Sparse Channel Estimation for MIMO OTFS/OTSM Systems Using Finite-Resolution ADCs

Anand Mehrotra, *Graduate Student Member, IEEE*, Suraj Srivastava, *Member, IEEE*, N. Shanmughanadha Reddy, Aditya Jagannatham, *Senior Member, IEEE*, and Lajos Hanzo, *Life Fellow, IEEE*

***Abstract*— Variational Bayesian learning (VBL)-based sparse channel state information (CSI) estimation is conceived for multiple input multiple output (MIMO) orthogonal time frequency space (OTFS) and for orthogonal time sequence multiplexing (OTSM)-based systems relying on low-resolution analog-to-digital convertors (ADCs). First, the CSI estimation model is developed for MIMO-OTFS systems considering quantized outputs. Then a novel VBL technique is developed for exploiting the inherent DD domain sparsity. Subsequently, an end-to-end system model is derived for MIMO-OTSM systems, once again, using only finite-resolution ADCs. Similar to OTFS systems, it is demonstrated that the channel is sparse in the delay-sequency (DS)-domain. Thus the sparse CSI estimation problem of the MIMO-OTSM system can also be solved using the VBL technique developed for its OTFS counterpart. A bespoke minimum mean square error (MMSE) receiver is developed for data detection, which unlike the conventional MMSE receiver also accounts for the quantization error. Finally, finite-resolution ADCs emerge as a solution, offering reduced costs and energy consumption amid the growing challenge posed by energy-intensive high-resolution ADCs in Next-Generation (NG) systems. The efficacy of the proposed techniques is validated by simulation results, surpassing the state-of-the-art and signalling a transition towards more sustainable communication technologies.**

***Index Terms*— OTFS, delay-Doppler, OTSM, delay-sequency, sparsity, channel estimation, finite-resolution ADCs.**

Received 17 April 2024; revised 10 August 2024 and 2 November 2024; accepted 10 November 2024. Date of publication 20 November 2024; date of current version 18 June 2025. The work of Suraj Srivastava was supported by IIT Jodhpur's Research Grant No. I/RIG/SUS/20240043.The work of Aditya Jagannatham was supported in part by the Qualcomm Innovation Fellowship; in part by the Qualcomm 6G UR Gift; and in part by the Arun Kumar Chair Professorship. Lajos Hanzo would like to acknowledge the financial support of the Engineering and Physical Sciences Research Council (EPSRC) projects under grant EP/Y037243/1, EP/W016605/1, EP/X01228X/1, EP/Y026721/1, EP/W032635/1 and EP/X04047X/1 as well as of the European Research Council's Advanced Fellow Grant QuantCom (Grant No. 789028). The associate editor coordinating the review of this article and approving it for publication was C.-K. Wen. *(Corresponding author: Lajos Hanzo.)*

Anand Mehrotra and Aditya Jagannatham are with the Department of Electrical Engineering, Indian Institute of Technology Kanpur, Kanpur, Uttar Pradesh 208016, India (e-mail: anandme@iitk.ac.in; adityaj@iitk.ac.in).

Suraj Srivastava is with the Department of Electrical Engineering, Indian Institute of Technology Jodhpur, Jheepasani 342030, India (e-mail: surajsri@iitj.ac.in).

N. Shanmughanadha Reddy is with Rakuten Mobile, Tokyo 158-0094, Japan (e-mail: nshanmugha123@gmail.com).

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

Digital Object Identifier 10.1109/TCOMM.2024.3502682

I. Introduction

THE recent advances in the field of wireless communication have created both new opportunities and challenges. Next-generation (NG) networks require high-rate, low-latency communication in high-mobility scenarios, which poses a significant challenge. In these conditions, the widely used orthogonal frequency division multiplexing (OFDM) scheme struggles due to degradations arising from the time-selective nature of wireless channels [1], [2], [3].

To overcome these deleterious effects, a pair of seminal modulation schemes have attracted considerable attention in recent research, termed as orthogonal Time-Frequency Space (OTFS) [4], [5], [6] and orthogonal time sequency multiplexing (OTSM) [7], [8]. While the former is a multi-carrier modulation scheme, the latter is a single-carrier modulation scheme. In contrast to the conventional time-frequency (TF) domain representation of OFDM, OTFS is a delay-Doppler (DD)-domain modulation technique. The DD representation is quasi-time-invariant, which beneficially simplifies the estimation of the wireless channel and enables OTFS to achieve a degree of resilience to high-Doppler. This property in turn, makes it well-suited for ultra-high mobility scenarios. By contrast, in an OTSM system, the information symbols are located in the delay-sequency (DS) domain, where sequency refers to the number of zero-crossings per unit time interval. OTSM, similar to OTFS, provides considerable performance gains over OFDM at a reduced complexity [7].

Another noteworthy advantage of OTFS and OTSM systems is that they can be readily implemented by relying on the existing OFDM-based systems upon appending suitable precoding and decoding blocks [7], [9], [10], which makes them eminently amenable for practical implementation. However, to reap the full benefits these modulation schemes, having accurate CSI estimation is essential. A concise literature survey of CSI estimation and detection techniques conceived for OTFS and OTSM systems is presented in the next section.

A. State-of-the-Art in OTFS CSI Estimation

Numerous OTFS CSI estimation schemes have been proposed by various authors since its inception. The early research in this area concentrated on using the principle of impulse-based CSI estimation [11], [12]. In [12], a complete OTFS frame was transmitted for CSI estimation, leading to

increased pilot overhead. To address this issue, an embedded pilot-based CSI estimation scheme was proposed in [11] and [13]. In this technique, the pilots and data were transmitted in the same OTFS frame, thus improving the bandwidth efficiency. However, this approach requires the placement of guard bands around pilots to separate the interference among the corresponding outputs, which prevents the system from achieving its full potential with respect to spectral efficiency. Furthermore, the threshold-based estimation technique[1] employed therein [11], [12], [13] requires empirical tuning, which renders the performance sensitive to this parameter.

Additionally, due to the small number of scatterers in the DD-domain, one of the key characteristics of the OTFS channel is its inherent sparsity, which can be exploited for improving the accuracy of CSI acquisition [14]. To achieve this goal Shen et al. [9] proposed an orthogonal matching pursuit (OMP)-based scheme for harnessing the sparsity of the OTFS channel. Along similar lines, a modified OMP technique was proposed by Li et al. [15] that targets CSI estimation in the TF domain. Li and Yu [16] also presented a sparsity-agnostic turbo MMSE algorithm based on graph theory. Despite significantly outperforming the traditional MMSE strategy, their state-of-the-art system still performs worse than the Bayesian learning (BL) approach.

The superior performance of BL-based algorithms is due to its ability to exploit the sparsity of the underlying channel [17], [18]. More specifically, the authors of [17] proposed a BL-based CSI estimation algorithm that can be applied in a system having arbitrary transmitter and receiver pulse shaping filters. Furthermore, to reduce the overhead and power requirements Zhao et al. [19] have proposed a modified BL algorithm. In order to exploit the row and group (RG) sparsity of the channel, the authors of [20] have proposed a sophisticated RG-BL algorithm. Recent advances by the authors of [21], [22], and [23] have explored an off-grid sparse Bayesian inference that updates the virtual delay-Doppler grid to improve channel estimation accuracy in OTFS systems, which enables their framework to consider fractional Doppler shifts. Thus, the above works have achieved substantial progress in developing Bayesian techniques that effectively exploit the sparse structure of the DD domain channel. A brief review of recent contributions to OTSM-modulated systems is presented next.

B. State-of-the-Art in OTSM CSI Estimation

OTSM is a relatively new modulation scheme compared to OTFS, which was proposed to provide as low-complexity alternative to OTFS. The initial contribution on CSI estimation for OTSM [7] proposed an embedded pilot-based technique, wherein the delay-sequency grids of both the data and pilot symbols comprise sufficient guard bands, while CSI estimation is performed in the delay-time (DT)-domain. However, the inherent sparsity of the channel is not exploited in this work. An important aspect of CSI estimation in both OTFS and OTSM-based systems is the placement of pilot symbols. In OFDM-based systems, the pilots are typically loaded in the frequency domain [26]. By contrast, in an OTFS-based system, pilots can be placed either in the DD-domain [17], [19], [25] or in the TF-domain [20], [24] or alternatively in the time-domain (TD). Similarly, for OTSM-based systems, the pilots can be positioned in the DS-domain [7], [8] or the DT-domain. The authors of [28] propose frequency-domain pilot-aided, time-domain pilot-aided, and time-domain training sequence-based approaches for CSI estimation. The drawback of transmitting pilots in the DD- or TF- or DS-domain is their extra processing, which can be avoided when the pilots are placed directly in the time-domain for OTFS and OTSM systems.

C. Identification of Challenges in CSI Estimation

Despite the significant progress made by above contributions in the arena of CSI estimation for OTFS and OTSM systems, some research gaps still remain. For instance, the bandwidth is underutilized in [11], [12], and [13], while the OMP-based algorithms proposed in [9] and [15] are highly sensitive to the stopping parameter. Bayesian learning [17], [19], [20], [24], [25] has already been widely explored in OTFS modulated systems, but its application in OTSM modulations systems has hitherto not been characterised. Another important consideration in NG systems is energy consumption, which can be limited by the use of finite-resolution ADCs. However, in practical implementations, quantization error introduced by the use of finite resolution ADCs is inevitable and can significantly impact the performance of OTFS/ OTSM CSI estimation. This impairment manifests itself as additional noise, which degrades the CSI estimation performance.

The signal processing algorithms using limited-resolution ADCs have been well explored using frequency domain (FD) pilots in the literature on OFDM-based systems [27], [26]. However, there is a paucity of research on them in the context of OTFS and OTSM CSI estimation, thus inspiring us to develop attractive techniques for them. While OTFS and OTSM systems typically do not face significant issues related to finite-resolution ADCs, this model specifically addresses situations where ADC resolution constraints are pertinent, such as in practical implementations with limited hardware capabilities. This is particularly relevant in scenarios with large antenna arrays and ultra-wideband systems. A comparison of the existing literature is presented in Table-I. The CSI estimation and quality of detection is crucial for high overall system performance. Hence, a brief review of detection schemes is presented next.

D. Review of OTFS/ OTSM Data Detection Techniques

In [29], the authors proposed a seminal maximal ratio combiner (MRC)-based detection scheme for OTFS systems. However, while it yields a good performance at a manageable complexity, a noteworthy drawback of their scheme is that it

[1] In threshold-based estimation [11], [13] the coefficient estimates depend on an empirically tuned threshold ϵ to obtain the desired number of coefficients. A high threshold may inadvertently miss some of the weak coefficients, while a low threshold may give rise to spurious coefficients. However, in the proposed algorithm the parameter ϵ serves as a convergence threshold to terminate the iterations, ensuring that the algorithm stops once a stable and acceptable solution is determined. In our proposed algorithm, typically, ϵ is chosen to be a small value, approximately in the range of 10^{-4} to 10^{-3}, to ensure convergence of the log-likelihood to a local/global maximum.

TABLE I
CONTRASTING OUR CONTRIBUTIONS TO THE LITERATURE

	[7]**	[14]	[18]	[20]	[22]	[24]	[25]	[26]*	[27]*	Proposed
MIMO System		✓	✓	✓			✓		✓	✓
DD-domain sparsity		✓	✓	✓	✓	✓	✓			✓
Fractional Doppler	✓	✓	✓	✓	✓		✓			✓
Pilots considered	TD	TD	DD	TF	DD	TF	DD	FD	FD	TD
Variational Bayesian learning								✓		✓
Quantized received signal								✓	✓	✓
Finite resolution ADCs								✓	✓	✓
Improved MMSE-based detection rule										✓
DS-domain input-output model	✓									✓

Note: * considers OFDM scenario, ** consider OTSM scenario

requires a high pilot overhead for achieving the optimal SER. Authors of [30] and [31] propose a potent detection rule based on the maximum likelihood (ML) principles for data decoding in OTFS systems, with [30] also presenting an explicit diversity analysis of their scheme. It can also be noted that the work in [31] considered only integer Doppler shifts. As a further innovation Li et al. [32] proposed a hybrid maximum aposteriori (MAP) detection algorithm for reducing the complexity. The authors of [33] introduced a low-complexity MMSE receiver for OTFS systems by exploiting the banded structure and performing LU factorization. An iterative least squares (LS)-based approach is propounded by the authors of [7] for OTSM-based systems. However, for quantized receive signals, the performance of all these schemes is sub-optimal, since they do not incorporate the quantization error, into their design which is prevalent in limited-resolution ADCs-based systems. In this regard a modified MMSE scheme is proposed by [34] for MIMO systems. However, it can not be applied directly to OTFS/OTSM systems. Thus, a bespoke MMSE receiver is conceived in this work. Again our contributions are contrasted to the literature in Table I. The next subsection provides an overview of the novel aspects of this treatise.

E. Contributions of the Paper

1) A novel end-to-end system model is developed for a quantized MIMO OTFS system employing finite-resolution ADCs at the receiver. Hence we go beyond the contributions [17], [20], [25].
2) The OTFS CSI estimation problem can be transformed into a sparse estimation problem. However, this work further converts the sparse estimation problem into a quantized sparse estimation problem using a uniform quantizer Q().
3) Several algorithms have been proposed in the literature to address quantized compressive sensing problems [26], [27], [35], [36], [37]. Within this context, the sparse formulation can be solved using an EM-based Sparse Bayesian Learning (SBL) framework, where the channel sparsity is controlled by hyperparameters assumed to follow a Gaussian distribution [14], [17], [18]. Moreover, the Bayesian model learns the entire aposteriori distribution of the unknown variable from quantized data rather than a single-point estimate. However, to address the quantized SBL problem, a novel Variational Bayesian Learning (VBL) technique is proposed, assuming that the hyperparameters follow a gamma distribution. The goal is to minimise the Kullback-Leibler (KL) divergence between the desired aposteriori density and the constrained family of distributions, ensuring tractability and convergence guarantees to a local optimum. Thus, we leverage the SBL framework and employ a VB approach to address these challenges specifically in context of OTFS systems.
4) Following this, a comprehensive system model is derived for a MIMO OTSM-modulated system utilising finite-resolution ADCs. Similar to the OTFS system, the DS-domain channel is also sparse, and thus, the VBL-based scheme developed above can also be effortlessly applied for CSI acquisition in this system as well.
5) In order to overcome the detrimental effects of quantization errors, a modified MMSE-based data detector is proposed, which incorporates the effect of quantization errors, thus leading to improved performance over the conventional MMSE detector and other linear detectors.

F. Organization of the Work

The paper's structure is as follows: Section-II outlines the MIMO OTFS system model for quantized received signals using finite resolution ADCs, and Section-III formulates a sparse estimation problem. A novel VBL algorithm is proposed for solving the sparse estimation problem in Section-IV. Subsequently the system model of a MIMO OTSM modulated system is derived in Section-V, once again using finite-resolution ADCs, followed by the formulation of the corresponding sparse estimation problem in Section-VI. The modified MMSE detection rule, that incorporates the effect of quantization errors is presented in Section-VII. The simulation results of Section-VIII confirm the efficiency of our approach and its performance is also compared to that of other existing schemes. Section-IX concludes the paper while the Appendix provides the proof of an ancillary result used in the paper.

G. Notation

Lowercase and uppercase letters in boldface represent column vectors and matrices, respectively. The elements of indices of vectors and matrices are indexed beginning with

0. The vector equivalent $\mathbf{a}$ of the matrix $\mathbf{A}$ is obtained via the operation $\text{vec}(\mathbf{A})$ and it is formed by stacking the columns of A to form a single-column vector. Similarly, $\text{vec}_{MN}^{-1}(\mathbf{a})$ denotes the corresponding inverse vectorization operation to obtain the original matrix. A well-known property of the vec $(\cdot)$ operator, vec $(\mathbf{ABC}) = \left(\mathbf{C}^T \otimes \mathbf{A}\right)$ vec $(\mathbf{B})$, for compatible matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}$, is used in the paper, where $\otimes$ denotes the matrix Kronecker product. The quantity $\mathbf{F}_K$ is the normalized discrete Fourier transform (DFT) matrix of size $K \times K$, with its i,jth element given by $\mathbf{F}_K(i,j) = \frac{1}{\sqrt{K}} e^{-j2\pi \frac{ij}{K}}$.

II. MIMO-OTFS System Model With Quantization

In OTFS-modulated systems, the symbols occupy a two dimensional DD-domain grid of size $M \times N$, where M and N represent the number of information symbols placed along the delay and Doppler axes. This results in a delay resolution of $\Delta\tau = 1/M\Delta f$ and a Doppler resolution of $\Delta\nu = 1/NT$, where T and Δf represent the time and frequency axes resolution, respectively, so that $T\Delta f = 1$. Let the number of transmitter antennas (TAs) be N_t and receiver antennas (RAs) be N_r. Let furthermore the matrix $\mathbf{X}_{\text{DD},t} \in \mathbb{C}^{M\times N}$ represent the symbol for the t_{th} TA in the DD-domain, each of which is chosen from an appropriate constellation with symbol power σ_d^2. At the transmitter, the DD-domain symbol matrix $\mathbf{X}_{\text{DD},t}$ is mapped to the TF-domain matrix $\mathbf{X}_{\text{TF},t}$ as $\mathbf{X}_{\text{TF},t} = \mathbf{F}_M \mathbf{X}_{\text{DD},t} \mathbf{F}_N^H$. The equivalent time-domain transmit signal, which is further sampled at the sampling rate of $\frac{M}{T}$, is obtained using the Heisenberg transform. Hence, the transmit signal matrix $\mathbf{S}_t \in \mathbb{C}^{M\times N}$ can be formulated as

$$\mathbf{S}_t = \mathbf{P}_{tx}\mathbf{F}_M^H \mathbf{X}_{\text{TF},t} = \mathbf{P}_{tx}\mathbf{X}_{\text{DD},t}\mathbf{F}_N^H. \tag{1}$$

This work uses CP-aided approach [38], which is also valid in fractional Doppler scenario [39], as extended by several recent works [14], [18]. $\mathbf{P}_{tx} \in \mathbb{C}^{M\times M}$ is the matrix of transmit pulse samples obtained from sampling the pulse $\xi_{tx}(t)$, and defined as $\mathbf{P}_{tx} = \text{diag}\left\{\xi_{tx}\left(\frac{pT}{M}\right)\right\}_{p=0}^{M-1}$. The transmit symbol vector $\mathbf{s}_t \in \mathbb{C}^{MN\times 1}$ is obtained by vectorizing $\mathbf{S}_t$ from (1) modelled as

$$\mathbf{s}_t = \text{vec}(\mathbf{S}_t) = \left(\mathbf{F}_N^H \otimes \mathbf{P}_{tx}\right)\mathbf{x}_{\text{DD},t}, \tag{2}$$

where $\mathbf{x}_{\text{DD},t} = \text{vec}\left(\mathbf{X}_{\text{DD},t}\right) \in \mathbb{C}^{MN\times 1}$. Following this, the time-domain signal $\mathbf{s}$ is subjected to cyclic prefix (CP) insertion. Fig 1 shows a schematic representation of the different steps involved in the modulation process described above. Subsequently, the modulated signal is transmitted through an underspread wireless channel that obeys $\max(\mathrm{l}_i) \ll M$, $\max(\mathrm{k}_i) \ll N$. The wideband doubly selective DD-domain channel $h_{r,t}(\tau,\nu)$ between the r-th RA and the t-th TA is formulated as

$$h_{r,t}(\tau,\nu) = \sum_{i=1}^{\rho} h_{i,r,t}\delta(\tau-\tau_i)\delta(\nu-\nu_i), \tag{3}$$

where $1 \le r \le N_r$ and $1 \le t \le N_t$. Here, $h_{i,r,t}$ represents the complex path gain associated with the i-th multipath component for the t-th TA and r-th RA.

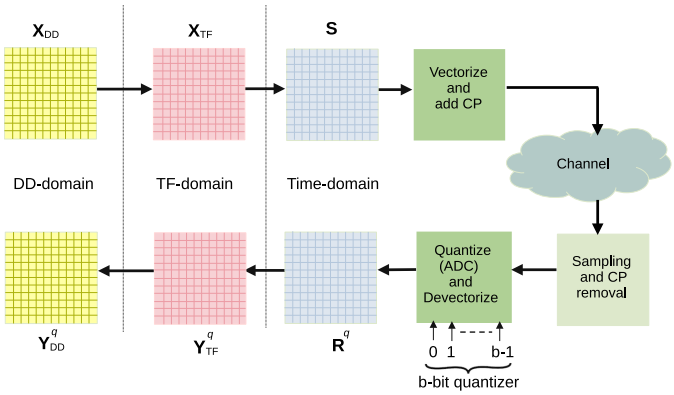

Fig. 1. Block diagram for OTFS modulated systems.

The parameters τ_i and ν_i denote the delay and Doppler shift of the i-th reflector, respectively. Note that in the channel model above, the antenna-variant delay can be ignored since both the number of antennas and bandwidth are small. Furthermore, the time-variant delay can also be ignored since both the bandwidth and duration are small. The quantity $\delta(\cdot)$ denotes the Dirac-delta function. As described in [13], [40], and [41] the delay resolution of $\Delta\tau = \frac{1}{M\Delta f}$ for a typical wide-band system is generally very fine. For instance, consider a system with parameters $M = 64$ and $\Delta f = 30$ kHz; the resultant delay resolution is $\Delta\tau = \frac{1}{M\Delta f} = 0.5208\mu$sec, which is reasonably small. Consequently, the delay shift associated with the i-th multipath component is assumed to be an integer multiple of this delay resolution, denoted as $\tau_i = l_i\Delta\tau$. The Doppler-shift ν_i for the i-th multi-path component is defined as $\nu_i = k_i\Delta\nu$, where k_i is the fractional index, which is given as $k_i = \text{round}(k_i) + \mathcal{K}_i$, for $|\mathcal{K}_i| < 0.5$.

At the receiver, sampling is conducted at the Nyquist rate to prevent problems associated with insufficient sampling. The signal is sampled with a duration of T/M. After sampling, the first L samples, which correspond to the transmitted cyclic prefix (CP), are discarded. The remaining samples are then quantized using finite-resolution ADCs. The received signal corresponding to the rth RA can be modeled as $\mathbf{r}_r \in \mathbb{C}^{M\times N}$ for $1 \le r \le N_r$ and it is given by

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

where the noise vector $\mathbf{w}_r \in \mathbb{C}^{MN\times 1}$ contains the additive white Gaussian noise samples, with covariance $\mathbb{E}(\mathbf{w}_r\mathbf{w}_r^H) = \sigma_w^2\mathbf{I}_{MN}$. $\mathbf{H}_{r,t} \in \mathbb{C}^{MN\times MN}$ is the channel matrix corresponding to the r-th RA and t-th TA, defined as

$$\mathbf{H}_{\mathbf{r},\mathbf{t}} = \sum_{i=0}^{\rho} h_{i,r,t}\left(\mathbf{\Pi}\right)^{l_i}\left(\mathbf{\Delta}_i\right), \tag{5}$$

where $\mathbf{\Pi}$ is the permutation matrix obtained by right circular shifting the identity matrix of order $MN \times MN$ and the matrix $\mathbf{\Delta} \in \mathbb{C}^{MN\times MN}$ is defined as $\Delta_i = \text{diag}\left\{1, \omega_i, \cdots, \omega_i^{MN-1}\right\}$, where $\omega_i = e^{j2\pi\frac{k_i}{MN}}$. It is important to note that the diagonal matrix $\mathbf{\Delta}$ differs from the scalar parameters $\Delta\tau$ and $\Delta\nu$, which represent delay and Doppler

resolution, respectively. At the receiver, a uniform quantizer is employed for quantizing the received signal $\mathbf{r}_r$. Notably, the DD-domain sparsity of the channel remains unaffected by the use of finite-resolution ADCs, as these are applied solely to the received samples yielding

$$\mathbf{r}_r^q = \mathcal{Q}\left(\sum_{t=1}^{N_t} \mathbf{H}_{r,t}\mathbf{s}_t + \mathbf{w}_r\right), \tag{6}$$

where the real and imaginary components are quantized separately due to the complex nature of the incoming signal vector $\mathbf{r}_r$.

Thus using the operators $\mathcal{R}\{\cdot\}$, $\mathcal{I}\{\cdot\}$ to obtain the real and imaginary parts, the k-th element in $\mathbf{r}_r$ can be expressed as

$$\mathbf{r}_r^q(k) = \mathcal{Q}\left(\mathbf{r}_r(k)\right) = \mathcal{Q}\Big(\mathcal{R}\{\mathbf{r}_r(k)\}\Big) + j\mathcal{Q}\Big(\mathcal{I}\{\mathbf{r}_r(k)\}\Big), \tag{7}$$

where $\mathcal{Q}(\cdot)$ is a b-bit quantizer function. The b-bit quantizer $\mathcal{Q}\left(\mathcal{C}\left\{\mathbf{r}_r(k)\right\}\right)$ for a scalar input $\mathcal{C}\left\{\mathbf{r}_r(k)\right\} \in \mathbb{R}$, where $\mathcal{C} \in \{\mathcal{R}, \mathcal{I}\}$ with $B = 2^b$ quantization levels is defined as follows,

$$\mathcal{Q}\Big(\mathcal{C}\left\{\mathbf{r}_k\right\}\Big) = \begin{cases} c_1, & \mathbf{r}_r(k) \in [u_0, u_1]; \\ c_2, & \mathbf{r}_r(k) \in (u_1, u_2]; \\ \cdots, & \cdots; \\ c_B, & \mathbf{r}_r(k) \in (u_{B-1}, u_B]. \end{cases} \tag{8}$$

The quantization thresholds are denoted by $u_0 < u_1 < \ldots < u_B$, and $c_1, \ldots, c_B$ are the quantizer output values. Specifically, we consider a simple mid-point uniform quantizer similar to [42], with the quantization level u_i, where $i = 0, 1, \cdots, B$, defined as

$$u_i = \left(\frac{-B}{2} + i\right)\Delta_q. \tag{9}$$

The quantization threshold c_i is set as

$$c_i = \frac{(u_{i-1} + u_i)}{2}, \tag{10}$$

where Δ_q is the quantization step size given by

$$\Delta_q = \frac{\max\left(\left|\mathcal{C}\left\{\mathbf{r}_r(k)\right\}\right|\right)}{2^b}, \tag{11}$$

where $k = 1, \cdots, MN$. The above equation in (6) can be expressed as

$$\mathbf{r}_r^q = \mathcal{Q}\left(\mathbf{H}_r\mathbf{s} + \mathbf{w}_r\right), \tag{12}$$

where $\mathbf{H}_r = [\mathbf{H}_{r,1}, \cdots, \mathbf{H}_{r,N_t}]$ and $\mathbf{s} = \left[\mathbf{s}_1^T, \mathbf{s}_2^T, \cdots, \mathbf{s}_{N_t}^T\right]^T$. Thus, stacking all the received output vectors $\mathbf{r}_r$ for $1 \le r \le N_r$ as $\left[(\mathbf{r}_1^q)^T, (\mathbf{r}_2^q)^T, \cdots, (\mathbf{r}_{N_r}^q)^T\right]^T \in \mathbb{C}^{MNN_r \times 1}$ one obtains $\mathbf{r}^q$, which can be expressed as

$$\mathbf{r^q} = \mathcal{Q}(\mathbf{H}\mathbf{s} + \mathbf{w}), \tag{13}$$

where $\mathbf{H} \in \mathbb{C}^{MNN_r \times MNN_t}$ is the time-domain MIMO-OTFS channel matrix. The quantization error is $\mathbf{e}_q = \mathbf{r}^q - \mathbf{r}$ and the i-th element of the error is given by $e_{q,i}$. Since, a uniform quantizer with step size Δ_q is employed, the error is uniformly distributed over the interval $(-\frac{\Delta_q}{2}, \frac{\Delta_q}{2})$, and it is expressed as $e_{q,i} \sim \mathcal{U}\left(-\frac{\Delta_q}{2}, \frac{\Delta_q}{2}\right)$ for $i = 1, 2, \cdots, MN_pN_r$. Therefore, the joint distribution of the quantization error $\mathbf{e}_q$ is given by

$$p_{e_q}(e_q) = \prod_{i=1}^{MN_pN_r} p_{e_i}(e_i) = \prod_{i=1}^{MN_pN_r} \frac{1}{\Delta_q} = \frac{1}{\Delta_q^{MN_pN_r}}.$$

The effective error of the system model is given by $\widetilde{\mathbf{e}} = \mathbf{e}_q + \mathbf{w}$. The total error for the i-th element is given as $\widetilde{e}_i = w_i + e_{q,i}$. The distribution of $\tilde{e}_i$ is given by the convolution of the Gaussian and uniform distributions given by $p_{\tilde{e}_i}(\widetilde{e}_i) = \int_{-\infty}^{\infty} p_{n_i}(z-u)p_{e_{q_i}}(u)\,du$. Upon harnessing the PDFs of w_i and e_{q_i}, one obtains

$$\begin{aligned} p_{\tilde{e}_i}(e_i) &= \frac{1}{\Delta_q}\int_{-\Delta_q/2}^{\Delta_q/2} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\widetilde{e}_i - u)^2}{2\sigma^2}\right) du \\ &= \frac{1}{2\Delta_q}\left(\text{erf}\left(\frac{\Delta_q - 2\widetilde{e}_i}{2\sqrt{2\sigma^2}}\right) + \text{erf}\left(\frac{\Delta_q + 2\widetilde{e}_i}{2\sqrt{2\sigma^2}}\right)\right). \end{aligned}$$

Therefore the joint distribution can be formulated as $p_{\tilde{e}}(\tilde{\mathbf{e}}) = [p_{\tilde{e}_i}(\tilde{e}_i)]^{MN_pN_r}$. The TF-domain demodulated symbol matrix $\mathbf{Y}_{\text{TF},r}^q \in \mathbb{C}^{M\times N}$ is formulated as $\mathbf{Y}_{\text{TF},r}^q = \mathbf{F}_M\mathbf{P}_{rx}\mathbf{R}_r^q$, where the quantized received signal matrix $\mathbf{R}_r^q \in \mathbb{C}^{M\times N}$ is obtained from the quantized received signal $\mathbf{r}_r^q$ as $\mathbf{R}_r^q = \text{vec}_{MN}^{-1}(\mathbf{r}_r^q)$. Upon vectorizing $\mathbf{Y}_{\text{TF},r}^q$, the demodulated symbol vector $\mathbf{y}_{\text{TF},r}^q \in \mathbb{C}^{MN\times 1}$ can be expressed as

$$\mathbf{y}_{\text{TF,r}}^q = \text{vec}\Big(\mathbf{F}_M\mathbf{P}_{rx}\mathbf{R}_r^q\Big) = \Big(\mathbf{I}_N \otimes \mathbf{F}_M\mathbf{P}_{rx}\Big)\mathbf{r}_r^q. \tag{14}$$

Similar to the transmit pulse shaping matrix, the received pulse shaping matrix $\mathbf{P}_{rx} \in \mathbb{C}^{M\times M}$ is obtained by sampling $\xi_{rx}(t)$ and it is given as $\mathbf{P}_{rx} = \text{diag}\left\{\xi_{rx}\left(\frac{pT}{M}\right)\right\}_{p=0}^{M-1}$. The demodulated DD-domain symbol matrix $\mathbf{Y}_{\text{DD},r}^q \in \mathbb{C}^{M\times N}$ is related to $\mathbf{Y}_{\text{TF},r}^q$ as $\mathbf{Y}_{\text{DD},r}^q = \mathbf{F}_M^H\mathbf{Y}_{\text{TF},r}^q\mathbf{F}_N$. Upon vectorizing $\mathbf{Y}_{\text{DD},r}^q$, one obtains

$$\begin{aligned} \mathbf{y}_{\text{DD},r}^q &= \text{vec}\Big(\mathbf{F}_M^H\mathbf{Y}_{\text{TF},r}^q\mathbf{F}_N\Big) = \text{vec}\Big(\mathbf{P}_{rx}\mathbf{R}_r^q\mathbf{F}_N\Big) \\ &= \Big(\mathbf{F}_N \otimes \mathbf{P}_{rx}\Big)\mathbf{r}_r^q. \end{aligned} \tag{15}$$

Substituting $\mathbf{r}_r^q$ from (6) and $\mathbf{s}_t$ from (2) into (15) yields the relationship

$$\begin{aligned} &\mathbf{y}_{\text{DD},r}^q \\ &= \Big(\mathbf{F}_N \otimes \mathbf{P}_{rx}\Big)\mathcal{Q}\left(\sum_{t=1}^{N_t} \mathbf{H}_{r,t}\mathbf{s}_t + \mathbf{w}_r\right) \\ &= \Big(\mathbf{F}_N \otimes \mathbf{P}_{rx}\Big)\mathcal{Q}\left(\sum_{t=1}^{N_t} \mathbf{H}_{r,t}\left(\mathbf{F}_N^H \otimes \mathbf{P}_{tx}\right)\mathbf{x}_{\text{DD},t} + \mathbf{w}_r\right). \end{aligned} \tag{16}$$

Stacking the received outputs $\mathbf{y}_{\text{DD},r}^q$ for $r = 1$ to N_r as $[(\mathbf{y}_{\text{DD},1}^q)^T, \cdots, (\mathbf{y}_{\text{DD},N_r}^q)^T]^T$ one obtains $\mathbf{y}_{\text{DD}}^q \in \mathbb{C}^{MNN_r\times 1}$ given as

$$\mathbf{y}_{\text{DD}}^q = (\mathbf{I}_{N_r} \otimes \mathbf{F}_N \otimes \mathbf{P}_{rx})\mathcal{Q}\left(\mathbf{H}(\mathbf{I}_{N_t} \otimes \mathbf{F}_N^H \otimes \mathbf{P}_{tx})\mathbf{x}_{\text{DD}} + \mathbf{w}\right). \tag{17}$$

The above expression can be further simplified as

$$\mathbf{y}_{\text{DD}}^q = (\mathbf{I}_{N_r} \otimes \mathbf{F}_N \otimes \mathbf{I}_M)\mathcal{Q}\left(\mathbf{H}(\mathbf{I}_{N_t} \otimes \mathbf{F}_N^H \otimes \mathbf{I}_M)\mathbf{x}_{\text{DD}} + \mathbf{w}\right), \tag{18}$$

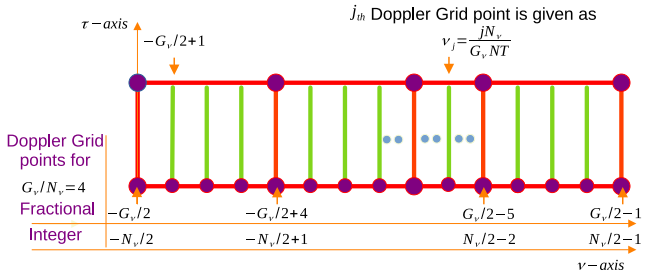

Fig. 2. Fractional and integer Doppler grid structure.

The simplification from (17) to (18) is attributed to the use of rectangular pulse shaping, where $\mathbf{P}_{tx} = \mathbf{P}_{rx} = \mathbf{I}_M$ and the function $\mathcal{Q}(.)$ is given by (8). Thus, the input DD-domain symbol vector $\mathbf{x}_{\text{DD}}$ is mapped to the output symbol vector $\widetilde{\mathbf{y}}^q_{\text{DD}} \in \mathbb{C}^{MNN_r \times 1}$ as

$$\widetilde{\mathbf{y}}^q_{\text{DD}} = \mathcal{Q}\Big(\widetilde{\mathbf{H}}\mathbf{x}_{\text{DD}} + \mathbf{w}\Big), \tag{19}$$

where $\widetilde{\mathbf{H}} = \mathbf{H}\left(\mathbf{I}_{N_t} \otimes \mathbf{F}_N^H \otimes \mathbf{I}_M\right)$ and $\tilde{\mathbf{y}}^q_{\text{DD}}$ is mapped to $\mathbf{y}^q_{\text{DD}}$ of (18) using the relationship $\widetilde{\mathbf{y}}^q_{\text{DD}} = \left(\mathbf{I}_{N_r} \otimes \mathbf{F}_N^H \otimes \mathbf{I}_M\right)\mathbf{y}^q_{\text{DD}}$. Data detection can now be performed using a conventional MMSE estimator as

$$\widehat{\mathbf{x}}^{\text{MMSE}}_{\text{DD}} = \left(\widetilde{\mathbf{H}}^H\widetilde{\mathbf{H}} + \frac{\sigma_d^2}{\sigma_w^2}\mathbf{I}_{MN}\right)^{-1}\widetilde{\mathbf{H}}^H\tilde{\mathbf{y}}^q_{\text{DD}}. \tag{20}$$

Note that the above equation in (19) represents a non-linear relationship between $\tilde{\mathbf{y}}^q_{\text{DD}}$ and $\mathbf{x}_{\text{DD}}$. Thus, a modified MMSE detector that incorporates the effect of quantization errors is introduced in Section-VII. The next section delves into our sparse estimation problem formulation.

III. Sparse Channel Model for MIMO OTFS Systems

Consider an under-spread wireless channel that obeys $\max(l_i) \leq M_\tau \leq\leq M$, $\max(k_i) \leq N_\nu \leq\leq N$, where the maximum delay and Doppler spreads are given by M_τ, N_ν, respectively. For introducing fractional Doppler, consider a grid of size $M_\tau \times G_\nu$, such that $G_\nu >> N_\nu$, i.e., each grid interval corresponding to integer Doppler shift is divided into multiple intervals. The jth Doppler-grid point, $-G_\nu/2 \leq j \leq G_\nu/2-1$, corresponds to a Doppler-shift of $\nu_j = \frac{jN_\nu}{G_\nu NT}$ Hz as shown in Fig. 2. Thus, the DD-domain channel $h(\tau,\nu)$ can be equivalently expressed as

$$h(\tau,\nu) = \sum_{i=0}^{M_\tau-1}\sum_{j=-G_\nu/2}^{G_\nu/2-1} h_{i,j}\delta(\tau-\tau_i)\delta(\nu-\nu_j),$$

where the (i,j)-th point on the delay-Doppler grid corresponds to the i-th delay tap τ_i and j-th Doppler tap ν_j, with $h_{i,j}$ being the associated path gain. It should be noted that only a small number of coefficients, $h_{i,j}$, viz $\rho\ (<< M_\tau G_\nu)$, of the total $M_\tau G_\nu$ are non-zero due to the existence of a few dominant reflectors in the wireless channel.

It is noteworthy that the parameter ν_j during channel estimation is defined as $\nu_j = \frac{jN_\nu}{G_\nu NT}$ Hz, which depends on the number of Doppler bins N of the OTFS frame. Hence the Doppler resolution of an integer Doppler case, i.e., for $N_\nu = G_\nu$, is still $\Delta\nu = \frac{1}{NT}$. As the precise values of the delay and Doppler parameters are unknown for each r,t, they are assumed to exist in the DD-grid of size $M_\tau \times G_\nu$ for an underspread channel and the equivalent channel matrix $\mathbf{H}_{r,t} \in \mathbb{C}^{MN_p \times MN_p}$ corresponding to the r-th RA, t-th TA can be expressed in matrix form as

$$\mathbf{H}_{r,t} = \sum_{i=0}^{M_\tau-1}\sum_{j=-G_\nu/2}^{G_\nu/2-1} h_{i,j,r,t}\left(\mathbf{\Pi}\right)^i\left(\mathbf{\Delta}_{i,j}\right), \tag{21}$$

where $\mathbf{\Delta}_{i,j} \in \mathbb{C}^{MN_p \times MN_p}$ is given by

$$\begin{cases} \text{diag}\left\{1, \bar{\omega}_j, \cdots, \bar{\omega}_j^{MN_p-i-1}, \bar{\omega}_j^{-i}, \cdots, \bar{\omega}_j^{-1}\right\}, & \text{if } i \neq 0, \\ \text{diag}\left\{1, \bar{\omega}_j, \cdots, \bar{\omega}_j^{MN_p-1}\right\}, & \text{for } i = 0, \end{cases} \tag{22}$$

where $\bar{\omega}_k = e^{j2\pi\frac{kN_\nu}{G_\nu MN}}$ and the total number of pilot symbols inserted in the time domain is MN_p. Thus, the quantized received signal $\mathbf{r}^q_r \in \mathbb{C}^{MN_p \times 1}$ corresponding to the rth RA in (6) can be expressed as

$$\mathbf{r}^q_r = \mathcal{Q}\left(\sum_{t=1}^{N_t}\sum_{i=0}^{M_\tau-1}\sum_{j=-G_\nu/2}^{G_\nu/2-1} h_{i,j,r,t}\left(\bar{\mathbf{\Pi}}\right)^i\left(\mathbf{\Delta}_{i,j}\right)\mathbf{s}_t + \mathbf{w}_r\right), \tag{23}$$

where $\mathbf{s}_t \in \mathbb{C}^{MN_p \times 1}$ is the time-domain transmitted pilot symbol. The above equation (23) can be expressed as

$$\mathbf{r}^q_r = \mathcal{Q}\left(\sum_{t=1}^{N_t}\sum_{i=0}^{M_\tau-1}\sum_{j=-G_\nu/2}^{G_\nu/2-1} h_{i,j,r,t}\boldsymbol{\psi}_{i,j,t} + \mathbf{w}_r\right), \tag{24}$$

where $\boldsymbol{\psi}_{i,j,t} = \left(\mathbf{\Pi}\right)^i\left(\mathbf{\Delta}_{i,j}\right)\mathbf{s}_t \in \mathbb{C}^{MN_p \times 1}$, which may be further simplified as

$$\mathbf{r}^q_r = \mathcal{Q}\left(\sum_{t=1}^{N_t}\mathbf{\Psi}_t\mathbf{h}_{r,t} + \mathbf{w}_r\right), \tag{25}$$

where the matrix $\mathbf{\Psi}_t \in \mathbb{C}^{MN_p \times M_\tau G_\nu}$ and vector $\mathbf{h}_{r,t}$ are defined as

$$\mathbf{\Psi}_t = \left[\bar{\psi}_{0,-G_\nu/2,t} \cdots \bar{\psi}_{0,G_\nu/2-1,t} \cdots \bar{\psi}_{M_\tau-1,G_\nu/2-1,t}\right]$$
$$\mathbf{h}_{r,t} = \left[h_{0,-G_\nu/2,r,t} \cdots h_{0,G_\nu/2-1,r,t} \cdots h_{M_\tau-1,G_\nu/2-1,r,t}\right]^T.$$

Furthermore, the above expression in (25) can be simplified as

$$\mathbf{r}^q_r = \mathcal{Q}\left(\widetilde{\mathbf{\Psi}}\mathbf{h}_r + \mathbf{w}_r\right), \tag{26}$$

where $\widetilde{\mathbf{\Psi}} = \left[\mathbf{\Psi}_1, \mathbf{\Psi}_2, \cdots, \mathbf{\Psi}_{N_t}\right] \in \mathbb{C}^{MN_p \times M_\tau G_\nu N_t}$ and $\mathbf{h}_r = \left[\mathbf{h}^T_{r,1}, \cdots, \mathbf{h}^T_{r,N_t}\right]^T \in \mathbb{C}^{M_\tau G_\nu N_t \times 1}$. By stacking the received vectors across all RAs, the received vector $\mathbf{r} \in \mathbb{C}^{MN_pN_r \times 1}$ is defined as $\mathbf{r}^q = [(\mathbf{r}^q_1)^T, (\mathbf{r}^q_2)^T, \cdots, (\mathbf{r}^q_{N_r})^T]^T$.

$$\mathbf{r}^q = \mathcal{Q}\Big(\mathbf{\Psi}\mathbf{h} + \mathbf{w}\Big), \tag{27}$$

where $\mathbf{h} = \left[(\mathbf{h}_1)^T, \cdots, (\mathbf{h}_{N_r})^T\right]^T \in \mathbb{C}^{M_\tau G_\nu N_t N_r \times 1}$. The sparse estimation problem formulated in (27) can now be solved using the powerful variational Bayesian learning (VBL) technique as shown next.

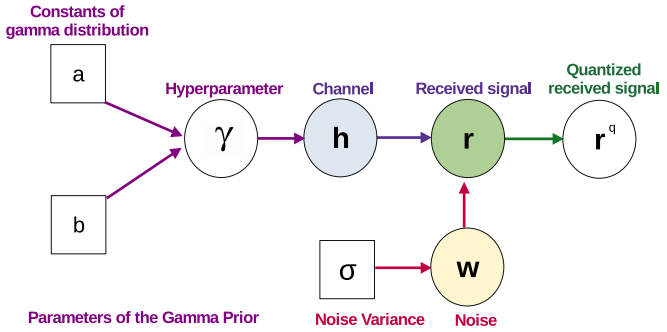

Fig. 3. Model for variational Bayesian learning.

IV. Variational Bayesian Learning Based CSI Acquisition

Consider the sparse estimation problem formulated in (27) for a quantized OTFS-based system. While the non-Bayesian techniques employ point-based estimation techniques for determining $\mathbf{h}$ from $\mathbf{r}_p^q$ that are not robust, the Bayesian approaches exploit the sparsity of $\mathbf{h}$ using a sparsity promoting prior. Consider the VBL model of Fig. 3, where a zero mean conditional complex Gaussian prior distribution is assumed for $\mathbf{h}$, defined as

$$p(\mathbf{h}|\boldsymbol{\gamma}) = \mathcal{CN}(\mathbf{0}, \boldsymbol{\Gamma}^{-1}) = \prod_{i=0}^{M_\tau G_\nu N_t N_r - 1} \frac{1}{(\pi\gamma_i)} \exp\left(-\frac{|\mathbf{h}(i)|^2}{\gamma_i}\right), \tag{28}$$

with $\boldsymbol{\Gamma} = \text{diag}(\boldsymbol{\gamma})$. Thus, a Gamma prior for the hyperparameter γ_i, which is a conjugate prior for the conditional pdf $p(h_i|\gamma_i)$ is defined as

$$\begin{aligned} p(\gamma_i) &= \text{Gamma}\ (\gamma_i; a, b) \\ &= \frac{b^a}{\Gamma(a)} \gamma_i^{a-1} \exp\{-b\gamma_i\},\ \gamma_i > 0,\ \forall i, \end{aligned} \tag{29}$$

where a, b are constants and $\Gamma(\cdot)$ denotes the Gamma function defined as $\Gamma(a) = \int_0^\infty t^{a-1}e^{-t}dt$. Therefore the joint distribution for the hyperparameter vector $p(\gamma)$ is given by

$$p(\boldsymbol{\gamma}) = \prod_{i=0}^{M_\tau G_\nu N_t N_r - 1} p\left(\gamma_i\right). \tag{30}$$

The internal variables for the VBL model are given by the set $\mathcal{A}$, i.e., $\mathcal{A} = \{\boldsymbol{\gamma}, \mathbf{h}, \mathbf{r}_p\}$ and the observed variable is $\mathbf{r}_p^q$ Furthermore, the quantity $\mathbf{h}$ can be estimated from $\mathbf{r}_p^q$ by computing the posterior marginal distribution $p\left(\mathbf{h} \mid \mathbf{r}_p^q\right)$, which requires the evaluation of the distribution of $\mathbf{r}_p^q$. However, the high dimensional integration of $p\left(\mathbf{r}_p^q\right) = \int p\left(\mathbf{r}_p^q, \mathcal{A}\right) d\mathcal{A}$ renders this problem mathematically intractable. In order to overcome this challenge, one can employ the VBL approach that identifies the variational distribution $q(\mathcal{A})$ that closely models the true posterior distribution $p\left(\mathcal{A} \mid \mathbf{r}_p^q\right)$ [43] as shown below. Consider the following identity

$$\ln p(\mathbf{r}_p^q) = \mathcal{L}(q) + \text{KL}(q||p), \tag{31}$$

where $\mathcal{L}(q)$ is the lower bound determined as

$$\mathcal{L}(q) = \int q(\mathcal{A}) \ln \frac{p(\mathbf{r}_p^q, \mathcal{A})}{q(\mathcal{A})}\, d\mathcal{A}, \tag{32}$$

and $\text{KL}(q\|p)$ is the Kullback-Leibler divergence between $p\left(\mathcal{A} \mid \mathbf{r}_p^q\right)$ and $q(\mathcal{A})$ defined as

$$\text{KL}(q||p) = -\int q(\mathcal{A}) \ln \frac{p(\mathcal{A}|\boldsymbol{r}_p^q)}{q(\mathcal{A})}\, d\mathcal{A}, \tag{33}$$

so that $\text{KL}(q\|p) \geq 0$. Thus, minimizing the KL divergence is equivalent to approximately choosing the distribution to represent $q(\mathcal{A})$ that maximizes $\mathcal{L}(q)$. To achieve an efficient and tractable approximation, consider the mean-field approximation, where the distinct variables in $q(\mathcal{A})$ are independent, yielding

$$q(\mathcal{A}) = \prod_i q\left(\mathcal{A}_i\right) = q(\boldsymbol{\gamma})q(\mathbf{h})q(\mathbf{r}_p). \tag{34}$$

Substituting (34) into (31), and maximizing $\mathcal{L}(q)$ as shown in Appendix A, one obtains

$$q(\mathcal{A}_i) \propto \exp\left\{\mathbb{E}_{\sim q(\mathcal{A}_i)}\left(\ln p(\mathbf{r}_p^q, \mathcal{A})\right)\right\}, \tag{35}$$

where $\mathbb{E}_{\sim q(\mathcal{A}_i)}$ is the expectation with respect to all factors, except for $q\left(\mathcal{A}_i\right)$. Thus, it can be observed from the above equation that $q\left(\mathcal{A}_i\right)$ is dependent on the expectation with respect to the other factors. The posterior density update of $q(\mathbf{r}_p)$, $q(\mathbf{h})$ and $q(\boldsymbol{\gamma})$ is obtained using [44], as described below.

Computation of $q\left(\boldsymbol{r}_p\right)$: The quantity $q\left(\boldsymbol{r}_p\right)$ can be obtained from (35) as

$$\begin{aligned} q(\mathbf{r}_p) &\propto \exp\left\{\ln p(\mathbf{r}_p^q|\mathbf{r}_p) + \mathbb{E}_{q(\mathbf{h})}\left(\ln p(\mathbf{r}_p|\mathbf{h})\right)\right\} \\ &\propto p(\mathbf{r}_p^q|\mathbf{r}_p) \exp\left\{\mathbb{E}_{q(\boldsymbol{h})}\left(\ln p(\mathbf{r}_p|\mathbf{h})\right)\right\}, \end{aligned} \tag{36}$$

where $p(\mathbf{r}_p|\mathbf{h})$ is distributed as

$$p(\mathbf{r}_p|\mathbf{h}) \sim \mathcal{CN}(\boldsymbol{\Psi}\mathbf{h}, \sigma^2\mathbf{I}_{MN}). \tag{37}$$

Given the observation data $\mathbf{r}_p$, the quantity $\mathbf{r}_p^q$ is distributed as $p(\mathbf{r}_p^q|\mathbf{r}_p) = \mathbb{1}\Big(\mathbf{r}_p \in (\boldsymbol{l}, \boldsymbol{u}]\Big)$, where l and u are the minimum and maximum values corresponding to the quantizer output $\mathbf{r}_p^q$, respectively, and $\mathbb{1}\left(\mathbf{r}_p \in (\boldsymbol{l}, \boldsymbol{u}]\right) := \mathbb{1}\left(\mathcal{R}\{\mathbf{r}_p\} \in (\mathcal{R}\{\boldsymbol{l}\}, \mathcal{R}\{\boldsymbol{u}\}]\right) \cdot \mathbb{1}\left(\mathcal{I}\{\mathbf{r}_p\} \in (\mathcal{I}\{\boldsymbol{l}\}, \mathcal{I}\{\boldsymbol{u}\}]\right)$. Substituting $p(\mathbf{r}_p|\mathbf{h})$ from (37) and $p(\mathbf{r}_p^q|\mathbf{r}_p)$ into (36), one obtains the expression of $q(\mathbf{r}_p)$ as

$$\begin{aligned} q(\mathbf{r}_p) &\propto \mathbb{1}\Big(\mathbf{r}_p \in (\boldsymbol{l}, \boldsymbol{u}]\Big) \exp\left\{\mathbb{E}_{q(\mathbf{h})}\left(-\frac{1}{\sigma^2}\|\mathbf{r}_p - \boldsymbol{\Psi}_p\mathbf{h}\|_2^2\right)\right\} \\ &\propto \mathbb{1}\Big(\mathbf{r}_p \in (\boldsymbol{l}, \boldsymbol{u}]\Big) \exp\left\{-\frac{1}{\sigma^2}\|\mathbf{r}_p - \boldsymbol{\Psi}_p\mathbb{E}_{q(\mathbf{h})}\left(\mathbf{h}\right)\|_2^2\right\}. \end{aligned} \tag{38}$$

The above expression depends on $q(\boldsymbol{h})$, which can be computed as shown below.

Computation $q(\mathbf{h})$: From (35) one obtains the relationship

$$q(\mathbf{h}) \propto \exp\left\{\mathbb{E}_{q(\mathbf{r}_p)}\Big(\ln p(\mathbf{r}_p|\mathbf{h})\Big) + \mathbb{E}_{q(\boldsymbol{\gamma})}\Big(\ln p(\mathbf{h}|\boldsymbol{\gamma})\Big)\right\}. \tag{39}$$

Substituting the expressions for $p(\mathbf{r}_p|\mathbf{h})$ from (37) and $p(\mathbf{h}|\boldsymbol{\gamma})$ from (28) into (39) yields

$$q(\mathbf{h}) \propto \exp\left\{\mathbb{E}_{q(\boldsymbol{r}_p)}\Big(-\frac{1}{\sigma^2}\|\mathbf{r}_p-\boldsymbol{\Psi}_p\mathbf{h}\|_2^2\Big)+\mathbb{E}_{q(\boldsymbol{\gamma})}\Big(-\mathbf{h}^H\boldsymbol{\Gamma}\mathbf{h}\Big)\right\}$$
$$\propto \exp\left\{-(\mathbf{h}-\boldsymbol{\mu})^H\boldsymbol{\Sigma}^{-1}(\mathbf{h}-\boldsymbol{\mu})\right\}. \quad (40)$$

This shows that $q(\mathbf{h})$ corresponds to the Gaussian distribution $\mathcal{CN}(\boldsymbol{\Sigma},\boldsymbol{\mu})$ with $\boldsymbol{\Sigma}$, $\boldsymbol{\mu}$ determined as

$$\boldsymbol{\Sigma} = \left(\frac{1}{\sigma_w^2}\boldsymbol{\Psi}_p^H\boldsymbol{\Psi}_p + \mathbb{E}_{q(\boldsymbol{\gamma})}(\boldsymbol{\Gamma})\right)^{-1},$$
$$\boldsymbol{\mu} = \frac{1}{\sigma_w^2}\boldsymbol{\Sigma}\boldsymbol{\Psi}_p^H\mathbb{E}_{q(\mathbf{r}_p)}(\mathbf{r}_p), \quad (41)$$

where $\mathbb{E}_{q(\mathbf{r}_p)}(\mathbf{r}_p)$ is derived in Appendix B. Thus we have $\mathbb{E}_{q(\mathbf{h})}(\mathbf{h}) = \boldsymbol{\mu}$. The above expression for $\boldsymbol{\Sigma}$ is dependent on $q(\boldsymbol{\gamma})$ which can be computed as shown below.

Computation of $q(\boldsymbol{\gamma})$: From (35), $q(\boldsymbol{\gamma})$ is obtained as

$$q(\boldsymbol{\gamma}) \propto \exp\left\{\mathbb{E}_{q(\mathbf{h})}\Big(\ln p(\mathbf{h}|\boldsymbol{\gamma})\Big) + \ln p(\boldsymbol{\gamma})\right\}. \quad (42)$$

Substituting $p(\mathbf{h}|\boldsymbol{\gamma})$ from (28) and $p(\boldsymbol{\gamma})$ from (29) in the above equation, one obtains

$$q(\boldsymbol{\gamma}) \propto \exp\left\{\mathbb{E}_{q(\mathbf{h})}\Big(-\boldsymbol{h}^H\boldsymbol{\Gamma}\boldsymbol{h}\Big) + \ln|\boldsymbol{\Gamma}| + \sum_{i=1}^{M_\tau G_\nu}\ln p(\gamma_i)\right\}. \quad (43)$$

The above expression can be simplified as the product of independent Gamma distributions, Gamma $\left(\gamma_i;\widetilde{a},\widetilde{b}\right)$ given as

$$q(\boldsymbol{\gamma}) \propto \prod_{i=1}^{M_\tau G_\nu N_t N_r}\exp\left\{-\mathbb{E}_{q(\mathbf{h})}\left(h_i^2\right)\gamma_i\right\}\gamma_i p(\gamma_i)$$
$$\propto \prod_{i=1}^{M_\tau G_\nu N_t N_r}\gamma_i^{(a+1)-1}\exp\left\{-\left(b+\mathbb{E}_{q(\boldsymbol{h})}\left(h_i^2\right)\gamma_i\right\}, \quad (44)$$

where $\widetilde{a} = a+1, \widetilde{b} = b+\mathbb{E}_{q(\mathbf{h})}(h_i^2)$. Thus, we have

$$\mathbb{E}_{q(\boldsymbol{\gamma})}(\gamma_i) = \frac{\tilde{a}}{\tilde{b}} = \frac{a+1}{b+\mathbb{E}_{q(\mathbf{h})}(h_i^2)}, \quad (45)$$

where, we have $a = b = 10^{-6}$ from [45], and upon using (41) $\mathbb{E}_{q(\mathbf{h})}(h_i^2)$ is computed as $\mathbb{E}_{q(\mathbf{h})}(h_i^2) = [\boldsymbol{\mu}\boldsymbol{\mu}^H + \boldsymbol{\Sigma}]_{i,i}$. Finally, $\mathbb{E}_{q(\boldsymbol{\gamma})}(\boldsymbol{\Gamma})$ is obtained as

$$\mathbb{E}_{q(\boldsymbol{\gamma})}(\boldsymbol{\Gamma}) = \text{diag}\left(\left[\mathbb{E}_{q(\boldsymbol{\gamma})}(\gamma_1),\ldots,\mathbb{E}_{q(\boldsymbol{\gamma})}(\gamma_{M_\tau G_\nu})\right]\right). \quad (46)$$

Determining $\boldsymbol{\mu}$ using the expression obtained in (41) requires the computation of $\mathbb{E}_{q(\mathbf{r}_p)}(\mathbf{r}_p)$, which is determined in (94) of Appendix B. One can now employ (40), (38) and (44), for iteratively updating the approximations $q\left(\mathbf{r}_p\right), q(\mathbf{h})$ and $q(\boldsymbol{\gamma})$, of the true posterior distributions $p\left(\mathbf{r}_p \mid \mathbf{r}_p^q\right), p\left(\mathbf{h} \mid \mathbf{r}_p^q\right)$ and $p\left(\boldsymbol{\gamma} \mid \mathbf{r}_p^q\right)$, respectively. Subsequently, the estimated channel is given by the posterior mean, i.e., $\widehat{\mathbf{h}}^{\text{VBL}} = \mathbb{E}_{q(\mathbf{h})}(\mathbf{h})$.

Algorithm 1 VBL for OTFS/OTSM Modulated MIMO Systems

Inputs: Received vector $\mathbf{r}_p^q$, pilot dictionary matrix $\boldsymbol{\Psi}_p$, residue ϵ and maximum iterations $N_{\max}$.
Initializations: $\widehat{\gamma}_i^{(0)} = 1, \forall\ 0 \le i \le M_\tau G_\nu N_t N_r - 1, \widehat{\boldsymbol{\Gamma}}^{(0)} = \mathbf{I}_{M_\tau G_\nu N_t N_r}$, $\widehat{\boldsymbol{\Gamma}}^{(-1)} = \mathbf{0}$, $m = 0$.
while $\left(\|\widehat{\boldsymbol{\Gamma}}^{(m)} - \widehat{\boldsymbol{\Gamma}}^{(m-1)}\|_F^2 \ge \epsilon \text{ and } m < N_{\max}\right)$ **do**
$m \leftarrow m+1$
Compute the *aposteriori* covariance and mean

$$\boldsymbol{\Sigma} = \left(\frac{1}{\sigma^2}\boldsymbol{\Psi}_p^H\boldsymbol{\Psi}_p + \mathbb{E}_{q(\boldsymbol{\gamma})}(\boldsymbol{\Gamma})\right)^{-1}$$
$$\boldsymbol{\mu} = \frac{1}{\sigma^2}\boldsymbol{\Sigma}\boldsymbol{\Psi}_p^H\mathbb{E}_{q(\mathbf{r}_p)}(\mathbf{r}_p)$$

Compute the hyperparameter estimates

$$\mathbb{E}_{q(\boldsymbol{\gamma})}(\gamma_i) = \frac{\tilde{a}}{\tilde{b}} = \frac{a+1}{b+\mathbb{E}_{q(\mathbf{h})}(h_i^2)}$$

Output: $\widehat{\mathbf{h}}^{\text{VBL}} = \boldsymbol{\mu}$.

Computation of $q(\mathbf{e})$: The approximate distribution of error $q(\mathbf{e})$ using a variational Bayesian approach is formulated as

$$q(\mathbf{e}_q) \propto \exp\left\{\mathbb{E}_{q(\mathbf{h})}\left(\ln p(\mathbf{r}_p^q|\mathbf{h},\mathbf{e}_q)\right)\right\}p(\mathbf{e}_q)$$
$$\propto \mathbb{1}\Big(\mathbf{e}_q \in (\widetilde{\boldsymbol{l}},\widetilde{\boldsymbol{u}})\Big)\exp\left\{-\frac{1}{2\sigma^2}\mathbb{E}_{q(\mathbf{h})}\|\mathbf{r}_p^q - \mathbf{e}_q - \boldsymbol{\Psi}_p\mathbf{h}\|^2\right\}$$
$$\propto \mathbb{1}\Big(\mathbf{e}_q \in (\widetilde{\boldsymbol{l}},\widetilde{\boldsymbol{u}})\Big)\exp\left\{-\frac{1}{2\sigma^2}\|\mathbf{e} - (\mathbf{r}_p^q - \boldsymbol{\Psi}_p\boldsymbol{\mu})\|^2\right\}$$

where, $p(\mathbf{e}_q) = \mathbb{1}\Big(\mathbf{e}_q \in (\widetilde{\boldsymbol{l}},\widetilde{\boldsymbol{u}})\Big)$, $\widetilde{\boldsymbol{l}} = \mathbf{r}_p^q - \boldsymbol{l}$ and $\widetilde{\boldsymbol{u}} = \mathbf{r}_p^q - \boldsymbol{u}$. Alternatively, it can be expressed in terms of step size as $p(\mathbf{e}_q) = \mathcal{U}(-\frac{\Delta_q}{2},\frac{\Delta_q}{2})$. Note that $q(\mathbf{e})$ is the product of PDFs of truncated Gaussian distributions. Thus, $\widehat{\mathbf{H}}_{\text{DD},r,t}$ can now be obtained from the channel coefficients $\widehat{\mathbf{h}}^{\text{VBL}}$ by using the relationship

$$\widehat{\mathbf{H}}_{\text{DD},r,t} = \Big(\mathbf{F}_N \otimes \mathbf{P}_{rx}\Big)\left(\sum_{i=0}^{M_\tau-1}\sum_{j=-G_\nu/2}^{G_\nu/2-1}h_{i,j,r,t}^{\text{VBL}}\left(\boldsymbol{\Pi}\right)^i\left(\boldsymbol{\Delta}_{i,j}\right)\right)\Big(\mathbf{F}_N^H \otimes \mathbf{P}_{tx}\Big). \quad (47)$$

Algorithm 1 concisely summarises the various steps in the VBL-based CSI estimation procedure for OTFS-modulated MIMO systems. The computational complexity of the proposed scheme for each iteration is approximately of the order of $O(M_\tau^2 G_\nu^2 N_t^2 M N_p)$. This complexity arises mainly due to the computation of the covariance term in (41). By comparison, the expectation maximization-based Bayesian learning approach has a similar complexity of $O(M_\tau^3 G_\nu^3 N_t^3)$, which results from the matrix inversion of size $M_\tau G_\nu N_t \times M_\tau G_\nu N_t$. Additionally, the worst-case complexity of the OMP-based scheme is $O(M^3 N_p^3 N_t^3)$, as dominated by the intermediate least squares estimate computation required in each iteration. After addressing the complexity of the proposed algorithm, the next paragraph delves into the efficiency analysis.

The efficiency E of the proposed framework can be mathematically evaluated by considering the overhead ϱ per frame and it is given as $E = 1 - \varrho$. Specifically, the proposed CSI estimation framework transmits MN_p pilot symbols within an OTFS frame of MN symbols. Consequently, the overhead is $\varrho = \frac{N_p}{N} = 0.25$ if $N_p = 8$, $N = 32$ and $\varrho = \frac{N_p}{N} = 0.125$ if $N_p = 4$, $N = 32$. However, for the embedded pilots scheme of [13], the overhead is given as $\varrho = \frac{(N_t M_\tau + M_\tau + N_t)(2G_\nu+1)}{MNN_t} \approx 0.2588$, where $N_t = 2$, $M_\tau = 6$, $G_\nu = 26$, $M = 64$, $N = 32$. The higher overhead of $N_p = 8$ can be attributed to using finite-resolution ADCs at the receiver, which requires a comparatively higher number of pilots for accurate CSI estimation. Interestingly, the quantized framework can also be extended to OTSM-modulated systems, as shown next.

V. MIMO OTSM System Model With Quantized Output

Consider an OTSM signal grid with frame duration $T_f = NT$ and bandwidth $BW = M\Delta f$, which satisfies $\Delta f = 1/T$, similar to the OTFS system defined in Section-II. The block diagram of the OTSM transceiver is shown in Fig 4. Let $\mathbf{X}_{\text{DS}} \in \mathbb{C}^{M\times N}$ denote the information symbol matrix which is placed in the delay sequency (DS)-domain, in contrast to the DD-domain of OTFS. The quantities N and M are the delay and sequency parameters, respectively. The symbols are chosen from an appropriate constellation with symbol power σ_d^2. Upon applying the Walsh-Hadamard transform (WHT) to each row of the DS-domain signal $\mathbf{X}_{\text{DS},t}$ for the tth TA, the DS-domain signal can be converted to the DT-domain signal $\mathbf{X}_{\text{DT},t}$, which is given as follows

$$\mathbf{X}_{\text{DT},t} = \mathbf{X}_{\text{DS},t} \cdot \mathbf{W}_N, \tag{48}$$

where $\mathbf{W}_N$ is the WHT matrix of size $N \times N$, whose (n, m)th element is given as

$$\mathbf{W}_N(n, m) = (1/\sqrt{N})W(n, m/N + 0.5/N). \tag{49}$$

The quantity W above denotes the Walsh function defined as $W(n, m/N + 0.5/N) \rightarrow \{-1, 1\}$. The time-domain signal $\widetilde{\mathbf{s}}_t \in \mathbb{C}^{MN\times 1}$ is derived from the delay-time (DT) domain signal by vectorizing the latter $\mathbf{X}_{\text{DT},t} \in \mathbb{C}^{M\times N}$ [7], [8].

$$\widetilde{\mathbf{s}}_t = \text{vec}(\mathbf{P}_{tx}\mathbf{X}_{\text{DT},t}), \tag{50}$$

where $\mathbf{P}_{tx} \in \mathbb{C}^{M\times M}$ is the transmitter pulse shaping filter. Substituting (48) into (50), one obtains the relationship

$$\widetilde{\mathbf{s}}_t = \text{vec}(\mathbf{P}_{tx}\mathbf{X}_{\text{DS},t}\mathbf{W}_N) = (\mathbf{W}_N^T \otimes \mathbf{P}_{tx})\text{vec}(\mathbf{X}_{\text{DS},t}). \tag{51}$$

Furthermore, after appending a CP of length L to $\widetilde{\mathbf{s}}_t$ it is transmitted through the under-spread DD-domain channel $h_{r,t}(\tau, \nu)$ given by (3). The DD-domain channel can be converted to the corresponding DT-domain channel $h_{r,t}(\tau, t)$ as

$$h_{r,t}(\tau, t) = \int h_{r,t}(\tau, \nu)\text{e}^{j2\pi\nu(t-\tau)}\, d\nu, \tag{52}$$

which can be further simplified by substituting $h_{r,t}(\tau, \nu)$ from (3) as

$$h_{r,t}(\tau, t) = \sum_{i=1}^{\rho} h_{i,r,t}e^{j2\pi\nu_i(t-\tau_i)}. \tag{53}$$

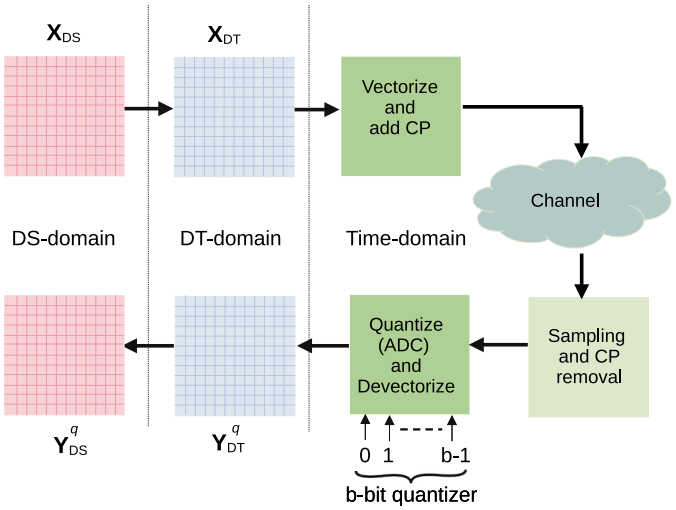

Fig. 4. Block diagram OTSM system model.

Hence, the sampled DT-domain channel $h_{r,t}[l, q]$ is given by

$$h_{r,t}[l, q] = h_{r,t}(\tau, t)|_{\tau=\frac{l}{M\Delta f}, t=\frac{q}{M\Delta f}}. \tag{54}$$

The above expression can be equivalently expressed in the discrete-time model as

$$h_{r,t}[l, q] = \sum_{i=1}^{p} h_{i,r,t}e^{j\frac{2\pi}{MN}k_i(q-l)}\delta(l - li), \tag{55}$$

where the delay index is $\tau_i = \frac{\ell_i}{M\Delta f}$, corresponding to the path delay i. Similar to OTFS systems, the path delays are integer multiples of the delay resolution of $\Delta\tau = \frac{1}{M\Delta f}$. Furthermore, the Doppler index is $\nu_i = \frac{k_i}{NT}$, where k_i may assume integer or fractional multiples of the Doppler resolution $\Delta\nu = \frac{1}{NT}$. At the receiver, the CP is removed from the received signal $\mathbf{r}$. The q_{th} element of $\mathbf{r}_r$ obtained during the demodulation process is mathematically modelled as

$$\widetilde{r}_r[q] = \sum_{t=1}^{N_t} h_{r,t}[l, q]\widetilde{s}_t[q - l] + \widetilde{w}_r[q], \tag{56}$$

where we have $0 \le q \le NM-1$. Upon stacking the outputs as the time-domain signal $\widetilde{\mathbf{r}}_r = \left[\widetilde{r}_r[0], \widetilde{r}_r[1], \cdots, \widetilde{r}_r[MN-1]\right]^T$, one obtains the input-output relationship

$$\widetilde{\mathbf{r}}_r = \sum_{t=1}^{N_t} \mathbf{H}_{r,t}\widetilde{\mathbf{s}}_t + \widetilde{\mathbf{w}}_r, \tag{57}$$

where the noise vector $\widetilde{\mathbf{w}}$ is the additive white Gaussian noise $\widetilde{\mathbf{w}} \in \mathbb{C}^{MN\times 1}$, with covariance $\mathbb{E}(\widetilde{\mathbf{w}}\widetilde{\mathbf{w}}^H) = \sigma_w^2\mathbf{I}_{MN}$. Let us quantize the received signal $\widetilde{\mathbf{r}}_r$ using the uniform quantizer $\mathcal{Q}(.)$, as defined for the OTFS counterpart in (8), Section-II. The quantized received signal $\widetilde{\mathbf{r}}_r^q \in \mathbb{C}^{MN\times 1}$ may then be expressed as

$$\widetilde{\mathbf{r}}_r^q = \mathcal{Q}\left(\sum_{t=1}^{N_t} \mathbf{H}_{r,t}\widetilde{\mathbf{s}}_t + \widetilde{\mathbf{w}}_r\right), \tag{58}$$

where the quantizer is applied to the real and imaginary parts of $\widetilde{\mathbf{r}}_r$ separately. Subsequently, applying the inverse vector

operator to the received time-domain signal $\widetilde{\mathbf{r}}_r^q$, one obtains the DT-domain signal $\mathbf{Y}_{\mathrm{DT},r}^q \in \mathcal{C}^{M\times N}$ given by

$$\mathbf{Y}_{\mathrm{DT},r}^q = \mathrm{vec}^{-1}(\tilde{\mathbf{r}}_r^q). \tag{59}$$

Application of the inverse Walsh Hadamard transform (IWHT) converts the DT-domain signal $\mathbf{Y}_{\mathrm{DT},r}^q$ to the DS-domain signal $\mathbf{Y}_{\mathrm{DS},r}^q$ that is given by

$$\mathbf{Y}_{\mathrm{DS},r}^q = \mathbf{P}_{rx}\mathbf{Y}_{\mathrm{DT},r}^q\mathbf{W}_N, \tag{60}$$

where $\mathbf{P}_{rx} \in \mathbb{C}^{M\times M}$ is the receiver pulse shaping filter. One can vectorize the DS-domain signal $\mathbf{y}_{\mathrm{DS}}^q$ to obtain the input-output model as

$$\begin{aligned}\mathbf{y}_{\mathrm{DS},r}^q &= \mathrm{vec}\left(\mathbf{Y}_{\mathrm{DS},r}^q\right) = \mathrm{vec}\left(\mathbf{P}_{rx}\mathbf{Y}_{\mathrm{DT},r}^q\mathbf{W}_N\right)\\ &= (\mathbf{W}_N^T\otimes\mathbf{P}_{rx})\,\mathrm{vec}(\mathbf{Y}_{\mathrm{DT},r}^q) = (\mathbf{W}_N^T\otimes\mathbf{P}_{rx})\widetilde{\mathbf{r}}_r^q.\end{aligned} \tag{61}$$

Upon substituting $\widetilde{\mathbf{r}}_r^q$ from (58) and $\widetilde{\mathbf{s}}_t$ from (51) into (61), one obtains

$$\mathbf{y}_{\mathrm{DS},r}^q = \left(\mathbf{W}_N^T\otimes\mathbf{P}_{rx}\right)\mathcal{Q}\left(\sum_{t=1}^{N_t}\mathbf{H}_{r,t}\left(\mathbf{W}_N^T\otimes\mathbf{P}_{rx}\right)\mathbf{x}_{\mathrm{DS},t}+\widetilde{\mathbf{w}}_r\right). \tag{62}$$

Stacking the received outputs $\mathbf{y}_{\mathrm{DS},r}^q$ as $[(\mathbf{y}_{\mathrm{DS},1}^q)^T,\cdots,(\mathbf{y}_{\mathrm{DS},N_r}^q)^T]^T$ one obtains $\mathbf{y}_{\mathrm{DS}}^q \in \mathbb{C}^{MNN_r\times 1}$ as

$$\mathbf{y}_{\mathrm{DS}}^q = (\mathbf{I}_{N_r}\otimes\mathbf{W}_N^T\otimes\mathbf{P}_{rx})\mathcal{Q}\left(\mathbf{H}(\mathbf{I}_{N_t}\otimes\mathbf{W}_N^T\otimes\mathbf{P}_{tx})\mathbf{x}_{\mathrm{DS}}+\widetilde{\mathbf{w}}\right). \tag{63}$$

The above expression in (63) can be further simplified for rectangular pulses as

$$\mathbf{y}_{\mathrm{DS}}^q = (\mathbf{I}_{N_r}\otimes\mathbf{W}_N^T\otimes\mathbf{I}_M)\mathcal{Q}\left(\mathbf{H}(\mathbf{I}_{N_t}\otimes\mathbf{W}_N^T\otimes\mathbf{I}_M)\mathbf{x}_{\mathrm{DS}}+\widetilde{\mathbf{w}}\right), \tag{64}$$

where $\mathbf{P}_{tx}=\mathbf{P}_{rx}=\mathbf{I}_M$ and the function $\mathcal{Q}(.)$ is given by (8). Thus, the input DD-domain symbol vector $\mathbf{x}_{\mathrm{DS}}$ is mapped to the output symbol vector $\widetilde{\mathbf{y}}_{\mathrm{DS}}^q \in \mathbb{C}^{MNN_r\times 1}$ as

$$\widetilde{\mathbf{y}}_{\mathrm{DS}}^q = \mathcal{Q}\left(\widetilde{\mathbf{H}}\mathbf{x}_{\mathrm{DS}}+\widetilde{\mathbf{w}}\right), \tag{65}$$

where $\widetilde{\mathbf{H}} = \mathbf{H}(\mathbf{I}_{N_t}\otimes\mathbf{W}_N^T\otimes\mathbf{I}_M)$, $\widetilde{\mathbf{y}}_{\mathrm{DS}}^q = (\mathbf{I}_{N_r}\otimes\mathbf{W}_N\otimes\mathbf{I}_M)\mathbf{y}_{\mathrm{DS}}^q$. The MMSE detection rule of recovering the symbol vector $\hat{\mathbf{x}}_{\mathrm{DS}}$ is given by

$$\hat{\mathbf{x}}_{\mathrm{DS}}^{\mathrm{MMSE}} = \left(\widetilde{\mathbf{H}}^H\widetilde{\mathbf{H}}+\frac{\sigma_d^2}{\sigma_w^2}\mathbf{I}_{MN}\right)^{-1}\widetilde{\mathbf{H}}^H\widetilde{\mathbf{y}}_{\mathrm{DS}}^q. \tag{66}$$

However, similar to OTFS systems, the performance of the above detector suffers from the quantization error introduced by the finite-resolution ADCs. Thus, a modified detection rule is proposed in Section- VII which incorporates this effect. The sparse CSI estimation problem of the quantized OTSM system is formulated next.

VI. Sparse Channel Model for MIMO OTSM Systems

Consider the DD-domain channel $h_{r,t}(\tau,\nu)$ as given in Section III. The DT-domain sparse channel can be expressed as

$$h_{r,t}(\tau,t) = \sum_{i=0}^{M_\tau-1}\sum_{j=-G_\nu/2}^{G_\nu/2-1} h_{i,j,r,t}\mathrm{e}^{j2\pi\nu_j(t-\tau_i)}. \tag{67}$$

The above channel can also be represented in a matrix form by (21), similar to OTFS systems. Substituting $\mathbf{H}_{r,t}$ from (21) into (58) and rearranging the received signal $\widetilde{\mathbf{r}}_r^q \in \mathbb{C}^{MN_p\times 1}$ can be obtained as

$$\widetilde{\mathbf{r}}_r^q = \mathcal{Q}\Big(\sum_{t=1}^{N_t}\sum_{i=0}^{M_\tau-1}\sum_{j=-G_\nu/2}^{G_\nu/2-1} h_{i,j,r,t}\left(\mathbf{\Pi}\right)^i\left(\mathbf{\Delta}_{i,j}\right)\widetilde{\mathbf{s}}_t+\mathbf{w}_r\Big), \tag{68}$$

where the total number of pilot symbols $\widetilde{\mathbf{s}}_t$ inserted in the time domain are MN_p. The above equation (68) can be expressed as

$$\widetilde{\mathbf{r}}_r^q = \mathcal{Q}\Big(\sum_{t=1}^{N_t}\sum_{i=0}^{M_\tau-1}\sum_{j=-G_\nu/2}^{G_\nu/2-1} h_{i,j,r,t}\widetilde{\boldsymbol{\psi}}_{i,j,t}+\mathbf{w}_r\Big), \tag{69}$$

where we have $\widetilde{\boldsymbol{\psi}}_{i,j,t} = \left(\mathbf{\Pi}\right)^i\left(\mathbf{\Delta}_{i,j}\right)\widetilde{\mathbf{s}}_t \in \mathbb{C}^{MN_p\times 1}$. The above relationship can be simplified as

$$\widetilde{\mathbf{r}}_r^q = \mathcal{Q}\left(\sum_{t=1}^{N_t}\widetilde{\mathbf{\Psi}}_t\mathbf{h}_{r,t}+\mathbf{w}_r\right), \tag{70}$$

where $\widetilde{\mathbf{\Psi}}_t \in \mathbb{C}^{MN_p\times M_\tau G_\nu}$ is the dictionary matrix and $\mathbf{h}_{r,t}$ is the channel coefficient vector corresponding to the r_{th} RA and t_{th} TA defined as

$$\widetilde{\mathbf{\Psi}}_t = \left[\widetilde{\psi}_{0,-G_\nu/2,t}\cdots\widetilde{\psi}_{0,G_\nu/2-1,t}\cdots\widetilde{\psi}_{M_\tau-1,G_\nu/2-1,t}\right]$$

$$\mathbf{h}_{r,t} = \left[h_{0,G_\nu/2,r,t}\cdots h_{0,G_\nu/2-1,r,t}\cdots h_{M_\tau-1,G_\nu/2-1,r,t}\right]^T.$$

Furthermore, the above expression in (70) can be simplified as

$$\widetilde{\mathbf{r}}_r^q = \mathcal{Q}\left(\widetilde{\mathbf{\Psi}}\mathbf{h}_r+\mathbf{w}_r\right), \tag{71}$$

where $\widetilde{\mathbf{\Psi}} \in \mathbb{C}^{MN_p\times M_\tau G_\nu N_t}$ obeys $\widetilde{\mathbf{\Psi}} = \left[\widetilde{\mathbf{\Psi}}_1,\widetilde{\mathbf{\Psi}}_2,\cdots,\widetilde{\mathbf{\Psi}}_{N_t}\right]$ and $\mathbf{h}_r \in \mathbb{C}^{M_\tau G_\nu N_t\times 1}$ is formed as $\left[\mathbf{h}_{r,1}^T,\cdots,\mathbf{h}_{r,N_t}^T\right]^T$.

By stacking the received vectors across all RAs, the received vector $\widetilde{\mathbf{r}}^q \in \mathbb{C}^{MN_pN_r\times 1}$ is defined as $\widetilde{\mathbf{r}}^q = \left[(\widetilde{\mathbf{r}}_1^q)^T,(\widetilde{\mathbf{r}}_2^q)^T,\cdots,(\widetilde{\mathbf{r}}_{N_r}^q)^T\right]^T$:

$$\widetilde{\mathbf{r}}^q = \mathcal{Q}\left(\widetilde{\mathbf{\Psi}}\mathbf{h}+\mathbf{w}\right), \tag{72}$$

where we have $\mathbf{h} = \left[(\mathbf{h}_1)^T,\cdots,(\mathbf{h}_{N_r})^T\right]^T \in \mathbb{C}^{M_\tau G_\nu N_t N_r\times 1}$.

The VBL-based sparse estimation procedure of Algorithm 1 can be applied to (72) for OTSM modulated systems to estimate $\mathbf{h}$, which can be further utilized to compute the

DS-domain channel $\widehat{\mathbf{H}}_{\mathrm{DS},r,t}$ between the rth RA and tth TA given by

$$\widehat{\mathbf{H}}_{\mathrm{DS},r,t} = \left(\mathbf{W}_N^T \otimes \mathbf{P}_{rx}\right)\left(\sum_{i=0}^{M_\tau-1}\sum_{j=-G_\nu/2}^{G_\nu/2-1} h_{i,j,r,t}^{\mathrm{VBL}}\,(\mathbf{\Pi})^i\,(\mathbf{\Delta}_{i,j})\right)\left(\mathbf{W}_N^T \otimes \mathbf{P}_{tx}\right). \tag{73}$$

The subsequent section delves into the OTFS and OTSM detection scheme.

VII. Detection of Data From Quantized Received Signals

Consider the end-to-end system models given by (19) and (65), respectively, for OTFS and OTSM-based systems. For simplicity of notation, the subscripts DD and DS are dropped, and DX are used hereafter since the procedure is similar for both systems. The generalized input-output relationship in these systems can be expressed as

$$\widetilde{\mathbf{y}}_{\mathrm{DX}}^q = \quad \mathcal{Q}(\mathbf{y}_{\mathrm{DX}}) = \mathcal{Q}(\widetilde{\mathbf{H}}\mathbf{x}_{\mathrm{DX}} + \mathbf{w}) = \mathbf{y}_{\mathrm{DX}} + \mathbf{e}, \tag{74}$$

where $\mathbf{x}_{\mathrm{DX}} = \mathrm{vec}(\mathbf{X}_{\mathrm{DX}})$, $\mathbb{E}(\mathbf{X}_{\mathrm{DX}}\mathbf{X}_{\mathrm{DX}}^H) = \sigma_d^2\mathbf{I}_M$. The vector $\mathbf{e}$ denotes the quantization error arising due to the use of finite resolution ADCs, where e_i represents its ith element and for a uniform quantizer, e_i and y_i^q satisfy the following properties

$$\mathrm{E}\left[e_i\right] = 0, \mathrm{E}\left[y_i^q e_i\right] = \mathrm{E}\left[y_i e_i + e_i^2\right] = 0. \tag{75}$$

The distortion factor d arising due to the quantization process is modelled as $d_i = \mathbb{E}\left[e_i^2\right]/c_{y_iy_i}$, where $c_{y_iy_i} = \mathrm{E}\left[y_i^2\right]$. For a uniform quantizer, $d_i = d = \Delta^2/12$. Substituting d_i into (75) yields $\mathrm{E}\left[y_i e_i\right] = -dc_{y_iy_i}$. The mean square error (MSE) for the estimate of $\mathbf{x}_{\mathrm{DX}}$ is defined as

$$\mathrm{MSE} = \mathbb{E}\left[\|\mathbf{x}_{\mathrm{DX}} - \hat{\mathbf{x}}_{\mathrm{DX}}\|_2^2\right] = \mathbb{E}\left[\|\mathbf{x}_{\mathrm{DX}} - \mathbf{G}_{\mathrm{DX}}\widetilde{\mathbf{y}}_{\mathrm{DX}}^q\|_2^2\right],$$

where the linear receiver $\mathbf{G}_{\mathrm{DX}}$ that minimizes the MSE, also termed the LMMSE receiver, is defined as

$$\mathbf{G}_{\mathrm{DX}} = \mathbf{C}_{\mathbf{x}\mathbf{y}^q}\mathbf{C}_{\mathbf{y}^q\mathbf{y}^q}^{-1}, \tag{76}$$

where $\mathbf{C}_{\mathbf{x}\mathbf{y}^q}$ is of size $(MN \times MN)$ defined as

$$\mathbf{C}_{\mathbf{x}\mathbf{y}^q} = \mathbb{E}\left[\mathbf{x}_{\mathrm{DX}}(\widetilde{\mathbf{y}}_{\mathrm{DX}}^q)^H\right] = \mathbb{E}[\mathbf{x}_{\mathrm{DX}}(\mathbf{y}_{\mathrm{DX}} + \mathbf{e})^H] = \mathbf{C}_{\mathbf{xy}} + \mathbf{C}_{\mathbf{xe}}. \tag{77}$$

The covariance matrices $\mathbf{C}_{\mathbf{xy}}$ and $\mathbf{C}_{\mathbf{xe}}$ are defined as

$$\mathbf{C}_{\mathbf{xy}} = \mathbb{E}\left[\mathbf{x}_{\mathrm{DX}}\mathbf{y}_{\mathrm{DX}}^H\right] = \mathbb{E}\left[\mathbf{x}_{\mathrm{DX}} \cdot (\widetilde{\mathbf{H}}\mathbf{x}_{\mathrm{DX}} + \mathbf{w})^H\right] = \sigma_d^2\widetilde{\mathbf{H}}^H \tag{78}$$

$$\begin{aligned}\mathbf{C}_{\mathbf{xe}} &= \mathbb{E}\left[\mathbf{x}_{\mathrm{DX}}\mathbf{e}^H\right] = \mathbb{E}_{\mathbf{y}}\left[\mathbb{E}\left[\mathbf{x}_{\mathrm{DX}}\mathbf{e}^H \mid \mathbf{y}_{\mathrm{DX}}\right]\right] \\ &= \mathbb{E}_{\mathbf{y}}\left[\mathbb{E}[\mathbf{x}_{\mathrm{DX}} \mid \mathbf{y}_{\mathrm{DX}}]\mathbb{E}\left[\mathbf{e}^H \mid \mathbf{y}_{\mathrm{DX}}\right]\right] \\ &\approx \mathbb{E}_{\mathbf{y}}\left[\mathbf{C}_{\mathbf{xy}}\mathbf{C}_{\mathbf{yy}}^{-1}\mathbf{y}_{\mathrm{DX}}\mathbb{E}\left[\mathbf{e}^H \mid \mathbf{y}_{\mathrm{DX}}\right]\right] \\ &= \mathbf{C}_{\mathbf{xy}}\mathbf{C}_{\mathbf{yy}}^{-1}\mathbb{E}\left[\mathbf{y}_{\mathrm{DX}}\mathbf{e}^H\right] = -d\mathbf{C}_{\mathbf{xy}} = -d\sigma_d^2\widetilde{\mathbf{H}}^H. \end{aligned} \tag{79}$$

Substituting (79) into (77) yields

$$\mathbf{C}_{\mathbf{x}\mathbf{y}^q} = (1-d)\,\sigma_d^2\widetilde{\mathbf{H}}^H. \tag{80}$$

From (76), the covariance matrix $\mathbf{C}_{\mathbf{y}^q\mathbf{y}^q} \in MN \times MN$ is determined as

$$\begin{aligned}\mathbf{C}_{\mathbf{y}^q\mathbf{y}^q} &= \mathbb{E}\left[\widetilde{\mathbf{y}}_{\mathrm{DX}}^q(\widetilde{\mathbf{y}}_{\mathrm{DX}}^q)^H\right] = \mathbb{E}\left[(\mathbf{y}_{\mathrm{DX}} + \mathbf{e})(\mathbf{y}_{\mathrm{DX}} + \mathbf{e})^H\right] \\ &= \mathbf{C}_{\mathbf{yy}} + \mathbf{C}_{\mathbf{ye}} + \mathbf{C}_{\mathbf{ye}}^H + \mathbf{C}_{\mathbf{ee}}. \end{aligned} \tag{81}$$

For computing the covariance matrix $\mathbf{C}_{\mathbf{y}^q\mathbf{y}^q}$, where various terms in the expressions above can be evaluated as shown next. The quantity $\mathbf{C}_{\mathbf{yy}}$ is obtained as

$$\begin{aligned}\mathbf{C}_{\mathbf{yy}} &= \mathbb{E}\left[\mathbf{y}_{\mathrm{DX}}\mathbf{y}_{\mathrm{DX}}^H\right] = \mathbb{E}\left[(\widetilde{\mathbf{H}}\mathbf{x}_{\mathrm{DX}} + \mathbf{w})(\widetilde{\mathbf{H}}\mathbf{x}_{\mathrm{DX}} + \mathbf{w})^H\right] \\ &= \mathbb{E}\left[\widetilde{\mathbf{H}}\mathbf{x}_{\mathrm{DX}}\mathbf{x}_{\mathrm{DX}}^H\widetilde{\mathbf{H}}^H + \widetilde{\mathbf{H}}\mathbf{x}_{\mathrm{DX}}\mathbf{w}^H + \mathbf{w}\mathbf{x}_{\mathrm{DX}}^H\widetilde{\mathbf{H}}^H + \mathbf{w}\mathbf{w}^H\right] \\ &= \sigma_d^2\widetilde{\mathbf{H}}\widetilde{\mathbf{H}}^H + \sigma_w^2\mathbf{I}_{MN}. \end{aligned} \tag{82}$$

Consider the ith element of $\mathbf{y}$ and jth element of vector $\mathbf{e}$. The covariance $c_{\mathbf{y}_i\mathbf{e}_j}$ corresponding to these quantities is given by

$$c_{\mathbf{y}_i\mathbf{e}_j} = \begin{cases} \mathbb{E}\left[y_i e_i\right] = -d\,c_{y_iy_i} & \text{if } i = j \\ \mathbb{E}\left[y_i e_j^*\right] = -d\,c_{y_iy_j} & \text{if } i \neq j, \end{cases} \tag{83}$$

where $\mathbb{E}\left[y_i e_j^*\right]$ is given as

$$\begin{aligned}\mathbb{E}\left[y_i e_j^*\right] &= \mathbb{E}_{y_j}\left[\mathbb{E}\left[y_i e_j^* \mid y_j\right]\right] = \mathbb{E}_{y_j}\left[\mathbb{E}\left[y_i \mid y_j\right]\mathbb{E}\left[e_j^* \mid y_j\right]\right] \\ &\approx \mathbb{E}_{y_j}\left[c_{y_iy_j}c_{y_jy_j}^{-1}y_j\mathbb{E}\left[e_j^* \mid y_j\right]\right] = c_{y_iy_j}c_{y_jy_j}^{-1}\mathbb{E}\left[y_j e_j^*\right] \end{aligned} \tag{84}$$

The linear estimator $r_{y_iy_j}r_{y_jy_j}^{-1}y_j$ and the Bayesian estimator $\mathrm{E}\left[y_i \mid y_j\right]$ are identical if the vector $\mathbf{y}$ is jointly Gaussian distributed and thus the above expression of the covariance matrix $\mathbf{C}_{\mathbf{ye}}$ can be approximated as

$$\mathbf{C}_{\mathbf{ye}} \approx -d\mathbf{C}_{\mathbf{yy}}. \tag{85}$$

Similarly, the elements of the covariance matrix $\mathbf{C}_{\mathbf{ee}}$ are computed as

$$c_{\mathbf{e}_i\mathbf{e}_j} = \begin{cases} \mathbb{E}\left[e_i e_i\right] = dc_{y_iy_i} & \text{if } i = j \\ \mathbb{E}\left[e_i e_j^*\right] = d^2c_{y_iy_j} & \text{if } i \neq j, \end{cases}$$

where the expression for $\mathbb{E}\left[e_i e_j^*\right]$ is derived as

$$\begin{aligned}\mathbb{E}\left[e_i e_j^*\right] &= \mathbb{E}_{y_j}\left[\mathbb{E}\left[e_i e_j^* \mid y_j\right]\right] = \mathbb{E}_{y_j}\left[\mathbb{E}\left[e_i \mid y_j\right]\mathbb{E}\left[e_j^* \mid y_j\right]\right] \\ &\approx \mathbb{E}_{y_j}\left[c_{e_iy_j}c_{y_jy_j}^{-1}y_j\mathbb{E}\left[e_j^* \mid y_j\right]\right] = c_{y_je_i}^*c_{y_jy_j}^{-1}\mathbb{E}\left[y_j e_j^*\right] \\ &= d^2c_{y_jy_i}^*. \end{aligned} \tag{86}$$

Thus, the covariance matrix of quantization error can be approximated as

$$\begin{aligned}\mathbf{C}_{\mathbf{ee}} &\approx d\,\mathrm{diag}\left(\mathbf{C}_{\mathbf{yy}}\right) + d^2\,\mathrm{nondiag}\left(\mathbf{C}_{\mathbf{yy}}\right) \\ &= d\mathbf{C}_{\mathbf{yy}} - (1-d)\,d\,\mathrm{nondiag}\left(\mathbf{C}_{\mathbf{yy}}\right). \end{aligned} \tag{87}$$

Upon substituting the parameters into (81) yields

$$\mathbf{C}_{\mathbf{y}^q\mathbf{y}^q} \approx (1-d)\left(\mathbf{C}_{\mathbf{yy}} - d\mathrm{nondiag}\left(\mathbf{C}_{\mathbf{yy}}\right)\right).$$

Thus, the linear receiver that minimizes the MSE is formulated as

$$\mathbf{G}_{\mathrm{DX}} \approx \mathbf{C}_{\mathbf{xy}}\Big(\mathbf{C}_{\mathbf{yy}} - d\ \ \mathrm{nondiag}(\mathbf{C}_{\mathbf{yy}})\Big)^{-1}. \tag{88}$$

The estimates of the DD-domain symbols $\hat{\mathbf{x}}_{\text{DD}}$ and DS-domain symbols $\hat{\mathbf{x}}_{\text{DS}}$ can now be obtained as

$$\hat{\mathbf{x}}_{\text{DD}} = \mathbf{G}_{\text{DD}}\widetilde{\mathbf{y}}^q, \qquad \hat{\mathbf{x}}_{\text{DS}} = \mathbf{G}_{\text{DS}}\widetilde{\mathbf{y}}^q. \tag{89}$$

Using (88), the LMMSE receiver $\mathbf{G}_{\text{DD}}$ can be obtained in the DD-domain for OTFS systems. Similarly, for OTSM systems in the DS-domain, the receiver is denoted as $\mathbf{G}_{\text{DS}}$.

VIII. Simulation Results

This section presents simulation results, where parameters used are listed in Table- II. In the simulations the extended vehicular A model (EVA) model [46] is employed, where the channel coefficients for each r, t are generated by multiplying the power profile of the EVA channel model with coefficients drawn as independent samples of the complex normal distribution $\mathcal{CN}(0, 1)$. In the simulations, $M = 64$ and $\Delta f = 30$ kHz are selected, yielding a bandwidth of $BW = M\Delta f = 1.920$ MHz. This bandwidth results in a delay resolution or symbol duration of $1/BW = \frac{1}{M\Delta f} = 0.5208\mu s$, which is consistent with the lower end of wideband system requirements. The maximum delay spread for the EVA-channel model, with a delay resolution of 0.5208μs, results in a maximum delay tap of $M_\tau = 6$. The total number of Monte iterations used is 1000. For the BL-based approach, the stopping parameters ϵ and N_{max} are set to 10^{-6} and 100. The performance metrics evaluated are the normalised mean square error (NMSE) and the symbol error rate (SER). The NMSE is defined as

$$\text{NMSE} = \frac{\|\widehat{\mathbf{H}}_{\text{DX}} - \mathbf{H}_{\text{DX}}\|^2}{\|\mathbf{H}_{\text{DX}}\|^2}. \tag{90}$$

A. NMSE Performance for OTFS Systems

The NMSE performance for the OTFS-modulated system is depicted in Fig. 5(a), where NMSE is defined by (90). Within this framework, $\widehat{\mathbf{H}}_{\text{DX}} = \widehat{\mathbf{H}}_{\text{DD}}$ is the CSI estimated in the DD-domain for the MIMO OTFS modulated system. The channel corresponding to the hypothetical receiver having perfectly known CSI is given as $\mathbf{H}_{\text{DX}} = \mathbf{H}_{\text{DD}}$. In this context, the proposed algorithm is compared with the Orthogonal matching pursuit (OMP) [9], focal underdetermined system solver (FOCUSS) [47] and traditional MMSE estimators. The performance is also evaluated against a VBL having unquantized outputs. The results are presented for $N_p = 4, 8$, where the total number of pilots in time-domain are MN_p.

An important observation from Fig 5(a) is that as the pilot SNR increases, the NMSE decreases. This is along expected lines since the impact of noise on the estimate decreases due to the increasing pilot power, which leads to an improvement in the CSI estimation. Furthermore, the Bayesian learning-based VBL technique conceived is seen to yield a remarkably improved performance as compared to the OMP, FOCUSS, and MMSE. The reduced performance of OMP can be attributed to its sensitivity to the stopping parameter, which also depends on the SNR value. Hence, the stopping parameter optimized for the mid-range SNR may not be optimal for higher SNRs. Furthermore, at high SNRs, the quantization error dominates the thermal noise at the receiver, and OMP does not take the quantization error into account; hence resulting in poor performance. The MMSE estimator employed is given as $\widehat{\mathbf{H}}_{MMSE} = \left(\Psi^H\mathbf{R}_w^{-1}\Psi + \mathbf{R}_h^{-1}\right)^{-1}\Psi^H\mathbf{R}_w^{-1}\mathbf{r}^q$, where $\mathbf{R}_h \in \mathbb{C}^{M_\tau G_\nu N_t \times M_\tau G_\nu N_t}$ denotes covariance matrix of vector $\mathbf{h}_r$. Since the true channel covariance is unknown, we assume $\mathbf{R}_h = \mathbf{I}_{M\tau G_\nu N_t}$. Additionally, only ρ out of the $M_\tau G_\nu$ elements of the vector $h_{r,t}$ are typically non-zero, resulting in a sparse channel. The conventional MMSE estimator does not take advantage of this sparsity, and the quantized observation vector leads to further suboptimality. Moreover, reducing the number of pilots from $N_p = 8$ to $N_p = 4$ degrades the NMSE performance. At high SNR, the impact is more severe owing to increased step size, which elevates the quantization noise, leading to performance erosion.

TABLE II
Parameters Used in OTFS and OTSM Systems

Carrier frequency f_c	10 GHz
Subcarrier spacing Δf	30 KHz
# of Delay-axis symbols M	64
# of Doppler-axis symbols N	32
# of TA N_t	2
# of RA N_r	2
Max. doppler spread M_τ	6
Max. delay spread N_ν	6
Modulation scheme	BPSK, QPSK
Pulse-shape	Rectangular

B. NMSE Performance of OTSM Systems

The NMSE of the OTSM-modulated system is given by (90), where $\widehat{\mathbf{H}}_{\text{DX}} = \widehat{\mathbf{H}}_{\text{DS}}$ is the CSI estimated in the DS-domain for OTSM modulated systems. The channel corresponding to the hypothetical receiver with perfectly known CSI is given as $\mathbf{H}_{\text{DX}} = \mathbf{H}_{\text{DS}}$. Fig. 6(a) shows the NMSE comparison of the different schemes competing for CSI acquisition in the OTSM system, which exhibits a trend similar to that of Fig. 5 seen for OTFS systems, wherein the proposed VBL scheme outperforms the other estimators, such as the OMP, FOCUSS, and MMSE, while achieving an NMSE close to that of the VBL applied to an unquantized received signal.

C. SER Performance of OTFS and OTSM Systems

Figures 5(b), 6(b) show the SER performance of the receivers constructed as per (89) using the CSI estimates obtained from the various estimation schemes considered. The system parameters considered are $M = 64$, $N = 32$, $N_p = 8$, $b = 6$, pilot SNR = 10dB. The modulation scheme considered is BPSK and QPSK. The conventional MMSE and modified MMSE are considered for comparison. The SER of the modified MMSE scheme-based detection rule for the proposed VBL technique, represented as VBL-M, is compared to that of the OMP and MMSE schemes for both the modulated systems. Moreover, the performance is benchmarked against that of a hypothetical receiver having a perfect CSI scenario. It can be observed from Fig. 5(b) and Fig. 6(b) that the SER

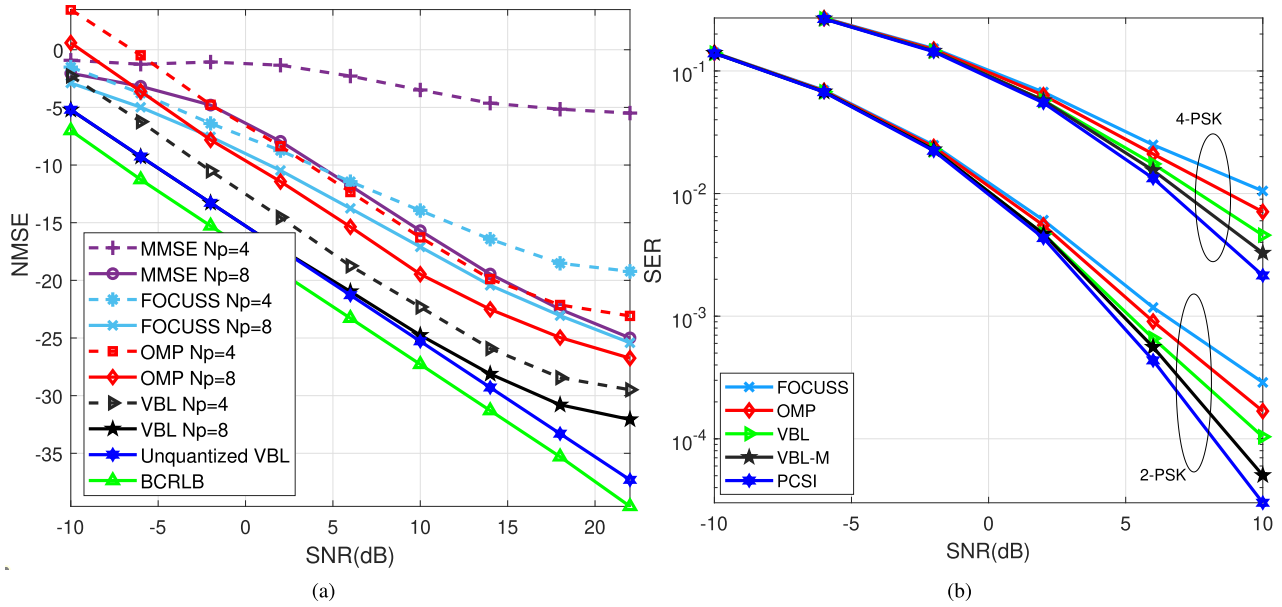

Fig. 5. OTFS system with $M = 64$, $N = 32$ (a) NMSE vs SNR with $b = 6$ and $N_p = [4, 8]$ (b) SER vs. SNR with $b = 6$ and $N_p = 8$.

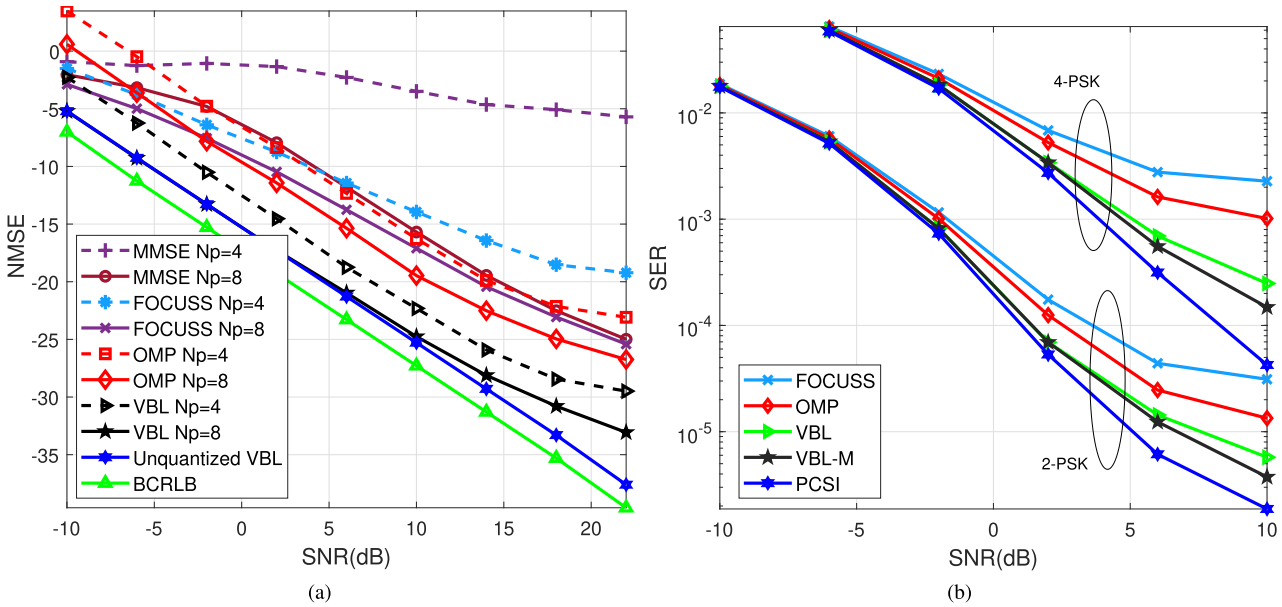

Fig. 6. OTSM system with $M = 64$, $N = 32$ (a) NMSE vs SNR with $b = 6$ and $N_p = [4, 8]$ (b) SER vs. SNR with $b = 6$ and $N_p = 8$.

performance of the VBL framework is substantially improved in comparison to the other algorithms listed above owing to the superior CSI estimates obtained by the proposed algorithm. This observation is in line with the previous discussions concerning the NMSE plots. Also, the modified MMSE detector VBL-M leads to a further performance improvement over its conventional counterpart since the latter does not account for the quantization error.

D. NMSE and SER Performance With Varying ADC Resolutions

The performance of the proposed VBL algorithm in terms of NMSE relative to the SNR for $M = 64$, $N = 32$ and the quantization bit values $b = [1, 2, \cdots, 8]$ in the MIMO OTFS and OTSM system is as illustrated in Figures 7(a) and 8(a), respectively. The performance is also compared to that of $b =$ inf bit, which represents an unquantized scenario. It is evident that as the number of bits is increased, the NMSE performance improves owing to an increase in the resolution and the accompanying reduction in the quantization error. Moreover, in both plots, the NMSE performance of the proposed CSI estimation schemes saturates at higher values of SNR. This is due to the fact that at high SNR, the quantization noise power increases due to the increased quantization step size. This effect is also observed in Figures 7(b) and 8(b), where the SER performance for $b = 4$ and $b = 6$ is analyzed, demonstrating a similar trend.

Furthermore, the NMSE performance in Figures 5(a) and 6(a) or 7(a) and 8(a) for OTFS and OTSM system models are comparable, as they both use the same

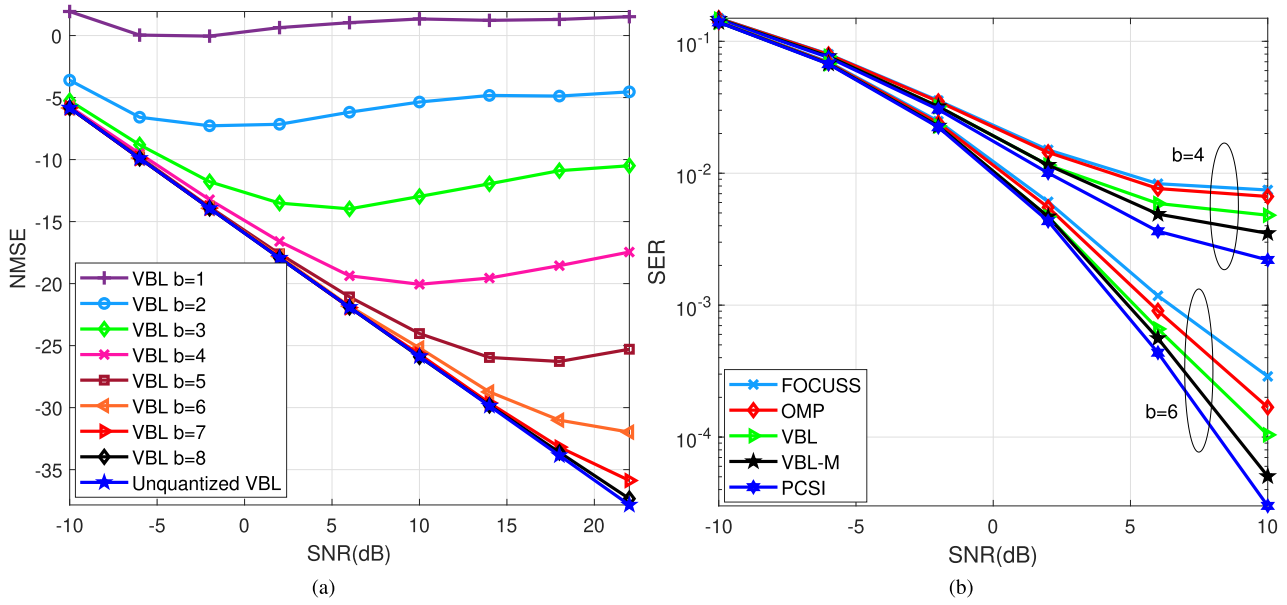

Fig. 7. OTFS system with $M = 64$, $N = 32$ (a) NMSE vs SNR with $b = [1 : 8]$, $N_p = 8$ (b) SER vs. SNR with $b = 4, 6$, $N_p = 8$.

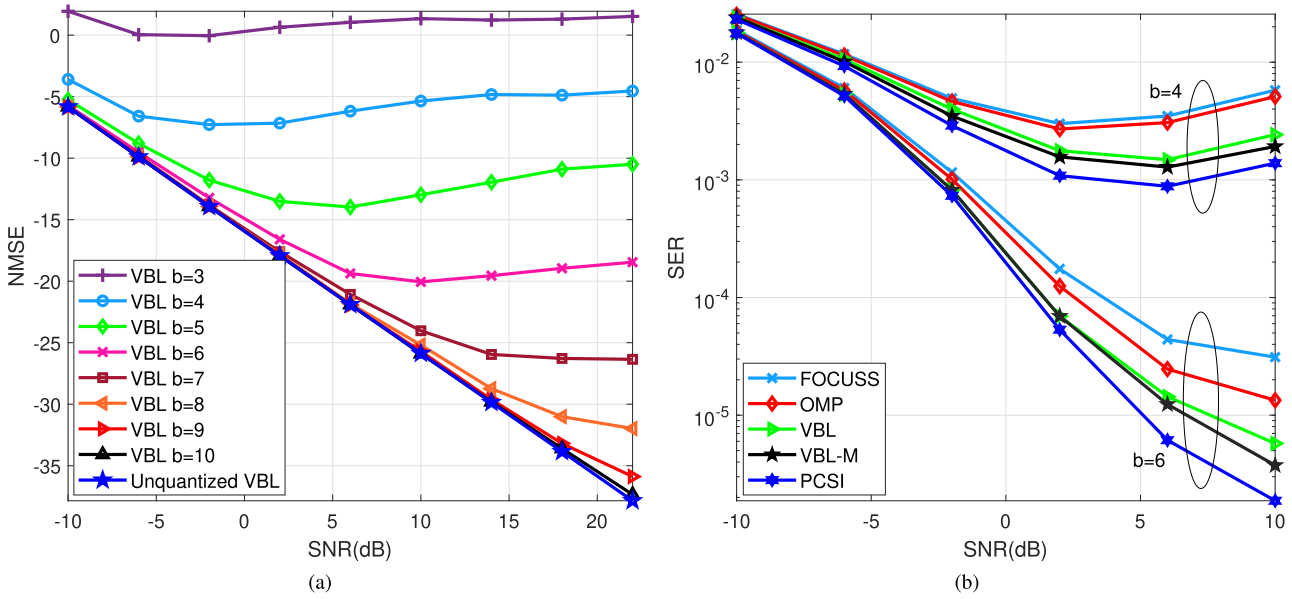

Fig. 8. OTSM system with $M = 64$, $N = 32$ (a) NMSE vs SNR with $b = [1 : 8]$, $N_p = 8$ (b) SER vs. SNR with $b = 4, 6$, $N_p = 8$.

time-domain pilots. The estimated and original channel coefficients undergo similar transformations when considering the DD-domain channels for OTFS, as described in (47) and DS-domain channels for OTSM systems, as described in (73). However, the SER performance of the OTSM-based system, which transmits information in the DS-domain, surpasses that of the OTFS-based system, which transmits in the DD-domain. This improvement is due to the use of the WHT transform [7], [8]. Additionally, the detection time for OTSM is shorter compared to OTFS. At the transmitter, the OTFS system requires converting information from the DD-domain to the TF-domain and then to the time domain. In contrast, OTSM converts DS-domain information to the DT-domain, which is then vectorized to the time domain. At the receiver, these steps are performed in reverse, making the latter process more efficient.

IX. Summary and Conclusion

Novel CSI estimation techniques have been conceived for OTFS- and OTSM-aided MIMO systems using finite-resolution ADCs. The end-to-end input-output models were developed for both systems. The CSI estimation problem of both was reformulated as sparse estimation problems and a novel Bayesian learning-based VBL algorithm was derived for CSI estimation in these systems that exploits the inherent sparsity in the DD-domain and DS-domain, respectively. Subsequently, a modified MMSE-based detector was proposed for incorporating the quantization error. This method was shown to attain a substantially enhanced detection performance when compared to the idealized unquantized MMSE counterpart. In conclusion, Finite-resolution ADCs emerged as a promising solution amidst the rising demand for high-resolution ADCs. Comprehensive simulations validated the effectiveness of

proposed techniques, showcasing enhanced performance in terms of NMSE and SER compared to existing algorithms. These results not only exceeded the state-of-the-art but also signify a transition towards sustainable communication technologies.

APPENDIX A
COMPUTATION OF $\mathcal{L}(q)$ AND $q(\mathcal{A}_i)$

Consider the lower bound $\mathcal{L}(q)$ given by (32). Substituting (34) into (32), one obtains

$$\begin{aligned}\mathcal{L}(q) &= \int \prod_i q(\mathcal{A}_i) \left\{ \ln p\left(\mathbf{r}_p^q, \mathcal{A}\right) - \sum_i \ln q(\mathcal{A}_i) \right\} d\mathcal{A} \\ &= \int q(\mathcal{A}_i) \left\{ \int \ln p\left(\mathbf{r}_p^q, \mathcal{H}\right) \prod_{j\neq i} q(\mathcal{A}_j)\, d\mathcal{A}_{\sim i} \right\} d\mathcal{A}_i \\ &\quad - \int q(\mathcal{A}_i) \ln q(\mathcal{A}_i)\, d\mathcal{A}_i + \text{ constant term} \\ &= \int q(\mathcal{A}_i) \ln \tilde{p}\left(\mathbf{r}_p^q, \mathcal{A}_i\right) d\mathcal{A}_i - \int q(\mathcal{A}_i) \ln q(\mathcal{A}_i)\, d\mathcal{A}_i \\ &\quad + \text{ constant term} \\ &= -\mathrm{KL}\left(q(\mathcal{A}_i) \,\|\tilde{p}\left(\mathbf{r}_p^q, \mathcal{A}_i\right)\right) + \text{ constant term}\,, \end{aligned} \tag{91}$$

where the term $\ln \tilde{p}\left(\mathbf{r}_p^q, \mathcal{A}_i\right)$ is determined as

$$\ln \tilde{p}(\boldsymbol{z}_p, \mathcal{A}_i) = \int \ln p\left(\mathbf{r}_p^q, \mathcal{A}\right) \prod_{j\neq i} q(\mathcal{A}_j)\, d\mathcal{A}_{\sim i}. \tag{92}$$

The integration above is performed with respect to all $q(\mathcal{A}_j), j \neq i$. The constant term does not depend on $\mathcal{A}_i$. Maximization of the quantity $\mathcal{L}(q)$ reduces to the minimization of the term $\mathrm{KL}\left(q(\mathcal{A}_i) \,\|\tilde{p}\left(\mathbf{r}_p^q, \mathcal{A}_i\right)\right)$, which is attained, when $q(\mathcal{A}_i)$ is equal to $\tilde{p}\left(\mathbf{r}_p^q, \mathcal{A}_i\right)$, as shown below

$$\begin{aligned} q(\mathcal{A}_i) &= \tilde{p}\left(\mathbf{r}_p^q, \mathcal{A}_i\right) \\ &= \frac{1}{Z} \exp\left\{ \int \ln p\left(\mathbf{r}_p^q, \mathcal{A}\right) \prod_{j\neq i} q(\mathcal{A}_j)\, d\mathcal{A}_{\sim i} \right\} \\ &\propto \exp\left\{ \mathbb{E}_{\sim q(\mathcal{A}_i)}\left(\ln p(\mathbf{r}_p^q, \mathcal{A}) \right) \right\}, \end{aligned} \tag{93}$$

where the expectation is with respect to $\sim q(\mathcal{A}_i)$, i.e., all $q(\mathcal{A}), j \neq i$. The quantity Z is the normalization factor required for obtaining a valid probability distribution.

APPENDIX B
COMPUTATION OF $\mathbb{E}_{q(\mathbf{r}_p)}(\mathbf{r}_p)$

Consider the real and imaginary parts of the k^{th} element of $\mathbf{r}_p$ given as $(\mathcal{R}\{r_{p,k}\})$, $(\mathcal{I}\{r_{p,k}\})$. Their expected values with respect to $q(\mathbf{r}_p)$ can be formulated as [44] and [48],

$$\begin{aligned} &\mathbb{E}_{q(\mathbf{r}_p)}(\mathcal{R}\{r_{p,k}\}) \\ &= \mathcal{R}\left\{ \left[\boldsymbol{\Psi}_{\boldsymbol{p}} \mathbb{E}_{q(\mathbf{h})}(\mathbf{h})\right]_k \right\} + \frac{\sigma_w}{\sqrt{2}} \mathcal{F}(\alpha_{real}, \beta_{real}), \\ &\mathbb{E}_{q(\boldsymbol{r}_p)}(\mathcal{I}\{r_{p,k}\}) \\ &= \mathcal{I}\left\{ \left[\boldsymbol{\Psi}_{\boldsymbol{p}} \mathbb{E}_{q(\mathbf{h})}(\mathbf{h})\right]_k \right\} + \frac{\sigma_w}{\sqrt{2}} \mathcal{F}(\alpha_{imag}, \beta_{imag}). \end{aligned}$$

where $\mathcal{F}(\alpha, \beta) = \frac{\phi(\alpha)-\phi(\beta)}{\Phi(\beta)-\Phi(\alpha)}$, $\alpha_{\mathcal{C}}$ and $\beta_{\mathcal{C}}$ are given as

$$\alpha_{\mathcal{C}} = \frac{\sqrt{2}}{\sigma_w}\left(\mathcal{C}\{l_k\} - \mathcal{C}\left\{ \left[\boldsymbol{\Psi}_p \mathbb{E}_{q(\mathbf{h})}(\mathbf{h})\right]_k \right\} \right),$$

$$\beta_{\mathcal{C}} = \frac{\sqrt{2}}{\sigma_w}\left(\mathcal{C}\{u_k\} - \mathcal{C}\left\{ \left[\boldsymbol{\Psi}_p \mathbb{E}_{q(\mathbf{h})}(\mathbf{h})\right]_k \right\} \right)$$ and $\phi(\cdot)$ denotes the standard normal distribution's probability density function, and its cumulative distribution function is given by $\Phi(\cdot)$. Combining the real and imaginary parts to obtain the complex quantity $r_{p,k}$ leads to the result $\mathbb{E}_{q(\mathbf{r}_p)}(r_{p,k}) = \mathbb{E}_{q(\mathbf{r}_p)}(\mathcal{R}\{r_{p,k}\}) + \mathrm{j}\mathbb{E}_{q(\mathbf{r}_p)}(\mathcal{I}\{r_{p,k}\})$. Upon stacking the values for $k = 1, \ldots, MN$, we obtain the quantity

$$\mathbb{E}_{q(\mathbf{r}_p)}(\mathbf{r}_p) = \left[\mathbb{E}_{q(\boldsymbol{r}_p)}(r_{p,1}), \mathbb{E}_{q(\boldsymbol{r}_p)}(r_{p,2}), \ldots, \mathbb{E}_{q(\boldsymbol{r}_p)}(r_{p,MN})\right]^T. \tag{94}$$

REFERENCES

[1] L. Hanzo, Y. Akhtman, J. Akhtman, L. Wang, and M. Jiang, *MIMO-OFDM for LTE, WiFi and WiMAX: Coherent Versus Non-Coherent and Cooperative Turbo Transceivers*. Hoboken, NJ, USA: Wiley, 2011.

[2] T. Strohmer and S. Beaver, "Optimal OFDM design for time-frequency dispersive channels," *IEEE Trans. Commun.*, vol. 51, no. 7, pp. 1111–1122, Jul. 2003.

[3] S. A. Matin and L. B. Milstein, "OFDM system performance, variability and optimality with design imperfections and channel impediments," *IEEE Trans. Veh. Technol.*, vol. 70, no. 1, pp. 381–397, Jan. 2021.

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

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

[6] S. S. Das and R. Prasad, *OTFS: Orthogonal Time Frequency Space Modulation A Waveform for 6G*, 1st ed. New York, NY, USA: River, Sep. 2022, p. 220, doi: 10.1201/9781003339021.

[7] T. Thaj, E. Viterbo, and Y. Hong, "Orthogonal time sequency multiplexing modulation: Analysis and low-complexity receiver design," *IEEE Trans. Wireless Commun.*, vol. 20, no. 12, pp. 7842–7855, Dec. 2021.

[8] T. Thaj and E. Viterbo, "Orthogonal time sequency multiplexing modulation," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2021, pp. 1–7.

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

[10] A. Farhang, A. RezazadehReyhani, L. E. Doyle, and B. Farhang-Boroujeny, "Low complexity modem structure for OFDM-based orthogonal time frequency space modulation," *IEEE Wireless Commun. Lett.*, vol. 7, no. 3, pp. 344–347, Jun. 2018.

[11] A. Mehrotra, R. K. Singh, S. Srivastava, and A. K. Jagannatham, "Channel estimation techniques for CP-aided OTFS systems relying on practical pulse shapes," in *Proc. IEEE Int. Conf. Signal Process. Commun. (SPCOM)*, Jul. 2022, pp. 1–5.

[12] A. Fish, S. Gurevich, R. Hadani, A. M. Sayeed, and O. Schwartz, "Delay-Doppler channel estimation in almost linear complexity," *IEEE Trans. Inf. Theory*, vol. 59, no. 11, pp. 7632–7644, Nov. 2013.

[13] P. Raviteja et al., "Embedded pilot-aided channel estimation for OTFS in delay-Doppler channels," *IEEE Trans. Veh. Technol.*, vol. 68, no. 5, pp. 4906–4917, May 2019.

[14] A. Mehrotra, S. Srivastava, S. Asifa, A. K. Jagannatham, and L. Hanzo, "Online Bayesian learning-aided sparse CSI estimation in OTFS modulated MIMO systems for ultra-high-Doppler scenarios," *IEEE Trans. Commun.*, vol. 72, no. 4, pp. 2182–2200, Apr. 2024.

[15] M. Li et al., "A new path division multiple access for the massive MIMO-OTFS networks," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 4, pp. 903–918, Apr. 2021.

[16] H. Li and Q. Yu, "Doubly-iterative sparsified MMSE turbo equalization for OTFS modulation," *IEEE Trans. Commun.*, vol. 71, no. 3, pp. 1336–1351, Mar. 2023.

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

[18] A. Mehrotra, S. Srivastava, A. K. Jagannatham, and L. Hanzo, "Data-aided CSI estimation using affine-precoded superimposed pilots in orthogonal time frequency space modulated MIMO systems," *IEEE Trans. Commun.*, vol. 71, no. 8, pp. 4482–4498, Aug. 2023.

[19] L. Zhao, W.-J. Gao, and W. Guo, "Sparse Bayesian learning of delay-Doppler channel for OTFS system," *IEEE Commun. Lett.*, vol. 24, no. 12, pp. 2766–2769, Dec. 2020.

[20] S. Srivastava, R. K. Singh, A. K. Jagannatham, and L. Hanzo, "Bayesian learning aided simultaneous row and group sparse channel estimation in orthogonal time frequency space modulated MIMO systems," *IEEE Trans. Commun.*, vol. 70, no. 1, pp. 635–648, Jan. 2022.

[21] Y. Shan, F. Wang, Y. Hao, J. Yuan, J. Hua, and Y. Xin, "Off-grid channel estimation using grid evolution for OTFS systems," *IEEE Trans. Wireless Commun.*, vol. 23, no. 8, pp. 9549–9565, Aug. 2024.

[22] Z. Wei et al., "Off-grid channel estimation with sparse Bayesian learning for OTFS systems," *IEEE Trans. Wireless Commun.*, vol. 21, no. 9, pp. 7407–7426, Sep. 2022.

[23] Q. Wang, Y. Liang, Z. Zhang, and P. Fan, "2D off-grid decomposition and SBL combination for OTFS channel estimation," *IEEE Trans. Wireless Commun.*, vol. 22, no. 5, pp. 3084–3098, May 2023.

[24] 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 Trans. Veh. Technol.*, vol. 70, no. 8, pp. 8343–8348, Aug. 2021.

[25] L. Zhao, J. Yang, Y. Liu, and W. Guo, "Block sparse Bayesian learning-based channel estimation for MIMO-OTFS systems," *IEEE Commun. Lett.*, vol. 26, no. 4, pp. 892–896, Apr. 2022.

[26] X. Cheng, B. Xia, K. Xu, and S. Li, "Bayesian channel estimation and data detection in oversampled OFDM receiver with low-resolution ADC," *IEEE Trans. Wireless Commun.*, vol. 20, no. 9, pp. 5558–5571, Sep. 2021.

[27] C. Studer and G. Durisi, "Quantized massive MU-MIMO-OFDM uplink," *IEEE Trans. Commun.*, vol. 64, no. 6, pp. 2387–2399, Jun. 2016.

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

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

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

[31] V. Kumar Singh, M. F. Flanagan, and B. Cardiff, "Maximum likelihood channel path detection and MMSE channel estimation in OTFS systems," in *Proc. IEEE 92nd Veh. Technol. Conf. (VTC-Fall)*, Nov. 2020, pp. 1–5.

[32] S. Li, "Hybrid MAP and PIC detection for OTFS modulation," *IEEE Trans. Veh. Technol.*, vol. 70, no. 7, pp. 7193–7198, Jul. 2021.

[33] S. Tiwari, S. S. Das, and V. Rangamgari, "Low complexity LMMSE receiver for OTFS," *IEEE Commun. Lett.*, vol. 23, no. 12, pp. 2205–2209, Dec. 2019.

[34] A. Mezghani, M.-S. Khoufi, and J. A. Nossek, "A modified MMSE receiver for quantized MIMO systems," in *Proc. ITG/IEEE WSA*, Jan. 2010, pp. 1–5.

[35] T. M. Lok and V. K.-W. Wei, "Channel estimation with quantized observations," in *Proc. IEEE Int. Symp. Inf. Theory*, Aug. 1998, p. 333.

[36] A. Mezghani, F. Antreich, and J. A. Nossek, "Multiple parameter estimation with quantized channel output," in *Proc. Int. ITG Workshop Smart Antennas (WSA)*, Feb. 2010, pp. 143–150.

[37] Y. Ding, S. Chiu, and B. D. Rao, "Bayesian channel estimation algorithms for massive MIMO systems with hybrid analog-digital processing and low-resolution ADCs," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 3, pp. 499–513, Jun. 2018.

[38] P. Raviteja, Y. Hong, E. Viterbo, and E. Biglieri, "Practical pulse-shaping waveforms for reduced-cyclic-prefix OTFS," *IEEE Trans. Veh. Technol.*, vol. 68, no. 1, pp. 957–961, Jan. 2018.

[39] H.-T. Sheng and W.-R. Wu, "Time-frequency domain channel estimation for OTFS systems," *IEEE Trans. Wireless Commun.*, vol. 23, no. 2, pp. 937–948, Feb. 2024.

[40] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

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

[42] C.-K. Wen, C.-J. Wang, S. Jin, K.-K. Wong, and P. Ting, "Bayes-optimal joint channel-and-data estimation for massive MIMO with low-precision ADCs," *IEEE Trans. Signal Process.*, vol. 64, no. 10, pp. 2541–2556, May 2016.

[43] J. Winn and C. M. Bishop, "Variational message passing," *J. Mach. Learn. Res.*, vol. 6, pp. 661–694, Apr. 2005.

[44] Z. Yang, L. Xie, and C. Zhang, "Variational Bayesian algorithm for quantized compressed sensing," *IEEE Trans. Signal Process.*, vol. 61, no. 11, pp. 2815–2824, Jun. 2013.

[45] C. M. Bishop and M. Tipping, "Variational relevance vector machines," 2013, *arXiv:1301.3838*.

[46] *Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) Radio Transmission and Reception*, document ETSI TS 136 300 V11.6.0, 3GPP Technical Specification Group Radio Access Network, 2013. [Online]. Available: https://www.3gpp.org

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

[48] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, vol. 1, 2nd ed., New York, NY, USA: Wiley, Oct. 1994.

Anand Mehrotra (Graduate Student Member, IEEE) received the B.Tech. and M.Tech. degrees in electronics and communication engineering from Dr. A.P.J. Abdul Kalam Technical University. He is currently pursuing the Ph.D. degree in electrical engineering with Indian Institute of Technology Kanpur, under the supervision of Prof. Aditya Jagannatham. With over three years of teaching experience, his research explores cutting-edge advancements in communication technologies, particularly innovations in 5G and beyond, aiming to make impactful contributions to the theory and applications of next-generation wireless networks.

Suraj Srivastava (Member, IEEE) received the Ph.D. degree in electrical engineering from Indian Institute of Technology Kanpur, India, and the M.Tech. degree in electronics and communication engineering from Indian Institute of Technology Roorkee, India. From October 2022 to November 2023, he was a Senior Lead Engineer with Qualcomm India Pvt. Ltd., where he was involved in the development of 5G-IoT modems. Previously, he was a Staff-I Systems Design Engineer with Broadcom Research India Pvt. Ltd. He was a Lead Engineer with Samsung Research India, where he worked on designing layer-2 of the 3G UMTS/WCDMA/HSDPA modems. In December 2023, he joined the Department of Electrical Engineering, Indian Institute of Technology Jodhpur, where he is currently an Assistant Professor. His research interests include quantum communications and computing, sparse signal processing in 5G wireless systems, mmWave and tera-hertz communication, OTFS, RadCom, optimization and machine learning, low-resolution MIMO communications, and IRS technology. He is a member of IET, IETE, and INAE-YA. He was awarded the Qualcomm Innovation Fellowship (QIF) twice for the years 2018 and 2022 and the Outstanding Ph.D. Thesis Award and the Silver Medal from IIT Kanpur.

N. Shanmughanadha Reddy received the bachelor's degree in electronics and communication engineering from Sri Venkateswara University, Tirupati, in 2019, and the master's degree in signal processing and communication from IIT Kanpur in 2023. Currently, he is a Software Engineer with Rakuten Mobile, Tokyo, Japan. His research interests include wireless modern communications, networking, and AI in communications.

Aditya Jagannatham (Senior Member, IEEE) received the bachelor's degree from Indian Institute of Technology, Bombay, and the M.S. and Ph.D. degrees from the University of California at San Diego, San Diego, CA, USA. From April 2007 to May 2009, he was employed as a Senior Wireless Systems Engineer with Qualcomm Inc., San Diego, where he was a part of the Qualcomm CDMA Technologies Division. He is currently a Professor with the Department of Electrical Engineering, IIT Kanpur, where he also holds the Arun Kumar Chair Professorship. His research interests include next-generation wireless cellular and WiFi networks, with a special emphasis on various 5G technologies, such as massive MIMO, mmWave MIMO, FBMC, NOMA, as well as emerging 6G technologies, such as OTFS, IRS, THz systems, and VLC. He has been twice awarded the P.K. Kelkar Young Faculty Research Fellowship for excellence in research, received multiple Qualcomm Innovation Fellowships in 2018 and 2022, the IIT Kanpur Excellence in Teaching Award, the CAL(IT)2 Fellowship at the University of California at San Diego, the Upendra Patel Achievement Award at Qualcomm San Diegom and the Qualcomm 6G UR India Gift.

Lajos Hanzo (Life Fellow, IEEE) received the Honorary Doctorate degree from the Technical University of Budapest in 2009 and the Honorary Doctorate degree from Edinburgh University in 2015. He is a Foreign Member of Hungarian Science Academy; a fellow of the Royal Academy of Engineering (FREng), IET, and EURASIP; and holds the IEEE Eric Sumner Technical Field Award. For further details please see http://www-mobile.ecs.soton.ac.uk, https://en.wikipedia.org/wiki/Lajos_Hanzo.