Reliable Transmission in FBMC-Based OTFS Systems With TF Domain Pilot-Aided Channel Estimation and Equalization

Ying Wang, Qiang Guo, Jianhong Xiang, *Member, IEEE*, and Yu Zhong

***Abstract*—Orthogonal Time Frequency Space (OTFS) modulation can effectively support high-mobility communication scenarios. However, Orthogonal Frequency Division Multiplexing (OFDM) based OTFS suffers from high spectrum leakage. Implementing channel estimation and equalization that simultaneously supports both OTFS and OFDM also faces challenges. In this paper, we adopt Filter Bank Multi-Carrier (FBMC) modulation as an alternative to OFDM and propose a TF-domain pilot-aided channel estimation and equalization scheme, which can improve spectrum leakage and enhance compatibility. Specifically, first, based on the core function of the prototype filter, we choose the Hermite prototype filter with symmetric properties to construct the FBMC-based OTFS system, enhancing adaptability to dynamic channels. Second, considering the dynamic characteristics of fast time-varying channels, we construct Bessel fitting or a priori information-assisted channel extrapolation mechanisms to achieve accurate tracking of channel parameters. Finally, we derive the criterion for determining the channel wide-sense stationarity time interval, which provides a basis for the update mechanism of prior information. Simulation results show that the proposed scheme can work robustly on doubly-selection channels. Compared to classical OTFS, FBMC-based OTFS significantly improves reliability in high mobility scenarios.**

***Index Terms*—OTFS, channel estimation and equalization, Bessel fitting, critical wide-sense stationary time interval, doubly selective channel.**

I. Introduction

NEXT-GENERATION wireless communications should fulfill the demand for ultra-high reliable transmission. The 3rd Generation Partnership Project has decided to continue adopting Orthogonal Frequency Division Multiplexing (OFDM) as the waveform for 5G [1]. Although this decision is significant for 4G compatibility, it is not the most efficient technology for next-generation wireless communications. It is well known that in high mobility communication scenarios, channels generally exhibit fast time-varying characteristics with high Doppler spread. In such high Doppler conditions, the performance of OFDM degrades [2]. Therefore, Orthogonal Time Frequency Space (OTFS) modulation has become an emerging technology for next-generation communication. OTFS can fulfill highly reliable wireless transmission for high mobility scenarios [3], [4]. However, classical OTFS is regarded as precoded OFDM or a collection of multiple OFDM signals [5], [6]. Thus, OFDM-based OTFS inherits some drawbacks from OFDM (e.g., high Out-of-Band (OOB) emissions). Compared to OFDM, Filter Bank Multi-Carrier (FBMC) modulation has superior spectral properties and is more robust in doubly selective channels [7]. We expect to achieve an orders-of-magnitude improvement in reliability and optimize other key features by switching the underlying strategy of OTFS from OFDM to FBMC. However, such a switch also introduces new challenges, such as an increased demand for time resources.

Although FBMC-based OTFS may have certain limitations, its research remains important because its ultra-high reliability and superior spectral properties exhibit potential use cases. In this paper, we focus on a pilot-aided channel estimation and equalization framework integrated in the Time-Frequency (TF) domain, which is a key component in achieving the ultra-high reliability of FBMC-based OTFS. It is worth noting that the proposed scheme is also compatible with classical OTFS (i.e., OFDM-based OTFS).

A. Related Work

Pishvaei et al. [8] designed and implemented an FBMC-based OTFS system. In addition, the authors also analyzed the Bit Error Rate (BER) and spectral efficiency of FBMC-based OTFS. However, since the focus is on system implementation and analysis, the authors do not provide implementation clues for key components (i.e., channel estimation and equalization). Deng et al. [9] conducted a systematic survey for OTFS and its variants. They pointed out that OTFS can be regarded as an FBMC system with Inverse Symplectic Fast Fourier Transform (ISFFT) as a precoder. However, the design of receiver algorithms, pilot schemes, and performance evaluation under high-mobility channels still lack systematic investigation. Shafie et al. [6] discussed the coexistence of OTFS modulation with OFDM-based communication systems

Received 15 October 2025; revised 30 December 2025; accepted 17 February 2026. Date of publication 26 February 2026; date of current version 4 March 2026. This work is supported in part by Fundamental Research Funds for the Central Universities, China [3072022CF0801], and Sichuan Provincial Key Laboratory of Agile Intelligent Computing Project Fund. The associate editor coordinating the review of this article and approving it for publication was Z. Wei. *(Corresponding author: Jianhong Xiang.)*

Ying Wang, Qiang Guo, and Jianhong Xiang are with the College of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China (e-mail: wangyingstu@163.com; guoqiang@hrbeu.edu.cn; xiangjianhong@hrbeu.edu.cn).

Yu Zhong is with the 10th Research Institute, China Electronics Technology Group Corporation, Chengdu 610036, China (e-mail: jade.zhong@hotmail.com).

Digital Object Identifier 10.1109/TCOMM.2026.3668148

and proposed an interference cancellation-based channel estimation technique. Channel estimation based on interference cancellation improves estimation accuracy. However, iterative solving of the maximum likelihood estimation and linear equations may lead to excessively high computational complexity. Specifically, in high Doppler spread or multipath scenarios, the step of iteration increases with the number of paths and the maximum Doppler index, leading to a significant increase in complexity. Chatterjee et al. [10] proposed a Non-Orthogonal Multiple Access (NOMA) scheme based on OTFS and developed a unified matrix framework that supports both OTFS and OFDM, facilitating algorithm development. Compared with traditional schemes, NOMA-based OTFS exhibits enhanced interference resilience. Liao et al. [11] proposed a pilot-assisted channel estimation scheme for SIMO-OFDM systems, which is applicable to generalize channel estimation in complex multipath environments and distributed antenna systems. However, the assumption of a quasi-static channel fails to account for the impact of rapid channel variations. Raviteja et al. [2] proposed an embedded pilot-aided channel estimation scheme. For integer and fractional Doppler shifts, they proposed different symbol arrangement methods and employed the threshold method for channel estimation. Although the authors mentioned the existence of an optimal threshold, the time-varying nature of the channel may cause the fixed threshold cannot be adapted to the environment. Liu et al. [12] proposed a Turbo BEM OTFS receiver, which is suitable for high-mobility communication channels Doppler-shift or Doppler-spread. However, the BEM order needs to be dynamically adjusted according to the degree of Doppler spread. Zhang et al. [4] proposed a cross-domain iterative OTFS receiver for sparse doubly selective channels, which alternately performs channel estimation and signal detection in the TF domain and Delay-Doppler (DD) domain. However, the performance of the Multi-Response Sparse Bayesian Learning (MSBL) algorithm is highly dependent on the sparsity assumption of the channel. If the number of multipaths increases or the sparsity dynamically fluctuates, leading to a significant increase in the number of non-zero paths, the estimation accuracy of MSBL will dramatically decrease. Thaj et al. [13] proposed an iterative Rake decision feedback equalizer for ZP-OTFS. The authors assume that the channel delay is an integer multiple of the sampling period. However, in high-resolution or multipath-dense scenarios, the wireless channel delay is typically fractional [14]. The assumption of integer delays may lead to model mismatches. Zhang et al. [15] proposed low-overhead OTFS transmission with frequency-domain or time-domain channel estimation. Both frequency-domain and time-domain channel estimation rely on interpolation algorithms. The time-domain phase shift interpolation assumes that the channel parameters vary linearly within the interpolation period, which makes it difficult to accurately track fast-varying channel parameters. Pilot-assisted Compressed Sensing (CS) channel estimation is the most popular approach. The reason is that CS theory can exploit the sparsity of the channel in the DD domain. For example: CE-BEM [16] and generalized CS BEM [17] applied to OTFS can effectively improve channel estimation performance. However, the choice of BEM order requires a trade-off between complexity and fitting accuracy. A too-low order will result in increased estimation errors, while a too-high order will increase computational complexity. The CS-based DD domain estimation exploits sparsity to reduce pilot overhead [18], but at the cost of increased reconstruction complexity. For example, Wang et al. [19] pointed out that the reconstruction algorithms in CS involve high-order matrix operations and hardware-unfriendly nonlinear functions, making the reconstruction complexity high and limiting the practical deployment.

Despite being widely regarded as a candidate technology for 6G, OTFS has not yet been adopted in current 5G systems, which still employ a variant of OFDM [1], [20] with data symbols mapped in TF domain. Thus, in the early deployment of 6G OTFS, it is necessary to explore transmission schemes that coordinate the TF and DD domains, to ensure compatibility with hybrid systems, i.e., the coexistence of OTFS and OFDM. Unlike the work in [12], [16], and [17], we propose a TF domain pilot-aided channel estimation scheme that does not require cluster pilots or BEM. In addition, unlike the work in [4] and [15], our channel estimation and equalization are performed in TF domain. Extrapolation vectors or matrices are carefully designed to improve estimation accuracy.

B. Contribution

In this paper, we propose a TF-domain Pilot-Aided Channel Estimation and Equalization (TF-PA-CEE) framework for OTFS systems, which is applicable to both OFDM-based OTFS and FBMC-based OTFS.[1] The implementation details of TF-PA-CEE are as follows: data mapping is performed in DD domain, while pilot insertion and reception processing are conducted in TF domain, thereby enabling the integration of OTFS channel estimation and equalization in TF domain. Our main contributions can be summarized as follows:

- To address the compatibility issues of channel estimation and equalization in hybrid systems, we propose a TF-PA-CEE framework. The core of the proposed scheme lies in constructing a channel extrapolation mechanism based on Bessel fitting or prior information assistance. The latter can provide higher estimation accuracy. Specifically, first, the channel dataset or the previous frame channel is used as prior information to calculate the channel correlation matrix. Second, according to the orthogonal projection theorem and the obtained channel correlation matrix, calculate the extrapolation matrix with a closed-form solution for the pilot positions. Finally, by weighted summation of the channel at each pilot position and the corresponding extrapolation matrix slices, the accurate estimation of the channel in the TF domain is achieved. Thereafter, TF domain equalization is performed to achieve time-varying channel distortion compensation.
- For the critical time interval problem of wide-sense stationary characteristics in time-varying channels, we derive a criterion for determining the wide-sense

[1]The detailed code can be downloaded at https://github.com/WangYeeng/TF-PA-CEE-for-FBMC-OTFS-master

stationary interval based on relative channel mean square error. First, based on the wide-sense stationary uncorrelated scattering assumption, we derive the autocorrelation function of multipath fading in the classical wireless channel model. Second, by introducing the assumptions of multipath asymptotics and short-term path delay invariance, we derive a closed-form expression for the channel frequency response correlation. Finally, by constraining the boundary conditions of the mean square error, we derive the theoretical upper bound for the wide-sense stationary time interval. The upper bound quantifies the effective period of the previous frame channel, providing a mathematical criterion for the update mechanism of prior information in channel estimation.

- We adopt the "Extended Vehicular A" and "TDL-A" channel models defined in the 3GPP 38.901 standard to verify the feasibility of the scheme. The results show that the proposed scheme can work robustly on doubly selective channels. On the other hand, with almost the same channel estimation accuracy, the reliability of FBMC-based OTFS is improved by about an order of magnitude compared to OFDM-based OTFS.

C. Outline and Notationn

The rest of this paper is organized as follows: Section II describes the OTFS system model based on FBMC and OFDM, including continuous and discrete time models, channel modeling, etc. Section III introduces the proposed TF-PA-CEE scheme. Section IV describes numerical results, including the analysis of the effective time interval for prior information, Normalized Mean Square Error (NMSE) and BER evaluation, as well as computational complexity analysis. The conclusion is drawn in Section V.

Notation: Bold uppercase and lowercase letters denote matrix and vector, respectively. $(\cdot)^{-1}$, $(\cdot)^{*}$, $(\cdot)^{T}$ and $(\cdot)^{H}$ denote the inverse, conjugate, transpose, and conjugate transpose operations, respectively. $\mathrm{diag}\{\mathbf{A}\}$ denotes the extraction of the diagonal elements of $\mathbf{A}$, while $\mathrm{diag}\{\mathbf{a}\}$ denotes the diagonal matrix with $\mathbf{a}$ on its diagonal. $\mathrm{vec}\{\cdot\}$ denotes the column vectorization operation for a matrix. $\mathbb{E}\{\cdot\}$ denotes expectation. $\mathbf{I}_N$ denotes the $N\times N$ unit matrix. The imaginary unit is denoted as $\mathrm{j}=\sqrt{-1}$. $\otimes$ denotes the Kronecker product. $\mathbb{C}$ and $\mathbb{R}$ denote the complex and real domains, respectively.

II. System Model

In the OTFS system, data symbols are placed on the DD grid. If the system has L subcarriers with a bandwidth of F and K time symbols with a duration of T, then the total bandwidth and total duration are LF and KT, respectively. The Quadrature Amplitude Modulation (QAM) data symbols, $d_{l',k'}\in\mathbb{C}$, $l'\in\mathbb{N}[0\ L-1]$, $k'\in\mathbb{N}[0\ K-1]$, are arranged on the DD lattice $\Lambda=\left\{\left(\frac{k'}{KT},\frac{l'}{LF}\right)\right\}$. Employing the ISFFT, the DD-domain symbol $d_{l',k'}$ is mapped to the TF lattice $\Lambda^{\perp}=\{(kT,lF)\}$, resulting in the TF-domain symbol $x_{l,k}\in\mathbb{C}$, $l\in\mathbb{N}[0\ L-1]$, $k\in\mathbb{N}[0\ K-1]$. $x_{l,k}$ can be calculated as

$$x_{l,k}=\frac{1}{\sqrt{LK}}\sum_{l'=0}^{L-1}\sum_{k'=0}^{K-1}d_{l',k'}\exp\left(\mathrm{j}2\pi\left(\frac{k'k}{K}-\frac{l'l}{L}\right)\right). \quad (1)$$

The symbol $x_{l,k}$ in the TF domain can be transmitted by various multi-carrier systems (e.g., OFDM and FBMC, etc.). Mathematically, the time-domain transmitted signal $s(t)$ of a multicarrier system can be expressed as [20], [21], and [22]

$$s(t)=\sum_{l=0}^{L-1}\sum_{k=0}^{K-1}x_{l,k}\underbrace{p(t-kT)\exp\left(\mathrm{j}2\pi lF(t-kT)\right)}_{g_{l,k}(t)}, \quad (2)$$

where $p(t)$ denotes the prototype pulse (i.e., the prototype filter), whose time-frequency shifted version $g_{l,k}(t)$ denotes the base pulse. The choice of the prototype pulse directly determines the properties of the multicarrier system. For example, rectangular pulse of duration T is employed in the OFDM system, which is the TF domain transmitted strategy for classical OTFS. Note that rectangular pulse is not localized in frequency, leading to OTFS with high OOB emission. To reduce OOB emission, we adopt the Hermite pulse, defined as [23]

$$\begin{aligned}&p(t)=\sqrt{F}\mathrm{e}^{-2\pi(tF)^2}\sum_{i=\{0,4,8,\cdots\}}c_i\mathrm{H}\left(i,2\sqrt{\pi}tF\right)\\&st.\quad\int_{-\infty}^{\infty}\mathrm{H}^2(i,t)\,\mathrm{e}^{-t^2}\mathrm{d}t=2^i\sqrt{\pi}i!\end{aligned}, \quad (3)$$

where the coefficients c_i can be found in [24]. Hermite pulse has the same shape in time and frequency, which allows us to exploit the symmetry to obtain good joint time-frequency localization. The Hermite pulse determines that the TF domain transmitted strategy of OTFS is replaced by FBMC. However, non-rectangular pulses lead to non-orthogonality [25], which reduces reliability. In addition, when QAM modulation (complex symbols) is adopted instead of Offset Quadrature Amplitude Modulation (OQAM) (interleaved transmission of real symbols), the FBMC fails to concentrate the interference in the imaginary domain [20]. Therefore, real orthogonality cannot be satisfied either. To maintain orthogonality in the complex domain, the TF lattice density needs to be reduced. Considering the scarcity of low-frequency resources [26], [27], we do not change the subcarrier bandwidth $F=\frac{1}{T}$ but instead expand the symbol duration to $2T$. This means that FBMC-based OTFS requires more time resources. However, FBMC does not employ CPs, and by optimizing the prototype pulse, the loss in spectral efficiency can be minimized.

The time-domain transmitted signal $s(t)$ passes through a time-varying multipath channel $h(t,\tau)$, and the received time-domain signal $r(t)$ is obtained, denoted as

$$r(t)=\int_{\mathbb{R}}h(t,\tau)\,s(t-\tau)\,\mathrm{d}\tau+\mathrm{n}(t), \quad (4)$$

where $\mathrm{n}(t)$ denotes noise. $h(t,\tau)$ is defined as [28] and [29]

$$h(t,\tau)=\frac{1}{\sqrt{N_p}}\sum_{i=1}^{N_p}a_i(t)\,\mathrm{e}^{\mathrm{j}(2\pi\nu_i t+\varphi_i)}\delta(\tau-\tau_i(t)), \quad (5)$$

where N_p denotes the number of multipaths. $a_i(t)$ and $\tau_i(t)$ denote the time-varying attenuation and time-varying delay of the ith th path, respectively. ν_i and φ_i denote the Doppler shift and random phase of the ith path, respectively. $\delta(\cdot)$ denotes

the Dirac delta function. Projecting the received signal $r(t)$ onto the base pulse $g_{l,k}(t)$, we obtain the TF domain received symbol $y_{l,k} \in \mathbb{C}$, denoted as

$$y_{l,k} = \int_{\mathbb{R}} r(t)\, g_{l,k}^{*}(t)\, \mathrm{d}t. \tag{6}$$

Note that Eq. (6) corresponds to matched filtering. Therefore, the receiver can optimize the received base pulses for better system performance [30]. For example, in CP-OFDM, the received base pulse removes the CP. In FBMC-based OTFS, we adopt the same base pulse at the transmitter and receiver. The received TF domain symbol $y_{l,k} \in \mathbb{C}$ can be converted to the DD domain by Symplectic Finite Fourier Transform (SFFT), denoted as

$$\bar{d}_{l',k'} = \frac{1}{\sqrt{LK}} \sum_{l=0}^{L-1} \sum_{k=0}^{K-1} y_{l,k} \exp\left(-\mathrm{j}2\pi \left(\frac{k'k}{K} - \frac{l'l}{L}\right)\right). \tag{7}$$

So far, we have only provided the downlink transmission process without including channel estimation and equalization.

It is worth noting that in practical applications, the signal is time discrete. The above continuous-time representation only provides analytical insights and cannot be physically implemented. Therefore, we study the discrete-time model and simplify the representation. According to Eq. (2), the main characteristic of a multicarrier system is the base pulse. We sample the base pulse at a rate of f_s and stack N_s sample values into a vector to obtain the discrete-time base pulse $\mathbf{g}_{l,k} \in \mathbb{C}^{N_s \times 1}$. In FBMC, the base pulse spans O symbol periods, where O is called the overlap factor. To ensure correct sampling, the center of the base pulse needs to be shifted by OT to align with the reference time point. Therefore, the time range for sampling is $\left[-\frac{OT}{2}, \frac{OT}{2} + (K-1)T\right)$. To represent the double integral in Eq. (2), we stack all the discrete basis pulses into matrix $\mathbf{G} = [\mathbf{g}_{0,0}, \cdots, \mathbf{g}_{L-1,K-1}] \in \mathbb{C}^{N_s \times LK}$ and stack all the TF domain symbols into vector $\mathbf{x} = [x_{0,0}, \cdots, x_{L-1,K-1}]^T \in \mathbb{C}^{LK \times 1}$. According to the rules of matrix algebra, the discrete transmitted signal $\mathbf{s} \in \mathbb{C}^{N_s \times 1}$ can be expressed as

$$\mathbf{s} = \mathbf{G}\mathbf{x}. \tag{8}$$

For digital communication systems, $h(t, \tau)$ is generally discretized into a tapped delay line model $h(n, m)$, denoted as

$$h(n, m) = \frac{1}{\sqrt{N_p}} \sum_{i=1}^{N_p} a_i(n)\, \mathrm{e}^{\mathrm{j}(2\pi \nu_i n / f_s + \varphi_i)} \delta(m - m_i), \tag{9}$$

where n is the discrete-time index, and m is the discrete delay index. $m_i = \lfloor \tau_i / f_s \rfloor$ is the discrete delay tap position of the ith path. $h(n, m)$ can be written in the banded time-varying convolution matrix $\boldsymbol{\mathcal{H}} \in \mathbb{C}^{N_s \times N_s}$, defined as [20] and [31]

$$[\boldsymbol{\mathcal{H}}]_{i,j} = h(i, i-j). \tag{10}$$

Thus, Eq. (6) can be expressed as

$$\mathbf{y} = \mathbf{G}^H \mathbf{r} = \mathbf{G}^H \boldsymbol{\mathcal{H}} \mathbf{G} \mathbf{x} + \mathbf{n}, \tag{11}$$

where $\mathbf{y} = [y_{0,0}, \cdots, y_{L-1,K-1}]^T \in \mathbb{C}^{LK \times 1}$ denotes the received TF domain symbol. $\mathbf{r} \in \mathbb{C}^{N_s \times 1}$ denotes the discrete received signal and $\mathbf{n} \sim \mathcal{CN}\left(0, P_n \mathbf{G}^H \mathbf{G}\right)$ the Gaussian noise with power P_n. To further simplify the representation, we let $\mathbf{H} = \mathbf{G}^H \boldsymbol{\mathcal{H}} \mathbf{G} \in \mathbb{C}^{LK \times LK}$. Thereby, Eq. (11) can be rewritten as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}. \tag{12}$$

Note that the orthogonality is described as $\mathbf{G}^H \mathbf{G} = \mathbf{I}_{LK}$ for both OFDM-based OTFS and FBMC-based OTFS. Applying ISFFT/SFFT to Eq. (12), we can realize OTFS transmission. In fact, the multicarrier modulation framework constructed in Eq. (12) is waveform independent. That is, when the prototype pulses satisfy the orthogonality condition, Eq. (12) can be generalized to characterize any multicarrier system. Thus, the design and implementation of OTFS are no longer limited to OFDM and can achieve better performance than classical OTFS. For example, FBMC-based OTFS can inherit the superior spectral properties of FBMC, see Fig. 1. Notably, OTFS based on the Zak transform (Zak-OTFS) exhibits a more compact time-frequency mapping property, which contributes to the suppression of OOB emissions. However, its spectral convergence performance is still lower than that of FBMC-based OTFS. On the other hand, the OOB improvement of Heisenberg OTFS is highly sensitive to the choice of window functions. When a rectangular window is employed, its power spectral density is identical to that of OFDM-based OTFS. However, when non-rectangular window functions are used, although spectral localization is improved, its TF domain transmission is not traditional rectangular window OFDM, but rather a multi-carrier systems evolved from OFDM [9]. In the Heisenburg OTFS shown in Fig. 1, we employed the sinc function with a Hann window.

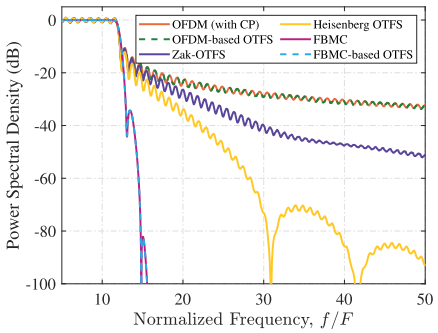

Fig. 1. Spectral properties for different systems. For FBMC, we consider the Hermite prototype filter with $O = 4$.

III. Channel Estimation and Equalization

Although the channel exhibits sparsity in DD domain, making it suitable for CS-based estimation [32], we chose to place the pilots and perform channel estimation and equalization in TF domain based on the feasibility of time or frequency domain estimation. For example, by reformulating OTFS as a precoded OFDM system, channel estimation can be performed in either time or frequency domain [15]. The well-known TF-domain estimation method targets Time-Varying Frequency Responses (TVFR) or channel gains on subcarriers. However, to address the compatibility issues in hybrid systems, we investigate and estimate the equivalent TF input-output relationship matrix $\mathbf{H} \in \mathbb{C}^{LK \times LK}$, see Eq.(12). The matrix

$\mathbf{H}$ can characterize all transmission properties. Specifically, the diagonal elements are approximately equal to TVFR, or equivalently the channel gain $h_{l,k}$ on individual subcarrier [33], whereas the off-diagonal elements are system-inherent or channel-induced interference.

A. Pilot Pattern and Mapping

According to the OFDM signal format in Long Term Evolution (LTE), we consider the diamond-shaped pilot pattern [34]. That is, the pilots are diagonally distributed in the time-frequency plane. This pilot pattern provides uniform coverage in the TF domain, which can reduce extrapolation errors, making it suitable for high mobility and high-frequency selective environments.

Assume that $|\mathcal{P}|$ and $|\mathcal{D}|$ denote the number of pilots and data, respectively, and that $|\mathcal{P}| + |\mathcal{D}| = LK$. Then $\mathbf{x}_{\mathcal{P}} \in \mathbb{C}^{|\mathcal{P}|\times 1}$ and $\mathbf{d}_{\mathcal{D}} \in \mathbb{C}^{|\mathcal{D}|\times 1}$ denote the pilot sequence of TF domain and the data symbols in the DD domain, respectively. Note that to estimate the channel information in the time-frequency domain, the pilot sequence cannot be placed in the DD domain. The reason is that the ISFFT will destroy the prior information in the pilots that characterizes the TF-domain channel. To convert $\mathbf{d}_{\mathcal{D}} \in \mathbb{C}^{|\mathcal{D}|\times 1}$ to the TF domain, we first arrange the data symbols in the two-dimensional DD domain $\Lambda = \left\{\left(\frac{k'}{(K-K_{\mathcal{P}})T}, \frac{l'}{(L-L_{\mathcal{P}})F}\right)\right\}$, denoted as

$$f : \{\mathbf{d}_{\mathcal{D}}, |\mathcal{D}| \times 1\} \to f : \{\mathbf{D}, L - L_{\mathcal{P}} \times K - K_{\mathcal{P}}\}, \quad (13)$$

where $f : \{\cdot\}$ denotes the matrix reshaping function. $L_{\mathcal{P}}$ and $K_{\mathcal{P}}$ are the number of subcarriers and symbols occupied by pilots, respectively. Then, by performing ISFFT on $\mathbf{D} \in \mathbb{C}^{L-L_{\mathcal{P}}\times K-K_{\mathcal{P}}}$, we can obtain the data symbols $\mathbf{x}_{\mathcal{D}} \in \mathbb{C}^{|\mathcal{D}|\times 1}$ in the TF domain, denoted as

$$\mathbf{x}_{\mathcal{D}} = \mathrm{vec}\left\{\mathrm{ISFFT}\left\{\mathbf{D}\right\}\right\}. \quad (14)$$

According to the pilot pattern, we can construct the pilot mapping matrix $\mathbf{C} \in \mathbb{R}^{LK\times LK}$. This matrix is not only used to correctly map the pilot and data symbols in the TF domain to their corresponding positions, but also to determine the corresponding basis pulse, thereby optimizing the signal processing process. $\mathbf{C} \in \mathbb{R}^{LK\times LK}$ can be calculated as

$$\mathbf{C}_{i,j} = \rho\mathbf{I}_{|\mathcal{P}|}, \quad \forall i \in \{i \,|\mathbf{\Pi}_i = 1\}, \; j \in [1, |\mathcal{P}|], \quad (15)$$

$$\mathbf{C}_{i,j} = \mathbf{I}_{|\mathcal{D}|}, \quad \forall i \in \{i \,|\mathbf{\Pi}_i = 0\}, \; j \in [|\mathcal{P}| + 1, |\mathcal{P}| + |D|], \quad (16)$$

where ρ is the power offset factor from pilot to data. $\mathbf{\Pi} \in \mathbb{R}^{LK\times 1}$ denotes the vectorized pilot pattern. To control the energy distribution of the pilot and data symbols, we normalize $\mathbf{C} \in \mathbb{R}^{LK\times LK}$, which is expressed as

$$\mathbf{C} = \frac{\mathbf{C}}{\sqrt{\frac{1}{LK}\sum_{i=1}^{LK}\left[\mathbf{C}\mathbf{C}^T\right]_{i,i}}}. \quad (17)$$

Thereby, the time-frequency lattice mapping of pilot and data symbols can be expressed as

$$\mathbf{x} = \mathbf{C}\left[\mathbf{x}_{\mathcal{P}}, \mathbf{x}_{\mathcal{D}}\right]^T, \quad (18)$$

where $\mathbf{x} \in \mathbb{C}^{LK\times 1}$ is the transmitted symbol that contains both data and pilots. The visualization mapping process for pilots and data symbols is shown in Fig. 2.

B. Pilot-Aided TF-Domain Channel Estimation

The Channel of Pilot Position (CoPP) can be obtained by minimizing the L2 norm, denoted as

$$\hat{h}_{l_{\mathcal{P}},k_{\mathcal{P}}} = \arg\min_{h_{l_{\mathcal{P}},k_{\mathcal{P}}}} \|y_{l_{\mathcal{P}},k_{\mathcal{P}}} - h_{l_{\mathcal{P}},k_{\mathcal{P}}} x_{l_{\mathcal{P}},k_{\mathcal{P}}}\|_2^2. \quad (19)$$

We stack all the estimated CoPPs in vector $\hat{\mathbf{h}}_{\mathcal{P}} = [\hat{h}_{l_1,k_1}, \cdots, \hat{h}_{l_{|\mathcal{P}|},k_{|\mathcal{P}|}}]^T \in \mathbb{C}^{|\mathcal{P}|\times 1}$ and obtain the Channel of Data Position (CoDP) by extrapolation, denoted as

$$\hat{h}_{l,k} = \boldsymbol{\xi}_{l,k}^H \hat{\mathbf{h}}_{\mathcal{P}}, \quad (20)$$

where $\boldsymbol{\xi}_{l,k} \in \mathbb{C}^{|\mathcal{P}|\times 1}$ denotes the extrapolation vector at the time-frequency position (l, k). $\boldsymbol{\xi}_{l,k} \in \mathbb{C}^{|\mathcal{P}|\times 1}$ depends on the extrapolation method. For example, linear extrapolation performs well in flat channel estimation. However, in doubly selective channels, the time-frequency two-dimensional selective fading leads to off-diagonal elements carrying significant interference components that cannot be ignored. This requires estimating the global matrix, rather than the simplified diagonalized vector [20]. Using CoPP, we can estimate $\mathbf{H}$ as

$$\hat{\mathbf{H}} = \sum_{i=1}^{|\mathcal{P}|} \mathbf{\Xi}_i \hat{h}_{l_i,k_i}, \; st. \; [\mathbf{\Xi}_i]_{\ell 1,\ell 2} = \left[\boldsymbol{\xi}_{\ell 1,\ell 2}^H\right]_i, \quad (21)$$

where $\mathbf{\Xi} \in \mathbb{C}^{LK\times LK\times|\mathcal{P}|}$ denotes the extrapolation matrix. $\ell 1 = l_1 + k_1 L \in [1, LK]$ and $\ell 2 = l_2 + k_2 L \in [1, LK]$. When $l = l_1 = l_2$ and $k = k_1 = k_2$, $\boldsymbol{\xi}_{\ell 1,\ell 2} = \boldsymbol{\xi}_{l,k}$. To obtain the optimal channel estimate, we consider the orthogonal projection theorem [35], which states that the estimation error must be orthogonal to the estimated value. According to the orthogonal projection theorem, we can obtain

$$\mathbb{E}\left\{\left(\mathbf{H} - \hat{\mathbf{H}}\right)\hat{\mathbf{H}}^H\right\} = \mathbf{0}. \quad (22)$$

That is

$$\mathbb{E}\left\{\mathbf{H}\hat{\mathbf{H}}^H\right\} = \mathbb{E}\left\{\hat{\mathbf{H}}\hat{\mathbf{H}}^H\right\}. \quad (23)$$

Substituting $\hat{\mathbf{H}} = \sum_{i=1}^{|\mathcal{P}|} \mathbf{\Xi}_i \hat{h}_{l_i,k_i}$ into Eq. (23), we obtain

$$\begin{aligned}\mathbb{E}\left\{\mathbf{H}\hat{\mathbf{H}}^H\right\} &= \mathbb{E}\left\{\mathbf{H}\left(\sum_{i=1}^{|\mathcal{P}|}\mathbf{\Xi}_i\hat{h}_{l_i,k_i}\right)^H\right\} \\ &= \sum_{i=1}^{|\mathcal{P}|}\mathbb{E}\left\{\mathbf{H}\hat{h}_{l_i,k_i}^H\right\}\mathbf{\Xi}_i^H \\ &= \sum_{i=1}^{|\mathcal{P}|}\mathbf{R}_{\mathbf{H},[\hat{\mathbf{h}}_{\mathcal{P}}]_i}\mathbf{\Xi}_i^H, \end{aligned} \quad (24)$$

and

$$\begin{aligned}\mathbb{E}\left\{\hat{\mathbf{H}}\hat{\mathbf{H}}^H\right\} &= \mathbb{E}\left\{\left(\sum_{i=1}^{|\mathcal{P}|}\mathbf{\Xi}_i\hat{h}_{l_i,k_i}\right)\left(\sum_{j=1}^{|\mathcal{P}|}\mathbf{\Xi}_j\hat{h}_{l_j,k_j}\right)^H\right\} \\ &= \sum_{i=1}^{|\mathcal{P}|}\sum_{j=1}^{|\mathcal{P}|}\mathbf{\Xi}_i\mathbb{E}\left\{\hat{h}_{l_i,k_i}\hat{h}_{l_j,k_j}^H\right\}\mathbf{\Xi}_j^H \\ &= \sum_{i=1}^{|\mathcal{P}|}\sum_{j=1}^{|\mathcal{P}|}\mathbf{\Xi}_i\mathbf{R}_{[\hat{\mathbf{h}}_{\mathcal{P}}]_{i,j}}\mathbf{\Xi}_j^H, \end{aligned} \quad (25)$$

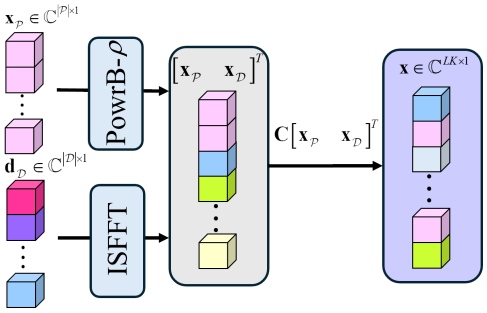

Fig. 2. Pilots and data symbols mapping process. PowerB-ρ is a power balancer, which can reduce the power difference.

where $\mathbf{R}_{\mathbf{H},[\hat{\mathbf{h}}_{\mathcal{P}}]_i} \triangleq \mathbb{E}\left\{\mathbf{H}\hat{h}_{l_i,k_i}^H\right\}$ is the cross-correlation matrix between $\mathbf{H}$ and the ith estimated component, while $\mathbf{R}_{[\hat{\mathbf{h}}_{\mathcal{P}}]_{i,j}} \triangleq \mathbb{E}\left\{\hat{h}_{l_i,k_i}\hat{h}_{l_j,k_j}^H\right\}$ is the cross-correlation between the ith and jth estimated components (which becomes the autocorrelation when $i=j$). Therefore, the orthogonality condition between the estimation error and the estimated value can be rewritten as

$$\sum_{i=1}^{|\mathcal{P}|}\mathbf{R}_{\mathbf{H},[\hat{\mathbf{h}}_{\mathcal{P}}]_i}\mathbf{\Xi}_i^H = \sum_{i=1}^{|\mathcal{P}|}\sum_{j=1}^{|\mathcal{P}|}\mathbf{\Xi}_i\mathbf{R}_{[\hat{\mathbf{h}}_{\mathcal{P}}]_{i,j}}\mathbf{\Xi}_j^H. \tag{26}$$

For computational convenience, we reshape $\mathbf{\Xi} \in \mathbb{C}^{LK\times LK\times|\mathcal{P}|}$ into $\tilde{\mathbf{\Xi}} \in \mathbb{C}^{L^2K^2\times|\mathcal{P}|}$. Thus, Eq. (26) can be rewritten in a compact form, i.e.

$$\mathbf{R}_{\mathbf{H},\hat{\mathbf{h}}_{\mathcal{P}}}\tilde{\mathbf{\Xi}}^H = \tilde{\mathbf{\Xi}}\mathbf{R}_{\hat{\mathbf{h}}_{\mathcal{P}}}\tilde{\mathbf{\Xi}}^H. \tag{27}$$

According to Eq. (27), $\tilde{\mathbf{\Xi}} \in \mathbb{C}^{L^2K^2\times|\mathcal{P}|}$ can be calculated as

$$\tilde{\mathbf{\Xi}} = \mathbf{R}_{\mathbf{H},\hat{\mathbf{h}}_{\mathcal{P}}}\left(\mathbf{R}_{\hat{\mathbf{h}}_{\mathcal{P}}}\right)^{-1}, \tag{28}$$

where $\mathbf{R}_{\mathbf{H},\hat{\mathbf{h}}_{\mathcal{P}}} \in \mathbb{C}^{L^2K^2\times|\mathcal{P}|}$ and $\mathbf{R}_{\hat{\mathbf{h}}_{\mathcal{P}}} \in \mathbb{C}^{|\mathcal{P}|\times|\mathcal{P}|}$. According to Eq. (24) and Eq. (25), the calculation of the extrapolation vector or matrix requires the utilization of channel correlation information. We provide two methods for calculating channel correlation information.

1) Bessel Fitting-Based Channel Correlation Matrix: We assume that the delay τ and Doppler shift ν are statistically independent. The time-frequency correlation function for the Jakes model can be defined as [36]

$$r_{Ht}(n_1-n_2) = J_0\left(2\pi\nu_{\max}T\frac{n_1-n_2}{N_s}\right), \tag{29}$$

$$r_{Hf}(l_1-l_2) = \mathrm{sinc}\left(\pi\tau_{\max}F(l_1-l_2)\right), \tag{30}$$

where $J_0(\cdot)$ denotes the zero-order Bessel function. $\nu_{\max}$ and $\tau_{\max}$ denote the maximum delay and maximum Doppler shift, respectively. Note that Eq. (30) provides the frequency correlation of $h_{l,k}$, while the time correlation can be calculated as

$$r_{\bar{H}t}(k_1-k_2) = \frac{1}{N_s^2}\sum_{n1=1}^{N_s}\sum_{n2=1}^{N_s} r_{Ht}(n_1-n_2+(k_1-k_2)N_s). \tag{31}$$

Thereby, the time-frequency correlation matrix $\mathbf{R}_{\mathbf{h}} \in \mathbb{R}^{LK\times LK}$ of the vectorized channel $\mathbf{h} \in \mathbb{C}^{LK\times 1}$ can be calculated as

$$\mathbf{R}_{\mathbf{h}} = \begin{bmatrix} r_{\bar{H}t}(0) & \cdots & r_{\bar{H}t}(1-K) \\ \vdots & \ddots & \vdots \\ r_{\bar{H}t}(K-1) & \cdots & r_{\bar{H}t}(0) \end{bmatrix} \otimes \begin{bmatrix} r_{Hf}(0) & \cdots & r_{Hf}(1-L) \\ \vdots & \ddots & \vdots \\ r_{Hf}(L-1) & \cdots & r_{Hf}(0) \end{bmatrix} \tag{32}$$

For the one-tap channel vector $\mathbf{h} \in \mathbb{C}^{LK\times 1}$ in the TF domain, the correlation matrices $\mathbf{R}_{\mathbf{h},\mathbf{h}_{\mathcal{P}}} \in \mathbb{R}^{LK\times|\mathcal{P}|}$ and $\mathbf{R}_{\mathbf{h}_{\mathcal{P}}} \in \mathbb{R}^{|\mathcal{P}|\times|\mathcal{P}|}$ can be obtained based on the appropriate elements of $\mathbf{R}_{\mathbf{h}} \in \mathbb{R}^{LK\times LK}$, primarily relying on the vectorized pilot index. Thereby, the extrapolation vector $\boldsymbol{\xi}_{l,k} \in \mathbb{C}^{|\mathcal{P}|\times 1}$ can be calculated as

$$\boldsymbol{\xi}_{l,k} = \left[\mathbf{R}_{\mathbf{h},\mathbf{h}_{\mathcal{P}}}\left(\mathbf{R}_{\mathbf{h}_{\mathcal{P}}}\right)^{-1}\right]_{:,\ell}^H, \tag{33}$$

where $\ell = l + kL \in [1, LK]$. It is worth noting that the above conclusions only apply to estimating one-tap channels and cannot estimate channel-induced interference.

2) A Priori Information-Assisted Channel Correlation Matrix: To estimate the channel-induced interference, we can utilize the channel dataset as prior information to calculate the correlation matrix. For example, the dataset used to train deep learning models [37]. In addition, by transmitting only one symbol in the first frame (i.e., $x=1$), we can obtain the corresponding time-varying convolution matrix $\boldsymbol{\mathcal{H}}$. Then, by using $\boldsymbol{\mathcal{H}}$ as prior information, we can estimate the channel state for each frame. The effectiveness of the prior information can be determined according to *Lemma* 1.

Lemma 1: In the channel model of Eq. (5), assuming that the upper bound of the mean square error for the multi-carrier system channel estimation is ζ, then the effective time interval between the symbol frame carrying the prior information $\boldsymbol{\mathcal{H}}$ and the symbol frame of the channel to be estimated is determined by the critical wide-sense stationary time interval $\mathcal{T}_{\max}$, defined as

$$\mathcal{T}_{\max} = \frac{J_0^{-1}\left(\sqrt{(1-\zeta/2)/P_h}\right)}{2\pi\nu_{\max}}, \tag{34}$$

where $J_0^{-1}(\cdot)$ is the inverse function of the zero-order Bessel function and P_h the channel power. *Proof:* The proof is provided in Appendix.

Note that under the strict constraints of *Lemma* 1, the desired channel correlation matrix can be directly obtained. In general, it is desirable for the constraints in *Lemma* 1 to be less stringent. That is, under the less stringent constraints of *Lemma* 1, $\mathcal{T}_{\max}$ can be sufficiently large. However, in doubly selective channels, $\boldsymbol{\mathcal{H}}$ cannot satisfy the less stringent constraints. Fortunately, the second-order statistical characteristics of fast time-varying channels vary slowly [38]. Thus, we can calculate the desired channel correlation matrix with the second-order statistic of $\boldsymbol{\mathcal{H}}$. Thus, the priori information does not need to be updated frequently. Note that prior information should be updated before the estimator becomes strictly

invalid. The strict invalid of an estimator can be assessed from two perspectives: the physical performance boundary and the theoretical limit.

1) From the physical performance perspective, when ζ exceeds the channel power, the estimation results are completely dominated by noise or errors, and the estimator becomes strictly invalid.
2) From the theoretical limit perspective, when the Cramér-Rao Lower Bound (CRLB) diverges due to extremely low SNR or model identifiability issues (for power-normalized channels, CRLB $= \frac{1}{\text{SNR}}$), parameter estimation is not theoretically realizable [39].

Assuming that $\boldsymbol{\mathcal{H}}$ satisfies the less stringent constraints of *Lemma* 1, then $\mathbf{R}_{\mathbf{H},[\mathbf{h}_\mathcal{P}]_i} \in \mathbb{C}^{L^2K^2\times 1}$ can be calculated as

$$\begin{aligned} &f:\left\{\mathbf{R}_{\boldsymbol{\mathcal{H}}}\left(\mathbf{g}_{\ell_{\mathcal{P}i}}^T\otimes\mathbf{g}_{\ell_{\mathcal{P}i}}^H\right)^H, N_s^2\times 1\right\}\\ &\rightarrow f:\left\{\mathbf{R}_{\boldsymbol{\mathcal{H}},[\mathbf{h}_\mathcal{P}]_i}, N_s\times N_s\right\}, \end{aligned} \tag{35}$$

$$\begin{aligned} &f:\left\{\mathbf{G}^H\mathbf{R}_{\boldsymbol{\mathcal{H}},[\mathbf{h}_\mathcal{P}]_i}\mathbf{G}, LK\times LK\right\}\\ &\rightarrow f:\left\{\mathbf{R}_{\mathbf{H},[\mathbf{h}_\mathcal{P}]_i}, L^2K^2\times 1\right\}, \end{aligned} \tag{36}$$

where $\mathbf{R}_{\boldsymbol{\mathcal{H}}} = \mathbb{E}\left\{\text{vec}\{\boldsymbol{\mathcal{H}}\}\,\text{vec}\{\boldsymbol{\mathcal{H}}\}^H\right\}$ is the second-order statistic of the time-varying convolution matrix. $\ell_\mathcal{P}$ is the vectorized pilot index. Considering all pilot indices, we can obtain the $\mathbf{R}_{\mathbf{H},\mathbf{h}_\mathcal{P}} \in \mathbb{C}^{L^2K^2\times|\mathcal{P}|}$. Notably, unlike in $\mathbf{R}_{\mathbf{H},\mathbf{h}_\mathcal{P}} \in \mathbb{C}^{L^2K^2\times|\mathcal{P}|}$, the autocorrelation components in $\mathbf{R}_{\mathbf{h}_\mathcal{P}} \in \mathbb{C}^{|\mathcal{P}|\times|\mathcal{P}|}$ can utilize prior information, such as noise power, to enhance estimation accuracy. $\mathbf{R}_{\mathbf{h}_\mathcal{P}} \in \mathbb{C}^{|\mathcal{P}|\times|\mathcal{P}|}$ is defined as

$$\mathbf{R}_{\mathbf{h}_\mathcal{P}} = \mathbb{E}\left\{\mathbf{h}_\mathcal{P}\mathbf{h}_\mathcal{P}^H\right\}. \tag{37}$$

When $\ell_{\mathcal{P}i}\neq\ell_{\mathcal{P}j}$, $[\mathbf{R}_{\mathbf{h}_\mathcal{P}}]_{\ell_{\mathcal{P}i},\ell_{\mathcal{P}j}}$ can be calculated as

$$\begin{aligned} [\mathbf{R}_{\mathbf{h}_\mathcal{P}}]_{\ell_{\mathcal{P}i},\ell_{\mathcal{P}j}} &= \mathbf{g}_{\ell_{\mathcal{P}i}}^H\mathbf{H}\mathbf{g}_{\ell_{\mathcal{P}i}}\left(\mathbf{g}_{\ell_{\mathcal{P}j}}^H\mathbf{H}\mathbf{g}_{\ell_{\mathcal{P}j}}\right)^H\\ &= \left(\mathbf{g}_{\ell_{\mathcal{P}i}}^T\otimes\mathbf{g}_{\ell_{\mathcal{P}i}}^H\right)\mathbf{R}_{\boldsymbol{\mathcal{H}}}\left(\mathbf{g}_{\ell_{\mathcal{P}j}}^T\otimes\mathbf{g}_{\ell_{\mathcal{P}j}}^H\right)^H. \end{aligned} \tag{38}$$

When $\ell_{\mathcal{P}i}=\ell_{\mathcal{P}j}$, $[\mathbf{R}_{\mathbf{h}_\mathcal{P}}]_{\ell_{\mathcal{P}i},\ell_{\mathcal{P}j}}$ can be calculated as

$$\begin{aligned} &[\mathbf{R}_{\mathbf{h}_\mathcal{P}}]_{\ell_{\mathcal{P}i},\ell_{\mathcal{P}i}}\\ &= \text{Tr}\left\{\left((\mathbf{GC})^T\otimes\mathbf{g}_{\ell_{\mathcal{P}i}}^H\right)\mathbf{R}_{\boldsymbol{\mathcal{H}}}\left((\mathbf{GC})^T\otimes\mathbf{g}_{\ell_{\mathcal{P}i}}^H\right)^H\right\}\\ &\quad+\frac{P_\text{n}}{\rho}\mathbf{g}_{\ell_{\mathcal{P}i}}^H\mathbf{g}_{\ell_{\mathcal{P}i}}, \end{aligned} \tag{39}$$

where $\mathbf{g}_{\ell_{\mathcal{P}i}} = [\mathbf{G}]_{:,\ell_{\mathcal{P}i}}$. Eq. (38) and Eq. (39) denote the cross-correlation and autocorrelation of $\mathbf{h}_\mathcal{P}$, respectively, which together construct $\mathbf{R}_{\mathbf{h}_\mathcal{P}} \in \mathbb{C}^{|\mathcal{P}|\times|\mathcal{P}|}$. Thereby, the extrapolation matrix $\boldsymbol{\Xi} \in \mathbb{C}^{LK\times LK\times|\mathcal{P}|}$ can be calculated as

$$\begin{aligned} &f:\left\{\mathbf{R}_{\mathbf{H},\mathbf{h}_\mathcal{P}}(\mathbf{R}_{\mathbf{h}_\mathcal{P}})^{-1}, L^2K^2\times|\mathcal{P}|\right\}\\ &\rightarrow f:\{\boldsymbol{\Xi}, LK\times LK\times|\mathcal{P}|\}. \end{aligned} \tag{40}$$

By combining the extrapolation matrix and the mathematical framework of Eq. (21), we can derive the global estimate of the TF-domain channel matrix $\mathbf{H}$. In addition, relying on the diagonal dominance of time-frequency doubly selective channels, the one-tap channels corresponding to each time-frequency lattice can be extracted directly from the main diagonal elements of the $\mathbf{H}$. For ease of understanding, the steps of the Priori Information-Assisted (PIA) cross-domain channel estimation are summarized in **Algorithm** 1. Note that although Stage 2 of **Algorithm** 1 performs computations using the compact form in Eq. (21), the underlying operations remain unchanged.

Algorithm 1 The PIA TF-Domain Channel Estimator

Input: $\mathbf{R}_{\boldsymbol{\mathcal{H}}}$, $\hat{\mathbf{h}}_\mathcal{P}$, $\mathbf{C}$, $\mathbf{G}$ and pilot index $\ell_\mathcal{P}$.
Output: Estimated channel $\hat{\mathbf{H}}$.

 # Stage 1: Obtaining the extrapolation matrix.
for $i=1,\cdots,|\mathcal{P}|$ **do**
  $\mathbf{g}_{\ell_{\mathcal{P}i}} \leftarrow [\mathbf{G}]_{:,\ell_{\mathcal{P}i}}$;
  $f:\left\{\mathbf{R}_{\boldsymbol{\mathcal{H}},[\mathbf{h}_\mathcal{P}]_i}, N_s\times N_s\right\}\leftarrow$
  $f:\left\{\mathbf{R}_{\boldsymbol{\mathcal{H}}}\left(\mathbf{g}_{\ell_{\mathcal{P}i}}^T\otimes\mathbf{g}_{\ell_{\mathcal{P}i}}^H\right)^H, N_s^2\times 1\right\}$;
  $f:\left\{\mathbf{R}_{\mathbf{H},[\mathbf{h}_\mathcal{P}]_i}, L^2K^2\times 1\right\}\leftarrow$
  $f:\left\{\mathbf{G}^H\mathbf{R}_{\boldsymbol{\mathcal{H}},[\mathbf{h}_\mathcal{P}]_i}\mathbf{G}, LK\times LK\right\}$.
  for $j=1,\cdots,|\mathcal{P}|$ **do**
    if $\ell_{\mathcal{P}i}\neq\ell_{\mathcal{P}j}$ **then**
      $\mathbf{g}_{\ell_{\mathcal{P}j}} \leftarrow [\mathbf{G}]_{:,\ell_{\mathcal{P}j}}$;
      $[\mathbf{R}_{\mathbf{h}_\mathcal{P}}]_{\ell_{\mathcal{P}i},\ell_{\mathcal{P}j}}\leftarrow$
      $\left(\mathbf{g}_{\ell_{\mathcal{P}i}}^T\otimes\mathbf{g}_{\ell_{\mathcal{P}i}}^H\right)\mathbf{R}_{\boldsymbol{\mathcal{H}}}\left(\mathbf{g}_{\ell_{\mathcal{P}j}}^T\otimes\mathbf{g}_{\ell_{\mathcal{P}j}}^H\right)^H$.
    else
      $[\mathbf{R}_{\mathbf{h}_\mathcal{P}}]_{\ell_{\mathcal{P}i},\ell_{\mathcal{P}i}}\leftarrow$
      $\text{tr}\left\{\left((\mathbf{GC})^T\otimes\mathbf{g}_{\ell_{\mathcal{P}i}}^H\right)\mathbf{R}_{\boldsymbol{\mathcal{H}}}\left((\mathbf{GC})^T\otimes\mathbf{g}_{\ell_{\mathcal{P}i}}^H\right)^H\right\}$
      $+\frac{P_\text{n}}{\rho}\mathbf{g}_{\ell_{\mathcal{P}i}}^H\mathbf{g}_{\ell_{\mathcal{P}i}}$.
    end
  end
end
 # Stage 2 Channel Estimation
$\tilde{\boldsymbol{\Xi}}\leftarrow\mathbf{R}_{\mathbf{H},\mathbf{h}_\mathcal{P}}(\mathbf{R}_{\mathbf{h}_\mathcal{P}})^{-1}$,
 # Compact form of Eq. (21).
$f:\left\{\hat{\mathbf{H}}, LK\times LK\right\}\leftarrow f:\left\{\tilde{\boldsymbol{\Xi}}\hat{\mathbf{h}}_\mathcal{P}, L^2K^2\times 1\right\}$.
return $\hat{\mathbf{H}}$.

C. Theoretical Performance Analysis

To qualitatively analyze the performance differences of channel estimation in different domains, we examine the Mean Squared Error (MSE) of channel estimation in both TF domain and DD domain. For notational simplicity, we set $\mathcal{L}=L^2K^2$ in the observation model, such that the TF domain channel is denoted as $\mathbf{h}_{TF}\in\mathbb{C}^\mathcal{L}$ and the DD domain channel as $\mathbf{h}_{DD}\in\mathbb{C}^\mathcal{L}$. Moreover, pilot symbols of the same type are employed in both TF domain and DD domain to ensure consistency in the comparison.

1) TF-Domain Estimation: The CoPP observation model (including noise) derived from Eq. (19) can be expressed as:

$$\hat{\mathbf{h}}_\mathcal{P} = \mathbf{S}\mathbf{h}_{TF}+\mathbf{n}_\mathcal{P},\quad \mathbf{n}_\mathcal{P}\sim\mathcal{CN}\left(0,P_\text{n}\mathbf{I}_\mathcal{P}\right), \tag{41}$$

where $\mathbf{S}\in\{0,1\}^{|\mathcal{P}|\times\mathcal{L}}$ is the pilot selection matrix derived from the pilot mapping matrix $\mathbf{C}$. Thus, $\mathbf{R}_{\mathbf{H},\mathbf{h}_\mathcal{P}}$ and $\mathbf{R}_{\mathbf{h}_\mathcal{P}}$ can

be simply calculated as

$$\mathbf{R}_{\mathbf{H},\mathbf{h}_{\mathcal{P}}} \approx \mathbb{E}\left\{\mathbf{h}_{TF}\mathbf{h}_{\mathcal{P}}^{H}\right\} = \mathbf{R}_{\mathbf{h}_{TF}}\mathbf{S}^{H}, \tag{42}$$

$$\mathbf{R}_{\mathbf{h}_{\mathcal{P}}} \approx \mathbb{E}\left\{\mathbf{h}_{\mathcal{P}}\mathbf{h}_{\mathcal{P}}^{H}\right\} = \mathbf{S}\mathbf{R}_{\mathbf{h}_{TF}}\mathbf{S}^{H} + P_{\mathrm{n}}\mathbf{I}_{\mathcal{P}}. \tag{43}$$

Based on the compact form in Eq. (21), the estimated value of $\mathbf{h}_{TF} \in \mathbb{C}^{\mathcal{L}}$ can be written as

$$\hat{\mathbf{h}}_{TF} = \mathbf{R}_{\mathbf{h}_{TF}}\mathbf{S}^{H}\left(\mathbf{S}\mathbf{R}_{\mathbf{h}_{TF}}\mathbf{S}^{H} + P_{\mathrm{n}}\mathbf{I}_{\mathcal{P}}\right)^{-1}\hat{\mathbf{h}}_{\mathcal{P}}. \tag{44}$$

The corresponding MSE is

$$\begin{aligned}\bar{e}_{\mathrm{MSE}}^{TF} &= \mathbb{E}\left\{\left\|\mathbf{h}_{TF} - \hat{\mathbf{h}}_{TF}\right\|_2^2\right\} \\ &= \mathrm{tr}\left\{\mathbf{R}_{\mathbf{h}_{TF}} - \mathbf{R}_{\mathbf{h}_{TF}}\mathbf{S}^{H}\left(\mathbf{S}\mathbf{R}_{\mathbf{h}_{TF}}\mathbf{S}^{H} + P_{\mathrm{n}}\mathbf{I}_{\mathcal{P}}\right)^{-1}\mathbf{S}\mathbf{R}_{\mathbf{h}_{TF}}\right\}.\end{aligned} \tag{45}$$

Given $\mathbf{R}_{\mathbf{h}}$, Eq. (45) achieves the optimal MSE.

2) DD-Domain Estimation: For a fair comparison, we conduct the analysis under the same observation model. If there exists a unitary transformation $\mathbf{U} \in \mathbb{C}^{\mathcal{L}\times\mathcal{L}}$ (e.g., the discrete matrix corresponding to the SFFT/ISFFT) that maps the TF domain channel to the DD domain, then $\mathbf{h}_{DD} \in \mathbb{C}^{\mathcal{L}}$ can be expressed as:

$$\mathbf{h}_{DD} = \mathbf{U}\mathbf{h}_{TF}. \tag{46}$$

The observation model in DD domain is

$$\hat{\mathbf{h}}_{\mathcal{P}} = \mathbf{S}\mathbf{U}^{H}\mathbf{h}_{DD} + \mathbf{n}_{\mathcal{P}} = \mathbf{D}\mathbf{h}_{DD} + \mathbf{n}_{\mathcal{P}}. \tag{47}$$

If the same estimation scheme as in Eq. (21) is adopted, the estimated channel $\hat{\mathbf{h}}_{DD}$ in DD domain can be written as

$$\hat{\mathbf{h}}_{DD} = \mathbf{R}_{\mathbf{h}_{DD}}\mathbf{D}^{H}\left(\mathbf{D}\mathbf{R}_{\mathbf{h}_{DD}}\mathbf{D}^{H} + P_{\mathrm{n}}\mathbf{I}_{\mathcal{P}}\right)^{-1}\hat{\mathbf{h}}_{\mathcal{P}}, \tag{48}$$

where $\mathbf{R}_{\mathbf{h}_{DD}} = \mathbb{E}\left\{\mathbf{U}\mathbf{h}_{TF}(\mathbf{U}\mathbf{h}_{TF})^{H}\right\} = \mathbf{U}\mathbf{R}_{\mathbf{h}_{TF}}\mathbf{U}^{H}$. We transform $\hat{\mathbf{h}}_{DD}$ back to the TF domain:

$$\hat{\mathbf{h}}_{DD\to TF} = \mathbf{U}^{H}\hat{\mathbf{h}}_{DD}. \tag{49}$$

According to Eq. (48), $\mathbf{R}_{\mathbf{h}_{DD}} = \mathbf{U}\mathbf{R}_{\mathbf{h}_{TF}}\mathbf{U}^{H}$ and $\mathbf{U}\mathbf{U}^{H} = \mathbf{I}$, we can rigorously prove that $\hat{\mathbf{h}}_{DD\to TF} = \hat{\mathbf{h}}_{TF}$, see Eq. (50), as shown at the bottom of the page.

Thus, the two are exactly equivalent, and the corresponding MSEs are also the same. This indicates that, given the same second-order prior $\mathbf{R}_{\mathbf{h}}$ or employing the same channel estimation scheme, the TF-domain and DD-domain estimations are theoretically equivalent.

3) Possible Differences: Many existing DD-domain estimation schemes exploit the sparsity of $\mathbf{h}_{DD}$ (i.e., $\mathbf{h}_{DD} \approx \mathbf{h}_s$) and perform sparse recovery using a predefined dictionary. Thus, the actual observation model becomes:

$$\mathcal{Y} = \mathbf{D}\mathbf{h}_{DD} + \mathbf{n}_{\mathcal{P}} = \mathbf{D}\left(\mathbf{h}_s + \mathbf{h}_r\right) + \mathbf{n}_{\mathcal{P}} \approx \mathbf{D}\mathbf{h}_s + \mathbf{n}_{\mathcal{P}}, \tag{51}$$

where $\mathbf{h}_r$ is the residual information that is not captured under the sparsity assumption. If a sparse reconstruction algorithm (e.g., BP, LASSO, or OMP) is adopted, the error upper bound can be expressed as:

$$\left\|\mathbf{h}_{DD} - \hat{\mathbf{h}}_{DD}\right\|_2 \le c_1\left\|\mathbf{h}_r\right\| + c_2\left\|\mathbf{n}_{\mathcal{P}}\right\|, \tag{52}$$

where c_1 and c_2 are related to the RIP condition of $\mathbf{D}$ and the employed sparse recovery algorithm. When $\|\mathbf{h}_r\|$ is large (model mismatch) or the condition of $\mathbf{D}$ is poor, the reconstruction error increases.

In addition, if sparse recovery erroneously classifies certain components as 'zero' or incorrectly identifies the support set, estimation based on this support set can lead to structurally significant errors. Errors in support set detection are typically 'discrete and nonlinear': even a small noise or approximation error can cause the recovery algorithm to select incorrect atoms, leading to a dramatic increase in estimation error. When recursive estimation is employed, the error is amplified according to the recursion matrix. Let the recursive matrix be $\mathbf{\Phi}$; then the stepwise error recursion is given by:

$$e_{t+1} = \mathbf{\Phi}e_t + \xi_t. \tag{53}$$

If the spectral radius $\rho(\mathbf{\Phi}) > 1$, it may lead to error accumulation or amplification. However, the proposed TF-PA-CEE scheme performs global extrapolation in one step, thus avoiding error propagation and chain amplification. Thus, the performance gain of TF-PA-CEE mainly stems from:

- More effective second-order prior information.
- Reduced error propagation in one-step extrapolation.
- Avoidance of information loss due to improper sparsity assumptions.

In addition, if the second-order prior is strictly updated according to *Lemma* 1, the extrapolation error of TF-PA-CEE is controllable. Although DD estimation can still be applied to FBMC-OTFS, the use of prototype filters (e.g., PHYDYAS, Hermite, etc.) introduces inherent imaginary interference in the TF domain. When switching to the DD domain, interference spreads to other DD domain bins, weakening the sparsity of the channel. In contrast, treating imaginary interference as one of the transmission properties and employing a TF estimation with OFDM pilot-compatible structures can better adapt to FBMC-OTFS. Thus, TF-domain estimation is preferable specifically for FBMC-OTFS and exhibits better consistency with OFDM-OTFS.

D. TF-Domain Equalization

Since both pilot deployment and channel estimation are performed in the TF domain, the equalization must be implemented synchronously in the TF domain, which ensures logical consistency in the signal processing chain. When the one-tap channel $h_{l,k}$ in the TF domain is estimated, simple one-tap

$$\begin{aligned}\hat{\mathbf{h}}_{DD\to TF} &= \mathbf{U}^{H}\mathbf{R}_{\mathbf{h}_{DD}}\mathbf{D}^{H}\left(\mathbf{D}\mathbf{R}_{\mathbf{h}_{DD}}\mathbf{D}^{H} + P_{\mathrm{n}}\mathbf{I}_{\mathcal{P}}\right)^{-1}\hat{\mathbf{h}}_{\mathcal{P}} = \mathbf{U}^{H}\left(\mathbf{U}\mathbf{R}_{\mathbf{h}_{TF}}\mathbf{U}^{H}\right)\mathbf{U}\mathbf{S}^{H}\left(\mathbf{S}\mathbf{U}^{H}\mathbf{U}\mathbf{R}_{\mathbf{h}_{TF}}\mathbf{U}^{H}\mathbf{S}^{H} + P_{\mathrm{n}}\mathbf{I}_{\mathcal{P}}\right)\hat{\mathbf{h}}_{\mathcal{P}} \\ &= \mathbf{R}_{\mathbf{h}_{TF}}\mathbf{S}^{H}\left(\mathbf{S}\mathbf{R}_{\mathbf{h}_{TF}}\mathbf{S}^{H} + P_{\mathrm{n}}\mathbf{I}_{\mathcal{P}}\right)^{-1}\hat{\mathbf{h}}_{\mathcal{P}} = \hat{\mathbf{h}}_{TF}.\end{aligned} \tag{50}$$

equalization can be used directly. The one-tap Zero Forced (ZF) equalization can be written as

$$\hat{x}_{l,k}^{ZF} = y_{l,k}/h_{l,k}. \tag{54}$$

Note that to maintain the integrity of the signal frame structure, the equalization operation must preserve the pilot symbols. The one-tap Minimum Mean Square Error (MMSE) equalization [33] can be written as

$$\hat{x}_{l,k}^{MMSE} = \frac{y_{l,k}h_{l,k}^*}{|h_{l,k}|^2 + P_{\mathrm{n}}} \frac{1}{\frac{1}{L}\sum_{l=1}^{L} 1/\left(1 + P_{\mathrm{n}}/|h_{l,k}|^2\right)}. \tag{55}$$

The second term of Eq. (55) is a scaling factor that guarantees an approximate unbiased estimation. On the other hand, according to the model in Eq. (12), the full-block MMSE equalization can be written as

$$\hat{\mathbf{x}} = \mathbf{H}^H\left(\mathbf{H}\mathbf{H}^H + P_{\mathrm{n}}\mathbf{R}_{\mathrm{n}}\right)^{-1}\mathbf{y}, \tag{56}$$

where $\mathbf{R}_{\mathrm{n}} = \mathbf{G}^H\mathbf{G}$ is the noise correlation matrix. Moreover, Maximum Likelihood (ML) detection is equivalent to

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\boldsymbol{\mathcal{S}}} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2, \tag{57}$$

where $\boldsymbol{\mathcal{S}}$ denotes the set of modulation symbols. For high-order constellations, to reduce complexity, Eq. (44) can be equivalently implemented using a sphere decoder. The detection criterion of the sphere decoder can be expressed as

$$\|\mathbf{y}' - \mathbf{R}\mathbf{x}\|^2 \le d^2, \ s.t. \ \mathbf{H} = \mathbf{Q}\mathbf{R}, \tag{58}$$

where $\mathbf{y}' = \mathbf{Q}^H\mathbf{y}$ and d is the search radius. $\mathbf{Q}$ and $\mathbf{R}$ denote the unitary matrix and the upper triangular matrix, respectively, resulting from the QR decomposition of $\mathbf{H}$. After removing all pilot symbols from the estimated symbol $\hat{\mathbf{x}} \in \mathbb{C}^{LK\times 1}$, the estimated data symbols $\hat{\mathbf{x}}_{\mathcal{D}} \in \mathbb{C}^{|\mathcal{D}|\times 1}$ in the TF domain can be obtained. Then, by performing the SFFT on $\hat{\mathbf{x}}_{\mathcal{D}} \in \mathbb{C}^{|\mathcal{D}|\times 1}$, we can recover the corresponding data symbols $\hat{\mathbf{d}}_{\mathcal{D}} \in \mathbb{C}^{|\mathcal{D}|\times 1}$ in the DD domain. One-tap equalizers, block equalizers and sphere decoder have different requirements for channel information. The proposed scheme can directly provide the required channel information for both, according to specific demands. Thus, our scheme is more flexible.

IV. Numerical Results

For numerical evaluation, we perform a series of experiments. We assume that all path delays fall within the continuous interval $[0, \tau_{\max}]$ and are discretized via uniform sampling. The Doppler shifts of the paths follow a continuous Jakes spectrum, i.e., $\nu_i = \nu_{\max}\cos\theta_i$ with $\theta_i \sim \mathcal{U}(0, 2\pi)$, which are also discretized through uniform sampling. Considering a single-antenna OTFS system, the Doppler shifts across different subcarriers (or antennas) are approximately identical. Thus, our experiments neglect the Doppler squint effect described in [40]. However, The modeling and compensation of Doppler squint under millimeter-wave massive MIMO conditions will serve as one of important directions in our future research. Unless stated otherwise, the simulation parameters are summarized in Table I.

TABLE I
Simulation Parameters

Parameter Name	Expression	Value
Carrier frequency	f_c	2.5GHz
Subcarrier spacing	F	15kHz
Number of subcarriers per frame	L	12
Number of symbols per frame	K	14
Number of pilots per frame	$\lvert\mathcal{P}\rvert$	12
Total number of parallel frames	N_{f}	4, 8
Total sampling rate	$5FL \times N_{\mathrm{f}}$	{3.6, 7.2}MHz
QAM Constellation	/	4, 16
Channel model	/	EVA, TDL-A
Number of resolvable paths	/	8, 16
Number of Rayleigh multipaths	N_p	200
Velocity	V	200km/h, 500km/h

Note that the above key parameters strictly comply with the technical specifications of 3GPP LTE Release 19 [33]. The sampling rate is set to 5 times the single-frame bandwidth, which not only satisfies the Nyquist sampling criterion (twice the bandwidth) but also aligns with the commonly adopted 3-5 times oversampling practice in engineering implementations. In OFDM-based OTFS, the cyclic prefix (CP) length is set to 4.7μs, which is derived from the maximum delay spread ($\tau_{\max} = 2.3\mu$s) of a typical urban channel model with a safety factor of 2.

In simulation settings, the Doppler shifts follow a continuous Jakes spectrum; thus, the mapping of Doppler shift onto the discrete Doppler grid is generally non-integer; The delays of all Rayleigh/Ricean channel paths are uniformly distributed within the maximum delay spread. Accordingly, the mappings of the 200 Rayleigh/Ricean paths onto the discrete delay grid are generally non-integer. This indicates that the energy of each physical path is distributed across multiple bins in DD domain instead of being confined to integer grid, which corresponds to fractional delay and fractional Doppler. In addition, although the system can only resolve finite delay indices, each index may include multiple physical paths with close delays which cannot be resolved. According to the central limit theorem, the large number of irresolvable paths leads to the channel exhibiting Rayleigh/Ricean fading properties. Therefore, the 200 Rayleigh multipaths are not individually resolved in the TF/DD domains, but are equivalently treated as Rayleigh fading coefficients over finite indices.

A. Analysis of Effective Frame Interval

Lemma 1 reveals the temporal decay law of prior information in time-varying channels. The longer the time interval between two frames, the worse the timeliness (i.e., the larger the mean square error). For quantitative analysis, we selected the Doppler shifts $\nu \in \{10, 40, 300, 1400\}$Hz, which correspond to mobility velocities of approximately $V \in \{5, 20, 130, 600\}$km/h, respectively. Such a parameter set covers from walking to high mobility scenarios.

Fig. 3a shows the mean square error corresponding to different frame intervals. From Fig. 3a, we observe that an increase in the frame interval $\mathcal{T}$ leads to an increase in the mean square error. That is, the timeliness of the prior information deteriorates, resulting in a decrease in the accuracy

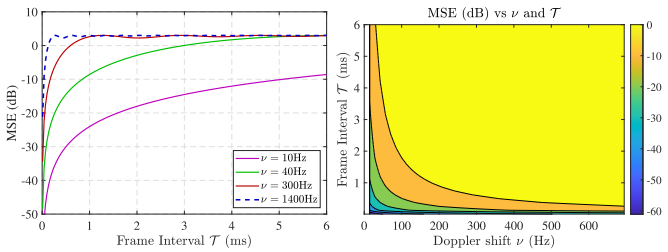

Fig. 3. (a) Frame interval versus mean square error; (b) Mean square error plot: MSE vs frame interval and Doppler shift.

of channel estimation. On the other hand, low Doppler shift ($\nu \leq 40$Hz, corresponding to $V \leq 20$km/h) allows us to adopt a prior information update period of $\mathcal{T} = 4$ms. In high-mobility communication scenarios ($\nu \geq 300$Hz, corresponding to $V \geq 130$km/h), the channel exhibits fast fading characteristics. To ensure the accuracy of channel estimation, the update period of prior information needs to be shortened to $\mathcal{T} = 1$ms. To simultaneously investigate the impact of frame interval and fast time-varying characteristics on the effectiveness of prior information, we treat $\mathcal{T}$ and ν as variables and analyze their effects on the mean square error.

Fig. 3b shows the mean square error corresponding to different frame intervals and Doppler frequency shifts. Based on the numerical results in Fig. 3b, we can select an appropriate frame interval to optimize the accuracy of channel estimation under different channel conditions. It is worth noting that channel estimation also incorporates CoPP information. We do not treat the prior information strictly as the channel information of other frames but rather extrapolate the channel state of other frames based on its statistical characteristics. Therefore, the constraints of *Lemma* 1 can be appropriately relaxed, or the upper bound of the mean square error can be relaxed to -3dB. Once the constraints of the Lemma are relaxed, we can adopt a larger update period.

B. Performance Analysis of NMSE

According to the constraints of *Lemma* 1, we choose the update period as $\mathcal{T} = 2$ms. To evaluate the accuracy of channel estimation, we choose the NMSE as the evaluation metric, denoted as

$$\bar{e}^2_{NMSE} = 10\lg\left(\left\|\hat{\mathbf{H}} - \mathbf{H}\right\|^2_{m2} / \|\mathbf{H}\|^2_{m2}\right), \tag{59}$$

where $\|\cdot\|_{m2}$ denotes the $m2$-norm of the matrix. For reference, the Cramér-Rao Low bound (CRLB) is provided as a baseline. Specifically, according to Eq. (5), let $h_i(t_n) = \frac{1}{\sqrt{N_p}} a_i e^{\mathrm{j}(2\pi\nu_i t_n + \varphi_i)} \delta_{t_n,\tau_i}$ denote the channel of ith path. Then, the Fisher information matrix of the unknown parameter $\{\nu_i, \tau_i, \varphi_i\}$ is given by:

$$\mathbf{F} = \begin{bmatrix} \mathbf{F}_{\nu\nu} & \mathbf{F}_{\nu\tau} & \mathbf{F}_{\nu\varphi} \\ \mathbf{F}_{\tau\nu} & \mathbf{F}_{\tau\tau} & \mathbf{F}_{\tau\varphi} \\ \mathbf{F}_{\varphi\nu} & \mathbf{F}_{\varphi\tau} & \mathbf{F}_{\varphi\varphi} \end{bmatrix}, \quad s.t.$$
$$[\mathbf{F}_{\nu\nu}]_{i,j} = \frac{2}{P_{\mathrm{n}}} \Re\left\{ \sum_{n_{\mathcal{P}}=1}^{|\mathcal{P}|} \frac{\partial h_i^*}{\partial \nu_i} \frac{\partial h_j}{\partial \nu_j} \right\}, \; i,j = 1, \cdots, N_p, \tag{60}$$

The partial derivative of $h_i(t_n)$ with respect to $\{\nu_i, \tau_i, \varphi_i\}$ can be calculated as

$$\begin{cases} \frac{\partial h_i(t_n)}{\partial \nu_i} = \frac{1}{\sqrt{N_p}} a_i \mathrm{j} 2\pi t_n e^{\mathrm{j}(2\pi\nu_i t_n + \varphi_i)} \delta_{t_n,\tau_i} \\ \frac{\partial h_i(t_n)}{\partial \tau_i} = -\frac{1}{\sqrt{N_p}} a_i e^{\mathrm{j}(2\pi\nu_i t_n + \varphi_i)} \delta'_{t_n,\tau_i} \\ \frac{\partial h_i(t_n)}{\partial \varphi_i} = \frac{1}{\sqrt{N_p}} a_i \mathrm{j} e^{\mathrm{j}(2\pi\nu_i t_n + \varphi_i)} \delta_{t_n,\tau_i}, \end{cases} \tag{61}$$

where δ'_{t_n,τ_i} is the derivative of Dirac delta function. According to the CRLB, the covariance lower bound of the unbiased estimator for parameter $\{\nu_i, \tau_i, \varphi_i\}$ is $\mathrm{Cov}(\{\hat{\nu}_i, \hat{\tau}_i, \hat{\varphi}_i\}) \geq \mathbf{F}^{-1}$. Thus, the variance lower bound of ith path parameter is given by:

$$\begin{cases} \mathrm{var}(\hat{\nu}_i) \geq \left[\mathbf{F}^{-1}\right]_{i,i} \\ \mathrm{var}(\hat{\tau}_i) \geq \left[\mathbf{F}^{-1}\right]_{N_p+i,N_p+i} \\ \mathrm{var}(\hat{\varphi}_i) \geq \left[\mathbf{F}^{-1}\right]_{2N_p+i,2N_p+i}. \end{cases} \tag{62}$$

Since $h_i(t_n)$ is a function of the parameter $\{\nu_i, \tau_i, \varphi_i\}$, we define all path parameters as:

$$\boldsymbol{\theta} = \left[\nu_i, \cdots, \nu_{N_p}, \tau_i, \cdots, \tau_{N_p}, \varphi_i, \cdots, \varphi_{N_p}\right]^T. \tag{63}$$

Then, the estimated value of the parameter is $\hat{\boldsymbol{\theta}} = \boldsymbol{\theta} + \Delta\boldsymbol{\theta}$. By defining $h(\boldsymbol{\theta}) = \sum_{i=1}^{N_p} h_i(\nu_i, \tau_i, \varphi_i)$, the first-order Taylor expansion of $h(\boldsymbol{\theta} + \Delta\boldsymbol{\theta})$ can be expressed as:

$$h(\boldsymbol{\theta} + \Delta\boldsymbol{\theta}) \approx h(\boldsymbol{\theta}) + \nabla_{\boldsymbol{\theta}} h(\boldsymbol{\theta})^T \Delta\boldsymbol{\theta}, \tag{64}$$

where $\nabla_{\boldsymbol{\theta}}$ is the parameter gradient, and $\nabla_{\boldsymbol{\theta}} h(\boldsymbol{\theta})^T \Delta\boldsymbol{\theta}$ can be calculated as:

$$\nabla_{\boldsymbol{\theta}} h(\boldsymbol{\theta})^T \Delta\boldsymbol{\theta} = \sum_{i=1}^{N_p} \left(\frac{\partial h(\boldsymbol{\theta})}{\partial \nu_i} \Delta\nu_i + \frac{\partial h(\boldsymbol{\theta})}{\partial \tau_i} \Delta\tau_i + \frac{\partial h(\boldsymbol{\theta})}{\partial \varphi_i} \Delta\varphi_i \right). \tag{65}$$

$$\mathrm{Cov}(\hat{\mathbf{h}}) \geq \sum_{i=1}^{N_p} \left(\frac{\partial \mathbf{h}}{\partial \nu_i} \frac{\partial \mathbf{h}^H}{\partial \nu_i} \left[\mathbf{F}^{-1}\right]_{i,i} + \frac{\partial \mathbf{h}}{\partial \tau_i} \frac{\partial \mathbf{h}^H}{\partial \tau_i} \left[\mathbf{F}^{-1}\right]_{N_p+i,N_p+i} + \frac{\partial \mathbf{h}}{\partial \varphi_i} \frac{\partial \mathbf{h}^H}{\partial \varphi_i} \left[\mathbf{F}^{-1}\right]_{2N_p+i,2N_p+i} \right). \tag{69}$$

$$\mathrm{NMSE}_{\mathrm{CRLB}} \geq \sum_{i=1}^{N_p} \left(\left\| \frac{\partial \mathbf{h}}{\partial \nu_i} \right\|^2 \left[\mathbf{F}^{-1}\right]_{i,i} + \left\| \frac{\partial \mathbf{h}}{\partial \tau_i} \right\|^2 \left[\mathbf{F}^{-1}\right]_{N_p+i,N_p+i} + \left\| \frac{\partial \mathbf{h}}{\partial \varphi_i} \right\|^2 \left[\mathbf{F}^{-1}\right]_{2N_p+i,2N_p+i} \right) / \sum_{i=1}^{N_p} \|\mathbf{h}_i\|^2. \tag{70}$$

$$\mathrm{NMSE}^{\mathcal{P}}_{\mathrm{CRLB}} \geq \sum_{i=1}^{N_p} \left(\left\| \mathbf{S} \frac{\partial \mathbf{h}}{\partial \nu_i} \right\|^2 \left[\mathbf{F}^{-1}\right]_{i,i} + \left\| \mathbf{S} \frac{\partial \mathbf{h}}{\partial \tau_i} \right\|^2 \left[\mathbf{F}^{-1}\right]_{N_p+i,N_p+i} + \left\| \mathbf{S} \frac{\partial \mathbf{h}}{\partial \varphi_i} \right\|^2 \left[\mathbf{F}^{-1}\right]_{2N_p+i,2N_p+i} \right) / \sum_{i=1}^{N_p} \|\mathbf{S}\mathbf{h}_i\|^2. \tag{71}$$

Define $\mathbf{h} \triangleq [h_1(\boldsymbol{\theta}), \cdots, h_{\mathcal{L}}(\boldsymbol{\theta})]$, then

$$\hat{\mathbf{h}} \approx \mathbf{h} + \mathbf{J}\Delta\boldsymbol{\theta}, \tag{66}$$

where $\mathbf{J} \in \mathbb{C}^{\mathcal{L}\times 3N_p}$ is Jacobian matrix, which can be calculated as

$$\mathbf{J} = \left[\frac{\partial \mathbf{h}}{\partial\{\nu_1, \cdots, \nu_{N_p}\}}, \frac{\partial \mathbf{h}}{\partial\{\tau_1, \cdots, \tau_{N_p}\}}, \frac{\partial \mathbf{h}}{\partial\{\varphi_1, \cdots, \varphi_{N_p}\}}\right]. \tag{67}$$

Thus, the linearized representation of $\hat{\mathbf{h}}$ with respect to the path parameters can be written as:

$$\hat{\mathbf{h}} \approx \mathbf{h} + \sum_{i=1}^{N_p} \frac{\partial \mathbf{h}}{\partial \nu_i}\Delta\nu_i + \frac{\partial \mathbf{h}}{\partial \tau_i}\Delta\tau_i + \frac{\partial \mathbf{h}}{\partial \varphi_i}\Delta\varphi_i. \tag{68}$$

Accordingly, the CRLB of the error covariance can be expressed as Eq. (69), shown at the bottom of the previous page. Thus, the final lower bound for NMSE can be calculated as Eq. (70), shown at the bottom of the previous page. To compute Eq. (70), we map the NMSE to the observable information CoPP at receiver. According to Eq. (41), $\hat{\mathbf{h}}_{\mathcal{P}} = \mathbf{Sh}$. Thereby, $\mathrm{NMSE}^{\mathcal{P}}_{\mathrm{CRLB}}$ can be calculated as Eq. (71), shown at the bottom of the previous page.

Note that $\mathbf{S}\frac{\partial \mathbf{h}}{\partial \nu_i} = \frac{\partial \mathbf{Sh}}{\partial \nu_i}$. Thus, the partial derivative of $\mathbf{Sh}$ with ith path can be calculated as

$$\begin{cases} \dfrac{\partial \mathbf{Sh}_i}{\partial \nu_i} = \mathrm{j}2\pi \mathrm{diag}\left\{t_1, \cdots, t_{|\mathcal{P}|}\right\}\mathbf{Sh}_i \\ \dfrac{\partial \mathbf{Sh}_i}{\partial \tau_i} = -(\mathbf{Sh}_i)' \\ \dfrac{\partial \mathbf{Sh}_i}{\partial \varphi_i} = \mathrm{j}\mathbf{Sh}_i, \end{cases} \tag{72}$$

where $(\mathbf{Sh}_i)'$ is the numerical derivative of the delay with respect to the sampling time. Thus, the squared norm at the pilot positions can be calculated as:

$$\begin{cases} \left\|\mathbf{S}\dfrac{\partial \mathbf{h}}{\partial \nu_i}\right\|^2 = 4\pi^2 \displaystyle\sum_{n_{\mathcal{P}}=1}^{|\mathcal{P}|} t_{n_{\mathcal{P}}}^2 |h_i(l_{n_{\mathcal{P}}}, k_{n_{\mathcal{P}}})|^2 \\ \left\|\mathbf{S}\dfrac{\partial \mathbf{h}}{\partial \tau_i}\right\|^2 \approx \displaystyle\sum_{n_{\mathcal{P}}=1}^{|\mathcal{P}|} \left|h_i{}'(l_{n_{\mathcal{P}}}, k_{n_{\mathcal{P}}})\right|^2 \\ \left\|\mathbf{S}\dfrac{\partial \mathbf{h}}{\partial \varphi_i}\right\|^2 = \displaystyle\sum_{n_{\mathcal{P}}=1}^{|\mathcal{P}|} |h_i(l_{n_{\mathcal{P}}}, k_{n_{\mathcal{P}}})|^2. \end{cases} \tag{73}$$

Thereby, the CRLB baseline can be obtained through CoPP. Since the pilot structure has a significant impact on channel estimation, it is necessary to compare it with other pilot placements to verify the effectiveness of the adopted scheme. To validate the practicality of the selected pattern, we compared it with other classical pilot patterns, including comb-type, block-type, and clustered structures.

Fig. 4a and Fig. 4b show the NMSE performance and spectral efficiency, respectively, under different pilot arrangements. The results in Fig. 4 show that the Diamond Pilot consistently achieves the best NMSE and spectral efficiency. Block Pilot only provides continuous estimation support in the time domain, whereas comb pilots primarily focus on frequency-domain estimation, resulting in suboptimal performance. Cluster pilots are unevenly distributed across time-frequency

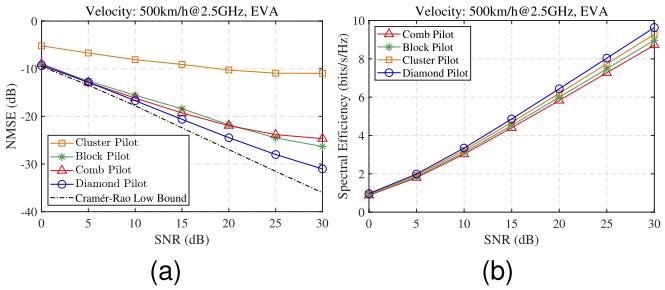

Fig. 4. Different pilot arrangements, total symbol frames, total symbol frames $N_{\mathrm{f}} = 8$. (a) NMSE performance. (b) spectral efficiency.

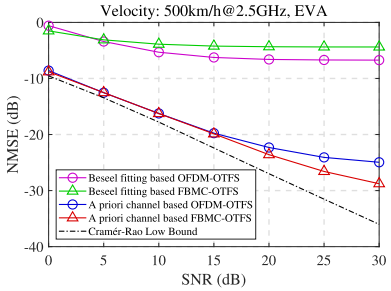

Fig. 5. NMSE plots for one-tap channel estimation: Bessel fitting method vs prior information-assisted method. Total number of symbol frames $N_{\mathrm{f}} = 8$.

resources, resulting in the worst NMSE performance. Note that the experiments in Fig. 4 are based on the priori information-assisted method.

Fig. 5 shows the NMSE performance of both methods in one-tap channel estimation. As shown in Fig. 5, within the SNR range of [0, 30] dB, the NMSE achieved by the prior-information-aided method significantly outperforms that of the Bessel-fitting approach. When Doppler shift is 1.16 kHz (500 km/h), the NMSE of the priori information-assisted method is in the range of [-8, -30] dB, while the Bessel-fit method only reaches [-2, −7] dB. Therefore, we subsequently focus more on the priori information-assisted channel estimation. In Fig. 4a and Fig. 5, the NMSE of the proposed scheme exhibits a significant gap from the CRLB in high-SNR. The reason is that interference dominates in high SNRs, which causes pilot-based interference estimation to be insufficiently accurate. Fortunately, interference can be eliminated through iteration. According to Eq. (11), the interference is described as the off-diagonal elements of $\mathbf{H}$. Thus, the iterative formula can be expressed as

$$\begin{aligned} &\mathbf{y}^{(i+1)} = \mathbf{y}^{(i)} - \left(\hat{\mathbf{H}}^{(i)} - \mathrm{diag}\left\{\hat{\mathbf{h}}^{(i)}\right\}\right)\mathbf{C}\left[\mathbf{x}_{\mathcal{P}}, \hat{\mathbf{x}}_{\mathcal{D}}\right]^T \\ &st.\ \hat{\mathbf{h}}^{(i)} = \mathrm{diag}\left\{\hat{\mathbf{H}}^{(i)}\right\}. \end{aligned} \tag{74}$$

Since both $\hat{\mathbf{H}}^{(i)}$ and $\hat{\mathbf{h}}^{(i)}$ depend on previous estimates, the strict nonlinear coupling makes it difficult to obtain the closed-form solution of Eq. (74). We adopt a first-order linearization approach to characterize error propagation and thereby analyze the convergence of the iteration. Under the linear approximation, the update of the iterative error can be approximately

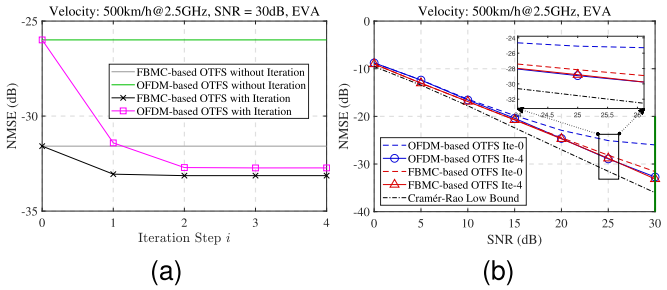

Fig. 6. Total number of symbol frames $N_{\mathrm{f}} = 8$. (a) Iteration step vs. NMSE for SNR = 30dB. (b) NMSE vs. SNR for no iteration and 4 iterations.

expressed as

$$\mathbf{E}^{(i+1)} \approx \left(\mathbf{I} - \mathbf{A}^{(i)}\right) \mathbf{E}^{(i)} + \boldsymbol{\varsigma}^{(i)}, \tag{75}$$

where $\mathbf{E}^{(i)} \triangleq \hat{\mathbf{H}}^{(i)} - \mathbf{H}$ and $\mathbf{A}^{(i)} = \left(\hat{\mathbf{H}}^{(i)} - \mathrm{diag}\left\{\hat{\mathbf{h}}^{(i)}\right\}\right) \mathbf{C}'$. $\mathbf{C}'$ is a constant matrix related to $\mathbf{C}$, $\mathbf{x}_{\mathcal{P}}$ and $\hat{\mathbf{x}}_{\mathcal{D}}$. $\boldsymbol{\varsigma}^{(i)}$ corresponds to the higher-order residual term. Neglecting $\boldsymbol{\varsigma}^{(i)}$, the dominant term is determined by the spectral properties of $\mathbf{I} - \mathbf{A}^{(i)}$. If $\left\|\mathbf{A}^{(i)}\right\| < 2$ and $\left\|\mathbf{I} - \mathbf{A}^{(i)}\right\| \leq \beta < 1$ are valid for all i, then the error $\mathbf{E}^{(i)}$ decreases according to β^i. When the error tends to zero, $\hat{\mathbf{H}}^{(i)} - \mathrm{diag}\left\{\hat{\mathbf{h}}^{(i)}\right\}$ can fully characterize the interference. Consequently, the $\mathbf{y}^{(i+1)}$ in Eq. (74) converges to the received symbols without interference. In the simulation data, for all i, it holds that $\left\|\mathbf{A}^{(i)}\right\| < 2$ and $\left\|\mathbf{I} - \mathbf{A}^{(i)}\right\| \leq \beta < 1$. Therefore, Eq. (74) shows a tendency to converge. Furthermore, replacing $y_{l_{\mathcal{P}},k_{\mathcal{P}}}$ in Eq. (19) with $y^{i+1}_{l_{\mathcal{P}},k_{\mathcal{P}}}$, we can obtain interference-free CoPP information, which allows for more accurate channel estimation.

Fig. 6a shows the NMSE for different number of iteration steps at SNR = 30dB. Fig. 6b shows the NMSE versus SNR for iterated and un-iterated. According to the results shown in Fig. 6a, we observe that OFDM-based OTFS requires two iterations for the NMSE to stabilize and converge to -33.46dB, whereas FBMC-based OTFS requires only one iteration. In particular, all systems achieve a substantial improvement in estimation accuracy after the first iteration, indicating that interference induced by the doubly selective channel can be largely mitigated through iterative processing. On the other hand, the NMSE performance quickly stabilizes as the number of iterations increases, indicating that three iterations are sufficient to achieve a stable NMSE performance, and further iterations are unnecessary. From Fig. 6, the NMSE improves by approximately 7 dB for OFDM-OTFS and by about 2 dB for FBMC-OTFS after iteration. Furthermore, after iterative interference cancellation, OFDM-based OTFS and FBMC-based OTFS exhibit almost identical channel estimation accuracy.

Due to the insufficient estimation accuracy of the Bessel fitting-based method observed in preliminary experiments, the following analysis focuses exclusively on the prior information-aided approach. On the other hand, the above results are obtained under the EVA channel model. To enhance the credibility of the proposed scheme, we provide additional simulation results based on the more commonly used TDL-A channel model, including different system dimensions, various user velocities, and different numbers of resolvable paths. It is worth noting that the resource block size specified in the 3GPP LTE standard is fixed. Therefore, to ensure that the system parameters comply with practical application standards, we adjust the system dimensions by changing the total number of transmitted symbol frames in parallel. Specifically, we considered the following scenarios: a 4-parallel frame with 8 resolvable paths at 200 km/h, a 4-parallel frame at 500 km/h, and an 8-parallel frame with 16 resolvable paths at 500 km/h. In addition, schemes such as CE-BEM [16], Embedded Pilot-Aided Channel Estimation (EPACE) [2], and Joint Channel Estimation and Data Detection (JCEDD) [41] are included in comparison with the proposed TF-PA-CEE to further validate its effectiveness.

Fig. 7 shows the NMSE performance of different estimation methods under various scenarios. According to the results in Fig. 7, the proposed scheme maintains the same trend conclusions as the existing schemes over the entire SNR range, while consistently achieving the best performance. This indicates that the proposed method can maintain stable estimation performance under different parameter settings. However, in Fig. 7, the NMSE of TF-PA-CEE sometimes falls below the CRLB. The reason is that the CRLB is derived from unbiased estimators without assuming any extra prior information, while in the simulations, second-order statistical priors of the channel are introduced. Introducing bias by exploiting prior information (e.g., channel statistics) to achieve NMSE superior to the unbiased estimation lower bound does not violate the CRLB theorem but rather results from the biased Bayesian estimation property [42]. In addition, to evaluate the estimation accuracy of different schemes under different pilot overheads, the NMSE versus pilot overhead is investigated at SNR= 10dB, see Fig. 7d.

A larger pilot overhead provides more available observation samples, improving the accuracy of channel estimation, but at the same time reduces the resources available for data transmission. As shown in Fig. 7d TF-PA-CEE outperforms other comparison schemes across the entire range of overhead. In the low-overhead region ($<$10%), its performance advantage is pronounced, making it highly suitable for communication scenarios with constrained pilot resources.

C. Performance Analysis of BER

Based on the proposed channel estimation results, we perform equalization under two different channel models and analyze the reliability of the communication link. To evaluate reliability, we adopt BER as the performance metric and compare one-tap equalizers, full-block MMSE equalizer, sphere decoder, Message Passing (MP) [43] and Expectation Propagation (EP) [44] detectors.

Fig. 8 shows the BER performance differences of FBMC-based OTFS under different equalizers. The experiment is performed in the non-iterative channel estimation framework with 4-QAM constellation mapping scheme. The results show that, under the same SNR conditions, the sphere decoder ($d = 2$) exhibits a significant advantage, with its BER curve approximately 1-2dB lower than that of EP detector. MP,

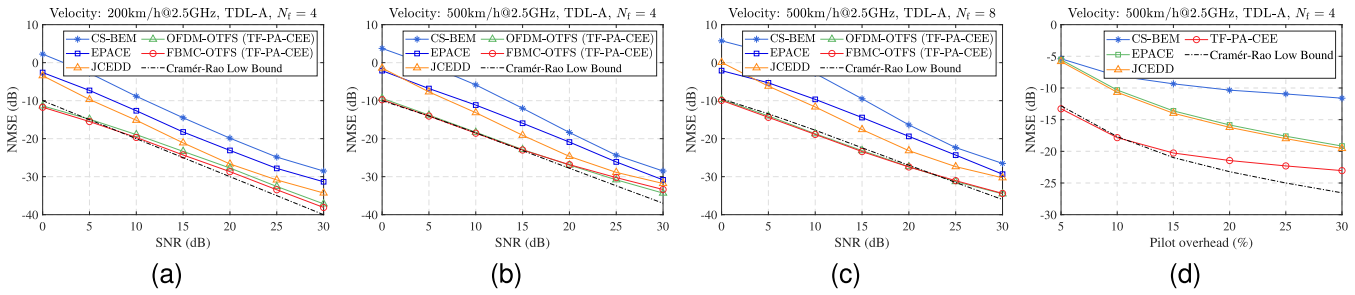

Fig. 7. (a) 4-parallel frame (corresponding to 8 resolvable paths) at 200 km/h. (b) 4-parallel frame (corresponding to 8 resolvable paths) at 500 km/h. (c) 8-parallel frame (corresponding to 16 resolvable paths) at 500 km/h. (d) NMSE versus pilot overhead under SNR= 10dB.

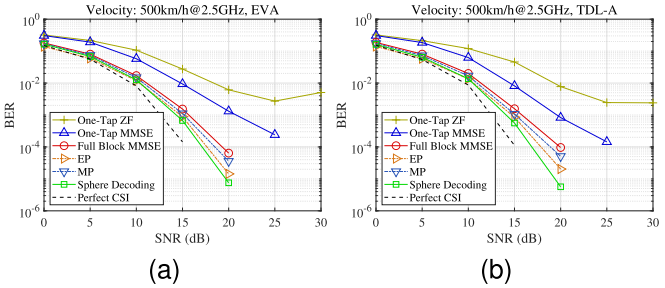

Fig. 8. BER versus SNR for FBMC-based OTFS with different equalizers, with $N_f = 8$. (a) EVA channel at 500 km/h. (b) TDL-A channel at 500 km/h.

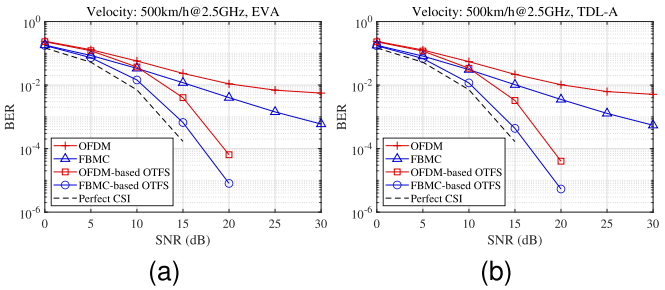

Fig. 9. BER performance with $N_f = 8$ and 4QAM, (a) Under the EVA channel; (b) Under the TDL-A channel.

EP, and the sphere decoder all satisfy the BER target threshold of 10^{-3} within the SNR range of [15, 30] dB, confirming suitability for FBMC-OTFS. Further, we extend to the multi-dimensional performance evaluation of OTFS, specifically exploring the impact of different multi-carrier mechanisms and constellation modulation orders on the BER. To eliminate the impact of equalizer differences on the evaluation results, we adopt the sphere decoder as the reference receiver and conduct experiments within an iterative channel estimation framework.

Fig. 9a and Fig. 9a respectively show the BER gains of OFDM-based OTFS and FBMC-based OTFS compared to conventional OFDM and FBMC under the EVA and TDL-A channel models. The results show that, at a target BER of 10^{-3}, OTFS achieves an SNR gain of approximately 10dB. Moreover, FBMC-based OTFS exhibits superior BER performance compared to OFDM-based OTFS, achieving an improvement of approximately one order of magnitude.

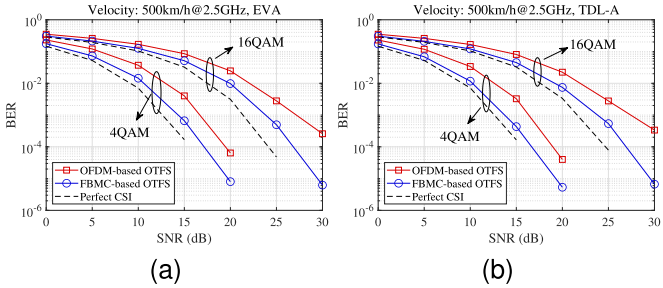

Fig. 10. (a) BER performance under the EVA channel with $N_f = 8$. (b) BER performance under the TDL-A channel with $N_f = 8$.

Furthermore, we investigate the BER performance of FBMC-based OTFS and OFDM-based OTFS under different modulation orders. Fig. 10a and Fig. 10b show the BER performance of OFDM-based OTFS and FBMC-based OTFS under the EVA and TDL-A channel models, respectively. The modulation orders are {4,16}-QAM. To avoid graphic redundancy and enhance visualization, the BER curve of the original multi-carrier system is selectively removed in Fig. 10. However, we should keep in mind that the reliability of original multicarrier systems is lower than that of their derived OTFS. The results show that FBMC-based OTFS maintains a stable BER advantage across different modulation orders.

As shown in the above results, in a high-mobility scenario ($V = 500$km/h), FBMC-based OTFS demonstrates a reliability advantage over OFDM-based OTFS, with an approximately one-order-of-magnitude improvement in BER performance. The reason is that the good time-frequency localization of the Hermite prototype filter effectively reduces the effect of time-frequency two-dimensional spreading in the doubly selective channel. Theoretical and experimental analyses indicate that the reliability of OTFS depends on the time-frequency joint resolution of the underlying multi-carrier modulation mechanism. Thus, by optimizing the TF-domain transmission strategy, OTFS can inherit the interference-resistant capability of multi-carrier modulation.

D. Analysis of Computational Complexity

In the practical evaluation of channel estimation algorithms, we consider the analysis of computational complexity. Due to the insufficient estimation accuracy of the Bessel fitting

TABLE II
SUMMARY OF COMPUTATIONAL COMPLEXITY

i	Calculation of the correlation matrix (updated N_u times)	$\mathcal{O}\left(3N_u\left\|\mathcal{P}\right\|N_s^4\right)$
ii	Calculating the Extrapolation Matrix and Estimating the CoDP	$\mathcal{O}\left(\left\|\mathcal{P}\right\|^3+L^2K^2\left(\left\|\mathcal{P}\right\|^2+\left\|\mathcal{P}\right\|\right)\right)$
iii	Iterative interference cancellation (M steps)	$\mathcal{O}\left(M\left(L^2K^2\left(\left\|\mathcal{P}\right\|+1\right)+L^3K^3\right)\right)$
iv	Total complexity	$\mathcal{O}\left(C_{\mathrm{i}}+C_{\mathrm{ii}}+C_{\mathrm{iii}}\right)\approx\mathcal{O}\left(N_s^4\right)$

method in preliminary experiments (NMSE > -10 dB), it fails to meet the requirements for high mobility scenarios. Therefore, we exclude the discussion of its complexity. Currently, we focus on the complexity of a priori information-assisted extrapolation. Note that we consider the number of complex multiplications as the metric. According to Eqs. (35) - (39), the dimension for $\mathbf{R}_{\boldsymbol{\mathcal{H}}}$ is $N_s^2\times N_s^2$. Thus, each calculation of the correlation matrix leads to a complexity of about $\mathcal{O}\left(N_s^4\right)$. The correlation matrix calculation for all pilot positions requires a total of $3\left|\mathcal{P}\right|$ times $\mathcal{O}\left(N_s^4\right)$. In addition, each update of the prior information requires recalculating the correlation matrix. In practical engineering implementation, $\boldsymbol{\mathcal{H}}$ is a sparse banded matrix even though $N=(O+K-1)N_{FFT}$ is large, with N_{FFT} denoting the number of FFT points. Thus, $\mathbf{R}_{\boldsymbol{\mathcal{H}}}$ is also sparse. Thereby, the computational complexity of obtaining the correlation matrix may be lower than $\mathcal{O}\left(3\left|\mathcal{P}\right|\left(N_s^4\right)\right)$. According to Eq. (40), the calculation of the extrapolation matrix involves the inversion of $\mathbf{R}_{\mathbf{h}_{\mathcal{P}}}\in\mathbb{C}^{\left|\mathcal{P}\right|\times\left|\mathcal{P}\right|}$ and multiplication with $\mathbf{R}_{\mathbf{H},\mathbf{h}_{\mathcal{P}}}\in\mathbb{C}^{L^2K^2\times\left|\mathcal{P}\right|}$, with computational complexities $\mathcal{O}\left(\left|\mathcal{P}\right|^3\right)$ and $\mathcal{O}\left(L^2K^2\left|\mathcal{P}\right|^2\right)$, respectively. According to Eq. (21), the complexity of estimating CoDP is $\mathcal{O}\left(\left|\mathcal{P}\right|L^2K^2\right)$. According to Eq. (74), the complexity of each iteration is $\mathcal{O}\left(L^2K^2+L^3K^3\right)$, and the complexity of estimating CoDP is similarly $\mathcal{O}\left(\left|\mathcal{P}\right|L^2K^2\right)$. In Table II, we summarize the complexity of prior information-assisted channel estimation at each stage. It is worth noting that the frame duration of OFDM-based OTFS is 1ms per frame, while the frame duration of FBMC-based OTFS is 2ms per frame. According to the parameters in Table I, OFDM-based OTFS requires updating the prior information 3 times, while FBMC-based OTFS requires updating it 6 times. For iterative interference cancellation, the number of iteration steps does not exceed 4. Thus, the total complexity of the prior information-aided channel estimation is approximately $\mathcal{O}\left(N_s^4\right)$. Surely, in the LTE standard, large bandwidths contain L and K large enough that cannot be ignored. In this case, the total complexity is $\mathcal{O}\left(C_{\mathrm{i}}+C_{\mathrm{ii}}+C_{\mathrm{iii}}\right)$.

To visualize the complexity of different schemes, we quantitatively analyze it by changing L. Note that in the 3GPP LTE frame structure, the number of time symbols within 1ms, $K=14$, remains constant. Therefore, we keep K constant and change the number of subcarriers L by increasing the number of parallel transmission frames. Fig. 11 shows the complexity of the different schemes. According to the results in Fig. 11, the proposed TF-PA-CEE method can reduce computational complexity by approximately 25% compared with JCEDD when iterative interference cancellation is not performed. Although the complexity of TF-PA-CEE (with interference cancellation) increases, two iterations are typically sufficient,see Fig. (6a), and the complexity remains significantly lower than that of OMP. Its core is performing interference cancellation through a series of linear operations with minimal iterations. Although the complexity increases when iterative interference cancellation is applied, this comes with a significant performance improvement. Therefore, we recommend adopting the non-iterative version in scenarios with stringent real-time requirements, while employing the iterative version in performance-prioritized scenarios to achieve a trade-off between complexity and performance.

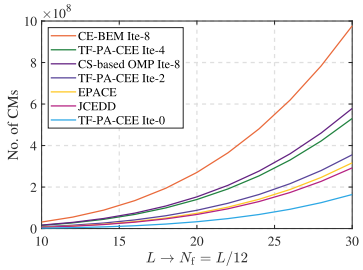

Fig. 11. Relationship between the number of Complex Multiplications (CMs) and the number of subcarriers for different systems.

V. CONCLUSION

In this paper, we have constructed an OTFS system based on FBMC and proposed a TF-PA-CEE scheme. The proposed channel estimation scheme achieves accurate tracking of dynamic channel parameters by constructing a Bessel fitting or a priori information-assisted extrapolation mechanism. In addition, the criterion for determining the channel wide-sense stationary time interval is derived to assist the prior information update mechanism. By following the principle of signal processing consistency, both channel estimation and equalization are implemented in the TF domain. The experimental data show that the proposed scheme provides good NMSE performance in high mobility scenarios. Compared to classical OTFS, FBMC-based OTFS achieves an improvement of approximately one order of magnitude in BER performance, while reducing OOB emissions to below -50dB.

It is worth noting that the performance analysis in this paper is based on simulations. In practical implementations, the following issues should be considered: (1) the multi-symbol filtering in FBMC-OTFS may result in high processing latency, and parallel architectures should be optimized to meet real-time requirements; (2) in high-mobility scenarios, the rapid variations of doubly selective channels increase the difficulty of updating prior information, for which adaptive

tracking algorithms can be considered; (3) hardware factors, such as power amplifier nonlinearity and quantization errors, may degrade system performance, and robustness should be enhanced at the algorithm design stage. The above issues are essential considerations for the practical implementation of the proposed scheme, and we plan to address them in future work.

APPENDIX
PROOF OF LEMMA 1 IN CONTINUOUS TIME MODELING

In the channel model of Eq. (5), the following assumptions need to be clearly specified: $a_i(t)$ follows Rayleigh fading and the time correlation is determined by the Doppler spectrum; For a short period of time, $\tau_i(t) \approx \tau_i(t+\mathcal{T})$; The time variation of the channel is primarily caused by the Doppler shift ν_i and the time-varying amplitude $a_i(t)$. The Channel Frequency Response (CFR) on subcarrier l can be defined as

$$H(t,l) = \frac{1}{\sqrt{N_p}} \sum_{i=1}^{N_p} a_i(t)\, \mathrm{e}^{\mathrm{j}(2\pi\nu_i t+\varphi_i)} \mathrm{e}^{-\mathrm{j}2\pi l F \tau_i(t)}. \tag{76}$$

The expression $H(t,l)$ is the Fourier transform of $h(t,\tau)$ with respect to τ, and the frequency is replaced by lF. If the time interval between two frames is $\mathcal{T}$, then the CFR correlation can be calculated as

$$\begin{aligned} R(\mathcal{T}) &= \mathbb{E}\left\{H(t,l)\, H^*(t+\mathcal{T},l)\right\} \\ &= \frac{1}{N_p} \sum_{i=1}^{N_p} \mathbb{E}\left\{a_i(t)\, a_i^*(t+\mathcal{T})\right\} \mathrm{e}^{\mathrm{j}2\pi\nu_i\mathcal{T}} \\ &\quad \times \mathrm{e}^{-\mathrm{j}2\pi l F(\tau_i(t)-\tau_i(t+\mathcal{T}))}. \end{aligned} \tag{77}$$

Since $\tau_i(t) \approx \tau_i(t+\mathcal{T})$, $R(\mathcal{T})$ can be simplified as

$$R(\mathcal{T}) = \frac{1}{N_p} \sum_{i=1}^{N_p} R_{a_i(t)}(\mathcal{T})\, \mathrm{e}^{\mathrm{j}2\pi\nu_i\mathcal{T}}, \tag{78}$$

where $R_{a_i(t)}(\mathcal{T})$ is the autocorrelation function of the Rayleigh fading for the ith path. If the signal undergoes scattering from Wide-Sense Stationary Uncorrelated Scattering (WSSUS) environments, then $R_{a_i(t)}(\mathcal{T})$ depends on the Jakes spectrum $S(\nu_i)$ [45]. $S(\nu_i)$ is denoted as

$$S(\nu_i) = \begin{cases} \dfrac{1}{\pi\nu_{\max}\sqrt{1-(\nu_i/\nu_{\max})^2}}, & |\nu_i| \le \nu_{\max} \\ 0, & otherwise. \end{cases} \tag{79}$$

According to the Wiener-Khinchin theorem, $R_{a_i(t)}(\mathcal{T})$ can be calculated as

$$\begin{aligned} R_{a_i(t)}(\mathcal{T}) &= P_h \int_{-\nu_{\max}}^{\nu_{\max}} S(\nu_i)\, \mathrm{e}^{\mathrm{j}2\pi\nu_i\mathcal{T}} \mathrm{d}\nu_i \\ &= \frac{2P_h}{\pi\nu_{\max}} \int_0^{\nu_{\max}} \frac{\cos(2\pi\nu_i\mathcal{T})}{\sqrt{1-(\nu_i/\nu_{\max})^2}} \mathrm{d}\nu_i \\ &\quad . \end{aligned} \tag{80}$$

Assuming that ν_i follows a uniform angular distribution (i.e., $\nu_i = \nu_{\max}\cos(\theta_i)$, where θ_i follows a uniform distribution), then

$$\begin{aligned} &R_{a_i(t)}(\mathcal{T}) \\ &= \frac{2P_h}{\pi\nu_{\max}} \int_{\frac{\pi}{2}}^{0} \frac{\cos(2\pi\nu_{\max}\mathcal{T}\cos(\theta_i))}{\sqrt{1-\cos^2(\theta_i)}} (-\nu_{\max}\sin(\theta_i))\, \mathrm{d}\theta_i \\ &= \frac{2P_h}{\pi} \int_0^{\frac{\pi}{2}} \cos(2\pi\nu_{\max}\mathcal{T}\cos(\theta_i))\, \mathrm{d}\theta_i \\ &= P_h J_0(2\pi\nu_{\max}\mathcal{T}). \end{aligned} \tag{81}$$

Assuming a sufficiently large transmission path (i.e., $N_p \to \infty$), then $R(\mathcal{T})$ can be calculated as

$$\begin{aligned} R(\mathcal{T}) &= P_h J_0(2\pi\nu_{\max}\mathcal{T}) \lim_{N_p\to\infty} \frac{1}{N_p} \sum_{i=1}^{N_p} \mathrm{e}^{\mathrm{j}2\pi\nu_i\mathcal{T}} \\ &= P_h J_0(2\pi\nu_{\max}\mathcal{T}) \frac{1}{2\pi} \int_0^{2\pi} \mathrm{e}^{\mathrm{j}2\pi\mathcal{T}\nu_{\max}\cos(\theta_i)} \mathrm{d}\theta_i \\ &= P_h J_0^2(2\pi\nu_{\max}\mathcal{T}). \end{aligned} \tag{82}$$

If $H(t,l)$ is directly used to predict $H(t+\mathcal{T},l)$ without relying on CoPP, then the estimated mean squared error is

$$\begin{aligned} &\bar{e}_{MSE} \\ &= \mathbb{E}\left\{\left|H(t+\mathcal{T},l) - H(t,l)\right|^2\right\} \\ &= \mathbb{E}\left\{\left| \frac{1}{\sqrt{N_p}} \sum_{i=1}^{N_p} a_i(t+\mathcal{T})\, \mathrm{e}^{\mathrm{j}(2\pi\nu_i(t+\mathcal{T})+\varphi_i)} \mathrm{e}^{-\mathrm{j}2\pi l F\tau_i(t+\mathcal{T})} \right.\right. \\ &\quad \left.\left. - \frac{1}{\sqrt{N_p}} \sum_{i=1}^{N_p} a_i(t)\, \mathrm{e}^{\mathrm{j}(2\pi\nu_i t+\varphi_i)} \mathrm{e}^{-\mathrm{j}2\pi l F\tau_i(t)} \right|^2 \right\} \\ &= \frac{1}{N_p} \mathbb{E}\left\{\left| \sum_{i=1}^{N_p} \left(a_i(t+\mathcal{T})\, \mathrm{e}^{\mathrm{j}2\pi\nu_i\mathcal{T}} - a_i(t)\right) \right.\right. \\ &\quad \left.\left. \times \mathrm{e}^{\mathrm{j}(2\pi\nu_i t+\varphi_i)} \mathrm{e}^{-\mathrm{j}2\pi l F\tau_i(t)} \right|^2 \right\}. \end{aligned} \tag{83}$$

The Eq. (83) can be expanded as

$$\begin{aligned} \bar{e}_{MSE} = \frac{1}{N_p} \mathbb{E}\left\{ \sum_{i=1}^{N_p} \sum_{j=1}^{N_p} \right. &\left(a_i(t+\mathcal{T})\, \mathrm{e}^{\mathrm{j}2\pi\nu_i\mathcal{T}} - a_i(t)\right) \\ &\times \left(a_j(t+\mathcal{T})\, \mathrm{e}^{\mathrm{j}2\pi\nu_j\mathcal{T}} - a_j(t)\right)^* \\ &\left. \times \mathrm{e}^{\mathrm{j}(2\pi(\nu_i-\nu_j)t+(\varphi_i-\varphi_j))} \mathrm{e}^{-\mathrm{j}2\pi l F(\tau_i(t)-\tau_j(t))} \right\}. \end{aligned} \tag{84}$$

For $\forall i = j$, The autocorrelation of the phase is $\mathrm{e}^{\mathrm{j}0} = 1$. Under the WSSUS conditions, the interval $\mathcal{T}$ only affects the autocorrelation of the gain $a_i(t)$, without destroying the independence between paths. Additionally, $a_i(t)$ is independent across different paths and has zero mean value. Thus, for $\forall i \neq j$, $\mathbb{E}\left\{a_i(t)\, a_j^*(t)\right\} = 0$. Thereby, the cross terms are zero. Eq. (84) can be simplified as

$$\bar{e}_{MSE} = \frac{1}{N_p} \sum_{i=1}^{N_p} \mathbb{E}\left\{\left|a_i(t+\mathcal{T})\, \mathrm{e}^{\mathrm{j}2\pi\nu_i\mathcal{T}} - a_i(t)\right|^2\right\}$$

$$= \frac{1}{N_p}\sum_{i=1}^{N_p}\mathbb{E}\left\{\left|a_i\left(t+\mathcal{T}\right)\right|^2+\left|a_i\left(t\right)\right|^2 - 2\Re\left\{a_i\left(t+\mathcal{T}\right)a_i^*\left(t\right)\mathrm{e}^{\mathrm{j}2\pi\nu_i\mathcal{T}}\right\}\right\}$$
$$= \frac{1}{N_p}\sum_{i=1}^{N_p}\left(1+1-2R\left(\mathcal{T}\right)\right)$$
$$= 2\left(1-R\left(\mathcal{T}\right)\right). \tag{85}$$

If we require $\bar{e}_{MSE}\leq\zeta$, then

$$R\left(\mathcal{T}\right)=P_hJ_0^2\left(2\pi\nu_{\max}\mathcal{T}\right)\geq 1-\zeta/2 \tag{86}$$

Thus, $\mathcal{T}\leq\frac{J_0^{-1}\left(\sqrt{(1-\zeta/2)/P_h}\right)}{2\pi\nu_{\max}}$.

REFERENCES

[1] *Study on New Radio Access Technology Physical Layer Aspect*, document TR 38.802, 3GPP, 2017. [Online]. Available: https://www.3gpp.org/ftp/Specs/archive/38_series/38.802/

[2] P. Raviteja, K. T. Phan, and Y. Hong, "Embedded pilot-aided channel estimation for OTFS in delay–Doppler channels," *IEEE Trans. Veh. Technol.*, vol. 68, no. 5, pp. 4906–4917, May 2019.

[3] H. Qu, G. Liu, L. Zhang, M. A. Imran, and S. Wen, "Low-dimensional subspace estimation of continuous-Doppler-spread channel in OTFS systems," *IEEE Trans. Commun.*, vol. 69, no. 7, pp. 4717–4731, Jul. 2021.

[4] Y. Zhang et al., "A cross-domain iterative OTFS receiver for sparse doubly selective channels," *IEEE Wireless Commun. Lett.*, vol. 13, no. 1, pp. 54–58, Jan. 2024.

[5] T. Zemen, M. Hofer, D. Löschenbrand, and C. Pacher, "Iterative detection for orthogonal precoding in doubly selective channels," in *Proc. IEEE 29th Annu. Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, Sep. 2018, pp. 1–7.

[6] A. Shafie, J. Yuan, P. Fitzpatrick, T. Sakurai, and Y. Fang, "On the coexistence of OTFS modulation with OFDM-based communication systems," *IEEE Trans. Commun.*, vol. 72, no. 11, pp. 6822–6838, Nov. 2024.

[7] R. Nissel, F. Ademaj, and M. Rupp, "Doubly-selective channel estimation in FBMC-OQAM and OFDM systems," in *Proc. IEEE 88th Veh. Technol. Conf. (VTC-Fall)*, Aug. 2018, pp. 1–5.

[8] S. M. Pishvaei, B. Mozaffari Tazehkand, and J. Pourrostam, "Design and performance evaluation of FBMC-based orthogonal time–frequency space (OTFS) system," *Phys. Commun.*, vol. 53, Aug. 2022, Art. no. 101723.

[9] Q. Deng, Y. Ge, and Z. Ding, "A unifying view of OTFS and its many variants," *IEEE Commun. Surveys Tuts.*, vol. 27, no. 6, pp. 3561–3586, Dec. 2025.

[10] A. Chatterjee, V. Rangamgari, S. Tiwari, and S. S. Das, "Nonorthogonal multiple access with orthogonal time–frequency space signal transmission," *IEEE Syst. J.*, vol. 15, no. 1, pp. 383–394, Mar. 2021.

[11] J.-C. Liao, W.-C. Chen, and C.-D. Chung, "Performance analysis of SIMO–OFDM systems with pilot-aided channel estimation," *IEEE Syst. J.*, vol. 17, no. 2, pp. 2366–2377, Jun. 2023.

[12] Y. Liu, Y. L. Guan, and D. González, "Turbo BEM OTFS receiver with optimized superimposed pilot power," *IEEE Trans. Commun.*, vol. 72, no. 1, pp. 601–617, Jan. 2024.

[13] T. Thaj and E. Viterbo, "Low complexity iterative rake decision feedback equalizer for zero-padded OTFS systems," *IEEE Trans. Veh. Technol.*, vol. 69, no. 12, pp. 15606–15622, Dec. 2020.

[14] S. E. Zegrar, A. S. Sümer, and H. Arslan, "Fractional delay and fractional Doppler estimation and mitigation framework in OTFS systems," *IEEE Trans. Veh. Technol.*, vol. 74, no. 2, pp. 2884–2896, Feb. 2025.

[15] H. Zhang, X. Huang, and J. A. Zhang, "Low-overhead OTFS transmission with frequency or time domain channel estimation," *IEEE Trans. Veh. Technol.*, vol. 73, no. 1, pp. 799–811, Jan. 2024.

[16] Q. Guo, H. Jiang, J. Xiang, and Y. Zhong, "A CS-BEM OTFS channel estimation approach for sparse continuous Doppler-spread channels," *IEEE Wireless Commun. Lett.*, vol. 13, no. 11, pp. 2985–2989, Nov. 2024.

[17] Q. Guo, H. Jiang, J. Xiang, and Y. Zhong, "Low complexity iterative channel estimation and detection based on pilot-assisted for ZP-OTFS," *IEEE Commun. Lett.*, vol. 29, no. 4, pp. 724–728, Apr. 2025.

[18] Y. Luo, Y. L. Guan, Y. Ge, D. González G, and C. Yuen, "A novel angle-delay-Doppler estimation scheme for AFDM-ISAC system in mixed near-field and far-field scenarios," *IEEE Internet Things J.*, vol. 12, no. 13, pp. 22669–22682, Jul. 2025.

[19] S. Wang, Y. Luo, P. Zuo, L. Pan, Y. Li, and Z. Sun, "In-memory analog solution of compressed sensing recovery in one step," *Sci. Adv.*, vol. 9, no. 50, p. 2908, Dec. 2023.

[20] R. Nissel, S. Schwarz, and M. Rupp, "Filter bank multicarrier modulation schemes for future mobile communications," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 8, pp. 1768–1782, Aug. 2017.

[21] Y. Wang, Q. Guo, J. Xiang, and Y. Liu, "Doubly selective channel estimation and equalization based on ICI/ISI mitigation for OQAM-FBMC systems," *Phys. Commun.*, vol. 59, Aug. 2023, Art. no. 102120.

[22] Y. Wang, Q. Guo, J. Xiang, L. Wang, and Y. Liu, "Bi-orthogonality recovery and MIMO transmission for FBMC systems based on non-sinusoidal orthogonal transformation," *Signal Process.*, vol. 219, Jun. 2024, Art. no. 109427.

[23] R. Haas and J.-C. Belfiore, "A time-frequency well-localized pulse for multiple carrier transmission," *Wireless Pers. Commun.*, vol. 5, no. 1, pp. 1–18, Jul. 1997.

[24] R. Nissel and M. Rupp, "On pilot-symbol aided channel estimation in FBMC-OQAM," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2016, pp. 3681–3685.

[25] P. Raviteja, Y. Hong, E. Viterbo, and E. Biglieri, "Practical pulse-shaping waveforms for reduced-cyclic-prefix OTFS," *IEEE Trans. Veh. Technol.*, vol. 68, no. 1, pp. 957–961, Jan. 2019.

[26] M. Giordani, M. Polese, A. Roy, D. Castor, and M. Zorzi, "A tutorial on beam management for 3GPP NR at mmWave frequencies," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 1, pp. 173–196, 1st Quart., 2019.

[27] W. Jiang et al., "Terahertz communications and sensing for 6G and beyond: A comprehensive review," *IEEE Commun. Surveys Tuts.*, vol. 26, no. 4, pp. 2326–2381, 2024.

[28] *Study on Channel Model for Frequencies From 0.5 to 100 GHz*, document TR 38.901, 3GPP, 2022. [Online]. Available: https://www.3gpp.org/ftp/Specs/archive/38_series/38.901/

[29] V. Pawar and K. Krishna Naik, "Blind multipath time varying channel estimation using recursive Cholesky update," *AEU - Int. J. Electron. Commun.*, vol. 70, no. 1, pp. 113–119, Jan. 2016.

[30] G. Matz, D. Schafhuber, K. Grochenig, M. Hartmann, and F. Hlawatsch, "Analysis, optimization, and implementation of low-interference wireless multicarrier systems," *IEEE Trans. Wireless Commun.*, vol. 6, no. 5, pp. 1921–1931, May 2007.

[31] Y. Wang, Q. Guo, J. Xiang, and Y. Zhong, "Pruned DCT precoding-based FBMC modulation: An SC-FDMA inspired approach," *Signal Process.*, vol. 231, Jun. 2025, Art. no. 109897.

[32] W. Shen, L. Dai, J. An, P. Fan, and R. W. Heath, "Channel estimation for orthogonal time frequency space (OTFS) massive MIMO," *IEEE Trans. Signal Process.*, vol. 67, no. 16, pp. 4204–4217, Aug. 2019.

[33] R. Nissel and M. Rupp, "Pruned DFT-spread FBMC: Low PAPR, low latency, high spectral efficiency," *IEEE Trans. Commun.*, vol. 66, no. 10, pp. 4811–4825, Oct. 2018.

[34] *Physical Channels and Modulation*, document TS 36.211, 3GPP, 2024. [Online]. Available: https://www.3gpp.org/ftp/Specs/archive/36_series/36.211/

[35] M. Ekstrom, "Realizable Wiener filtering in two dimensions," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-30, no. 1, pp. 31–40, Feb. 1982.

[36] P. Hoeher, S. Kaiser, and P. Robertson, *Pilot-Symbol-Aided Channel Estimation in Time and Frequency*. Boston, MA, USA: Springer, 1997, pp. 169–178.

[37] H. Fu, W. Si, and I.-M. Kim, "Deep learning-based joint pilot design and channel estimation for OFDM systems," *IEEE Trans. Commun.*, vol. 71, no. 8, pp. 4577–4590, Aug. 2023.

[38] T. Takahashi, A. Tölli, S. Ibi, and S. Sampei, "Low-complexity large MIMO detection via layered belief propagation in beam domain," *IEEE Trans. Wireless Commun.*, vol. 21, no. 1, pp. 234–249, Jan. 2022.

[39] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, 1993. [Online]. Available: https://www.amazon.com/exec/obidos/ASIN/0133457117/

[40] M. Duan, P. Zhang, S. Zhang, Y. Ge, O. A. Dobre, and C. Yuen, "Channel estimation and hybrid precoding for massive MIMO-OTFS system with doubly squint," *IEEE Trans. Commun.*, vol. 73, no. 10, pp. 9388–9404, Oct. 2025.

[41] M. Li, S. Zhang, Y. Ge, F. Gao, and P. Fan, "Joint channel estimation and data detection for hybrid RIS aided millimeter wave OTFS systems," *IEEE Trans. Commun.*, vol. 70, no. 10, pp. 6832–6848, Oct. 2022.

[42] P. Stoica and R. L. Moses, "On biased estimators and the unbiased Cramér–Rao lower bound," *Signal Process.*, vol. 21, no. 4, pp. 349–350, Dec. 1990.
[43] Y. Ge, Q. Deng, P. C. Ching, and Z. Ding, "OTFS signaling for uplink NOMA of heterogeneous mobility users," *IEEE Trans. Commun.*, vol. 69, no. 5, pp. 3147–3161, May 2021.
[44] F. Long, K. Niu, and J. Lin, "Low complexity block equalizer for OTFS based on expectation propagation," *IEEE Wireless Commun. Lett.*, vol. 11, no. 2, pp. 376–380, Feb. 2022.
[45] P. Hoeher, "A statistical discrete-time model for the WSSUS multipath channel," *IEEE Trans. Veh. Technol.*, vol. 41, no. 4, pp. 461–468, Apr. 1992.

Ying Wang received the B.S. degree in communication engineering from Henan Engineering University, Henan, China, in 2021, and the degree from Harbin Engineering University (HUE), Harbin, China, in 2021, where he is currently pursuing the Ph.D. degree in information and communication engineering. He has authored/co-authored ten technical articles in major international journals. He has received two Chinese invention patents. His main research interests include 4G/5G communications, massive MIMO, multicarrier modulation, waveform design, and signal processing in wireless communications.

Qiang Guo received the M.S. and Ph.D. degrees from Harbin Engineering University (HEU), Harbin, China, in 2003 and 2007, respectively. He is an Expert in accreditation for Chinese Education Ministry and Postgraduate Education, and a fellow of Ukrainian Engineering Academy. He is currently a Professor with HEU, where he is also the Leader of the Ocean Renaissance Academic Team. He has authored/co-authored more than 30 technical articles in major international journals, and one published a scholarly book by Science Publisher. His main research interests include radar detection and countermeasures, and satellite communications.

Jianhong Xiang (Member, IEEE) received the M.S. and Ph.D. degrees from Harbin Engineering University (HEU), Harbin, China, in 2006 and 2009, respectively. He is currently a Professor with HEU, where he is also the Executive Deputy Director of the National Experimental Teaching Demonstration Center for Electrical and Electronics. He is the Vice President for the Digital Economy Research Association of Heilongjiang Province. He has authored/co-authored more than 30 technical papers and has received more than ten Chinese invention patents. His primary research interests include communications countermeasures, artificial intelligence, and satellite communications.

Yu Zhong received the B.S. degree in applied electronic technology from Guilin University of Electronic Science and Technology, China, in 2001. In 2002, he was with the 10th Research Institute, China Electronics Technology Group Corporation, where he is currently a Senior Engineer, working as a Field Master. His research interests include signal processing and wireless communications.