OTFS Transceiver Design and Sparse Doubly-Selective CSI Estimation in Analog and Hybrid Beamforming Aided mmWave MIMO Systems

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

***Abstract*—Orthogonal time frequency space (OTFS) waveform based millimeter wave (mmWave) MIMO systems are capable of achieving high data rates in high-mobility scenarios. Hence, transceivers are designed for both analog beamforming (AB) and hybrid beamforming (HB), where we commence by deriving the delay-Doppler (DD)-domain input-output relationship considering a delay-Doppler-angular domain channel model. Subsequently, a novel two-stage procedure is developed for transmit beamformer (TBF)/ precoder (TPC) and receiver combiner (RC) design, and for estimating the DD-domain's equivalent channel state information (CSI). The key feature of the proposed framework is that the RF TBF/ TPC and RC design maximizes the directional beamforming gains. It is also demonstrated that the low-dimensional baseband CSI of the DD-domain becomes sparse for mmWave-AB MIMO OTFS systems, and block-sparse for mmWave-HB MIMO OTFS systems. Subsequently, Bayesian learning (BL) and block-sparse BL (BS-BL) solutions are developed for improved CSI estimation. We also derive the Bayesian Cramer-Rao lower bounds (BCRLB) for benchmarking the mean-squared-error (MSE) of the CSI estimates. Finally, our simulation results demonstrate the improved efficacy of the proposed transceiver designs and confirm the enhanced CSI estimation performance of the BL-based schemes over other competing sparse signal recovery schemes.**

***Index Terms*—OTFS, mmWave, high-Doppler, analog beamforming, hybrid precoding, delay-Doppler-angular domain channel, block sparsity, CSI estimation, BCRLB**

I. Introduction

The next generation 6G cellular networks are expected to support high-speed communication in ultra-high-Doppler scenarios, where the mobile velocities can go up to 400-500 km/h for high-speed trains (HSTs) [1], [2], or even up to 800-1000 km/h for air-plane users. The impact of Doppler further increases at high carrier frequencies, such as in the millimeter wave (mmWave) regime and beyond [3], [4]. In such scenarios, the performance of conventional multicarrier modulation techniques, such as orthogonal frequency division multiplexing (OFDM), degrades significantly due to the resultant inter-carrier-interference (ICI) arising from the large Doppler-shifts. In this context, the recently developed orthogonal time frequency space (OTFS) modulation technique [5]–[7], has garnered a lot of attention from the wireless research community due to its superior capability of handling high-Doppler doubly-selective wireless channels. In contrast to existing solutions, the wireless channel in OTFS is modeled in the delay-Doppler (DD)-domain and the information symbols are multiplexed by direct mapping to the DD-domain grid [8], [9], which renders the transformed wireless channel relatively static over a longer observation interval [4], [5]. This makes the estimation of doubly-selective channels substantially easier than the existing approaches in next-generation wireless scenarios.

In line with the ongoing research in OTFS, a challenging problem is to explore its applicability for the mmWave band, which holds significant promise due to its bandwidth reserves. However, the realization of a robust mmWave communication link critically hinges on the employment of multiple-input multiple-output (MIMO) beamforming arrays that can successfully overcome the high propagation losses and signal blockage [3]. In such transceivers, it is practically infeasible to implement a dedicated radio frequency (RF) chain for each antenna element, which has led to the development of novel architectures based on analog beamforming (AB) and hybrid beamforming (HB) especially for the mmWave regime. The amalgamation of mmWave-AB and mmWave-HB MIMO architectures with the OTFS waveform has the potential of successfully exploiting the massive bandwidth available in the mmWave band to support high data rates, especially in the challenging high-Doppler scenarios that arise due to the high carrier frequency. And again, these novel architectures also enjoy the gains gleaned from highly directional beamforming, thanks to the large antenna array based mmWave MIMO deployment, together with low hardware and signal processing complexity due to the hybrid architecture.

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. A. Chockalingam acknowledges the support from J. C. Bose National Fellowship, Department of Science and Technology, Government of India. 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)

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). A. Chockalingam is with the Department of Electrical Communication Engineering, Indian Institute of Science, Bengaluru 560012, India (e-mail: achockal@iisc.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).

However, these advantages of the mmWave MIMO OTFS-based systems necessitate optimal design of the transceiver relying on the RF transmit beamformer (TBF)/ precoder (TPC) and receiver combiner (RC). This in turn requires accurate DD-domain channel state information (CSI). Furthermore, the CSI is also required at the receiver for reliable detection. Hence, the transceiver design and CSI estimation play a key role in realizing the high performance gains promised by mmWave MIMO OTFS systems, which form the prime focus of this treatise.

A. Literature Review

In the context of mmWave communication, the initial studies [3], [10], [11] focused mainly on channel modeling, propagation characteristics, antenna design and link budget analysis at different carrier frequencies. Alkhateeb *et al.* [12] developed a procedure for low-complexity sparse channel estimation and multi-resolution codebook design for determining the TPC weights of mmWave MIMO systems, incorporating also the hybrid signal processing architecture of such systems. Ayach *et al.* [13] proposed a novel technique that exploited the spatially sparse nature of the mmWave MIMO channels to formulate the hybrid TPC/ RC design as a sparse reconstruction problem, which was subsequently solved using the simultaneous orthogonal matching pursuit (SOMP) algorithm. Along similar lines, the authors of [14], [15] also exploited the sparsity of the mmWave MIMO channel to develop techniques that achieved significantly improved CSI estimation performance at a reduced pilot overhead using the OMP and sparse Bayesian learning (SBL) techniques, respectively. It must however be noted that [12]–[15] only consider narrowband flat-fading channel estimation, whereas a typical mmWave MIMO channel is frequency-selective. To overcome this drawback, the authors of [16], [17] developed OMP-based sparse channel estimation techniques both for single-carrier and multicarrier wideband mmWave MIMO systems.

However, the above early seminal contributions did not account for the time variation of the mmWave MIMO channel arising due to mobility. The authors of [18]–[20] considered a doubly-selective temporally-correlated block-fading channel in their wideband mmWave MIMO systems, wherein the time-variation of the gains of the multipath components has been modeled by a first-order autoregressive (AR) process. However, the AR model employed requires knowledge of the temporal-correlation coefficient of the channel, which has been popularly evaluated using Jake's model [21], [22]. Furthermore, their application is limited to low-mobility scenarios, wherein the temporal-correlation is high, which justifies the assumption of a doubly-selective block-fading channel. The recent paper [23] by Gao *et al.* presented a novel three-stage scheme for the estimation of a doubly-selective mmWave MIMO channel considering both realistic multipath delays as well as Doppler shifts. A key drawback of all these existing approaches for doubly-selective CSI estimation is that they are based on the conventional time-frequency (TF)-domain representation of the channel, which varies over the entire TF-grid. This in turn requires channel estimation to be carried out more frequently, leading to significantly higher transmission and computational overheads. Furthermore, their ability to handle ultra-high-Doppler use cases is also limited. The recently proposed OTFS modulation addresses precisely this problem and has gained immense popularity due to its capability of overcoming the effects of high-Doppler. A brief review of the relevant OTFS literature is presented next.

OTFS, which is a revolutionary new modulation technique, was originally proposed by Hadani *et al.* in [5]–[7], which clearly illustrated its advantages over the well-known OFDM waveform for a doubly-selective channel. Subsequently, the authors of [24] derived the end-to-end DD-domain relationship in such a system considering ideal bi-orthogonal and practical rectangular pulse shapes. Several contributions, such as [8], [25]–[27], exploit the input-output model derived in [24] for CSI estimation in OTFS systems, and use either training impulses or embedded pilots. By contrast, a few recent studies [28]–[30] additionally exploit the DD-domain sparsity of the channel, thus achieving superior CSI estimation performance in comparison to the previous works. However, an impediment of the DD-domain CSI estimation techniques is that they require a DD-domain guard interval around the pilot location in order to avoid interference from the data symbols belonging to the same OTFS frame. Thus, the pilot overhead of these schemes is typically very high. This problem is further exacerbated in MIMO OTFS systems, especially for a large number of antennas, since one has to place multiple guard intervals at each transmit antenna. The recent proposal in [31] and its MIMO extension [32] have addressed this issue by considering TF-domain pilots, where the pilot symbols of all the transmit antennas are transmitted in a common TF resource block, thereby leading to a significant reduction in the pilot overhead.

A detailed overview of OTFS and various research opportunities have been given in the recent survey papers [4], [33]. Interested readers can also look at the following references for a detailed understanding of OTFS. The author of [34] derives OTFS modulation from first principles. First, using the continuous ZAK transform, the author has derived the orthonormal-basis for approximately time- and band-limited signals, which are also simultaneously localized in the DD-domain. Next, the OTFS waveform is obtained from the orthonormal-basis derived. On the other hand, the pre-print [35] derives the OTFS model using the discrete ZAK transform, which is motivated by its digital implementation similar to OFDM. Both [34] and [35] demonstrate that OTFS can be efficiently implemented independently without necessitating OFDM as its underlying block. The authors of [36] compare OTFS and OFDM in terms of their achievable rates considering the signal constellation and detector's soft-outputs in the presence of channel sparsity. More specifically, the authors design a pilot arrangement and the associated channel estimation algorithm for sparse DD-domain channels. Furthermore, they also target pilot overhead optimization in order to maximize the achievable rate. OTFS has also been demonstrated to have significant potential in enabling joint radar and communication [37], since the radar parameters, such as range and velocity, can be readily mapped to the delay and Doppler components, respectively, of the

TABLE I
KEY ATTRIBUTES OF THIS TREATISE IN COMPARISON TO THE EXISTING BL-BASED SPARSE DD-DOMAIN CSI ESTIMATION TECHNIQUES

	[31]	[32]	[41]	[30]	Our
mmWave MIMO architectures					✓
DDA-domain channel model					✓
DD-domain sparsity	✓	✓	✓	✓	✓
Angular-domain sparsity					✓
EM for hyperparameter est.	✓	✓	✓	✓	✓
Time-domain pilot insertion					✓
BCRLB	✓				✓
Block-/ group-sparsity		✓			✓
Fractional-Doppler		✓	✓	✓	✓
Interference-free pilot outputs	✓	✓			✓

CSI obtained for communication systems. The authors of [38] study the problem of window design for OTFS modulation to improve the performance of channel estimation and data detection considering a fractional-Doppler scenario. It has been demonstrated that windowing in the TF-domain can potentially increase the effective channel sparsity in the DD-domain for improved CSI estimation. Furthermore, it can also provide new degrees of freedom for further improving the detection performance. Another interesting contribution [39] analyzes the error performance of coded OTFS systems, which demonstrates a significant performance improvement over coded OFDM modulation in high-mobility channels, considering various channel codes, such as classical convolutional codes and state-of-the-art LDPC codes.

Although the contributions reviewed above and the references therein already form a rich literature on OTFS, there is a dearth of contributions which consider the high-frequency mmWave band and the associated challenges discussed above. In fact, while some works such as [5], [40], demonstrate the performance of a SISO OTFS system in the mmWave regime, none of them have explored mmWave MIMO OTFS systems, which holds the key for achieving ultra-high data-rates. Furthermore, there are significant challenges to be overcome for realizing the full potential of OTFS in mmWave MIMO systems using large antenna arrays. To this end, a list of our novel contributions is presented next.

B. Contributions of the Paper

1) This paper begins with deriving the DD-domain input-output relationship for OTFS-based mmWave-AB MIMO systems. Toward this end, a delay-Doppler-angular (DDA)-domain channel model has been developed, wherein each multipath component is characterized by its delay, Doppler, AoA and AoD components.
2) A two-stage procedure is developed for RF TBF/ RC design and for estimating the equivalent DD-domain CSI. The RF TBF/ RC pair is designed to have weights maximizing their directional gain. Subsequently, a sparse signal recovery based formulation has been developed for jointly estimating the equivalent complex path-gain of the multipath components along with their delay- and Doppler-shifts. To this end, a Bayesian learning (BL)-based procedure is also harnessed for estimating the resultant low-dimensional sparse DD-domain CSI.
3) Next, we extend the OTFS-aided transceiver design and CSI estimation procedure for mmWave-HB MIMO systems, which can support multiple parallel data streams for spatial multiplexing. Here, the key idea is to employ OTFS modulation/ demodulation at each RF chain, based on which the end-to-end system model is derived for signal detection.
4) The CSI estimation procedure developed for a mmWave-AB MIMO OTFS system is subsequently extended to a mmWave-HB MIMO OTFS system. Interestingly, the DD-domain CSI for this scenario follows a block-sparse structure, for which a novel block-sparse BL (BS-BL) approach is developed for exploiting this specific sparse structure. Furthermore, the Bayesian Cramer-Rao lower bound (BCRLB) is also developed for benchmarking the mean-squared-error (MSE) of the CSI estimates obtained using the proposed techniques.
5) It can be readily observed from Table-I that the added contributions of this proposed work are significant, since none of the existing papers in the literature comprehensively develop OTFS transceiver designs and sparse DD-domain CSI estimation techniques for analog and hybrid beamforming aided mmWave MIMO systems by exploiting the delay-Doppler-angular (DDA)-domain sparsity. The key novelty of our current study is to couple the beam-alignment procedure with DD-domain sparse CSI estimation in an efficient fashion, so that the overhead and feedback required is minimal.

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. The trace operator is denoted by $\mathrm{Tr}(\cdot)$, while $\mathbb{E}\{\cdot\}$ denotes the statistical expectation. The probability density function (pdf) of a complex Gaussian random vector having a mean vector of $\boldsymbol{\mu}$ and covariance matrix of $\boldsymbol{\Sigma}$ is denoted by $\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. $\otimes$ denotes the Kronecker product of two matrices.

II. OTFS-BASED MMWAVE ANALOG BEAMFORMING

Consider a mmWave-AB-based MIMO OTFS system having N_t transmit antennas (TAs) and N_r receive antennas (RAs). As shown in Fig. 1(a), all the TAs and RAs are fed through a single RF-chain using a network of digitally controlled phase-shifters [3]. Thus, we have a RF TBF $\mathbf{f}_{\mathrm{RF}} \in \mathbb{C}^{N_t \times 1}$ at the transmitter, while at the receiver, we have a RF RC $\mathbf{w}_{\mathrm{RF}} \in \mathbb{C}^{N_r \times 1}$. Note that the elements of the RF TBF and RC in a mmWave-AB MIMO system have a constant magnitude, which can be constrained to $|\mathbf{f}_{\mathrm{RF}}(i)| = \frac{1}{\sqrt{N_t}}, 1 \leq i \leq N_t$, and $|\mathbf{w}_{\mathrm{RF}}(j)| = \frac{1}{\sqrt{N_r}}, 1 \leq i \leq N_r$. Without loss of generality, let $\mathbf{H}(\tau, \nu) \in \mathbb{C}^{N_r \times N_t}$ represent a doubly-dispersive mmWave

MIMO channel, which is a 2D-function of the delay τ and Doppler ν. The signal $r(t)$ received at the output of the RF RC, in response to the input signal $s(t)$ transmitted by the transmit RF chain, is expressed as

$$r(t) = \int_{\tau}\int_{\nu} \mathbf{w}_{\text{RF}}^{H}\mathbf{H}(\tau,\nu)\mathbf{f}_{\text{RF}}s(t-\tau)e^{j2\pi\nu(t-\tau)}\mathrm{d}\tau\mathrm{d}\nu + \mathbf{w}_{\text{RF}}^{H}\tilde{\mathbf{w}}(t), \quad (1)$$

where $\tilde{\mathbf{w}}(t) \in \mathbb{C}^{N_r\times 1}$ denotes the spatially and temporally uncorrelated additive white Gaussian noise (AWGN) process of power σ^2. For simplicity, we introduce $w(t) = \mathbf{w}_{\text{RF}}^{H}\tilde{\mathbf{w}}(t)$, which is also temporally uncorrelated and has an average power σ^2. The procedure of obtaining an OTFS modulated waveform $s(t)$ at the transmitter is described next.

To this end, let Δf (Hz) denote the subcarrier spacing and T (seconds) represent the symbol duration, so that $T\Delta f = 1$. Furthermore, let M and N signify the number of symbols placed along the frequency and time axes, respectively, in the TF-grid. These TF-domain symbols are obtained via a suitable mapping from an equivalent DD-grid, where M and N represent the number of symbols placed along the delay and Doppler axes, respectively. Thus, the resultant OTFS system has a bandwidth $M\Delta f$ and frame duration NT. Accordingly, the delay and Doppler axes of the DD-domain grid are sampled at integer multiples of $\Delta\tau = \frac{1}{M\Delta f}$ and $\Delta\nu = \frac{1}{NT}$.

A. OTFS Modulation

The OTFS system at the transmitter first places an information symbol matrix $\mathbf{X}_{\text{DD}} \in \mathbb{C}^{M\times N}$ in the DD-grid, where the element $\mathbf{X}_{\text{DD}}(l,k)$ denote the DD-domain symbol placed at the delay index $l, 0 \le l \le M-1$, and Doppler index $k, 0 \le k \le N-1$. Subsequently, the transmitter employs the inverse symplectic finite Fourier transform (ISFFT) for mapping these DD-domain symbols onto the TF-grid. This operation is mathematically given by $\mathbf{X}_{\text{TF}} = \mathbf{F}_M\mathbf{X}_{\text{DD}}\mathbf{F}_N^H \in \mathbb{C}^{M\times N}$ [24], [42], where $\mathbf{F}_M$ and $\mathbf{F}_N$ represent the discrete Fourier transform (DFT) matrices of orders M and N, respectively. Here $\mathbf{X}_{\text{TF}}$ signifies the corresponding TF-domain symbol matrix, where the (m,n)th element $\mathbf{X}_{\text{TF}}(m,n)$ denotes the signal corresponding to the subcarrier-index m and time-index n. Subsequently, the time-domain signal $s(t)$ is obtained as follows. Let $p_{\text{tx}}(t)$ denote the transmit pulse of duration T. The discrete Heisenberg transform of the symbols $\mathbf{X}_{\text{TF}}(m,n)$, parameterized by the transmit pulse shaping filter $p_{\text{tx}}(t)$ [5], can be expressed as

$$s(t) = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1}\mathbf{X}_{\text{TF}}(m,n)p_{\text{tx}}(t-nT)e^{j2\pi m\Delta f(t-nT)}. \quad (2)$$

The DDA-domain model for the mmWave MIMO channel is described next.

B. DDA-Domain mmWave MIMO Wireless Channel Model

The DD-domain representation of the mmWave MIMO wireless channel is given by

$$\mathbf{H}(\tau,\nu) = \sum_{p=1}^{L_p}\mathbf{H}_p\delta(\tau-\tau_p)\delta(\nu-\nu_p), \quad (3)$$

where L_p signifies the number of reflectors/ multipath components. The quantities τ_p and ν_p denote the delay- and Doppler-shift, respectively, introduced by the the pth multipath component, whereas $\mathbf{H}_p \in \mathbb{C}^{N_r\times N_t}$ represents a matrix comprised of the complex path gains between each transmit-receive antenna pair. Furthermore, $\delta(\cdot)$ in (3) represents the Dirac-delta function. As described in [3], [11], the angular domain representation of the matrix $\mathbf{H}_p$ can be formulated as $\mathbf{H}_p = \sqrt{\frac{N_tN_r}{L_p}}h_p\mathbf{a}_r(\theta_p)\mathbf{a}_t^H(\phi_p)$, where θ_p and ϕ_p represent the AoA and AoD, respectively, corresponding to the pth multipath component, whereas h_p signifies its complex path-gain. The response of the transmit and receive antenna arrays can be characterized by their array steering vectors $\mathbf{a}_t(\cdot) \in \mathbb{C}^{N_t\times 1}$ and $\mathbf{a}_r(\cdot) \in \mathbb{C}^{N_r\times 1}$, respectively, which for a uniform linear array (ULA), are defined as

$$\mathbf{a}_t(\phi_p) = \frac{1}{\sqrt{N_t}}\left[1, e^{-j\frac{2\pi}{\lambda}d_t\cos(\phi_p)}, \ldots, e^{-j\frac{2\pi}{\lambda}(N_t-1)d_t\cos(\phi_p)}\right]^T, \quad (4)$$

$$\mathbf{a}_r(\theta_p) = \frac{1}{\sqrt{N_r}}\left[1, e^{-j\frac{2\pi}{\lambda}d_r\cos(\theta_p)}, \ldots, e^{-j\frac{2\pi}{\lambda}(N_r-1)d_r\cos(\theta_p)}\right]^T, \quad (5)$$

where the quantities λ, d_r, and d_t denote the signal wavelength, as well as the RA and TA spacings, respectively. Upon substituting $\mathbf{H}(\tau,\nu)$ from (3) into (1), we obtain

$$\begin{aligned} r(t) &= \sqrt{\frac{N_tN_r}{L_p}}\sum_{p=1}^{L_p}\Big[h_p\mathbf{w}_{\text{RF}}^H\mathbf{a}_r(\theta_p)\mathbf{a}_t^H(\phi_p)\mathbf{f}_{\text{RF}}s(t-\tau_p) \\ &\qquad e^{j2\pi\nu_p(t-\tau_p)}\Big] + w(t) \\ &= \sum_{p=1}^{L_p}\tilde{h}_p s(t-\tau_p)e^{j2\pi\nu_p(t-\tau_p)} + w(t), \end{aligned} \quad (6)$$

where $\tilde{h}_p = \sqrt{\frac{N_tN_r}{L_p}}h_p\mathbf{w}_{\text{RF}}^H\mathbf{a}_r(\theta_p)\mathbf{a}_t^H(\phi_p)\mathbf{f}_{\text{RF}}$. The Doppler-shift ν_p corresponding to the pth multipath component can be expressed as $\nu_p = \frac{j_p}{NT}$, where j_p can be written as $j_p = \text{round}(j_p) + \kappa_{\nu_p}$. Here, the quantity κ_{ν_p}, with $|\kappa_{\nu_p}| < \frac{1}{2}$, represents fractional-Doppler. On the other hand, as described in [22], [24], [27], 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 τ_p as $\tau_p = \frac{l_p}{M\Delta f}$, where l_p is an integer.

C. OTFS Demodulator

Let $p_{\text{rx}}(t)$ denote the receiver pulse shaping filter of duration T. The received signal $r(t)$ is first processed using a matched filter as $Y(f,t) = \int_{t'} r(t')p_{\text{rx}}^*(t'-t)e^{-j2\pi f(t'-t)}dt'$, which is then sampled at integer multiples of the subcarrier spacing Δf and symbol duration T to get $\mathbf{Y}_{\text{TF}}(m,n) = Y(f,t)|_{f=m\Delta f, t=nT}$. In the mathematics literature, these operations are commonly referred to as the discrete Wigner transform [5]. Thus, the matrix $\mathbf{Y}_{\text{TF}} \in \mathbb{C}^{M\times N}$ comprises the TF-domain demodulated symbols. Finally, the DD-domain demodulated OTFS signal, denoted by $\mathbf{Y}_{\text{DD}} \in \mathbb{C}^{M\times N}$, is obtained by performing the SFFT of the TF-domain demodulated

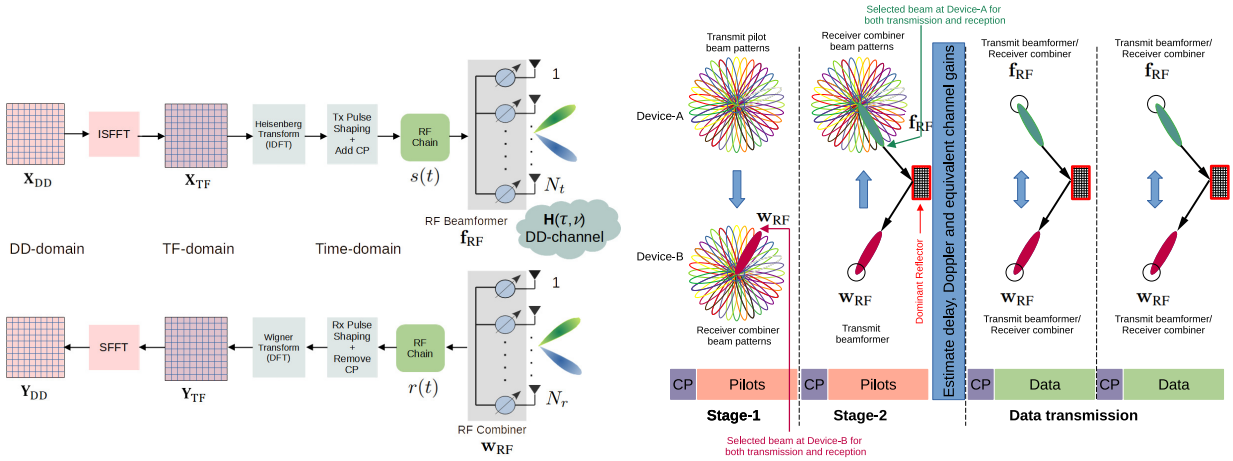

Fig. 1. (a) Architecture of the OTFS-based mmWave-AB MIMO transceiver; (b) Training frame structure of the proposed CSI estimation procedure for mmWave-AB MIMO systems.

signal $\mathbf{Y}_{\text{TF}}$, which is expressed as $\mathbf{Y}_{\text{DD}} = \mathbf{F}_M^H \mathbf{Y}_{\text{TF}} \mathbf{F}_N$ [24], [42]. The next subsection derives the end-to-end DD-domain relationship between the transmit and receive signals.

D. Input-Output Relationship of the mmWave-AB MIMO OTFS Systems

Since the transmit pulse shaping filter $p_{\text{tx}}(t)$ is assumed to be of duration T, the transmit signal $s_n(t), nT \leq t \leq (n+1)T$, corresponding to the nth symbol duration can be expressed using (2) as $s_n(t) = p_{\text{tx}}(t - nT) \sum_{m=0}^{M-1} \mathbf{X}_{\text{TF}}(m,n) e^{j2\pi m \Delta f (t-nT)}$. Sampling this signal at the Nyquist rate $\frac{M}{T}$, followed by rearranging the resultant M samples to form the vector $\mathbf{s}_n$, one obtains the model $\mathbf{s}_n = \mathbf{P}_{\text{tx}} \mathbf{F}_M^H \mathbf{x}_{\text{TF},n}$, where $\mathbf{P}_{\text{tx}} = \text{diag}\left\{ p_{\text{tx}}\left(\frac{pT}{M}\right) \right\}_{p=0}^{M-1}$ and $\mathbf{x}_{\text{TF},n}$ denotes the nth column of the matrix $\mathbf{X}_{\text{TF}}$. This can be interpreted as performing an M-point IFFT over each column of the TF-domain symbol matrix $\mathbf{X}_{\text{TF}}$, followed by transmit pulse-shaping. Thus, the time-domain symbol matrix $\mathbf{S} = [\mathbf{s}_0, \mathbf{s}_1, \cdots, \mathbf{s}_{N-1}] \in \mathbb{C}^{M \times N}$ can be determined as $\mathbf{S} = \mathbf{P}_{\text{tx}} \mathbf{F}_M^H \mathbf{X}_{\text{TF}} = \mathbf{P}_{\text{tx}} \mathbf{X}_{\text{DD}} \mathbf{F}_N^H$, since we have $\mathbf{X}_{\text{TF}} = \mathbf{F}_M \mathbf{X}_{\text{DD}} \mathbf{F}_N^H$. Thus, the MN-samples of the transmit signal $s(t)$, i.e., $\left\{ s(t)|_{t=\frac{pT}{M}} \right\}_{p=0}^{MN-1}$, can be obtained by vectorizing the matrix $\mathbf{S}$ as

$$\mathbf{s} = \text{vec}\left(\mathbf{S}\right) = \left(\mathbf{F}_N^H \otimes \mathbf{P}_{\text{tx}}\right) \mathbf{x}_{\text{DD}} \in \mathbb{C}^{MN \times 1}, \tag{7}$$

where $\mathbf{x}_{\text{DD}} = \text{vec}\left(\mathbf{X}_{\text{DD}}\right)$. Furthermore, similar to cyclic prefix (CP)-based block-transmission systems, the vector $\mathbf{s}$ can be extended by a CP of length L in order to eliminate the inter-block interference. The amalgamation of a mmWave hybrid beamforming architecture with OTFS-aided physical layer waveforms has an additional advantage in terms of lower PAPR in comparison to the traditional mmWave-HB MIMO OFDM systems, especially when $N << M$ [43].

After removing the CP at the output of the RF RC, the qth received sample $r(q)$ obtained as $r(q) = r(t)|_{t=\frac{qT}{M}}, 0 \leq q \leq MN-1$, is succinctly given by

$$r(q) = \sum_{p=1}^{L_p} \tilde{h}_p s\left([q - i_p]_{MN}\right) e^{j2\pi \frac{j_p (q-i_p)}{MN}} + w(q), \tag{8}$$

where $w(q) = w(t)|_{t=\frac{qT}{M}}$ and $[\cdot]_{MN}$ denotes the modulo-MN operation. Let $\mathbf{\Pi} \in \mathbb{C}^{MN \times MN}$ denote a permutation matrix and $\mathbf{\Delta}_p \in \mathbb{C}^{MN \times MN}$ represent a diagonal matrix defined as

$$\mathbf{\Delta}_p = \begin{cases} \text{diag}\left\{1, \omega_p, \cdots, \omega_p^{MN-i_p-1}, \omega_p^{-i_p}, \cdots, \omega_p^{-1}\right\}, \\ \qquad\qquad\qquad\qquad \text{if } i_p \neq 0, \\ \text{diag}\left\{1, \omega_p, \cdots, \omega_p^{MN-1}\right\}, \text{for } i_p = 0, \end{cases}$$

where $\omega_p = e^{j2\pi \frac{j_p}{MN}}$. 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 $r(q)$ and the noise process $w(q)$, given as $\mathbf{r} = [r(0), r(1), \cdots, r(MN-1)]^T$, $\mathbf{w} = [w(0), w(1), \cdots, w(MN-1)]^T$. Employing these notations, and using the relationship given in (8), the received signal vector $\mathbf{r}$ can be expressed as

$$\mathbf{r} = \mathbf{H}\mathbf{s} + \mathbf{w}, \tag{9}$$

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

$$\mathbf{H} = \sum_{p=1}^{L_p} \tilde{h}_p \left(\mathbf{\Pi}\right)^{i_p} \mathbf{\Delta}_p. \tag{10}$$

Let the vector $\mathbf{r}_n \in \mathbb{C}^{M \times 1}$ be comprised of the M-received samples corresponding to the nth symbol duration, i.e., $\mathbf{r}_n = [r_n(0), r_n(1), \cdots, r_n(M-1)]^T$, where the quantity $r_n(q), 0 \leq q \leq M-1$, represents the qth sample of the nth symbol duration. Now, considering the receive pulse-shaping filter $p_{\text{rx}}(t)$ also to be of duration T, the TF-demodulated symbol $\mathbf{Y}_{\text{TF}}(m,n)$ can be expressed in terms of the received samples $r_n(q)$ as $\mathbf{Y}_{\text{TF}}(m,n) = p_{\text{rx}}^*\left(\frac{mT}{M}\right) \sum_{q=0}^{M-1} r_n(q) e^{-j2\pi \frac{qm}{M}}$. Furthermore, let $\mathbf{y}_{\text{TF},n} \in \mathbb{C}^{M \times 1}$ denote the nth column of the matrix $\mathbf{Y}_{\text{TF}}$. This can be formulated as $\mathbf{y}_{\text{TF},n} = \mathbf{F}_M \mathbf{P}_{\text{rx}} \mathbf{r}_n$, where $\mathbf{P}_{\text{rx}} = \text{diag}\left\{ p_{\text{rx}}^*\left(\frac{qT}{M}\right) \right\}_{q=0}^{M-1}$. Upon concatenating

$\mathbf{y}_{\text{TF},n}, 0 \leq n \leq N-1$, the matrix $\mathbf{Y}_{\text{TF}}$ can be expressed as $\mathbf{Y}_{\text{TF}} = [\mathbf{y}_{\text{TF},0}, \mathbf{y}_{\text{TF},1}, \cdots, \mathbf{y}_{\text{TF},N-1}] = \mathbf{F}_M \mathbf{P}_{\text{rx}} \mathbf{R}$, where $\mathbf{R} \in \mathbb{C}^{M \times N}$ is comprised of the time-domain samples as $\mathbf{R} = [\mathbf{r}_0, \mathbf{r}_1, \cdots, \mathbf{r}_{N-1}]$. Finally, vectorizing $\mathbf{Y}_{\text{DD}}$ as $\mathbf{y}_{\text{DD}} = \text{vec}(\mathbf{Y}_{\text{DD}})$, and in turn, utilizing (9) and (7), we obtain an end-to-end relationship of the mmWave-AB MIMO OTFS system in the DD-domain as

$$\mathbf{y}_{\text{DD}} = \mathbf{H}_{\text{DD}} \mathbf{x}_{\text{DD}} + \mathbf{w}_{\text{DD}}, \tag{11}$$

where the matrix $\mathbf{H}_{\text{DD}} \in \mathbb{C}^{MN \times MN}$ is expressed as $\mathbf{H}_{\text{DD}} = (\mathbf{F}_N \otimes \mathbf{P}_{rx}) \mathbf{H} \left(\mathbf{F}_N^H \otimes \mathbf{P}_{\text{tx}}\right) \in \mathbb{C}^{MN \times MN}$. Furthermore, the DD-domain noise vector $\mathbf{w}_{\text{DD}} \in \mathbb{C}^{MN \times 1}$ of (11) is related to the time-domain noise $\mathbf{w}$ of (9) as $\mathbf{w}_{\text{DD}} = (\mathbf{F}_N \otimes \mathbf{P}_{\text{rx}}) \mathbf{w}$. Hence, its covariance matrix $\mathbf{R}_{w,\text{DD}} \in \mathbb{C}^{MN \times MN}$ is given by $\mathbf{R}_{w,\text{DD}} = \mathbb{E}\left[\mathbf{w}_{\text{DD}} \mathbf{w}_{\text{DD}}^H\right] = \sigma^2 \left[\mathbf{I}_N \otimes \left(\mathbf{P}_{\text{rx}} \mathbf{P}_{\text{rx}}^H\right)\right]$.

It is important to note that the subsequent symbol detection procedure using (11) relies on the availability of the CSI $\mathbf{H}_{\text{DD}}$, which from (10), in turn, depends on the equivalent path-gains $\tilde{h}_p$, and delay and Doppler indices i_p and j_p, respectively, for all the L_p multipath components. Furthermore, the equivalent path gains $\tilde{h}_p, 1 \leq p \leq L_p$, as described in (6), depend on the true path-gains h_p, and the choice of the RF TBF $\mathbf{f}_{\text{RF}}$ and RC $\mathbf{w}_{\text{RF}}$. Hence, it can be readily observed that the CSI acquisition, and the RF TBF and RC design in a mmWave-AB MIMO OTFS system are intricately intertwined. In fact, the joint CSI estimation and RF TBF/RC design problem is challenging. Therefore, we develop a 2-stage procedure for the same, which forms the focus of the next section.

III. RF Beamformer/ Combiner Design and Sparse Channel Estimation for mmWave-AB MIMO OTFS Systems

We now continue by conceiving a two-stage procedure for designing the RF TBF $\mathbf{f}_{\text{RF}}$ and RC $\mathbf{w}_{\text{RF}}$, while also carrying out sparse estimation of the delay τ_p, Doppler ν_p and the equivalent path-gain $\tilde{h}_p$ of the multipath components in an mmWave-AB MIMO OTFS system. The training frame structure of the proposed procedure is described in Fig. 1(b). To this end, consider a set of feasible angular grid-points $\Phi_{\text{A}} = \{\phi_a : \phi_a \in [0, \pi], 1 \leq a \leq G_{\text{A}}\}$ for Device-A and a similar set $\Phi_{\text{B}} = \{\theta_b : \theta_b \in [0, \pi], 1 \leq b \leq G_{\text{B}}\}$ at Device-B, where the grid-sizes G_{A} and G_{B} obey $(G_{\text{A}}, G_{\text{B}}) \geq \max\{N_t, N_r\}$. Typically, G_{A} and G_{B} are set to $2N_t$ and $2N_r$ [3], [14], [16], respectively, as also considered in our simulation results for achieving high angular resolution of the estimated AoD/ AoA. These sets of angular grid-points are chosen according to the following conditions [14], [15]

$$\begin{aligned} \cos(\phi_a) &= \frac{2}{G_{\text{A}}}(a-1) - 1, 1 \leq a \leq G_{\text{A}}, \\ \cos(\theta_b) &= \frac{2}{G_{\text{B}}}(b-1) - 1, 1 \leq b \leq G_{\text{B}}. \end{aligned} \tag{12}$$

The codebooks $\mathbf{C}_{\text{A}} \in \mathbb{C}^{N_t \times G_{\text{A}}}$ and $\mathbf{C}_{\text{B}} \in \mathbb{C}^{N_r \times G_{\text{B}}}$, also commonly referred to as dictionary matrices, comprise the array's steering vectors corresponding to the sets Φ_{A} and Φ_{B}, respectively, which are given as

$$\begin{aligned} \mathbf{C}_{\text{A}} &= [\mathbf{a}_t(\phi_1), \mathbf{a}_t(\phi_2) \ldots, \mathbf{a}_t(\phi_{G_{\text{A}}})], \\ \mathbf{C}_{\text{B}} &= [\mathbf{a}_r(\theta_1), \mathbf{a}_r(\theta_2) \ldots, \mathbf{a}_r(\theta_{G_{\text{B}}})]. \end{aligned} \tag{13}$$

A. Stage-1 ($\mathbf{w}_{RF}$ Design)

Let $s_{1,\mathcal{P},i}$ and $\mathbf{f}_{\text{RF},i} \in \mathbb{C}^{N_t \times 1}, 0 \leq i \leq N_{p,1} - 1$, denote the ith baseband pilot symbol and training beam, respectively, transmitted by Device-A, where $N_{p,1}$ represents the number of pilots in Stage-1. Each element of the TBF pilot beam $\mathbf{f}_{\text{RF},i}$ can be set as $\mathbf{C}_{\text{A}}(j,k), 1 \leq j \leq N_t, 1 \leq k \leq G_{\text{A}}$, where the indices j and k are randomly selected. Considering rectangular transmit and receive pulse shapes, the received pilot output $\mathbf{r}_1(q) \in \mathbb{C}^{N_r \times 1}$ after addition/removal of the CP can be expressed as

$$\mathbf{r}_1(q) = \sqrt{\frac{N_t N_r}{L_p}} \sum_{p=1}^{L_p} \Bigg[h_p \mathbf{a}_r(\theta_p) \mathbf{a}_t^H(\phi_p) \bar{\mathbf{s}}_{1,\mathcal{P},[q-i_p]_{N_{p,1}}} e^{j2\pi \frac{j_p(q-i_p)}{MN}} \Bigg] + \mathbf{w}_1(q), \tag{14}$$

where the quantity $\bar{\mathbf{s}}_{1,\mathcal{P},i} \in \mathbb{C}^{N_t \times 1}$ signifies the pilot vector transmitted by Device-A at time-instant i, which is defined as $\bar{\mathbf{s}}_{1,\mathcal{P},i} = \mathbf{f}_{\text{RF},i} s_{1,\mathcal{P},i}$, whereas $\mathbf{w}_1(q) \in \mathbb{C}^{N_r \times 1}$ represents the noise. Device-B combines the received pilot vectors $\mathbf{r}_1(q), 0 \leq q \leq N_{p,1} - 1$, using the RC beams of the codebook $\mathbf{C}_{\text{B}}$. This operation is described as follows. Concatenating $\mathbf{r}_1(q), 0 \leq q \leq N_{p,1} - 1$, one obtains the received pilot matrix $\mathbf{R}_1 \in \mathbb{C}^{N_r \times N_{p,1}}$ as $\mathbf{R}_1 = [\mathbf{r}_1(0), \mathbf{r}_1(1), \cdots, \mathbf{r}_1(N_{p,1}-1)]$. Device-B processes the received pilot matrix $\mathbf{R}_1$ using the codebook matrix $\mathbf{C}_{\text{B}}$ to obtain the correlation matrix $\mathbf{\Psi}_1 \in \mathbb{C}^{G_{\text{B}} \times N_{p,1}}$ as $\mathbf{\Psi}_1 = \mathbf{C}_{\text{B}}^H \mathbf{R}_1$. Note that the matrix $\mathbf{\Psi}_1$ comprises the correlations of the G_{B} RC beams with the $N_{p,1}$ received pilots. Thus, the index of the RC beam, which is maximally correlated with the received signal, can be readily obtained as $b_{\text{opt}} = \underset{b=1,\cdots,G_{\text{B}}}{\arg\max} \left\|\mathbf{\Psi}_1(b,:)\right\|^2$, and the optimal RC $\mathbf{w}_{\text{RF}}^{\text{opt}}$ is given by $\mathbf{w}_{\text{RF}}^{\text{opt}} = \mathbf{C}_{\text{B}}(:, b_{\text{opt}})$. When the grid-sizes G_{A} and G_{B} are very large, one can employ a multi-resolutional hierarchical codebook, as considered in [3], [12], which performs coarse and fine beam adaptation in several stages. Here, an adaptive search is performed over the AoA/ AoD starting with wide beams in the early stages and narrowing down the search based on the outputs in the later stages so as to focus only in the most promising directions. Another interesting future goal can be to jointly estimate the dominant AoA and AoD, similar to [44] proposed for OFDM.

B. Stage-2 ($\mathbf{f}_{RF}$ Design and Equivalent Sparse DD-domain CSI Estimation)

In Stage-2, Device-B employs the RF TBF $\left(\mathbf{w}_{\text{RF}}^{\text{opt}}\right)^*$ obtained from Stage-1 for transmission of the baseband pilot signal $s_{2,\mathcal{P},i}, 0 \leq i \leq N_{p,2} - 1$. Thus, the transmit pilot vector $\bar{\mathbf{s}}_{2,\mathcal{P},i} \in \mathbb{C}^{N_r \times 1}$ is given by $\bar{\mathbf{s}}_{2,\mathcal{P},i} = \left(\mathbf{w}_{\text{RF}}^{\text{opt}}\right)^* s_{2,\mathcal{P},i}$. Once again, considering rectangular transmit and receive pulse shapes, and after addition and removal of the CP, the pilot output $\mathbf{r}_2(q) \in \mathbb{C}^{N_t \times 1}, 0 \leq q \leq N_{p,2} - 1$, at Device-A is

given by

$$\mathbf{r}_2(q) = \sqrt{\frac{N_t N_r}{L_p}} \sum_{p=1}^{L_p} \Bigg[h_p \mathbf{a}_t^*(\phi_p) \mathbf{a}_r^T(\theta_p) \bar{\mathbf{s}}_{2,\mathcal{P},[q-i_p]_{N_{p,2}}} e^{j2\pi \frac{j_p(q-i_p)}{MN}} \Bigg] + \mathbf{w}_2(q), \quad (15)$$

where the quantity $\mathbf{w}_2(q) \in \mathbb{C}^{N_t \times 1}$ denotes the noise. Let the concatenated received pilot matrix $\mathbf{R}_2 \in \mathbb{C}^{N_t \times N_{p,2}}$ be defined as $\mathbf{R}_2 = [\mathbf{r}_2(0), \mathbf{r}_2(1), \cdots, \mathbf{r}_2(N_{p,2}-1)]$. Device-A now computes the correlation matrix $\mathbf{\Psi}_2 \in \mathbb{C}^{G_\text{A} \times N_{p,2}}$ using the codebook matrix $\mathbf{C}_\text{A}$ as $\mathbf{\Psi}_2 = \mathbf{C}_\text{A}^T \mathbf{R}_2$. The optimal TBF $\mathbf{f}_\text{RF}^\text{opt}$ can be obtained as $\mathbf{f}_\text{RF}^\text{opt} = \mathbf{C}_\text{A}(:, a_\text{opt})$, where the index a_opt is given by $a_\text{opt} = \underset{a=1,\cdots,G_\text{A}}{\arg\max} \left\| \mathbf{\Psi}_2(a,:) \right\|^2$. Hence, the key novelty of the proposed beam alignment protocol is as follows. Note that the $N_{p,1}$ Stage-1 pilot beam patterns at Device-A employ random phases at the phase-shifters, which help exciting various angular modes of the mmWave MIMO channel with a much lower number of pilot beams in comparison to those obtained by the conventional DFT-codebook. Furthermore, it is important to note that the outputs of all the $N_{p,1}$ pilot beams are jointly processed for TBF/ RC beam-selection, i.e. while $\mathbf{w}_\text{RF}$ Design at Device-B. At Stage-2, Device-B employs this selected beam $\mathbf{w}_\text{RF}$ for sending $N_{p,2}$ pilot symbols, and all the $N_{p,2}$ pilot outputs are jointly processed for $\mathbf{f}_\text{RF}$ Design at Device-A. Hence, joint processing of pilot outputs in their respective stages make our beam-alignment procedure more reliable. Also note that the $N_{p,2}$ Stage-2 pilot outputs have been directly employed for the DD-domain equivalent CSI estimation, which is of significant importance, since one does not need any additional stage for DD-domain CSI estimation. Due to AoA/ AoD reciprocity in both TDD and FDD modes [45], [46], the beam pattern selected for RC at any device can also be used as TBF. Hence no feedback is required for these beam patterns.

Subsequently, the pilot output combined using $\mathbf{f}_\text{RF}^\text{opt}$ can be employed for estimating the equivalent channel of the DD-domain. Note that the desired combined pilot output $\mathbf{r}_\mathcal{P} \in \mathbb{C}^{N_{p,2} \times 1}$ is obtained as $\mathbf{r}_\mathcal{P} = [\mathbf{\Psi}_2(a_\text{opt}, :)]^T$, and its qth element, denoted by $\mathbf{r}_\mathcal{P}(q)$, can be expressed as

$$\mathbf{r}_\mathcal{P}(q) = \sum_{p=1}^{L_p} \bar{h}_p s_{2,\mathcal{P},[q-i_p]_{N_{p,2}}} e^{j2\pi \frac{j_p(q-i_p)}{MN}} + w_2(q), \quad (16)$$

where $\bar{h}_p$ represents the equivalent channel for the pth multi-path component, defined as

$$\bar{h}_p = \sqrt{\frac{N_t N_r}{L_p}} h_p \left(\mathbf{f}_\text{RF}^\text{opt}\right)^T \mathbf{a}_t^*(\phi_p) \mathbf{a}_r^T(\theta_p) \left(\mathbf{w}_\text{RF}^\text{opt}\right)^*, \quad (17)$$

and the combined noise sample $w_2(q)$ is obtained as $w_2(q) = \left(\mathbf{f}_\text{RF}^\text{opt}\right)^T \mathbf{w}_2(q)$. Furthermore, stacking the transmit baseband pilots $s_{2,\mathcal{P},i}, 0 \le i \le N_{p,2}-1$, as $\mathbf{s}_\mathcal{P} = \left[s_{2,\mathcal{P},0},\ s_{2,\mathcal{P},1}, \cdots, s_{2,\mathcal{P},N_{p,2}-1}\right]^T \in \mathbb{C}^{N_{p,2} \times 1}$, the pilot output $\mathbf{r}_\mathcal{P}$ can be formulated as

$$\mathbf{r}_\mathcal{P} = \left[\sum_{p=1}^{L_p} \bar{h}_p \left(\bar{\mathbf{\Pi}}\right)^{i_p} \bar{\mathbf{\Delta}}_{i_p,j_p} \right] \mathbf{s}_\mathcal{P} + \mathbf{w}_\mathcal{P}, \quad (18)$$

Algorithm 1 BL-based CSI estimation in AB-based mmWave MIMO OTFS systems

Input: Pilot output $\mathbf{r}_\mathcal{P}$, dictionary matrix $\mathbf{\Psi}$, stopping parameters η and S_max

Initialization: $\widehat{\gamma}_i^{(0)} = 1, \forall\ i \Rightarrow \widehat{\mathbf{\Gamma}}^{(0)} = \mathbf{I}_{M_s G_s}$, counter $j = 0$ and $\widehat{\mathbf{\Gamma}}^{(-1)} = \mathbf{0}_{M_s G_s \times M_s G_s}$

while $\left(\left\| \widehat{\mathbf{\Gamma}}^{(j)} - \widehat{\mathbf{\Gamma}}^{(j-1)} \right\|_F^2 \ge \eta \ \text{ and } j < S_\text{max} \right)$ **do**

1) $j = j + 1$
2) **E-step:** Compute the *a posteriori* covariance and mean

$$\mathbf{\Sigma}^{(j)} = \left[\sigma^{-2} \mathbf{\Psi}^H \mathbf{\Psi} + \left(\widehat{\mathbf{\Gamma}}^{(j-1)} \right)^{-1} \right]^{-1},$$
$$\boldsymbol{\mu}^{(j)} = \sigma^{-2} \mathbf{\Sigma}^{(j)} \mathbf{\Psi}^H \mathbf{r}_\mathcal{P}$$

3) **M-step:** Update the hyperparameters

for $i = 1, 2, \cdots, M_s G_s$ **do**

$$\widehat{\gamma}_i^{(j)} = \mathbf{\Sigma}^{(j)}(i,i) + \left| \boldsymbol{\mu}^{(j)}(i) \right|^2$$

end for

$\widehat{\mathbf{\Gamma}}^{(j)} = \text{diag}\left(\widehat{\gamma}_1^{(j)}, \widehat{\gamma}_2^{(j)}, \cdots, \widehat{\gamma}_{M_s G_s}^{(j)} \right)$

end while

Output: $\widehat{\mathbf{h}}_\text{BL} = \left[\mathbf{\Psi}^H \mathbf{\Psi} + \sigma^2 \left(\widehat{\mathbf{\Gamma}}^{(j)} \right)^{-1} \right]^{-1} \mathbf{\Psi}^H \mathbf{r}_\mathcal{P}$

where $\mathbf{w}_\mathcal{P} \in \mathbb{C}^{N_{p,2} \times 1}$ represents the noise vector, which is obtained by stacking the noise samples $w_2(q)$ as $\mathbf{w}_\mathcal{P} = [w_2(0), w_2(1), \cdots, w_2(N_{p,2}-1)]^T$, $\bar{\mathbf{\Pi}}$ denotes a permutation matrix of size $N_{p,2} \times N_{p,2}$ and $\bar{\mathbf{\Delta}}_{i_p,j_p} \in \mathbb{C}^{N_{p,2} \times N_{p,2}}$ is a diagonal matrix, defined as

$$\bar{\mathbf{\Delta}}_{i_p,j_p} = \begin{cases} \text{diag}\left\{1, \omega_p, \cdots, \omega_p^{N_{p,2}-i_p-1}, \omega_p^{-i}, \cdots, \omega_p^{-1}\right\}, \\ \qquad\qquad\qquad\qquad \text{if } i_p \neq 0, \\ \text{diag}\left\{1, \omega_p, \cdots, \omega_p^{N_{p,2}-1}\right\}, \text{for } i_p = 0. \end{cases}$$

Using (18), one can now formulate the DD-domain sparse CSI estimation model as follows.

Let the integer taps M_s and N_s denote the maximum delay- and Doppler-spread of the channel, which obey $M_s << M$ and $N_s << N$, for a typical under-spread channel. For sparse representation, one can now consider a virtual Doppler-grid of size $G_s (>> N_s)$ on the Doppler axis, where the jth Doppler-grid point, $0 \le j \le G_s - 1$, corresponds to a Doppler-shift of $\nu_j = \frac{jN_s}{G_s NT}$ Hz. Note that the integer-Doppler tap corresponding to ν_j is $\left[\text{round}\left(\frac{jN_s}{G_s}\right)\right]$, whereas the fractional-Doppler is given by $\left[\frac{jN_s}{G_s} - \text{round}\left(\frac{jN_s}{G_s}\right)\right]$. Let $\mathbf{H}_{i,j}$ denote the path-gain matrix associated with the ith delay-tap and jth virtual Doppler-tap, the DD-domain sparse representation of the mmWave MIMO wireless channel is given by

$$\mathbf{H}(\tau, \nu) = \sum_{i=0}^{M_s-1} \sum_{j=0}^{G_s-1} \mathbf{H}_{i,j} \delta(\tau - \tau_i) \delta(\nu - \nu_j), \quad (19)$$

since only L_p matrices $\mathbf{H}_{i,j}$ out of $M_s G_s$ are non-zero, which correspond to the active delay-Doppler indices. Exploiting

this property, the channel estimation model of (18) can be reformulated as

$$\mathbf{r}_{\mathcal{P}} = \sum_{i=0}^{M_s-1}\sum_{j=0}^{G_s-1}\left[\left(\bar{\mathbf{\Pi}}\right)^{i}\bar{\mathbf{\Delta}}_{i,j}\mathbf{s}_{\mathcal{P}}\right]h_{i,j} + \mathbf{w}_{\mathcal{P}}, \tag{20}$$

where the quantity $h_{i,j}$ signifies the equivalent beamformed channel similar to (17). Substituting the quantity $\boldsymbol{\psi}_{i,j} = \left(\bar{\mathbf{\Pi}}\right)^{i}\bar{\mathbf{\Delta}}_{i,j}\mathbf{s}_{\mathcal{P}}$ into (20), the sparse DD-domain CSI estimation model for the mmWave-AB MIMO OTFS system is expressed as

$$\mathbf{r}_{\mathcal{P}} = \mathbf{\Psi}\mathbf{h} + \mathbf{w}_{\mathcal{P}}, \tag{21}$$

where the columns of the dictionary matrix $\mathbf{\Psi} \in \mathbb{C}^{N_{p,2}\times M_sG_s}$ are denoted by $\left\{\boldsymbol{\psi}_{i,j}\right\}_{i,j=0}^{M_s-1,G_s-1}$. Note that the CSI vector $\mathbf{h} \in \mathbb{C}^{M_sG_s\times 1}$, which contains the coefficients $\{h_{i,j}\}_{i,j=0}^{M_s-1,G_s-1}$ with a similar indexing order, is sparse in nature, since only a few coefficients L_p out of M_sG_s are non-zero. The MMSE channel estimate for the linear model of (21), is given by

$$\widehat{\mathbf{h}}_{\text{MMSE}} = \left(\mathbf{\Psi}^H\mathbf{R}_w^{-1}\mathbf{\Psi} + \mathbf{R}_h^{-1}\right)^{-1}\mathbf{\Psi}^H\mathbf{R}_w^{-1}\mathbf{r}_{\mathcal{P}}, \tag{22}$$

where $\mathbf{R}_w = \mathbb{E}\left[\mathbf{w}_{\mathcal{P}}\mathbf{w}_{\mathcal{P}}^H\right] \in \mathbb{C}^{N_{p,2}\times N_{p,2}}$ denotes the covariance matrix of the noise vector $\mathbf{w}_{\mathcal{P}}$ and $\mathbf{R}_h = \mathbb{E}\left[\mathbf{h}\mathbf{h}^H\right] \in \mathbb{C}^{M_sG_s\times M_sG_s}$ represents the covariance matrix of the DD-domain CSI vector $\mathbf{h}$. Here, the noise covariance $\mathbf{R}_w$ obeys $\mathbf{R}_w = \sigma^2\mathbf{I}_{N_{p,2}}$, whereas the channel covariance $\mathbf{R}_h$ is typically unknown.

To this end, we develop a BL-based framework, which employs a parameterized Gaussian prior as $p(\mathbf{h};\mathbf{\Gamma}) = \mathcal{CN}(\mathbf{0};\mathbf{\Gamma})$ [47], i.e., the unknown channel covariance matrix $\mathbf{R}_h$ has been modeled as $\mathbf{R}_h = \mathbf{\Gamma} = \text{diag}\left(\gamma_1,\gamma_2,\cdots,\gamma_{M_sG_s}\right)$, where $\gamma_i \in \mathbb{R}_+$ represents the hyperparameter signifying the prior variance of the ith element. Subsequently, these hyperparameters are estimated iteratively using the well-established expectation-maximization (EM) procedure. Upon convergence of the EM iterations, the BL-based estimate is obtained by substituting the estimate $\widehat{\mathbf{\Gamma}}$ of the hyperparameter matrix into the MMSE estimate of (22). The various steps of the proposed BL-based framework for the sparse DD-domain CSI estimation of our mmWave-AB MIMO OTFS system have been described in Algorithm-1, which is similar to the one derived in our preprint [48] for a Tera-Hertz hybrid MIMO system for exploiting the angular-sparsity.

In summary, Stage-1 designs the TBF/ RC $\mathbf{w}_{\text{RF}}$ for Device-B, followed by TBF/ RC $\mathbf{f}_{\text{RF}}$ design for Device-A in Stage-2. Furthermore, the Stage-2 pilot output $\mathbf{r}_{\mathcal{P}}$ described by the channel estimation model of (21) is also utilized for estimating the delay, Doppler and equivalent complex-valued path gains of all the multipath components. Subsequently, employing the $\mathbf{f}_{\text{RF}}$ and $\mathbf{w}_{\text{RF}}$ designed, several OTFS data-frames are transmitted, as shown in Fig. 1(b). The symbol detection for these data-frames is performed using the DD-domain input-output relationship derived in (11), which relies on the estimated CSI obtained from Stage-2.

Remark-1: Note that the pilot beams generated by setting random phases at phase-shifters are employed only at Device-A exclusively for the Stage-1 pilot beams. By contrast, the TBF/ RC beams at both Devices and in both Stages are selected from the codebooks $\mathbf{C}_{\text{A}}$ and $\mathbf{C}_{\text{B}}$, which align the receiver beam directions along the dominant AoA/ AoD. This is because the pilot beam patterns employed at Device-A in Stage-1 play a role similar to probing beams, since the actual AoAs/ AoDs are not known initially. Therefore, during Stage-1 probing, a broader beam direction generated by setting random phases at the phase-shifters has been preferred. Employing narrow pilot beams, for example those obtained from a DFT-codebook, may lead to very low SNR if the true AoD of the signal does not match with the beamformerâĂŹs main beam.

IV. OTFS-BASED MMWAVE HYBRID BEAMFORMING

Consider an OTFS-based mmWave-HB MIMO system having N_t TAs, N_r RAs and $N_{\text{RF}} << \min(N_t,N_r)$ RF chains, as shown in Fig. 2(a). The mmWave-HB MIMO system is comprised of an analog RF TPC[1] $\mathbf{F}_{\text{RF}} \in \mathbb{C}^{N_t\times N_{\text{RF}}}$, which maps the symbols of the N_{RF} RF chains to the N_t TAs. Similarly, at the receiver, we have an RF RC $\mathbf{W}_{\text{RF}} \in \mathbb{C}^{N_r\times N_{\text{RF}}}$, which maps the signal gleaned from the N_r RAs back to the N_{RF} RF chains. Thus, the signal $\mathbf{r}(t) \in \mathbb{C}^{N_{\text{RF}}\times 1}$ received at the output of the RF RC, in response to the signal $\mathbf{s}(t) \in \mathbb{C}^{N_{\text{RF}}\times 1}$ transmitted from the transmit RF chains, is expressed as

$$\mathbf{r}(t) = \sqrt{\frac{N_tN_r}{L_p}}\sum_{p=1}^{L_p}\Big[h_p\mathbf{W}_{\text{RF}}^H\mathbf{a}_r\left(\theta_p\right)\mathbf{a}_t^H\left(\phi_p\right)\mathbf{F}_{\text{RF}}\mathbf{s}(t-\tau_p) e^{j2\pi\nu_p(t-\tau_p)}\Big] + \mathbf{W}_{\text{RF}}^H\tilde{\mathbf{w}}(t),$$

where $\tilde{\mathbf{w}}(t)$ signifies the receiver noise, similar to (1). Furthermore, considering a block of MN samples, denoted by $\mathbf{s}(q) = \mathbf{s}(t)|_{t=\frac{qT}{M}}, 0 \le q \le MN-1$, followed by the CP attachment and removal process, the received samples $\mathbf{r}(q), 0 \le q \le MN-1$, at the output of the RF RC, are given by

$$\mathbf{r}(q) = \sum_{p=1}^{L_p}\mathbf{W}_{\text{RF}}^H\mathbf{H}_p\mathbf{F}_{\text{RF}}\mathbf{s}\big([q-i_p]_{MN}\big)e^{j2\pi\frac{j_p(q-i_p)}{MN}} + \mathbf{w}(q), \tag{23}$$

where $\mathbf{H}_p$ represents the pth tap of the mmWave MIMO channel, as described by Eq. (3), and $\mathbf{w}(q) = \mathbf{W}_{\text{RF}}^H\tilde{\mathbf{w}}(t)|_{t=\frac{qT}{M}}$ denotes the sampled noise vector. The DD-domain end-to-end system model of the mmWave-HB MIMO OTFS system is derived next.

As shown in Fig. 2(a), the mmWave-HB MIMO OTFS system performs OTFS modulation and demodulation at each transmit and receive RF chain, respectively. Let $\mathbf{X}_{\text{DD},u} \in \mathbb{C}^{M\times N}, 1 \le u \le N_{\text{RF}}$, denote the DD-domain symbol matrix corresponding to the uth transmit RF chain and let $\mathbf{Y}_{\text{DD},v} \in \mathbb{C}^{M\times N}, 1 \le v \le N_{\text{RF}}$, represent the DD-domain symbol matrix corresponding to the vth receive RF chain. Let $\widetilde{\mathbf{H}}_p = \mathbf{W}_{\text{RF}}^H\mathbf{H}_p\mathbf{F}_{\text{RF}} \in \mathbb{C}^{N_{\text{RF}}\times N_{\text{RF}}}$ represent the equivalent MIMO channel tap. Following the procedures described in Section-II-D, the vectorized output $\mathbf{y}_{\text{DD},v} = \text{vec}(\mathbf{Y}_{\text{DD},v}) \in$

[1]For HB MIMO systems, we have used TPC, which stands for transmit precoder, instead of TBF, i.e. transmit beamformer.

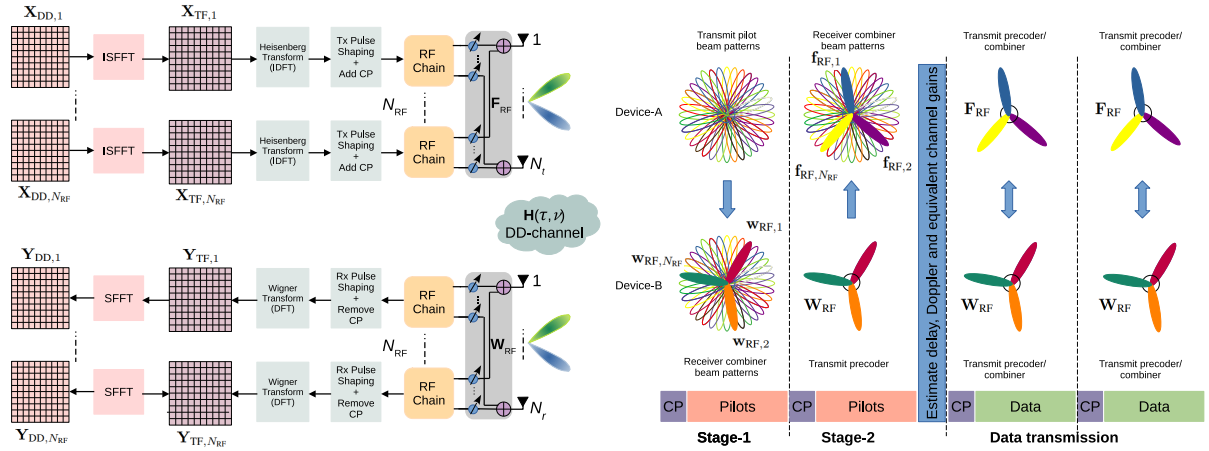

Fig. 2. (a) Architecture of the OTFS-based mmWave-HB MIMO transceiver; (b) Training frame structure of the proposed CSI estimation procedure for mmWave-HB MIMO systems.

$\mathbb{C}^{MN\times 1}$ and input $\mathbf{x}_{\text{DD},u} = \text{vec}(\mathbf{X}_{\text{DD},u}) \in \mathbb{C}^{MN\times 1}$ are related as

$$\mathbf{y}_{\text{DD},v} = \sum_{u=1}^{N_{\text{RF}}} \mathbf{H}_{\text{DD},v,u}\mathbf{x}_{\text{DD},u} + \mathbf{w}_{\text{DD},v}. \tag{24}$$

In the above, $\mathbf{w}_{\text{DD},v} \in \mathbb{C}^{MN\times 1}$ is given by $\mathbf{w}_{\text{DD},v} = (\mathbf{F}_N \otimes \mathbf{P}_{\text{rx}})\,\mathbf{w}_v$, where $\mathbf{w}_v \in \mathbb{C}^{MN\times 1}$ comprises the noise samples corresponding to the vth receive RF chain after RF combining, and $\mathbf{H}_{\text{DD},v,u} \in \mathbb{C}^{MN\times MN}$, similar to (11), follows

$$\mathbf{H}_{\text{DD},v,u} = \sum_{p=1}^{L_p} (\mathbf{F}_N \otimes \mathbf{P}_{\text{r}x}) \left[\tilde{h}_{p,v,u}\,(\mathbf{\Pi})^{i_p}\,\mathbf{\Delta}_p\right] \left(\mathbf{F}_N^H \otimes \mathbf{P}_{\text{tx}}\right), \tag{25}$$

with $\tilde{h}_{p,v,u}$ denoting the (v,u)th element of the matrix $\widetilde{\mathbf{H}}_p$, i.e., $\tilde{h}_{p,v,u} = \widetilde{\mathbf{H}}_p(v,u)$. Subsequently, upon stacking the vectors $\mathbf{y}_{\text{DD},v}, 1 \le v \le N_{\text{RF}}$, as $\bar{\mathbf{y}}_{\text{DD}} = \left[\mathbf{y}_{\text{DD},1}^T, \mathbf{y}_{\text{DD},2}^T, \cdots, \mathbf{y}_{\text{DD},N_{\text{RF}}}^T\right]^T \in \mathbb{C}^{MNN_{\text{RF}}\times 1}$, the DD-domain input-output model for the mmWave-HB MIMO OTFS system can be formulated as

$$\bar{\mathbf{y}}_{\text{DD}} = \bar{\mathbf{H}}_{\text{DD}}\bar{\mathbf{x}}_{\text{DD}} + \bar{\mathbf{w}}_{\text{DD}}, \tag{26}$$

where $\bar{\mathbf{x}}_{\text{DD}} = \left[\mathbf{x}_{\text{DD},1}^T, \cdots, \mathbf{x}_{\text{DD},N_{\text{RF}}}^T\right]^T \in \mathbb{C}^{MNN_{\text{RF}}\times 1}$, $\bar{\mathbf{w}}_{\text{DD}} = \left[\mathbf{w}_{\text{DD},1}^T, \cdots, \mathbf{w}_{\text{DD},N_{\text{RF}}}^T\right]^T \in \mathbb{C}^{MNN_{\text{RF}}\times 1}$ represent the stacked vectors of the transmit DD-domain symbols of all the transmit RF chains and stacked noise vectors of all the receive RF chains, respectively. The end-to-end channel matrix $\bar{\mathbf{H}}_{\text{DD}} \in \mathbb{C}^{MNN_{\text{RF}}\times MNN_{\text{RF}}}$ of the mmWave-HB MIMO OTFS system is given by

$$\begin{aligned}\bar{\mathbf{H}}_{\text{DD}} &= \begin{bmatrix} \mathbf{H}_{\text{DD},1,1} & \mathbf{H}_{\text{DD},1,2} & \cdots & \mathbf{H}_{\text{DD},1,N_{\text{RF}}} \\ \mathbf{H}_{\text{DD},2,1} & \mathbf{H}_{\text{DD},2,2} & \cdots & \mathbf{H}_{\text{DD},2,N_{\text{RF}}} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{H}_{\text{DD},N_{\text{RF}},1} & \mathbf{H}_{\text{DD},N_{\text{RF}},2} & \cdots & \mathbf{H}_{\text{DD},N_{\text{RF}},N_{\text{RF}}} \end{bmatrix} \\ &= (\mathbf{I}_{N_{\text{RF}}} \otimes \mathbf{F}_N \otimes \mathbf{P}_{\text{r}x}) \left[\sum_{p=1}^{L_p} \widetilde{\mathbf{H}}_p \otimes \left((\mathbf{\Pi})^{i_p}\,\mathbf{\Delta}_p\right)\right] \\ &\quad \left(\mathbf{I}_{N_{\text{RF}}} \otimes \mathbf{F}_N^H \otimes \mathbf{P}_{\text{tx}}\right). \end{aligned} \tag{27}$$

The next section develops a procedure for designing the RF TPC $\mathbf{F}_{\text{RF}}$, RC $\mathbf{W}_{\text{RF}}$, and for estimating the MIMO channel tap $\widetilde{\mathbf{H}}_p$ along with its delay index i_p and Doppler index j_p.

V. RF Precoder/ Combiner Design and Sparse Channel Estimation for mmWave-HB MIMO OTFS Systems

This section designs the RF TPC/ RC followed by estimation of the sparse DD-domain equivalent channel. As seen in Fig. 2(b), in Stage-1, the pilot beams are transmitted by Device-A and the RF RC $\mathbf{W}_{\text{RF}}$ is designed for Device-B by selecting the N_{RF} dominant beams from its codebook $\mathbf{C}_{\text{B}}$. Subsequently, in Stage-2, Device-B transmits the pilot symbols employing the RF TPC $\mathbf{W}_{\text{RF}}$ obtained from the Stage-1. The RF RC $\mathbf{F}_{\text{RF}}$ for Device-A is computed by selecting the N_{RF} dominant beams from its codebook $\mathbf{C}_{\text{A}}$, followed by computing the equivalent MIMO channel tap $\widetilde{\mathbf{H}}_p$, its delay index i_p and Doppler index j_p exploiting a block-sparse signal recovery technique. One can clearly identify here that the key difference with respect to the mmWave-AB MIMO OTFS system is in the selection of N_{RF} dominant beams followed by exploiting the special block-sparse structure, as explained later in this section.

A. Stage-1 ($\mathbf{W}_{RF}$ Design)

Let the ith baseband pilot vector and training RF TPC employed by Device-A be denoted by $\mathbf{s}_{1,\mathcal{P},i} \in \mathbb{C}^{N_{\text{RF}}\times 1}$ and $\mathbf{F}_{\text{RF},i} \in \mathbb{C}^{N_t\times N_s}, 0 \le i \le N_{p,1}-1$, respectively. The elements of the training RF TPC $\mathbf{F}_{\text{RF},i}$ are set as $\mathbf{C}_{\text{A}}(j,k), 1 \le j \le N_t, 1 \le k \le G_{\text{A}}$, where the indices j and k are randomly selected for each i. Thus, the pilot vector $\bar{\mathbf{s}}_{1,\mathcal{P},i} \in \mathbb{C}^{N_t\times 1}$ transmitted by Device-A at time-instant i is given as $\bar{\mathbf{s}}_{1,\mathcal{P},i} = \mathbf{F}_{\text{RF},i}\mathbf{s}_{1,\mathcal{P},i}$. The pilot output $\mathbf{r}_1(q) \in \mathbb{C}^{N_r\times 1}$ after the CP removal, i.e., for $0 \le q \le N_{p,1}-1$, can be

expressed as

$$\mathbf{r}_1(q) = \sqrt{\frac{N_t N_r}{L_p}} \sum_{p=1}^{L_p} \Bigg[h_p \mathbf{a}_r(\theta_p) \mathbf{a}_t^H(\phi_p) \bar{\mathbf{s}}_{1,\mathcal{P},[q-i_p]_{N_{p,1}}} e^{j2\pi \frac{j_p(q-i_p)}{MN}} \Bigg] + \mathbf{w}_1(q), \quad (28)$$

where $\mathbf{w}_1(q)$ represents the noise. Device-B now computes the correlation matrix $\mathbf{\Psi}_1$ as $\mathbf{\Psi}_1 = \mathbf{C}_{\text{B}}^H \mathbf{R}_1$, where the output pilot matrix obeys $\mathbf{R}_1 = [\mathbf{r}_1(0), \mathbf{r}_1(1), \cdots, \mathbf{r}_1(N_{p,1}-1)]$. The N_{RF} dominant beams used for designing the RF RC $\mathbf{W}_{\text{RF}}$ are selected as follows. Let the quantities ψ_b be defined as $\psi_b = \|\mathbf{\Psi}_1(b,:)\|^2$, which are arranged in a decreasing order as $\psi_{b_1} \geq \psi_{b_2} \geq \cdots \geq \psi_{b_{G_{\text{B}}}}$. Furthermore, let the set $\mathcal{B}$ be constructed as $\mathcal{B} = \{b_1, b_2, \cdots, b_{N_{\text{RF}}}\}$, which corresponds to the N_{RF} indices of the combining beams $\mathbf{a}_r(\theta_b)$ in the codebook $\mathbf{C}_{\text{B}}$ that are highly correlated with the received pilots $\mathbf{r}_1(q)$. The optimal RC $\mathbf{W}_{\text{RF}}^{\text{opt}}$ is then given by $\mathbf{W}_{\text{RF}}^{\text{opt}} = \mathbf{C}_{\text{B}}(:, \mathcal{B})$.

B. Stage-2 ($\mathbf{F}_{RF}$ Design and Equivalent Sparse DD-domain CSI Estimation)

Device-B now configures the RF TPC as $\left(\mathbf{W}_{\text{RF}}^{\text{opt}}\right)^*$ for transmitting the baseband pilot vectors $\mathbf{s}_{2,\mathcal{P},i} \in \mathbb{C}^{N_{\text{RF}} \times 1}, 0 \leq i \leq N_{p,2}-1$. After CP removal, the pilot output $\mathbf{r}_2(q) \in \mathbb{C}^{N_t \times 1}$ at Device-A is given by

$$\mathbf{r}_2(q) = \sqrt{\frac{N_t N_r}{L_p}} \sum_{p=1}^{L_p} \Bigg[h_p \mathbf{a}_t^*(\phi_p) \mathbf{a}_r^T(\theta_p) \bar{\mathbf{s}}_{2,\mathcal{P},[q-i_p]_{N_{p,2}}} e^{j2\pi \frac{j_p(q-i_p)}{MN}} \Bigg] + \mathbf{w}_2(q), \quad (29)$$

where the quantity $\bar{\mathbf{s}}_{2,\mathcal{P},i} \in \mathbb{C}^{N_r \times 1}$ is formulated as $\bar{\mathbf{s}}_{2,\mathcal{P},i} = \left(\mathbf{W}_{\text{RF}}^{\text{opt}}\right)^* \mathbf{s}_{2,\mathcal{P},i}$ and $\mathbf{w}_2(q) \in \mathbb{C}^{N_t \times 1}$ denotes the noise. Device-A then computes the correlation matrix $\mathbf{\Psi}_2 \in \mathbb{C}^{G_{\text{A}} \times N_{p,2}}$ as $\mathbf{\Psi}_2 = \mathbf{C}_{\text{A}}^T \mathbf{R}_2$, where $\mathbf{R}_2 \in \mathbb{C}^{N_t \times N_{p,2}}$ denotes the concatenated received pilot matrix defined as $\mathbf{R}_2 = [\mathbf{r}_2(0), \mathbf{r}_2(1), \cdots, \mathbf{r}_2(N_{p,2}-1)]$. Let the quantities ψ_a be arranged as $\psi_{a_1} \geq \psi_{a_2} \geq \cdots \geq \psi_{a_{G_{\text{A}}}}$, where $\psi_a = \|\mathbf{\Psi}_2(a,:)\|^2$. The optimal RC $\mathbf{F}_{\text{RF}}^{\text{opt}}$ for Device-A is given by $\mathbf{F}_{\text{RF}}^{\text{opt}} = \mathbf{C}_{\text{A}}(:, \mathcal{A})$, where the set $\mathcal{A}$ is formulated as $\mathcal{A} = \{a_1, a_2, \cdots, a_{N_{\text{RF}}}\}$.

The pilot output $\mathbf{R}_{\mathcal{P}} \in \mathbb{C}^{N_{p,2} \times N_{\text{RF}}}$ corresponding to the optimal RF RC $\mathbf{F}_{\text{RF}}^{\text{opt}}$ is obtained as

$$\mathbf{R}_{\mathcal{P}} = [\mathbf{r}_{\mathcal{P},1}, \mathbf{r}_{\mathcal{P},2}, \cdots, \mathbf{r}_{\mathcal{P},N_{\text{RF}}}] = [\mathbf{\Psi}_2(\mathcal{A},:)]^T, \quad (30)$$

where the vth column of $\mathbf{R}_{\mathcal{P}}$, denoted by $\mathbf{r}_{\mathcal{P},v} \in \mathbb{C}^{N_{p,2} \times 1}$ is comprised of $N_{p,2}$ pilot outputs corresponding to the vth RF chain. Let the equivalent baseband MIMO channel corresponding to the pth tap be defined as $\bar{\mathbf{H}}_p = \sqrt{\frac{N_t N_r}{L_p}} h_p \left(\mathbf{F}_{\text{RF}}^{\text{opt}}\right)^T \mathbf{a}_t^*(\theta_p) \mathbf{a}_r^T(\phi_p) \left(\mathbf{W}_{\text{RF}}^{\text{opt}}\right)^* \in \mathbb{C}^{N_{\text{RF}} \times N_{\text{RF}}}$, and let $\bar{h}_{p,v,u}$ denote its (v,u)th element. Constructing the baseband transmit pilot vector $\mathbf{s}_{\mathcal{P},u} \in \mathbb{C}^{N_{p,2} \times 1}$ corresponding to the uth transmit RF chain as $\mathbf{s}_{\mathcal{P},u} = \left[\mathbf{s}_{2,\mathcal{P},0}(u), \mathbf{s}_{2,\mathcal{P},1}(u), \cdots, \mathbf{s}_{2,\mathcal{P},N_{p,2}-1}(u)\right]^T$, the pilot output $\mathbf{r}_{\mathcal{P},v}$ at the vth receive RF chain can be formulated as

$$\mathbf{r}_{\mathcal{P},v} = \sum_{u=1}^{N_{\text{RF}}} \left[\sum_{p=1}^{L_p} \bar{h}_{p,v,u} \left(\bar{\mathbf{\Pi}}\right)^{i_p} \bar{\mathbf{\Delta}}_{i_p,j_p} \right] \mathbf{s}_{\mathcal{P},u} + \mathbf{w}_{\mathcal{P},v}, \quad (31)$$

where $\mathbf{w}_{\mathcal{P},v} \in \mathbb{C}^{N_{p,2} \times 1}$ represents the noise vector comprising the $N_{p,2}$ samples of the combined noise at the vth RF chain. Using (31), the DD-domain sparse CSI estimation model of estimating the equivalent channel $\bar{\mathbf{H}}_p$, similar to (20), can be formulated as

$$\mathbf{r}_{\mathcal{P},v} = \sum_{u=1}^{N_{\text{RF}}} \sum_{i=0}^{M_s-1} \sum_{j=0}^{G_s-1} \left[\left(\bar{\mathbf{\Pi}}\right)^i \bar{\mathbf{\Delta}}_{i,j} \mathbf{s}_{\mathcal{P},u} \right] h_{i,j,v,u} + \mathbf{w}_{\mathcal{P},v}, \quad (32)$$

where the quantity $h_{i,j,v,u}$ denotes the equivalent complex path-gain associated with the ith delay-grid, jth Doppler-grid between the vth receive RF chain and the uth transmit RF chain. Furthermore, upon substituting $\boldsymbol{\psi}_{i,j,u} = \left(\bar{\mathbf{\Pi}}\right)^i \bar{\mathbf{\Delta}}_{i,j} \mathbf{s}_{\mathcal{P},u}$ into (32), and then concatenating $\boldsymbol{\psi}_{i,j,u}, 0 \leq i \leq M_s-1, 0 \leq j \leq G_s-1$, $1 \leq u \leq N_{\text{RF}}$, for constructing the matrix $\bar{\mathbf{\Psi}} \in \mathbb{C}^{N_{p,2} \times M_s G_s N_{\text{RF}}}$, one can express the sparse DD-domain CSI estimation model for the vth output RF chain as

$$\mathbf{r}_{\mathcal{P},v} = \bar{\mathbf{\Psi}} \bar{\mathbf{h}}_v + \mathbf{w}_{\mathcal{P},v}, \quad (33)$$

where the vector $\bar{\mathbf{h}}_v \in \mathbb{C}^{M_s G_s N_{\text{RF}} \times 1}$ stacks the coefficients $h_{i,j,v,u}, \forall\ i,j,u$, in an appropriate indexing order. Furthermore, concatenating the quantities $\bar{\mathbf{h}}_v$ and $\mathbf{w}_{\mathcal{P},v}$ for all the RF chains as

$$\bar{\mathbf{H}} = \left[\bar{\mathbf{h}}_1, \bar{\mathbf{h}}_2, \cdots, \bar{\mathbf{h}}_{N_{\text{RF}}}\right] \in \mathbb{C}^{M_s G_s N_{\text{RF}} \times N_{\text{RF}}},$$
$$\mathbf{W}_{\mathcal{P},v} = [\mathbf{w}_{\mathcal{P},1}, \mathbf{w}_{\mathcal{P},2}, \cdots, \mathbf{w}_{\mathcal{P},N_{\text{RF}}}] \in \mathbb{C}^{N_{p,2} \times N_{\text{RF}}}, \quad (34)$$

the resultant estimation model of the DD-domain CSI of a mmWave-HB MIMO OTFS system is given as

$$\mathbf{R}_{\mathcal{P}} = \bar{\mathbf{\Psi}} \bar{\mathbf{H}} + \mathbf{W}_{\mathcal{P}}, \quad (35)$$

where the concatenated output pilot matrix $\mathbf{R}_{\mathcal{P}}$ has been defined in (30). Ignoring the output correlation across different RF chains, one can now acquire the MMSE estimate $\widehat{\mathbf{H}}_{\text{MMSE}} \in \mathbb{C}^{M_s G_s N_{\text{RF}} \times N_{\text{RF}}}$ of the DD-domain CSI matrix $\bar{\mathbf{H}}$ as

$$\widehat{\mathbf{H}}_{\text{MMSE}} = \left(\bar{\mathbf{\Psi}}^H \bar{\mathbf{\Psi}} + \sigma^2 \bar{\mathbf{R}}_h^{-1}\right)^{-1} \bar{\mathbf{\Psi}}^H \mathbf{R}_{\mathcal{P}}, \quad (36)$$

where $\bar{\mathbf{R}}_h \in \mathbb{C}^{M_s G_s N_{\text{RF}} \times M_s G_s N_{\text{RF}}}$ models the covariance matrix of the CSI vector $\bar{\mathbf{h}}_v$, which is also typically unknown and has to be estimated from the pilot output $\mathbf{R}_{\mathcal{P}}$. One can once again employ an enhanced BL-based framework for iteratively learning the prior covariance matrix $\bar{\mathbf{R}}_h$, which in turn yields the estimate $\widehat{\mathbf{H}}_{\text{MMSE}}$. To this end, it is important to understand the sparse structure inherent in the equivalent CSI matrix $\bar{\mathbf{H}}$, which is described next.

Without loss of generality, let us stack the coefficients $h_{i,j,v,u}$ of the equivalent CSI $\bar{\mathbf{h}}_v$ of (33), so that the last index u changes the fastest and the first index i changes the slowest. Therefore, as shown in Fig. 3, the CSI vector $\bar{\mathbf{h}}_v$ is comprised of the sub-vectors $\bar{\mathbf{h}}_{i,j,v} \in \mathbb{C}^{N_{\text{RF}} \times 1}$ as

$$\bar{\mathbf{h}}_v = \big[\bar{\mathbf{h}}_{0,0,v}^T, \bar{\mathbf{h}}_{0,1,v}^T, \cdots, \bar{\mathbf{h}}_{0,G_s-1,v}^T, \cdots, \bar{\mathbf{h}}_{i,j,v}^T, \cdots, \bar{\mathbf{h}}_{M_s-1,0,v}^T, \bar{\mathbf{h}}_{M_s-1,1,v}^T, \cdots, \bar{\mathbf{h}}_{M_s-1,G_s-1,v}^T\big]^T,$$

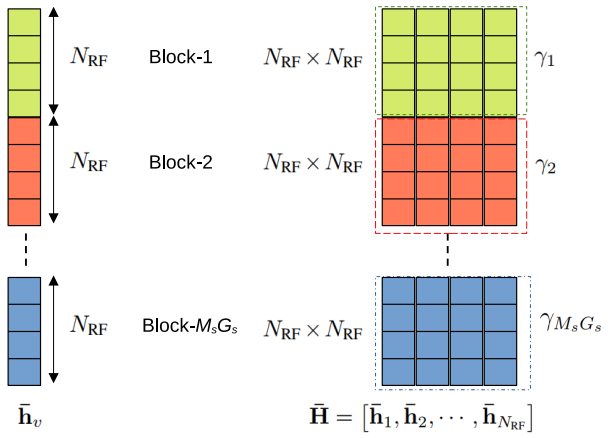

Fig. 3. Block-sparse structure of the DD-domain equivalent CSI of the mmWave-HB MIMO OTFS system.

where we have $\bar{\mathbf{h}}_{i,j,v} = \left[h_{i,j,v,1}, h_{i,j,v,2}, \cdots, h_{i,j,v,N_{\mathrm{RF}}}\right]^T$. Note that if the channel tap corresponding to the ith delay- and jth Doppler-index is non-zero, the N_{RF} elements of the sub-vector $\bar{\mathbf{h}}_{i,j,v}$ are potentially non-zero. On the other hand, if there is no multipath component present corresponding to the ith delay- and jth Doppler-index, all the N_{RF} elements of the sub-vector $\bar{\mathbf{h}}_{i,j,v}$ are zero. Thus, it can be readily observed that the CSI vector $\bar{\mathbf{h}}_v$ has M_sG_s blocks, where each block corresponds to a sub-vector $\bar{\mathbf{h}}_{i,j,v}$ having N_{RF} elements. Since each block can be entirely zero or non-zero, and typically, only L_p out of M_sG_s blocks are non-zero due to the limited number of multipath components, the vector $\bar{\mathbf{h}}_v$ has a block-sparse (BS) structure, as depicted in Fig. 3. Coming next to the concatenated CSI matrix $\bar{\mathbf{H}}$ of (34), since the locations of the non-zero delay- and Doppler-indices do not change for different antennas/ RF chains, the vectors $\bar{\mathbf{h}}_v, 1 \leq v \leq N_{\mathrm{RF}}$, possess an identical sparsity profile. This leads to an interesting structure of the matrix $\bar{\mathbf{H}}$, where the elements present in an $N_{\mathrm{RF}} \times N_{\mathrm{RF}}$ sub-matrix, i.e., a block of rows, become either all zero or non-zero. Incorporating this specific sparse structure in the BL-based framework can further enhance the accuracy of the estimated CSI.

In this context, we develop a BS-BL technique for exploiting the block-sparsity of the equivalent DD-domain CSI of the mmWave-HB MIMO OTFS system, which employs the parameterized prior covariance $\bar{\mathbf{R}}_h$ as $\bar{\mathbf{R}}_h = (\mathbf{\Gamma} \otimes \mathbf{I}_{N_{\mathrm{RF}}})$, for each BS vector $\bar{\mathbf{h}}_v$. This implies that the parameterized Gaussian prior $p\left(\bar{\mathbf{H}}; \mathbf{\Gamma}\right)$ for the DD-domain CSI matrix $\bar{\mathbf{H}}$ can be expressed as $p\left(\bar{\mathbf{H}}; \mathbf{\Gamma}\right) = \prod_{v=1}^{N_{\mathrm{RF}}} p\left(\bar{\mathbf{h}}_v; \mathbf{\Gamma}\right)$, where $p\left(\bar{\mathbf{h}}_v, \mathbf{\Gamma}\right) = \mathcal{CN}\left(\mathbf{0}, \mathbf{\Gamma} \otimes \mathbf{I}_{N_{\mathrm{RF}}}\right)$. Note that the prior above assigns an identical hyperparameter to each element of the $N_{\mathrm{RF}} \times N_{\mathrm{RF}}$ size block of the matrix $\bar{\mathbf{H}}$, as shown in Fig. 3. Thus, thanks to BS-sparsity, we only have to estimate M_sG_s hyperparameters, which is significantly lower than the number of elements $M_sG_sN_{\mathrm{RF}}^2$ in the CSI matrix $\bar{\mathbf{H}}$, hence improving the estimation accuracy. These hyperparameters can be once again estimated iteratively using the EM procedure. The various steps of the proposed BS-BL framework for the sparse DD-domain CSI estimator have been described in Algorithm-2, which follows from our paper [49]. Due to lack of space, the derivation of BCRLB corresponding to the MSE of the estimate of the DD-domain CSI matrix $\bar{\mathbf{H}}$ is given in our technical report [50].

Algorithm 2 BS-BL based BS-sparse CSI estimation in mmWave-HB MIMO OTFS systems

Input: Pilot output $\mathbf{R}_{\mathcal{P}}$, dictionary matrix $\bar{\mathbf{\Psi}}$, stopping parameters η and $S_{\max}$

Initialization: $\widehat{\gamma}_i^{(0)} = 1, \forall\, i \Rightarrow \widehat{\mathbf{\Gamma}}^{(0)} = \mathbf{I}_{M_sG_s}$, counter $j = 0$, $\widehat{\mathbf{\Gamma}}^{(-1)} = \mathbf{0}$

while $\left(\left\|\widehat{\mathbf{\Gamma}}^{(j)} - \widehat{\mathbf{\Gamma}}^{(j-1)}\right\|_F^2 \geq \eta \;\text{ and }\; j < S_{\max}\right)$ **do**

1) $j = j + 1$
2) **E-step:** Compute the *a posteriori* covariance and mean

$$\bar{\mathbf{\Sigma}}^{(j)} = \left[\sigma^{-2}\bar{\mathbf{\Psi}}^H\bar{\mathbf{\Psi}} + \left(\widehat{\mathbf{\Gamma}}^{(j-1)}\right)^{-1} \otimes \mathbf{I}_{N_{\mathrm{RF}}}\right]^{-1},$$

$$\mathbf{\mathcal{M}}^{(j)} = \sigma^{-2}\bar{\mathbf{\Sigma}}^{(j)}\bar{\mathbf{\Psi}}^H\mathbf{R}_{\mathcal{P}}$$

3) **M-step:** Update the hyperparameters

for $i = 1, 2, \cdots, M_sG_s$ **do**

$$\widehat{\gamma}_i^{(j)} = \frac{1}{N_{\mathrm{RF}}^2} \sum_{k=[(i-1)N_{\mathrm{RF}}+1]}^{iN_{\mathrm{RF}}} \sum_{l=1}^{N_{\mathrm{RF}}} \mathbf{\mathcal{M}}^{(j)}(k,l) + \frac{1}{N_{\mathrm{RF}}} \sum_{k=[(i-1)N_{\mathrm{RF}}+1]}^{iN_{\mathrm{RF}}} \left|\bar{\mathbf{\Sigma}}^{(j)}(k,k)\right|^2$$

end for

$\widehat{\mathbf{\Gamma}}^{(j)} = \mathrm{diag}\left(\widehat{\gamma}_1^{(j)}, \widehat{\gamma}_2^{(j)}, \cdots, \widehat{\gamma}_{M_sG_s}^{(j)}\right)$

end while

Output: $\widehat{\mathbf{H}}_{\text{BS-BL}} = \left[\bar{\mathbf{\Psi}}^H\bar{\mathbf{\Psi}} + \sigma^2\left(\widehat{\mathbf{\Gamma}}^{(j)}\right)^{-1} \otimes \mathbf{I}_{N_{\mathrm{RF}}}\right]^{-1}\bar{\mathbf{\Psi}}^H\mathbf{R}_{\mathcal{P}}$

TABLE II
SIMULATION PARAMETERS

Parameter	System-I	System-II
Carrier frequency (GHz)	28	60
Subcarrier spacing (KHz) (Δf)	625	1562.5
Symbols on delay-axis (M)	32	64
Symbols on Doppler-axis (N)	32	64
TAs/RAs (N_t, N_r)	$(32, 32)$	$(32, 32)$
Pilots in time-domain $(N_{p,1}, N_{p,2})$	$(50, 150)$	$(100, 300)$
Samples in CP (L)	16	24
Multipath components (L_p)	5	5
Angular grid-sizes $(G_{\mathrm{A}}, G_{\mathrm{B}})$	$(64, 64)$	$(64, 64)$
Delay-grid size (M_s)	16	16
Doppler-grid size (G_s)	32	64
Modulation scheme	8-PSK	8-PSK
Pulse-shape	Rectangular	Rectangular

VI. Simulation Results

In this section, the performance of the proposed transceiver design and BL-based sparse CSI estimation schemes is quantified for our mmWave-AB and mmWave-HB MIMO OTFS systems. The performance is also compared to that of other state-of-the-art sparse signal recovery techniques, such as OMP and the FOCal underdetermined system solver (FOCUSS) [51], and benchmarked against the BCRLBs derived

in our technical report [50] for mmWave-AB and mmWave-HB MIMO OTFS systems, respectively. To this end, the normalized MSE (NMSE) of the equivalent CSI is defined as $\frac{\left\|\widehat{\mathbf{H}}_{\mathrm{DD}}-\mathbf{H}_{\mathrm{DD}}\right\|_F^2}{\left\|\mathbf{H}_{\mathrm{DD}}\right\|_F^2}$ for mmWave-AB MIMO OTFS systems, and as $\frac{\left\|\widehat{\bar{\mathbf{H}}}_{\mathrm{DD}}-\bar{\mathbf{H}}_{\mathrm{DD}}\right\|_F^2}{\left\|\bar{\mathbf{H}}_{\mathrm{DD}}\right\|_F^2}$ for mmWave-HB MIMO OTFS systems. The resultant symbol-error-rates (SERs) have also been illustrated for demonstrating the end-to-end performance of the proposed transceiver designs followed by sparse CSI estimation techniques. The SNR in decibels (dB) is defined as $10\log_{10}\left(\frac{1}{\sigma^2}\right)$. Table-II shows the detailed system parameters for System-I and System-II considered in our simulations, unless specified otherwise.

Let us consider System-I, as described in Table-I of the paper, for which the bandwidth is $B = M\Delta f = 20$ MHz. For this system, the sampling interval evaluates to $T_s = \frac{1}{B} = 5 \times 10^{-8}$ sec. The size of a packet comprising Stage-1 and Stage-2 pilot symbols $(N_{p,1} + N_{p,2})$, one OTFS data frame of MN symbols, and their cyclic prefixes (CPs) of length L each, evaluates to $N_{p,1} + N_{p,2} + MN + 3L = 1272$, which leads to a packet duration of the proposed mmWave MIMO OTFS System-I as 6.36×10^{-5} sec. Now considering a velocity of 200 Km/h, the distance covered by the mobile in the packet duration calculated above evaluates to 0.353 cm, which is negligible with respect to a typical 5G-cell radius of approximately 200 meters. This example clarifies that the AoAs and AoDs of the multipath components remain approximately constant for several such packets (~ 10). Therefore, the beams selected for the TBF and RC can be employed for several OTFS data frames. Furthermore, it is also important to note that the AoA/ AoD reciprocity holds in both TDD and FDD mode [45], [46]. Due to this key property of the mmWave MIMO channel, the beam pattern selected for the TBF can also be used at the RC. Hence no feedback is required for these beam patterns.

A. mmWave-AB MIMO OTFS Systems

Fig. 4(a) and Fig. 4(b) demonstrate the NMSE versus SNR performance of the competing schemes for the mmWave-AB MIMO OTFS System-I and System-II, respectively. Note that the DD-domain sparse CSI estimation problem under consideration is underdetermined, since the number of pilots $N_{p,2}$ is less than the length M_sG_s of the low-dimensional equivalent CSI. For this scenario, it can be readily observed that the proposed BL-based sparse CSI estimation scheme significantly outperforms both OMP and FOCUSS. The poor performance of the OMP is attributed to its sensitivity to the stopping parameter as well as to the dictionary matrix $\mathbf{\Psi}$. Furthermore, the poor performance of FOCUSS is due to its sensitivity to the regularization parameter, which ultimately leads to convergence deficiencies. On the other hand, for the BL-based techniques, convergence is guaranteed to a sparse solution due to the well-established properties of the EM algorithm and BL cost function [47]. Furthermore, the NMSE of the BL technique is also seen to be very close to the BCRLB. This is significant, since the BCRLB has been plotted considering a hypothetical scenario where the DD-domain profile, i.e., the sparse locations in the equivalent CSI $\mathbf{h}$, is perfectly known, whereas the proposed BL technique does not need any such knowledge. This demonstrates the robust performance of the BL technique in comparison to the other competing sparse signal recovery algorithms. The improved CSI estimation performance of the proposed BL-based technique is also reflected in the SER versus SNR plot of Fig. 4(c). It can be readily observed that the proposed BL technique significantly outperforms the other competing schemes and achieves an SER close to that of the PCSI scenario. This is attributed to its improved CSI estimation accuracy.

B. mmWave-HB MIMO OTFS Systems

Turning our attention to the performance of the proposed BS-BL technique designed for our mmWave-HB MIMO OTFS systems, these have been demonstrated in Fig. 5(a) and Fig. 5(b) in terms of the NMSE and SER, respectively. The performance is also compared to that of the BL, M-BL, OMP, BS-OMP and MFOCUSS, and benchmarked against the BCRLB. Once again, the sparse DD-domain CSI estimation model of Eq. (35) becomes highly underdetermined, since one has to estimate the $M_sG_sN_{\mathrm{RF}}^2$ channel coefficients in the equivalent low-dimensional CSI $\bar{\mathbf{H}}$ from only $N_{p,2}N_{\mathrm{RF}}$ pilot outputs of the received pilot matrix $\mathbf{R}_{\mathcal{P}}$. However, as depicted in Fig. 3, the proposed BS-BL technique assigns an identical hyperparameter to each element of the $[N_{\mathrm{RF}} \times N_{\mathrm{RF}}]$-size block in order to exploit the inherent block-sparse nature of the CSI, thus requiring only M_sG_s hyperparameters for estimation. This significantly improves the CSI estimation, as seen from Fig. 5(a), where the proposed BS-BL technique outperforms all the other competing schemes and also achieves the BCRLB. The inferior performance of the BL and M-BL techniques with respect to the BS-BL is attributed to the fact that these schemes only exploit sparsity and row-sparsity, respectively, but not block-sparsity, hence estimating $M_sG_sN_{\mathrm{RF}}$ hyperparameters. Furthermore, since the M-BL scheme estimates these hyperparameters jointly for all the N_{RF} columns of $\bar{\mathbf{H}}$, it yields a superior NMSE with respect to the BL technique. Similarly, the BS-OMP technique, which exploits the BS-sparsity, yields a lower NMSE in comparison to the conventional OMP. Finally, the improved CSI estimation performance of the BS-BL technique is also reflected in its enhanced SER in Fig. 5(b), which is also very close to that of the PCSI scenario. This further demonstrates the enhanced efficacy of the proposed hybrid transceiver design developed in Section-V for mmWave-HB MIMO OTFS systems.

Fig. 5(c) demonstrates the NMSE of the channel frequency response (CFR) for mmWave-HB MIMO OFDM systems using the relevant state-of-the-art solutions, such as G-SBL [18], SOMP [17] and DSDS [23]. In particular, the CSI estimation performance of [18] and [17] is poor, since they ignore the effect of Doppler arising within an OFDM symbol. Due to this, their channel estimation models become inconsistent in the presence of Doppler, and their CSI estimation performance degrades further upon increasing the Doppler. It

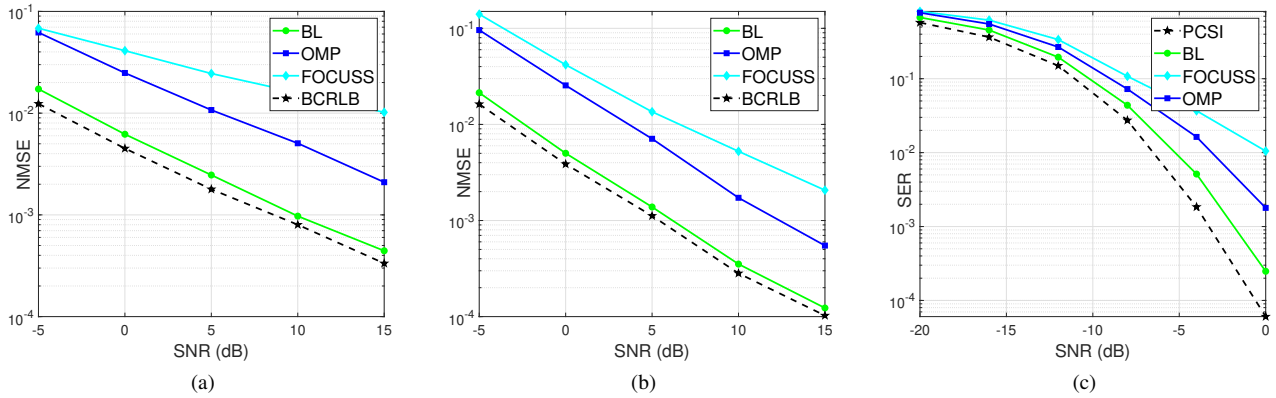

Fig. 4. (a) NMSE versus SNR performance of the mmWave-AB MIMO OTFS System-I; (b) NMSE versus SNR performance of the mmWave-AB MIMO OTFS System-II; (c) SER versus SNR performance of the mmWave-AB MIMO OTFS System-I.

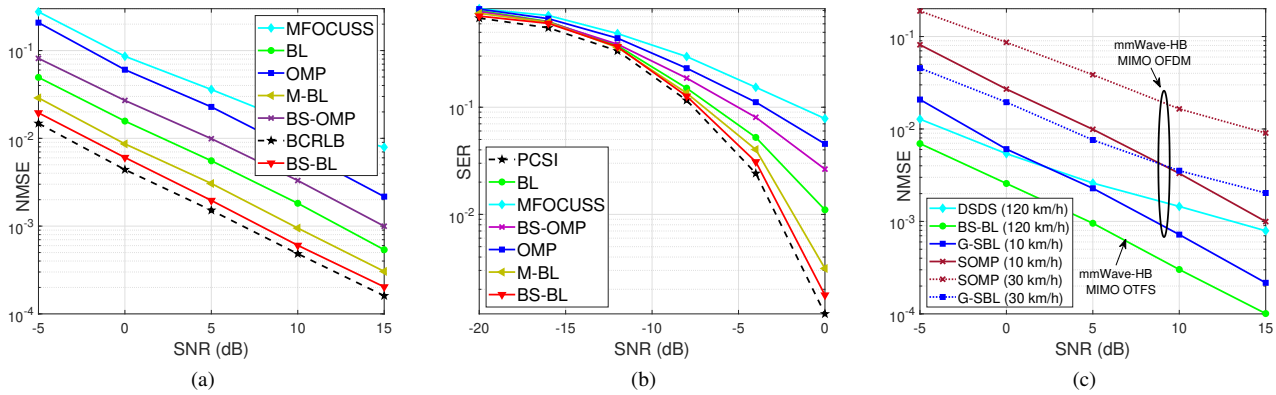

Fig. 5. (a) NMSE versus SNR performance of the mmWave-HB MIMO OTFS System-I with $N_{\text{RF}} = 2$ RF chains; (b) SER versus SNR performance of the mmWave-HB MIMO OTFS System-I with $N_{\text{RF}} = 2$ RF chains; (c) NMSE performance comparison of the mmWave-HB MIMO OTFS System-II with respect to the corresponding mmWave-HB MIMO OFDM system.

is also important to note that both [18] and [17] exploit only the angular-domain sparsity of the mmWave MIMO channel. The improved performance of the 3-stage scheme developed in [23] is attributed to the fact that it accounts for the Doppler in the channel estimation model derived and also exploits the associated delay-domain sparsity along with the angular-sparsity. However, it does not exploit the Doppler-domain sparsity. Finally, the superior performance of our proposed CSI estimation techniques for the mmWave MIMO OTFS systems is achieved via exploiting the joint sparsity of the delay, Doppler and angular domains.

C. Pilot Overhead

We would also like to demonstrate that the overhead of the proposed transceiver design and DD-domain CSI estimation procedure is very low. This is because Stage-1 designs TBF/RC for Device-A using $N_{p,1}$ pilot beams, while Stage-2 designs TBF/RC for Device-B using $N_{p,2}$ pilot symbols, and also estimates the DD-domain equivalent CSI from the output of these $N_{p,2}$ pilot symbols. Therefore, the proposed 2-stage procedure completes both TBF/ RC optimization as well as DD-domain CSI estimation using $(N_{p,1}+N_{p,2})$ pilots, where $N_{p,1} << N_{p,2} << MN$. Let us now consider the Stage-1, Stage-2 pilot symbols together with one OTFS data frame. The pilot overhead ρ evaluates to $\rho = \frac{N_p}{MN+N_p}$, where $N_p = N_{p,1} + N_{p,2}$. For the System-1 considered in our simulations, the pilot overhead ρ evaluates to 0.16, which is not high. And in fact, the actual pilot overhead is lower than this, since one can transmit several OTFS data frames on the same aligned beams before the DDA-domain CSI changes. For example, as demonstrated at the beginning of our simulation results, one can readily transmit approximately 10 OTFS data frames before performing the next beam-alignment and DD-domain CSI estimation procedure. This further reduces the pilot overhead to $\rho = \frac{N_p}{10MN+N_p} \approx 0.019$, i.e. only 1.9%, which is negligible. In addition, it is also important to note that the AoA/ AoD reciprocity holds in both TDD and FDD modes [45], [46]. Due to this key property of the mmWave MIMO channel, the beam pattern selected for the TBF can also be used as the RC. Hence no feedback is required for these beam patterns.

Coming next to the overhead comparison of the proposed method with the existing ones, to the best of our knowledge, none of the existing contributions have explored

the mmWave MIMO OTFS systems considering analog and hybrid beamforming architectures. Although some existing beam-alignment works, such as [44], are indeed suitable for mmWave MIMO OFDM systems, and they perform TBF/ RC optimization in a single stage, it is not immediately clear how this can be extended to mmWave MIMO OTFS systems. More specifically, [44] considers only the AoA/ AoD plus delay of each multipath component, but it ignores their Doppler. It seems that extending a similar beam-alignment strategy to mmWave MIMO OTFS systems considering Doppler will require further improvements in the channel estimation model derived, and it may also need additional stages/ pilots for estimating the beamformed channel coefficients for efficient data detection. By contrast, other OTFS-specific solutions, such as those in [27], [30], [36], [41], do indeed consider fully-digital sub-6 GHz MIMO systems, which do not perform the beam-alignment step. In particular, for mmWave MIMO OTFS systems relying on analog/ hybrid beamforming, which employ a pilot-embedded OTFS data frame similar to [27], [30], [36], [41], the procedure of coupling the beam-alignment with the DD-domain CSI estimation is not described. Note that the TBF/ RC design requires the transmission of several probing beams, which excite various possible angular modes of the mmWave MIMO channel. Setting one probing beam per pilot-embedded data OTFS frame for achieving this is not practical, since it will require an overhead of several OTFS frames for beam-alignment. By contrast, using multiple probing beams within an OTFS frame is equivalent to varying the complex-valued DD-domain path gains of the multipath components within an OTFS frame, which violates the fundamental assumptions routinely stipulated for OTFS systems. Hence, a fair comparison of the overhead for the proposed beam-alignment and DD-domain CSI estimation procedure requires further investigation of mmWave MIMO OTFS systems by the research community. In fact, it can be an interesting future problem to develop a single-stage solution, which designs the TBF/ RC at both the devices along with estimating the equivalent DD-domain CSI.

VII. Conclusions

We conceived mmWave-AB and mmWave-HB MIMO OTFS architectures, which have the potential to yield high data rates in high-mobility scenarios. The end-to-end DD-domain relationships were derived for both these mmWave MIMO architectures considering a delay-Doppler-angular domain channel model, followed by developing a two-stage procedure for designing the RF TBF/ TPC and RF RC, and estimating the equivalent low-dimensional DD-domain sparse CSI using BL-based techniques. The BCRLB was also derived for benchmarking the MSE of the DD-domain CSI estimates. The SER achieved using the proposed two-stage procedure is seen to be close to that of the perfect CSI scenario, which demonstrates the efficacy of the technique. It is also demonstrated that the BL and BS-BL techniques proposed for CSI estimation efficiently exploit the sparsity and block-sparsity, respectively, and require a significantly lower pilot overhead thanks to the key results in compressive sensing theory, which allows us to recover a sparse signal from a very small number of measurements. Furthermore, the proposed CSI estimation framework is eminently suitable for large antenna arrays, since it transmits pilots in a dedicated time-frequency resource block without necessitating multiple DD-domain guard intervals to be placed within the same OTFS frame. A possible future extension of this work can be estimating the AoA, AoD, delay and Doppler jointly in a single stage.

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] 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.

[4] 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.

[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, ArXiv:1808.00519*, 2018.

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

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

[9] C. Heil, "A discrete ZAK transform," 1989.

[10] T. S. Rappaport, R. W. Heath Jr, R. C. Daniels, and J. N. Murdock, *Millimeter wave wireless communications*. Pearson Education, 2015.

[11] I. A. Hemadeh, K. Satyanarayana, M. El-Hajjar, and L. Hanzo, "Millimeter-wave communications: Physical channel models, design considerations, antenna constructions, and link-budget," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 2, pp. 870–913, 2017.

[12] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 831–846, 2014.

[13] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 13, no. 3, pp. 1499–1513, 2014.

[14] J. Lee, G.-T. Gil, and Y. H. Lee, "Channel estimation via orthogonal matching pursuit for hybrid MIMO systems in millimeter wave communications," *IEEE Transactions on Communications*, vol. 64, no. 6, pp. 2370–2386, 2016.

[15] S. Srivastava, A. Mishra, A. Rajoriya, A. K. Jagannatham, and G. Ascheid, "Quasi-static and time-selective channel estimation for block-sparse millimeter wave hybrid MIMO systems: Sparse Bayesian learning (SBL) based approaches," *IEEE Transactions on Signal Processing*, vol. 67, no. 5, pp. 1251–1266, 2018.

[16] K. Venugopal, A. Alkhateeb, N. G. Prelcic, and R. W. Heath, "Channel estimation for hybrid architecture-based wideband millimeter wave systems," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 9, pp. 1996–2009, 2017.

[17] J. Rodríguez-Fernández, N. González-Prelcic, K. Venugopal, and R. W. Heath, "Frequency-domain compressive channel estimation for frequency-selective hybrid millimeter wave MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 17, no. 5, pp. 2946–2960, 2018.

[18] S. Srivastava, C. S. K. Patro, A. K. Jagannatham, and L. Hanzo, "Sparse, group-sparse and online Bayesian learning aided channel estimation for doubly-selective mmwave hybrid MIMO OFDM systems," *IEEE Transactions on Communications*, 2021.

[19] S. Srivastava, P. Singh, A. K. Jagannatham, A. Karandikar, and L. Hanzo, "Bayesian learning-based doubly-selective sparse channel estimation for millimeter wave hybrid MIMO-FBMC-OQAM systems," *IEEE Transactions on Communications*, vol. 69, no. 1, pp. 529–543, 2020.

[20] S. Srivastava, J. Nath, and A. K. Jagannatham, "Data aided quasistatic and doubly-selective CSI estimation using affine-precoded superimposed pilots in millimeter wave MIMO-OFDM systems," *IEEE Transactions on Vehicular Technology*, 2021.

[21] W. C. Jakes and D. C. Cox, *Microwave mobile communications*. Wiley-IEEE Press, 1994.

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

[23] S. Gao, X. Cheng, and L. Yang, "Estimating doubly-selective channels for hybrid mmwave massive MIMO systems: A doubly-sparse approach," *IEEE Transactions on Wireless Communications*, vol. 19, no. 9, pp. 5703–5715, 2020.

[24] 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.

[25] 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.

[26] 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.

[27] 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.

[28] 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.

[29] 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.

[30] 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.

[31] 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.

[32] ——, "Bayesian learning aided simultaneous row and group sparse channel estimation in orthogonal time frequency space modulated MIMO systems," *IEEE Transactions on Communications*, 2021.

[33] 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.

[34] S. K. Mohammed, "Derivation of OTFS modulation from first principles," *IEEE transactions on vehicular technology*, vol. 70, no. 8, pp. 7619–7636, 2021.

[35] F. Lampel, A. Alvarado, and F. M. Willems, "Orthogonal time frequency space modulation: A discrete Zak transform approach," *arXiv preprint arXiv:2106.12828*, 2021.

[36] L. Gaudio, G. Colavolpe, and G. Caire, "OTFS vs. OFDM in the presence of sparsity: A fair comparison," *IEEE Transactions on Wireless Communications*, 2021.

[37] L. Gaudio, M. Kobayashi, G. Caire, and G. Colavolpe, "On the effectiveness of OTFS for joint radar parameter estimation and communication," *IEEE Transactions on Wireless Communications*, vol. 19, no. 9, pp. 5951–5965, 2020.

[38] Z. Wei, W. Yuan, S. Li, J. Yuan, and D. W. K. Ng, "Transmitter and receiver window designs for orthogonal time-frequency space modulation," *IEEE transactions on communications*, vol. 69, no. 4, pp. 2207–2223, 2021.

[39] 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*, vol. 20, no. 9, pp. 6033–6048, 2021.

[40] G. Surabhi, M. K. Ramachandran, and A. Chockalingam, "OTFS modulation with phase noise in mmWave communications," in *2019 IEEE 89th Vehicular Technology Conference (VTC2019-Spring)*. IEEE, 2019, pp. 1–5.

[41] 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.

[42] 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.

[43] G. Surabhi, R. M. Augustine, and A. Chockalingam, "Peak-to-average power ratio of OTFS modulation," *IEEE Communications Letters*, vol. 23, no. 6, pp. 999–1002, 2019.

[44] X. Song, S. Haghighatshoar, and G. Caire, "A scalable and statistically robust beam alignment technique for millimeter-wave systems," *IEEE Transactions on Wireless Communications*, vol. 17, no. 7, pp. 4792–4805, 2018.

[45] Y. Ding and B. D. Rao, "Dictionary learning-based sparse channel representation and estimation for FDD massive MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 17, no. 8, pp. 5437–5451, 2018.

[46] K. Hugl, K. Kalliola, J. Laurila *et al.*, "Spatial reciprocity of uplink and downlink radio channels in FDD systems," in *Proc. COST*, vol. 273, no. 2. Citeseer, 2002, p. 066.

[47] 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.

[48] S. Srivastava, A. Tripathi, N. Varshney, A. K. Jagannatham, and L. Hanzo, "Hybrid transceiver design for Tera-Hertz MIMO systems relying on Bayesian learning aided sparse channel estimation," *arXiv preprint arXiv:2109.09664*, 2021.

[49] S. Srivastava, M. P. Suradkar, and A. K. Jagannatham, "BSBL-based block-sparse channel estimation for affine precoded OSTBC MIMO-OFDM systems," in *2020 National Conference on Communications (NCC)*. IEEE, 2020, pp. 1–6.

[50] S. Srivastava, R. K. Singh, and A. K. Jagannatham, "Technical report: Transceiver design and sparse doubly-selective CSI estimation in OTFS aided analog and hybrid beamforming mmWave MIMO." IIT Kanpur, Tech. Rep., 2021. [Online]. Available: http://www.iitk.ac.in/mwn/documents/MWNLab_TR_mmWave_MIMO_OTFS.pdf.

[51] 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.