

# Bayesian Learning Aided Simultaneous Row and Group Sparse Channel Estimation in Orthogonal Time Frequency Space Modulated MIMO systems

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

**Abstract**—A sparse channel state information (CSI) estimation model is proposed for reducing the pilot overhead of orthogonal time frequency space (OTFS) modulation aided multiple-input multiple-output (MIMO) systems. Explicitly, the pilots are directly transmitted over the time-frequency (TF)-domain grid for estimating the delay-Doppler (DD)-domain CSI that leads to a reduction of the pilot overhead, training duration and pre-processing complexity. Furthermore, it completely avoids placing multiple DD-domain guard intervals corresponding to each transmit antenna within the same OTFS frame, while keeping the training duration flexible, hence increasing the bandwidth efficiency. A unique benefit of the proposed CSI estimation model is that it can efficiently handle fractional Dopplers also. The resultant DD-domain CSI becomes simultaneously row and group (RG)-sparse. To exploit this compelling property, an orthogonal matching pursuit (OMP)-based RG-OMP technique is developed, conveniently complemented by an enhanced Bayesian learning (BL)-based RG-BL framework, both of which substantially outperform the state-of-the-art methods. Furthermore, low-complexity linear detectors are designed for the ensuing data detection phase, which directly employ the estimated DD-domain sparse CSI, without assuming any further knowledge concerning the number of dominant multipath components. Finally, simulation results are provided to demonstrate performance improvement of the proposed BL-based schemes over the OMP and the state-of-the-art schemes.

**Index Terms**—OTFS, delay-Doppler domain channel, simultaneous sparsity, fractional Doppler, channel estimation, high-mobility

## I. INTRODUCTION

Next generation wireless technologies are envisioned to support high data rates in scenarios where the relative mobile velocities can be up to 400-500 km/h for high-speed trains (HSTs) [1], [2], vehicle to vehicle and to everything (V2V

and V2X) [3], [4]. Communications become even more challenging for air-plane users, since their speeds are in the range of 800-1000 km/h [5]. The 4G and 5G cellular systems have adopted orthogonal frequency division multiplexing (OFDM) modulation as their waveform, which requires doubling the pilot overhead every time the Doppler-frequency is doubled. It is widely recognized that high mobility renders the wide-band wireless channel time as well as frequency-selective, i.e. doubly-selective [6]–[9]. As a result, multipath propagation exhibiting a high delay-spread leads to inter-symbol interference (ISI) [8], [10], whereas the high mobile velocity leads to significant inter-carrier interference (ICI) owing to the large Doppler-shift [6], [11]. Thus, given the continuously evolving and diverse range of use cases and applications, it becomes imperative to explore novel modulation, coding and multiple access (MA) techniques that are resilient both to the delay- and to the Doppler-shifts introduced by the wireless channel.

To this end, a novel delay-Doppler (DD)-domain modulation technique, originally proposed by Hadani *et al.* in [7], [12], [13], termed orthogonal time frequency space (OTFS), has gained significant popularity. As the terminology implies, the resultant basis functions of OTFS are orthogonal in the time-frequency (TF)-domain, which has been shown to outperform conventional multicarrier modulation techniques in high-Doppler channels. An important aspect of the OTFS theory is that it incorporates the key principle of DD-domain representation of the wireless channel [6], [12], [14], which leads to a significant reduction in the channel's dimensionality, to a value that is typically equal to the number of dominant reflectors ( $\sim 10 - 15$ ), thus potentially yielding the *sparse* representation of the channel [6]. The conventional TF-domain representation of the multipath wireless channel is time-varying at a rate inversely proportional to the *coherence-time* [10], which in turn depends on the mobile node's relative velocity and on the carrier's frequency [6], [10]. This phenomenon renders closed-loop channel estimation difficult to achieve. By contrast, the equivalent DD-domain representation of the wireless channel is nearly time-invariant across the entire bandwidth over a much longer observation duration, since the velocity and distance remain approximately constant for at least a few milliseconds. Together, the sparsity and slow variability of the DD-domain channel state information (CSI) facilitate its reliable estimation, which can be carried out at a potentially reduced pilot overhead [6], [15], [16], in comparison to conventional OFDM-style TF-domain procedures.

L. Hanzo would like to acknowledge the financial support of the Engineering and Physical Sciences Research Council projects EP/P034284/1 and EP/P003990/1 (COALESCE) as well as of the European Research Council's Advanced Fellow Grant QuantCom (Grant No. 789028). 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.

S. Srivastava, R. K. Singh and A. K. Jagannatham are with the Department of Electrical Engineering, Indian Institute of Technology Kanpur, Kanpur, U.P. 208016, India (e-mail: ssvrast@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).

Recently, OTFS has also been explored for non-orthogonal multiple access (NOMA) systems in [17]–[19], which are highly suitable to support massive connectivity in the next generation wireless systems. However, these advantages together with the improved end-to-end BER performance of the OTFS-based transceiver require accurate DD-domain CSI at the receiver for reliable detection. Hence, CSI estimation plays a central role in realizing the potential gains promised by the OTFS systems [15], [16], [20]–[23]. A brief review of the existing contributions is presented next in this context.

#### A. Review of Existing Contributions on OTFS CSI Estimation

Early contributions [7], [24] proposed impulse-based CSI estimation schemes for single-input single-output (SISO) OTFS systems. As described in [6], [12], [14], considering perfect *biorthogonal* transmit-receive (Tx-Rx) pulse shaping filters, the end-to-end DD-domain input-output relationship is given by the 2D-circular convolution between the transmit DD-domain signal and the DD-domain channel. Hence, the transmission of a training-impulse followed by thresholding yields an estimate of the DD-domain channel impulse response (CIR) [6]. This principle is extended to CSI estimation in multiple-input multiple-output (MIMO) OTFS systems in [20]. Here, the training-impulses corresponding to the different transmit antennas (TAs) are placed sufficiently apart obeying suitable DD-domain guards in order to separate their CIRs at the receiver. Thus, a key drawback of the proposed MIMO OTFS CSI estimation scheme therein is that it requires an entire OTFS frame for pilot transmission, which significantly reduces the spectral efficiency.

The authors of [16] propose embedded pilot (EP)-based CSI estimation schemes for SISO, MIMO as well as multiuser (MU) MIMO OTFS systems. Here, the DD-domain pilots, as well as the data and guard symbols are appropriately arranged in order to avoid interference between data and pilots at the receiver, followed by the thresholding method for CSI estimation. The key novelty of the proposed technique is that both CSI estimation and data detection are achieved within the same OTFS-frame. In a related contribution [25], Murali and Chockalingam consider pseudo-random noise (PN)-based pilot sequences in the DD-domain to estimate the CIR of a SISO OTFS system. A key drawback of all the training-impulse and embedded-pilot techniques reviewed above is that they typically require a higher pilot signal-to-noise ratio, thereby leading to an increase in the PAPR of the system. Furthermore, the performance of these schemes also depends to a great extent on the choice of a suitable threshold, which has to be empirically tuned for the best performance.

Alternatively, the innovative contributions [15], [21], [23], [26], [27] exploit the DD-domain sparsity of the wireless channel by conceiving an interesting formulation of the DD-domain CSI estimation model as a sparse signal recovery problem. These schemes have demonstrated superior CSI estimation performance in comparison to the previously discussed training impulse and embedded pilot techniques, since they leverage the sparsity of the underlying DD-domain channel. Shen *et al.* [15] develop the downlink CSI estimation model

for a massive MIMO OTFS system. Since the number of dominant reflectors in such a channel is typically very small, the massive MIMO OTFS channel exhibits 3D sparsity along the delay, Doppler and angular domains. Finally, a novel 3D-simultaneous orthogonal matching pursuit (3D-SOMP)-based technique is proposed in [15] to solve the resultant sparse CSI estimation problem, which extracts the 3D support of each multipath component in the delay, Doppler and angular dimensions. As a further development, the authors of [23] consider the uplink of an OTFS-based MA and model the DD-domain CSI estimation paradigm as a sparse signal recovery problem. The authors then suitably adapt the OMP and modified subspace pursuit (MSP) algorithms for sparse CSI estimation. To this end, Zhao *et al.* [21], proposed a novel pilot pattern, characterized by the absence of a DD-domain-guard band between the pilots and data. This pilot-data frame structure is successfully exploited in their work to formulate a sparse channel estimation problem for SISO OTFS systems. Another recent contribution [27] formulates a sparse signal recovery problem for estimating the original DD domain channel of the SISO OTFS system and employs the sparse Bayesian learning (SBL) framework. Initially, a 1D off-grid sparse signal recovery problem has been formulated based on a sampling grid defined in the DD-domain, which jointly estimates the delay and Doppler shifts. Subsequently, its low complexity version based on a 2D off-grid sparse signal recovery problem has been formulated for decoupling the estimation of the delay and Doppler shifts.

While good progress has been achieved by the contributions reviewed above, a few shortcomings remain. Prominently, almost all of the existing treatises employ a DD-domain-guard for pilot placement, which leads to a reduction of the resultant spectral efficiency. This problem becomes even more challenging in a MIMO OTFS system due to the multiple DD-guards that are required corresponding to different TAs. Hence, to the best of our knowledge, there is a paucity of research on the development of DD-domain sparse CSI estimation schemes for point-to-point MIMO OTFS systems. Furthermore, techniques for DD-domain sparse CSI estimation using a practical rectangular pulse is rarely seen. It is also a common practice for the subsequent data detection procedure to consider the non-linear iterative message passing (MP) detector [14] or its low complexity version, i.e., the variational Bayes (VB) detector [28]. In this context, zero-forcing (ZF) or minimum mean squared error (MMSE)-based linear detectors can be an attractive choice in these systems. Finally, it is also desirable to develop a framework which can efficiently estimate the fractional Dopplers. These gaps in the existing research motivate us to develop novel DD-domain sparse CSI estimation schemes for MIMO OTFS systems that are capable of overcoming the shortcomings of the existing schemes in the literature. The novel contributions of this paper are itemized next and are also boldly contrasted to those of the existing literature in Table-I. An overview of OTFS as well as various challenges and research opportunities have been discussed in a recent survey [29].

TABLE I  
SUMMARY OF LITERATURE SURVEY ON OTFS CHANNEL ESTIMATION (**R**: REQUIRED, **NR**: NOT REQUIRED)

	[22]	[20]	[15]	[16]	[21]	[23]	[26]	[27]	Proposed
MIMO		✓		✓					✓
DD-domain sparsity			✓		✓	✓	✓	✓	✓
Simultaneous row-group Sparsity									✓
Practical pulse shape			✓	✓	✓		✓		✓
Linear detector							✓		✓
Flexible pilot overhead							✓	✓	✓
Fractional Dopplers								✓	✓
Delay-Doppler-angular channel									✓
Requirement of DD-guards	R	R	R	R	NR	NR	NR	NR	NR

### B. Contributions of the Paper

- 1) By considering arbitrary Tx-Rx pulse shapes and fractional Dopplers, this paper first derives an end-to-end input-output model for MIMO OTFS systems. Based on this, a novel sparse channel estimation model is proposed, in which the pilots are transmitted in the TF-domain for reducing the pilot overhead, training duration and its pre-processing complexity. A key contribution of the proposed sparse channel estimation model is that it can efficiently estimate the fractional Dopplers.
- 2) Interestingly, it is also demonstrated that the DD-domain CSI of the MIMO OTFS system exhibits simultaneous row-group (RG)-sparsity. Another significant advantage of the proposed channel estimation model is that it completely avoids having multiple DD-domain-guards between the pilots and data symbols within the same OTFS frame, while facilitating a flexible pilot duration, hence increasing the spectral efficiency.
- 3) Initially, the conventional LS and MMSE schemes are developed for estimating the DD-domain CSI. Subsequently, novel RG-OMP and RG-BL techniques are derived for exploiting the simultaneous RG-sparsity of the MIMO OTFS DD-domain CSI.
- 4) For data detection, low-complexity linear detectors are derived, which directly employ the DD-domain CSI obtained from the sparse estimation schemes proposed, and do not require any knowledge concerning the number of dominant multipath components.

### C. Organization of the Work

Section-II derives the MIMO OTFS system model, followed by the proposed sparse channel estimation model in Section-III. The RG-OMP based scheme is developed toward the end of Section-III, while Section-IV develops the proposed RG-BL framework. This is followed by our simulation results in Section-V and Section-VI concludes the paper.

### D. Notation

Boldface lower case and upper case letters denote column vectors and matrices, respectively. The indices of the vectors and matrices elements start from 0. The vector equivalent of the matrix  $\mathbf{A}$  is denoted by  $\text{vec}(\mathbf{A})$ , which is formed

by stacking the columns to form a single column vector. Similarly,  $\text{vec}^{-1}(\mathbf{a})$  denotes the corresponding inverse vectorization operation to provide the original matrix. A well-known property of the  $\text{vec}(\cdot)$  operator, given by  $\text{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A}) \text{vec}(\mathbf{B})$ , is used in the paper, where  $\otimes$  denotes the Kronecker product of two matrices.

## II. OTFS SYSTEM MODEL

Consider an OTFS system having the frame duration of  $T_f = NT$  and bandwidth of  $B = M\Delta f$ , where  $T$  (seconds) denotes the symbol duration and  $\Delta f$  (Hz) represents the sub-carrier spacing, such that  $T\Delta f = 1$ . The quantities  $N$  and  $M$  represent the number of symbols along the time and frequency axes in the corresponding TF-grid, denoted by  $\Lambda_{\text{TF}}$ . OTFS is a 2D modulation technique, which places the information symbols in an equivalent DD-domain grid, denoted by  $\Lambda_{\text{DD}}$ , where the delay and Doppler axes are sampled at integer multiples of  $\Delta\tau = \frac{1}{B}$  and  $\Delta\nu = \frac{1}{T_f}$ , respectively, as demonstrated in Fig. 1(a). Thus, the DD-domain grid can be represented as  $\Lambda_{\text{DD}} = \{(l\Delta\tau, k\Delta\nu)\}_{l,k} = \left\{\left(\frac{l}{M\Delta f}, \frac{k}{NT}\right)\right\}_{l,k}$ , where the delay index  $l$  and Doppler index  $k$  obey  $l = 0, 1, \dots, M-1$  and  $k = 0, 1, \dots, N-1$ . Similarly, the corresponding TF-grid is given by  $\Lambda_{\text{TF}} = \{(m\Delta f, nT)\}_{m,n}$ , where the subcarrier index  $m$  follows  $m = 0, 1, \dots, M-1$ , whereas the symbol time index  $n$  obeys  $n = 0, 1, \dots, N-1$ . The signal processing operations of an OTFS system are described next.

### A. OTFS Modulation

Let  $\mathbf{X}_{\text{DD}} \in \mathbb{C}^{M \times N}$  denote a matrix of information symbols, where  $\mathbf{X}_{\text{DD}}(l, k)$  represents the symbol placed at delay index  $l$  and Doppler index  $k$ . The transmitter first maps these DD-domain symbols to the TF-domain by using the inverse symplectic finite Fourier transform (ISFFT) as follows

$$\mathbf{X}_{\text{TF}}(m, n) = \frac{1}{\sqrt{NM}} \sum_{l=0}^{M-1} \sum_{k=0}^{N-1} \mathbf{X}_{\text{DD}}(l, k) e^{j2\pi\left(\frac{nk}{N} - \frac{ml}{M}\right)}, \quad (1)$$

where the matrix  $\mathbf{X}_{\text{TF}} \in \mathbb{C}^{M \times N}$  denotes a TF-domain symbol matrix whose  $(m, n)$ th element, denoted by  $\mathbf{X}_{\text{TF}}(m, n)$ , represents the symbol to be transmitted over the  $m$ th subcarrier

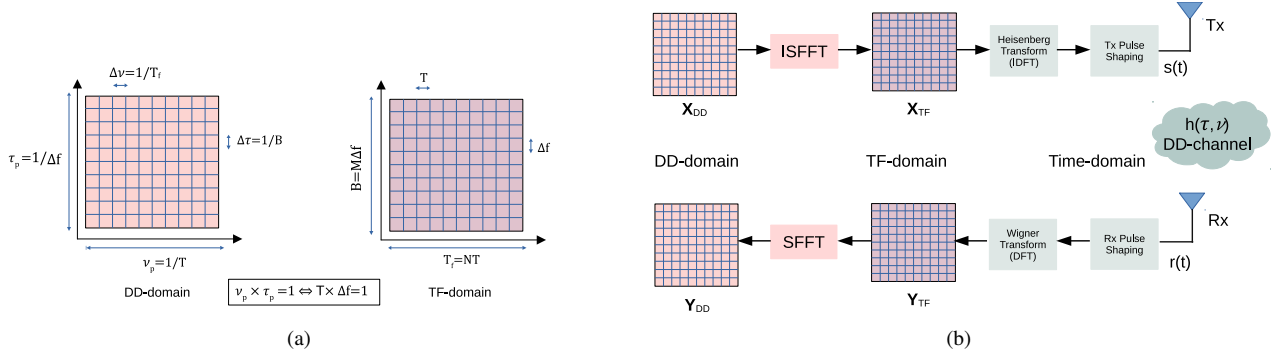


Fig. 1. (a) Relationship between various parameters of the DD-domain and TF-domain grids; (b) Architecture of the OTFS transceiver.

in the  $n$ th symbol duration. The matrix form of (1) can be equivalently written as

$$\mathbf{X}_{TF} = \mathbf{F}_M \mathbf{X}_{DD} \mathbf{F}_N^H, \quad (2)$$

where  $\mathbf{F}_N \in \mathbb{C}^{N \times N}$  represents the discrete Fourier transform (DFT) matrix, whose  $(i, j)$ th element is described as  $\mathbf{F}_N(i, j) = \frac{1}{\sqrt{N}} e^{-j2\pi \frac{ij}{N}}$ . Subsequently, any multicarrier modulation technique, such as OFDM, maps the 2D symbol sequence of the matrix  $\mathbf{X}_{TF}$  to a 1D sequence in the time-domain. Let  $p_{tx}(t)$  denote the transmit pulse of duration  $T$ , which is repeated  $N$ -times in the OTFS frame duration  $T_f$ . The time-domain signal  $s(t)$  is obtained by performing the Heisenberg transform of the symbols  $\mathbf{X}_{TF}(m, n)$  placed over the TF-grid  $\Lambda_{TF}$ , which is given by

$$s(t) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \mathbf{X}_{TF}(m, n) p_{tx}(t - nT) e^{j2\pi m \Delta f (t - nT)}. \quad (3)$$

Upon sampling the transmit signal  $s(t)$  at the Nyquist rate  $\frac{M}{T}$ , i.e. at the sampling interval  $\frac{T}{M}$ , the  $q$ th transmit sample  $s(q), 0 \leq q \leq MN - 1$ , is given by  $s(q) = s(t)|_{t=\frac{qT}{M}}$ . The  $MN$ -samples of the transmit signal  $s(t)$  in the form of a vector, denoted by  $\mathbf{s} \in \mathbb{C}^{MN \times 1}$ , can be formulated as follows. Let  $\mathbf{P}_{tx} \in \mathbb{C}^{M \times M}$  denotes a diagonal matrix, whose diagonal entries contain the  $M$ -samples of the transmit pulse  $p_{tx}(t)$ , i.e.,  $\mathbf{P}_{tx} = \text{diag} \left\{ p_{tx} \left( \frac{pT}{M} \right) \right\}_{p=0}^{M-1}$ . Thus, the transmit signal matrix  $\mathbf{S} \in \mathbb{C}^{M \times N}$  can be obtained as

$$\mathbf{S} = \mathbf{P}_{tx} \mathbf{F}_M^H \mathbf{X}_{TF} = \mathbf{P}_{tx} \mathbf{X}_{DD} \mathbf{F}_N^H, \quad (4)$$

where (4) follows upon substituting  $\mathbf{X}_{TF}$  from (2). Thus, the vector  $\mathbf{s}$  can be formulated as

$$\mathbf{s} = \text{vec}(\mathbf{S}) = (\mathbf{F}_N^H \otimes \mathbf{P}_{tx}) \mathbf{x}_{DD}, \quad (5)$$

where  $\mathbf{x}_{DD} = \text{vec}(\mathbf{X}_{DD})$ . Finally, prior to transmission, similar to OFDM, a cyclic prefix (CP) of length  $L$  is appended to  $\mathbf{s}$  in order to remove the inter-frame interference at the receiver. The various steps in the modulation process described above are shown schematically in Fig. 1(b).

### B. DD-domain Wireless Channel Model

Let  $h(\tau, \nu)$  denote the DD-domain representation of the wireless channel, which is a 2D function of the delay variable

$\tau$  and Doppler variable  $\nu$ . Since typically there are only a few dominant reflectors that constitute a wireless channel, it can be parameterized by only a small number of elements [6], [12], [14], which correspond to the delay- and Doppler-shifts introduced by those dominant reflectors. Let  $\tau_i$  and  $\nu_i$  represent the delay- and Doppler-shifts, whereas  $h_i$  denotes the complex path gain introduced by the  $i$ th reflector. The DD-domain wireless channel can be formulated as [6], [14]

$$h(\tau, \nu) = \sum_{i=1}^{L_p} h_i \delta(\tau - \tau_i) \delta(\nu - \nu_i), \quad (6)$$

where  $L_p$  denotes the number of dominant reflectors/ multipath components and  $\delta(\cdot)$  represents the Dirac-delta function. The Doppler-shift  $\nu_i$  associated with the  $i$ th multipath component can be expressed as  $\nu_i = \frac{k_i}{NT}$ , where  $k_i$  can be decomposed as  $k_i = \text{round}(k_i) + \kappa_{\nu_i}$ . Here, the quantity  $\kappa_{\nu_i}$ , with  $|\kappa_{\nu_i}| < \frac{1}{2}$ , represents fractional-Doppler. On the other hand, as described in [10], [14], [16], one does not have to consider fractional delays, since for a typical wideband system the delay resolution  $\Delta\tau = \frac{1}{M\Delta f}$  is small enough for considering the delay-shift  $\tau_i$  as  $\tau_i = \frac{l_i}{M\Delta f}$ , where  $l_i$  is an integer. Furthermore, we also have  $k_{\max} = \max(k_i) \ll N$  and  $l_{\max} = \max(l_i) \ll M$ , for a typical ‘under-spread’ wireless channel [14], [16], [30]. The signal  $r(t)$  at the output of the wireless channel obeys

$$r(t) = \int \int h(\tau, \nu) s(t - \tau) e^{j2\pi\nu(t - \tau)} d\tau d\nu + w(t), \quad (7)$$

where  $w(t)$  denotes the AWGN of average power  $\sigma^2$ . Substituting  $h(\tau, \nu)$  from (6) into the above equation, we obtain  $r(t) = \sum_{i=1}^{L_p} h_i s(t - \tau_i) e^{j2\pi\nu_i(t - \tau_i)} + w(t)$ , whose  $p$ th sample  $r(p) = r(t)|_{t=\frac{pT}{M}}$  is given by

$$r(p) = \sum_{i=1}^{L_p} h_i s(p - l_i) e^{j2\pi \frac{k_i(p - l_i)}{MN}} + w(p), \quad (8)$$

where  $w(p) = w(t)|_{t=\frac{pT}{M}}$ . Furthermore, by removing the initial  $L$ -samples that correspond to the cyclic prefix, the received samples  $r(p), 0 \leq p \leq MN - 1$ , obey

$$r(p) = \sum_{i=1}^{L_p} h_i e^{j2\pi \frac{k_i(p - l_i)}{MN}} s([p - l_i]_{MN}) + w(p), \quad (9)$$

where  $[\cdot]_{MN}$  denotes the modulo- $MN$  operation.

Let  $\mathbf{\Pi} \in \mathbb{C}^{MN \times MN}$  denote a standard permutation matrix and  $\mathbf{\Delta}_i \in \mathbb{C}^{MN \times MN}$  represent a diagonal matrix, which is defined as

$$\begin{cases} \text{diag} \left\{ 1, \omega_i, \dots, \omega_i^{MN-l_i-1}, \omega_i^{-l_i}, \dots, \omega_i^{-1} \right\}, & \text{if } l_i \neq 0, \\ \text{diag} \left\{ 1, \omega_i, \dots, \omega_i^{MN-1} \right\}, & \text{for } l_i = 0, \end{cases}$$

where  $\omega_i = e^{j2\pi \frac{k_i}{MN}}$ . Furthermore, let  $\mathbf{r} \in \mathbb{C}^{MN \times 1}$  and  $\mathbf{w} \in \mathbb{C}^{MN \times 1}$  comprise the  $MN$ -samples of the received signal  $r(t)$  and the noise process  $w(t)$ , given as

$$\begin{aligned} \mathbf{r} &= [r(0), r(1), \dots, r(MN-1)]^T, \\ \mathbf{w} &= [w(0), w(1), \dots, w(MN-1)]^T. \end{aligned} \quad (10)$$

Employing these notations, and using the relationship given in (9), the received signal vector  $\mathbf{r}$  can be expressed as

$$\mathbf{r} = \mathbf{H}\mathbf{s} + \mathbf{w}, \quad (11)$$

where the matrix  $\mathbf{H} \in \mathbb{C}^{MN \times MN}$  is obtained as

$$\mathbf{H} = \sum_{i=1}^{L_p} h_i (\mathbf{\Pi})^{l_i} \mathbf{\Delta}_i. \quad (12)$$

### C. OTFS Demodulator

As shown in Fig. 1(b), the received signal  $r(t)$  is processed by a filter matched to the receiver pulse  $p_{\text{rx}}(t)$  of duration  $T$ , which is repeated  $N$ -times, as

$$Y(f, t) = \int p_{\text{rx}}^*(t' - t) r(t') e^{-j2\pi f(t' - t)} dt'. \quad (13)$$

The output of the matched filter is sampled at integer multiples of the symbol duration  $T$  and subcarrier spacing  $\Delta f$ :

$$\begin{aligned} \mathbf{Y}_{\text{TF}}(m, n) &= Y(f, t)|_{f=m\Delta f, t=nT} \\ &= \int p_{\text{rx}}^*(t' - nT) r(t') e^{-j2\pi m\Delta f(t' - nT)} dt', \end{aligned} \quad (14)$$

where  $\mathbf{Y}_{\text{TF}} \in \mathbb{C}^{M \times N}$  denotes the TF-demodulated symbol matrix, whose  $(m, n)$ th element denotes the symbol corresponding to the  $m$ th subcarrier and  $n$ th symbol duration. This is referred to as the discrete Wigner transform [12], [14]. Let the received sample matrix  $\mathbf{R} \in \mathbb{C}^{M \times N}$  be defined as  $\mathbf{R} = \text{vec}^{-1}(\mathbf{r})$ . The TF-symbol matrix  $\mathbf{Y}_{\text{TF}}$  is given by

$$\mathbf{Y}_{\text{TF}} = \mathbf{F}_M \mathbf{P}_{\text{rx}} \mathbf{R}. \quad (15)$$

Finally, the demodulated DD-domain OTFS signal  $\mathbf{Y}_{\text{DD}} \in \mathbb{C}^{M \times N}$  can be obtained by the SFFT of the TF-domain demodulated signal  $\mathbf{Y}_{\text{TF}}$  as

$$\mathbf{Y}_{\text{DD}}(l, k) = \frac{1}{\sqrt{NM}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \mathbf{Y}_{\text{TF}}(m, n) e^{-j2\pi \left( \frac{nk}{N} - \frac{ml}{M} \right)}. \quad (16)$$

The above equation can be compactly written as  $\mathbf{Y}_{\text{DD}} = \mathbf{F}_M^H \mathbf{Y}_{\text{TF}} \mathbf{F}_N = \mathbf{P}_{\text{rx}} \mathbf{R} \mathbf{F}_N$ . Its equivalent vector  $\mathbf{y}_{\text{DD}} = \text{vec}(\mathbf{Y}_{\text{DD}}) = (\mathbf{F}_N \otimes \mathbf{P}_{\text{rx}}) \mathbf{r}$ , where upon substituting  $\mathbf{r}$  from (11) and  $\mathbf{s}$  from (5), the end-to-end DD-domain input-output model of the SISO OTFS system is obtained as

$$\mathbf{y}_{\text{DD}} = \mathbf{H}_{\text{DD}} \mathbf{x}_{\text{DD}} + \mathbf{v}_{\text{DD}}, \quad (17)$$

where  $\mathbf{H}_{\text{DD}} = (\mathbf{F}_N \otimes \mathbf{P}_{\text{rx}}) \mathbf{H} (\mathbf{F}_N^H \otimes \mathbf{P}_{\text{tx}}) \in \mathbb{C}^{MN \times MN}$  and  $\mathbf{v}_{\text{DD}} = (\mathbf{F}_N \otimes \mathbf{P}_{\text{rx}}) \mathbf{w} \in \mathbb{C}^{MN \times 1}$ . Finally, the covariance matrix  $\mathbf{R}_{\mathbf{v}, \text{DD}} = \mathbb{E}[\mathbf{v}_{\text{DD}} \mathbf{v}_{\text{DD}}^H] \in \mathbb{C}^{MN \times MN}$  of the noise obeys  $\mathbf{R}_{\mathbf{v}, \text{DD}} = \sigma^2 [\mathbf{I}_N \otimes (\mathbf{P}_{\text{rx}} \mathbf{P}_{\text{rx}}^H)]$ .

### D. MIMO OTFS System Model

Consider now a MIMO OTFS system having  $N_t$  TAs and  $N_r$  receive antennas (RAs). The DD-domain wireless channel  $h_{r,t}(\tau, \nu)$ ,  $1 \leq r \leq N_r$ ,  $1 \leq t \leq N_t$ , between the  $r$ -th RA and  $t$ -th TA, is given by

$$h_{r,t}(\tau, \nu) = \sum_{i=1}^{L_p} h_{i,r,t} \delta(\tau - \tau_i) \delta(\nu - \nu_i), \quad (18)$$

where  $h_{i,r,t}$  denotes the complex baseband channel coefficient corresponding to the  $i$ th reflector. Let  $\mathbf{X}_{\text{DD},t} \in \mathbb{C}^{M \times N}$  denote the DD-domain symbol matrix to be transmitted from the  $t$ -th TA and let  $\mathbf{Y}_{\text{DD},r} \in \mathbb{C}^{M \times N}$  represent the received symbol matrix at the  $r$ -th RA. Similar to the previous subsections, the input-output model at the  $r$ -th RA is given by

$$\mathbf{y}_{\text{DD},r} = \sum_{t=1}^{N_t} \mathbf{H}_{\text{DD},r,t} \mathbf{x}_{\text{DD},t} + \mathbf{v}_{\text{DD},r}, \quad (19)$$

where  $\mathbf{y}_{\text{DD},r} = \text{vec}(\mathbf{Y}_{\text{DD},r}) \in \mathbb{C}^{MN \times 1}$ ,  $\mathbf{x}_{\text{DD},t} = \text{vec}(\mathbf{X}_{\text{DD},t}) \in \mathbb{C}^{MN \times 1}$  and

$$\mathbf{H}_{\text{DD},r,t} = \sum_{i=1}^{L_p} h_{i,r,t} (\mathbf{F}_N \otimes \mathbf{P}_{\text{rx}}) \mathbf{\Pi}_i \mathbf{\Delta}_i (\mathbf{F}_N^H \otimes \mathbf{P}_{\text{tx}}). \quad (20)$$

The quantity  $\mathbf{v}_{\text{DD},r} \in \mathbb{C}^{MN \times 1}$  is given by  $\mathbf{v}_{\text{DD},r} = (\mathbf{F}_N \otimes \mathbf{P}_{\text{rx}}) \mathbf{w}_r$ , where  $\mathbf{w}_r \in \mathbb{C}^{MN \times 1}$  is comprised of the noise samples at the  $r$ -th RA. Stacking the received observation vectors  $\mathbf{y}_{\text{DD},r}$  for  $1 \leq r \leq N_r$ , as  $\bar{\mathbf{y}}_{\text{DD}} = [\mathbf{y}_{\text{DD},1}^T, \mathbf{y}_{\text{DD},2}^T, \dots, \mathbf{y}_{\text{DD},N_r}^T]^T \in \mathbb{C}^{MNN_r \times 1}$ , the end-to-end input-output model of the MIMO OTFS system is given by

$$\bar{\mathbf{y}}_{\text{DD}} = \bar{\mathbf{H}}_{\text{DD}} \bar{\mathbf{x}}_{\text{DD}} + \bar{\mathbf{v}}_{\text{DD}}, \quad (21)$$

where  $\bar{\mathbf{x}}_{\text{DD}} = [\mathbf{x}_{\text{DD},1}^T, \mathbf{x}_{\text{DD},2}^T, \dots, \mathbf{x}_{\text{DD},N_t}^T]^T \in \mathbb{C}^{MNN_t \times 1}$  represents the stacked vector of the transmit DD-domain symbols of all the TAs and  $\bar{\mathbf{v}}_{\text{DD}} = [\mathbf{v}_{\text{DD},1}^T, \mathbf{v}_{\text{DD},2}^T, \dots, \mathbf{v}_{\text{DD},N_r}^T]^T \in \mathbb{C}^{MNN_r \times 1}$  denotes the stacked noise vectors at the receiver. The DD-domain MIMO OTFS channel matrix  $\bar{\mathbf{H}}_{\text{DD}} \in \mathbb{C}^{MNN_r \times MNN_t}$  is given by

$$\bar{\mathbf{H}}_{\text{DD}} = \begin{bmatrix} \mathbf{H}_{\text{DD},1,1} & \mathbf{H}_{\text{DD},1,2} & \cdots & \mathbf{H}_{\text{DD},1,N_t} \\ \mathbf{H}_{\text{DD},2,1} & \mathbf{H}_{\text{DD},2,2} & \cdots & \mathbf{H}_{\text{DD},2,N_t} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{H}_{\text{DD},N_r,1} & \mathbf{H}_{\text{DD},N_r,2} & \cdots & \mathbf{H}_{\text{DD},N_r,N_t} \end{bmatrix}. \quad (22)$$

Finally, given the DD-domain CIR of the wireless channel, one can construct the matrix  $\bar{\mathbf{H}}_{\text{DD}}$ , which can then be employed for data detection. To this end, considering the information symbols having an average power of unity, the MMSE-based linear detector is formulated as

$$\bar{\mathbf{x}}_{\text{DD}}^{\text{MMSE}} = (\bar{\mathbf{H}}_{\text{DD}}^H \bar{\mathbf{R}}_{\mathbf{v}, \text{DD}}^{-1} \bar{\mathbf{H}}_{\text{DD}} + \mathbf{I}_{MNN_t})^{-1} \bar{\mathbf{H}}_{\text{DD}}^H \bar{\mathbf{R}}_{\mathbf{v}, \text{DD}}^{-1} \bar{\mathbf{y}}_{\text{DD}}, \quad (23)$$

where  $\bar{\mathbf{R}}_{v,DD} = \mathbb{E}[\bar{\mathbf{v}}_{DD}\bar{\mathbf{v}}_{DD}^H] \in \mathbb{C}^{MN_p \times MN_p}$  denotes the covariance matrix of the noise  $\bar{\mathbf{v}}_{DD}$ , which obeys  $\bar{\mathbf{R}}_{v,DD} = (\mathbf{I}_{N_p} \otimes \mathbf{R}_{v,DD})$ . Therefore, as mentioned earlier, CSI estimation plays a key role in deriving the benefits of the OTFS system, which forms the focus of the subsequent section.

### III. SPARSE CHANNEL ESTIMATION MODEL FOR MIMO OTFS SYSTEMS

Let the integer taps  $M_\tau$  and  $N_\nu$  denote the maximum delay- and Doppler-spread of the channel, which obey  $l_{\max} < M_\tau \ll M$  and  $k_{\max} < N_\nu \ll N$ , for a typical under-spread channel. For sparse representation, one can now consider a 2-dimensional DD-domain grid of size  $G_\tau$  on the delay-axis and  $G_\nu$  on the Doppler axis. As described in the previous section, since one does not have to consider fractional-delays, it is sufficient to set  $G_\tau = M_\tau$  and use the delay grid points on the integer-delay taps. Thus, the  $i$ th delay-grid point,  $0 \leq i \leq M_\tau$ , represents a delay of  $\tau_i = \frac{i}{M_\tau \Delta f}$  seconds. On the other hand, in order to capture fractional-Dopplers, one has to set  $G_\nu \gg N_\nu$ , as illustrated in Fig. 2(a), where the  $j$ th Doppler-grid point,  $0 \leq j \leq G_\nu$ , corresponds to a Doppler-shift of  $\nu_j = \frac{jN_\nu}{G_\nu NT}$  Hz. Note that the integer-Doppler tap corresponding to  $\nu_j$  is  $\left\lceil \text{round}\left(\frac{jN_\nu}{G_\nu}\right) \right\rceil$ , whereas the fractional-Doppler is given by  $\left[ \frac{jN_\nu}{G_\nu} - \text{round}\left(\frac{jN_\nu}{G_\nu}\right) \right]$ .

Let  $h_{i,r,t}^j$  denotes the path gain associated with the  $i$ th delay-tap and  $j$ th Doppler-tap between the  $r$ -th RA and  $t$ -th TA. The sparse representation of the DD-domain channel  $h_{r,t}(\tau, \nu)$  of (18) can be formulated as

$$h_{r,t}(\tau, \nu) = \sum_{i=0}^{M_\tau} \sum_{j=0}^{G_\nu} h_{i,r,t}^j \delta(\tau - \tau_i) \delta(\nu - \nu_j). \quad (24)$$

Note that since there are only a few dominant reflectors in the wireless channel, only a few coefficients  $L_p \ll [(M_\tau + 1)(G_\nu + 1)]$  are non-zero among the  $[(M_\tau + 1)(G_\nu + 1)]$  coefficients  $h_{i,r,t}^j$ . These coefficients signify the path gains of the non-zero DD-domain indices  $(i, j)$  of the dominant reflectors between the  $r$ -th RA and  $t$ -th TA.

Let  $\mathbf{X}_{\text{TF},\mathcal{P},t} \in \mathbb{C}^{M \times N_p}$  denote the transmit pilot symbol matrix placed on the TF-grid at the  $t$ -th TA, as shown in Fig. 2(b), where  $N_p$  denotes the number of pilot symbols placed along the time-axis. Thus, according to (4), the matrix  $\mathbf{S}_{\mathcal{P},t} \in \mathbb{C}^{M \times N_p}$  comprised of the time-domain pilot samples is obtained as  $\mathbf{S}_{\mathcal{P},t} = \mathbf{P}_{\text{tx}} \mathbf{F}_M^H \mathbf{X}_{\text{TF},\mathcal{P},t}$ . Furthermore, upon employing a procedure similar to (7)-(9), the time-domain pilot signal  $\mathbf{r}_{\mathcal{P},r} \in \mathbb{C}^{MN_p \times 1}$ , after the removal of the CP at the  $r$ -th RA is given by

$$\mathbf{r}_{\mathcal{P},r} = \sum_{t=1}^{N_t} \bar{\mathbf{H}}_{r,t} \mathbf{s}_{\mathcal{P},t} + \mathbf{w}_{\mathcal{P},r}, \quad (25)$$

where  $\mathbf{s}_{\mathcal{P},t} = \text{vec}(\mathbf{S}_{\mathcal{P},t}) \in \mathbb{C}^{MN_p \times 1}$  and  $\mathbf{w}_{\mathcal{P},r} \in \mathbb{C}^{MN_p \times 1}$  represents the noise vector. The matrix  $\bar{\mathbf{H}}_{r,t} \in \mathbb{C}^{MN_p \times MN_p}$  can be formulated as

$$\bar{\mathbf{H}}_{r,t} = \sum_{i=0}^{M_\tau} \sum_{j=0}^{G_\nu} h_{i,r,t}^j (\bar{\mathbf{\Pi}})^i \bar{\mathbf{\Delta}}_{i,j}, \quad (26)$$

where  $\bar{\mathbf{\Pi}}$  denotes a permutation matrix of size  $MN_p \times MN_p$ , whereas  $\bar{\mathbf{\Delta}}_{i,j} \in \mathbb{C}^{MN_p \times MN_p}$  represents a diagonal matrix, defined as

$$\begin{cases} \text{diag} \left\{ 1, \bar{\omega}_j, \dots, \bar{\omega}_j^{MN_p-i-1}, \bar{\omega}_j^{-i}, \dots, \bar{\omega}_j^{-1} \right\}, & \text{if } i \neq 0, \\ \text{diag} \left\{ 1, \bar{\omega}_j, \dots, \bar{\omega}_j^{MN_p-1} \right\}, & \text{for } i = 0, \end{cases}$$

where  $\bar{\omega}_k = e^{j2\pi \frac{kN_\nu}{G_\nu MN_p}}$ . Now, let the received pilot sample matrix  $\mathbf{R}_{\mathcal{P},r} \in \mathbb{C}^{M \times N_p}$  be arranged as  $\mathbf{R}_{\mathcal{P},r} = \text{vec}^{-1}(\mathbf{r}_{\mathcal{P},r})$ . Using this notation, the demodulated TF-domain pilot output matrix  $\mathbf{Y}_{\text{TF},\mathcal{P},r} \in \mathbb{C}^{M \times N_p}$  at the  $r$ -th RA is given similar to (15) as  $\mathbf{Y}_{\text{TF},\mathcal{P},r} = \mathbf{F}_M \mathbf{P}_{\text{rx}} \mathbf{R}_{\mathcal{P},r}$ . The equivalent vectorized form is expressed as

$$\mathbf{y}_{\text{TF},\mathcal{P},r} = \text{vec}(\mathbf{Y}_{\text{TF},\mathcal{P},r}) = [\mathbf{I}_{N_p} \otimes (\mathbf{F}_M \mathbf{P}_{\text{rx}})] \mathbf{r}_{\mathcal{P},r}. \quad (27)$$

In the above, upon substituting  $\mathbf{r}_{\mathcal{P},r}$  from (25), the quantity  $\mathbf{y}_{\text{TF},\mathcal{P},r}$  can be written as

$$\mathbf{y}_{\text{TF},\mathcal{P},r} = \sum_{t=1}^{N_t} (\mathbf{I}_{N_p} \otimes \mathbf{F}_M \mathbf{P}_{\text{rx}}) \bar{\mathbf{H}}_{r,t} (\mathbf{I}_{N_p} \otimes \mathbf{P}_{\text{tx}} \mathbf{F}_M^H) \mathbf{x}_{\text{TF},\mathcal{P},t} + \mathbf{v}_r,$$

where we have  $\mathbf{x}_{\text{TF},\mathcal{P},t} = \text{vec}(\mathbf{X}_{\text{TF},\mathcal{P},t}) \in \mathbb{C}^{MN_p \times 1}$  and  $\mathbf{v}_r = (\mathbf{I}_{N_p} \otimes \mathbf{F}_M \mathbf{P}_{\text{rx}}) \mathbf{w}_{\mathcal{P},r} \in \mathbb{C}^{MN_p \times 1}$ . In the above, the noise covariance  $\mathbf{R}_v = \mathbb{E}[\mathbf{v}_r \mathbf{v}_r^H] \in \mathbb{C}^{MN_p \times MN_p}$  obeys  $\mathbf{R}_v = \sigma^2 [\mathbf{I}_{N_p} \otimes (\mathbf{F}_M \mathbf{P}_{\text{rx}} \mathbf{P}_{\text{tx}}^H \mathbf{F}_M^H)]$ . Substituting  $\bar{\mathbf{H}}_{r,t}$  into the above, the pilot output  $\mathbf{y}_{\text{TF},\mathcal{P},r}$  can be expressed as

$$\mathbf{y}_{\text{TF},\mathcal{P},r} = \sum_{t=1}^{N_t} \mathbf{\Omega}_t \mathbf{h}_{r,t} + \mathbf{v}_r. \quad (28)$$

In the above model, the dictionary matrix  $\mathbf{\Omega}_t \in \mathbb{C}^{MN_p \times [(M_\tau+1)(G_\nu+1)]}$  is given by

$$\mathbf{\Omega}_t = \left[ \omega_{0,t}^0, \omega_{0,t}^1, \dots, \omega_{0,t}^{G_\nu}, \dots, \omega_{M_\tau,t}^0, \omega_{M_\tau,t}^1, \dots, \omega_{M_\tau,t}^{G_\nu} \right], \quad (29)$$

where the quantity  $\omega_{i,t}^j \in \mathbb{C}^{MN_p \times 1}$  obeys

$$\omega_{i,t}^j = (\mathbf{I}_{N_p} \otimes \mathbf{F}_M \mathbf{P}_{\text{rx}}) (\bar{\mathbf{\Pi}})^i \bar{\mathbf{\Delta}}_{i,j} (\mathbf{I}_{N_p} \otimes \mathbf{P}_{\text{tx}} \mathbf{F}_M^H) \mathbf{x}_{\text{TF},\mathcal{P},t},$$

and the vector  $\mathbf{h}_{r,t} \in \mathbb{C}^{[(M_\tau+1)(G_\nu+1)] \times 1}$  is defined as

$$\left[ h_{0,r,t}^0, h_{0,r,t}^1, \dots, h_{0,r,t}^{G_\nu}, \dots, h_{M_\tau,r,t}^0, h_{M_\tau,r,t}^1, \dots, h_{M_\tau,r,t}^{G_\nu} \right]^T. \quad (30)$$

The expression in (28) can now be rewritten as

$$\mathbf{y}_{\text{TF},\mathcal{P},r} = \tilde{\mathbf{\Omega}} \mathbf{h}_r + \mathbf{v}_r, \quad (31)$$

where the dictionary matrix  $\tilde{\mathbf{\Omega}} \in \mathbb{C}^{MN_p \times [(M_\tau+1)(G_\nu+1)N_t]}$  and the stacked CSI vector  $\mathbf{h}_r \in \mathbb{C}^{[(M_\tau+1)(G_\nu+1)N_t] \times 1}$  are given by

$$\tilde{\mathbf{\Omega}} = [\mathbf{\Omega}_1, \mathbf{\Omega}_2, \dots, \mathbf{\Omega}_{N_t}], \quad \mathbf{h}_r = [\mathbf{h}_{r,1}^T, \mathbf{h}_{r,2}^T, \dots, \mathbf{h}_{r,N_t}^T]^T. \quad (32)$$

Furthermore, upon concatenating the pilot outputs  $\mathbf{y}_{\text{TF},\mathcal{P},r}$ , for  $1 \leq r \leq N_r$ , the resultant multiple measurement vector

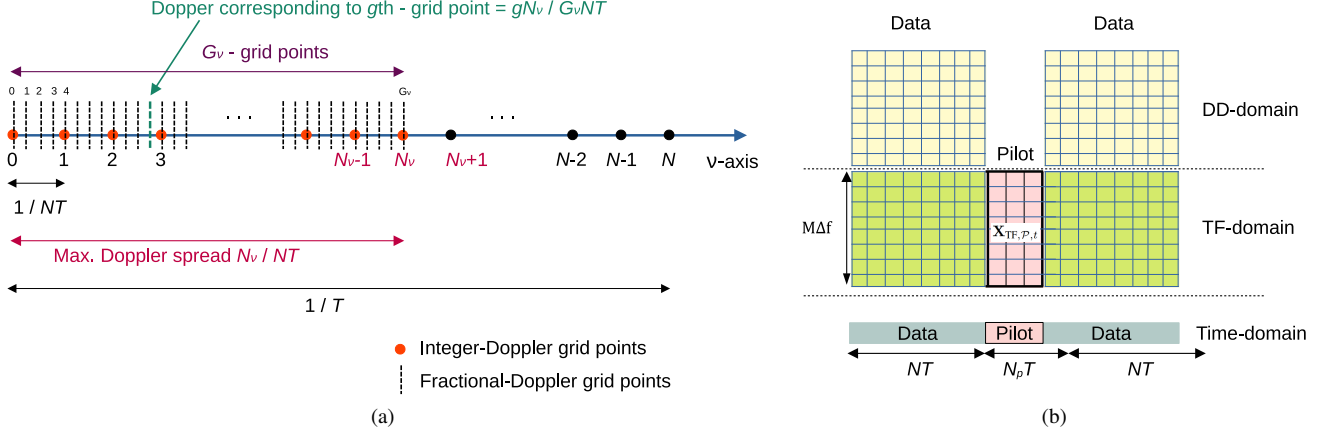


Fig. 2. (a) Fractional-Doppler grid; (b) Training frame structure of the proposed channel estimation procedure with  $N_p \ll N$ .

(MMV)-based sparse channel estimation model is obtained as

$$\begin{aligned} \tilde{\mathbf{Y}}_{\text{TF},\mathcal{P}} &= [\mathbf{y}_{\text{TF},\mathcal{P},1}, \mathbf{y}_{\text{TF},\mathcal{P},2}, \dots, \mathbf{y}_{\text{TF},\mathcal{P},N_r}] \in \mathbb{C}^{M N_p \times N_r} \\ &= \tilde{\mathbf{\Omega}} \tilde{\mathbf{H}} + \tilde{\mathbf{V}}, \end{aligned} \quad (33)$$

where the effective channel  $\tilde{\mathbf{H}} \in \mathbb{C}^{[(M_\tau+1)(G_\nu+1)N_t] \times N_r}$  and noise  $\tilde{\mathbf{V}} \in \mathbb{C}^{M N_p \times N_r}$  are given by

$$\tilde{\mathbf{H}} = [\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_{N_r}], \quad \tilde{\mathbf{V}} = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{N_r}]. \quad (34)$$

For (33), the conventional MMSE estimate  $\hat{\mathbf{H}}_{\text{MMSE}} \in \mathbb{C}^{[(M_\tau+1)(G_\nu+1)N_t] \times N_r}$  is given by

$$\hat{\mathbf{H}}_{\text{MMSE}} = \left( \tilde{\mathbf{\Omega}}^H \mathbf{R}_v^{-1} \tilde{\mathbf{\Omega}} + \tilde{\mathbf{R}}_h^{-1} \right)^{-1} \tilde{\mathbf{\Omega}}^H \mathbf{R}_v^{-1} \tilde{\mathbf{Y}}_{\text{TF},\mathcal{P}}, \quad (35)$$

where  $\tilde{\mathbf{R}}_h \in \mathbb{C}^{[(M_\tau+1)(G_\nu+1)N_t] \times [(M_\tau+1)(G_\nu+1)N_t]}$  denotes the covariance matrix of the channel vector  $\mathbf{h}_r$ . Since the true channel covariance is not known, the matrix  $\tilde{\mathbf{R}}_h$  is set to  $\mathbf{I}_{[(M_\tau+1)(G_\nu+1)N_t]}$  for the conventional sparsity-agnostic MMSE estimator. On the other hand, as described after (24), typically, only  $L_p$  out of  $[(M_\tau+1)(G_\nu+1)]$  elements of the vector  $\mathbf{h}_{r,t}$  are non-zero. Hence, the DD-domain CSI matrix  $\tilde{\mathbf{H}}$ , which in turn is comprised of  $\mathbf{h}_{r,t}, 1 \leq r \leq N_r, 1 \leq t \leq N_t$ , is *sparse* in nature. It is important to note that the conventional LS and MMSE channel estimation schemes, although appealingly simple, they fail to exploit the sparse nature of the CSI  $\tilde{\mathbf{H}}$  arising due to limited number of multipath components, which is a unique characteristic of the wireless channel representation in the DD-domain. Leveraging this aspect can lead to a significant improvement in the quality of the CSI estimate obtained. Hence, the next subsection develops an RG-OMP-based sparse channel estimation framework to exploit the sparsity.

To this end, let us now first describe the sparse structure inherent in the MIMO CSI matrix  $\tilde{\mathbf{H}}$ . To simplify the notation, let the total grid-size  $G$  be defined as  $G = [(M_\tau+1)(G_\nu+1)]$ . Note that the DD-domain CSI  $\mathbf{h}_{r,t}$  is sparse, and more importantly, from the channel model of (24), the delay and Doppler shifts corresponding to the multipath components do not change for the TA and RA indices. This implies that the vectors  $\mathbf{h}_{r,t}, 1 \leq r \leq N_r, 1 \leq t \leq N_t$ , share an identical sparsity profile. It follows from (32) that the vector

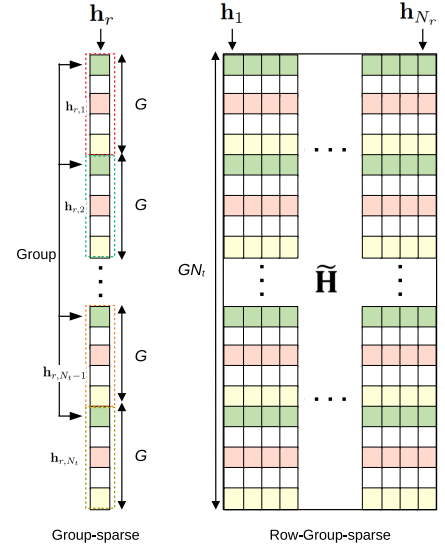


Fig. 3. Group-sparse and simultaneous row-group sparse structures of the DD-domain MIMO OTFS channel, where  $G = [(M_\tau+1)(G_\nu+1)]$ .

$\mathbf{h}_r$ , which stacks  $\mathbf{h}_{r,t}$  for  $1 \leq t \leq N_t$ , exhibits a *group-sparse* structure, since the  $i$ th group,  $0 \leq i \leq G-1$ , denoted by the set  $\left\{ \mathbf{h}_r \left( [(t-1)G] + i \right) \right\}_{t=1}^{N_t}$ , becomes either all zero or all non-zero, as described in Fig. 3. Furthermore, from (34), and as also depicted in Fig. 3, the  $i$ th group of rows denoted by the row indices  $\left\{ [(t-1)G] + i \right\}_{t=1}^{N_t}$ , becomes either all zero or all non-zero. Thus, the DD-domain CSI matrix  $\tilde{\mathbf{H}}$  is simultaneously row- and group-sparse in nature, i.e., there are only a few, typically  $L_p$  ( $\ll G$ ) groups, each having  $N_t$  rows, which are non-zero.

We now develop an OMP-based sparse channel estimation framework for exploiting this RG sparsity. The proposed RG-OMP scheme is described in Algorithm-1. The key steps that differ here from the conventional OMP technique are as follows. In each iteration, Step-2, 3 and 4 determine the group-index  $j$  that is maximally correlated to the previous residue  $\mathbf{R}_{i-1}$ . Subsequently, Step-5 selects all the  $N_t$  indices of the

**Algorithm 1** RG-OMP-based sparse channel estimation in MIMO OTFS systems

**Input:** Dictionary matrix  $\tilde{\Omega}$ , observation matrix  $\tilde{\mathbf{Y}}_{\text{TF},\mathcal{P}}$  and stopping parameter  $\epsilon_t$

**Initialization:**  $\mathcal{I} = [ ]$ , residue  $\mathbf{R}_{-1} = \mathbf{0}_{MN_p \times N_r}$ ,  $\mathbf{R}_0 = \tilde{\mathbf{Y}}_{\text{TF},\mathcal{P}}$ ,  $\hat{\mathbf{H}}_{\text{RG-OMP}} = \mathbf{0}_{GN_t \times N_r}$ ,  $\tilde{\Omega}^{\mathcal{I}} = [ ]$  and set counter  $i = 0$   
**while** ( $\| \mathbf{R}_{i-1} \|_F^2 - \| \mathbf{R}_i \|_F^2 \geq \epsilon_t$ ) **do**

- 1)  $i = i + 1$
- 2)  $\Psi = \tilde{\Omega}^H \mathbf{R}_{i-1}$
- 3)  $\varphi = \text{diag}(\Psi \Psi^H)$
- 4)  $j = \arg \max_{k=0,1,\dots,G-1} \sum_{t=1}^{N_t} \varphi[(t-1)G + k]$
- 5)  $\mathcal{J} = \{[(t-1)G + j]_{t=1}^{N_t}\}$
- 6)  $\mathcal{I} = \mathcal{I} \cup \mathcal{J}$
- 7)  $\tilde{\Omega}^{\mathcal{I}} = \tilde{\Omega}(:, \mathcal{I})$
- 8)  $\hat{\mathbf{H}}^i = (\tilde{\Omega}^{\mathcal{I}})^\dagger \tilde{\mathbf{Y}}_{\text{TF},\mathcal{P}}$
- 9)  $\mathbf{R}_i = \tilde{\mathbf{Y}}_{\text{TF},\mathcal{P}} - \tilde{\Omega}^{\mathcal{I}} \hat{\mathbf{H}}^i$

**end while**

$\hat{\mathbf{H}}_{\text{RG-OMP}}(\mathcal{I}, :) = \hat{\mathbf{H}}^i$

**Output:**  $\hat{\mathbf{H}}_{\text{RG-OMP}}$

group chosen. Thus, it can be observed that in each iteration, the RG-OMP algorithm selects the  $N_t$  columns of the matrix  $\tilde{\Omega}$ , which correspond to a group, thus yielding a row-group sparse estimate and in turn, performance improvement over the conventional OMP that selects only one column of the dictionary matrix in each iteration. However, the proposed RG-OMP scheme also suffers from the drawbacks of OMP, namely: error propagation and sensitivity to the choice of dictionary matrix and stopping threshold. Therefore, to overcome these drawbacks, the next section develops an enhanced RG-sparse BL (RG-BL) framework for CSI estimation in the MIMO OTFS system.

#### IV. RG-BL BASED SPARSE CHANNEL ESTIMATION IN MIMO OTFS SYSTEMS

The proposed RG-BL framework is based on the Bayesian principle, which begins by assigning the following parameterized Gaussian prior to the DD-domain channel  $\mathbf{h}_{r,t}$  [31]:

$$f(\mathbf{h}_{r,t}; \Gamma) = \prod_{i=0}^{G-1} \frac{1}{(\pi \gamma_i)} \exp\left(-\frac{|\mathbf{h}_{r,t}(i)|^2}{\gamma_i}\right). \quad (36)$$

Here, the quantity  $\gamma_i$ ,  $0 \leq i \leq G-1$ , denotes the hyperparameter and the matrix  $\Gamma = \text{diag}\{\gamma_i\}_{i=0}^{G-1} \in \mathbb{R}^{G \times G}$  represents the hyperparameter matrix. Note that the *a priori* covariance  $\mathbf{R}_h$  of the vector  $\mathbf{h}_{r,t}$  obeys  $\mathbf{R}_h = \Gamma$ , which is unknown to start with. Since the sparsity profile of  $\mathbf{h}_{r,t}$  is identical for  $1 \leq r \leq N_r, 1 \leq t \leq N_t$ , the prior corresponding to the MIMO CSI matrix  $\tilde{\mathbf{H}}$  is given by

$$f(\tilde{\mathbf{H}}; \Gamma) = \prod_{r=1}^{N_r} \prod_{t=1}^{N_t} \prod_{i=0}^{G-1} \frac{1}{(\pi \gamma_i)} \exp\left(-\frac{|\mathbf{h}_{r,t}(i)|^2}{\gamma_i}\right). \quad (37)$$

Note that to exploit the RG-sparsity, the hyperparameter  $\gamma_i$  is assigned to all the elements of the  $i$ th row-group, indexed

by the row indices  $\{(t-1)G + i\}_{t=1}^{N_t}$ , which simultaneously become either all zero or all non-zero. Thus, for estimating the  $[GN_t \times N_r]$ -element matrix  $\tilde{\mathbf{H}}$ , the proposed RG-BL technique employs only  $G$  hyperparameters, which is rendered possible via exploiting its special RG-sparse structure. Furthermore, the *a priori* covariance matrix  $\mathbf{R}_h$  of the vector  $\mathbf{h}_r$  follows  $\tilde{\mathbf{R}}_h = (\mathbf{I}_{N_t} \otimes \Gamma)$ . The MMSE estimate  $\mathcal{M} \in \mathbb{C}^{GN_t \times N_r}$  and its error covariance matrix  $\tilde{\Sigma} \in \mathbb{C}^{GN_t \times GN_t}$  of the CSI  $\tilde{\mathbf{H}}$  using (35) can be evaluated as [32]

$$\mathcal{M} = \tilde{\Sigma} \tilde{\Omega}^H \mathbf{R}_v^{-1} \tilde{\mathbf{Y}}_{\text{TF},\mathcal{P}}, \quad \tilde{\Sigma} = \left[ \tilde{\Omega}^H \mathbf{R}_v^{-1} \tilde{\Omega} + (\mathbf{I}_{N_t} \otimes \Gamma^{-1}) \right]^{-1}. \quad (38)$$

Thus, the MMSE estimate  $\mathcal{M}$  necessitates the estimation of the hyperparameter matrix  $\Gamma$ . Moreover, it can be observed from (37) that all the components of the  $i$ th group tend to zero, as the associated hyperparameter  $\gamma_i \rightarrow 0$  [31].

To this end, it is desirable to select the matrix  $\hat{\Gamma}$  that maximizes the log-Bayesian evidence  $\log \left[ f(\tilde{\mathbf{Y}}_{\text{TF},\mathcal{P}}; \Gamma) \right]$ , which evaluates to  $\log \left[ f(\tilde{\mathbf{Y}}_{\text{TF},\mathcal{P}}; \Gamma) \right] = c_1 - N_r \log \left[ \det(\tilde{\Sigma}_y) \right] - \sum_{r=1}^{N_r} \mathbf{y}_{\text{TF},\mathcal{P},r}^H (\tilde{\Sigma}_y)^{-1} \mathbf{y}_{\text{TF},\mathcal{P},r}$ , where  $\tilde{\Sigma}_y = \mathbf{R}_v + \tilde{\Omega}(\mathbf{I}_{N_t} \otimes \Gamma) \tilde{\Omega}^H \in \mathbb{C}^{MN_p \times MN_p}$  denotes the covariance matrix of the received pilot vector  $\mathbf{y}_{\text{TF},\mathcal{P},r}$  and the constant  $c_1 = -MN_p N_r \log(\pi)$ . Since this log-Bayesian evidence maximization with respect to the hyperparameter matrix  $\Gamma$  is a non-concave problem [31], the resultant optimization problem becomes intractable. In such a scenario, the expectation-maximization (EM) method is an ideal tool for maximizing the cost function in each iteration. Additionally, the EM procedure guarantees convergence to a local optimum [32]. Therefore, the proposed RG-BL framework employs the EM technique for DD-domain sparse CSI estimation in MIMO OTFS systems. The key steps of this procedure are derived next. Let the complete information set be constructed as  $\{\tilde{\mathbf{Y}}_{\text{TF},\mathcal{P}}, \tilde{\mathbf{H}}\}$ . Let  $\hat{\Gamma}^{(j-1)} = \text{diag}\{\hat{\gamma}_i^{(j-1)}\}_{i=0}^{G-1} \in \mathbb{R}^{G \times G}$  denote the estimate of  $\Gamma$  in the  $(j-1)$ st EM-iteration. The procedure to update the estimate  $\hat{\Gamma}^{(j)}$  in the  $j$ th EM-iteration is described in Theorem-1 presented next.

**Theorem 1.** Given  $\hat{\gamma}_i^{(j-1)}$ ,  $0 \leq i \leq G-1$ , the hyperparameter update  $\hat{\gamma}_i^{(j)}$  in the  $j$ th EM iteration, which maximizes the conditional-expectation of the log-likelihood function of the complete information set  $\{\tilde{\mathbf{Y}}_{\text{TF},\mathcal{P}}, \tilde{\mathbf{H}}\}$ , denoted by

$$\mathcal{L}(\Gamma | \hat{\Gamma}^{(j-1)}) = \mathbb{E}_{\tilde{\mathbf{H}} | \tilde{\mathbf{Y}}_{\text{TF},\mathcal{P}}, \hat{\Gamma}^{(j-1)}} \left\{ \log \left[ f(\tilde{\mathbf{Y}}_{\text{TF},\mathcal{P}}, \tilde{\mathbf{H}}; \Gamma) \right] \right\}, \quad (39)$$

is given by

$$\hat{\gamma}_i^{(j)} = \frac{1}{N_r N_t} \sum_{r=1}^{N_r} \sum_{t=1}^{N_t} \left| \mathcal{M}^{(j)}[(t-1)N_t + i, r-1] \right|^2 + \frac{1}{N_t} \sum_{t=1}^{N_t} \tilde{\Sigma}^{(j)}[(t-1)N_t + i, (t-1)N_t + i], \quad (40)$$

where the quantities  $\mathcal{M}^{(j)} = \tilde{\Sigma}^{(j)} \tilde{\Omega}^H \mathbf{R}_v^{-1} \tilde{\mathbf{Y}}_{\text{TF},\mathcal{P}}$  and  $\tilde{\Sigma}^{(j)} =$



**Algorithm 2** RG-BL based sparse channel estimation in MIMO OTFS systems

**Input:** Observation matrix  $\tilde{\mathbf{Y}}_{\text{TF},\mathcal{P}}$ , dictionary matrix  $\tilde{\mathbf{\Omega}}$ , noise covariance matrix  $\mathbf{R}_v$ , stopping parameters  $\epsilon$  and  $N_{\max}$

**Initialization:**  $\hat{\gamma}_i^{(0)} = 1, \forall 0 \leq i \leq G-1, \hat{\mathbf{\Gamma}}^{(0)} = \mathbf{I}_G, \hat{\mathbf{\Gamma}}^{(-1)} = \mathbf{0}$ , set counter  $j = 0$

**while**  $\left( \left\| \hat{\mathbf{\Gamma}}^{(j)} - \hat{\mathbf{\Gamma}}^{(j-1)} \right\|_F^2 \geq \epsilon \text{ and } j < N_{\max} \right)$  **do**

1)  $j = j + 1$

2) **E-step:** Compute the *a posteriori* covariance and mean

$$\begin{aligned} \mathcal{M}^{(j)} &= \tilde{\mathbf{\Sigma}}^{(j)} \tilde{\mathbf{\Omega}}^H \mathbf{R}_v^{-1} \tilde{\mathbf{Y}}_{\text{TF},\mathcal{P}}, \\ \tilde{\mathbf{\Sigma}}^{(j)} &= \left[ \tilde{\mathbf{\Omega}}^H \mathbf{R}_v^{-1} \tilde{\mathbf{\Omega}} + \left( \mathbf{I}_{N_t} \otimes \left( \hat{\mathbf{\Gamma}}^{(j-1)} \right)^{-1} \right) \right]^{-1} \end{aligned}$$

3) **M-step:** Compute the hyperparameter estimates **for**  $i = 0, 1, \dots, G-1$  **do**

$$\begin{aligned} \hat{\gamma}_i^{(j)} &= \frac{1}{N_r N_t} \sum_{r=1}^{N_r} \sum_{t=1}^{N_t} \left| \mathcal{M}^{(j)} [(t-1)N_t + i, r-1] \right|^2 \\ &+ \frac{1}{N_t} \sum_{t=1}^{N_t} \tilde{\mathbf{\Sigma}}^{(j)} [(t-1)N_t + i, (t-1)N_t + i] \end{aligned}$$

**end for**

**end while**

**Output:**  $\hat{\mathbf{H}}_{\text{RG-BL}} = \mathcal{M}^{(j)}$

$$\left[ \tilde{\mathbf{\Omega}}^H \mathbf{R}_v^{-1} \tilde{\mathbf{\Omega}} + \left( \mathbf{I}_{N_t} \otimes \left( \hat{\mathbf{\Gamma}}^{(j-1)} \right)^{-1} \right) \right]^{-1}.$$

*Proof.* Proof is given in Appendix-A  $\square$

The EM procedure described above is iterated for a maximum of  $N_{\max}$  iterations or until  $\| \hat{\mathbf{\Gamma}}^{(j)} - \hat{\mathbf{\Gamma}}^{(j-1)} \|_F^2 \leq \epsilon$ , whichever is achieved earlier, where the stopping parameters  $\epsilon$  and  $N_{\max}$  are suitably chosen. Upon convergence, the RG-BL-based estimate  $\hat{\mathbf{H}}_{\text{RG-BL}}$  of the DD-domain RG-sparse CSI  $\bar{\mathbf{H}}$  is obtained as the converged *a-posteriori* mean, i.e.,  $\hat{\mathbf{H}}_{\text{RG-BL}} = \mathcal{M}^{(j)}$ , whose resultant estimation uncertainty is characterized by its error covariance matrix  $\tilde{\mathbf{\Sigma}}^{(j)}$ . Subsequently, the estimate  $\hat{\mathbf{H}}_{\text{RG-BL}}$  is used to obtain the estimate of the matrix  $\bar{\mathbf{H}}_{\text{DD}}$  that can be employed in the linear detector of Eq. (23). Note that in this work, although we do not consider the CSI estimation uncertainty, the resultant NMSE and BER, as illustrated in Section-V, are very close to their respective performance benchmarks. This justifies the efficacy of the proposed CSI estimation schemes.

The estimate  $\hat{\mathbf{H}}_{\text{DD}}$  can be obtained as follows. First, using the relationships given in (32) and (34), the estimate of the element  $h_{i,r,t}^j$  of the vector  $\mathbf{h}_{r,t}$  of (30) is obtained as

$$\hat{h}_{i,r,t}^j = \hat{\mathbf{H}}_{\text{RG-BL}} [i(G_\nu + 1) + j + (t-1)G, r-1].$$

Next, the estimate  $\hat{h}_{i,r,t}^j$  is used for constructing the estimate

of the matrix  $\mathbf{H}_{\text{DD},r,t}$  of (20) as

$$\hat{\mathbf{H}}_{\text{DD},r,t} = (\mathbf{F}_N \otimes \mathbf{P}_{\text{rx}}) \left[ \sum_{i=0}^{M_\tau} \sum_{j=0}^{G_\nu} \hat{h}_{i,r,t}^j (\bar{\mathbf{\Pi}})^i \bar{\mathbf{\Delta}}_{i,j} \right] (\mathbf{F}_N^H \otimes \mathbf{P}_{\text{tx}}). \quad (41)$$

Finally, employing (22), the estimate of the CSI  $\bar{\mathbf{H}}_{\text{DD}}$  can be constructed using the estimates of its component matrices  $\mathbf{H}_{\text{DD},r,t}$  derived above. Thus, it can be clearly observed that in contrast to the existing contributions, such as [6], [14], our framework proposed for the DD-domain CSI estimation and the subsequent data detection does not require any knowledge of the number of dominant multipath components, which renders it eminently suitable for practical MIMO OTFS systems.

#### A. Delay-Doppler-Angular (DDA)-domain Channel Model

One can also employ an angular-domain geometric channel model [10] for the MIMO OTFS systems formulated as:

$$\mathbf{H}(\tau, \nu) = \sum_{i=1}^{L_p} \mathbf{H}_i \delta(\tau - \tau_i) \delta(\nu - \nu_i),$$

where the MIMO channel tap  $\mathbf{H}_i \in \mathbb{C}^{N_r \times N_t}$  is expressed in terms of the angle-of-arrival (AoA)  $\theta_i$  and angle-of-departure (AoD)  $\phi_i$  of the  $i$ th multipath component as  $\mathbf{H}_i = \alpha_i \mathbf{a}_r(\theta_i) \mathbf{a}_t^H(\phi_i)$ . In the model above,  $\alpha_i$  represents the complex-valued path gain, and  $\mathbf{a}_r(\theta_i) \in \mathbb{C}^{N_r \times 1}$  and  $\mathbf{a}_t(\phi_i) \in \mathbb{C}^{N_t \times 1}$  signify the receive and transmit array response vectors, respectively, which are defined as

$$\begin{aligned} \mathbf{a}_r(\theta_i) &= \left[ 1, e^{-j \frac{2\pi}{\lambda} d_r \cos(\theta_i)}, \dots, e^{-j \frac{2\pi}{\lambda} d_r (N_r-1) \cos(\theta_i)} \right], \\ \mathbf{a}_t(\phi_i) &= \left[ 1, e^{-j \frac{2\pi}{\lambda} d_t \cos(\phi_i)}, \dots, e^{-j \frac{2\pi}{\lambda} d_t (N_t-1) \cos(\phi_i)} \right], \end{aligned}$$

where  $\lambda$  denotes the wavelength, and  $d_r$  and  $d_t$  represent the receive and transmit antenna spacings, respectively. Employing the angular-domain channel model above, the quantity  $h_{i,r,t}$  of Eq. (18), which represents the channel coefficient for the  $i$ th multipath component between the  $r$ -th RA and  $t$ -th TA, is expressed as  $h_{i,r,t} = \mathbf{H}_i(r, t)$ . Since the MIMO OTFS system model and channel estimation schemes developed in Section-II and Section-III, respectively, consider a general scenario, the above angular domain channel model has been employed in our simulations in Section-V for illustrating the performance of the proposed approaches.

#### B. Computational Complexities

The key results of our complexity analysis are summarized below. The computational complexity order of the RG-BL technique may be shown to be  $\mathcal{O}(G^3 N_t^3)$ , which arises due to the matrix inversion of size- $[GN_t \times GN_t]$ . On the other hand, the worst-case complexity order of the RG-OMP scheme is seen to be  $\mathcal{O}(M^3 N_p^3)$ , which arises due to the intermediate LS estimate required in each iteration. Finally, the computational cost of the conventional EP-based [16] CSI estimator is seen to be on the the order of  $\mathcal{O}(NM_\tau N_t)$ , since it does not involve any matrix multiplication/ inversion operations. However, as discussed later in our simulation results, the performance of

TABLE II  
SIMULATION PARAMETERS OF SYSTEM-I

Parameter	Value
Carrier frequency (GHz)	4
Subcarrier spacing (KHz) ( $\Delta f$ )	15
# of symbols along delay-axis ( $M$ )	32
# of symbols along Doppler-axis ( $N$ )	32
# of pilots along time-axis ( $N_p$ )	8
# of samples in CP ( $L$ )	16
# of dominant reflectors ( $L_p$ )	5
Max. delay-spread ( $M_\tau$ )	16
Max. Doppler-spread ( $N_\nu$ )	8
Doppler-grid size $G_\nu$	8
Pulse-shape	Rectangular

TABLE III  
DD-PROFILE OF THE WIRELESS CHANNEL OF SYSTEM-I

Path-index( $i$ )	1	2	3	4	5
$\tau_i$ ( $\mu$ s)	2.08	4.16	6.24	8.32	10.41
$\nu_i$ (Hz)	0	470	940	1410	1880
Speed (Km/h)	0	126.9	253.8	380.7	507.6

the EP-based technique is significantly poor in comparison to our proposed techniques and its pilot overhead is also very high. Due to lack of space, the detailed derivations of the computational complexities of our various schemes have been relegated to our technical report [33].

## V. SIMULATION RESULTS

This section demonstrates the performance of the proposed BL-based scheme and compares it to that of the OMP [34] and FOCUSS [35] based sparse signal recovery techniques as well as to the state-of-the-art EP-based scheme of [16] for the DD-domain CSI estimation of the OTFS systems in terms of the normalized MSE (NMSE), defined as  $NMSE = \frac{\|\hat{\mathbf{H}}_{DD} - \mathbf{H}_{DD}\|_F^2}{\|\mathbf{H}_{DD}\|_F^2}$  and pilot overhead. Note that for a fair comparison, the pilot power of each impulse in the EP-based scheme is set as the total power of the  $MN_p$  pilot symbols of our proposed scheme. This is followed by illustrating the bit-error-rate (BER) performance of the linear MMSE detectors designed using the available CSI estimates. For the BL-based approach, the stopping parameters  $\epsilon$  and  $N_{max}$  are set to  $10^{-6}$  and 50, respectively. The SNR in decibels (dB) is defined as  $10 \log_{10} \left( \frac{1}{\sigma^2} \right)$ . Table-II and Table-III show the detailed system and channel parameters of a sub-6 GHz band channel in a high mobility scenario, termed as System-I, whereas Table-IV and Table-V summarize the same for a high-frequency millimeter wave band channel, termed as System-II, in the low- and average-mobility scenarios considered in our simulations, unless specified otherwise.

Fig. 4(a) and 4(b) compare the NMSE performance of the proposed BL-based scheme to that of the other competing approaches for the SISO OTFS System-I and System-II models, respectively. From the figure, we can readily observe

TABLE IV  
SIMULATION PARAMETERS OF SYSTEM-II

Parameter	Value
Carrier frequency (GHz)	28
Subcarrier spacing (KHz) ( $\Delta f$ )	78.125
# of symbols along delay-axis ( $M$ )	128
# of symbols along Doppler-axis ( $N$ )	128
# of pilots along time-axis ( $N_p$ )	16
# of samples in CP ( $L$ )	32
# of dominant reflectors ( $L_p$ )	5
Max. delay-spread ( $M_\tau$ )	32
Max. Doppler-spread ( $N_\nu$ )	16
Doppler-grid size $G_\nu$	16
Pulse-shape	Rectangular

TABLE V  
DD-PROFILE OF THE WIRELESS CHANNEL OF SYSTEM-II

Path-index( $i$ )	1	2	3	4	5
$\tau_i$ ( $\mu$ s)	0.3	1	1.7	2.4	3.1
$\nu_i$ (Hz)	0	610	1220	2440	3660
Speed (Km/h)	0	23.5	47	94	141

a significant NMSE improvement of the proposed RG-BL scheme in comparison to the MMSE, MFOCUSS, EP and RG-OMP approaches. Since the conventional MMSE and EP-based schemes do not exploit sparsity of the DD-domain CSI, they yield a poor NMSE. The poor performance of the RG-OMP can be attributed to its sensitivity to the stopping parameter as well as to the dictionary matrix, whereas the same for the FOCUSS arises due to its convergence deficiencies and sensitivity to the regularization parameter [31]. Thus, the performance of the competing sparse recovery algorithms, such as the RG-OMP and MFOCUSS, is not as robust as that of the BL-based approach.

Fig. 5(a) compares the NMSE of the proposed and existing sparse CSI estimation schemes for the setting of  $N_t = 2$  and  $N_r \in \{2, 4\}$  in the MIMO OTFS System-II. Observe from the Fig. 5(a) that the proposed RG-BL and RG-OMP techniques, specifically designed for MIMO OTFS CSI estimation, outperform the conventional OMP [34], BL [31] and EP-based techniques. This is owing to the fact that the RG-BL and RG-OMP techniques exploit both the row and group-sparsity of the DD-domain CSI  $\tilde{\mathbf{H}}$ , whereas the conventional BL and OMP techniques exploit only the sparsity. This demonstrates the robust performance of the RG-BL. Furthermore, its performance is seen to improve upon increasing the number of RAs  $N_r$ , a fact which is justified by the availability of more pilot outputs in  $\tilde{\mathbf{Y}}_{TF, \mathcal{P}}$  and exploiting the row sparsity. By contrast, the performance of the BL, OMP and EP-based schemes remain unchanged upon increasing the number of RAs  $N_r$ . This is attributed to the fact the BL and OMP do not exploit the row-sparsity of the DD-domain CSI  $\tilde{\mathbf{H}}$ , whereas the EP-based scheme is unable to leverage the common support of the DD-domain CSI across all the TA-RA pairs.

The robust performance of the proposed RG-BL technique

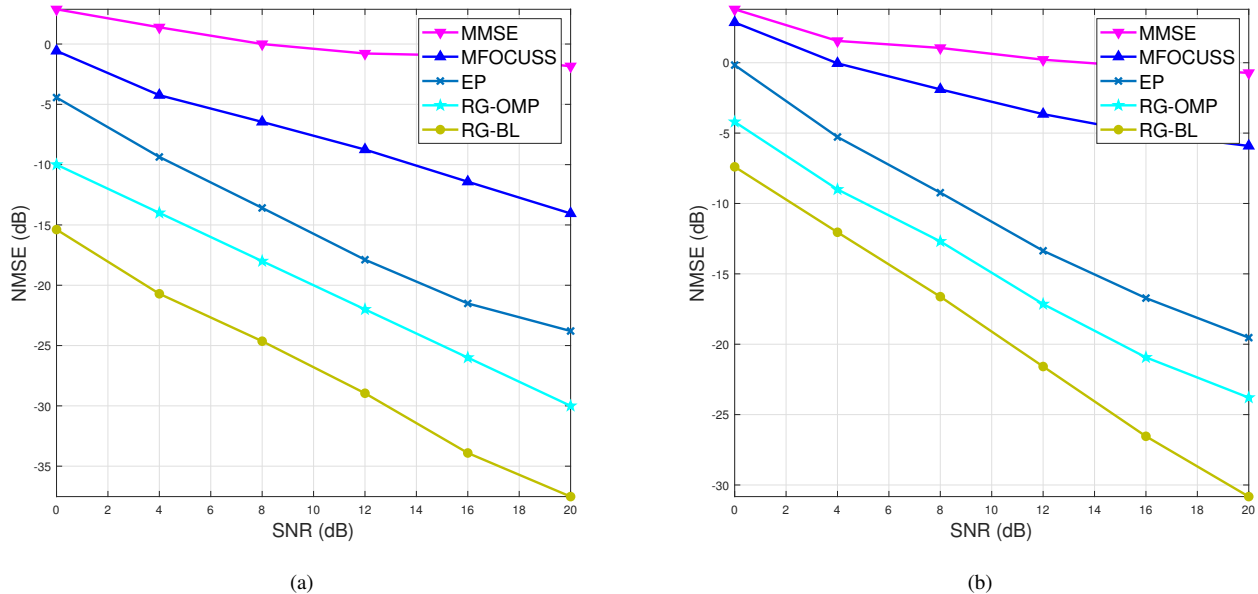


Fig. 4. NMSE versus SNR performance of the SISO OTFS for (a) System-I; (b) System-II.

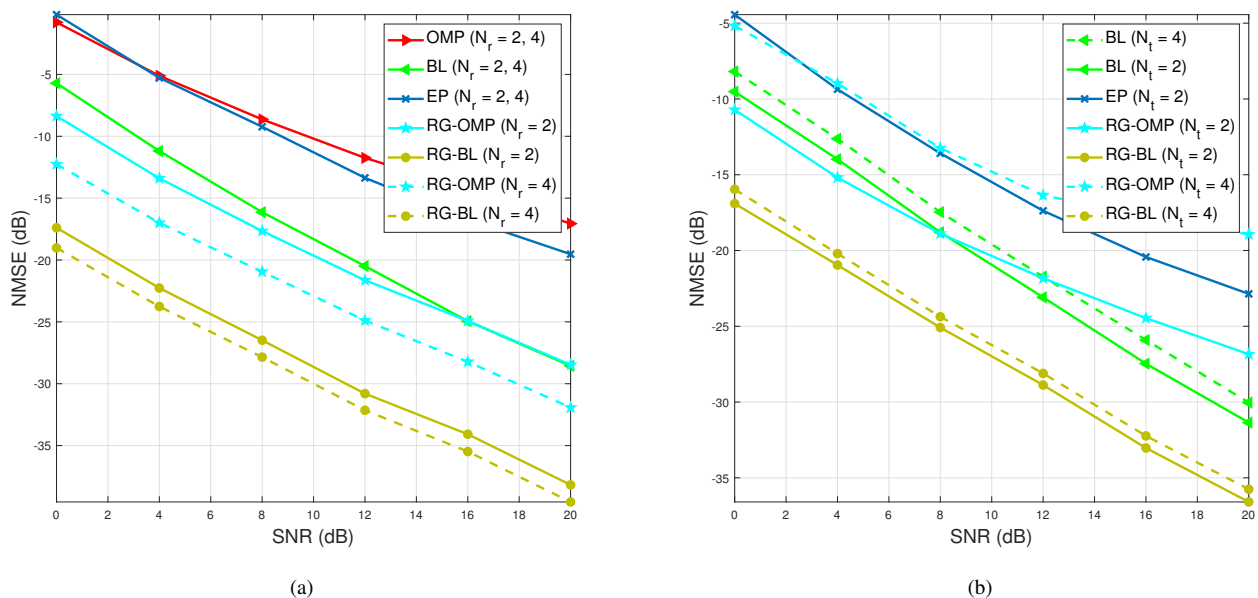


Fig. 5. NMSE versus SNR performance of the MIMO OTFS for (a) System-II with  $N_t = 2$ ,  $N_r \in \{2, 4\}$ ; (b) System-I with  $N_r = 2$ ,  $N_t \in \{2, 4\}$ .

is further demonstrated by considering a challenging scenario, where  $N_t > N_r$ . Fig. 5(b) compares the NMSE performance of the various sparse CSI estimation schemes for MIMO OTFS System-I having  $N_r = 2$  and  $N_t \in \{2, 4\}$ . Here, the CSI estimation becomes challenging for  $N_t = 4$ , since using the estimation model of (33), one has to infer  $GN_t N_r = 1224$  channel coefficients from only  $MN_p N_r = 512$  pilot outputs, resulting in a heavily ‘under-determined’ or ‘ill-posed’ estimation problem. It can be readily observed that in this scenario, the NMSE of the OMP- and FOCUSS-based schemes degrades significantly, whereas the proposed

RG-BL technique continues to perform best. The significant performance improvement of the proposed RG-BL technique is attributed to the fact that as shown in (37), it assigns only  $G$  distinct hyperparameters to characterize the *a priori* pdf of the  $[GN_t \times N_r]$ -element matrix  $\tilde{\mathbf{H}}$ . Thus, the RG-BL scheme has to estimate much fewer hyperparameters, namely only  $G$ , from the  $[MN_p \times N_r]$ -element pilot output matrix  $\tilde{\mathbf{Y}}_{\text{TF}, \mathcal{P}}$ , yielding an enhanced estimate of the hyperparameters, and in turn of the CSI matrix  $\tilde{\mathbf{H}}$ .

Finally, Fig. 6(a) portrays the BER achieved using the linear MMSE detectors derived from all the contending sparse recov-

ery schemes. For this, the estimate of the matrix  $\widehat{\mathbf{H}}_{\text{DD}}$  has been constructed from its equivalent DD-domain CSI estimates, like  $\widehat{\mathbf{H}}_{\text{RG-OMP}}$  or  $\widehat{\mathbf{H}}_{\text{RG-BL}}$ , using the procedure described right before Section-IV-A. Observe from Fig. 6(a) that owing to its improved estimation accuracy, the BER achieved using the detectors derived from the BL-based channel estimates is markedly better than that of its OMP and FOCUSS counterparts. Moreover, the proposed RG-BL technique is also seen to approach the BER of a hypothetical receiver having perfect CSI (PCSI), which demonstrates its enhanced ability of channel recovery.

Coming now to the required pilot overhead, note that for the proposed CSI estimation framework,  $MN_p$  symbols are pilots in a block of  $M(N + N_p)$  symbols. This leads to the normalized pilot overhead of  $\rho = \frac{N_p}{N+N_p}$ . On the other hand, the pilot overhead for the state-of-the-art EP-based technique [16] is given by  $\rho_{\text{EP}} \approx \frac{(N_t M_\tau + M_\tau + N_t)(2N_\nu + 1)}{MN}$  for integer-Doppler scenario, and it is even higher for fractional-Doppler case. Table-VI compares the pilot overhead of the proposed CSI estimation framework to the EP-based scheme. It can be readily observed that our proposed sparse CSI estimation framework has a significantly low pilot overhead in comparison to the EP-based scheme. At the same time, from the NMSE plots of Fig. 4 and 5, one can observe a significantly improved NMSE of the proposed RG-BL scheme in comparison to the EP-based method considering an identical total pilot power. The high pilot overhead of the EP-based scheme is attributed to the fact that it requires multiple DD-domain guards around the pilot symbols of each TA, and hence, its pilot overhead increases upon increasing  $N_t$ . However, our proposed CSI estimation framework completely avoids this by providing a common overlapping pilot duration for all the TAs. Interestingly, for the OTFS System-I, the pilot overhead of the EP-based scheme is significantly high even for  $N_t = 1$ , and it becomes quite impractical for  $N_t = 2$  and beyond. On the other hand, to a large extent, the pilot overhead remains constant for the proposed channel estimation framework upon increasing the number of TAs, thereby making it practically appealing for large-scale antenna scenarios.

Another unique aspect of the proposed channel estimation model is that the pilot symbols  $N_p$  can be flexibly adjusted to strike a suitable estimation accuracy versus pilot overhead trade-off. In Fig. 6(b), the NMSE performance of the various competing sparse CSI estimation frameworks is compared by varying the pilot overhead. The SNR for this simulation is set to 4 dB. It can be observed that the NMSE performance improves upon increasing the pilot overhead  $\rho$ , a direct consequence of the fact that the number of pilot measurements  $MN_p N_r$ , which is essentially the number of elements in the matrix  $\mathbf{Y}_{\text{TF},\rho}$ , increases upon increasing  $N_p$ . Moreover, it is also observed that the NMSE of the RG-BL technique with  $\rho = 0.2$  is lower than that of the RG-OMP with  $\rho = 0.33$ . Hence, for a desired level of estimation accuracy, the RG-BL scheme can lead to a significantly reduced pilot overhead. Thus, the improved CSI estimation performance, reduced pilot overhead along with low-complexity linear detectors requiring no prior knowledge of the number of dominant reflectors make

the proposed RG-BL scheme ideally suited for implementation in practical MIMO OTFS systems.

Fig. 7(a) and 7(b) illustrate the NMSE and BER performance, respectively, of the proposed schemes considering fractional Dopplers. For these simulations, the fractional Doppler  $\kappa_{\nu_i}$  for each multipath component has been uniformly generated in the range  $(-\frac{1}{2}, \frac{1}{2})$ . As described in Section-III, in order to handle this scenario, the Doppler grid-size  $G_\nu$  is set much higher than the maximum integer-Doppler tap  $N_\nu$ . The performance of the proposed schemes is also benchmarked against the integer-Doppler case. It can be readily observed that for the fractional-Doppler scenario, the performance of the proposed schemes is degraded slightly compared to the integer-Doppler scenario, but it is still very close to that of their integer-Doppler counterparts. This performance degradation is due to the fact that the true Doppler shifts  $k_i$  of all the multipath components differ from the Doppler-grid points that are used to construct the dictionary matrix  $\widehat{\mathbf{\Omega}}$ . As expected, the performance of the proposed techniques improves upon increasing the Doppler grid-size  $G_\nu$ , since it enhances the resolution of the grid. Furthermore, even in this scenario, the RG-BL technique continues to perform best. This demonstrates that the proposed channel estimation model and schemes are eminently suitable for practical OTFS systems.

Fig. 8(a) and 8(b) demonstrate the NMSE and BER, respectively, of the proposed schemes considering a discrete prolate spheroidal (DPS) [30], [36] waveform for the transmit pulse-shape  $p_{\text{tx}}(t)$ , which has a significantly reduced out-of-band power in comparison to the rectangular waveform. Note that for a rectangular waveform, all the time-domain transmit symbols are scaled identically. On the other hand, due to the structure of the DPS pulse-shape, the symbols lying on both the edges of the pulse duration  $T$  are scaled down significantly, which degrades the resultant NMSE and BER performance. Furthermore, it can be observed that the BER performance for the DPS waveform degrades significantly even for perfect CSI knowledge. However, the degradation in NMSE is not significant, which makes the proposed channel estimation frameworks suitable for all kinds of practical pulse-shapes.

## VI. CONCLUSIONS

This paper proposed a novel simultaneous row- and group-sparse channel estimation model considering fractional Dopplers and arbitrary Tx-Rx pulse shapes for MIMO OTFS systems, wherein the TF-domain pilots are transmitted to estimate the DD-domain CSI of the underlying wireless channel. In this context, both the novel RG-OMP, as well as the enhanced RG-BL frameworks were developed, which exploit the DD-domain RG-sparsity for improved CSI estimation. The proposed CSI estimation framework significantly reduces the pilot overhead, training duration and pilot pre-processing complexity, while completely eliminating the need for placing multiple DD-domain guards between the pilot and data symbols corresponding to each TA within the same OTFS frame. Additionally, it has a flexible training duration and it can also efficiently handle fractional-Doppler scenarios. All of these attributes of the proposed framework lead to increased spectral

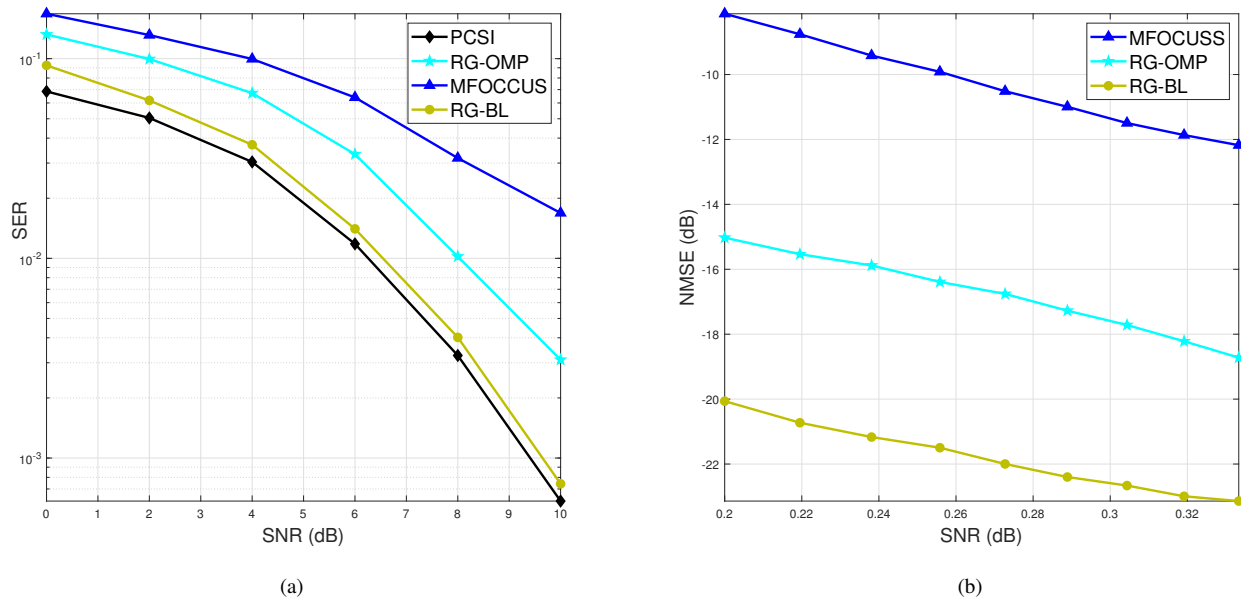


Fig. 6. (a) BER versus SNR performance of the MIMO OTFS System-I with  $N_t = N_r = 2$ ; (b) NMSE versus pilot overhead  $\rho$  for the SISO OTFS System-I.

TABLE VI

PILOT OVERHEAD COMPARISON IN THE FORM  $(\rho, \rho_{EP})$ , WHERE  $\rho$  AND  $\rho_{EP}$  DENOTE THE PILOT OVERHEAD OF THE PROPOSED AND EP-BASED SCHEMES, RESPECTIVELY, AND 'NA' REPRESENTS 'NOT APPLICABLE'

Number of TAs ( $N_t$ )	1	2	3	4
System-I	(0.2, 0.54)	(0.2, 0.83)	(0.2, NA)	(0.2, NA)
System-II	(0.11, 0.13)	(0.11, 0.197)	(0.11, 0.26)	(0.11, 0.33)

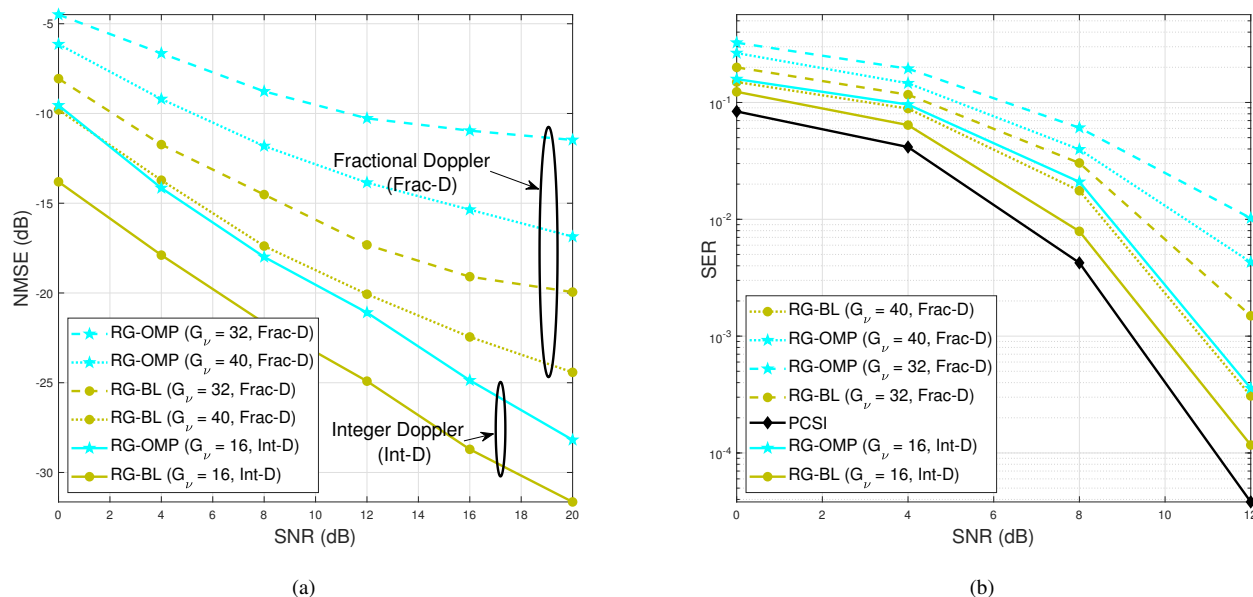


Fig. 7. (a) NMSE versus SNR and (b) BER versus SNR performance, with integer and fractional Dopplers.

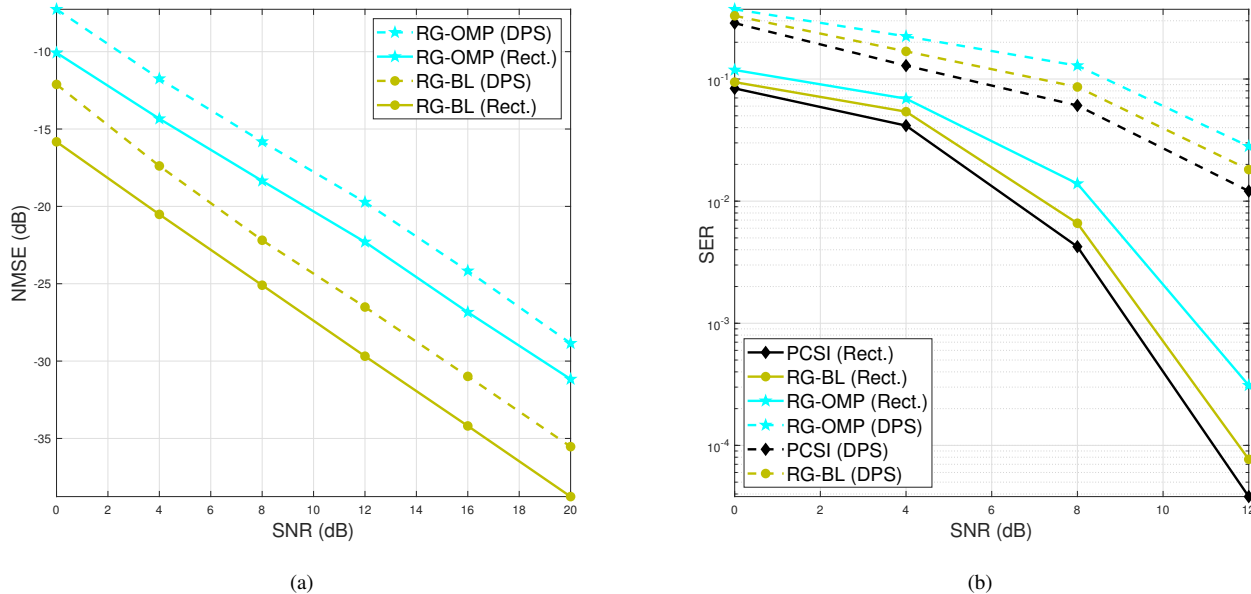


Fig. 8. (a) NMSE versus SNR and (b) BER versus SNR performance, with rectangular and DPS pulse shapes.

efficiency. Furthermore, low-complexity linear detectors are also derived for the data detection phase, which directly employ the sparse CSI estimated by the proposed schemes, without requiring any further knowledge of the number of dominant multipath components of the channel. Our simulation results demonstrated the performance benefits of the proposed schemes over the conventional MMSE scheme as well as over other state-of-the-art sparse recovery schemes.

#### APPENDIX A PROOF OF THEOREM-1

The expectation-step (E-step) evaluates  $\mathcal{L}(\mathbf{\Gamma}|\hat{\mathbf{\Gamma}}^{(j-1)})$  as

$$\mathbb{E}\left\{\log\left[f\left(\tilde{\mathbf{Y}}_{\text{TF},\mathcal{P}}|\tilde{\mathbf{H}}\right)\right]\right\} + \mathbb{E}\left\{\log\left[f\left(\tilde{\mathbf{H}};\mathbf{\Gamma}\right)\right]\right\}. \quad (42)$$

Subsequently, the maximization-step (M-step) maximizes  $\mathcal{L}(\mathbf{\Gamma}|\hat{\mathbf{\Gamma}}^{(j-1)})$  with respect to  $\mathbf{\Gamma}$  as

$$\hat{\mathbf{\Gamma}}^{(j)} = \arg \max_{\mathbf{\Gamma}} \mathbb{E}\left\{\log\left[f\left(\tilde{\mathbf{H}};\mathbf{\Gamma}\right)\right]\right\}, \quad (43)$$

since the first term within the  $\mathbb{E}\{\cdot\}$  in (42) is independent of  $\mathbf{\Gamma}$ . Upon substituting  $\log\left[f\left(\tilde{\mathbf{H}};\mathbf{\Gamma}\right)\right]$  from (37), and ignoring the terms that do not depend on  $\mathbf{\Gamma}$ , one obtains

$$\hat{\mathbf{\Gamma}}^{(j)} \equiv \arg \max_{\mathbf{\Gamma}} \sum_{i=0}^{G-1} \left[ -N_r N_t \log(\gamma_i) - \sum_{r=1, t=1}^{N_r, N_t} \frac{\mathbb{E}\{|\mathbf{h}_{r,t}(i)|^2\}}{\gamma_i} \right].$$

Differentiating the objective function with respect to  $\gamma_i$  and setting it equal to zero yields the estimates  $\hat{\gamma}_i^{(j)}$  as

$$\hat{\gamma}_i^{(j)} = \frac{1}{N_r N_t} \sum_{r=1}^{N_r} \sum_{t=1}^{N_t} \mathbb{E}_{\tilde{\mathbf{H}}|\tilde{\mathbf{Y}}_{\text{TF},\mathcal{P}};\hat{\mathbf{\Gamma}}^{(j-1)}} \left\{ |\mathbf{h}_{r,t}(i)|^2 \right\}. \quad (44)$$

In order to evaluate the conditional expectation  $\mathbb{E}_{\tilde{\mathbf{H}}|\tilde{\mathbf{Y}}_{\text{TF},\mathcal{P}};\hat{\mathbf{\Gamma}}^{(j-1)}}\{\cdot\}$ , one can employ the *a posteriori* pdf  $f\left(\tilde{\mathbf{H}}|\tilde{\mathbf{Y}}_{\text{TF},\mathcal{P}};\hat{\mathbf{\Gamma}}^{(j-1)}\right)$  of  $\tilde{\mathbf{H}}$ , which is evaluated as [32]

$$f\left(\tilde{\mathbf{H}}|\tilde{\mathbf{Y}}_{\text{TF},\mathcal{P}};\hat{\mathbf{\Gamma}}^{(j-1)}\right) = \mathcal{CN}\left(\mathcal{M}^{(j)}, \tilde{\mathbf{\Sigma}}^{(j)}\right). \quad (45)$$

In the above expression, the *a posteriori* covariance matrix  $\tilde{\mathbf{\Sigma}}^{(j)} \in \mathbb{C}^{GN_t \times GN_t}$  is obtained by setting  $\mathbf{\Gamma} = \hat{\mathbf{\Gamma}}^{(j-1)}$  in the expression of the MMSE error covariance  $\tilde{\mathbf{\Sigma}}$  of (38), whereas the *a posteriori* mean  $\mathcal{M}^{(j)} \in \mathbb{C}^{GN_t \times N_r}$  is obtained by substituting  $\tilde{\mathbf{\Sigma}} = \tilde{\mathbf{\Sigma}}^{(j)}$  into the expression of  $\mathcal{M}$  of (38). Thus, the conditional expectation in (44) obeys

$$\mathbb{E}\left\{|\mathbf{h}_{r,t}(i)|^2\right\} = \left| \mathcal{M}^{(j)}[(t-1)N_t + i, r-1] \right|^2 + \tilde{\mathbf{\Sigma}}^{(j)}[(t-1)N_t + i, (t-1)N_t + i],$$

which upon substituting into (44), yields the desired hyperparameter update of (40).

#### REFERENCES

- [1] F. Hasegawa, A. Taira, G. Noh, B. Hui, H. Nishimoto, A. Okazaki, A. Okamura, J. Lee, and I. Kim, "High-speed train communications standardization in 3GPP 5G NR," *IEEE Communications Standards Magazine*, vol. 2, no. 1, pp. 44–52, 2018.
- [2] Y. Liu, C.-X. Wang, and J. Huang, "Recent developments and future challenges in channel measurements and models for 5G and beyond high-speed train communication systems," *IEEE Communications Magazine*, vol. 57, no. 9, pp. 50–56, 2019.
- [3] S. Chen, J. Hu, Y. Shi, Y. Peng, J. Fang, R. Zhao, and L. Zhao, "Vehicle-to-everything (V2X) services supported by LTE-based systems and 5G," *IEEE Communications Standards Magazine*, vol. 1, no. 2, pp. 70–76, 2017.
- [4] K. Wevers and M. Lu, "V2X communication for ITS-from IEEE 802.11p towards 5G," *IEEE 5G Tech Focus*, vol. 1, no. 2, pp. 5–10, 2017.
- [5] J. Zhang, T. Chen, S. Zhong, J. Wang, W. Zhang, X. Zuo, R. G. Maunder, and L. Hanzo, "Aeronautical Ad-Hoc networking for the internet-above-the-clouds," *Proceedings of the IEEE*, vol. 107, no. 5, pp. 868–911, 2019.

- [6] 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.
- [7] R. Hadani and A. Monk, "OTFS: A new generation of modulation addressing the challenges of 5G," *ArXiv preprint, ArXiv:1802.02623*, 2018.
- [8] L. Hanzo, M. Münster, B. Choi, and T. Keller, *OFDM and MC-CDMA for broadband multi-user communications, WLANs and broadcasting*. John Wiley & Sons, 2005.
- [9] L. Hanzo, Y. Akhtman, J. Akhtman, L. Wang, and M. Jiang, *MIMO-OFDM for LTE, WiFi and WiMAX: Coherent versus non-coherent and cooperative turbo transceivers*. John Wiley & Sons, 2011.
- [10] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.
- [11] W. C. Jakes and D. C. Cox, *Microwave mobile communications*. Wiley-IEEE Press, 1994.
- [12] 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.
- [13] 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, ArXiv:1808.00519*, 2018.
- [14] 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.
- [15] W. Shen, L. Dai, J. An, P. Fan, and R. W. Heath, "Channel estimation for orthogonal time frequency space (OTFS) massive MIMO," *IEEE Transactions on Signal Processing*, vol. 67, no. 16, pp. 4204–4217, 2019.
- [16] 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.
- [17] Z. Ding, R. Schober, P. Fan, and H. V. Poor, "OTFS-NOMA: An efficient approach for exploiting heterogenous user mobility profiles," *IEEE Transactions on Communications*, vol. 67, no. 11, pp. 7950–7965, 2019.
- [18] K. Deka, A. Thomas, and S. Sharma, "OTFS-SCMA: A code-domain NOMA approach for orthogonal time frequency space modulation," *IEEE Transactions on Communications*, 2021.
- [19] Y. Ge, Q. Deng, P. Ching, and Z. Ding, "OTFS signaling for uplink NOMA of heterogeneous mobility users," *IEEE Transactions on Communications*, 2021.
- [20] 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.
- [21] L. Zhao, W.-J. Gao, and W. Guo, "Sparse Bayesian learning of delay-Doppler channel for OTFS system," *IEEE Communications Letters*, vol. 24, no. 12, pp. 2766–2769, 2020.
- [22] P. Raviteja, K. T. Phan, Y. Hong, and E. Viterbo, "Embedded delay-Doppler channel estimation for orthogonal time frequency space modulation," in *2018 IEEE 88th Vehicular Technology Conference (VTC-Fall)*. IEEE, 2018, pp. 1–5.
- [23] O. K. Rasheed, G. Surabhi, and A. Chockalingam, "Sparse delay-Doppler channel estimation in rapidly time-varying channels for multiuser OTFS on the uplink," in *2020 IEEE 91st Vehicular Technology Conference (VTC-Spring)*. IEEE, 2020, pp. 1–5.
- [24] S. S. Rakib and R. Hadani, "Orthogonal time frequency space modulation system," Mar. 27 2018, US Patent 9,929,783.
- [25] K. Murali and A. Chockalingam, "On OTFS modulation for high-Doppler fading channels," in *2018 Information Theory and Applications Workshop (ITA)*. IEEE, 2018, pp. 1–10.
- [26] 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.
- [27] Z. Wei, W. Yuan, S. Li, J. Yuan, and D. W. K. Ng, "Off-grid channel estimation with sparse Bayesian learning for OTFS systems," *arXiv preprint arXiv:2101.05629*, 2021.
- [28] W. Yuan, Z. Wei, J. Yuan, and D. W. K. Ng, "A simple variational bayes detector for orthogonal time frequency space (OTFS) modulation," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 7, pp. 7976–7980, 2020.
- [29] 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*, 2021.
- [30] 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.
- [31] D. P. Wipf and B. D. Rao, "Sparse Bayesian learning for basis selection," *IEEE Transactions on Signal processing*, vol. 52, no. 8, pp. 2153–2164, 2004.
- [32] S. M. Kay, *Fundamentals of Statistical Signal Processing*. Prentice Hall PTR, 1993.
- [33] S. Srivastava, R. K. Singh, A. K. Jagannatham, and L. Hanzo, "Technical report: Bayesian learning aided simultaneous row and group sparse channel estimation in orthogonal time frequency space modulated MIMO systems." IIT Kanpur, Tech. Rep., 2021. [Online]. Available: [http://www.iitk.ac.in/mwn/documents/MWNLab\\_TR\\_BL\\_MIMO\\_OTFS.pdf](http://www.iitk.ac.in/mwn/documents/MWNLab_TR_BL_MIMO_OTFS.pdf).
- [34] T. T. Cai and L. Wang, "Orthogonal matching pursuit for sparse signal recovery with noise," *IEEE Transactions on Information theory*, vol. 57, no. 7, pp. 4680–4688, 2011.
- [35] I. F. Gorodnitsky and B. D. Rao, "Sparse signal reconstruction from limited data using FOCUSS: A re-weighted minimum norm algorithm," *IEEE Transactions on Signal Processing*, vol. 45, no. 3, pp. 600–616, 1997.
- [36] D. Slepian, "Prolate spheroidal wave functions, Fourier analysis, and uncertainty—v: The discrete case," *Bell System Technical Journal*, vol. 57, no. 5, pp. 1371–1430, 1978.