Delay-Doppler and Angular Domain 4D-Sparse CSI Estimation in OTFS Aided MIMO Systems

Suraj Srivastava, *Member, IEEE,* Rahul Kumar Singh, Aditya K. Jagannatham *Member, IEEE* and Lajos Hanzo *Life Fellow, IEEE*

***Abstract*—A convenient delay, Doppler and angular-(DDA) domain representation of the multiple-input multiple-output (MIMO) wireless channel is conceived for deriving the end to end relationship in the delay-Doppler (DD)-domain for orthogonal time frequency space (OTFS)-based communications. Subsequently, a time-domain pilot based model is developed for estimating the DDA-domain channel state information (CSI) of our MIMO OTFS system. The key differentiating feature of the CSI estimation model derived is its ability to exploit the 4-dimensional (4D)-sparsity arising in the DDA-domain, given the limited number of dominant scatterers. Furthermore, the training overhead of the proposed framework is low, and the pilot placement is quite flexible, necessitating no guard-interval. Finally, an orthogonal matching pursuit (OMP) framework is employed for 4D-sparse CSI acquisition, followed by deriving the Oracle minimum mean squared error (Oracle-MMSE) and its Bayesian Cramer-Rao lower bound (BCRLB). Our simulation results confirm the improved CSI estimation performance attained over the benchmarks.**

***Index Terms*—Angular sparsity, channel estimation, OTFS, delay-Doppler, BCRLB, high-mobility**

I. Introduction

Orthogonal frequency division multiplexing (OFDM) is the dominant physical layer waveform in the operational 4G/ 5G cellular systems. Although immensely successful at attaining high data rates, it can lead to severe performance degradation in high-mobility use cases [1]–[3]. This can be attributed to the inter-carrier-interference arising due to the excessive Doppler shifts arising in the face of high mobile velocities and carrier frequencies [4]. Thus, it is imperative to conceive delay-Doppler (DD)-domain resilient modulation techniques. As a benefit of its enhanced performance in high Doppler wireless channels, orthogonal time frequency space (OTFS) signalling [1], [5]–[8], has gained popularity over conventional multicarrier modulation techniques. However, accurate knowledge of the DD-domain channel state information (CSI) is the key to achieving a superior performance in OTFS systems [9], [10]. Therefore, novel CSI estimation techniques specifically tailored for achieving a high CSI accuracy at a low pilot overhead are required for reaping the advantages of OTFS. A brief description of the state-of-the-art is given next.

A. K. Jagannatham would like to acknowledge the research supported in part by the Science and Engineering Research Board (SERB), Department of Science and Technology, Government of India, in part by the Space Technology Cell, IIT Kanpur, in part by the IIMA IDEA Telecom Centre of Excellence, in part by the Qualcomm Innovation Fellowship, and in part by the Arun Kumar Chair Professorship. L. Hanzo would like to acknowledge the financial support of the Engineering and Physical Sciences Research Council projects EP/W016605/1 and EP/P003990/1 (COALESCE) as well as of the European Research Council's Advanced Fellow Grant QuantCom (Grant No. 789028)

S. Srivastava, R. K. Singh and A. K. Jagannatham are with the Department of Electrical Engineering, Indian Institute of Technology Kanpur, Kanpur, UP 208016, India (e-mail: ssrivast@iitk.ac.in, rahulkz@iitk.ac.in, adityaj@iitk.ac.in).

L. Hanzo is with the School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, U.K. (e-mail: lh@ecs.soton.ac.uk).

A. Literature Survey

The existing treatises, which address channel estimation in OTFS systems, can be broadly classified into two categories: 1) Approaches that exploit the 2-dimensional (2D)-circular convolution based relationship [1], [9], [11] between the transmit symbols and the channel in the DD-domain; 2) Alternative schemes that benefit from the limited number of dominant multipath components, i.e., the DD-domain sparsity [10], [12], [13]. The initial solutions of [1], [11] developed impulse-based CSI estimation techniques for SISO OTFS systems, which were then extended to MIMO OTFS in [14]. These schemes employ a complete OTFS frame for CSI estimation, which is their key disadvantage. Subsequently, Raviteja *et al.* [9] developed an embedded pilot (EP) aided channel estimator, where the training and data symbols in the DD-domain are separated via suitable guard intervals. However, these schemes do not leverage the DD-domain sparsity. Furthermore, they need both a high pilot signal-to-noise ratio (SNR) and high pilot overheads for estimating the CSI accurately. By contrast, only a few recent authors [10], [12], [13] have formulated the CSI estimation problem as a compressive sensing (CS) problem, which also exploit the DD-domain sparsity for attaining an enhanced CSI accuracy at low pilot overheads. Furthermore, as described in [10], [15], [16], the MIMO wireless channel also exhibits sparsity in the angular domain due to the presence of only a few non-negligible angles of arrival (AoAs) and departure (AoDs). However, to the best of our knowledge, the existing OTFS treatises have not exploited this 4D-sparsity arising in the delay, Doppler and angular (DDA)-domain. This motivates us to derive the DD-domain input-output system and CSI estimation models, which exploit this 4D-sparsity for improving the existing contributions. Our novel contributions are presented next.

B. Contributions

- Considering a compact DDA-domain representation of the MIMO wireless channel and generalized transceiver

pulse shaping filters, an input-output model is derived for MIMO OTFS systems.
- Subsequently, we develop a CSI estimation model for efficiently exploiting the 4D-sparsity, where time-domain pilot symbols are transmitted for estimating the DDA-domain CSI. Then, an orthogonal matching pursuit (OMP) framework has been employed for acquiring the 4D-sparse CSI.
- The unique feature of our proposed CSI estimation framework is that its pilot placement is highly flexible leading to a significantly reduced pilot overhead.
- Furthermore, the Oracle minimum mean squared error (Oracle-MMSE) and Bayesian Cramer-Rao lower bound (BCRLB) benchmarks are also derived for the proposed estimation framework.

The recent contribution [17] suggested a novel integrated sensing and communication transmission architecture based on spatially-spread OTFS modulation, where spatial-spreading/ de-spreading operations facilitate angular-domain discretization. The key similarities between the proposed work and [17] is that both exploit DDA-domain channel modeling and angular-domain sparsity for cyclic prefix (CP)-aided OTFS systems. However, they are fundamentally different in the following aspects. The proposed solution considers a conventional point-to-point MIMO OTFS system from a communication perspective, where the AoAs and AoDs of the multipath components are different. This results in 4D-sparsity in the DDA-domain. By contrast, [17] considers a spatially-spread OTFS-based co-located radar/ base station and single antenna users for joint sensing and communication, where the AoAs and AoDs of the multipath components are identical for radar. Furthermore, [17] also considers a large antenna array at the base station and beam-tracking from a radar perspective, while our design does not consider these aspects, since our focus is on pure communication.

C. Notation

Boldface lower case and upper case letters denote column vectors and matrices, respectively. The quantity $\mathrm{diag}(a_0, a_1, \cdots, a_{N-1})$ represents a diagonal matrix having the principal diagonal elements given by $a_0, a_1, \cdots, a_{N-1}$, and $\mathbf{I}_N$ denotes the $N \times N$ identity matrix. Superscripts $\mathbf{A}^T$, $\mathbf{A}^H$, $\mathbf{A}^*$ and $\mathbf{A}^{-1}$ denote the transpose, Hermitian, conjugation and inverse respectively. The vector equivalent of the matrix $\mathbf{A}$ is denoted by $\mathrm{vec}(\mathbf{A})$, which is formed by stacking the columns to form a single column vector. $\otimes$ denotes the Kronecker product of two matrices. $\mathcal{U}(a, b)$ denotes uniform distribution between a and b.

II. MIMO OTFS System Model

A. OTFS Modulation

Let M and N denote the number of symbols placed on the delay and Doppler axes, respectively. Furthermore, let the subcarrier spacing of the underlying multicarrier modulation be denoted by Δf, whereas T represent the symbol duration, so that $T\Delta f = 1$. Consider a MIMO OTFS system having N_t transmit antennas (TAs) and N_r receive antennas (RAs). Let $\mathbf{X}_{\mathrm{DD},t} \in \mathbb{C}^{M\times N}$ represent the DD-domain information symbol matrix corresponding to the tth TA. The corresponding time-frequency (TF)-domain symbol matrix $\mathbf{X}_{\mathrm{TF},t} \in \mathbb{C}^{M\times N}$ is obtained as $\mathbf{X}_{\mathrm{TF},t} = \mathbf{F}_M \mathbf{X}_{\mathrm{DD},t} \mathbf{F}_N^H$ [18], where $\mathbf{F}_M$ and $\mathbf{F}_N$ are the normalized discrete Fourier transform (DFT) matrices of order M and N, respectively. Let $p_{\mathrm{tx}}(t)$ be the pulse shaping filter response of duration T at the transmitter. The transmit signal matrix $\mathbf{S}_{\mathrm{T},t} \in \mathbb{C}^{M\times N}$ in the time domain is given by $\mathbf{S}_{\mathrm{T},t} = \mathbf{P}_{\mathrm{tx}} \mathbf{F}_M^H \mathbf{X}_{\mathrm{TF},t} = \mathbf{P}_{\mathrm{tx}} \mathbf{X}_{\mathrm{DD},t} \mathbf{F}_N^H$, where $\mathbf{P}_{\mathrm{tx}} = \mathrm{diag}\left\{p_{\mathrm{tx}}\left(\frac{qT}{M}\right)\right\}_{q=0}^{M-1} \in \mathbb{C}^{M\times M}$. Finally, these MN-samples of the transmit signal matrix $\mathbf{S}_{\mathrm{T},t}$ are vectorized as

$$\mathbf{s}_{\mathrm{T},t} = \mathrm{vec}\left(\mathbf{S}_{\mathrm{T},t}\right) = \left(\mathbf{F}_N^H \otimes \mathbf{P}_{\mathrm{tx}}\right) \mathbf{x}_{\mathrm{DD},t}, \tag{1}$$

where the symbol vector $\mathbf{x}_{\mathrm{DD},t}$ obeys $\mathbf{x}_{\mathrm{DD},t} = \mathrm{vec}\left(\mathbf{X}_{\mathrm{DD},t}\right)$.

B. DDA-domain Channel Model

The DDA-domain wireless channel $\mathbf{H}(\tau, \nu, \theta, \phi) \in \mathbb{C}^{N_r\times N_t}$ can be formulated as [4], [15], [16]

$$\mathbf{H}(\tau, \nu, \theta, \phi) = \sum_{p=1}^{L_p} \alpha_p \mathbf{a}_r\left(\theta\right) \mathbf{a}_t^H\left(\phi\right) \delta(\tau - \tau_p)\delta(\nu - \nu_p) \times \delta(\theta - \theta_p)\delta(\phi - \phi_p), \tag{2}$$

where $\delta(\cdot)$ represents the Dirac-delta function, the 5-tuple $(\alpha_p, \tau_p, \nu_p, \theta_p, \phi_p)$ signifies the complex-valued path gain, delay, Doppler, AoA and AoD of the pth multipath component, and L_p is the number of dominant multipath components. Note that due to the presence of only a few dominant reflectors, L_p is typically very small. The vectors $\mathbf{a}_r\left(\theta_p\right) \in \mathbb{C}^{N_r\times 1}$ and $\mathbf{a}_t\left(\phi_p\right) \in \mathbb{C}^{N_t\times 1}$ represent the array steering vectors at the RAs and TAs corresponding to the AoA θ_p and AoD ϕ_p as described in [4], [15], [16]. For convenience, one can denote the MIMO channel $\mathbf{H}_p \in \mathbb{C}^{N_r\times N_t}$ corresponding to the pth multipath component as

$$\mathbf{H}_p = \alpha_p \mathbf{a}_r\left(\theta_p\right) \mathbf{a}_t^H\left(\phi_p\right), \tag{3}$$

and its (r, t)th element as $h_{p,r,t}$. Furthermore, by appropriately selecting M and N, the DD-domain parameters can be closely approximated as $\tau_p = \frac{i_p}{M\Delta f}, \nu_p = \frac{j_p}{NT}$, where the indices $i_p (<< M)$ and $j_p (<< N)$ are integers. A CP of L time-domain samples is added to the transmit signal $\mathbf{s}_{\mathrm{T},t}$ of (1). On removal of the CP from the output at the rth RA, the time-domain signal $\mathbf{y}_{\mathrm{T},r}$ comprised of MN-samples is expressed as [18]

$$\mathbf{y}_{\mathrm{T},r} = \sum_{t=1}^{N_t} \mathbf{H}_{r,t} \mathbf{s}_{\mathrm{T},t} + \mathbf{w}_r, \tag{4}$$

where $\mathbf{w}_r \in \mathbb{C}^{MN\times 1}$ represents the complex additive white Gaussian noise of mean zero and variance σ^2, and $\mathbf{H}_{r,t} \in \mathbb{C}^{MN\times MN}$ is expressed as

$$\mathbf{H}_{r,t} = \sum_{p=1}^{L_p} h_{p,r,t} \left(\mathbf{\Pi}\right)^{i_p} \left(\mathbf{\Delta}\right)^{j_p}. \tag{5}$$

Here, $\mathbf{\Pi} \in \mathbb{C}^{MN\times MN}$ denotes a permutation matrix and $\mathbf{\Delta} = \mathrm{diag}\left\{\omega^q\right\}_{q=0}^{MN-1} \in \mathbb{C}^{MN\times MN}$ with $\omega = e^{j2\pi\frac{1}{MN}}$.

C. OTFS Demodulation

Let $\mathbf{Y}_{\mathrm{T},r} = \mathrm{vec}^{-1}(\mathbf{y}_{\mathrm{T},r}) \in \mathbb{C}^{M\times N}$ denote the time-domain received sample matrix. Employing a receiver pulse shaping filter $p_{\mathrm{rx}}(t)$ of duration T, the TF-demodulated symbol matrix $\mathbf{Y}_{\mathrm{TF},r} \in \mathbb{C}^{M\times N}$ is obtained as $\mathbf{Y}_{\mathrm{TF},r} = \mathbf{F}_M \mathbf{P}_{\mathrm{rx}} \mathbf{Y}_{\mathrm{T},r}$, where $\mathbf{P}_{\mathrm{rx}} = \mathrm{diag}\left\{p^*_{\mathrm{rx}}\left(\frac{qT}{M}\right)\right\}_{q=0}^{M-1}$ [18]. Subsequently, the demodulated DD-domain OTFS signal $\mathbf{Y}_{\mathrm{DD},r} \in \mathbb{C}^{M\times N}$ is expressed as $\mathbf{Y}_{\mathrm{DD},r} = \mathbf{F}_M^H \mathbf{Y}_{\mathrm{TF},r} \mathbf{F}_N = \mathbf{P}_{\mathrm{rx}} \mathbf{Y}_{\mathrm{T},r} \mathbf{F}_N$. Alternatively, its vectorized representation $\mathbf{y}_{\mathrm{DD},r}$ is given as

$$\mathbf{y}_{\mathrm{DD},r} = \mathrm{vec}\left(\mathbf{Y}_{\mathrm{DD},r}\right) = \left(\mathbf{F}_N \otimes \mathbf{P}_{\mathrm{rx}}\right) \mathbf{y}_{\mathrm{T},r} \in \mathbb{C}^{MN\times 1}. \quad (6)$$

In the above, upon replacing $\mathbf{y}_{\mathrm{T},r}$ from (4), and in turn substituting $\mathbf{s}_{\mathrm{T},t}$ from (1), we have

$$\mathbf{y}_{\mathrm{DD},r} = \sum_{t=1}^{N_t} \mathbf{H}_{\mathrm{DD},r,t} \mathbf{x}_{\mathrm{DD},t} + \mathbf{v}_{\mathrm{DD},r}, \quad (7)$$

where the quantity $\mathbf{H}_{\mathrm{DD},r,t} \in \mathbb{C}^{MN\times MN}$ can be expressed as

$$\mathbf{H}_{\mathrm{DD},r,t} = \left(\mathbf{F}_N \otimes \mathbf{P}_{\mathrm{r}x}\right) \mathbf{H}_{r,t} \left(\mathbf{F}_N^H \otimes \mathbf{P}_{\mathrm{tx}}\right), \quad (8)$$

and $\mathbf{v}_{\mathrm{DD},r} = \left(\mathbf{F}_N \otimes \mathbf{P}_{\mathrm{rx}}\right) \mathbf{w}_r \in \mathbb{C}^{MN\times 1}$. Finally, the end-to-end MIMO OTFS system model is formulated as

$$\mathbf{y}_{\mathrm{DD}} = \mathbf{H}_{\mathrm{DD}} \mathbf{x}_{\mathrm{DD}} + \mathbf{v}_{\mathrm{DD}}, \quad (9)$$

where we have $\mathbf{y}_{\mathrm{DD}} = \left[\mathbf{y}^T_{\mathrm{DD},1}, \cdots, \mathbf{y}^T_{\mathrm{DD},N_r}\right]^T$, $\mathbf{x}_{\mathrm{DD}} = \left[\mathbf{x}^T_{\mathrm{DD},1}, \cdots, \mathbf{x}^T_{\mathrm{DD},N_t}\right]^T$ and $\mathbf{v}_{\mathrm{DD}} = \left[\mathbf{v}^T_{\mathrm{DD},1}, \cdots, \mathbf{v}^T_{\mathrm{DD},N_r}\right]^T$. The DD-domain MIMO OTFS channel matrix $\mathbf{H}_{\mathrm{DD}} \in \mathbb{C}^{MNN_r\times MNN_t}$ is given by

$$\mathbf{H}_{\mathrm{DD}} = \mathrm{blkmtx}\left(\left\{\mathbf{H}_{\mathrm{DD},r,t}\right\}_{r=1,t=1}^{N_r,N_t}\right), \quad (10)$$

where $\mathrm{blkmtx}(\cdot)$ formulates a block-matrix having $\mathbf{H}_{\mathrm{DD},r,t}$ in the rth row and tth column block. Now, upon substituting $\mathbf{H}_{\mathrm{DD},r,t}$ from (8) into (10), we have

$$\mathbf{H}_{\mathrm{DD}} = \left(\mathbf{I}_{N_r} \otimes \mathbf{F}_N \otimes \mathbf{P}_{\mathrm{r}x}\right) \bar{\mathbf{H}} \left(\mathbf{I}_{N_t} \otimes \mathbf{F}_N^H \otimes \mathbf{P}_{\mathrm{tx}}\right), \quad (11)$$

where $\bar{\mathbf{H}}$ is formulated as

$$\bar{\mathbf{H}} = \mathrm{blkmtx}\left(\left\{\mathbf{H}_{r,t}\right\}_{r=1,t=1}^{N_r,N_t}\right). \quad (12)$$

Furthermore, upon substituting $\mathbf{H}_{r,t}$ from (5) into (12), and employing the relationship given in (3), we have

$$\begin{aligned}\bar{\mathbf{H}} &= \sum_{p=1}^{L_p} \left[\mathbf{H}_p \otimes \left(\mathbf{\Pi}^{i_p} \mathbf{\Delta}^{j_p}\right)\right] \\ &= \sum_{p=1}^{L_p} \alpha_p \left[\left(\mathbf{a}_r\left(\theta_p\right) \mathbf{a}_t^H\left(\phi_p\right)\right) \otimes \left(\mathbf{\Pi}^{i_p} \mathbf{\Delta}^{j_p}\right)\right]. \end{aligned} \quad (13)$$

Finally, substituting (13) into (11) yields the resultant expression of the end-to-end MIMO OTFS channel $\mathbf{H}_{\mathrm{DD}}$. The next section describes the proposed DDA-domain sparse channel estimation scheme.

III. Proposed 4D-Sparse Channel Estimation

Let G_τ, G_ν, G_r and G_t denote the grid sizes employed in the delay, Doppler, AoA and AoD domains, respectively, for formulating the DDA-domain sparse representation of the MIMO channel. More specifically, the delay grid $\mathcal{G}(\tau)$ and Doppler grid $\mathcal{G}(\nu)$ are defined as $\mathcal{G}(\tau) = \{\tau_i : \tau_i = \frac{i}{M\Delta f}\}_{i=0}^{G_\tau-1}, \mathcal{G}(\nu) = \{\nu_j : \nu_j = \frac{j}{NT}\}_{j=0}^{G_\nu-1}$. Similarly, the AoA grid $\mathcal{G}(\theta)$ and AoD grid $\mathcal{G}(\phi)$ are defined as $\mathcal{G}(\theta) = \{\theta_k : \theta_k = k\frac{\pi}{G_r}\}_{k=1}^{G_r}, \mathcal{G}(\phi) = \{\phi_l : \phi_l = l\frac{\pi}{G_t}\}_{l=1}^{G_t}$. Let $\alpha_{i,j,k,l}$ represent the complex-valued path gain corresponding to the delay-index i, Doppler-index j, AoA-index k and AoD-index l. The DDA-domain channel of (2) can be expressed as

$$\begin{aligned}\mathbf{H}(\tau,\nu,\theta,\phi) = &\sum_{i=0}^{G_\tau-1}\sum_{j=0}^{G_\nu-1}\sum_{k=1}^{G_r}\sum_{l=1}^{G_t} \alpha_{i,j,k,l} \mathbf{a}_r\left(\theta\right) \mathbf{a}_t^H\left(\phi\right) \\ &\times \delta(\tau-\tau_i)\delta(\nu-\nu_j)\delta(\theta-\theta_k)\delta(\phi-\phi_l).\end{aligned}$$

For convenience, one can represent the MIMO channel $\mathbf{H}_{i,j,k,l}$ corresponding to the (i,j,k,l)th grid-point as

$$\mathbf{H}_{i,j,k,l} = \alpha_{i,j,k,l} \mathbf{a}_r\left(\theta_k\right) \mathbf{a}_t^H\left(\phi_l\right), \quad (14)$$

whose the (r,t)th element is denoted by $h_{i,j,k,l,r,t}$.

For the purpose of CSI estimation, each TA employs N_p pilot symbols, which are placed directly in the time-domain. Let the pilot vector transmitted by the tth TA be denoted by $\mathbf{s}_{\mathcal{P},t} \in \mathbb{C}^{N_p\times 1}$. Upon CP addition/ removal, the pilot vector $\mathbf{y}_{\mathcal{P},r} \in \mathbb{C}^{N_p\times 1}$ received at the rth RA is given by

$$\mathbf{y}_{\mathcal{P},r} = \sum_{t=1}^{N_t} \mathbf{H}_{\mathcal{P},r,t} \mathbf{s}_{\mathcal{P},t} + \mathbf{w}_{\mathcal{P},r}, \quad (15)$$

where $\mathbf{w}_{\mathcal{P},r} \in \mathbb{C}^{N_p\times 1}$ denotes the noise vector. In the above, the matrix $\mathbf{H}_{\mathcal{P},r,t} \in \mathbb{C}^{N_p\times N_p}$ is formulated as

$$\mathbf{H}_{\mathcal{P},r,t} = \sum_{i,j,k,l} h_{i,j,k,l,r,t} \left(\widetilde{\mathbf{\Pi}}\right)^i \left(\widetilde{\mathbf{\Delta}}_i\right)^j, \quad (16)$$

where $\widetilde{\mathbf{\Pi}}$ represents the standard permutation matrix of size $N_p \times N_p$ and $\widetilde{\mathbf{\Delta}}_i \in \mathbb{C}^{N_p\times N_p}$ is defined as

$$\widetilde{\mathbf{\Delta}}_i = \begin{cases} \mathrm{diag}\left\{1, \omega, \cdots, \omega^{N_p-i-1}, \omega^{-i}, \cdots, \omega^{-1}\right\}, \text{if } i \neq 0, \\ \mathrm{diag}\left\{1, \omega, \cdots, \omega^{N_p-1}\right\}, \text{for } i = 0. \end{cases}$$

Stacking the various quantities as $\mathbf{y}_{\mathcal{P}} = \left[\mathbf{y}^T_{\mathcal{P},1}, \cdots, \mathbf{y}^T_{\mathcal{P},N_r}\right]^T$, $\mathbf{s}_{\mathcal{P}} = \left[\mathbf{s}^T_{\mathcal{P},1}, \cdots, \mathbf{s}^T_{\mathcal{P},N_t}\right]^T$ and $\mathbf{w}_{\mathcal{P}} = \left[\mathbf{w}^T_{\mathcal{P},1}, \cdots, \mathbf{w}^T_{\mathcal{P},N_r}\right]^T$, the resultant channel estimation model is formulated as

$$\mathbf{y}_{\mathcal{P}} = \mathbf{H}_{\mathcal{P}} \mathbf{s}_{\mathcal{P}} + \mathbf{w}_{\mathcal{P}}, \quad (17)$$

where the quantity $\mathbf{H}_{\mathcal{P}} \in \mathbb{C}^{N_pN_r\times N_pN_t}$ is given by

$$\mathbf{H}_{\mathcal{P}} = \mathrm{blkmtx}\left(\left\{\mathbf{H}_{\mathcal{P},r,t}\right\}_{r=1,t=1}^{N_r,N_t}\right). \quad (18)$$

Now, upon substituting $\mathbf{H}_{\mathcal{P},r,t}$ from (16) into the above equation, and in turn utilizing the relationship in (14), one can express the matrix $\mathbf{H}_{\mathcal{P}}$, similar to (13), as

$$\mathbf{H}_{\mathcal{P}} = \sum_{i,j,k,l} \alpha_{i,j,k,l} \left[\left(\mathbf{a}_r\left(\theta_k\right) \mathbf{a}_t^H\left(\phi_l\right)\right) \otimes \left(\widetilde{\mathbf{\Pi}}^i \widetilde{\mathbf{\Delta}}_i^j\right)\right]. \quad (19)$$

Algorithm 1 OMP-based sparse CSI estimation

Input: Dictionary matrix $\mathbf{\Phi}$, observation vector $\mathbf{y}_{\mathcal{P}}$

Initialization: $\mathcal{I} = [\ ]$, residue $\mathbf{r}_{-1} = \mathbf{0}_{N_p N_r \times 1}$, $\mathbf{r}_0 = \mathbf{y}_{\mathcal{P}}$, $\widehat{\boldsymbol{\alpha}}_{\text{OMP}} = \mathbf{0}_{G_\tau G_\nu G_r G_t \times 1}$, $\mathbf{\Phi}^{\mathcal{I}} = [\ ]$, set counter $i = 0$

while $\left(\|\mathbf{r}_i\|^2 \geq \sigma^2 N_p N_r\right)$ **do**

1) $i = i + 1$
2) $j = \underset{k=1,\cdots,G_\tau G_\nu G_r G_t}{\arg\max} \left|\mathbf{\Phi}^H(:,k)\mathbf{r}_{i-1}\right|$
3) $\mathcal{I} = \mathcal{I} \cup j$
4) $\mathbf{\Phi}^{\mathcal{I}} = \mathbf{\Phi}(:,\mathcal{I})$
5) $\widehat{\boldsymbol{\alpha}}^i_{\text{LS}} = \left(\mathbf{\Phi}^{\mathcal{I}}\right)^{\dagger} \mathbf{y}_{\mathcal{P}}$
6) $\mathbf{r}_i = \mathbf{y}_{\mathcal{P}} - \mathbf{\Phi}^{\mathcal{I}} \widehat{\boldsymbol{\alpha}}^i_{\text{LS}}$

end while

$\widehat{\boldsymbol{\alpha}}_{\text{OMP}}(\mathcal{I}) = \widehat{\boldsymbol{\alpha}}^i_{\text{LS}}$

Output: $\widehat{\boldsymbol{\alpha}}_{\text{OMP}}$

Furthermore, upon substituting $\mathbf{H}_{\mathcal{P}}$ into (17), we obtain

$$\mathbf{y}_{\mathcal{P}} = \sum_{i,j,k,l} \boldsymbol{\phi}_{i,j,k,l}\ \alpha_{i,j,k,l} + \mathbf{w}_{\mathcal{P}}, \tag{20}$$

where the quantity $\boldsymbol{\phi}_{i,j,k,l} \in \mathbb{C}^{N_p N_r \times 1}$ obeys:

$$\boldsymbol{\phi}_{i,j,k,l} = \left[\left(\mathbf{a}_r(\theta_k)\, \mathbf{a}_t^H(\phi_l)\right) \otimes \left(\widetilde{\mathbf{\Pi}}^i \widetilde{\mathbf{\Delta}}_i^j\right)\right] \mathbf{s}_{\mathcal{P}}.$$

Note that the number of multi-path components L_p is very small. Therefore, very few, i.e., L_p, out of $G_\tau G_\nu G_r G_t$, elements in the set $\{\alpha_{i,j,k,l}\}$ are non-zero. Hence, one can now formulate the MIMO OTFS CSI estimation problem of Eq. (20) as the 4D-sparse signal recovery problem:

$$\mathbf{y}_{\mathcal{P}} = \mathbf{\Phi}\boldsymbol{\alpha} + \mathbf{w}_{\mathcal{P}}, \tag{21}$$

where the dictionary matrix $\mathbf{\Phi} \in \mathbb{C}^{N_p N_r \times G_\tau G_\nu G_r G_t}$ is comprised of $\boldsymbol{\phi}_{i,j,k,l}$ as its columns, whereas the vector $\boldsymbol{\alpha} \in \mathbb{C}^{G_\tau G_\nu G_r G_t \times 1}$ is a sparse vector, which is comprised of the corresponding coefficients $\alpha_{i,j,k,l}$. To this end, the least-square (LS) and minimum mean square error (MMSE)-based CSI estimation techniques, which although are very popular for conventional MIMO systems, fail to exploit the DDA-domain sparsity. Hence, the next subsection employs a low-complexity OMP-based CSI estimation technique for MIMO OTFS systems, which exploits this sparsity for enhanced accuracy.

A. OMP-Based Sparse CSI Estimation

The popular OMP technique of solving the sparse signal recovery problem of (21) is presented in Algorithm-1. In Step-2 of the ith iteration, the previous residue vector $\mathbf{r}_{i-1}$ is correlated with the columns of the dictionary matrix $\mathbf{\Phi}$. Subsequently, the column index j is obtained, which is highly aligned with the residue $\mathbf{r}_{i-1}$. Thus, this index j corresponds to a potentially active multipath component. Step-3 and Step-4 update the basis-set $\mathcal{I}$ and the basis-matrix $\mathbf{\Phi}^{\mathcal{I}}$ by adding the selected index j and the corresponding column of $\mathbf{\Phi}$, respectively. In Step-5, an LS solution $\widehat{\boldsymbol{\alpha}}^i_{\text{LS}}$ is determined, which is employed in Step-6 for updating the residue $\mathbf{r}_i$. Furthermore, this work considers a beneficial stopping rule that is free from any threshold tuning. As seen under the while-condition of Algorithm-1, the proposed OMP algorithm stops, when the l_2-norm-square of the residue becomes smaller than the noise power, viz. $\text{Trace}\left(\mathbb{E}\left[\mathbf{w}_{\mathcal{P}}\mathbf{w}_{\mathcal{P}}^H\right]\right) = \sigma^2 N_p N_r$. The OMP algorithm described above is seen to converge faster and to a more accurate solution with the aid of this rule. Finally, the estimate $\widehat{\boldsymbol{\alpha}}_{\text{OMP}}$ can be employed for constructing the estimate $\widehat{\mathbf{H}}_{\text{DD}}$ using Eq. (11) and (13).

TABLE I
SIMULATION PARAMETERS

	System-I	**System-II**
Carrier frequency (GHz)	4	28
Subcarrier spacing (KHz) (Δf)	15	40
# of TAs (N_t)	4	4
# of RAs (N_r)	4	8
# of subcarriers (M)	32	16
# of OFDM symbols (N)	16	32
Pilot length (N_p)	300	150
CP length (L)	8	6
# of multipath components (L_p)	$\mathcal{U}(5,8)$	$\mathcal{U}(5,8)$
Grid-size along delay-axis (G_τ)	6	8
Grid-size along Doppler-axis (G_ν)	6	8
Grid-size along AoA (G_r)	8	8
Grid-size along AoD (G_t)	8	8

Note that we perform the DDA-domain CSI estimation using the discrete-time input-output relationship of (15). Furthermore, the DDA-domain relationship of (9) is utilized exclusively for the end-to-end symbol detection. Thus, in our work, the CSI estimation and data transmission procedures are performed in two separate frames. Interestingly, since we employ the discrete-time input-output relationship of (15) for DDA-domain CSI estimation, the proposed CSI estimation framework can also be readily employed in an OFDM system. However, the conventional OFDM systems have no mechanism to exploit the estimated Dopplers of the multipath components. By contrast, exploiting this is the key motivation of developing the proposed 4D-sparse CSI estimation technique for MIMO OTFS systems, since it can readily exploit the estimated Doppler values of the multipath components for overcoming the time-selectivity of mobile wireless channel. The key challenge here is to additionally exploit the AoA and AoD information of the multipath components along with both the delay as well as the Doppler for developing the end-to-end system and channel estimation models. To the best of our knowledge, these have hitherto not been explored in the OTFS literature. Hence, the proposed designs are ideally suited for MIMO OTFS systems.

IV. SIMULATION RESULTS

To illustrate the performance of the proposed 4D-sparse CSI estimation framework, this section considers a pair of different MIMO OTFS systems, as parameterzied in Table-I. The delay and Doppler values of the channel are assumed to coincide with their integer grid-points, whereas both integer and fractional angular indices are considered for AoAs and AoDs. The performance of the proposed OMP framework is compared to that of the popular FOCUSS-based [19] sparse signal recovery scheme and also to the 'sparsity-agnostic'

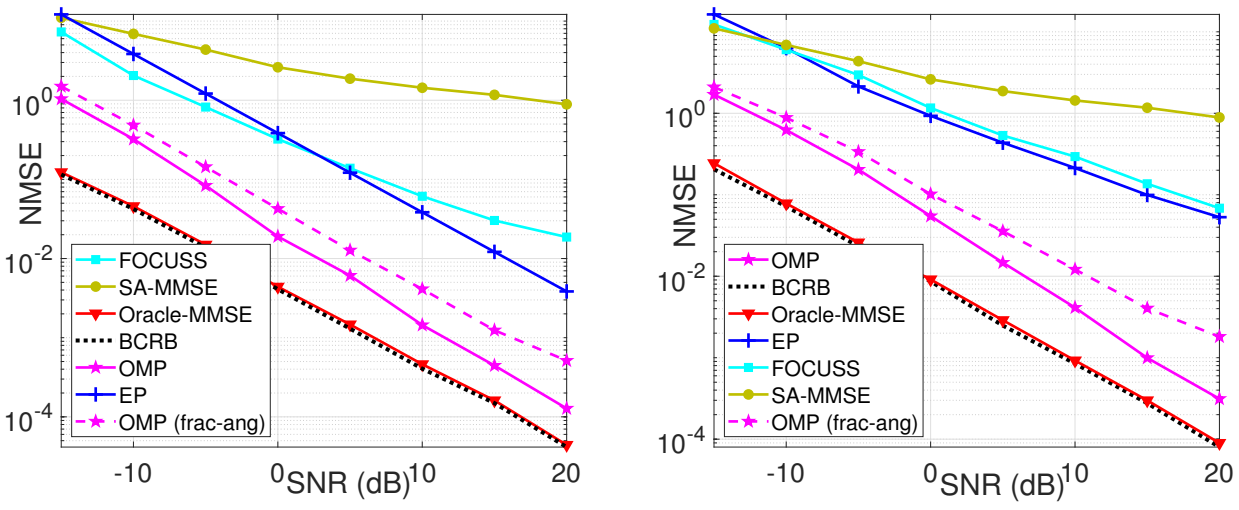

Fig. 1. NMSE vs. SNR comparison for (a) System-I, (b) System-II ('frac-ang' signifies fractional angular indices).

MMSE (SA-MMSE), which is the conventional MMSE estimator relying on the correlation matrix set to $\mathbf{I}_{G_\tau G_\nu G_r G_t}$. The performance is also compared to the hypothetical Oracle-MMSE and BCRLB benchmarks, which can be derived assuming the *a priori* knowledge of the delay, Doppler and angular parameters of the underlying MIMO channel. Let $\mathcal{H}$ denote the set comprising the indices of the non-zero elements in the sparse vector $\boldsymbol{\alpha}$, and $\boldsymbol{\Phi}_{\mathrm{o}} = \boldsymbol{\Phi}(:, \mathcal{H})$. The Oracle-MMSE estimator is formulated as

$$\widehat{\boldsymbol{\alpha}}_{\text{O-MMSE}} = \left(\boldsymbol{\Phi}_{\mathrm{o}}^H \boldsymbol{\Phi}_{\mathrm{o}} + \sigma^2 \mathbf{I}_{L_p}\right)^{-1} \boldsymbol{\Phi}_{\mathrm{o}}^H \mathbf{y}_{\mathcal{P}}. \tag{22}$$

Furthermore, the BCRLB corresponding to the estimate $\widehat{\boldsymbol{\alpha}}_{\text{O-MMSE}}$ is obtained as [20]

$$\text{MSE} \geq \text{Tr}\left[\sigma^2 \left(\boldsymbol{\Phi}_{\mathrm{o}}^H \boldsymbol{\Phi}_{\mathrm{o}} + \sigma^2 \mathbf{I}_{L_p}\right)^{-1}\right]. \tag{23}$$

Furthermore, the state-of-the-art EP-based estimator of [9] is also considered for comparison. The SNR in decibels (dB) is defined as $10\log_{10}\left(\frac{1}{\sigma^2}\right)$, whereas the normalized MSE (NMSE) is defined as $\frac{\left\|\widehat{\mathbf{H}}_{\text{DD}} - \mathbf{H}_{\text{DD}}\right\|_F^2}{\left\|\mathbf{H}_{\text{DD}}\right\|_F^2}$.

Fig. 1(a) and 1(b) plot the NMSE achieved by the various competing schemes with respect to the SNR. Firstly, we note that the dictionary matrices $\boldsymbol{\Phi}$ corresponding to System-I and System-II are of sizes $[1200 \times 2304]$ and $[1200 \times 4096]$, respectively, which result in highly under-determined channel estimation models, since there are only 1200 pilot observations, whereas the lengths of the parameters $\boldsymbol{\alpha}$ to be estimated are $\{2304, 4096\}$. Observe from both these plots that the proposed OMP framework yields a reasonably low NMSE for this 'ill-posed' CSI estimation scenario. This can be attributed to the key property of the CS algorithms, which can recover sparse signals from much fewer observations. Furthermore, the proposed framework has a significantly improved NMSE with respect to the FOCUSS [19], EP [9] and SA-MMSE schemes. The dependency of the FOCUSS-based CS technique on the regularization parameter as well as the resultant convergence errors lead to its degraded performance. The poor performance of both the EP-based scheme and of the conventional SA-MMSE estimator is attributed to the fact that they do not leverage the sparsity. By contrast, the proposed OMP algorithm does not employ any regularization parameter and also exploits the 4D-sparse structure of the DDA-domain CSI leading to the best NMSE amongst all. Furthermore, the NMSE of the OMP scheme approaches that of the Oracle-MMSE and BCRLB benchmarks, which are derived under the perfect knowledge of the DDA-domain sparsity profile, whereas the proposed OMP framework does not require this knowledge. The improved CSI estimation performance of the OMP framework is also seen in its resultant symbol error rate (SER) versus SNR plot of Fig. 2, where the MMSE-based linear detectors have been constructed using the estimated CSI. The SER achieved using the detector designed from the OMP-based estimated CSI is

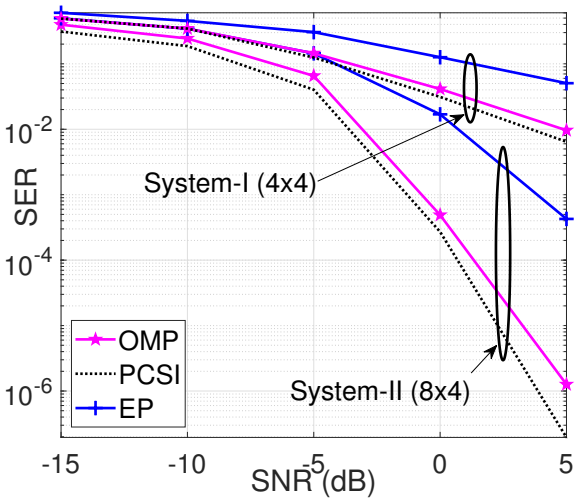

Fig. 2. SER vs. SNR comparison for System-I.

TABLE II
PILOT OVERHEAD COMPARISON: (PROPOSED, EP), 'NA' REPRESENTS 'NOT APPLICABLE'

N_t	**2**	**4**
System-I	$(0.369, 0.507)$	$(0.369, 0.863)$
System-II	$(0.226, 0.863)$	$(0.226, \text{NA})$

seen to be close to that designed using the perfect CSI (PCSI).

Finally, Table-II compares the pilot overhead of the proposed framework to that of the EP-based method of [9]. Note that for our method, the pilot overhead is given by $\frac{N_p}{MN+N_p}$. By contrast, for the EP-based technique, it is given by $\frac{(N_tG_\tau+G_\tau+N_t)(2G_\nu+1)}{MN}$, which arises due to the requirement of multiple guard symbols. Furthermore, it can also be readily observed that the pilot overhead of [14] is already significantly higher in comparison to [9], since the former does not place any data symbols along with pilots in the same OTFS frame, while the latter does this intelligently. It can be readily observed from Table-II that the pilot overhead of our proposed framework is significantly lower than that of the EP-based scheme of [9], since our estimation model does not require any DD-domain guard intervals, except for the CP addition in time-domain. Furthermore, the lower pilot overhead is also attributed to the compressive sensing (CS)-based problem formulation. Interestingly, upon increasing the number of transmit antennas N_t, and the DD-domain spread of the wireless channel, the EP-based technique becomes highly inefficient, since its pilot overhead tends to 100%. This is due to the fact that it cannot accommodate any data symbol along with the pilot and guard symbols in the same OTFS frame. Thus, the enhanced NMSE and SER performance compared to EP, as demonstrated in Fig. 1 and 2, coupled with lower pilot overhead of the proposed 4D-sparse CSI estimation framework, as illustrated in Table-II, make it eminently suitable for practical implementation in MIMO OTFS systems.

V. CONCLUSIONS

This paper conceived a delay-Doppler-angular domain representation of the wireless channel for deriving the end-to-end relationship in a MIMO OTFS system, followed by developing a time-domain pilot aided channel estimation model to exploit the inherent 4D-sparse structure. The proposed sparsity based OMP framework was seen to achieve superior NMSE of the estimated channel both in comparison to the state-of-the-art EP as well as to that conventional FOCUSS and MMSE schemes. Furthermore, the bandwidth efficiency of the proposed OMP-based 4D-sparse channel estimation model was seen to be significantly higher than that of the conventional MIMO OTFS CSI estimation schemes, since its pilot overhead is very low. A promising future direction is to extend this work for the millimeter wave band and large-antenna regimes, where the angular-sparsity is more dominant. Furthermore, one can also extend this work to handle fractional-Dopplers by constructing a virtual Doppler-grid as considered in [21].

REFERENCES

[1] R. Hadani and A. Monk, "OTFS: A new generation of modulation addressing the challenges of 5G," *ArXiv preprint:1802.02623*, 2018.

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

[3] M. Ramachandran, G. Surabhi, and A. Chockalingam, "OTFS: A new modulation scheme for high-mobility use cases," *Journal of the Indian Institute of Science*, pp. 1–22, 2020.

[4] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[5] R. Hadani, S. Rakib, M. Tsatsanis, A. Monk, A. J. Goldsmith, A. F. Molisch, and R. Calderbank, "Orthogonal time frequency space modulation," in *2017 IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE, 2017, pp. 1–6.

[6] R. Hadani, S. Rakib, S. Kons, M. Tsatsanis, A. Monk, C. Ibars, J. Delfeld, Y. Hebron, A. J. Goldsmith, A. F. Molisch *et al.*, "Orthogonal time frequency space modulation," *ArXiv preprint:1808.00519*, 2018.

[7] Z. Wei, W. Yuan, S. Li, J. Yuan, G. Bharatula, R. Hadani, and L. Hanzo, "Orthogonal time-frequency space modulation: A promising next-generation waveform," *IEEE Wireless Communications*, vol. 28, no. 4, pp. 136–144, 2021.

[8] S. Li, J. Yuan, W. Yuan, Z. Wei, B. Bai, and D. W. K. Ng, "Performance analysis of coded OTFS systems over high-mobility channels," *IEEE Transactions on Wireless Communications*, 2021.

[9] P. Raviteja, K. T. Phan, and Y. Hong, "Embedded pilot-aided channel estimation for OTFS in delay-Doppler channels," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 5, pp. 4906–4917, 2019.

[10] S. Srivastava, R. K. Singh, A. K. Jagannatham, and L. Hanzo, "Bayesian learning aided simultaneous row and group sparse channel estimation in orthogonal time frequency space modulated MIMO systems," *IEEE Transactions on Communications*, 2021.

[11] S. S. Rakib and R. Hadani, "Orthogonal time frequency space modulation system," Mar. 27 2018, US Patent 9,929,783.

[12] L. Zhao, W. Gao, and W. Guo, "Sparse Bayesian learning of delay-Doppler channel for OTFS system," *IEEE Communications Letters*, 2020.

[13] S. Srivastava, R. K. Singh, A. K. Jagannatham, and L. Hanzo, "Bayesian learning aided sparse channel estimation for orthogonal time frequency space modulated systems," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 8, pp. 8343–8348, 2021.

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

[15] R. W. Heath, N. Gonzalez-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 3, pp. 436–453, 2016.

[16] A. M. Sayeed, T. Sivanadyan, K. Liu, and S. Haykin, "Wireless communication and sensing in multipath environments using multi-antenna transceivers," in *Handbook on Array Processing and Sensor Networks*. Wiley Online Library, 2010, pp. 115–170.

[17] S. Li, W. Yuan, C. Liu, Z. Wei, J. Yuan, B. Bai, and D. W. K. Ng, "A novel ISAC transmission framework based on spatially-spread orthogonal time frequency space modulation," *IEEE Journal on Selected Areas in Communications*, 2022.

[18] P. Raviteja, Y. Hong, E. Viterbo, and E. Biglieri, "Practical pulse-shaping waveforms for reduced-cyclic-prefix OTFS," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 1, pp. 957–961, 2018.

[19] I. F. Gorodnitsky and B. D. Rao, "Sparse signal reconstruction from limited data using FOCUSS: A re-weighted minimum norm algorithm," *IEEE Trans. on Signal Processing*, vol. 45, no. 3, pp. 600–616, 1997.

[20] H. L. Van Trees and K. L. Bell, "Bayesian bounds for parameter estimation and nonlinear filtering/tracking," *AMC*, vol. 10, p. 12, 2007.

[21] Z. Wei, W. Yuan, S. Li, J. Yuan, and D. W. K. Ng, "Off-grid channel estimation with sparse Bayesian learning for OTFS systems," *IEEE Transactions on Wireless Communications*, 2022.