Angularly Sparse Channel Estimation in Dual-Wideband Tera-Hertz (THz) Hybrid MIMO Systems Relying on Bayesian Learning

Abhisha Garg *Graduate Student Member, IEEE,* Suraj Srivastava, *Member, IEEE,* Nimish Yadav, Aditya K. Jagannatham, *Senior Member, IEEE* and Lajos Hanzo, *Life Fellow, IEEE*

Abstract—Bayesian learning aided massive antenna array based THz MIMO systems are designed for *spatial-wideband* and *frequency-wideband* scenarios, collectively termed as the *dual-wideband* channels. Essentially, numerous antenna modules of the THz system result in a significant delay in the transmission/ reception of signals in the time-domain across the antennas, which leads to spatial-selectivity. As a further phenomenon, the wide bandwidth of THz communication results in substantial variation of the effective angle of arrival/ departure (AoA/ AoD) with respect to the subcarrier frequency. This is termed as the *beam squint effect*, which renders the channel state information (CSI) estimation challenging in such systems. To address this problem, initially, a pilot-aided (PA) Bayesian learning (PA-BL) framework is derived for the estimation of the Terahertz (THz) MIMO channel that relies exclusively on the pilot beams transmitted. Since the framework designed can successfully operate in an ill-posed model, it can verifiably lead to reduced pilot transmissions in comparison to conventional methodologies. The above paradigm is subsequently extended to additionally incorporate data symbols to derive a Data-Aided (DA) BL approach that performs joint data detection and CSI estimation. We will demonstrate that it is capable of improving the dual-wideband channel's estimate, despite further reducing the training overhead. The Bayesian Cramér-Rao bounds (BCRLBs) are also obtained for explicitly characterizing the lower bounds on the mean squared error (MSE) of the PA-BL and DA-BL frameworks. Our simulation results show the improved normalized MSE (NMSE) and bit-error rate (BER) performance of the proposed estimation schemes and confirm that they approach their respective BCRLB benchmarks.

Index Terms—Terahertz, MIMO, dual-wideband, pilot-aided, data-aided, beam squint, Bayesian learning

I. INTRODUCTION

Terahertz (THz) communication, relying on the (0.1-10) THz band, has the promise of achieving the ultra-high data rates targeted by next generation networks [1]. When compared to the mmWave band that ranges from (30-100) GHz, the THz band can provide spectral windows with bandwidths in the range of 10 GHz or even higher [2]. This can in turn lead to the development of rapid data transfer capabilities links with rates in the range of several hundreds of Gbps [3]. However, it must be noted that the successful achievement of THz communication is enormously challenging, as the THz signal suffers from high atmospheric losses ranging from 5dB to 100dB [4] and path losses ranging from 70dB to 140 dB [2], which can be attributed to the absorption by gases and the high carrier frequencies, respectively. In addition, THz signals are prone to blockage by physical barriers such as walls, human bodies etc [5]. Multiple-input multiple-output (MIMO) technology, which is capable of generating narrow directional beams having high array gains, is extremely well-suited for mitigating the losses occurring in the THz band. Fortuitously, numerous antenna elements can be embedded at both the transceiver ends in a THz system within a small form-factor [6]. At this juncture, it should be noted that the conventional transceiver architecture that necessitates a separate radio frequency (RF) chain per antenna can lead to a an excessive increase in the hardware complexity required to carry out the signal processing operations pertaining to modulation and demodulation [7]. A popular solution recommended in the latest research literature [8], [9] to overcome this impediment is the novel hybrid signal processing architecture, wherein the signal processing operations are effectively split between the RF and baseband stages. This is in stark contrast to an orthodox transceiver design wherein the entire signal processing burden is borne by the baseband processor. The hybrid architecture coupled with large MIMO arrays can enormously increase the signal gain via optimal beamforming, while also limiting the hardware complexity through joint RF and baseband signal processing.

Moreover, a typical wideband THz channel is frequency-selective in nature, due to the multipath induced delay spread, which is a characteristic feature of wireless signal propagation [10]. This phenomenon is termed as frequency-wideband effect. A pair of crucial factors contributing to frequency selectivity in the THz band, even for LoS scenarios, are high atmospheric absorption and free-space path losses [11]. The heightened atmospheric absorption primarily arises from water vapor molecules. Furthermore, due to frequency-selective nature of the absorption coefficient, the THz band is char-

acterized by ultra-wideband transmission windows (TWs) and absorption coefficient peak regions (ACPRs) [2]. Additionally, the presence of ACPRs in the THz band leads to pulse broadening in the time domain, presenting practical communication challenges within this frequency range. Owing to the short wavelength of the THz signal, a large antenna array leads to a progressively increasing phase shift in the received/ transmitted signals across the constituent antennas, corresponding to commensurate delays in the time domain [12]. This is termed as the spatial-wideband effect. This leads to the frequency dependent nature of the array response [13]. A general THz system that is exposed to both frequency-wideband and spatial-wideband effects is termed as a dual-wideband system, which is considered in our work. Furthermore, a practical wideband THz system suffers from the beam squint problem that can be attributed to the frequency dependent nature of the array response [13]. The availability of accurate channel state information (CSI), which plays a key role in beamforming, is of paramount importance in such systems. However, conventional techniques such as the least-squares (LS) and linear minimum mean squared error (LMMSE) estimators, require an excessive pilot overhead, owing to the large number of antennas, which reduces the spectral efficiency (SE). Thus, efficient CSI estimation techniques, which exploit the properties of the dual-wideband THz MIMO channel, hold the key for the successful realization of reliable high speed communication in these systems which also considers the beam squint effect. A concise overview of the relevant prior research is provided next.

A. Review of the existing literature

Jornet and Akyildiz [14] developed a novel THz channel model accounting also for the molecular absorption and path loss, which was successfully leveraged to determine the capacity of THz systems. In [15], the principles of molecular absorption were combined with radioactive transfer theory to evaluate the path loss of THz signals for propagation distances in the range of a centimeter to a few meters, followed by the formulation of the THz channel model. Lin et al., [16], developed a general array-of-subarrays architecture for a hybrid THz MIMO system, complemented by an innovative algorithm for adaptive distance-aware beamforming, which requires knowledge of the CSI. Therefore, it is important to note at this juncture that the end-to-end performance of a THz band hybrid MIMO system is crucially contingent on the availability of accurate CSI. This motivates the design of superior channel estimation techniques for THz MIMO applications, which can potentially maximize the performance. The well-known traditional channel estimation approaches, such as the LS and MMSE, perform poorly in the THz regime owing to their exclusive pilot overhead, which consequently leads to a low spectral efficiency [17]. Nevertheless, a significant point to note is that the THz frequency range exhibits a strong directional propagation characteristic owing to using a large antenna array, hence resulting in a sparsely populated multi-path channel in the angular domain. Naturally, sparse signal recovery techniques, which have gained popularity in recent research, can lead to verifiably improved CSI estimates in THz

MIMO transceivers. In this context, the authors of [18] design a superimposed pilot modulation technique, followed by sparse Bayesian learning (SBL)-based approaches for the estimation of the frequency-selective wideband channel. More specifically, they proposed a multiple-measurement vector (MMV)-based SBL algorithm for joint CSI estimation and data-detection. Venugopal et al., in their cutting-edge work [19], proposed techniques for the estimation of a frequency-selective mmWave wideband channel in both single-carrier frequency domain equalization (SC-FDE) and orthogonal frequency-division multiplexing (OFDM)-based mmWave systems. In their work, the CSI estimation model was formulated as a sparse recovery problem, which was subsequently solved via a time-frequency estimation algorithm. The authors of [20] developed a block least-mean squares (BLMS) sparse adaptive framework for estimating the channel in a mmWave MIMO SC-FDE system. As a further advance, Gao et al., in their pioneering treatise [21], proposed a sparse beamspace CSI estimation technique for the estimation of a wideband mmWave MIMO system having lens antenna arrays. The novel scheme proposed in their work is based on successive support detection (SSD), which is inspired by the successive interference cancellation (SIC) principle of multi-user systems. Wang et al. [22] conceived new CSI estimation techniques for both frequency division duplex (FDD) as well as time division duplex (TDD) mmWave massive MIMO systems. Their framework also considers the spatial-wideband and frequency-wideband effects inherent in such systems and exploits the sparsity both in the angular and delay domains, while comprehensively covering both the uplink as well as the downlink. An attractive sensing-based high-resolution technique was designed in [23] that can successfully extract the physical parameters of the uplink CSI, viz., AoAs, time delays and complex channel gains, while accounting for the beam squint effect in mmWave hybrid MIMO systems. This is achieved by adaptively updating the dictionary matrix. A few of the latest treatises have also considered the impact of the spatial-wideband and frequency-wideband effects in massive MIMO THz channel estimation, which is of considerable practical importance. Dovelos et al. [24] model an OFDM-based wideband THz system using a uniform planar array (UPA) under the spatial-frequency wideband effect. A sparse beamspace representation of the wideband THz channel is formulated in their work, followed by the application of the orthogonal matching pursuit (OMP) and the generalized simultaneous OMP (GSOMP) algorithm popular for the acquisition of the sparse channel. Chou et al., [25] addressed the spatial-wideband and frequency-wideband effect with respect to the estimation of a time-varying MIMO-OFDM channel in sub-Terahertz massive MIMO systems. The authors in their cutting-edge work developed a multiple measurement vector least squares compressed sensing (MMV-LS-CS) based CSI estimation technique, which utilizes the MMV-LS to recover the channel employing the previous channel support and the MMV-CS to detect any time-varying components in the residual signal. Although their pioneering work incorporates time-selectivity in sub-THz systems, no joint channel estimation and data detection is considered, which can lead to a remarkable enhancement in performance.

Table I
BOLDLY CONTRASTING OUR CONTRIBUTION TO THE LITERATURE

Features	[6]	[8]	[9]	[13]	[17]	[22]	[24]	[25]	[26]	**This paper**
THz hybrid MIMO	✓						✓	✓		✓
Dual-Wideband effect			✓			✓		✓		✓
Reflection losses	✓						✓	✓		✓
Molecular absorption losses	✓						✓	✓		✓
Diffused-ray modeling	✓	✓								✓
BCRLB	✓	✓		✓			✓		✓	✓
SC-FDE system										✓
Data-Aided CSI acquisition										✓

To fill this knowledge gap, we conceive a novel DA Bayesian learning framework for reliable sparse CSI estimation and data-detection in a THz MIMO system, while considering also the dual-wideband effect. In Table I we boldly contrast our contributions to the literature. The next section briefly describes the main contributions of our work in more details.

B. Contributions of this work

1) A hybrid frequency-dependent THz MIMO channel model is developed by considering the dual-wideband effects, which additionally incorporates the absorption, reflection as well as free space losses for a uniform linear array (ULA) at the base station (BS) serving a single user. Notably, our THz channel model incorporates the spatial wideband effect to construct a frequency-dependent model of the array response, while also considering the frequency-wideband effect, which accounts for multipath propagation. Existing literature on THz channel estimation, such as [24] and [25], have solely focused on the spatial-wideband effect and did not include the frequency-wideband effect by considering multipath delays and the sampled version of transmit pulse shaping filter response. Therefore, the novel contribution of our work lies in considering the dual wideband effects, which are missing in the existing literature such as [27], [28]. Moreover, an efficient zero-padded (ZP) block structure is proposed to transform the time domain (TD) signal processing problem into an equivalent frequency domain (FD) one, which remarkably reduces the complexity of THz MIMO CSI estimation as well as signal detection. Additionally, this novel transmission structure requires very few pilot beams to excite the numerous angular modes of the channel.

2) A sparse signal recovery problem is formulated for the THz MIMO channel. A pair novel pilot-aided Bayesian learning (PA-BL) and data-aided Bayesian learning (DA-BL)-based methodologies are proposed for estimating the sparse beamspace dual-wideband channel. Since it can operate in ill-posed scenerios, it requires significantly fewer pilot transmissions in comparison to the traditional LS and MMSE schemes.

3) The Bayesian Cramér-Rao lower bounds (BCRLB) are derived for the mean-square error (MSE) of both the PA and DA estimators proposed in the THz regime. These serve as excellent benchmarks to gauge the performance of the proposed algorithm.

4) Simulation results are presented for quantifying the CSI estimation and data-detection performance achieved in the dual-wideband THz MIMO system considered. It is shown that the normalized MSEs (NMSEs) of the PA-BL and DA-BL algorithms approach their respective BCRLBs, which demonstrates their efficiency. This is especially significant considering the fact that they do not require prior knowledge of either the statistics or of the support of the sparse CSI.

C. Organization of the paper

The paper is structured as follows. Section II derives the system model of a hybrid massive MIMO system in the THz regime utilizing the SC-FDE framework. Section III motivates and models the dual-wideband effect in the above system and also characterizes the loss induced by the beam squint effect. Section IV determines the model for sparse CSI estimation in the wideband THz system. Section V describes the PA-BL approach proposed for CSI estimation, with the requisite mathematical foundation. Next, Section VI chronicles the DA-BL approach, which yields an improved performance over the previous paradigm. Section VII outlines our simulation results, while Section VIII concludes the paper.

D. Notation

The following notation is used throughout this paper. Letters in upper-case boldface such as $\mathbf{A}$ represent matrices, while letters in lower-case boldface denote vectors. Superscripts $(.)^T$, $(.)^*$, $(.)^{-1}$, $(.)^H$, and $(.)^\dagger$ represent the transpose, conjugate, inverse, Hermitian and pseudo-inverse operators, respectively. The quantity blkdiag $(\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_N)$ represents the block-diagonal matrix with matrices $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_N$ on the principal diagonal, which are not necessarily square. The vectorization operation that stacks the columns of any matrix $\mathbf{A}$ is represented by vec$(\mathbf{A})$, while the inverse of the same operation is represented by vec$^{-1}(\mathbf{a})$. The operators $*$ and $\otimes$ represent the linear convolution and Kronecker product, respectively, while Tr$(.)$ represents the matrix trace operator. The following standard property of the vec$(.)$ operation, vec$(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A})vec(\mathbf{B})$, is also used in the paper. The operators $\| \cdot \|_2$ and $\| \cdot \|_F$ represent the l_2-norm and Frobenius norm, respectively, of vectors and matrices. The expectation operator is denoted by $\mathbb{E}\{.\}$, while the absolute value is denoted by $| \cdot |$; the symmetric complex Gaussian distribution with mean $\boldsymbol{\mu}$ and variance $\boldsymbol{\Sigma}$ is represented by $\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.
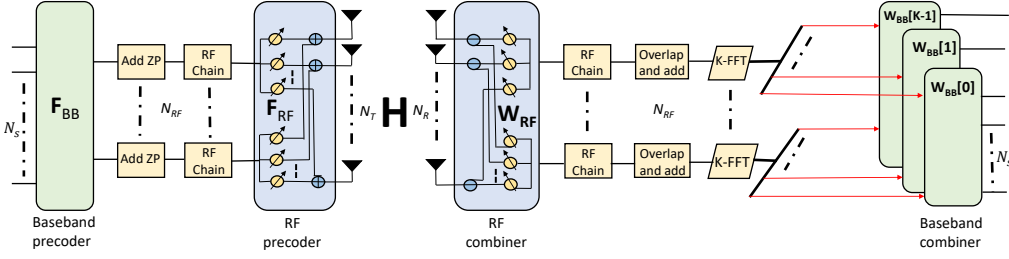
Figure 1. Schematic representation of SC-FDE-based wideband THz hybrid MIMO system

II. SC-FDE BASED THz HYBRID MIMO SYSTEM MODEL

Consider a single carrier wideband THz hybrid MIMO system that transmits N_s data streams from a transmitter equipped with N_T transmit antennas (TAs), to a receiver having N_R receive antennas (RAs). There are N_{RF} RF chains at the transmitter and the receiver, where $N_s \leq N_{RF} \ll \min(N_T, N_R)$. The THz hybrid MIMO transmitter is equipped with a baseband precoder in cascade with an RF precoder, represented by the matrices $\mathbf{F}_{BB} \in \mathbb{C}^{N_{RF} \times N_s}$ and $\mathbf{F}_{RF} \in \mathbb{C}^{N_T \times N_{RF}}$, respectively. In a similar manner, at the RAs, the hybrid processing module is comprised of an RF beamformer $\mathbf{W}_{RF} \in \mathbb{C}^{N_R \times N_{RF}}$ in cascade with a baseband combiner $\mathbf{W}_{BB} \in \mathbb{C}^{N_{RF} \times N_s}$. The hybrid signal processing architecture of the THz MIMO system is shown in Fig. 1. Since the RF transmit precoder (TPC) matrix $\mathbf{F}_{RF}$ and RF receiver combiner (RC) matrix $\mathbf{W}_{RF}$ are constrained to exclusively comprise digitally controlled phase-shifters, having elements obeying $|\mathbf{F}_{RF}(i,j)| = \frac{1}{\sqrt{N_T}}$, $|\mathbf{W}_{RF}(i,j)| = \frac{1}{\sqrt{N_R}}$ [19]. Furthermore, $\mathbf{F}_{RF}$ and $\mathbf{W}_{RF}$ can be represented as $\mathbf{F}_{RF} = [\mathbf{f}_1, \mathbf{f}_2, \cdots, \mathbf{f}_{N_{RF}}]$ and $\mathbf{W}_{RF} = [\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_{N_{RF}}]$, respectively. Considering a frequency-selective wideband channel in the THz band comprised of L delay taps, the lth delay tap can be expressed by the complex channel matrix $\mathbf{H}_l$, $0 \leq l \leq L - 1$, of size $N_R \times N_T$. The symbol vector that is transmitted at time instant t is denoted by $\mathbf{s}(t) \in \mathbb{C}^{N_s \times 1}$. After baseband combining, the output signal vector at the tth time instant is given by

$$\check{\mathbf{y}}(t) = \mathbf{W}_{BB}^H \mathbf{W}_{RF}^H [\mathbf{H}_t * \mathbf{x}(t)] + \mathbf{W}_{BB}^H \mathbf{W}_{RF}^H \mathbf{v}(t), \quad (1)$$

where $\mathbf{x}(t) = \mathbf{F}_{RF} \mathbf{F}_{BB} \mathbf{s}(t) \in \mathbb{C}^{N_T \times 1}$ represents the precoded symbol vector and $\mathbf{v}(t) \in \mathbb{C}^{N_R \times 1}$ represents the additive noise vector that follows the distribution $\mathcal{CN}(\mathbf{0}_{N_R \times 1}, \sigma_v^2 \mathbf{I}_{N_R})$. The operator $*$ represents linear convolution. The equivalent FD model of this system is developed next.

A. SC-FDE channel estimation model

Again Fig. 1 depicts the architecture of a hybrid THz MIMO system relying on FD signal processing. Each transmit frame duration in the system is divided into training and data phases, as shown in Fig. 2. The training phase is comprised of M blocks with each individual block consisting of N_p pilot vectors. Let $\mathbf{u}_m^{(p)} \in \mathbb{C}^{N_{RF} \times 1}$ denote the pth complex pilot vector, which follows the relationship $0 \leq p \leq N_p - 1$ for every mth block [20]. Prior to transmitting the pilot sequence on an individual RF chain, $L - 1$ zeros are appended in each

mth block to generate a zero-padded (ZP) block of length $K = N_p + L - 1$. Similarly, zero-matrices of size $N_R \times N_T$ are appended to the MIMO channel taps $\mathbf{H}_l$ to obtain a block of length K. Thus, the zero-padded pilot vectors and THz MIMO channel taps can be described in a compact fashion as

$$\left\{ \mathbf{u}_m^{(q)} \right\}_{q=0}^{K-1} = \left\{ \mathbf{u}_m^{(0)}, \mathbf{u}_m^{(1)}, \cdots, \mathbf{u}_m^{(N_p-1)}, \underbrace{\mathbf{0}, \cdots, \mathbf{0}}_{L-1} \right\} \text{ and}$$

$$\{\mathbf{H}_q\}_{q=0}^{K-1} = \left\{ \mathbf{H}_0, \mathbf{H}_1, \cdots, \mathbf{H}_{L-1}, \underbrace{\mathbf{0}, \cdots, \mathbf{0}}_{N_p-1} \right\}, \text{ respectively.}$$

For each block m, the training RF TPC and RC are represented as $\mathbf{F}_{RF,m} \in \mathbb{C}^{N_T \times N_{RF}}$ and $\mathbf{W}_{RF,m} \in \mathbb{C}^{N_R \times N_{RF}}$, respectively. Thus, the output signal vectors $\widetilde{\mathbf{y}}_m(q) \in \mathbb{C}^{N_{RF} \times 1}$ corresponding to the mth block follow the model given by

$$\widetilde{\mathbf{y}}_m(q) = \mathbf{W}_{RF,m}^H (\mathbf{H}_q \circledast_K \widetilde{\mathbf{u}}_m^{(q)}) + \mathbf{W}_{RF,m}^H \mathbf{v}_m(q). \quad (2)$$

There $\widetilde{\mathbf{u}}_m^{(q)} = \mathbf{F}_{RF,m} \mathbf{u}_m^{(q)}$, $\mathbf{v}_m(q)$ denotes the complex white AWGN noises which follows the distribution $\mathcal{CN}(\mathbf{0}_{N_r \times 1}, \sigma_v^2 \mathbf{I}_{N_r})$, and $\circledast_K$ is the circular convolution operation between sequences of length K. Consider the mth TD output block defined as $\{\widetilde{\mathbf{y}}_m(q)\}_{q=0}^{K-1} = \{\widetilde{\mathbf{y}}_m(0), \widetilde{\mathbf{y}}_m(1), \cdots, \widetilde{\mathbf{y}}_m(K-1)\}$. At the output, employing the K-point fast fourier transform (FFT) operation, the FD representation that corresponds to the hybrid MIMO system in the THz band may be obtained as

$$\{\mathbf{y}_m[k]\}_{k=0}^{K-1} = \text{FFT}(\{\widetilde{\mathbf{y}}_m(q)\}_{q=0}^{K-1})$$
$$= \{\mathbf{y}_m[0], \mathbf{y}_m[1], \cdots, \mathbf{y}_m[K-1]\}. \quad (3)$$

Thus, the received FD output vector $\mathbf{y}_m[k] \in \mathbb{C}^{N_{RF} \times 1}$ at the kth subcarrier can be expressed as:

$$\mathbf{y}_m[k] = \mathbf{W}_{RF,m}^H \mathbf{H}[k] \mathbf{F}_{RF,m} \mathbf{u}_m[k] + \mathbf{W}_{RF,m}^H \check{\mathbf{v}}_m[k], \quad (4)$$

where $\mathbf{H}[k]$ depicts the THz MIMO channel as frequency response (CFR), whose model is detailed in the next section. The vector $\mathbf{u}_m[k] \in \mathbb{C}^{N_{RF} \times 1}$ denotes the kth FFT point of the pilot sequence $\mathbf{u}_m^{(q)}$, and $\check{\mathbf{v}}_m[k] \in \mathbb{C}^{N_R \times 1}$ represents the kth FFT point of the noise sequence $\mathbf{v}_m^{(q)}$. Upon exploiting the property of vec(.) operator as described in Section-I-C, (4) can be recast as

$$\mathbf{y}_m[k] = \underbrace{(\mathbf{u}_m^T[k] \mathbf{F}_{RF,m}^T \otimes \mathbf{W}_{RF,m}^H)}_{\boldsymbol{\Phi}_m[k] \in \mathbb{C}^{N_{RF} \times N_R N_T}} \mathbf{h}[k] + \mathbf{v}_m[k], \quad (5)$$

where $\mathbf{h}[k] = \text{vec}(\mathbf{H}[k]) \in \mathbb{C}^{N_T N_R \times 1}$ represents the vectorized form of the THz MIMO channel model at the kth subcarrier, $\boldsymbol{\Phi}_m[k]$ represents the sensing matrix for the mth block of
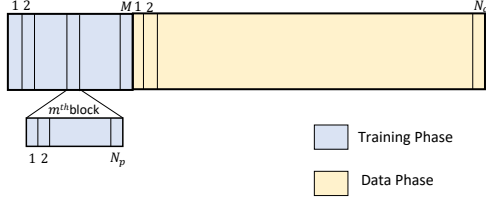
Figure 2. Frame structure employed in the wideband THz hybrid MIMO system using SC-FDE.

the kth subcarrier, while $\mathbf{v}_m[k] = \mathbf{W}_{RF,m}^H \check{\mathbf{v}}_m[k] \in \mathbb{C}^{N_{RF} \times 1}$ is the noise vector at the output of the RF combiner of the kth subcarrier. Note also that,

$$\mathbb{E}\left\{\mathbf{v}_m[k]\mathbf{v}_m^H[k]\right\} = \mathbb{E}\left\{\mathbf{W}_{RF,m}^H \check{\mathbf{v}}_m[k]\left(\mathbf{W}_{RF,m}^H \check{\mathbf{v}}_m[k]\right)^H\right\}$$
$$= \sigma_v^2 K \mathbf{W}_{RF,m}^H \mathbf{W}_{RF,m} = \mathbf{R}_m. \quad (6)$$

The next section describes the THz channel incorporating the dual-wideband effect.

III. THz Channel Model with dual-wideband effect

We now develop an explicit model for the THz MIMO channel incorporating *dual wideband* effect, comprising of the spatial and frequency wideband effects. We commence by describing the spatial wideband effect.

A. Spatial-wideband effect

Let us assume that a signal arrives at a ULA having N identical antennas, indexed as $\{1, 2, \cdots, N\}$, at the azimuth angle θ. The frequency-dependent array steering vector $\mathbf{a}(\theta, f)$ of this system is given by

$$\mathbf{a}(\theta, f) = \frac{1}{\sqrt{N}}\left[1, e^{-j\frac{2\pi}{\lambda}d_c \cos\theta}, \cdots, e^{-j(N-1)\frac{2\pi}{\lambda}d_c \cos\theta}\right]^T. \quad (7)$$

The quantity $\lambda = \frac{c}{f}$ represents the wavelength as a function of frequency and d_c is the fixed inter-antenna spacing, that is typically set as $d_c = \frac{c}{2 \times f_c}$. In general, the transceiver is designed based on the central carrier frequency f_c, which is an approximation of the frequency over the dynamic frequency range. Therefore, the approximation $d_c \approx \frac{\lambda}{2}$ only holds for a narrowband channel, i.e., when the bandwidth B obeys $B \ll f_c$. However, this condition does not hold for a wideband signal spanning a wide frequency range. Let $\triangle\varphi_{n,k}$ denote the spatial phase offset for the nth antenna with the AoA/AoD equal to θ at subcarrier k. Let f_k represent the frequency of the k-th subcarrier that is defined as $f_k \triangleq f_c + \left(k - \frac{K+1}{2}\right)\frac{B}{K}$. From (7), the expression for $\triangle\varphi_{n,k}$ is given by

$$\triangle\varphi_{n,k} = \frac{2\pi}{\lambda_k}(n-1)d_c \cos\theta = \frac{2\pi}{\frac{c}{f_k}}(n-1)\frac{c}{f_c} \times \frac{1}{2}\cos\theta$$
$$= \pi(n-1)\delta_k \cos\theta, \quad (8)$$

where $\delta_k = \frac{f_k}{f_c}$ represents the relative frequency of the subcarrier. Thus, the effective spatial AoA $\hat{\theta}(f_k)$ at subcarrier frequency f_k is given by

$$\hat{\theta}(f_k) = \arccos(\delta_k \cos\theta). \quad (9)$$

Note that the spatial phase offset $\triangle\varphi_{n,k}$ varies by the subcarrier index as seen in Equation (8), which implies that the effective spatial AoA/ AoD $\hat{\theta}(f_k)$ as defined in Equation (9) also varies with respect to the subcarrier index k. Thus, the spatial phase offset depends on the subcarrier frequency. Due to the large antenna aperture in the THz band, the propagation delay between the first and last antenna elements is non-negligible, hence the transmit symbol reaches different antennas with different delays. This is termed the *spatial-wideband* effect. This spatial wideband effect is aggravated in the THz band due to having numerous antenna elements, coupled with the large bandwidth. By the same token, the variation in the frequency across the subcarriers leads to variations in the effective AoA/ AoD across the frequency band, which is known as the *beam squint* effect. Thus, the beam squint effect is the manifestation of the spatial wideband effect in the frequency domain and therefore, the two phenomena are analogous [22], [29]. Moreover, the magnitude of the beam squint is directly proportional to the extent of spatial wideband effect occurring in such systems, as discussed in Theorem-1 of [23]. Note that a majority of the literature ignores the above effect. Additionally, the THz channel experiences the frequency wideband effect as a consequence of the multi-path delay spread. Therefore, a combination of spatial and frequency wideband effects in THz hybrid MIMO systems is termed the dual-wideband effect. From (9), one can further see that the effective spatial AoA/ AoD $\hat{\theta}(f_k)$ varies with respect to the subcarrier frequency, which depends on the relative frequency δ_k. Hence, the normalized array gain $G(\theta, f_k)$ [24] for a wideband THz system for the kth subcarrier and angular direction θ is $G(\theta, f_k) = |\tilde{\mathbf{a}}^H(\theta, f_c)\tilde{\mathbf{a}}(\theta, f_k)|^2$. The next section develops a model for the wideband THz channel suffering from the dual-wideband effect, which arises due to a combination of the spatial-wideband and frequency-wideband effects.

B. THz channel model with dual wideband effect

The array steering vector of the wideband THz channel exhibiting beam squint, characterized by frequency dependent angles, is given by

$$\tilde{\mathbf{a}}(\theta, f_k) = \frac{1}{\sqrt{N}}\left[1, e^{-j\pi\delta_k \cos(\theta)}, \cdots, e^{-j(N-1)\pi\delta_k \cos(\theta)}\right]^T. \quad (10)$$

Thus, the dual wideband channel at the kth subcarrier, where $k = 1, \cdots, K$ can be modeled as

$$\mathbf{H}[k] = \mathbf{H}^{\text{LoS}}[k] + \mathbf{H}^{\text{NLoS}}[k], \quad (11)$$

where the LoS and NLoS components $\mathbf{H}^{\text{LoS}}[k]$, $\mathbf{H}^{\text{NLoS}}[k]$ are defined as

$$\mathbf{H}^{\text{LoS}}[k] = \sqrt{N_T N_R}\alpha_{\text{L}}(f_k, d) \times \beta_{\tau_{\text{L}}} \times A_T A_R \tilde{\mathbf{a}}_R(\phi_{\text{L}}^R, f_k)$$
$$\tilde{\mathbf{a}}_T^H(\phi_{\text{L}}^T, f_k), \quad (12)$$

$$\mathbf{H}^{\text{NLoS}}[k] = \sqrt{\frac{N_T N_R}{N_{\text{NLoS}} N_{ray}}} \sum_{z=1}^{N_{\text{NLoS}}} \sum_{j=1}^{N_{ray}} \alpha_{z,j}(f_k, d) \times \beta_{\tau_{z,j}} \times$$
$$A_T A_R \tilde{\mathbf{a}}_R(\phi_{z,j}^R, f_k)\tilde{\mathbf{a}}_T^H(\phi_{z,j}^T, f_k). \quad (13)$$

The quantity $\beta_{\tau_{z,j}}$ is defined as $\beta_{\tau_{z,j}} = \sum_{l=0}^{K-1} p(lT_s - \tau_{z,j})e^{-j\frac{2\pi kl}{K}}, \forall k,l$. The quantities N_R and N_T represent the number of RAs and TAs, N_{ray} is the number of diffused rays from a scatterer, while N_{NLoS} represents the number of clusters/ dominant scatterers present in the propagation environment. Note that, here a cluster represents a group of rays that arrive at the receiver around the same time and from the same direction [30], which arise due to reflection and scattering from a dominant scatterer. The pulse shaping filter is represented by $p(.)$ and the sampling time by T_s, whereas A_R and A_T denote the RA and TA gains. The quantities $\tau_{(.)}$, $\phi_{(.)}^R$, $\phi_{(.)}^T$, $\alpha(.)$ represent the delay, AoA, AoD and complex path gain for the lth delay tap, the models for which are shown in the next section.

1) Calculation of the complex-path gains: The typical THz wireless channel incurs a high loss owing to various factors such as molecular absorption losses and molecular noise due to water vapour, among others. These aspects can be incorporated in the modeling of the complex-path gains $\alpha_L(f_k, d)$, $\alpha_{z,j}(f_k, d)$, which are a function of both the sub-carrier frequency f_k and transmission distance d. The complex path gain $\alpha_{(.)}(f_k, d)$ can be characterized by its magnitude and angle components as $\alpha_{(.)}(f_k, d) = |\alpha_{(.)}(f_k, d)|e^{j\omega}$. According to [14], the path-gain of the LoS component can be modeled as

$$|\alpha_{\text{L}}(f_k, d)|^2 = L_{\text{spread}}(f_k, d)L_{\text{abs}}(f_k, d), \tag{14}$$

where $L_{\text{spread}}(f_k, d)$ represents the free-space loss, while $L_{\text{abs}}(f_k, d)$ represents the molecular absorption loss [24], at frequency f_k, which are respectively given by

$$L_{\text{spread}}(f_k, d) = \left(\frac{c}{4\pi f_k d}\right)^2, L_{\text{abs}}(f_k, d) = e^{-k_{\text{abs}}(f_k)d}. \tag{15}$$

The quantity $k_{\text{abs}}(f_k)$ represents the molecular absorption coefficient that is given $k_{\text{abs}}(f_k) = \sum_{a,g} k_{\text{abs}}^{a,g}(f_k)$, where $k_{\text{abs}}^{a,g}(f_k)$ represents the measure of radiation absorption by the ath isotopologue of gas g. The parameter $k_{\text{abs}}^{a,g}(f_k)$ can be calculated from the HITRAN database given in [14]. Similarly, for the complex path gain of the zth NLoS component, the magnitude of the jth diffuse-ray is mathematically given by

$$|\alpha_{z,j}(f_k, d)|^2 = \Gamma_{z,j}^2(f_k)L_{\text{spread}}(f_k, d)L_{\text{abs}}(f_k, d), \tag{16}$$

where $\Gamma_{z,j}(f_k)$ represents the first-order reflection coefficient for the jth diffuse ray of the zth NLoS cluster [6], which is defined as the product of the Fresnel Reflection Coefficient $\gamma_{z,j}(f_k)$ and of the Rayleigh roughness factor $\varrho_{z,j}(f_k)$ [31], that are respectively given by

$$\Gamma_{z,j}(f_k) = \gamma_{z,j}(f_k)\varrho_{z,j}(f_k),$$

$$\gamma_{z,j}(f_k) = \frac{Z(f_k)\cos(\theta_{\text{in}_{z,j}}) - Z_0\cos(\theta_{\text{ref}_{z,j}})}{Z(f_k)\cos(\theta_{\text{in}_{z,j}}) + Z_0\cos(\theta_{\text{ref}_{z,j}})},$$

$$\varrho_{z,j}(f_k) = e^{-\frac{1}{2}\left(\frac{4\pi f\sigma\cos\left(\theta_{\text{in}_{z,j}}\right)}{c}\right)^2}, \tag{17}$$

where $\theta_{\text{in}_{z,j}}$ denotes the angle of incidence, $\theta_{\text{ref}_{z,j}} = \sin^{-1}\left(\sin(\theta_{\text{in}_{z,j}})\frac{Z(f_k)}{Z_0}\right)$ represents the angle of refraction, $Z_0 = 377\Omega$ refers to the intrinsic impedance of free space and $Z(f_k)$ represents the characteristic impedance of the reflecting

medium given [31] by $Z(f_k) = \sqrt{\frac{\mu_0}{\varepsilon_0\left(n^2 - (\frac{\varsigma c}{4\pi f_k})^2 - j\frac{2n\varsigma c}{4\pi f_k}\right)}}$. The constants μ_0 and ε_0 represent the free-space permeability and permittivity respectively, n denotes the index of refraction and ς represents the absorption coefficient of the reflecting medium. Note that, the traditional CSI estimation techniques such as LS/ MMSE cannot exploit the spatial-sparsity of the THz MIMO channel that occurs due to the higher link directionality. The virtual channel model for the THz system, that enables improved estimation, is described in the subsequent section.

IV. SPARSE CHANNEL ESTIMATION MODEL FOR AN SC-FDE THz SYSTEM

As discussed in Section II-A, the output vector at the RC for the mth block at the kth subcarrier is given by

$$\mathbf{y}_m[k] = \mathbf{\Phi}_m[k]\mathbf{h}[k] + \mathbf{v}_m[k]. \tag{18}$$

To derive an equivalent model [32] for the kth subcarrier, we vertically stack the outputs $\mathbf{y}_m[k]$ for all the blocks, i.e, $1 \leq m \leq M$

$$\underbrace{\begin{bmatrix} \mathbf{y}_1[k] \\ \vdots \\ \mathbf{y}_M[k] \end{bmatrix}}_{\mathbf{y}_p[k]} = \underbrace{\begin{bmatrix} \mathbf{\Phi}_1[k] \\ \vdots \\ \mathbf{\Phi}_M[k] \end{bmatrix}}_{\mathbf{\Phi}_p[k]} \mathbf{h}[k] + \underbrace{\begin{bmatrix} \mathbf{v}_1[k] \\ \vdots \\ \mathbf{v}_M[k] \end{bmatrix}}_{\mathbf{v}_p[k]}, \tag{19}$$

where, $\mathbf{y}_p[k] \in \mathbb{C}^{MN_{RF} \times 1}$ represents the stacked pilot output vector, $\mathbf{\Phi}_p[k] \in \mathbb{C}^{MN_{RF} \times N_R N_T}$ denotes the stacked pilot sensing matrix and $\mathbf{v}_p[k] \in \mathbb{C}^{MN_{RF} \times 1}$ represents the stacked noise corresponding to all the M blocks for subcarrier k. Thus, the stacked system model at the combiner output is given by

$$\mathbf{y}_p[k] = \mathbf{\Phi}_p[k]\mathbf{h}[k] + \mathbf{v}_p[k]. \tag{20}$$

At this point, one can formulate the traditional ZF and MMSE estimators for the wideband THz MIMO channel at each subcarrier as

$$\hat{\mathbf{h}}_{\text{LS}}[k] = (\mathbf{\Phi}_p^H[k]\mathbf{\Phi}_p[k])^{-1}\mathbf{\Phi}_p^H[k]\mathbf{y}_p[k] \tag{21}$$

$$\hat{\mathbf{h}}_{\text{MMSE}}[k] = (\mathbf{R}_h^{-1}[k] + \mathbf{\Phi}_p^H[k]\mathbf{R}_{v_p}^{-1}\mathbf{\Phi}_p[k])^{-1}\mathbf{\Phi}_p^H[k]\mathbf{y}_p[k], \tag{22}$$

where, the noise covariance matrix $\mathbf{R}_{v_p} = \sigma_v^2 K$ blkdiag $\left(\left\{\mathbf{W}_{RF,m}^H \mathbf{W}_{RF,m}\right\}_{m=1}^M\right)$ and $\mathbf{R}_{v_p} \in \mathbb{C}^{MN_{RF} \times MN_{RF}}$. Furthermore, $\mathbf{R}_h[k] = \mathbb{E}[\mathbf{h}[k]\mathbf{h}^H[k]] \in \mathbb{C}^{N_R N_T \times N_R N_T}$ denotes the channel covariance matrix. Note that the conventional LS and MMSE-based estimators require excessive pilot overheads in THz MIMO systems since they require an over-determined model, i.e, $MN_{RF} \geq N_T N_R$. This substantially degrades the spectral efficiency. Moreover, these orthodox techniques listed above do not leverage the specific properties of the THz channel. An important observation in this context is the fact that the directional nature of propagation in the THz band leads to an angularly sparse multipath channel [6]. In such scenarios, the recently developed sparse signal processing paradigm can yield excellent results for channel estimation in THz band hybrid MIMO systems, even in ill-posed scenarios, wherein we have $MN_{RF} \ll N_T N_R$. This feature may lead to substantial

enhancement in the accuracy of channel estimation as well as bandwidth-efficiency. Therefore, the following section presents a sparse channel estimation model in THz band communication.

A. Sparse channel estimation model for a wideband THz hybrid MIMO system

Let G_T and G_R represent the angular grid-sizes, which obey the relationship $(G_T, G_R) \geq \max(N_T, N_R)$. The transmit and receive angular grids Θ_T and Θ_R, respectively, are created by considering the directional-cosines to be uniformly spaced between -1 to 1, mathematically described as

$$\Theta_T = \left\{ \theta_t : \cos(\theta_t) = \frac{2}{G_T}(t-1) - 1, 1 \leq t \leq G_T \right\},$$
$$(23)$$

$$\Theta_R = \left\{ \theta_r : \cos(\theta_r) = \frac{2}{G_R}(r-1) - 1, 1 \leq r \leq G_R \right\}.$$
$$(24)$$

Let $\mathbf{A}_T(\Theta_T, f_k) \in \mathbb{C}^{N_T \times G_T}$ and $\mathbf{A}_R(\Theta_R, f_k) \in \mathbb{C}^{N_R \times G_R}$ represent the dictionary matrices for the transmit and receive array manifold vectors using the angular grids Θ_T and Θ_R [22], which are modeled as

$$\mathbf{A}_T(\Theta_T, f_k) = [\tilde{\mathbf{a}}_T(\theta_1, f_k), \tilde{\mathbf{a}}_T(\theta_2, f_k), \cdots, \tilde{\mathbf{a}}_T(\theta_{G_T}, f_k)],$$
$$(25)$$

$$\mathbf{A}_R(\Theta_R, f_k) = [\tilde{\mathbf{a}}_R(\theta_1, f_k), \tilde{\mathbf{a}}_R(\theta_2, f_k), \cdots, \tilde{\mathbf{a}}_R(\theta_{G_R}, f_k)].$$
$$(26)$$

Because of the particular selection of the AoD and AoA grids, the array response dictionary matrices [18] satify the conditions

$$\mathbf{A}_T(\Theta_T, f_k)\mathbf{A}_T^H(\Theta_T, f_k) = \frac{G_T}{N_T}\mathbf{I}_{N_T},$$

$$\mathbf{A}_R(\Theta_R, f_k)\mathbf{A}_R^H(\Theta_R, f_k) = \frac{G_R}{N_R}\mathbf{I}_{N_R}.$$
$$(27)$$

It is noteworthy that the number of multipath components is substantially lower in a THz system due to the reflections, molecular absorption and free-space losses, as described earlier. This fact, coupled with the presence of only a few highly-directional beams, renders the THz MIMO channel $\mathbf{H}_b[k]$ angularly sparse in nature. Subsequently, the THz MIMO channel $\mathbf{H}[k]$ can be represented using the virtual channel model [32] as

$$\mathbf{H}[k] = \mathbf{A}_R(\Theta_R, f_k)\mathbf{H}_b[k]\mathbf{A}_T^H(\Theta_T, f_k),$$
$$(28)$$

where $\mathbf{H}_b[k] \in \mathbb{C}^{G_R \times G_T}$ is the beamspace CFR matrix that pertains to $\mathbf{H}[k]$. Exploiting the property of the vec(.) operator described in Section-I-C, the vectorized THz MIMO channel can be represented as

$$\mathbf{h}[k] = \text{vec}(\mathbf{H}[k]) = [\mathbf{A}_T^*(\Theta_T, f_k) \otimes \mathbf{A}_R(\Theta_R, f_k)]\,\mathbf{h}_b[k],$$
$$(29)$$

where $\mathbf{h}_b[k] = \text{vec}(\mathbf{H}_b[k]) \in \mathbb{C}^{G_R G_T \times 1}$. Let $\mathbf{\Psi}[k] = [\mathbf{A}_T^*(\Theta_T, f_k) \otimes \mathbf{A}_R(\Theta_R, f_k)]$ denote the *sparsifying-dictionary*, where $\mathbf{\Psi}[k] \in \mathbb{C}^{N_R N_T \times G_R G_T}$. Thus, upon

substituting (29) in to (20), the expression for the output vector can be recast as

$$\mathbf{y}_p[k] = \underbrace{\mathbf{\Phi}_p[k]\mathbf{\Psi}[k]}_{\tilde{\mathbf{\Psi}}_p[k]}\mathbf{h}_b[k] + \mathbf{v}_p[k],$$
$$(30)$$

where $\tilde{\mathbf{\Psi}}_p[k] \in \mathbb{C}^{MN_{RF} \times G_R G_T}$ represents an equivalent sensing matrix. It is noteworthy that other sparse estimation techniques, such as Basis-Pursuit [33], FOCUSS [34], and OMP [35], suffer from structural and convergence-related issues. For instance, in the case of Basis-Pursuit, there is no guarantee that the sparsest solution aligns with the global minimum of its cost function, leading to structural errors. Similarly, the FOCUSS algorithm frequently converges to suboptimal local minima, resulting in convergence errors. Additionally, the OMP algorithm presents challenges; a low cessation criterion leads to a larger number of iterations and more non-zero components in the sparse vector, causing structural errors. Conversely, a high cessation criterion results in fewer iterations, failing to capture all the dominant modes of the channel and causing convergence errors. By contrast, the proposed Bayesian learning methods assure global convergence with a high probability due to the inherent attributes of their cost function and the use of the EM algorithm. Furthermore, in a noiseless scenario, the global minimum of the Bayesian learning cost function corresponds to the sparsest solution [36]. Even in a noisy scenario, all the local minima are sparse and have improved representation over non-sparse solutions. The next section describes the framework that utilizes Bayesian learning to estimate $\mathbf{h}_b[k]$, that intends to exploit the sparsity of the THz MIMO channel as described above.

V. PA-BL FOR SPARSE CHANNEL ESTIMATION IN THZ MIMO SYSTEMS

This section introduces a Pilot Aided-Bayesian Learning (PA-BL) technique for CSI estimation in dual-wideband channel. This can also be applied for the case of an under-determined system, i.e, $MN_{RF} \ll G_T G_R$ [18]. At first, the parameterized Gaussian prior $f(\mathbf{h}_b[k]; \mathbf{\Gamma}_k)$ is allocated to the kth pilot subcarrier for the sparse vector $\mathbf{h}_b[k]$ [36], which is given by

$$f(\mathbf{h}_b[k]; \mathbf{\Gamma}_k) = \prod_{i=1}^{G_R G_T} (\pi\gamma_{k,i})^{-1}\exp\left(-\frac{|\mathbf{h}_b[k](i)|^2}{\gamma_{k,i}}\right),$$
$$(31)$$

where $\gamma_{k,i}$ represents ith hyper-parameter for the kth subcarrier; $1 \leq i \leq G_R G_T$, and the matrix $\mathbf{\Gamma}_k \in \mathbb{R}^{G_R G_T \times G_R G_T}$ is the diagonal matrix of hyper-parameters is given by $\mathbf{\Gamma}_k = \text{diag}(\gamma_{k,1}, \gamma_{k,2}, \cdots, \gamma_{k,G_R G_T})$. The MMSE estimate of $\hat{\mathbf{h}}_b[k]$ is given by

$$\hat{\mathbf{h}}_b[k] = \left(\tilde{\mathbf{\Psi}}_p^H[k]\mathbf{R}_{v_p}^{-1}\tilde{\mathbf{\Psi}}_p[k] + \mathbf{\Gamma}_k^{-1}\right)^{-1}\tilde{\mathbf{\Psi}}_p^H[k]\mathbf{R}_{v_p}^{-1}\mathbf{y}_p[k].$$
$$(32)$$

It can be noted that MMSE relies on the hyper-parameter matrix $\mathbf{\Gamma}_k$. Therefore, estimating $\mathbf{\Gamma}_k$ holds the key toward determining the sparse channel $\hat{\mathbf{h}}_b[k]$. The log-likelihood

$\log\left[f\left(\mathbf{y}_p[k];\boldsymbol{\Gamma}_k\right)\right]$ of the hyper-parameter matrix $\boldsymbol{\Gamma}_k$ can be formulated as

$$\log\left[f\left(\mathbf{y}_p[k];\boldsymbol{\Gamma}_k\right)\right] = c_1 - \log\left[\det\left(\mathbf{R}_{\mathbf{y}_p}[k]\right)\right] - \mathbf{y}_p^H[k]\mathbf{R}_{\mathbf{y}_p}^{-1}[k]\mathbf{y}_p[k], \quad (33)$$

where $c_1 = -MN_{RF}\log(\pi)$ and the pilot covariance matrix $\mathbf{R}_{\mathbf{y}_p}[k] \in \mathbb{C}^{MN_{RF}\times MN_{RF}}$ is given by

$$\mathbf{R}_{\mathbf{y}_p}[k] = \mathbb{E}\left\{\mathbf{y}_p[k]\mathbf{y}_p^H[k]\right\}. \quad (34)$$

Substituting equation (30) in to (34) yields the pilot covariance matrix in the form of

$$\mathbf{R}_{\mathbf{y}_p}[k] = \mathbf{R}_{v_p} + \tilde{\boldsymbol{\Psi}}_p[k]\boldsymbol{\Gamma}_k\tilde{\boldsymbol{\Psi}}_p^H[k], \quad (35)$$

where $\mathbf{v}_p[k] \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_{v_p})$. Maximization of the log-likelihood with respect to $\boldsymbol{\Gamma}_k$ is mathematically intractable. Thus, the Expectation Maximization (EM) algorithm is a suitable choice for iterative maximization, as it ensures convergence to a local optimum [37]. The EM-based procedure of determining the sparse channel estimate is described below. Let the estimate of the ith hyperparameter at the kth subcarrier obtained from the $(j-1)$st EM iteration be represented as $\hat{\gamma}_{k,i}^{(j-1)}$ and $\hat{\boldsymbol{\Gamma}}_k^{(j-1)}$ denotes the hyperparameter matrix at the kth subcarrier, which is further defined as $\hat{\boldsymbol{\Gamma}}_k^{(j-1)} = \mathrm{diag}\left(\hat{\gamma}_{k,1}^{(j-1)}, \hat{\gamma}_{k,2}^{(j-1)}, \cdots, \hat{\gamma}_{k,G_RG_T}^{(j-1)}\right)$. The steps involved in updating the estimate $\hat{\boldsymbol{\Gamma}}_k^{(j)}$ in the jth iteration of the EM algorithm can be outlined as follows. Let the ith hyper-parameter estimate obtained in iteration $(j-1)$st for the kth subcarrier be denoted by $\hat{\gamma}_{k,i}^{(j-1)}$. The hyper-parameter update $\hat{\gamma}_{k,i}^{(j)}$ in the jth EM-iteration is obtained by maximizing the log-likelihood function $\mathcal{L}\left(\boldsymbol{\Gamma}_k|\hat{\boldsymbol{\Gamma}}_k^{(j-1)}\right)$, defined as

$$\mathcal{L}\left(\boldsymbol{\Gamma}_k|\hat{\boldsymbol{\Gamma}}_k^{(j-1)}\right) = \mathbb{E}_{\mathbf{h}_b[k]|\mathbf{y}_p[k];\hat{\boldsymbol{\Gamma}}_k^{(j-1)}}\left\{\log f\left(\mathbf{y}_p[k],\mathbf{h}_b[k];\boldsymbol{\Gamma}_k\right)\right\}. \quad (36)$$

The conditional expectation of the log-likelihood function $\mathcal{L}\left(\boldsymbol{\Gamma}_k|\hat{\boldsymbol{\Gamma}}_k^{(j-1)}\right)$ in the E-step is given

$$\begin{aligned}\mathcal{L}\left(\boldsymbol{\Gamma}_k|\hat{\boldsymbol{\Gamma}}_k^{(j-1)}\right) &= \mathbb{E}_{\mathbf{h}_b[k]|\mathbf{y}_p[k];\hat{\boldsymbol{\Gamma}}_k^{(j-1)}}\left\{\log f\left(\mathbf{y}_p[k],\mathbf{h}_b[k];\boldsymbol{\Gamma}_k\right)\right\}\\ &= \mathbb{E}_{\mathbf{h}_b[k]|\mathbf{y}_p[k];\hat{\boldsymbol{\Gamma}}_k^{(j-1)}}\left\{\log\left[f(\mathbf{y}_p[k]|\mathbf{h}_b[k])\right]\right\} + \\ &\quad \mathbb{E}_{\mathbf{h}_b[k]|\mathbf{y}_p[k];\hat{\boldsymbol{\Gamma}}_k^{(j-1)}}\left\{\log\left[f(\mathbf{h}_b[k];\boldsymbol{\Gamma}_k)\right]\right\}. \end{aligned} \quad (37)$$

The expression enclosed within the first expectation operator in the aforementioned equation can be simplified as

$$\begin{aligned}\log f(\mathbf{y}_p[k]|\mathbf{h}_b[k]) &= -MN_{RF}\log(\pi) - \log\left[\det(\mathbf{R}_{v_p})\right] - \\ &\left(\mathbf{y}_p[k] - \tilde{\boldsymbol{\Psi}}_p[k]\mathbf{h}_b[k]\right)^H\mathbf{R}_{v_p}^{-1}\left(\mathbf{y}_p[k] - \tilde{\boldsymbol{\Psi}}_p[k]\mathbf{h}_b[k]\right).\end{aligned} \quad (38)$$

It can be seen that the above expression is independent of the hyper-parameter matrix $\boldsymbol{\Gamma}_k$, which implies that it can be overlooked in the ensuing M-step of the CSI estimation algorithm. Thus, the equivalent optimization problem of determining the beamspace CSI estimate can be formulated

$$\hat{\boldsymbol{\Gamma}}_k^{(j)} = \arg\max_{\boldsymbol{\Gamma}_k} \mathbb{E}_{\mathbf{h}_b[k]|\mathbf{y}_p[k];\hat{\boldsymbol{\Gamma}}_k^{(j-1)}}\left\{\log\left[f\left(\mathbf{h}_b[k];\boldsymbol{\Gamma}_k\right)\right]\right\}. \quad (39)$$

On substituting $f(\mathbf{h}_b[k];\boldsymbol{\Gamma}_k)$ from equation (31) into the optimization objective, one can observe that the resultant maximization problem can be separated in reference to the hyper-parameters, and the individual optimization problems are given as

$$\hat{\gamma}_{k,i}^{(j)} = \arg\max_{\gamma_{k,i}}\left[-\log(\gamma_{k,i}) - \frac{\mathbb{E}_{\mathbf{h}_b[k]|\mathbf{y}_p[k];\hat{\boldsymbol{\Gamma}}_k^{(j-1)}}\left\{|\mathbf{h}_b[k](i)|^2\right\}}{\gamma_{k,i}}\right]. \quad (40)$$

On solving the aforementioned equation, the estimates $\hat{\gamma}_{k,i}^{(j)}$ are determined as

$$\hat{\gamma}_{k,i}^{(j)} = \mathbb{E}_{\mathbf{h}_b[k]|\mathbf{y}_p[k];\hat{\boldsymbol{\Gamma}}_k^{(j-1)}}\left\{|\mathbf{h}_b[k](i)|^2\right\}, \quad (41)$$

where, the *a posteriori* distribution of $\mathbf{h}_b[k]$, required to simplify the conditional expectation $\mathbb{E}_{\mathbf{h}_b[k]|\mathbf{y}_p[k];\hat{\boldsymbol{\Gamma}}_k^{(j-1)}}\{.\}$ above, can be determined as

$$f\left(\mathbf{h}_b[k]|\mathbf{y}_p[k];\hat{\boldsymbol{\Gamma}}_k^{(j-1)}\right) = \mathcal{CN}\left(\boldsymbol{\mu}_b^{(j)}[k], \mathbf{R}_b^{(j)}[k]\right). \quad (42)$$

The quantity $\boldsymbol{\mu}_b^{(j)}[k] \in \mathbb{C}^{G_RG_T\times 1}$ represents the *a posteriori* mean, while the quantity $\mathbf{R}_b^{(j)}[k] \in \mathbb{C}^{G_RG_T\times G_RG_T}$ represents the covariance matrix. These expressions are given below

$$\boldsymbol{\mu}_b^{(j)}[k] = \mathbf{R}_b^{(j)}[k]\tilde{\boldsymbol{\Psi}}_p^H[k]\mathbf{R}_{v_p}^{-1}\mathbf{y}_p[k],$$

$$\mathbf{R}_b^{(j)}[k] = \left[\tilde{\boldsymbol{\Psi}}_p^H[k]\mathbf{R}_{v_p}^{-1}\tilde{\boldsymbol{\Psi}}_p[k] + \left(\hat{\boldsymbol{\Gamma}}_k^{(j-1)}\right)^{-1}\right]^{-1}. \quad (43)$$

Using the *a posteriori* PDF from (42), the expression in (41) for the hyper-parameter estimate reduces to

$$\hat{\gamma}_{k,i}^{(j)} = \mathbf{R}_b^{(j)}[k](i,i) + |\boldsymbol{\mu}_b^{(j)}[k](i)|^2. \quad (44)$$

Algorithm 1 provides an overview of the sparse channel estimation technique for THz hybrid MIMO systems, employing the PA-BL framework [6]. The procedure is repeated until the hyperparameter estimate converges, i.e, $\|\hat{\boldsymbol{\Gamma}}_k^{(j)} - \hat{\boldsymbol{\Gamma}}_k^{(j-1)}\|_F^2 < \epsilon_0$ or the number of iterations reached its maximum affordable value $K_{\max}$. The PA-BL based channel estimate after convergence is given by

$$\hat{\mathbf{h}}_b[k] = \boldsymbol{\mu}_b^{(j)}[k]. \quad (45)$$

A. Bayesian Cramér-Rao lower bound for pilot-aided learning

Consider the parameterized Gaussian prior assigned to the ith hyper-parameter of the kth pilot subcarrier corresponding to $\mathbf{h}_b[k]$, as mentioned in (31). The corresponding PDF is given by

$$f(\mathbf{h}_b[k];\boldsymbol{\Gamma}_k) = \frac{1}{(\pi)^{G_RG_T}\det(\boldsymbol{\Gamma}_k)}\exp\left(-\mathbf{h}_b^H[k]\boldsymbol{\Gamma}_k^{-1}\mathbf{h}_b[k]\right). \quad (46)$$

The log-likelihood of the beamspace channel $\mathbf{h}_b[k]$ is obtained via applying the log function to both sides of the above equation, yielding

$$\begin{aligned}\mathcal{L}\left(f\left[\mathbf{h}_b[k];\boldsymbol{\Gamma}_k\right]\right) = &-G_RG_T\log(\pi) - \log\left[\det(\boldsymbol{\Gamma}_k)\right] - \\ &\mathbf{h}_b^H[k]\boldsymbol{\Gamma}_k^{-1}\mathbf{h}_b[k].\end{aligned} \quad (47)$$

The conditional PDF of the pilot output vector $\mathbf{y}_p[k]$, as defined in equation (30), is given by the complex normal

Algorithm 1 PA-BL for sparse channel estimation in THz MIMO systems

---

Input: Pilot output $\mathbf{y}_p[k]$, equivalent sensing matrix $\tilde{\boldsymbol{\Psi}}_p[k]$, noise covariance $\mathbf{R}_{v_p}$, dictionary matrices $\mathbf{A}_R(\Theta_R, f_k)$ and $\mathbf{A}_T(\Theta_T, f_k)$, stopping parameters ϵ and $K_{\max}$

Initialization: $\gamma_{k,i}^{(0)} = 1, \forall\, 1 \leq i \leq G_R G_T$ where $1 \leq k \leq K$ $\Rightarrow \hat{\boldsymbol{\Gamma}}_k^{(0)} = \mathbf{I}_{G_R G_T}$, $\hat{\boldsymbol{\Gamma}}_k^{(-1)} = \mathbf{0}$ and counter $j = 0$

while $\left(\| \hat{\boldsymbol{\Gamma}}_k^{(j)} - \hat{\boldsymbol{\Gamma}}_k^{(j-1)} \|_F > \epsilon \text{ and } j < K_{\max} \right)$ **do**

  1) $j \leftarrow j + 1$

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

$$\mathbf{R}_b^{(j)}[k] = \left[\tilde{\boldsymbol{\Psi}}_p^H[k] \mathbf{R}_{v_p}^{-1} \tilde{\boldsymbol{\Psi}}_p[k] + \left(\hat{\boldsymbol{\Gamma}}_k^{(j-1)} \right)^{-1} \right]^{-1}$$

$$\boldsymbol{\mu}_b^{(j)}[k] = \mathbf{R}_b^{(j)}[k] \tilde{\boldsymbol{\Psi}}_p^H[k] \mathbf{R}_{v_p}^{-1} \mathbf{y}_p[k]$$

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

    for $i = 1, \cdots, G_R G_T$ **do**

$$\hat{\gamma}_{k,i}^{(j)} = \mathbf{R}_b^{(j)}[k](i,i) + |\boldsymbol{\mu}_b^{(j)}[k](i)|^2$$

    end for

end while

$\hat{\mathbf{h}}_b[k] = \boldsymbol{\mu}_b^{(j)}[k]$

Output: $\widehat{\mathbf{H}}_{\text{PA-BL}}[k] = \mathbf{A}_R(\Theta_R, f_k) \text{vec}^{-1}(\hat{\mathbf{h}}_b[k]) \mathbf{A}_T^H(\Theta_T, f_k)$

---

distribution $\mathcal{CN}\left(\tilde{\boldsymbol{\Psi}}_p[k] \mathbf{h}_b[k], \mathbf{R}_{v_p} \right)$. Hence, the expression of the conditional PDF is

$$f\left(\mathbf{y}_p[k] | \mathbf{h}_b[k] \right) = \vartheta[k] \exp\Big(- \left(\mathbf{y}_p[k] - \tilde{\boldsymbol{\Psi}}_p[k] \mathbf{h}_b[k] \right)^H$$
$$\mathbf{R}_{v_p}^{-1} \left(\mathbf{y}_p[k] - \tilde{\boldsymbol{\Psi}}_p[k] \mathbf{h}_b[k] \right) \Big), \tag{48}$$

where $\vartheta[k] = \frac{1}{(\pi)^{MN_{RF}} |\mathbf{R}_{v_p}|}$. Taking the logarithm of both sides of the equation above gives us

$$\mathcal{L}\left(\mathbf{y}_p[k] | \mathbf{h}_b[k] \right) = -MN_{RF} \log(\pi) - \log\left[\det(\mathbf{R}_{v_p}) \right] -$$
$$\left(\mathbf{y}_p[k] - \tilde{\boldsymbol{\Psi}}_p[k] \mathbf{h}_b[k] \right)^H \mathbf{R}_{v_p}^{-1} \left(\mathbf{y}_p[k] - \tilde{\boldsymbol{\Psi}}_p[k] \mathbf{h}_b[k] \right). \tag{49}$$

Let $\mathbf{J}[k] \in \mathbb{C}^{G_R G_T \times G_R G_T}$ denote the Bayesian Fisher Information Matrix (FIM), formulated as

$$\mathbf{J}[k] = \mathbf{J}_p[k] + \mathbf{J}_b[k], \tag{50}$$

where $\mathbf{J}_p[k] \in \mathbb{C}^{G_R G_T \times G_R G_T}$ denotes the FIM associated with the pilot output $\mathbf{y}_p[k]$, while $\mathbf{J}_b[k] \in \mathbb{C}^{G_R G_T \times G_R G_T}$ represents the FIM corresponding to the *a prior* information of the beamspace CSI $\mathbf{h}_b[k]$. The above quantities can be formulated as

$$\mathbf{J}_p[k] = -\mathbb{E}_{\mathbf{y}_p[k], \mathbf{h}_b[k]} \left\{ \frac{\partial^2 \mathcal{L}(\mathbf{y}_p[k] | \mathbf{h}_b[k])}{\partial \mathbf{h}_b[k] \partial \mathbf{h}_b^H[k]} \right\},$$

$$\mathbf{J}_b[k] = -\mathbb{E}_{\mathbf{h}_b[k]} \left\{ \frac{\partial^2 \mathcal{L}(\mathbf{h}_b[k]; \boldsymbol{\Gamma}_k)}{\partial \mathbf{h}_b[k] \partial \mathbf{h}_b^H[k]} \right\}. \tag{51}$$

Upon using the conditional PDF in (49), the FIM $\mathbf{J}_p[k]$ can be derived as

$$\mathbf{J}_p[k] = \tilde{\boldsymbol{\Psi}}_p^H[k] \mathbf{R}_{v_p}^{-1} \tilde{\boldsymbol{\Psi}}_p[k]. \tag{52}$$

Similarly, the FIM associated with $\mathbf{h}_b[k]$ is obtained by using the *a prior* PDF from equation (46) in equation (51). The corresponding expression for $\mathbf{J}_b[k]$ can be formulated as

$$\mathbf{J}_b[k] = \hat{\boldsymbol{\Gamma}}_k^{-1}. \tag{53}$$

Using the expressions in (52) and (53), the Bayesian FIM is given by

$$\mathbf{J}[k] = \tilde{\boldsymbol{\Psi}}_p^H[k] \mathbf{R}_{v_p}^{-1} \tilde{\boldsymbol{\Psi}}_p[k] + \hat{\boldsymbol{\Gamma}}_k^{-1}. \tag{54}$$

Therefore, the BCRLB that pertains to the MSE of the estimate $\hat{\mathbf{h}}_b[k]$ can be expressed as

$$\text{MSE}(\hat{\mathbf{h}}_b[k]) = \mathbb{E}\left\{ \| \hat{\mathbf{h}}_b[k] - \mathbf{h}_b[k] \|^2 \right\} \geq \text{Tr}\left\{ \mathbf{J}^{-1}[k] \right\}$$
$$= \text{Tr}\left\{ \left(\tilde{\boldsymbol{\Psi}}_p^H[k] \mathbf{R}_{v_p}^{-1} \tilde{\boldsymbol{\Psi}}_p[k] + \hat{\boldsymbol{\Gamma}}_k^{-1} \right)^{-1} \right\}. \tag{55}$$

Upon exploiting the relationship between the beamspace representation of the THz MIMO channel and its vectorized version, the BCRLB for the estimated CSI is finally given by

$$\text{MSE}(\hat{\mathbf{H}}[k]) \geq \text{Tr}\left\{ \boldsymbol{\Psi}[k] \mathbf{J}^{-1}[k] \boldsymbol{\Psi}^H[k] \right\}. \tag{56}$$

A drawback of THz systems is the fact that the received signal is severely affected by noise in the training phase. Therefore, it is difficult to obtain an accurate CSI estimate using exclusively pilot symbols [38]. In order to overcome this limitation, the DA-BL framework is proposed next which leverages the data along with the pilot symbols for enhancing the estimation accuracy.

VI. DA-BL For Sparse Channel Estimation in THz Hybrid MIMO systems

The data-model is developed next for our THz MIMO system, which facilitates joint channel estimation and data-detection. Let $\mathbf{x}_{d,m}[k]$ denote the data symbol vector whose elements are derived from an appropriate constellation. The data model corresponding to the k-th subcarrier is given by

$$\mathbf{y}_{d,m}[k] = \mathbf{W}_{RF,m}^H \mathbf{H}[k] \mathbf{F}_{RF,m} \mathbf{x}_{d,m}[k] + \mathbf{W}_{RF,m}^H \mathbf{v}_{d,m}[k], \tag{57}$$

where we have $\mathbf{y}_{d,m}[k] \in \mathbb{C}^{N_{RF} \times 1}, \mathbf{W}_{RF,m} \in \mathbb{C}^{N_R \times N_{RF}}$ and $\mathbf{F}_{RF,m} \in \mathbb{C}^{N_T \times N_{RF}}$. The construction of $\mathbf{W}_{RF,m}$ and $\mathbf{F}_{RF,m}$ is explained in Section VII-A. In (57) $\mathbf{x}_{d,m}[k] \in \mathbb{C}^{N_{RF} \times 1}$ and $\mathbf{v}_{d,m}[k] \in \mathbb{C}^{N_R \times 1}$ represents the K-point FFT of $\mathbf{v}_{d,m}$, where $\mathbf{v}_{d,m} \sim \mathcal{CN}(\mathbf{0}_{N_R \times 1}, \sigma_d^2 \mathbf{I}_{N_R})$. Let $\tilde{\mathbf{v}}_{d,m}[k] = \mathbf{W}_{RF,m}^H \mathbf{v}_{d,m}[k]$ represent the equivalent noise after receiver combining and $\mathbf{H}_{\text{eq},m}[k] = \mathbf{W}_{RF,m}^H \mathbf{H}[k] \mathbf{F}_{RF,m} \in \mathbb{C}^{N_{RF} \times N_{RF}}$ represent the equivalent channel matrix. Using equation (57), one can readily derive the expressions for the conventional ZF and MMSE based detectors, which are respectively given by

$$\hat{\mathbf{x}}_{d,m}^{\text{ZF}}[k] = \left(\mathbf{H}_{\text{eq},m}^H[k] \mathbf{H}_{\text{eq},m}[k] \right)^{-1} \mathbf{H}_{\text{eq},m}^H[k] \mathbf{y}_{d,m}[k], \tag{58}$$

$$\hat{\mathbf{x}}_{d,m}^{\text{MMSE}}[k] = \left(\mathbf{H}_{\text{eq},m}^H[k] \mathbf{H}_{\text{eq},m}[k] + \mathbf{R}_{d,m} \right)^{-1} \mathbf{H}_{\text{eq},m}^H[k] \mathbf{y}_{d,m}[k], \tag{59}$$

where $\mathbf{R}_{d,m}$ represents the noise covariance matrix for k-th subcarrier, which is defined as

$$\mathbf{R}_{d,m} = \mathbb{E}\left\{ \tilde{\mathbf{v}}_{d,m}[k]\tilde{\mathbf{v}}_{d,m}^H[k] \right\} = \sigma_d^2 K \mathbf{W}_{RF,m}^H \mathbf{W}_{RF,m}. \quad (60)$$

The CSI estimate obtained in the previous section using the PA-BL algorithm is used for obtaining the initial data symbol estimate of the MMSE estimator. Note that during the estimation process, $\mathbf{W}_{RF,m}^H$ is determined as per Section VII-A, while $\mathbf{W}_{BB,m}^H[k]$ is set as the identity matrix. During the following data detection phase, $\mathbf{W}_{BB,m}^H[k]$ is designed as the LMMSE equalizer $\left(\mathbf{H}_{\mathrm{eq},m}^H[k]\mathbf{H}_{\mathrm{eq},m}[k] + \mathbf{R}_{d,m} \right)^{-1} \mathbf{H}_{\mathrm{eq},m}^H[k]$. By expliting the characterstics of the vec(.) operator as discussed in Section-I-C, one can recast the system model in equation (57) as

$$\mathbf{y}_{d,m}[k] = \boldsymbol{\Phi}_{d,m}[k]\mathbf{h}[k] + \tilde{\mathbf{v}}_{d,m}[k], \quad (61)$$

where the equivalent sensing matrix $\boldsymbol{\Phi}_{d,m}[k] \in \mathbb{C}^{N_{RF} \times N_T N_R}$ is given by

$$\boldsymbol{\Phi}_{d,m}[k] = (\mathbf{x}_{d,m}^T[k]\mathbf{F}_{RF,m}^T \otimes \mathbf{W}_{RF,m}^H). \quad (62)$$

Finally, upon stacking the outputs $\mathbf{y}_{d,m}[k]$ across all the data blocks, one obtains

$$\underbrace{\begin{bmatrix} \mathbf{y}_{d,1}[k] \\ \vdots \\ \mathbf{y}_{d,N_d}[k] \end{bmatrix}}_{\mathbf{y}_d[k]} = \underbrace{\begin{bmatrix} \boldsymbol{\Phi}_{d,1}[k] \\ \vdots \\ \boldsymbol{\Phi}_{d,N_d}[k] \end{bmatrix}}_{\boldsymbol{\Phi}_d[k]} \mathbf{h}[k] + \underbrace{\begin{bmatrix} \tilde{\mathbf{v}}_{d,1}[k] \\ \vdots \\ \tilde{\mathbf{v}}_{d,N_d}[k] \end{bmatrix}}_{\mathbf{v}_d[k]}, \quad (63)$$

where $\mathbf{y}_d[k] \in \mathbb{C}^{N_d N_{RF} \times 1}$, $\boldsymbol{\Phi}_d[k] \in \mathbb{C}^{N_d N_{RF} \times N_T N_R}$ and $\mathbf{v}_d[k] \in \mathbb{C}^{N_d N_{RF} \times 1}$. Thus, the noise covariance matrix obeys $\mathbf{R}_{v_d} \in \mathbb{C}^{N_d N_{RF} \times N_d N_{RF}}$ where $\mathbf{R}_{v_d} = \sigma_d^2 K \, \mathrm{blkdiag}\left(\left\{ \mathbf{W}_{RF,m}^H \mathbf{W}_{RF,m} \right\}_{m=1}^M \right)$. Eventually, to develop the joint CSI estimation model and data-detection, one can further concatenate the outputs corresponding to the data and pilot models i.e $\mathbf{y}_d[k]$ and $\mathbf{y}_p[k]$ respectively [18], from equation (63) and (19) as

$$\underbrace{\begin{bmatrix} \mathbf{y}_d[k] \\ \mathbf{y}_p[k] \end{bmatrix}}_{\mathbf{y}[k] \, \in \, \mathbb{C}^{(N_d+M)N_{RF} \times 1}} = \underbrace{\begin{bmatrix} \boldsymbol{\Phi}_d[k] \\ \boldsymbol{\Phi}_p[k] \end{bmatrix}}_{\boldsymbol{\Phi}[k] \, \in \, \mathbb{C}^{(N_d+M)N_{RF} \times N_T N_R}} \mathbf{h}[k] + \underbrace{\begin{bmatrix} \mathbf{v}_d[k] \\ \mathbf{v}_p[k] \end{bmatrix}}_{\mathbf{v}[k] \, \in \, \mathbb{C}^{(N_d+M)N_{RF} \times 1}}. \quad (64)$$

Furthermore, the equivalent sensing matrix for the data-aided CSI estimation model is given by $\tilde{\boldsymbol{\Psi}}[k] = \boldsymbol{\Phi}[k]\boldsymbol{\Psi}[k] \in \mathbb{C}^{(N_d+M)N_{RF} \times G_R G_T}$. The covariance matrix $\mathbf{R}_v \in \mathbb{C}^{(N_d+M)N_{RF} \times (N_d+M)N_{RF}}$ of the noise vector $\mathbf{v}[k]$ is

$$\mathbf{R}_v = \mathrm{blkdiag}\left(\mathbf{R}_{v_d}, \mathbf{R}_{v_p} \right). \quad (65)$$

The proposed DA-BL scheme estimates the hyper-parameter matrix $\boldsymbol{\Gamma}_k$ using the EM procedure described below. Let the entire information set for this be defined as $\boldsymbol{\eta}_k = \{\mathbf{X}_d[k], \boldsymbol{\Gamma}_k\}$, where $\mathbf{X}_d[k] \in \mathbb{C}^{N_{RF} \times N_d}$ obtained by horizontally concatenating the transmit data vectors across all blocks and the parameter estimate at iteration $(j-1)$ be given as

$$\hat{\boldsymbol{\eta}}_k^{(j-1)} = \left\{ \hat{\mathbf{X}}_d^{(j-1)}[k], \hat{\boldsymbol{\Gamma}}_k^{(j-1)} \right\}, \quad (66)$$

Algorithm 2 BL-based channel estimation and robust data detection for THz MIMO systems

Input: Observation vector $\mathbf{y}[k]$, Data output $\mathbf{Y}_d[k]$, Pilot sensing matrix $\boldsymbol{\Phi}_p[k]$, Equivalent sensing matrix $\tilde{\boldsymbol{\Psi}}[k]$, Noise Covariance $\mathbf{R}_v$, dictionary matrices $\mathbf{A}_R(\Theta_R, f_k)$ and $\mathbf{A}_T(\Theta_T, f_k)$, stopping parameters ϵ and $J_{\max}$

Initialization: $\hat{\mathbf{X}}_d^{(0)}[k] = \hat{\mathbf{X}}_{d,p}^{\mathrm{MMSE}}[k]$ estimated using the CSI obtained from PA-BL algorithm, counter $j = 0$

while$(\| \hat{\boldsymbol{\Gamma}}_k^{(j)} - \hat{\boldsymbol{\Gamma}}_k^{(j-1)} \|_F > \epsilon)$ and $(j < J_{\max})$ **do**

    $j \leftarrow j + 1$

    E-step: Compute the *aposteriori* covariance and mean

$$\boldsymbol{\Sigma}^{(j)}[k] = \left[(\tilde{\boldsymbol{\Psi}}^{(j-1)}[k])^H \mathbf{R}_v^{-1} \tilde{\boldsymbol{\Psi}}^{(j-1)}[k] + \left(\hat{\boldsymbol{\Gamma}}_k^{(j-1)} \right)^{-1} \right]^{-1}$$

$$\hat{\mathbf{h}}_b^{(j)}[k] = \boldsymbol{\Sigma}^{(j)}[k](\tilde{\boldsymbol{\Psi}}^{(j-1)}[k])^H \mathbf{R}_v^{-1} \mathbf{y}[k]$$

    M-step1: Update the hyperparameters

$$\hat{\gamma}_{k,i}^{(j)} = \boldsymbol{\Sigma}^{(j)}[k](i,i) + |\hat{\mathbf{h}}_b^{(j)}[k](i)|^2; \quad 1 \le i \le G_R G_T$$

    M-step2:

     1) Evaluate $\overline{\mathbf{Y}}_d[k]$ and $\hat{\mathcal{H}}_{eq}^{(j)}[k]$ using equation (76)

     2) Update the data estimate $\hat{\mathbf{X}}_d^{(j)}[k]$ using (82)

     3) Demodulate the estimated data and use it to update $\tilde{\boldsymbol{\Psi}}^{(j)}[k]$

end while

$\hat{\mathbf{h}}_b[k] = \hat{\mathbf{h}}_b^{(j)}[k]$

Output: $\hat{\mathbf{H}}_{\mathrm{DA\text{-}BL}}[k] = \mathbf{A}_R(\Theta_R, f_k)\mathrm{vec}^{-1}(\hat{\mathbf{h}}_b[k])\mathbf{A}_T^H(\Theta_T, f_k)$, $\hat{\mathbf{X}}_d[k] = \hat{\mathbf{X}}_d^{(j)}[k]$

---

where $\hat{\mathbf{X}}_d^{(j-1)}[k]$ denotes the estimate of the data in the $(j-1)$st iteration, whereas $\hat{\boldsymbol{\Gamma}}_k^{(j-1)}$ denotes the hyper-parameter matrix obtained in the $(j-1)$st iteration. The conditional expectation of the log-likelihood function $\mathcal{L}(\boldsymbol{\eta}_k|\hat{\boldsymbol{\eta}_k}^{(j-1)})$ calculated in the E-step for the jth iteration is given by

$$\mathcal{L}(\boldsymbol{\eta}_k|\hat{\boldsymbol{\eta}_k}^{(j-1)}) = \mathbb{E}_{\mathbf{h}_b[k]|\mathbf{y}[k];\hat{\boldsymbol{\eta}_k}^{(j-1)}} \left\{ \log\left[p\left(\mathbf{y}[k], \mathbf{h}_b[k]; \boldsymbol{\eta}_k \right) \right] \right\}$$
$$= \mathbb{E}\left\{ \log\left[p\left(\mathbf{y}[k]|\mathbf{h}_b[k]; \mathbf{X}_d[k] \right) \right] \right\} +$$
$$\mathbb{E}\left\{ \log\left[p\left(\mathbf{h}_b[k]; \boldsymbol{\Gamma}_k \right) \right] \right\}. \quad (67)$$

By simplifying the first term in the above expectation, we obtain

$$\mathcal{L}(p\left(\mathbf{y}[k]|\mathbf{h}_b[k]; \mathbf{X}_d[k] \right)) = -(N_d + M)N_{RF}\log(\pi) -$$
$$\log\left[\det(\mathbf{R}_v) \right] - \left(\mathbf{y}[k] - \tilde{\boldsymbol{\Psi}}[k]\mathbf{h}_b[k] \right)^H \mathbf{R}_v^{-1} \times$$
$$\left(\mathbf{y}[k] - \tilde{\boldsymbol{\Psi}}[k]\mathbf{h}_b[k] \right). \quad (68)$$

Observe that the initial term relies only on the data matrix $\mathbf{X}_d[k]$ through the matrix $\tilde{\boldsymbol{\Psi}}[k]$. In a similar fashion, one can see that the second term is solely determined by the hyper-parameter matrix $\boldsymbol{\Gamma}_k$

$$\mathcal{L}(p\left[\mathbf{h}_b[k]; \boldsymbol{\Gamma}_k \right]) = -G_R G_T \log(\pi) - \log[\det(\boldsymbol{\Gamma}_k)] -$$
$$\mathbf{h}_b^H[k]\boldsymbol{\Gamma}_k^{-1}\mathbf{h}_b[k]. \quad (69)$$

Thus, the joint maximization of the log-likelihood $\mathcal{L}(\boldsymbol{\eta}_k|\hat{\boldsymbol{\eta}_k}^{(j-1)})$ with respect to $\boldsymbol{\eta}_k$ in the M-step breaks

down to two independent maximization steps, one with respect to $\mathbf{X}_d[k]$ and the other with $\mathbf{\Gamma}_k$. The optimization problem in the M-step corresponding to the hyper-parameter update is given by

$$\hat{\gamma}_{k,i}^{(j)} = \arg \max_{\gamma_{k,i} \geq 0} \mathbb{E}_{\mathbf{h}_b[k]|\mathbf{y}[k];\hat{\boldsymbol{\eta}}_k^{(j-1)}} \{\log [p(\mathbf{h}_b[k]; \mathbf{\Gamma}_k)]\}$$
$$\forall\ 1 \leq i \leq G_R G_T. \quad (70)$$

Along similar lines to the PA-BL technique, the estimate of $\gamma_{k,i}$ in iteration j is given by

$$\hat{\gamma}_{k,i}^{(j)} = \mathbb{E}_{\mathbf{h}_b[k]|\mathbf{y}[k];\hat{\mathbf{\Gamma}}_k^{(j-1)}} \{|\mathbf{h}_b[k](i)|^2\}$$
$$= \mathbf{\Sigma}^{(j)}[k](i,i) + |\hat{\mathbf{h}}_b^{(j)}[k](i)|^2, \quad (71)$$

where the *a posteriori* mean vector $\hat{\mathbf{h}}_b^{(j)}[k] \in \mathbb{C}^{G_R G_T \times 1}$ is formulated as

$$\hat{\mathbf{h}}_b^{(j)}[k] = \mathbf{\Sigma}^{(j)}[k] \left(\tilde{\mathbf{\Psi}}^{(j-1)}[k]\right)^H \mathbf{R}_v^{-1} \mathbf{y}[k], \quad (72)$$

and its covariance matrix $\mathbf{\Sigma}^{(j)}[k] \in \mathbb{C}^{G_R G_T \times G_R G_T}$ is given by the expression

$$\mathbf{\Sigma}^{(j)}[k] = \left[(\tilde{\mathbf{\Psi}}^{(j-1)}[k])^H \mathbf{R}_v^{-1} \tilde{\mathbf{\Psi}}^{(j-1)}[k] + \left(\hat{\mathbf{\Gamma}}_k^{(j-1)}\right)^{-1} \right]^{-1}. \quad (73)$$

The matrix $\tilde{\mathbf{\Psi}}^{(j-1)}[k]$ that appears in the above expression depends on the data symbol matrix $\hat{\mathbf{X}}_d^{(j-1)}[k]$ in the $(j-1)$st iteration. Therefore, the M-step of the optimization procedure toward determining the data matrix update $\hat{\mathbf{X}}_d^{(j)}[k]$ can be formulated as

$$\hat{\mathbf{X}}_d^{(j)}[k] = \arg \max_{\mathbf{X}_d[k]} \mathbb{E}\left\{ \log [p(\mathbf{y}_d[k] \mid \mathbf{h}_b[k]; \mathbf{X}_d[k])] \right\}. \quad (74)$$

Using equation (57), the above maximization objective can be reformulated as seen in equation (77). This can be further simplified as seen in (78)-(80). After further simplification, (74) can be reframed as

$$\hat{\mathbf{X}}_d^{(j)}[k] = \arg \min_{\mathbf{X}_d[k]} \left\{ \left\| \overline{\mathbf{Y}}_d[k] - \widehat{\boldsymbol{\mathcal{H}}}_{eq}^{(j)}[k] \mathbf{X}_d[k] \right\|_F^2 \right\}, \quad (75)$$

where, we have

$$\overline{\mathbf{Y}}_d[k] = \begin{bmatrix} \mathbf{Y}_d[k] \\ \mathbf{0} \end{bmatrix}, \quad \widehat{\boldsymbol{\mathcal{H}}}_{eq}^{(j)}[k] = \begin{bmatrix} \hat{\mathbf{H}}_{eq}^{(j)}[k] \\ \left(\mathbf{\Xi}^{(j)}[k]\right)^{\frac{1}{2}} \end{bmatrix}, \quad (76)$$

and $\mathbf{Y}_d[k] \in \mathbb{C}^{N_{RF} \times N_d}$ is obtained by horizontally concatenating the received data vectors across all blocks. The quantity $\hat{\mathbf{H}}_{eq}^{(j)}[k] \in \mathbb{C}^{N_{RF} \times N_{RF}}$ above is defined as

$$\hat{\mathbf{H}}_{eq}^{(j)}[k] = \mathbb{E}\left\{ \mathbf{W}_{RF}^H \mathbf{H}[k] \mathbf{F}_{RF} \right\}$$
$$= \mathbf{W}_{RF}^H \mathbf{A}_R(\Theta_R, f_k) \text{vec}^{-1}\left(\hat{\mathbf{h}}_b^{(j)}[k]\right) \mathbf{A}_T^H(\Theta_T, f_k) \mathbf{F}_{RF}, \quad (81)$$

and the procedure of dividing $\mathbf{\Xi}^{(j)}[k]$ is detailed in Appendix-A. The ZF based detector used for updating the data symbol matrix is obtained as,

$$\hat{\mathbf{X}}_d^{\text{ZF}(j)}[k] = \left\{ \left(\widehat{\boldsymbol{\mathcal{H}}}_{eq}^{(j)}[k]\right)^H \widehat{\boldsymbol{\mathcal{H}}}_{eq}^{(j)}[k] \right\}^{-1} \left(\widehat{\boldsymbol{\mathcal{H}}}_{eq}^{(j)}[k]\right)^H \overline{\mathbf{Y}}_d[k]. \quad (82)$$

The nearest neighbour detection rule is subsequently used to map each element in the matrix above to a symbol in the transmit constellation. The E-step and the M-step described above are executed iteratively until the algorithm converges. The various steps of the DA-BL framework are summarized in Algorithm 2. The BCRLB for DA-BL framework is derived in Appendix B. In general, for a sparse signal recovery problem, when considering an over-complete dictionary, the performance of the proposed Bayesian learning schemes such as PA-BL and DA-BL seem to remain consistent even with a reduced number of observations. By contrast, the performance of non-Bayesian learning techniques such as Basis-Pursuit, FOCUSS, OMP degrade significantly, as observed in our simulation results.

A. Computational Complexity and Memory Occupancy

This subsection derives the computational complexity of the proposed PA-BL and DA-BL techniques for THz hybrid MIMO systems. Due to space constraints, the comprehensive calculations of the computational complexities for both the proposed and existing THz hybrid MIMO channel estimation methods are provided in our technical report [39]. The key findings from Tables I to III of the technical report are discussed next. It is evident from Table I that the OMP technique exhibits a worst-case complexity order of $\mathcal{O}(M^3 N_{RF}^3)$, attributed to the necessity of an intermediate least squares (LS) estimate at each iteration. As shown in Table II, the computational complexity of the PA-BL technique is of the order $\mathcal{O}(M N_{RF}^3 + G_R^3 G_T^3)$, which arises due to the matrix inversion of size $[G_R G_T \times G_R G_T]$ for computing the a posteriori covariance $\mathbf{R}_b^{(j)}[k]$. On the other hand, from Table III, one can see that the computational complexity of the proposed DA-BL technique is of the order $\mathcal{O}\left((N_d + M)^3 N_{RF}^3 + G_R^3 G_T^3\right)$, which is higher than that of the PA-BL and OMP-based techniques. However, the performance, as shown in the simulation results of Section VII, indicates superior performance for DA-BL. The OMP technique has the lowest complexity, but its performance is also poor. Furthermore, due to space constraints, we have similarly derived the memory occupancy for different channel estimation schemes, including OMP, PA-BL, and DA-BL in Table IV of our technical report [39]. It is evident from the memory and computational complexity analysis, that the proposed DA-BL technique has the highest memory requirement and computational complexity, but it also exhibits superior performance. Hence, one can see a clear trade-off performance vs. computational cost/ memory between the proposed and existing CSI estimation schemes.

VII. SIMULATION RESULTS

This section provides the performance evaluation of the proposed PA-BL and DA-BL techniques conceived for THz MIMO systems in the face of the dual-wideband effect. The amplitudes of the path coefficients corresponding to the LoS and NLoS elements are evaluated using equations (14) and (16), respectively, while the corresponding phase shifts are produced using a probability distribution that is uniform over the range (-π,π]. The coefficient of molecular absorption is

$$\hat{\mathbf{X}}_d^{(j)}[k] = \arg \max_{\mathbf{X}_d[k]} \mathbb{E}_{\mathbf{H}_{\text{eq}}[k]|\mathbf{Y}_d[k];\hat{\mathbf{X}}_d^{(j-1)}[k]} \left\{ \log \left[p\left(\mathbf{Y}_d[k] \mid \mathbf{H}_{\text{eq}}[k]; \mathbf{X}_d[k] \right) \right] \right\} \tag{77}$$

$$= \arg \min_{\mathbf{X}_d[k]} \mathbb{E}_{\mathbf{H}_{\text{eq}}[k]|\mathbf{Y}_d[k];\hat{\mathbf{X}}_d^{(j-1)}[k]} \left\{ \left\| \mathbf{Y}_d[k] - \mathbf{H}_{\text{eq}}[k]\mathbf{X}_d[k] \right\|_F^2 \right\} \tag{78}$$

$$= \arg \min_{\mathbf{X}_d[k]} \left\{ \left\| \mathbf{Y}_d[k] - \hat{\mathbf{H}}_{\text{eq}}^{(j)}[k]\mathbf{X}_d[k] \right\|_F^2 + \left\| \left(\mathbf{\Xi}^{(j)}[k] \right)^{\frac{1}{2}} \mathbf{X}_d[k] \right\|_F^2 \right\} \tag{79}$$

$$= \arg \min_{\mathbf{X}_d[k]} \left\{ \left\| \begin{bmatrix} \mathbf{Y}_d[k] \\ \mathbf{0}_{N_{RF} \times N_d} \end{bmatrix} - \begin{bmatrix} \hat{\mathbf{H}}_{\text{eq}}^{(j)}[k] \\ \left(\mathbf{\Xi}^{(j)}[k] \right)^{\frac{1}{2}} \end{bmatrix} \mathbf{X}_d[k] \right\|_F^2 \right\} \tag{80}$$

calculated using equation (15), and its parameters have been taken from the HITRAN database [40]. The THz MIMO channel incorporating the dual-wideband effect is generated using $N_{cl} = 4$ clusters, namely one LoS and $N_{\text{NLoS}} = 3$ clusters. The carrier frequency utilized is $f_c = 1$ THz and the transmission distance considered is $d = 10$ metres. We consider an indoor office scenario, where the molecular composition of air is given as water vapour 1%, oxygen 20.9% and nitrogen 78.1%. The system temperature and pressure, respectively, are considered to be 1 atm and 296K for calculating the molecular absorption loss using equation (15). The AoAs/AoDs associated with each multipath component have been generated from a Laplacian distribution, where the multipath component's angle exhibits a standard deviation of $\frac{1}{10}$ radian from the mean, which in turn is picked with uniform probability from an angular grid. Three reflecting media have been considered with the standard deviations of their roughness set as $\sigma \in \{0.05, 0.13, 0.15\}$ mm [31]. The transmit and receive antenna gains are set to $A_T = 31$ dB, $A_R = 31$ dB, respectively. Two THz MIMO systems are considered for simulations and their parameters are given in Table II. The grid sizes are chosen so that $G_T, G_R \geq \max(N_T, N_R)$. Both the systems have the number of pilot blocks M set as $M \in \{30, 50, 70, 90\}$. The spacings at the transmit and receive antenna ends are taken to be $d_t = d_r = \lambda/2$. Moreover, we consider the utilization of a raised cosine pulse shaping filter with a roll-off factor of 0.80 and an up-sampling factor of 20. The constituent elements of the transmitted pilot symbol vector $\mathbf{u}_{m,n}^{(p)}$ are chosen at unity average power from a 8-PSK constellation.

A. Wideband THz hybrid MIMO channel estimation

The RF TPC and RC components, which are frequency-independent as shown in equation (2), are generated as [20]

$$\mathbf{F}_{RF,m}(i,j) = \frac{1}{\sqrt{N_t}} \exp(j\varphi_{i,j}),$$

$$\mathbf{W}_{RF,m}(i,j) = \frac{1}{\sqrt{N_r}} \exp(j\psi_{i,j}). \tag{83}$$

The phases $\varphi_{i,j}, \psi_{i,j}$ for the set (i,j) are selected randomly with a uniform probability distribution across the set $\mathcal{A} = \left\{ 0, \frac{2\pi}{2^{N_Q}}, \cdots, \frac{(2^{N_Q}-1)2\pi}{2^{N_Q}} \right\}$ with $N_Q = 4$ denoting the

Table II
SIMULATION PARAMETERS FOR THE DUAL-WIDEBAND THZ SYSTEM

Parameter	System-I	System-II
# of TAs (N_T)	32	64
# of RAs (N_R)	32	64
# of RF chains (N_{RF})	6	12
# of diffused ray (N_{ray})	1	1
# of subcarriers (K)	16	32
# of data vectors (N_d)	100	200
Transmit angular grid size	64	128
Receive angular grid size	64	128

angle quantization parameter. Fig. 3(a) and (b) compares the normalized MSE (NMSE) of the channel estimates for the proposed PA-BL technique and the OMP technique for a dual-wideband THz MIMO channel, with parameters as per for Systems I and II. The tuning parameter for the OMP [35] algorithm which impacts the convergence of the algorithm is set as $\epsilon_o = 1$ for both the systems, while the stopping criterion and the upper limit on the number of PA-BL iterations are defined as $\epsilon = 1$ and $K_{max} = 20$, respectively. The NMSE metric is defined as NMSE $= \frac{\sum_{k=0}^{K-1} \|\hat{\mathbf{H}}[k] - \mathbf{H}[k]\|_F^2}{\sum_{k=0}^{K-1} \|\mathbf{H}[k]\|_F^2}$. Observe from Fig. 3(a) and (b), that the NMSE performance improves significantly as one increases the number of training blocks M. The explanation for this lies in the use of greater number of pilots which leads to more reliable CSI estimates. It is evident from the plot that the proposed PA-BL technique yields a considerably improved performance over the OMP algorithm for the entire SNR range under consideration. The existing OMP approach performs poorly owing to its susceptibility to changes in the stopping criterion and dictionary matrix. As a consequence, even slight modifications to the cessation criterion or deviations from the dictionary matrix constraints can result in severe structural and convergence problems. The OMP approach also suffers from error propagation, because errors in the index selection cannot be corrected in subsequent iterations. It can be seen that the NMSE of the proposed PA-BL scheme improves by approximately ≈ 30 dB, when the SNR is varied from -10 dB to 20 dB. Upon increasing the number of pilot blocks M, the performance of the PA-BL scheme improves considerably. Interestingly, for a higher number of training frames such as $M = 90$, the proposed PA-BL algorithm performs close to the corresponding BCRLB. This
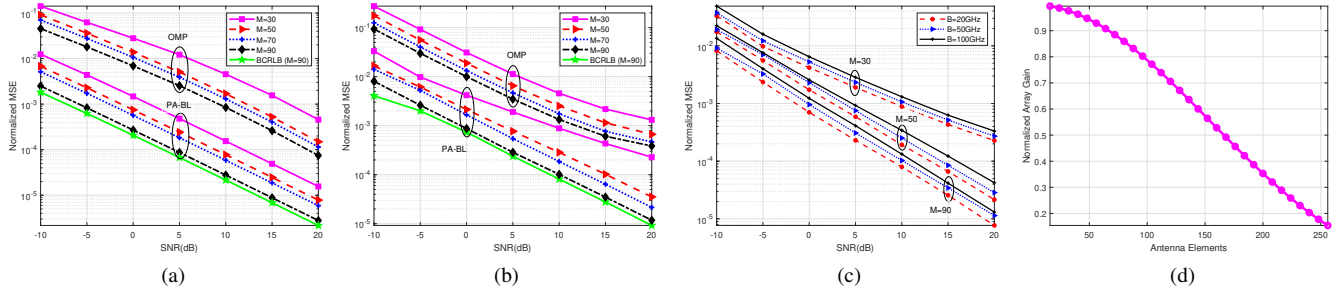
Figure 3. Performance of the proposed PA-BL technique with the existing OMP technique for (a)System-I (b)System-II. (c) NMSE performance of the proposed PA-BL algorithm versus SNR, for various values of the bandwidth $B \in \{20, 50, 100\}$ GHz, (d) Normalized array gain versus number of antenna elements with bandwidth $B = 20$GHz at $f_c = 1$ THz.

is of crucial importance, as the BCRLB is calculated under the assumption of perfect knowledge of the AoAs/AoDs, which is an ideal case. In stark contrast, the PA-BL technique does not require any such prior information. Further, as demonstrated in [36], upon convergence of the PA-BL algorithm, only the hyperparameters corresponding to the locations of the active AoA/ AoD pairs retain significant values, while the rest of the hyperparameters are driven to zero. By contrast, the performance of OMP and many other sparse recovery algorithms depends highly on the cessation criteria, which may be different for different propagation scenarios, thus leading to performance degradation, if not set carefully. Hence the proposed PA-BL algorithm is capable of learning the propagation environment in different network deployments. Fig. 3(c), shows the performance of the PA-BL technique for increasing signal bandwidth B. As discussed in Section III-C, when the user bandwidth increases, the impact of the beam squint progressively increases, which in turn degrades the performance of CSI estimation. It is apparent from Fig. 3(c) that the increased NMSE of the channel estimate at ultra-high bandwidths in the range $B = 100$ GHz in a dual-wideband THz system can be offset by augmenting the number of pilot blocks M. Moreover, the proposed PA-BL technique is seen to yield robust performance over a wide bandwidth range from 20 GHz to 100 GHz, even without any compensation for the dual-wideband effect, which is convenient for practical implementation. The proposed PA-BL framework is able to estimate the $N_R \in \{32, 64\} \times N_T \in \{32, 64\}$ dual-wideband THz MIMO channel using $M \in \{30, 50, 70, 90\}$ pilot blocks and $N_{RF} \in \{6, 16\}$, with corresponding MN_{RF} values of $\{180, 300, 420, 540\}$ for $N = 6$ and $\{480, 800, 1120, 1440\}$ for $N = 16$. Note that all these values satisfy $MN_{RF} \ll N_T N_R$, where $N_T N_R = 1024$ for System-I and $N_T N_R = 4096$ for System-II. This leads to an under-determined system in Equation (30), wherein the conventional LS/MMSE estimators cannot be applied for CSI estimation. In fact, the LS and MMSE require $M \approx 130$ pilot blocks for System-I and $M \approx 250$ pilot blocks for System-II, which increases the pilot overhead and reduces the spectral efficiency. This clearly illustrates the unique advantage of the proposed novel scheme for THz CSI estimation over the orthodox estimation paradigms. Fig. 3(d) depicts the variation of the normalized array gain with respect to the number of

antenna elements N_R. As N_R increases, the spatial wideband effect in the THz system exhibits a significant impact due to the notable delay between the first and last antenna elements, leading to a decrease in the array gain. Consequently, it can be inferred that as $N_R \to \infty$, $G(\theta, f_k) \to 0$. Furthermore, Fig. 4(c) illustrates the variation of the normalized array gain with respect to the bandwidth B. A large bandwidth induces a significant shift in the relative subcarrier frequency $\delta_k = \frac{f_k}{f_c}$, leading to a significant degradation in normalized array gain. The results for the DA-BL framework that can lead to a further improvement in the performance of the CSI estimation are presented next.

B. DA-BL wideband THz hybrid MIMO channel estimation

Fig. 4 and 5 demonstrate the performance improvement of the proposed PA-BL and DA-BL based techniques over traditional LS and other sparse estimation techniques such as OMP [35] and FOCUSS [34] along with a hypothetical scenario, where the true AoAs and AoDs are known, i.e, a Genie-aided transceiver for THz hybrid MIMO systems with parameter settings corresponding to System-I and System-II, as shown in Table II of the paper. Using the concatenated model for all the M blocks as discussed in the context of equation (64), the hybrid TPC $\mathbf{F}_{RF}$ and hybrid RC $\mathbf{W}_{RF}$ are chosen to be block-independent, and are set as described in equation (83). As discussed, the performance of the traditional LS-based scheme is inferior, because the technique does not leverage the sparsity occurring in the THz MIMO channel. The performance of the OMP algorithm is adversely affected by its sensitivity to the selection of the stopping criteria and the choice of the dictionary matrix. Furthermore, FOCUSS suffers from convergence-related problems and it is highly sensitive to the regularization parameter. On the other hand, the PA-BL framework leverages sparsity through EM-based hyperparameter estimation. Moreover, it eliminates the need for tuning or regularization parameters and it is robust to the choice of dictionary matrix. However, the exclusive reliance on pilot symbols makes it challenging to attain a precise CSI estimate. Therefore, the proposed DA-BL demonstrates a superior performance compared to other sparse estimation techniques due to its ability to utilize the data symbols coupled with a minimal pilot overhead. Furthermore the DA-BL framework beneficially leverages the data symbols in addition
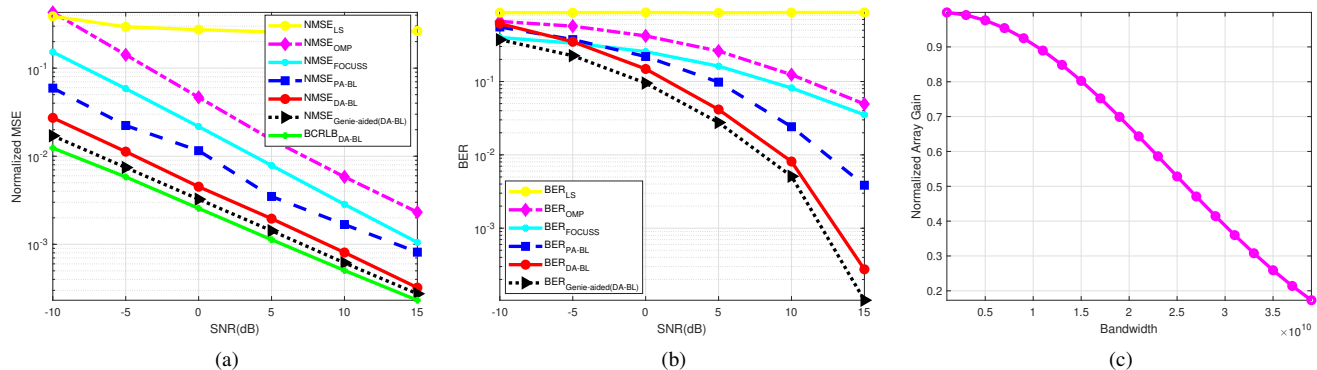
Figure 4. (a) NMSE versus SNR comparison of the proposed PA-BL and DA-BL techniques with the Genie-aided estimator, LS, OMP, FOCUSS and the BCRLB, (b) BER versus SNR comparison of the proposed PA-BL and DA-BL techniques with the Genie-aided estimator for System-I. (c) Normalized array gain versus bandwidth with $N_T = 128$ and $N_R = 128$ at $f_c = 1$THz.

to the pilot subcarriers to improve the channel estimate by iteratively maximizing the joint likelihood, as discussed in Section VI. The number of data vectors N_d that contain the information symbols for the DA-BL framework of System I and II is set as $N_d = \{100, 200\}$, respectively. The number of Monte Carlo iterations for the DA-BL framework is considered to be $DL_{max} = 20$ with the threshold of $\epsilon = 1$ as discussed in Algorithm 2. Additionally, it is clear from the figure that the performance of the proposed DA-BL framework approaches that of the Genie-aided system in the high SNR range, which evidences the efficiency of the proposed scheme. To benchmark the performance, the BCRLB is also plotted, as determined in Section V-A. The accuracy of data detection at the receiver using the MMSE receiver for the proposed PA-BL, DA-BL techniques and Genie-aided transceiver has been demonstrated by plotting the BER in Fig. 4(b) and similarly, for System-II, in Fig. 5 (b). From 4(b) one can further observe that there is performance improvement of approximately 4 dB is attained for the DA-BL framework over the PA-BL technique across the SNR range of -10 dB to 15 dB. For both the settings, the DA-BL-based receiver can be seen to yield a markedly improved performance in comparison to its PA-BL counterpart. Furthermore, the performance of the former, when System-II is considered, is also extremely close to that of a Genie-aided solution, and almost overlaps in the high-SNR regime, as it is clearly seen in Fig. 5(a). Fig. 5(c) illustrates the impact of the number of data-aided iterations on NMSE performance for SNR set as 10 dB. Importantly, the performance of the PA-BL framework and Genie-aided estimator are insensitive to the number of data decoding iterations and are thus constant in this figure. One can also observe that the NMSE of the proposed DA-BL scheme achieves that of the Genie-aided receiver for as few as 8 DA-BL iterations. This is of considerable importance, as the DA-BL framework achieves this without the need for prior knowledge of the true AoA/ AoD. This underscores its suitability for practical implementation in THz systems, where prior information is typically unavailable. Another noteworthy aspect is that, similar to the PA-BL receiver considered previously, for the setting with $M = 30$ pilot blocks and $N_d = 100$ data vectors,

the relation $(M + N_d)N_{RF} = 780 \ll N_T N_R = 1024$ once again holds true for the DA-BL receiver in System-I, which results in an under-determined estimation model as per Equation (30). Thus, this clearly shows that even in such a scenario, wherein the conventional LS and MMSE estimators are ineffective, the dual-wideband channel can be estimated precisely, lower complexity and a notably smaller training overhead, employing the proposed Data-Aided framework. Moreover, the Bayesian learning algorithm has demonstrated superior performance compared to other sparse recovery techniques in scenarios where the spatial wideband effect is negligible [6], [18]. However, when the spatial wideband effect is significant, both the CSI estimation accuracy and the resultant BER performance degrade. This deterioration is attributed to the subcarrier-dependent nature of the transmit and receive dictionary matrices i.e $\mathbf{A}_T(\Theta_T, f_k)$ and $\mathbf{A}_R(\Theta_R, f_k)$ respectively. To address this issue, we seek a sparse recovery technique, which is robust to the choice of dictionary matrix and ensures convergence. In this context, the PA-BL algorithm emerges as an ideal choice. Additionally, to enhance the estimation accuracy and mitigate the impact of the spatial wideband effect, this work also proposes a data-aided Bayesian learning technique. This technique iteratively incorporates data symbol estimates along with the development of a robust ZF based data symbol detector. Notably, other sparse signal recovery techniques such as OMP, FOCUSS and LASSO perform poorly in the presence of the spatial-wideband effect. Furthermore, there is no systematic approach to formulate their data-aided extensions. Thus, the relative performance improvement of the Bayesian learning based schemes over OMP/ FOCUSS/ LASSO is significant in the presence of the spatial-wideband effect. In short, the excellent CSI estimation performance, coupled with lower overheads and robust data-detection, render the proposed frameworks eminently suitable for application in dual-wideband THz MIMO systems.

VIII. CONCLUSIONS

New channel estimation paradigms considering angular domain sparsity in practical THz MIMO wideband channels have been conceived. The THz channel was modeled rigorously
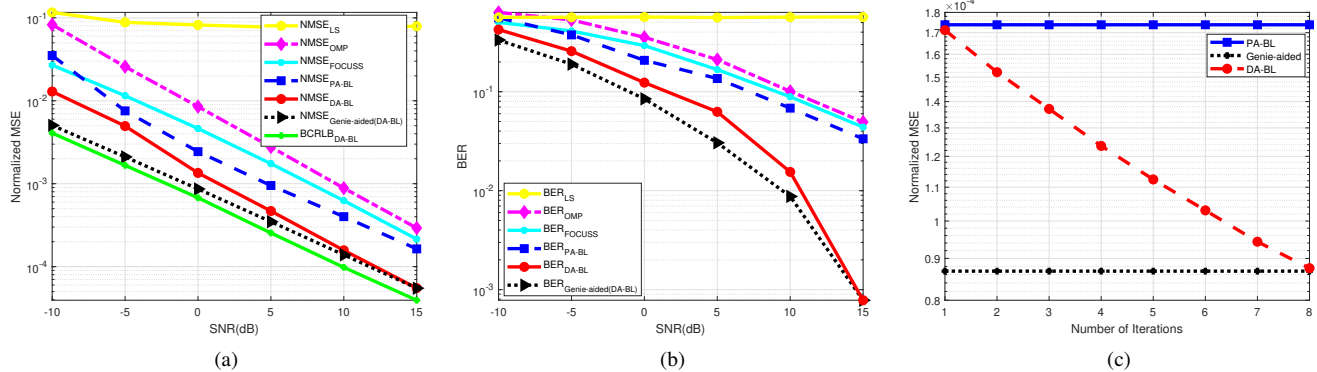
Figure 5. (a) NMSE versus SNR comparison of the proposed PA-BL and DA-BL techniques with the Genie-aided estimator, LS, OMP, FOCUSS and the BCRLB, (b) BER versus SNR comparison of the proposed PA-BL and DA-BL techniques with the Genie-aided estimator for System-II. (c) NMSE of the DA-BL scheme with increasing number of iterations of DA-BL framework.

using parameters derived from the HITRAN database, considering the free-space losses in addition to the dual-wideband effect naturally occurring in such systems. An SC-FDE based I/O model was developed to beneficially transform the TD model to its equivalent FD representation. Subsequently, exploiting the sparsity of the dual-wideband THz channel, novel Bayesian learning approaches were proposed for CSI recovery. To begin with, the PA-BL methodology was conceived for channel estimation, that uses only pilot symbols, followed by the DA-BL approach for superior performance, that leverages the concatenated pilot and data outputs. By explicitly deriving the BCRLBs, we characterize the estimation performance of PA-BL and DA-BL frameworks. A comprehensive set of simulation results was presented that evidences the NMSE and BER performance benefits of the various schemes formulated in this paper. In particular, the NMSE of the DA-BL paradigm was seen to closely follow that of the Genie-aided transceiver and the BCRLB references at high SNRs, a feature that shows the advantages of practical deployment of the different algorithms devised.

APPENDIX A
CALCULATION OF THE MATRIX $\boldsymbol{\Xi}^{(j)}[k]$

The optimization cost function in (78) can be simplified as
$$\mathbb{E}\left\{\|\mathbf{Y}_d[k] - \mathbf{H}_{\text{eq}}[k]\mathbf{X}_d[k]\|_F^2\right\} =$$

$$\text{Tr}\bigg\{\mathbf{Y}_d^H[k]\mathbf{Y}_d[k] - \mathbf{Y}_d^H[k]\hat{\mathbf{H}}_{\text{eq}}^{(j)}[k]\mathbf{X}_d[k] -$$

$$\mathbf{X}_d^H[k]\left(\hat{\mathbf{H}}_{\text{eq}}^{(j)}[k]\right)^H \mathbf{Y}_d[k] - \mathbf{X}_d^H[k]\mathbb{E}\left\{\mathbf{H}_{\text{eq}}^H[k]\mathbf{H}_{\text{eq}}[k]\right\}\mathbf{X}_d[k]\bigg\},$$
$$(84)$$

where $\hat{\mathbf{H}}_{\text{eq}}^{(j)}[k]$ is given by (81). Also the quantity $\mathbb{E}\left\{\mathbf{H}_{\text{eq}}^H[k]\mathbf{H}_{\text{eq}}[k]\right\}$ is expressed as

$$\mathbb{E}\left\{\mathbf{H}_{\text{eq}}^H[k]\mathbf{H}_{\text{eq}}[k]\right\} = \left(\hat{\mathbf{H}}_{\text{eq}}^{(j)}[k]\right)^H \hat{\mathbf{H}}_{\text{eq}}^{(j)}[k] + \boldsymbol{\Xi}^{(j)}[k]. \quad (85)$$

The procedure for computation of the matrix $\boldsymbol{\Xi}^{(j)}[k]$ is described next. Let $\mathbf{h}_{\text{eq}}[k] \in \mathbb{C}^{N_{RF}^2 \times 1}$ be the vectorized form of the equivalent channel matrix given by $\mathbf{h}_{\text{eq}}[k] = \text{vec}\left(\mathbf{H}_{\text{eq}}[k]\right)$ and its estimate $\hat{\mathbf{h}}_{\text{eq}}^{(j)}[k]$ be defined as $\hat{\mathbf{h}}_{\text{eq}}^{(j)}[k] = \text{vec}\left(\hat{\mathbf{H}}_{\text{eq}}^{(j)}[k]\right)$.

Its error covariance matrix $\boldsymbol{\Sigma}_{\text{eq}}^{(j)}[k] \in \mathbb{C}^{N_{RF}^2 \times N_{RF}^2}$ can be derived by substituting the value of $\hat{\mathbf{H}}_{\text{eq}}^{(j)}[k]$ from (81) given

$$\boldsymbol{\Sigma}_{\text{eq}}^{(j)}[k] = \widetilde{\boldsymbol{\Phi}}[k]\boldsymbol{\Sigma}^{(j)}[k]\widetilde{\boldsymbol{\Phi}}^H[k], \quad (86)$$

where $\widetilde{\boldsymbol{\Phi}}[k] = \left[\mathbf{F}_{RF}^T \otimes \mathbf{W}_{RF}^H\right]\boldsymbol{\Psi}[k] \in \mathbb{C}^{N_{RF}^2 \times G_R G_T}$. Subsequently, the (a,b)th element of the matrix $\boldsymbol{\Xi}^{(j)}[k]$ is defined

$$\boldsymbol{\Xi}^{(j)}[k](a,b) = \text{Tr}\big[\boldsymbol{\Sigma}_{\text{eq}}^{(j)}[k](\tilde{a} - N_{RF} + 1 : \tilde{a},$$
$$\tilde{b} - N_{RF} + 1 : \tilde{b})\big], \quad (87)$$

where $\tilde{a} = aN_{RF}$ and $\tilde{b} = bN_{RF}$.

APPENDIX B
BAYESIAN CRAMÉR-RAO LOWER BOUND FOR DATA-AIDED LEARNING

Let $\mathbf{J}_{\text{DA}}[k] \in \mathbb{C}^{G_R G_T \times G_R G_T}$ denote the combined Bayesian FIM, which can be evaluated as
$$\mathbf{J}_{\text{DA}}[k] = \mathbf{J}_{d,\text{DA}}[k] + \mathbf{J}_{b,\text{DA}}[k], \quad (88)$$

where $\mathbf{J}_{d,\text{DA}}[k] \in \mathbb{C}^{G_R G_T \times G_R G_T}$ denotes the FIM associated with the data output, while $\mathbf{J}_{b,\text{DA}}[k]$ represents the FIM corresponding to the prior information of the beamspace CSI $\mathbf{h}_b[k]$. The above quantities can be computed as

$$\mathbf{J}_{d,\text{DA}}[k] = -\mathbb{E}_{\mathbf{y}[k], \mathbf{h}_b[k]}\left\{\frac{\partial^2 \mathcal{L}(\mathbf{y}[k]|\mathbf{h}_b[k]; \mathbf{X}_d[k])}{\partial \mathbf{h}_b[k]\partial \mathbf{h}_b^H[k]}\right\}$$

$$\mathbf{J}_{b,\text{DA}}[k] = -\mathbb{E}_{\mathbf{h}_b[k]}\left\{\frac{\partial^2 \mathcal{L}(\mathbf{h}_b[k]; \boldsymbol{\Gamma}_k)}{\partial \mathbf{h}_b[k]\partial \mathbf{h}_b^H[k]}\right\}. \quad (89)$$

Using the conditional PDF as obtained in (67), the FIM $\mathbf{J}_{d,\text{DA}}[k]$ can be obtained as
$$\mathbf{J}_{d,\text{DA}}[k] = \tilde{\boldsymbol{\Psi}}[k]\mathbf{R}_v^{-1}\tilde{\boldsymbol{\Psi}}[k]. \quad (90)$$

Similarly, as obtained in the PA-BL, the FIM associated with $\mathbf{h}_b[k]$ is obtained using the prior PDF from equation (68), and the corresponding expression for $\mathbf{J}_b[k]$ can be obtained as
$$\mathbf{J}_{b,\text{DA}}[k] = \hat{\boldsymbol{\Gamma}}_k^{-1}. \quad (91)$$

Using the expressions in (90) and (91), the Bayesian FIM can be represented as
$$\mathbf{J}_{\text{DA}}[k] = \tilde{\boldsymbol{\Psi}}^H[k]\mathbf{R}_v^{-1}\tilde{\boldsymbol{\Psi}}[k] + \hat{\boldsymbol{\Gamma}}_k^{-1}. \quad (92)$$

Thus, the lower bound on the MSE of the estimate $\hat{\mathbf{h}}_b[k]$, can be obtained as

$$\mathrm{MSE}(\hat{\mathbf{h}}_b[k]) = \mathbb{E}\left\{ \| \hat{\mathbf{h}}_b[k] - \mathbf{h}_b[k] \|^2 \right\} \geq \mathrm{Tr}\left\{ \mathbf{J}_{\mathrm{DA}}^{-1}[k] \right\}$$
$$= \mathrm{Tr}\left\{ \left(\tilde{\boldsymbol{\Psi}}^H[k]\mathbf{R}_v^{-1}\tilde{\boldsymbol{\Psi}}[k] + \hat{\boldsymbol{\Gamma}}_k^{-1} \right)^{-1} \right\}.$$
(93)

Exploiting the relationship between the beamspace representation of the THz MIMO channel and its vectorized version, the BCRLB for the estimated CSI is finally given by

$$\mathrm{MSE}(\hat{\mathbf{H}}[k]) \geq \mathrm{Tr}\left\{ \boldsymbol{\Psi}[k]\mathbf{J}_{\mathrm{DA}}^{-1}[k]\boldsymbol{\Psi}^H[k] \right\}.$$
(94)

REFERENCES

[1] I. F. Akyildiz, C. Han, Z. Hu, S. Nie, and J. M. Jornet, "Terahertz band communication: An old problem revisited and research directions for the next decade," *IEEE Transactions on Communications*, vol. 70, no. 6, pp. 4250–4285, 2022.

[2] A. Shafie, N. Yang, C. Han, J. M. Jornet, M. Juntti, and T. Kurner, "Terahertz communications for 6G and beyond wireless networks: Challenges, key advancements, and opportunities," *IEEE Network (Early Access)*, pp. 1–8, 2022.

[3] H. Sarieddeen, N. Saeed, T. Y. Al-Naffouri, and M.-S. Alouini, "Next generation Terahertz communications: A rendezvous of sensing, imaging, and localization," *IEEE Communications Magazine*, vol. 58, no. 5, pp. 69–75, 2020.

[4] X. Yang, M. Li, Y. Wang, and Y. Zeng, "Atmospheric absorption loss estimation of Terahertz-wave band satellite ground detection," in *Proceedings of the 18th IEEE International Conference on Communication Technology*, 2018, pp. 741–744.

[5] H. Sarieddeen, M.-S. Alouini, and T. Y. Al-Naffouri, "An overview of signal processing techniques for Terahertz communications," *Proceedings of the IEEE*, vol. 109, no. 10, pp. 1628–1665, 2021.

[6] S. Srivastava, A. Tripathi, N. Varshney, A. K. Jagannatham, and L. Hanzo, "Hybrid transceiver design for Tera-hertz MIMO systems relying on Bayesian learning aided sparse channel estimation," *IEEE Transactions on Wireless Communications*, vol. 22, no. 4, pp. 2231–2245, 2023.

[7] A. Faisal, H. Sarieddeen, H. Dahrouj, T. Y. Al-Naffouri, and M.-S. Alouini, "Ultra-massive MIMO systems at Terahertz bands: Prospects and challenges," *IEEE Vehicular Technology Magazine*, vol. 15, no. 4, pp. 33–42, 2020.

[8] Y.-P. Lin, "Hybrid MIMO-OFDM beamforming for wideband mmWave channels without instantaneous feedback," *IEEE Transactions on Signal Processing*, vol. 66, no. 19, pp. 5142–5151, 2018.

[9] H. Li, M. Li, Q. Liu, and A. L. Swindlehurst, "Dynamic hybrid beamforming with low-resolution PSs for wideband mmWave MIMO-OFDM systems," *IEEE Journal on Selected Areas in Communications*, no. 9, pp. 2168–2181, 2020.

[10] B. Ning, Z. Tian, Z. Chen, C. Han, J. Yuan, and S. Li, "Prospective beamforming technologies for ultra-massive MIMO in Terahertz communications: A tutorial," *arXiv preprint arXiv:2107.03032*, 2021.

[11] H. Elayan, O. Amin, R. M. Shubair, and M.-S. Alouini, "Terahertz communication: The opportunities of wireless technology beyond 5G," in *2018 International Conference on Advanced Communication Technologies and Networking (CommNet)*. IEEE, 2018, pp. 1–5.

[12] L. Dai, J. Tan, Z. Chen, and H. V. Poor, "Delay-phase precoding for wideband THz massive MIMO," *IEEE Transactions on Wireless Communications*, vol. 21, no. 9, pp. 7271–7286, 2022.

[13] M. Jian, F. Gao, Z. Tian, S. Jin, and S. Ma, "Angle-domain aided UL/DL channel estimation for wideband mmwave massive MIMO systems with beam squint," *IEEE Transactions on Wireless Communications*, vol. 18, no. 7, pp. 3515–3527, 2019.

[14] J. M. Jornet and I. F. Akyildiz, "Channel modeling and capacity analysis for electromagnetic wireless nanonetworks in the Terahertz band," *IEEE Transactions on Wireless Communications*, vol. 10, no. 10, pp. 3211–3221, 2011.

[15] A. Saeed, O. Gurbuz, and M. A. Akkas, "Terahertz communications at various atmospheric altitudes," *Physical Communication*, vol. 41, p. 101113, 2020.

[16] C. Lin and G. Y. L. Li, "Terahertz communications: An array-of-subarrays solution," *IEEE Communications Magazine*, vol. 54, no. 12, pp. 124–131, 2016.

[17] Y. Zhang, Y. Huo, D. Wang, X. Dong, and X. You, "Channel estimation and hybrid precoding for distributed phased arrays based MIMO wireless communications," *IEEE Transactions on Vehicular Technology*, no. 11, pp. 12 921–12 937, 2020.

[18] S. Srivastava, J. Nath, and A. K. Jagannatham, "Data aided quasistatic and doubly-selective CSI estimation using affine-precoded superimposed pilots in millimeter wave MIMO-OFDM systems," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 7, pp. 6983–6998, 2021.

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

[20] S. Srivastava, P. Sharma, S. Dwivedi, A. K. Jagannatham, and L. Hanzo, "Fast block LMS based estimation of angularly sparse channels for single-carrier wideband millimeter wave hybrid MIMO systems," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 1, pp. 666–681, 2021.

[21] X. Gao, L. Dai, S. Zhou, A. M. Sayeed, and L. Hanzo, "Wideband beamspace channel estimation for millimeter-wave MIMO systems relying on lens antenna arrays," *IEEE Transactions on Signal Processing*, no. 18, pp. 4809–4824, 2019.

[22] B. Wang, F. Gao, S. Jin, H. Lin, and G. Y. Li, "Spatial-and frequency-wideband effects in millimeter-wave massive MIMO systems," *IEEE Transactions on Signal Processing*, vol. 66, no. 13, pp. 3393–3406, 2018.

[23] B. Wang, M. Jian, F. Gao, G. Y. Li, and H. Lin, "Beam squint and channel estimation for wideband mmWave massive MIMO-OFDM systems," *IEEE Transactions on Signal Processing*, vol. 67, no. 23, pp. 5893–5908, 2019.

[24] K. Dovelos, M. Matthaiou, H. Q. Ngo, and B. Bellalta, "Channel estimation and hybrid combining for wideband Terahertz massive MIMO systems," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 6, pp. 1604–1620, 2021.

[25] T.-H. Chou, N. Michelusi, D. J. Love, and J. V. Krogmeier, "Dual-Wideband Time-Varying Sub-Terahertz Massive MIMO Systems: A Compressed Training Framework," *arXiv preprint arXiv:2201.00992*, 2022.

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

[27] Z. Sha and Z. Wang, "Channel estimation and equalization for Terahertz receiver with RF impairments," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 6, pp. 1621–1635, 2021.

[28] A. M. Elbir, W. Shi, K. V. Mishra, and S. Chatzinotas, "Federated Learning for THz Channel Estimation," *arXiv preprint arXiv:2207.06017*, 2022.

[29] M. Wang, F. Gao, N. Shlezinger, M. F. Flanagan, and Y. C. Eldar, "A block sparsity based estimator for mmWave massive MIMO channels with beam squint," *IEEE Transactions on Signal Processing*, vol. 68, pp. 49–64, 2019.

[30] S. Ju, Y. Xing, O. Kanhere, and T. S. Rappaport, "Millimeter wave and sub-terahertz spatial statistical channel model for an indoor office building," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 6, pp. 1561–1575, 2021.

[31] R. Piesiewicz, C. Jansen, D. Mittleman, T. Kleine-Ostmann, M. Koch, and T. Kurner, "Scattering analysis for the modeling of THz communication systems," *IEEE Transactions on Antennas and Propagation*, vol. 55, no. 11, pp. 3002–3009, 2007.

[32] J. Rodriguez-Fernandez and N. Gonzalez-Prelcic, "Channel estimation for frequency-selective mmwave MIMO systems with beam-squint," in *Proceedings of the IEEE Global Communications Conference (GLOBECOM)*, 2018, pp. 1–6.

[33] C. Ekanadham, D. Tranchina, and E. P. Simoncelli, "Recovery of sparse translation-invariant signals with continuous basis pursuit," *IEEE transactions on signal processing*, vol. 59, no. 10, pp. 4735–4744, 2011.

[34] S. F. Cotter, B. D. Rao, K. Engan, and K. Kreutz-Delgado, "Sparse solutions to linear inverse problems with multiple measurement vectors," *IEEE Transactions on Signal Processing*, vol. 53, no. 7, pp. 2477–2488, 2005.

[35] A. Kulkarni and T. Mohsenin, "Low overhead architectures for OMP compressive sensing reconstruction algorithm," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 64, no. 6, pp. 1468–1480, 2017.

[36] D. P. Wipf and B. D. Rao, "Sparse Bayesian learning for basis selection," *IEEE Transactions on Signal processing*, vol. 52, no. 8, pp. 2153–2164, 2004.

[37] A. Mishra, A. Rajoriya, A. K. Jagannatham, and G. Ascheid, "Sparse Bayesian learning-based channel estimation in millimeter wave hybrid MIMO systems," in *Proceedings of the 18th IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2017, pp. 1–5.

[38] S. Alzeer, S. Almatrudi, H. A. Bukhari, Y. Han, Y. Ding, and B. D. Rao, "Millimeter wave channel estimation using data-aided DoA estimation," in *Proceedings of the 52nd IEEE Asilomar Conference on Signals, Systems, and Computers*, 2018, pp. 1567–1571.

[39] A. Garg, S. Srivastava, N. Yadav, A. K. Jagannatham, and L. Hanzo. "Technical Report: Angularly Sparse Channel Estimation in Dual-Wideband Tera-Hertz (THz) Hybrid MIMO Systems Relying on Bayesian Learning" IIT Kanpur, Tech. Rep., 2023. [Online]. Available: https://www.iitk.ac.in/mwn/documents/MWNLab_TR_THz2023.pdf

[40] C. Hill. (1999) HITRAN Online Interface to the HITRAN Database. [Online]. Available: https://hitran.org/

Nimish Yadav received the B.Tech. degree in Electronics Engineering from Kamla Nehru Institute of Technology, Sultanpur, India, in 2020 and the M.Tech degree in Signal Processing and Communication from Department of Electrical Engineering, Indian Institute of Technology Kanpur, Kanpur, India, in 2022. He is currently working with Qualcomm India Pvt. Ltd., Chennai, India, designated as a Hardware Systems Engineer in Wireless LANs. His research interests include sparse signal processing and THz communication. He was awarded with the Bronze medal from Dr. A.P.J. Abdul Kalam Technical University, Uttar Pradesh, India for academic excellence.

Abhisha Garg (Graduate Student Member, IEEE) received the B.Tech degree in Electronics & Communication from Graphic Era University, Dehradun, India, in 2016, and the M.Tech degree in Signal Processing from Guru Gobind Singh Indraprastha University, New Delhi, India, in 2020. She is currently working toward the Ph.D. degree with the Department of Electrical Engineering, Indian Institute of Technology Kanpur, Kanpur, India. Her area of research includes THz communications, sparse signal processing, machine learning and its applications. She was awarded with the Gold medal from Graphic Era University, Dehradun, India for academic excellence. She was also awarded the Young Scientist award from International Science Community Association (ISCA), in 2017. She was a recipient of IEEE Signal Processing Society (SPS) scholarship, 2023. She was also a finalist for the Qualcomm Innovation Fellowship (QIF), 2023.

Aditya K. Jagannatham (Senior Member, IEEE) received the bachelor's degree from the 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 (QCT) 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 (QIF 2018, 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 Diego and the Qualcomm 6G UR India gift.

Suraj Srivastava (Member, IEEE) received the M.Tech. degree in Electronics and Communication Engineering from Indian Institute of Technology Roorkee, India, in 2012, and Ph.D. degree in Electrical Engineering from Indian Institute of Technology Kanpur, India, in 2022. From July 2012 to November 2013, he was employed as a Staff-I systems design engineer with Broadcom Research India Pvt. Ltd., Bangalore, and from November 2013 to December 2015, he was employed as a lead engineer with Samsung Research India, Bangalore where he worked on developing layer-2 of the 3G UMTS/WCDMA/HSDPA modem. He worked as a senior lead engineer in Qualcomm India Pvt. Ltd., Bangalore, from October 2022 to November 2023. His research interests include applications of Sparse Signal Processing in 5G Wireless Systems, mmWave and Tera-Hertz Communication, Orthogonal Time-Frequency Space (OTFS), Joint Radar and Communication (RadCom), Optimization and Machine Learning. He was awarded Qualcomm Innovation Fellowship (QIF) in year 2018 and 2022 from Qualcomm. He was also awarded Outstanding Ph.D. Thesis and Outstanding Teaching Assistant awards from IIT Kanpur. He is currently an Assistant Professor at the Department of Electrical Engineering, Indian Institute of Technology Jodhpur.

Lajos Hanzo (Life Fellow, IEEE) received Honorary Doctorates from the Technical University of Budapest (2009) and Edinburgh University (2015). He is a Foreign Member of the Hungarian Science-Academy, Fellow of the Royal Academy of Engineering (FREng), of the IET, of 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.