IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received September 11, 2019, accepted September 30, 2019, date of publication October 10, 2019,
date of current version October 24, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2946881

Cluster-Based Fast Time-Varying MIMO Channel Fading Prediction in the High-Speed Scenario

YUXIANG ZHANG, (Student Member, IEEE), JIANHUA ZHANG, (Senior Member, IEEE), AND LI YU, (Student Member, IEEE)

State Key Lab of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China

Corresponding author: Jianhua Zhang (jhzhang@bupt.edu.cn)

This work was supported in part by the Beijing Municipal Science and Technology Commission Project under Grant Z181100003218007, in part by the National Key Research and Development Program of China under Grant 2018YFF0301201, in part by the Key Project of State Key Lab of Networking and Switching Technology under Grant NST20180105, and in part by the China Scholarship Council (CSC) under Grant 201806470015.

ABSTRACT Accurate channel state information (CSI) is important for the coherent detection of multiple-input-multiple-output (MIMO) system. Especially in a high-speed scenario, fast time-varying CSI gotten by the conventional channel estimation schemes tend to be out of date, thus tracking and predicting CSI are more attractive and indispensable. Motivated by those, a time-varying MIMO channel fading prediction framework is proposed in this paper. The principle behind our scheme is that the cluster-based fading channel has the spatial consistency property, which means the small-scale parameters of channel clusters evolve continuously and smoothly in the time and spatial domains. Thus CSI can be tracked and predicted within several or tens of wavelengths. The proposed scheme is composed of an extended Bayesian Estimation Kalman Filter to track the time-varying CSI evolution, and a Cluster Drifting Based Prediction to obtain the small-scale parameters of channel clusters. The performance of the proposed scheme is simulatively verified by a standard clustered-based channel model.

INDEX TERMS MIMO channel fading prediction, high-speed scenario, spatial consistency, cluster drifting.

I. INTRODUCTION

Accurate channel state information (CSI) is important for coherent detection, thus achieving the high performance and providing quality of service support in the fifth Generation (5G) and beyond 5G (B5G) multiple-input-multiple-output (MIMO) systems [1]–[4]. In the TDD system, it is feasible to obtain CSI by channel reciprocity through uplink channel estimation. However, the CSI may become outdated due to the estimation, processing, and propagation delay [5] which will degrade the system performance seriously [5], [6]. Especially in 5G and B5G systems, high-speed scenarios become important, like users still want to get the high data rate in high-speed vehicles and trains [7]–[10]. In these cases, the channel fading between the base station (BS) and high-speed mobile user (MS) fluctuates dramatically within a short time period [6]. As a key technique to improve the accuracy of CSI, the channel fading prediction becomes an important research issue in a high-speed scenario.

The associate editor coordinating the review of this manuscript and approving it for publication was Meng-Lin Ku.

The state-of-the-art channel fading prediction schemes have been reviewed in [5], [11]. One category of widely adopted schemes are non-parametric schemes, including the autoregressive modeling approach [12]–[14] and basis expansion approach [15], [16]. Pre-known autocorrelation functions or basis functions are usually required for these two approaches. However, they are more suitable to be adopted in a single-input-single-output system without the consideration of spatial characteristics. Another widely-used scheme is parametric channel model (PCM) approach, which is based on the sum-of-sinusoids principle [17] and spatial channel models [18], [19]. With spatial characteristic fully considered in these models, the PCM approach is quite appropriate for the time-varying MIMO channel fading prediction [5].

In the conventional PCM approach [5], [20], a MS is assumed with a relatively low velocity, and hence the time-varying CSI can be represented by several time-invariant small-scale parameters. With the account of this parametric representation, it is feasible to predict the time-varying MIMO channel fading. However, when the MS is moving with a high speed, all the small-scale parameters, including

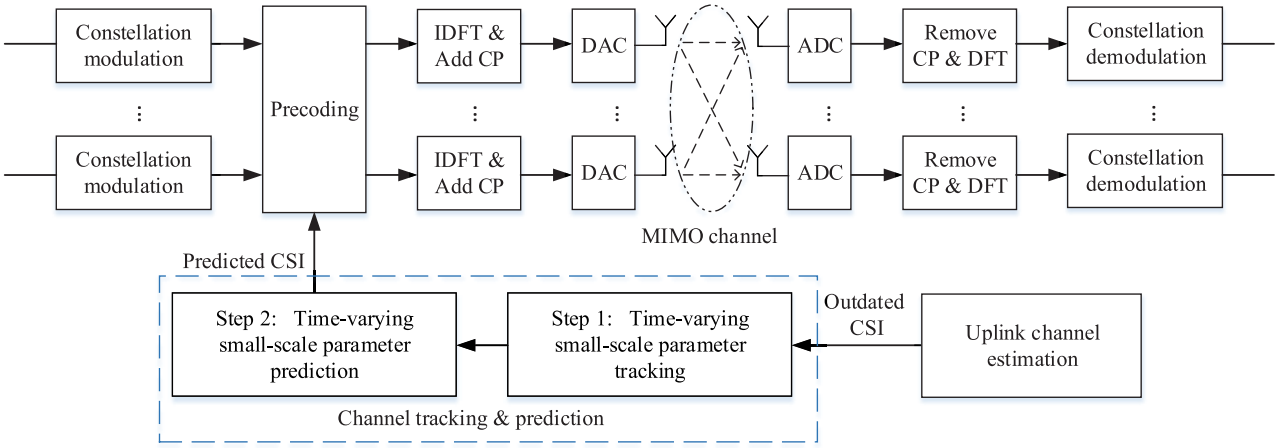


FIGURE 1. System model for a MIMO-OFDM downlink channel in the TDD mode with channel tracking and prediction.

Doppler frequency, angles of arrival (AoA) and department (AoD), and time delay are time-varying within a short time. Therefore, it becomes a challenging issue to combine the MIMO channel fading prediction with time-varying small-scale parameters. In [5], [21], the authors propose schemes with the parameter changing rates preset or pre-known by the algorithms. In [22], [23], the authors propose a tracking algorithm based on Kalman Filter to track the time-varying small-scale parameter variations. More importantly in [23], Bayesian estimation is embedded in the tracking process, which is useful to estimate the model order number. Recently, some research work on channel prediction for massive MIMO and millimeter wave has been published in [24], [25]. However, they mainly focus on the narrow-band system and time-invariant small-scale parameters, respectively.

To the best of our knowledge, the previous studies on channel fading prediction mainly focus on the static or quasi-static scenario, in which the small-scale parameters are assumed to be time-invariant. However, more attention should be paid to the case of time-varying small-scale parameters in the high-speed scenario. Since the small-scale parameters are estimated independently in different time instants (snapshots), it is intractable to combine spatial consistency property (SCP) with the conventional PCM channel fading prediction scheme. Also, without the prior information of time-varying small-scale parameters evolution, the predicted CSI gotten by conventional approaches is not accurate.

Thanks to the Geometry-based Stochastic Channel Models (GBSMs) [18] and the cluster based channel SCP [26], [27], the small-scale parameters of channel clusters evolve continuously and smoothly in the time and spatial domains. Also, the cluster visibility region (VR) [19] is used to describe the birth and death length of each cluster [28]–[30]. In one VR, the corresponding clusters are active, the small-scale parameters of which evolve continuously and smoothly. Therefore, it is potential to utilize the SCP to track and predict the variation trend of time-varying small-scale parameters within one VR.

In this paper, we present a time-varying MIMO channel fading prediction framework in a high-speed scenario. For better illustration, the proposed prediction framework in a MIMO-OFDM system with a TDD mode is shown in Fig. 1 (Page 3). In the tracking stage, the Bayesian Estimation Kalman Filter (BEKF) algorithm in [23] is adopted and extended to track the time-varying small-scale parameters with the consideration of the cluster-based channel SCP. As an important indicator in the process of Kalman tracking, the Cramér–Rao lower bounds (CRLBs) of the tracking parameters are derived more accurately. In the prediction stage, it is assumed that the channel fading prediction horizon is smaller than one VR, and hence the SCP is guaranteed within this horizon. Also, with this assumption, a Cluster Drifting Based Prediction (CDBP) algorithm is proposed to predict the MIMO channel fading with the time-varying small-scale parameters.

The rest of the paper is organized as follows: in Section II, the signal model is given with a parametric representation of MIMO channel fading by GBSMs. The extended BEKF and the proposed CDBP algorithms are introduced in Sections III and IV, respectively. Simulation results are presented in Section V. Finally, Section VI concludes the paper.

For description convenience, we give explanations on some notations used in the paper. $\otimes$ and $\odot$ denote Kronecker and Khatri–Rao Products, respectively. $(\cdot)^T$ and $(\cdot)^H$ stand for transpose and conjugate transpose, respectively. $\mathbf{I}_n$ denotes identity matrix with order n. The matrix and column vector are denoted by the bold uppercase and lowercase letters, respectively. $\mathbb{E}_x[\cdot]$ denotes the expectation of the parameter x.

II. SIGNAL MODEL

Since GBSM [18], [19] has been widely used as the standardized channel model [31]–[33] in the fourth Generation (4G) and 5G communication systems, it can reflect the realistic time-varying MIMO channel fading characteristics rather than make a theoretical assumption. Therefore, in this paper, GBSM is adopted to make a parametric representation for the

realistic CSI, which is similar with that in [34], [35]. Assume that the receiver (Rx) and transmitter (Tx)[1] are equipped with antenna arrays of U and S elements, which are located at $\mathbf{d}_{rx,1},\ldots,\mathbf{d}_{rx,U}\in\mathbb{R}^3$ and $\mathbf{d}_{tx,1},\ldots,\mathbf{d}_{tx,S}\in\mathbb{R}^3$ with respect to an arbitrary reference point. We stack the received signals of all antenna elements at Rx and Tx in one vector and the equivalent baseband MIMO channel model is expressed as (1)

$$\mathbf{h}(t,\tau)=\sum_{l=1}^{L}w_l(t)\cdot\mathbf{c}_{tx}\left(\boldsymbol{\psi}_{tx,l}(t)\right)\otimes\mathbf{c}_{rx}\left(\boldsymbol{\psi}_{rx,l}(t)\right)\cdot\delta\left(\tau-\tau_l(t)\right),\quad(1)$$

where $\mathbf{h}(t,\tau)$ is the stacked vector-representation of time-varying channel impulse response (CIR)[2], with t and l denoting the indexes of the snapshot and physical path[3]. The complex magnitude $w_l(t)=\alpha_l(t)\cdot e^{j\varphi_l(t)}$ with $\alpha_l(t)$ denoting the magnitude caused by large-scale fading and $\varphi_l(t)$ denoting the phase change. Note that both initial phase and the impact of time-varying Doppler frequency are included in $\varphi_l(t)$. $\mathbf{c}_{rx}\left(\boldsymbol{\psi}_{rx,l}(t)\right)$ and $\mathbf{c}_{tx}\left(\boldsymbol{\psi}_{tx,l}(t)\right)$ are the antenna steering vector of Rx and Tx, with $\boldsymbol{\psi}_{rx,l}(t)$ and $\boldsymbol{\psi}_{tx,l}(t)$ denoting AoA and AoD. $\boldsymbol{\psi}_{pos,l}(t)=\left[\phi_{pos,l}(t),\theta_{pos,l}(t)\right]$ with pos representing rx or tx and $\phi_{rx,l}(t)$, $\theta_{rx,l}(t)$, $\phi_{tx,l}(t)$, $\theta_{tx,l}(t)$ denote the azimuth angle of arrival (AAoA), elevation angle of arrival (EAoA), azimuth angle of departure (AAoD), elevation angle of departure (EAoD), respectively. $\delta(\tau-\tau_l(t))$ denotes the Dirac function of time delay $\tau_l(t)$.

Without regard to the coupling between different antenna elements, the antenna steering vector can be further expressed as

$$\mathbf{c}_{pos}\left(\psi_{pos,l}(t)\right)=\left[c_{pos,1}\left(\psi_{pos,l}(t)\right),\ldots,c_{pos,p}\left(\psi_{pos,l}(t)\right),\ldots,c_{pos,P}\left(\psi_{pos,l}(t)\right)\right]^T,\quad(2)$$

where p represents the index of Rx or Tx antenna element. With the assumption of the plane wave propagation, (2) is rewritten as

$$\mathbf{c}_{pos}\left(\boldsymbol{\psi}_{pos,l}(t)\right)=\mathbf{a}\left(\psi_{pos,l}(t)\right)\odot\exp\left(j2\pi\lambda^{-1}\mathbf{D}_{pos}^T\cdot\mathbf{e}\left(\boldsymbol{\psi}_{pos,l}(t)\right)\right),\quad(3)$$

where λ, $\mathbf{e}\left(\boldsymbol{\psi}_{pos,l}(t)\right)$ are the wavelength and the unit wave vector in the direction of $\boldsymbol{\psi}_{pos,l}(t)$. $\mathbf{D}_{pos}=\left[\mathbf{d}_{pos,1},\ldots,\mathbf{d}_{pos,P}\right]$ with $\mathbf{d}_{pos,p}$ denoting the location vector of each antenna element. $\mathbf{a}\left(\boldsymbol{\psi}_{pos,l}(t)\right)$ is the complex antenna gain vector, in which the responses of elements are all set to 1 in this paper.

[1]Since the downlink channel is considered in this paper, the transmitter and receiver denote BS and MS respectively.

[2]CIR is a mathematical expression for MIMO channel fading.

[3]CIR is composed of several clusters and each cluster consists of a group of physical paths with similar propagation characteristics [18], [31]–[33]. However, limited by the system resolution (i.e., antenna aperture and number [36]) and algorithm precision, each cluster can be seen as several independent paths in the process of channel fading tracking and prediction.

The discrete Fourier transform (DFT) of the CIR in (1) is given as

$$\mathbf{g}(t,f_m)=\sum_{l=1}^{L}w_l(t)\cdot\mathbf{c}_{tx}\left(\boldsymbol{\psi}_{tx,l}(t)\right)\otimes\mathbf{c}_{rx}\left(\boldsymbol{\psi}_{rx,l}(t)\right)\cdot e^{-j2\pi f_m\tau_l(t)},\quad(4)$$

where f_m is the mth frequency bin (subcarrier frequency in the OFDM system) with $m=0,1,\ldots,M-1$.

We stack $\mathbf{g}(t,f_m)$ with $m=0,1,\ldots,M-1$ in one vector and get $\mathbf{g}(t,f)=\left[\mathbf{g}^T(t,f_0),\ldots,\mathbf{g}^T(t,f_M)\right]^T$. With Kronecker product, we have

$$\mathbf{g}(t,f)=\sum_{l=1}^{L}w_l(t)\cdot\mathbf{s}_{\tau_l,\boldsymbol{\psi}_{rx,l},\boldsymbol{\psi}_{tx,l}}(t),\quad(5)$$

$$\mathbf{s}_{\tau_l,\boldsymbol{\psi}_{rx,l},\boldsymbol{\psi}_{tx,l}}(t)=\mathbf{u}(\tau_l(t))\otimes\mathbf{c}_{tx}\left(\boldsymbol{\psi}_{tx,l}(t)\right)\otimes\mathbf{c}_{rx}\left(\boldsymbol{\psi}_{rx,l}(t)\right),\quad(6)$$

where the time-varying small-scale parameter set is given by $\left[\alpha_l(t),\varphi_l(t),\tau_l(t),\boldsymbol{\psi}_{rx,l}(t),\boldsymbol{\psi}_{tx,l}(t)\right]^T$, and the frequency steering vector is $\mathbf{u}(\tau_l(t))=[e^{-j2\pi f_0\tau_l(t)},\ldots,e^{-j2\pi f_{M-1}\tau_l(t)}]^T$.

Considering the additive noise in the signal model, we express the received signal as

$$\mathbf{y}(t,f)=\sum_{l=1}^{L}w_l(t)\cdot\mathbf{s}_{\tau_l,\boldsymbol{\psi}_{rx,l},\boldsymbol{\psi}_{tx,l}}(t)+\boldsymbol{\varepsilon}(t,f),\quad(7)$$

where $\boldsymbol{\varepsilon}(t,f)$ is a complex circular symmetric white noise with the variance σ^2. It is worth noting that $\mathbf{y}(t,f)$ can be regarded as the uplink estimated CSI with noise. For presentation simplicity, the antenna polarization is not considered in the signal model. Besides, we only consider the azimuth plane and thus, the angular parameter $\boldsymbol{\psi}_{pos,l}(t)$ is simplified as $\phi_{pos,l}(t)$. The case of antenna polarization and elevation angles can be extended.

III. CHANNEL FADING TRACKING IN MOBILE PROPAGATION ENVIRONMENT

As presented in the part of introduction, the time-varying small-scale parameter are tracked based on cluster-based channel SCP. In this section, we show details of the time-varying small-scale parameter tracking algorithm in Step 1 of Fig. 1. Let the observable parameter vector $\boldsymbol{\xi}(t)=[\boldsymbol{\alpha}^T(t),\boldsymbol{\varphi}^T(t),\boldsymbol{\tau}^T(t),\boldsymbol{\phi}_{rx}^T(t),\boldsymbol{\phi}_{tx}^T(t),\sigma^2]^T$ and the non-observable parameter vector $\Delta\boldsymbol{\xi}(t)=[\Delta\boldsymbol{\alpha}^T(t),\Delta\boldsymbol{\varphi}^T(t),\Delta\boldsymbol{\tau}^T(t),\Delta\boldsymbol{\phi}_{rx}^T(t),\Delta\boldsymbol{\phi}_{tx}^T(t)]^T$, where $\boldsymbol{\alpha}(t)$, $\boldsymbol{\varphi}(t)$, $\boldsymbol{\tau}(t)$, $\boldsymbol{\phi}_{rx}(t)$ and $\boldsymbol{\phi}_{tx}(t)$ are the L-dimensional parameter vectors of all physical paths. $\Delta\boldsymbol{\alpha}(t)$, $\Delta\boldsymbol{\varphi}(t)$, $\Delta\boldsymbol{\tau}(t)$, $\Delta\boldsymbol{\phi}_{rx}(t)$ and $\Delta\boldsymbol{\phi}_{tx}(t)$ are the corresponding L-dimensional parameter change vectors. In the following, the state-space model adopted in the tracking method is introduced, which is proved to be robust to track the parameters in [22], [23], [37].

A. STATE-SPACE MODEL

Limited by the frequency bandwidth and antenna aperture, the small-scale parameters such as time delay, AoA

and AoD evolve very slowly within a specific distance. As described in Section IV-B1, they are highly-correlated in several continuous snapshots due to the specific scatterers in the propagation environment. However, as shown in (1), the time-varying channel fading is superimposed by several independent physical paths with different small-scale parameters and it varies dramatically in a very short distance (i.e., several wavelengths). As a useful tool, state-space model is utilized to build the association of the parameters in adjacent time snapshots. As introduced in [37], it can be described as a Guassian-Markov model. The tracking parameter vector is given as $\boldsymbol{\gamma}(t) = \left[\boldsymbol{\xi}^T(t), \Delta\boldsymbol{\xi}^T(t)\right]^T$. And then the time-evolution equation of the state-space model is expressed as

$$\boldsymbol{\gamma}(t) = \mathbf{F} \cdot \boldsymbol{\gamma}(t-1) + \mathbf{w}_{tr}(t), \tag{8}$$

where $\mathbf{w}_{tr}(t)$ is the white normal distributed noise and $\mathbf{F}$ is the state transition matrix preset according to the state-space model. Limited by the space, only the expression of $\mathbf{F}$ is shown in Appendix A. The detailed illustration and principle are given in [37].

B. BAYESIAN ESTIMATION AND KALMAN FILTER TRACKING

To estimate and track the time-varying small-scale parameters, the BEKF algorithm in [23] is selected in the tracking stage. Compared with the traditional snapshot-based parameter estimation algorithms in [5], [20], [34], [38], the parameter correlation property between adjacent snapshots is considered in the parameter estimation process of BEKF. Therefore, the BEKF makes it possible to build an evolution relationship of the parameters between adjacent snapshots, that is to track the evolution of parameters with time. Also within the BEKF, the Variational Bayesian Space-Alternating Generalized Expectation-Maximization (VB-SAGE) algorithm [38] is embedded to estimate the small-scale parameters in each snapshot. Thanks to the VB-SAGE, the channel fading can be decomposed and represented by several small-scale parameters given in Section II. Also with VB-SAGE, it is feasible to obtain the model order, (i.e., the path number), which is useful to remove the false path and add the new path in the tracking stage. In the following, we first introduce the VB-SAGE algorithm and then illustrate the principle of BEKF algorithm in detail.

1) VB-SAGE ALGORITHM

VB-SAGE algorithm is the extension of SAGE algorithm [34] which can estimate the probability distribution of parameters within the Bayesian framework. Different from SAGE algorithm, we can make a statistical hypothesis with the estimated probability distribution by the VB-SAGE algorithm. The decision theory is utilized to judge the path number. Usually, in the initialization of VB-SAGE algorithm, a rough estimation of the parameters are obtained. After the initialization, the parameters are adjusted further. A threshold value (related to Signal-Noise-Ratio (SNR)) is preset and the path number is also determined according to this value. When the power of the estimated path is lower than the threshold value, it will be deleted. On the contrary, it will be added in the estimated paths. Due to space limit, we only give the computing results of estimated parameter probability distributions. More detailed theoretical derivation and algorithm description can be referred in [38].

In order to simplify our analysis, we assume the parameters of different paths are independent. The final form of w_l, τ_l, $\phi_{rx,l}$ and $\phi_{tx,l}$ can be computed by maximizing the logarithm function of posterior probability, which are expressed in (9), (10), (11), as shown at the bottom of this page, where η_l and ς_l are the sparsity prior and posterior variance of w_l. For simplicity, the time indicator variable is omitted and $\mathbf{y}(t, f)$ is expressed as $\mathbf{y}$.

2) BEKF TRACKING ALGORITHM

BEKF is firstly initialized by the VB-SAGE algorithm and then the stages of Kalman Filter prediction and VB-SAGE based model order update are implemented in an iterative way. Similar in [23], the flow graph of the modified BEKF is shown in Fig. 2. As illustrated in Section IV, we focus on utilizing the time evolution trend gotten in the tracking stage to make the channel fading prediction. Thus, it is assumed that no birth and death of paths exist in the tracking process in this paper. It can be referred in [23] to get the detailed illustration

$$\ell(\mathbf{y}; w_l) = -\left(\mathbf{y} - \sum_{l=1}^{L} w_l \cdot \mathbf{s}_{\tilde{\tau}_l, \tilde{\phi}_{rx,l}, \tilde{\phi}_{tx,l}}\right)^H \tilde{\sigma}^{-2} \left(\mathbf{y} - \sum_{l=1}^{L} w_l \cdot \mathbf{s}_{\tilde{\tau}_l, \tilde{\phi}_{rx,l}, \tilde{\phi}_{tx,l}}\right) - {w_l}^H \tilde{\eta}_l w_l. \tag{9}$$

$$\ell\left(\mathbf{y}; \tau_l, \phi_{rx,l}, \phi_{tx,l}\right) = -\left(\mathbf{y} - \sum_{l=1}^{L} \tilde{w}_l \cdot \mathbf{s}_{\tau_l, \phi_{rx,l}, \phi_{tx,l}}\right)^H \tilde{\sigma}^{-2} \left(\mathbf{y} - \sum_{l=1}^{L} \tilde{w}_l \cdot \mathbf{s}_{\tau_l, \phi_{rx,l}, \phi_{tx,l}}\right) \tag{10}$$

$$\begin{aligned}\ell\left(\mathbf{y}; \sigma^2\right) = &-\log\left(\left(\sigma^2\right)^N\right) - \left(\mathbf{y} - \sum_{l=1}^{L} \tilde{w}_l \cdot \mathbf{s}_{\tilde{\tau}_l, \tilde{\phi}_{rx,l}, \tilde{\phi}_{tx,l}}\right)^H \sigma^{-2} \left(\mathbf{y} - \sum_{l=1}^{L} w_l \cdot \mathbf{s}_{\tilde{\tau}_l, \tilde{\phi}_{rx,l}, \tilde{\phi}_{tx,l}}\right) \\ &- \sum_{l=1}^{L} \sigma^{-2} \mathbf{s}^H_{\tilde{\tau}_l, \tilde{\phi}_{rx,l}, \tilde{\phi}_{tx,l}} \mathbf{s}_{\tilde{\tau}_l, \tilde{\phi}_{rx,l}, \tilde{\phi}_{tx,l}} \tilde{\varsigma}_l\end{aligned} \tag{11}$$

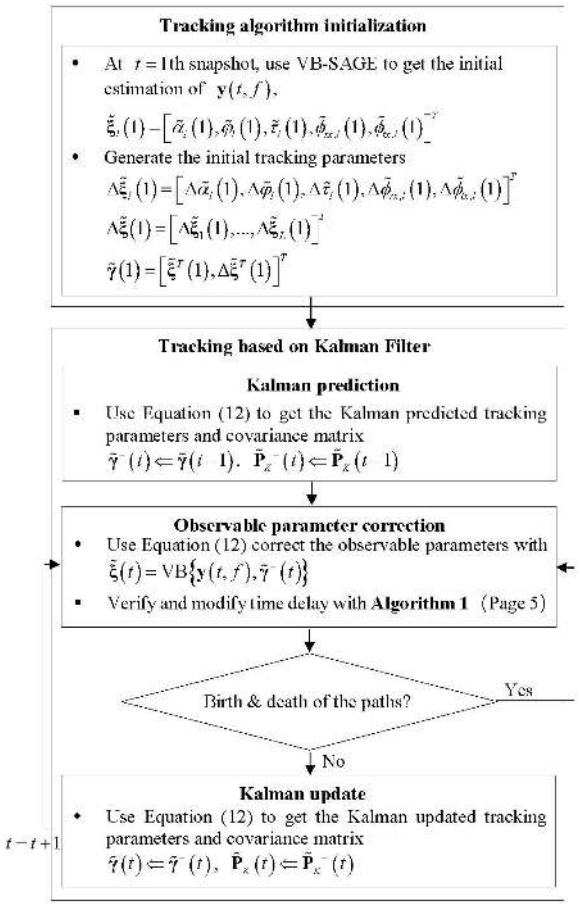


FIGURE 2. Graph flow of BEKF algorithm.

and simulation results for the birth and death of paths (model order change detection) in the tracking stage.

By utilizing the VB-SAGE algorithm, we get the linear state space of the estimated parameters $\tilde{\boldsymbol{\gamma}}(t)$, which can be tracked by Kalman Filter. Note that we add $\sim$ on variables to indicate the estimated values by the algorithm. Therefore, embedding VB-SAGE into Kalman Filter, we can get the following parameter tracking and update equations in (12).

$$\begin{aligned}
\tilde{\boldsymbol{\gamma}}^-(t) &= \mathbf{F}\cdot\tilde{\boldsymbol{\gamma}}(t-1)\\
\tilde{\mathbf{P}}_K^-(t) &= \mathbf{F}\cdot\tilde{\mathbf{P}}_K(t-1)\,\mathbf{F}^H+\mathbf{Q}\\
\mathbf{K}_K &= \tilde{\mathbf{P}}_K^-(t)\cdot\mathbf{H}^T\left(\mathbf{H}\cdot\tilde{\mathbf{P}}_K^-(t)\cdot\mathbf{H}^T+\mathbf{R}\right)^{-1}\\
\tilde{\boldsymbol{\gamma}}(t) &= \tilde{\boldsymbol{\gamma}}^-(t)+\mathbf{K}_K\left(\mathrm{VB}\left\{\mathbf{y}(t,f),\tilde{\boldsymbol{\gamma}}^-(t)\right\}-\mathbf{H}\tilde{\boldsymbol{\gamma}}^-(t)\right)\\
\tilde{\mathbf{P}}_K(t) &= (\mathbf{I}-\mathbf{K}\cdot\mathbf{H})\cdot\tilde{\mathbf{P}}_K^-(t),
\end{aligned}\tag{12}$$

where the superscript $^-$ denotes the estimated value obtained by the step of Kalman Filter prediction. $\tilde{\boldsymbol{\gamma}}^-(t)$ and $\tilde{\mathbf{P}}_K^-(t)$ are the predicted tracking parameter vector and the covariance matrix, while $\tilde{\boldsymbol{\gamma}}(t)$ and $\tilde{\mathbf{P}}_K(t)$ are updated results by the VB-SAGE algorithm. The meaning of transition matrix $\mathbf{F}$ is same with that in (8). $\mathbf{Q}$ is the covariance matrix of the tracking model noise. $\mathbf{H}$ is the measurement matrix of Kalman Filter, which builds an association between the tracking and observable parameters. That is

$$\tilde{\boldsymbol{\xi}}(t)=\mathbf{H}\cdot\tilde{\boldsymbol{\gamma}}(t).\tag{13}$$

$\mathrm{VB}\left\{\mathbf{y}(t,f),\tilde{\boldsymbol{\gamma}}^-(t)\right\}$ is used to estimate the parameters with the initial value of $\tilde{\boldsymbol{\gamma}}^-(t)$. In this step, the parameters between two adjacent snapshots are associated automatically. $\mathbf{R}$ is the measurement noise in Kalman Filter, which can be asymptotically computed by Cramér–Rao lower bound associated to Fisher information matrix. Compared with the derivation in [23], we give a more accurate expression in Appendix B. More details can be found in [22], [23].

Algorithm 1 Time Delay Verification and Modification

Input: CIR with noise $\mathbf{y}(t,f)$ at the tth snapshot, the predicted tracking parameter vector $\tilde{\boldsymbol{\gamma}}^-(t)$, and the updated tracking parameter vector $\tilde{\boldsymbol{\gamma}}(t)$

Output: Updated $\tilde{\boldsymbol{\gamma}}^-(t)$ and $\tilde{\boldsymbol{\gamma}}(t)$

Step 1. Use VB-SAGE to estimate parameters with the same path number L of $\tilde{\boldsymbol{\gamma}}(t)$

$\tilde{\boldsymbol{\xi}}_{VB}(t)=\mathrm{VB}\{\mathbf{y}(t,f)\}$

$\tilde{\boldsymbol{\xi}}_{VB}(t)=\left[\tilde{\boldsymbol{\alpha}}_{VB}^T(t),\tilde{\boldsymbol{\varphi}}_{VB}^T(t),\tilde{\boldsymbol{\tau}}_{VB}^T(t),\tilde{\boldsymbol{\phi}}_{VB,rx}^T(t),\tilde{\sigma}_{VB}^2\right]^T$

Step 2. Sort $\tilde{\boldsymbol{\tau}}_{VB}(t)$, $\tilde{\boldsymbol{\tau}}^-(t)$, $\tilde{\boldsymbol{\tau}}(t)$ with descending order

$\left[\tilde{\boldsymbol{\tau}}_{sort,VB}(t),ind_{VB}\right]=\mathrm{sort}\{\tilde{\boldsymbol{\tau}}_{VB}(t)\}$

$\left[\tilde{\boldsymbol{\tau}}_{sort}^-(t),ind^-\right]=\mathrm{sort}\left\{\tilde{\boldsymbol{\tau}}^-(t)\right\}$

$\left[\tilde{\boldsymbol{\tau}}_{sort}(t),i\tilde{n}d\right]=\mathrm{sort}\{\tilde{\boldsymbol{\tau}}(t)\}$

$\tilde{\boldsymbol{\xi}}_{sort,VB}(t)=\mathrm{sort}\left\{\tilde{\boldsymbol{\xi}}_{VB}(t),ind_{VB}\right\}$

$\tilde{\boldsymbol{\xi}}_{sort}^-(t)=\mathrm{sort}\left\{\tilde{\boldsymbol{\xi}}^-(t),ind^-\right\}$

$\tilde{\boldsymbol{\xi}}_{sort}(t)=\mathrm{sort}\left\{\tilde{\boldsymbol{\xi}}(t),i\tilde{n}d\right\}$

Step 3. Judge the abrupt change of time delay

if $\tilde{\boldsymbol{\tau}}_{sort,VB}(t)=\tilde{\boldsymbol{\tau}}_{sort}(t)$ **then**

$\tilde{\boldsymbol{\tau}}_{sort}(t)=\tilde{\boldsymbol{\tau}}_{sort}(t)$

else

Find $l_{ac}\in\mathbb{L}_{ac}$ with $\tilde{\tau}_{sort,VB,l_{ac}}(t)\neq\tilde{\tau}_{sort,l_{ac}}^-(t)$

Update $\tilde{\boldsymbol{\tau}}_{sort,\mathbb{L}_{ac}}^-(t)$ with (14a) and (14b)

Update $\tilde{\boldsymbol{\xi}}^-(t)$ and $\tilde{\boldsymbol{\gamma}}^-(t)$ with $\tilde{\boldsymbol{\tau}}_{sort,ind_{per}}^-(t)$ and ind^-

Update $\tilde{\boldsymbol{\gamma}}(t)$ with $\mathrm{VB}\left\{\mathbf{y}(t,f),\tilde{\boldsymbol{\gamma}}^-(t)\right\}$

end if

Limited by the bandwidth in the communication system, the resolution of time delay is usually not high enough to vary smoothly with time, which makes $\tau_l(t)$ remain static in several sampling snapshots and then change abruptly in a specific snapshot. This makes it intractable to track time delay with Kalman Filter accurately. Considering this issue, we modify the time delay tracking method. As shown in Fig. 2, a step of observable parameter correction is added and it is described in Algorithm 1.

In Algorithm 1, $\boldsymbol{\xi}_{VB}(t)=\mathrm{VB}\{\mathbf{y}(t,f)\}$ means that using VB-SAGE algorithm to obtain the parameter estimation of $\mathbf{y}(t,f)$ directly. $\boldsymbol{\tau}_{sort,VB}(t)$, $\tilde{\boldsymbol{\tau}}_{sort}^-(t)$ and $\tilde{\boldsymbol{\tau}}_{sort}(t)$ are the time delays obtained by VB-SAGE, Kalman prediction step and Kalman update step in a descending order. ind_{VB}, ind^-

and $\tilde{ind}$ are corresponding sorting index vectors. sort $\{\xi, ind\}$ denotes to rearrange the parameters with the order ind. $\tilde{\boldsymbol{\xi}}_{sort,VB}(t)$, $\tilde{\boldsymbol{\xi}}^{-}_{sort}(t)$ and $\tilde{\boldsymbol{\xi}}_{sort}(t)$ are the sorting results of $\tilde{\boldsymbol{\xi}}_{VB}(t)$, $\tilde{\boldsymbol{\xi}}^{-}(t)$ and $\tilde{\boldsymbol{\xi}}(t)$ with order ind_{VB}, ind^{-} and $\tilde{ind}$. In step 3, due to the extremely small part of the multi-paths with abruptly changing time delay, the original tracking and updated paths can be associated by exhaustive search.

$$\arg\min_{ind_{per}} Z = \mathrm{MCD}\left\{\tilde{\boldsymbol{\xi}}_{sort,VB,\mathbb{L}_{ac}}(t), \tilde{\boldsymbol{\xi}}_{sort,ind_{per}}(t)\right\},$$

$$\text{s.t. } \tilde{\boldsymbol{\xi}}_{sort,ind_{per}}(t) = \mathrm{sort}\left\{\tilde{\boldsymbol{\xi}}_{sort,\mathbb{L}_{ac}}(t), ind_{per}\right\}, \quad (14a)$$

$$\tilde{\boldsymbol{\tau}}^{-}_{sort,ind_{per}}(t) = \tilde{\boldsymbol{\tau}}_{sort,VB,\mathbb{L}_{ac}}(t), \quad (14b)$$

where MCD $\{\cdot\}$ denotes to compute the multi-path component distance (MCD) defined in [39]. ind_{per} is a permutation of $\mathbb{L}_{ac}$ which denotes the path set with the time delay of abrupt change. As shown in (14a) and (14b), the original tracking paths are sorted with different permutations of $\mathbb{L}_{ac}$. And then the MCD values between the original tracking and updated paths can be calculated in case of different permutations. Since MCD distance is calculated with the consideration of all the parameters for each physical path, a small MCD value means the parameters of the original tracking and updated paths are similar and therefore the order with the smallest MCD value can be utilized to associate the original tracking and updated paths.

IV. CHANNEL FADING PREDICTION IN MOBILE PROPAGATION ENVIRONMENT

In this section, we will introduce the small-scale parameter prediction in Step 2 of Fig. 1. The traditional and the proposed CDBP prediction schemes are both presented.

A. TRADITIONAL PCM PREDICTION METHOD

In traditional PCM prediction method, the small-scale parameters $\tilde{\boldsymbol{\xi}}_{l}(t)$ are assumed to be time-invariant in the prediction horizon. In this case, the time-invariant parameters can be easily estimated by snapshot-based algorithms, such as subspace-based ESPRIT [20], multidimensional ESPRIT [5], expectation-maximum-based SAGE [34], and VB-SAGE [38] algorithms. Only the path phases evolve with time in terms of time-invariant Doppler frequencies in the traditional PCM method in [5], [20]. But, in the high-speed scenario, Rx moves with a much larger distance in a short time period. In this case, the assumption of time-invariant small-scale parameters is not reasonable, which degrades the prediction accuracy seriously in terms of time-varying parameters. Since the parameters are estimated independently in each snapshot, it is intractable to track and predict the variation of the small-scale parameters with time in traditional PCM method.

B. PROPOSED PARAMETER TRACKING AND PREDICTION METHOD

In this subsection, we introduce the multi-bounce model which is the basis of the proposed prediction method. With the snapshot-associated small-scale parameters obtained by BEKF in Section III-B, we can utilize the time-evolution trend of the parameters to improve the prediction accuracy, especially in the high-speed scenario.

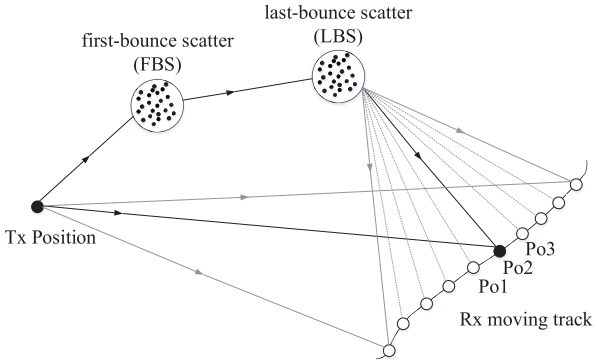


FIGURE 3. Sketch of cluster drifting multi-bounce model.

1) MULTI-BOUNCE MODEL

According to [19], VR is defined to describe the birth and death of clusters. When Rx moves into a VR, the corresponding cluster becomes active and the parameters evolve smoothly with the consideration of the channel SCP. Since the prediction horizon is usually smaller than a VR, the predicted parameters are highly-correlated with those in the previous time snapshots. Therefore, it is possible to predict the time-varying small-scale parameters in a VR. Similar as that in [40], the multi-bounce model used in the paper is illustrated in Fig. 3. In this model, the locations of Tx, the first-bounce scatter and the last-bounce scatter are assumed to be static when Rx is moving along a trajectory as introduced in [28]. This assumption is reasonable when we consider the distance between BS and MS is much larger than the prediction horizon. And also the clusters with significant impact on the CIR are only associated with some main scatterers. Based on the assumption above, the complex magnitude, time delay, AoA and AoD of the line-of-sight (LoS) path vary with Rx moving. We call the above assumed prediction model as *cluster drifting*. The drifting length, i.e., the lifetime of the VR cluster has been researched in [29] and [30].

2) CLUSTER DRIFTING BASED PREDICTION ALGORITHM

With BEKF algorithm introduced in Section III-B, the time-varying small-scale parameters in the tracking horizon can be obtained. Also with the multi-bounce model in IV-B1, it is potential to utilize the tracking small-scale parameters to estimate the positions of FBSs and LBSs in Fig. 3. Then, based on the principle of geometric-based channel model in [31]–[33], the channel fading can be calculated utilizing the positions of Tx, Rx, FBSs and LBSs. However, this geometry-based method requires very accurate position information. Since the phase of each path varies dramatically in the wavelength level, the channel prediction accuracy may be degraded severely with a little error of position information.

Therefore, this geometric based method is not considered in this paper and it is more useful in the channel modeling rather than channel prediction.

We hope to predict the channel fading with the tracking small-scale parameters directly. With the SCP guaranteed in a VR, the smoothly time-varying parameters can be fitted by a specific model and it can be used to predict the parameters in the prediction horizon. It is worth mentioning that since the SCP is assumed for each cluster, it is guaranteed for all the paths corresponding to each cluster. Although the multi-bounce model in Fig. 3 is based on the cluster, the SCP can be utilized to predict the parameters of the paths in each cluster independently. This simplifies the steps without clustering in the prediction progress. In this paper, polynomial fitting is utilized to catch the time-evolution trend of different parameters, and the model order validation is also adopted to improve the accuracy of the fitting model. With the consideration of tracking the parameters including complex magnitude $w_l(t)$, time delay $\tau_l(t)$, and angles $\phi_{rx,l}(t)$, $\phi_{tx,l}(t)$, we give the graph flow of the prediction model in Fig. 4. In the following, the prediction schemes for different parameters are introduced and discussed.

Since the prediction range is very small compared with the large distance between Tx and Rx, i.e., usually several wavelengths, AoD of the LoS path and the path magnitude vary much smaller compared with other time-varying parameters. In this case, the impact of pass loss and AoD on CSI can be ignored in the step of channel prediction. Accordingly, the magnitude α_l and AoD $\phi_{tx,l}(t)$ are assumed to be constant within the prediction region, and only the time-evolution trend of $\varphi_l(t)$, $\tau_l(t)$, and $\phi_{rx,l}(t)$ need to be fitted and predicted, which are only associated with the last-bounce scatterer (LBS) as shown in Fig. 3.

However, due to the random uniformly distributed property of $\varphi_l(t)$ as described in [37], it is hard to fit and predict the time-evolution trend. In the BEKF algorithm, the parameter change variable $\Delta\varphi_l(t)$ is considered to help track $\varphi_l(t)$ for its slowly-varying property as shown in the empirical density of the path phase change in [37]. Also, the time-varying Doppler phase information is implied in $\Delta\varphi_l(t)$, and is utilized to predict the time-varying trend of $\varphi_l(t)$ in the prediction step.

Another reason for selecting $\Delta\varphi_l(t)$ as the prediction variable is its symmetric property under the assumption of the multi-bounce model in Section IV-B1. For the position of Tx, LBS and FBS are assumed to be static, there might exist positions Po1, Po2 and Po3 with Po1 and Po3 being symmetrical about Po2 as shown in Fig. 3. In this case, the distance between LBS and Po1 is equal to that between LBS and Po3, which means the phase difference change from Po1 to Po2 is just the negative of that from Po2 to Po3. This characteristic can be used to improve the prediction accuracy and range. With the analysis given above, the prediction steps of $\Delta\varphi_l(t)$, $\tau_l(t)$, and $\phi_{rx,l}(t)$ are given in the following.

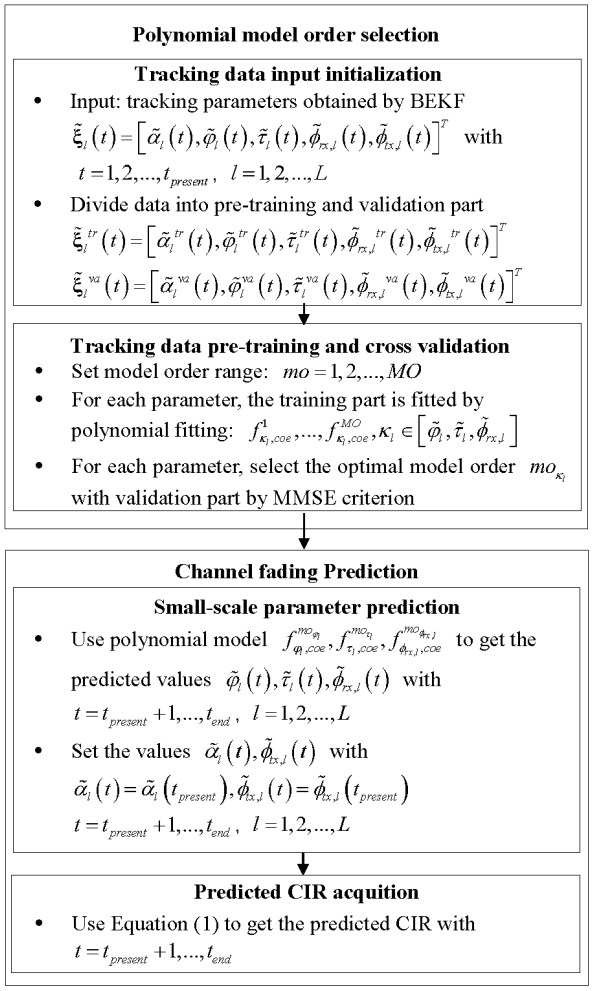


FIGURE 4. Graph flow of CDBP algorithm.

Prediction of $\varphi_l(t)$***:***

- $\Delta\varphi_l(t)=f^{mo}_{\Delta\varphi_l,coe}(t)$, $t=t_{present}+1,\ldots,t_{end}$.
- if there exist t_0 satisfying
  $\Delta\varphi_l(t_0)\cdot\Delta\varphi_l(t_0+1)<0$, $t=t_{present}+1,\ldots,t_{end}-1$,
  then
  $\Delta\varphi_l(t)=-\Delta\varphi_l(2t_0-t)$, $t=t_0+1,\ldots,t_{end}$.

Prediction of $\tau_l(t)$***,*** $\phi_{rx,l}(t)$***,*** $\alpha_l(t)$ ***and*** $\phi_{tx,l}(t)$***:***

- $\phi_{rx,l}(t)=f^{mo}_{\phi_{rx,l},coe}(t)$, $t=t_{present}+1,\ldots,t_{end}$,
  $\tau_l(t)=f^{mo}_{\tau_l,coe}(t)$, $t=t_{present}+1,\ldots,t_{end}$,
  $\alpha_l(t)=\alpha_l\left(t_{present}\right)$, $t=t_{present+1},\ldots,t_{end}$,
  $\phi_{tx,l}(t)=\phi_{tx,l}\left(t_{present}\right)$, $t=t_{present+1},\ldots,t_{end}$.

where $t_{present}$ and t_{end} are the index of the present and the last prediction time snapshots. t_0 is the snapshot of Po2 in Fig. 3. $f^{mo}_{\Delta\varphi_l,coe}$, $f^{mo}_{\phi_{rx,l},coe}$ and $f^{mo}_{\tau_l,coe}$ are the optimum polynomial fitting model given in Fig. 4. The tracking parameters obtained by BEKF are divided into training and testing parts. With the models of order range $mo=1,2,\ldots,MO$ trained by training data, the optimum model order mo is verified and selected by the testing data.

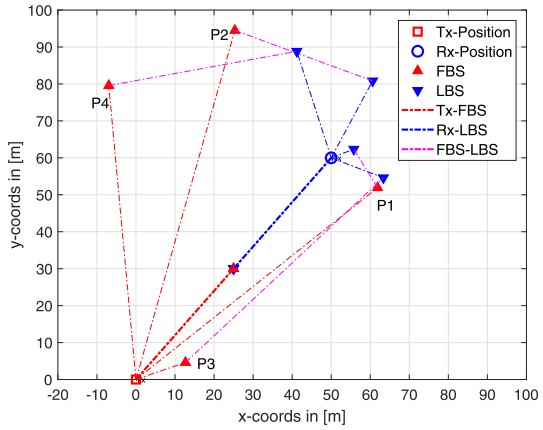


FIGURE 5. Sketch of simulated scenario generated by channel simulation model in [40]: Tx is located at [0, 0]. Rx is moving along a linear trajectory from [50, 60] to [60, 60]. 16-element ULA and 16-element UCA antennas with isotropic radiators of vertical polarization are equipped at Tx and Rx respectively.

V. SIMULATION RESULTS

In order to verify the performance, the proposed channel fading prediction framework has been implemented and tested by a GBSM based channel simulation model in [40]. Since the channel SCP is considered in this model, it can substantially support the simulation of time-varying MIMO channel fading. And also with the account of the prediction horizon smaller than one VR, it is assumed that no birth and death of the clusters exist in the simulation of tracking and prediction stages. We consider the urban micro-cell NLoS scenario in [32]. For only azimuth angle is considered in this paper, all the model parameters retain the default values other than the elevation angles to be set as zeros. The sketch of the generated scenario is given in Fig. 5. Tx is located at origin and Rx is moving along a ten-meter trajectory in a far field condition. According to [40], the parameters change smoothly and the SCP is guaranteed in this small segment. Some other specific simulation parameters are given in Table 1. Since the moving distance of Rx is directly proportional to the mobile velocity (e.g., 4.2λ in the case of 150 km/h and 3 GHz within 10 ms) with a specific channel sampling rate, the length of prediction horizon is used to verify the performance of the prediction algorithm in this paper. Better performance in a larger prediction horizon by the proposed method means it can achieve better performance in terms of higher moving velocity within a specific time period. Also as shown in Table 1, the sample density (per $\frac{\lambda}{2}$) decreases with the mobile velocity increasing in terms of a fixed channel sampling rate. To show the performance clearly, only four NLoS clusters (with one path in each cluster, P1, P2, P3, P4 in Fig. 5) are generated in the simulation.

Since a MIMO-OFDM downlink system with the TDD mode is considered in this paper, the channel prediction error with the CSI obtained by uplink channel estimation and the bit-error-rate (BER) performance at Rx in the downlink system are used to test the proposed scheme. Also, for the

TABLE 1. Channel model simulation configurations.

Parameter	Value
Carrier Frequency	3 GHz
Bandwidth	100 MHz
Time Delay Resolution	10 ns
Frequency bin number	64
Tx Antenna	ULA with 16 isotropic radiators in vertical polarization
Rx Antenna	UCA with 16 isotropic radiators in vertical polarization
Sample Density	5, 10, 15 per $\frac{\lambda}{2}$
Mobile Velocity	150, 100, 50 km/h

CSI obtained from uplink channel estimation is usually not perfect, the performance is tested at different uplink SNR. The uplink SNR of estimated CSI is defined as $\mathrm{SNR}_{uplink} = E_t\left[\|\mathbf{g}(t,f)\|_2^2\right] / \left(\sigma^2 USM\right)$, where M is the frequency bin number, σ^2, U, and S are defined same as those in Section II. As given in [32], the delay spread of the urban micro-cell NLoS scenario is about 129 ns, which means the coherent bandwidth is about 7.76 MHz. Thus in the simulation, $M = 64$ is selected to guarantee the channel to be frequency-flat in each frequency bin.

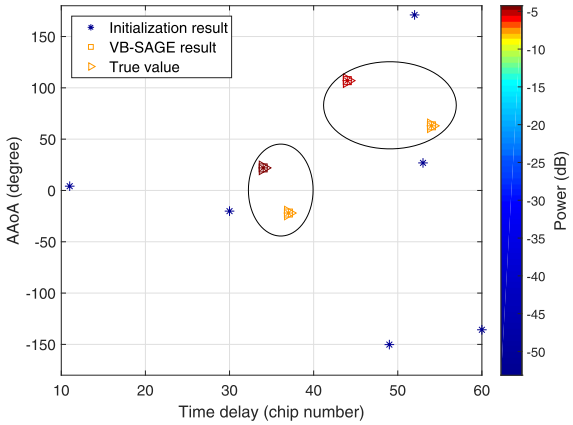


FIGURE 6. The compared results of parameters among the initialization, VB-SAGE algorithm and the channel model.

First, the results of the estimated parameters obtained by the initialization and VB-SAGE algorithm in the first snapshot is compared with the true values obtained from the channel model in Fig. 6. As can be seen from Fig. 6, the parameters are estimated roughly in the initialization step, in which many paths with very small power are also obtained. However, these parameters are estimated further in the step of VB-SAGE algorithm and the paths with very small power are deleted according to a preset threshold SNR (7 dB in the simulation). Therefore, with the step of VB-SAGE algorithm, the accurate path number can be obtained and these estimated paths can be tracked and predicted independently.

Then, a visualization of the parameters evolving with time is shown in Fig. 7. The negative and positive parts of x-axis (in wavelength) correspond to the parameter tracking and prediction horizons respectively. As can be seen in Fig. 7,

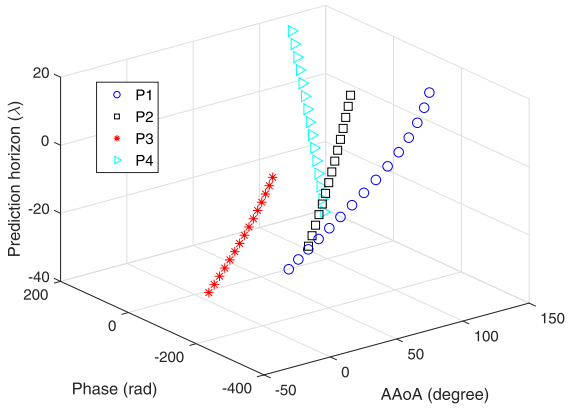


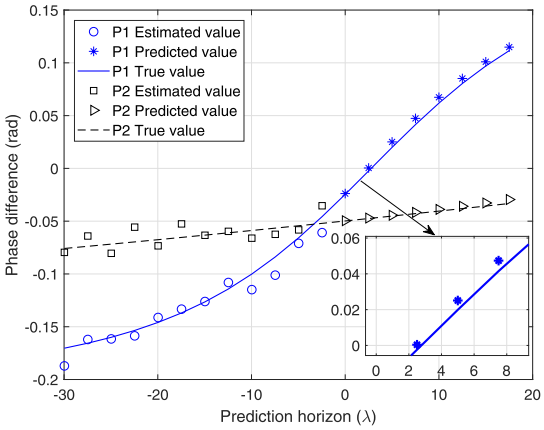
FIGURE 7. Visualization of time-varying parameter evolution for all paths, with uplink SNR of estimated CSI equal to 15 dB.

the parameters vary greatly with time. Especially for P1, almost the difference of 100 degree for AAoA can be observed within the range of about 50 λ. This is because the LBS of P1 is very close to the Rx, and the LBS has a significant impact on the CSI in a very short distance. In this case, it is not reasonable to assume the parameters to be time-invariant. Therefore, it is necessary to track the evolution trend of the time-varying parameters.

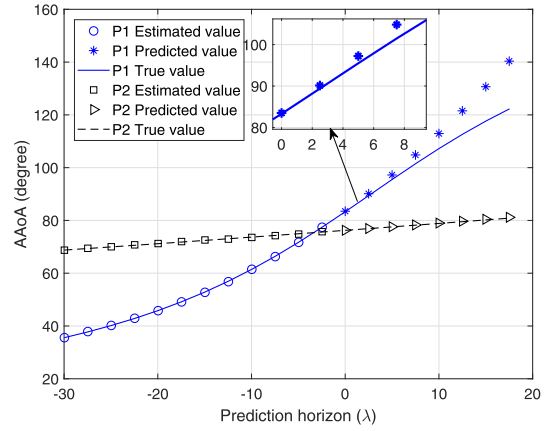
Then the results of parameter tracking and prediction are shown in Fig. 8. The paths P1 and P2 are selected as the examples. The estimated and predicted values denote those obtained from the BEKF tracking and CBDP algorithms, respectively. The true value is generated from the channel simulation model in [40]. In the tracking horizon, it can be observed that the values of AAoAs vary with time and this lead to the time-varying Doppler frequencies (especially for P1 in Fig 5 and 8). With the time-varying parameter tracking and fitting, the prediction scheme performs well in the prediction horizon between 0 and 10 λ, especially after 2.5 λ where the phase difference values are opposite to that before 2.5 λ.

Next, we show the errors between the CSI estimated and predicted by the proposed scheme and the target CSI value generated from the channel model in the tracking and prediction horizons. The errors are measured by normalized square error (NSE) as shown in Fig. 9. Also, the NSE lower bound (NSELB) is plotted for comparison. The calculations of NSE and NSELB are given in Appendix C. With the criterion similar to [41], −13 dB is used to judge if the predicted sample is useful in the channel prediction. It is observed that the proposed CDBP algorithm performs much better than the traditional method in terms of the useful prediction horizon length. This is intuitive in Fig. 8 where the parameters (phase difference, AAoA and time delay) vary with time and our algorithm can track the time-evolution trend and obtain a more accurate prediction performance.

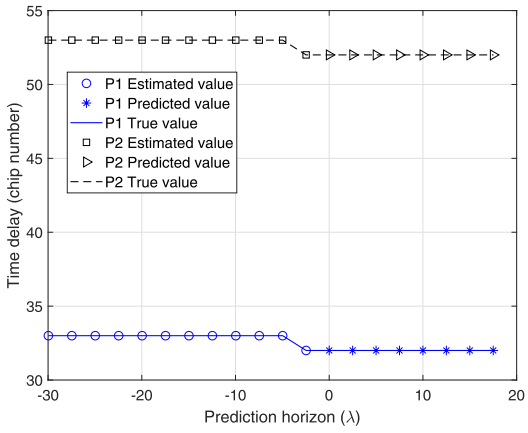
To verify the impact of uplink SNR of estimated CSI on the proposed prediction scheme, the NSE performances at different uplink SNR of estimated CSI are given in Fig. 10.



(a) Phase difference



(b) AAoA



(c) Time delay

FIGURE 8. Results of estimated values by BEKF and predicted values by CDBP for P1 and P2 paths, with uplink SNR of estimated CSI equal to 15 dB.

As expected, it can be observed that NSE decreases with uplink SNR increasing for both the traditional and the proposed prediction schemes. However, note that the NSE performances in the prediction horizon are very similar at 15 and 20 dB, which is limited by the precision of parameter extraction in VB-SAGE algorithm.

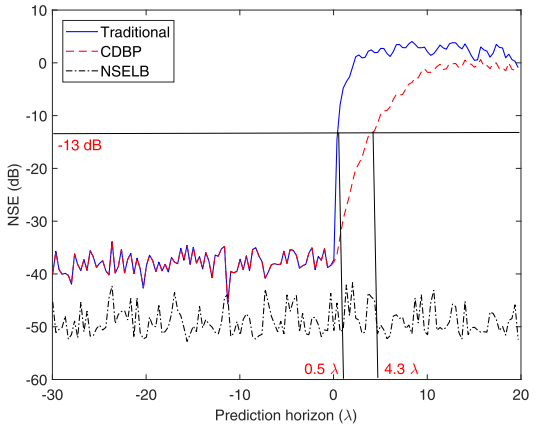


FIGURE 9. NSE versus prediction horizon, with uplink SNR of estimated CSI equal to 15 dB.

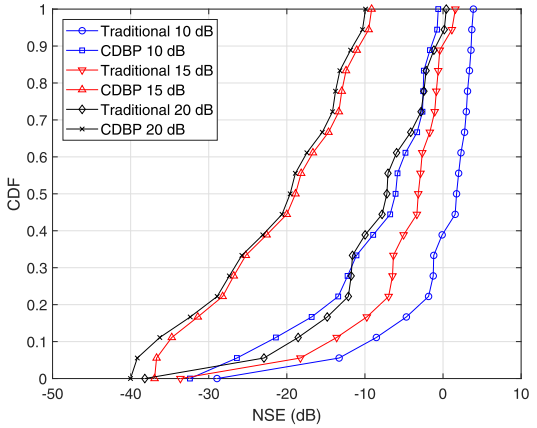


FIGURE 10. The CDF of NSE for different uplink SNR of estimated CSI within 6 λ prediction horizon.

Also, the BER performances versus the prediction horizon length at different received SNR in a MIMO-OFDM downlink system are presented in Fig. 11. The predicted CSI obtained by the traditional and the proposed CDBP schemes and the perfect CSI generated by the channel simulation model are used to conduct the precoding at Tx as shown in Fig. 1. In the case of the traditional method, the BER performances degrade seriously due to the severe mismatch between the precoding weights and the time-varying CSI. However by the CDBP scheme, the BER reduces by about up to three orders of magnitude in the 2 λ prediction horizon at 20 dB downlink received SNR. Although both the traditional and the proposed CDBP schemes become worse with the prediction horizon length increasing, the CDBP scheme performs much better especially in the case of higher downlink received SNR.

Finally, the proposed channel prediction method is adopted in the case of 19 clusters (with one path in each cluster), which is same with the value defined in urban micro-cell NLoS scenario in [32]. Similar with the case of 4 paths,

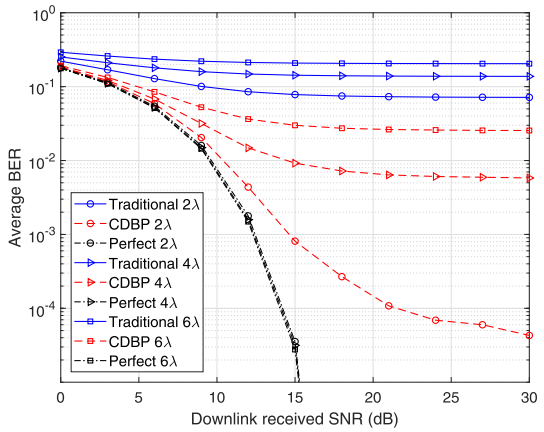


FIGURE 11. BER vs prediction horizon of 2, 4, and 6 λ with uplink SNR of estimated CSI equal to 20 dB.

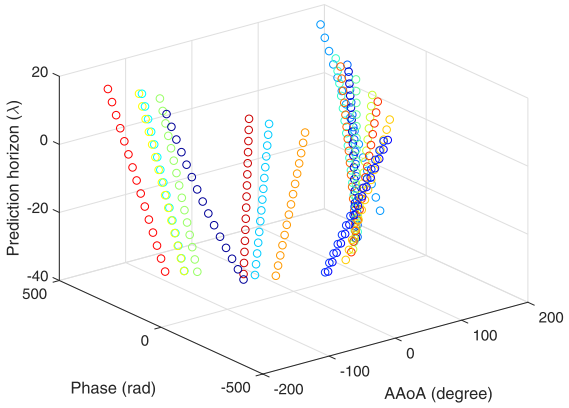


FIGURE 12. Visualization of time-varying parameter evolution for all paths, with uplink SNR of estimated CSI equal to 20 dB, in the case of 19 clusters.

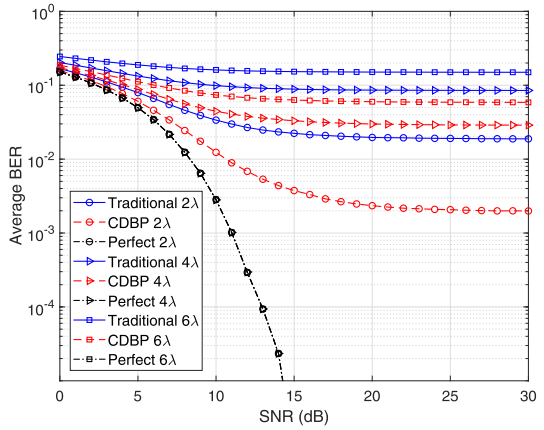


FIGURE 13. BER vs prediction horizon of 2, 4, and 6 λ with uplink SNR of estimated CSI equal to 20 dB, in the case of 19 clusters.

a visualization of the parameters evolving with time is shown in Fig. 12. And then the BER performances versus the prediction horizon length in the case of 19 clusters are shown in Fig. 13. The results in Fig. 13 are consistent with them

in Fig. 11. The proposed CDBP scheme performs better than the traditional scheme in all prediction horizon lengths. On the one hand, due to the increase of the cluster number, it becomes more difficult to estimate and track the small-scale parameters of all paths accurately. And also the error introduced by the fitting model might become large. For example, the error floor in the case of CDBP 2 λ in Fig. 13 is much higher than that in Fig. 11. On the other hand, the power of each independent cluster becomes smaller with the fixed SNR (20 dB). Therefore, the error introduced by the parameter estimation, tracking and fitting model in the case of 19 clusters also become smaller compared with that in the case of 4 clusters, especially for the traditional method in the case of 2 λ in Fig. 13.

VI. CONCLUSION

In this paper, we have presented a framework of time-varying MIMO channel fading prediction in a high-speed scenario. The proposed framework consists of time-varying small-scale parameter tracking and prediction. In the tracking stage, the parameters are estimated and tracked at present and previous snapshots by the BEKF algorithm. Then a specific fitting model (polynomial fitting) is used to extract the time-evolution trends of different parameters. Model order verification has been adopted to avoid the under-or-over fitting. In the prediction stage, under the assumption of cluster-based channel spatial consistency property in the prediction horizon, the parameters are predicted using the fitting model, and the predicted CIRs are obtained by the predicted parameters. Simulation results have shown that the proposed channel fading prediction framework tracks and predicts the time-varying small-scale parameters efficiently, thus achieving an improved system performance.

APPENDIX A
EXPRESSION OF F

First, the discrete-time state transition equation for a single path is given as

$$\begin{bmatrix} \alpha_l(t) \\ \varphi_l(t) \\ \tau_l(t) \\ \phi_{rx,l}(t) \\ \phi_{tx,l}(t) \\ \sigma^2(t) \\ \Delta\alpha_l(t) \\ \Delta\varphi_l(t) \\ \Delta\tau_l(t) \\ \Delta\phi_{rx,l}(t) \\ \Delta\phi_{tx,l}(t) \end{bmatrix} = \begin{bmatrix} \alpha_l(t-1) + \Delta\alpha_l(t-1) \\ \varphi_l(t-1) + \Delta\varphi_l(t-1) \\ \tau_l(t-1) + \Delta\tau_l(t-1) \\ \phi_{rx,l}(t-1) + \Delta\phi_{rx,l}(t-1) \\ \phi_{tx,l}(t-1) + \Delta\phi_{tx,l}(t-1) \\ \sigma^2(t-1) \\ \Delta\alpha_l(t-1) \\ \Delta\varphi_l(t-1) \\ \Delta\tau_l(t-1) \\ \Delta\phi_{rx,l}(t-1) \\ \Delta\phi_{tx,l}(t-1) \end{bmatrix}, \tag{15}$$

Thus the transition matrix $\mathbf{F}_l$ of the tracking parameter vector for a single path can be easily obtained from (15). Therefore, the transition matrix $\mathbf{F}$ is expressed as $\mathbf{F} = \mathbf{F}_l \otimes \mathbf{I}_L$.

APPENDIX B
DERIVATION OF R

For VB-SAGE is unbiased Bayesian Estimator, the Kalman Filter measurement additive noise, i.e., mean squared error (MSE) can be lower bounded by Cramér–Rao bound, which is computed by the inverse of the Fisher Information matrix [42]. First, we give the simplified form of $\mathbf{R}$:

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_{\boldsymbol{\alpha}} & 0 & 0 & 0 & 0 & 0 \\ 0 & \mathbf{R}_{\boldsymbol{\varphi}} & 0 & 0 & 0 & 0 \\ 0 & 0 & \mathbf{R}_{\boldsymbol{\tau}} & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{R}_{\boldsymbol{\phi}_{rx}} & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{R}_{\boldsymbol{\phi}_{tx}} & 0 \\ 0 & 0 & 0 & 0 & 0 & \mathbf{R}_{\sigma^2} \end{bmatrix}, \tag{16}$$

where the diagonal submatrices $\mathbf{R}_{\boldsymbol{\alpha}}$, $\mathbf{R}_{\boldsymbol{\varphi}}$, $\mathbf{R}_{\boldsymbol{\tau}}$, $\mathbf{R}_{\phi_{rx}}$, $\mathbf{R}_{\phi_{tx}}$ and $\mathbf{R}_{\sigma^2}$ are the covariance matrixes. They can be computed by the Fisher Information Matrixes $\mathcal{I}(\boldsymbol{\alpha})$, $\mathcal{I}(\boldsymbol{\varphi})$, $\mathcal{I}(\boldsymbol{\tau})$, $I(\boldsymbol{\phi}_{rx})$, $I(\boldsymbol{\phi}_{rx})$ and $\mathcal{I}(\sigma^2)$ respectively. The time indicator t is dropped off. The elements of them are given as

$$\mathcal{I}(\boldsymbol{\kappa})_{i,k} = -\mathbb{E}_{\boldsymbol{\kappa}}\left[\frac{\partial \ell(\mathbf{y};\boldsymbol{\kappa})}{\partial\kappa_i \partial\kappa_j}\right], \tag{17}$$

where $\boldsymbol{\kappa} \in [\boldsymbol{\alpha}, \boldsymbol{\varphi}, \boldsymbol{\tau}, \boldsymbol{\phi}_{rx}, \boldsymbol{\phi}_{tx}]$, the logarithm function of posterior probability $\ell(\mathbf{y};\boldsymbol{\kappa})$ of α_l, φ_l, τ_l, $\phi_{rx,l}$, $\phi_{tx,l}$ and σ^2 are given in (9), (10) and (11) respectively.

Let $\Gamma = \mathrm{diag}\{\boldsymbol{\eta}\}$, where $\boldsymbol{\eta} = [\eta_1, \ldots, \eta_L]$ and η_l is same as that in (9). Beginning with (9) and using $w_l = \alpha_l \cdot e^{j\varphi_l}$, we have the derivation of $\mathcal{I}(\boldsymbol{\alpha})$ and $\mathcal{I}(\boldsymbol{\varphi})$ in (18) and (19), as shown at the next page. In the following, we give the derivation of $\mathcal{I}(\boldsymbol{\tau})$, $\mathcal{I}(\boldsymbol{\phi}_{rx})$ and $\mathcal{I}(\boldsymbol{\phi}_{tx})$ in (20), as shown at the next page, where $\iota \in [\tau, \phi_{rx}, \phi_{tx}]$.

For only azimuth angles are considered, we let $\mathbf{D}_{pos} = [\mathbf{dx}, \mathbf{dy}]^T$, where $\mathbf{dx}$ and $\mathbf{dy}$ are the decomposition of the antenna element location vector on two orthogonal directions in azimuth plane. Combining (3) and (6), we have

$$\frac{\partial}{\partial\phi_{rx,i}}\left\{\left(\sum_{l=1}^{L}\mathbf{s}_{\tau_l,\phi_{rx,l},\phi_{tx,l}}w_l\right)\right\} = \mathbf{u}(\tau_i) \otimes \mathbf{c}(\phi_{tx,i}) \otimes \left(\frac{\partial\mathbf{c}(\phi_{rx,i})}{\partial\phi_{rx,i}}\right) \cdot w_i, \tag{22a}$$

$$\frac{\partial\mathbf{c}(\phi_{rx,i})}{\partial\phi_{rx,i}} = \mathbf{c}(\phi_{rx,i}) \odot \left(-\mathbf{dx}\cdot\sin\phi_{rx,i} + \mathbf{dy}\cdot\cos\phi_{rx,i}\right). \tag{22b}$$

And also we can give the expression for Tx in the similar way.

Let $\mathbf{f} = [f_0, \ldots, f_{M-1}]^T$. Using (6), we have

$$\frac{\partial}{\partial\tau_i}\left\{\left(\sum_{l=1}^{L}\mathbf{s}_{\tau_l,\phi_{rx,l},\phi_{tx,l}}w_l\right)\right\} = \frac{\partial\mathbf{u}(\tau_i)}{\partial\tau_i} \otimes \mathbf{c}(\phi_{tx,i}) \otimes \mathbf{c}(\phi_{rx,i}) \cdot w_i, \tag{23a}$$

$$\frac{\partial\mathbf{u}(\tau_i)}{\partial\tau_i} = \left[\mathbf{u}(\tau_i) \odot (-j2\pi\mathbf{f})\right]. \tag{23b}$$

$$\mathcal{I}(\boldsymbol{\alpha})_{i,k} = -\mathbb{E}_{\boldsymbol{\alpha}}\left[\frac{\partial \ell(\mathbf{y};\alpha)}{\partial\alpha_i\partial\alpha_k}\right] = -\mathbb{E}_{\alpha}\left[\frac{\partial}{\partial\alpha_k}\left\{\frac{\partial w_i^*}{\partial\alpha_i}\mathbf{s}^H_{\tau_l,\phi_{rx,l},\phi_{tx,l}}\sigma^{-2}\mathbf{y} - \frac{\partial w_i^*}{\partial\alpha_i}\mathbf{s}^H_{\tau_l,\phi_{rx,l},\phi_{tx,l}}\sigma^{-2}\left(\sum_{l=1}^{L} w_l\cdot\mathbf{s}_{\tau_l,\phi_{rx,l},\phi_{tx,l}}\right)\right.\right.$$
$$\left.\left.-\left(\sum_{l=1}^{L} w_l\cdot\mathbf{s}_{\tau_l,\phi_{rx,l},\phi_{tx,l}}\right)^H\sigma^{-2}\mathbf{s}_{\tau_l,\phi_{rx,l},\phi_{tx,l}}\frac{\partial w_i}{\partial\alpha_i} + \mathbf{y}^H\sigma^{-2}\mathbf{s}_{\tau_l,\phi_{rx,l},\phi_{tx,l}}\frac{\partial w_i}{\partial\alpha_i} - \sum_{l=1}^{L}\frac{\partial w_i^*}{\partial\alpha_i}\Gamma_{i,l}w_l - \sum_{l=1}^{L} w_l^*\Gamma_{l,i}\frac{\partial w_i}{\partial\alpha_i}\right\}\right]$$
$$= -\mathbb{E}_{\alpha}\left[-\frac{\partial w_i^*}{\partial\alpha_i}\mathbf{s}^H_{\tau_l,\phi_{rx,l},\phi_{tx,l}}\sigma^{-2}\mathbf{s}_{\tau_l,\phi_{rx,l},\phi_{tx,l}}\frac{\partial w_k}{\partial\alpha_k} - \frac{\partial w_i^*}{\partial\varphi_i}\Gamma_{i,k}\frac{\partial w_k}{\partial\varphi_k} - \frac{\partial w_k^*}{\partial\varphi_k}\Gamma_{k,i}\frac{\partial w_i}{\partial\varphi_i}\right]$$
$$= \mathbb{E}_{\alpha}\left[\mathbb{R}\left\{2e^{-j\varphi_i}\mathbf{s}^H_{\tau_l,\phi_{rx,l},\phi_{tx,l}}\sigma^{-2}\mathbf{s}_{\tau_l,\phi_{rx,l},\phi_{tx,l}}e^{j\varphi_k} + 2e^{-j\varphi_i}\Gamma_{i,k}e^{j\varphi_k}\right\}\right] \tag{18}$$

$$\mathcal{I}(\boldsymbol{\varphi})_{i,k} = -\mathbb{E}_{\boldsymbol{\varphi}}\left[\frac{\partial \ell(\mathbf{y};\varphi)}{\partial\varphi_i\partial\varphi_k}\right] = -\mathbb{E}_{\varphi}\left[\frac{\partial}{\partial\varphi_k}\left\{\frac{\partial w_i^*}{\partial\varphi_i}\mathbf{s}^H_{\tau_l,\phi_{rx,l},\phi_{tx,l}}\sigma^{-2}\mathbf{y} - \frac{\partial w_i^*}{\partial\varphi_i}\mathbf{s}^H_{\tau_l,\phi_{rx,l},\phi_{tx,l}}\sigma^{-2}\left(\sum_{l=1}^{L} w_l\cdot\mathbf{s}_{\tau_l,\phi_{rx,l},\phi_{tx,l}}\right)\right.\right.$$
$$\left.\left.-\left(\sum_{l=1}^{L} w_l\cdot\mathbf{s}^H_{\tau_l,\phi_{rx,l},\phi_{tx,l}}\right)^H\sigma^{-2}\mathbf{s}_{\tau_l,\phi_{rx,l},\phi_{tx,l}}\frac{\partial w_i}{\partial\varphi_i} + \mathbf{y}^H\sigma^{-2}\mathbf{s}_{\tau_l,\phi_{rx,l},\phi_{tx,l}}\frac{\partial w_i}{\partial\varphi_i} - \sum_{l=1}^{L}\frac{\partial w_i^*}{\partial\varphi_i}\Gamma_{i,l}w_l - \sum_{l=1}^{L} w_l^*\Gamma_{l,i}\frac{\partial w_i}{\partial\varphi_i}\right\}\right]$$
$$= -\mathbb{E}_{\varphi}\left[-\frac{\partial w_i^{*H}}{\partial\varphi_i}\mathbf{s}^H_{\tau_l,\phi_{rx,l},\phi_{tx,l}}\sigma^{-2}\mathbf{s}_{\tau_l,\phi_{rx,l},\phi_{tx,l}}\frac{\partial w_k}{\partial\varphi_k} - \frac{\partial w_k^*}{\partial\varphi_k}\mathbf{s}^H_{\tau_l,\phi_{rx,l},\phi_{tx,l}}\sigma^{-2}\mathbf{s}_{\tau_l,\phi_{rx,l},\phi_{tx,l}}\frac{\partial w_i}{\partial\varphi_i} - \frac{\partial w_i^*}{\partial\varphi_i}\Gamma_{i,k}\frac{\partial w_k}{\partial\varphi_k} - \frac{\partial w_k^*}{\partial\varphi_k}\Gamma_{k,i}\frac{\partial w_i}{\partial\varphi_i}\right]$$
$$= \mathbb{E}_{\varphi}\left[\mathbb{R}\left\{2\alpha_i\alpha_k e^{-j\varphi_i}\mathbf{s}^H_{\tau_l,\phi_{rx,l},\phi_{tx,l}}\sigma^{-2}\mathbf{s}_{\tau_l,\phi_{rx,l},\phi_{tx,l}}e^{j\varphi_k} + 2\alpha_i\alpha_k e^{-j\varphi_i}\Gamma_{i,k}e^{j\varphi_k}\right\}\right] \tag{19}$$

$$\mathcal{I}(\iota)_{i,k} = -\mathbb{E}_{\iota}\left[\frac{\partial \ell(\mathbf{y};\iota)}{\partial\iota_i\partial\iota_k}\right] = -\mathbb{E}_{\iota}\left[\frac{\partial}{\partial\iota_k}\left\{\frac{\partial}{\partial\iota_i}\left\{\left(\sum_{l=1}^{L}\mathbf{s}_{\tau_l,\phi_{rx,l},\phi_{tx,l}}w_l\right)^H\right\}\sigma^{-2}\mathbf{y} + \mathbf{y}^H\sigma^{-2}\frac{\partial}{\partial\iota_i}\left\{\left(\sum_{l=1}^{L}\mathbf{s}_{\tau_l,\phi_{rx,l},\phi_{tx,l}}w_l\right)\right\}\right.\right.$$
$$\left.\left.-\frac{\partial\ell}{\partial\iota_i}\left\{\left(\sum_{l=1}^{L}\mathbf{s}_{\tau_l,\phi_{rx,l},\phi_{tx,l}}w_l\right)^H\right\}\sigma^{-2}\left(\sum_{l=1}^{L}\mathbf{s}_{\tau_l,\phi_{rx,l},\phi_{tx,l}}w_l\right) - \left(\sum_{l=1}^{L}\mathbf{s}_{\tau_l,\phi_{rx,l},\phi_{tx,l}}w_l\right)^H\sigma^{-2}\frac{\partial}{\partial\iota_i}\left\{\left(\sum_{l=1}^{L}\mathbf{s}_{\tau_l,\phi_{rx,l},\phi_{tx,l}}w_l\right)\right\}\right]$$
$$= -\mathbb{E}_{\iota}\left[\frac{\partial}{\partial\iota_i\partial\iota_k}\left\{\left(\sum_{l=1}^{L}\mathbf{s}_{\tau_l,\phi_{rx,l},\phi_{tx,l}}w_l\right)^H\right\}\sigma^{-2}\left(\mathbf{y} - \sum_{l=1}^{L}\mathbf{s}_{\tau_l,\phi_{rx,l},\phi_{tx,l}}w_l\right) + \left(\mathbf{y} - \sum_{l=1}^{L}\mathbf{s}_{\tau_l,\phi_{rx,l},\phi_{tx,l}}w_l\right)^H\sigma^{-2}\right.$$
$$\cdot\frac{\partial}{\partial\iota_i\partial\iota_k}\left\{\left(\sum_{l=1}^{L}\mathbf{s}_{\tau_l,\phi_{rx,l},\phi_{tx,l}}w_l\right)\right\} - \frac{\partial}{\partial\iota_i}\left\{\left(\sum_{l=1}^{L}\mathbf{s}_{\tau_l,\phi_{rx,l},\phi_{tx,l}}w_l\right)^H\right\}\sigma^{-2}\frac{\partial}{\partial\iota_k}\left\{\left(\sum_{l=1}^{L}\mathbf{s}_{\tau_l,\phi_{rx,l},\phi_{tx,l}}w_l\right)\right\}$$
$$\left.-\frac{\partial}{\partial\iota_k}\left\{\left(\sum_{l=1}^{L}\mathbf{s}_{\tau_l,\phi_{rx,l},\phi_{tx,l}}w_l\right)^H\right\}\sigma^{-2}\frac{\partial}{\partial\iota_i}\left\{\left(\sum_{l=1}^{L}\mathbf{s}_{\tau_l,\phi_{rx,l},\phi_{tx,l}}w_l\right)\right\}\right]$$
$$= \mathbb{E}_{\iota}\left[\mathbb{R}\left\{2\frac{\partial}{\partial\iota_i}\left\{\left(\sum_{l=1}^{L}\mathbf{s}_{\tau_l,\phi_{rx,l},\phi_{tx,l}}w_l\right)^H\right\}\sigma^{-2}\frac{\partial}{\partial\iota_k}\left\{\left(\sum_{l=1}^{L}\mathbf{s}_{\tau_l,\phi_{rx,l},\phi_{tx,l}}w_l\right)\right\}\right\}\right] \tag{20}$$

$$\mathcal{I}\left(\sigma^2\right) = -\mathbb{E}_{\sigma^2}\left[\frac{\partial^2\ell\left(\mathbf{y};\sigma^2\right)}{\partial\left(\sigma^2\right)^2}\right]$$
$$= N\left(\sigma^2\right)^{-2} - \left(\mathbf{y} - \sum_{l=1}^{L}\mathbf{s}_{\tau_l,\phi_{rx,l},\phi_{tx,l}}w_l\right)^H\left(\sigma^2\right)^{-3}\left(\mathbf{y} - \sum_{l=1}^{L}\mathbf{s}_{\tau_l,\phi_{rx,l},\phi_{tx,l}}w_l\right) - \left(\sigma^2\right)^{-3}\sum_{l=1}^{L}\mathbf{s}^H_{\tau_l,\phi_{rx,l},\phi_{tx,l}}\mathbf{s}_{\tau_l,\phi_{rx,l},\phi_{tx,l}}\varsigma l \tag{21}$$

Therefore, $\mathcal{I}(\boldsymbol{\tau})$, $\mathcal{I}(\boldsymbol{\phi_{rx}})$ and $\mathcal{I}(\boldsymbol{\phi_{tx}})$ are computed by (3), (6) and (20).

At last, $\mathcal{I}(\sigma^2)$ is expressed in (21), as shown at the previous page.

APPENDIX C
DERIVATION OF NSE AND NSELB

First, the tracking or prediction error of CSI is given by $\mathbf{e}(t,f) = \tilde{\mathbf{g}}(t,f) - \mathbf{g}(t,f)$, where $\tilde{\mathbf{g}}(t,f)$ and $\mathbf{g}(t,f)$ represent the estimated CSI gotten by tracking or prediction and the real one gotten by channel model. Thus, the NSE of estimated CSI is expressed as $NSE = \frac{|\mathbf{e}(t,f)|_2^2}{|\mathbf{g}(t,f)|_2^2}$, where $|\cdot|_2$ denotes the Euclidean norm.

Then, according to [43], the NSE is lower bounded as

$$NSELB \geq \frac{Tr\left\{\mathbf{D}_g \mathbf{R} \mathbf{D}_g^H\right\}}{|\mathbf{g}(t,f)|_2^2}, \tag{24}$$

where $Tr\{\cdot\}$ denotes the trace of a matrix and $\mathbf{R}$ is calculated according to Appendix B. And $\mathbf{D}_g$ is given as

$$\mathbf{D}_g = [\frac{\partial \mathbf{g}(t,f)}{\partial \boldsymbol{\alpha}^T}, \frac{\partial \mathbf{g}(t,f)}{\partial \boldsymbol{\varphi}^T}, \frac{\partial \mathbf{g}(t,f)}{\partial \boldsymbol{\tau}^T}, \frac{\partial \mathbf{g}(t,f)}{\partial \boldsymbol{\phi}_{rx}^T}, \frac{\partial \mathbf{g}(t,f)}{\partial \boldsymbol{\phi}_{tx}^T}, \frac{\partial \mathbf{g}(t,f)}{\partial \sigma^2}]. \tag{25}$$

REFERENCES

[1] M. Shafi, A. F. Molisch, P. J. Smith, T. Haustein, P. Zhu, P. De Silva, F. Tufvesson, A. Benjebbour, and G. Wunder, "5G: A tutorial overview of standards, trials, challenges, deployment and practice," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 6, pp. 1201–1221, Jun. 2017.

[2] S. Chen and J. Zhao, "The requirements, challenges, and technologies for 5G of terrestrial mobile telecommunication," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 36–43, May 2014.

[3] A. Gupta and E. R. K. Jha, "A survey of 5G network: Architecture and emerging technologies," *IEEE Access*, vol. 3, pp. 1206–1232, Jul. 2015.

[4] Q. Ye, J. Li, K. Qu, W. Zhuang, X. S. Shen, and X. Li, "End-to-end quality of service in 5G networks: Examining the effectiveness of a network slicing framework," *IEEE Trans. Veh. Technol.*, vol. 13, no. 2, pp. 65–74, Jun. 2018.

[5] R. O. Adeogun, "Channel prediction for mobile MIMO wireless communication systems," Ph.D. dissertation, School Eng. Comput. Sci., Univ. Wellington, Wellington, New Zealand, 2015.

[6] A. Duel-Hallen, "Fading channel prediction for mobile radio adaptive transmission systems," *Proc. IEEE*, vol. 95, no. 12, pp. 2299–2313, Dec. 2007.

[7] C.-X. Wang, F. Haider, X. Gao, X.-H. You, Y. Yang, D. Yuan, H. M. Aggoune, H. Haas, S. Fletcher, and E. Hepsaydir, "Cellular architecture and key technologies for 5G wireless communication networks," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 122–130, Feb. 2014.

[8] K. Guan, B. Ai, B. Peng, D. He, G. Li, J. Yang, Z. Zhong, and T. Kürner, "Towards realistic high-speed train channels at 5G millimeter-wave band—Part I: Paradigm, significance analysis, and scenario reconstruction," *IEEE Trans. Veh. Technol.*, vol. 67, no. 10, pp. 9112–9128, Oct. 2018.

[9] T. Zhou, H. Li, Y. Wang, L. Liu, and C. Tao, "Channel modeling for future high-speed railway communication systems: A survey," *IEEE Access*, vol. 7, pp. 52818–52826, 2019.

[10] H. Jiang, Z. Zhang, J. Dang, and L. Wu, "A novel 3-D massive MIMO channel model for vehicle-to-vehicle communication environments," *IEEE Trans. Commun.*, vol. 66, no. 1, pp. 79–90, Jan. 2018.

[11] M. S. Kumar and S. Kirthiga, "Review of parametric radio channel prediction schemes for MIMO system," in *Proc. IEEE ICCPCT*, Mar. 2016, pp. 1–5.

[12] A. Heidari, A. K. Khandani, and D. Mcavoy, "Adaptive modelling and long-range prediction of mobile fading channels," *IET Commun.*, vol. 4, no. 1, pp. 39–50, 2010.

[13] A. Duel-Hallen, S. Hu, and H. Hallen, "Long-range prediction of fading signals: Enabling adaptive transmission for mobile radio channels," *IEEE Signal Process. Mag.*, vol. 17, no. 3, pp. 62–75, May 2000.

[14] T. Eyceoz, A. Duel-Hallen, and H. Hallen, "Deterministic channel modeling and long range prediction of fast fading mobile radio channels," *IEEE Commun. Lett.*, vol. 2, no. 9, pp. 254–256, Sep. 1998.

[15] A. M. Sayeed, A. Sendonaris, and B. Aazhang, "Multiuser detection in fast-fading multipath environments," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 9, pp. 1691–1701, Dec. 1998.

[16] T. Zemen, C. F. Mecklenbrauker, and B. H. Fleury, "Time-variant channel prediction using time-concentrated and band-limited sequences," in *Proc. IEEE ICC*, vol. 12, Jun. 2006, pp. 5660–5665.

[17] M. Patzold, *Mobile Fading Channels*. New York, NY, USA: Wiley, 2002.

[18] A. F. Molisch, H. Asplund, R. Heddergott, M. Steinbauer, and T. Zwick, "The COST259 directional channel model—Part I: Overview and methodology," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 3421–3433, Dec. 2006.

[19] L. Liu, C. Oestges, J. Poutanen, K. Haneda, P. Vainikainen, F. Quitin, F. Tufvesson, and P. De Doncker, "The COST 2100 MIMO channel model," *IEEE Wireless Commun.*, vol. 19, no. 6, pp. 92–99, Dec. 2012.

[20] J. B. Andersen, J. Jensen, S. H. Jensen, and F. Frederiksen, "Prediction of future fading based on past measurements," in *Proc. IEEE VTC*, vol. 1, Sep. 1999, pp. 151–155.

[21] Y. Wang, Y. Wang, S. Zhang, and H. Cen, "Channel tracking and transmission design in 5G large-scale MIMO system" *IEEE Access*, vol. 7, pp. 62032–62041, 2019.

[22] T. Jost, W. Wang, U.-C. Fiebig, and F. Pérez-Fontán, "Detection and tracking of mobile propagation channel paths," *IEEE Trans. Antennas Propag.*, vol. 60, no. 10, pp. 4875–4883, Oct. 2012.

[23] M. Mittermayer, "Detection and tracking of mobile channel impulse responses," M.S. thesis, Dept. Statist., Univ. Munich, Munich, Germany, Jul. 2014.

[24] W. Peng, M. Zou, and T. Jiang, "Channel prediction in time-varying massive MIMO environments," *IEEE Access*, vol. 5, pp. 23938–23946, 2017.

[25] C. Lv, J.-C. Lin, and Z. Yang, "Channel prediction for millimeter wave MIMO-OFDM communications in rapidly time-varying frequency-selective fading channels," *IEEE Access*, vol. 7, pp. 15183–15195, 2019.

[26] K. Guan, G. Li, D. He, L. Wang, B. Ai, R. He, Z. Zhong, L. Tian, and J. Dou, "Spatial consistency of dominant components between ray-tracing and stochastic modeling in 3GPP high-speed train scenarios," in *Proc. 11th Eur. Conf. Antennas Propag. (EUCAP)*, Mar. 2017, pp. 3182–3186.

[27] S. Ju and T. S. Rappaport, "Millimeter-wave extended NYUSIM channel model for spatial consistency," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2018, pp. 1–6.

[28] T. Jost, W. Wang, U.-C. Fiebig, and F. Perez-Fontan, "Movement of equivalent scatterers in geometry-based stochastic channel models," *IEEE Antennas Wireless Propag. Lett.*, vol. 11, pp. 555–558, May 2012.

[29] N. Czink, R. Tian, S. Wyne, F. Tufvesson, J.-P. Nuutinen, J. Ylitalo, E. Bonek, and A. F. Molisch, "Tracking time-variant cluster parameters in MIMO channel measurements," in *Proc. ChinaCOM*, Aug. 2007, pp. 1147–1151.

[30] C. Wang, J. Zhang, L. Tian, M. Liu, and Y. Wu, "The spatial evolution of clusters in massive MIMO mobile measurement at 3.5 GHz," in *Proc. IEEE VTC Spring*, Jun. 2017, pp. 1–6.

[31] *Guidelines for Evaluation of Radio Interface Technologies for IMT-Advanced*, document ITU-R M.2135-1, Dec. 2009.

[32] *Study on 3D Channel Model for LTE*, document TR 36.873, Version 12.2.0, 3GPP, Jun. 2015. [Online]. Available: http://www.3gpp.org/dynareport/36873.html

[33] (Jun. 2018). *Study on Channel Model for Frequencies From 0.5 to 100 GHz*. [Online]. Available: http://www.3gpp.org/DynaReport/38901.html

[34] B. H. Fleury, M. Tschudin, R. Heddergott, D. Dahlhaus, and K. Ingeman Pedersen, "Channel parameter estimation in mobile radio environments using the SAGE algorithm," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 3, pp. 434–450, Mar. 1999.

[35] J. Zhang, Y. Zhang, Y. Yu, R. Xu, Q. Zheng, and P. Zhang, "3-D MIMO: How much does it meet our expectations observed from channel measurements?" *IEEE J. Sel. Areas Commun.*, vol. 35, no. 8, pp. 1887–1903, Aug. 2017.
[36] J. Zhang, Z. Zheng, Y. Zhang, J. Xi, X. Zhao, and G. Gui, "3D MIMO for 5G NR: Several observations from 32 to massive 256 antennas based on channel measurement," *IEEE Commun. Mag.*, vol. 56, no. 3, pp. 62–70, Mar. 2018.
[37] J. Salmi, A. Richter, and V. Koivunen, "Detection and tracking of MIMO propagation path parameters using state-space approach," *IEEE Trans. Signal Process.*, vol. 57, no. 4, pp. 1538–1550, Apr. 2009.
[38] D. Shutin and B. H. Fleury, "Sparse variational Bayesian SAGE algorithm with application to the estimation of multipath wireless channels," *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 3609–3623, Aug. 2011.
[39] M. Steinbauer, H. Ozcelik, H. Hofstetter, and C. F. Mecklenbrauker, "How to quantify multipath separation," *IEICE Trans. Commun.*, vol. 85, no. 3, pp. 552–557, Mar. 2002.
[40] S. Jaeckel, L. Raschkowski, K. Börner, and L. Thiele, "Quadriga-quasi deterministic radio channel generator, user manual and documentation," Fraunhofer Heinrich Hertz Inst., Berlin, Germany, Tech. Rep. v2.0.0, 2017.
[41] T. Svantesson and A. L. Swindlehurst, "A performance bound for prediction of MIMO channels," *IEEE Trans. Signal Process.*, vol. 54, no. 2, pp. 520–529, Feb. 2006.
[42] E. L. Lehmann and G. Casella, *Theory of Point Estimation*. New York, NY, USA: Wiley, 1983.
[43] M. D. Larsen, A. L. Swindlehurst, and T. Svantesson, "Performance bounds for MIMO-OFDM channel estimation," *IEEE Trans. Signal Process.*, vol. 57, no. 5, pp. 1901–1916, May 2009.

YUXIANG ZHANG received the B.S. degree in electronic information engineering from the Dalian University of Technology, in 2014, where he is currently pursuing the Ph.D. degree in information and communication engineering with the Beijing University of Posts and Telecommunications. His current research interests include channel modeling, massive MIMO, millimeter wave communication, beamforming, and so on.

JIANHUA ZHANG received the B.S. degree from the North China University of Technology, in 1994, and the Ph.D. degree from the Beijing University of Posts and Telecommunications (BUPT), in 2003, where she is currently a Professor. She has published more than 70 journal articles and nearly 200 conference papers. Her research interests include artificial intelligence, data mining, especially in massive MIMO and millimeter wave channel modeling, channel emulator, OTA test, and so on. She received the China Communications Best Paper Award at 2016 and shared VTC 2015 Spring, JCN 2009 Best Paper Awards. She continuously contributed to channel model standards from ITU-R M.2135 to 3GPP 36.873,900/901 and she was the Drafting Group (DG) Chairwoman of ITU-R IMT-2020 Channel Model.

LI YU received the B.S. degree from Tianjin University, in 2011, and the M.S. degree from the Beijing University of Posts and Telecommunications, Beijing, China, in 2014, where he is currently pursuing the Ph.D. degree. His current research interests include channel modeling, channel prediction, 3D MIMO, machine learning, and so on.