$.
9: Update the search scopes to $\{(\tau,\phi,\theta)\,|\mathcal{I}_{i-1} \le (\tau,\phi,\theta) \le \mathcal{I}_i\}$ and $\{(\tau,\phi,\theta)\,|\mathcal{I}_i \le (\tau,\phi,\theta) \le \mathcal{I}_{i+1}\}$.
10: Execute CSO algorithm in these two search scopes respectively and obtain two new global optimal solutions $(\tau_{i1},\phi_{i1},\theta_{i1})$ and $(\tau_{i2},\phi_{i2},\theta_{i2})$.
11: Add these two new global optimal solutions to set $\mathcal{I}$ and sort $\mathcal{I}$ in ascending order. Then, the smallest fitness value element $(\tau_{\min},\phi_{\min},\theta_{\min})$ in set $\mathcal{I}$ is the coordinate of another spectral peak.
12: **end for**
13: Calculate the path gain $\hat{\boldsymbol{\alpha}}(n)$ according to (31).
14: Calculate the channel matrix $\hat{\mathbf{H}}_m(n)$ according to (32).

D. Channel Estimator for UPA

Take (10) into (14) as (33) shown at the bottom of this page. Here, $\mathbf{A}_B^x = [\mathbf{a}_B^x(\theta_1^x),\cdots,\mathbf{a}_B^x(\theta_L^x)] \in \mathbb{C}^{Q_a^x \times L}$ and $\mathbf{A}_B^z = [\mathbf{a}_B^z(\theta_1^z),\cdots,\mathbf{a}_B^z(\theta_L^z)] \in \mathbb{C}^{Q_a^z \times L}$. Stacking U_r vectors $\mathbf{y}_{m,i}(n)$ in a column, an U_rQ_r vector is obtained as (34) shown at the top of next page.

To accurately estimate the transmission delays of mmWave channels, it is necessary to acquire the characteristics of transmission delays from the signals on all selected subcarriers. Then, (34) can be reconstructed as (17) and an MU_rQ_r vector is obtained as

$$\begin{aligned}\mathbf{y}(n) &= \left[\mathbf{y}_1^T(n),\cdots,\mathbf{y}_M^T(n)\right]^T \\ &= \underbrace{\mathbf{E} \odot \left(\mathbf{F}^H\mathbf{A}_U\right) \odot \left(\mathbf{W}_x^H\mathbf{A}_B^x\right) \odot \left(\mathbf{W}_z^H\mathbf{A}_B^z\right)}_{\triangleq \mathbf{B}} \boldsymbol{\alpha}(n) + \mathbf{n}(n),\end{aligned} \tag{35}$$

where the l-th column of spatial matrix $\mathbf{B}$ is shown as (36) at the top of next page.

The following procedure is the same as Algorithm 1, whose four-dimensional spatial spectrum of the proposed beamspace MUSIC algorithm for UPA is calculated by

$$P(\tau,\phi,\theta^x,\theta^z) = \frac{1}{\left|\mathbf{b}(\tau,\phi,\theta^x,\theta^z)^H\,\mathbf{U}_n\right|}. \tag{37}$$

Then, the estimates of τ_l, ϕ_l, θ_l^x and θ_l^z denoted as $\hat{\tau}_l$, $\hat{\phi}_l$, $\hat{\theta}_l^x$ and $\hat{\theta}_l^z$ are output, and the estimates of μ_l and ν_l are respectively $\hat{\mu}_l = \arcsin\left(\frac{\lambda}{d}\hat{\theta}_l^z\right)$ and $\hat{\nu}_l = \arccos\left(\frac{\lambda}{d\cos\hat{\mu}_l}\hat{\theta}_l^x\right)$.

Remark 3: (22) depends on the assumption that $\mathbf{B}$ has a full-column rank, which is ensured by $L < MU_rQ_r$. Then, the channel estimation accuracy of the proposed channel estimator is constrained by the number of RF chains. Due to the sparse nature of mmWave channels, L is generally small [32], [33]. Thus, $L < MU_rQ_r$ can be easily satisfied.

IV. Cramér–Rao Bound

Cramér–Rao bound (CRB) is a lower bound on the variance of any unbiased estimator derived under a specific model, which can provide a benchmark for evaluating the accuracy of different parameter estimators [40], [41]. In this paper, the estimates of transmission delays, AOAs and AODs are first acquired by the proposed beamspace MUSIC algorithm, and then the path gains are estimated by the least-squares method. Due to the affect of zero-mean i.i.d. Gaussian noise, the proposed beamspace MUSIC algorithm can produce maximum likelihood estimates, and the least square estimation is a maximum likelihood estimation when the estimation error is an i.i.d. Gaussian random variable. Thus, the proposed channel parameter estimator is essentially a maximum likelihood estimator, which is asymptotically unbiased and asymptotically achieves the CRB results for enough data samples.

From (17), it can be found that the CRB results of θ_l, ϕ_l and τ_l are derived from N independent identically distributed

$$\begin{aligned}\mathbf{y}_{m,i}(n) &= \sum_{l=1}^{L} \alpha_l(n)\, e_m(\tau_l) \left(\mathbf{f}_i^H\mathbf{a}_U(\phi_l)\right) \left(\mathbf{W}^H\mathbf{a}_B^x(\theta_l^x) \otimes \mathbf{a}_B^z(\theta_l^z)\right) + \mathbf{n}_{m,i}(n) \\ &= \sum_{l=1}^{L} \alpha_l(n)\, e_m(\tau_l) \left(\mathbf{f}_i^H\mathbf{a}_U(\phi_l)\right) \left(\mathbf{W}_x \otimes \mathbf{W}_z\right)^H \left(\mathbf{a}_B^x(\theta_l^x) \otimes \mathbf{a}_B^z(\theta_l^z)\right) + \mathbf{n}_{m,i}(n) \\ &= \sum_{l=1}^{L} \alpha_l(n)\, e_m(\tau_l) \left(\mathbf{f}_i^H\mathbf{a}_U(\phi_l)\right) \left(\mathbf{W}_x^H\mathbf{a}_B^x(\theta_l^x) \otimes \mathbf{W}_z^H\mathbf{a}_B^z(\theta_l^z)\right) + \mathbf{n}_{m,i}(n) \\ &= \mathbf{e}_m \odot \left(\mathbf{f}_i^H\mathbf{A}_U\right) \odot \left(\mathbf{W}_x^H\mathbf{A}_B^x\right) \odot \left(\mathbf{W}_z^H\mathbf{A}_B^z\right) \boldsymbol{\alpha}(n) + \mathbf{n}_{m,i}(n),\end{aligned} \tag{33}$$

$$\mathbf{y}_m(n) = \left[\mathbf{y}_{m,1}^T(n), \cdots, \mathbf{y}_{m,U_r}^T(n)\right]^T = \mathbf{e}_m \odot \left(\mathbf{F}^H \mathbf{A}_U\right) \odot \left(\mathbf{W}_x^H \mathbf{A}_B^x\right) \odot \left(\mathbf{W}_z^H \mathbf{A}_B^z\right) \boldsymbol{\alpha}(n) + \mathbf{n}_m(n). \tag{34}$$

$$\mathbf{b}(\tau_l, \phi_l, \theta_l^x, \theta_l^z) = \mathbf{e}(\tau_l) \otimes \left(\mathbf{F}^H \mathbf{a}_U(\phi_l)\right) \otimes \left(\mathbf{W}_x^H \mathbf{a}_B^x(\theta_l^x)\right) \otimes \left(\mathbf{W}_z^H \mathbf{a}_B^z(\theta_l^z)\right). \tag{36}$$

TABLE I
INITIALIZATION PARAMETERS OF CSO ALGORITHM

Maximum Number of Iterations	100
Chicken Swarm Size	100
Rooster Proportion	0.15
Hen Proportion	0.7
Chick Proportion	0.15
Mother Hen Proportion	0.5
Chicken Swarm Dimension	3
Chicken Swarm Update Interval	5
Lower Search Bound	0,0,0
Upper Search Bound	$1e-7, 0.8481, 1.0654$

observations, and the CRB result of $\alpha_l(n)$ is derived from one independent identically distributed observation. Hence, due to the independent observation additivity of CRB, the CRB result of θ_l can be calculated as

$$\mathrm{CRB}(\theta_l) = \left[\left(N\boldsymbol{\Omega}_n(\boldsymbol{\kappa})\right)^{-1}\right]_{l,l} = \frac{1}{N}\left[\boldsymbol{\Omega}_n(\boldsymbol{\kappa})^{-1}\right]_{l,l}. \tag{38}$$

Similarly, the CRB results of ϕ_l, τ_l and $\alpha_l(n)$ are respectively calculated as

$$\begin{aligned}
\mathrm{CRB}(\phi_l) &= \frac{1}{N}\left[\boldsymbol{\Omega}_n(\boldsymbol{\kappa})^{-1}\right]_{l+L\times 1, l+L\times 1}, \\
\mathrm{CRB}(\tau_l) &= \frac{1}{N}\left[\boldsymbol{\Omega}_n(\boldsymbol{\kappa})^{-1}\right]_{l+L\times 2, l+L\times 2}, \\
\mathrm{CRB}(\alpha_l(n)) &= \left[\boldsymbol{\Omega}_n(\boldsymbol{\kappa})^{-1}\right]_{l+L\times 3, l+L\times 3}.
\end{aligned} \tag{39}$$

The detailed calculation process of $\boldsymbol{\Omega}_n(\boldsymbol{\kappa})$ is demonstrated in Appendix.

V. SIMULATION RESULTS

Consider the mmWave MIMO-OFDM system where the BS is equipped with Q_a = 16 antennas and Q_r = 8 RF chains, and the MS is equipped with U_a = 8 antennas and U_r = 4 RF chains. Both of them are equipped with ULAs. Consider the channel model with L = 2 paths [13], where the path gain of each path is a random variable following $\alpha_l(n) \overset{i.i.d.}{\sim} \mathcal{CN}(0,1)$ [26], the transmission delays, physical AODs and physical AOAs are respectively generated with uniform distribution in $\left[0, 10^{-7}\right]$ seconds [22], $[0, 0.8481]$ and $[0, 1.0654]$. The total number of OFDM subcarriers is M_0 = 128, where the first M = 8 subcarriers are selected for channel estimation. The sampling rate is f_s = 0.32GHz [22], and the total number of time blocks is N = 40. The signal-to-noise ratio (SNR) is defined as $\|\mathbf{y}(n) - \mathbf{n}(n)\|_F^2 / \|\mathbf{n}(n)\|_F^2$. Besides, the initialization parameters of CSO algorithm are shown in TABLE I, and the estimation accuracy of these channel parameters and channel matrix is measured by the normalized mean square error (NMSE), i.e.,

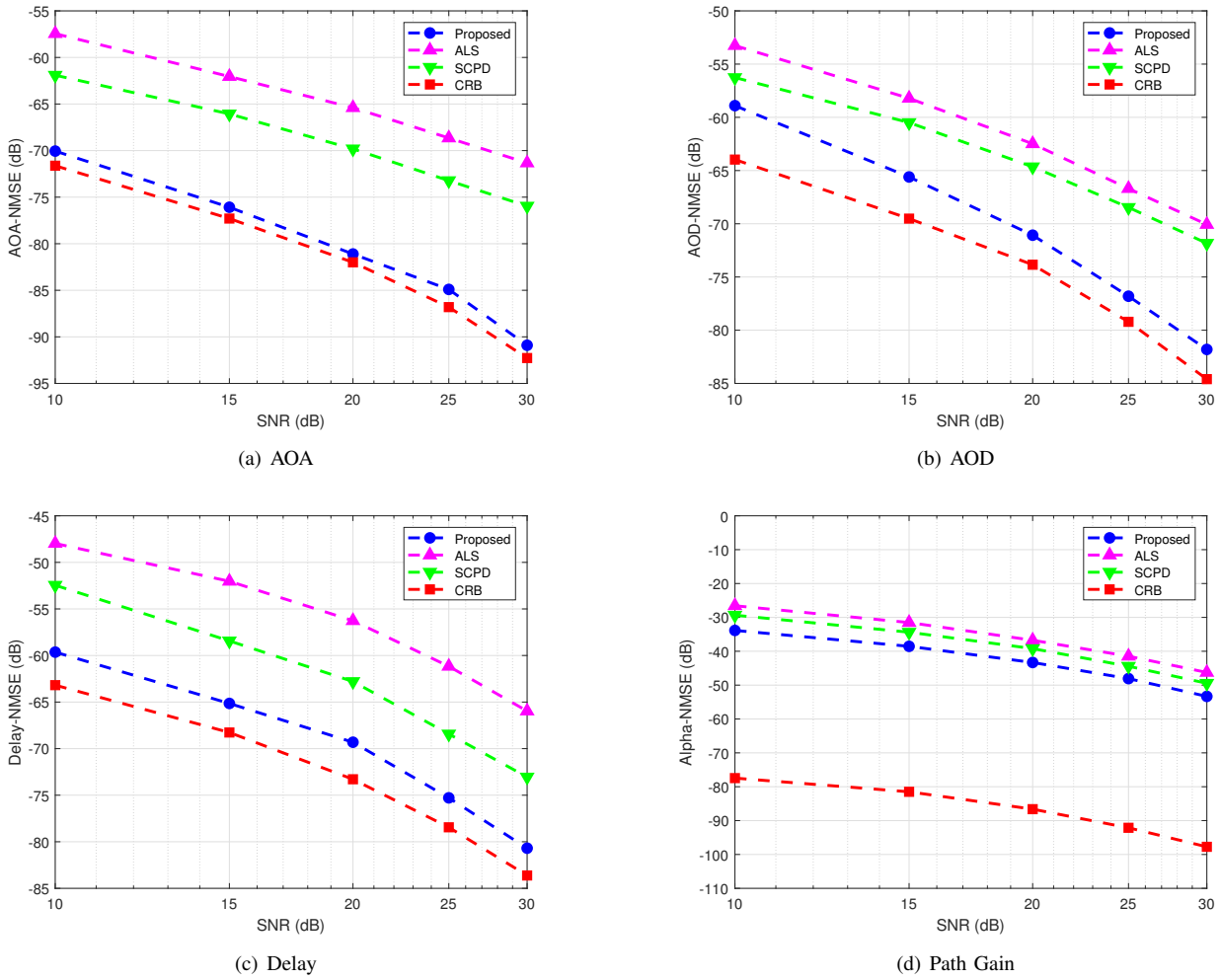
$$\begin{aligned}
\mathrm{NMSE}_\zeta &= \frac{\left\|\boldsymbol{\zeta} - \hat{\boldsymbol{\zeta}}\right\|_F^2}{\|\boldsymbol{\zeta}\|_F^2}, \\
\mathrm{NMSE}_\alpha &= \frac{\sum_{n=1}^{N} \|\boldsymbol{\alpha}(n) - \hat{\boldsymbol{\alpha}}(n)\|_F^2}{\sum_{n=1}^{N} \|\boldsymbol{\alpha}(n)\|_F^2}, \\
\mathrm{NMSE}_H &= \frac{\sum_{m=1}^{M}\sum_{n=1}^{N} \left\|\mathbf{H}_m(n) - \hat{\mathbf{H}}_m(n)\right\|_F^2}{\sum_{m=1}^{M}\sum_{n=1}^{N} \|\mathbf{H}_m(n)\|_F^2},
\end{aligned} \tag{40}$$

where $\boldsymbol{\zeta}$ can be replaced by $\boldsymbol{\tau} = [\tau_1, \cdots, \tau_L]^T$, $\boldsymbol{\phi} = [\phi_1, \cdots, \phi_L]^T$, $\boldsymbol{\theta} = [\theta_1, \cdots, \theta_L]^T$, $\boldsymbol{\nu} = [\nu_1, \cdots, \nu_L]^T$ and $\boldsymbol{\mu} = [\mu_1, \cdots, \mu_L]^T$.

Fig.4 illustrated the NMSE and CRB curves of these four channel parameters under different SNRs, where the proposed channel parameter estimator was compared with ALS [21] and SCPD [22]. It can be found that the estimation performance of these channel parameters improved exponentially with the increase of SNR. Specifically, the gap between NMSE and CRB curves of path gain was large because the LS method was susceptible to noise, while the gap between NMSE and CRB curves of AOA was very small. Besides, it can be found that the performance of the proposed channel parameter estimator was better than that of ALS and SCPD. These results indicated that the proposed channel parameter estimator was effective and highly accurate.

Fig.5 illustrated the performance of the proposed channel estimator under different numbers of paths. It can be found that the channel estimation accuracy of the proposed channel estimator decreased continuously with the increase of path number, because the spatial spectrum of MUSIC algorithm appeared spectrum ambiguity when the number of paths was large, i.e., the spectral peaks corresponding to different paths interfered with each other, which resulted in the difficulty of distinguishing every spectral peak on the spatial spectrum. Moreover, it can be found that the performance of the proposed channel estimator was still high when the number of paths was large, which indicated that the proposed channel estimator had better robustness.

Fig.6 illustrated the performance of the proposed channel estimator under different numbers of RF chains equipped at the BS, where L = 1 and the number of antennas equipped at the BS was twice that of RF chains. It can be found that



(a) AOA (b) AOD

(c) Delay (d) Path Gain

Fig. 4. Performance comparison of channel parameters under different SNRs

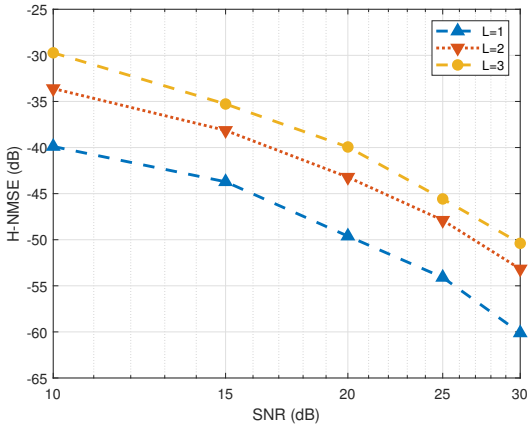


Fig. 5. Performance comparison of channel matrix under different numbers of paths

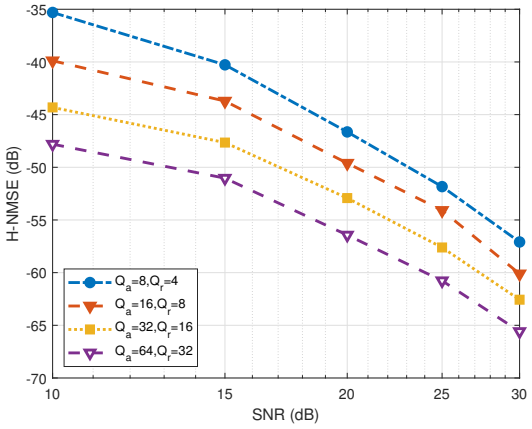
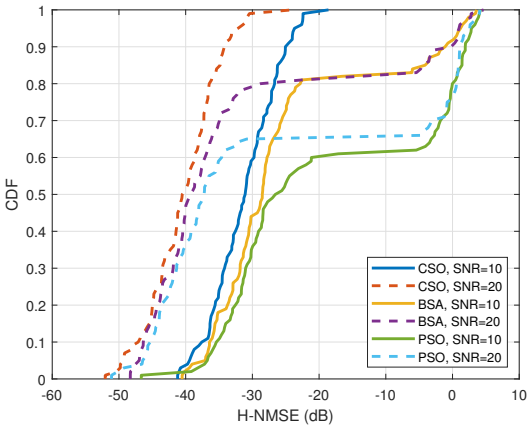


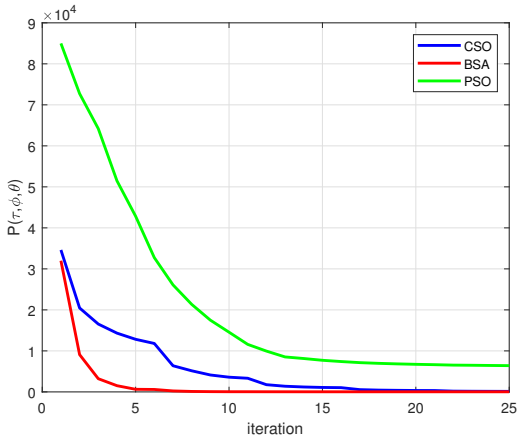
Fig. 6. Performance comparison of channel matrix under different numbers of RF chains equipped at the BS

the NMSE curves of the proposed channel estimator decreased continuously with the increase of the number of RF chains, and the performance of the proposed channel estimator was still high when the number of antennas was small or large, which indicated that the proposed channel estimator was suitable for both small-scale and large-scale MIMO-OFDM systems.

Fig.7(a) showed the cumulative distribution function (CDF) curves of different multi-spectral peak search methods under

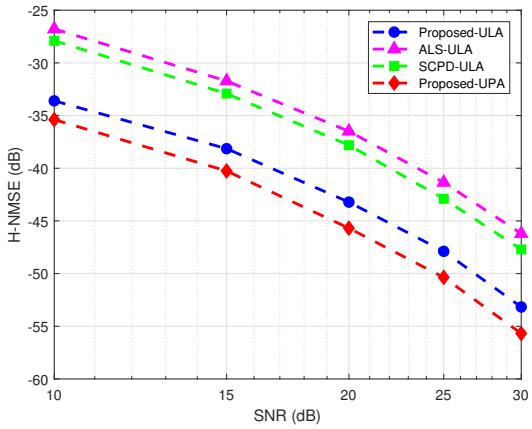


(a) Accuracy

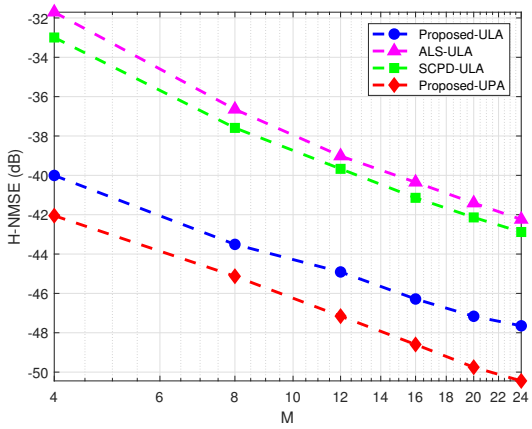


(b) Convergence Speed

Fig. 7. Performance comparison of different optimization algorithms



(a) NMSE comparison of different SNRs



(b) NMSE comparison of different subcarrier numbers

Fig. 8. Performance comparison of channel matrix under different SNRs and numbers of subcarriers

different SNRs, where the abscissa was the NMSE of channel matrix. It can be found that the proposed multi-spectral peak search method had a higher probability on the smaller NMSE of channel matrix than other methods respectively based on the BSA [42] and PSO [43], even when the SNR was small. Moreover, Fig.7(b) showed the convergence speed of different optimization algorithms, where the abscissa and ordinate were respectively the number of iterations and the value of spatial spectrum. It can be found that the convergence speed of CSO was slightly slower than that of BSA and much faster than that of PSO. These results indicated that the proposed multi-spectral peak search method was more accurate and better for solving the channel estimation problem studied in this paper.

Consider the mmWave MIMO-OFDM system where the BS and MS are respectively equipped with UPA and ULA. Here, $Q_a^x = 8$ antennas and $Q_r^x = 4$ RF chains of BS are in the $\mathcal{X}$ direction, $Q_a^z = 8$ antennas and $Q_r^z = 4$ RF chains of BS are in the $\mathcal{Z}$ direction. The e-angle and a-angle of each path are respectively generated with uniform distribution in $[0, 0.8481]$ and $[0, 0.5119]$.

Fig.8 respectively illustrated the NMSE curves of channel matrix under different SNRs and numbers of subcarriers. It can be found that the NMSE curves of these three channel estimation methods constantly decreased with the increase of SNR or the number of subcarriers, and the estimation accuracy of the proposed channel estimator was much higher than that of ALS and SCPD, even when the number of subcarriers or SNR was small. These results indicated that the proposed channel estimator had better channel estimation performance and can effectively reduce the number of subcarriers selected for channel estimation.

Specially, it can be found from Fig.8 that the performance of the proposed channel estimator for UPA was better than that for ULA, even when the number of subcarriers or SNR was small, because the total number of UPA antennas equipped at the BS was very large. According to (26) and (37), it can be found that the spatial spectrums of the proposed beamspace MUSIC algorithm for the systems with UPA and ULA were respectively four-dimensional and three-dimensional, which meant that the proposed multi-spectral peak search method had better robustness to cope with the increase of the dimension of spatial spectrum.

VI. Conclusion

In this paper, the channel parameter estimator for mmWave hybrid MIMO-OFDM systems was proposed. Specifically, the estimates of AOAs, AODs and transmission delays were simultaneously acquired by the proposed beamspace MUSIC algorithm, and the LS method was employed to estimate the path gains. To avoid the high complexity of traditional spectral peak search method, a multi-spectral peak search method was skillfully designed based on the CSO algorithm. Then, the proposed channel estimator was extended to more actual systems with UPAs, which had better performance. Finally, the CRB results of these four channel parameters were derived to indicate the optimal asymptotically achievable performance of the proposed channel parameter estimator. The simulation results showed that the proposed channel estimator had better estimation accuracy and robustness in both small-scale and large-scale MIMO-OFDM systems.

Appendix

Consider the reconstructed signal $\mathbf{y}(n) \in \mathbb{C}^{MU_rQ_r}, n = 1, \cdots, N$ in (17) and define $\boldsymbol{\kappa} = \left[\boldsymbol{\theta}^T, \boldsymbol{\phi}^T, \boldsymbol{\tau}^T, \boldsymbol{\alpha}(n)^T\right]$. Then, the log-likelihood function of $\boldsymbol{\kappa}$ at the n-th block is

$$L_n(\boldsymbol{\kappa}) = -MU_rQ_r \ln\left(\pi\sigma^2\right) - \frac{1}{\sigma^2}\left\|\mathbf{y}(n) - \mathbf{B}\boldsymbol{\alpha}(n)\right\|_F^2, \quad (41)$$

and the complex Fisher information matrix (FIM) for $\boldsymbol{\kappa}$ is given by

$$\boldsymbol{\Omega}_n(\boldsymbol{\kappa}) = \mathbb{E}\left[\left(\frac{\partial L_n(\boldsymbol{\kappa})}{\partial \boldsymbol{\kappa}}\right)^H \left(\frac{\partial L_n(\boldsymbol{\kappa})}{\partial \boldsymbol{\kappa}}\right)\right]. \quad (42)$$

To calculate $\boldsymbol{\Omega}_n(\boldsymbol{\kappa})$, the partial derivative of $L_n(\boldsymbol{\kappa})$ with respect to $\boldsymbol{\kappa}$ should be first computed. The partial derivative of $L_n(\boldsymbol{\kappa})$ with respect to θ_l can be calculated as

$$\frac{\partial L_n(\boldsymbol{\kappa})}{\partial \theta_l} = \frac{1}{\sigma^2}\left\{\boldsymbol{\alpha}(n)^H \frac{\partial \mathbf{B}^H}{\partial \theta_l}\mathbf{n}(n) + \mathbf{n}(n)^H \frac{\partial \mathbf{B}}{\partial \theta_l}\boldsymbol{\alpha}(n)\right\}, \quad (43)$$

where

$$\frac{\partial \mathbf{B}}{\partial \theta_l} = \mathbf{E} \odot \left(\mathbf{F}^H \mathbf{A}_U\right) \odot \left(\mathbf{W}^H \frac{\partial \mathbf{A}_B}{\partial \theta_l}\right) \triangleq \mathbf{B}_{\theta_l}, \quad (44)$$

whose only l-th column is a non-zero vector defined as $\mathbf{b}_{\theta_l} \in \mathbb{C}^{MU_rQ_r}$, because

$$\frac{\partial \mathbf{A}_B}{\partial \theta_l} = \left[\mathbf{0}_{Q_a\times 1}, \cdots, \frac{\partial \mathbf{a}_B(\theta_l)}{\partial \theta_l}, \cdots, \mathbf{0}_{Q_a\times 1}\right],$$
$$\frac{\partial \mathbf{a}_B(\theta_l)}{\partial \theta_l} = \mathbf{D}_\theta \mathbf{a}_B(\theta_l),$$
$$\mathbf{D}_\theta = j2\pi \mathrm{diag}(0, 1, \cdots, Q_a - 1).$$

It can be found that the l-th row of $\frac{\partial \mathbf{A}_B^H}{\partial \theta_l}$ is equal to $\frac{\partial \mathbf{a}_B^H(\theta_l)}{\partial \theta_l}$ given by

$$\frac{\partial \mathbf{a}_B^H(\theta_l)}{\partial \theta_l} = \mathbf{a}_B^H(\theta_l)\mathbf{D}_\theta^H = \left(\frac{\partial \mathbf{a}_B(\theta_l)}{\partial \theta_l}\right)^H.$$

Therefore, $\frac{\partial \mathbf{A}_B^H}{\partial \theta_l} = \left(\frac{\partial \mathbf{A}_B}{\partial \theta_l}\right)^H$ and $\frac{\partial \mathbf{B}^H}{\partial \theta_l} = \left(\frac{\partial \mathbf{B}}{\partial \theta_l}\right)^H$. Then, (43) can be rewritten as

$$\begin{aligned}\frac{\partial L_n(\boldsymbol{\kappa})}{\partial \theta_l} &= 2\,\mathrm{Re}\left\{\frac{1}{\sigma^2}\alpha_l(n)\,\mathbf{n}(n)^H \mathbf{b}_{\theta_l}\right\} \\ &= 2\,\mathrm{Re}\left\{\frac{1}{\sigma^2}\mathbf{e}_l^T\boldsymbol{\alpha}(n)\,\mathbf{n}(n)^H \mathbf{B}_{\theta d}\mathbf{e}_l\right\} \\ &= 2\,\mathrm{Re}\left\{\mathbf{e}_l^T\mathbf{N}^\theta\mathbf{e}_l\right\},\end{aligned} \quad (45)$$

where $\mathrm{Re}\{\cdot\}$ is the real part of a complex number, $\mathbf{e}_l \in \mathbb{R}^{L\times 1}$ is a canonical vector whose only l-th element is non-zero, $\mathbf{B}_{\theta d} = [\mathbf{b}_{\theta_1}, \cdots, \mathbf{b}_{\theta_L}] \in \mathbb{C}^{MU_rQ_r\times L}$ and

$$\mathbf{N}^\theta = \frac{1}{\sigma^2}\boldsymbol{\alpha}(n)\,\mathbf{n}(n)^H \mathbf{B}_{\theta d}. \quad (46)$$

Similarly, the partial derivative of $L_n(\boldsymbol{\kappa})$ with respect to ϕ_l, τ_l and $\alpha_l(n)$ can be respectively calculated as

$$\begin{aligned}\frac{\partial L_n(\boldsymbol{\kappa})}{\partial \phi_l} &= 2\,\mathrm{Re}\left\{\mathbf{e}_l^T\mathbf{N}^\phi\mathbf{e}_l\right\}, \\ \frac{\partial L_n(\boldsymbol{\kappa})}{\partial \tau_l} &= 2\,\mathrm{Re}\left\{\mathbf{e}_l^T\mathbf{N}^\tau\mathbf{e}_l\right\}, \\ \frac{\partial L_n(\boldsymbol{\kappa})}{\partial \alpha_l(n)} &= \mathbf{N}^\alpha\mathbf{e}_l,\end{aligned} \quad (47)$$

where

$$\begin{aligned}\mathbf{N}^\phi &= \frac{1}{\sigma^2}\boldsymbol{\alpha}(n)\,\mathbf{n}(n)^H \mathbf{B}_{\phi d}, \\ \mathbf{N}^\tau &= \frac{1}{\sigma^2}\boldsymbol{\alpha}(n)\,\mathbf{n}(n)^H \mathbf{B}_{\tau d}, \\ \mathbf{N}^\alpha &= \frac{1}{\sigma^2}\mathbf{n}(n)^H \mathbf{B},\end{aligned} \quad (48)$$

where $\mathbf{B}_{\phi d} = [\mathbf{b}_{\phi_1}, \cdots, \mathbf{b}_{\phi_L}]$ and $\mathbf{B}_{\tau d} = [\mathbf{b}_{\tau_1}, \cdots, \mathbf{b}_{\tau_L}]$, $\mathbf{b}_{\phi_l}$ and $\mathbf{b}_{\tau_l}$ are respectively the only non-zero column vectors of $\mathbf{B}_{\phi_l}$ and $\mathbf{B}_{\tau_l}$ respectively calculated by

$$\begin{aligned}\mathbf{B}_{\phi_l} &= \mathbf{E} \odot \left(\mathbf{F}^H \frac{\partial \mathbf{A}_U}{\partial \phi_l}\right) \odot \left(\mathbf{W}^H \mathbf{A}_B\right), \\ \mathbf{B}_{\tau_l} &= \frac{\partial \mathbf{E}}{\partial \tau_l} \odot \left(\mathbf{F}^H \mathbf{A}_U\right) \odot \left(\mathbf{W}^H \mathbf{A}_B\right),\end{aligned} \quad (49)$$

where

$$\frac{\partial \mathbf{A}_U}{\partial \phi_l} = \left[\mathbf{0}_{U_a\times 1}, \cdots, \frac{\partial \mathbf{a}_U(\phi_l)}{\partial \phi_l}, \cdots, \mathbf{0}_{U_a\times 1}\right],$$
$$\frac{\partial \mathbf{a}_U(\phi_l)}{\partial \phi_l} = \mathbf{D}_\phi \mathbf{a}_U(\phi_l),$$
$$\mathbf{D}_\phi = j2\pi \mathrm{diag}(0, 1, \cdots, U_a - 1),$$

and

$$\frac{\partial \mathbf{E}}{\partial \tau_l} = \left[\mathbf{0}_{M\times 1}, \cdots, \frac{\partial \mathbf{e}(\tau_l)}{\partial \tau_l}, \cdots, \mathbf{0}_{M\times 1}\right],$$
$$\frac{\partial \mathbf{e}(\tau_l)}{\partial \tau_l} = \mathbf{D}_\tau \mathbf{e}(\tau_l),$$
$$\mathbf{D}_\tau = -j2\pi\Delta_f \mathrm{diag}(1, 2, \cdots, M).$$

Regarding the diagonal elements of $\mathbf{\Omega}_n(\boldsymbol{\kappa})$, the (l_1,l_2)-th $(l_1,l_2=1,\cdots,L)$ element of $\mathbb{E}\left[\left(\frac{\partial L_n(\boldsymbol{\kappa})}{\partial \boldsymbol{\theta}}\right)^H\left(\frac{\partial L_n(\boldsymbol{\kappa})}{\partial \boldsymbol{\theta}}\right)\right]$ is given by

$$\begin{aligned}&\mathbb{E}\left[\left(\frac{\partial L_n(\boldsymbol{\kappa})}{\partial \theta_{l_1}}\right)^*\left(\frac{\partial L_n(\boldsymbol{\kappa})}{\partial \theta_{l_2}}\right)\right]\\&=4\mathbb{E}\left[\mathrm{Re}\left\{\mathbf{e}_{l_1}^T\mathbf{N}^\theta\mathbf{e}_{l_1}\right\}^*\mathrm{Re}\left\{\mathbf{e}_{l_2}^T\mathbf{N}^\theta\mathbf{e}_{l_2}\right\}\right]\\&=4\mathbb{E}\left[\mathrm{Re}\left\{\mathbf{N}^\theta(l_1,l_1)\right\}\mathrm{Re}\left\{\mathbf{N}^\theta(l_2,l_2)\right\}\right]\\&=\mathbb{E}\left[\left(\mathbf{N}^\theta(l_1,l_1)^*+\mathbf{N}^\theta(l_1,l_1)\right)\right.\\&\qquad\left.\left(\mathbf{N}^\theta(l_2,l_2)^*+\mathbf{N}^\theta(l_2,l_2)\right)\right],\end{aligned}\tag{50}$$

where $\mathbf{N}^\theta(l_1,l_1)$ represents the (l_1,l_1)-th element of $\mathbf{N}^\theta$. If $\mathbf{n}^\theta=\mathrm{vec}\left(\mathbf{N}^\theta\right)$, there is

$$\mathbf{n}^\theta=\frac{1}{\sigma^2}\mathbf{B}_{\theta d}^T\otimes\boldsymbol{\alpha}(n)\,\mathbf{n}(n)^*.\tag{51}$$

It can be seen from (51) that $\mathbf{n}^\theta$ is the linear transformation of $\mathbf{n}(n)^*$, and $\mathbf{n}(n)\sim\mathcal{CN}\left(0,\sigma^2\mathbf{I}_{MU_rQ_r}\right)$. Therefore, $\mathbf{n}^\theta$ also follows the i.i.d. complex Gaussian distribution, whose covariance matrix $\mathbf{C}_\theta\in\mathbb{C}^{L^2\times L^2}$ is

$$\begin{aligned}\mathbf{C}_\theta&=\mathbb{E}\left[\mathbf{n}^\theta\left(\mathbf{n}^\theta\right)^H\right]\\&=\frac{1}{\sigma^4}\left(\mathbf{B}_{\theta d}^T\otimes\boldsymbol{\alpha}(n)\right)\mathbb{E}\left[\mathbf{n}(n)^*\left(\mathbf{n}(n)^*\right)^H\right]\left(\mathbf{B}_{\theta d}^T\otimes\boldsymbol{\alpha}(n)\right)^H\\&=\frac{1}{\sigma^4}\left(\mathbf{B}_{\theta d}^T\otimes\boldsymbol{\alpha}(n)\right)\sigma^2\mathbf{I}_{MU_rQ_r}\left(\mathbf{B}_{\theta d}^*\otimes\boldsymbol{\alpha}(n)^H\right)\\&=\frac{1}{\sigma^2}\left(\mathbf{B}_{\theta d}^T\mathbf{B}_{\theta d}^*\right)\otimes\left(\boldsymbol{\alpha}(n)\,\boldsymbol{\alpha}(n)^H\right),\end{aligned}\tag{52}$$

and second-order moment is $\mathbf{M}_\theta=\mathbb{E}\left[\mathbf{n}^\theta\left(\mathbf{n}^\theta\right)^T\right]=\mathbf{0}_{L^2\times L^2}$. Thus, (50) can be rewritten as

$$\begin{aligned}&\mathbb{E}\left[\left(\frac{\partial L_n(\boldsymbol{\kappa})}{\partial \theta_{l_1}}\right)^*\left(\frac{\partial L_n(\boldsymbol{\kappa})}{\partial \theta_{l_2}}\right)\right]\\&=\mathbb{E}\left[\left(\mathbf{n}^\theta(p)^*+\mathbf{n}^\theta(p)\right)\left(\mathbf{n}^\theta(q)^*+\mathbf{n}^\theta(q)\right)\right]\\&=\mathbf{C}_\theta(p,q)^*+\mathbf{C}_\theta(p,q)\\&=2\,\mathrm{Re}\left\{\mathbf{C}_\theta(p,q)\right\},\end{aligned}\tag{53}$$

where $\mathbf{n}^\theta(p)$ represents the p-th element of $\mathbf{n}^\theta$, $\mathbf{C}_\theta(p,q)$ represents the (p,q)-th element of $\mathbf{C}_\theta$, $p=(l_1-1)\times L+l_1$ and $q=(l_2-1)\times L+l_2$. Similarly,

$$\begin{aligned}&\mathbb{E}\left[\left(\frac{\partial L_n(\boldsymbol{\kappa})}{\partial \phi_{l_1}}\right)^*\left(\frac{\partial L_n(\boldsymbol{\kappa})}{\partial \phi_{l_2}}\right)\right]=2\,\mathrm{Re}\left\{\mathbf{C}_\phi(p,q)\right\},\\&\mathbb{E}\left[\left(\frac{\partial L_n(\boldsymbol{\kappa})}{\partial \tau_{l_1}}\right)^*\left(\frac{\partial L_n(\boldsymbol{\kappa})}{\partial \tau_{l_2}}\right)\right]=2\,\mathrm{Re}\left\{\mathbf{C}_\tau(p,q)\right\},\\&\mathbb{E}\left[\left(\frac{\partial L_n(\boldsymbol{\kappa})}{\partial \alpha_{l_1}(n)}\right)^*\left(\frac{\partial L_n(\boldsymbol{\kappa})}{\partial \alpha_{l_2}(n)}\right)\right]=\mathbf{C}_\alpha(l_1,l_2)^*,\end{aligned}\tag{54}$$

where

$$\begin{aligned}\mathbf{C}_\phi&=\frac{1}{\sigma^2}\left(\mathbf{B}_{\phi d}^T\mathbf{B}_{\phi d}^*\right)\otimes\left(\boldsymbol{\alpha}(n)\,\boldsymbol{\alpha}(n)^H\right),\\\mathbf{C}_\tau&=\frac{1}{\sigma^2}\left(\mathbf{B}_{\tau d}^T\mathbf{B}_{\tau d}^*\right)\otimes\left(\boldsymbol{\alpha}(n)\,\boldsymbol{\alpha}(n)^H\right),\\\mathbf{C}_\alpha&=\frac{1}{\sigma^2}\mathbf{B}^T\mathbf{B}^*.\end{aligned}\tag{55}$$

For the non-diagonal elements of $\mathbf{\Omega}_n(\boldsymbol{\kappa})$, the (l_1,l_2)-th element of $\mathbb{E}\left[\left(\frac{\partial L_n(\boldsymbol{\kappa})}{\partial \boldsymbol{\theta}}\right)^H\left(\frac{\partial L_n(\boldsymbol{\kappa})}{\partial \boldsymbol{\phi}}\right)\right]$ is given by

$$\begin{aligned}&\mathbb{E}\left[\left(\frac{\partial L_n(\boldsymbol{\kappa})}{\partial \theta_{l_1}}\right)^*\left(\frac{\partial L_n(\boldsymbol{\kappa})}{\partial \phi_{l_2}}\right)\right]\\&=4\mathbb{E}\left[\mathrm{Re}\left\{\mathbf{e}_{l_1}^T\mathbf{N}^\theta\mathbf{e}_{l_1}\right\}^*\mathrm{Re}\left\{\mathbf{e}_{l_2}^T\mathbf{N}^\phi\mathbf{e}_{l_2}\right\}\right]\\&=4\mathbb{E}\left[\mathrm{Re}\left\{\mathbf{N}^\theta(l_1,l_1)\right\}\mathrm{Re}\left\{\mathbf{N}^\phi(l_2,l_2)\right\}\right]\\&=\mathbb{E}\left[\left(\mathbf{N}^\theta(l_1,l_1)^*+\mathbf{N}^\theta(l_1,l_1)\right)\right.\\&\qquad\left.\left(\mathbf{N}^\phi(l_2,l_2)^*+\mathbf{N}^\phi(l_2,l_2)\right)\right]\\&=2\,\mathrm{Re}\left\{\mathbf{C}_{\theta\phi}(p,q)\right\},\end{aligned}\tag{56}$$

where

$$\mathbf{C}_{\theta\phi}=\frac{1}{\sigma^2}\left(\mathbf{B}_{\theta d}^T\mathbf{B}_{\phi d}^*\right)\otimes\left(\boldsymbol{\alpha}(n)\,\boldsymbol{\alpha}(n)^H\right).\tag{57}$$

Similarly,

$$\begin{aligned}&\mathbb{E}\left[\left(\frac{\partial L_n(\boldsymbol{\kappa})}{\partial \theta_{l_1}}\right)^*\left(\frac{\partial L_n(\boldsymbol{\kappa})}{\partial \tau_{l_2}}\right)\right]=2\,\mathrm{Re}\left\{\mathbf{C}_{\theta\tau}(p,q)\right\},\\&\mathbb{E}\left[\left(\frac{\partial L_n(\boldsymbol{\kappa})}{\partial \phi_{l_1}}\right)^*\left(\frac{\partial L_n(\boldsymbol{\kappa})}{\partial \tau_{l_2}}\right)\right]=2\,\mathrm{Re}\left\{\mathbf{C}_{\phi\tau}(p,q)\right\},\end{aligned}\tag{58}$$

where

$$\begin{aligned}\mathbf{C}_{\theta\tau}&=\frac{1}{\sigma^2}\left(\mathbf{B}_{\theta d}^T\mathbf{B}_{\tau d}^*\right)\otimes\left(\boldsymbol{\alpha}(n)\,\boldsymbol{\alpha}(n)^H\right),\\\mathbf{C}_{\phi\tau}&=\frac{1}{\sigma^2}\left(\mathbf{B}_{\phi d}^T\mathbf{B}_{\tau d}^*\right)\otimes\left(\boldsymbol{\alpha}(n)\,\boldsymbol{\alpha}(n)^H\right).\end{aligned}\tag{59}$$

Besides,

$$\begin{aligned}&\mathbb{E}\left[\left(\frac{\partial L_n(\boldsymbol{\kappa})}{\partial \theta_{l_1}}\right)^*\left(\frac{\partial L_n(\boldsymbol{\kappa})}{\partial \alpha_{l_2}(n)}\right)\right]=\mathbf{C}_{\theta\alpha}(p,l_2)^*,\\&\mathbb{E}\left[\left(\frac{\partial L_n(\boldsymbol{\kappa})}{\partial \phi_{l_1}}\right)^*\left(\frac{\partial L_n(\boldsymbol{\kappa})}{\partial \alpha_{l_2}(n)}\right)\right]=\mathbf{C}_{\phi\alpha}(p,l_2)^*,\\&\mathbb{E}\left[\left(\frac{\partial L_n(\boldsymbol{\kappa})}{\partial \tau_{l_1}}\right)^*\left(\frac{\partial L_n(\boldsymbol{\kappa})}{\partial \alpha_{l_2}(n)}\right)\right]=\mathbf{C}_{\tau\alpha}(p,l_2)^*,\end{aligned}\tag{60}$$

where

$$\begin{aligned}\mathbf{C}_{\theta\alpha}&=\frac{1}{\sigma^2}\left(\mathbf{B}_{\theta d}^T\otimes\boldsymbol{\alpha}(n)\right)\mathbf{B}^*,\\\mathbf{C}_{\phi\alpha}&=\frac{1}{\sigma^2}\left(\mathbf{B}_{\phi d}^T\otimes\boldsymbol{\alpha}(n)\right)\mathbf{B}^*,\\\mathbf{C}_{\tau\alpha}&=\frac{1}{\sigma^2}\left(\mathbf{B}_{\tau d}^T\otimes\boldsymbol{\alpha}(n)\right)\mathbf{B}^*.\end{aligned}\tag{61}$$

Since $\mathbf{C}_{\theta\phi}$ reflects the relationship between two different parameters, the corresponding elements of $\mathbf{C}_{\theta\phi}$ are non-zero only when the two parameters correspond to the same path. Therefore, $\mathbf{C}_{\theta\phi}$ should be a block diagonal matrix. Similarly, $\mathbf{C}_{\theta\tau}$, $\mathbf{C}_{\theta\alpha}$, $\mathbf{C}_{\phi\tau}$, $\mathbf{C}_{\phi\alpha}$ and $\mathbf{C}_{\tau\alpha}$ should also be block diagonal matrices.

REFERENCES

[1] X. You, C. -X. Wang, J. Huang *et al.*, "Towards 6G wireless communication networks: vision, enabling technologies, and new paradigm shifts," *Sci China Inf Sci*, vol. 64, no. 1, pp. 1-74, 2021, doi: 10.1007/s11432-020-2955-6.

[2] P. Zhu, H. Lin, J. Bao, J. Li and D. Wang, "Beam Tracking for Distributed Millimeter-Wave Massive MIMO Systems Based on the Unscented Kalman Filter," in *IEEE Wireless Communications Letters*, vol. 11, no. 4, pp. 712-716, April 2022, doi: 10.1109/LWC.2022.3140852.

[3] P. Z. Zhang, C. Yi, B. S. Yang *et al.*, "In-building coverage of millimeter-wave wireless networks from channel measurement and modeling perspectives," *Sci China Inf Sci*, vol. 63, no. 8, 2020.

[4] Q. Xue, X. Fang, M. Xiao and L. Yan, "Multiuser Millimeter Wave Communications With Nonorthogonal Beams," in *IEEE Transactions on Vehicular Technology*, vol. 66, no. 7, pp. 5675-5688, July 2017.

[5] J. Xu, X. Wang, P. Zhu and X. You, "Privacy-Preserving Channel Estimation in Cell-Free Hybrid Massive MIMO Systems," in *IEEE Transactions on Wireless Communications*, vol. 20, no. 6, pp. 3815-3830, June 2021, doi: 10.1109/TWC.2021.3053770.

[6] Y. Hu, J. Zhan, Z. H. Jiang, C. Yu and W. Hong, "An Orthogonal Hybrid Analog–Digital Multibeam Antenna Array for Millimeter-Wave Massive MIMO Systems," in *IEEE Transactions on Antennas and Propagation*, vol. 69, no. 3, pp. 1393-1403, March 2021.

[7] L. Zhao, M. Li, C. Liu, S. V. Hanly, I. B. Collings and P. A. Whiting, "Energy Efficient Hybrid Beamforming for Multi-User Millimeter Wave Communication With Low-Resolution A/D at Transceivers," in *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 9, pp. 2142-2155, Sept. 2020.

[8] R. W. Heath, N. González-Prelcic, S. Rangan, W. Roh and A. M. Sayeed, "An Overview of Signal Processing Techniques for Millimeter Wave MIMO Systems," in *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 3, pp. 436-453, April 2016.

[9] H. Kim, G. -T. Gil and Y. H. Lee, "Two-Step Approach to Time-Domain Channel Estimation for Wideband Millimeter Wave Systems With Hybrid Architecture," in *IEEE Transactions on Communications*, vol. 67, no. 7, pp. 5139-5152, July 2019.

[10] Z. Guo, X. Wang and W. Heng, "Millimeter-Wave Channel Estimation Based on 2-D Beamspace MUSIC Method," in *IEEE Transactions on Wireless Communications*, vol. 16, no. 8, pp. 5384-5394, Aug. 2017.

[11] J. Lee, G. -T. Gil and Y. H. Lee, "Channel Estimation via Orthogonal Matching Pursuit for Hybrid MIMO Systems in Millimeter Wave Communications," in *IEEE Transactions on Communications*, vol. 64, no. 6, pp. 2370-2386, June 2016.

[12] C. Hu, L. Dai, T. Mir, Z. Gao and J. Fang, "Super-Resolution Channel Estimation for MmWave Massive MIMO With Hybrid Precoding," in *IEEE Transactions on Vehicular Technology*, vol. 67, no. 9, pp. 8954-8958, Sept. 2018.

[13] Baoguo Yang, K. B. Letaief, R. S. Cheng and Zhigang Cao, "Channel estimation for OFDM transmission in multipath fading channels based on parametric channel modeling," in *IEEE Transactions on Communications*, vol. 49, no. 3, pp. 467-479, March 2001.

[14] C. Studer and E. G. Larsson, "PAR-Aware Large-Scale Multi-User MIMO-OFDM Downlink," in *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 303-313, February 2013.

[15] L. You, X. Gao, A. L. Swindlehurst and W. Zhong, "Channel Acquisition for Massive MIMO-OFDM With Adjustable Phase Shift Pilots," in *IEEE Transactions on Signal Processing*, vol. 64, no. 6, pp. 1461-1476, March15, 2016.

[16] X. Lin, S. Wu, C. Jiang, L. Kuang, J. Yan and L. Hanzo, "Estimation of Broadband Multiuser Millimeter Wave Massive MIMO-OFDM Channels by Exploiting Their Sparse Structure," in *IEEE Transactions on Wireless Communications*, vol. 17, no. 6, pp. 3959-3973, June 2018.

[17] W. Tong, W. Xu, F. Wang, J. Shang, M. Pan and J. Lin, "Deep Learning Compressed Sensing-Based Beamspace Channel Estimation in mmWave Massive MIMO Systems," in *IEEE Wireless Communications Letters*, vol. 11, no. 9, pp. 1935-1939, Sept. 2022.

[18] B. Wang, M. Jian, F. Gao, G. Y. Li and H. Lin, "Beam Squint and Channel Estimation for Wideband mmWave Massive MIMO-OFDM Systems," in *IEEE Transactions on Signal Processing*, vol. 67, no. 23, pp. 5893-5908, Dec. 2019.

[19] J. Zhao, J. Liu, F. Gao, W. Jia and W. Zhang, "Gridless Compressed Sensing Based Channel Estimation for UAV Wideband Communications With Beam Squint," in *IEEE Transactions on Vehicular Technology*, vol. 70, no. 10, pp. 10265-10277, Oct. 2021.

[20] Z. Zhou, J. Fang, L. Yang, H. Li, Z. Chen and S. Li, "Channel Estimation for Millimeter-Wave Multiuser MIMO Systems via PARAFAC Decomposition," in *IEEE Transactions on Wireless Communications*, vol. 15, no. 11, pp. 7501-7516, Nov. 2016.

[21] Z. Zhou, J. Fang, L. Yang, H. Li, Z. Chen and R. S. Blum, "Low-Rank Tensor Decomposition-Aided Channel Estimation for Millimeter Wave MIMO-OFDM Systems," in *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 7, pp. 1524-1538, July 2017.

[22] Y. Lin, S. Jin, M. Matthaiou and X. You, "Tensor-Based Channel Estimation for Millimeter Wave MIMO-OFDM With Dual-Wideband Effects," in *IEEE Transactions on Communications*, vol. 68, no. 7, pp. 4218-4232, July 2020.

[23] P. Stoica and R. L. Moses, *Spectral Analysis of Signals*. Upper Saddle River, NJ, USA: Prentice-Hall, 2005.

[24] R. Schmidt, "Multiple emitter location and signal parameter estimation," in *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276-280, March 1986.

[25] M. Wang, F. Gao, S. Jin and H. Lin, "An Overview of Enhanced Massive MIMO With Array Signal Processing Techniques," in *IEEE Journal of Selected Topics in Signal Processing*, vol. 13, no. 5, pp. 886-901, Sept. 2019.

[26] F. Dong, W. Wang, Z. Huang and P. Huang, "High-Resolution Angle-of-Arrival and Channel Estimation for mmWave Massive MIMO Systems With Lens Antenna Array," in *IEEE Transactions on Vehicular Technology*, vol. 69, no. 11, pp. 12963-12973, Nov. 2020.

[27] K. F. Masood, R. Hu, J. Tong, J. Xi, Q. Guo and Y. Yu, "A Low-Complexity Three-Stage Estimator for Low-Rank mmWave Channels," in *IEEE Transactions on Vehicular Technology*, vol. 70, no. 6, pp. 5920-5931, June 2021.

[28] Y. Cheng, J. Du, M. He, Z. Mo, P. Zhang and X. Wang, "A Fast Tensor-Based Channel Estimation Method in mmWave Massive MIMO-OFDM Systems," *2021 IEEE 21st International Conference on Communication Technology (ICCT)*, 2021, pp. 243-247.

[29] J. Song, J. Choi and D. J. Love, "Common Codebook Millimeter Wave Beam Design: Designing Beams for Both Sounding and Communication With Uniform Planar Arrays," in *IEEE Transactions on Communications*, vol. 65, no. 4, pp. 1859-1872, April 2017.

[30] X. Wu, X. Yang, S. Ma, B. Zhou and G. Yang, "Hybrid Channel Estimation for UPA-Assisted Millimeter-Wave Massive MIMO IoT Systems," in *IEEE Internet of Things Journal*, vol. 9, no. 4, pp. 2829-2842, Feb. 2022.

[31] R. Zhang, L. Cheng, S. Wang, Y. Lou, W. Wu and D. W. K. Ng, "Tensor Decomposition-Based Channel Estimation for Hybrid mmWave Massive MIMO in High-Mobility Scenarios," in *IEEE Transactions on Communications*, vol. 70, no. 9, pp. 6325-6340, Sept. 2022.

[32] J. Du, M. Han, Y. Chen, L. Jin and F. Gao, "Tensor-Based Joint Channel Estimation and Symbol Detection for Time-Varying mmWave Massive MIMO Systems," in *IEEE Transactions on Signal Processing*, vol. 69, pp. 6251-6266, 2021.

[33] S. Rangan, T. S. Rappaport and E. Erkip, "Millimeter-Wave Cellular Wireless Networks: Potentials and Challenges," in *Proceedings of the IEEE*, vol. 102, no. 3, pp. 366-385, March 2014.

[34] L. You, X. Gao, G. Y. Li, X. -G. Xia and N. Ma, "BDMA for Millimeter-Wave/Terahertz Massive MIMO Transmission With Per-Beam Synchronization," in *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 7, pp. 1550-1563, July 2017.

[35] A. Abdallah and M. M. Mansour, "Efficient Angle-Domain Processing for FDD-Based Cell-Free Massive MIMO Systems," in *IEEE Transactions on Communications*, vol. 68, no. 4, pp. 2188-2203, April 2020.

[36] A. Wang, R. Yin and C. Zhong, "Channel Estimation for Uniform Rectangular Array Based Massive MIMO Systems With Low Complexity," in *IEEE Transactions on Vehicular Technology*, vol. 68, no. 3, pp. 2545-2556, March 2019.

[37] B. Nadler, "Nonparametric Detection of Signals by Information Theoretic Criteria: Performance Analysis and an Improved Estimator," in *IEEE Transactions on Signal Processing*, vol. 58, no. 5, pp. 2746-2756, May 2010.

[38] K. Wang, Z. Li, H. Cheng and K. Zhang, "Mutation chicken swarm optimization based on nonlinear inertia weight," *2017 3rd IEEE International Conference on Computer and Communications (ICCC)*, 2017, pp. 2206-2211.

[39] X. Liang, D. Kou and L. Wen, "An Improved Chicken Swarm Optimization Algorithm and its Application in Robot Path Planning," in *IEEE Access*, vol. 8, pp. 49543-49550, 2020.

[40] M. Trinh-Hoang, M. Viberg and M. Pesavento, "Cramér-Rao Bound for DOA Estimators Under the Partial Relaxation Framework: Derivation and Comparison," in *IEEE Transactions on Signal Processing*, vol. 68, pp. 3194-3208, 2020.

[41] L. Khamidullina, I. Podkurkov and M. Haardt, "Conditional and Unconditional Cramér-Rao Bounds for Near-Field Localization in Bistatic MIMO Radar Systems," in *IEEE Transactions on Signal Processing*, vol. 69, pp. 3220-3234, 2021.

[42] E. H. Houssein, M. M. Ahmed, M. A. Elaziz, A. A. Ewees and R. M. Ghoniem, "Solving Multi-Objective Problems Using Bird Swarm Algorithm," in *IEEE Access*, vol. 9, pp. 36382-36398, 2021.

[43] K. Alkebsi and W. Du, "A Fast Multi-Objective Particle Swarm Optimization Algorithm Based on a New Archive Updating Mechanism," in *IEEE Access*, vol. 8, pp. 124734-124754, 2020.

Pengcheng Zhu received the B.S. and M.S. degrees in electrical engineering from Shandong University, Jinan, China, in 2001 and 2004, respectively, and the Ph.D. degree in communication and information science from Southeast University, Nanjing, China, in 2009. He is now a Professor with the National Mobile Communications Research Laboratory, Southeast University. His research interests lie in the areas of wireless communications and mobile networks, including 5G/6G mobile communication systems, massive MIMO, ultra-reliable and low latency communications (URLLC), and mmWave communications.

Huixin Lin was born in Jilin, China, in 1998. She received the B.S. degree in electronic information engineering from Northeastern University, China, in 2020. She is currently pursuing the M.S. degree in communication engineering with the National Mobile Communications Research Laboratory, Southeast University, Nanjing, China. Her research interests include distributed massive MIMO, channel estimation, as well as mmWave communications.

Jiamin Li received the B.S. and M.S. degrees in communication and information systems from Hehai University, Nanjing, China, in 2006 and 2009, respectively, and the Ph.D. degree in information and communication engineering from Southeast University, Nanjing, in 2014. He joined the National Mobile Communications Research Laboratory, Southeast University, in 2014, where he has been an Associate Professor since 2019. His research interests include massive MIMO, distributed antenna systems, and multi-objective optimization.

Dongming Wang received the B.S. degree from Chongqing University of Posts and Telecommunications in 1999, the M.S. degree from Nanjing University of Posts and Telecommunications in 2002, and the Ph.D. degree from Southeast University, China, in 2006. He joined the National Mobile Communications Research Laboratory, Southeast University, in 2006, where he is currently a Professor. His research interests include signal processing for wireless communications, large-scale distributed MIMO systems (cell-free massive MIMO), and open radio access network (ORAN).

Xiaohu You received his Master and Ph.D. Degrees from Southeast University, Nanjing, China, in Electrical Engineering in 1985 and 1988, respectively. Since 1990, he has been working with National Mobile Communications Research Laboratory at Southeast University, where he holds the rank of director and professor. His research interests include mobile communication systems, signal processing and its applications. He has contributed over 100 IEEE journal papers and 2 books in the areas of adaptive signal processing, neural networks and their applications to communication systems. From 1999 to 2002, he was the Principal Expert of the C3G Project, responsible for organizing China's 3G Mobile Communications R&D Activities. From 2001-2006, he was the Principal Expert of the China National 863 Beyond 3G Future Project. Since 2013, he has been the Principal Investigator of China National 863 5G Project. Professor You served as the general chairs of IEEE WCNC 2013 and IEEE VTC 2016. Now he is Secretary General of the Future Forum, vice Chair of China IMT-2020 Promotion Group, vice Chair of China National Mega Project on New Generation Mobile Network. He was the recipient of the National 1st Class Invention Prize in 2011, and he was selected as IEEE Fellow in the same year due to his contributions to development of mobile communications in China.