

Wireless Channel Map Enabled Instantaneous Channel State Information Acquisition in High-Mobility Scenarios

Yinglan Bu, Cheng-Xiang Wang[✉], Fellow, IEEE, Chen Huang[✉], Member, IEEE,
Shuaifei Chen[✉], Member, IEEE, Junling Li[✉], Member, IEEE, Jianghan Ji, and Yunfei Chen[✉], Fellow, IEEE

Abstract—High-mobility and large-bandwidth applications at high frequencies make the acquisition of doubly-selective channels costly and complex, as the fast fading channel causes a surge in pilot overheads and inter-carrier interference. Accurate estimation of the complex channel gain (CG) and the carrier frequency offset (CFO) is required to support subsequent high-accuracy channel prediction that alleviates the heavy pilot burden. The emerging technology of the wireless channel map (WCM) can approximately reproduce the actual propagation environment digitally with customized channel parameters, offering an opportunity to accurately estimate the complex CG and CFO. In this paper, a WCM-based prior distribution construction method and a joint complex CG and CFO estimation algorithm are proposed. Specifically, a parameterized linear estimation problem for the complex CG is generated based on a nonuniform delay-domain off-grid channel representation, and with more realistic prior distributions constructed by the knowledge from the WCM-provided angular-delay power spectrum density, the joint estimation problem is solved under the Bayesian inference framework. Simulation results demonstrate the superiority of the proposed algorithm, with a better performance in terms of estimation accuracy and bit error rates (BER) than existing baselines. It is also verified that the proposed WCM-based algorithm is highly adaptable to different WCM precision and robust to different user speeds.

Index Terms—Channel estimation, compressed sensing, high mobility, inter-carrier interference, wireless channel map.

Received 14 April 2025; revised 29 August 2025 and 1 January 2026; accepted 28 January 2026. Date of publication 10 February 2026; date of current version 19 February 2026. This work was supported by the National Natural Science Foundation of China (NSFC) under Grants 62394290, 62394291, and 62394294, and the Research Fund of National Mobile Communications Research Laboratory, Southeast University, under Grant 2026A05. The associate editor coordinating the review of this article and approving it for publication was S. Majhi. (*Corresponding authors:* Cheng-Xiang Wang; Chen Huang.)

Yinglan Bu, Cheng-Xiang Wang, and Junling Li are with the National Mobile Communications Research Laboratory, School of Information Science and Engineering, Southeast University, Nanjing 211189, China, and also with the Purple Mountain Laboratories, Nanjing 211111, China (e-mail: ylbu@seu.edu.cn; chxwang@seu.edu.cn; junlingli@seu.edu.cn).

Chen Huang and Shuaifei Chen are with the Purple Mountain Laboratories, Nanjing 211111, China, and also with the National Mobile Communications Research Laboratory, School of Information Science and Engineering, Southeast University, Nanjing 211189, China (e-mail: huangchen@pmlabs.com.cn; shuaifeichen@seu.edu.cn).

Jianghan Ji is with the National Mobile Communications Research Laboratory, School of Information Science and Engineering, Southeast University, Nanjing 211189, China (e-mail: ji_jh@seu.edu.cn).

Yunfei Chen is with the Department of Engineering, University of Durham, DH1 3LE Durham, U.K. (e-mail: yunfei.chen@durham.ac.uk).

Digital Object Identifier 10.1109/TCOMM.2026.3663282

I. INTRODUCTION

HIGH frequencies and large bandwidths enable high transmission rates. The centimeter- and millimeter-waves have already been incorporated in the third generation partnership project (3GPP) standards [1]. Besides, multiple emerging communication technologies, such as satellite and high-speed train communications, become necessary components in the future three-dimensional (3D) coverage communication networks [2], [3], making mobility a prominent feature. Their implementation imposes great challenges on channel acquisition, as the severe Doppler frequency shift (DFS) induced by the high mobility of mobile user equipment (UE) and moving scatterers will lead to rapid channel aging and inter-carrier interference (ICI) in multi-carrier systems. In fact, ICI can also provide additional Doppler diversity to improve the system performance [4]. However, this requires support of full-dimensional channel information in space-time-frequency domains, and will impose large pilot overhead with frequent channel sounding. Good estimates of key channel parameters—the complex channel gain (CG) and carrier frequency offset (CFO)—can enable subsequent high-accuracy model-based channel predictions [5], and then help alleviate the heavy pilot burden by extending the channel sounding period.

The CFO contains two main components: the frequency mismatch between transmitter and receiver oscillators and the DFS. The CFO can be obtained by blind estimation with special signal structures [6] or be measured by redundant samples [7], [8], [9], [10], [11]. Efficient approaches to estimating the complex CG and CFO jointly have also been extensively studied. In [7] and [8], the joint estimation algorithms employed maximum likelihood estimators (MLE) to solve the joint estimation problem with expectation-maximization (EM) methods and coordinate descent optimization, respectively. In [9], another MLE for the joint estimation was proposed and solved by oblique non-orthogonal projection. Furthermore, extended Kalman and particle filtering methods were also employed to solve such a joint estimation problem in [10] and [11], respectively. However, in these works, CFOs being considered only referred to the oscillator frequency mismatch, while DFSs were mostly encompassed in the time-varying channels and therefore cannot be extracted for the subsequent

time-domain channel prediction. Furthermore, current joint estimation algorithms were often proposed based on simplified channel models, such as the Jake's channel model, which cannot reflect the actual wireless propagation environment.

Channel sparsity can be employed to reduce computational complexity and increase estimation accuracy. For example, by constructing Bayesian networks with hidden Markov prior, the structured channel sparsity in antenna and angle domains was analyzed and utilized to improve the estimation accuracy and efficiency in [12] and [13], respectively. For high-mobility scenarios, a distributed compressive sensing (DCS) method and its improvements were proposed in [14], [15], and [16] to reformulate the doubly-selective channel estimation as a delay-domain sparse coefficient recovery problem to achieve high estimation accuracy with simultaneous orthogonal matching pursuit and smoothing. Though the DCS method is easy to implement, it cannot estimate more than one random variable, and the obtained channel is only available for adjacent symbols. Another method called sparse Bayesian learning (SBL) can achieve higher estimation accuracy with a combination of multipath sparsity priors in some specific domains [17]. However, the continuous parameters of multipath components, such as the angles, need to be discretized for the analysis of sparsity in the corresponding domain, as the actual number of channel paths is unknown. This will impose another problem of limited resolution. Therefore, a model-based off-grid channel representation was proposed in [18] to resolve the sparse multipath in the angle domain using the geometry-based stochastic channel model (GBSM). A joint estimation algorithm with such an off-grid representation was proposed in [19] based on SBL to simultaneously obtain the complex CG and DFS of narrow-band fast-fading channels, and a Doppler compensation for such a scenario was also proposed by leveraging massive MIMO beamforming. However, only two-dimensional (2D) terminal trajectories are considered in current GBSM-based solutions, while the mobility of scatterers is often ignored. For doubly-selective channels with ICI, the joint estimation problem of the doubly-selective channel and CFO containing DFSs is yet to be solved.

Using more precise channel prior information, better estimation performance can be achieved [20]. Additional prior knowledge can help improve the channel estimation performance, as verified in [21] that more accurate channel estimation and higher transmission efficiencies can be achieved with a channel power map. Based on the integrated interpolation and data/model dual-driven method, previous research in [22] and [23] have constructed static channel power and channel gain maps with accuracy proved by the mean square error (MSE) performance, respectively. With the emergence of integrated sensing and communications (ISAC), a wireless channel map (WCM) that can depict the dynamic wireless propagation environment was proposed in [24], which can provide necessary prior information in the form of customized channel parameters. In [25], a detailed WCM construction method was proposed based on semi-deterministic channel modeling to enable an accurate location-based annotation of dynamic large-scale channel parameters through environmental parameter calibration based on channel mea-

surement data and ambient scatterer identification/localization. For static clusters like buildings or the LOS component that only relates to the UE position, ray tracing was employed and trained offline [26]. The multipath components from dynamic clusters are generated according to a cluster-based channel modeling method, with dynamic clusters identified and located by the point cloud and computer vision method. In [27], the authors further proved that such a method can achieve a digital large-scale channel parameter mapping, such as the delay PSD, that is highly consistent with actual measurement data.

In this paper, a WCM-based joint complex CG and CFO estimation algorithm is proposed. A WCM-based prior distribution construction method is also proposed to enhance the accuracy of the joint estimation and the bit error rate (BER) performance. The main contributions of this paper can be summarized as follows.

- By using more practical channel models, the joint complex CG and CFO estimation for doubly selective channels with consideration of ICI is formulated as a dynamic compressed sensing problem based on a nonuniform off-grid channel representation. Different from 2D terminal trajectories considered in the previous literature, such a joint estimation problem can cope with 3D terminal trajectories and the mobility of ambient scatterers. Besides, unlike previous studies where the CFO contains only the local oscillator instability, the CFO estimation in our scheme also includes DFSs that are usually distinct on different channel paths, which can make preparation for the subsequent model-based channel prediction.
- With the WCM-provided angular-delay power spectrum density (PSD) and the terminal mobility information, a prior distribution construction method for critical channel parameters of the complex CG and CFO is proposed, and can extract key prior knowledge, such as angle-delay domain channel sparsity and possible multipath directions from the WCM-provided ergodic channel information, thus achieving a more accurate description of the dynamic propagation environment. Based on more accurate prior distributions, the joint estimation is solved with MM algorithms under the Bayesian inference framework, and an efficient joint complex CG and CFO estimation algorithm is proposed accordingly.
- The proposed WCM-based algorithm is demonstrated to achieve higher estimation accuracy and lower BER than previous approaches without the WCM by simulation results, even when the WCM is lossy. Impacts of critical parameters, such as the WCM precision and UE speed, on the performance of the proposed algorithm are analyzed. Besides, the proposed algorithm is shown to be robust to different UE speeds, and the additional Doppler diversity gain at high speeds brought by the full-dimensional channel estimation considering ICI is confirmed and analyzed. It is also proven that, with prior distributions that are more consistent with actual propagation environments, better communication performance can be achieved.

The rest of this paper is organized as follows. The pilot transmission model considering the ICI and nonuniform off-grid channel representation is introduced in Section II. Details

TABLE I
DEFINITIONS OF ABBREVIATIONS IN THIS PAPER

Abbreviations	Definitions	Abbreviations	Definitions
CG	Channel gain	CFO	Carrier frequency offset
DFS	Doppler frequency shift	ICI	Inter-carrier interference
DCS	Distributed compressive sensing	WCM	Wireless channel map
UE	User equipment	BS	Base station
GBSM	Geometry-based stochastic channel model	OFDM	Orthogonal frequency-division multiplexing
LCS	Local coordinate system	GCS	Global coordinate system
(N)LOS	(Non-)line-of-sight	SBL	Sparse Bayesian learning
CIR	Channel impulse response	CTF	Channel transfer function
AoD	Azimuth angle of departure	EoD	Elevation angle of departure
MM	Minimization-maximization	PS	Parameter set
BER	Bit error rate	NMSE	Normalized mean square error

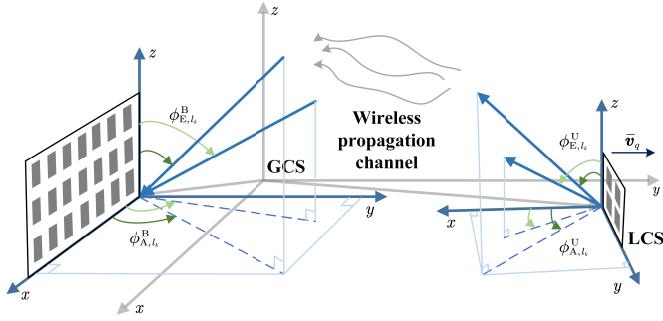


Fig. 1. Illustration of the considered wireless propagation channel.

of WCM-based prior distribution construction for critical channel parameters are given in Section III. In Section IV, the parameterized linear estimation problem for complex CGs is generated based on the reformulated pilot transmission model, and the proposed joint estimation algorithm is elaborated. Simulation results and conclusions are given in Section V and VI, respectively. Definitions of main abbreviations are given in Table I.

Notations: $(\cdot)^{-1}$, $(\cdot)^H$, and $(\cdot)^T$ denote inverse, conjugate transpose, and transpose, respectively. Bold lowercase indicates vectors, while bold uppercase indicates matrices. I_N denotes an $N \times N$ identity matrix. $\|\cdot\|$ denotes the Frobenius norm. $|\mathcal{A}|$ takes the cardinality of set \mathcal{A} . $|x|$ takes the absolute value of x . $\lceil \cdot \rceil$ denotes round up. \otimes and \odot denote the Kronecker product and Hadamard product, respectively. $[\cdot]_{a=1, \dots, A}$ and $[\cdot]_{a \in \mathcal{A}}$ both represent concatenation along the first dimension. $[x]_a$ and $[X]_{a,b}$ denote the a -th element in x and the (a, b) -th element in X , respectively. $[x]_{a_1:a_2}$ denotes the vector consisting of the a_1 -th to a_2 -th elements in x , $[X]_{a_1:a_2, b_1:b_2}$ denotes the matrix consisting of the intersection elements of rows a_1 to a_2 and columns b_1 to b_2 in X . $\text{supp}(\cdot)$ denotes the nonzero element index set. $\text{diag}(\cdot)$ denotes the diagonal operator.

II. SYSTEM MODEL

A. Doubly-Selective Channel Model

We investigate the joint complex CG and CFO estimation problem in high-mobility scenarios for wideband MIMO systems, in which the base station (BS) and UE are both equipped with a half-wavelength uniform planar array (UPA), and the centimeter-wave band is applied. The BS UPA has $M = M_2 \times M_1$ elements, with M_2 in each column and M_1 in

each row. The UE UPA has $N = N_2 \times N_1$ elements, with N_2 in each column and N_1 in each row. The wireless propagation channel and UPAs equipped at both sides are illustrated in Fig. 1. The BS array is placed on the xoz plane of the BS-side local coordinate system (LCS), while the UE array is placed on the yoz plane of the UE-LCS, both employing a right-hand system. The conversion from LCS coordinates to the global coordinate system (GCS) coordinates can be found in [28]. Since UE postures can be captured by a variety of built-in sensing devices, we assume that the BS is informed of the UE posture in a timely manner.

An orthogonal frequency-division multiplexing (OFDM) system with a bandwidth of f_s is considered. Each OFDM symbol contains P subcarriers with a subcarrier spacing of $\Delta f = \frac{f_s}{P}$. The sampling period and time duration of valid data within a symbol are denoted as T_s and T_d , respectively. After the data block is transformed by a normalized P -point inverse fast Fourier transform, a CP is inserted in the time domain with a length of T_{GI} , containing P_{GI} samples, to avoid the inter-symbol interference caused by frequency-selective fading, and is set larger than the maximum multipath delay. After passing through the transmit filter, wireless propagation channel, and receive filter, the temporal sequence will be removed from the CP and then transformed back into frequency-domain samples by a P -point fast Fourier transform. The entire symbol duration can be calculated as $T_{\text{symb}} = T_d + T_{GI} = (P + P_{GI})T_s$. Assume that the channel is quasi-static during a time block containing multiple channel uses [29], [30], which indicates that multipath components, including the multipath angles, delays, and power, stay unchanged because the physical position of a mobile device moving in a wide space generally varies more slowly than the transient channel. It is also assumed that the UE velocity and scatterer mobility remain unchanged within a time block. The line-of-sight (LOS) scenario is considered. The frequency responses of the transmit and receive filters are set to one at the flat region.

Referring to the pervasive channel model proposed in [31] and [32], the channel impulse response (CIR) between the m -th transmit antenna and n -th receive antenna for the i -th symbol in the q -th time block can be denoted as

$$h_{m,n,q,i}(t, \tau) = \sum_{l_q=0}^{L_q} \rho_{l_q} e^{j k_f \left((\hat{r}_{l_q}^U)^T \bar{d}_n^U \right)}$$

$$\times e^{jk_f \left((\hat{r}_{l_q}^B)^T \bar{d}_m^B \right)} e^{jk_f (\nu_{l_q}^D t)} \delta(\tau - \tau_{l_q} T_s), \quad (1)$$

where $k_f = \frac{2\pi}{\lambda_c}$ represents the wavenumber in free space, λ_c denotes the corresponding wavelength of the carrier frequency f_c , δ stands for the Dirac delta function, τ_{l_q} is the normalized delay with respect to T_s , \bar{d}_n^U and \bar{d}_m^B denote element positions in the LCSs of UE and BS arrays, respectively. It should be noticed that τ_{l_q} is not necessarily an integer. There are $L_q + 1$ paths at the q -th time block. Spherical unit vectors of path l_q in the UE and BS LCSs are denoted as $\hat{r}_{l_q}^U$ and $\hat{r}_{l_q}^B$, respectively. If $l_q = 0$ represents the LOS component, $\hat{r}_{l_q}^U$ and $\hat{r}_{l_q}^B$ can be written as

$$\hat{r}_{l_q}^U = \left[\sin \phi_{E,l_q}^U \cos \phi_{A,l_q}^U, \sin \phi_{E,l_q}^U \sin \phi_{A,l_q}^U, \cos \phi_{E,l_q}^U \right]^T, \quad (2)$$

$$\hat{r}_{l_q}^B = \left[\sin \phi_{E,l_q}^B \cos \phi_{A,l_q}^B, \sin \phi_{E,l_q}^B \sin \phi_{A,l_q}^B, \cos \phi_{E,l_q}^B \right]^T, \quad (3)$$

where $\phi_{A,l_q}^U \in (-\frac{\pi}{2}, \frac{\pi}{2})$, $\phi_{A,l_q}^B \in (0, \pi)$ and $\phi_{E,l_q}^U / \phi_{E,l_q}^B \in (0, \pi)$ denote the azimuth angle of departure (AoD) and the elevation angle of departure (EoD) of path l_q at the UE/BS side, respectively. Besides, ρ_{l_q} denotes the channel complex envelope on path l_q , which is assumed to be unchanged within a time block. The DFS on path l_q can be denoted as [28]

$$\nu_{l_q}^D = \left(\hat{r}_{l_q}^{U,GCS} \right)^T \bar{v}_q + \nu_{l_q}^S (1 - \delta(l_q)), \quad (4)$$

where $\bar{v}_q = v_q [\sin \psi_{E,q} \cos \psi_{A,q}, \sin \psi_{E,q} \sin \psi_{A,q}, \cos \psi_{E,q}]^T$ denotes the UE velocity, v_q is the UE moving speed, $\hat{r}_{l_q}^{U,GCS}$ corresponds to $\hat{r}_{l_q}^U$ in the GCS, $\psi_{A,q}$ and $\psi_{E,q}$ denote the azimuth and elevation angles of the UE traveling direction, respectively. From the GBSM, we know that the DFS consists of two components, one caused by the mobile terminal and the other caused by moving scatterers. The latter, as described in the 3GPP TR 38.901 [28], can be modeled as $\nu_{l_q}^S = 2\varpi_{l_q} v_{l_q}$, where v_{l_q} is a random variable evenly distributed between the negative and positive v_{scatt} , v_{scatt} denotes the maximum clutter speed, ϖ_{l_q} is a random variable representing the mobility of scatterers and follows a Bernoulli distribution with mean ζ , and ζ captures the statistical degree of scatterer mobility in the propagation environment. After sampling the channel with a time-domain spacing of T_s , the discrete CG of path l_q can be written as

$$\begin{aligned} \alpha_{m,n,q,i}^{l_q}(l) &= \rho_{l_q} e^{jk_f \left((\hat{r}_{l_q}^U)^T \bar{d}_n^U \right)} e^{jk_f \left((\hat{r}_{l_q}^B)^T \bar{d}_m^B \right)} \\ &\times e^{jk_f \nu_{l_q}^D (T_s l + (i-1) T_{symb})}, \end{aligned} \quad (5)$$

$$\begin{aligned} \left(\hat{r}_{l_q}^U \right)^T \bar{d}_n^U &= \sin \phi_{E,l_q}^U \sin \phi_{A,l_q}^U (n_1 - 1) d_c^U + \cos \phi_{E,l_q}^U \\ &\times (n_2 - 1) d_r^U, \end{aligned} \quad (6)$$

$$\begin{aligned} \left(\hat{r}_{l_q}^B \right)^T \bar{d}_m^B &= \sin \phi_{E,l_q}^B \cos \phi_{A,l_q}^B (m_1 - 1) d_c^B + \cos \phi_{E,l_q}^B \\ &\times (m_2 - 1) d_r^B, \end{aligned} \quad (7)$$

where d_c^B/d_c^U and d_r^B/d_r^U correspond to the distance between columns and rows at the BS/UE antenna array, respec-

tively. We further define equivalent angels along vertical and horizontal directions as $\varphi_{l_q}^B = \cos \phi_{E,l_q}^B \frac{d_r^B}{\lambda_c}$, $\vartheta_{l_q}^B = \sin \phi_{E,l_q}^B \cos \phi_{A,l_q}^B \frac{d_c^B}{\lambda_c}$, $\varphi_{l_q}^U = \cos \phi_{E,l_q}^U \frac{d_r^U}{\lambda_c}$, $\vartheta_{l_q}^U = \sin \phi_{E,l_q}^U \sin \phi_{A,l_q}^U \frac{d_c^U}{\lambda_c}$, to describe path directions in the 3D space. Steering vectors at both sides can be rewritten as

$$\mathbf{b}_{l_q}^B = \mathbf{u} \left(\varphi_{l_q}^B, M_2 \right)^T \otimes \mathbf{u} \left(\vartheta_{l_q}^B, M_1 \right)^T, \quad (8)$$

$$\mathbf{b}_{l_q}^U = \mathbf{u} \left(\varphi_{l_q}^U, N_2 \right)^T \otimes \mathbf{u} \left(\vartheta_{l_q}^U, N_1 \right)^T, \quad (9)$$

where $\mathbf{u}(\cdot)$ is defined as

$$\mathbf{u}(\theta, C) = \left[1, e^{j2\pi\theta}, \dots, e^{j2\pi\theta(C-1)} \right]^T \in \mathbb{C}^C, \quad (10)$$

and is a periodic function with a period of 1 about θ . Then, the CG in (5) can be rewritten in a matrix form as

$$\bar{H}_{q,i}^{l_q}(l) = \rho_{l_q} e^{jk_f \nu_{l_q}^D (T_s l + (i-1) T_{symb})} \left(\mathbf{b}_{l_q}^B \right)^T \mathbf{b}_{l_q}^U. \quad (11)$$

For better analyzing the relationship between the angular-delay PSD and path direction, we define an N -point discrete Fourier transform (DFT) matrix as

$$\mathbf{U}^U = \mathbf{U}_{el}^U \otimes \mathbf{U}_{az}^U \in \mathbb{C}^{N \times N}, \quad (12)$$

$$\mathbf{U}_{el}^U = \frac{1}{\sqrt{N_2}} \left[\mathbf{u} \left(\frac{2a-1}{2N_2} - 0.5, N_2 \right)^T \right]_{a=1, \dots, N_2} \in \mathbb{C}^{N_2 \times N_2}, \quad (13)$$

$$\mathbf{U}_{az}^U = \frac{1}{\sqrt{N_1}} \left[\mathbf{u} \left(\frac{2b-1}{2N_1} - 0.5, N_1 \right)^T \right]_{b=1, \dots, N_1} \in \mathbb{C}^{N_1 \times N_1}. \quad (14)$$

An M -point DFT matrix $\mathbf{U}^B = \mathbf{U}_{el}^B \otimes \mathbf{U}_{az}^B \in \mathbb{C}^{M \times M}$ for the other side is defined in the same way. The CG $\bar{H}_{q,i}^{l_q}$ can be mapped into the angle domain as [33]

$$\begin{aligned} \tilde{H}_{q,i}^{l_q}(l) &= \mathbf{U}^B \bar{H}_{q,i}^{l_q}(l) \left(\mathbf{U}^U \right)^T \\ &= \mathbf{A}_{l_q} e^{jk_f \nu_{l_q}^D (T_s l + (i-1) T_{symb})}, \end{aligned} \quad (15)$$

where $\mathbf{A}_{l_q} = \rho_{l_q} \left(\mathbf{a}_{l_q}^B \right)^T \mathbf{a}_{l_q}^U$, $\mathbf{a}_{l_q}^B = \mathbf{b}_{l_q}^B \left(\mathbf{U}^B \right)^T$, and $\mathbf{a}_{l_q}^U = \mathbf{b}_{l_q}^U \left(\mathbf{U}^U \right)^T$, $\mathbf{a}_{l_q}^B$ and $\mathbf{a}_{l_q}^U$ represent the mapping of steering vectors to virtual directions. Consequently, CIRs between all antenna pairs in (1) can be stacked into an angle-domain CIR matrix as

$$\tilde{H}_{q,i}(l, \tau) = \sum_{l_q=0}^{L_q} \mathbf{A}_{l_q} e^{jk_f \left(\nu_{l_q}^D (T_s l + (i-1) T_{symb}) \right)} \delta(\tau - \tau_{l_q} T_s). \quad (16)$$

B. Nonuniform Delay-Domain Off-Grid Basis

The angle-domain channel representation with an off-grid basis can flexibly deal with actual continuous angles in a discretized way [18]. However, for better exploiting the full-dimensional insights in 3D space, UPAs are deployed on both sides, so that the beam domain to be sampled contains two dimensions of azimuth and elevation. Therefore, the number of

grid points must increase quadratically to improve the beam-domain resolution, which results in an overwhelming amount of computation.

For transmission systems at high frequencies, the small sampling period, which is inversely proportional to the large bandwidth, offers an opportunity for high multipath resolution in the delay domain. However, the high resolution with a large number of sampling points still results in a high computational burden. It can often be observed from delay PSDs that the power of multipath in the delay domain mainly concentrates near the LOS component and several strong non-line-of-sight (NLOS) components, and the PSD generally decreases as the delay increases, since those NLOS paths with large delays travel longer distances and experience more fading, thus arriving with lower power [32]. With the average delay $\mu_{\tau,q}$ and root mean square delay spread $\sigma_{\tau,q}$, the range where multipath power mainly concentrates in the delay domain can be known, and thus a nonuniform grid is proposed with scales set as

$$\bar{\tau}_k = \frac{(\mu_{\tau,q} + \sigma_{\tau,q})}{T_s \lceil \eta_\tau P_\tau \rceil} k, \quad (17)$$

when $k=0, \dots, \lceil \eta_\tau P_\tau \rceil - 1$, and when $k=\lceil \eta_\tau P_\tau \rceil, \dots, P_\tau - 1$,

$$\bar{\tau}_k = \frac{\left(\mu_{\tau,q} + \sigma_{\tau,q} + \frac{(T_{\text{GI}} - (\mu_{\tau,q} + \sigma_{\tau,q})(k - \lceil \eta_\tau P_\tau \rceil))}{P_\tau - \lceil \eta_\tau P_\tau \rceil} \right)}{T_s}.$$

We sprinkle more points in $[0, \mu_{\tau,q} + \sigma_{\tau,q}]$ with an adjustable proportion η_τ , and compress the number of points covering $[0, T_{\text{GI}}]$ to $P_\tau \ll P_{\text{GI}}$. This allows for a high resolution with fewer grid points and thus greatly reduces the computational size. The grid spacing can then be calculated as

$$\Delta \bar{\tau}_k = [\Delta \bar{\tau}]_k = \begin{cases} \bar{\tau}_{k+1} - \bar{\tau}_k, & k = 0, \dots, P_\tau - 2 \\ P_{\text{GI}} - \bar{\tau}_{P_\tau - 1}, & k = P_\tau - 1. \end{cases} \quad (18)$$

If $\bar{\tau}_k$ is the point closest to τ_{l_q} , i.e., $k_{l_q} = \arg \min_k |\bar{\tau}_k - \tau_{l_q}|$, we can then rewrite τ_{l_q} as

$$\tau_{l_q} = \bar{\tau}_{k_{l_q}} + \Delta \tau_{q,k_{l_q}}, \quad (19)$$

where $\Delta \tau_{q,k}$ denotes the off-grid gap and is set as

$$\Delta \tau_{q,k} = \begin{cases} \tau_{l_q} - \bar{\tau}_{k_{l_q}}, & k = k_{l_q}, \forall l_q = 0, \dots, L_q \\ 0, & \text{otherwise.} \end{cases} \quad (20)$$

Therefore, the CG matrix in (16) can be rewritten as

$$\tilde{\mathbf{H}}_{q,i}^D(l, k) = \bar{\mathbf{A}}_{q,k} e^{j k_f (\bar{\nu}_{q,k}^D (T_s l + (i-1) T_{\text{symb}}))}, \quad (21)$$

with $\bar{\mathbf{A}}_{q,k}$ and $\bar{\nu}_{q,k}^D$ set to

$$\bar{\mathbf{A}}_{q,k}, \bar{\nu}_{q,k}^D = \begin{cases} \mathbf{A}_{l_q}, \nu_{l_q}^D, & k = k_{l_q}, \forall l_q = 0, \dots, L_q \\ \mathbf{0}_{M \times N}, 0, & \text{otherwise.} \end{cases} \quad (22)$$

The nonuniform off-grid channel representation separates the CG and DFS of multipath, which enables the following compressed sensing solutions for the joint estimation problem.

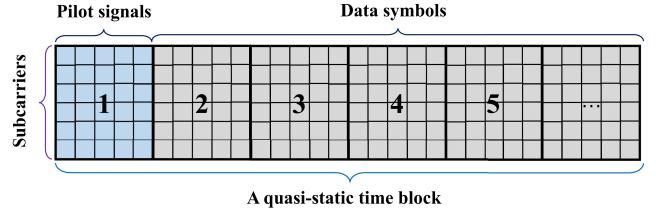


Fig. 2. Illustration of the pilot pattern.

C. Transmission Model

Accurate CFO estimation is necessary for subsequent model-based channel prediction and can be realized jointly with complex CG estimation through pilot training. The pilot signals are inserted between data symbols at some intervals within each time block, and are placed across all subcarriers, as shown in Fig. 2. Normally, in a time block, it is sufficient to insert pilot signals at the beginning symbol, and channels of the subsequent symbols can be obtained by the model-based channel prediction based on the estimated multipath parameters. Let \mathcal{I}_P denote the symbol index set of pilots, $I_P = |\mathcal{I}_P|$. At the i -th symbol, the pilot vector at the p -th sample is denoted as $\mathbf{s}_{q,i,p} \in \mathbb{C}^N$. The i -th received symbol at the BS can be written with the channel transfer function (CTF) as

$$\mathbf{y}_{q,i} = \mathbf{H}_{q,i} \mathbf{s}_{q,i} + \mathbf{n}_{q,i}, \quad (23)$$

$$\mathbf{H}_{q,i} = \begin{bmatrix} \mathbf{H}_{q,i,1,1} & \cdots & \mathbf{H}_{q,i,1,P} \\ \vdots & \ddots & \vdots \\ \mathbf{H}_{q,i,P,1} & \cdots & \mathbf{H}_{q,i,P,P} \end{bmatrix}, \quad (24)$$

$$\mathbf{s}_{q,i} = \left[(\mathbf{s}_{q,i,1})^T, \dots, (\mathbf{s}_{q,i,P})^T \right]^T, \quad (25)$$

where $\mathbf{n}_{q,i} \sim \mathcal{CN}(\mathbf{0}, \sigma_q^2 \mathbf{I}_{MP})$ denotes the additive white Gaussian noise (AWGN) on each receive antenna and subcarrier, with σ_q^2 indicating the noise power, \mathbf{H}_{q,i,p_1,p_2} represents the CTF of the OFDM MIMO channel between subcarriers and is given in (26), as shown at the bottom of the next page. Besides, by denoting the local oscillator frequency mismatch as $\bar{\nu}_{q,k}^O$, the total CFO can be written as $\bar{\nu}_{q,k} = \bar{\nu}_{q,k}^D + \bar{\nu}_{q,k}^O$, and the equivalent time-varying channel can then be denoted as

$$\tilde{\mathbf{H}}_{q,i}(l, k) = \tilde{\mathbf{H}}_{q,i}^D(l, k) e^{j k_f (\bar{\nu}_{q,k}^O (T_s l + (i-1) T_{\text{symb}}))} = \bar{\mathbf{A}}_{q,k} e^{j k_f (\bar{\nu}_{q,k}^O (T_s l + (i-1) T_{\text{symb}}))}. \quad (27)$$

With $\mathbf{U}_{NP}^U = \mathbf{I}_P \otimes (\mathbf{U}^U)^*$ and $\mathbf{U}_{NP}^B = \mathbf{I}_P \otimes (\mathbf{U}^B)^H$, (24) can be rewritten as

$$\mathbf{H}_{q,i} = \mathbf{U}_{NP}^B \underbrace{\begin{bmatrix} \tilde{\mathbf{H}}_{q,i,1,1} & \cdots & \tilde{\mathbf{H}}_{q,i,1,P} \\ \vdots & \ddots & \vdots \\ \tilde{\mathbf{H}}_{q,i,P,1} & \cdots & \tilde{\mathbf{H}}_{q,i,P,P} \end{bmatrix}}_{\tilde{\mathbf{H}}_{q,i}^{\text{ICI}}} \mathbf{U}_{NP}^U. \quad (28)$$

Thus, the received frequency-domain samples are rewritten as

$$\mathbf{y}_{q,i} = \mathbf{U}_{NP}^B \tilde{\mathbf{H}}_{q,i}^{\text{ICI}} \bar{\mathbf{s}}_{q,i} + \mathbf{n}_{q,i}, \quad (29)$$

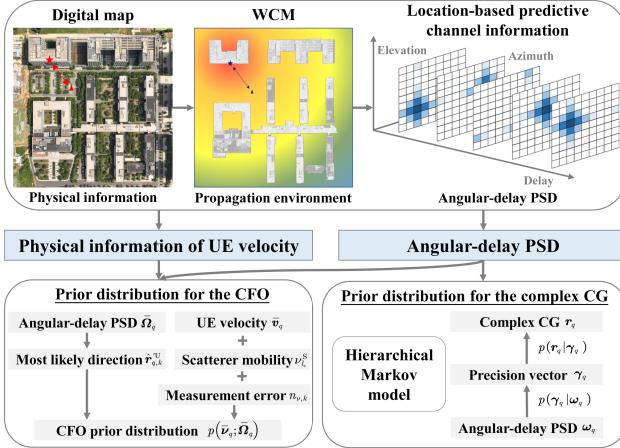


Fig. 3. Illustration of the WCM and its application in prior distribution construction for critical channel parameters.

where $\bar{s}_{q,i} = \mathbf{U}_{NP}^H \mathbf{s}_{q,i}$. Left-multiply both sides of (29) by (\mathbf{U}_{NP}^B) , we can have

$$\tilde{\mathbf{y}}_{q,i} = (\mathbf{U}_{NP}^B)^H \mathbf{y}_{q,i} = \tilde{\mathbf{H}}_{q,i}^{\text{ICI}} \bar{s}_{q,i} + \tilde{\mathbf{n}}_{q,i}, \quad (30)$$

where $\tilde{\mathbf{n}}_{q,i} = (\mathbf{U}_{NP}^B)^H \mathbf{n}_{q,i}$ still follows $\mathcal{CN}(\mathbf{0}, \sigma_q^2 \mathbf{I}_{MP})$, since \mathbf{U}_{NP}^B is a unitary matrix.

III. WCM-BASED PRIOR DISTRIBUTION CONSTRUCTION

A. WCM Precision

In order to quantitatively evaluate the accuracy of the pre-constructed channel information from the WCM, an accuracy key performance indicator (KPI) called cosine similarity is defined as

$$r_a = \frac{\mathbf{a}^H \hat{\mathbf{a}}}{\|\mathbf{a}\| \|\hat{\mathbf{a}}\|}, \quad (31)$$

where $\hat{\mathbf{a}}$ is the WCM-provided large-scale parameter and \mathbf{a} is the actual one during transmission. Though the dynamic WCM cannot perfectly reproduce the actual propagation environment digitally, it can still realize an ergodic result by accumulating channel information within a certain period of time. As shown in Fig. 3, some critical parameters, such as channel power distributions in the angle or delay domains, can then be known approximately.

B. Prior Distribution Construction for Complex CGs

The SBL method adjusts the sparsity of estimates by regulating parameterized prior distributions and is able to incorporate prior knowledge through the hierarchical Markov model [19]. The WCM is capable of providing loads of helpful prior knowledge for efficient channel estimation, e.g.,

PSDs indicating power distributions and sparsity in different domains. Though PSDs offered by the WCM are only approximations, they can still help improve the efficiency and accuracy of Bayesian learning, which will be shown.

The WCM-provided angular-delay PSD that is location-based and time-irrelevant is denoted as $\bar{\Omega}_q(k) = \mathbb{E}\{\bar{\mathbf{A}}_{q,k} \odot (\bar{\mathbf{A}}_{q,k})^*\}$. The multipath components can be regarded as independent of each other since they are randomly generated in a complicated propagation environment. Therefore, \mathbf{r}_q can be considered as a zero-mean complex Gaussian random vector with

$$p(\mathbf{r}_q | \gamma_q) = \mathcal{CN}(\mathbf{r}_q | \mathbf{0}, \text{diag}(\gamma_q)^{-1}), \quad (32)$$

where $\gamma_q = [\gamma_{q,s}]_{s=1,\dots,MNP_\tau}$ is the precision vector and serves as a penalty term to control the sparsity of \mathbf{r}_q . As elements in γ_q are inverses of element variances in \mathbf{r}_q , if $\gamma_{q,s}$ takes an infinite value, $[\mathbf{r}_q]_s$ will approach 0. The conditional prior for γ_q is written as

$$p(\gamma_q | \omega_q) = \prod_{s=1}^{MNP_\tau} \Gamma(\gamma_{q,s}; 1 + c_{1,q,s}^r, c_{2,q,s}^r), \quad (33)$$

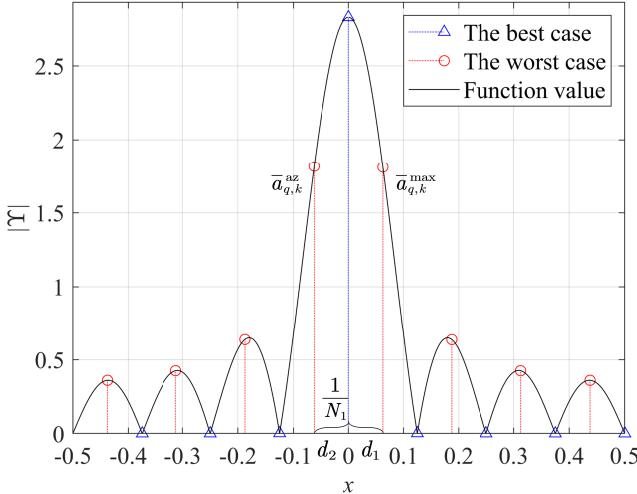
where $\omega_q = \text{vec}([\bar{\Omega}_q(k)]_{k=0,\dots,P_\tau-1})$, and $\Gamma(\gamma_{q,s}; 1 + c_{1,q,s}^r, c_{2,q,s}^r)$ is a Gamma hyperprior, $c_{1,q,s}^r$ and $c_{2,q,s}^r$ correspond to the shape and rate parameters, respectively. For all $s \in \{1, 2, \dots, MNP_\tau\}$, $c_{1,q,s}^r > -1$ and $c_{2,q,s}^r > 0$. A three-layer Markov model is employed to incorporate prior information from the WCM into the parameterized prior distribution of \mathbf{r}_q , as can be seen in Fig. 3, the relationship between $c_{1,q,s}^r$ and $c_{2,q,s}^r$ is specifically regulated by ω_q as $\frac{c_{1,q,s}^r + 1}{c_{2,q,s}^r} = [\omega_q]_s^{-1}$, normally we set $c_{1,q,s}^r \rightarrow 0$, $c_{2,q,s}^r \rightarrow 0$ to have a broader hyperprior.

C. Prior Distribution Construction for CFOs

According to (4), the DFS of each path consists of two components: the spherical unit vector representing the path direction and the UE velocity. Estimating azimuth and elevation angles of multipath directly will lead to an excessive computational burden. The multipath angles can be extracted from angle-domain CGs, as the ratio of channel projections of a path onto virtual directions by the DFT matrix is closely related to the path angle. The virtual direction with the maximum projection value approaches the actual path direction the most. At the fixed b -th azimuth and a -th elevation direction, $\mathbf{a}_{l_q}^U$ in (15) can be decomposed as

$$\begin{aligned} \mathbf{a}_{l_q}^U((a-1)N_1 + b) \\ = \left(\left[\mathbf{U}_{\text{el}}^U \right]_{a,:} \mathbf{u}(\varphi_{l_q}^U, N_2) \right) \left(\left[\mathbf{U}_{\text{az}}^U \right]_{b,:} \mathbf{u}(\vartheta_{l_q}^U, N_1) \right) \end{aligned}$$

$$\mathbf{H}_{q,i,p_1,p_2} = (\mathbf{U}^B)^H \underbrace{\left(\frac{1}{P} \sum_{k=0}^{P_\tau-1} \sum_{l=0}^{P-1} \tilde{\mathbf{H}}_{q,i}(l,k) e^{-j \frac{2\pi}{P} (\bar{\tau}_k + \Delta\tau_{q,k})(p_2-1)} e^{j \frac{2\pi}{P} l(p_2-p_1)} \right)}_{\tilde{\mathbf{H}}_{q,i,p_1,p_2}} (\mathbf{U}^U)^* \quad (26)$$

Fig. 4. Amplitude of function Υ versus x ($C = 8$).

$$= \Upsilon \left(N_2, \frac{2a-1}{2N_2} + \varphi_{l_q}^U - 0.5 \right) \Upsilon \left(N_1, \frac{2b-1}{2N_1} + \vartheta_{l_q}^U - 0.5 \right), \quad (34)$$

where $\Upsilon(C, x) = \frac{\sin(C\pi x)}{\sqrt{C} \sin(\pi x)} e^{(jx(C-1))}$, and $|\Upsilon(C, x)|$ is a periodic function with period 1 about x . The amplitude of the function $\Upsilon(C, x)$ within a single period is shown in Fig. 4. Take the equivalent elevation angle as an example. If $\varphi_{l_q}^U$ falls exactly on one of the fixed virtual directions, the amplitude of the function Υ can reach the maximum, as shown in the best case in Fig. 4. In fact, the ratio of the first two greatest angle-domain CG amplitudes on path l_q is determined by $\varphi_{l_q}^U$. Though the exact CG matrix $\bar{\mathbf{A}}_{q,k}$ can hardly be acquired before estimation, we can still obtain the most probable path directions in a statistical sense according to the angular-delay PSD offered by the WCM, and this part of the calculation can be delegated to the WCM without incurring an extra computational burden to online estimation.

We define $[\bar{\mathbf{A}}'_{q,k}]_{m,n} = [\bar{\boldsymbol{\Omega}}_q(k)]_{m,n}^{1/2}$ to represent the ergodic angular-delay spectrum. Assume that the multi-path power concentrates on L'_q delay-domain grid points with indices of $\{k_{l_q} | l_q = 0, \dots, L'_q\}$, and k_{l_q} belongs to $\text{supp}(\{\bar{\boldsymbol{\Omega}}_q(k)\}_{k=1,..,P_\tau})$. Let $\bar{a}_{q,k}^{\max}$ denote the maximum element amplitude in $\bar{\mathbf{A}}'_{q,k}$, and it is in the c_{\max} -th column of $\bar{\mathbf{A}}'_{q,k}$. We define that $b_1 = \text{mod}(c_{\max} - 1, N_1) + 1$ and $a_1 = \lfloor \frac{c_{\max}}{N_1} \rfloor + 1$. Among all elements that are in the same role as the maximum amplitude element and whose column number belongs to $\{N_1(a_1 - 1) + 1, N_1(a_1 - 1) + 2, \dots, N_1a_1\}$, the second largest element amplitude is denoted as $\bar{a}_{q,k}^{az}$, and its column number is c_{submax}^{az} . We further define that $b_2 = \text{mod}(c_{\text{submax}}^{az} - 1, N_1) + 1$. If the first two largest amplitudes are equal, b_1 takes the smaller number, and b_2 takes the larger one. Among all elements that are in the same role as the maximum amplitude element and whose column number belongs to $\{b_1, b_1 + N_1, \dots, b_1 + (N_2 - 1)N_1\}$, the second largest element amplitude is denoted as $\bar{a}_{q,k}^{\text{el}}$, and its column number is $c_{\text{submax}}^{\text{el}}$. We further define that $a_2 = \lfloor \frac{c_{\text{submax}}^{\text{el}}}{N_1} \rfloor + 1$. If the first two largest amplitudes are equal, a_1 takes the smaller number, and a_2 takes the larger one. If the actual

azimuth/elevation direction of a path is consistent with one of the fixed azimuth/elevation directions, $b_1 = b_2/a_1 = a_2$. Based on the periodicity of the function $|\Upsilon(C, x)|$, we have the following theorem.

Theorem 1: With the WCM-provided angular-delay PSD and critical element indexes introduced above, $\varphi_{q,k}^U$ and $\vartheta_{q,k}^U$ on point $k = k_{l_q}$ can be written as

$$\varphi_{q,k}^U = \begin{cases} \frac{N_2 - 2a_1 + 1}{2N_2} - \frac{\eta_{\text{el}}}{N_2(\eta_{\text{el}} + 1)}, \\ a_1 - a_2 \in (-N_2 + 1, 0], a_1 - a_2 = N_2 - 1 \\ \frac{N_2 - 2a_1 + 1}{2N_2} + \frac{\eta_{\text{el}}}{N_2(\eta_{\text{el}} + 1)}, \\ a_1 - a_2 \in (0, N_2 - 1), a_1 - a_2 = -N_2 + 1 \end{cases}, \quad (35)$$

$$\vartheta_{q,k}^U = \begin{cases} \frac{N_1 - 2b_1 + 1}{2N_1} - \frac{\eta_{az}}{N_1(\eta_{az} + 1)}, \\ b_1 - b_2 \in (-N_1 + 1, 0], b_1 - b_2 = N_1 - 1 \\ \frac{N_1 - 2b_1 + 1}{2N_1} + \frac{\eta_{az}}{N_1(\eta_{az} + 1)}, \\ b_1 - b_2 \in (0, N_1 - 1), b_1 - b_2 = -N_1 + 1 \end{cases}, \quad (36)$$

when N_1 and N_2 are very large. Here, $\eta_{\text{el}} = \frac{\bar{a}_{q,k}^{\text{el}}}{\bar{a}_{q,k}^{\max}}$ and $\eta_{az} = \frac{\bar{a}_{q,k}^{az}}{\bar{a}_{q,k}^{\max}}$ denote ratios between the first two maximum amplitudes in elevation and azimuth directions, respectively.

Proof: Detailed derivations can be found in Appendix A. ■

With the inferred $\varphi_{q,k}^U$ and $\vartheta_{q,k}^U$, we have

$$\hat{\mathbf{r}}_{q,k}^{\text{U}} = \left[\sqrt{1 - \left(\vartheta_{q,k}^U \frac{\lambda_c}{d_c^{\text{U}}} \right)^2} - \left(\vartheta_{q,k}^U \frac{\lambda_c}{d_c^{\text{U}}} \right)^2, \vartheta_{q,k}^U \frac{\lambda_c}{d_c^{\text{U}}}, \vartheta_{q,k}^U \frac{\lambda_c}{d_r^{\text{U}}} \right]^T \quad (37)$$

to obtain $\hat{\mathbf{r}}_{q,k}^{\text{U,GCS}}$ on point $k = k_{l_q}$ by LCS-GCS conversion. Such an approximate linear solution can greatly alleviate the computational burden. Taking into account the oscillator frequency mismatch as well as localization and calculation errors of the aforementioned parameters, the composition of the CFO on point $k = k_{l_q}$ can be expressed as

$$\bar{\nu}_{q,k} = \begin{cases} \left(\hat{\mathbf{r}}_{q,k}^{\text{U,GCS}} \right)^T \bar{\mathbf{v}}_q + n_{\nu,k}, k = k_0 \\ \left(\hat{\mathbf{r}}_{q,k}^{\text{U,GCS}} \right)^T \bar{\mathbf{v}}_q + \nu_{q,k}^S + n_{\nu,k}, k = k_1, \dots, k_{L'_q} \end{cases}, \quad (38)$$

where $n_{\nu,k} \sim \mathcal{N}(0, \sigma_{\nu,k}^2)$ covers all mismatches and errors. With such a composition, an accurate prior distribution for the CFO can be obtained, as shown in the following theorem.

Theorem 2: With the knowledge of the most probable path directions and the UE velocity provided by the WCM, the prior distribution of CFO $\bar{\nu}_q$ can be written as

$$p(\bar{\nu}_q; \bar{\boldsymbol{\Omega}}_q) = \prod_{k=0}^{P_{\text{GI}}-1} f_q(\bar{\nu}_{q,k}; \bar{\boldsymbol{\Omega}}_q), \quad (39)$$

where function f_q is defined as (40), as shown at the bottom of the next page, with $C_{0,q}(k) = \left(\hat{\mathbf{r}}_{q,k}^{\text{U,GCS}} \right)^T \bar{\mathbf{v}}_q$ and $Q(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt$.

Proof: Detailed derivations can be found in Appendix B. ■

IV. JOINT COMPLEX CG AND CFO ESTIMATION

A. Model Reformulation and Problem Generation

From the composition of $\tilde{\mathbf{H}}_{q,i}^{\text{ICI}}$ in (28), $\tilde{\mathbf{H}}_{q,i}^{\text{ICI}}$ is determined by the multipath delay $\Delta\tau_q = [\Delta\tau_{q,0}, \dots, \Delta\tau_{q,P_\tau-1}]^\top$, CFO $\bar{\nu}_q = [\bar{\nu}_{q,0}, \dots, \bar{\nu}_{q,P_\tau-1}]^\top$, and angle-delay domain CG $\bar{\mathbf{A}}_q = [(\bar{\mathbf{A}}_{q,0})^\top, \dots, (\bar{\mathbf{A}}_{q,P_\tau-1})^\top]^\top$. The CTF $\tilde{\mathbf{H}}_{q,i,p_1,p_2}$ can be reformulated as

$$\tilde{\mathbf{H}}_{q,i,p_1,p_2} = \frac{1}{P} \mathbf{Z}_{q,i,p_1,p_2} \bar{\mathbf{A}}_q, \quad (41)$$

$$\begin{aligned} \mathbf{Z}_{q,i,p_1,p_2} &= \left[e^{-j\frac{2\pi}{P}(\bar{\tau}_0 + \Delta\tau_{q,0})(p_2-1)} \xi_{q,i,p_1,p_2}^0 (\bar{\nu}_{q,0}) \mathbf{I}_M, \dots, \right. \\ &\quad \left. e^{-j\frac{2\pi}{P}(\bar{\tau}_{P_\tau-1} + \Delta\tau_{q,P_\tau-1})(p_2-1)} \xi_{q,i,p_1,p_2}^{P_\tau-1} (\bar{\nu}_{q,P_\tau-1}) \mathbf{I}_M \right], \end{aligned} \quad (42)$$

$$\xi_{q,i,p_1,p_2}^k (\bar{\nu}_{q,k}) = \sum_{l=0}^{P-1} e^{jk_f(\bar{\nu}_{q,k}(T_s l + (i-1)T_{\text{symb}}))} e^{j\frac{2\pi}{P}l(p_2-p_1)}. \quad (43)$$

Using (28) and (41), (30) can be rewritten as

$$\begin{aligned} \tilde{\mathbf{y}}_{q,i} &= \tilde{\mathbf{H}}_{q,i}^{\text{ICI}} \bar{\mathbf{s}}_{q,i} + \tilde{\mathbf{n}}_{q,i} \\ &= \frac{1}{P} \mathcal{K}_{q,i} (\Delta\tau_q, \bar{\nu}_q) \mathbf{r}_q + \tilde{\mathbf{n}}_{q,i}, \end{aligned} \quad (44)$$

where $\mathbf{r}_q = \text{vec}(\bar{\mathbf{A}}_q)$,

$$\mathcal{K}_{q,i} (\Delta\tau_q, \bar{\nu}_q) = \left[(\mathcal{K}_{q,i,1} (\Delta\tau_q, \bar{\nu}_q))^\top, \dots, (\mathcal{K}_{q,i,P} (\Delta\tau_q, \bar{\nu}_q))^\top \right]^\top, \quad (45)$$

$$\mathcal{K}_{q,i,p_1} (\Delta\tau_q, \bar{\nu}_q) = \sum_{p_2=1}^P (\bar{\mathbf{s}}_{q,i,p_2})^\top \otimes \mathbf{Z}_{q,i,p_1,p_2}, \quad (46)$$

and $\bar{\mathbf{s}}_{q,i} = \left[(\bar{\mathbf{s}}_{q,i,1})^\top, \dots, (\bar{\mathbf{s}}_{q,i,P})^\top \right]^\top$. By concatenating pilot symbols in the same time block, we can have

$$\tilde{\mathbf{y}}_q = \frac{1}{P} \mathcal{K}_q (\Delta\tau_q, \bar{\nu}_q) \mathbf{r}_q + \tilde{\mathbf{n}}_q, \quad (47)$$

where $\tilde{\mathbf{y}}_q = [\tilde{\mathbf{y}}_{q,i}]_{i \in \mathcal{I}_P}$, $\mathcal{K}_q (\Delta\tau_q, \bar{\nu}_q) = [\mathcal{K}_{q,i} (\Delta\tau_q, \bar{\nu}_q)]_{i \in \mathcal{I}_P}$, $\tilde{\mathbf{n}}_q = [\tilde{\mathbf{n}}_{q,i}]_{i \in \mathcal{I}_P}$, and $\tilde{\mathbf{n}}_q$ is still a zero mean complex Gaussian random vector with a covariance matrix of $\sigma_q^2 \mathbf{I}_{MPI_P}$. A Gamma hyperprior of $\alpha_q \sim \Gamma(1 + c_1^\alpha, c_2^\alpha)$ is formulated for the unknown noise power, with $c_1^\alpha, c_2^\alpha \rightarrow 0$, $c_1^\alpha > -1$ and $c_2^\alpha > 0$, $\alpha_q = \sigma_q^{-2}$ represents the noise precision.

We can see that the original pilot transmission model has been transformed into a measurement model for \mathbf{r}_q , and \mathbf{r}_q is a

sparse MNP_τ -dimensional random vector representing the CG in the angle-delay domain. As there are still unknown random variables—the delay offset $\Delta\tau_q$ and CFO $\bar{\nu}_q$ contained in the measurement matrix \mathcal{K}_q , (47) actually leads to a parameterized linear estimation problem that cannot be solved by normal compressed sensing algorithms like matching pursuit directly. However, with the aid of the WCM, we can still recover the concerned CGs and CFOs in the Bayesian framework. Our goal is to determine the parameter set (PS) $\Xi_q = \{\alpha_q, \Delta\tau_q, \bar{\nu}_q, \gamma_q\}$, and recover \mathbf{r}_q using $\tilde{\mathbf{y}}_q$ and \mathcal{K}_q . The maximum a prior (MAP) estimator for Ξ_q is formulated as

$$\hat{\Xi}_q = \arg \max_{\Xi_q} \ln p(\Xi_q, \tilde{\mathbf{y}}_q). \quad (48)$$

As a commonly used method for solving parameterized linear estimation problems, the minorization-maximization (MM) algorithm is employed in this paper to obtain the optimal PS Ξ_q iteratively with proper surrogate function settings.

B. Sparse CG Recovery

When the PS is fixed, from the precision vector γ_q , the support set for \mathbf{r}_q , which is denoted by $\mathcal{S}_q^\gamma = \text{supp}(\{\gamma_{q,s}^{-1}\}_{s=1, \dots, MNP_\tau})$, can be obtained. Considering the better performance under low SNRs, the minimum mean square error (MMSE) method is adopted, and the CG estimate can be written as

$$\begin{aligned} \left(\hat{\mathbf{r}}_q^{\text{MMSE}} \right)_{\mathcal{S}_q^\gamma} &= P \mathbf{R}_{\text{rr},q} (\mathcal{K}_q^{\text{spar}})^H \left(\mathcal{K}_q^{\text{spar}} \mathbf{R}_{\text{rr},q} (\mathcal{K}_q^{\text{spar}})^H \right. \\ &\quad \left. + P^2 \sigma_q^2 \mathbf{I}_{MPI_P} \right)^{-1} \tilde{\mathbf{y}}_q, \end{aligned} \quad (49)$$

where $\mathbf{R}_{\text{rr},q} = \text{diag}((\gamma_q)_{\mathcal{S}_q^\gamma})^{-1}$, $\mathcal{K}_q^{\text{spar}} = (\mathcal{K}_q (\Delta\hat{\tau}_q, \hat{\bar{\nu}}_q))_{\mathcal{S}_q^\gamma}$, $\Delta\hat{\tau}_q$ and $\hat{\bar{\nu}}_q$ are in the fixed PS. As the channel multipath is often sparsely distributed in the delay domain, the dimensions of the matrices involved in the calculation are greatly reduced.

C. WCM-Based SBL With MM Algorithms

From (47), $\tilde{\mathbf{y}}_q$ is determined by the channel and AWGN and can thus be regarded as a complex Gaussian random vector with a distribution of

$$p(\tilde{\mathbf{y}}_q | \mathbf{r}_q, \Xi_q) = \frac{\exp\left(-\alpha_q \|\tilde{\mathbf{y}}_q - \frac{1}{P} \mathcal{K}_q (\Delta\tau_q, \bar{\nu}_q) \mathbf{r}_q\|^2\right)}{(\pi \alpha_q^{-1})^{MPI_P}}. \quad (50)$$

A more tractable lower bound is employed here in the MM algorithm to replace the E-step of the original EM method, and

$$f_q(\bar{\nu}_{q,k}; \bar{\Omega}_q) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma_{\nu,k}^2} \exp\left(-\frac{(\bar{\nu}_{q,k} - C_{0,q}(k))^2}{2\sigma_{\nu,k}^2}\right), & k = k_0 \\ \frac{\zeta}{4v_{\text{scatt}}} \left(Q\left(\frac{2v_{\text{scatt}} - \bar{\nu}_{q,k} + C_{0,q}(k)}{\sigma_{\nu,k}}\right) - Q\left(\frac{C_{0,q}(k) - 2v_{\text{scatt}} - \bar{\nu}_{q,k}}{\sigma_{\nu,k}}\right) \right) + \frac{(1-\zeta)}{2\sqrt{2\pi}\sigma_{\nu,k}^2} \exp\left(-\frac{(\bar{\nu}_{q,k} - C_{0,q}(k))^2}{2\sigma_{\nu,k}^2}\right), & k = k_1, \dots, k_{L'_q} \\ \frac{1}{\sqrt{2\pi}\sigma_{\nu,k}^2} \exp\left(-\frac{\bar{\nu}_{q,k}^2}{2\sigma_{\nu,k}^2}\right), & \text{otherwise} \end{cases}, \quad (40)$$

in the M-step, the surrogate function is maximized. According to the Jensen's inequality, for any given $\dot{\Xi}_q$, the surrogate function of

$$\mathcal{U}(\Xi_q | \dot{\Xi}_q) = \int p(r_q | \tilde{y}_q, \dot{\Xi}_q) \ln \frac{p(r_q, \tilde{y}_q, \Xi_q)}{p(r_q | \tilde{y}_q, \dot{\Xi}_q)} d_{r_q} \quad (51)$$

is always a lower bound to the real value function $\ln p(\Xi_q, \tilde{y}_q)$ at $\dot{\Xi}_q$ [18]. Therefore, the original problem of (48) can be tackled by solving

$$\Xi_q^{(t)} = \arg \max_{\Xi_q} \mathcal{U}(\Xi_q | \Xi_q^{(t-1)}). \quad (52)$$

From the Bayes' theorem, we can have [17]

$$p(r_q | \tilde{y}_q, \Xi_q) = \frac{\exp \left(- (r_q - \mu_q)^H \Sigma_q^{-1} (r_q - \mu_q) \right)}{|\pi \Sigma_q|}, \quad (53)$$

where the mean and covariance matrix are written as

$$\mu_q(\Xi_q) = \frac{\alpha_q}{P} \Sigma_q(\Xi_q) (\mathcal{K}_q(\Delta\tau_q, \bar{\nu}_q))^H \tilde{y}_q, \quad (54)$$

$$\Sigma_q(\Xi_q) = \left(\frac{\alpha_q}{P^2} (\mathcal{K}_q(\Delta\tau_q, \bar{\nu}_q))^H (\mathcal{K}_q(\Delta\tau_q, \bar{\nu}_q)) + \text{diag}(\gamma_q) \right)^{-1}. \quad (55)$$

Though variables in problem (52) are coupled together, we can still search for the optimal solution by the block coordinate descent method.

1) *Update for $\bar{\nu}_q$ at the t-*Th* Iteration:* With $\{\alpha_q, \Delta\tau_q, \gamma_q\} = \{\alpha_q^{(t-1)}, \Delta\tau_q^{(t-1)}, \gamma_q^{(t-1)}\}$ and $\dot{\Xi}_q^{(t)} = \{\alpha_q^{(t-1)}, \Delta\tau_q^{(t-1)}, \bar{\nu}_q^{(t-1)}, \gamma_q^{(t-1)}\}$, (51) can be rewritten as

$$\begin{aligned} & \mathcal{U}(\alpha_q^{(t-1)}, \Delta\tau_q^{(t-1)}, \bar{\nu}_q, \gamma_q^{(t-1)} | \dot{\Xi}_q^{(t)}) \\ &= \int p(r_q | \tilde{y}_q, \dot{\Xi}_q^{(t)}) \ln p(\tilde{y}_q | r_q, \alpha_q^{(t-1)}, \Delta\tau_q^{(t-1)}, \bar{\nu}_q) d_{r_q} \\ &+ \ln p(\bar{\nu}_q) \end{aligned} \quad (56)$$

$$\begin{aligned} &= -\alpha_q^{(t-1)} C_{1,q} \left(\Delta\tau_q^{(t-1)}, \bar{\nu}_q, \dot{\Xi}_q^{(t)} \right) + \ln p(\bar{\nu}_q), \\ &C_{1,q} \left(\Delta\tau_q, \bar{\nu}_q, \dot{\Xi}_q \right) \\ &= \frac{1}{P^2} \text{tr} \left(\mathcal{K}_q(\Delta\tau_q, \bar{\nu}_q) \Sigma_q(\dot{\Xi}_q) (\mathcal{K}_q(\Delta\tau_q, \bar{\nu}_q))^H \right) \\ &+ \left\| \frac{1}{P} \mathcal{K}_q(\Delta\tau_q, \bar{\nu}_q) \mu_q(\dot{\Xi}_q) - \tilde{y}_q \right\|^2, \end{aligned} \quad (57)$$

by ignoring terms irrelevant to $\bar{\nu}_q$. Although $\bar{\nu}_q$ is deeply coupled in \mathcal{K}_q , and it is hard to determine the convexity

or concavity of (56) with respect to $\bar{\nu}_q$, we can still find a suboptimal solution along the gradient denoted by

$$\mathbf{g}_{\nu,q}(\bar{\nu}_q) = [g_{\nu,q,0}(\bar{\nu}_q), \dots, g_{\nu,q,P_\tau-1}(\bar{\nu}_q)]^T, \quad (58)$$

where $g_{\nu,q,k}(\bar{\nu}_q)$ is derived as (59), as shown at the bottom of the page, with components defined as

$$\begin{aligned} \mathcal{C}_{q,i,p_1}^{\bar{\nu}} &= \frac{1}{P} \sum_{n=1}^N \sum_{k=0}^{P_\tau-1} \sum_{p_2=1}^P [\bar{s}_{q,i,p_2}]_n \xi_{q,i,p_1,p_2}^k(\bar{\nu}_q, k) \\ &\times e^{-j \frac{2\pi}{P} ((\bar{\tau}_k + \Delta\tau_{q,k}^{(t-1)}) (p_2-1) + l(p_1-p_2))} \mu_{q,k,n}(\dot{\Xi}_q^{(t)}) - \tilde{y}_{q,i,p_1}, \end{aligned} \quad (60)$$

$$\begin{aligned} \mathcal{C}_{q,i,p_1,k}^{\bar{\nu}} &= \frac{j k_f}{P} \sum_{n=1}^N \sum_{p_2=1}^P \sum_{l=0}^{P-1} (T_s l + (i-1) T_{\text{symb}}) [\bar{s}_{q,i,p_2}]_n \\ &\times e^{-j \frac{2\pi}{P} ((\bar{\tau}_k + \Delta\tau_{q,k}^{(t-1)}) (p_2-1) + l(p_1-p_2))} \\ &\times e^{j k_f \bar{\nu}_{q,k} (T_s l + (i-1) T_{\text{symb}})} \mu_{q,k,n}(\dot{\Xi}_q^{(t)}), \end{aligned} \quad (61)$$

$$\begin{aligned} \mathcal{D}_{q,i,p_1,k_1,n_1}^{\bar{\nu}} &= \sum_{n_2=1}^N \sum_{k_2=1}^{P_\tau} \sum_{p_2=1}^P [\bar{s}_{q,i,p_2}]_{n_2}^* \left(\xi_{q,i,p_1,p_2}^{k_2-1}(\bar{\nu}_{q,k_2-1}) \right)^* \\ &\times e^{j \frac{2\pi}{P} ((\bar{\tau}_{k_2-1} + \Delta\tau_{q,k_2-1}^{(t-1)}) (p_2-1))} \Sigma_{q,k_1,n_1,k_2,n_2}^T(\dot{\Xi}_q^{(t)}), \end{aligned} \quad (62)$$

$$\begin{aligned} \mathcal{K}_{q,i,p_1,k_1,n_1}^{\bar{\nu}} &= \sum_{p_2=1}^P \sum_{l=0}^{P-1} j k_f (T_s l + (i-1) T_{\text{symb}}) [\bar{s}_{q,i,p_2}]_{n_1} \\ &\times e^{-j \frac{2\pi}{P} ((\bar{\tau}_{k_1-1} + \Delta\tau_{q,k_1-1}^{(t-1)}) (p_2-1) + l(p_1-p_2))} \\ &\times e^{j k_f \bar{\nu}_{q,k_1-1} (T_s l + (i-1) T_{\text{symb}})} \mathbf{I}_M, \end{aligned} \quad (63)$$

and (64), as shown at the bottom of the next page. It is defined that

$$\tilde{y}_{q,i,p_1} = [\tilde{y}_{q,i}]_{(p_1-1)M+1:p_1M}, \quad (65)$$

$$\mu_{q,k,n}(\dot{\Xi}_q) = [\mu_q(\dot{\Xi}_q)]_{((n-1)P_\tau+k)M+1:((n-1)P_\tau+k+1)M}, \quad (66)$$

$$\begin{aligned} \Sigma_{q,k_1,n_1,k_2,n_2}(\dot{\Xi}_q) &= [\Sigma_q(\dot{\Xi}_q)]_{((n_1-1)P_\tau+k_1-1)M+1:((n_1-1)P_\tau+k_1)M, \dots} \\ &\quad ((n_2-1)P_\tau+k_2-1)M+1:((n_2-1)P_\tau+k_2)M \end{aligned} \quad (67)$$

Referring to the two-point step size gradient method proposed by Barzilai and Borwein, we search for the optimal solution

$$\begin{aligned} g_{\nu,q,k}(\bar{\nu}_q) &= \mathcal{B}_{q,k}^{\bar{\nu}} - 2\alpha_q^{(t-1)} \sum_{i=1}^{I_p} \sum_{p_1=1}^P \text{Re} \left[(\mathcal{C}_{q,i,p_1}^{\bar{\nu}})^H \mathcal{C}_{q,i,p_1,k}^{\bar{\nu}} \right] \\ &- \frac{2\alpha_q^{(t-1)}}{P^2} \left(\sum_{i=1}^{I_p} \sum_{p_1=1}^P \sum_{n=1}^N \text{Re} \left[\text{vec}^T(\mathcal{D}_{q,i,p_1,k+1,n}^{\bar{\nu}}) \text{vec}(\mathcal{K}_{q,i,p_1,k+1,n}^{\bar{\nu}}) \right] \right) \end{aligned} \quad (59)$$

for $\bar{\nu}_{q,k}$ along the gradient ascent with a step size of

$$\varepsilon_{\nu}^{(t)} = \frac{\left(\bar{\nu}_q^{(t-1)} - \bar{\nu}_q^{(t-2)}\right)^T \left(\mathbf{g}_{\nu,q}\left(\bar{\nu}_q^{(t-1)}\right) - \mathbf{g}_{\nu,q}\left(\bar{\nu}_q^{(t-2)}\right)\right)}{\left\|\mathbf{g}_{\nu,q}\left(\bar{\nu}_q^{(t-1)}\right) - \mathbf{g}_{\nu,q}\left(\bar{\nu}_q^{(t-2)}\right)\right\|^2}, \quad (68)$$

which is determined by the combined information of the current t -th iteration and the last. Thus, the t -th iteration can be expressed as

$$\bar{\nu}_q^{(t)} = \bar{\nu}_q^{(t-1)} + \varepsilon_{\nu}^{(t)} \mathbf{g}_{\nu,q}\left(\bar{\nu}_q^{(t-1)}\right), \quad (69)$$

and $\bar{\nu}_{q,k}$ can only take values in $[-v_q - 2v_{\text{scatt}}, v_q + 2v_{\text{scatt}}]$.

2) *Update for α_q at the t -Th Iteration:* Similarly, with $\{\Delta\tau_q, \bar{\nu}_q, \gamma_q\} = \{\Delta\tau_q^{(t-1)}, \bar{\nu}_q^{(t)}, \gamma_q^{(t-1)}\}$ and $\dot{\Xi}_q^{\alpha} = \{\alpha_q^{(t-1)}, \Delta\tau_q^{(t-1)}, \bar{\nu}_q^{(t)}, \gamma_q^{(t-1)}\}$, the objective in (52) can be rewritten as

$$\begin{aligned} \mathcal{U}\left(\alpha_q, \Delta\tau_q^{(t-1)}, \bar{\nu}_q^{(t)}, \gamma_q^{(t-1)} \mid \dot{\Xi}_q^{\alpha}\right) \\ = \int p\left(\mathbf{r}_q \mid \tilde{\mathbf{y}}_q, \dot{\Xi}_q^{\alpha}\right) \ln p\left(\tilde{\mathbf{y}}_q \mid \mathbf{r}_q, \alpha_q, \Delta\tau_q^{(t-1)}, \bar{\nu}_q^{(t)}\right) d\mathbf{r}_q \\ + \ln p(\alpha_q) \\ = (c_1^{\alpha} + MPI_P) \ln \alpha_q \\ - \left(c_2^{\alpha} + C_{1,q}\left(\Delta\tau_q^{(t-1)}, \bar{\nu}_q^{(t)}, \dot{\Xi}_q^{\alpha}\right)\right) \alpha_q, \end{aligned} \quad (70)$$

after removing terms irrelevant with α_q . Since (70) is a concave function, with the first-order necessary condition for optimality, the optimal solution for α_q can be written as

$$\alpha_q^{(t)} = \frac{c_1^{\alpha} + MPI_P}{c_2^{\alpha} + C_{1,q}\left(\Delta\tau_q^{(t-1)}, \bar{\nu}_q^{(t)}, \dot{\Xi}_q^{\alpha}\right)}. \quad (71)$$

3) *Update for γ_q at the t -Th Iteration:* With $\{\alpha_q, \Delta\tau_q, \bar{\nu}_q, \gamma_q\} = \{\alpha_q^{(t)}, \Delta\tau_q^{(t-1)}, \bar{\nu}_q^{(t)}\}$ and $\dot{\Xi}_q^{\gamma} = \{\alpha_q^{(t)}, \Delta\tau_q^{(t-1)}, \bar{\nu}_q^{(t)}, \gamma_q^{(t-1)}\}$, (51) can be rewritten as

$$\begin{aligned} \mathcal{U}\left(\alpha_q^{(t)}, \Delta\tau_q^{(t-1)}, \bar{\nu}_q^{(t)}, \gamma_q \mid \dot{\Xi}_q^{\gamma}\right) \\ = \int p\left(\mathbf{r}_q \mid \tilde{\mathbf{y}}_q, \dot{\Xi}_q^{\gamma}\right) \ln p\left(\mathbf{r}_q \mid \gamma_q\right) d\mathbf{r}_q + \ln p(\gamma_q) \\ = \sum_{s=1}^{MNP\tau} (c_{1,s} + 1) \ln \gamma_{q,s} - \sum_{s=1}^{MNP\tau} c_{2,s} \gamma_{q,s} \\ - \text{tr}\left(\text{diag}(\gamma_q) \left(\boldsymbol{\Sigma}_q\left(\dot{\Xi}_q^{\gamma}\right) + \boldsymbol{\mu}_q\left(\dot{\Xi}_q^{\gamma}\right) \boldsymbol{\mu}_q\left(\dot{\Xi}_q^{\gamma}\right)^H\right)\right), \end{aligned} \quad (72)$$

$$\mathcal{B}'_{q,k}^{\bar{\nu}} = \begin{cases} \frac{C_{0,q}(k) - \bar{\nu}_{q,k}}{\sigma_{\nu}^2}, k = k_0 \\ \zeta \sigma_{\nu}^2 \left(e^{-\frac{(\bar{\nu}_{q,k} + 2v_{\text{scatt}} - C_{0,q}(k))^2}{2\sigma_{\nu}^2}} - e^{-\frac{(\bar{\nu}_{q,k} - 2v_{\text{scatt}} - C_{0,q}(k))^2}{2\sigma_{\nu}^2}} \right) e^{\frac{(\bar{\nu}_{q,k} - C_{0,q}(k))^2}{2\sigma_{\nu}^2} + 2v_{\text{scatt}}(1-\zeta)(C_{0,q}(k) - \bar{\nu}_{q,k})} \\ \frac{\sqrt{2\pi}\zeta\sigma_{\nu}^3 \left(Q\left(\frac{2v_{\text{scatt}} - \bar{\nu}_{q,k} + C_{0,q}(k)}{\sigma_{\nu}}\right) - Q\left(\frac{-2v_{\text{scatt}} - \bar{\nu}_{q,k} + C_{0,q}(k)}{\sigma_{\nu}}\right)\right) e^{\frac{(\bar{\nu}_{q,k} - C_{0,q}(k))^2}{2\sigma_{\nu}^2} + 2v_{\text{scatt}}\sigma_{\nu}^2(1-\zeta)}}{2\sigma_{\nu}^2}, k = k_1, \dots, k_{L_q^{\alpha}} \\ -\frac{\bar{\nu}_{q,k}}{\sigma_{\nu}^2}, \text{otherwise} \end{cases} \quad (64)$$

by ignoring terms not related to γ_q . Similarly, since (72) is concave, the maximum point can be derived as

$$\gamma_{q,s}^{(t)} = \frac{c_{1,q,s}^{\alpha} + 1}{c_{2,q,s}^{\alpha} + \left[\boldsymbol{\Sigma}_q\left(\dot{\Xi}_q^{\gamma}\right) + \boldsymbol{\mu}_q\left(\dot{\Xi}_q^{\gamma}\right) \boldsymbol{\mu}_q\left(\dot{\Xi}_q^{\gamma}\right)^H\right]_{s,s}}. \quad (73)$$

Algorithm 1 WCM-Based Joint CG and CFO Estimation

Input: $\tilde{\mathbf{y}}_q$, $\{\tilde{s}_{q,i}\}_{i \in \mathcal{I}_P}$, $\boldsymbol{\omega}_q$ and $C_{0,q}$ from the WCM.

Output: $\hat{\Xi}_q$ and $\hat{\mathbf{r}}_q^{\text{MMSE}}$.

Procedures:

- 1) Initialize $\bar{\nu}_q$ and γ_q according to $C_{0,q}$ and $\boldsymbol{\omega}_q$, respectively. Initialize $\hat{\mathbf{r}}_q^{\text{MMSE}} = \mathbf{0}_{MNP\tau}$, $\Delta\tau_q = 0$, the iteration counter $t = 1$, and set α_q to an arbitrary constant.
 - 2) Update $\bar{\nu}_q^{(t)}$ according to (69), $\alpha_q^{(t)}$ according to (71), $\gamma_q^{(t)}$ according to (73), and $\Delta\tau_q^{(t)}$ according to (81).
 - 3) Repeat step 2 until the stop criteria are satisfied or the maximum iteration is reached.
 - 4) Obtain the support set of \mathcal{S}_q^{γ} according to the estimated precision vector $\hat{\gamma}_q$, recover the non-zero elements in $\hat{\mathbf{r}}_q^{\text{MMSE}}$ by (49) with $\hat{\Xi}_q$.
-

4) *Update for $\Delta\tau_q$ at the t -Th Iteration:* There are methods for obtaining multipath delays, such as the rotational invariance techniques and delay tracking module proposed in [34]. In this paper, we propose another solution to estimate the multipath parameters altogether within the same Bayesian framework. As multipath delays are generated randomly and independently, we suppose that the off-grid delay offsets are uniformly distributed, and $\Delta\tau_q$ follows a noninformative prior distribution $\Delta\tau_q \sim U\left(\prod_{k=0}^{P\tau-1} [0, \Delta\bar{\tau}_k]\right)$. With $\{\alpha_q, \bar{\nu}_q, \gamma_q\} = \{\alpha_q^{(t)}, \bar{\nu}_q^{(t)}, \gamma_q^{(t)}\}$ and $\dot{\Xi}_q^{\tau} = \{\alpha_q^{(t)}, \Delta\tau_q^{(t-1)}, \bar{\nu}_q^{(t)}, \gamma_q^{(t-1)}\}$, (51) can be rewritten as

$$\begin{aligned} \mathcal{U}\left(\alpha_q^{(t)}, \Delta\tau_q, \bar{\nu}_q^{(t)}, \gamma_q^{(t)} \mid \dot{\Xi}_q^{\tau}\right) \\ = \int p\left(\mathbf{r}_q \mid \tilde{\mathbf{y}}_q, \dot{\Xi}_q^{\tau}\right) \ln p\left(\tilde{\mathbf{y}}_q \mid \mathbf{r}_q, \alpha_q^{(t)}, \Delta\tau_q, \bar{\nu}_q^{(t)}\right) d\mathbf{r}_q \\ + \ln p(\Delta\tau_q) \\ = -\alpha_q^{(t)} C_{1,q}\left(\Delta\tau_q, \bar{\nu}_q^{(t)}, \dot{\Xi}_q^{\tau}\right), \end{aligned} \quad (74)$$

after removing terms not related to $\Delta\tau_q$. The same as $\bar{\nu}_q$, $\Delta\tau_q$ is also deeply coupled in the measurement matrix, but

we can still obtain a suboptimal solution by minimizing $C_{1,q}(\Delta\tau_q, \bar{\nu}_q^{(t)}, \dot{\Xi}_q^\tau)$ along its gradient denoted by

$$\mathbf{g}_{\Delta\tau,q}(\Delta\tau_q) = [g_{\Delta\tau,q,0}(\Delta\tau_{q,k}), \dots, g_{\Delta\tau,q,P_\tau-1}(\Delta\tau_{q,k})]^\top. \quad (75)$$

It can be derived that

$$\begin{aligned} g_{\Delta\tau,q,k}(\Delta\tau_{q,k}) &= 2 \sum_{i=1}^{I_P} \sum_{p_1=1}^P \operatorname{Re} \left[(\mathcal{C}_{q,i,p_1}^\tau)^\text{H} \mathcal{C}'_{q,i,p_1,k} \right] \\ &\quad + \frac{2}{P^2} \sum_{p_1=1}^P \sum_{i=1}^{I_P} \\ &\quad \sum_{n=1}^N \operatorname{Re} \left[\operatorname{vec}^\text{T} (\mathcal{D}_{q,i,p_1,k+1,n}^\tau) \operatorname{vec} (\mathcal{K}'_{q,i,p_1,k+1,n}) \right], \end{aligned} \quad (76)$$

with components denoted as

$$\begin{aligned} \mathcal{C}_{q,i,p_1}^\tau &= \frac{1}{P} \sum_{n=1}^N \sum_{k=0}^{P_\tau-1} \sum_{p_2=1}^P [\bar{s}_{q,i,p_2}]_n \xi_{q,i,p_1,p_2}^k \left(\bar{\nu}_{q,k}^{(t)} \right) \\ &\quad \times e^{-j\frac{2\pi}{P}(\bar{\tau}_k + \Delta\tau_{q,k})(p_2-1)} \boldsymbol{\mu}_{q,k,n} \left(\dot{\Xi}_q^\tau \right) - \tilde{\mathbf{y}}_{q,i,p_1}, \end{aligned} \quad (77)$$

$$\begin{aligned} \mathcal{C}'_{q,i,p_1,k} &= -j \frac{2\pi}{P^2} \sum_{n=1}^N \sum_{p_2=1}^P (p_2-1) [\bar{s}_{q,i,p_2}]_n \xi_{q,i,p_1,p_2}^k \left(\bar{\nu}_{q,k}^{(t)} \right) \\ &\quad \times e^{-j\frac{2\pi}{P}(\bar{\tau}_k + \Delta\tau_{q,k})(p_2-1)} \boldsymbol{\mu}_{q,k,n} \left(\dot{\Xi}_q^\tau \right), \end{aligned} \quad (78)$$

$$\begin{aligned} \mathcal{D}_{q,i,p_1,k_1,n_1}^\tau &= \sum_{k_2=1}^{P_\tau} \sum_{n_2=1}^N \sum_{p_2=1}^P [\bar{s}_{q,i,p_2}]_{n_2}^* \left(\xi_{q,i,p_1,p_2}^{k_2-1} \left(\bar{\nu}_{q,k_2-1}^{(t)} \right) \right)^* \\ &\quad \times e^{(j\frac{2\pi}{P}(\bar{\tau}_{k_2-1} + \Delta\tau_{q,k_2-1})(p_2-1))} \boldsymbol{\Sigma}_{q,k_1,n_1,k_2,n_2}^\text{T} \left(\dot{\Xi}_q^\tau \right), \end{aligned} \quad (79)$$

$$\begin{aligned} \mathcal{K}'_{q,i,p_1,k_1,n_1}^\tau &= -j \frac{2\pi}{P} \sum_{p_2=1}^P (p_2-1) [\bar{s}_{q,i,p_2}]_{n_1} \xi_{q,i,p_1,p_2}^{k_1-1} \left(\bar{\nu}_{q,k_1-1}^{(t)} \right) \\ &\quad \times e^{(-j\frac{2\pi}{P}(\bar{\tau}_{k_1-1} + \Delta\tau_{q,k_1-1})(p_2-1))} \mathbf{I}_M. \end{aligned} \quad (80)$$

Then the t -th update for $\Delta\tau_q$ can be written as

$$\Delta\tau_q^{(t)} = \Delta\tau_q^{(t-1)} - \varepsilon_\tau^{(t)} \mathbf{g}_{\Delta\tau,q} \left(\Delta\tau_q^{(t-1)} \right), \quad (81)$$

with a step size defined in

$$\begin{aligned} \varepsilon_\tau^{(t)} &= \frac{\left(\Delta\tau_q^{(t-1)} - \Delta\tau_q^{(t-2)} \right)^\text{T}}{\left\| \mathbf{g}_{\Delta\tau,q} \left(\Delta\tau_q^{(t-1)} \right) - \mathbf{g}_{\Delta\tau,q} \left(\Delta\tau_q^{(t-2)} \right) \right\|^2} \\ &\quad \times \left(\mathbf{g}_{\Delta\tau,q} \left(\Delta\tau_q^{(t-1)} \right) - \mathbf{g}_{\Delta\tau,q} \left(\Delta\tau_q^{(t-2)} \right) \right). \end{aligned} \quad (82)$$

Since $\Delta\tau_{q,k} \in [0, 1]$, if $\Delta\tau_{q,k}^{(t)}$ falls below or exceeds the boundary, it will take the lower or upper bounds, respectively. The algorithm is summarized in Algorithm 1.

As in the calculation of step sizes for updating $\bar{\nu}_q^{(t)}$ and $\Delta\tau_q^{(t)}$, gradients of twice iterations before are needed, the step size in the first step will be set to $\frac{1}{100}$ of the corresponding value intervals, and the direction is determined by the positive or negative signs of gradients.

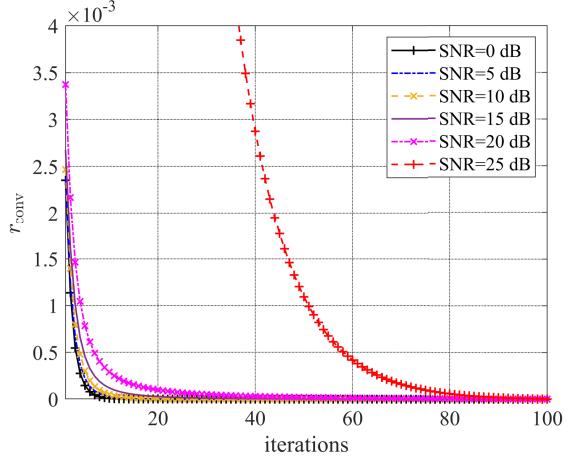


Fig. 5. Convergence performance of the proposed algorithm with different SNR settings.

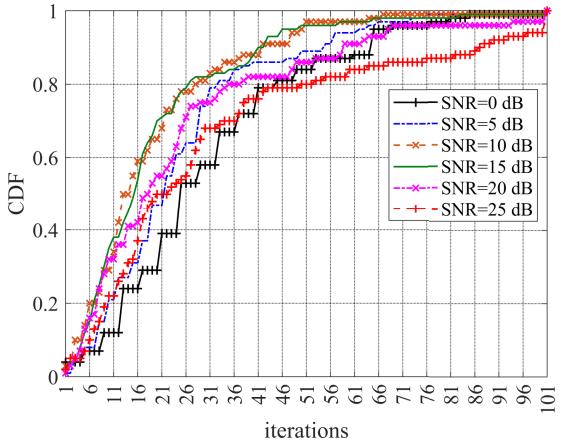


Fig. 6. Cumulative distribution function (CDF) of the number of iterations at convergence.

D. Convergence and Complexity Analysis

During the iteration of the proposed algorithm, the objective likelihood function is non-decreasing with the increase of t , since the surrogate function (51), as a lower bound, is non-decreasing in each iteration. In fact, they can only gradually approach an upper limit of 0. To validate the convergence of the proposed algorithm, the convergence accuracy, which is defined as the objective function value difference $\Delta\mathcal{U}^{(t)} = \mathcal{U}^{\text{opt}} - \mathcal{U}^{(t)}$ normalized by the “optimal” function value \mathcal{U}^{opt} , has been examined [35]. The average convergence accuracy after 60 Monte Carlo tests, i.e., $r_{\text{conv}} = \mathbb{E} \{ \Delta\mathcal{U}^{(t)} / \mathcal{U}^{\text{opt}} \}$, is demonstrated in Fig. 5, which validates the correctness of the proposed algorithm by showing the convergence accuracy of 10^{-8} . In addition, it can be found that the higher the SNR, the slower the convergence speed, since more iterations are needed for a higher estimation accuracy result. We also simulated the number of iterations when the proposed algorithm converges. As shown in Fig. 6, the proposed algorithm can converge within 100 iterations.

The computational complexity of the proposed algorithm mainly comes from the update of the PS. The complexity in calculating $\boldsymbol{\Sigma}_q$ is $\mathcal{O}(I_P M^3 N^2 P_\tau^2 P)$. The complexity

TABLE II
COMPLEXITY AND FUNCTIONALITY COMPARISON

Algorithm	Complexity	Functionality
The proposed algorithm	$\mathcal{O}(I_P M^3 N^2 P_\tau^2 P)$	Obtaining current CSI and multipath parameters for the subsequent channel prediction
The conventional method of DCS	$\mathcal{O}(M^3 N^2 G_P (L_q)^3)$	Obtaining current CSI, pilot signals are inserted in each symbol

for updating μ_q , α_q , and $\Delta\tau_q$ are the same as that for updating Σ_q , and the complexity for updating $\bar{\nu}_q$ and γ_q are $\mathcal{O}(I_P N^2 P^2 P_\tau)$ and $\mathcal{O}(MNP_\tau)$, respectively. Normally, I_P is set to 1, and the overall complexity of the proposed algorithm in MIMO systems is $\mathcal{O}(I_P M^3 N^2 P_\tau^2 P)$. The comparison of the complexity and functionality of the proposed algorithm and the baseline is given in Table II, where G_P indicates the number of effective pilot samples in a symbol. It is worth noting that, though the complexity of the proposed algorithm is higher, based on the extracted multipath parameters, high-precision model-based channel prediction can be carried out within a quasi-static time block, thereby increasing the channel sounding period effectively, and the channel state information (CSI) within an entire time block can be obtained through a single estimation with only one pilot symbol. Thus, the computational complexity of the proposed algorithm can be averaged over each symbol, making the computational complexity for a single channel acquisition comparable to that of the traditional ones.

V. RESULTS AND ANALYSIS

In this section, the performance of the proposed algorithm is analyzed and compared with that of baselines. The centimeter-wave channel at 10 GHz is simulated using parameters from [28], and the rural macro cell scenario is considered. The bandwidth f_s is set to 100 MHz, the number of grid points P_τ is set to 200, and the adjustable grid proportion η_τ is set to 0.9. The noise PSD is set to -174 dBm/Hz. The normalized mean square error (NMSE) is employed as a performance metric. Under M_c Monte Carlo tests, the NMSE of the estimation for \mathbf{a} is calculated as NMSE_a (dB) = $10 \lg \left(\frac{1}{MN M_c} \sum_{m_c=1}^{M_c} \frac{\|\hat{\mathbf{a}}_{m_c} - \mathbf{a}_{m_c}\|^2}{\|\mathbf{a}_{m_c}\|^2} \right)$, where $\hat{\mathbf{a}}_{m_c}$ and \mathbf{a}_{m_c} are the m_c -th estimate and real value, respectively. Three baselines are considered for performance comparisons: the DCS method [16], the SBL method [17], and the MMSE with a perfect PS [36]. By substituting the perfect CFOs and delays into \mathcal{K}_q , an optimal baseline can be obtained by performing MMSE estimation according to (47). The UE speed is set between 150 km/h and 350 km/h, the mean of scatterer mobility ζ is set between 0.2 and 0.4, where 0.2 is suggested in the standard 3GPP TR 38.901 [28], and the maximum clutter speed v_{scatt} is set between 1 and 21 m/s, covering the moving speeds from stride to vehicular. The quadrature phase-shift keying (QPSK) is employed for BER performance comparison. Both MMSE and zero-forcing (ZF) equalizers are considered. The equalizer is applied to the entire channel matrix $\tilde{\mathbf{H}}_{q,i}^{\text{ICI}}$ according to the transmission model in (30) to fully utilize the antenna and Doppler diversity gains. The WCM precision is denoted as r_ω , representing the similarity between the WCM-provided angular-delay PSD and the real value.

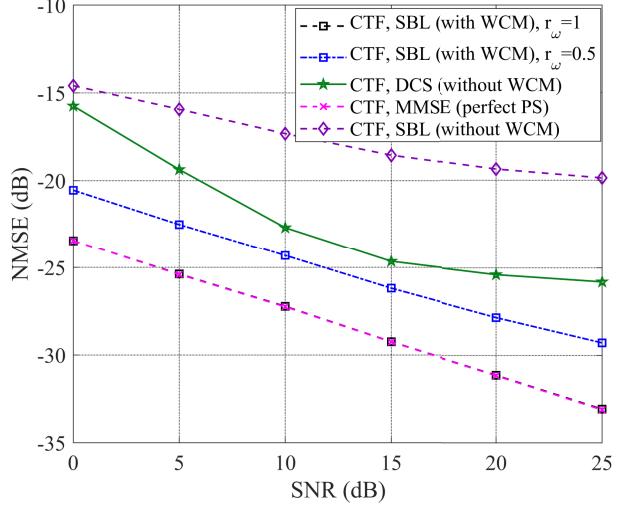


Fig. 7. NMSEs of the CTF estimation of the proposed algorithm and baselines under different SNRs ($v_q = 350$ km/h, $\zeta = 0.2$, and $v_{\text{scatt}} = 1$ m/s).

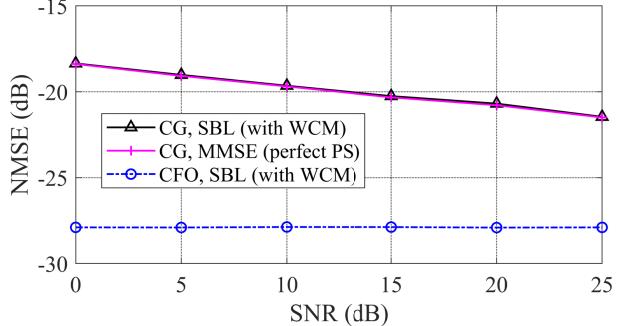


Fig. 8. NMSEs of the complex CG and CFO estimations of the proposed algorithm and baseline under different SNRs ($v_q = 350$ km/h, $\zeta = 0.2$, $v_{\text{scatt}} = 1$ m/s, and $r_\omega = 1$).

A. NMSE Performance

The NMSE performance of the CTF estimation of the proposed algorithm in comparison with baselines is shown in Fig. 7. Under different SNR settings, the proposed algorithm (SBL with the WCM) always outperforms the DCS and SBL without the WCM, no matter when r_ω equals 1 or drops to 0.5. With the aid of pre-constructed channel information provided by the WCM, the proposed WCM-based algorithm can reduce the CTF estimation error to less than 13% of that of the SBL without the WCM when $r_\omega = 1$. In Fig. 8, the NMSEs of the CG and CFO estimations in comparison with the baseline are shown. When $r_\omega = 1$, the NMSE of the CG estimation for the proposed algorithm is almost identical to that of the MMSE with a perfect PS. Additionally, the CFO estimation accuracy is nearly the same for different SNR settings. This is because, with finite time samples, the phase shift caused by mobility is a periodic function of the CFO, as shown in (16). The CFO obtained after convergence is only the equivalent optimal

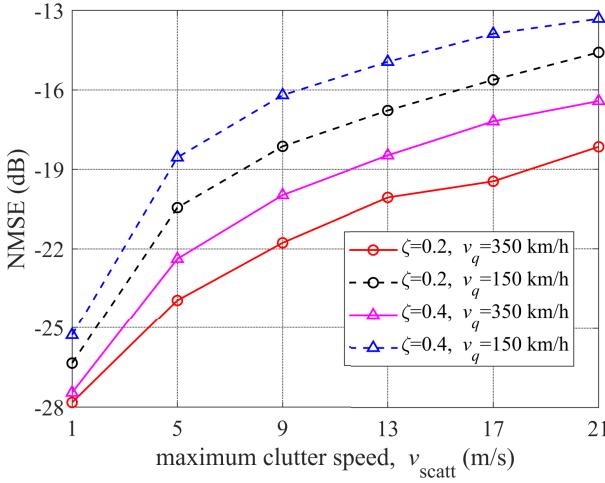


Fig. 9. NMSEs of the CFO versus the maximum clutter speed (SNR = 10 dB, $r_\omega = 1$ m/s).

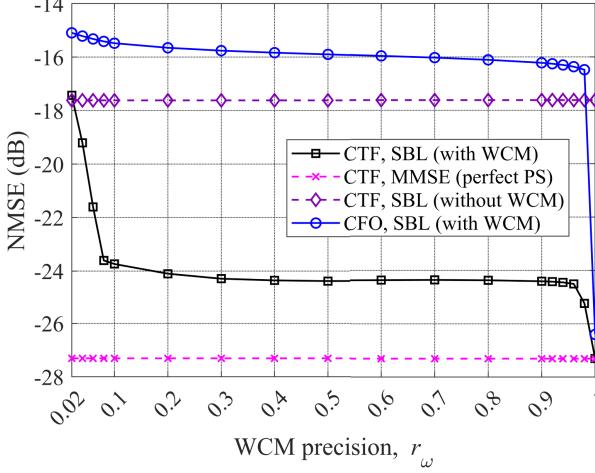


Fig. 10. NMSEs of the CTF and CFO versus WCM precision ($v_q = 150 \text{ km/h}$, $\zeta = 0.2$, $v_{\text{scatt}} = 1 \text{ m/s}$, and SNR = 10 dB).

solution, and its NMSE depends more on the initialization, i.e., the WCM precision. It is worth noting that such a solution has only a slight impact on the final NMSE performance of the CTF estimation, since the obtained CFO estimation is an equivalent optimal solution.

The CFO estimation performance for different mobility settings is also shown in Fig. 9. When the maximum clutter speed \$v_{\text{scatt}}\$ increases, the CFO NMSE goes up and gradually flattens, and as the moving scatterer ratio \$\zeta\$ increases, the CFO estimation accuracy declines correspondingly as well. This is all due to the increasing uncertainty of ambient moving scatterers. The larger the values of \$v_{\text{scatt}}\$ and \$\zeta\$ are, the more moving scatterers with faster speeds are in the propagation environment. It can also be noticed that the proposed algorithm can achieve a better CFO estimation with higher UE speeds, because the DFS induced by terminal mobility contributes more to the CFO, and an accurate DFS estimation based on the WCM leads to a lower CFO NMSE at higher UE speeds.

We further investigate the effect of WCM precision \$r_\omega\$ on the NMSE performance. As can be seen in Fig. 10, when WCM precision decreases, the CFO estimation accuracy of the proposed algorithm first drops and then remains at an

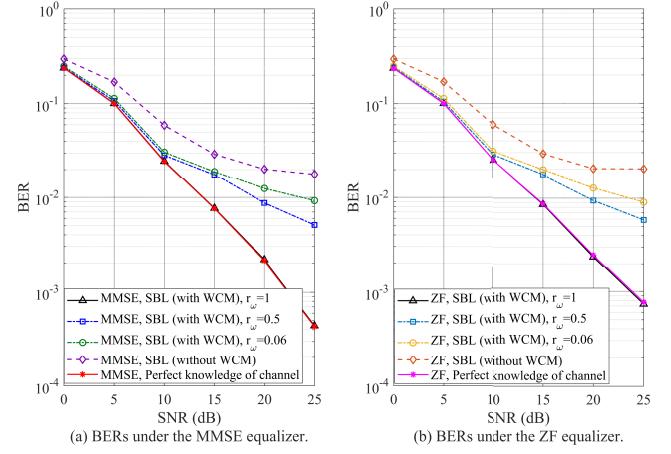


Fig. 11. BERs of the proposed algorithm and baselines with (a) the MMSE equalizer and (b) the ZF equalizer under different SNRs ($v_q = 350 \text{ km/h}$, $\zeta = 0.2$, and $v_{\text{scatt}} = 1 \text{ m/s}$).

acceptable level. Though the CFO estimation accuracy is affected by the initialization, the redundancy for calculation errors set up in the CFO prior distributions still makes it adaptive to different WCM precision. Similarly, as the WCM precision decreases, the CTF estimation accuracy of the proposed algorithm slightly decreases first, but can still remain at a good level until the WCM precision drops below 0.1. The proposed algorithm shows good adaptability to different WCM precision, and given a fixed channel estimation accuracy threshold, the WCM only needs to reach the minimum accuracy bound. For example, a CTF NMSE below -24.3 dB only requires the WCM precision greater than 0.4. The CTF estimation is less affected by the degradation of WCM precision, because the optimal solution of the CFO estimation may not be unique, while that of the CTF estimation is almost unique and therefore less affected by the prior distribution settings. Additionally, comparing the results in Fig. 10 and Fig. 7 when SNR=10 dB and \$r_\omega = 1\$, we can see that the increasing UE speed has little effect on the CTF NMSE. Since the estimated CTF matrix covers full-dimensional channel information including ICI, the proposed algorithm is highly robust to different UE speeds.

B. BER Performance

In addition to the NMSE, the BER performance of the proposed algorithm is also compared and analyzed. We can see from Fig. 11 that, for either the MMSE or ZF equalizers, the BERs of the proposed algorithm with different WCM precision are always lower than that of the SBL without the WCM, which indicates the benefits from the WCM. When $r_\omega = 1$, the BER performance of the proposed algorithm can reach that with perfect knowledge of the channel. The proposed algorithm also shows excellent adaptability to different WCM precision in BER performance at low SNRs, even when the WCM accuracy drops to 0.5, the decrease of the BER performance is still very limited compared to that when $r_\omega = 1$. Overall, the advantage of the MMSE equalizer over that of the ZF is more obvious when the channel estimation is more accurate, i.e., when the WCM precision is higher. It can also be found that with more accurate channel prior information,

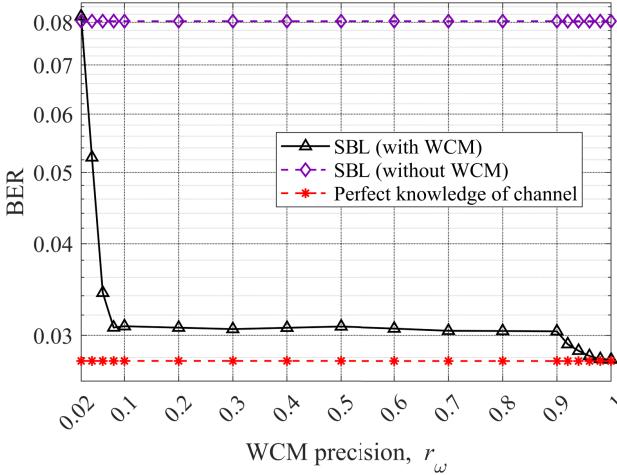


Fig. 12. BERs of the proposed algorithm versus WCM precision under the ZF equalizer ($v_q = 150$ km/h, $\zeta = 0.2$, $v_{\text{scatt}} = 1$ m/s, and SNR = 10 dB).

better transmission performance can be reached, especially at high SNRs.

In Fig. 12, the impact of WCM precision on the BER performance of the proposed algorithm is also investigated. As the WCM precision declines, the BER performance experiences a slight decrease and then stabilizes at a near-optimal value until the WCM precision drops below 0.1. The BER performance of the proposed algorithm always outperforms that of the SBL without the WCM. The trend of the BER performance changing with WCM precision is consistent with that of the CTF NMSE because the BER performance is basically determined by the CTF estimation accuracy under the same SNR. Such results also reflect the significance of WCM-provided prior information of the propagation environment in improving transmission performance, even if the provided information is lossy. Besides, by comparing results in Fig. 12 with those in Fig. 11 (b) when SNR=10 dB and $r_\omega = 1$, it can be found that the BER performance is better at higher speeds, due to the additional Doppler diversity gain [37]. For example, when SNR=10 dB and $r_\omega = 1$, the BER at 150 km/h is 0.0277, while at 350 km/h it is 0.0251.

VI. CONCLUSION

In this paper, a WCM-based joint complex CG and CFO estimation algorithm has been proposed for doubly-selective channels in high-mobility scenarios. A WCM-based prior distribution construction method for the complex CG and CFO has also been proposed. Based on the nonuniform off-grid channel representation, the joint estimation problem has been analyzed and solved by sparse signal recovery using SBL and MM algorithms. Simulation results have demonstrated the superiority of the proposed algorithm with its adaptability to different WCM precision and the capability of achieving better NMSE and BER performance than the SBL without the WCM and DCS. When SNR=10 dB, compared with SBL without the WCM, the proposed algorithm can reduce the channel estimation error from -17.3 dB to -27.3 dB, achieving a tenfold improvement in estimation accuracy. It has been demonstrated that the proposed algorithm can achieve effective multipath parameter extraction, which is necessary for the

subsequent model-based channel prediction. It has also been demonstrated that the proposed algorithm is highly robust to different UE speeds in full-dimensional channel information acquisition and can achieve even better BER performance at higher speeds due to additional Doppler diversity. It has also been verified that with more accurate prior knowledge of the propagation environment, higher communication performance can be achieved.

APPENDIX A CALCULATION OF MULTIPATH DIRECTIONS

Take the equivalent azimuth angle as an example. The virtual direction with the maximum amplitude $\bar{a}_{q,k}^{\max}$ approaches the closest to the true path direction. As shown in Fig. 4, the interval between two virtual directions is fixed at $\frac{1}{N_1}$, it can thus be derived that

$$d_1 = \frac{1}{\pi} \arctan \left(\frac{\eta_{az} \sin \left(\frac{\pi}{N_1} \right)}{1 + \eta_{az} \cos \left(\frac{\pi}{N_1} \right)} \right), \quad (83)$$

with the amplitude ratio $\eta_{az} = \left| \frac{\bar{a}_{q,k}}{\bar{a}_{q,k}^{\max}} \right|$. When N_1 gets large, (83) can be simplified as

$$d_1 \xrightarrow{N_1 \rightarrow \infty} \frac{\eta_{az}}{N_1 (1 + \eta_{az})}. \quad (84)$$

From (34), we can know that the actual value of $\vartheta_{q,k}^U$ can be seen as how far the virtual azimuth directions have been moved from their original positions $\frac{2b_1-1}{2N_1} - 0.5$. As $\vartheta_{q,k}^U$ takes values between $[-0.5, 0.5]$, when $b_1 < b_2$ or $b_1 - b_2 = N_1 - 1$, $\frac{2b_1-1}{2N_1} + \vartheta_{q,k}^U - 0.5$ should be equal to the negative d_1 , and when $b_1 > b_2$ or $b_1 - b_2 = -N_1 + 1$, $\frac{2b_1-1}{2N_1} + \vartheta_{q,k}^U - 0.5$ should be equal to the positive d_1 , then (36) can be obtained. The equivalent elevation angle $\varphi_{q,k}^U$ is calculated in the same way.

APPENDIX B PRIOR DISTRIBUTION CONSTRUCTION FOR $\bar{\nu}_q$

For $k = k_0$, it can be derived that $p(\bar{\nu}_{q,k_0}; \bar{\Omega}_q) = \mathcal{N}(\bar{\nu}_{q,k_0}; C_{0,q}(k_0), \sigma_{\nu,k_0}^2)$. When $k \notin \{k_0, \dots, k_{L'_q}\}$, we have $p(\bar{\nu}_{q,k}; \bar{\Omega}_q) = \mathcal{N}(\bar{\nu}_{q,k}; C_{0,q}(k), \sigma_{\nu,k}^2)$. If $Z_{q,k} = \varpi_{q,k} v_{q,k}$, we can know that $Z_{q,k} \in [-v_{\text{scatt}}, v_{\text{scatt}}]$, then the cumulative distribution function of $Z_{q,k}$ will be

$$\begin{aligned} F_{Z_{q,k}}(z_{q,k}) &= P(\varpi_{q,k} = 1) P(v_{q,k} \leq z_{q,k}) + P(\varpi_{q,k} = 0) P(0 \leq z_{q,k}) \\ &= \frac{z_{q,k} + v_{\text{scatt}}}{2v_{\text{scatt}}} \zeta + (1 - \zeta) \mathcal{X}(z_{q,k}), \end{aligned} \quad (85)$$

where \mathcal{X} stands for the step function, $z_{q,k} \in [-v_{\text{scatt}}, v_{\text{scatt}}]$. The probability density function (PDF) of $Z_{q,k}$ can thus be written as

$$\begin{aligned} p(Z_{q,k}) &= f_{Z_{q,k}}(z_{q,k}) \\ &= \begin{cases} \frac{\zeta}{2v_{\text{scatt}}} + (1 - \zeta) \delta(z_{q,k}), & z_{q,k} \in [-v_{\text{scatt}}, v_{\text{scatt}}] \\ 0, & \text{otherwise.} \end{cases} \end{aligned} \quad (86)$$

As $C_{0,q}(k) + n_{\nu,k}$ and $\nu_{q,k}^S$ from (38) can be viewed independent of each other, $p(\bar{\nu}_{q,k}; \bar{\Omega}_q)$ with $k \in \{k_1, \dots, k_{L'_q}\}$ can be calculated as

$$p(\bar{\nu}_{q,k}; \bar{\Omega}_q) = p(2Z_{q,k}) * p(C_{0,q}(k) + n_{\nu,k}), \quad (87)$$

where the former can be calculated by the method of random variable transformation, and the latter is a Gaussian distribution with a mean of $C_{0,q}(k)$. With all the above, (40) can then be obtained.

REFERENCES

- [1] C.-X. Wang et al., "On the road to 6G: Visions, requirements, key technologies and testbeds," *IEEE Commun. Surveys Tuts.*, vol. 25, no. 2, pp. 905–974, 2nd Quart., 2023.
- [2] C.-X. Wang, Y. Yang, J. Huang, X. Gao, T. J. Cui, and L. Hanzo, "Electromagnetic information theory: Fundamentals and applications for 6G wireless communication systems," *IEEE Wireless Commun.*, vol. 31, no. 5, pp. 279–286, Oct. 2024.
- [3] C.-X. Wang et al., "Modeling, capacity studies, antenna and system designs for 6G/B6G 3-D continuous-space radio channels enabled by electromagnetic information theory," *IEEE Commun. Surveys Tuts.*, vol. 28, pp. 1–63, 2026.
- [4] J. Wu and P. Fan, "A survey on high mobility wireless communications: Challenges, opportunities and solutions," *IEEE Access*, vol. 4, pp. 450–476, 2016.
- [5] F. Luo, *Machine Learning for Future Wireless Communications*. Hoboken, NJ, USA: Wiley, 2019.
- [6] T. Roman and V. Koivunen, "Subspace method for blind CFO estimation for OFDM systems with constant modulus constellations," in *Proc. IEEE VTC*, vol. 2, Jun. 2005, pp. 1253–1257.
- [7] E. P. Simon, L. Ros, H. Hijazi, and M. Ghogho, "Joint carrier frequency offset and channel estimation for OFDM systems via the EM algorithm in the presence of very high mobility," *IEEE Trans. Signal Process.*, vol. 60, no. 2, pp. 754–765, Feb. 2012.
- [8] H. Abdzadeh-Ziabari, W.-P. Zhu, and M. N. S. Swamy, "Joint maximum likelihood timing, frequency offset, and doubly selective channel estimation for OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 67, no. 3, pp. 2787–2791, Mar. 2018.
- [9] P. Muneer and S. M. Sameer, "Iterative joint carrier frequency offset and doubly selective channel estimation in high-mobility MIMO-OFDMA uplink using oblique projection," *IEEE Trans. Veh. Technol.*, vol. 65, no. 9, pp. 7110–7121, Sep. 2016.
- [10] E. P. Simon, H. Hijazi, L. Ros, M. Berbineau, and P. Degauque, "Joint estimation of carrier frequency offset and channel complex gains for OFDM systems in fast time-varying vehicular environments," in *Proc. IEEE ICC Workshops*, vol. 57, May 2010, pp. 1–5.
- [11] H. Hijazi, A. Dakhllallah, S. A. Hajj, A. C. A. Ghouwayel, and A. Dhayni, "Joint CFO and time-varying channel estimation by particle filtering in OFDM systems," in *Proc. ICCIT*, vol. 57, Jun. 2013, pp. 241–245.
- [12] Y. Zhu, H. Guo, and V. K. N. Lau, "Bayesian channel estimation in multi-user massive MIMO with extremely large antenna array," *IEEE Trans. Signal Process.*, vol. 69, pp. 5463–5478, 2021.
- [13] 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.
- [14] P. Cheng et al., "Channel estimation for OFDM systems over doubly selective channels: A distributed compressive sensing based approach," *IEEE Trans. Commun.*, vol. 61, no. 10, pp. 4173–4185, Oct. 2013.
- [15] Q. Qin, L. Gui, B. Gong, X. Ren, and W. Chen, "Structured distributed compressive channel estimation over doubly selective channels," *IEEE Trans. Broadcast.*, vol. 62, no. 3, pp. 521–531, Sep. 2016.
- [16] B. Gong, L. Gui, Q. Qin, X. Ren, and W. Chen, "Block distributed compressive sensing-based doubly selective channel estimation and pilot design for large-scale MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 66, no. 10, pp. 9149–9161, Oct. 2017.
- [17] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *J. Mach. Learn. Res.*, vol. 1, pp. 211–244, Jun. 2001.
- [18] J. Dai, A. Liu, and V. K. N. Lau, "FDD massive MIMO channel estimation with arbitrary 2D-array geometry," *IEEE Trans. Signal Process.*, vol. 66, no. 10, pp. 2584–2599, May 2018.
- [19] G. Liu, A. Liu, R. Zhang, and M. Zhao, "Angular-domain selective channel tracking and Doppler compensation for high-mobility mmWave massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 20, no. 5, pp. 2902–2916, May 2021.
- [20] S. M. Kay, *Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, 1993.
- [21] S. Chen et al., "Channel map-based angle domain multiple access for cell-free massive MIMO communications," *IEEE J. Sel. Topics Signal Process.*, vol. 19, no. 2, pp. 366–380, Mar. 2025.
- [22] H. Sun and J. Chen, "Integrated interpolation and block-term tensor decomposition for spectrum map construction," *IEEE Trans. Signal Process.*, vol. 72, pp. 3896–3911, 2024.
- [23] X. Xu and Y. Zeng, "How much data is needed for channel knowledge map construction?," *IEEE Trans. Wireless Commun.*, vol. 23, no. 10, pp. 13011–13021, Oct. 2024.
- [24] J. Li, C.-X. Wang, C. Huang, T. Qi, and T. Wu, "Digital twin online channel modeling: Challenges, principles, and applications," *IEEE Veh. Technol. Mag.*, vol. 20, no. 1, pp. 94–103, Mar. 2025.
- [25] T. Qi, C. Huang, J. Shi, J. Li, S. Chen, and C. Wang, "A novel dynamic channel map for 6G MIMO communications," in *Proc. IEEE ICCC*, Sep. 2024, pp. 809–814.
- [26] S. Yang, C.-X. Wang, Y. Wang, J. Huang, Y. Zhou, and E.-H.-M. Aggoun, "An efficient pre-processing method for 6G dynamic ray-tracing channel modeling," *IEEE Trans. Veh. Technol.*, vol. 74, no. 5, pp. 6941–6953, May 2025.
- [27] T. Qi et al., "A novel 6G dynamic channel map based on a hybrid channel model," *IEEE Trans. Veh. Technol.*, early access, Aug. 26, 2025, doi: [10.1109/TVT.2025.3602807](https://doi.org/10.1109/TVT.2025.3602807).
- [28] *Study on Channel Model for Frequencies From 0.5 To 100 GHz (Release 19)*, document TR 38.901, Sep. 2025.
- [29] W. Guo, W. Zhang, P. Mu, F. Gao, and H. Lin, "High-mobility wideband massive MIMO communications: Doppler compensation, analysis and scaling laws," *IEEE Trans. Wireless Commun.*, vol. 18, no. 6, pp. 3177–3191, Jun. 2019.
- [30] J. Ma, S. Zhang, H. Li, F. Gao, and S. Jin, "Sparse Bayesian learning for the time-varying massive MIMO channels: Acquisition and tracking," *IEEE Trans. Commun.*, vol. 67, no. 3, pp. 1925–1938, Mar. 2019.
- [31] C.-X. Wang, Z. Lv, X. Gao, X. You, Y. Hao, and H. Haas, "Pervasive wireless channel modeling theory and applications to 6G GBSMs for all frequency bands and all scenarios," *IEEE Trans. Veh. Technol.*, vol. 71, no. 9, pp. 9159–9173, Sep. 2022.
- [32] C. Wang et al., "An enhanced 6G pervasive channel model towards standardization," *Scientia Sinica Informationis*, vol. 55, no. 6, p. 1486, Jun. 2025.
- [33] A. M. Sayeed, "Deconstructing multiantenna fading channels," *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2563–2579, Oct. 2002.
- [34] B. Yang, K. B. Letaief, R. S. Cheng, and Z. Cao, "Channel estimation for OFDM transmission in multipath fading channels based on parametric channel modeling," *IEEE Trans. Commun.*, vol. 49, no. 3, pp. 467–479, Mar. 2001.
- [35] S. Chen, J. Zhang, E. Björnson, Ö. T. Demir, and B. Ai, "Energy-efficient cell-free massive MIMO through sparse large-scale fading processing," *IEEE Trans. Wireless Commun.*, vol. 22, no. 12, pp. 9374–9389, Dec. 2023.
- [36] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO networks: Spectral, energy, and hardware efficiency," *Found. Trends Signal Process.*, vol. 11, nos. 3–4, pp. 154–655, 2017.
- [37] J. Wu and Y. R. Zheng, "Oversampled orthogonal frequency division multiplexing in doubly selective fading channels," *IEEE Trans. Commun.*, vol. 59, no. 3, pp. 815–822, Mar. 2011.



Yinglan Bu received the B.E. degree in communication engineering from Zhengzhou University, China, in 2020. She is currently pursuing the Ph.D. degree with the National Mobile Communications Research Laboratory, School of Information Science and Engineering, Southeast University, China. Her research interests include channel estimation, optimization theory, and 6G wireless 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 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 was with Heriot-Watt University, Edinburgh, U.K., from 2005 to 2018, where he was promoted to a Professor in 2011. He has been a Professor with Southeast University, Nanjing, China, since 2018. He is currently the Dean of the School of Information Science and Engineering. He is also a Professor with Pervasive Communication Research Center, Purple Mountain Laboratories, Nanjing, China. He has authored four books, three book chapters, and over 660 papers in refereed journals and conference proceedings, including 31 highly cited papers. He has also delivered 39 invited keynote speeches/talks and 23 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 European Academy of Sciences and Arts (EASA), a fellow of the Royal Society of Edinburgh (FRSE) and IET, an IEEE Communications Society Distinguished Lecturer in 2019 and 2020, and a Highly-Cited Researcher recognized by Clarivate Analytics. He received the IEEE Neal Shepherd Memorial Best Propagation Paper Award in 2024. He also received 19 Best Paper Awards from the international conferences. He has served as the TPC chair and the general chair for more than 30 international conferences. He is currently an Executive Editorial Committee Member of IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. He has served as an Editor for over sixteen international journals, including IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, from 2007 to 2009, IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, from 2011 to 2017, and IEEE TRANSACTIONS ON COMMUNICATIONS, from 2015 to 2017. He was a Guest Editor of IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, IEEE TRANSACTIONS ON BIG DATA, and IEEE TRANSACTIONS ON COGNITIVE COMMUNICATIONS AND NETWORKING.



Shuaifei Chen (Member, IEEE) received the B.S. degree in communication engineering and the Ph.D. degree in information and communication engineering from Beijing Jiaotong University, China, in 2018 and 2023, respectively.

From 2019 to 2020, he visited the Department of Communication Systems, Linköping University, Sweden. From 2021 to 2022, he visited the Division of Communication Systems, KTH Royal Institute of Technology, Sweden. He is currently a Research Associate Professor with the Purple Mountain Laboratories, Nanjing, China, and also a Post-Doctoral Researcher with the National Mobile Communications Research Laboratory, School of Information Science and Engineering, Southeast University, Nanjing, China. Since January 2025, he has been an External Supervisor with Southeast University. His research interests include signal processing and resource allocation for wireless communications, cell-free and ultra-massive MIMO, and digital twin-aided 6G multiple antenna technologies.. He also serves as a Technical Program Committee (TPC) Member for several conferences, including IEEE ICC and IEEE GLOBECOM. He was recognized as an Exemplary Reviewer of IEEE TRANSACTIONS ON COMMUNICATIONS in 2021 and IEEE COMMUNICATIONS LETTERS in 2023. He was selected in the Outstanding Postdoctoral Fellow Program in Jiangsu in 2024 and the Young Elite Scientists Sponsorship Program by Jiangsu Association for Science and Technology (JAST) in 2025. He was a recipient of the Best Paper Award from IEEE ICC 2025. He serves as an Associate Editor for IEEE OPEN JOURNAL OF THE COMMUNICATIONS SOCIETY.



Junling Li (Member, IEEE) received the B.S. degree from Tianjin University, Tianjin, China, the M.S. degree from Beijing University of Posts and Telecommunications, Beijing, China, in 2013 and 2016, respectively, and the Ph.D. degree from the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada, in 2020. She was a Joint Post-Doctoral Research Fellow with Shenzhen Institute of Artificial Intelligence and Robotics for Society (AIRS), University of Waterloo, and The Chinese University of Hong

Kong, Shenzhen, from 2020 to 2022. She is currently an Associate Professor with the National Mobile Communications Research Laboratory, Southeast University, Nanjing, China. Her research interests include digital twin, channel modeling, game theory, and machine learning-based channel prediction. She received the Best Paper Award at IEEE ICCC 2019, ICCT 2023, and ICC 2025.



Jianghan Ji received the B.E. degree in information engineering from Southeast University, Nanjing, China, in 2023, where she is currently pursuing the M.Sc. degree with the National Mobile Communications Research Laboratory. Her research interests include channel estimation, and channel map.



Yunfei Chen (Fellow, IEEE) received the B.E. and M.E. degrees in electronics engineering from Shanghai Jiao Tong University, Shanghai, China, in 1998 and 2001, respectively, and the Ph.D. degree from the University of Alberta, Edmonton, AB, Canada, in 2006. He is currently a Professor with the Department of Engineering, University of Durham, U.K. His research interests include wireless communications, performance analysis, joint radar communications designs, cognitive radios, wireless relaying, and energy harvesting.