Received 24 July 2024; revised 5 September 2024; accepted 15 September 2024. Date of publication 18 September 2024;
date of current version 27 September 2024.

Digital Object Identifier 10.1109/OJCOMS.2024.3463202

Joint Channel and Data Estimation via Parametric Bilinear Inference for OTFS Demodulation

KENTA ITO [1] (Graduate Student Member, IEEE), TAKUMI TAKAHASHI [1] (Member, IEEE), KENGO FURUTA [1] (Graduate Student Member, IEEE), SHINSUKE IBI [2] (Senior Member, IEEE), AND GIUSEPPE THADEU FREITAS DE ABREU [3] (Senior Member, IEEE)

[1]Graduate School of Engineering, Osaka University, Suita 565-0871, Japan

[2]Faculty of Science and Engineering, Doshisha University, Kyotanabe 610-0394, Japan

[3]School of Computer Science and Engineering, Constructor University Bremen, 28759 Bremen, Germany

CORRESPONDING AUTHOR: T. TAKAHASHI (e-mail: takahashi@comm.eng.osaka-u.ac.jp)

This work was supported in part by JSPS KAKENHI under Grant JP23K13335, Grant JP23K22754, and Grant JP23K20935; and in part by MIC/FORWARD under Grant JPMI240710001.

ABSTRACT This paper proposes a novel joint channel and data estimation (JCDE) algorithm via parametric bilinear Gaussian belief propagation (PBiGaBP) for orthogonal time frequency space (OTFS) demodulation. In doubly-selective fading channels, since the Doppler shift breaks the orthogonality between subcarriers, the signal demodulation process requires estimating the path gains and data symbols from the high-dimensional signals that are multiplexed in the frequency-time (FT) domain. To obtain highly accurate channel state information (CSI) using a typical channel estimation scheme based only on reference signals, an increase in pilot overhead to suppress inter-symbol interference (ISI) is unavoidable. In addition, large-scale matrix operations based on the size of OTFS-equivalent channels also pose problems in terms of computational cost. To address these issues, we focus on the fact that OTFS demodulation can be formulated as a large-scale parametric bilinear inference (PBI) problem, and solve it using the Gaussian belief propagation (GaBP) approach, which enables an approximate implementation of the sum-product algorithm (SPA), based on the central limit theorem (CLT), to design a low-complexity and high-accuracy JCDE algorithm using short pilots. Simulation results show that the proposed PBiGaBP-based JCDE algorithm outperforms the state-of-the-art (SotA) schemes and approaches the performance of idealized (genie-aided) scheme in terms of mean square error (MSE) and bit error rate (BER) performances.

INDEX TERMS OTFS demodulation, high-mobility communications, joint channel and data estimation, parametric bilinear inference, Gaussian belief propagation.

I. INTRODUCTION

THE FIFTH-GENERATION (5G) advanced and future sixth-generation (6G) wireless networks are strongly anticipated by industry as platforms that provide reliable connections for a variety of high-mobility terminals, such as autonomous vehicles, unmanned aerial vehicles (UAVs), high-speed trains (HSTs), and so on [1], [2], [3]. Under such high-mobility scenarios, the wireless channel is known to suffer from doubly-selective fading with frequency-selectivity due to multipath propagation with different delays and time-selectivity due to Doppler shift caused by the mobility of the wireless terminals and/or scatters around them [4]. Orthogonal frequency-division multiplexing (OFDM), which is used in earlier 5G wireless releases, is robust against frequency-selectivity but vulnerable to time-selectivity. Thus, if the current OFDM system is used as-is in high-mobility communication scenarios, *e.g.*, when assuming a speed of 300 km/h or more, the Doppler shift causes the orthogonality among subcarriers to collapse, making it difficult to maintain stable wireless communications [5], [6], [7].

To address this issue, orthogonal time frequency space (OTFS) modulation [2], [8], which is robust against doubly-selective fading channels, has attracted much attention. In

the OTFS modulation, data symbols are placed in the delay-Doppler (DD) domain, multiplexed on a two-dimensional orthogonal basis function, and then transmitted as signals in the frequency-time (FT)-domain. Owning to this operation, OTFS modulation allows for both time and frequency diversity and provides more robust communication than the OFDM alternative [9]. In addition, robust equalization against doubly-selective fading can be achieved by exploiting the sparsity of the DD-domain channel when the long-term statistics, *i.e.*, delay and Doppler parameters, are available as a (pre-estimated) prior knowledge [10], [11].

However, since the orthogonality between subcarriers cannot be used in the OTFS demodulation in a doubly-selective fading environment, the equalization process becomes complicated due to severe inter-symbol interference (ISI) among the pilot and data symbols. In order to eliminate the effect of ISI and demodulate data symbols with high accuracy, it is essential to obtain highly accurate channel state information (CSI). In addition, the size of the equivalent channel matrix for observing data symbols increases as a result of symbol spreading, making practical implementation difficult without a low-complexity demodulation algorithm [12], [13], [14], [15].

A well-known channel estimation scheme for OTFS systems has been proposed that uses an OTFS frame consisting only of periodically transmitted pilot symbols while taking advantage of the sparsity found in DD-domain channels [12]. However, this scheme inevitably deteriorates system performance due to increased pilot overhead. In addition, when the Doppler shift is large, the data demodulation accuracy is degraded due to the aged CSI because of the time variability of channel between frames. To reduce these inconveniences, an embedded pilot (EP) scheme was proposed in [13] that mixes pilot and data symbols in a single frame; however, to avoid ISI between pilot and data, a number of zero symbols corresponding to the maximum delay and maximum Doppler shift are inserted. Not only is transmission efficiency reduced due to zero padding, but appropriate frame design according to wireless channel conditions is also an issue.

To avoid this reduction in transmission efficiency, the authors of [14] proposed a superimposed pilot (SP) scheme based on the integrated design of OTFS frames and demodulation schemes, in which pilot and data symbols are superimposed in the power domain within a single frame and then transmitted, and the receiver uses a joint channel and data estimation (JCDE) algorithm to exploit tentatively detected data symbols as *soft* pilots for channel estimation, thereby achieving accurate demodulation while suppressing ISI [14], [16]. However, the computational cost required for the JCDE-based demodulation increases rapidly with frame size; hence, it is not necessarily the most suitable form for OTFS signals, which tends to be high-dimensional.

In light of the above, this paper designs a JCDE algorithm based on a low-complexity Bayesian message passing algorithm (MPA) [17], [18] motivated by the fact that an OTFS signal model can be formulated into large-scale parametric bilinear inference (PBI) problem [19], in which two unknown variables (*i.e.*, channel gains and data symbols) with a symmetric linear regression structure are estimated from multi-dimensional observations. Note that in the context of data detection for OTFS signals, it has been shown in [20], [21] that MPAs based on belief propagation (BP) that utilize the ISI structure of a doubly-selective channel can achieve reliable demodulation performance.

The most common Bayesian MPA tailored for large-scale PBI is the parametric bilinear generalized approximate message passing (PBiGAMP) algorithm [19], which is derived by approximating the sum-product algorithm (SPA), designed based on the PBI formulation, using the framework for deriving Bayesian optimal MPAs in the large-system limit condition[1] proposed in [22], and notably, possesses linear complexity with respect to each dimension of the system. Its operating principle is an *extrinsic* value exchange on a graphical model, achieved by using an *Onsager* correction term [23], [24] that predicts and decouples the self-feedback across iterative processes, and it is empirically known that these MPA converge asymptotically to the Bayesian optimal solution as the system size increases in the case of independent and identically distributed (i.i.d.) sub-Gaussian observations.[2] Thus, if the system size is not sufficiently enough, the Onsager correction term does not work properly, and the estimation accuracy of PBiGAMP is significantly reduced. Since it is necessary to exchange extrinsic values via a structural matrix, the derivation process relies heavily on large-system limit approximation, making it difficult to use PBiGAMP as-is for realistic-sized systems. In many studies on signal demodulation using PBiGAMP (*e.g.*, [25], [26]), its poor estimation performance is compensated for using error correction that assumes strong channel coding.

As an MPA that achieves accurate PBI in realistic system sizes, bilinear adaptive vector approximate message passing (BAd-VAMP) [27], [28] has been proposed, which relaxes the maximum likelihood (ML) estimation problem using the expectation consistent (EC) approximation [29] and introduces an adaptive adjustment of the variance of unknown variables. However, since BAd-VAMP consists of channel estimation via least square (LS) and data detection via vector approximate message passing (VAMP) [30], [31], it is inevitable that the computational complexity will increase due to high-dimensional matrix inversion operations, which ultimately compromises the scalability of the system. In addition, the LS-based channel estimation cannot take into account the reliability of symbol estimates, so there is also the problem that the estimation accuracy drops significantly when using non-orthogonal pilots.

Bilinear unitary approximate message passing (BiUAMP), which uses only a whitening process based on singular value decomposition (SVD) of equivalent channel matrix has also been proposed in [32]; however, its algorithmic structure is

[1]The system limit condition, where input dimensions M and U and output dimension N are infinity for fixed rates $\rho_1 = N/M$ and $\rho_2 = N/U$.

[2]Note that an MPA that can be rigorously proven to be Bayesian optimal in PBI has not been presented, and remains an open problem.

essentially the same as VAMP (SVD form) [30], and it does not offer the same scalability potential of PBiGAMP.

Based on the above, we propose a novel low-complexity JCDE algorithm via parametric bilinear Gaussian belief propagation (PBiGaBP) [33][3] to achieve more accurate PBI, with the same order of complexity as PBiGAMP. The core idea of the proposed method is to relax the approximation in the derivation of the algorithm. Gaussian belief propagation (GaBP) [17], [18] relies only on scalar Gaussian approximation (SGA) based on mild central limit theorem (CLT), in contrast to the much harder asymptotic conditions required by GAMP. Indeed, it has been shown in [34] and [35] that the GAMP-based linear/bilinear inference algorithms are proven to be systematically derived from a rigorous approximation of the GaBP-based algorithms in the large-system limit condition. Thus, the GaBP approach corresponds to a generalization of the GAMP approach based on classical BP regime in finite-size systems, and its extrinsic value exchange rule, which subtracts self-feedback across iterations from each message individually, enables estimation performance to be maintained even in realistic-sized systems [36], [37], [38], [39].

To validate the efficacy of the GaBP approach to design the JCDE algorithm for OTFS demodulation, this paper extends our conference paper [33] by presenting a detailed process flow of the JCDE algorithm via PBiGaBP, providing a detailed simulation analysis of estimation performance compared to the state-of-the-art (SotA) alternatives such as PBiGAMP and BAd-VAMP, and then presenting a detailed computational complexity analysis.

Our contributions are summarized as follows:

- Inspired by [14], the OTFS demodulation consisting of channel estimation and data symbol estimation is formulated as a PBI problem by incorporating all the transmit (TX) and receive (RX) processing, *i.e.*, symbol spreading/inverse spreading (domain conversion) and cyclic prefix (CP) insertion/removal, into an equivalent channel matrix. In addition, based on its structure, we consider an appropriate signal domain in which the Bayesian JCDE algorithm should be performed in terms of a Gram matrix of channel.
- A novel JCDE algorithm via PBiGaBP for OTFS demodulation is presented, wherein the large-scale PBI is performed based on the GaBP approach that relies only on the SGA in conformity with mild CLT, whose underlying assumptions are much softer than the large-system limit assumption on which the PBiGAMP relies heavily. In our method, the ISI is gradually suppressed through iterative processing using an interference cancellation mechanism; therefore, there is no need to methodically optimize the design or placement of pilot symbols to account for delay and Doppler shift.
- To confirm the validity of the proposed PBiGaBP-based JCDE algorithm for OTFS demodulation, the proposed method is evaluated for various system parameters. The numerical results are presented in terms of normalized mean square error (NMSE) of estimated quantities and bit error rate (BER). In addition to evaluating the estimation performance, the computational cost required for each iteration step is assessed in terms of the number of real multiplication operations. The results show that the proposed method outperforms the SotA alternatives for all system parameters used, and approaches the performance of an idealized scheme in which channel coefficients are perfectly known.

The remainder of this paper is organized as follows: in Section II, the mathematical notation and signal model are presented and then a basic channel estimation technique based on linear minimum mean square error (MMSE) filtering is reviewed. In Section III the new JCDE algorithm via PBiGaBP for OTFS demodulation is proposed. In Section IV, computer simulations are conducted to validate the efficacy of the proposed method. In Section V, a brief summary of this paper is presented.

II. PRELIMINARIES

A. MATHEMATICAL NOTATION

The following notation is used throughout, unless when otherwise specified. Sets of real and complex numbers are denoted by $\mathbb{R}$ and $\mathbb{C}$, respectively. Vectors and matrices are denoted by lower- and upper-case bold-face fonts, respectively. The conjugate, transpose, and conjugate transpose operators are denoted by $\cdot^*$, $\cdot^{\mathrm{T}}$, and $\cdot^{\mathrm{H}}$, respectively. The imaginary unit is denoted by $\mathrm{j} = \sqrt{-1}$. Random variables and their outcomes are denoted in sans serif and italic fonts, respectively, as in a and a, such that the conditional probability density function (PDF) and the conditional expectation of the outcome a of a, given the occurrence b of b are respectively denoted by $p_{\mathsf{a}|\mathsf{b}}(a|b)$ and $\mathbb{E}_{\mathsf{a}}[a|b]$. The complex Gaussian distribution with mean a and variance b is denoted by $\mathcal{CN}(a, b)$. The $a \times a$ square identity matrix is denoted by $\boldsymbol{I}_a$. The $a \times b$ zero matrix is denoted by $\boldsymbol{0}_{a\times b}$. The $a \times a$ discrete Fourier transform (DFT) matrix is denoted by F_a. The vector formed by vertically stacking the column vectors of a matrix $\boldsymbol{A}$ is denoted by $\boldsymbol{a} = \mathrm{vec}[\boldsymbol{A}]$. The inverse operation is defined as $\boldsymbol{A} = \mathrm{vec}^{-1}[\boldsymbol{a}]$ The diagonal matrix constructed by placing the elements of a vector $\boldsymbol{a}$ on its main diagonal is denoted by $\mathrm{diag}[\boldsymbol{a}]$. The (i, j)-th element and the sub-matrix consisting of columns a_1 to a_2 of a matrix $\boldsymbol{A}$ are respectively denoted by $[\boldsymbol{A}]_{i,j}$ and $[\boldsymbol{A}]_{a_1:a_2}$. The Kronecker product operation is denoted by $\otimes$. The floor function and Dirac delta function are respectively denoted by $\lfloor\cdot\rfloor$ and $\delta(\cdot)$. Finally, we omit the integrating variable differential from all integrals and use the notation $\sum_{i\neq j}^{I} a_i \triangleq \sum_{i=1}^{I} a_i - a_j$.

B. SIGNAL MODEL

Consider a signal demodulation problem consisting of channel estimation and data symbol detection in an OTFS

[3]This conference paper is an earlier version of this article, which was presented at the IEEE CCNC 2024.

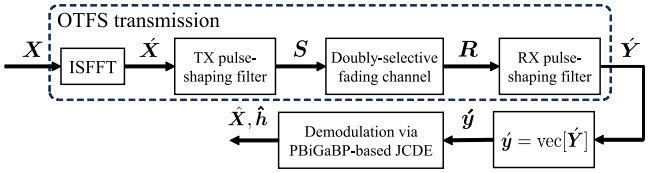


FIGURE 1. OTFS system block diagram, including the proposed PBiGaBP-based JCDE block.

communication system in which the transmitter and receiver have a single antenna.[4] Fig. 1 shows the block diagram of the OTFS system, including the signal demodulation block via the proposed PBiGaBP-based JCDE algorithm. The following sections will explain each block in detail.

1) TRANSMITTER

First, data symbols placed in the DD domain are transformed into the FT domain symbols by using inverse symplectic finite Fourier transform (ISFFT). The TX symbol matrix $\boldsymbol{X}$ of the size $K \times L$ in the DD domain is expressed as

$$\boldsymbol{X} \triangleq [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_l, \ldots, \boldsymbol{x}_L] \in \mathbb{C}^{K\times L}, \tag{1a}$$

$$\boldsymbol{x}_l \triangleq \left[x_{1,l}, \ldots, x_{k,l}, \ldots, x_{K,l}\right]^{\mathrm{T}} \in \mathbb{C}^{K\times 1}, \tag{1b}$$

where $x_{k,l}$ is a TX symbol with average power density $\mathbb{E}_{\mathbf{x}_{k,l}}[|x_{k,l}|^2] = E_{\mathrm{s}}$.

In the DD domain, k and l represent delay- and Doppler-index, respectively. Denoting a pilot symbol matrix and a data symbol matrix by $\boldsymbol{X}_{\mathrm{p}} \in \mathbb{C}^{K\times L}$ and $\boldsymbol{X}_{\mathrm{d}} \in \mathbb{C}^{K\times L}$, respectively, $\boldsymbol{X}$ can be expressed as

$$\boldsymbol{X} = \boldsymbol{X}_{\mathrm{p}} + \boldsymbol{X}_{\mathrm{d}}, \tag{2}$$

where $\boldsymbol{X}_{\mathrm{p}}$ has nonzero elements only at the indices corresponding to the pilot-assigned resources and $\boldsymbol{X}_{\mathrm{d}}$ has nonzero elements only at the indices corresponding to the data-assigned resources. The nonzero element of $\boldsymbol{X}_{\mathrm{d}}$, *i.e.*, the data symbol, represents one of Q-quadrature amplitude modulation (QAM) constellation points $\mathcal{X} \triangleq \{\chi_1, \ldots, \chi_q, \ldots, \chi_Q\}$.

The FT-domain TX symbol matrix is given by the DD–FT conversion via the ISFFT as

$$\acute{\boldsymbol{X}} = \boldsymbol{F}_K \boldsymbol{X} \boldsymbol{F}_L^{\mathrm{H}}, \tag{3}$$

where notice that in the FT domain k and l represent frequency- and time-indexes, respectively.

Denoting the number of frequency bins in the system bandwidth by K, the delay-time (DT)-domain TX symbol matrix $\acute{\boldsymbol{S}} \in \mathbb{C}^{K\times L}$ can be expressed as [32]

$$\acute{\boldsymbol{S}} = \tilde{\boldsymbol{G}}_{\mathrm{T}} \boldsymbol{F}_K^{\mathrm{H}} \boldsymbol{F}_K \boldsymbol{X} \boldsymbol{F}_L^{\mathrm{H}} \triangleq \boldsymbol{G}_{\mathrm{T}} \boldsymbol{F}_K \boldsymbol{X} \boldsymbol{F}_L^{\mathrm{H}}, \tag{4}$$

where $\tilde{\boldsymbol{G}}_{\mathrm{T}} \in \mathbb{C}^{K\times K}$ is a TX pulse-shaping filter, and the TX filter matrix is defined as $\boldsymbol{G}_{\mathrm{T}} \triangleq \tilde{\boldsymbol{G}}_{\mathrm{T}} \boldsymbol{F}_K^{\mathrm{H}}$.

[4]Note that the method addressed in this paper can be easily extended to spatially multi-input multi-output (MIMO) setting where the transmitter and receiver have multiple antennas.

According to the OFDM regime, a CP of length C is inserted into every vector of $\acute{\boldsymbol{S}}$ as

$$\boldsymbol{S} = \boldsymbol{A}_+ \acute{\boldsymbol{S}} = \boldsymbol{A}_+ \boldsymbol{G}_{\mathrm{T}} \boldsymbol{F}_K \boldsymbol{X} \boldsymbol{F}_L^{\mathrm{H}} \in \mathbb{C}^{(C+K)\times L}, \tag{5}$$

with the CP insertion matrix

$$\boldsymbol{A}_+ \triangleq \left[[\boldsymbol{I}_K]_{(K-C)+1:K}, \boldsymbol{I}_K\right]^{\mathrm{T}} \in \mathbb{R}^{(C+K)\times K}. \tag{6}$$

The TX matrix $\boldsymbol{S}$ consisting of L column vectors of size $(K+C)\times 1$ is processed by parallel-to-serial (P/S) transformation, converting to a column vector of size $(K+C)L\times 1$. Using an identity $\mathrm{vec}[\boldsymbol{ABC}] = (\boldsymbol{C}^{\mathrm{T}} \otimes \boldsymbol{A})\mathrm{vec}[\boldsymbol{B}]$, (5) can be rewritten as

$$\boldsymbol{s} \triangleq \mathrm{vec}[\boldsymbol{S}] = \left(\boldsymbol{F}_L^{\mathrm{H}} \otimes \boldsymbol{A}_+ \boldsymbol{G}_{\mathrm{T}} \boldsymbol{F}_K\right)\boldsymbol{x}, \tag{7}$$

where $\boldsymbol{x} \triangleq \mathrm{vec}\{\boldsymbol{X}\}$. Note that (7) holds because the DFT matrix $\boldsymbol{F}_L$ is an unitary matrix, *i.e.*, $\boldsymbol{F}_L = \boldsymbol{F}_L^{\mathrm{T}}$.

2) DOUBLY-SELECTIVE FADING CHANNEL

At the receiver, the RX vector $\boldsymbol{r} \in \mathbb{C}^{(C+K)L\times 1}$ is obtained through the doubly-selective fading channel as

$$\boldsymbol{r} = \boldsymbol{H}\boldsymbol{s} + \boldsymbol{n}, \tag{8}$$

where $\boldsymbol{n} \in \mathbb{C}^{(C+K)L\times 1}$ is a additive white Gaussian noise (AWGN) vector, each element of which obeys $\mathcal{CN}(0, N_0)$. The channel matrix is expressed as [10], [40]

$$\boldsymbol{H} = \sum_{u=1}^{U} h_u \boldsymbol{\Pi}^{\lfloor \frac{\tau_u}{\Delta t} \rceil} \boldsymbol{\Delta}_u, \tag{9}$$

where $\boldsymbol{\Delta}_u$ and $\boldsymbol{\Pi}$ are respectively defined as follows:

$$\boldsymbol{\Delta}_u \triangleq \mathrm{diag}\left[\rho_u^0, \rho_u^1, \ldots, \rho_u^{(C+K)L-1}\right], \quad \rho_u = e^{\mathrm{j}2\pi \nu_u \Delta t}, \tag{10}$$

$$\boldsymbol{\Pi} \triangleq \begin{bmatrix} 0 & \cdots & 0 & 1 \\ 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \end{bmatrix}, \tag{11}$$

where U is the number of paths, $u \in \mathcal{U} \triangleq \{1, \ldots, U\}$, and $h_u \sim \mathcal{CN}(0, \phi)$, with $\phi = 1/U$, is the complex path gain of the u-th path. $\boldsymbol{\Delta}$ and $\boldsymbol{\Pi}$ represent a Doppler-shift diagonal matrix and a delay-shift matrix of the size $(C+K)L$, respectively, with the Doppler frequency and delay time of the u-th multi-path component, ν_u and τ_u.

The time resolution is defined as $\Delta t \triangleq \frac{1}{f_s} = \frac{1}{K\Delta f}$, where f_s and Δf denote the sampling rate and subcarrier spacing, respectively.

3) RECEIVER

First, the RX matrix $\boldsymbol{R}$ of the size $(K+C)\times L$ is obtained by processing the serial-to-parallel (S/P) transformation to RX vector $\boldsymbol{r}$ in (8), *i.e.*, $\boldsymbol{R} = \mathrm{vec}^{-1}[\boldsymbol{r}]$, and then the CP inserted in every block is removed as

$$\acute{\boldsymbol{R}} = \boldsymbol{A}_- \boldsymbol{R} \in \mathbb{C}^{K\times L}, \tag{12}$$

with the CP removal matrix

$$\boldsymbol{A}_- \triangleq \left[\boldsymbol{0}_{K\times C}, \boldsymbol{I}_K\right] \in \mathbb{R}^{K\times(C+K)}. \tag{13}$$

Next, applying an RX filter $\boldsymbol{G}_{\mathrm{R}}$ to $\acute{\boldsymbol{R}}$ in (12), the FT-domain RX symbol matrix $\acute{\boldsymbol{Y}}$ can be expressed as

$$\acute{\boldsymbol{Y}} = \boldsymbol{G}_{\mathrm{R}}\acute{\boldsymbol{R}} \in \mathbb{C}^{K\times L}, \tag{14}$$

$$\boldsymbol{G}_{\mathrm{R}} \triangleq \boldsymbol{G}_{\mathrm{T}}^{\mathrm{H}} = \boldsymbol{F}_K\tilde{\boldsymbol{G}}_T^{\mathrm{H}}. \tag{15}$$

Finally, with the FT–DD conversion via the SFFT, the DD-domain RX symbol matrix $\acute{\boldsymbol{Y}}$ is obtained as

$$\boldsymbol{Y} = \boldsymbol{F}_K^{\mathrm{H}}\acute{\boldsymbol{Y}}\boldsymbol{F}_L. \tag{16}$$

4) PROBLEM FORMULATION

Similar to (7), the vector representation of the DD-domain RX symbol matrix $\boldsymbol{Y}$ in (16) is obtained as

$$\boldsymbol{y} \triangleq \mathrm{vec}[\boldsymbol{Y}] = (\boldsymbol{F}_L \otimes \boldsymbol{I}_K)\big(\boldsymbol{I}_L \otimes \boldsymbol{F}_K^{\mathrm{H}}\big)\acute{\boldsymbol{y}}, \tag{17}$$

where $\acute{\boldsymbol{y}}$ is the FT-domain RX symbol vector, *i.e.*,

$$\begin{aligned}\acute{\boldsymbol{y}} &\triangleq \mathrm{vec}\{\acute{\boldsymbol{Y}}\} = (\boldsymbol{I}_L \otimes \boldsymbol{G}_{\mathrm{R}})\big(\boldsymbol{I}_L \otimes \boldsymbol{A}_{\_}\big)\boldsymbol{r}\\ &= \boldsymbol{W}_{\mathrm{R}}\boldsymbol{H}\boldsymbol{W}_{\mathrm{T}}\boldsymbol{x} + \boldsymbol{v},\end{aligned} \tag{18}$$

with

$$\boldsymbol{W}_{\mathrm{R}} \triangleq \boldsymbol{I}_L \otimes \boldsymbol{G}_{\mathrm{R}}\boldsymbol{A}_{\_}, \tag{19a}$$

$$\boldsymbol{W}_{\mathrm{T}} \triangleq \boldsymbol{F}_L^{\mathrm{H}} \otimes \boldsymbol{A}_{+}\boldsymbol{G}_{\mathrm{T}}\boldsymbol{F}_K, \tag{19b}$$

$$\boldsymbol{v} \triangleq \boldsymbol{W}_{\mathrm{R}}\boldsymbol{n}. \tag{19c}$$

For reasons discussed later in Section II-C, in this paper, channel estimation and data detection are realized by directly equalizing the FT-domain RX vector $\acute{\boldsymbol{y}}$ in (18); therefore, the SFFT process in (16) is not necessary.

Substituting (9) into (18) yields

$$\begin{aligned}\acute{\boldsymbol{y}} &= \boldsymbol{W}_{\mathrm{R}}\left(\sum_{u=1}^{U} h_u\boldsymbol{\Pi}^{\lfloor\frac{\tau_u}{\Delta t}\rfloor}\boldsymbol{\Delta}_u\right)\boldsymbol{W}_{\mathrm{T}}\boldsymbol{x} + \boldsymbol{v}\\ &= \sum_{u=1}^{U} h_u\boldsymbol{\Gamma}_u\boldsymbol{x} + \boldsymbol{v},\end{aligned} \tag{20}$$

where the effects of long-term statistics, *i.e.*, delay and Doppler shift, are summarized as structural matrix

$$\boldsymbol{\Gamma}_u \triangleq \boldsymbol{W}_{\mathrm{R}}\boldsymbol{\Pi}^{\lfloor\frac{\tau_u}{\Delta t}\rfloor}\boldsymbol{\Delta}_u\boldsymbol{W}_{\mathrm{T}}, \tag{21}$$

and the number of TX and RX dimensions are respectively defined as $M \triangleq KL$ and $N \triangleq KL$.[5]

According to [14], [41], under doubly-selective fading channels, it can be assumed that the delay and Doppler parameters are constant during the multiple OTFS frame intervals, while the path may change rapidly. This fact implies that delay and Doppler parameters estimated in the previous OTFS frame can be used in subsequent OTFS frames for channel estimation and data detection. Furthermore, in integrated sensing and communication (ISAC) scenarios, it is also possible to separately estimate the delay and Doppler parameters using radar sensing technology [42]. Inspired by the above works, there are many studies that investigate channel coefficient estimation based on the assumption that delay and Doppler are estimated in advance [14], [32], and our study is positioned as a development based on these previous studies. Therefore, in this paper, the long-term statistical structure $\boldsymbol{\Gamma}_u, \forall u$, is assumed to be known (pre-estimated in advance), and from here we will focus on estimating the path gains (channel coefficients) $h_u, \forall u$, and data symbols.

From (20), the FT-domain RX vector $\acute{\boldsymbol{y}}$ can be reformulated as a PBI [19], [27] with a symmetric linear regression structure, wherein two linear inference problems with respect to the channel vector $\boldsymbol{h} \triangleq [h_1, \ldots, h_u, \ldots, h_U]^{\mathrm{T}}$ and the TX symbol vector $\boldsymbol{x}$ are included as follows:

$$\acute{\boldsymbol{y}} = \boldsymbol{\Omega}\boldsymbol{h} + \boldsymbol{v} = \boldsymbol{\Lambda}\boldsymbol{x} + \boldsymbol{v}, \tag{22}$$

where $\boldsymbol{\Omega}$ and $\boldsymbol{\Lambda}$ can be interpreted as equivalent observation matrices for $\boldsymbol{h}$ and $\boldsymbol{x}$, respectively, which are obtained as

$$\boldsymbol{\Omega} \triangleq [\boldsymbol{\Gamma}_1\boldsymbol{x}, \ldots, \boldsymbol{\Gamma}_u\boldsymbol{x}, \ldots, \boldsymbol{\Gamma}_U\boldsymbol{x}] \in \mathbb{C}^{N\times U}, \tag{23a}$$

$$\boldsymbol{\Lambda} \triangleq \sum_{u=1}^{U} h_u\boldsymbol{\Gamma}_u \in \mathbb{C}^{N\times M}. \tag{23b}$$

From (20), the n-th element of $\acute{\boldsymbol{y}}$ can be expressed as

$$\acute{y}_n = \sum_{u=1}^{U}\sum_{m=1}^{M} h_u\gamma_{u,nm}x_m + v_n, \tag{24}$$

where $\gamma_{u,nm} \triangleq [\boldsymbol{\Gamma}_u]_{nm}$.

C. APPROPRIATE SIGNAL DOMAIN FOR BAYESIAN PBI

In this subsection, we explain why the FT-domain signal in (18) is used instead of the DD-domain signal in (17) when the Bayesian JCDE algorithm is executed. Similar to the conventional PBiGAMP, the proposed method gradually improves estimation accuracy by solving the two linear regression problems in (22) based on low-complexity matched filters (MFs) in an iterative manner. Unlike the MMSE filter, which includes an inverse matrix operation, it is difficult for the MF to sufficiently suppress the correlation between observations; hence, in order to achieve accurate estimation, it is necessary to perform the filtering process in an appropriate signal domain where instantaneous correlation is as small as possible [43]. This RX correlation structure in each signal domain can be visually confirmed by a Gram matrix of the corresponding equivalent observation matrix [44].

Fig. 2 compares the intensity of the elements in the DD-domain and FT-domain Gram matrices of equivalent observation matrices, which are respectively obtained as

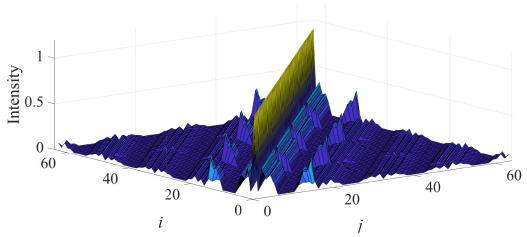
$$\boldsymbol{G}^{\mathcal{DD}} = \check{\boldsymbol{\Lambda}}\check{\boldsymbol{\Lambda}}^{\mathrm{H}}, \tag{25a}$$

$$\boldsymbol{G}^{\mathcal{FT}} = \boldsymbol{\Lambda}\boldsymbol{\Lambda}^{\mathrm{H}}, \tag{25b}$$

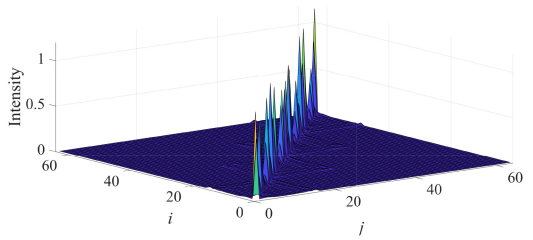
where from (17) and (23b), the DD-domain observation matrix can be expressed as

$$\check{\boldsymbol{\Lambda}} = \big(\boldsymbol{F}_L \otimes \boldsymbol{F}_K^{\mathrm{H}}\big)\boldsymbol{\Lambda}. \tag{26}$$

[5]In practical systems, please note that M and N do not always have the same value due to processes that change the number of signal dimensions, such as oversampling at the transmitter and/or receiver.



(a) Intensity of the elements in the Gram matrix of the DD-domain observation matrix, *i.e.*, $\boldsymbol{G}^{\mathcal{DD}} = \check{\boldsymbol{\Lambda}}\check{\boldsymbol{\Lambda}}^{\mathrm{H}}$.



(b) Intensity of the elements in the Gram matrix of the FT-domain observation matrix, *i.e.*, $\boldsymbol{G}^{\mathcal{FT}} = \boldsymbol{\Lambda}\boldsymbol{\Lambda}^{\mathrm{H}}$.

FIGURE 2. Intensity of the elements in the Gram matrices, where the OTFS system configuration was set to $(K, L, U) = (16, 4, 6)$, and the other parameters are the same as the numerical result in Section IV.

As shown in Fig. 2(a), the fixed correlation structure brought about by $\boldsymbol{F}_L \otimes \boldsymbol{F}_K^{\mathrm{H}}$ emphasizes the off-diagonal elements of the DD-domain Gram matrix in a regular pattern, resulting in large instantaneous correlations among RX observations. In contrast, as shown in Fig. 2(b), the FT-domain Gram matrix approaches a diagonal matrix in high-dimensional OTFS systems due to the high randomness of the observation matrix. This fact implies that when designing low-complexity Bayesian PBI based on MF, it is necessary to design an algorithm based on FT-domain signals.

D. MMSE FILTERING-BASED CHANNEL ESTIMATION

Before moving on to the proposed method, in this subsection we briefly review the conventional pilot-based linear MMSE channel estimation method [14]. In addition, for later convenience, let $\mathcal{M}_{\mathrm{p}}$ denote the set of indices to which pilot symbols are assigned and $\mathcal{M}_{\mathrm{d}}$ denote the set of indices to which data symbols are assigned.

Focusing on the estimation of $\boldsymbol{h}$, (22) can be rewritten by distinguishing between the pilot and data parts as

$$\acute{\boldsymbol{y}} = \boldsymbol{\Omega}_{\mathrm{p}}\boldsymbol{h} + \boldsymbol{\Omega}_{\mathrm{d}}\boldsymbol{h} + \boldsymbol{v}, \tag{27}$$

with

$$\boldsymbol{\Omega}_{\mathrm{p}} \triangleq \left[\boldsymbol{\Gamma}_1\boldsymbol{x}_{\mathrm{p}}, \ldots, \boldsymbol{\Gamma}_u\boldsymbol{x}_{\mathrm{p}}, \ldots, \boldsymbol{\Gamma}_U\boldsymbol{x}_{\mathrm{p}}\right] \in \mathbb{C}^{N\times U}, \tag{28a}$$

$$\boldsymbol{\Omega}_{\mathrm{d}} \triangleq \left[\boldsymbol{\Gamma}_1\boldsymbol{x}_{\mathrm{d}}, \ldots, \boldsymbol{\Gamma}_u\boldsymbol{x}_{\mathrm{d}}, \ldots, \boldsymbol{\Gamma}_U\boldsymbol{x}_{\mathrm{d}}\right] \in \mathbb{C}^{N\times U}, \tag{28b}$$

where we define $\boldsymbol{x}_{\mathrm{p}} \triangleq \mathrm{vec}[\boldsymbol{X}_{\mathrm{p}}]$ and $\boldsymbol{x}_{\mathrm{d}} \triangleq \mathrm{vec}[\boldsymbol{X}_{\mathrm{d}}]$, respectively. Since $\boldsymbol{x}_{\mathrm{d}}$ is unknown when estimating $\boldsymbol{h}$, the second term in (27) is addressed as the equivalent noise.

Accordingly, the linear MMSE filter can be obtained by solving the following minimization problem:

$$\boldsymbol{W} = \arg\min_{\boldsymbol{W}'} J(\boldsymbol{W}'), \ J(\boldsymbol{W}') \triangleq \mathbb{E}_{\mathbf{h},\mathbf{v},\mathbf{x}_{\mathrm{d}}}\left[\left|\boldsymbol{h} - \boldsymbol{W}'\acute{\boldsymbol{y}}\right|^2\right]. \tag{29}$$

Solving for $\frac{\partial J(\boldsymbol{W}')}{\partial \boldsymbol{W}'^{\mathrm{H}}} = 0$, the resultant linear MMSE filter can be expressed as

$$\boldsymbol{W} = \phi\boldsymbol{\Omega}_{\mathrm{p}}^{\mathrm{H}}\left[\phi\boldsymbol{\Omega}_{\mathrm{p}}\boldsymbol{\Omega}_{\mathrm{p}}^{\mathrm{H}} + \boldsymbol{\Theta}\right]^{-1}, \tag{30}$$

with

$$\boldsymbol{\Theta} = \phi\sum_{u=1}^{U}\boldsymbol{\Gamma}_u\boldsymbol{\Psi}\boldsymbol{\Gamma}_u^{\mathrm{H}} + N_0\boldsymbol{I}_N, \tag{31}$$

where the diagonal elements of $\boldsymbol{\Psi} \triangleq \mathrm{diag}[\psi_1, \ldots, \psi_M] \in \mathbb{C}^{M\times M}$ are given by

$$\psi_m = \begin{cases} E_{\mathrm{s}} & m \in \mathcal{M}_{\mathrm{d}} \\ 0 & \text{otherwise.} \end{cases} \tag{32}$$

By applying the linear MMSE filter to the FT-domain RX symbol vector $\acute{\boldsymbol{y}}$, the estimate is obtained by

$$\hat{\boldsymbol{h}} = \boldsymbol{W}\acute{\boldsymbol{y}}, \tag{33}$$

where its mean square error (MSE) is computed by

$$J(\boldsymbol{W}) = \mathrm{tr}\left[\left[\frac{1}{\phi}\cdot\boldsymbol{I}_U + \boldsymbol{\Omega}_{\mathrm{p}}^{\mathrm{H}}\boldsymbol{\Theta}^{-1}\boldsymbol{\Omega}_{\mathrm{p}}\right]^{-1}\right]. \tag{34}$$

From (34), we can see that the linear MMSE estimation requires high-dimensional matrix inversion operations according to the size of the observation matrix. In addition, for highly accurate estimation, it is necessary to reduce the ratio of data symbols in the OTFS frame, which inevitably reduces the system capacity. It is also worth noting that the performance of MMSE filtering-based estimation is consistent with the optimal maximum *a-posteriori* (MAP) estimation when the channel coefficients follow a Gaussian distribution, and therefore this performance presents a lower bound reference that can be achieved when only pilot symbols are used, in terms of NMSE.

III. BILINEAR GAUSSIAN BELIEF PROPAGATION

In this section, we describe the JCDE algorithm via PBiGaBP based on the FT-domain system model in (24), which detects the intended data symbol matrix $\boldsymbol{X}_{\mathrm{d}}$ and estimates the channel vector $\boldsymbol{h}$, out of the FT-domain RX vector $\acute{\boldsymbol{y}}$, the pilot matrix $\boldsymbol{X}_{\mathrm{p}}$, and the long-term statistics $\boldsymbol{\Gamma}_u, \forall u$.

From Bayes' rule, the posterior joint PDF of $\boldsymbol{h}$ and $\boldsymbol{x}_{\mathrm{d}}$, given $\acute{\boldsymbol{y}}$ and $\boldsymbol{x}_{\mathrm{p}}$, can be expressed as

$$p_{\mathbf{h},\mathbf{x}_{\mathrm{d}}|\acute{\mathbf{y}},\mathbf{x}_{\mathrm{p}}}(\boldsymbol{h}, \boldsymbol{x}_{\mathrm{d}}|\acute{\boldsymbol{y}}, \boldsymbol{x}_{\mathrm{p}}) = \frac{p_{\acute{\mathbf{y}}|\mathbf{h},\mathbf{x}_{\mathrm{p}},\mathbf{x}_{\mathrm{d}}}(\acute{\boldsymbol{y}}|\boldsymbol{h}, \boldsymbol{x}_{\mathrm{p}}, \boldsymbol{x}_{\mathrm{d}})p_{\mathbf{h}}(\boldsymbol{h})p_{\mathbf{x}_{\mathrm{p}}}(\boldsymbol{x}_{\mathrm{p}})p_{\mathbf{x}_{\mathrm{d}}}(\boldsymbol{x}_{\mathrm{d}})}{p_{\acute{\mathbf{y}},\mathbf{x}_{\mathrm{p}}}(\acute{\boldsymbol{y}}, \boldsymbol{x}_{\mathrm{p}})}$$

$$\propto \left[\prod_{n=1}^{N} p_{\acute{\mathsf{y}}_n|\mathsf{z}_n}(\acute{y}_n|z_n)\right]\left[\prod_{u=1}^{U} p_{\mathsf{h}_u}(h_u)\right] \times \left[\prod_{m\in\mathcal{M}_\mathrm{p}} \delta\left(x_m - x_m^\mathrm{p}\right) \prod_{m\in\mathcal{M}_\mathrm{d}} p_{\mathsf{x}_m}\left(x_m^\mathrm{d}\right)\right], \tag{35}$$

where we define $z_n \triangleq \sum_{u=1}^{U}\sum_{m=1}^{M} h_u\gamma_{u,nm}x_m$, $x_m^\mathrm{p} \triangleq [\boldsymbol{x}_\mathrm{p}]_{m,1}$, and $x_m^\mathrm{d} \triangleq [\boldsymbol{x}_\mathrm{d}]_{m,1}$. Fig. 3 shows the tripartite FG consisting of factor nodes (FNs) (square) and two variable nodes (VNs) (circle), which correspond to the RX symbols and unknown variables (*i.e.,* channel coefficients and data symbols), respectively. As shown in the figure, the factor graph (FG) for PBI has a symmetrical structure with two VNs; hence, the effects of self-feedback propagating from each of them interact with each other and are severely emphasized. Therefore, in order to design an MPA that achieves stable convergence behavior, a mechanism that decouples the effects of self-feedback with high accuracy is required.

In the BP-based MPAs, the estimation reliability can be gradually improved by iteratively exchanging the *beliefs* (*i.e.*, likelihood information reflecting estimation reliability) and *soft replicas* (*i.e.*, tentative estimates) on the FG. The PBiGaBP-based JCDE algorithm consists of three processes: soft interference canceler (soft IC), belief generator (BG), and soft replica generator (soft RG). In the soft IC, the interference component is removed from the RX signal using the soft replicas of $\boldsymbol{x}$ and $\boldsymbol{h}$ generated in the previous iteration step. Next, based on the SGA of the residual interference component, the BG computes the likelihood information (beliefs) of $\boldsymbol{x}$ and $\boldsymbol{h}$ from the output of the soft IC. Finally, the soft RG generates the soft replicas from the beliefs based on the conditional expectation.

Since the PBiGaBP algorithm propagates different beliefs for each RX index n, the soft replica of x_m^d and h_u are defined as $\hat{x}_{n,m}^\mathrm{d}, \forall n$, and $\hat{h}_{n,u}, \forall n$, respectively, and their MSEs are respectively defined as

$$\psi_{n,m}^{x} \triangleq \mathbb{E}_{x_m}\left[|\tilde{x}_{n,m}|^2\right],\ \tilde{x}_{n,m} \triangleq x_m^\mathrm{d} - \hat{x}_{n,m}^\mathrm{d},\ \forall n, \tag{36a}$$

$$\psi_{n,u}^{h} \triangleq \mathbb{E}_{h_u}\left[|\tilde{h}_{n,u}|^2\right],\ \tilde{h}_{n,u} \triangleq h_u - \hat{h}_{n,u},\ \forall n, \tag{36b}$$

where the quantities $\tilde{x}_{n,m}$ and $\tilde{h}_{n,u}$ denote the estimation errors, respectively. Accordingly, from (2), the soft replica of x_m can be expressed as

$$\hat{x}_{n,m} \triangleq \hat{x}_{n,m}^\mathrm{d} + x_m^\mathrm{p},\ \forall n, \tag{37}$$

and its MSE is $\psi_{n,m}^{x}, \forall n$. For convenience, channel estimation and data detection are explained in separate subsections in order, but in practice they are processed in parallel.

A. CHANNEL ESTIMATION

Let us consider the estimation of an arbitrary channel coefficient h_u, assuming that the soft replicas, *i.e.*, $\hat{x}_{n,m}, \forall n, m$, and $\hat{h}_{n,u}, \forall n, m$, were obtained in the previous iteration.

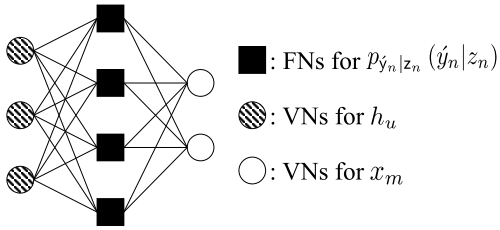


FIGURE 3. The FG for PBI ($N = 4$, $M = 2$, $U = 3$).

1) SOFT IC

The estimation process starts with the soft IC process with the aid of the soft replicas. At the first iteration ($t = 1$), since $\hat{x}_{n,m}^\mathrm{d} = 0, \forall n, m$, the soft replica is only set with pilot symbols, *i.e.*, $\hat{x}_m^\mathrm{p}, \forall m$. The soft IC process for the n-th RX symbol $\acute{y}_n$ can be expressed as

$$\begin{aligned}\acute{z}_{u,n}^{h} &= \acute{y}_n - \sum_{i\neq u}^{U}\sum_{m=1}^{M} \hat{h}_{n,i}\gamma_{i,nm}\hat{x}_{n,m} \\ &= h_u\sum_{m=1}^{M}\gamma_{u,nm}x_m \\ &\quad + \sum_{i\neq u}^{U}\sum_{m=1}^{M}\gamma_{i,nm}\left(h_i x_m - \hat{h}_{n,i}\hat{x}_{n,m}\right) + v_n. \end{aligned}\tag{38}$$

Under the large-system conditions, the residual interference-plus-noise component in (38) can be approximated as a complex Gaussian variable in conformity with CLT; this is referred to as SGA. Under the SGA, the conditional PDF of $\acute{z}_{u,n}^{h}$, given h_u, can be expressed as

$$p_{\acute{\mathsf{y}}_{n,u}^{h}|h_u}\left(\acute{z}_{u,n}^{h}|h_u\right) \propto \exp\left[-\frac{|\acute{z}_{u,n}^{h} - \hat{\omega}_{nu}h_u|^2}{\xi_{u,n}^{h}}\right], \tag{39}$$

where the effective gain for h_u is obtained as

$$\hat{\omega}_{nu} = \sum_{m=1}^{M}\gamma_{u,nm}\hat{x}_{n,m}, \tag{40}$$

and the effective variance is computed as

$$\begin{aligned}\xi_{u,n}^{h} &= \mathbb{E}_{\{\mathsf{h}_i, i\in\mathcal{U}\setminus u\},\{x_m,\forall m\}}\left[\left|\acute{\mathsf{z}}_{u,n}^{h} - \hat{\omega}_{nu}h_u\right|^2 \mid h_u\right] \\ &\approx \sum_{m=1}^{M}\psi_{n,m}^{x}\left|\hat{\lambda}_{u,nm}\right|^2 + \sum_{i\neq u}^{U}\psi_{n,i}^{h}\sum_{m=1}^{M}\psi_{n,m}^{x}|\gamma_{i,nm}|^2 \\ &\quad + \sum_{i\neq u}^{U}\psi_{n,i}^{h}|\hat{\omega}_{ni}|^2 + \phi\sum_{m=1}^{M}\psi_{n,m}^{x}|\gamma_{u,nm}|^2 + N_0, \end{aligned}\tag{41}$$

with

$$\hat{\lambda}_{u,nm} = \sum_{i\neq u}^{U}\hat{h}_{n,i}\gamma_{i,nm}. \tag{42}$$

Note that in (41), the instantaneous quantities h_u and $|h_u|^2$ are not available; therefore, by approximately replacing these

with long-term statistics, *i.e.*, mean $\mathbb{E}_{\mathsf{h}_u}[h_u] = 0$ and variance $\mathbb{E}_{\mathsf{h}_u}[|h_u|^2] = \phi$, it becomes possible to derive the message-passing rules in the closed-form [35].

2) BG

Assuming high-precision SGA of the effective noise components in $\tilde{y}^h_{n,u}, \forall n$, the beliefs corresponding to h_u are combined over all the RX indices except for the n-th RX index, which results in the extrinsic belief for h_u. This is expressed as

$$p_{\hat{\mathsf{q}}_{n,u}|\mathsf{h}_u}(\hat{q}_{n,u}|h_u) = \prod_{i\neq n}^{N} p_{\tilde{y}^h_{u,i}|\mathsf{h}_u}\left(\tilde{y}^h_{u,i}|h_u\right) \propto \exp\left[-\frac{|h_u - \hat{q}_{n,u}|^2}{\psi^q_{n,u}}\right], \tag{43}$$

where $\hat{q}_{n,u}$ and $\psi^q_{n,u}$ are the expected value and variance of h_u in the extrinsic belief domain, which are respectively computed as

$$\hat{q}_{n,u} = \psi^q_{n,u}\sum_{i\neq n}^{N}\frac{\hat{\omega}^*_{iu}\tilde{y}^h_{u,i}}{\xi^h_{u,i}}, \quad \psi^q_{n,u} = \left(\sum_{i\neq n}^{N}\frac{|\hat{\omega}_{iu}|^2}{\xi^h_{u,i}}\right)^{-1}. \tag{44}$$

In the extrinsic combining operation (43), the self-feedback across iterations is decoupled by directly removing the belief corresponding to the n-th RX index, *i.e.*, $\tilde{y}^h_{u,n}$, from the belief propagated to the soft RG, *i.e.*, $\hat{q}_{n,u}$. The resultant message passing derived based on this classical BP regime significantly suppresses the correlation between $\acute{y}_n$ and $\hat{q}_{n,u}$, allowing to avoid self-noise regression on the FG even when the system size is not so large [36], [37], [38], [39]. In addition, it is worth mentioning that this extrinsic belief generation mechanism was proven to be rigorously replaced by the *Onsager* correction term in the large-system limit [34].

Thus, the difference in the mechanism that suppresses the self-feedback across iterations is the biggest difference between the GaBP and generalized approximate message passing (GAMP) approaches, and is the main reason why the GaBP-based algorithm is more robust than GAMP in terms of system size limitations and belief correlation due to structural matrices [35].

3) SOFT RG

Assuming that the effective noise components in $\hat{q}_{n,u}$ are not correlated with each other under high-precision SGA, using Bayes' rule, the soft replica of h_u can be in general obtained from the element-wise conditional expectation, given the extrinsic belief $\hat{q}_{n,u}$, as

$$\hat{h}_{n,u} = \int_{h_u} h_u \frac{p_{\hat{\mathsf{q}}_{n,u}|\mathsf{h}_u}(\hat{q}_{n,u}|h_u)p_{\mathsf{h}_u}(h_u)}{\int_{h'_u} p_{\hat{\mathsf{q}}_{m,u}|\mathsf{h}_u}(\hat{q}_{n,u}|h'_u)p_{\mathsf{h}_u}(h'_u)}, \tag{45}$$

where h_u obeys $\mathcal{CN}(0,\phi)$. Using the Gaussian-PDF multiplication rule [45] in (45), the soft replica $\hat{h}_{n,u}$ and its MSE can be respectively rewritten as

$$\hat{h}_{n,u} = \mathbb{E}_{\mathsf{h}_u}\left[h_u|\hat{q}_{n,u},\psi^q_{n,u}\right] = \frac{\phi\hat{q}_{n,u}}{\psi^q_{n,u}+\phi}, \tag{46a}$$

$$\psi^h_{n,u} = \mathbb{E}_{\mathsf{h}_u}\left[|\tilde{h}_{n,u}|^2|\hat{q}_{n,u},\psi^q_{n,u}\right] = \frac{\phi\psi^q_{n,u}}{\psi^q_{n,u}+\phi}. \tag{46b}$$

B. DATA DETECTION

Next, let us consider the data detection. The above discussion on channel estimation can be applied in the same manner to the detection of an arbitrary $x_m, m \in \mathcal{M}_{\mathrm{d}}$.

Similar to (38), the soft IC process for the n-th RX symbol $\acute{y}_n$ can be expressed as

$$\tilde{y}^x_{m,n} = \acute{y}_n - \sum_{u=1}^{U}\sum_{i\neq m}^{M}\hat{h}_{n,u}\gamma_{u,ni}\hat{x}_{n,i}. \tag{47}$$

Similar to (39), under the SGA of the residual interference-plus-noise component in (47), the effective gain for x_m is obtained as

$$\hat{\lambda}_{nm} = \sum_{u=1}^{U}\hat{h}_{n,u}\gamma_{u,nm}, \tag{48}$$

and the effective variance is computed as

$$\xi^x_{m,n} \approx \sum_{u=1}^{U}\psi^h_{n,u}|\hat{\omega}_{m,nu}|^2 + \sum_{u=1}^{U}\psi^x_{n,u}\sum_{i\neq u}^{U}\psi^h_{n,i}|\gamma_{i,nm}|^2 + \sum_{i\neq m}^{M}\psi^x_{n,i}|\hat{\lambda}_{ni}|^2 + E_{\mathrm{s}}\sum_{u=1}^{U}\psi^h_{n,u}|\gamma_{u,nm}|^2 + N_0, \tag{49}$$

with

$$\hat{\omega}_{m,nu} = \sum_{i\neq m}^{M}\gamma_{u,ni}\hat{x}_{n,i}. \tag{50}$$

Similar to (41), the instantaneous quantities x_m and $|x_m|^2$ are approximately replaced by the true mean $\mathbb{E}_{\chi_q}[\chi_q] = 0$ and variance $\mathbb{E}_{\chi_q}[|\chi_q|^2] = E_{\mathrm{s}}$.

In a similar manner as in (43), the extrinsic belief for x_m can be expressed as

$$p_{\hat{\mathsf{r}}_{n,m}|\mathsf{x}_m}(\hat{r}_{n,m}|x_m) \propto \exp\left[-\frac{|x_m - \hat{r}_{n,m}|^2}{\psi^r_{n,m}}\right], \tag{51}$$

with

$$\hat{r}_{n,m} = \psi^r_{n,m}\sum_{i\neq n}^{N}\frac{\hat{\lambda}^*_{im}\tilde{y}^x_{m,i}}{\xi^x_{m,i}}, \quad \psi^r_{n,m} = \left(\sum_{i\neq n}^{N}\frac{|\hat{\lambda}_{im}|^2}{\xi^x_{m,i}}\right)^{-1}. \tag{52}$$

Similar to (45), using Bayes' rule, the soft replica of x_m and its MSE can generally be obtained from the symbol-wise conditional expectation, given $\hat{r}_{n,m}$, as

$$\hat{x}^{\mathrm{d}}_{n,m} = \mathbb{E}_{\mathsf{x}_m}\left[x_m|\hat{r}_{n,m},\psi^r_{n,m}\right] = \sum_{\chi_q\in\mathcal{X}}\frac{\chi_q p_{\hat{\mathsf{r}}_{n,m}|\mathsf{x}_m}(\hat{r}_{n,m}|\chi_q)p_{\mathsf{x}_m}(\chi_q)}{\sum_{\chi'_q\in\mathcal{X}}p_{\hat{\mathsf{r}}_{n,m}|\mathsf{x}_m}\left(\hat{r}_{n,m}|\chi'_q\right)p_{\mathsf{x}_m}\left(\chi'_q\right)}, \tag{53a}$$

$$\psi^x_{n,m} = \mathbb{E}_{\mathsf{x}_m}\left[|\tilde{x}_{n,m}|^2|\hat{r}_{n,m},\psi^r_{n,m}\right]$$

$$= \sum_{\chi_q \in \mathcal{X}} \frac{\left|\chi_q - \hat{x}^{\mathrm{d}}_{n,m}\right|^2 p_{\hat{\mathsf{r}}_{n,m}|\mathsf{x}_m}\left(\hat{r}_{n,m}|\chi_q\right) p_{\mathsf{x}_m}\left(\chi_q\right)}{\sum_{\chi'_q \in \mathcal{X}} p_{\hat{\mathsf{r}}_{n,m}|\mathsf{x}_m}\left(\hat{r}_{n,m}|\chi'_q\right) p_{\mathsf{x}_m}\left(\chi'_q\right)}. \quad (53b)$$

When the number of iterations reaches the predetermined value T, $\hat{x}^{\mathrm{d}}_m$ is hard-detected as

$$\hat{x}^{\mathrm{d}}_m = \arg\max_{\chi_q \in \mathcal{X}} \left|\chi_q - \hat{r}_m\right|^2, \quad (54)$$

with

$$\hat{r}_m = \psi^r_m \sum_{n=1}^{N} \frac{\hat{\lambda}^*_{nm} \acute{\tilde{y}}^x_{m,n}}{\xi^x_{m,n}}, \quad \psi^r_m = \left(\sum_{n=1}^{N} \frac{|\hat{\lambda}_{nm}|^2}{\xi^x_{m,n}}\right)^{-1}. \quad (55)$$

C. LLR COMPUTATION FOR CHANNEL-CODED SYSTEMS

In channel-coded systems, at the final iteration of the JCDE algorithm ($t = T$), bit-wise log-likelihood ratios (LLRs) must be computed as the input to the channel decoder. Under the SGA of the interference-plus-noise component in $\acute{\tilde{y}}^x_{m,n}, \forall n$, in a similar manner to (51), the posterior belief for x_m can be expressed as

$$p_{\hat{\mathsf{r}}_m|\mathsf{x}_m}\left(\hat{r}_m|x_m\right) \propto \exp\left[-\frac{\left|x_m - \hat{r}_m\right|^2}{\psi^r_m}\right]. \quad (56)$$

When x_m consists of N_s code bits, $c_1, \ldots, c_{n_s}, \ldots, c_{N_s}$, the bit-wise LLR corresponding to the n_s-th bit c_{n_s} is obtained as

$$\beta_{n_s}(x_m) = \ln\left(\frac{\sum_{\chi_q \in \mathcal{X}|c_{n_s}=1} p_{\hat{\mathsf{r}}_m|\mathsf{x}_m}\left(\hat{r}_m|\chi_q\right)}{\sum_{\chi'_q \in \mathcal{X}|c_{n_s}=0} p_{\hat{\mathsf{r}}_m|\mathsf{x}_m}\left(\hat{r}_m|\chi'_q\right)}\right), \quad (57)$$

where $\mathcal{X}|c_{n_s} = c$ is a set consisting of all candidate symbols such that c_{n_s} is $c \in \{0, 1\}$.

D. ALGORITHM DESCRIPTION

The JCDE algorithm based on PBiGaBP exploits the estimated data symbols as effective pilot sequences, which can significantly reduce the pilot overhead while improving the estimation reliability compared to conventional two-stage estimation. However, when the ratio of pilot symbols in the OTFS frame is extremely small, the estimation accuracy in the initial iterations deteriorates, resulting in a significant degradation of the iterative convergence property. Therefore, this paper attempts to further improve the estimation accuracy by introducing a belief damping method [46].

The pseudo-code of the proposed JCDE algorithm via PBiGaBP described above is given in Algorithm 1. Lines 11, 12, 21, and 22 correspond to the belief damping procedure, which aims to avoid the iterative estimation being trapped at a local optimum, especially at the early stage of the iterations by allowing an inertial update of the quantities, where $\eta \in [0, 1]$ is the damping factor.

Algorithm 1 PBiGaBP Algorithm With Belief Damping

Input: $\boldsymbol{y}$, $\boldsymbol{x}_{\mathrm{p}}$, T (Num. of iterations), η
Output: $\left\{\beta_{n_s}(x_m), \forall m \in \mathcal{M}_{\mathrm{d}}, n_s\right\}$

/* Initialization */
1: $\forall n, m \in \mathcal{M}_{\mathrm{p}} : \psi^x_{n,m}(1) = 0$
2: $\forall n, m \in \mathcal{M}_{\mathrm{d}} : \psi^x_{n,m}(1) = E_{\mathrm{s}}$
3: $\forall n, m : \hat{x}^{\mathrm{d}}_{n,m}(1) = 0$
4: $\forall n, u : \hat{h}_{n,u}(1) = 0, \ \psi^h_{n,u}(1) = \phi$
/* JCDE Estimation */
5: **for** $t = 1$ to T **do**
6: $\quad \forall n, m : \hat{x}_{n,m}(t) = \hat{x}^{\mathrm{d}}_{n,m}(t) + x^{\mathrm{p}}_m$
/* Channel estimation: $\forall u$ */
7: $\quad \forall n : \hat{\omega}_{nu}(t) = \sum_{m=1}^{M} \gamma_{u,nm} \hat{x}_{n,m}(t)$
8: $\quad \forall n, m : \hat{\lambda}_{u,nm}(t) = \sum_{i \neq u}^{U} \hat{h}_{n,i}(t) \gamma_{i,nm}$
9: $\quad \forall n : \acute{\tilde{y}}^h_{u,n}(t) = \acute{y}_n - \sum_{i \neq u}^{U} \hat{h}_{n,i}(t) \hat{\omega}_{ni}(t)$
10: $\quad \forall n : \xi^h_{u,n}(t) = \sum_{m=1}^{M} \psi^x_{n,m}(t) \left|\hat{\lambda}_{u,nm}(t)\right|^2$
$+ \sum_{i \neq u}^{U} \psi^h_{n,i}(t) \left(\left|\hat{\omega}_{ni}(t)\right|^2 + \sum_{m=1}^{M} \psi^x_{n,m}(t) \left|\gamma_{i,nm}\right|^2\right)$
$+ \phi \sum_{m=1}^{M} \psi^x_{n,m}(t) \left|\gamma_{u,nm}\right|^2 + N_0$
11: $\quad \forall n : \rho^h_{n,u}(t) = \eta \frac{\left|\hat{\omega}_{nu}(t)\right|^2}{\xi^h_{u,n}(t)} + (1 - \eta) \rho^h_{n,u}(t-1)$
12: $\quad \forall n : \mu^h_{n,u}(t) = \eta \frac{(\hat{\omega}_{nu}(t))^* \acute{\tilde{y}}^h_{u,n}(t)}{\xi^h_{u,n}(t)} + (1 - \eta) \mu^h_{n,u}(t-1)$
13: $\quad \forall n : \psi^q_{n,u}(t) = \left(\sum_{i \neq n}^{N} \rho^h_{i,u}(t)\right)^{-1}$
14: $\quad \forall n : \hat{q}_{n,u}(t) = \psi^q_{n,u}(t) \sum_{i \neq n}^{N} \mu^h_{i,u}(t)$
15: $\quad \forall n : \hat{h}_{n,u}(t+1) = \mathbb{E}_{\mathsf{h}_u}\left[h_u | \hat{q}_{n,u}(t), \psi^q_{n,u}(t)\right]$
16: $\quad \forall n : \psi^h_{n,u}(t+1) = \mathbb{E}_{\mathsf{h}_u}\left[|\tilde{h}_{n,u}|^2 | \hat{q}_{n,u}(t), \psi^q_{n,u}(t)\right]$
/* Data estimation: $\forall m \in \mathcal{M}_{\mathrm{d}}$ */
17: $\quad \forall n : \hat{\lambda}_{nm}(t) = \sum_{u=1}^{U} \hat{h}_{n,u}(t+1) \gamma_{u,nm}$
18: $\quad \forall n, u : \hat{\omega}_{m,nu}(t) = \sum_{i \neq m}^{M} \gamma_{u,ni} \hat{x}_{n,i}(t)$
19: $\quad \forall n : \acute{\tilde{y}}^x_{m,n}(t) = \acute{y}_n - \sum_{i \neq m}^{M} \hat{\lambda}_{ni}(t) \hat{x}_{n,i}(t)$
20: $\quad \forall n : \xi^x_{m,n}(t) = \sum_{u=1}^{U} \psi^h_{n,u}(t+1) \left|\hat{\omega}_{m,nu}(t)\right|^2$
$+ \sum_{i \neq m}^{M} \psi^x_{n,i}(t) (\left|\hat{\lambda}_{ni}(t)\right|^2$
$+ \sum_{u=1}^{U} \psi^h_{n,u}(t+1) \left|\gamma_{u,ni}\right|^2)$
$+ E_{\mathrm{s}} \sum_{u=1}^{U} \psi^h_{n,u}(t+1) \left|\gamma_{u,nm}\right|^2 + N_0$
21: $\quad \forall n : \rho^x_{n,m}(t) = \eta \frac{\left|\hat{\lambda}_{nm}(t)\right|^2}{\xi^x_{m,n}(t)} + (1 - \eta) \rho^x_{n,m}(t-1)$
22: $\quad \forall n : \mu^x_{n,m}(t) = \eta \frac{(\hat{\lambda}_{nm}(t))^* \acute{\tilde{y}}^x_{m,n}(t)}{\xi^h_{m,n}(t)} + (1 - \eta) \mu^x_{n,m}(t-1)$
23: $\quad \forall n : \psi^r_{n,m}(t) = \left(\sum_{i \neq n}^{N} \rho^x_{i,m}(t)\right)^{-1}$
24: $\quad \forall n : \hat{r}_{n,m}(t) = \psi^r_{n,m}(t) \sum_{i \neq n}^{N} \mu^x_{i,m}(t)$
25: $\quad \forall n : \hat{x}^{\mathrm{d}}_{n,m}(t+1) = \mathbb{E}_{\mathsf{x}_m}\left[x_m | \hat{r}_{n,m}(t), \psi^r_{n,m}(t)\right]$
26: $\quad \forall n : \psi^x_{n,m}(t+1) = \mathbb{E}_{\mathsf{x}_m}\left[|\tilde{x}_{n,m}|^2 | \hat{r}_{n,m}(t), \psi^r_{n,m}(t)\right]$
27: **end for**
28: $\forall n_s$: Compute bit-wise LLR $\beta_{n_s}(x_m)$ via (57).
// Termination

IV. NUMERICAL RESULTS

Computer simulations were conducted to validate the performance of the proposed JCDE receiver in OTFS

TABLE 1. Simulation parameters.

Channel Model	Doubly Selective Fading Channel
Frequency Resource (K)	32
Time Resource (L)	16
CP Length (C)	8
TX filter ($\tilde{\boldsymbol{G}}_{\mathrm{T}}$)	Identity matrix $\boldsymbol{I}_K$
Subcarrier spacing (Δf)	15 kHz
Maximum Doppler Frequency (f_{d})	750 Hz

systems. The simulation parameters are summarized in Table 1.

The number of propagation paths was set to $U = 6$. The time-delay interval for each channel path was assumed to be $\frac{1}{K\Delta f}$, and each path has a different Doppler shift, where the Doppler frequency was assumed to be uniformly distributed within $[-750, 750]$ Hz [47], [48], [49]. The time and frequency synchronizations were assumed to be perfect. Unless otherwise specified, the long-term statistics of the delay and Doppler parameters were assumed to be estimated without error in advance, and thus the long-term statistical structure $\boldsymbol{\Gamma}_u, \forall u$, was assumed to be known. The Gray-coded 4QAM and 16QAM were employed for symbol mapping. The belief damping was applied to all iterative demodulation schemes, where the damping factor η was set to 1.0 at $t = 1$ and 0.5 otherwise. The number of iterations was set to $T = 16$ and 32 for 4QAM and 16QAM, respectively.

A. FRAME STRUCTURE OF OTFS

Without loss of generality, in this paper, we construct the TX symbol vector $\boldsymbol{x}$ in (18) as

$$\boldsymbol{x} = \underbrace{\begin{bmatrix} \boldsymbol{\sigma} \\ \boldsymbol{0}_{M_{\mathrm{d}}\times 1} \end{bmatrix}}_{\triangleq \boldsymbol{x}_{\mathrm{p}}} + \underbrace{\begin{bmatrix} \boldsymbol{0}_{M_{\mathrm{p}}\times 1} \\ \mathcal{X}^{M_{\mathrm{d}}\times 1} \end{bmatrix}}_{\triangleq \boldsymbol{x}_{\mathrm{d}}} \in \mathbb{C}^{M\times 1}, \tag{58}$$

where $\boldsymbol{\sigma} \in \mathbb{C}^{M_{\mathrm{p}}\times 1}$ is the pilot symbol vector, where $|\mathcal{M}_{\mathrm{p}}| = M_{\mathrm{p}}$ is the number of pilot symbols. When employing the Zadoff-chu sequence as the pilot symbol vector, we have

$$[\boldsymbol{\sigma}]_{m,1} = \sqrt{E_{\mathrm{s}}} \exp\left[\mathrm{j}\frac{\pi m(m-1)}{M_{\mathrm{p}}}\right]. \tag{59}$$

B. BER PERFORMANCE IN UNCODED CASE

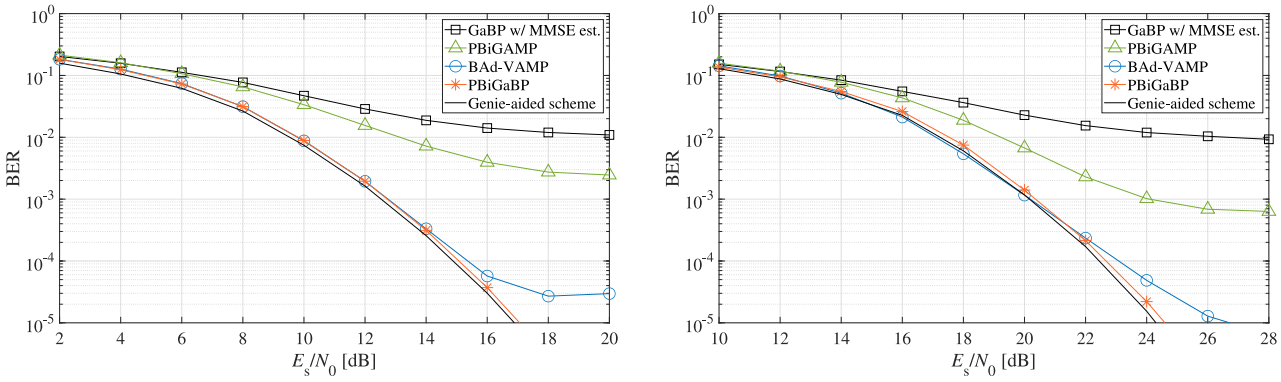
The first set of results is shown in Fig. 3, where the performances in terms of BER as a function of the E_{s}/N_0, of the following uncoded OTFS systems, are compared:

- *GaBP w/ MMSE est:* Two-stage receiver consisting of MMSE filtering-based channel estimation presented in Section II-D and GaBP-based data detection [18]. Provides a reference performance to verify the gain achieved by the JCDE mechanism, which enables the use of tentatively detected data symbols as soft pilots for channel estimation.
- *PBiGAMP:* SotA JCDE receiver based on PBiGAMP [19], where the belief damping is introduced as in [50]. Note that the adaptive damping is not applied.[6] Provide a performance comparison with the most common low-complexity JCDE algorithm for massive PBI.
- *BAd-VAMP:* SotA JCDE receiver based on BAd-VAMP [27], wherein matrix inversion operations are required. The belief damping is introduced as in [27], [52]. Provide a performance comparison with the high-complexity JCDE algorithm, which achieves accurate PBI even with realistic system sizes.
- *PBiGaBP:* Proposed JCDE receiver based on PBiGaBP presented in Algorithm 1.
- *Genie-aided scheme:* Idealized scheme in which the perfect CSI is known at the receiver. Provides an absolute lower bound the proposed method can achieve.

The results in Fig. 4 show the BER performance in OTFS systems with $(K, L) = (32, 16)$, *i.e.*, $N = M = 512$ and $U = 6$. The symbol mapping rules used were Gray-coded 4QAM in Fig. 4(a) and 16QAM in Fig. 4(b), respectively. The number of pilot symbols was set to $M_{\mathrm{p}} = 32$ for 4QAM and $M_{\mathrm{p}} = 64$ for 16QAM, respectively.

The two-stage demodulation scheme, "GaBP w/ MMSE est.," suffers from high-level error floors due to fatal errors in channel estimation using the linear MMSE filter. On the one hand, "PBiGAMP" improves channel estimation accuracy by exploiting the estimated data symbols as the effective pilot symbols and achieves improved data detection performance. However, "PBiGAMP" is unable to sufficiently decouple the self-feedback due to the difference between the actual OTFS system and the large-system limit condition, resulting in a significant performance gap compared to "Genie-aided scheme." On the other hand, "BAd-VAMP," which uses data detection based on VAMP, can significantly improve detection performance compared to "PBiGAMP." However, the noise enhancement due to LS estimation causes the channel estimation accuracy to deteriorate in the early iterations; hence, we can see that the BER performance deteriorates, especially when the pilot rate is low ($M_{\mathrm{p}}/M = 32/512 \approx 0.06$). The most attractive feature is that the proposed "PBiGaBP" can significantly improve the detection performance compared to "PBiGAMP" with the same complexity order and approach the Genie-aided performance without suffering from error floors. Remarkably, the degradation at BER $= 10^{-4}$ is less than 1.0 dB for both system setups.

[6] Adaptive damping enables stable convergence of PBiGAMP, but the number of iterations required for convergence deeply depends on the system parameters [45]. In fact, in order to achieve stable convergence behavior, the number of iterations needs to be set to a fairly large number [51], which is undesirable from an implementation perspective.



(a) Gray-coded 4QAM, $M_{\rm p} = 32$. (b) Gray-coded 16QAM, $M_{\rm p} = 64$

FIGURE 4. BER performance of uncoded OTFS systems with respect to SNR, where $(K, L, U) = (32, 16, 6)$.

C. NMSE PERFORMANCE

Let us shift our focus to the channel estimation accuracy of proposed method to observe the behavior of the JCDE algorithms in more detail. This paper evaluates the channel estimation capability in terms of NMSE performance, where NMSE is defined as NMSE $\triangleq \mathbb{E}_{\mathrm{h}_u}[|h_u - \hat{h}_u|^2]/\mathbb{E}_{\mathrm{h}_u}[|h_u|^2]$. In addition to "PBiGAMP," "BAd-VAMP," and "PBiGaBP," the following performances are compared:

- *MMSE:* Baseline MMSE filtering-based channel estimation presented in Section II-D. Provides a lower bound performance for channel estimation using only pilot symbols in this simulation setup.
- *Genie-aided MMSE:* Idealized scheme in which the channel vector is estimated by MMSE filtering with the perfect knowledge of data symbols (*i.e.*, equivalent to an OTFS frame being composed entirely of pilot symbols). Provides an absolute lower bound in terms of the NMSE performance [24], [37].

Fig. 5 shows the NMSE performance, where the system parameters are the same as those in Fig. 4. As expected from the results in Fig. 4, severe estimation errors occur in "MMSE," indicating that the equivalent noise from the data component in (27) makes it difficult to improve the NMSE performance significantly even in the high $E_{\rm s}/N_0$ region. In addition, while "PBiGAMP" is able to improve the BER performance compared to "GaBP w/ MMSE est." in Fig. 4, Fig. 5 shows an area of degraded performance compared to "MMSE." This is because the convergence behavior of "PBiGAMP" is unstable for realistic system sizes, so channel vector estimates may diverge according to the instantaneous parameter realization, which deteriorates the NMSE performance. As shown in Fig. 5(b), when the pilot resources are sufficient and $E_{\rm s}/N_0$ is high, *i.e.*, the initial convergence of "PBiGAMP" is stable, the divergence behavior is suppressed and the performance improvement from "MMSE" is confirmed. On the other hand, "BAd-VAMP" can improve channel estimation accuracy in the low $E_{\rm s}/N_0$ region due to the LS-based channel estimation and JCDE mechanism; however, the performance degradation from the lower bound is still significant in the high $E_{\rm s}/N_0$ region shown in Fig. 5(b). In contrast, "PBiGaBP" can achieve stable estimation and approach the Genie-aided performance in the high $E_{\rm s}/N_0$ region by exploiting estimated data symbols as soft pilot symbols based on the GaBP approach. For "PBiGaBP" without matrix inversion operations, there is a gap between the performance of "PBiGaBP" and Genie-aided MMSE in Fig. 5(b) because the achievable estimation accuracy is inherently limited by the system size [53]. However, as can be seen from Fig. 4, the effect on the BER performance is small.

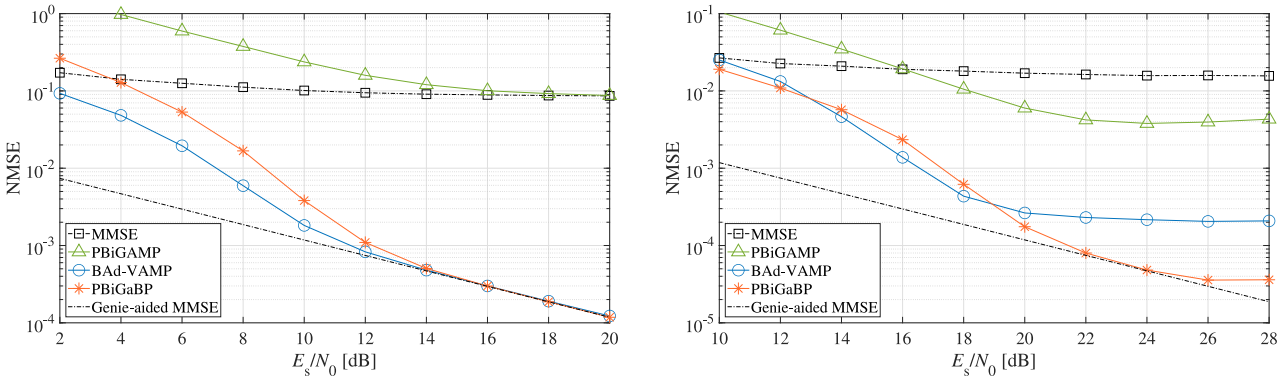
D. ROBUSTNESS TO CHANGES IN PILOT LENGTH

Next, we focus on the robustness of the proposed method to changes in frame configuration. Fig. 6 shows the BER performance as a function of $\kappa = M_{\rm p}/M$ when M is fixed to 512, where the other system parameters are the same as those in Fig. 4. The $E_{\rm s}/N_0$ is fixed at 16 dB and 23 dB in Figs. 6(a) and 6(b), respectively.

In both configurations, the two-stage demodulation scheme "GaBP w/ MMSE est." must have sufficient pilot resources in the OTFS frame to reduce the gap with "Genie-aided performance." In addition, when the system size is insufficient, the estimation accuracy of "PBiGAMP" is highly dependent on the ratio of pilot symbols κ, and when κ is small, it is difficult to make effective use of the estimated data symbols. In contrast, the message update rule of "PBiGaBP," which depends only on the SGA based on CLT, allows it to significantly improve its robustness with respect to pilot length due to its small dependence on the large-system limit approximation. More specifically, "PBiGaBP" approaches the Genie-aided performance with approximately $\kappa = 0.06$ for 4QAM mapping and $\kappa = 0.12$ for 16QAM mapping, respectively, and is expected to increase the system capacity compared to the alternatives.

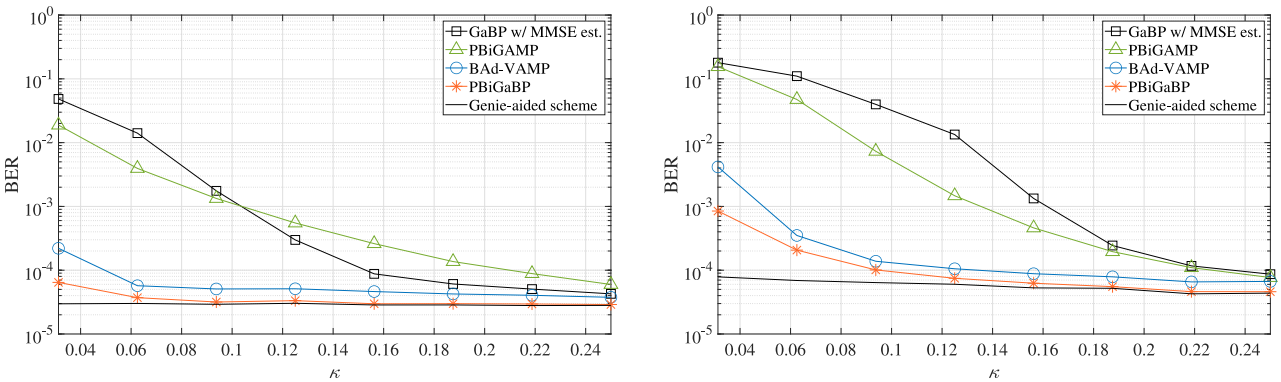
E. ROBUSTNESS TO DOPPLER PARAMETER ERRORS

Up to this point, the delay and Doppler parameters that make up the long-term statistical structure $\boldsymbol{\Gamma}_u, \forall u$, have been assumed to be known, but in practice these parameters are estimated in advance, and errors occur according to



(a) Gray-coded 4QAM, $M_{\mathrm{p}} = 32$. (b) Gray-coded 16QAM, $M_{\mathrm{p}} = 64$.

FIGURE 5. NMSE performance of uncoded OTFS systems with respect to signal-to-noise ratio (SNR), where $(K, L, U) = (32, 16, 6)$.



(a) Gray-coded 4QAM at $E_{\mathrm{s}}/N_0 = 16$ dB. (b) Gray-coded 16QAM at $E_{\mathrm{s}}/N_0 = 23$ dB.

FIGURE 6. BER performance of uncoded OTFS systems with respect to $\kappa = M_{\mathrm{p}}/M$, where $(K, L, U) = (32, 16, 6)$.

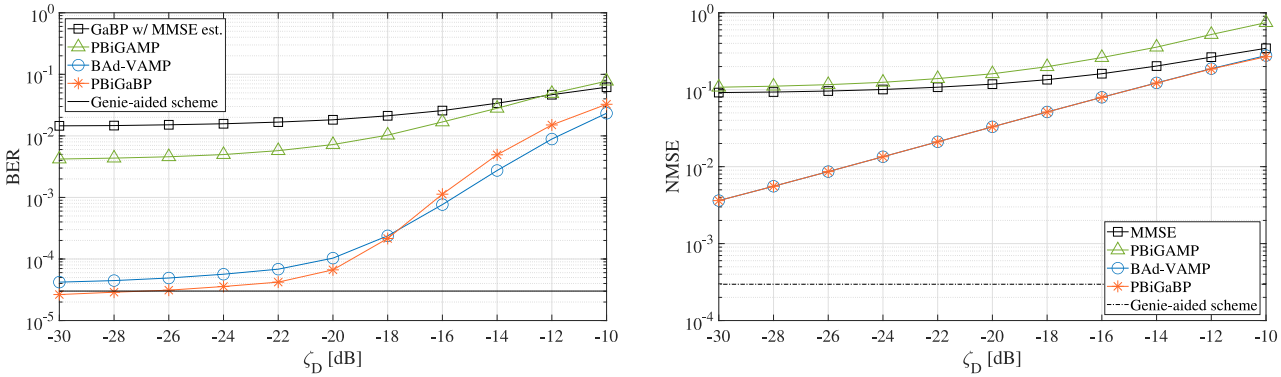
the estimation accuracy. According to [14], the sampling time resolution can be assumed to be sufficiently high to accurately approximate the path delay to the nearest sampling point in wideband wireless communications systems. In fact, the normalized delays are typically assumed to be integer-valued (*i.e.*, no fractional components) in the current wireless communication systems because the sampling frequency is sufficiently high relative to the path delay, and the path delay can be quantized with high resolution [54], [55]. In addition, it is worth noting that it is always possible to increase the time resolution by oversampling [56]. In contrast, since it is generally difficult to obtain sufficient resolution for Doppler shift, the impact of the estimation error of the Doppler parameter becomes the dominant factor affecting the accuracy of the channel estimation and data detection [57], [58]. Therefore, we focus on the impact of the estimation error of the Doppler parameter.

Fig. 7 shows the BER and NMSE performances of OTFS systems as a function of the NMSE of Doppler parameter (frequency) ν_u, *i.e.*,

$$\zeta_{\mathrm{D}} \triangleq \frac{\mathbb{E}_{\nu_u}\left[\left|\nu_u - \hat{\nu}_u\right|^2\right]}{\mathbb{E}_{\nu_u}\left[\left|\nu_u\right|^2\right]}, \tag{60}$$

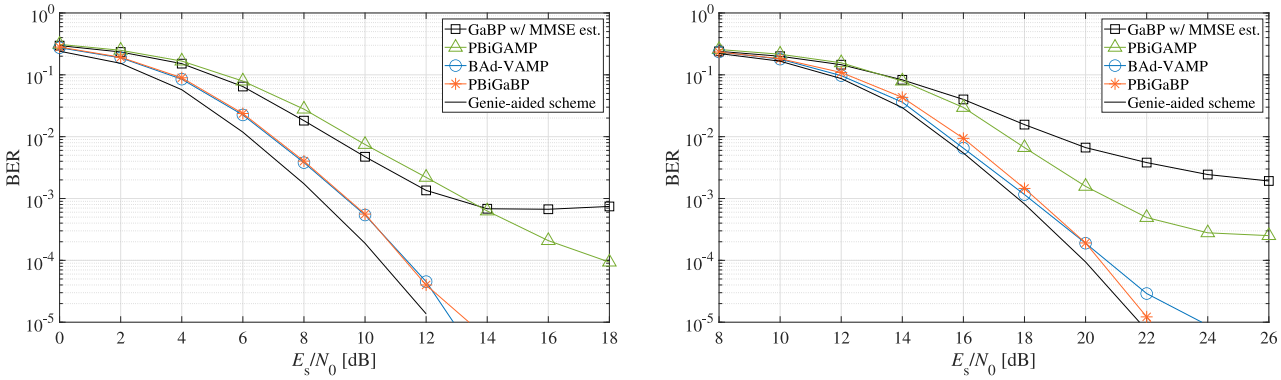
where the estimated parameter is expressed as $\hat{\nu}_u = \nu_u + \tilde{\nu}_u$, $\tilde{\nu}_u \sim \mathcal{CN}(0, N_\nu)$ and N_ν is the variance of estimation error $\tilde{\nu}_u$ corresponding to ζ_{D}. The E_{s}/N_0 is fixed at 16 dB, and the system parameters are the same as those in Figs. 4(a) and 5(a).

From both results, we can see that as the Doppler parameter deviates from the true value, the estimation accuracy of the proposed method and all the benchmark algorithms decreases with the same trend. In Fig. 7(b), the NMSE performance of the channel coefficients for "PBiGaBP" and "BAd-VAMP" decreases linearly with ζ_{D}. This is because the effect of the Doppler parameter error is expressed as Gaussian noise in conformity with CLT, and in the signal model, it can be modeled as extra AWGN. Consequently, Fig. 7(a) shows that "BAd-VAMP," which robustly operates even in the low E_{s}/N_0 region by matrix inversion operation, can slightly improve the detection accuracy compared to the proposed method when the Doppler parameter error is large. The most attractive feature is that the proposed "PBiGaBP" is almost as robust as "BAd-VAMP" against errors in the Doppler parameter and always maintains a significant performance improvement compared to the SotA "PBiGAMP," which implies the efficacy of the proposed method in more realistic scenarios.



(a) BER performances. (b) NMSE performances.

FIGURE 7. BER and NMSE performance of uncoded OTFS systems at $E_s/N_0 = 16$ dB with respect to the NMSE of the Dopper parameter ζ_D, where $(K, L, U, M_p) = (32, 16, 6, 32)$ and Gray-coded 4QAM is used.



(a) Gray-coded 4QAM, $M_p = 32$, coding ratio: $1/2$. (b) Gray-coded 16QAM, $M_p = 64$, coding ratio: $2/3$.

FIGURE 8. BER performance of coded OTFS systems with respect to SNR, where $(K, L, U) = (32, 16, 6)$.

F. BER PERFORMANCE IN CODED CASE

Assuming that the long-term statistical structure is known again, we demonstrate the validity of the proposed method in channel-coded systems. Fig. 8 shows the BER performance in the channel-coded system, where the other system parameters are the same as those in Fig. 4. Using low-density parity-check (LDPC) code that conforms to the 5G new radio (NR) specification, the coding rate was set to $1/2$ for 4QAM and $2/3$ for 16QAM, respectively. The SPA is used as the soft decision decoder, and error correction by the channel decoder is performed only once after demodulation.

As shown in Fig. 8, due to poor channel estimation accuracy, "GaBP w/ MMSE est." and "PBiGAMP" are unable to fully demonstrate the error correction capability, and as a result, a decrease in accuracy is inevitable. In contrast, "PBiGaBP" achieved a BER of 10^{-5} in both configurations, and the performance degradation from the "Genie-aided scheme" is suppressed to around 1.0 dB at BER $= 10^{-4}$. The slight deviation from "Genie-aided scheme" is due to the LLR calculated from the estimated channel deviating from the ideal LLR assumed by the SPA.

G. COMPLEXITY ANALYSIS

Finally, the computational cost of each JCDE algorithm is evaluated in terms of the number of real multiplication operations required to demodulate data symbols. To evaluate the approximate number of real multiplication operations, we adopt the following basic assumptions: [59]

- A multiplication of a complex matrix $\boldsymbol{A}_1 \in \mathbb{C}^{d_1 \times d_2}$ with a complex matrix $\boldsymbol{A}_2 \in \mathbb{C}^{d_2 \times d_3}$ requires $d_1 d_2 d_3$ multiplications (of complex numbers).
- An inversion of a complex matrix $\boldsymbol{A}_3 \in \mathbb{C}^{d_1 \times d_1}$ requires approximately $2d_1^3 - d_1^2$ multiplications and d_1^2 divisions (of complex numbers).
- A division of complex numbers requires 1 complex multiplication (division and multiplication are assumed to have the same complexity).
- A complex multiplication requires 4 real multiplications.

Fig. 9 shows the number of real multiplication operations as a function of the number of frequency resources K, where the compression ratio is fixed at $\kappa = M_p/M = 1/8$, and the other system parameters are the same as in Fig. 4(b). First, "BAd-VAMP" with matrix inversion operations requires $\mathcal{O}(N^3)$ complexity per iteration, which significantly increases the computational cost compared to "PBiGAMP". In contrast, all the computations of the PBiGaBP algorithm, *i.e.*, the number of multiplication, division, subtraction, and addition operations, are scalar-by-scalar; therefore, the computational

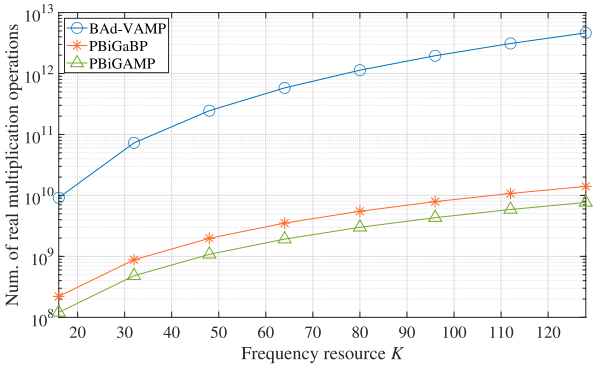


FIGURE 9. The approximate number of real multiplication operations as a function of the frequency resource K.

complexity of our proposed JCDE algorithm is of order $\mathcal{O}(MNU)$ per iteration, which is similar to that of the conventional PBiGAMP-based method.

This fact can be confirmed by the results in Fig. 9, which shows that the proposed method can improve estimation accuracy without significantly increasing the computational cost. More specifically, at $K = 64$, "PBiGaBP" can operate at about twice the computational cost of "PBiGAMP". It should be noted that although the complexity order of PBiGAMP and PBiGaBP is similar, PBiGAMP, which can reduce the number of beliefs propagated on the FG by the large-system limit approximation, achieves a lower computational cost from the perspective of more practical processing costs. However, considering that PBiGAMP cannot operate properly at real system sizes and is extremely vulnerable to channel correlation, and that PBiGaBP greatly improves performance and can achieve the BER performance close to the lower reference in many scenarios, we can conclude that the proposed method can achieve an excellent trade-off between estimation performance and computational cost.

V. CONCLUSION

We proposed a novel low-complexity JCDE algorithm via PBiGaBP for OTFS transmission systems over doubly-selective fading channels. The proposed method does not require matrix operations; thus, it has linear computational complexity with respect to N, M, and U, respectively, *i.e.*, $\mathcal{O}(NMU)$. The JCDE for the OTFS demodulation is reformulated as the PBI problem, and then the message-passing rule is designed based on the GaBP framework, which depends only on the SGA based on CLT. Simulation results demonstrate the efficacy of the proposed method in terms of BER and NMSE performances in various scenarios. The results revealed that the proposed scheme significantly outperforms the SotA counterparts and asymptotically approaches the performance of the idealized Genie-aided scheme even when the pilot ratio in the OTFS frame is low.

REFERENCES

[1] Y. Liu, C.-X. Wang, J. Huang, J. Sun, and W. Zhang, "Novel 3-D nonstationary mmWave massive MIMO channel models for 5G high-speed train wireless communications," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 2077–2086, Mar. 2019.

[2] Z. Wei et al., "Orthogonal time-frequency space modulation: A promising next-generation waveform," *IEEE Wireless Commun.*, vol. 28, no. 4, pp. 136–144, Aug. 2021.

[3] H. Chang et al., "A novel nonstationary 6G UAV-to-ground wireless channel model with 3-D arbitrary trajectory changes," *IEEE Internet Things J.*, vol. 8, no. 12, pp. 9865–9877, Jun. 2021.

[4] F. Hlawatsch and G. Matz, *Wireless Communications Over Rapidly Time-Varying Channels*, 1st ed. Cambridge, MA, USA: Academic, 2011.

[5] Y. Zhao and S.-G. Haggman, "Sensitivity to Doppler shift and carrier frequency errors in OFDM systems-the consequences and solutions," in *Proc. Veh. Technol. Conf. (VTC)*, 1996, pp. 1564–1568.

[6] T. Wang, J. G. Proakis, E. Masry, and J. R. Zeidler, "Performance degradation of OFDM systems due to Doppler spreading," *IEEE Trans. Wireless Commun.*, vol. 5, no. 6, pp. 1422–1432, Jun. 2006.

[7] V. Vahidi and E. Saberinia, "OFDM high speed train communication systems in 5G cellular networks," in *Proc. 15th IEEE Annu. Consum. Commun. Netw. Conf. (CCNC)*, 2018, pp. 1–6.

[8] R. Hadani et al., "Orthogonal time frequency space modulation," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, 2017, pp. 1–6.

[9] G. D. Surabhi, R. M. Augustine, and A. Chockalingam, "On the diversity of uncoded OTFS modulation in doubly-dispersive channels," *IEEE Trans. Wireless Commun.*, vol. 18, no. 6, pp. 3049–3063, Jun. 2019.

[10] 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.

[11] S. Srivastava, R. K. Singh, A. K. Jagannatham, A. Chockalingam, and L. Hanzo, "OTFS transceiver design and sparse doubly-selective CSI estimation in analog and hybrid beamforming aided mmWave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 21, no. 12, pp. 10902–10917, Dec. 2022.

[12] M. K. Ramachandran and A. Chockalingam, "MIMO-OTFS in high-Doppler fading channels: Signal detection and channel estimation," in *Proc. IEEE GLOBECOM*, 2018, pp. 206–212.

[13] 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.

[14] H. B. Mishra, P. Singh, A. K. Prasad, and R. Budhiraja, "OTFS channel estimation and data detection designs with superimposed pilots," *IEEE Trans. Wireless Commun.*, vol. 21, no. 4, pp. 2258–2274, Apr. 2022.

[15] N. Hashimoto, N. Osawa, K. Yamazaki, and S. Ibi, "Channel estimation and equalization for CP-OFDM-based OTFS in fractional Doppler channels," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, 2021, pp. 1–7.

[16] R. Ouchikh, A. Aïssa-El-Bey, T. Chonavel, and M. Djeddou, "Iterative channel estimation and data detection algorithm for OTFS modulation," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, 2022, pp. 5263–5267.

[17] Y. Kabashima, "A CDMA multiuser detection algorithm on the basis of belief propagation," *J. Phy. A, Math. General*, vol. 36, no. 43, pp. 11111–11121, Oct. 2003.

[18] T. Takahashi, S. Ibi, and S. Sampei, "Design of adaptively scaled belief in multi-dimensional signal detection for higher-order modulation," *IEEE Trans. Commun.*, vol. 67, no. 3, pp. 1986–2001, Mar. 2019.

[19] J. T. Parker and P. Schniter, "Parametric bilinear generalized approximate message passing," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 4, pp. 795–808, Jun. 2016.

[20] P. Raviteja, K. T. Phan, Y. Hong, and E. Viterbo, "Interference cancellation and iterative detection for orthogonal time frequency space modulation," *IEEE Trans. Wireless Commun.*, vol. 17, no. 10, pp. 6501–6515, Oct. 2018.

[21] X. Li and W. Yuan, "OTFS detection based on gaussian mixture distribution: A generalized message passing approach," *IEEE Commun. Lett.*, vol. 28, no. 1, pp. 178–182, Jan. 2024.

[22] S. Rangan, "Generalized approximate message passing for estimation with random linear mixing," in *Proc. IEEE Int. Symp. Inf. Theory*, 2011, pp. 2168–2172.

[23] D. L. Donoho, A. Maleki, and A. Montanari, "Message-passing algorithms for compressed sensing," *Proc. Nat. Acad. Sci.*, vol. 106, no. 45, pp. 18914–18919, 2009.

[24] M. Bayati and A. Montanari, "The dynamics of message passing on dense graphs, with applications to compressed sensing," 2011, *arXiv:1001.3448*.

[25] P. Sun, Z. Wang, and P. Schniter, "Joint channel-estimation and equalization of single-carrier systems via bilinear AMP," *IEEE Trans. Signal Process.*, vol. 66, no. 10, pp. 2772–2785, May 2018.

[26] Y. Ma, N. Wu, J. A. Zhang, B. Li, and L. Hanzo, "Parametric bilinear iterative generalized approximate message passing reception of FTN multi-carrier signaling," *IEEE Trans. Commun.*, vol. 69, no. 12, pp. 8443–8458, Dec. 2021.

[27] S. Sarkar, A. K. Fletcher, S. Rangan, and P. Schniter, "Bilinear recovery using adaptive vector-AMP," *IEEE Trans. Signal Process.*, vol. 67, no. 13, pp. 3383–3396, Jul. 2019.

[28] X. Meng and J. Zhu, "Bilinear adaptive generalized vector approximate message passing," *IEEE Access*, vol. 7, pp. 4807–4815, 2019.

[29] A. Fletcher, M. Sahraee-Ardakan, S. Rangan, and P. Schniter, "Expectation consistent approximate inference: Generalizations and convergence," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, 2016, pp. 190–194.

[30] S. Rangan, P. Schniter, and A. K. Fletcher, "Vector approximate message passing," *IEEE Trans. Inf. Theory*, vol. 65, no. 10, pp. 6664–6684, Oct. 2019.

[31] K. Takeuchi, "Rigorous dynamics of expectation-propagation-based signal recovery from unitarily invariant measurements," *IEEE Trans. Inf. Theory*, vol. 66, no. 1, pp. 368–386, Jan. 2020.

[32] X. Yang, H. Li, Q. Guo, J. A. Zhang, X. Huang, and Z. Cheng, "Sensing aided uplink transmission in OTFS ISAC with joint parameter association, channel estimation and signal detection," *IEEE Trans. Veh. Technol.*, vol. 73, no. 6, pp. 9109–9114, Jun. 2024.

[33] K. Furuta, T. Takahashi, K. Ito, and S. Ibi, "Joint channel and data estimation via Bayesian parametric bilinear inference for OTFS transmission," in *Proc. IEEE 20th Consum. Commun. Netw. Conf. (CCNC)*, 2024, pp. 887–892.

[34] R. Tamaki, K. Ito, T. Takahashi, S. Ibi, and S. Sampei, "Suppression of self-noise feedback in GAMP for highly correlated large MIMO detection," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2022, pp. 1300–1305.

[35] K. Ito, T. Takahashi, S. Ibi, and S. Sampei, "Bilinear gaussian belief propagation for massive MIMO detection with non-orthogonal pilots," *IEEE Trans. Commun.*, vol. 72, no. 2, pp. 1045–1061, Feb. 2024.

[36] T. Takahashi, H. Iimori, K. Ishibashi, S. Ibi, and G. T. F. De Abreu, "Bayesian bilinear inference for joint channel tracking and data detection in millimeter-wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 23, no. 9, pp. 11136–11153, Sep. 2024.

[37] T. Takahashi, H. Iimori, K. Ando, K. Ishibashi, S. Ibi, and G. T. F. De Abreu, "Bayesian receiver design via bilinear inference for cell-free massive MIMO with low-resolution ADCs," *IEEE Trans. Wireless Commun.*, vol. 22, no. 7, pp. 4756–4772, Jul. 2023.

[38] H. Iimori, T. Takahashi, K. Ishibashi, G. T. F. De Abreu, D. G. González, and O. Gonsa, "Joint activity and channel estimation for extra-large MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 21, no. 9, pp. 7253–7270, Sep. 2022.

[39] H. Iimori, T. Takahashi, K. Ishibashi, G. T. F. De Abreu, and W. Yu, "Grant-free access via bilinear inference for cell-free MIMO with low-coherence pilots," *IEEE Trans. Wireless Commun.*, vol. 20, no. 11, pp. 7694–7710, Nov. 2021.

[40] Y. Yue, J. Shi, Z. Li, J. Hu, and Z. Tie, "Model-driven deep learning assisted detector for OTFS with channel estimation error," *IEEE Commun. Lett.*, vol. 28, no. 4, pp. 842–846, Apr. 2024.

[41] G. Matz, "On non-WSSUS wireless fading channels," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2465–2478, Sep. 2005.

[42] W. Yuan, Z. Wei, S. Li, J. Yuan, and D. W. K. Ng, "Integrated sensing and communication-assisted orthogonal time frequency space transmission for vehicular networks," *IEEE J. Sel. Topics Signal Process.*, vol. 15, no. 6, pp. 1515–1528, Nov. 2021.

[43] 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.

[44] T. Takahashi, S. Ibi, and S. Sampei, "Design of criterion for adaptively scaled belief in iterative large MIMO detection," *IEICE Trans. Commun.*, vol. E102-B, no. 2, pp. 285–297, 2019.

[45] J. T. Parker, P. Schniter, and V. Cevher, "Bilinear generalized approximate message passing—Part I: Derivation," *IEEE Trans. Signal Process.*, vol. 62, no. 22, pp. 5839–5853, Nov. 2014.

[46] A. Chockalingam and B. S. Rajan, *Large MIMO Systems*. New York, NY, USA: Cambridge Univ. Press, 2014.

[47] Y. Liu, Y. L. Guan, and D. G. González, "BEM OTFS receiver with superimposed pilots over channels with Doppler and delay spread," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2022, pp. 2411–2416.

[48] C. Yang, J. Wang, Z. Pan, and S. Shimamoto, "Delay-Doppler frequency domain-aided superimposing pilot OTFS channel estimation based on deep learning," in *Proc. IEEE 96th Veh. Technol. Conf. (VTC)*, 2022, pp. 1–6.

[49] S. R. Mattu and A. Chockalingam, "Learning based delay-Doppler channel estimation with interleaved pilots in OTFS," in *Proc. IEEE 96th Veh. Technol. Conf. (VTC)*, 2022, pp. 1–6.

[50] J. T. Parker, P. Schniter, and V. Cevher, "Bilinear generalized approximate message passing—Part II: Applications," *IEEE Trans. Signal Process.*, vol. 62, no. 22, pp. 5854–5867, Nov. 2014.

[51] T. Fujitsuka and K. Takeuchi, "Pilot decontamination in massive MIMO uplink via approximate message-passing," *IEICE Trans. Fundam. Electron., Commun. Comput. Sci.*, vol. E103-A, no. 12, pp. 1356–1366, 2020.

[52] J. Céspedes, P. M. Olmos, M. Sánchez-Fernández, and F. Perez-Cruz, "Expectation propagation detection for high-order high-dimensional MIMO systems," *IEEE Trans. Commun.*, vol. 62, no. 8, pp. 2840–2849, Aug. 2014.

[53] R. Hayakawa and K. Hayashi, "Discreteness-aware approximate message passing for discrete-valued vector reconstruction," *IEEE Trans. Signal Process.*, vol. 66, no. 24, pp. 6443–6457, Dec. 2018.

[54] Y. Hong, T. Thaj, and E. Viterbo, *Delay-Doppler Communications: Principles and Applications*. Amsterdam, The Netherlands: Academic, 2022.

[55] A. Bemani, N. Ksairi, and M. Kountouris, "Affine frequency division multiplexing for next generation wireless communications," *IEEE Trans. Wireless Commun.*, vol. 22, no. 11, pp. 8214–8229, Nov. 2023.

[56] P. Priya, E. Viterbo, and Y. Hong, "Low complexity MRC detection for OTFS receiver with oversampling," *IEEE Trans. Wireless Commun.*, vol. 23, no. 2, pp. 1459–1473, Feb. 2024.

[57] S. P. Muppaneni, S. R. Mattu, and A. Chockalingam, "Channel and radar parameter estimation with fractional delay-Doppler using OTFS," *IEEE Commun. Lett.*, vol. 27, no. 5, pp. 1392–1396, May 2023.

[58] F. Liu, Z. Yuan, Q. Guo, Z. Wang, and P. Sun, "Message passing-based structured sparse signal recovery for estimation of OTFS channels with fractional Doppler shifts," *IEEE Trans. Wireless Commun.*, vol. 20, no. 12, pp. 7773–7785, Dec. 2021.

[59] S. Yoon and C.-B. Chae, "Low-complexity MIMO detection based on belief propagation over pairwise graphs," *IEEE Trans. Veh. Technol.*, vol. 63, no. 5, pp. 2363–2377, Jun. 2014.

KENTA ITO (Graduate Student Member, IEEE) received the B.E. degree in electrical engineering from Toyama University, Japan, in 2019, and the M.E. degree in communication engineering from Osaka University, Japan, in 2021, where he is currently pursuing the Ph.D. degree with the Graduate School of Engineering. His research interests include belief propagation, compressed sensing, signal processing, and wireless communications.

TAKUMI TAKAHASHI (Member, IEEE) received the B.E., M.E., and Ph.D. degrees in communication engineering from Osaka University, Osaka, Japan, in 2016, 2017, and 2019, respectively. From 2018 to 2019, he was a Visiting Researcher with the Centre for Wireless Communications, University of Oulu, Finland. In 2019, he joined the Graduate School of Engineering, Osaka University as an Assistant Professor. His current research interests include Bayesian inference, belief propagation, signal processing, and wireless communications. He received the 80th Best Paper Award from IEICE, and the 2019 and 2023 Best Paper Awards from IEICE Communication Society. He was certified as an Exemplary Reviewer of IEEE WIRELESS COMMUNICATIONS LETTERS in 2023.

KENGO FURUTA (Graduate Student Member, IEEE) received the B.E. and M.E. degrees in communication engineering from Osaka University, Osaka, Japan, in 2022 and 2024 respectively, where he is currently pursuing the Ph.D. degree with the Graduate School of Engineering. His research interests include sparse Bayesian learning, compressed sensing, sparse signal reconstruction, and wireless communication.

SHINSUKE IBI (Senior Member, IEEE) received the B.E. degree in advanced engineering from the Suzuka College of Technology, Japan, in 2002, and the M.E. and Ph.D. degrees in communication engineering from Osaka University, Japan, in 2004 and 2006, respectively. From 2005 to 2006, he was a Visiting Researcher with the Centre for Wireless Communications, University of Oulu, Finland. In 2006, he joined the Graduate School of Engineering, Osaka University. From 2010 to 2011, he was a Visiting Researcher with the University of Southampton, U.K. He moved to Doshisha University in 2019, where he is currently a Professor with the Faculty of Science and Engineering. His research interests include EXIT-based coding theory, iterative detection, digital signal processing, cognitive radio, and communication theory. He received the 64th, 71st, and 80th Best Paper Awards from IEICE, the 2017, 2018, 2019, and 2023 Best Paper Awards from IEICE Communication Society, and the 24th Telecom System Technology Award from the Telecommunication Advancement Foundation.

GIUSEPPE THADEU FREITAS DE ABREU (Senior Member, IEEE) received the B.Eng. degree in electrical engineering and the specialization (Latu Sensu) degree in telecommunications engineering from the Universidade Federal da Bahia, Salvador, Brazil, in 1996 and 1997, respectively, and the M.Eng. and D.Eng. degrees in physics, electrical and computer engineering from Yokohama National University, Japan, in March 2001 and March 2004, respectively. He was a Postdoctoral Fellow and later an Adjunct Professor (Docent) of Statistical Signal Processing and Communications Theory with the Department of Electrical and Information Engineering, University of Oulu, Finland, from 2004 to 2006 and from 2006 to 2011, respectively. Since 2011, he has been a Professor of Electrical Engineering with Jacobs University Bremen, Germany. From April 2015 to August 2018, he was a Full Professor with the Department of Computer and Electrical Engineering, Ritsumeikan University, Japan. His research interests include communications and signal processing, including communications theory, estimation theory, statistical modeling, wireless localization, cognitive radio, wireless security, MIMO systems, ultrawideband and millimeter wave communications, full-duplex and cognitive radio, compressive sensing, energy harvesting networks, random networks, connected vehicles networks, and joint communications and sensing. He was a recipient of the Uenohara Award by Tokyo University in 2000 for his master's thesis work. He was a co-recipient of best paper awards at several international conferences, and was awarded the JSPS, Heiwa Nakajima, and NICT Fellowships (twice) in 2010, 2013, 2015, and 2018, respectively. He served as an Associate Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2009 to 2014 and for the IEEE TRANSACTIONS ON COMMUNICATIONS from 2014 to 2017. He was an Executive Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2018 to 2021 and an Editor of the IEEE COMMUNICATIONS LETTERS from 2021 to 20124. He has also been an Editor to the IEEE SIGNAL PROCESSING LETTERS since 2021.