IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received September 3, 2017, accepted October 10, 2017, date of publication November 1, 2017,
date of current version November 14, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2766091

Channel Prediction in Time-Varying Massive MIMO Environments

WEI PENG, (Senior Member, IEEE), MENG ZOU, AND TAO JIANG, (Senior Member, IEEE)
School of Electronics Information and Communications, Huazhong University of Science and Technology, Wuhan 430074, China

Corresponding author: Wei Peng (pengwei@hust.edu.cn)

This work was supported in part by the National Science Foundation of China under Grant 61301129 and Grant 61771214, the National Science Foundation for Distinguished Young Scholars of China under Grant 61325004 and Grant 61631015, the National High Technology Development 863 Program of China under Grant 2014AA01A704, and the Innovative Project of Shenzhen City in China under Grant JCYJ20170307171931096.

ABSTRACT The massive MIMO channel is characterized by non-stationarity and fast variation, thereby the channel state information obtained by traditional methods will be outdated and the system performance will be degraded. In this paper, we propose a channel prediction algorithm in massive MIMO environments. First, considering the channel characteristics, we propose a first-order Taylor expansion-based predictive channel modeling method. Then, a channel prediction algorithm consisting of the estimation stage and prediction stage is proposed and the interval of effective prediction (IEP) is derived. The performance of the proposed algorithm is testified by numerical simulations. It is shown that, within the IEP, a reliable channel prediction can be obtained with low computational complexity.

INDEX TERMS Massive MIMO, fast-varying, non-stationary, channel prediction, first-order Taylor expansion.

I. INTRODUCTION

Massive multiple-input multiple-output (MIMO) technology is well known for its numerous merits, such as high spectrum and energy efficiency, large capacity, and simple transceiver design [1]–[3]. In the meantime, massive MIMO faces challenges [4] including realistic system modeling, high-dimensional channel estimation, heterogeneous user scheduling, and so on. Especially, accurate channel state information (CSI) serves as the priori knowledge for techniques like precoding and coherent detection, thus, an effective channel estimation is of great importance for practical applications of the massive MIMO technology [4], [5].

Traditionally, channel estimation is obtained in a time division duplex (TDD) mode, where pilots are transmitted in the uplink transmission and the estimated CSI is then feedback in the reverse link [5], [6]. However, ultra-high bands including the millimeter and terahertz have been considered for wireless communications [7], [8]. In ultra-high bands, the channel coherent time is significantly reduced and becomes shorter than the pilot transmission time. As a result, the CSI acquisition based on uplink estimation and downlink feedback will provide outdated information, thus seriously degrade the system performance. In this circumstance, reliable channel prediction to forecast the channel variation is necessary.

Considering the channel prediction problem in massive MIMO environments, channel modeling should be put in the first place. Different from the traditional MIMO system with normal-size antennas, massive MIMO has large-size antenna arrays so that the distance between the transmitter and receiver may be shorter than the Rayleigh distance, which is defined as $2L^2/\lambda$, where L and λ are the antenna array dimension and the carrier wavelength, respectively [9]. As a result, the far field assumption [10] is not applicable and the near field modeling [11] should be used instead. In addition, due to the mobility of terminals together with limited but changing physical scatters, non-stationarity is also envisioned by the massive MIMO channel [9], [12].

Existing researches on channel prediction usually adopt the autoregressive (AR) predictive modeling [13], [14] or deterministic parameters based modeling [15]. Assuming a wide sense stationary channel, AR modeling based channel prediction calculates the channel state through its previous states and the temporal channel correlation has to be evaluated through training [14], [16]. However, in the non-stationary and fast-varying environment, the temporal correlation of channel coefficients are difficult to be obtained and the AR modeling based prediction is not applicable. Similarly, the deterministic parameters based modeling is applied under

2169-3536 © 2017 IEEE. Translations and content mining are permitted for academic research only.
Personal use is also permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

the assumption of quasi-static channel, where the number of propagation paths is invariant, thus is also not applicable. Therefore, more effective prediction modeling is necessary.

In this paper, we focus on the channel prediction problem in time-varying massive MIMO environments. Firstly, we propose a channel modeling method for prediction based on the first-order Taylor expansion (FIT). It is then verified that the proposed FIT modeling has a relatively small residue error. Next, a channel prediction algorithm consisting of estimation stage (E-stage) and prediction stage (P-stage) is proposed based on the FIT modeling. In addition, the interval of effective prediction (IEP) is analyzed, it is found that the IEP is inversely proportional to $f_{\max}^2 \cos^2\theta$, where $f_{\max}$ is the maximum Doppler frequency and θ is the angel between the waveform departure direction and the moving direction of the mobile terminal. The effectiveness of the proposed algorithm is verified by simulation results. It is shown that, within IEP, a reliable channel prediction can be obtained, and the computational complexity is only marginally increased compared with the traditional channel estimation, especially when the number of antennas is large.

The rest of the paper is organized as follows. Section II introduces the system model. In Section III, we first present the traditional AR prediction modeling, then the FIT modeling is proposed. The channel prediction algorithm based on the FIT modeling is proposed in Section IV. Numerical results are then presented in Section V. Finally, the paper is concluded in Section VI.

II. SYSTEM MODEL

A. SIGNAL REPRESENTATION

Considering a time-varying massive MIMO system, there are M_T and M_R antennas at the transmitter and receiver, respectively, where $M_R \geq M_T$ and uniform linear antenna arrays (ULAs) are used for simplicity. During one transmission consisting of N time slots ($N \geqslant M_T$), the transmitter sends an $M_T \times N$ dimensional matrix of data $\mathbf{X}$, and the matrix form of the base-band received signals is given as

$$\mathrm{Y} = \mathbf{HX} + \mathbf{Z}, \tag{1}$$

where $\mathbf{Y}$ is an $M_R \times N$ matrix of received signals, $\mathbf{Z} \in \mathbb{C}^{M_R \times N}$ is the matrix of independent and identically (i.i.d) distributed additive noise, its elements follow the complex Gaussian distribution with zero mean and variance σ_n^2. $\mathbf{H} \in \mathbb{C}^{M_R \times M_T}$ is the matrix of time-varying channel fading coefficients.

B. NON-STATIONARY & TIME-VARYING MASSIVE MIMO CHANNEL

In this paper, a three-dimension two-cylinder geometry-based stochastic channel model (3D-GBSM) [9], [17] is employed, as shown in Fig. 1. This channel model is selected since it has been proved to have close agreements with 3GPP channel measurements in terms of the statistical properties.

In Fig. 1, the transmitter represents the mobile station (MS) moving at a velocity $\vec{v}$, and the receiver denotes the immovable base station (BS). The scatters around the transmitter and receiver are equivalently represented by the scatters spreading on the surface of the two cylinders. In the 3D-GBSM, the channel coefficient is composed of the line-of-sight (LOS) and non-line-of-sight (NLOS) components, yielding a Rician distributed variable. Accordingly, the channel coefficient between the m^{th} $(m = 1, 2, \cdots, M_T)$ transmitting antenna and the q^{th} $(q = 1, 2, \cdots, M_R)$ receiving antenna at moment t is given as

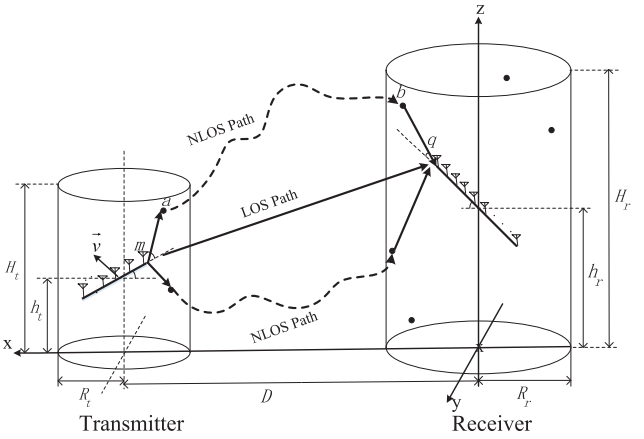

FIGURE 1. The 3D-GBSM model for massive MIMO channel.

$$h_{mq}(t) = h_{mq}^{\text{LOS}}(t) + \sum_{l=1}^{L} h_{mq}^{l}(t), \tag{2}$$

where $h_{mq}^{\text{LOS}}(t)$ is the channel fading on the LOS path, $h_{mq}^{l}(t)$ is the channel fading on the l^{th} NLOS path, and L is the number of the NLOS paths. Every NLOS path is represented by a double-bounced propagation [9], [17], i.e. the transmission will be reflected by the equivalent scatters around the transmitter and receiver, denoted by a and b, respectively, as shown in Fig. 1.

The LOS and NLOS fading coefficients can be given as

$$h_{mq}^{\text{LOS}}(t) = \sqrt{\frac{K}{K+1}} \exp\left\{ j\left(\psi_0 + \frac{2\pi}{\lambda} |\overrightarrow{mq}| + 2\pi f_{\max} t \frac{\langle \overrightarrow{mq} \cdot \vec{v} \rangle}{|\overrightarrow{mq}| \cdot |\vec{v}|} \right) \right\}, \tag{3}$$

and

$$h_{mq}^{l}(t) = \sqrt{P_l} \exp\left\{ j\left[\psi_0 + \frac{2\pi}{\lambda} \left(|\overrightarrow{ma}| + |\overrightarrow{bq}| + |ab| \right) + 2\pi f_{\max} t \frac{\langle \overrightarrow{ma} \cdot \vec{v} \rangle}{|\overrightarrow{ma}| \cdot |\vec{v}|} \right] \right\}, \tag{4}$$

respectively, where K is the Rician factor, $j = \sqrt{-1}$, ψ_0 is the initial phase of the waveform, λ is the carrier wavelength, and $f_{\max}$ denotes the maximum Doppler frequency. $P_l (l = 1, 2, \cdots, L)$ is the power of the l^{th} NLOS path where $\sum_{l=1}^{L} P_l = \frac{1}{K+1}$. $|\vec{v}|$ denotes the absolute velocity of the transmitter. $\overrightarrow{ma}$ represents the propagation from the

transmitting antenna to the equivalent scatter a. ab represents the propagation from a to the equivalent scatter b, where the random link $|ab| \geqslant \left|\overrightarrow{ab}\right|$. $\overrightarrow{bq}$ is the propagation from b to the receiving antenna. Besides, $\langle\cdot\rangle$ represents the inner product of two vectors, and $|\cdot|$ denotes the vector module.

III. MODELING FOR CHANNEL VARIATION

In this section, we first introduce the traditional AR modeling for the time-varying channel and reveal that it is not applicable in massive MIMO environments. Then we propose a FIT modeling, which will be used in the proposed channel prediction.

A. TRADITIONAL AR MODELING

For the time-varying but wide sense stationary channel coefficients, an AR process can be used to represent their variation [16], [18]–[21]. Namely, the channel between the m^{th} transmitting antenna and the q^{th} receiving antenna (hereafter, the antenna indexes m and q are omitted for simplicity) at moment t can be represented by its p previous states [13], [21], given as

$$h(t) = \sum_{n=1}^{p} d_{p,n}(t)h(t-j) + w_p(t), \tag{5}$$

where p is the AR order. $h(t-1), h(t-2), \cdots, h(t-p)$ are p previously obtained channel states, $d_{p,n}(t)$ represents the slowly-changing AR coefficients whose computation requires the knowledge of the time domain channel correlation [13], and $w_p(t)$ is the residual error. Generally speaking, a higher order AR model has a smaller residual error while the computation of the AR coefficients will be more complicated. Therefore, researchers usually adopt the first-order and second-order AR models, i.e., AR(1) and AR(2), given as

$$h(t) = d_{1,1}h(t-1) + w_1(t), \tag{6}$$

and

$$h(t) = d_{2,1}h(t-1) + d_{2,2}h(t-2) + w_2(t), \tag{7}$$

respectively. When considering the traditional MIMO channel in the stationary environment, AR coefficients can be calculated by the temporal correlations of $h(t)$. Specifically, following the Jakes model, AR(1) and AR(2) coefficients are obtained as [21]–[23]

$$d_{1,1} = J_0(2\pi f_{\max} T), \tag{8}$$

and

$$d_{2,1} = 2r_d \cos(2\pi f_{\max} T), \quad d_{2,2} = r_d^2, \tag{9}$$

respectively, where $J_0(\cdot)$ denotes the zeroth order Bessel function, T represents the symbol duration, r_d is the pole radius corresponding to the steepness of the power spectrum peaks.

Note that $f_{\max}$ is usually not available. In addition, the calculation of the correlation-dependant AR coefficients become very difficult in the non-stationary environments. Therefore, the AR modeling is not suitable for massive MIMO environments.

B. PROPOSED FIT MODELING

Assume that $h(t)$ is n-th order derivable, the Taylor expansion of $h(t)$ is given as

$$h(t) = \frac{h(t_0)}{0!} + \frac{1}{1!}\frac{\partial h(t_0)}{\partial t}(t-t_0) + \cdots + \frac{1}{n!}\frac{\partial^n h(t_0)}{\partial^n t}(t-t_0)^n + \cdots, \tag{10}$$

where t_0 is in the adjacent region of moment t, and $\partial^n h(t_0)/\partial t$ is the n-th order derivative. Since the high-order power function $(t-t_0)^n$ approximates to zero when $t \to t_0$, (10) can be simplified as a first-order Taylor expansion (FIT), given as

$$h(t) = h(t_0) + \frac{\partial h(t_0)}{\partial t}(t-t_0) + w_{\text{FIT}}(t), \tag{11}$$

where $w_{\text{FIT}}(t)$ is the residual error, given as

$$w_{\text{FIT}}(t) = \frac{1}{2!}\frac{\partial^2 h(t_0)}{\partial^2 t}(t-t_0)^2 + \frac{1}{3!}\frac{\partial^3 h(t_0)}{\partial^2 t}(t-t_0)^3 + \cdots + \frac{1}{n!}\frac{\partial^n h(t_0)}{\partial^2 t}(t-t_0)^n + \cdots. \tag{12}$$

In the following, the residual error of FIT modeling is compared with that of AR(1) and AR(2) modeling. A massive MIMO system with $M_T = 10$ and $M_R = 64$ is considered. The antennas at both the transmitter and receiver are separated by half wavelength. The distance between the transmitter and receiver is $D = 200m$. The radii of the cylinders on both transmitter and receiver sides are $R_r = R_t = 25m$. The heights of the antenna arrays are $h_t = 1.5$m and $h_r = 32$m, respectively. The initial phase ψ_0 is a random variable with uniform distribution over $[0, 2\pi)$. The velocity of the MS is $|\vec{v}| = 3m/s$ and the sampling period is $T_s = 66.7\mu s$. The equivalent scatters are uniformly distributed on the surfaces of the two cylinders with a density of $0.1/m^2$. The normalized mean square error (NMSE) is used to evaluate the accuracy of channel modeling, given as

$$\text{NMSE}(t) = \frac{E\left\{|w(t)|^2\right\}}{E\left\{|h(t)|^2\right\}} \tag{13}$$

where $E\{\cdot\}$ means the expectation, $w(t)$ represents $w_1(t)$, $w_2(t)$ and $w_{\text{FIT}}(t)$ for AR(1), AR(2) and FIT modeling methods, respectively. The simulation results are shown in Fig. 2.

It can be seen from Fig. 2 that the FIT modeling and the AR(2) modeling have the similar NMSE performance while the AR(1) modeling has a significant error. Note that, as aforementioned, the coefficients of AR(2) modeling is difficult to be obtained in practice. Thus we use the FIT modeling to predict the time-varying channel coefficients in massive MIMO environments.

IV. PROPOSED CHANNEL PREDICTION

In this section, we propose a channel prediction algorithm based on the FIT model. Assume that the first-order derivative $\partial h(t)/\partial t$ remains unchanged during a time internal, which is called the internal of effective prediction (IEP).

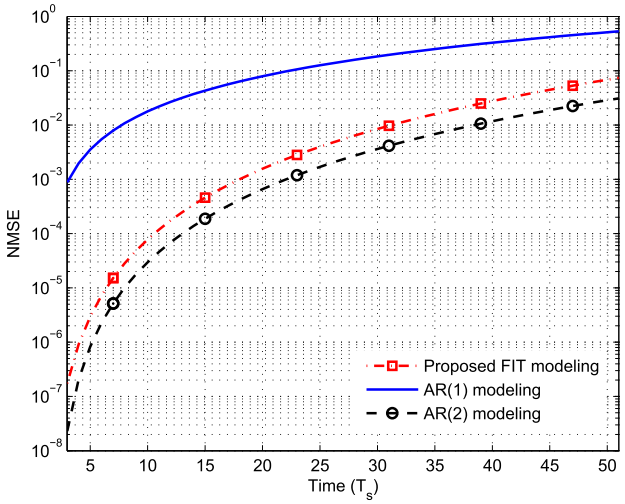

FIGURE 2. The NMSE of the proposed FIT modeling and AR(1)/AR(2) modeling.

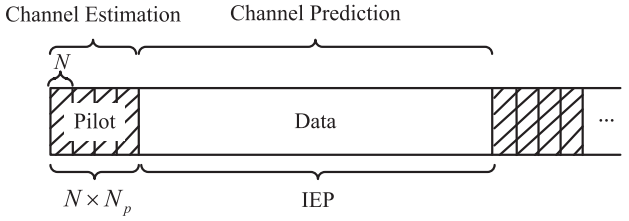

FIGURE 3. Structure of the transmitting sequence.

Within the IEP, the proposed channel prediction consists of two stages: E-stage and P-stage, as shown in Fig. 3. The transmitting sequence consists of pilot and data signals, where the $N \times N_p$-length pilot is adopted to produce N_p estimated channel states. Then, the estimated information will be used for channel prediction within the IEP. In the following, the channel estimation and prediction will be presented in detail, then the IEP will be analyzed.

A. CHANNEL ESTIMATION AND PREDICTION

Given (1), for the N-length pilot sequence $\mathbf{X}(t)$, where $t = t_0, t_1, \cdots, t_{N_p-1}$, the channel estimation can be obtained by a least square (LS) estimator, given as

$$\hat{\mathbf{H}}_{LS}(t) = \mathbf{Y}(t) \times \text{pinv}(\mathbf{X}(t)), \tag{14}$$

where pinv$(\cdot)$ denotes the pseudo-inverse operation. Note that, the LS estimator is selected for its simplicity. However, it can be straight forward replaced by the minimum mean square error (MMSE) estimator or any other channel estimators.

Denoting $\hat{h}_{\text{LS}}(t)$ as an element of $\hat{\mathbf{H}}_{\text{LS}}(t)$, according to the FIT model in (11), the prediction of $h(t)$ can be realized by determining $h(t_0)$ and $\partial h(t_0)/\partial t$. Let $h(t_0) = a_0$ and $\partial h(t_0)/\partial t = a_1$ for simplicity, the predicted channel coefficient at moment t is given as

$$\tilde{h}(t) = a_0 + a_1(t - t_0). \tag{15}$$

Given $\hat{h}_{\text{LS}}(t_0), \hat{h}_{\text{LS}}(t_1), \cdots, \hat{h}_{\text{LS}}(t_{N_p-1})$, a_0 and a_1 can be obtained by solving the minimization problem as

$$\langle \hat{a}_0, \hat{a}_1 \rangle = \arg\min_{a_0, a_1 \in \mathbb{R}} \sum_{i=0}^{N_p-1} \left\| \hat{h}_{\text{LS}}(t_i) - \tilde{h}(t_i) \right\|^2. \tag{16}$$

Employing the linear regression method [24], the problem in (16) is transformed into solving the linear equations, given as

$$\begin{bmatrix} \sum_{i=0}^{N_p-1} 1 & \sum_{i=0}^{N_p-1} g_i \\ \sum_{i=0}^{N_p-1} g_i & \sum_{i=0}^{N_p-1} g_i^2 \end{bmatrix} \begin{bmatrix} \hat{a}_0 \\ \hat{a}_1 \end{bmatrix} = \begin{bmatrix} \sum_{i=0}^{N_p-1} \hat{h}_{\text{LS}}(t_i) \\ \sum_{i=0}^{N_p-1} g_i \hat{h}_{\text{LS}}(t_i) \end{bmatrix}, \tag{17}$$

where $g_i = t_i - t_0$, and the solution to (17) is

$$\hat{a}_1 = \frac{N_p \sum_{i=0}^{N_p-1} g_i \hat{h}_{\text{LS}}(t_i) - \sum_{i=0}^{N_p-1} g_i \sum_{i=0}^{N_p-1} \hat{h}_{\text{LS}}(t_i)}{N_p \sum_{i=0}^{N_p-1} g_i^2 - \left(\sum_{i=0}^{N_p-1} g_i^2 \right)^2}, \tag{18}$$

and

$$\hat{a}_0 = \frac{\sum_{i=0}^{N_p-1} \hat{h}_{\text{LS}}(t_i) - \hat{a}_1 \sum_{i=0}^{N_p-1} g_i}{N_p}. \tag{19}$$

Substituting (18) and (19) into (15), we have

$$\begin{aligned} \tilde{h}(t) = & \frac{\sum_{i=0}^{N_p-1} \hat{h}_{\text{LS}}(t_i) - \hat{a}_1 \sum_{i=0}^{N_p-1} g_i}{N_p} \\ & + \frac{N_p \sum_{i=0}^{N_p-1} g_i \hat{h}_{\text{LS}}(t_i) - \sum_{i=0}^{N_p-1} g_i \sum_{i=0}^{N_p-1} \hat{h}_{\text{LS}}(t_i)}{N_p \sum_{i=0}^{N_p-1} g_i^2 - \left(\sum_{i=0}^{N_p-1} g_i^2 \right)^2} (t - t_0). \end{aligned} \tag{20}$$

In summary, the main steps to realize the proposed FIT model based channel prediction (FIT-CP) are listed in Algorithm 1.

Algorithm 1 FIT-CP Algorithm

Require: $\mathbf{Y}(t_i), \mathbf{X}(t_i), t_i = t_0, t_1, \cdots, t_{N_p-1}$

Ensure: $\tilde{\mathbf{H}}(t_n), t_n = t_{N_p}, t_{N_p+1}, \cdots$

Channel Estimation:

1: for $t_i = t_0, t_1, \cdots, t_{N_p-1}$

2: Obtain the LS channel estimation $\hat{h}_{\text{LS}}(t_i)$ following (14);

3: end for

Channel Prediction:

4: Obtain $\hat{a}_0, \hat{a}_1$ following (18), (19);

5: Predict the channel coefficients $\tilde{h}(t_n)$ following (20);

6: **return** $\tilde{\mathbf{H}}(t_n)$;

B. ANALYSIS ON THE INTERVAL OF EFFECTIVE PREDICTION

As aforementioned, the proposed FIT modeling has a small NMSE within the IEP, where $\partial h(t)/\partial t$ varies little and can be regarded as a constant. In this subsection, the length of the IEP is analyzed.

According to (2), $\partial h(t)/\partial t$ is given as

$$\frac{\partial h(t)}{\partial t} = \frac{\partial h^{\mathrm{LOS}}(t)}{\partial t} + \sum_{l=1}^{L} \frac{\partial h^{l}(t)}{\partial t}. \tag{21}$$

It is clear that (21) consists of the LOS and NLOS derivatives. Taking the LOS derivative as an example, it is calculated as

$$\frac{\partial h^{\mathrm{LOS}}(t)}{\partial t} = 2\pi j\sqrt{\frac{K}{K+1}}\left(\frac{1}{\lambda}\frac{\partial|\overline{mq}|}{\partial t} + f_{\max}\frac{\langle\overline{mq}\cdot\vec{v}\rangle}{|\overline{mq}|\cdot|\vec{v}|}\right) \times \exp\left\{j\left[\phi_0 + \frac{2\pi}{\lambda}|\overline{mq}| + 2\pi f_{\max} t\frac{\langle\overline{mq}\cdot\vec{v}\rangle}{|\overline{mq}|\cdot|\vec{v}|}\right]\right\}. \tag{22}$$

For mathematical convenience, the logarithmic form of (22) is used, given as

$$\ln\left(\frac{\partial h^{\mathrm{LOS}}(t)}{\partial t}\right) = \ln\left(2\pi j\sqrt{\frac{K}{K+1}}\right) + \ln\left(\frac{1}{\lambda}\frac{\partial|\overline{mq}|}{\partial t} + f_{\max}\frac{\langle\overline{mq}\cdot\vec{v}\rangle}{|\overline{mq}|\cdot|\vec{v}|}\right) + j\left[\phi_0 + \frac{2\pi}{\lambda}|\overline{mq}| + 2\pi f_{\max} t\frac{\langle\overline{mq}\cdot\vec{v}\rangle}{|\overline{mq}|\cdot|\vec{v}|}\right]. \tag{23}$$

The variation of $\ln\left(\partial h^{\mathrm{LOS}}(t)/\partial t\right)$ during time interval Δt is

$$\Delta\ln\left(\frac{\partial h^{\mathrm{LOS}}(t)}{\partial t}\right) = \ln\left(\frac{\partial h^{\mathrm{LOS}}(t+\Delta t)}{\partial t}\right) - \ln\left(\frac{\partial h^{\mathrm{LOS}}(t)}{\partial t}\right). \tag{24}$$

Substituting (23) into (24) and then get the absolute value of (24), we have (the detailed derivation can be found in the Appendix)

$$\left|\Delta\ln\left(\frac{\partial h^{\mathrm{LOS}}(t)}{\partial t}\right)\right| \approx \beta_0\Delta t, \tag{25}$$

where

$$\beta_0 = \left\{\frac{2\pi}{\lambda}|\vec{v}| + 2\pi f_{\max}\frac{\langle\overline{mq}\cdot\vec{v}\rangle}{|\overline{mq}|\cdot|\vec{v}|}\right\}. \tag{26}$$

Theorem 1: If $\ln x - \ln y = \varepsilon$ and $y \neq 0$, then

$$x - y = (e^{\varepsilon} - 1)y. \tag{27}$$

Proof: If $\ln x - \ln y = \varepsilon$, we can get

$$e^{\ln x - \ln y} = \frac{x}{y} = e^{\varepsilon}, \tag{28}$$

therefore $x - y = (e^{\varepsilon} - 1)y$. ■

According to (26) and ***Theorem 1***, we can get the variation of $\partial h(t)/\partial t$ during the time interval Δt, given as

$$\frac{\partial h^{\mathrm{LOS}}(t+\Delta t)}{\partial t} - \frac{\partial h^{\mathrm{LOS}}(t)}{\partial t} = \left[\exp(\beta_0\Delta t) - 1\right]\frac{\partial h^{\mathrm{LOS}}(t)}{\partial t}. \tag{29}$$

Substituting (22) into (29), the absolute value of the variation in (29) can be obtained as

$$\left|\frac{\partial h^{\mathrm{LOS}}(t+\Delta t)}{\partial t} - \frac{\partial h^{\mathrm{LOS}}(t)}{\partial t}\right| = \left|\left[\exp(\beta_0\Delta t) - 1\right]2\pi j\sqrt{\frac{K}{K+1}}\left(\frac{1}{\lambda}\frac{\partial|\overline{mq}|}{\partial t} + f_{\max}\frac{\langle\overline{mq}\cdot\vec{v}\rangle}{|\overline{mq}|\cdot|\vec{v}|}\right)\right|. \tag{30}$$

In the practical system, a threshold value δ_{LOS} can be used, namely, (22) is regarded as a constant when the variation of $\partial h(t)/\partial t$ is smaller than δ_{LOS}. Thus the length of IEP, denoted by T_{IEP}, is determined as

$$\left|\frac{\partial h^{\mathrm{LOS}}(t+T_{\mathrm{IEP}})}{\partial t} - \frac{\partial h^{\mathrm{LOS}}(t)}{\partial t}\right| = \delta_{\mathrm{LOS}} \tag{31}$$

Substituting (30) into (31), we have

$$T_{\mathrm{IEP}} = \frac{1}{\beta_0}\ln\left[1 + \frac{\delta_{\mathrm{LOS}}}{2\pi\sqrt{\frac{K}{K+1}}\left(\frac{1}{\lambda}\frac{\partial|\overline{mq}|}{\partial t} + f_{\max}\frac{\langle\overline{mq}\cdot\vec{v}\rangle}{|\overline{mq}|\cdot|\vec{v}|}\right)}\right]. \tag{32}$$

A closed-form expression of T_{IEP} can then be obtained by substituting (26) into (32), given as

$$T_{\mathrm{IEP}} = \ln\left[1 + \frac{\delta_{\mathrm{LOS}}}{2\pi\sqrt{\frac{K}{K+1}}\left(\frac{1}{\lambda}\frac{\partial|\overline{mq}|}{\partial t} + f_{\max}\frac{\langle\overline{mq}\cdot\vec{v}\rangle}{|\overline{mq}|\cdot|\vec{v}|}\right)}\right] \times\left(\frac{2\pi}{\lambda}|\vec{v}| + 2\pi f_{\max}\frac{\langle\overline{mq}\cdot\vec{v}\rangle}{|\overline{mq}|\cdot|\vec{v}|}\right)^{-1}. \tag{33}$$

For simplicity, we define

$$\cos\theta = \frac{\langle\overline{mq}\cdot\vec{v}\rangle}{|\overline{mq}|\cdot|\vec{v}|}, \tag{34}$$

where θ is the angle between the departure direction $\overline{mq}$ and the moving direction $\vec{v}$. Then, (33) can be simplified and we have

$$T_{\mathrm{IEP}} \propto \frac{\delta_{\mathrm{LOS}}}{f_{\max}^2\cos^2\theta}. \tag{35}$$

Now it is cleat that T_{IEP} is inversely proportional to $f_{\max}^2\cos^2\theta$.

The analysis on the NLOS component is similar to the LOS component, and the derivation is omitted for brevity.

V. NUMERICAL RESULTS

In this section, Monte Carlo simulations are carried out to testify the performance of the proposed channel prediction approach (referred to as FIT-CP). The traditional channel estimation without prediction [25] (referred to as traditional CE) and channel estimation based on interpolation (referred to as interpolation-based CE) are also simulated for comparison. Some representative simulation results are presented in terms of the prediction accuracy, the predictive trends and the computational complexity.

A. PREDICTION ACCURACY

To evaluate the accuracy of the proposed channel predictor, we use the MSE as an indicator, given as

$$\mathrm{MSE} = E\left\{\frac{1}{N_m}\sum_{t=N_p}^{N_p+N_m-1}\left\|\tilde{h}(t)-h(t)\right\|^2\right\}. \tag{36}$$

where N_m is the prediction length, $N_m + N_p \leq T_{\mathrm{IEP}}$.

Firstly, the MSE performance of the proposed FIT-CP is calculated as a function of the pilot ratio which is defined as $\rho = N_p/(N_p+N_m)$, and the results are shown in Fig. 4, where the signal to noise power ratio (SNR) is set as 10 dB.

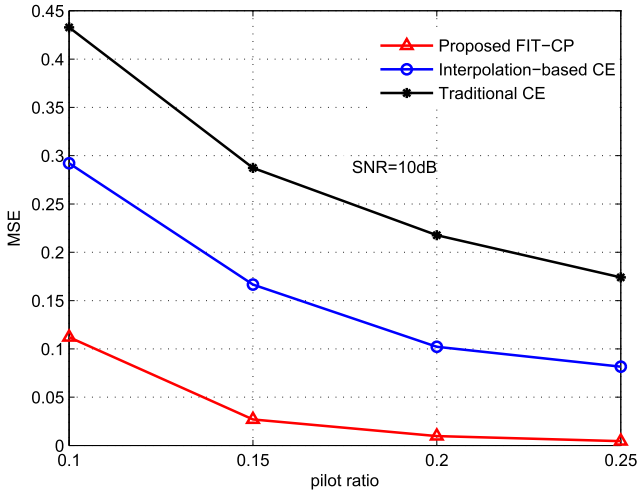

FIGURE 4. MSE versus pilot ratio.

It is observed that the MSE performance of the proposed FIT-CP algorithm improves as the pilot ratio increases. This is natural since that when longer pilot sequence is used, the estimation of a_0 and a_1 in (18) and (19) is more accurate, thus the channel prediction error is reduced. It is also observed that the MSE performance of the proposed FIT-CP is significantly better than the interpolation-based CE, while they both outperform the traditional CE. Therefore, the proposed channel prediction algorithm is effective in reducing the CSI acquisition error.

Then, the effect of the SNR on channel prediction is evaluated, and the results are shown in Fig. 5, where the pilot ratio is set as $\rho = 0.15$. By increasing the SNR, an MSE of 10^{-2} can be obtained by the proposed FIT-CP when

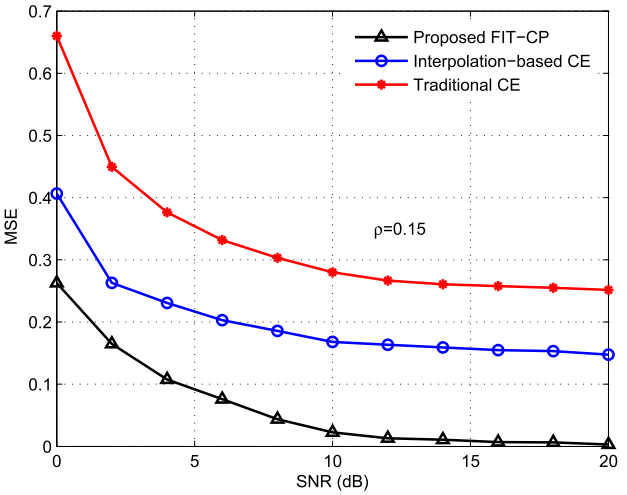

FIGURE 5. The MSE versus SNR.

SNR = 6 dB, where an error floor of MSE $\approx$ 0.15 exists for the interpolation-based CE. It should be noted that, the interpolation-based CE actually gives delayed CSI since it has to be carried out after the LS channel estimation in the next pilot transmission. Therefore, the proposed FIT-CP is superior in terms of both the accuracy and timeliness.

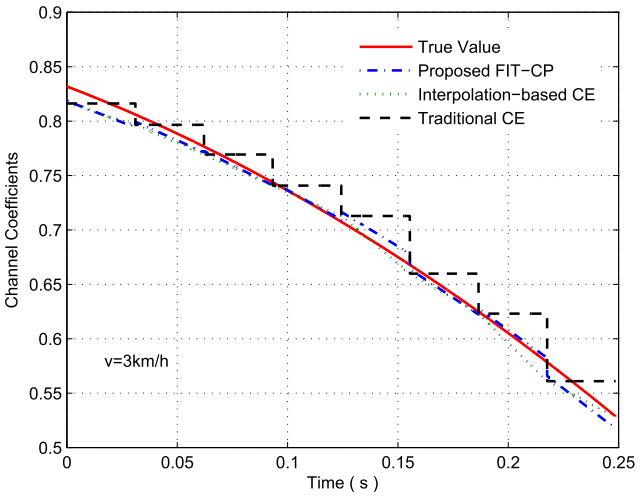

FIGURE 6. Prediction trends of the proposed FIT-CP in a short period.

B. PREDICTIVE TRENDS

To illustrate the predictive trends of the proposed FIT-CP, we change the channel variation by varying the speed of the MS. The short-time results are shown in Fig. 6, where $|\vec{v}| = 3$km/h is used, so that $f_{\max}T = 0.0013$. The long time results are shown in Fig. 7, where $|\vec{v}| = 10$ km/h and $f_{\max}T = 0.0043$.

It can be observed that, when the channel coefficient is monotonically increasing/decreasing in a short time, the proposed FIT-CP can trace the channel variation smoothly and yield a reliable channel prediction. However, when the channel coefficient undergoes significant fluctuations during a long period of time, the predicted channel coefficient

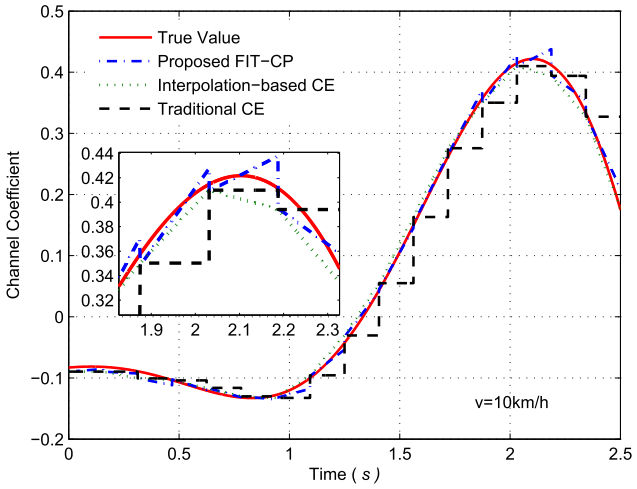

FIGURE 7. Prediction trends of the proposed FIT-CP in a long period.

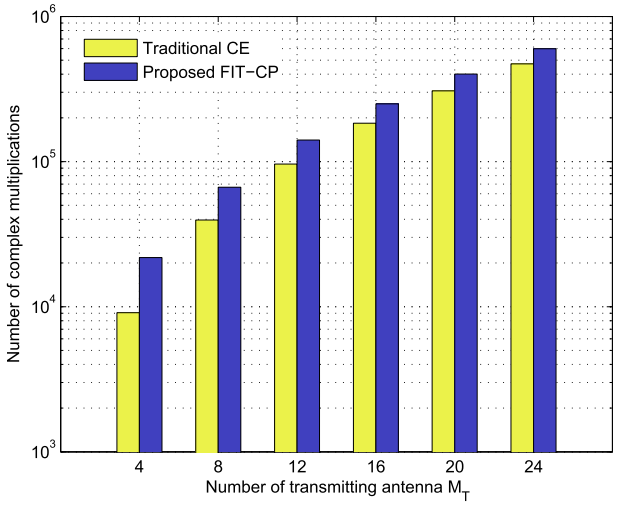

FIGURE 8. Computational complexity of the proposed channel predictor.

vibrates with flutter interference. Note that, it is also indicated by (35) that T_{IEP} reduces dramatically as the velocity increases, therefore, more sophisticated signal processing techniques should be employed to improve the prediction performance. E.g., the joint channel prediction and data detection can be a solution, where the detected data can be used as an extension of the pilot signal, thus the pilot ratio is equivalently increased and an improved prediction accuracy can be obtained.

C. COMPUTATIONAL COMPLEXITY

Last but not least, the computational complexity of the proposed channel predictor is evaluated. Note that, the proposed channel prediction consists of E-stage and P-stage, thus its computational complexity can be obtained as the sum complexity of these two stages.

Taking the number of complex multiplications as the measurement, and assuming that $N = M_T$, the computational complexity of the E-stage is dominated by the pseudo matrix inversion of $\mathbf{X}(t)$, which takes $\mathcal{O}\left(2M_T^3 + 2M_T^2 + M_T\right)$ complex multiplications [26], [27]. Thus, the computational complexity of one LS channel estimation is

$$C_{\text{LS}} = \mathcal{O}\left(2M_T^3 + (2 + M_R)M_T^2 + M_T\right). \tag{37}$$

For the P-stage, the computation complexity is dominated by the calculation of $\hat{a}_0$ and $\hat{a}_1$ using the linear regression method, given as

$$C_{\text{LR}} = \mathcal{O}\left((2N_p + N_m)M_R M_T\right). \tag{38}$$

Therefore, the computational complexity of the proposed FIT-CP is

$$C_{\text{FIT-CP}} = N_p C_{\text{LS}} + C_{\text{LR}}. \tag{39}$$

Substituting (37) and (38) into (39), we have

$$\begin{aligned} C_{\text{FIT-CP}} = \mathcal{O}\Big(&N_P\left(2M_T^3 + (2 + M_R)M_T^2 + M_T\right) \\ &+ \left(2N_p + N_m\right)M_R M_T\Big) \end{aligned} \tag{40}$$

For comparison, the computation complexity of the traditional CE is evaluated, given as

$$\begin{aligned} &C_{\text{Traditional CE}} \\ &= \mathcal{O}\left((1.5N_p + 0.5)M_T^3 + ((0.5 + M_R)N_p + 1.5)M_T^2 + M_T\right). \end{aligned} \tag{41}$$

According to (40) and (41), the computation complexities of the proposed FIT-CP and the traditional CE are plotted in Fig. 8, where $N_p = 8$ and $N_m = 32$. It is shown that, only marginal complexity increase is required by the proposed channel predictor compared to the traditional channel estimator. In addition, this additional complexity becomes negligible as the number of the antennas increases. Thus, the proposed channel predictor is computationally efficient.

VI. CONCLUSION

In this paper, the time-varying massive MIMO channel has been considered and a channel prediction algorithm has been proposed. Firstly, considering the non-stationary and fast-varying properties, FIT channel modeling has been proposed. Then, based on the FIT, the channel prediction approach is proposed, which consists of the E-stage and P-stage. The performance of the proposed channel predictor is testified by simulation results. It has been shown that, a reliable channel prediction can be obtained with low computational complexity. It has also been shown that, in a long period of time with the presence of high mobility, the proposed channel predictor may suffer from flutter interference, and more sophisticated signal processing techniques, such as the joint channel prediction and data detection, should be considered as subjects for the future research.

APPENDIX

This appendix presents the derivation of (25). According to (23), $\ln\left(\partial h^{\text{LOS}}(t)/\partial t\right)$ consists of three independent

components, given as

$$\ln\left(\frac{\partial h^{\mathrm{LOS}}(t)}{\partial t}\right)=\underbrace{\ln\left(2\pi j\sqrt{\frac{K}{K+1}}\right)}_{①}+\underbrace{\ln\left(\frac{1}{\lambda}\frac{\partial|\overrightarrow{mq}|}{\partial t}+f_{\max}\frac{\langle\overrightarrow{mq}\cdot\vec{v}\rangle}{|\overrightarrow{mq}|\cdot|\vec{v}|}\right)}_{②}+\underbrace{j\left[\phi_0+\frac{2\pi}{\lambda}|\overrightarrow{mq}|+2\pi f_{\max}t\frac{\langle\overrightarrow{mq}\cdot\vec{v}\rangle}{|\overrightarrow{mq}|\cdot|\vec{v}|}\right]}_{③}. \tag{A-1}$$

It is clear that component ① is a constant. In the following, components ② and ③ will be analyzed, respectively.

Denoting the initial coordinates of the transmitter and receiver as (x_0, y_0, z_0) and (x_1, y_1, z_1), respectively. The velocity of transmitter is $\vec{v} = (v_x, v_y, v_z)$ and $|\vec{v}| = \sqrt{v_x^2+v_y^2+v_z^2}$. Therefore,

$$\overrightarrow{mq}=\left(x_0-x_1+v_x\Delta t, y_0-y_1+v_y\Delta t, z_0-z_1+v_z\Delta t\right), \tag{A-2}$$

and

$$\langle\overrightarrow{mq}\cdot\vec{v}\rangle = v_x(x_0-x_1+v_x\Delta t)+v_y(y_0-y_1+v_y\Delta t)+v_z(z_0-z_1+v_z\Delta t). \tag{A-3}$$

According to (A-2) and (A-3), the derivative $\partial|\overrightarrow{mq}|/\partial t$ is rewritten as

$$\frac{\partial\,|\overrightarrow{mq}|}{\partial t}=\frac{\langle\overrightarrow{mq}\cdot\vec{v}\rangle}{|\overrightarrow{mq}|}. \tag{A-4}$$

Substituting (34) into (A-4), it is clear that $\partial|\overrightarrow{mq}|/\partial t=\cos\theta$. Therefore, the component ② can be rewritten as

$$\ln\left(\frac{1}{\lambda}\frac{\partial|\overrightarrow{mq}|}{\partial t}+f_{\max}\frac{\langle\overrightarrow{mq}\cdot\vec{v}\rangle}{|\overrightarrow{mq}|\cdot|\vec{v}|}\right)=\ln\left[\left(f_{\max}+\frac{|\vec{v}|}{\lambda}\right)\cos\theta\right]. \tag{A-5}$$

Assuming that the transmitting antenna moves to a new location m' at time $t+\Delta t$, and $\left\langle\overrightarrow{m'q}\cdot\vec{v}\right\rangle/\left(|\overrightarrow{m'q}|\cdot|\vec{v}|\right)=\cos\theta'$. Since $|\overrightarrow{mm'}|\ll D$, $\cos\theta\approx\cos\theta'$. Thus, the component ② can be treated as a constant. Therefore, the variation of $\ln\left(\partial h^{\mathrm{LOS}}(t+\Delta t)/\partial t\right)$ is determined by component ③, given as

$$\ln\left(\frac{\partial h^{\mathrm{LOS}}(t+\Delta t)}{\partial t}\right)-\ln\left(\frac{\partial h^{\mathrm{LOS}}(t)}{\partial t}\right)=j\left\{\frac{2\pi}{\lambda}(|\overrightarrow{m'q}|-|\overrightarrow{mq}|)+2\pi f_{\max}\Delta t\frac{\langle\overrightarrow{mq}\cdot\vec{v}\rangle}{|\overrightarrow{mq}|\cdot|\vec{v}|}\right\}. \tag{A-6}$$

Since $|\overrightarrow{m'q}|-|\overrightarrow{mq}|\le|\overrightarrow{m'm}|$, (A-6) is approximated as

$$\ln\left(\frac{\partial h^{\mathrm{LOS}}(t+\Delta t)}{\partial t}\right)-\ln\left(\frac{\partial h^{\mathrm{LOS}}(t)}{\partial t}\right)\approx j\left\{\frac{2\pi}{\lambda}|\overrightarrow{m'm}|+2\pi f_{\max}\Delta t\frac{\langle\overrightarrow{mq}\cdot\vec{v}\rangle}{|\overrightarrow{mq}|\cdot|\vec{v}|}\right\}. \tag{A-7}$$

Substituting $|\overrightarrow{m'm}|=|\vec{v}|\Delta t$ into (A-7), we have

$$\beta_0=\left\{\frac{2\pi}{\lambda}|\vec{v}|+2\pi f_{\max}\frac{\langle\overrightarrow{mq}\cdot\vec{v}\rangle}{|\overrightarrow{mq}|\cdot|\vec{v}|}\right\}. \tag{A-8}$$

Therefore, the derivation of (25) is completed.

REFERENCES

[1] F. Rusek *et al.*, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[2] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[3] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742–758, Oct. 2014.

[4] H. Xie, F. Gao, and S. Jin, "An overview of low-rank channel estimation for massive MIMO systems," *IEEE Access*, vol. 4, pp. 7313–7321, 2016.

[5] M. Biguesh and A. B. Gershman, "Training-based MIMO channel estimation: A study of estimator tradeoffs and optimal training signals," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 884–893, Mar. 2006.

[6] D. Hu, L. He, and X. Wang, "Semi-blind pilot decontamination for massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 1, pp. 525–536, Jan. 2016.

[7] A. Hirata and M. Yaita, "Ultrafast terahertz wireless communications technologies," *IEEE Trans. THz Sci. Technol.*, vol. 5, no. 6, pp. 1128–1132, Nov. 2015.

[8] X. Gao, L. Dai, Y. Zhang, T. Xie, X. Dai, and Z. Wang, "Fast channel tracking for terahertz beamspace massive MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 66, no. 7, pp. 5689–5696, Jul. 2017.

[9] S. Wu, C.-X. Wang, E.-H. M. Aggoune, M. M. Alwakeel, and Y. He, "A non-stationary 3-D wideband twin-cluster model for 5G massive MIMO channels," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1207–1218, Jun. 2014.

[10] A. D. Yaghjian, "Approximate formulas for the far field and gain of open-ended rectangular waveguide," *IEEE Trans. Antennas Propag.*, vol. AP-32, no. 4, pp. 378–384, Apr. 1984.

[11] A. D. Yaghjian, "An overview of near-field antenna measurements," *IEEE Trans. Antennas Propag.*, vol. AP-34, no. 1, pp. 30–45, Jan. 1986.

[12] S. Wu, C. X. Wang, H. Haas, E. H. M. Aggoune, M. M. Alwakeel, and B. Ai, "A non-stationary wideband channel model for massive MIMO communication systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 3, pp. 1434–1446, Mar. 2015.

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

[14] K. E. Baddour and N. C. Beaulieu, "Autoregressive modeling for fading channel simulation," *IEEE Trans. Wireless Commun.*, vol. 4, no. 4, pp. 1650–1662, Jul. 2005.

[15] L. Fan, Q. Wang, Y. Huang, and L. Yang, "Performance analysis of low-complexity channel prediction for uplink massive MIMO," *IET Commun.*, vol. 10, no. 14, pp. 1744–1751, Sep. 2016.

[16] S. M. Kay, *Modern Spectral Estimation*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1988.

[17] Y. Xie, B. Li, X. Zuo, M. Yang, and Z. Yan, "A 3D geometry-based stochastic model for 5G massive MIMO channels," in *Proc. IEEE 11th Int. Conf. Heterogeneous Netw. Quality, Rel., Secur. Robustness (QSHINE)*, Aug. 2015, pp. 216–222.

[18] B.-S. Chen, C.-Y. Yang, and W.-J. Liao, "Robust fast time-varying multipath fading channel estimation and equalization for MIMO-OFDM systems via a fuzzy method," *IEEE Trans. Veh. Technol.*, vol. 61, no. 4, pp. 1599–1609, May 2012.

[19] Z. Liu, X. Ma, and G. B. Giannakis, "Space-time coding and Kalman filtering for time-selective fading channels," *IEEE Trans. Commun.*, vol. 50, no. 2, pp. 183–186, Feb. 2002.
[20] E. Karami and M. Shiva, "Blind multi-inputŰmulti-output channel tracking using decision-directed maximum-likelihood estimation," *IEEE Trans. Veh. Technol.*, vol. 56, no. 3, pp. 1447–1454, May 2007.
[21] C. Komninakis, C. Fragouli, A. H. Sayed, and R. D. Wesel, "Multi-input multi-output fading channel tracking and equalization using Kalman estimation," *IEEE Trans. Signal Process.*, vol. 50, no. 5, pp. 1065–1076, May 2002.
[22] A. Mishra, R. Gayathri, and A. K. Jagannatham, "Random parameter EM-based Kalman filter (REKF) for joint symbol detection and channel estimation in fast fading STTC MIMO systems," *IEEE Signal Process. Lett.*, vol. 21, no. 6, pp. 766–770, Jun. 2014.
[23] P. H. Y. Wu and A. Duel-Hallen, "Multiuser detectors with disjoint Kalman channel estimators for synchronous CDMA mobile radio channels," *IEEE Trans. Commun.*, vol. 48, no. 5, pp. 752–756, May 2000.
[24] G. A. F. Seber and A. J. Lee, *Linear Regression Analysis*, vol. 936. Hoboken, NJ, USA: Wiley, 2012.
[25] Y. Chai, F. Hu, and L. Jin, "A novel time interpolation channel estimation for IEEE802.11ac system," in *Proc. 6th IEEE Int. Conf. Softw. Eng. Service Sci. (ICSESS)*, Sep. 2015, pp. 722–725.
[26] M. Arakawa, "Computational workloads for commonly used signal processing kernels," MIT, Cambridge, MA, USA, Tech. Rep. SPR-9, 2006.
[27] R. Hunger, *Floating Point Operations in Matrix-Vector Calculus*. Munich, Germany: Munich Univ. Tech., Inst. Circuit Theory Signal Processing, 2005.

WEI PENG (M'07–SM'11) received the Ph.D. degree from the Hong Kong University. She is currently an Associate Professor with the School of Electronics Information and Communications, Huazhong University of Science and Technology, Wuhan, China. She was an Assistant Professor with Tohoku University, Japan, from 2009 to 2013. Her research interests include wireless signal processing, network design and optimization.

MENG ZOU is currently pursuing the Master degree with the Huazhong University of Science and Technology.

TAO JIANG (M'06–SM'10) received the Ph.D. degree in information and communication engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2004. He is currently a Distinguished Professor with the School of Electronics Information and Communications, Huazhong University of Science and Technology, Wuhan, China. From 2004 to 2007, he was with some universities, such as Brunel University and the University of Michigan-Dearborn, respectively. He has authored or co-authored about 300 technical papers in major journals and conferences and 9 books/chapters in the areas of communications and networks. He served or is serving as an symposium technical program committee member of some major IEEE conferences, including INFOCOM, GLOBECOM, and ICC. He was invited to serve as the TPC Symposium Chair for the IEEE GLOBECOM 2013, the IEEEE WCNC 2013, and ICCC 2013. He served or is serving as an Associate Editor of some technical journals in communications, including the IEEE TRANSACTIONS ON SIGNAL PROCESSING, the IEEE Communications Surveys and Tutorials, the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, the IEEE INTERNET OF THINGS JOURNAL, and he is an Associate Editor-In-Chief of China Communications. He was a recipient of the NSFC for Distinguished Young Scholars Award in 2013. He received the Most Cited Chinese Researchers announced by Elsevier in 2014, 2015, and 2016.