






Time-Varying Channel Estimation Scheme for Uplink MU-MIMO in 6G Systems

Jianhao Wang, *Student Member, IEEE*, Wensheng Zhang , *Member, IEEE*, Yunfei Chen , *Senior Member, IEEE*, Zhi Liu , *Member, IEEE*, Jian Sun , *Member, IEEE*, and Cheng-Xiang Wang , *Fellow, IEEE*

Abstract—Channel estimation is a challenging issue for millimeter wave (mmWave) and massive multiple-input multiple-output (MIMO) in the future sixth generation (6G) wireless systems, where the conventional estimation schemes may fail to track the fast varying channels, especially in high-speed mobile scenarios. In this paper, a novel tensor-based uplink channel estimation scheme is proposed for multi-user MIMO (MU-MIMO) systems over time-varying channels. In the proposed scheme, a low-overhead pilot transmission scheme is designed to track the varying channel. The received uplink signal at the base station (BS) is formulated as a third-order tensor which admits a CANDECOMP/PARAFAC (CP) model. The CP decomposition issue is then solved using blind matrix decomposition, in which the special structures of signals in the time dimension and the matrix subspace are utilized. By exploiting low-rank structure of the signal tensor, the channel parameters (angles of arrival/departure, path gains, and Doppler shifts) are estimated from the factor matrices. Moreover, the proposed scheme is theoretically analyzed and it is guaranteed with low pilot overhead. Simulation results verify that the proposed scheme can outperform the compressed sensing (CS) based scheme and the iteration-based scheme in terms of accuracy and stability. The uniqueness of the proposed scheme is also verified in our simulation.

Index Terms—CP decomposition, essentially unique, MU-MIMO, tensor, time-varying channel estimation.

I. INTRODUCTION

MASSIVE multiple-input multiple-output (MIMO) and millimeter-wave (mmWave) are key techniques in the

Manuscript received 11 January 2022; revised 8 June 2022; accepted 12 July 2022. Date of publication 21 July 2022; date of current version 14 November 2022. This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grants 62071276, 61960206006, and 61771293, in part by the Key Technologies R&D Program of Jiangsu (Prospective and Key Technologies for Industry) under Grants BE2022067 and BE2022067-1, in part by the King Abdullah University of Science and Technology Research Funding under Grant ORA-2021-CRG10-4696, in part by the Natural Science Foundation of Shandong Province under Grants ZR2020MF002, ZR2019ZD05, and ZR2020LZH013, and in part by EU H2020 RISE TESTBED2 Project under Grant 872172. The review of this article was coordinated by Prof. Mugen Peng. (Corresponding authors: Wensheng Zhang; Cheng-Xiang Wang.)

Jianhao Wang, Wensheng Zhang, Zhi Liu, and Jian Sun are with the School of Information Science and Engineering, Shandong Provincial Key Lab of Wireless Communication Technologies, Shandong University, Qingdao, Shandong 266237, China (e-mail: wangjianhao@mail.sdu.edu.cn; zhangwsh@sdu.edu.cn; liuzhi@sdu.edu.cn; sunjian@sdu.edu.cn).

Yunfei Chen is with the School of Engineering, University of Warwick, CV4 7AL Coventry, U.K. (e-mail: Yunfei.Chen@warwick.ac.uk).

Cheng-Xiang Wang is with the National Mobile Communications Research Laboratory, School of Information Science and Engineering, Southeast University, Nanjing, Jiangsu 210096, China, and also with the Purple Mountain Laboratories, Nanjing 211111, China (e-mail: chxwang@seu.edu.cn).

Digital Object Identifier 10.1109/TVT.2022.3192902

fifth-generation (5G) wireless communication systems and will still play an important role in future sixth-generation (6G) systems [1]–[6]. Communications at the mmWave band (30–300 GHz) can offer gigabit-per-second data rate by exploiting the large bandwidth, and the extremely short wavelength makes it possible to integrate a large number of antennas into small devices, which can provide beamforming gain to compensate for the significant attenuation and signal absorption [7]. In traditional MIMO systems, each antenna element needs to be supported by a radio frequency (RF) chain in order to support fully-digital beamforming. However, this is too expensive for massive MIMO due to the large number of antenna elements [8]. To address this issue, a hybrid beamforming architecture has been proposed, aiming to reduce the number of RF chains by combining analog and digital beamforming [9]. Despite using a smaller number of RF chains, the performance of hybrid beamforming architecture has been shown to be similar to the fully-digitally ones [10]. This is because the sparse nature of the mmWave makes the channel matrix low-rank [11]. However, to provide sufficient beamforming gain, the accurate channel state information (CSI) is required for precoding at the transmitter and combining at the receiver.

Due to the large number of antennas and high training overhead, channel estimation is a very challenging issue for massive MIMO. Many research efforts have been devoted to addressing this problem, and different approaches have been exploited to estimate the mmWave channel [12]–[26]. In [12], the deep convolutional neural network (CNN) was utilized to estimate the channel for mmWave orthogonal frequency division multiplexing (OFDM) systems. The deep CNN is good at extracting the characteristics to reduce the pilot overhead and estimate the channel more accurately. Other schemes employed the known sparse characteristics of mmWave channel to design estimation algorithms. For example [13]–[18], the estimation of mmWave channel was formulated as a sparse problem and solved using the compressive sensing (CS) based method. An adaptive algorithm was introduced for narrow-band mmWave channel in [13]. The authors divided the estimation into multilayer corresponding to multi-resolution codebook and the algorithm begins with a coarse grid and only the chosen sub-range was refined in the next stage. The process ends until achieving the desired resolution. In order to reduce the coherence of angle grids, a non-uniformly quantized scheme to construct redundant dictionary was proposed in [14]. Besides, a lot of algorithms have been proposed to deal with the k -sparse problem with high probability [27],

such as basis pursuit (BP) [28] and orthogonal matching pursuit (OMP) [29]. However, the CS-based algorithms need iteration or exhaustive search to achieve the desired resolution, which is too slow to track the rapid variation of mmWave channel. Furthermore, the CS methods are based on the assumption that the angles of arrival (AoAs) and angles of departure (AoDs) are quantized which lead to quantization errors. Although the quantization error can be reduced via increasing the number of quantization levels, the size of the measurement matrix may increase, requiring more pilots and higher complexity to recover the parameters.

Tensor theory is a powerful math tool to deal with high-dimensional data and has been widely used in signal processing [20]–[25]. The downlink channel estimation was investigated in [20] and a conjugate-invariant training sequence was designed to construct the tensor model. However, this study did not consider the hybrid beamforming architecture widely used in mmWave systems. In [21] and [22], the author converted multi-user channels and frequency selective channels into a third-order tensor and the channel parameters were estimated using iteration and one-dimensional search. In addition, low-rank tensor model can also be combined with traditional methods. A scheme combining CS with low-rank tensor to solve the multi-way sparse signal recovery problem was introduced in [25]. However, the CANDECOMP/PARAFAC (CP) decomposition in these schemes were solved by alternating least-squares (ALS) algorithm which is not guaranteed to converge to the global optimum, i.e., the success rate cannot be guaranteed.

The aforementioned works have considered the static channel. However, high-speed low-latency communications (HSLLC) is one of potential application scenarios of 6G with massive MIMO in the mmWave bands [2]. Due to higher Doppler shift, the coherence time could be much shorter for the mmWave systems than that for frequencies below 6 GHz [30]. The time-varying channel has been studied in [18], [24], where the estimation was decoupled into two separated stages including angle recovery followed by gain estimation, and an adaptive angle estimation algorithm was proposed for searching. However, it requires exclusive overhead for feedback downlink channel from the receiver to the base station (BS). To the best of our knowledge, there are few works investigating the estimation for uplink MU-MIMO systems over time-varying channels. Such channels may appear in some important applications such as the communication between unmanned aerial vehicle (UAV) group and ground station (GS) [31], the communication between high-speed trains (HSTs) and GS [32]. This motivates us to develop efficient channel estimation algorithms for time-varying uplink mmWave systems.

In this paper, we propose a new channel estimation framework for uplink mmWave systems over time-varying channels. Comparing with previous works, this work focuses more on the characteristics of time-varying channels and the relationship between the number of paths and the uniqueness of CP decomposition. Moreover, the special structure of the signal is further explored. Two assumptions are made: 1) the pilot transmission between mobile stations (MS) keeps synchronization; 2) the velocities of MSs are distinct and keep constant in a very short period of

time. The main contributions of this paper are summarized as follows:

- We design a novel pilot transmission scheme which can make better use of time domain information to track the rapid variations of mmWave channel. This scheme only inserts pilots into the header slots, and uses the characteristics of time-varying channels to predict the channel state on the subsequent slots. Therefore, lower pilot overhead is achieved than the transmission scheme in [18].
- We organize the received signals at the BS as a third-order low-rank tensor by leveraging the above transmission scheme and channel sparsity in the angle domain, and the tensor can fit a CP model. The three dimensions of tensor signal are space domain, time domain and coding domain. The uniqueness of the CP model can ensure that the decomposition does not change the column structure of the factor matrices which contain CSI. Furthermore, the angle pairing is not needed in this scheme since all factor matrices share the same permutation matrix. Considering multiple dimensions also makes our estimation more accurate.
- We further explore the Vandermonde structure of the third-factor matrix when the Doppler shifts remain constant during one frame. By exploiting the Vandermonde structure and the subspace method, we propose to utilize an algorithm which only uses basic linear algebra to solve the CP decomposition. Compared with the traditional iterative method, the proposed method has better robustness and accuracy. Furthermore, we compare the success rate of the proposed method with that of the ALS-based method [33], which can verify that the proposed method achieves better performance.
- Under the uniqueness condition, we derive the relationship between pilot overhead and the number of channel paths, which verifies that the proposed method can achieve substantial training overhead reduction. Besides, we analyze the performance of the proposed method both in single-path channels and multi-path channels for uplink MU-MIMO systems.

The rest of the paper is organized as follows. In Section II, we provide notations and some basic definitions of tensor decomposition. Section III introduces the model for the time-varying channel and signal model for MU-MIMO systems. The tensor-based channel estimation algorithm, uniqueness condition, and computational complexity of the proposed algorithm are analyzed in Section IV. Simulation results are presented in Section V. Finally, the conclusions are drawn in Section VI.

II. NOTATIONS AND PRELIMINARIES

A. Notations

Lowercase letters, boldface lowercase letters, boldface capital letters, and Euler script letters denote scalars, vectors, matrices, and tensors, respectively, e.g., x , \mathbf{x} , \mathbf{X} , \mathcal{X} . Superscripts $(\cdot)^T$, $(\cdot)^*$, $(\cdot)^H$, $(\cdot)^{-1}$, and $(\cdot)^\dagger$ denote the transpose, conjugate, Hermitian transpose, matrix inverse, and pseudo inverse, respectively. Subscript $[\mathbf{x}]_i$, $[\mathbf{X}]_{i,j}$, $[\mathcal{X}]_{i,j,k}$, \mathbf{X}_r , $[\mathbf{X}]_{:,m:n}$ denote the i -th entry of \mathbf{x} , element (i, j) of \mathbf{X} , element (i, j, k) of third-order tensor \mathcal{X} ,

the r -th column of \mathbf{X} and submatrix of \mathbf{X} from the m -th to the n -th rows, respectively. The $\mathbf{X}^{(n)}$ denotes the n -th matrix in a sequence. We use $\|\cdot\|_F$ and $\|\cdot\|_2$ for Frobenius norm and l_2 -norm. The $\tilde{\mathbf{X}}$ denotes an estimate of \mathbf{X} . The symbols \otimes , \odot , $*$, \circ denote Kronecker, Khatri-Rao, Hadamard, and outer product, respectively. We use $D(\mathbf{x})$ and $d(\mathbf{X})$ to denote the diagonal matrix formed by \mathbf{x} and the diagonal element vector of \mathbf{X} , and $\angle(z)$ denotes the operation to extract the phase of the complex number z . The $n \times n$ identity matrix is denoted by \mathbf{I}_n .

B. Tensor Preliminaries

A tensor in this paper is a multidimensional array [33]. Vectors and matrices can be viewed as the first-order tensor and second-order tensor, respectively. Let $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times \dots \times I_N}$ denote the N -th order tensor. We briefly provide some preliminaries on tensor and CP decomposition. More details can be found in [33], [34].

Definition 1 (rank-one tensor): The N -th order tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times \dots \times I_N}$ is a rank-one tensor if it can be represented as the outer product of N vectors $\{\mathbf{a}^n \in \mathbb{C}^{I_n}\}_{n=1}^N$, i.e.,

$$\mathcal{X} = \mathbf{a}^{(1)} \circ \mathbf{a}^{(2)} \circ \dots \circ \mathbf{a}^{(N)}. \quad (1)$$

Definition 2 (tensor rank): The rank of tensor \mathcal{X} , denoted by $\text{rank}(\mathcal{X})$, is defined as the minimum number of rank-one tensors needed to form \mathcal{X} as their sum.

Definition 3 (CP decomposition): The CP decomposition factorizes a tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times \dots \times I_N}$ into R rank-one tensors, which can be written as

$$\mathcal{X} = \sum_{r=1}^R \mathbf{a}_r^{(1)} \circ \mathbf{a}_r^{(2)} \circ \dots \circ \mathbf{a}_r^{(N)} \quad (2)$$

where R denotes the rank of tensor. The CP decomposition can be concisely written as

$$\mathcal{X} = [\mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \dots, \mathbf{A}^{(N)}]. \quad (3)$$

Definition 4 (unfolding): The unfolding of tensor \mathcal{X} in the n -th mode is denoted by $\mathbf{X}_{(n)}$, which is the process of reorganizing the elements of \mathcal{X} into $\mathbf{X}_{(n)}$. The element (i_1, i_2, \dots, i_N) of \mathcal{X} maps to matrix element (i_n, j) of $\mathbf{X}_{(n)}$, where $j = 1 + \sum_{k=1, k \neq n}^N (i_k - 1)J_k$ with $J_k = \prod_{m=1, m \neq k}^{k-1} I_m$. By exploiting factor matrices, $\mathbf{X}_{(n)}$ can be written as

$$\mathbf{X}_{(n)} = \mathbf{A}^{(n)} \left(\mathbf{A}^{(N)} \odot \dots \odot \mathbf{A}^{(n+1)} \odot \mathbf{A}^{(n-1)} \odot \dots \odot \mathbf{A}^{(1)} \right)^T. \quad (4)$$

Definition 5 (k -rank): The k -rank of a matrix \mathbf{X} is denoted by $k_{\mathbf{X}}$, which is the largest value R such that any R columns of \mathbf{X} are linearly independent.

Definition 6 (essentially unique): The essential uniqueness means the factor matrices of CP decomposition are unique up to permutation and scaling ambiguity, i.e., if $\mathcal{X} = [\tilde{\mathbf{A}}^{(1)}, \tilde{\mathbf{A}}^{(2)}, \dots, \tilde{\mathbf{A}}^{(N)}]$ holds, there exists a permutation matrix Π and diagonal scaling matrices $\{\Delta^{(n)}\}_{n=1}^N$, such that $\{\tilde{\mathbf{A}}^{(n)} = \mathbf{A}^{(n)} \Pi \Delta^{(n)}\}_{n=1}^N$ and $\Pi_{n=1}^N \Delta^{(n)} = \mathbf{I}_R$.

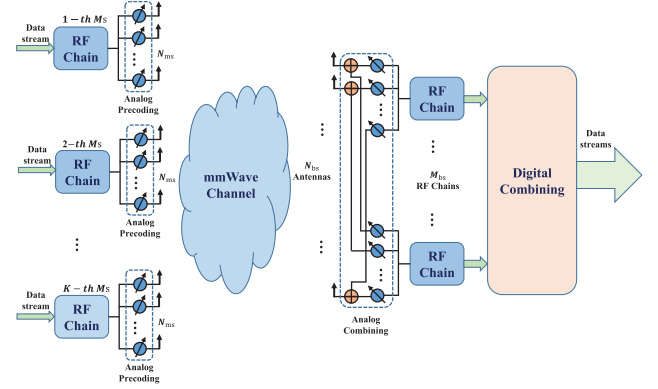


Fig. 1. The uplink mmWave MU-MIMO system with hybrid combining at the BS and analog precoder at MSs.

III. SYSTEM MODEL AND PROBLEM STATEMENT

A. Time-Varying Channel

We consider an uplink MU-MIMO mmWave system consisting of one BS and K MSs, as shown in Fig. 1. To facilitate the hardware implementation, a fully connected hybrid combiner structure [4] is employed by the BS and analog precoder is employed by each MS. The BS is equipped with N_{BS} antennas and M_{BS} RF chains with $N_{BS} > M_{BS}$, and each MS is equipped with N_{MS} antennas and there is only one RF chain. The MSs have different velocities, and we aim to jointly estimate the CSI of all MSs at the BS. By leveraging the sparsity nature of mmWave channels, we assume a geometric channel model between the BS and the k -th MS, which consists of L_k scatterers. Compared with the phase fluctuation caused by the Doppler shift, the path gains $(\alpha_{l,k}(p))$, AoAs $(\theta_{l,k}(p))$ and AoDs $(\varphi_{l,k}(p))$ vary negligibly in high-mobility scenarios, since they are from the large-scale scattering environment and the influence of mobility is limited in a short period of time [4]. Thus the variations of channel in a short time can be regarded as only affected by Doppler shift, and it is reasonable to assume that the path gains and angles keep constant during one frame [18], [24], i.e.,

$$\alpha_{l,k}(p) = \alpha_{l,k}, \theta_{l,k}(p) = \theta_{l,k}, \varphi_{l,k}(p) = \varphi_{l,k}. \quad (5)$$

Under this assumption, the time-varying channel $\mathbf{H}_k(t) \in \mathbb{C}^{N_{BS} \times N_{MS}}$ at the t -th time slot of the k -th MS can be written as [3], [4]

$$\mathbf{H}_k(t) = \sum_{l=1}^{L_k} \alpha_{l,k} e^{j2\pi f_{l,k}^d P_s T_s t} \mathbf{a}_{BS}(\theta_{l,k}) \mathbf{a}_{MS}^H(\varphi_{l,k}) \quad (6)$$

where $\alpha_{l,k} \sim \mathcal{CN}(0, \sigma_\alpha^2)$, $\theta_{l,k}$, and $\varphi_{l,k}$ denote the complex path gain, AoA, and AoD of the k -th MS's l -th path, respectively, f^d and T_s denote Doppler shift and sampling period, respectively. The steering vectors of the BS and MS antenna array are represented by $\mathbf{a}_{BS}(\theta) \in \mathbb{C}^{N_{BS} \times 1}$ and $\mathbf{a}_{MS}(\varphi) \in \mathbb{C}^{N_{MS} \times 1}$. For a uniform linear array (ULA), the vectors are given by

$$\mathbf{a}_{BS}(\theta) \triangleq \frac{1}{\sqrt{N_{BS}}} \left[1, e^{j\frac{2\pi}{\lambda} d \sin(\theta)}, \dots, e^{j\frac{2\pi}{\lambda} d (N_{BS}-1) \sin(\theta)} \right]^T$$

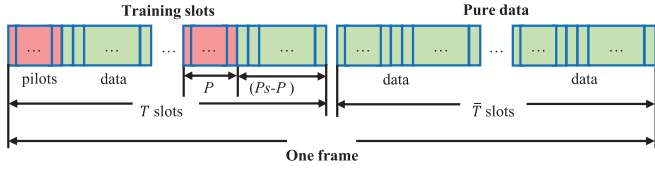


Fig. 2. The transmission structure for uplink time-varying channel.

$$\mathbf{a}_{MS}(\varphi) \triangleq \frac{1}{\sqrt{N_{MS}}} \left[1, e^{j\frac{2\pi}{\lambda}d \sin(\varphi)}, \dots, e^{j\frac{2\pi}{\lambda}d(N_{MS}-1) \sin(\varphi)} \right]^T \quad (7)$$

where d denotes the antenna spacing between two adjacent elements and λ denotes the signal wavelength, typically $d = \frac{\lambda}{2}$.

B. Signal Model

A novel pilot transmission scheme is designed for time-varying channel. Unlike the design in IEEE 802.11a where the estimation only uses a known preamble in the frame header for the whole frame, we insert pilots into the first several slots of each frame. Specifically, a transmission frame is comprised of $T + \bar{T}$ time slots. The first T slots are used to track the channel and the CSI of the remaining \bar{T} slots can be dynamically predicted with the channel model and the estimated parameters. In this design, each time slot is divided into P_s intervals. In the first T time slots, the top P ($P \ll P_s$) intervals are chosen for training and the rest for data transmission. The coherence time needs to be greater than P intervals. The MSs transmit the pilot symbols precoded by the precoding vector $\mathbf{f}_{k,p} \in \mathbb{C}^{N_{MS} \times 1}$, $p = 1, \dots, P$ at the p -th intervals. Without loss of generality, we set the pilot symbol $s_k(p) = 1$. The signals can be received simultaneously by associating with $\mathbf{w}_m \in \mathbb{C}^{N_{BS} \times 1}$, $m = 1, \dots, M_{BS}$ at the BS. The transmission scheme is illustrated in Fig. 2. In the stage of channel estimation, we do not consider the design of RF precoder and combiner. Since $\mathbf{f}_{k,p}$ and \mathbf{w}_m are implemented by analog phase shifters, their entries contain constant magnitudes, i.e. $|\mathbf{f}_{k,p}|_i = 1$, $|\mathbf{w}_m|_i = 1$, $\forall i$. Therefore, the received signal in the m -th RF chain at the p -th intervals of the t -th time slot can be written as

$$y_{m,t,p} = \sum_{k=1}^K \mathbf{w}_m^H \mathbf{H}_k(t) \mathbf{f}_{k,p} + n_{m,t,p} \quad (8)$$

where $n_{m,t,p}$ is zero mean, circularly symmetric complex Gaussian noise. These received signals can be put in a vector as

$$\mathbf{y}_{m,t} = \sum_{k=1}^K \mathbf{w}_m^H \mathbf{H}_k(t) \mathbf{F}_k + \mathbf{n}_{m,t} \quad (9)$$

and in a matrix as

$$\begin{aligned} \mathbf{Y}_t &= \sum_{k=1}^K \mathbf{W}^H \mathbf{H}_k(t) \mathbf{F}_k + \mathbf{N}_t \\ &= \sum_{k=1}^K \sum_{i=1}^{L_k} \alpha_{i,k} e^{j2\pi f_l^d T_s P_s t} \mathbf{W}^H \mathbf{a}_{BS}(\theta_{i,k}) \mathbf{a}_{MS}^H(\varphi_{i,k}) \mathbf{F}_k + \mathbf{N}_t \end{aligned}$$

$$\begin{aligned} &= \sum_{l=1}^L \alpha_l e^{j2\pi f_l^d T_s P_s t} \mathbf{W}^H \mathbf{a}_{BS}(\theta_l) \mathbf{a}_{MS}^H(\varphi_l) \bar{\mathbf{F}} + \mathbf{N}_t \\ &= \sum_{l=1}^L e^{j2\pi f_l^d T_s P_s t} \cdot \hat{\mathbf{a}}_{BS}(\theta_l) \circ \hat{\mathbf{a}}_{MS}(\varphi_l) + \mathbf{N}_t \end{aligned} \quad (10)$$

where the parameters are defined as

$$\mathbf{W} \triangleq [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_{M_{BS}}] \in \mathbb{C}^{N_{BS} \times M_{BS}} \quad (11a)$$

$$\mathbf{F}_k \triangleq [\mathbf{f}_{k,1}, \mathbf{f}_{k,2}, \dots, \mathbf{f}_{k,P}] \in \mathbb{C}^{N_{MS} \times P} \quad (11b)$$

$$\hat{\mathbf{a}}_{BS}(\theta_l) \triangleq \alpha_l \mathbf{W}^H \mathbf{a}_{BS}(\theta_l) \in \mathbb{C}^{M_{BS} \times 1} \quad (11c)$$

$$\hat{\mathbf{a}}_{MS}(\varphi_l) \triangleq \bar{\mathbf{F}}^T \mathbf{a}_{MS}^*(\varphi_l) \in \mathbb{C}^{P \times 1}. \quad (11d)$$

Correspondingly $L = \sum_{k=1}^K L_k$, $\bar{\mathbf{F}}$ represents \mathbf{F}_k when the l -th path belongs to k -th MS, where $l = \sum_{j=1}^{k-1} L_j + i$. In (10), the received signal matrix at the t -th time slot can be formed as a weighted sum of a set of rank-one outer products. Thus, the low-rank structure tensor \mathcal{Y} which admits the CP decomposition can be obtained by concatenating the received signal matrices in T time slots, where \mathbf{Y}_t , $t = 0, \dots, (T-1)$ are slices of \mathcal{Y} . We have

$$\begin{aligned} \mathcal{Y} &= \sum_{l=1}^L \hat{\mathbf{a}}_{BS}(\theta_l) \circ \hat{\mathbf{a}}_{MS}(\varphi_l) \circ \boldsymbol{\beta}_l + \mathcal{N} \\ &= [\mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \mathbf{A}^{(3)}] + \mathcal{N} \end{aligned} \quad (12)$$

where $\boldsymbol{\beta}_l \triangleq [1, e^{j2\pi f_l^d T_s P_s}, \dots, e^{j2\pi f_l^d T_s (T-1) P_s}]^T \in \mathbb{C}^{T \times 1}$, and $\mathcal{N} \in \mathbb{C}^{M_{BS} \times P \times T}$ denotes the noise tensor. Our target is to estimate the key parameters $\{f_l^d, \alpha_l, \theta_l, \varphi_l\}_{l=1}^L$ from the received tensor. To do this, the third-order tensor \mathcal{Y} can be factorized into three factor matrices, and the essentially unique (Definition 6) can be ensured under mild conditions. In (12), the factor matrices can be defined as

$$\begin{aligned} \mathbf{A}^{(1)} &\triangleq [\hat{\mathbf{a}}_{BS}(\theta_1), \hat{\mathbf{a}}_{BS}(\theta_2), \dots, \hat{\mathbf{a}}_{BS}(\theta_L)] \\ \mathbf{A}^{(2)} &\triangleq [\hat{\mathbf{a}}_{MS}(\varphi_1), \hat{\mathbf{a}}_{MS}(\varphi_2), \dots, \hat{\mathbf{a}}_{MS}(\varphi_L)] \\ \mathbf{A}^{(3)} &\triangleq [\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_L] \end{aligned} \quad (13)$$

Note that the entries of each column in $\mathbf{A}^{(3)}$ constitute a geometric series. Thus, $\mathbf{A}^{(3)}$ is a Vandermonde matrix with distinct generators $\{z_l = e^{j2\pi f_l^d T_s P_s}\}_{l=1}^L$.

IV. CHANNEL ESTIMATION FOR TIME-VARYING MU-MIMO SYSTEMS

In this section, the tensor-based channel estimation algorithm will be given. The first step is to solve the CP decomposition, from which we can get three factor matrices. Then, the channel parameters will be calculated from the three factor matrices. To have a better comparison, we will firstly introduce the traditional

Algorithm 1: ALS-Related Algorithm.

Require: the received signal tensor $\mathcal{Y} \in \mathbb{C}^{M_{BS} \times P \times T}$, maximum iterations J_{iter} , regularization parameter μ and convergence threshold ε .

- 1: Generate random initial factor matrices $\mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \mathbf{A}^{(3)}, err_{old}$.
- 2: **for** $i = 1 : J_{iter}$ **do**
- 3: calculate $\mathbf{A}_{i+1}^{(1)}, \mathbf{A}_{i+1}^{(2)}, \mathbf{A}_{i+1}^{(3)}$ from (15).
- 4: $\tilde{\mathcal{Y}}_{i+1} = \llbracket \mathbf{A}_{i+1}^{(1)}, \mathbf{A}_{i+1}^{(2)}, \mathbf{A}_{i+1}^{(3)} \rrbracket$
- 5: $err_{new} = \|\tilde{\mathcal{Y}}_{i+1} - \mathcal{Y}\|_F^2 / \|\tilde{\mathcal{Y}}_i\|_F^2$.
- 6: **if** $|err_{old} - err_{new}| < \varepsilon$ **then**
- 7: **break**;
- 8: **end if**
- 9: $err_{old} = err_{new}$;
- 10: **end for**
- 11: **return** $\tilde{\mathbf{A}}^{(1)}, \tilde{\mathbf{A}}^{(2)}, \tilde{\mathbf{A}}^{(3)}$.

method and then give the subspace based method to solve the CP model. Besides, the condition of essential uniqueness is analyzed, based on which we can derive the relationship between pilots overhead and the number of channel paths. For simplicity, the noise term is omitted in the following derivation.

A. CP Decomposition

1) *ALS:* To perform the CP decomposition, (12) can be formulated into the following problem

$$\min_{\mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \mathbf{A}^{(3)}} \|\mathcal{Y} - \tilde{\mathcal{Y}}\|_F^2 \quad (14)$$

where $\tilde{\mathcal{Y}} = \llbracket \mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \mathbf{A}^{(3)} \rrbracket$. The most popular way to compute the optimization problem in (14) is ALS algorithm [33], [34], which iteratively updates factor matrices by fixing all but one matrix. However, this iterative method has three issues: 1) The convergence rate is unpredictable; 2) The convergence cannot be guaranteed to the global optimum; 3) The quality of the initialization seriously affects the final solution. For the first issue, the regularized ALS (RALs) adds a corresponding regularization term to each sub-problem to increase the convergence rate, known as the proximal point modification of the Gauss-Seidel method (PGS) [35]. The sub-problems solved by iteration can be represented as

$$\begin{aligned} \mathbf{A}_{i+1}^{(1)} &\leftarrow \arg \min_{\mathbf{A}^{(1)}} \|\mathbf{Y}_{(1)} - \mathbf{A}_i^{(1)} (\mathbf{A}_i^{(3)} \odot \mathbf{A}_i^{(2)})^T\|_F^2 + \mu \|\mathbf{A}_i^{(1)}\|_F^2 \\ \mathbf{A}_{i+1}^{(2)} &\leftarrow \arg \min_{\mathbf{A}^{(2)}} \|\mathbf{Y}_{(2)} - \mathbf{A}_i^{(2)} (\mathbf{A}_i^{(3)} \odot \mathbf{A}_i^{(1)})^T\|_F^2 + \mu \|\mathbf{A}_i^{(2)}\|_F^2 \\ \mathbf{A}_{i+1}^{(3)} &\leftarrow \arg \min_{\mathbf{A}^{(3)}} \|\mathbf{Y}_{(3)} - \mathbf{A}_i^{(3)} (\mathbf{A}_{i+1}^{(2)} \odot \mathbf{A}_{i+1}^{(1)})^T\|_F^2 + \mu \|\mathbf{A}_i^{(3)}\|_F^2 \end{aligned} \quad (15)$$

where $\{\mathbf{A}_{i+1}^{(n)}\}_{n=1}^3$ are the solutions in the i -th iteration and μ represents the regularization parameter. The ALS-related algorithm is presented in Algorithm 1.

2) *Subspace Based Method:* Tensor decomposition is matrix decomposition in essence. Therefore, we can use the structural

characteristics of the matrix to solve CP decomposition by matrix decomposition. By leveraging the Vandermonde structure of $\mathbf{A}^{(3)}$, we propose to use linear algebra [36] to solve CP decomposition. Firstly, we choose a pair (K_1, L_1) subjected to $K_1 + L_1 = T + 1$ (such as $K_1 = \lfloor (T + 1)/2 \rfloor, L_1 = T + 1 - K_1$), and define a cyclic selection matrix

$$\mathbf{J}_l = [\mathbf{0}_{K_1 \times (l-1)} \quad \mathbf{I}_{K_1} \quad \mathbf{0}_{K_1 \times (L_1-1)}] \in \mathbb{C}^{K_1 \times T}. \quad (16)$$

Left multiplying \mathbf{J}_l by $\mathbf{A}^{(3)}$ can choose the submatrix of $\mathbf{A}^{(3)}$ from the l -th to the $(l + K_1)$ -th rows, i.e., $[\mathbf{A}^{(3)}]_{l:l+K_1}$. According to Definition 4, the unfolding of \mathcal{Y} in the first dimension can be written as

$$\mathbf{Y}_{(1)} = \mathbf{A}^{(1)} (\mathbf{A}^{(3)} \odot \mathbf{A}^{(2)})^T \in \mathbb{C}^{M_{BS} \times TP}. \quad (17)$$

Thus, we can expand the dimension of $\mathbf{Y}_{(1)}$ as

$$\begin{aligned} \mathbf{Y}^s &= [(\mathbf{J}_1 \otimes \mathbf{I}_P) \mathbf{Y}_{(1)}^T, \dots, (\mathbf{J}_{L_1} \otimes \mathbf{I}_P) \mathbf{Y}_{(1)}^T] \\ &= \left[\left((\mathbf{J}_1 \mathbf{A}^{(3)}) \odot \mathbf{A}^{(2)} \right) \mathbf{A}^{(1)T}, \dots, \left((\mathbf{J}_{L_1} \mathbf{A}^{(3)}) \odot \mathbf{A}^{(2)} \right) \mathbf{A}^{(1)T} \right]. \end{aligned} \quad (18)$$

Since $\mathbf{A}^{(3)}$ is a Vandermonde matrix, we have

$$\mathbf{J}_l \mathbf{A}^{(3)} = \mathbf{J}_l \mathbf{A}^{(3)} \mathbf{D} \left(\left[\mathbf{A}^{(3)} \right]_{l,:} \right). \quad (19)$$

Equation (18) can be rewritten as

$$\begin{aligned} \mathbf{Y}^s &= \left[\left((\mathbf{J}_1 \mathbf{A}^{(3)} \mathbf{D}([\mathbf{A}^{(3)}]_{1,:})) \odot \mathbf{A}^{(2)} \right) \mathbf{A}^{(1)T}, \dots, \right. \\ &\quad \left. \left((\mathbf{J}_1 \mathbf{A}^{(3)} \mathbf{D}([\mathbf{A}^{(3)}]_{L_1,:})) \odot \mathbf{A}^{(2)} \right) \mathbf{A}^{(1)T} \right] \\ &= \left[\left((\mathbf{J}_1 \mathbf{A}^{(3)}) \odot \mathbf{A}^{(2)} \mathbf{D}([\mathbf{A}^{(3)}]_{1,:}) \right) \mathbf{A}^{(1)T}, \dots, \right. \\ &\quad \left. \left((\mathbf{J}_1 \mathbf{A}^{(3)}) \odot \mathbf{A}^{(2)} \mathbf{D}([\mathbf{A}^{(3)}]_{L_1,:}) \right) \mathbf{A}^{(1)T} \right] \\ &= (\mathbf{A}^{(K_1,3)} \odot \mathbf{A}^{(2)}) \left[\mathbf{D}([\mathbf{A}^{(3)}]_{1,:}) \mathbf{A}^{(1)T}, \dots, \right. \\ &\quad \left. \mathbf{D}([\mathbf{A}^{(3)}]_{L_1,:}) \mathbf{A}^{(1)T} \right] \\ &= (\mathbf{A}^{(K_1,3)} \odot \mathbf{A}^{(2)}) (\mathbf{A}^{(L_1,3)} \odot \mathbf{A}^{(1)})^T \end{aligned} \quad (20)$$

where $\mathbf{A}^{(K_1,3)}$ denotes $[\mathbf{A}^{(3)}]_{1:K_1,:}$. Consider the blind matrix (20), what we know is that the \mathbf{Y}^s is expanded by the observed data. The $(\mathbf{A}^{(K_1,3)} \odot \mathbf{A}^{(2)})$ and $(\mathbf{A}^{(L_1,3)} \odot \mathbf{A}^{(1)})$ are of full column rank and row rank, respectively. Let $\mathbf{Y}^s = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^H$ denote the truncated Singular Value Decomposition (SVD) of \mathbf{Y}^s . Since the columns of \mathbf{U} span the same subspace as the columns of \mathbf{Y}^s , there exists a nonsingular matrix $\mathbf{M} \in \mathbb{C}^{L \times L}$ satisfying

$$\mathbf{U} \mathbf{M} = \mathbf{A}^{(K_1,3)} \odot \mathbf{A}^{(2)} \quad (21a)$$

$$\mathbf{V}^* \mathbf{\Sigma} \mathbf{M}^{-T} = \mathbf{A}^{(L_1,3)} \odot \mathbf{A}^{(1)}. \quad (21b)$$

For convenience, we define

$$\begin{aligned} \mathbf{U}_1 &\triangleq [\mathbf{U}]_{1:(K_1-1)P_s} \in \mathbb{C}^{K_1(P-1) \times L} \\ \mathbf{U}_2 &\triangleq [\mathbf{U}]_{P+1:K_1P_s} \in \mathbb{C}^{K_1(P-1) \times L} \end{aligned} \quad (22a)$$

$$\begin{aligned} \underline{\mathbf{A}}^{(K_1,3)} &\triangleq \left[\mathbf{A}^{(K_1,3)} \right]_{1:K_1-1,;} \in \mathbb{C}^{(K_1-1) \times L} \\ \overline{\mathbf{A}}^{(K_1,3)} &\triangleq \left[\mathbf{A}^{(K_1,3)} \right]_{2:K_1,;} \in \mathbb{C}^{(K_1-1) \times L}. \end{aligned} \quad (22b)$$

Since $\mathbf{A}^{(K_1,3)}$ is a Vandermonde matrix, we have

$$\overline{\mathbf{A}}^{(K_1,3)} = \underline{\mathbf{A}}^{(K_1,3)} \mathbf{Z} \quad (23)$$

where $\mathbf{Z} = \mathbf{D}([z_1, z_2, \dots, z_L])$, and $\{z_l = e^{j2\pi f_l^d T_s P_s}\}_{l=1}^L$ are generators of $\mathbf{A}^{(3)}$. By combining (21) and (23), we can obtain

$$\mathbf{U}_2 \mathbf{M} = \mathbf{U}_1 \mathbf{M} \mathbf{Z} \quad (24)$$

$$\mathbf{U}_1^\dagger \mathbf{U}_2 = \mathbf{M} \mathbf{Z} \mathbf{M}^{-1} \triangleq \hat{\mathbf{Z}}. \quad (25)$$

We note that \mathbf{Z} is a diagonal matrix, whose diagonal elements are $\{z_l\}_{l=1}^L$. It can be noticed that (25) is a typical similarity transformation. The Doppler shifts can be extracted from the eigenvalues of $\hat{\mathbf{Z}}$. Let $\hat{\mathbf{Z}} = \mathbf{Q} \mathbf{\Lambda} \mathbf{Q}^{-1}$ denote Eigenvalue Decomposition (EVD), and we have

$$d(\mathbf{\Lambda}) = d(\mathbf{Z}) \mathbf{\Pi} \quad (26a)$$

$$\mathbf{Q} = \mathbf{M} \mathbf{\Delta} \mathbf{\Pi} \quad (26b)$$

where $\mathbf{\Pi}$ and $\mathbf{\Delta}$ represent the permutation matrix and diagonal scaling matrix, respectively. Thus, the eigenvalues of $\hat{\mathbf{Z}}$ are the diagonal elements of \mathbf{Z} with permutation ambiguity. By exploiting the estimated generators, we can recover the Vandermonde matrix $\mathbf{B}^{(3)}$ from

$$\mathbf{B}^{(3)} = \mathbf{A}^{(3)} \mathbf{\Pi}. \quad (27)$$

The next step is to find $\mathbf{A}^{(2)}$ up to permutation and scaling ambiguity. According to the properties of Kronecker product, we can get

$$\left(\frac{\mathbf{b}_r^{(K_1,3)H}}{\mathbf{b}_r^{(K_1,3)H} \mathbf{b}_r^{(K_1,3)}} \otimes \mathbf{I}_P \right) (\mathbf{b}_r^{(K_1,3)} \otimes \mathbf{b}_r^{(2)}) = \mathbf{b}_r^{(2)} \quad (28)$$

where $\mathbf{b}_r^{(K_1,3)}$ is the r -th column of $\mathbf{B}^{(K_1,3)}$. Since $\mathbf{b}_r^{(3)}$ is generated by $[\mathbf{\Lambda}]_{r,r}$, we obtain $\mathbf{b}_r^{(K_1,3)H} \mathbf{b}_r^{(K_1,3)} = K_1$. By combining with (21a), (26b), and (27), we get

$$\mathbf{b}_r^{(K_1,3)} \otimes \mathbf{b}_r^{(2)} = \mathbf{U} \mathbf{q}_r. \quad (29)$$

Then, (28) can be rewritten as

$$\left(\frac{\mathbf{b}_r^{(K_1,3)H}}{K_1} \otimes \mathbf{I}_P \right) \mathbf{U} \mathbf{q}_r = \mathbf{b}_r^{(2)}. \quad (30)$$

Thus, we can recover $\mathbf{B}^{(2)} = [\mathbf{b}_1^{(2)}, \dots, \mathbf{b}_L^{(2)}]$, which satisfies

$$\mathbf{B}^{(2)} = \mathbf{A}^{(2)} \mathbf{\Delta} \mathbf{\Pi}. \quad (31)$$

The procedure to get $\mathbf{A}^{(1)}$ is the same as $\mathbf{A}^{(2)}$. We have

$$\left(\frac{\mathbf{b}_r^{(L_1,3)H}}{\mathbf{b}_r^{(L_1,3)H} \mathbf{b}_r^{(L_1,3)}} \otimes \mathbf{I}_{M_{BS}} \right) (\mathbf{b}_r^{(L_1,3)} \otimes \mathbf{b}_r^{(1)}) = \mathbf{b}_r^{(1)}. \quad (32)$$

By combining (21b), (26), and (27), we can get

$$\left(\mathbf{b}_r^{(L_1,3)} \otimes \mathbf{b}_r^{(1)} \right) = \mathbf{V}^* \mathbf{\Sigma} [\mathbf{Q}^{-T}]_{:,r}. \quad (33)$$

Thus (32) can be rewritten as

$$\left(\frac{\mathbf{b}_r^{(L_1,3)H}}{L_1} \otimes \mathbf{I}_{M_{BS}} \right) \mathbf{V}^* \mathbf{\Sigma} [\mathbf{Q}^{-T}]_{:,r} = \mathbf{b}_r^{(1)}. \quad (34)$$

We can arrange the $\mathbf{B}^{(1)} = [\mathbf{b}_1^{(1)}, \dots, \mathbf{b}_L^{(1)}]$, which satisfies

$$\mathbf{B}^{(1)} = \mathbf{A}^{(1)} \mathbf{\Delta}^{-1} \mathbf{\Pi}. \quad (35)$$

Consequently, we have obtained $\{\mathbf{B}^{(n)}\}_{n=1}^3$ as the estimated factor matrices of \mathcal{Y} up to permutation and scaling ambiguity. We do not need to pair the AoA, AoD and gain for each propagation path as [26], because the factor matrices share the same permutation matrix $\mathbf{\Pi}$. The ambiguity will be eliminated in the following subsection.

B. Parameters Estimation and Ambiguity Elimination

We observe from (13) that the columns of the factor matrices are related to the AoAs and AoDs. Furthermore, the scaling and permutation ambiguity do not affect the structure of the columns of factor matrices. Therefore, the AoAs and AoDs can be estimated from the columns of $\mathbf{B}^{(1)}$ and $\mathbf{B}^{(2)}$. In fully-digital beamforming architecture, \mathbf{W} is a square matrix and the angles can be identified by a rotationally invariant technique after calculating the inverse of \mathbf{W} . However, we consider a hybrid beamforming architecture in which the number of RF chains is smaller than the number of antennas. It is an underdetermined system of equations when we try to eliminate the \mathbf{W} from $\mathbf{B}^{(1)}$. A cosine correlation based method can be used to handle this problem as

$$\tilde{\theta}_r = \arg \max_{\theta_r} \frac{|\mathbf{b}_r^{(1)H} (\mathbf{W}^H \tilde{\mathbf{a}}_{BS}(\theta_r))|}{\|\mathbf{b}_r^{(1)}\|_2 \|(\mathbf{W}^H \tilde{\mathbf{a}}_{BS}(\theta_r))\|_2} \quad (36)$$

$$\tilde{\varphi}_r = \arg \max_{\varphi_r} \frac{|\mathbf{b}_r^{(2)H} (\tilde{\mathbf{F}}^T \tilde{\mathbf{a}}_{MS}(\varphi_r))|}{\|\mathbf{b}_r^{(2)}\|_2 \|(\tilde{\mathbf{F}}^T \tilde{\mathbf{a}}_{MS}(\varphi_r))\|_2}. \quad (37)$$

The maximization problems (36)–(37) can also be effectively solved by performing an one-dimensional search and we can achieve high accuracy by using dense search grids. Then, the Doppler shifts can be derived from the generators of $\mathbf{B}^{(3)}$ as

$$\tilde{f}_r^d = \frac{1}{2\pi T_s P_s} \angle (\mathbf{\Lambda}_{r,r}). \quad (38)$$

Using the estimated AoAs, we define

$$\mathbf{C}^{(1)} = \mathbf{W}^H [\mathbf{a}_{BS}(\tilde{\theta}_1), \mathbf{a}_{BS}(\tilde{\theta}_2), \dots, \mathbf{a}_{BS}(\tilde{\theta}_L)]. \quad (39)$$

Then one has

$$\begin{aligned} \mathbf{Y}_{(3)} &= \mathbf{A}^{(3)} \left(\mathbf{A}^{(2)} \odot \mathbf{A}^{(1)} \right)^T \in \mathbb{C}^{M_{BS} \times TP} \\ &\approx \mathbf{B}^{(3)} \left(\mathbf{B}^{(2)} \odot \mathbf{B}^{(1)} \right)^T \\ &= \mathbf{B}^{(3)} \left(\mathbf{B}^{(2)} \odot \left(\mathbf{C}^{(1)} \mathbf{\Gamma} \right) \right)^T \end{aligned}$$

Algorithm 2: Tensor-Subspace Based Channel Estimation Algorithm.

- Require:** the received signal tensor $\mathcal{Y} \in \mathbb{C}^{M_{BS} \times P \times T}$.
- 1: Choose a pair (K_1, L_1) subject to $K_1 + L_1 = T + 1$, and rearrange \mathcal{Y} as $\mathbf{Y}^s \in \mathbb{C}^{K_1 P \times L_1 M_{BS}}$ by (20).
 - 2: Calculate the Truncated-SVD of \mathbf{Y}^s as $\mathbf{Y}^s = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^H$.
 - 3: Select submatrices \mathbf{U}_1 and \mathbf{U}_2 by (22a) and calculate the EVD as $\mathbf{U}_1^\dagger \mathbf{U}_2 = \mathbf{Q} \mathbf{\Lambda} \mathbf{Q}^{-1}$.
 - 4: Consider the diagonal elements of $\mathbf{\Lambda}$ as the generators to reconstruct $\mathbf{B}^{(3)}$.
 - 5: Reconstruct $\mathbf{B}^{(2)}$ and $\mathbf{B}^{(1)}$ by (30) and (34), respectively.
 - 6: Estimate the AoAs $\{\tilde{\theta}_r\}_{r=1}^L$, AoDs $\{\tilde{\varphi}_r\}_{r=1}^L$, Doppler shifts and path gains $\{\tilde{f}_r^d\}_{r=1}^L$ by (36), (37), (38), and (42), respectively.
 - 7: **return** channel parameters $\{\tilde{\theta}_r, \tilde{\varphi}_r, \tilde{f}_r^d, \tilde{\alpha}_r\}_{r=1}^L$ and channel matrices $\{\tilde{\mathbf{H}}_k\}_{k=1}^K$.
-

$$= \mathbf{B}^{(3)} \mathbf{\Gamma} \left(\mathbf{B}^{(2)} \odot \mathbf{C}^{(1)} \right)^T \quad (40)$$

where $\mathbf{\Gamma}$ is a diagonal matrix whose diagonal elements are path gains which can be written as

$$\mathbf{\Gamma} \triangleq \mathbf{D}([\alpha_1, \dots, \alpha_L]). \quad (41)$$

Since $\text{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A})\text{vec}(\mathbf{B})$ holds, the gains can be obtained by linear square estimation as

$$\text{vec}(\mathbf{\Gamma}) = \left(\left(\mathbf{B}^{(2)} \odot \mathbf{C}^{(1)} \right) \otimes \mathbf{B}^{(3)} \right)^\dagger \text{vec}(\mathbf{Y}_{(3)}). \quad (42)$$

The proposed tensor-based channel estimation algorithm is summarized in Algorithm 2.

C. Uniqueness Analysis

In this section, we discuss the uniqueness condition of CP decomposition. Unlike matrix factorization, the CP decomposition is essentially unique under a mild condition. The most general conclusion on uniqueness is firstly discussed by Kruskal [37], depending on the k-rank (Definition 5). Specifically, we have the following theorem:

Theorem 1 (Kruskal condition [37]): Let $\mathcal{X} = [\mathbf{A}, \mathbf{B}, \mathbf{C}]$ be a CP solution which decomposes the third-order tensor \mathcal{X} into R rank-one tensor. The solution is essentially unique if $k_{\mathbf{A}} + k_{\mathbf{B}} + k_{\mathbf{C}} \geq 2R + 2$.

Theorem 2 (Vandermonde k-rank [38]): A Vandermonde matrix $\mathbf{V} \in \mathbb{C}^{n \times m}$ with distinct nonzero generators is not only full rank but is also full k-rank, such as $k_{\mathbf{V}} = \text{rank}(\mathbf{V}) = \min(n, m)$.

The proof for Theorem 1 can be found in [39]. From the above theorem, if the following condition holds

$$k_{\mathbf{A}^{(1)}} + k_{\mathbf{A}^{(2)}} + k_{\mathbf{A}^{(3)}} \geq 2L + 2 \quad (43)$$

then the $[\mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \mathbf{A}^{(3)}]$ is an essentially unique solution of \mathcal{Y} in the noiseless case. Following this, $[\mathbf{B}^{(1)}, \mathbf{B}^{(2)}, \mathbf{B}^{(3)}]$ is

guaranteed to be another solution of \mathcal{Y} with the form of (27), (31), and (35). Since $\mathbf{A}^{(3)} \in \mathbb{C}^{T \times L}$ is a Vandermonde matrix with distinct nonzero generators, according to Theorem 2, the k-rank of $\mathbf{A}^{(3)}$ is

$$k_{\mathbf{A}^{(3)}} = \min(T, L). \quad (44)$$

Considering the k-rank of $\mathbf{A}^{(1)}$ and $\mathbf{A}^{(2)}$, where the AoAs and AoDs follow a uniform distribution, and each entry of \mathbf{F}_k and \mathbf{W} is chosen uniformly from a unit circle. We can conclude that the $k_{\mathbf{A}^{(1)}}$ is equal to the smaller number of its column and row with probability one [22], i.e.,

$$k_{\mathbf{A}^{(1)}} = \min(M_{BS}, L) \quad (45)$$

$$k_{\mathbf{A}^{(2)}} = \min(P, L). \quad (46)$$

Therefore, the condition in (43) can be expressed as

$$\min(M_{BS}, L) + \min(P, L) + \min(T, L) \geq 2L + 2. \quad (47)$$

Assuming $M_{BS} \geq L$, we only need to ensure $P + T \geq L + 2$ when $T \leq L$ or $P \leq L$ to meet the condition in (45). As shown in pilot transmission scheme, we need to transmit $P \times T$ pilots symbols in one frame. With the determined L , the minimum overhead can be derived by the following linear programming problem

$$\begin{aligned} & \min_{T, P \in \mathbb{N}^+} TP \\ & \text{s.t.} \quad \begin{cases} P + T \geq L + 2 \\ T \geq 2 \\ P \geq 2 \end{cases} \end{aligned} \quad (48)$$

where \mathbb{N}^+ denotes positive integer. Since L is small in mmWave channel [4], the uniqueness condition is easy to satisfy with small T and P . It will be verified in our simulation experiments (Figs. 10 and 11), and properly increasing pilot overhead will improve the estimation performance.

Theorem 3 (k-rank of Khatri-Rao Product [38]): Given $\mathbf{A} \in \mathbb{C}^{I \times F}$ and $\mathbf{B} \in \mathbb{C}^{J \times F}$, if $k_{\mathbf{A}} + k_{\mathbf{B}} \geq F + 1$ and $IJ \geq F$, then $\mathbf{B} \odot \mathbf{A}$ is full column k-rank.

The aforementioned analysis uses the assumption that all angles are distinct. However, if there are two propagation paths with distinct AoAs but the same or very close AoDs, the condition in (43) cannot be satisfied since $k_{\mathbf{A}^{(2)}} = 1$. In our work, a factor matrices transformation method can be adopted to deal with this. We note that \mathbf{Y}^s in (20) can be regarded as the unfolding form of a new tensor and the new factor matrices are $\mathbf{A}^{(K_1, 3)} \odot \mathbf{A}^{(2)}$, $\mathbf{A}^{(L_1, 3)}$, and $\mathbf{A}^{(1)}$. According to Theorem 3, we have $k_{\mathbf{A}^{(K_1, 3)} \odot \mathbf{A}^{(2)}} = L$ when $k_{\mathbf{A}^{(K_1, 3)}} \geq L$, that is $K_1 \geq L$. Because of the Vandermonde structure, we have $k_{\mathbf{A}^{(L_1, 3)}} = \min(L_1, L)$. Therefore, the new factor matrices not only contain all the CSI but also easily satisfy the kruskal condition

$$k_{\mathbf{A}^{(K_1, 3)} \odot \mathbf{A}^{(2)}} + k_{\mathbf{A}^{(L_1, 3)}} + k_{\mathbf{A}^{(1)}} \geq 2L + 2. \quad (49)$$

D. Computational Complexity Analysis

In practice, channel estimation needs to be fast enough to track the time-varying channel. Thus, it is necessary to analyze the

complexity of the proposed channel estimation algorithm. Here, the main computationally intensive operations will be analyzed in terms of complex multiplications. For brevity, the intermediate variables K_1 and L_1 are considered to be of the same complexity level as $\mathcal{O}(T)$. In the stage of CP decomposition, the expansion operation in step 1 takes $\mathcal{O}(T^3 P^2 M_{BS})$ flops; the SVD and EVD in steps 2-3 have the complexity $\mathcal{O}(T^3 P M_{BS})$ and $\mathcal{O}(L^3)$, respectively; the reconstruction of $\mathbf{B}^{(2)}$ and $\mathbf{B}^{(1)}$ require $\mathcal{O}(P^2 T L^2)$ and $\mathcal{O}(M_{BS}^2 T L^2)$ flops. In the stage of parameter estimation, the one-dimensional search of AoAs and AoDs need $\mathcal{O}(M_{BS} N_{BS} G L)$ and $\mathcal{O}(P N_{MS} G L)$, respectively. Finally, the operation of pseudo-inverse to calculate path gains costs $\mathcal{O}(T^2 P^2 M_{BS}^2 L^2)$ flops.

To sum up, the total complexity of the proposed method is given by $\mathcal{O}(T^2 P^2 M_{BS}^2 L^2 + G L (M_{BS} N_{BS}) + T^3 P M_{BS} (P + M_{BS}) + T L^2 (P^2 + M_{BS}^2))$. In order to improve the estimation accuracy of angle, the angles grids always have $G > N_{BS} N_{MS}$. Meanwhile, the values of T , P , M_{BS} , and L are smaller than N_{BS} and N_{MS} . By comparison, the estimation schemes for time-varying channel in [18] and [24] have a complexity of $\mathcal{O}(L G^2)$ and $\mathcal{O}(L N_{BS} N_{MS} G)$. Hence, the proposed scheme is computationally more efficient.

V. SIMULATION RESULTS AND DISCUSSION

The performance of the proposed algorithm is evaluated via computer simulation with the following settings. We consider a system with one BS and $K=3$ MSs. The BS and MSs are equipped with ULA antennas with $N_{BS}=64$ and $N_{MS}=16$, respectively. The number of RF chain at the BS is set to $M_{BS}=10$ and only one at the MS. For each channel realization, the path gains follow a circularly symmetric Gaussian distribution $\alpha_k \sim \mathcal{CN}(0, \sigma_\alpha^2=1)$, and the AoAs/AoDs are uniformly distributed in $[-\frac{\pi}{2}, \frac{\pi}{2}]$. The carrier frequency is set to $f_c = 28$ GHz since the 28 GHz band is considered as one of the typical bands for highspeed wireless transmission [30]. The sampling period is $T_s=0.1$ μ s. The signal-to-noise ratio (SNR) is defined as $\text{SNR} = 10 \log_{10} \left(\frac{\sigma_s^2}{\sigma_n^2} \right)$, where σ_n^2 is the noise variance. The estimation accuracy of channel parameters is measured by mean square error (MSE) as

$$\text{MSE}(\mathbf{x}) = \frac{1}{N_{iter}} \sum_{n=1}^{N_{iter}} \|\mathbf{x} - \tilde{\mathbf{x}}\|_2^2. \quad (50)$$

The performance of channel matrices are measured by normalized mean square error (NMSE)

$$\text{NMSE}(\mathbf{H}) = \frac{1}{N_{iter}} \sum_{n=1}^{N_{iter}} \sum_{k=1}^K \frac{\|\mathbf{H}_k - \tilde{\mathbf{H}}_k\|_F^2}{\|\mathbf{H}_k\|_F^2} \quad (51)$$

where $\tilde{\mathbf{H}}_k$ represents the estimation channel of the k -th MS. In our simulations, all the results are averaged over $N_{iter}=5000$ Monte-Carlo trials.

Firstly, we examine the performance of proposed algorithm in a single-path channel (each MS can contribute only one

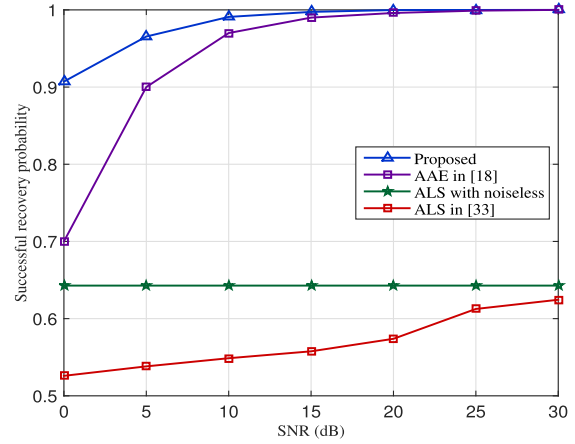


Fig. 3. Comparison among different algorithms in terms of the successful recovery probability of angles versus SNR.

distinguishable propagation path), denoted by $L_k = [1, 1, 1]$. The relative velocities of the three MSs are chosen as $[30, 90, 150]$ m/s with $f^d = [2800, 8400, 14000]$ Hz. When the carrier frequency is 28 GHz, a maximum speed of 150 m/s results in the coherence time of 71.4 μ s, which means that the coherence time contains 714 symbols. We set $P=8$, $T=12$, $P_s=200$. In Fig. 3, we examine the successful recovery probability of AoDs versus SNR when the estimated angles satisfy $|\hat{\varphi} - \varphi| \leq 3\%$. The factor matrices of ALS are randomly initialized and $J_{iter} = 200$, $\varepsilon = 10^{-4}$. From this figure, we can see that the proposed method achieves better performance than the iteration-based method. It is about 40% better than ALS algorithm at SNR = 30 dB. This is because the proposed method only uses linear algebraic operations which is more stable. However, the ALS algorithm cannot be guaranteed to converge to a global minimum or even difficult to converge in some special cases. This phenomenon is called swamp [35] which may require thousands of iterations to converge. Moreover, the success rate of the ALS scheme is less than 70% even in the noiseless case. Compared with the Adaptive Angle Estimation (AAE) [18], the proposed scheme has similar performance in the high SNR region (SNR ≥ 15 dB) and achieves an SNR gain of 5 dB in the low SNR region.

In Fig. 4, we show the performance curves of the channel parameters measured by MSE. It indicates that the proposed method achieves excellent performance when SNR ≥ 10 dB. In fact, a few failed estimations affect the overall performance when SNR ≤ 10 dB. In addition, we can see the AoAs estimation achieve better performance than AoDs. This is because that the BS with more antennas ($N_{BS}=64$) than MSs ($N_{MS}=16$) provides higher gain. The performance of AoD estimation can be closer to AoAs when we set $N_{MS}=32$. However, the number of antennas does not affect the performance of path gains. We do three hundred simulations to examine the performance of Doppler shifts for each SNR (0:0.3:30) and the results are shown in Fig. 5. We can see that with the increase of SNR, the distribution of estimated value is closer to the actual value. But even under the condition of SNR = 0 dB, the proposed method

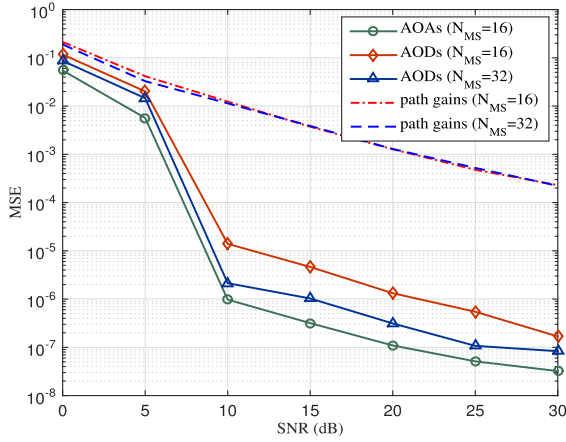


Fig. 4. MSE performance for proposed method in terms of the parameter estimation versus SNR.

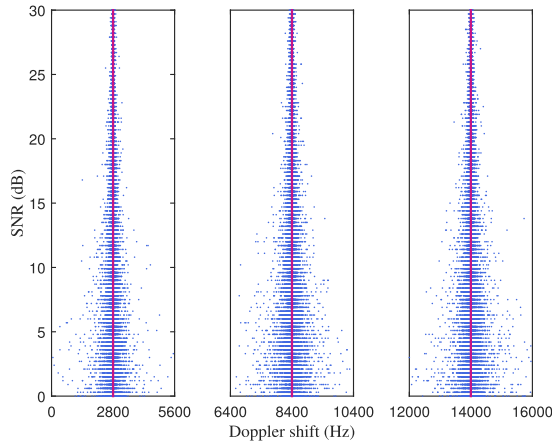


Fig. 5. The distribution of estimated Doppler shifts versus SNR.

still achieves good performance and almost all estimates have a deviation of less than 10%. However, as the Doppler shift increases, the variance of our estimation also increases.

In the designed pilot transmission scheme, the top P intervals of the first T slots are used to track the time-varying channel. To reduce overhead, the time-varying gains of last \bar{T} slots can be predicted by channel model and previous estimated results. We define time-varying gain as $\rho(t) = \alpha_l e^{j2\pi f_l^d T_s P_s t}$. In Fig. 6, we show the MSE performance of estimated time-varying gains ($0 \leq t < T$) and predicted time-varying gains ($T \leq t < T + \bar{T}$) when $\mathbf{L}_k = [2, 2, 2]$, $T = 12$, $\bar{T} = 18$. From the figure, we can see that the scheme achieves better performance in the training stage and reaches the optimum at the $\frac{T}{2}$ -th slots. This is because that the phase of $\rho(t)$ changes uniformly in $[0, T]$. The performance of the prediction stage decreases slowly over time due to the cumulative effect of errors. Moreover, we compare the performance under different P . We find that when P is greater than a certain value, it is difficult for us to improve the performance by increasing P , which can provide a reference for us to determine the number of pilots. In this scheme, the pilot overhead is about $\frac{PT}{(T+\bar{T})P_s}$. In Fig. 7, we further compare the NMSE performance among the conventional CS-OMP method [16], the improved CS-SD method [17], and the proposed channel

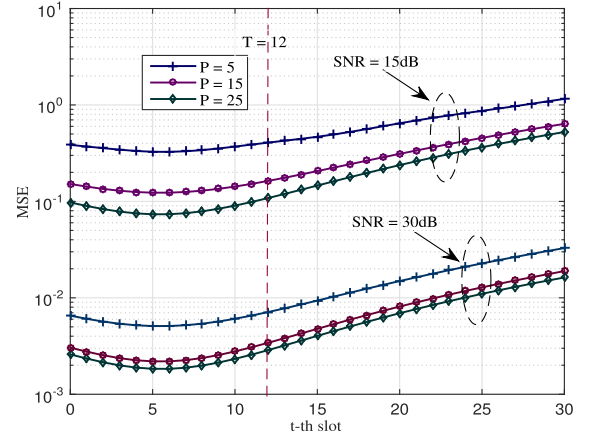


Fig. 6. MSE performance of estimated time-varying gains ($0 \leq t < T$) and predicted time-varying gains ($T \leq t < T + \bar{T}$).

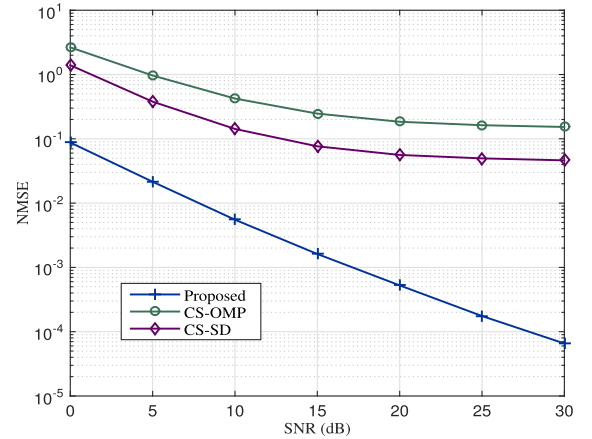


Fig. 7. NMSE performance comparison among CS-OMP, CS-SD and proposed channel estimation.

estimation scheme. To implement the CS-based methods, we use 6 bits to quantize the channel parameters. It is clear to see that the proposed method outperforms the CS-based schemes. Moreover, the proposed method yields accurate estimation even in the low SNR region. This is just because that we perform a truncated SVD in step 2 of Algorithm 2, which only keeps the main largest L singular values as signal component and the rests are discarded as noise.

We evaluate the NMSE performance when the channels have multiple paths. As shown in Fig. 8, the proposed method achieves the best performance when $\mathbf{L}_k = [1, 1, 1]$. In this case, the mmWave channels only have one line-of-sight (LOS) path which leads to focusing energy and reducing complexity. Nonetheless, the proposed method also obtains satisfied performance. For example, to achieve the $\text{NMSE} = 10^{-3}$, the SNRs required by the proposed method are 17 dB, 20 dB, 25 dB, and 30 dB for single, two, three and four paths of each MSs. In practice, real-world measurements in New York [40] confirm that the number of clusters is very small at 28 GHz bands (typically 1-5 in their measurement).

Next, we investigate the bit error rate (BER) performance when the estimated channel is used for beamforming. Under

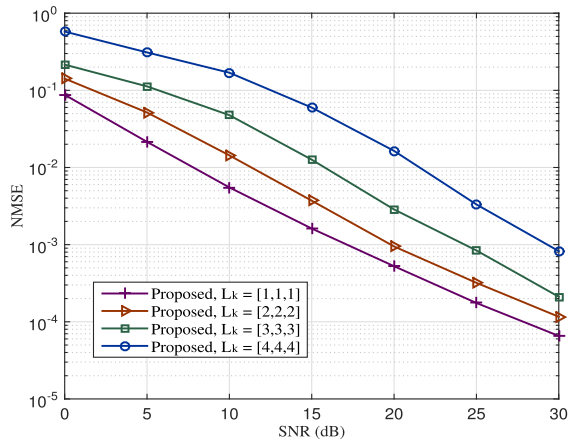


Fig. 8. NMSE performance comparison when each MS has single path, two paths, three paths and four paths of the channel.

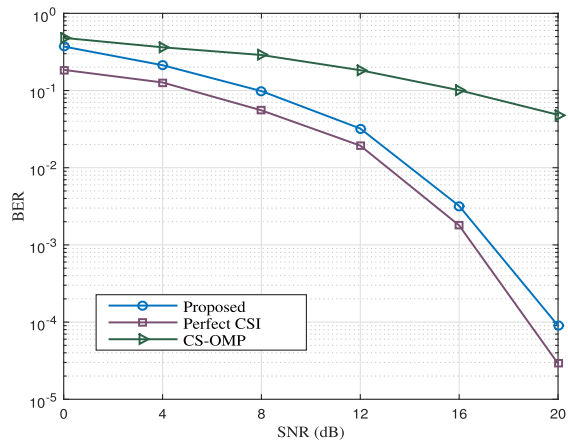


Fig. 9. BER performance of the proposed scheme and CS-OMP method.

the channel reciprocity of time division duplex (TDD) protocol, the downlink channel can be regarded as the transpose of the uplink channel. In our system, the downlink channel of the k -th MS is given by $\mathbf{H}_k^d = \tilde{\mathbf{H}}_k^T$. To fully achieve the spatial multiplexing gain, the downlink transmission strategy in [41] is adopted. After estimating the channel, the digital precoder and analog beamforming/combining vectors are designed by hybrid beamforming technique in [42]. The transmit symbols are randomly generated and modulated by 16-QAM. Fig. 9 depicts the BERs of different algorithms versus SNR. From the figure, it is clear that the proposed scheme outperforms the CS-OMP method and the curve of perfect CSI is given as the lower bound. In such a system, the accuracy of channel estimation, the interference between MSs, and the performance of hybrid beamforming technique can affect the BER performance.

Finally, in order to verify the essential unique condition of the proposed method in (47), we depict the NMSEs curves against the number of interval in a time slot and the number of time slots in a training frame in Figs. 10 and 11. We choose the two-path channel ($L=6$) and four-path channel ($L=12$) to verify our conclusions. From Fig. 10, we can see that the proposed method does not work for four-path channel when $P=3$. This is because that $P=3$ cannot meet the Kruscal's condition when $M_{BS} =$

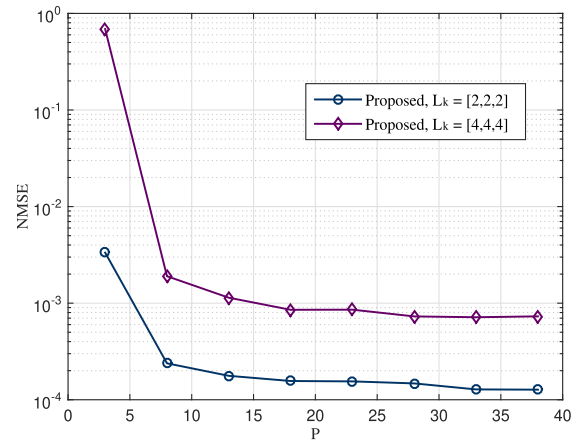


Fig. 10. NMSE performance against the number of interval for pilot transmission in one time slots when $T = 12$.

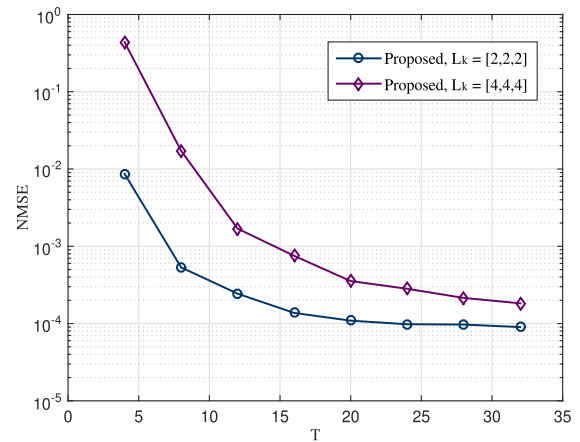


Fig. 11. NMSE performance against the number of time slot in one training frame when $P = 8$.

$10, T=8$, and $L=12$. The performance is not so good for two-path channel when $P=3$, but it is still a successful estimation. We observe that the NMSE performance turns to saturation when $P=8$. Then we set $P=8$ to investigate the performance against T . Similarly, the proposed method fails to work for four-path channel when $T=4$, $M_{BS}=10$, $P=8$, and $L=12$ since the Kruscal's condition is not satisfied. Despite the performance is not so good, it still achieves the $\text{NMSE}=10^{-2}$, and the NMSE performance turns to saturation when $T=12$. In consequence, we can obtain good performance when set $P=8, T=12$.

VI. CONCLUSION

In this paper, we have proposed a tensor based channel estimator for uplink MU-MIMO systems over time varying channels in mmWave communication. By exploiting the sparsity of mmWave channel, we have formulated the estimation problem as a sparse signal recovery problem. Specifically, we have designed a novel pilot transmission scheme to track the time-varying channel. The received signal at the BS has been concatenated as a third-order tensor which fits a low-rank tensor model, then we have used the subspace-based algorithm to calculate the factor matrices. Finally, we have estimated the angles,

path gains and Doppler shifts from the factor matrices. Simulation results have verified that the proposed channel estimator is more stable than ALS-based method and outperform existing CS-based methods such as CS-OMP and CS-SD. The designed pilot transmission scheme can achieve good performance and low overhead. Furthermore, the relationship between pilot overhead and the number of paths have been derived and verified by simulations, which demonstrated the proposed method can accurately estimate the time-varying mmWave channel with low pilot overhead. For the future work, the proposed scheme can be evaluated with the specific standardized channel models for practicality.

REFERENCES

- [1] C.-X. Wang, J. Bian, J. Sun, W. Zhang, and M. Zhang, "A survey of 5G channel measurements and models," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 4, pp. 3142–3168, Oct.–Dec. 2018.
- [2] X. You, C.-X. Wang, and J. Huang, "Towards 6G wireless communication networks: Vision, enabling technologies, and new paradigm shifts," *Sci. China Inf. Sci.*, vol. 64, 2021, Art. no. 110301.
- [3] R. W. Heath, N. Gonzalez-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 436–453, Apr. 2016.
- [4] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter-wave cellular wireless networks: Potentials and challenges," *Proc. IEEE*, vol. 102, no. 3, pp. 366–385, Mar. 2014.
- [5] J. Huang, C.-X. Wang, H. Chang, J. Sun, and X. Gao, "Multi-frequency multi-scenario millimeter wave MIMO channel measurements and modeling for B5G wireless communication systems," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 9, pp. 2010–2025, Sep. 2020.
- [6] J. Bian, C.-X. Wang, X. Gao, X. You, and M. Zhang, "A general 3D non-stationary wireless channel model for 5G and beyond," *IEEE Trans. Wireless Commun.*, vol. 20, no. 5, pp. 3211–3224, May 2021.
- [7] X. Wang et al., "Millimeter wave communication: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 3, pp. 1616–1653, Jul.–Sep. 2018.
- [8] D. Zhang, Y. Wang, X. Li, and W. Xiang, "Hybridly connected structure for hybrid beamforming in mmWave massive MIMO systems," *IEEE Trans. Commun.*, vol. 66, no. 2, pp. 662–674, Feb. 2018.
- [9] I. Ahmed et al., "A survey on hybrid beamforming techniques in 5G: Architecture and system model perspectives," *IEEE Commun. Surv. Tut.*, vol. 20, no. 4, pp. 3060–3097, Oct.–Dec. 2018.
- [10] A. Alkhateeb, J. Mo, N. Gonzalez-Prelcic, and R. W. Heath, "MIMO precoding and combining solutions for millimeter-wave systems," *IEEE Commun. Mag.*, vol. 52, no. 12, pp. 122–131, Dec. 2014.
- [11] I. A. Hemadeh, K. Satyanarayana, M. El-Hajjar, and L. Hanzo, "Millimeter-wave communications: Physical channel models, design considerations, antenna constructions, and link-budget," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 2, pp. 870–913, Apr.–Jun. 2018.
- [12] P. Dong, H. Zhang, G. Y. Li, I. S. Gaspar, and N. NaderiAlizadeh, "Deep CNN-Based channel estimation for mmWave massive MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 5, pp. 989–1000, Sep. 2019.
- [13] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 831–846, Oct. 2014.
- [14] 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.
- [15] M. Ke, Z. Gao, Y. Wu, X. Gao, and R. Schober, "Compressive sensing-based adaptive active user detection and channel estimation: Massive access meets massive MIMO," *IEEE Trans. Signal Process.*, vol. 68, pp. 764–779, 2020.
- [16] R. Méndez-Rial, C. Rusu, A. Alkhateeb, N. González-Prelcic, and R. W. Heath, "Channel estimation and hybrid combining for mmWave: Phase shifters or switches?," in *Proc. Inf. Theory Appl. Workshop*, 2015, pp. 90–97.
- [17] X. Gao, L. Dai, S. Han, C. I, and X. Wang, "Reliable beamspace channel estimation for millimeter-wave massive MIMO systems with lens antenna array," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 6010–6021, Sep. 2017.
- [18] Q. Qin, L. Gui, P. Cheng, and B. Gong, "Time-varying channel estimation for millimeter wave multiuser MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 67, no. 10, pp. 9435–9448, Oct. 2018.
- [19] K. Venugopal, A. Alkhateeb, N. G. Prelcic, and R. W. Heath, "Channel estimation for hybrid architecture-based wideband millimeter wave systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 1996–2009, Sep. 2017.
- [20] C. Qian, X. Fu, and N. D. Sidiropoulos, "Algebraic channel estimation algorithms for FDD massive MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 5, pp. 961–973, Sep. 2019.
- [21] Z. Zhou, J. Fang, L. Yang, H. Li, Z. Chen, and S. Li, "Channel estimation for millimeter-wave multiuser MIMO systems via PARAFAC decomposition," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7501–7516, Nov. 2016.
- [22] 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," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1524–1538, Jul. 2017.
- [23] Z. Lin, T. Lv, W. Ni, J. A. Zhang, and R. P. Liu, "Tensor-based multi-dimensional wideband channel estimation for mmWave hybrid cylindrical arrays," *IEEE Trans. Commun.*, vol. 68, no. 12, pp. 7608–7622, Dec. 2020.
- [24] L. Cheng, G. Yue, D. Yu, Y. Liang, and S. Li, "Millimeter wave time-varying channel estimation via exploiting block-sparse and low-rank structures," *IEEE Access*, vol. 7, pp. 123355–123366, 2019.
- [25] N. D. Sidiropoulos and A. Kyrillidis, "Multi-way compressed sensing for sparse low-rank tensors," *IEEE Signal Process. Lett.*, vol. 19, no. 11, pp. 757–760, Nov. 2012.
- [26] W. Ma, C. Qi, and G. Y. Li, "High-resolution channel estimation for frequency-selective mmwave massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 19, no. 5, pp. 3517–3529, May 2020.
- [27] M. Rani, S. B. Dhok, and R. B. Deshmukh, "A systematic review of compressive sensing: Concepts, implementations and applications," *IEEE Access*, vol. 6, pp. 4875–4894, 2018.
- [28] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Rev.*, vol. 43, no. 1, pp. 129–159, 2001.
- [29] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.
- [30] H. Song, X. Fang, and Y. Fang, "Millimeter-wave network architectures for future high-speed railway communications: Challenges and solutions," *IEEE Wireless Commun.*, vol. 23, no. 6, pp. 114–122, Dec. 2016.
- [31] H. Chang et al., "A novel non-stationary 6G UAV-to-ground wireless channel model with 3D arbitrary trajectory changes," *IEEE Internet Things J.*, vol. 8, no. 12, pp. 9865–9877, Jun. 2021.
- [32] Y. Liu, C.-X. Wang, J. Huang, J. Sun, and W. Zhang, "Novel 3-D non-stationary mmWave massive MIMO channel models for 5G high-speed train wireless communications," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 2077–2086, Mar. 2019.
- [33] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, 2009.
- [34] N. D. Sidiropoulos, L. De Lathauwer, X. Fu, K. Huang, E. E. Papalexakis, and C. Faloutsos, "Tensor decomposition for signal processing and machine learning," *IEEE Trans. Signal Process.*, vol. 65, no. 13, pp. 3551–3582, Jul. 2017.
- [35] N. Li, "Variants of ALS on tensor decompositions and applications," *Diss. Thesis Gradworks*, vol. 51, no. 3, pp. 455–500, 2013.
- [36] M. Sørensen and L. De Lathauwer, "Blind signal separation via tensor decomposition with vandermonde factor: Canonical polyadic decomposition," *IEEE Trans. Signal Process.*, vol. 61, no. 22, pp. 5507–5519, Nov. 2013.
- [37] J. B. Kruskal, "Three-way arrays: Rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics," *Linear Algebra Appl.*, vol. 18, no. 2, pp. 95–138, 1977.
- [38] N. Sidiropoulos and X. Liu, "Identifiability results for blind beamforming in incoherent multipath with small delay spread," *IEEE Trans. Signal Process.*, vol. 49, no. 1, pp. 228–236, Jan. 2001.
- [39] A. Stegeman and N. D. Sidiropoulos, "On kruskal's uniqueness condition for the candcomp/parafac decomposition," *Linear Algebra Appl.*, vol. 420, no. 2/3, pp. 540–552, 2006.

- [40] M. R. Akdeniz et al., "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1164–1179, Jun. 2014.
- [41] X. Gao, L. Dai, Z. Gao, T. Xie, and Z. Wang, "Precoding for mmWave massive MIMO," in *mmWave Massive MIMO*, S. Mumtaz, J. Rodriguez, and L. Dai, Eds. Cambridge, MA, USA: Academic Press, 2017, ch. 5, pp. 79–111. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/B9780128044186000054>
- [42] O. E. 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.



Jianhao Wang (Student Member, IEEE) received the B.Sc. degree in electronic information engineering from Jiangsu University, Zhenjiang, China, in 2019, and the M.Sc. degree in information and communication engineering from Shandong University, Jinan, China, in 2022. He is currently with Huawei Technologies Company, Ltd., Nanjing, China. His research interests include tensor theory, B5G wireless communications, millimeter wave, and massive MIMO.



Wensheng Zhang (Member, IEEE) received the M.E. degree in electrical engineering from Shandong University, Jinan, China, in 2005, and the Ph.D. degree in electrical engineering from Keio University, Tokyo, Japan, in 2011. In 2011, he joined the School of Information Science and Engineering, Shandong University, where he is currently an Associate Professor. He was a Visiting Scholar with the University of Oulu, Oulu, Finland, in 2010, and with the University of Arkansas, Fayetteville, AR, USA, in 2019. His research interests include tensor computing, random

matrix theory, and intelligent B5G wireless communications.



Yunfei Chen (Senior Member, IEEE) received the B.E. and M.E. degrees in electronics engineering from Shanghai Jiaotong University, Shanghai, China, in 1998 and 2001, respectively, and the Ph.D. degree from the University of Alberta, Edmonton, AB, Canada, in 2006. He is currently an Associate Professor with the University of Warwick, Coventry, U.K. His research interests include wireless communications, cognitive radios, wireless relaying, and energy harvesting.



Zhi Liu (Member, IEEE) received the Ph.D. degree from the Institute of Image Processing and Pattern Recognition, Shanghai Jiao Tong University, Shanghai, China, in 2008. He is currently a Professor with the School of Information Science and Engineering, Shandong University, Jinan, China. He is the Head of Intelligent Information Processing Group. His current research interests include the applications of computational intelligence to linked multicomponent Big Data systems, medical images in the neurosciences, multimodal human computer interaction,

remote sensing image processing, content-based image retrieval, semantic modeling, data processing, classification, and data mining.



Jian Sun (Member, IEEE) received the B.Sc. degree in applied electronic technology, the M.Eng. degree in measuring and testing technologies and instruments, and the Ph.D. degree in communication and information systems from Zhejiang University, Hangzhou, China, in 1996, 1999, and 2005, respectively. From 2005 to 2018, he was a Lecturer with the School of Information Science and Engineering, Shandong University, Jinan, China. Since 2018, he has been an Associate Professor. In 2008, he was a Visiting Scholar with the University of California San Diego, San Diego, CA, USA. In 2011, he was a Visiting Scholar with Heriot-Watt University, U.K., supported by U.K.–China Science Bridges: Research and Development on (B)4G Wireless Mobile Communications Project. His current research interests include signal processing for wireless communications, channel sounding and modeling, joint communications and sensing, and software defined radio.



Cheng-Xiang Wang (Fellow, IEEE) received the B.Sc. and M.Eng. degrees in communication and information systems from Shandong University, Jinan, China, in 1997 and 2000, respectively, and the Ph.D. degree in wireless communications from Aalborg University, Aalborg, Denmark, in 2004. He was a Research Assistant with the Hamburg University of Technology, Hamburg, Germany, from 2000 to 2001, a Visiting Researcher with Siemens AG Mobile Phones, Munich, Germany, in 2004, and a Research Fellow with the University of Agder, Grimstad, Norway, from 2001 to 2005. Since 2005, he has been with Heriot-Watt University, Edinburgh, U.K., where he was promoted to a Professor in 2011. In 2018, he joined Southeast University, Nanjing, China, as a Professor. He is also a part-time Professor with Purple Mountain Laboratories, Nanjing, China. He has authored four books, three book chapters, and more than 470 papers in refereed journals and conference proceedings, including 26 highly cited papers. He has delivered 25 invited keynote speeches/talks and 14 tutorials in international conferences. His current research interests include wireless channel measurements and modeling, 6G wireless communication networks, and electromagnetic information theory. He is a Member of Academia Europaea, The Academy of Europe, a Member of the European Academy of Sciences and Arts, a Fellow of the Royal Society of Edinburgh, IET, and China Institute of Communications.

In 2019 and 2020, he was an IEEE Communications Society Distinguished Lecturer, a Highly-Cited Researcher recognized by Clarivate Analytics during 2017–2020, and one of the most cited Chinese Researchers recognized by Elsevier in 2021. He is currently an Executive Editorial Committee Member of IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. He was the Editor of more than ten international journals, including the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, from 2007 to 2009, IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, from 2011 to 2017, and IEEE TRANSACTIONS ON COMMUNICATIONS, from 2015 to 2017. He was the Guest Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, Special Issue on Vehicular Communications and Networks Lead Guest Editor, Special Issue on Spectrum and Energy Efficient Design of Wireless Communication Networks, and Special Issue on Airborne Communication Networks. He was also the Guest Editor of IEEE TRANSACTIONS ON BIG DATA, Special Issue on Wireless Big Data, and the Guest Editor of IEEE TRANSACTIONS ON COGNITIVE COMMUNICATIONS AND NETWORKING, Special Issue on Intelligent Resource Management for 5G and Beyond. He was a TPC Member, the TPC Chair, and General Chair of more than 80 international conferences. He was the recipient of 14 Best Paper Awards from IEEE GLOBECOM 2010, IEEE ICCT 2011, ITST 2012, IEEE VTC 2013Spring, IWCMC 2015, IWCMC 2016, IEEE/CIC ICC 2016, WPMC 2016, WOCC 2019, IWCMC 2020, WCSP 2020, CSPS2021, WCSP 2021, and also the 2020–2022 AI 2000 Most Influential Scholar Award Honourable Mention in recognition of his outstanding and vibrant contributions in the field of Internet of Things.