

# Channel Acquisition for HF Skywave Massive MIMO-OFDM Communications

Ding Shi<sup>1</sup>, Graduate Student Member, IEEE, Linfeng Song<sup>1</sup>, Graduate Student Member, IEEE, Wenqi Zhou, Student Member, IEEE, Xiqi Gao<sup>1</sup>, Fellow, IEEE, Cheng-Xiang Wang<sup>1</sup>, Fellow, IEEE, and Geoffrey Ye Li<sup>2</sup>, Fellow, IEEE

**Abstract**—In this paper, we investigate channel acquisition for high frequency (HF) skywave massive multiple-input multiple-output (MIMO) communications with orthogonal frequency division multiplexing (OFDM) modulation. We first introduce the concept of triple beams (TBs) in the space-frequency-time (SFT) domain and establish a TB based channel model using sampled triple steering vectors. With the established channel model, we then investigate the optimal channel estimation and pilot design for pilot segments. Specifically, we find the conditions that allow pilot reuse among multiple user terminals (UTs), which significantly reduces pilot overhead and increases the number of UTs that can be served. Moreover, we propose a channel prediction method for data segments based on the estimated TB domain channel. To reduce the complexity, we formulate the channel estimation as a statistical inference problem and then obtain the channel by the proposed constrained Bethe free energy minimization (CBFEM) based channel estimation algorithm, which can be implemented with low complexity by exploiting the structure of the TB matrix together with the chirp z-transform (CZT). Simulation results demonstrate the superior performance of the proposed channel acquisition approach.

**Index Terms**—Massive MIMO-OFDM, HF skywave communications, pilot reuse, channel estimation, channel prediction.

## I. INTRODUCTION

HIGH frequency (HF) communications, whose frequency range is usually from 3 MHz to 30 MHz, can provide worldwide coverage through skywave propagation. Compared

Manuscript received 4 April 2022; revised 17 August 2022 and 23 October 2022; accepted 3 November 2022. Date of publication 1 December 2022; date of current version 12 June 2023. This work was supported by the National Key Research and Development Program of China under Grant 2018YFB1801103, the Jiangsu Province Basic Research Project under Grant BK20192002, the National Natural Science Foundation under Grant 61960206006, and the Key Research and Development Plan of Jiangsu Province (Industry Foresight and Key Core Technologies) under Grant BE2022067. An earlier version of this paper was presented in part at the IEEE GLOBECOM 2022 [1]. The associate editor coordinating the review of this article and approving it for publication was R. Dinis. (*Corresponding author: Xiqi Gao.*)

Ding Shi, Linfeng Song, Wenqi Zhou, Xiqi Gao, and Cheng-Xiang Wang are with the National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China, and also with the Purple Mountain Laboratories, Nanjing 211111, China (e-mail: shiding@seu.edu.cn; songlf@seu.edu.cn; wqzhou@seu.edu.cn; xqgao@seu.edu.cn; chxwang@seu.edu.cn).

Geoffrey Ye Li is with the Department of Electrical and Electronic Engineering, Imperial College London, SW7 2AZ London, U.K. (e-mail: geoffrey.li@imperial.ac.uk).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TWC.2022.3223100>.

Digital Object Identifier 10.1109/TWC.2022.3223100

with satellite communications, an alternative for global coverage, HF communications can be flexibly deployed with relatively low cost and are robust to jamming [2], [3]. However, the low data transmission rate of HF communications makes it hard to compete with satellite communications. Therefore, there has been some research on applying the multiple-input multiple-output (MIMO) technique to traditional point-to-point HF communications [4], [5], [6], where certain performance gains can be achieved.

In the past decade, massive MIMO has been widely studied for terrestrial cellular communications. It can significantly improve the spectrum and power efficiencies by deploying a large number of antennas at the base station (BS) and simultaneously serving a number of user terminals (UTs) in the same time-frequency resource [7]. Recently, it has been introduced into the HF skywave communications in [8], where a spatial-beam based wideband model for HF skywave massive MIMO channels within the orthogonal frequency division multiplexing (OFDM) transmission framework has been developed and the asymptotic achievable sum-rate is obtained when using the minimum mean-squared error (MMSE) based uplink (UL) receiver and downlink (DL) precoder with perfect channel state information (CSI) at the BS. It is demonstrated in [8] that massive MIMO can also vastly improve the performance of HF skywave communications.

The acquisition of the CSI is essential for massive MIMO to harvest the significant performance gain, which has been well investigated in terrestrial cellular massive MIMO systems [9], [10], [11], [12], [13], [14], [15], [16], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26], [27], [28], [29]. The conventional orthogonal pilot signaling and channel estimation methods, such as least squares (LS) and MMSE [9], are not suitable for the massive MIMO system due to the overwhelming pilot overhead and the forbidding computational complexity. By exploiting the constraint of the angle spread in the massive MIMO, the pilot reuse schemes have been proposed to reduce the pilot overhead, where UTs with non-overlapping angle domain channels are allowed to share the same pilot [10], [11], [12], [13]. For these pilot schemes, the pilot scheduling can be performed based on channel statistics [10], [11] or the spatial information acquired by the spatial rotation enhanced discrete Fourier transform (DFT) method [12], [13]. On the other hand, due to the limited local scatters, the terrestrial massive MIMO channel exhibits sparsity [30], which enables channel

estimation through sparse signal recovery. Thus sparse channel estimation in [31] has been proposed to reduce the pilot overhead. In [14], [15], [16], and [17], the greedy algorithms iteratively identify the sparse support and reconstruct the channel. Besides, statistical inference methods have also been developed for efficient sparse channel recovery. The expectation-maximization (EM) Gaussian-mixture approximate message passing (AMP) algorithm [18] and the turbo-orthogonal AMP algorithm [19] exploit the channel sparsity in the angle domain to estimate massive MIMO channels. For massive MIMO-OFDM communications, the sparse angle-delay domain channel can be obtained by EM-generalized AMP, EM-vector AMP algorithms [20] and the structured turbo compressive sensing (CS) algorithm [21]. Recently, constrained Bethe free energy minimization (CBFEM) unifies different message passing algorithms into a single optimization framework in [22]. CBFEM transforms the statistical inference problem into an optimization one with a clear objective function, where different constraints will lead to different solving algorithms, such as the expectation propagation (EP) variant [22], the hierarchical hybrid message passing [23], and the AMP with nearest neighbor sparsity pattern learning algorithm [24]. Moreover, channel estimation approaches highly depend on the channel models. The spatial-beam based channel model has been widely used in [10], [18], and [25] for channel estimation and also applied to the signal detection [32] and precoding [33]. An accurate spatial-beam based channel model using finely sampled steering vectors in [26], [27], [28], and [29] can improve channel estimation performance.

In this paper, we investigate channel acquisition for HF skywave massive MIMO-OFDM communications. We introduce the concept of the triple beams (TBs) in space-frequency-time (SFT) domain and use it to derive a TB based channel model. Based on the TB based channel model, we can design pilots and develop the channel estimation algorithm to accurately estimate the HF skywave massive MIMO-OFDM channels. The main contributions are summarized as follows.

- We derive a TB based channel model from the physical principle of HF skywave channels using sampled triple steering vectors in the SFT domain, each of which corresponds to a physical TB consisting of the spatial-beam, the frequency-beam, and the temporal-beam, pointing towards the sampled directional cosine, delay, and Doppler frequency, respectively.
- Based on the proposed channel model, we first investigate the optimal estimation of channels at pilot segments and find the conditions to minimize the normalized mean-squared error (NMSE) of channel estimates. Specifically, we show that UTs with overlapping TB domain channels should be allocated pilot sequences with different phase shift factors while UTs with non-overlapping TB domain channels can reuse the same pilot sequence. Furthermore, the pilot is designed, including UT grouping and pilot scheduling, and the channel prediction method for data segments is established based on the estimated TB domain channel.
- We formulate the channel estimation as a statistical inference problem, which can be solved within the

optimization framework of CBFEM. We then acquire the channel by the proposed CBFEM based channel estimation algorithm, which can be implemented with low complexity by exploiting the structure of the TB matrix together with the chirp z-transform (CZT).

The rest of the paper is organized as follows. In Section II, we derive the channel model for HF skywave massive MIMO-OFDM communications. In Section III, we investigate the optimal channel estimation for pilot segments and develop a pilot design. Moreover, a channel prediction method for data segments is presented. In Section IV, the CBFEM based channel estimation algorithm and its low-complexity implementation are presented. Section V provides simulation results and the paper is concluded in Section VI.

*Notations:* The uppercase (lowercase) boldface letters denote matrices (vectors). The superscripts  $(\cdot)^*$ ,  $(\cdot)^T$ ,  $(\cdot)^H$  and  $\text{tr}\{\cdot\}$  denote the conjugate, transpose, conjugated-transpose, and matrix trace operations, respectively.  $j = \sqrt{-1}$  denotes the imaginary unit.  $\|\mathbf{x}\|$  denotes the  $\ell_2$ -norm of  $\mathbf{x}$ , and  $\|\mathbf{X}\|_F$  denotes the Frobenius norm of  $\mathbf{X}$ .  $|\Psi|$  denotes the cardinality of set  $\Psi$ .  $\Psi_1 \times \Psi_2$  denotes the Cartesian product of sets  $\Psi_1$  and  $\Psi_2$ .  $\lceil x \rceil$  denotes the smallest integer that is not less than  $x$ , while  $\lfloor x \rfloor$  denotes the largest integer that is not greater than  $x$ .  $\circ$  and  $\otimes$  denote the Hadamard and the Kronecker product operators, respectively.  $[\mathbf{x}]_a$  and  $[\mathbf{x}]_{a:b}$  denote the  $a$ -th element and elements  $a$  to  $b$  of vector  $\mathbf{x}$ , respectively.  $[\mathbf{X}]_{a,b}$ ,  $[\mathbf{X}]_{:,b}$ , and  $[\mathbf{X}]_{a:b,c:d}$  denote the  $(a,b)$ -th element, the  $b$ -th column, and rows  $a$  to  $b$  and columns  $c$  to  $d$  of  $\mathbf{X}$ , respectively.  $\mathbb{E}\{\cdot\}$ ,  $\text{Var}\{\cdot\}$ ,  $\mathbb{H}\{\cdot\}$  and  $\mathbb{D}\{\cdot\}$  denote the statistical expectation, variance, entropy, and relative entropy, respectively.  $\text{diag}\{\mathbf{x}\}$  denotes the diagonal matrix with  $\mathbf{x}$  along its main diagonal, and  $\text{diag}\{\mathbf{X}_1, \dots, \mathbf{X}_N\}$  forms the block-diagonal matrix with  $\mathbf{X}_1, \dots, \mathbf{X}_N$ .  $\mathbf{0}$  denote the all-zero vector or matrix.  $\mathbf{I}_N$  denotes the identity matrix of dimension  $N$  while  $\mathbf{I}_{N \times G}$  denotes the matrix composed of the first  $G$  ( $\leq N$ ) columns of  $\mathbf{I}_N$ .  $\mathbf{X} \succ \mathbf{0}$  ( $\mathbf{X} \succeq \mathbf{0}$ ) denotes that  $\mathbf{X}$  is Hermitian positive definite (semi-definite).  $\mathcal{CN}(\mathbf{x}; \mathbf{a}, \mathbf{A})$  denotes the circular symmetric complex Gaussian probability density function with mean  $\mathbf{a}$  and covariance  $\mathbf{A}$ .

## II. HF SKYWAVE MASSIVE MIMO-OFDM CHANNEL MODEL

After introducing the configuration of the HF skywave massive MIMO-OFDM system in this section, we present the concept of the triple beams in the SFT domain and establish a triple-beam based channel model from the physical principle of HF skywave channels by using sampled triple steering vectors.

### A. System Configuration

We consider a HF skywave massive MIMO-OFDM system. The BS is equipped with a uniform linear array (ULA) with  $M$  antennas and serves  $U$  single-antenna UTs. Let  $N_c$  denote the number of subcarriers,  $N_g$  denotes the length of the cyclic prefix (CP),  $\Delta f$  denotes the subcarrier spacing, and  $T_s = \frac{1}{N_c \Delta f}$  denotes the sampling interval. We assume that  $N_v$  valid

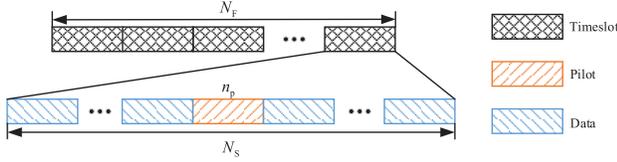


Fig. 1. The frame structure.

subcarriers are used to transmit data and pilots with the index set  $\mathcal{K} = \{k_0, k_1, \dots, k_{N_v-1}\}$ .

The HF skywave massive MIMO-OFDM system operates in time-division duplex (TDD) mode with frame structure shown in Fig. 1. Each frame consists of  $N_F$  timeslots and each timeslot consists of  $N_S$  OFDM symbols, thus the number of OFDM symbols in each frame is  $N = N_F N_S$ . In each timeslot, the  $n_p$ -th OFDM symbol is the pilot used for UL training and the other OFDM symbols are used for UL/DL data transmission.

In HF skywave communications, the carrier frequency  $f_c$  varies with the ionospheric conditions at different times. Therefore, we set inter-antenna spacing  $d$  according to the highest system operating frequency  $f_o$ , that is,  $d = \lambda_o/2$ , where  $\lambda_o = c/f_o$  is the wavelength, and  $c$  is the speed of light. This is different from the traditional massive MIMO communications, where the antenna spacing is usually half the wavelength of the carrier.

### B. Triple-Beam Based Channel Model

Let  $x_{u,n,k}$  denote the transmitted data of UT  $u$  on the  $n$ -th OFDM symbol at the  $k$ -th subcarrier, where  $n \in \{0, 1, \dots, N-1\}$  and  $k \in \mathcal{K}$ . After OFDM modulation, the analog baseband signal on the  $n$ -th OFDM symbol transmitted by UT  $u$  with the CP can be expressed as

$$\bar{x}_{u,n}(t) = \sum_{k=k_0}^{k_{N_v-1}} x_{u,n,k} e^{j2\pi k \Delta f t}, \quad -N_g T_s \leq t - n T_{\text{sym}} < N_c T_s, \quad (1)$$

where  $T_{\text{sym}} = (N_c + N_g) T_s$  is the time duration of an OFDM symbol including CP. At the BS, the received analog baseband signal on the  $n$ -th OFDM symbol at the  $m$ -th antenna can be expressed as

$$\bar{y}_{m,n}(t) = \sum_{u=0}^{U-1} \int_{-\infty}^{\infty} \bar{h}_{u,m}(t, \tau) \bar{x}_{u,n}(t - \tau) d\tau + \bar{z}_{m,n}(t), \quad (2)$$

where  $m \in \{0, 1, \dots, M-1\}$ ,  $\bar{z}_{m,n}(t)$  is the additive white Gaussian noise and  $\bar{h}_{u,m}(t, \tau)$  is the time-varying channel impulse response between UT  $u$  and the  $m$ -th antenna of the BS. The channel impulse response can be expressed as

$$\bar{h}_{u,m}(t, \tau) = \sum_{p=0}^{P_u-1} \gamma_{u,p} e^{j2\pi \nu_{u,p} t} e^{-j2\pi f_c m \Delta \tau \Omega_{u,p}} \times \delta(\tau - \tau_{u,p} - m \Delta \tau \Omega_{u,p}), \quad (3)$$

where  $P_u$  is the number of paths between UT  $u$  and the BS,  $\gamma_{u,p}$ ,  $\nu_{u,p}$ , and  $\Omega_{u,p}$  is the complex-valued gain, the Doppler frequency, and the directional cosine of the  $p$ -th path of UT  $u$ , respectively,  $\tau_{u,p}$  is the  $p$ -th propagation path delay between UT  $u$  and the first antenna of the BS, and  $\Delta \tau = d/c$ . In (3), the complex-valued gain can be expressed as  $\gamma_{u,p} = \beta_{u,p} e^{j\varphi_{u,p}}$ , where  $\beta_{u,p}$  and  $\varphi_{u,p}$  are the gain and initial phase, respectively, and  $\varphi_{u,p}$  is uniformly distributed over  $[0, 2\pi)$ . The directional cosine is defined as  $\Omega_{u,p} \triangleq \sin \theta_{u,p}^{\text{az}} \cos \theta_{u,p}^{\text{el}}$ , where  $\theta_{u,p}^{\text{az}}$  and  $\theta_{u,p}^{\text{el}}$  are the azimuth angle of arrival (AoA) and elevation AoA, respectively. Note that the propagation delay across the antenna array (i.e.,  $m \Delta \tau \Omega_{u,p}$  in the delta function) is considered due to the spatial-wideband effect caused by the equipped large-scale antenna array and wider bandwidth compared with traditional HF communications [8], [34].

We assume that the channel state keeps constant within an OFDM symbol and varies symbol by symbol due to the Doppler effect. After the OFDM demodulation, the received data at the  $k$ -th subcarrier on the  $n$ -th OFDM symbol at the  $m$ -th antenna is given by

$$y_{m,n,k} = \sum_{u=0}^{U-1} h_{u,m,n,k}^{\text{SFT}} x_{u,n,k} + z_{m,n,k}, \quad (4)$$

where  $z_{m,n,k}$  is the additive white Gaussian noise with distribution  $\mathcal{CN}(z_{m,n,k}; 0, \sigma_z^2)$ , and  $h_{u,m,n,k}^{\text{SFT}}$  is channel frequency response at the  $k$ -th subcarrier on the  $n$ -th OFDM symbol between UT  $u$  and the  $m$ -th antenna at the BS, which is given by

$$\begin{aligned} h_{u,m,n,k}^{\text{SFT}} &= \int \bar{h}_{u,m}(n T_{\text{sym}}, \tau) e^{-j2\pi k \Delta f \tau} d\tau \\ &= \sum_{p=0}^{P_u-1} \gamma_{u,p} e^{j2\pi \nu_{u,p} n T_{\text{sym}}} e^{-j2\pi f_c m \Delta \tau \Omega_{u,p}} \\ &\quad \times e^{-j2\pi k \Delta f \tau_{u,p}} e^{-j2\pi k \Delta f m \Delta \tau \Omega_{u,p}}. \end{aligned} \quad (5)$$

We consider the whole space-frequency domain channel over  $N$  OFDM symbols between UT  $u$  and the BS, which is referred to as the space-frequency-time (SFT) domain channel vector,  $\mathbf{h}_u^{\text{SFT}}$ , with element  $h_{u,m,n,k}^{\text{SFT}}$  of index  $(n M N_v + (k - k_0) M + m)$ . We denote

$$\mathbf{v}_k(\Omega) \triangleq \left[ 1, e^{-j2\pi(f_c + k \Delta f) \Delta \tau \Omega}, \dots, e^{-j2\pi(f_c + k \Delta f)(M-1) \Delta \tau \Omega} \right]^T \in \mathbb{C}^{M \times 1}, \quad (6)$$

$$\mathbf{u}(\tau) \triangleq \left[ e^{-j2\pi k_0 \Delta f \tau}, e^{-j2\pi k_1 \Delta f \tau}, \dots, e^{-j2\pi k_{N_v-1} \Delta f \tau} \right]^T \in \mathbb{C}^{N_v \times 1}, \quad (7)$$

$$\mathbf{d}(\nu) \triangleq \left[ 1, e^{j2\pi \nu T_{\text{sym}}}, \dots, e^{j2\pi \nu (N-1) T_{\text{sym}}} \right]^T \in \mathbb{C}^{N \times 1}, \quad (8)$$

as steering vectors in the space domain, the frequency domain, and the time domain, pointing towards directional cosine  $\Omega$ , delay  $\tau$ , and Doppler frequency  $\nu$ , respectively. Note that the steering vectors in the space domain are different for different subcarriers due to the spatial-wideband effect. Denote

$$\mathbf{p}(\Omega, \tau, \nu) \triangleq \mathbf{d}(\nu) \otimes (\mathbf{v}(\Omega) \circ (\mathbf{u}(\tau) \otimes \mathbf{e}_M)) \in \mathbb{C}^{M N_v N \times 1}, \quad (9)$$

where  $\mathbf{e}_M \triangleq [1, 1, \dots, 1]^T \in \mathbb{C}^{M \times 1}$ , and

$$\mathbf{v}(\Omega) \triangleq \left[ \mathbf{v}_{k_0}(\Omega)^T, \mathbf{v}_{k_1}(\Omega)^T, \dots, \mathbf{v}_{k_{N_v-1}}(\Omega)^T \right]^T \in \mathbb{C}^{MN_v \times 1}. \quad (10)$$

Therefore, the SFT channel can be expressed as

$$\mathbf{h}_u^{\text{SFT}} = \sum_{p=0}^{P_u-1} \gamma_{u,p} \mathbf{p}(\Omega_{u,p}, \tau_{u,p}, \nu_{u,p}) \in \mathbb{C}^{MN_v N \times 1}. \quad (11)$$

In this physical channel model,  $\mathbf{p}(\Omega_{u,p}, \tau_{u,p}, \nu_{u,p})$  represents a triple steering vector pointing towards the channel parameters  $(\Omega_{u,p}, \tau_{u,p}, \nu_{u,p})$ . Moreover, the complex-valued gain of a path,  $\gamma_{u,p}$ , is also the weight of the triple steering vector. Based on this physical channel model, we next derive a statistical channel model by sampling the triple steering vectors, which is referred to as the triple-beam based channel model and will be used in the channel acquisition.

We assume that all UTs are synchronized. Note that the parameters  $(\Omega_{u,p}, \tau_{u,p}, \nu_{u,p})$  of each path for each UT are limited within sets  $\mathcal{B}_{\text{an}} = \{\Omega | \Omega \in [-1, 1]\}$ ,  $\mathcal{B}_{\text{de}} = \{\tau | \tau \in [0, \tau_{\text{max}}]\}$  and  $\mathcal{B}_{\text{do}} = \{\nu | \nu \in [-\nu_{\text{max}}, \nu_{\text{max}}]\}$ , respectively, where  $\tau_{\text{max}} \triangleq \frac{N_g}{N_c \Delta f}$  is the maximum delay spread, and  $\nu_{\text{max}} \triangleq \frac{N_d}{2NT_{\text{sym}}}$  is the maximum Doppler [35], [36]. We uniformly divide these sets into multiple disjoint subsets as

$$\mathcal{B}_{\text{an}} = \bigcup_{n_{\text{an}}=0}^{N_{\text{an}}-1} \Lambda_{n_{\text{an}}}^{\text{an}}, \quad \Lambda_{n_{\text{an}}}^{\text{an}} = \left[ \frac{2n_{\text{an}} - N_{\text{an}}}{N_{\text{an}}}, \frac{2n_{\text{an}} - N_{\text{an}} + 2}{N_{\text{an}}} \right), \quad (12a)$$

$$\mathcal{B}_{\text{de}} = \bigcup_{n_{\text{de}}=0}^{N_{\text{de}}-1} \Lambda_{n_{\text{de}}}^{\text{de}}, \quad \Lambda_{n_{\text{de}}}^{\text{de}} = \left[ \frac{N_{\tau} n_{\text{de}}}{N_{\text{de}} N_v \Delta f}, \frac{N_{\tau} (n_{\text{de}} + 1)}{N_{\text{de}} N_v \Delta f} \right), \quad (12b)$$

$$\mathcal{B}_{\text{do}} = \bigcup_{n_{\text{do}}=0}^{N_{\text{do}}-1} \Lambda_{n_{\text{do}}}^{\text{do}}, \quad \Lambda_{n_{\text{do}}}^{\text{do}} = \left[ \frac{N_d (n_{\text{do}} - \frac{N_{\text{do}}}{2})}{N_{\text{do}} N T_{\text{sym}}}, \frac{N_d (n_{\text{do}} - \frac{N_{\text{do}}}{2} + 1)}{N_{\text{do}} N T_{\text{sym}}} \right), \quad (12c)$$

where  $N_{\tau} \triangleq \frac{N_v N_g}{N_c}$ , and  $\frac{2}{N_{\text{an}}}$ ,  $\frac{N_{\tau}}{N_{\text{de}} N_v \Delta f}$  and  $\frac{N_d}{N_{\text{do}} N T_{\text{sym}}}$  are the intervals in directional cosine, delay and Doppler frequency, which can be flexibly adjusted by properly selecting  $N_{\text{an}} = \left\lceil F_{\text{an}} M \frac{f_c}{f_o} \right\rceil$ ,  $N_{\text{de}} = F_{\text{de}} N_{\tau}$ , and  $N_{\text{do}} = F_{\text{do}} N_d$ , where  $F_{\text{an}}$ ,  $F_{\text{de}}$ , and  $F_{\text{do}}$  are referred to as fine factors.

Denote  $\Gamma_u \triangleq \{(\Omega_{u,0}, \tau_{u,0}, \nu_{u,0}), (\Omega_{u,1}, \tau_{u,1}, \nu_{u,1}), \dots, (\Omega_{u,P_u-1}, \tau_{u,P_u-1}, \nu_{u,P_u-1})\}$  as the path parameter set of UT  $u$ , and let  $\Lambda_{n_{\text{an}}, n_{\text{de}}, n_{\text{do}}} \triangleq \Lambda_{n_{\text{an}}}^{\text{an}} \times \Lambda_{n_{\text{de}}}^{\text{de}} \times \Lambda_{n_{\text{do}}}^{\text{do}}$ . The SFT domain channel in (11) can be rewritten as

$$\mathbf{h}_u^{\text{SFT}} = \sum_{n_{\text{an}}=0}^{N_{\text{an}}-1} \sum_{n_{\text{de}}=0}^{N_{\text{de}}-1} \sum_{n_{\text{do}}=0}^{N_{\text{do}}-1} \sum_{(\Omega_{u,p}, \tau_{u,p}, \nu_{u,p}) \in \Gamma_u \cap \Lambda_{n_{\text{an}}, n_{\text{de}}, n_{\text{do}}}} \gamma_{u,p} \times \mathbf{p}(\Omega_{u,p}, \tau_{u,p}, \nu_{u,p}). \quad (13)$$

We approximate the triple steering vector  $\mathbf{p}(\Omega_{u,p}, \tau_{u,p}, \nu_{u,p})$  for  $(\Omega_{u,p}, \tau_{u,p}, \nu_{u,p}) \in \Gamma_u \cap \Lambda_{n_{\text{an}}, n_{\text{de}}, n_{\text{do}}}$  as the sampled triple steering vector  $\mathbf{p}(\Omega_{n_{\text{an}}}, \tau_{n_{\text{de}}}, \nu_{n_{\text{do}}})$ , where  $\Omega_{n_{\text{an}}} = \frac{2n_{\text{an}} - N_{\text{an}}}{N_{\text{an}}}$ ,

<sup>1</sup>For fixed  $f_o$ , the angular resolution of the antenna array is proportional to  $f_c$ .

$\tau_{n_{\text{de}}} = \frac{N_{\tau} n_{\text{de}}}{N_{\text{de}} N_v \Delta f}$ , and  $\nu_{n_{\text{do}}} = \frac{N_d (n_{\text{do}} - N_{\text{do}}/2)}{N_{\text{do}} N T_{\text{sym}}}$ . This approximation tends to be accurate as the numbers of divided subsets  $N_{\text{an}}$ ,  $N_{\text{de}}$  and  $N_{\text{do}}$  tend to be relatively large. We then define a matrix  $\mathbf{P}$  that consists of sampled triple steering vectors and its  $(n_{\text{do}} N_{\text{an}} N_{\text{de}} + n_{\text{de}} N_{\text{an}} + n_{\text{an}})$ -th column is  $\mathbf{p}(\Omega_{n_{\text{an}}}, \tau_{n_{\text{de}}}, \nu_{n_{\text{do}}})$ , which represents the triple steering vector pointing towards the sampled directional cosine  $\Omega_{n_{\text{an}}}$ , delay  $\tau_{n_{\text{de}}}$  and Doppler frequency  $\nu_{n_{\text{do}}}$ . Moreover, according to (9),  $\mathbf{P}$  can be expressed as

$$\mathbf{P} \triangleq \mathbf{D} \otimes (\mathbf{V} (\mathbf{U} \otimes \mathbf{I}_{N_{\text{an}}})) \in \mathbb{C}^{MN_v N \times N_{\text{an}} N_{\text{de}} N_{\text{do}}}, \quad (14)$$

where  $\mathbf{D} \in \mathbb{C}^{N \times N_{\text{do}}}$  with  $[\mathbf{D}]_{i,j} = e^{j2\pi i \frac{N_d (j - N_{\text{do}}/2)}{N_{\text{do}} N}}$ ,  $\mathbf{U} \in \mathbb{C}^{N_v \times N_{\text{de}}}$  with  $[\mathbf{U}]_{i,j} = e^{-j2\pi (i+k_0) \frac{N_{\tau} j}{N_{\text{de}} N_v}}$ ,  $\mathbf{V} = \text{diag}\{\mathbf{V}_{k_0}, \dots, \mathbf{V}_{k_{N_v-1}}\}$ , and  $\mathbf{V}_k \in \mathbb{C}^{M \times N_{\text{an}}}$  with  $[\mathbf{V}_k]_{i,j} = e^{-j2\pi (f_c + k \Delta f) i \Delta \tau \frac{2j - N_{\text{an}}}{N_{\text{an}}}}$ . Specially, when the spatial-wideband effect is not considered,  $\mathbf{V}_k$  will be reduced to  $\tilde{\mathbf{V}} \in \mathbb{C}^{M \times N_{\text{an}}}$  with  $[\tilde{\mathbf{V}}]_{i,j} = e^{-j2\pi f_c i \Delta \tau \frac{2j - N_{\text{an}}}{N_{\text{an}}}}$ , which is independent of the subcarrier, and (14) can be rewritten as  $\mathbf{P} = \mathbf{D} \otimes \mathbf{U} \otimes \tilde{\mathbf{V}}$  in this case.

Then the SFT domain channel in (13) can be approximated as

$$\mathbf{h}_u^{\text{SFT}} = \mathbf{P} \mathbf{h}_u^{\text{TB}}, \quad (15)$$

where  $\mathbf{h}_u^{\text{TB}} \in \mathbb{C}^{N_{\text{an}} N_{\text{de}} N_{\text{do}} \times 1}$  can be expressed as

$$[\mathbf{h}_u^{\text{TB}}]_{n_{\text{do}} N_{\text{an}} N_{\text{de}} + n_{\text{de}} N_{\text{an}} + n_{\text{an}}} \triangleq \sum_{(\Omega_{u,p}, \tau_{u,p}, \nu_{u,p}) \in \Gamma_u \cap \Lambda_{n_{\text{an}}, n_{\text{de}}, n_{\text{do}}}} \gamma_{u,p}. \quad (16)$$

In the channel representation of (15), all the UTs share the same set of sampled triple steering vectors. Each sampled triple steering vector corresponds to a physical triple-beam in the SFT domain. Therefore, we refer to it as the *triple-beam (TB) based channel model*.  $\mathbf{h}_u^{\text{TB}}$  is the TB domain channel vector and  $\mathbf{P}$  is referred to as the TB matrix. Furthermore, the SFT domain channel covariance matrix of UT  $u$  can be given by

$$\begin{aligned} \mathbf{R}_u^{\text{SFT}} &= \mathbb{E} \left\{ \mathbf{h}_u^{\text{SFT}} (\mathbf{h}_u^{\text{SFT}})^H \right\} = \mathbf{P} \mathbb{E} \left\{ \mathbf{h}_u^{\text{TB}} (\mathbf{h}_u^{\text{TB}})^H \right\} \mathbf{P}^H \\ &= \mathbf{P} \mathbf{R}_u^{\text{TB}} \mathbf{P}^H \in \mathbb{C}^{MN_v N \times MN_v N}, \end{aligned} \quad (17)$$

where  $\mathbf{R}_u^{\text{TB}} \triangleq \mathbb{E} \left\{ \mathbf{h}_u^{\text{TB}} (\mathbf{h}_u^{\text{TB}})^H \right\} \in \mathbb{C}^{N_{\text{an}} N_{\text{de}} N_{\text{do}} \times N_{\text{an}} N_{\text{de}} N_{\text{do}}}$  is the TB domain channel covariance matrix of UT  $u$ , also referred to as the TB domain statistical CSI. We assume that TB domain channel coefficients follow independent complex Gaussian distributions with zero mean and different variances. Hence,  $\mathbf{R}_u^{\text{TB}}$  is a diagonal matrix with the  $(n_{\text{do}} N_{\text{an}} N_{\text{de}} + n_{\text{de}} N_{\text{an}} + n_{\text{an}})$ -th diagonal element expressed as  $\sum_{(\Omega_{u,p}, \tau_{u,p}, \nu_{u,p}) \in \Gamma_u \cap \Lambda_{n_{\text{an}}, n_{\text{de}}, n_{\text{do}}}} \beta_{u,p}^2$ . Due to the limited propagation path number, Doppler spread, and small angle spread, the TB domain channel typically exhibits sparsity in HF skywave massive MIMO-OFDM communication. Hence, compared with  $\mathbf{R}_u^{\text{SFT}}$ , the number of non-zero elements in  $\mathbf{R}_u^{\text{TB}}$  is substantially smaller since it is a diagonal matrix and most of the elements on the diagonal are approximately zero. Moreover,  $\mathbf{R}_u^{\text{TB}}$  varies slowly relative to the communication timescale [34], [36], [37]. Therefore, there are sufficient

resources to acquire  $\mathbf{R}_u^{\text{TB}}$ . For instance, the low-complexity method in [38] can estimate the spatial-frequency beam domain statistical CSI directly without involving the instantaneous channel, which can be straightforwardly extended to the TB based channels. Thus we assume that the TB domain statistical CSI  $\mathbf{R}_u^{\text{TB}}$  is available at the BS in the rest of the paper.

In our proposed TB based channel model, we generalize the concept of the spatial-beam in massive MIMO [8], [10] to the TB consisting of the spatial-beam, the frequency-beam and the temporal-beam. Moreover, when  $f_o = f_c$ ,  $N_{\text{an}} = M$ ,  $N_{\text{de}} = N_\tau$ ,  $N_{\text{do}} = N_d$ , the spatial-wideband effect is ignored, the columns of  $\mathbf{P}$  will become orthogonal. Besides, we can set  $N_{\text{an}} > M$ ,  $N_{\text{de}} > N_\tau$ , and  $N_{\text{do}} > N_d$  to make sampling intervals in (12) smaller and obtain a higher resolution of the angle, delay, and Doppler frequency, respectively. Although the orthogonality among the columns in  $\mathbf{P}$  no longer holds under such a setting, it makes the channel model more accurate and ensures more accurate channel acquisition. Moreover, the TB based channel model can also be applied in generic massive MIMO communications and wireless channels with different carrier frequencies, different propagation scenarios, and different system configurations although we focus on the HF skywave channels in this work.

### III. CHANNEL ACQUISITION AND PILOT DESIGN

In this section, we first investigate optimal channel estimation for pilot segments. Then we find the conditions to minimize the NMSE of the channel estimate and develop a pilot design, including UT grouping and pilot scheduling. Moreover, we propose a method to predict the channel at data segments based on the estimated TB domain channel.

#### A. Channel Estimation for Pilot Segment

We first consider the estimation of the SFT domain channel at pilot segments. According to (15) and the frame structure shown in Fig. 1, the SFT domain channel at pilot segments can be expressed as

$$\mathbf{h}_u^{\text{SFT,P}} = \tilde{\mathbf{P}}\mathbf{h}_u^{\text{TB}} \in \mathbb{C}^{MN_v N_F \times 1}, \quad (18)$$

where

$$\begin{aligned} \tilde{\mathbf{P}} &\triangleq (\mathbf{\Theta} \otimes \mathbf{I}_{N_v} \otimes \mathbf{I}_M) \mathbf{P} \\ &= \tilde{\mathbf{D}} \otimes (\mathbf{V} (\mathbf{U} \otimes \mathbf{I}_{N_{\text{an}}})) \in \mathbb{C}^{MN_v N_F \times N_{\text{an}} N_{\text{de}} N_{\text{do}}}, \end{aligned} \quad (19)$$

$\mathbf{\Theta} \in \mathbb{C}^{N_F \times N}$  with the  $n_F$ -th row being the  $(n_F N_S + n_p)$ -th row of  $\mathbf{I}_N$ , and  $\tilde{\mathbf{D}} = \mathbf{\Theta} \mathbf{D} \in \mathbb{C}^{N_F \times N_{\text{do}}}$ .

In each timeslot in Fig. 1, the channel estimation should be performed to obtain the CSI using the pilot. We combine the current timeslot with the previous  $N_F - 1$  timeslots to form a complete frame as in Fig. 1, where the current timeslot refers to the last timeslot in the whole frame. In this case, we need to save the received signals at pilot segments in the previous  $N_F - 1$  timeslots and perform channel estimation together with the received signal at the pilot segment in the current timeslot.

Let  $\mathbf{x}_u^{\text{p}} \in \mathbb{C}^{N_v \times 1}$  denote the transmitted pilot of UT  $u$  at valid subcarriers. The received signal  $\mathbf{y} \in \mathbb{C}^{MN_v N_F \times 1}$  at the

BS can be given by

$$\mathbf{y} = \sum_{u=0}^{U-1} \mathbf{X}_u \mathbf{h}_u^{\text{SFT,P}} + \mathbf{z} = \mathbf{X} \mathbf{h}^{\text{SFT,P}} + \mathbf{z}, \quad (20)$$

where  $\mathbf{h}^{\text{SFT,P}} \triangleq \left[ (\mathbf{h}_0^{\text{SFT,P}})^{\text{T}}, (\mathbf{h}_1^{\text{SFT,P}})^{\text{T}}, \dots, (\mathbf{h}_{U-1}^{\text{SFT,P}})^{\text{T}} \right]^{\text{T}} \in \mathbb{C}^{MN_v N_F U \times 1}$ ,  $\mathbf{X}_u \triangleq \mathbf{I}_{N_F} \otimes (\text{diag}\{\mathbf{x}_u^{\text{p}}\} \otimes \mathbf{I}_M) \in \mathbb{C}^{MN_v N_F \times MN_v N_F}$ ,  $\mathbf{X} \triangleq [\mathbf{X}_0, \mathbf{X}_1, \dots, \mathbf{X}_{U-1}] \in \mathbb{C}^{MN_v N_F \times MN_v N_F U}$ , and  $\mathbf{z} \in \mathbb{C}^{MN_v N_F \times 1}$  is the additive white Gaussian noise with distribution  $\mathcal{CN}(\mathbf{z}; \mathbf{0}, \sigma_z^2 \mathbf{I}_{MN_v N_F})$ . Substituting (18) into (20), we can obtain that

$$\mathbf{y} = \mathbf{X} (\mathbf{I}_U \otimes \tilde{\mathbf{P}}) \mathbf{h}^{\text{TB}} + \mathbf{z} = \mathbf{A} \mathbf{h}^{\text{TB}} + \mathbf{z}, \quad (21)$$

where  $\mathbf{h}^{\text{TB}} \triangleq \left[ (\mathbf{h}_0^{\text{TB}})^{\text{T}}, (\mathbf{h}_1^{\text{TB}})^{\text{T}}, \dots, (\mathbf{h}_{U-1}^{\text{TB}})^{\text{T}} \right]^{\text{T}} \in \mathbb{C}^{N_{\text{an}} N_{\text{de}} N_{\text{do}} U \times 1}$  and  $\mathbf{A} \triangleq \mathbf{X} (\mathbf{I}_U \otimes \tilde{\mathbf{P}})$ .

According to the relationship between the SFT domain channel and the TB domain channel shown in (18), the estimation of  $\mathbf{h}^{\text{SFT,P}}$  can be converted to the estimation of  $\mathbf{h}^{\text{TB}}$ . Since the TB domain channel  $\mathbf{h}^{\text{TB}}$  is distributed as the complex Gaussian distribution, the optimal estimate of  $\mathbf{h}^{\text{TB}}$  is the MMSE estimate, which is given by

$$\hat{\mathbf{h}}^{\text{TB}} = \mathbf{R}^{\text{TB}} \mathbf{A}^{\text{H}} (\mathbf{A} \mathbf{R}^{\text{TB}} \mathbf{A}^{\text{H}} + \sigma_z^2 \mathbf{I}_{MN_v N_F})^{-1} \mathbf{y}, \quad (22)$$

where  $\mathbf{R}^{\text{TB}} = \mathbb{E} \left\{ \mathbf{h}^{\text{TB}} (\mathbf{h}^{\text{TB}})^{\text{H}} \right\} = \text{diag} \{ \mathbf{R}_0^{\text{TB}}, \mathbf{R}_1^{\text{TB}}, \dots, \mathbf{R}_{U-1}^{\text{TB}} \}$ .

The estimate of  $\mathbf{h}^{\text{SFT,P}}$  can be expressed as

$$\begin{aligned} \hat{\mathbf{h}}^{\text{SFT,P}} &= (\mathbf{I}_U \otimes \tilde{\mathbf{P}}) \hat{\mathbf{h}}^{\text{TB}} \\ &= \mathbf{R}^{\text{SFT,P}} \mathbf{X}^{\text{H}} (\mathbf{X} \mathbf{R}^{\text{SFT,P}} \mathbf{X}^{\text{H}} + \sigma_z^2 \mathbf{I}_{MN_v N_F})^{-1} \mathbf{y}, \end{aligned} \quad (23)$$

where

$$\begin{aligned} \mathbf{R}^{\text{SFT,P}} &= (\mathbf{I}_U \otimes \tilde{\mathbf{P}}) \mathbf{R}^{\text{TB}} (\mathbf{I}_U \otimes \tilde{\mathbf{P}})^{\text{H}} \\ &= \mathbb{E} \left\{ \mathbf{h}^{\text{SFT,P}} (\mathbf{h}^{\text{SFT,P}})^{\text{H}} \right\} \\ &= \text{diag} \left\{ \mathbf{R}_0^{\text{SFT,P}}, \mathbf{R}_1^{\text{SFT,P}}, \dots, \mathbf{R}_{U-1}^{\text{SFT,P}} \right\}. \end{aligned} \quad (24)$$

Note that (23) is also the MMSE estimate of  $\mathbf{h}^{\text{SFT,P}}$ .

As an important measure for channel estimation, the NMSE of  $\hat{\mathbf{h}}^{\text{SFT,P}}$  is defined as

$$\begin{aligned} \text{NMSE} &\triangleq \frac{\mathbb{E} \left\{ \left\| \mathbf{h}^{\text{SFT,P}} - \hat{\mathbf{h}}^{\text{SFT,P}} \right\|^2 \right\}}{\mathbb{E} \left\{ \left\| \mathbf{h}^{\text{SFT,P}} \right\|^2 \right\}} \\ &= \sum_{u=0}^{U-1} \frac{1}{MN_v N_F U \vartheta_u} \text{tr} \left\{ \mathbf{R}_u^{\text{SFT,P}} - \mathbf{R}_u^{\text{SFT,P}} \mathbf{X}_u^{\text{H}} \mathbf{C}^{-1} \mathbf{X}_u \mathbf{R}_u^{\text{SFT,P}} \right\}, \end{aligned} \quad (25)$$

where  $\mathbf{C} \triangleq \sum_{u=0}^{U-1} \mathbf{X}_u \mathbf{R}_u^{\text{SFT,P}} \mathbf{X}_u^{\text{H}} + \sigma_z^2 \mathbf{I}_{MN_v N_F}$  and  $\vartheta_u \triangleq \sum_{p=0}^{P_u-1} \beta_{u,p}^2$  is the large-scale fading between the BS and UT  $u$ .

From (25), the NMSE of  $\hat{\mathbf{h}}^{\text{SFT,P}}$  can be regarded as the average of the NMSE of each UT. For UT  $u$ , the term  $\sum_{u=0}^{U-1} \mathbf{X}_u \mathbf{R}_u^{\text{SFT,P}} \mathbf{X}_u^{\text{H}}$  in  $\mathbf{C}$  implies that the NMSE of each UT

is related to not only its own statistical CSI but also that of other UTs, which is referred to as the inter-UT interference.

### B. Pilot Design

We next discuss the pilot design for the optimal channel estimation. Motivated by the widely used phase shift pilots [39], the transmitted pilot of UT  $u$  is given by

$$\mathbf{x}_u^P = \sigma_p \mathbf{x}_c \circ \begin{bmatrix} e^{-j2\pi k_0 \frac{N_\tau \phi_u}{N_{de} N_V}, e} & e^{-j2\pi k_1 \frac{N_\tau \phi_u}{N_{de} N_V}, \dots, e} & e^{-j2\pi k_{N_V-1} \frac{N_\tau \phi_u}{N_{de} N_V}} \end{bmatrix}^T, \quad (26)$$

where  $\sigma_p$  is the square root of the pilot transmit power,  $\phi_u \in \{0, N_{de}, \dots, (\lfloor N_V/N_\tau \rfloor - 1)N_{de}\}$  is referred to as the phase shift factor and  $\mathbf{x}_c$  is a sequence with unity modulus elements. A good choice for  $\mathbf{x}_c$  is the Zadoff-Chu (ZC) sequence [40], which has been extensively used in the fifth generation (5G) wireless networks [41]. Note that  $\tilde{\mathbf{P}}$  in (19) can be rewritten as

$$\tilde{\mathbf{P}} = \bar{\mathbf{P}} (\mathbf{I}_{N_{do}} \otimes \mathbf{I}_{\lfloor N_V/N_\tau \rfloor N_{de} \times N_{de}} \otimes \mathbf{I}_{N_{an}}), \quad (27)$$

where

$$\bar{\mathbf{P}} \triangleq \tilde{\mathbf{D}} \otimes (\mathbf{V}(\bar{\mathbf{U}} \otimes \mathbf{I}_{N_{an}})) \in \mathbb{C}^{M N_V N_F \times N_{an} \lfloor N_V/N_\tau \rfloor N_{de} N_{do}}, \quad (28)$$

and  $\bar{\mathbf{U}} \in \mathbb{C}^{N_V \times \lfloor N_V/N_\tau \rfloor N_{de}}$  with  $[\bar{\mathbf{U}}]_{i,j} = e^{-j2\pi(i+k_0)\frac{N_\tau j}{N_{de} N_V}}$ .

It can be checked that  $\mathbf{X}_u \tilde{\mathbf{P}} = \sigma_p \mathbf{X}_c \bar{\mathbf{P}} \mathbf{S}_u$ , where  $\mathbf{X}_c \triangleq \mathbf{I}_{N_F} \otimes (\text{diag}\{\mathbf{x}_c\} \otimes \mathbf{I}_M)$  and

$$\mathbf{S}_u \triangleq (\mathbf{I}_{N_{do}} \otimes [\mathbf{0}_{N_{de} \times \phi_u}, \mathbf{I}_{N_{de}}, \mathbf{0}_{N_{de} \times ((\lfloor N_V/N_\tau \rfloor - 1)N_{de} - \phi_u)}])^T \otimes \mathbf{I}_{N_{an}} \quad (29)$$

is the selection matrix depending on  $\phi_u$ . Thus  $\mathbf{A}$  can be rewritten as

$$\mathbf{A} = \sigma_p \mathbf{X}_c \bar{\mathbf{P}} [\mathbf{S}_0, \dots, \mathbf{S}_{U-1}]. \quad (30)$$

Next, we properly allocate all pilot sequences to each UT and optimize the channel estimation performance. To this end, we have the following theorem, proved in Appendix A, which provides the criterion for the UT grouping and pilot scheduling.

*Theorem 1:* When  $M, N_V, N_F \rightarrow \infty$ , the minimum value of NMSE is given by

$$\text{NMSE}_{\min} = \sum_{u=0}^{U-1} \frac{1}{M N_V N_F U \vartheta_u} \text{tr} \left\{ \mathbf{R}_u^{\text{SFT},P} - \mathbf{R}_u^{\text{SFT},P} \left( \mathbf{R}_u^{\text{SFT},P} + \frac{\sigma_z^2}{\sigma_p^2} \mathbf{I}_{M N_V N_F} \right)^{-1} \mathbf{R}_u^{\text{SFT},P} \right\}, \quad (31)$$

provided that, for  $\forall u, u'$  and  $u \neq u'$ , one of the following conditions is satisfied 1)

- 1)  $\phi_u \neq \phi_{u'}$ ,  $\mathbf{R}_u^{\text{TB}} \mathbf{R}_{u'}^{\text{TB}} \neq \mathbf{0}$ ,
- 2)  $\mathbf{R}_u^{\text{TB}} \mathbf{R}_{u'}^{\text{TB}} = \mathbf{0}$ .

From Theorem 1, when the NMSE achieves the minimum value, the NMSE of the channel estimate of each UT is only related to its own statistical CSI and is not related to other UTs, which means that the inter-UT interference is eliminated. Furthermore,  $\mathbf{R}_u^{\text{TB}} \mathbf{R}_{u'}^{\text{TB}} \neq \mathbf{0}$  means that there is an overlap between the TB domain channels of UTs  $u$  and  $u'$  while

$\mathbf{R}_u^{\text{TB}} \mathbf{R}_{u'}^{\text{TB}} = \mathbf{0}$  implies that the TB domain channels of UTs  $u$  and  $u'$  are non-overlapping. Therefore, condition 1) implies that to minimize the NMSE, UTs with overlapping TB domain channels should be allocated pilot sequences with different phase shift factors while condition 2) implies that UTs with non-overlapping TB domain channels can reuse the same pilot sequence. Such a result is intuitive because we need pilot sequences with different pilot shift factors to separate UTs with overlapping TB domain channels along the delay-dimension at the BS while UTs with non-overlapping TB domain channels can be directly separated at the BS, thus the same pilot sequence can be reused to reduce pilot overhead and increases the number of UTs that can be served. The above result is also consistent with findings in conventional massive MIMO literature [10], [12].

Next, we define the degree of channel overlap between UTs  $u$  and  $u'$  as

$$\rho_{u,u'} \triangleq \frac{\text{tr} \{ \mathbf{R}_u^{\text{TB}} \mathbf{R}_{u'}^{\text{TB}} \}}{\| \mathbf{R}_u^{\text{TB}} \|_F \| \mathbf{R}_{u'}^{\text{TB}} \|_F}. \quad (32)$$

It can be directly checked that  $0 \leq \rho_{u,u'} \leq 1$ . In practical systems, the conditions in Theorem 1 are hard to be well satisfied due to limited  $M$ ,  $N_V$ , and  $N_F$ . However, the power of the TB domain channel of each UT is typically concentrated on a narrow support and the channel overlap degree among UTs with different supports of TB domain channel can be small. Therefore, we can divide all UTs into  $S$  groups ( $S \leq \lfloor N_V/N_\tau \rfloor$ ) according to the following UT grouping criterion.

- i). The degree of channel overlap between any two UTs in the same group should be as low as possible.
- ii). UTs with a high degree of channel overlap should be allocated to different groups.

Meanwhile, we allocate pilot sequences with different phase shift factors to each UT group. With UT grouping and pilot scheduling, the inter-UT interference can be suppressed and the channel estimation performance can be improved.

Motivated by the hierarchical clustering [42], we propose a UT grouping algorithm shown in Algorithm 1 according to the UT grouping criterion. Specifically, we first place each UT in a group of its own. Then in each iteration, we search for a pair of UT groups, where the average channel overlap degree between UTs in these two groups is the smallest, and combine them into one UT group. The iteration is terminated when the current number of groups is equal to  $S$ .

Note that the TB domain statistical CSI includes three dimensions, i.e., the spatial-beam dimension, the frequency-beam dimension, and the temporal-beam dimension. In the above discussion, all these three dimensions are utilized for UT grouping and the computational complexity of  $\rho_{u,u'}$  in (32) is  $\mathcal{O}(N_{an} N_{de} N_{do})$ . Typically, the angle spread of the HF skywave channel is small [43], [44], [45], which implies that the TB domain channel of many UTs can be non-overlapping along the spatial-beam dimension, and we can just use the spatial-beam dimension of the TB domain statistical CSI to perform UT grouping. To this end, we first define

$$\mathbf{R}_u^{\text{B}} \triangleq \sum_{i=0}^{N_{de} N_{do} - 1} \sum_{j=0}^{N_{de} N_{do} - 1} [\mathbf{R}_u^{\text{TB}}]_{\substack{i N_{an} : (i+1) N_{an} - 1 \\ j N_{an} : (j+1) N_{an} - 1}} \in \mathbb{C}^{N_{an} \times N_{an}}. \quad (33)$$

**Algorithm 1** UT Grouping Algorithm

---

**Input:** The number of UTs  $U$ ; the number of groups to divide  $S$ ; the degree of overlap between UTs  $\{\rho_{u,u'}, u, u' = 0, 1, \dots, U-1\}$

**Initialize:** The index set of UT groups  $\Psi = \{0, 1, \dots, U-1\}$ ; the initial UT grouping result  $\Upsilon_s = \{s\}, s \in \Psi$

**while**  $|\Psi| > S$  **do**

$\{s_1, s_2\} = \arg \min_{s \in \Psi, s' \in \Psi \setminus s} \sum_{u \in \Upsilon_s, u' \in \Upsilon_{s'}} \frac{\rho_{u,u'}}{|\Upsilon_s| |\Upsilon_{s'}|}$

$\Upsilon_{s_1} \leftarrow \Upsilon_{s_1} \cup \Upsilon_{s_2}$

$\Psi \leftarrow \Psi \setminus s_2$

**end**

**Output:** UT grouping result  $\{\Upsilon_s, s \in \Psi\}$

---

Then we have the following theorem, proved in Appendix B.

**Theorem 2:** For arbitrary  $N_v$  and  $N_F$ , when  $M \rightarrow \infty$ , the NMSE in (25) can be rewritten as

$$\text{NMSE} = \sum_{u=0}^{U-1} \frac{1}{MN_v N_F U \vartheta_u} \text{tr} \left\{ \mathbf{R}_u^{\text{SFT}, \text{P}} - \mathbf{R}_u^{\text{SFT}, \text{P}} \mathbf{X}_u^{\text{H}} \bar{\mathbf{C}}_u^{-1} \mathbf{X}_u \mathbf{R}_u^{\text{SFT}, \text{P}} \right\}, \quad (34)$$

where

$$\bar{\mathbf{C}}_u \triangleq \mathbf{X}_u \mathbf{R}_u^{\text{SFT}, \text{P}} \mathbf{X}_u^{\text{H}} + \sum_{u' \in \mathcal{I}_u} \mathbf{X}_{u'} \mathbf{R}_{u'}^{\text{SFT}, \text{P}} \mathbf{X}_{u'}^{\text{H}} + \sigma_z^2 \mathbf{I}_{MN_v N_F}, \quad (35)$$

and  $\mathcal{I}_u \triangleq \{u' | u' \neq u, \mathbf{R}_{u'}^{\text{B}}, \mathbf{R}_u^{\text{B}} \neq \mathbf{0}\}$ .

By comparing  $\bar{\mathbf{C}}_u$  and  $\mathbf{C}$ , the NMSE of each UT is only related to the statistical CSI of its own and UTs with overlapping TB domain channels along the spatial-beam dimension when only  $M \rightarrow \infty$ . Therefore, when the TB domain channels of UTs  $u$  and  $u'$  are non-overlapping along the spatial-beam dimension, i.e.,  $u' \notin \mathcal{I}_u$ , they can reuse the same pilot sequence without affecting the NMSE performance.

We define channel overlap degree in the spatial-beam dimension between UTs  $u$  and  $u'$  as

$$\tilde{\rho}_{u,u'} \triangleq \frac{\text{tr} \left\{ \mathbf{R}_u^{\text{B}} \mathbf{R}_{u'}^{\text{B}} \right\}}{\|\mathbf{R}_u^{\text{B}}\|_{\text{F}} \|\mathbf{R}_{u'}^{\text{B}}\|_{\text{F}}}. \quad (36)$$

It is worth noting that when we only utilize the spatial-beam dimension of TB domain statistical CSI to perform UT grouping, the computational complexity of  $\tilde{\rho}_{u,u'}$  is  $\mathcal{O}(N_{\text{an}})$ , which is much smaller than that of  $\rho_{u,u'}$ . Moreover, UTs with overlapping TB domain channels along the spatial-beam dimension can still be allocated pilot sequences with different phase shift factors to suppress the inter-UT interference. The corresponding UT grouping and pilot scheduling method are similar to the case of using the whole TB domain statistical CSI, in which only  $\rho_{u,u'}$  needs to be replaced by  $\tilde{\rho}_{u,u'}$ .

### C. Channel Prediction for Data Segment

With the channel estimates for pilot segments, we can directly use it for DL transmit design or UL signal detection in the slow fading channel, where the channel changing over one

time slot can be ignored. On the other hand, in the fast fading channel, it is not appropriate to directly apply the channel at pilot segments to data segments [46].

If the channel varies symbol by symbol, we can try to predict the channel at the data segment using the relationship between the SFT domain channel and the TB domain channel. Specifically, according to (15), the estimate of the SFT domain channel of the  $N_F$  timeslots for all UTs can be given by

$$\hat{\mathbf{h}}^{\text{SFT}} = (\mathbf{I}_U \otimes \mathbf{P}) \hat{\mathbf{h}}^{\text{TB}}, \quad (37)$$

where  $\hat{\mathbf{h}}^{\text{SFT}} \triangleq \left[ \left( \hat{\mathbf{h}}_0^{\text{SFT}} \right)^{\text{T}}, \left( \hat{\mathbf{h}}_1^{\text{SFT}} \right)^{\text{T}}, \dots, \left( \hat{\mathbf{h}}_{U-1}^{\text{SFT}} \right)^{\text{T}} \right]^{\text{T}} \in \mathbb{C}^{MN_v NU \times 1}$ , and  $\hat{\mathbf{h}}_u^{\text{SFT}} \in \mathbb{C}^{MN_v N \times 1}$  is the estimate of the SFT domain channel of the whole frame for UT  $u$ . Note that the channel of the data segment to be predicted is in the current timeslot and has been contained in  $\hat{\mathbf{h}}^{\text{SFT}}$ . Hence, the predicted space-frequency domain channel corresponding to the  $n_S$ -th OFDM symbol of UT  $u$  in the current timeslot is given by

$$\hat{\mathbf{h}}_{u, n_S}^{\text{SF}} = \left[ \hat{\mathbf{h}}_u^{\text{SFT}} \right]_{\substack{((N_F-1)N_S+n_S)MN_v: \\ ((N_F-1)N_S+n_S+1)MN_v-1}} \in \mathbb{C}^{MN_v \times 1}, \quad (38)$$

where  $n_S = 0, 1, \dots, N_S - 1$ .

In summary, to acquire the CSI of the current timeslot, we first combine the current timeslot and the previous timeslots into a complete frame and utilize the received signal at pilot segments to obtain the channel estimation result  $\hat{\mathbf{h}}^{\text{SFT}, \text{P}}$  and  $\hat{\mathbf{h}}^{\text{TB}}$ . Then we use the estimated TB domain channel,  $\hat{\mathbf{h}}^{\text{TB}}$ , to predict the channel at the data segment according to (37) and (38).

## IV. CBFEM BASED CHANNEL ESTIMATION

To reduce the complexity, we formulate the channel estimation as a statistical inference problem and develop a CBFEM based channel estimation algorithm. Then we exploit the structure of the TB matrix and the CZT to further reduce its computational complexity.

### A. Algorithm

Although the MMSE estimation in (22) can achieve the optimal channel estimation performance, its complexity is unaffordable in practical systems. Since the statistical CSI,  $\mathbf{R}^{\text{TB}}$ , is available, (21) can be rewritten as  $\mathbf{y} = \bar{\mathbf{A}} \bar{\mathbf{h}}^{\text{TB}} + \mathbf{z}$ , where  $\bar{\mathbf{h}}^{\text{TB}} \in \mathbb{C}^{N_{\text{ave}}^{\text{TB}} U \times 1}$  consists of non-zero elements in  $\mathbf{h}^{\text{TB}}$ ,  $\bar{\mathbf{A}} \in \mathbb{C}^{MN_v N_F \times N_{\text{ave}}^{\text{TB}} U}$  consists of the columns of  $\mathbf{A}$  corresponding to the non-zero positions, and  $N_{\text{ave}}^{\text{TB}}$  is the average number of non-zero elements of  $\mathbf{h}_u^{\text{TB}}$  for each UT. Thus the complexity of (22) is  $\mathcal{O} \left( (MN_v N_F)^3 + (MN_v N_F)^2 N_{\text{ave}}^{\text{TB}} U \right)$ . Moreover, with the matrix inversion lemma, the MMSE estimate of  $\bar{\mathbf{h}}^{\text{TB}}$  can be rewritten as  $\bar{\mathbf{h}}_{\text{MMSE}}^{\text{TB}} = \bar{\mathbf{R}}^{\text{TB}} \left( \bar{\mathbf{A}}^{\text{H}} \bar{\mathbf{A}} \bar{\mathbf{R}}^{\text{TB}} + \sigma_z^2 \mathbf{I} \right)^{-1} \bar{\mathbf{A}}^{\text{H}} \mathbf{y}$ , where  $\bar{\mathbf{R}}^{\text{TB}}$  is a diagonal matrix consisting of the non-zero diagonal elements of  $\mathbf{R}^{\text{TB}}$ . In this case, the complexity becomes  $\mathcal{O} \left( (N_{\text{ave}}^{\text{TB}} U)^3 + (N_{\text{ave}}^{\text{TB}} U)^2 MN_v N_F \right)$ . Therefore, the complexity of MMSE estimation is  $\mathcal{O} \left( \min \left\{ (MN_v N_F)^3 + (MN_v N_F)^2 N_{\text{ave}}^{\text{TB}} U, (N_{\text{ave}}^{\text{TB}} U)^3 + (N_{\text{ave}}^{\text{TB}} U)^2 MN_v N_F \right\} \right)$ .

Although the TB domain channel exhibit sparsity in HF skywave massive MIMO-OFDM communications, when the number of UTs  $U$  is relatively large, the large  $N_{\text{ave}}^{\text{TB}}U$  will make the complexity of MMSE estimation still unbearable in practical systems. Therefore, we develop a low-complexity channel estimation algorithm by formulating the channel estimation as a statistical inference problem.

Many statistical inference methods can be used to solve this problem, such as message passing algorithms [18], [19], [20], [21], [22], [23], [24], [47]. A novel technique, named CBFEM, can unify different message passing algorithms into a single optimization framework, which has a clear objective function and can derive different solving algorithms through different constraints. Therefore, we propose to use the CBFEM technique to estimate  $\mathbf{h}^{\text{TB}}$  by transforming the statistical inference problem into an optimization one and derive the CBFEM based channel estimation algorithm by solving this problem.

From (21), we have

$$\begin{aligned} p(\mathbf{h}^{\text{TB}}, \mathbf{w} | \mathbf{y}) &= \frac{1}{p(\mathbf{y})} p(\mathbf{y} | \mathbf{w}) p(\mathbf{w} | \mathbf{h}^{\text{TB}}) p(\mathbf{h}^{\text{TB}}) \\ &= \frac{1}{p(\mathbf{y})} \prod_{i=0}^{MN_{\text{v}}N_{\text{F}}-1} p(y_i | w_i) \prod_{i=0}^{MN_{\text{v}}N_{\text{F}}-1} p(w_i | \mathbf{h}^{\text{TB}}) \\ &\quad \prod_{j=0}^{N_{\text{an}}N_{\text{de}}N_{\text{do}}U-1} p(h_j^{\text{TB}}), \end{aligned} \quad (39)$$

where  $\mathbf{w} \triangleq \mathbf{A}\mathbf{h}^{\text{TB}} \in \mathbb{C}^{MN_{\text{v}}N_{\text{F}} \times 1}$  is the auxiliary vector,  $p(y_i | w_i) = \mathcal{CN}(y_i; w_i, \sigma_z^2)$ ,  $p(w_i | \mathbf{h}^{\text{TB}}) = \delta(w_i - \mathbf{a}_i \mathbf{h}^{\text{TB}})$ ,  $p(h_j^{\text{TB}}) = \mathcal{CN}(h_j^{\text{TB}}; 0, [\mathbf{R}^{\text{TB}}]_{j,j})$ ,  $\mathbf{a}_i$  is the  $i$ -th row of  $\mathbf{A}$ ,  $h_j^{\text{TB}}$  is the  $j$ -th element of  $\mathbf{h}^{\text{TB}}$ , and  $y_i$  and  $w_i$  are the  $i$ -th element of  $\mathbf{y}$  and  $\mathbf{w}$ , respectively.

Variational Bayesian inference [48], [49] can be utilized to find a trial belief  $b(\mathbf{h}^{\text{TB}}, \mathbf{w})$  to approximate a posterior probability density  $p(\mathbf{h}^{\text{TB}}, \mathbf{w} | \mathbf{y})$  from a specific probability density family  $\mathcal{Q}$  by minimizing the variational free energy  $F_V(b)$ , i.e.,  $b(\mathbf{h}^{\text{TB}}, \mathbf{w}) = \arg \min_{b(\mathbf{h}^{\text{TB}}, \mathbf{w}) \in \mathcal{Q}} F_V(b)$ , and  $F_V(b)$  is defined as

$$F_V(b) = \mathbb{D}\{b(\mathbf{h}^{\text{TB}}, \mathbf{w}) \| p(\mathbf{h}^{\text{TB}}, \mathbf{w} | \mathbf{y})\} - \ln p(\mathbf{y}), \quad (40)$$

where  $-\ln p(\mathbf{y})$  is termed Helmholtz free energy. We use the Bethe approximation to limit the range of probability density family  $\mathcal{Q}$  by introducing factor beliefs and variable beliefs [50]. Denote  $b_{y,i}(w_i)$ ,  $b_{w,i}(w_i, \mathbf{h}^{\text{TB}})$ , and  $b_{h,j}(h_j^{\text{TB}})$  as the factor beliefs of  $p(y_i | w_i)$ ,  $p(w_i | \mathbf{h}^{\text{TB}})$ , and  $p(h_j^{\text{TB}})$ , respectively. Let  $q_{w,i}(w_i)$  and  $q_{h,j}(h_j^{\text{TB}})$  denote the variable beliefs of  $w_i$  and  $h_j^{\text{TB}}$ , respectively. According to the Bethe approximation, we have

$$\begin{aligned} b(\mathbf{h}^{\text{TB}}, \mathbf{w}) &= \frac{\prod_{i=0}^{MN_{\text{v}}N_{\text{F}}-1} b_{y,i}(w_i) b_{w,i}(w_i, \mathbf{h}^{\text{TB}}) \prod_{j=0}^{N_{\text{an}}N_{\text{de}}N_{\text{do}}U-1} b_{h,j}(h_j^{\text{TB}})}{\prod_{i=0}^{MN_{\text{v}}N_{\text{F}}-1} q_{w,i}(w_i) \prod_{j=0}^{N_{\text{an}}N_{\text{de}}N_{\text{do}}U-1} (q_{h,j}(h_j^{\text{TB}}))^{MN_{\text{v}}N_{\text{F}}}}, \\ \forall b(\mathbf{h}^{\text{TB}}, \mathbf{w}) &\in \mathcal{Q}. \end{aligned} \quad (41)$$

Moreover, the Bethe approximation needs to fulfill the marginalization consistency constraints

$$b_{y,i}(w_i) = \int b_{w,i}(w_i, \mathbf{h}^{\text{TB}}) d\mathbf{h}^{\text{TB}} = q_{w,i}(w_i), \quad (42a)$$

$$b_{h,j}(h_j^{\text{TB}}) = \int b_{w,i}(w_i, \mathbf{h}^{\text{TB}}) dw_i d\mathbf{h}_{h_j^{\text{TB}}}^{\text{TB}} = q_{h,j}(h_j^{\text{TB}}), \quad (42b)$$

where the normalization and non-negative constraints are omitted for brevity since they hold for any valid belief. According to [51], we can relax the constraints in (42) into the first-order and second-order moment consistency constraints as follows

$$\mathbb{E}\{w_i | b_{y,i}\} = \mathbb{E}\{w_i | b_{w,i}\} = \mathbb{E}\{w_i | q_{w,i}\} \quad (43a)$$

$$\mathbb{E}\{h_j^{\text{TB}} | b_{h,j}\} = \mathbb{E}\{h_j^{\text{TB}} | b_{w,i}\} = \mathbb{E}\{h_j^{\text{TB}} | q_{h,j}\} \quad (43b)$$

$$\mathbb{E}\{|w_i|^2 | b_{y,i}\} = \mathbb{E}\{|w_i|^2 | b_{w,i}\} = \mathbb{E}\{|w_i|^2 | q_{w,i}\} \quad (43c)$$

$$\begin{aligned} \mathbb{E}\{|h_j^{\text{TB}}|^2 | b_{h,j}\} &= \frac{\sum_{i=0}^{MN_{\text{v}}N_{\text{F}}-1} \mathbb{E}\{|h_j^{\text{TB}}|^2 | b_{w,i}\}}{MN_{\text{v}}N_{\text{F}}} \\ &= \mathbb{E}\{|h_j^{\text{TB}}|^2 | q_{h,j}\} \end{aligned} \quad (43d)$$

where (43d) is an approximate second-order moment consistency constraint due to the fact that the belief  $b_{w,i}$  contains the variable  $h_j^{\text{TB}}$  for any  $i$ .

Substituting (41) into (40), we can obtain the Bethe free energy as follows

$$\begin{aligned} F_B(b) &= \sum_{i=0}^{MN_{\text{v}}N_{\text{F}}-1} \mathbb{D}\{b_{y,i} \| p(y_i | w_i)\} + \sum_{i=0}^{MN_{\text{v}}N_{\text{F}}-1} \mathbb{D}\{b_{w,i} \| p(w_i | \mathbf{h}^{\text{TB}})\} \\ &\quad + \sum_{j=0}^{N_{\text{an}}N_{\text{de}}N_{\text{do}}U-1} \mathbb{D}\{b_{h,j} \| p(h_j^{\text{TB}})\} + \mathbb{H}\{q_{w,i}\} \\ &\quad + \sum_{j=0}^{N_{\text{an}}N_{\text{de}}N_{\text{do}}U-1} MN_{\text{v}}N_{\text{F}} \mathbb{H}\{q_{h,j}\}, \end{aligned} \quad (44)$$

Finally, we transform the channel estimation into the following CBFEM question

$$\min_{\{b_{y,i}\}, \{b_{w,i}\}, \{b_{h,j}\}, \{q_{w,i}\}, \{q_{h,j}\}} F_B(b) \quad \text{s.t. (43)}. \quad (45)$$

We can utilize Lagrange multipliers to solve the CBFEM question, which is given in Appendix C. The resulting CBFEM based channel estimation algorithm is summarized in Algorithm 2, where  $b_h \triangleq \prod_{j=0}^{N_{\text{an}}N_{\text{de}}N_{\text{do}}U-1} b_{h,j}$ ,  $\mathbf{e} \triangleq [1, 1, \dots, 1]^T \in \mathbb{C}^{N_{\text{an}}N_{\text{de}}N_{\text{do}}U \times 1}$  and  $(\cdot)^{\circ-1}$  denotes the element-wise inverse. Moreover, during the iterations, the damping factor can be leveraged to ensure the convergence of the algorithm [22].

### B. Low-Complexity Implementation

The computational complexity of each iteration of the proposed CBFEM based channel estimation algorithm mainly comes from steps 7 and 8. In this subsection, we focus on these two steps and further reduce their complexity.

**Algorithm 2** CBFEM Based Channel Estimation Algorithm

**Input:**  $\mathbf{A}$ ,  $\mathbf{y}$ ,  $\sigma_p^2$ ,  $\sigma_z^2$ ,  $p(\mathbf{h}^{\text{TB}}) = \mathcal{CN}(\mathbf{h}^{\text{TB}}; \mathbf{0}, \mathbf{R}^{\text{TB}})$

1 **Initialize:**  $b_h = p(\mathbf{h}^{\text{TB}})$ ,  $\boldsymbol{\eta}^{h,b_h} = \mathbf{0}$

2 **repeat**

3  $\boldsymbol{\eta}^{h,b_w} = -\text{Var}\{\mathbf{h}^{\text{TB}} | b_h\}^{\circ-1} - \frac{1}{MN_v N_F} \boldsymbol{\eta}^{h,b_h}$

4  $\boldsymbol{\eta}^{h,b_h} =$   
 $MN_v N_F \left( \left( \sigma_p^2 \mathbf{e}^{\text{T}}(\boldsymbol{\eta}^{h,b_w})^{\circ-1} \sigma_z^2 \right) \sigma_p^{-2} \mathbf{e} - (\boldsymbol{\eta}^{h,b_w})^{\circ-1} \right)^{\circ-1}$

5  $\tilde{\boldsymbol{\tau}}^{h,b_w} = \mathbb{E}\{\mathbf{h}^{\text{TB}} | b_h\} \circ \text{Var}\{\mathbf{h}^{\text{TB}} | b_h\}^{\circ-1}$

6  $\boldsymbol{\kappa} = \tilde{\boldsymbol{\tau}}^{h,b_w} \circ (\boldsymbol{\eta}^{h,b_w})^{\circ-1}$

7  $\boldsymbol{\psi} = \mathbf{A}\boldsymbol{\kappa}$

8  $\boldsymbol{\varpi} = \frac{1}{MN_v N_F \sigma_p^2} \mathbf{A}^{\text{H}}(\mathbf{y} + \boldsymbol{\psi}) - \boldsymbol{\kappa}$

9  $b_h \propto p(\mathbf{h}^{\text{TB}}) \mathcal{CN}(\mathbf{h}^{\text{TB}}; \boldsymbol{\varpi}, \text{diag}\{- (\boldsymbol{\eta}^{h,b_h})^{\circ-1}\})$

10 **until** the termination condition is fulfilled;

**Output:**  $\hat{\mathbf{h}}^{\text{TB}} = \mathbb{E}\{\mathbf{h}^{\text{TB}} | b_h\}$

Recalling (28) and (30), steps 7 and 8 of the CBFEM based channel estimation algorithm can be rewritten as

$$\boldsymbol{\psi} = \sigma_p \mathbf{X}_c \left( \tilde{\mathbf{D}} \otimes (\mathbf{V} (\bar{\mathbf{U}} \otimes \mathbf{I}_{N_{\text{an}}})) \right) [\mathbf{S}_0, \dots, \mathbf{S}_{U-1}] \boldsymbol{\kappa}, \quad (46)$$

$$\boldsymbol{\varpi} = \frac{1}{MN_v N_F \sigma_p} [\mathbf{S}_0, \dots, \mathbf{S}_{U-1}]^{\text{H}} \left( \tilde{\mathbf{D}}^{\text{H}} \otimes ((\bar{\mathbf{U}}^{\text{H}} \otimes \mathbf{I}_{N_{\text{an}}}) \mathbf{V}^{\text{H}}) \right) \times \mathbf{X}_c^{\text{H}}(\mathbf{y} + \boldsymbol{\psi}) - \boldsymbol{\kappa}. \quad (47)$$

Since  $\mathbf{V}_k$  and  $\tilde{\mathbf{D}}$  are not DFT matrices, fast Fourier transform (FFT) cannot be directly used. In this case, the CZT [52] can be utilized to reduce the computational complexity of (46) and (47). Specifically,  $\mathbf{V}_k$  and  $\tilde{\mathbf{D}}$  can be rewritten as

$$\mathbf{V}_k = \text{diag}\left\{ \boldsymbol{\xi}_{W_{S,k}, N_{\text{an}}, 0, M} \right\} \mathbf{F}_{N_{(S)} \times M}^{\text{H}} \times \text{diag}\left\{ \tilde{\boldsymbol{\xi}}_{W_{S,k}, M, N_{\text{an}}, N_{(S)}} \right\} \mathbf{F}_{N_{(S)} \times N_{\text{an}}} \text{diag}\left\{ \boldsymbol{\xi}_{W_{S,k}, 0, 0, N_{\text{an}}} \right\}, \quad (48a)$$

$$\tilde{\mathbf{D}} = \text{diag}\left\{ \boldsymbol{\xi}_{W_{\text{T}}, N_{\text{do}}, \frac{n_p N_{\text{do}}}{N_S}, N_F} \right\} \mathbf{F}_{N_{(T)} \times N_F}^{\text{H}} \times \text{diag}\left\{ \tilde{\boldsymbol{\xi}}_{W_{\text{T}}, N_F, N_{\text{do}}, N_{(T)}} \right\} \mathbf{F}_{N_{(T)} \times N_{\text{do}}} \text{diag}\left\{ \boldsymbol{\xi}_{W_{\text{T}}, -\frac{2n_p}{N_S}, 0, N_{\text{do}}} \right\}, \quad (48b)$$

respectively, where  $\mathbf{F}_N$  denotes the  $N$ -dimensional unitary DFT matrix,  $\mathbf{F}_{N \times G}$  denotes the matrix composed of the first  $G$  ( $\leq N$ ) columns of  $\mathbf{F}_N$ ,  $N_{(S)}$  and  $N_{(T)}$  are integers greater than or equal to  $M + N_{\text{an}} - 1$  and  $N_F + N_{\text{do}} - 1$ , respectively.  $W_{S,k} \triangleq e^{-j \frac{4\pi (J_c + k \Delta f) \Delta \tau}{N_{\text{an}}}}$ ,  $W_F \triangleq e^{-j \frac{2\pi N_{\text{r}}}{N_{\text{de}} N_v}}$ ,  $W_T \triangleq e^{j \frac{2\pi N_d N_S}{N_{\text{do}} N}}$ ,  $\boldsymbol{\xi}_{W, \alpha_1, \alpha_2, N} \in \mathbb{C}^{N \times 1}$  with  $[\boldsymbol{\xi}_{W, \alpha_1, \alpha_2, N}]_i = W^{(i^2 - \alpha_1 i - \alpha_2)/2}$ , and  $\tilde{\boldsymbol{\xi}}_{W, N_1, N_2, N} \triangleq \sqrt{N} \mathbf{F}_N \left[ \boldsymbol{\xi}_{W, 0, 0, N_1}^{\text{H}}, \mathbf{0}, \boldsymbol{\xi}_{W, 2(N_2-1), -(N_2-1)^2, N_2-1}^{\text{H}} \right]^{\text{T}} \in \mathbb{C}^{N \times 1}$ .

With (46), (47) and (48), steps 7 and 8 can be implemented efficiently and the complexities are respectively reduced to  $\mathcal{O}(N_{\text{an}} N_{\text{do}} \bar{N}_{\text{de}} \log_2 \bar{N}_{\text{de}} + N_v N_{\text{do}} N_{(S)} \log_2$

TABLE I  
SIMULATION PARAMETERS

Parameter	Value
Carrier frequency $f_c$	16 MHz
Subcarrier spacing $\Delta f$	250 Hz
Number of subcarriers $N_c$	2048
Length of CP $N_g$	512
Number of valid subcarriers $N_v$	1536
Number of BS antennas $M$	128
BS antenna spacing $d$	9 m
Size of frame ( $N_F, N_S, n_p$ )	(8, 14, 6)
Number of UTs $U$	64
Velocity of UTs $v_u$	30 / 100 / 250 km/h

$N_{(S)} + MN_v N_{(T)} \log_2 N_{(T)})$  and  $\mathcal{O}(N_v N_F N_{(S)} \log_2 N_{(S)} + N_{\text{an}} N_F \bar{N}_{\text{de}} \log_2 \bar{N}_{\text{de}} + N_{\text{an}} \bar{N}_{\text{de}} N_{(T)} \log_2 N_{(T)})$ , where  $\bar{N}_{\text{de}} \triangleq \lfloor N_v / N_{\text{r}} \rfloor N_{\text{de}}$ .

## V. SIMULATION RESULTS

In this section, we provide simulation results to illustrate the performance of the proposed channel acquisition approach for HF skywave massive MIMO-OFDM communications. The main simulation parameters are given in Table I. We generate UTs at a distance of 2000 km from the BS, and azimuth angles of the UTs seen from the BS are generated in the interval  $(-90^\circ, 90^\circ)$ . A commercial ray-tracing software, Proplab-Pro version 3.1, is utilized to generate the realistic HF skywave channel for each UT independently [53]. Specifically, we use the software to get parameters of each propagation path between the BS and each UT, including the path gain, the propagation delay, the azimuth and elevation AoA. The maximum Doppler comes from the ionosphere and UT mobility, i.e.,  $\nu_{\text{max}} = \frac{\nu_{\text{iono}}}{2} + \frac{v_u}{c} f_c$ , where  $\nu_{\text{iono}}$  is the Doppler spread imposed by the ionosphere and  $v_u$  is the velocity of UT  $u$ . In the simulation,  $\nu_{\text{iono}}$  is set to be 0.5 Hz for moderate ionospheric conditions at mid-latitude regions [54], the Doppler frequency of each path is randomly generated within the range  $[-\nu_{\text{max}}, \nu_{\text{max}}]$ , and  $N_d$  is set to be 8 according to the maximum Doppler. Then using the generated path parameters and randomly generated initial phases for paths, the SFT domain channel of each UT is generated according to (5) and the statistical CSI  $\mathbf{R}^{\text{TB}}$  is acquired by the method proposed in [38].

Note that  $\hat{\mathbf{h}}^{\text{SFT}, p}$  consists of the estimated spatial-frequency domain channels of pilot segments in  $N_F$  timeslots and only the spatial-frequency domain channel of the last timeslot (i.e., the current timeslot) is needed for DL transmit design or UL signal detection. Therefore, in addition to the NMSE of all timeslots defined in (25), we will also show the NMSE between the estimated channel and the real channel in the current timeslot. The signal-to-noise ratio (SNR) in the simulation refers to the received SNR, whose range is set to be  $[-10, 20]$  dB. Such a range is realizable in practice since each UT only needs about 50 W of transmit power to achieve a SNR of up to 20 dB under the given simulation parameters. We refer to the UT grouping algorithm based on the channel overlap degree (32) and (36) as TB-UG and B-UG, respectively, and

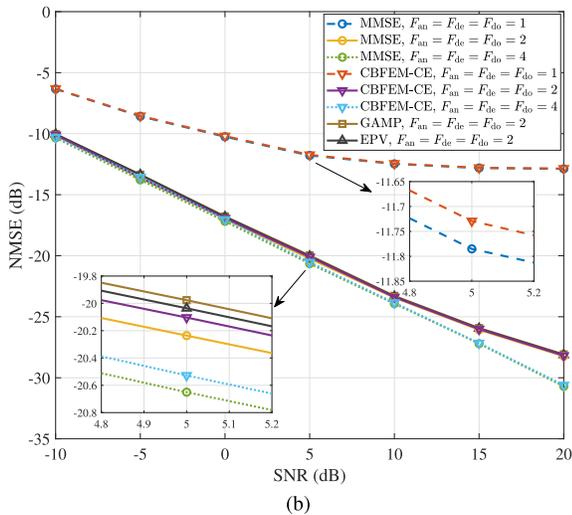
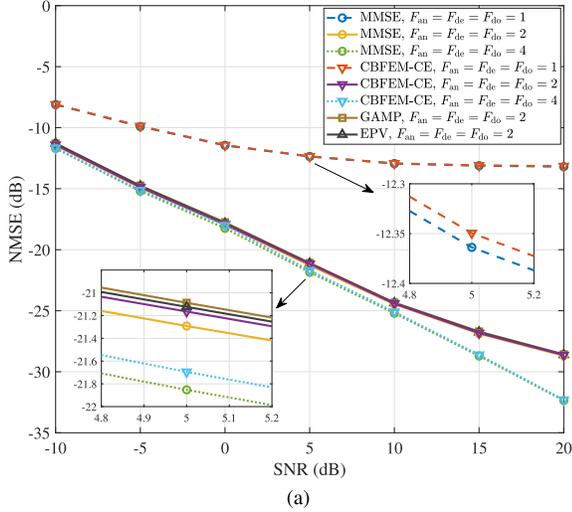


Fig. 2. The NMSE performance of different algorithms with different fine factors. (a) NMSE of all timeslots; (b) NMSE of current timeslot.

the number of groups to divide is set to be  $S = \lfloor N_v / N_\tau \rfloor$ . The performance of random UT grouping (Random-UG) is also provided for comparison, and we refer to the CBFEM based channel estimation algorithm as the CBFEM-CE algorithm.

We first compare the NMSE performance of different algorithms with different fine factors. The MMSE estimation, the generalized approximate message passing (GAMP) [47], and the expectation propagation variant (EPV) [22] algorithms are included along with the proposed CBFEM-CE algorithm. Since the simulation parameters in Table I will make the complexity of the MMSE estimation unbearable, we use a set of smaller parameters only in this simulation. Specifically, we set  $M = 64$ ,  $N_v = 128$ , and the number of UTs is set to be 32. The simulation results are shown in Fig. 2, where the UT grouping algorithm is TB-UG and the UT velocity is 100 km/h. From the figure, larger fine factors can achieve better performance due to the improved accuracy of the proposed TB based channel model. Furthermore, the

TABLE II  
COMPUTATIONAL COMPLEXITIES OF DIFFERENT CHANNEL ESTIMATION ALGORITHMS

Algorithm	Complexity
MMSE	$\mathcal{O}\left(\min\left\{(MN_v N_F)^3 + (MN_v N_F)^2 N_{\text{ave}}^{\text{TB}} U, (N_{\text{ave}}^{\text{TB}} U)^3 + (N_{\text{ave}}^{\text{TB}} U)^2 MN_v N_F\right\}\right)$
CBFEM-CE (per iteration)	$\mathcal{O}\left((N_{\text{an}} N_{\text{do}} + N_{\text{an}} N_F) \bar{N}_{\text{de}} \log_2 \bar{N}_{\text{de}} + (N_v N_{\text{do}} + N_v N_F) N_{\text{(S)}} \log_2 N_{\text{(S)}} + (MN_v + N_{\text{an}} \bar{N}_{\text{de}}) N_{\text{(T)}} \log_2 N_{\text{(T)}}\right)$
GAMP / EPV (per iteration)	$\mathcal{O}(MN_v N_F N_{\text{ave}}^{\text{TB}} U)$

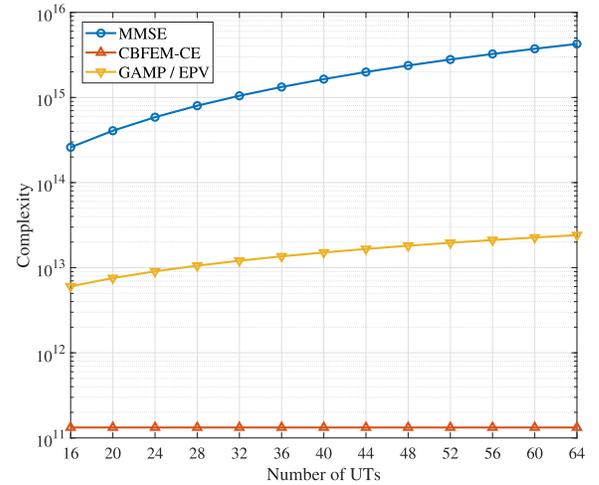


Fig. 3. The complexities of different algorithms versus the number of UTs.

performance of CBFEM-CE, GAMP, and EPV algorithms can approach that of the optimal MMSE estimation with a very small gap. Moreover, the proposed CBFEM-CE algorithm can be implemented with much lower complexity by using the CZT and the structure of the TB matrix.

The computational complexity of different channel estimation algorithms are summarized in Table II and are plotted in Fig. 3 under different numbers of UTs, where  $F_{\text{an}} = F_{\text{de}} = F_{\text{do}} = 2$  and the number of iterations for CBFEM-CE, GAMP, and EPV algorithms are set to 300 (the number of iterations required is different for different SNRs and system configurations). From the figure, the complexity of MMSE is the highest due to the matrix-inversion. On the other hand, owing to the utilization of the structure of the TB matrix and the CZT, the complexity of CBFEM-CE algorithm is the lowest among all the algorithms.

Fig. 4 shows the convergence performance of different algorithms under different SNRs, where the UT velocity is 100 km/h,  $F_{\text{an}} = F_{\text{de}} = F_{\text{do}} = 2$ , and the UT grouping algorithm is TB-UG. From the figure, the NMSE performance of the CBFEM-CE algorithm improves as the number of iterations increases until the convergence. Moreover, the CBFEM-CE algorithm has similar convergence behavior to that of GAMP and EPV algorithms with much lower complexity.

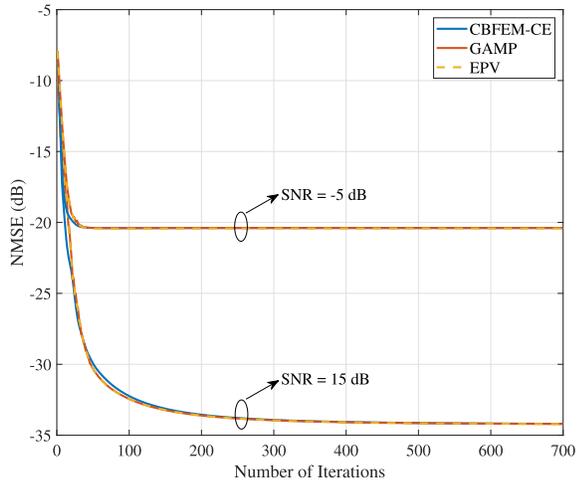


Fig. 4. The convergence performance of different algorithms under different SNRs.

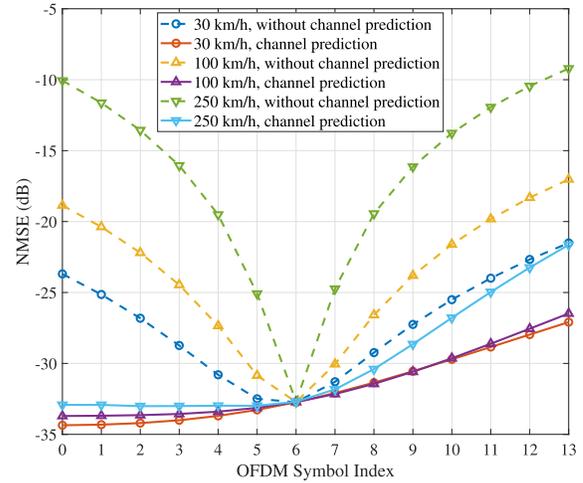


Fig. 6. The NMSE performance versus OFDM symbol index under different UT velocities.

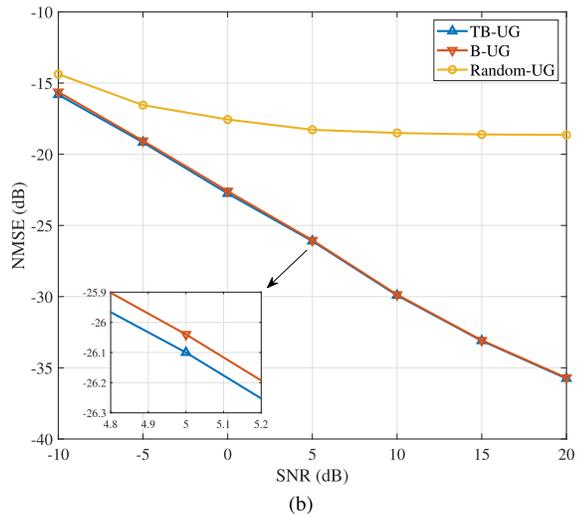
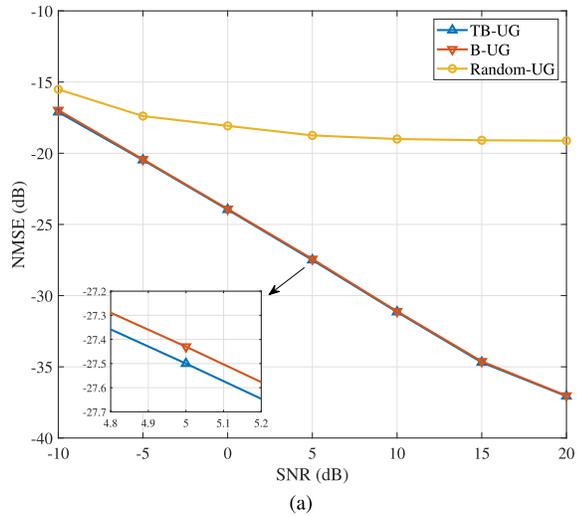


Fig. 5. The NMSE performance of different UT grouping algorithms. (a) NMSE of all timeslots; (b) NMSE of current timeslot.

Fig. 5 compares the NMSE performance of different UT grouping algorithms, where the UT velocity is 100 km/h,

$F_{an} = F_{de} = F_{do} = 2$ , and the channel estimation algorithm is CBFEM-CE. From the figure, the proposed UT grouping algorithms are crucial to the channel estimation and have a significant performance gain over random-UG, especially when the SNR is high. This is because there is severe inter-UT interference when the number of UTs is large, which can be effectively tackled by the proposed UT grouping algorithms. Moreover, the performance of TB-UG is almost the same as that of B-UG while the computational complexity of  $\tilde{\rho}_{u,u'}$  used in B-UG is much lower than that of  $\rho_{u,u'}$  used in TB-UG.

Fig. 6 shows the NMSE performance of the proposed channel prediction method under different UT velocities, where the frame structure is as in Fig. 1, SNR = 15 dB, the UT grouping algorithm is TB-UG,  $F_{an} = F_{de} = F_{do} = 2$ , and the channel estimation algorithm is CBFEM-CE. The NMSE here refers to the NMSE between the channel obtained by channel prediction and the real channel of each OFDM symbol in the current timeslot. We observe that when we directly apply the estimated channel at the pilot segment to the data segment, the NMSE performance will rapidly degrade as the delay between the pilot symbol and data symbols increases, especially when the UT velocity is fast. However, the proposed channel prediction method can acquire a more accurate CSI of the data segment due to the utilization of the estimated TB domain channel, which is conducive to DL transmit design and UL signal detection.

## VI. CONCLUSION

In this paper, we have investigated the channel acquisition for HF skywave massive MIMO-OFDM communications. We first established a TB based channel model using sampled triple steering vectors, each of which corresponds to a physical TB in the SFT domain. Then, based on the channel model, we investigated the optimal channel estimation for pilot segments and revealed the conditions for minimizing the NMSE of the channel estimate. We showed that UTs with overlapping TB domain channels should be allocated pilot sequences with different phase shift factors while UTs with

non-overlapping TB domain channels can reuse the same pilot sequence. Moreover, the pilot design was given, including UT grouping and pilot scheduling, and the channel prediction method for the data segment was provided based on the TB domain channel estimate. To reduce the complexity of the channel estimation, we developed the CBFEM based channel estimation algorithm and its low-complexity implementation using CZT based on the structure of the TB matrix. Simulation results verified the validity of the proposed CSI acquisition approach.

#### APPENDIX A PROOF OF THEOREM 1

We start by presenting a lemma that is required in the subsequent proof.

*Lemma 1:* When one of the conditions in Theorem 1 satisfies, we have

$$\lim_{M, N_v, N_F \rightarrow \infty} \frac{1}{MN_v N_F} \mathbf{R}_u^{\text{SFT}, \text{P}} \mathbf{X}_u^{\text{H}} \mathbf{X}_{u'} \mathbf{R}_{u'}^{\text{SFT}, \text{P}} = \mathbf{0}. \quad (49)$$

*Proof:* According to (17) and the fact that  $\mathbf{X}_u \tilde{\mathbf{P}} = \sigma_p \mathbf{X}_c \bar{\mathbf{P}} \mathbf{S}_u$  we have

$$\begin{aligned} & \lim_{M, N_v, N_F \rightarrow \infty} \frac{1}{MN_v N_F} \mathbf{R}_u^{\text{SFT}, \text{P}} \mathbf{X}_u^{\text{H}} \mathbf{X}_{u'} \mathbf{R}_{u'}^{\text{SFT}, \text{P}} \\ &= \lim_{M, N_v, N_F \rightarrow \infty} \frac{\sigma_p^2}{MN_v N_F} \underbrace{\tilde{\mathbf{P}} \mathbf{R}_u^{\text{TB}} \mathbf{S}_u^{\text{H}} \bar{\mathbf{P}}^{\text{H}} \mathbf{X}_c^{\text{H}} \mathbf{X}_c \bar{\mathbf{P}} \mathbf{S}_{u'}^{\text{TB}} \mathbf{R}_{u'}^{\text{TB}} \tilde{\mathbf{P}}^{\text{H}}}_{\Phi}. \end{aligned} \quad (50)$$

Then we have

$$\begin{aligned} & \lim_{M, N_v, N_F \rightarrow \infty} \frac{\sigma_p^2}{MN_v N_F} [\Phi]_{a,b} \\ &= \lim_{M, N_v, N_F \rightarrow \infty} \sigma_p^2 [\mathbf{R}_u^{\text{TB}}]_{a,a} [\mathbf{R}_{u'}^{\text{TB}}]_{b,b} \\ & \quad \times \underbrace{\frac{1}{N_F} \sum_{n_F=0}^{N_F-1} e^{-j2\pi(n_F N_S + n_p) \frac{N_d(n_{\text{do}} - n'_{\text{do}})}{N_{\text{do}} N}}}_{\alpha_T(n_{\text{do}}, n'_{\text{do}})} \\ & \quad \times \underbrace{\frac{1}{N_v} \sum_{k=k_0}^{k_{N_v}-1} e^{j2\pi k \frac{N_\tau(n_{\text{de}} - n'_{\text{de}} + \phi_u - \phi_{u'})}{N_{\text{de}} N_v}}}_{\alpha_F(n_{\text{de}}, n'_{\text{de}}, \phi_u, \phi_{u'})} \\ & \quad \times \underbrace{\frac{1}{M} \sum_{m=0}^{M-1} e^{j2\pi(f_c + k\Delta f)m\Delta\tau \frac{2(n_{\text{an}} - n'_{\text{an}})}{N_{\text{an}}}}}_{\alpha_S(n_{\text{an}}, n'_{\text{an}}, k)}, \end{aligned} \quad (51)$$

where

$$\begin{aligned} n_{\text{do}} &= \lfloor a/(N_{\text{de}} N_{\text{an}}) \rfloor, \quad n'_{\text{do}} = \langle \lfloor a/N_{\text{an}} \rfloor \rangle_{N_{\text{de}}}, \quad n_{\text{an}} = \langle a \rangle_{N_{\text{an}}}, \\ n'_{\text{do}} &= \lfloor b/(N_{\text{de}} N_{\text{an}}) \rfloor, \quad n'_{\text{de}} = \langle \lfloor b/N_{\text{an}} \rfloor \rangle_{N_{\text{de}}}, \quad n'_{\text{an}} = \langle b \rangle_{N_{\text{an}}}, \end{aligned}$$

and  $\langle \cdot \rangle_N$  denotes the modulo- $N$  operation.

When  $\mathbf{R}_u^{\text{TB}} \mathbf{R}_{u'}^{\text{TB}} \neq \mathbf{0}$ , there is an overlap between TB domain channels of UTs  $u$  and  $u'$ , thus we let  $\phi_u \neq \phi_{u'}$  so that  $n_{\text{de}} - n'_{\text{de}} + \phi_u - \phi_{u'} \neq 0$  and  $\alpha_F(n_{\text{de}}, n'_{\text{de}}, \phi_u, \phi_{u'}) = 0$  when  $N_v \rightarrow \infty$ . When  $\mathbf{R}_u^{\text{TB}} \mathbf{R}_{u'}^{\text{TB}} = \mathbf{0}$ , there is no overlap between TB domain channels of UTs  $u$  and  $u'$ , which means that when both  $[\mathbf{R}_u^{\text{TB}}]_{a,a}$  and  $[\mathbf{R}_{u'}^{\text{TB}}]_{b,b}$  are not equal to 0, we have

$a \neq b$ , thus at least one of  $n_{\text{do}} - n'_{\text{do}} \neq 0$ ,  $n_{\text{de}} - n'_{\text{de}} + \phi_u - \phi_{u'} \neq 0$ , and  $n_{\text{an}} - n'_{\text{an}} \neq 0$  satisfies. In this case, at least one of  $\alpha_T(n_{\text{do}}, n'_{\text{do}})$ ,  $\alpha_F(n_{\text{de}}, n'_{\text{de}}, \phi_u, \phi_{u'})$  and  $\alpha_S(n_{\text{an}}, n'_{\text{an}}, k)$  is equal to 0 when  $M, N_v, N \rightarrow \infty$ . In summary, when one of the conditions in Theorem 1 satisfies,

$$\lim_{M, N_v, N \rightarrow \infty} \frac{\sigma_p^2}{MN_v N_F} \Phi = \mathbf{0}. \quad (52)$$

Substituting (52) into (50), we obtain (49). This completes the proof of Lemma 1.

Next, we define that

$$\tilde{\mathbf{C}}_u \triangleq \mathbf{X}_u \mathbf{R}_u^{\text{SFT}, \text{P}} \mathbf{X}_u^{\text{H}} + \sigma_z^2 \mathbf{I}_{MN_v N_F}. \quad (53)$$

Due to the fact that  $\mathbf{C} \succ \mathbf{0}$ ,  $\tilde{\mathbf{C}}_u \succ \mathbf{0}$ , and  $\mathbf{C} - \tilde{\mathbf{C}}_u \succeq \mathbf{0}$ , we can obtain that  $\tilde{\mathbf{C}}_u^{-1} - \mathbf{C}^{-1} \succeq \mathbf{0}$  [55]. Therefore,  $\mathbf{R}_u^{\text{SFT}, \text{P}} \mathbf{X}_u^{\text{H}} (\tilde{\mathbf{C}}_u^{-1} - \mathbf{C}^{-1}) \mathbf{X}_u \mathbf{R}_u^{\text{SFT}, \text{P}} \succeq \mathbf{0}$ , and

$$\text{NMSE} \sum_{u=0}^{U-1} \frac{1}{MN_v N_F U \vartheta_u} \text{tr} \left\{ \mathbf{R}_u^{\text{SFT}, \text{P}} - \mathbf{R}_u^{\text{SFT}, \text{P}} \mathbf{X}_u^{\text{H}} \tilde{\mathbf{C}}_u^{-1} \mathbf{X}_u \mathbf{R}_u^{\text{SFT}, \text{P}} \right\}. \quad (54)$$

According to Lemma 1, when one of the conditions in Theorem 1 satisfies, we have

$$\begin{aligned} & \lim_{M, N_v, N_F \rightarrow \infty} \frac{1}{MN_v N_F} \mathbf{R}_u^{\text{SFT}, \text{P}} \mathbf{X}_u^{\text{H}} \mathbf{C} \mathbf{X}_u \mathbf{R}_u^{\text{SFT}, \text{P}} \\ &= \lim_{M, N_v, N_F \rightarrow \infty} \frac{1}{MN_v N_F} \mathbf{R}_u^{\text{SFT}, \text{P}} \mathbf{X}_u^{\text{H}} \tilde{\mathbf{C}}_u \mathbf{X}_u \mathbf{R}_u^{\text{SFT}, \text{P}} \\ &\Rightarrow \lim_{M, N_v, N_F \rightarrow \infty} \frac{1}{MN_v N_F} \text{tr} \left\{ \mathbf{R}_u^{\text{SFT}, \text{P}} - \mathbf{R}_u^{\text{SFT}, \text{P}} \mathbf{X}_u^{\text{H}} \mathbf{C}^{-1} \mathbf{X}_u \mathbf{R}_u^{\text{SFT}, \text{P}} \right\} \\ &= \lim_{M, N_v, N_F \rightarrow \infty} \frac{1}{MN_v N_F} \text{tr} \left\{ \mathbf{R}_u^{\text{SFT}, \text{P}} - \mathbf{R}_u^{\text{SFT}, \text{P}} \mathbf{X}_u^{\text{H}} \tilde{\mathbf{C}}_u^{-1} \mathbf{X}_u \mathbf{R}_u^{\text{SFT}, \text{P}} \right\}. \end{aligned} \quad (55)$$

Substituting (55) into (25), the equality in (54) holds and the NMSE is reduced to the minimum value in (31). This completes the proof.

#### APPENDIX B PROOF OF THEOREM 2

We start by presenting a lemma that is required in the subsequent proof.

*Lemma 2:* For arbitrary  $N_v$  and  $N_F$ , when  $u' \notin \mathcal{I}_u$ , we have

$$\lim_{M \rightarrow \infty} \frac{1}{MN_v N_F} \mathbf{R}_u^{\text{SFT}, \text{P}} \mathbf{X}_u^{\text{H}} \mathbf{X}_{u'} \mathbf{R}_{u'}^{\text{SFT}, \text{P}} = \mathbf{0}. \quad (56)$$

*Proof:* When  $N_v$  and  $N_F$  are arbitrary and only  $M \rightarrow \infty$ , we can obtain that

$$\lim_{M \rightarrow \infty} \frac{1}{MN_v N_F} \mathbf{R}_u^{\text{SFT}, \text{P}} \mathbf{X}_u^{\text{H}} \mathbf{X}_{u'} \mathbf{R}_{u'}^{\text{SFT}, \text{P}} = \lim_{M \rightarrow \infty} \frac{\sigma_p^2}{MN_v N_F} \tilde{\mathbf{P}} \tilde{\Phi} \tilde{\mathbf{P}}^{\text{H}}, \quad (57)$$

where

$$\begin{aligned} & \lim_{M \rightarrow \infty} \frac{\sigma_p^2}{MN_v N_F} [\tilde{\Phi}]_{a,b} \\ &= \lim_{M \rightarrow \infty} \sigma_p^2 \sum_{i=0}^{N_{\text{de}} N_{\text{do}} - 1} \sum_{j=0}^{N_{\text{de}} N_{\text{do}} - 1} [\mathbf{R}_u^{\text{TB}}]_{a, i N_{\text{an}} + n_{\text{an}}} [\mathbf{R}_{u'}^{\text{TB}}]_{j N_{\text{an}} + n'_{\text{an}}, b} \\ & \quad \times \alpha_T(n''_{\text{do}}, n'''_{\text{do}}) \alpha_F(n''_{\text{de}}, n'''_{\text{de}}, \phi_u, \phi_{u'}) \alpha_S(n_{\text{an}}, n'_{\text{an}}, k), \end{aligned} \quad (58)$$

where

$$n''_{\text{do}} = \lfloor i/N_{\text{de}} \rfloor, \quad n''_{\text{de}} = \langle i \rangle_{N_{\text{de}}}, \quad n'''_{\text{do}} = \lfloor j/N_{\text{de}} \rfloor, \quad n'''_{\text{de}} = \langle j \rangle_{N_{\text{de}}},$$

and  $n_{\text{an}}, n'_{\text{an}}, \alpha_{\text{T}}(n''_{\text{do}}, n'''_{\text{do}}), \alpha_{\text{F}}(n''_{\text{de}}, n'''_{\text{de}}, \phi_u, \phi_{u'})$  and  $\alpha_{\text{S}}(n_{\text{an}}, n'_{\text{an}}, k)$  are similar to the definitions in Appendix A.

When  $u' \notin \mathcal{I}_u$ , the TB domain channel of UTs  $u$  and  $u'$  is non-overlapping along the spatial-beam dimension, which means that when both  $[\mathbf{R}_u^{\text{TB}}]_{a,i N_{\text{an}}+n_{\text{an}}}$  and  $[\mathbf{R}_{u'}^{\text{TB}}]_{j N_{\text{an}}+n'_{\text{an}},b}$  are not equal to 0, we have  $n_{\text{an}} \neq n'_{\text{an}}$ , thus  $\alpha_{\text{S}}(n_{\text{an}}, n'_{\text{an}}, k) = 0$  when  $M \rightarrow \infty$ . Therefore, we have

$$\lim_{M \rightarrow \infty} \frac{\sigma_{\text{p}}^2}{MN_{\text{v}}N_{\text{F}}} \tilde{\Phi} = \mathbf{0}. \quad (59)$$

Substituting (59) into (57), we obtain (56). This completes the proof of Lemma 2.

Then we have in (60), shown at the bottom of the page, where (a) follows from Lemma 2. Substituting (60) into (25), we can obtain (34). This concludes the proof.

### APPENDIX C DERIVATION OF CBFEM BASED CHANNEL ESTIMATION ALGORITHM

The Lagrange function of the (45) can be given by

$$\begin{aligned} L_{\text{B}} = & F_{\text{B}} \\ & + \sum_{i=0}^{MN_{\text{v}}N_{\text{F}}-1} 2\text{Re} \left\{ \left( \tau_i^{w,b_y} \right)^* \left( \mathbb{E} \{ w_i | q_{w,i} \} - \mathbb{E} \{ w_i | b_{y,i} \} \right) \right\} \\ & + \sum_{i=0}^{MN_{\text{v}}N_{\text{F}}-1} 2\text{Re} \left\{ \left( \tau_i^{w,b_w} \right)^* \left( \mathbb{E} \{ w_i | q_{w,i} \} - \mathbb{E} \{ w_i | b_{w,i} \} \right) \right\} \\ & + \sum_{i=0}^{MN_{\text{v}}N_{\text{F}}-1} \sum_{j=0}^{N_{\text{an}}N_{\text{de}}N_{\text{do}}U-1} 2\text{Re} \left\{ \left( \tau_{i,j}^{h,b_w} \right)^* \left( \mathbb{E} \{ h_j^{\text{TB}} | q_{h,j} \} - \mathbb{E} \{ h_j^{\text{TB}} | b_{w,i} \} \right) \right\} \\ & + \sum_{j=0}^{N_{\text{an}}N_{\text{de}}N_{\text{do}}U-1} 2\text{Re} \left\{ \left( \tau_j^{h,b_h} \right)^* \left( \mathbb{E} \{ h_j^{\text{TB}} | q_{h,j} \} - \mathbb{E} \{ h_j^{\text{TB}} | b_{h,j} \} \right) \right\} \\ & + \sum_{i=0}^{MN_{\text{v}}N_{\text{F}}-1} \eta_i^{w,b_y} \left( \mathbb{E} \{ |w_i|^2 | q_{w,i} \} - \mathbb{E} \{ |w_i|^2 | b_{y,i} \} \right) \\ & + \sum_{i=0}^{MN_{\text{v}}N_{\text{F}}-1} \eta_i^{w,b_w} \left( \mathbb{E} \{ |w_i|^2 | q_{w,i} \} - \mathbb{E} \{ |w_i|^2 | b_{w,i} \} \right) \\ & + \sum_{j=0}^{N_{\text{an}}N_{\text{de}}N_{\text{do}}U-1} \eta_j^{h,b_w} \left( MN_{\text{v}}N_{\text{F}} \mathbb{E} \{ |h_j^{\text{TB}}|^2 | q_{h,j} \} - \sum_{i=0}^{MN_{\text{v}}N_{\text{F}}-1} \mathbb{E} \{ |h_j^{\text{TB}}|^2 | b_{w,i} \} \right) \\ & + \sum_{j=0}^{N_{\text{an}}N_{\text{de}}N_{\text{do}}U-1} \eta_j^{h,b_h} \left( \mathbb{E} \{ |h_j^{\text{TB}}|^2 | q_{h,j} \} - \mathbb{E} \{ |h_j^{\text{TB}}|^2 | b_{h,j} \} \right). \end{aligned} \quad (61)$$

Then by setting the first-order derivatives of (61) for each belief equal to zeros, a series of fixed-point equations can be obtained as follows.

$$b_{y,i} \propto p(y_i | w_i) \mathcal{CN} \left( w_i; -\frac{\tau_i^{w,b_y}}{\eta_i^{w,b_y}}, -\frac{1}{\eta_i^{w,b_y}} \right), \quad (62)$$

$$q_{w,i} \propto \mathcal{CN} \left( w_i; -\frac{\tau_i^{w,b_y} + \tau_i^{w,b_w}}{\eta_i^{w,b_y} + \eta_i^{w,b_w}}, -\frac{1}{\eta_i^{w,b_y} + \eta_i^{w,b_w}} \right), \quad (63)$$

$$b_{w,i} \propto p(w_i | \mathbf{h}^{\text{TB}}) \mathcal{CN} \left( w_i; -\frac{\tau_i^{w,b_w}}{\eta_i^{w,b_w}}, -\frac{1}{\eta_i^{w,b_w}} \right) \\ \times \prod_{j=0}^{N_{\text{an}}N_{\text{de}}N_{\text{do}}U-1} \mathcal{CN} \left( h_j^{\text{TB}}; -\frac{\tau_{i,j}^{h,b_w}}{\eta_j^{h,b_w}}, -\frac{1}{\eta_j^{h,b_w}} \right), \quad (64)$$

$$q_{h,j} \propto \mathcal{CN} \left( h_j^{\text{TB}}; -\frac{\sum_{i=0}^{MN_{\text{v}}N_{\text{F}}-1} \tau_{i,j}^{h,b_w} + \tau_j^{h,b_h}}{MN_{\text{v}}N_{\text{F}}\eta_j^{h,b_w} + \eta_j^{h,b_h}}, -\frac{MN_{\text{v}}N_{\text{F}}}{MN_{\text{v}}N_{\text{F}}\eta_j^{h,b_w} + \eta_j^{h,b_h}} \right), \quad (65)$$

$$b_{h,j} \propto p(h_j^{\text{TB}}) \mathcal{CN} \left( h_j^{\text{TB}}; -\frac{\tau_j^{h,b_h}}{\eta_j^{h,b_h}}, -\frac{1}{\eta_j^{h,b_h}} \right). \quad (66)$$

Next, for constraints (43), let (XX-1), (XX-2), (XX-3) denote the equation between the first and the second term, the second and the third term, the first and third term, respectively, for brevity. For example, (43a-1) denotes  $\mathbb{E} \{ w_i | b_{y,i} \} = \mathbb{E} \{ w_i | b_{w,i} \}$ , (43a-2) denotes  $\mathbb{E} \{ w_i | b_{w,i} \} = \mathbb{E} \{ w_i | q_{w,i} \}$  and (43a-3) denotes  $\mathbb{E} \{ w_i | b_{y,i} \} = \mathbb{E} \{ w_i | q_{w,i} \}$ . Let  $a_{ij}$  denote the  $(i, j)$ -th element of  $\mathbf{A}$ . Note that the modulus of each element of  $\mathbf{A}$  is the same, i.e.,  $\sigma_{\text{p}}$ .

According to (43b-3) and (43d-3), we can obtain that

$$\eta_j^{h,b_w} = -\frac{1}{\text{Var} \{ h_j^{\text{TB}} | b_{h,j} \}} - \frac{\eta_j^{h,b_h}}{MN_{\text{v}}N_{\text{F}}}. \quad (67)$$

According to (43a-2) and (43c-2), we can obtain that

$$\eta_i^{w,b_y} = \left( \sum_{j=0}^{N_{\text{an}}N_{\text{de}}N_{\text{do}}U-1} \frac{\sigma_{\text{p}}^2}{\eta_j^{h,b_w}} \right)^{-1} \triangleq \tilde{\eta}^{w,b_y}. \quad (68)$$

According to (43a-3), (43c-3) and  $p(y_i | w_i) \propto \mathcal{CN}(y_i; w_i, \sigma_z^2)$ , we can obtain that  $\eta_i^{w,b_w} = -1/\sigma_z^2$  and  $\tau_i^{w,b_w} = y_i/\sigma_z^2$ . Then according to (43b-2), (43d-2) and (68), we can obtain that

$$\eta_j^{h,b_h} = \left( \left( \tilde{\eta}^{w,b_y} \right)^{-1} - \sigma_z^2 - \frac{1}{MN_{\text{v}}N_{\text{F}}\eta_j^{h,b_w}} \right)^{-1}, \quad (69)$$

$$\begin{aligned} & \lim_{M \rightarrow \infty} \frac{1}{MN_{\text{v}}N_{\text{F}}} \mathbf{R}_u^{\text{SFT,P}} \mathbf{X}_u^{\text{H}} \mathbf{C} \mathbf{X}_u \mathbf{R}_u^{\text{SFT,P}} \stackrel{(a)}{=} \lim_{M \rightarrow \infty} \frac{1}{MN_{\text{v}}N_{\text{F}}} \mathbf{R}_u^{\text{SFT,P}} \mathbf{X}_u^{\text{H}} \bar{\mathbf{C}}_u \mathbf{X}_u \mathbf{R}_u^{\text{SFT,P}} \\ & \Rightarrow \lim_{M \rightarrow \infty} \frac{1}{MN_{\text{v}}N_{\text{F}}} \text{tr} \{ \mathbf{R}_u^{\text{SFT,P}} - \mathbf{R}_u^{\text{SFT,P}} \mathbf{X}_u^{\text{H}} \mathbf{C}^{-1} \mathbf{X}_u \mathbf{R}_u^{\text{SFT,P}} \} \\ & = \lim_{M \rightarrow \infty} \frac{1}{MN_{\text{v}}N_{\text{F}}} \text{tr} \{ \mathbf{R}_u^{\text{SFT,P}} - \mathbf{R}_u^{\text{SFT,P}} \mathbf{X}_u^{\text{H}} \bar{\mathbf{C}}_u^{-1} \mathbf{X}_u \mathbf{R}_u^{\text{SFT,P}} \}, \end{aligned} \quad (60)$$

$$\tau_j^{h,b_h} = \sum_{i=0}^{MN_v N_F - 1} \frac{a_{ij}^* (y_i + \psi_i) - \sigma_p^2 \frac{\tau_{i,j}^{h,b_w}}{\eta_j^{h,b_w}}}{(\tilde{\eta}^{w,b_y})^{-1} - \sigma_z^2 - \frac{\sigma_p^2}{\eta_j^{h,b_w}}} = \sum_{i=0}^{MN_v N_F - 1} \mu_{i,j}, \quad (70)$$

where

$$\psi_i \triangleq \sum_{j=0}^{N_{an} N_{de} N_{do} U - 1} a_{ij} \frac{\tau_{i,j}^{h,b_w}}{\eta_j^{h,b_w}}, \quad (71)$$

$$\mu_{i,j} \triangleq - \frac{a_{ij}^* (y_i + \psi_i) - \sigma_p^2 \frac{\tau_{i,j}^{h,b_w}}{\eta_j^{h,b_w}}}{(\tilde{\eta}^{w,b_y})^{-1} - \sigma_z^2 - \frac{\sigma_p^2}{\eta_j^{h,b_w}}}. \quad (72)$$

According to (69) and (70), we have

$$\varpi_j \triangleq - \frac{\tau_j^{h,b_h}}{\eta_j^{h,b_h}} = \frac{1}{MN_v N_F \sigma_p^2} \sum_{i=0}^{MN_v N_F - 1} \left( a_{ij}^* (y_i + \psi_i) - \sigma_p^2 \frac{\tau_{i,j}^{h,b_w}}{\eta_j^{h,b_w}} \right) \quad (73)$$

According to (43b-3) and (43d-3), we can obtain that

$$\begin{aligned} \sum_{i=0}^{MN_v N_F - 1} \tau_{i,j}^{h,b_w} &= \frac{MN_v N_F \mathbb{E} \{ h_j^{TB} | b_{h,j} \}}{\text{Var} \{ h_j^{TB} | q_{h,j} \}} - \tau_j^{h,b_h} \\ &= \frac{MN_v N_F \mathbb{E} \{ h_j^{TB} | b_{h,j} \}}{\text{Var} \{ h_j^{TB} | b_{h,j} \}} - \tau_j^{h,b_h}. \end{aligned} \quad (74)$$

According to (70), (74) and the constraint  $\mathbb{E} \{ h_j^{TB} | b_{w,i} \} = \mathbb{E} \{ h_j^{TB} | b_{w,i'} \}$ , we can obtain that

$$\tau_{i,j}^{h,b_w} = \frac{\mathbb{E} \{ h_j^{TB} | b_{h,j} \}}{\text{Var} \{ h_j^{TB} | b_{h,j} \}} - \tau_j^{h,b_h} + \sum_{i'=0, i' \neq i}^{MN_v N_F - 1} \mu_{i',j}. \quad (75)$$

In HF skywave massive MIMO-OFDM communications,  $MN_v N_F$  is usually very large. Hence we approximate that  $\sum_{i'=0, i' \neq i}^{MN_v N_F - 1} \mu_{i',j} \approx \sum_{i=0}^{MN_v N_F - 1} \mu_{i,j}$ . The approximation here is conducive to the derivation of the low-complexity implementation, and the resulting channel estimation algorithm still has a satisfactory performance, as shown in simulation results. Then  $\tau_{i,j}^{h,b_w}$  is rewritten as

$$\tau_{i,j}^{h,b_w} = \frac{\mathbb{E} \{ h_j^{TB} | b_{h,j} \}}{\text{Var} \{ h_j^{TB} | b_{h,j} \}} \triangleq \tilde{\tau}_j^{h,b_w}. \quad (76)$$

We define  $\eta_j^{h,b_w}$ ,  $\eta_j^{h,b_h}$ ,  $\tilde{\tau}_j^{h,b_w}$  and  $\varpi$ , whose  $j$ -th element are  $\eta_j^{h,b_w}$ ,  $\eta_j^{h,b_h}$ ,  $\tilde{\tau}_j^{h,b_w}$  and  $\varpi_j$ , respectively. In addition, we define  $\psi$ , whose  $i$ -th element is  $\psi_i$ . Then (66), (67), (68), (69), (71), (73) and (76) can be rewritten in vector and matrix form and constitute the CBFEM based channel estimation algorithm.

## REFERENCES

- [1] D. Shi, L. Song, W. Zhou, X. Q. Gao, C.-X. Wang, and G. Y. Li, "Channel estimation for HF skywave massive MIMO-OFDM with triple-beam based channel model," in *Proc. IEEE GLOBECOM*, Rio de Janeiro, Brazil, Dec. 2022.
- [2] J. Wang, G. Ding, and H. Wang, "HF communications: Past, present, and future," *China Commun.*, vol. 15, no. 9, pp. 1-9, 2018.
- [3] M. Uysal and M. Heidarpour, "Cooperative communication techniques for future-generation HF radios," *IEEE Commun. Mag.*, vol. 50, no. 10, pp. 56-63, Oct. 2012.
- [4] S. D. Gunashekar, E. M. Warrington, and S. Salous, "Investigations into the feasibility of multiple input multiple output techniques within the HF band: Preliminary results," *Radio Sci.*, vol. 44, no. 1, pp. 1-15, Feb. 2009.
- [5] Y. Erhel, D. Lemur, M. Oger, J. Le Masson, and F. Marie, "Evaluation of ionospheric HF MIMO channels: Two complementary circular polarizations reduce correlation," *IEEE Antennas Propag. Mag.*, vol. 58, no. 6, pp. 38-48, Dec. 2016.
- [6] S. B. Sahay, K. C. Bhagwat, and P. R. J. Mohan, "Exploitation of MIMO techniques for reliable HF communication," in *Proc. Int. Conf. Signal Process. Commun. (SPCOM)*, Bangalore, India, Jul. 2012, pp. 1-4.
- [7] 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.
- [8] X. Yu, A.-A. Lu, X. Gao, G. Y. Li, G. Ding, and C.-X. Wang, "HF skywave massive MIMO communication," *IEEE Trans. Wireless Commun.*, vol. 21, no. 4, pp. 2769-2785, Apr. 2022.
- [9] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 264-273, Mar. 2013.
- [10] L. You et al., "Pilot reuse for massive MIMO transmission over spatially correlated Rayleigh fading channels," *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3352-3366, Jun. 2015.
- [11] L. You et al., "Pilot reuse for vehicle-to-vehicle underlay massive MIMO transmission," *IEEE Trans. Veh. Technol.*, vol. 69, no. 5, pp. 5693-5697, May 2020.
- [12] H. Xie, F. Gao, S. Zhang, and S. Jin, "A unified transmission strategy for TDD/FDD massive MIMO systems with spatial basis expansion model," *IEEE Trans. Veh. Technol.*, vol. 66, no. 4, pp. 3170-3184, Apr. 2017.
- [13] H. Xie, F. Gao, S. Jin, J. Fang, and Y.-C. Liang, "Channel estimation for TDD/FDD massive MIMO systems with channel covariance computing," *IEEE Trans. Wireless Commun.*, vol. 17, no. 6, pp. 4206-4218, Jun. 2018.
- [14] X. Rao and V. K. N. Lau, "Compressive sensing with prior support quality information and application to massive MIMO channel estimation with temporal correlation," *IEEE Trans. Signal Process.*, vol. 63, no. 18, pp. 4914-4924, Sep. 2015.
- [15] Z. Gao, L. Dai, W. Dai, B. Shim, and Z. Wang, "Structured compressive sensing-based spatio-temporal joint channel estimation for FDD massive MIMO," *IEEE Trans. Commun.*, vol. 64, no. 2, pp. 601-617, Feb. 2016.
- [16] X. Rao and V. K. N. Lau, "Distributed compressive CSIT estimation and feedback for FDD multi-user massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3261-3271, Jun. 2014.
- [17] M. Masood, L. H. Afify, and T. Y. Al-Naffouri, "Efficient coordinated recovery of sparse channels in massive MIMO," *IEEE Trans. Signal Process.*, vol. 63, no. 1, pp. 104-118, Jan. 2015.
- [18] C.-K. Wen, S. Jin, K.-K. Wong, J.-C. Chen, and P. Ting, "Channel estimation for massive MIMO using Gaussian-mixture Bayesian learning," *IEEE Trans. Wireless Commun.*, vol. 14, no. 3, pp. 1356-1368, Mar. 2015.
- [19] A. Liu, L. Lian, V. K. N. Lau, and X. Yuan, "Downlink channel estimation in multiuser massive MIMO with hidden Markovian sparsity," *IEEE Trans. Signal Process.*, vol. 66, no. 18, pp. 4796-4810, Sep. 2018.
- [20] J. Mo, P. Schniter, and R. W. Heath, Jr., "Channel estimation in broadband millimeter wave MIMO systems with few-bit ADCs," *IEEE Trans. Signal Process.*, vol. 66, no. 5, pp. 1141-1154, Mar. 2018.
- [21] X. Kuai, L. Chen, X. Yuan, and A. Liu, "Structured turbo compressed sensing for downlink massive MIMO-OFDM channel estimation," *IEEE Trans. Wireless Commun.*, vol. 18, no. 8, pp. 3813-3826, Aug. 2019.
- [22] D. Zhang, X. Song, W. Wang, G. Fettweis, and X. Gao, "Unifying message passing algorithms under the framework of constrained Bethe free energy minimization," *IEEE Trans. Wireless Commun.*, vol. 20, no. 7, pp. 4144-4158, Jul. 2021.
- [23] X. Liu, W. Wang, X. Song, X. Gao, and G. Fettweis, "Sparse channel estimation via hierarchical hybrid message passing for massive MIMO-OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 20, no. 11, pp. 7118-7134, Nov. 2021.
- [24] X. Lin, S. Wu, C. Jiang, L. Kuang, J. Yan, and L. Hanzo, "Estimation of broadband multiuser millimeter wave massive MIMO-OFDM channels by exploiting their sparse structure," *IEEE Trans. Wireless Commun.*, vol. 17, no. 6, pp. 3959-3973, Jun. 2018.

- [25] Z. Gao, L. Dai, Z. Wang, and S. Chen, "Spatially common sparsity based adaptive channel estimation and feedback for FDD massive MIMO," *IEEE Trans. Signal Process.*, vol. 63, no. 23, pp. 6169–6183, Dec. 2015.
- [26] K. Dovelos, M. Matthaiou, H. Q. Ngo, and B. Bellalta, "Channel estimation and hybrid combining for wideband terahertz massive MIMO systems," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 6, pp. 1604–1620, Jun. 2021.
- [27] S. Xia, Y. Shi, Y. Zhou, and X. Yuan, "Reconfigurable intelligent surface for massive connectivity: Joint activity detection and channel estimation," *IEEE Trans. Signal Process.*, vol. 69, pp. 5693–5707, 2021.
- [28] H. Liu, X. Yuan, and Y.-J.-A. Zhang, "Matrix-calibration-based cascaded channel estimation for reconfigurable intelligent surface assisted multiuser MIMO," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2621–2636, Nov. 2020.
- [29] K. Venugopal, A. Alkhateeb, N. González-Prelcic, and R. W. Heath, Jr., "Channel estimation for hybrid architecture-based wideband millimeter wave systems," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 1996–2009, Sep. 2017.
- [30] Y. Zhou, M. Herdin, A. M. Sayeed, and E. Bonek, "Experimental study of MIMO channel statistics and capacity via the virtual channel representation," Univ. Wisconsin-Madison, Madison, WI, USA, Tech. Rep., Feb. 2007.
- [31] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak, "Compressed channel sensing: A new approach to estimating sparse multipath channels," *Proc. IEEE*, vol. 98, no. 6, pp. 1058–1076, Jun. 2010.
- [32] L. Gu, W. Wang, W. Zhong, and X. Gao, "Message-passing detector for uplink massive MIMO systems based on energy spread transform," in *Proc. IEEE 27th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, Sep. 2016, pp. 1–6.
- [33] A.-A. Lu, X. Gao, W. Zhong, C. Xiao, and X. Meng, "Robust transmission for massive MIMO downlink with imperfect CSI," *IEEE Trans. Commun.*, vol. 67, no. 8, pp. 5362–5376, Aug. 2019.
- [34] B. Wang, F. Gao, S. Jin, H. Lin, and G. Y. Li, "Spatial- and frequency-wideband effects in millimeter-wave massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 66, no. 13, pp. 3393–3406, May 2018.
- [35] W. Shen, L. Dai, J. An, P. Fan, and R. W. Heath, Jr., "Channel estimation for orthogonal time frequency space (OTFS) massive MIMO," *IEEE Trans. Signal Process.*, vol. 67, no. 16, pp. 4204–4217, Aug. 2019.
- [36] D. Shi et al., "Deterministic pilot design and channel estimation for downlink massive MIMO-OTFS systems in presence of the fractional Doppler," *IEEE Trans. Wireless Commun.*, vol. 20, no. 11, pp. 7151–7165, Nov. 2021.
- [37] L. You, X. Q. Gao, A. L. Swindlehurst, and W. Zhong, "Channel acquisition for massive MIMO-OFDM with adjustable phase shift pilots," *IEEE Trans. Signal Process.*, vol. 64, no. 6, pp. 1461–1476, Mar. 2016.
- [38] A.-A. Lu, Y. Chen, and X. Gao, "2D beam domain statistical CSI estimation for massive MIMO uplink," 2022, *arXiv:2207.04695*.
- [39] G. Wunder, S. Stefanatos, A. Flinthe, I. Roth, and G. Caire, "Low-overhead hierarchically-sparse channel estimation for multiuser wideband massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 18, no. 4, pp. 2186–2199, Apr. 2019.
- [40] D. C. Chu, "Polyphase codes with good periodic correlation properties," *IEEE Trans. Inf. Theory*, vol. IT-18, no. 4, pp. 531–532, Jul. 1972.
- [41] E. Dahlman, S. Parkvall, and J. Skold, *5G NR: The Next Generation Wireless Access Technology*. New York, NY, USA: Academic, 2018.
- [42] J. Oyelade et al., "Data clustering: Algorithms and its applications," in *Proc. Int. Conf. Comput. Sci. Appl. (ICCSA)*, Saint Petersburg, Russia, Jul. 2019, pp. 71–81.
- [43] A. L. Moustakas, G. C. Alexandropoulos, A. Polydoros, I. Kaddas, and I. Dagres, "Impact of imperfect channel estimation in HF OFDM-MIMO communications," in *Proc. IEEE 30th Annu. Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, Sep. 2019, pp. 1–6.
- [44] M. Balser and W. B. Smith, "Some statistical properties of pulsed oblique HF ionospheric transmissions," *J. Res. Nat. Bur. Standards-D, Radio Propag.*, vol. 66D, no. 6, pp. 721–730, Nov. 1962.
- [45] N. Maslin, *HF Communications: A Systems Approach*. London, U.K.: Pitman, 1987.
- [46] K. T. Truong and R. W. Heath, "Effects of channel aging in massive MIMO systems," *J. Commun. Netw.*, vol. 15, no. 4, pp. 338–351, 2013.
- [47] S. Rangan, "Generalized approximate message passing for estimation with random linear mixing," in *Proc. IEEE Int. Symp. Inf. Theory*, Saint Petersburg, Russia, Jul. 2011, pp. 2168–2172.
- [48] D. G. Tzikas, A. C. Likas, and N. P. Galatsanos, "The variational approximation for Bayesian inference," *IEEE Signal Process. Mag.*, vol. 25, no. 6, pp. 131–146, Jan. 2008.
- [49] B. J. Frey, *Graphical Models for Machine Learning and Digital Communication*. Cambridge, MA, USA: MIT Press, 1998.
- [50] J. S. Yedidia, W. T. Freeman, and Y. Weiss, "Constructing free-energy approximations and generalized belief propagation algorithms," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2282–2312, Jul. 2005.
- [51] T. Heskes and O. Zoeter, "Expectation propagation for approximate inference in dynamic Bayesian networks," in *Proc. Uncertainty Artif. Intell.*, Calgary, AB, Canada, 2002, pp. 216–223.
- [52] L. Rabiner, R. Schafer, and C. Rader, "The chirp z-transform algorithm," *IEEE Trans. Audio Electroacoust.*, vol. AU-17, no. 2, pp. 86–92, Jun. 1969.
- [53] Solar Terrestrial Dispatch. *Proplab-Pro Version 3*. Accessed: Oct. 18, 2018. [Online]. Available: <http://www.spacew.com/proplab/index.html>
- [54] *Testing of HF Modems With Bandwidths of up to About 12 KHz Using Ionospheric Channel Simulators*, document ITU-R F.1487, 2000. [Online]. Available: <https://www.itu.int/rec/R-REC-F.1487-0-200005-1/en>
- [55] G. A. Seber, *A Matrix Handbook for Statisticians*. Hoboken, NJ, USA: Wiley, 2008.



**Ding Shi** (Graduate Student Member, IEEE) received the B.E. degree in information science and engineering from Southeast University, Nanjing, China, in 2018, where he is currently pursuing the Ph.D. degree with the National Mobile Communications Research Laboratory. His research interests include signal processing and massive MIMO communications.



**Linfeng Song** (Graduate Student Member, IEEE) received the B.E. degree in radio wave propagation and antenna from Wuhan University, Wuhan, China, in 2020. He is currently pursuing the M.E. degree with the National Mobile Communications Research Laboratory, Southeast University, Nanjing, China. His research interests include massive MIMO communications and HF communications.



**Wenqi Zhou** (Student Member, IEEE) received the B.Sc. degree in electrical engineering and automation from Shandong Construction University, China, in 2001, and the M.Eng. degree in control theory and control engineering from Shandong University, China, in 2006. He is currently pursuing the Ph.D. degree with Southeast University, China. His research interests include 6G scenario classification and identification and AI-based predictive channel modeling.



**Xiqi Gao** (Fellow, IEEE) received the Ph.D. degree in electrical engineering from Southeast University, Nanjing, China, in 1997.

He joined the Department of Radio Engineering, Southeast University, in April 1992. Since May 2001, he has been a Professor of information systems and communications. From September 1999 to August 2000, he was a Visiting Scholar at Massachusetts Institute of Technology, Cambridge; and Boston University, Boston, MA, USA. From August 2007 to July 2008, he has visited the Darmstadt

University of Technology, Darmstadt, Germany, as a Humboldt Scholar. His current research interests include broadband multicarrier communications, massive MIMO wireless communications, satellite communications, optical wireless communications, information theory, and signal processing for wireless communications. Dr. Gao received the Science and Technology Awards of the State Education Ministry of China in 1998, 2006, and 2009; the National Technological Invention Award of China in 2011; the Science and Technology Award of Jiangsu Province of China in 2014; and the 2011 IEEE Communications Society Stephen O. Rice Prize Paper Award in the field of communications theory. From 2007 to 2012, he served as the Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. From 2009 to 2013, he served as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING. From 2015 to 2017, he served as the Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS.



**Cheng-Xiang Wang** (Fellow, IEEE) received the B.Sc. and M.Eng. degrees in communication and information systems from Shandong University, China, in 1997 and 2000, respectively, and the Ph.D. degree in wireless communications from Aalborg University, Denmark, in 2004.

He was a Research Assistant with the Hamburg University of Technology, Hamburg, Germany, from 2000 to 2001; a Visiting Researcher with Siemens AG Mobile Phones, Munich, Germany, in 2004; and a Research Fellow with the University of Agder, Grimstad, Norway, from 2001 to 2005. He has been with Heriot-Watt University, Edinburgh, U.K., since 2005, where he was promoted to a Professor in 2011. In 2018, he joined Southeast University, Nanjing, China, as a Professor. He is also a part-time Professor with the Purple Mountain Laboratories, Nanjing. He has authored four books, three book chapters, and more than 480 papers in refereed journals and conference proceedings, including 27 highly cited papers. He has also delivered 24 invited keynote speeches/talks and 14 tutorials in international conferences. His current research interests include wireless channel measurements and modeling, 6G wireless communication networks, and electromagnetic information theory.

Dr. Wang is a member of the Academia Europaea (The Academy of Europe) and the European Academy of Sciences and Arts (EASA); and a fellow of the Royal Society of Edinburgh (FRSE), IET, and China Institute of Communications (CIC). He is a Highly-Cited Researcher recognized by Clarivate Analytics from 2017 to 2020 and one of the most cited Chinese Researchers recognized by Elsevier in 2021. He is currently an Executive Editorial Committee Member of the *IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS*. He has served as a TPC member, a TPC chair, and a general chair for more than 80 international conferences. He received 15 Best Paper Awards from IEEE GLOBECOM 2010, IEEE ICCT 2011, ITST 2012, IEEE VTC 2013 Spring, IWCMC 2015, IWCMC 2016, IEEE/CIC ICC 2016, WPMC 2016, WOCC 2019, IWCMC 2020, WCSP 2020, CSPS2021, WCSP 2021, and IEEE/CIC ICC 2022. He also received the 2020–2022 “AI 2000 Most Influential Scholar Award Honourable Mention” in recognition of his outstanding and vibrant contributions in the field of the Internet of Things. He has served as the Editor for over ten international journals, including the *IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS*, from 2007 to 2009, the *IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY*, from 2011 to 2017, and the *IEEE TRANSACTIONS ON COMMUNICATIONS*, from 2015 to 2017. He was a Guest Editor of the *IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS*, Special Issue on Vehicular Communications and Networks (a Lead Guest Editor), Special Issue on Spectrum and Energy Efficient Design of Wireless Communication Networks, and Special Issue on Airborne Communication Networks. He was also a Guest Editor of the *IEEE TRANSACTIONS ON BIG DATA*, Special Issue on Wireless Big Data. He is a Guest Editor of the *IEEE TRANSACTIONS ON COGNITIVE COMMUNICATIONS AND NETWORKING*, Special Issue on Intelligent Resource Management for 5G and Beyond. He is an IEEE Communications Society Distinguished Lecturer from 2019 to 2020.



**Geoffrey Ye Li** (Fellow, IEEE) is currently a Chair Professor at the Imperial College London, U.K. Before joining Imperial in 2020, he was a Professor at the Georgia Institute of Technology, USA, for 20 years and a Principal Technical Staff Member with AT&T Labs–Research in NJ, USA, for five years. His general research interests include statistical signal processing and machine learning for wireless communications. In the related areas, he has published over 600 journal and conference papers. In addition to over 40 granted patents and several

books. His publications have been cited over 56,000 times with an H-index over 110. He has been recognized as a Highly Cited Researcher, by Thomson Reuters, almost every year.

He was awarded a IET fellow for his contributions to signal processing for wireless communications. He has won several prestigious awards from IEEE Signal Processing, Vehicular Technology, and Communications Societies, including IEEE ComSoc Edwin Howard Armstrong Achievement Award in 2019. He also received the 2015 Distinguished ECE Faculty Achievement Award from Georgia Tech. He has organized and chaired many international conferences, including the Technical Program Vice-Chair of the IEEE ICC'03; the General Co-Chair of the IEEE GlobalSIP'14, the IEEE VTC'19 Fall, the IEEE SPAWC'20, and IEEE VTC'22 Fall. He has been involved in editorial activities for over 20 technical journals, including the founding Editor-in-Chief of *IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS* Special Series on ML in Communications and Networking.