Sparse, Group-Sparse and Online Bayesian Learning Aided Channel Estimation for Doubly-Selective mmWave Hybrid MIMO OFDM Systems

Suraj Srivastava, *Graduate Student Member, IEEE,* Ch Suraj Kumar Patro, Aditya K. Jagannatham, *Member, IEEE,* Lajos Hanzo, *Fellow, IEEE*

***Abstract*—Sparse, group-sparse and online channel estimation is conceived for millimeter wave (mmWave) multiple-input multiple-output (MIMO) orthogonal frequency division multiplexing (OFDM) systems. We exploit the angular sparsity of the mmWave channel impulse response (CIR) to achieve improved estimation performance. First a sparse Bayesian learning (SBL)-based technique is developed for the estimation of each individual subcarrier's quasi-static channel, which leads to an improved performance versus complexity trade-off in comparison to conventional channel estimation. Then a novel group-sparse Bayesian learning (G-SBL) scheme is conceived for reducing the channel estimation mean square error (MSE). The salient aspect of our G-SBL technique is that it exploits the frequency-domain (FD) correlation of the channel's frequency response (CFR), while transmitting pilots on only a few subcarriers, thus it has a reduced pilot overhead. A low complexity (LC) version of G-SBL, termed LCG-SBL, is also developed that reduces the computational cost of the G-SBL significantly. Subsequently, an online G-SBL (O-SBL) variant is designed for the estimation of doubly-selective mmWave MIMO OFDM channels, which has low processing delay and exploits temporal correlation as well. This is followed by the design of a hybrid transmit precoder and receive combiner, which can operate directly on the estimated beamspace domain CFRs, together with a limited channel state information (CSI) feedback. Our simulation results confirms the accuracy of the analysis.**

***Index Terms*—mmWave, MIMO, OFDM, channel estimation, sparse Bayesian learning, Cramer-Rao bound, hybrid signal processing.**

I. Introduction

Millimeter Wave (mmWave) wireless technology [1], [2], has emerged as one of the promising candidates for next-generation networks. Orthogonal frequency division multiplexing (OFDM) [3] is eminently suitable for mmWave MIMO systems by providing resilience to the multipath distortion and inter-symbol-interference (ISI), which are capable of accommodating numerous antenna elements in a shirt-pocket-sized handset. However, the large antenna arrays of mmWave MIMO systems lead to a high hardware complexity and power consumption due to the high sampling rate of ADC/ DACs and owing to the large number of power amplifiers (PAs). For tackling this challenge, the hybrid MIMO signal processing architecture that requires a much reduced number of radio-frequency (RF) chains in comparison to the number of antennas, has gained significant popularity as a viable solution for mmWave MIMO OFDM systems [4]–[6]. In contrast to conventional communication systems where the bulk of the signal processing operations such as precoding and combining are performed exclusively in the baseband, in such a hybrid MIMO OFDM transceiver, the signal processing tasks are partitioned between the analog and digital domains. However, the success of this novel architecture in mmWave MIMO OFDM systems depends critically on the accuracy of the channel state information (CSI). Hence, the accuracy of channel estimation determines the attainable gains of mmWave MIMO OFDM systems [4], [6]. A brief review of the contributions that address this problem is described next.

A. Review of Existing Works in mmWave MIMO OFDM Channel Estimation

The conventional least squares (LS) and minimum mean squared error (MMSE) approaches are not well-suited for mmWave hybrid MIMO OFDM systems due to the low SNR and large sizes of the antenna arrays at both the transmitter and receiver. Furthermore, the conventional approaches are also inefficient and lead to sub-optimal performance since they ignore the angular-sparsity of the mmWave MIMO channel [1], which can be jointly attributed to the reduced scattering and diffraction effects in the mmWave regime as well as to the highly focused beam impinging from the large antenna arrays. The authors of [7], [8] proposed various beam training strategies for CSI acquisition, where pilot beams are used for estimating the angles of arrival and departure (AoA/ AoD) pairs of the multipath components. Although this technique has an appealingly low complexity for a limited number of users and for coarse angular resolutions, it requires a large amount of feedback that increases substantially for finer angular resolutions. Sparse signal estimation schemes, such as those in [9]–[11], offer sophisticated CSI acquisition alternatives, and are

L. Hanzo would like to acknowledge the financial support of the Engineering and Physical Sciences Research Council projects EP/P034284/1 and EP/P003990/1 (COALESCE) as well as of the European Research Council's Advanced Fellow Grant QuantCom (Grant No. 789028). A. K. Jagannatham would like to acknowledge the research supported in part by the Science and Engineering Research Board (SERB), Department of Science and Technology, Government of India, in part by the Space Technology Cell, IIT Kanpur, in part by the IIMA IDEA Telecom Centre of Excellence, in part by the Qualcomm Innovation Fellowship, and in part by the Arun Kumar Chair Professorship.

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

C. S. K. Patro is with Qualcomm, Hyderabad, Telangana 500081, India (email: cpatro@qti.qualcomm.com).

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

well-suited for mmWave channel estimation. The foundation of these propositions is the equivalent representation of the mmWave MIMO channel in the beamspace domain, which is known to be *sparse*, and is hence amenable to sparse signal recovery using a low pilot overhead. This paradigm has only recently been extended to frequency-selective channels in [6], wherein the authors conceived a scheme for sparse estimation of a wideband mmWave MIMO channel via sparse representation of the concatenated frequency-selective mmWave MIMO channel in a single-carrier system. The associated sparse signal recovery problem therein is solved using the orthogonal matching pursuit (OMP) scheme that performs greedy selection of the columns of the dictionary matrix with the aim of obtaining a sparse approximation of the pilot vector observed [6]. This concept has then also been extended to mmWave MIMO OFDM systems using the simultaneous OMP (SOMP) technique of [4]. However, the performance of the OMP technique, as well as of its offshoot, the SOMP, is heavily reliant both on the choice of the dictionary matrix and of the stopping criterion, which renders them susceptible to convergence errors, in turn leading to performance erosion. Furthermore, the framework of [4] advocates the transmission of pilot beams on all the subcarriers, hence leading to spectral inefficiency and to potentially excessive processing delays. In this context, sparse Bayesian learning (SBL) [12] has been shown to lead to improved sparse estimation performance over several existing approaches, such as OMP [13], FOCUSS [14], Basis Pursuit (BP) [15] etc. SBL has been subsequently extended in [16] and [17] to exploit simultaneous sparsity as well as temporal correlation arising from multiple measurements. SBL has been exploited for sparse channel estimation of a quasi-static narrowband flat-fading mmWave hybrid MIMO system in [18]. This has been subsequently extended to a time-selective sparse mmWave MIMO channel using SBL-based hierarchical Bayesian Kalman filtering in [10]. The authors of [19] and [20] developed an SBL-based time-domain (TD) sparse channel estimation schemes for a single-carrier (SC) hybrid wideband mmWave MIMO system, considering quasi-static frequency-selective and doubly-selective fading scenarios. However, none of the existing contributions explore the SBL paradigm for quasi-static frequency-selective and doubly-selective mmWave MIMO OFDM systems, where the channel's frequency response (CFR) is group-sparse, and also exhibits frequency- and time-domain (FD and TD) correlation. This motivates the development of novel sparse mmWave MIMO OFDM channel estimation techniques and hybrid precoder/ combiner designs for circumventing the drawbacks of the existing schemes listed above. The contributions of the paper are succinctly summarized next. In Table-I, we boldly contrast our contributions to the state-of-the-art.

B. Contributions

1) This work proposes an SBL-based *group-sparse* channel estimation paradigm for channel estimation in mmWave hybrid MIMO OFDM systems. The proposed G-SBL scheme performs joint estimation of the mmWave MIMO OFDM CFRs using only a few pilot subcarriers, which leads to a reduced pilot overhead and improved spectral efficiency. Furthermore, another interesting aspect of this scheme is that it exploits the FD correlation of the mmWave MIMO OFDM CFR, which is novel in the context of the mmWave literature.
2) An efficient low complexity (LC) version of G-SBL termed as LCG-SBL is also developed, which significantly cuts down the computational cost of G-SBL by replacing the inversion of a large-dimensional matrix by the inversion of two smaller matrices. This results in a substantial complexity reduction, because the complexity of matrix inversion is related to the cube of its dimensions.
3) Next, an online G-SBL (O-SBL)-based CSI estimation scheme is developed for doubly-selective mmWave MIMO OFDM systems, which exploits both the FD and TD correlation, while simultaneously also guaranteeing convergence to sparse estimates upon proper initialization. In contrast to the family of block processing schemes, such as SBL-KF [20] and HBKF [10], the proposed O-SBL technique employs the sequential linear MMSE (LMMSE) principle for updating the CSI in each training frame. Thus, it has a low complexity as well as low processing delay, which makes it appealing for practical implementation. To the best of our knowledge, the frame-wise CSI tracking aspect of the O-SBL is novel in the context of mmWave MIMO OFDM systems.
4) Finally, a simplified technique is developed for joint hybrid transceiver design across all the subcarriers, which directly employs the beamspace channel estimates obtained via the proposed estimation techniques. The proposed algorithm requires only limited CSI of the beamspace channel, namely the non-zero coefficients and their respective indices, which substantially reduces the feedback required. Furthermore, in contrast to the existing SOMP [21] and MSBL [22] based designs, the proposed hybrid transceiver design is non-iterative, and hence computationally efficient.
5) Simulation results are presented for comparing the MSE as well as the spectral efficiency of the proposed schemes to the existing schemes and to the analytical benchmarks.

C. Notations

The following notation is used throughout the paper. Boldface small and capital alphabets are used to denote vectors and matrices, respectively. $\mathbf{X}\left(\mathcal{S},:\right)$ and $\mathbf{X}\left(:,\mathcal{S}\right)$ denote the submatrices formed by extracting the rows and columns respectively, of the matrix $\mathbf{X}$, corresponding to indices given in the set $\mathcal{S}$. $\mathcal{CN}\left(\boldsymbol{\mu},\boldsymbol{\Sigma}\right)$ denotes a complex Gaussian random vector with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. The vector obtained by stacking the columns of matrix $\mathbf{X}$ is denoted by $\mathrm{vec}\left(\mathbf{X}\right)$ and $\mathrm{vec}^{-1}(\mathbf{x})$ represents a matrix obtained by the corresponding inverse vectorization operation. $\mathbf{X}\otimes\mathbf{Y}$ denotes the Kronecker product of matrices $\mathbf{X}$ and $\mathbf{Y}$. The following property of the $\mathrm{vec}\left(\cdot\right)$ operator and the Kroneker product [23] is used in the paper:

$$\mathrm{vec}\left(\mathbf{ABC}\right)=\left(\mathbf{C}^{T}\otimes\mathbf{A}\right)\mathrm{vec}\left(\mathbf{B}\right). \tag{1}$$

TABLE I: Our contributions in contrast to the state-of-the-art

	[9]	[18]	[10]	[6]	[4]	[19]	[20]	**Proposed**
Wideband CSI acquisition	×	×	×	✓	✓	✓	✓	✓
Multicarrier (OFDM)	×	×	×	×	✓	×	×	✓
Group/ block/ simultaneous sparsity	×	×	✓	×	✓	✓	✓	✓
Pilots on a few subcarriers	×	×	×	×	×	×	×	✓
Exploiting FD correlation	×	×	×	×	×	×	×	✓
Time-frequency Interpolation	×	×	×	×	×	×	×	✓
Quasi-static CSI acquisition	✓	✓	✓	✓	✓	✓	×	✓
Time-selective CSI acquisition	×	×	✓	×	×	×	✓	✓
Frame-wise CSI updation	×	×	×	×	×	×	×	✓
Low processing delay	×	×	×	×	×	×	×	✓
BCRLB bounds	×	✓	✓	×	×	✓	✓	✓
Limited CSI feedback	×	×	×	×	×	×	×	✓
Hybrid transceiver design based on estimated beamspace domain CSI	×	×	✓	×	×	×	×	✓

II. mmWave MIMO OFDM system model

Consider a mmWave hybrid MIMO OFDM system having N_T transmit antennas (TAs), N_R receive antennas (RAs) and K subcarriers. The transmitter and receiver employ N_{RF}^T and N_{RF}^R RF chains, respectively, in order to transmit N_s data streams, where $N_s \leq \min\left(N_{RF}^T, N_{RF}^R\right) \ll \min\left(N_T, N_R\right)$. We consider a frequency-selective mmWave MIMO channel having L delay taps, with the lth delay tap represented by the matrix $\mathbf{H}_l \in \mathbb{C}^{N_R \times N_T}, \forall\ 0 \leq l \leq L-1$. On each subcarrier k, the mmWave MIMO system, employs a cascade of the frequency-selective baseband precoder $\mathbf{F}_{BB}[k] \in \mathbb{C}^{N_{RF}^T \times N_s}$ followed by the frequency-flat RF precoder $\mathbf{F}_{RF} \in \mathbb{C}^{N_T \times N_{RF}^T}$. Similarly, at the receiver, the received signal is processed by a cascade of the frequency-flat RF combiner $\mathbf{W}_{RF} \in \mathbb{C}^{N_R \times N_{RF}^R}$ followed by the frequency-selective baseband combiner $\mathbf{W}_{BB}[k] \in \mathbb{C}^{N_{RF}^R \times N_s}$ on each subcarrier. The input symbols of each RF chain at the transmitter are processed using K-point inverse fast Fourier transforms (IFFTs) followed by zero-padding (ZP) of length L. Similarly, at the receiver, the output block of length $K+L$ of each RF chain is processed using the *overlap and add* technique [24] followed by a K-point FFT operation. The choice of ZP is preferred instead of the traditional cyclic prefix (CP) in mmWave MIMO OFDM systems in order to enable reconfiguration of the analog circuitry [4], [6]. The signal $\mathbf{y}[k] \in \mathbb{C}^{N_s \times 1}$ at the output of the baseband combiner is given as

$$\begin{aligned}\mathbf{y}[k] = \mathbf{W}_{BB}^H[k]\,\mathbf{W}_{RF}^H\mathbf{H}[k]\,\mathbf{F}_{RF}\mathbf{F}_{BB}[k]\,\mathbf{x}[k] \\ + \mathbf{W}_{BB}^H[k]\,\mathbf{W}_{RF}^H\mathbf{n}[k],\end{aligned}$$

where $\mathbf{x}[k] \in \mathbb{C}^{N_s \times 1}$ is the vector of transmit symbols and $\mathbf{n}[k] \in \mathbb{C}^{N_R \times 1}$ denotes the complex additive white Gaussian noise (AWGN) distributed as $\mathcal{CN}\left(\mathbf{0}_{N_R \times 1}, \sigma^2\mathbf{I}_{N_R}\right)$. The matrix $\mathbf{H}[k] \in \mathbb{C}^{N_R \times N_T}$ represents the mmWave MIMO OFDM channel matrix corresponding to the kth subcarrier and is determined by the K-point FFT of the channel taps as

$$\mathbf{H}[k] = \sum_{l=0}^{L-1} \mathbf{H}_l\ e^{-j\frac{2\pi k}{K}l}. \tag{2}$$

Let us now consider a quasi-static frequency-selective mmWave MIMO channel that is assumed to be constant for a block of several OFDM symbols. This is extended to doubly-selective, i.e., time as well as frequency-selective channels, in Section-V. For the purpose of channel estimation, the transmitter is assumed to employ M training frames. Let $\mathbf{F}_{RF,m} \in \mathbb{C}^{N_T \times N_{RF}^T}$, $\mathbf{W}_{RF,m} \in \mathbb{C}^{N_R \times N_{RF}^R}$ and $\mathbf{s}_m[k] \in \mathbb{C}^{N_{RF}^T \times 1}$ denote the RF precoder, RF combiner and pilot symbol vector, respectively, during the mth training frame for the kth subcarrier. The pilot vector $\mathbf{y}_m[k] \in \mathbb{C}^{N_{RF}^R \times 1}$ at the output of the RF combiner for the mth training frame is given as

$$\mathbf{y}_m[k] = \mathbf{W}_{RF,m}^H\mathbf{H}[k]\,\mathbf{F}_{RF,m}\mathbf{s}_m[k] + \mathbf{W}_{RF,m}^H\mathbf{n}_m[k], \tag{3}$$

where the noise vector $\mathbf{n}_m[k] \in \mathbb{C}^{N_R \times 1} \sim \mathcal{CN}\left(\mathbf{0}_{N_R \times 1}, \sigma^2\mathbf{I}_{N_R}\right)$. By exploiting the Kronecker property of the $\mathrm{vec}(\cdot)$ operator from (1), the model for $\mathbf{y}_m[k]$ can be recast as

$$\mathbf{y}_m[k] = \underbrace{\left(\mathbf{s}_m^T[k]\,\mathbf{F}_{RF,m}^T \otimes \mathbf{W}_{RF,m}^H\right)}_{\mathbf{\Phi}_m[k]}\underbrace{\mathrm{vec}\left(\mathbf{H}[k]\right)}_{\mathbf{h}[k]} + \mathbf{n}_{c,m}[k], \tag{4}$$

where $\mathbf{\Phi}_m[k] \in \mathbb{C}^{N_{RF}^R \times N_R N_T}$ and $\mathbf{n}_{c,m}[k] = \mathbf{W}_{RF,m}^H\mathbf{n}_m[k] \in \mathbb{C}^{N_{RF}^R \times 1}$ denotes the output noise vector for the mth training frame that is distributed as $\mathcal{CN}(\mathbf{0}, \mathbf{R}_m)$ with $\mathbf{R}_m = \sigma^2\mathbf{W}_{RF,m}^H\mathbf{W}_{RF,m} \in \mathbb{C}^{N_{RF}^R \times N_{RF}^R}$. Thus it can be seen that the noise at the output of the RF combiner is colored. Naturally, one has to take this into consideration in order to obtain a reliable estimate of the mmWave MIMO OFDM channel, and has thus been incorporated in the G-SBL, LCG-SBL and O-SBL schemes proposed in Section-IV-A, Section-IV-B and Section-V, respectively. The quantity $\mathbf{h}[k] \in \mathbb{C}^{N_R N_T \times 1}$ denotes the equivalent mmWave MIMO OFDM channel vector for the kth subcarrier. Concatenating the vectors $\mathbf{y}_m[k]$ for the various training frames $m, 1 \leq m \leq M$, the equivalent system model can be obtained as

$$\underbrace{\begin{bmatrix}\mathbf{y}_1[k]\\ \vdots \\ \mathbf{y}_M[k]\end{bmatrix}}_{\mathbf{y}[k]} = \underbrace{\begin{bmatrix}\mathbf{\Phi}_1[k]\\ \vdots \\ \mathbf{\Phi}_M[k]\end{bmatrix}}_{\mathbf{\Phi}[k]}\mathbf{h}[k] + \underbrace{\begin{bmatrix}\mathbf{n}_{c,1}[k]\\ \vdots \\ \mathbf{n}_{c,M}[k]\end{bmatrix}}_{\mathbf{n}_c[k]}, \tag{5}$$

where $\mathbf{y}[k] \in \mathbb{C}^{MN_{RF}^R \times 1}, \mathbf{\Phi}[k] \in \mathbb{C}^{MN_{RF}^R \times N_R N_T}$ and the additive noise $\mathbf{n}_c[k] \in \mathbb{C}^{MN_{RF}^R \times 1}$ is distributed as $\mathbf{n}_c[k] \sim \mathcal{CN}\left(\mathbf{0}_{MN_{RF}^R \times 1}, \mathbf{R}\right)$. The covariance matrix $\mathbf{R} \in$

$\mathbb{C}^{MN_{RF}^R \times MN_{RF}^R}$ of the equivalent noise vector $\mathbf{n}_c[k]$ is a block diagonal matrix comprised of the matrices $\mathbf{R}_m, \forall 1 \leq m \leq M$, on its principal diagonal. For the above estimation model, the conventional 'sparsity-agnostic' techniques such as LS and MMSE schemes require the sensing matrix $\mathbf{\Phi}[k]$ to be overdetermined, i.e., $MN_{RF}^R \geq N_R N_T$. This implies that the minimum number of training frames required for the conventional techniques is $M = \dfrac{N_R N_T}{N_{RF}^R}$, which leads to a significant overhead due to the fact that $N_{RF}^R \ll \min(N_T, N_R)$, and N_T, N_R are typically high in a mmWave MIMO system. Moreover, these traditional approaches do not exploit the sparsity of the mmWave MIMO channel [4], [6], even though doing so typically lead to a substantial improvement in the channel estimation performance. Given this motivation, the next section develops our sparse channel model for a frequency-selective mmWave MIMO channel.

III. Sparse Channel Model for the mmWave MIMO OFDM System

Using the clustered channel model in [10], [21], the delay tap $\mathbf{H}_l, \forall\ 0 \leq l \leq L-1$, of a typical frequency-selective mmWave MIMO channel can be readily modeled as

$$\mathbf{H}_l = \beta \sum_{i=1}^{N_{\text{cl}}} \sum_{j=1}^{N_{\text{ray},i}} \alpha_{ij} p(lT_s - \tau_{ij})\, \mathbf{a}_R(\phi_{ij})\, \mathbf{a}_T^H(\theta_{ij}), \quad (6)$$

where $\beta = \sqrt{\dfrac{N_T N_R}{\sum_{i=1}^{N_{\text{cl}}} N_{\text{ray},i}}}$, N_{cl} and $N_{\text{ray},i}$ denote the number of clusters and rays in the ith cluster, respectively. The quantities θ_{ij}, ϕ_{ij}, α_{ij} and τ_{ij} represent the AoD, AoA, complex channel gain and delay, respectively, of the jth ray in the ith cluster, and $p(\tau)$ is the response of the pulse shaping filter. The quantity T_s denotes the sampling period. The vectors $\mathbf{a}_R(\phi_{ij}) \in \mathbb{C}^{N_R \times 1}$ and $\mathbf{a}_T(\theta_{ij}) \in \mathbb{C}^{N_T \times 1}$ denote the receive and transmit array response vectors corresponding to the AoA ϕ_{ij} and AoD θ_{ij}, respectively, and are given by the following expressions

$$\mathbf{a}_R(\phi_{ij}) = \frac{1}{\sqrt{N_R}}\left[1, e^{-j\frac{2\pi}{\lambda} d_R \cos\phi_{ij}}, \ldots, e^{-j\frac{2\pi}{\lambda}(N_R-1) d_R \cos\phi_{ij}}\right]^T, \quad (7)$$

$$\mathbf{a}_T(\theta_{ij}) = \frac{1}{\sqrt{N_T}}\left[1, e^{-j\frac{2\pi}{\lambda} d_T \cos\theta_{ij}}, \ldots, e^{-j\frac{2\pi}{\lambda}(N_T-1) d_T \cos\theta_{ij}}\right]^T, \quad (8)$$

where λ denotes wavelength of the mmWave signal, while d_R and d_T are the antenna spacings of the arrays at the receiver and transmitter, respectively. The lth mmWave MIMO channel tap $\mathbf{H}_l$ can be expressed in the compact form of

$$\mathbf{H}_l = \mathbf{A}_R \mathbf{H}_{D,l} \mathbf{A}_T^H, \quad (9)$$

where $\mathbf{A}_R \in \mathbb{C}^{N_R \times N_{\text{ray}}}$, $\mathbf{A}_T \in \mathbb{C}^{N_T \times N_{\text{ray}}}$ denote the concatenated matrices of the receive and transmit array response vectors, defined as $\mathbf{A}_R = \left[\{\mathbf{a}_R(\phi_{ij})\}_{i=1,j=1}^{N_{\text{cl}},N_{\text{ray},i}}\right]$, $\mathbf{A}_T = \left[\{\mathbf{a}_T(\theta_{ij})\}_{i=1,j=1}^{N_{\text{cl}},N_{\text{ray},i}}\right]$, and $N_{\text{ray}} = \sum_{i=1}^{N_{\text{cl}}} N_{\text{ray},i}$. The diagonal matrix $\mathbf{H}_{D,l} = \text{diag}\left(\{K_{ij}\}_{i=1,j=1}^{N_{\text{cl}},N_{\text{ray},i}}\right) \in \mathbb{C}^{N_{\text{ray}} \times N_{\text{ray}}}$, where $K_{ij} = \beta \alpha_{ij} p(lT_s - \tau_{ij})$. The sparse channel model for the mmWave MIMO OFDM system is developed next.

Let us consider partitions of the AoA and AoD spaces spanning the interval $[0, \pi)$ with grids Φ_R and Θ_T, which are comprised of $G_R, G_T \geq \max\{N_T, N_R\}$ angles, respectively. These quantized angles $\{\phi_g \in \Phi_R, \forall\ 1 \leq g \leq G_R\}$ and $\{\theta_g \in \Theta_T, \forall\ 1 \leq g \leq G_T\}$ are chosen as per the following conditions described in [9],

$$\cos(\phi_g) = \frac{2}{G_R}(g-1) - 1, \forall\ 1 \leq g \leq G_R,$$
$$\cos(\theta_g) = \frac{2}{G_T}(g-1) - 1, \forall\ 1 \leq g \leq G_T. \quad (10)$$

The transmit and receive array response dictionary matrices $\mathbf{A}_T(\Theta_T) \in \mathbb{C}^{N_T \times G_T}$ and $\mathbf{A}_R(\Phi_R) \in \mathbb{C}^{N_R \times G_R}$ are obtained by concatenating the array response vectors corresponding to angular grids Θ_T and Φ_R, respectively, as $\mathbf{A}_T(\Theta_T) = [\mathbf{a}_T(\theta_1), \mathbf{a}_T(\theta_2), \ldots, \mathbf{a}_T(\theta_{G_T})]$, $\mathbf{A}_R(\Phi_R) = [\mathbf{a}_R(\phi_1), \mathbf{a}_R(\phi_2), \ldots, \mathbf{a}_R(\phi_{G_R})]$. The beamspace representation of the mmWave MIMO channel tap $\mathbf{H}_l$ can be obtained as

$$\mathbf{H}_l \approx \mathbf{A}_R(\Phi_R)\, \mathbf{H}_{b,l} \mathbf{A}_T^H(\Theta_T), \quad (11)$$

where $\mathbf{H}_{b,l} \in \mathbb{C}^{G_R \times G_T}$ denotes the equivalent beamspace channel matrix corresponding to $\mathbf{H}_l$, as described in [9]. Using the model above, the mmWave MIMO OFDM channel for the kth subcarrier can be expressed as

$$\mathbf{H}[k] = \mathbf{A}_R(\Phi_R)\, \mathbf{H}_b[k]\, \mathbf{A}_T^H(\Theta_T), \quad (12)$$

where $\mathbf{H}_b[k] \in \mathbb{C}^{G_R \times G_T}$ denotes the beamspace channel matrix of $\mathbf{H}[k]$ that is defined as

$$\mathbf{H}_b[k] = \sum_{l=0}^{L-1} \mathbf{H}_{b,l}\ e^{-j\frac{2\pi k}{K} l}. \quad (13)$$

As described in [1], [2], due to the weak scattering and diffraction in the mmWave regime owing to increased attenuation from blockages, coupled with the highly directional nature of signal propagation, typically there is only a few spatial multipath components in the mmWave MIMO channel. Therefore, the beamspace channel matrix $\mathbf{H}_{b,l}$ is *sparse*, i.e., only a few of its elements are significant, with the rest being close to zero. Furthermore, from (11), since the AoAs/ AoDs corresponding to all the delay taps $\mathbf{H}_l$ are identical, the locations of these significant coefficients in their corresponding beamspace representation $\mathbf{H}_{b,l}$ coincide. Another interesting observation from (13) is that the subcarrier beamspace channel matrix $\mathbf{H}_b[k]$ is also sparse, and share a common sparsity profile with the TD beamspace channel matrix $\mathbf{H}_{b,l}$.

From (12), using the well known property of the $\text{vec}(\cdot)$ operator and matrix Kronecker product given in (1), the vectorized channel $\mathbf{h}[k] = \text{vec}(\mathbf{H}[k])$ can be described as

$$\mathbf{h}[k] = \mathbf{\Psi} \mathbf{h}_b[k], \quad (14)$$

where $\mathbf{h}_b[k] = \text{vec}(\mathbf{H}_b[k]) \in \mathbb{C}^{G_R G_T \times 1}$ is the sparse beamspace channel vector corresponding to the kth subcarrier and $\mathbf{\Psi} = [\mathbf{A}_T^*(\Theta_T) \otimes \mathbf{A}_R(\Phi_R)] \in \mathbb{C}^{N_R N_T \times G_R G_T}$ is the

sparsifying dictionary matrix. Substituting (14) into (5) yields the sparse estimation model for the kth subcarrier as

$$\mathbf{y}\left[k\right]=\bar{\mathbf{\Phi}}\left[k\right]\mathbf{h}_b\left[k\right]+\mathbf{n}_c\left[k\right], \tag{15}$$

where $\bar{\mathbf{\Phi}}\left[k\right]=\mathbf{\Phi}\left[k\right]\mathbf{\Psi}\in\mathbb{C}^{MN_{RF}^{R}\times G_RG_T}$ denotes the *equivalent sensing matrix*. To this end, the SBL-based per-subcarrier approach conceived for channel estimation is highlighted first briefly. While this is clearly suboptimal, it serves not only to describe the preliminaries of SBL, but also to better motivate the more sophisticated joint G-SBL, LCG-SBL and online G-SBL (O-SBL) schemes developed in the later sections, which exploit both the group-sparsity as well as the FD and TD correlation.

IV. SBL-based Sparse Channel Estimation in mmWave MIMO OFDM Systems

SBL is a Bayesian method conceived for sparse signal recovery, which initially assigns the following parameterized Gaussian prior to the beamspace channel vector $\mathbf{h}_b\left[k\right]$ [12]

$$p\left(\mathbf{h}_b\left[k\right];\mathbf{\Gamma}\right)=\prod_{i=1}^{G_RG_T}\left(\pi\gamma_i\right)^{-1}\exp\left(-\frac{\left|\mathbf{h}_b\left[k\right](i)\right|^2}{\gamma_i}\right), \tag{16}$$

where $\gamma_i\in\mathbb{R}_+$ denotes the hyperparameter of $\mathbf{h}_b\left[k\right](i)$, the ith element of $\mathbf{h}_b\left[k\right]$, and $\mathbf{\Gamma}\in\mathbb{R}_+^{G_RG_T\times G_RG_T}$ denotes the diagonal matrix that is comprised of the hyperparameters γ_i on its principal diagonal. Using the standard results for MMSE estimation from [25], the MMSE estimate $\boldsymbol{\mu}_k\in\mathbb{C}^{G_RG_T\times 1}$ and the corresponding error covariance matrix $\mathbf{\Sigma}_k\in\mathbb{C}^{G_RG_T\times G_RG_T}$ of the beamspace channel vector $\mathbf{h}_b\left[k\right]$, corresponding to the prior in (16), can be formulated as

$$\begin{aligned}\boldsymbol{\mu}_k&=\mathbf{\Sigma}\bar{\mathbf{\Phi}}^H\left[k\right]\mathbf{R}^{-1}\mathbf{y}\left[k\right],\\ \mathbf{\Sigma}_k&=\left(\mathbf{\Gamma}^{-1}+\bar{\mathbf{\Phi}}^H\left[k\right]\mathbf{R}^{-1}\bar{\mathbf{\Phi}}\left[k\right]\right)^{-1}.\end{aligned} \tag{17}$$

Observe from above that the MMSE estimate $\boldsymbol{\mu}_k$ depends on the hyperparameter matrix $\mathbf{\Gamma}$ through the error covariance matrix $\mathbf{\Sigma}_k$. Hence, the beamspace channel estimation problem of mmWave hybrid MIMO OFDM systems effectively reduces to the estimation of the hyperparameters γ_i. As described in [12], the iterative expectation-maximization (EM) is well-suited for this task. The EM-based update equation for the estimation of the hyperparameters is described next. Let $\widehat{\mathbf{\Gamma}}^{(p-1)}$ denote the estimate of the hyperparameter matrix $\mathbf{\Gamma}$ in the $(p-1)$st EM iteration. The update $\widehat{\gamma}_i^{(p)}$ for the pth EM iteration can be obtained as [10], [12]

$$\widehat{\gamma}_i^{(p)}=\left|\boldsymbol{\mu}_k^{(p)}\left(i\right)\right|^2+\mathbf{\Sigma}_k^{(p)}\left(i,i\right), \tag{18}$$

where the quantities $\boldsymbol{\mu}_k^{(p)}$ and $\mathbf{\Sigma}_k^{(p)}$ is obtained by substituting $\mathbf{\Gamma}=\widehat{\mathbf{\Gamma}}^{(p-1)}$ into (17). The EM-update equation described above is repeated until the Frobenius norm-square of the difference of the hyperparameter matrix estimates for successive iterations falls below a suitable threshold ϵ, i.e. $\left\|\widehat{\mathbf{\Gamma}}^{(p)}-\widehat{\mathbf{\Gamma}}^{(p-1)}\right\|_F^2<\epsilon$ or the maximum number of EM iterations $p_{\max}$ is reached. Upon convergence, the SBL-based beamspace channel estimate $\widehat{\mathbf{h}}_b\left[k\right]$ for the mmWave MIMO OFDM system is given as $\widehat{\mathbf{h}}_b\left[k\right]=\widehat{\boldsymbol{\mu}}_k$, where $\widehat{\boldsymbol{\mu}}_k$ denotes the converged *a posteriori* mean. This can be finally employed to obtain the mmWave MIMO channel matrix corresponding to the kth subcarrier as $\widehat{\mathbf{H}}\left[k\right]=\mathrm{vec}^{-1}\left(\mathbf{\Psi}\widehat{\boldsymbol{\mu}}_k\right)$.

A. G-SBL based Group-Sparse Channel Estimation in mmWave MIMO OFDM Systems

The SBL procedure developed above performs channel estimation over each subcarrier k and leads to improved performance by exploiting the sparsity in the beamspace domain in comparison to the conventional LS/ MMSE estimators. However, the performance of this scheme can be significantly further improved via the joint estimation of the beamspace channel vector over all the pilot subcarriers. Let $\mathcal{K}_p=\left\{k_1,k_2,\cdots,k_{K_p}\right\}$ denote the set of subcarrier indices loaded with pilot symbols, where $K_p\leq K$ denotes the number of pilot subcarriers. Furthermore, consider identical pilot symbols across all the pilot subcarriers, i.e., $\mathbf{s}_m\left[k\right]=\mathbf{s}_m,\forall k\in\mathcal{K}_p$. It follows from (4) that we have:

$$\mathbf{\Phi}_m\left[k\right]=\mathbf{\Phi}_m=\left(\mathbf{s}_m^T\mathbf{F}_{RF,m}^T\otimes\mathbf{W}_{RF,m}^H\right), \tag{19}$$

which, thanks to (5), implies that the dictionary matrix obeys $\mathbf{\Phi}\left[k\right]=\mathbf{\Phi}=\left[\mathbf{\Phi}_1^T,\mathbf{\Phi}_2^T,\cdots,\mathbf{\Phi}_M^T\right]^T$. Therefore, the simultaneous-sparse channel estimation model for the mmWave MIMO OFDM system can be formulated as

$$\mathbf{Y}=\bar{\mathbf{\Phi}}\mathbf{H}_b+\mathbf{N}, \tag{20}$$

where the matrices $\mathbf{Y}\in\mathbb{C}^{MN_{RF}^{R}\times K_p}$, $\mathbf{H}_b\in\mathbb{C}^{G_RG_T\times K_p}$ and $\mathbf{N}\in\mathbb{C}^{MN_{RF}^{R}\times K_p}$ are constructed as

$$\begin{aligned}\mathbf{Y}&=\left[\mathbf{y}\left[k_1\right],\ldots,\mathbf{y}\left[k_{K_p}\right]\right],\\ \mathbf{H}_b&=\left[\mathbf{h}_b\left[k_1\right],\ldots,\mathbf{h}_b\left[k_{K_p}\right]\right],\\ \mathbf{N}&=\left[\mathbf{n}_c\left[k_1\right],\ldots,\mathbf{n}_c\left[k_{K_p}\right]\right],\end{aligned}$$

and the dictionary matrix obeys $\bar{\mathbf{\Phi}}=\mathbf{\Phi}\mathbf{\Psi}$. As described previously in Section-III, the columns of the concatenated beamspace channel matrix $\mathbf{H}_b$ share a common sparsity-profile. Hence, the matrix $\mathbf{H}_b$ is *simultaneous-sparse* in nature. The equivalent model for the estimation of the simultaneous-sparse matrix $\mathbf{H}_b$ above can be derived as

$$\mathbf{y}=\widetilde{\mathbf{\Phi}}\mathbf{h}_b+\mathbf{n}, \tag{21}$$

where $\mathbf{y}=\mathrm{vec}\left(\mathbf{Y}^T\right)\in\mathbb{C}^{MN_{RF}^{R}K_p\times 1}$, $\mathbf{n}=\mathrm{vec}\left(\mathbf{N}^T\right)\in\mathbb{C}^{MN_{RF}^{R}K_p\times 1}$ and the dictionary matrix obeys $\widetilde{\mathbf{\Phi}}=\left(\bar{\mathbf{\Phi}}\otimes\mathbf{I}_{K_p}\right)\in\mathbb{C}^{MN_{RF}^{R}K_p\times G_RG_TK_p}$. Furthermore, the covariance matrix $\widetilde{\mathbf{R}}\in\mathbb{C}^{MN_{RF}^{R}K_p\times MN_{RF}^{R}K_p}$ of the concatenated noise vector $\mathbf{n}$ can be obtained as $\widetilde{\mathbf{R}}\triangleq\mathbb{E}\left\{\mathbf{n}\mathbf{n}^H\right\}=\left(\mathbf{R}\otimes\mathbf{I}_{K_p}\right)$. The concatenated beamspace channel vector $\mathbf{h}_b=\mathrm{vec}\left(\mathbf{H}_b^T\right)\in\mathbb{C}^{G_RG_TK_p\times 1}$ exhibits an interesting structure. Let the ith group $\mathbf{h}_b^i\in\mathbb{C}^{K_p\times 1},\forall 1\leq i\leq G_RG_T$, of $\mathbf{h}_b$ be expressed as $\mathbf{h}_b^i=\mathbf{h}_b\left[(i-1)K_p+1:iK_p\right]$. Owing to the simultaneous-sparse structure of the beamspace channel $\mathbf{h}_b[k]$ across the K_p pilot subcarriers, it can be noted that all the K_p coefficients within each group $\mathbf{h}_b^i$ of the vector $\mathbf{h}_b$ are likely to be either all zero or all non-zero. Therefore, the vector $\mathbf{h}_b$ has a *group-sparse* structure. The G-SBL

framework, which leverages the *group-sparsity* inherent in the concatenated beamspace channel $\mathbf{h}_b$ can now be developed as follows. To this end, the prior corresponding to the beamspace vector $\mathbf{h}_b$ is formulated as [17]

$$p\left(\mathbf{h}_b;\mathbf{\Gamma},\mathbf{G}_c\right)=\prod_{i=1}^{G_RG_T} p\left(\mathbf{h}_b^i;\gamma_i,\mathbf{G}_c\right), \tag{22}$$

where the prior $p\left(\mathbf{h}_b^i;\gamma_i,\mathbf{G}_c\right)$ assigned to the ith group $\mathbf{h}_b^i$ is given by

$$p\left(\mathbf{h}_b^i;\gamma_i,\mathbf{G}_c\right)=\frac{1}{(\pi\gamma_i)^{K_p}\det(\mathbf{G_c})}\ \exp\left(-\frac{\left(\mathbf{h}_b^i\right)^H\mathbf{G}_c^{-1}\mathbf{h}_b^i}{\gamma_i}\right). \tag{23}$$

In the above, the matrix $\mathbf{G}_c\in\mathbb{C}^{K_p\times K_p}$ represents the FD correlation, and is assumed to be unknown along with the hyperparameters γ_i. The Bayesian evidence $\log p\left(\mathbf{y};\mathbf{\Gamma},\mathbf{G}_c\right)$ can once again be maximized using the EM framework in order to estimate the hyperparameter matrix $\mathbf{\Gamma}$ and the correlation matrix $\mathbf{G}_c$. The key steps in this process are given below.

Let $\widehat{\mathbf{\Gamma}}^{(p-1)}$ and $\widehat{\mathbf{G}}_c^{(p-1)}$ denote the estimate of the hyperparameter matrix $\mathbf{\Gamma}$ and correlation matrix $\mathbf{G}_c$ in the $(p-1)$st EM iteration. The E-step in the pth iteration determines the expected value of the log-likelihood $\mathcal{L}\left(\mathbf{\Gamma},\mathbf{G}_c\mid\widehat{\mathbf{\Gamma}}^{(p-1)},\widehat{\mathbf{G}}_c^{(p-1)}\right)$ of the complete data set $\{\mathbf{y},\mathbf{h}_b\}$ as

$$\mathcal{L}\left(\mathbf{\Gamma},\mathbf{G}_c\mid\widehat{\mathbf{\Gamma}}^{(p-1)},\widehat{\mathbf{G}}_c^{(p-1)}\right)= \mathbb{E}_{\mathbf{h}_b|\mathbf{y};\widehat{\mathbf{\Gamma}}^{(p-1)},\widehat{\mathbf{G}}_c^{(p-1)}}\Big\{\log\ p\big(\mathbf{y},\mathbf{h}_b;\mathbf{\Gamma},\mathbf{G}_c\big)\Big\}. \tag{24}$$

Upon applying Bayes' rule in (24) and ignoring the term $\log p\big(\mathbf{y}\mid\mathbf{h}_b\big)$ that does not depend on the hyperparameter matrix $\mathbf{\Gamma}$ and the correlation matrix $\mathbf{G}_c$, the subsequent M-step for maximization of $\mathcal{L}\left(\mathbf{\Gamma},\mathbf{G}_c\mid\widehat{\mathbf{\Gamma}}^{(p-1)},\widehat{\mathbf{G}}_c^{(p-1)}\right)$ with respect to $\mathbf{\Gamma}$ and $\mathbf{G}_c$ can be formulated as

$$\left(\widehat{\mathbf{\Gamma}}^{(p)},\widehat{\mathbf{G}}_c^{(p)}\right)=\arg\max_{\mathbf{\Gamma},\mathbf{G}_c}\mathbb{E}\Big\{\log\ p\big(\mathbf{h}_b;\mathbf{\Gamma},\mathbf{G}_c\big)\Big\}. \tag{25}$$

It follows from (22) and (23) that the M-step with respect to hyperparameter decouples for each γ_i, and is given as

$$\begin{aligned}\widehat{\gamma}_i^{(p)}&=\arg\max_{\gamma_i}\mathbb{E}_{\mathbf{h}_b|\mathbf{y};\widehat{\mathbf{\Gamma}}^{(p-1)},\widehat{\mathbf{G}}_c^{(p-1)}}\Big\{\log\ p\big(\mathbf{h}_b^i;\gamma_i,\widehat{\mathbf{G}}_c^{(p-1)}\big)\Big\}.\\ &\equiv\arg\max_{\gamma_i}\mathbb{E}_{\mathbf{h}_b|\mathbf{y};\widehat{\mathbf{\Gamma}}^{(p-1)},\widehat{\mathbf{G}}_c^{(p-1)}}\Big\{-K_p\log(\gamma_i)\\ &\quad-\frac{1}{\gamma_i}\mathrm{Tr}\left(\left(\widehat{\mathbf{G}}_c^{(p-1)}\right)^{-1}\mathbf{h}_b^i\left(\mathbf{h}_b^i\right)^H\right)\Big\}.\end{aligned} \tag{26}$$

Once again, upon differentiating the objective function in (26) with respect to γ_i and setting it equal to zero, one obtains the hyperparameter update equation for the pth EM iteration as

$$\widehat{\gamma}_i^{(p)}=\frac{1}{K_p}\mathrm{Tr}\left(\left(\widehat{\mathbf{G}}_c^{(p-1)}\right)^{-1}\mathbb{E}\left\{\mathbf{h}_b^i\left(\mathbf{h}_b^i\right)^H\right\}\right). \tag{27}$$

The *a posteriori* pdf $p\left(\mathbf{h}_b\mid\mathbf{y};\widehat{\mathbf{\Gamma}}^{(p-1)},\widehat{\mathbf{G}}_c^{(p-1)}\right)=\mathcal{CN}\left(\widetilde{\boldsymbol{\mu}}^{(p)},\widetilde{\mathbf{\Sigma}}^{(p)}\right)$ can be evaluated using

$$\begin{aligned}\widetilde{\boldsymbol{\mu}}^{(p)}&=\widetilde{\mathbf{\Sigma}}^{(p)}\widetilde{\mathbf{\Phi}}^H\widetilde{\mathbf{R}}^{-1}\mathbf{y},\\ \widetilde{\mathbf{\Sigma}}^{(p)}&=\left(\left(\widehat{\mathbf{\Gamma}}^{(p-1)}\otimes\widehat{\mathbf{G}}_c^{(p-1)}\right)^{-1}+\widetilde{\mathbf{\Phi}}^H\widetilde{\mathbf{R}}^{-1}\widetilde{\mathbf{\Phi}}\right)^{-1}.\end{aligned} \tag{28}$$

Employing the above *a posteriori* pdf of $\mathbf{h}_b$ in (27) yields the following estimate of the hyperparameter $\widehat{\gamma}_i^{(p)}$

$$\widehat{\gamma}_i^{(p)}=\frac{1}{K_p}\mathrm{Tr}\left(\left(\widehat{\mathbf{G}}_c^{(p-1)}\right)^{-1}\left(\widetilde{\mathbf{\Sigma}}_i^{(p)}+\widetilde{\boldsymbol{\mu}}_i^{(p)}\left(\widetilde{\boldsymbol{\mu}}_i^{(p)}\right)^H\right)\right), \tag{29}$$

where $\widetilde{\boldsymbol{\mu}}_i^{(p)}\in\mathbf{C}^{K_p\times 1}$ and $\widetilde{\mathbf{\Sigma}}_i^{(p)}\in\mathbb{C}^{K_p\times K_p}$ are defined as

$$\begin{aligned}\widetilde{\boldsymbol{\mu}}_i^{(p)}&=\widetilde{\boldsymbol{\mu}}^{(p)}\left[(i-1)K_p+1:iK_p\right],\\ \widetilde{\mathbf{\Sigma}}_i^{(p)}&=\widetilde{\mathbf{\Sigma}}^{(p)}\left[(i-1)K_p+1:iK_p,(i-1)K_p+1:iK_p\right].\end{aligned}$$

Note that $\widetilde{\boldsymbol{\mu}}_i^{(p)}$ and $\widetilde{\mathbf{\Sigma}}_i^{(p)}$ denote the *a posteriori* mean and covariance of the ith group $\mathbf{h}_b^i$ of the group-sparse beamspace channel vector $\mathbf{h}_b$. Similarly, as shown in Appendix-A, the update $\widehat{\mathbf{G}}_c^{(p)}$ of the correlation matrix is obtained as

$$\widehat{\mathbf{G}}_c^{(p)}=\frac{1}{G_RG_T}\sum_{i=1}^{G_RG_T}\frac{\widetilde{\mathbf{\Sigma}}_i^{(p)}+\widetilde{\boldsymbol{\mu}}_i^{(p)}\left(\widetilde{\boldsymbol{\mu}}_i^{(p)}\right)^H}{\widehat{\gamma}_i^{(p)}}. \tag{30}$$

Upon convergence, the G-SBL estimate $\widehat{\mathbf{h}}_b$ of the concatenated beamspace channel $\mathbf{h}_b$ is given by the converged *a posteriori* mean $\widetilde{\boldsymbol{\mu}}^{(p)}$. The estimate $\widehat{\mathbf{h}}_b[k_i]$ of the beamspace channel vector for the k_ith pilot subcarrier is given by the ith row of the $K_p\times G_RG_T$-dimensional matrix $\widehat{\mathbf{H}}_b^T$ obtained by reshaping the vector $\widehat{\mathbf{h}}_b$, i.e., $\widehat{\mathbf{H}}_b^T=\mathrm{vec}^{-1}\left(\widehat{\mathbf{h}}_b\right)$. The estimate $\widehat{\mathbf{H}}\left[k_i\right]\in\mathbb{C}^{N_R\times N_T}$ of the corresponding channel matrix is obtained as $\widehat{\mathbf{H}}\left[k_i\right]=\mathrm{vec}^{-1}\left(\mathbf{\Psi}\widehat{\mathbf{h}}_b[k_i]\right)$. Finally, one can estimate the mmWave MIMO OFDM channel matrices on the remaining subcarriers as follows. Observe that using the estimate $\widehat{\mathbf{H}}_b\in\mathbb{C}^{G_RG_T\times K_p}$ of the beamspace channel matrix $\mathbf{H}_b$, the estimate $\widehat{\mathbf{H}}_{\mathcal{K}_p}$ of the concatenated channel matrix $\mathbf{H}_{\mathcal{K}_p}=\left[\mathbf{h}\left[k_1\right],\mathbf{h}\left[k_2\right],\cdots,\mathbf{h}\left[k_{K_p}\right]\right]\in\mathbb{C}^{N_RN_T\times K_p}$ for the pilot subcarriers can be derived as $\widehat{\mathbf{H}}_{\mathcal{K}_p}=\mathbf{\Psi}\widehat{\mathbf{H}}_b$. Let $\mathbf{F}\in\mathbb{C}^{K\times L}$ denote the truncated DFT matrix obtained from the $K\times K$-element DFT matrix. Let $\mathbf{F}_p=\mathbf{F}\left(\mathcal{K}_p,:\right)\in\mathbb{C}^{K_p\times L}$ denote the submatrix of $\mathbf{F}$ containing only the rows corresponding to the pilot subcarriers. The estimate $\widehat{\mathbf{H}}$ of the concatenated channel $\mathbf{H}=\left[\mathbf{h}\left[1\right],\mathbf{h}\left[2\right],\cdots,\mathbf{h}\left[K\right]\right]\in\mathbb{C}^{N_RN_T\times K}$ across all the subcarriers can be obtained as $\widehat{\mathbf{H}}=\left(\mathbf{F}\mathbf{F}_p^{\dagger}\widehat{\mathbf{H}}_{\mathcal{K}_p}^T\right)^T$. Considering now the kth column of $\widehat{\mathbf{H}}$, the estimate $\widehat{\mathbf{H}}\left[k\right]$ of the channel matrix $\mathbf{H}\left[k\right]$ for any non-pilot subcarrier k can be succinctly stated as $\widehat{\mathbf{H}}\left[k\right]=\mathrm{vec}^{-1}\left(\widehat{\mathbf{H}}\left(:,k\right)\right)$. A concise summary of the sequence of steps in the G-SBL procedure for mmWave MIMO OFDM channel estimation is presented in Algorithm-1.

B. Low Complexity (LC) G-SBL (LCG-SBL) based Group-Sparse Channel Estimation

As seen from the detailed discussions in the previous section, the G-SBL algorithm requires the inversion of a $[G_R G_T K_p \times G_R G_T K_p]$-size matrix $\widetilde{\mathbf{\Sigma}}^{(p)}$ for the evaluation of the a posteriori covariance, which renders it computationally complex. Therefore, in order to simplify the procedure, inspired by the T-MSBL technique of [17], this subsection derives the LCG-SBL algorithm for group-sparse channel estimation. The covariance matrix $\widetilde{\mathbf{\Sigma}}^{(p)}$ of (28) can be approximated as [17]

$$\begin{aligned}
&\widetilde{\mathbf{\Sigma}}^{(p)} \\
&= \left(\left(\widehat{\mathbf{\Gamma}}^{(p-1)} \right)^{-1} \otimes \left(\widehat{\mathbf{G}}_c^{(p-1)} \right)^{-1} + \left(\bar{\mathbf{\Phi}}^H \mathbf{R}^{-1} \bar{\mathbf{\Phi}} \right) \otimes \mathbf{I}_{K_p} \right)^{-1} \\
&\approx \left(\left(\widehat{\mathbf{\Gamma}}^{(p-1)} \right)^{-1} + \bar{\mathbf{\Phi}}^H \mathbf{R}^{-1} \bar{\mathbf{\Phi}} \right)^{-1} \otimes \widehat{\mathbf{G}}_c^{(p-1)}, \qquad (31)
\end{aligned}$$

where the approximation above is close to equality either at high SNR or at low FD correlation. Using this, the quantity $\widetilde{\mathbf{\Sigma}}_i^{(p)}$ of (29) and (30) can be approximated as $\widetilde{\mathbf{\Sigma}}_i^{(p)} \approx \mathbf{\Sigma}^{(p)}(i,i) \widehat{\mathbf{G}}_c^{(p-1)}$, where $\mathbf{\Sigma}^{(p)} = \left(\left(\widehat{\mathbf{\Gamma}}^{(p-1)} \right)^{-1} + \bar{\mathbf{\Phi}}^H \mathbf{R}^{-1} \bar{\mathbf{\Phi}} \right)^{-1}$. Along the similar lines, the quantity $\widetilde{\boldsymbol{\mu}}^{(p)}$ of (28) can be approximated as [17]

$$\begin{aligned}
\widetilde{\boldsymbol{\mu}}^{(p)} &= \widetilde{\mathbf{\Sigma}}^{(p)} \left[\left(\bar{\mathbf{\Phi}}^H \mathbf{R}^{-1} \right) \otimes \mathbf{I}_{K_p} \right] \mathbf{y} \\
&\approx \left[\left(\mathbf{\Sigma}^{(p)} \bar{\mathbf{\Phi}}^H \mathbf{R}^{-1} \right) \otimes \mathbf{I}_{K_p} \right] \mathrm{vec}\left(\mathbf{Y}^T \right) = \mathrm{vec}\left(\left(\widehat{\mathbf{H}}_b^{(p)} \right)^T \right), \qquad (32)
\end{aligned}$$

where $\widehat{\mathbf{H}}_b^{(p)} = \mathbf{\Sigma}^{(p)} \bar{\mathbf{\Phi}}^H \mathbf{R}^{-1} \mathbf{Y}$. Therefore, the quantity $\widetilde{\boldsymbol{\mu}}_i^{(p)}$ of (29) and (30) can be approximated as $\widetilde{\boldsymbol{\mu}}_i^{(p)} \approx \widehat{\mathbf{h}}_{b,i}^{(p)}$, where $\widehat{\mathbf{h}}_{b,i}^{(p)}$ denotes the ith row of the matrix $\widehat{\mathbf{H}}_b^{(p)}$, i.e., $\widehat{\mathbf{h}}_{b,i}^{(p)} = \widehat{\mathbf{H}}_b^{(p)}(i,:)$. Substituting these approximations in (29) and (30) yield

$$\widehat{\gamma}_i^{(p)} = \mathbf{\Sigma}^{(p)}(i,i) + \frac{1}{K_p} \left(\widehat{\mathbf{h}}_{b,i}^{(p)} \right)^H \left(\widehat{\mathbf{G}}_c^{(p-1)} \right)^{-1} \widehat{\mathbf{h}}_{b,i}^{(p)}, \qquad (33)$$

$$\begin{aligned}
\widehat{\mathbf{G}}_c^{(p)} = \frac{1}{G_R G_T} \Bigg[&\sum_{i=1}^{G_R G_T} \frac{\mathbf{\Sigma}^{(p)}(i,i)}{\widehat{\gamma}_i^{(p)}} \widehat{\mathbf{G}}_c^{(p-1)} \\
&+ \sum_{i=1}^{G_R G_T} \frac{1}{\widehat{\gamma}_i^{(p)}} \widehat{\mathbf{h}}_{b,i}^{(p)} \left(\widehat{\mathbf{h}}_{b,i}^{(p)} \right)^H \Bigg]. \qquad (34)
\end{aligned}$$

However, we suggest following robust update rule for the correlation matrix [17]:

$$\widetilde{\mathbf{G}}_c^{(p)} = \sum_{i=1}^{G_R G_T} \frac{1}{\widehat{\gamma}_i^{(p)}} \widehat{\mathbf{h}}_{b,i}^{(p)} \left(\widehat{\mathbf{h}}_{b,i}^{(p)} \right)^H + \eta \mathbf{I}_{K_p}, \widehat{\mathbf{G}}_c^{(p)} = \frac{\widetilde{\mathbf{G}}_c^{(p)}}{\| \widetilde{\mathbf{G}}_c^{(p)} \|_F}. \qquad (35)$$

As described in our technical report [26], the BCRB for the MSE of the estimate $\widehat{\mathbf{H}}$ can be expressed as

$$\mathrm{MSE}\left(\widehat{\mathbf{H}} \right) \geq \mathrm{Tr}\left\{ \widetilde{\mathbf{\Psi}} \mathbf{J}_B^{-1} \widetilde{\mathbf{\Psi}}^H \right\}, \qquad (36)$$

Algorithm 1: G-SBL for mmWave hybrid MIMO OFDM channel estimation

Input: Observation $\mathbf{y} \in \mathbb{C}^{MN_{RF}^R K_p \times 1}$, equivalent sensing matrix $\widetilde{\mathbf{\Phi}} \in \mathbb{C}^{MN_{RF}^R K_p \times G_R G_T K_p}$, noise covariance matrix $\widetilde{\mathbf{R}} \in \mathbb{C}^{MN_{RF}^R K_p \times MN_{RF}^R K_p}$, stopping parameters ϵ and $p_{\max}$

Output: $\widehat{\mathbf{H}}[k], \forall 0 \leq k \leq K-1$

Initialization: $\widehat{\mathbf{\Gamma}}^{(0)} = \mathbf{I}_{G_R G_T}$, $\widehat{\mathbf{\Gamma}}^{(-1)} = \mathbf{0}_{G_R G_T \times G_R G_T}$, $\widehat{\mathbf{G}}_c^{(0)} = \mathbf{I}_{K_p}$ and counter $p = 0$

while $\left(\left\| \widehat{\mathbf{\Gamma}}^{(p)} - \widehat{\mathbf{\Gamma}}^{(p-1)} \right\|_F^2 > \epsilon \;\&\&\; p < p_{\max} \right)$ **do**

  $p \leftarrow p + 1$

  E-step: Evaluate *a posteriori* mean $\widetilde{\boldsymbol{\mu}}^{(p)}$ and covariance $\widetilde{\mathbf{\Sigma}}^{(p)}$ using (28)

  M-step: Evaluate estimate of the hyperparameters $\widehat{\gamma}_i^{(p)}$ and correlation matrix $\widehat{\mathbf{G}}_c^{(p)}$ using (29) and (30)

  $\widehat{\mathbf{\Gamma}}^{(p)} = \mathrm{diag}\left(\widehat{\gamma}_1^{(p)}, \widehat{\gamma}_2^{(p)}, \ldots, \widehat{\gamma}_{G_R G_T}^{(p)} \right)$

end

$\widehat{\mathbf{h}}_b = \widetilde{\boldsymbol{\mu}}^{(p)}$

return: Obtain $\widehat{\mathbf{H}}[k]$ using the procedure described after Eq. (30)

where we have $\widetilde{\mathbf{\Psi}} = (\mathbf{\Psi} \otimes \mathbf{I}_K) \in \mathbb{C}^{N_R N_T K \times G_R G_T K}$ and $\mathbf{J}_B = \widetilde{\mathbf{\Phi}}^H \widetilde{\mathbf{R}}^{-1} \widetilde{\mathbf{\Phi}} + (\mathbf{\Gamma} \otimes \mathbf{G}_c)^{-1}$ [27].

C. Convergence Results of the SBL-based Algorithms

As shown in [12], the related sparse signal recovery techniques, such as basis-pursuit [15] and FOCUSS [14], suffer from significant shortcomings. For instance, for basis pursuit, the global minimum of its cost function may not necessarily coincide with the sparsest solution, which leads to *structural* errors. On the other hand, the FOCUSS algorithm has the tendency to frequently converge to suboptimal local minima, leading to *convergence* deficiencies. By contrast, global convergence is guaranteed with high probability for the SBL technique due to the well-established properties of the SBL cost function and using the EM algorithm. More specifically, reference [12] demonstrates explicitly that for a noiseless model, when the sensing matrix $\bar{\mathbf{\Phi}}[k]$ satisfies the unique representation property (URP), i.e., any subsets of MN_{RF}^R columns of $\bar{\mathbf{\Phi}}[k]$ are linearly independent, the global minimum of the SBL cost function is achieved at the sparsest solution. Furthermore, interestingly, even for the noisy scenario, all the local minima, termed as the *degenerate* solutions, are also sparse. Thus, they are better than any non-sparse solution. Furthermore, the convergence of G-SBL and LCG-SBL can be confirmed as discussed below, thanks to the result in Theorem-1 of [17]. In the limit, when the noise variance $\sigma^2 \to 0$ and $\left(\sum_{i=1}^{N_{\mathrm{cl}}} N_{\mathrm{ray},i} \right) < MN_{RF}^R$, the estimate $\widehat{\mathbf{h}}_b$ equals to $\mathbf{h}_b$ with probability 1. Interestingly, this holds, regardless of the estimated covariance $\widehat{\mathbf{G}}_c$.

V. Online G-SBL (O-SBL) for doubly-selective Sparse mmWave MIMO OFDM channel estimation

The previous sections considered the estimation of a quasi-static mmWave MIMO OFDM channel. However, practical mmWave MIMO channels are temporally correlated in nature [10], [20], which leads to time- and frequency-selectivity, representing a doubly-selective mmWave MIMO OFDM channel. This section develops the pertinent model, followed by a frame-wise online channel estimation procedure for the same.

A. Doubly-selective mmWave MIMO OFDM Channel Model

The model in (6) for a quasi-static channel, can now be readily extended to model the time-selective channel matrix $\mathbf{H}_{l,n}$ for the lth delay tap in the nth block, as shown below

$$\mathbf{H}_{l,n} = \beta \sum_{i=1}^{N_{\rm cl}} \sum_{j=1}^{N_{{\rm ray},i}} \alpha_{ij,n} p\left(lT_s - \tau_{ij}\right) \mathbf{a}_R\left(\phi_{ij}\right) \mathbf{a}_T^H\left(\theta_{ij}\right), \tag{37}$$

where $\alpha_{ij,n}$ denotes the complex channel gain corresponding to the jth ray in the ith cluster and nth block. The time evolution of the channel gain can be modeled using a first order autoregressive (AR-1) process [20], [28]

$$\alpha_{ij,n} = \rho \alpha_{ij,n-1} + \sqrt{1-\rho^2} w_{ij,n}, \tag{38}$$

where the quantity ρ denotes the TD correlation coefficient. Similar to [28], the correlation coefficient ρ can be evaluated from Jake's well-established model for the wireless channel as $\rho = J_0\left(2\pi f_D T_B\right)$, where J_0 is the zeroth order Bessel function of first kind and f_D denotes the maximum Doppler frequency. The model noise process obeys $w_{ij,n} \sim \mathcal{CN}\left(0, \sigma_w^2\right)$, which is also termed as the innovation noise and it is assumed to be independent of $\alpha_{ij,n-1}, \forall n$. Similar to $\mathbf{H}_b$ of (20), one can also define the concatenated time-selective mmWave MIMO OFDM beamspace channel $\mathbf{H}_{b,n} \in \mathbb{C}^{G_R G_T \times K_p}$, the time-variation of which can be modeled as

$$\mathbf{H}_{b,n} = \rho \mathbf{H}_{b,n-1} + \sqrt{1-\rho^2} \mathbf{W}_n. \tag{39}$$

The innovation noise matrix $\mathbf{W}_n \in \mathbb{C}^{G_R G_T \times K_p}$ above and the concatenated channel matrix $\mathbf{H}_{b,n}$ are both *simultaneous-sparse* in nature and share a common sparsity-profile. Similar to (3), let $\mathbf{y}_{n,m}[k_i] \in \mathbb{C}^{N_{RF}^R \times 1}$ and $\mathbf{n}_{n,m}[k_i] \in \mathbb{C}^{N_R \times 1}$ denote the received output and noise vector for the k_ith pilot subcarrier in the mth training frame and nth block. The channel estimation model for the mth training frame in the nth block is given as

$$\mathbf{Y}_{n,m} = \bar{\mathbf{\Phi}}_m \mathbf{H}_{b,n} + \mathbf{N}_{n,m}, \tag{40}$$

where $\mathbf{Y}_{n,m} \in \mathbb{C}^{N_{RF}^R \times K_p}$ denotes the concatenated output matrix across all the pilot subcarriers in the mth training frame, which is defined as $\mathbf{Y}_{n,m} = \left[\mathbf{y}_{n,m}[k_1], \mathbf{y}_{n,m}[k_2], \ldots, \mathbf{y}_{n,m}\left[k_{K_p}\right]\right]$ and $\mathbf{N}_{n,m} \in \mathbb{C}^{N_{RF}^R \times K_p}$ is obtained by a similar concatenation of the combined noise vectors $\mathbf{W}_{RF,m}^H \mathbf{n}_{n,m}[k_i]$ across all the pilot subcarriers. Employing (19), the dictionary matrix $\bar{\mathbf{\Phi}}_m \in \mathbb{C}^{N_{RF}^R \times G_R G_T}$ defined as $\bar{\mathbf{\Phi}}_m = \mathbf{\Phi}_m \mathbf{\Psi}$, is designed to be identical across all the pilot subcarriers in order to exploit the group-sparsity. An equivalent model for the group-sparse estimation can be derived as

$$\mathbf{y}_{n,m} = \widetilde{\mathbf{\Phi}}_m \mathbf{h}_{b,n} + \mathbf{n}_{n,m}, \tag{41}$$

where $\mathbf{y}_{n,m} = \mathrm{vec}\left(\mathbf{Y}_{n,m}^T\right) \in \mathbb{C}^{N_{RF}^R K_p \times 1}, \mathbf{n}_{n,m} = \mathrm{vec}\left(\mathbf{N}_{n,m}^T\right) \in \mathbb{C}^{N_{RF}^R K_p \times 1}$ and the dictionary matrix obeys $\widetilde{\mathbf{\Phi}}_m = \left(\bar{\mathbf{\Phi}}_m \otimes \mathbf{I}_{K_p}\right) \in \mathbb{C}^{N_{RF}^R K_p \times G_R G_T K_p}$. Furthermore, the covariance matrix $\widetilde{\mathbf{R}}_m \in \mathbb{C}^{N_{RF}^R K_p \times N_{RF}^R K_p}$ of the concatenated noise vector $\mathbf{n}_{n,m}$ can be obtained as $\widetilde{\mathbf{R}}_m \triangleq \mathbb{E}\left\{\mathbf{n}_{n,m}\mathbf{n}_{n,m}^H\right\} = \left(\mathbf{R}_m \otimes \mathbf{I}_{K_p}\right)$, whereas the group-sparse beamspace channel vector $\mathbf{h}_{b,n} = \mathrm{vec}\left(\mathbf{H}_{b,n}^T\right) \in \mathbb{C}^{G_R G_T K_p \times 1}$. The online O-SBL procedure of estimating the doubly-selective sparse mmWave MIMO OFDM channel $\mathbf{h}_{b,n}$ is developed next.

B. O-SBL for Online Doubly-selective Channel Estimation

The proposed O-SBL scheme begins with assigning the following parameterized Gaussian prior to the doubly-selective beamspace channel vector $\mathbf{h}_{b,n}$

$$p\left(\mathbf{h}_{b,n}; \mathbf{\Gamma}_n, \mathbf{G}_{c,n}\right) = \prod_{i=1}^{G_R G_T} \frac{1}{\left(\pi\gamma_{i,n}\right)^{K_p} \det\left(\mathbf{G}_{c,n}\right)} \exp\left(-\frac{\left(\mathbf{h}_{b,n}^i\right)^H \mathbf{G}_{c,n}^{-1} \mathbf{h}_{b,n}^i}{\gamma_i}\right), \tag{42}$$

where $\mathbf{h}_{b,n}^i$ represents the ith group of the vector $\mathbf{h}_{b,n}$. Let $\widehat{\mathbf{h}}_{b,n}^{m-1} \in \mathbb{C}^{G_R G_T K_p \times 1}$ and $\mathbf{\Sigma}_n^{m-1} \in \mathbb{C}^{G_R G_T K_p \times G_R G_T K_p}$ denote the LMMSE estimate of the beamspace channel matrix $\mathbf{h}_{b,n}$ and the associated error covariance matrix, respectively, obtained from the $(m-1)$st training frame in the nth block. Using the sequential LMMSE procedure [25], the estimate $\widehat{\mathbf{h}}_{b,n}^m$ and its error covariance $\mathbf{\Sigma}_n^m$ in the mth training frame can be recursively updated as

$$\begin{aligned} \widehat{\mathbf{h}}_{b,n}^m &= \widehat{\mathbf{h}}_{b,n}^{m-1} + \mathbf{K}_{n,m}\left(\mathbf{y}_{n,m} - \widetilde{\mathbf{\Phi}}_m \widehat{\mathbf{h}}_{b,n}^{m-1}\right), \\ \mathbf{\Sigma}_n^m &= \left(\mathbf{I}_{G_R G_T K_p} - \mathbf{K}_{n,m}\widetilde{\mathbf{\Phi}}_m\right)\mathbf{\Sigma}_n^{m-1}, \end{aligned} \tag{43}$$

$$\mathbf{K}_{n,m} = \mathbf{\Sigma}_n^{m-1}\widetilde{\mathbf{\Phi}}_m^H\left(\widetilde{\mathbf{\Phi}}_m \mathbf{\Sigma}_n^{m-1} \widetilde{\mathbf{\Phi}}_m^H + \widetilde{\mathbf{R}}_m\right)^{-1}. \tag{44}$$

Thus, the quantity $\widehat{\mathbf{h}}_{b,n}^M$ denotes the estimate of the concatenated beamspace channel and $\mathbf{\Sigma}_n^M$ represents the associated error covariance matrix obtained by processing all the M training frames in the nth block. The corresponding estimate of the hyperparameter matrix $\widehat{\mathbf{\Gamma}}_n = \mathrm{diag}\left(\{\widehat{\gamma}_{i,n}\}_{i=1}^{G_R G_T}\right)$ and the correlation matrix $\widehat{\mathbf{G}}_{c,n}$ at the end of the nth block, similar to (29) and (30), can be obtained using the following update equation

$$\widehat{\gamma}_{i,n} = \frac{1}{K_p}\mathrm{Tr}\left(\left(\widehat{\mathbf{G}}_{c,n-1}\right)^{-1}\left(\mathbf{\Sigma}_{i,n}^M + \widehat{\mathbf{h}}_{b,i,n}^M\left(\widehat{\mathbf{h}}_{b,i,n}^M\right)^H\right)\right), \tag{45}$$

$$\widehat{\mathbf{G}}_{c,n} = \frac{1}{G_R G_T}\sum_{i=1}^{G_R G_T} \frac{\mathbf{\Sigma}_{i,n}^M + \widehat{\mathbf{h}}_{b,i,n}^M\left(\widehat{\mathbf{h}}_{b,i,n}^M\right)^H}{\widehat{\gamma}_{i,n}}, \tag{46}$$

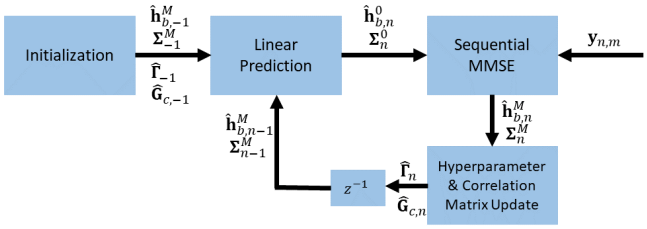


Fig. 1: Block diagram showing steps involved in sequential LMMSE based online O-SBL algorithm

where $\widehat{\mathbf{h}}_{b,i,n}^{M} = \widehat{\mathbf{h}}_{b,n}^{M}\left[(i-1)K_p+1:iK_p\right]$ and $\mathbf{\Sigma}_{i,n}^{M} = \mathbf{\Sigma}_{n}^{M}\left[(i-1)K_p+1:iK_p,(i-1)K_p+1:iK_p\right]$. Using the principles of optimal linear prediction [25], the quantities $\widehat{\mathbf{h}}_{b,n+1}^{0}$ and $\mathbf{\Sigma}_{n+1}^{0}$ for the $(n+1)$st block can be initialized as

$$\widehat{\mathbf{h}}_{b,n+1}^{0} = \rho\widehat{\mathbf{h}}_{b,n}^{M},\ \mathbf{\Sigma}_{n+1}^{0} = \rho^2\mathbf{\Sigma}_{n}^{M} + \left(1-\rho^2\right)\left(\widehat{\mathbf{\Gamma}}_n \otimes \widehat{\mathbf{G}}_{c,n}\right). \tag{47}$$

Note that in $\mathbf{\Sigma}_{n+1}^{0}$ above, the covariance matrix of the innovation incorporates $\widehat{\mathbf{\Gamma}}_n$, which captures the spatial group-sparsity inherent in the beamspace channel $\mathbf{h}_{b,n}$. Therefore, this leads to faster convergence of the online algorithm for sparse channel estimation. A block diagram of the proposed O-SBL approach is given in Fig. 1. The various steps conceived for the estimation of the sparse doubly-selective beamspace channel along with a suitable initialization procedure are succinctly stated in Algorithm-2. Finally, it is worth noting that the proposed approach is MMSE-optimal, and more importantly, of *online* nature, since it sequentially processes the output pilot $\mathbf{y}_{n,m}$ corresponding to each individual training frame m. Therefore, it has a significantly lower computational complexity and processing delay in comparison to the family of block processing schemes such as SBL and G-SBL described previously as well as those in the existing literature.

Let $\mathbf{H}_n$ and $\widehat{\mathbf{H}}_n$ denote the time-selective extensions of the concatenated channel $\mathbf{H}$ and its estimate $\widehat{\mathbf{H}}$, respectively. As described in our technical report [26], the BCRB of the estimate $\widehat{\mathbf{H}}_n$ can be formulated as

$$\text{MSE}\left(\widehat{\mathbf{H}}_n\right) \geq \text{Tr}\left(\widetilde{\mathbf{\Psi}}\mathbf{J}_{B,n}^{-1}\widetilde{\mathbf{\Psi}}^{H}\right), \tag{48}$$

where $\mathbf{J}_{B,n}$ can be computed recursively [29] as $\mathbf{J}_{B,n} = \left(\left(1-\rho^2\right)\left(\mathbf{\Gamma}\otimes\mathbf{G}_c\right)+\rho^2\mathbf{J}_{B,n-1}^{-1}\right)^{-1} + \widetilde{\mathbf{\Phi}}^{H}\widetilde{\mathbf{R}}^{-1}\widetilde{\mathbf{\Phi}}$. Furthermore, it follows from the established results of the existing contributions [12], [16], [17] that the converged hyperparameter vector $\widehat{\gamma}^{(p)} = \text{diag}\left(\widehat{\mathbf{\Gamma}}^{(p)}\right)$ employed in the initialization procedure of the proposed O-SBL technique is guaranteed to converge to a sparse vector, since the original techniques are guaranteed to converge for arbitrary initialization of the hyperparameters. Furthermore, since the O-SBL technique essentially computes the *a posteriori* mean $\widehat{\mathbf{h}}_{b,n}^{M}$ of the beamspace channel $\mathbf{h}_{b,n}$ in an online fashion, the sparse initialization of the hyperparameter matrix $\widehat{\mathbf{\Gamma}}_{-1}$ and the covariance matrix $\mathbf{\Sigma}_{-1}^{M}$ also guarantees gleaning sparse estimates from O-SBL. To justify this fact, let $\mathbf{h}_{b,n}^{i}$ denote the ith group of $\mathbf{h}_{b,n}$, which obeys $\mathbf{h}_{b,n}^{i} = \mathbf{h}_{b,n}\left[(i-1)K_p+1:iK_p\right]$. Mathematically, if $\widehat{\gamma}_{i,-1}=0$, i.e. if the prior variance is zero, one can readily infer that

$$\text{Prob}\left(\mathbf{h}_{b,n}^{i}(k)=0|\mathbf{y}_0,\cdots\mathbf{y}_n;\widehat{\gamma}_{i,-1}=0\right)=1,\forall\, 1\leq k\leq K_p.$$

Algorithm 2: O-SBL for mmWave hybrid MIMO OFDM systems

Input: Observation $\mathbf{y}_{n,m}\in\mathbb{C}^{N_{RF}^{R}K_p\times 1}$, sensing matrix $\widetilde{\mathbf{\Phi}}_m\in\mathbb{C}^{N_{RF}^{R}K_p\times G_RG_TK_p}$, noise covariance matrix $\widetilde{\mathbf{R}}_m\in\mathbb{C}^{N_{RF}^{R}K_p\times N_{RF}^{R}K_p}$, correlation coefficient ρ, converged estimate $\widehat{\mathbf{\Gamma}}^{(p)}$ and $\widehat{\mathbf{G}}_c^{(p)}$ of the hyperparameter matrix $\mathbf{\Gamma}_n$ and $\mathbf{G}_{c,n}$ obtained from the 0th block using LCG-SBL

Output: $\widehat{\mathbf{H}}_n[k], \forall 0\leq k\leq K-1$

Initialization: $\widehat{\mathbf{\Gamma}}_{-1}=\widehat{\mathbf{\Gamma}}^{(p)}$, $\widehat{\mathbf{G}}_{c,-1}=\widehat{\mathbf{G}}_c^{(p)}$, $\widehat{\mathbf{h}}_{b,-1}^{M}=\mathbf{0}_{G_RG_TK_p\times 1}$, $\mathbf{\Sigma}_{-1}^{M}=\widehat{\mathbf{\Gamma}}^{(p)}\otimes\widehat{\mathbf{G}}_c^{(p)}$

for $n=0,1,2,\ldots$ **do**

  Initialize $\widehat{\mathbf{h}}_{b,n}^{0}$ and $\mathbf{\Sigma}_{n}^{0}$ using (47)

  for $m=1,2,\ldots,M$ **do**

    Compute $\mathbf{K}_{n,m}$ using (44), and update $\widehat{\mathbf{h}}_{b,n}^{m}$ and $\mathbf{\Sigma}_{n}^{m}$ using (43)

  end

  Update $\widehat{\gamma}_{i,n}$ and $\widehat{\mathbf{G}}_{c,n}$ using (45) and (46), respectively

  $\widehat{\mathbf{\Gamma}}_n=\text{diag}\left(\widehat{\gamma}_{1,n},\widehat{\gamma}_{2,n},\ldots,\widehat{\gamma}_{G_RG_T,n}\right)$

  return: Obtain $\widehat{\mathbf{H}}_n[k]$ using the procedure described after Eq. (30)

end

C. Computational Complexity Analysis

Due to lack of space, the detailed derivations of the computational complexities of the proposed and existing mmWave MIMO OFDM channel estimation techniques are given in our technical report [26]. The key implications of the results are discussed below. As shown in Table-I of the technical report [26], the computation complexity of SBL is of the order $\mathcal{O}\left(KG_R^3G_T^3+KM\left(N_{RF}^{R}\right)^3\right)$. It is evident from Table-II of the technical report [26], the joint processing of the K_p pilot subcarriers in the G-SBL procedure leads to a higher complexity of the order $\mathcal{O}\left(G_R^3G_T^3K_p^3+M\left(N_{RF}^{R}\right)^3\right)$, which arises due to the fact that it necessitates the inversion of the $[G_RG_TK_p\times G_RG_TK_p]$-dimensional matrix for the computation of $\widetilde{\mathbf{\Sigma}}^{(p)}$. Thus, SBL has a lower complexity, but it also has poor performance, as described in Section-VII, since it does not exploit the group-sparsity of the mmWave MIMO OFDM channel. It follows from Table-III of the technical report [26] that the proposed LCG-SBL technique has a complexity order of $\mathcal{O}\left(G_R^3G_T^3+M\left(N_{RF}^{R}\right)^3+G_RG_TK_p^3\right)$. This is due to the fact that it requires inversion of a smaller $[G_RG_T\times G_RG_T]$-dimensional matrix for computing $\mathbf{\Sigma}^{(p)}$. Moreover, as shown in Section-VII, its NMSE performance is close to that of the

G-SBL. Hence, the proposed LCG-SBL technique is efficient as well as of low complexity. Finally, as derived in Table-V of the technical report [26], the worst case complexity order of the SOMP is $\mathcal{O}\left(G_R^2 G_T^2 K + M^3 \left(N_{RF}^R\right)^3\right)$, which is lower compared to the SBL-based approaches. However, as shown in our simulation results, its performance is significantly poor.

For doubly-selective sparse mmWave MIMO OFDM channel estimation, Algorithm-1 of the technical report [26] develops the SBL-KF [20] technique. As it is evident from Table-VI of the technical report [26], its computational cost is $\mathcal{O}\left(G_R^2 G_T^2 K_p^3 M N_{RF}^R + M^3 \left(N_{RF}^R\right)^3 K_p^3\right)$, where the order term $\mathcal{O}\left(M^3 \left(N_{RF}^R\right)^3 K_p^3\right)$ appears due to inversion of $[M N_{RF}^R K_p \times M N_{RF}^R K_p]$-element matrix while computing $\mathbf{K}_n$. The proposed O-SBL technique described in Section-V-B of our paper completely obviates this through an online estimation procedure, which requires the inversion of only an $[N_{RF}^R K_p \times N_{RF}^R K_p]$-element matrix for the evaluation of the Kalman gain $\mathbf{K}_{n,m}$ in (44). Thus, as evaluated in Table-IV of the technical report [26], its block-wise complexity order is $\mathcal{O}\left(G_R^2 G_T^2 K_p^3 M N_{RF}^R + M \left(N_{RF}^R\right)^3 K_p^3\right)$, which may be deemed moderate, making it attractive for practical implementation.

VI. Hybrid Precoder Design for mmWave Hybrid MIMO OFDM Systems

This section develops the framework for data transmission in the mmWave MIMO OFDM system employing the CSI estimate obtained using the various schemes presented in this paper. The existing contributions, such as [21], [22], assume perfect CSI for designing the RF precoder and combiner for a frequency-flat channel. Furthermore, they either consider the feasible array response vectors to be perfectly known or employ an array response dictionary matrix constructed from the quantized angular-grid for representing the RF precoder $\mathbf{F}_{RF}$. To the best of our knowledge, none of the existing works have directly employed the estimate $\widehat{\mathbf{h}}_b$ of the underlying beamspace channel to design the RF precoder in a mmWave MIMO-OFDM system, which is naturally the most suitable approach, given the availability of the beamspace domain channel estimates. The proposed hybrid transceiver design addresses this open problem.

To this end, the choice of a suitable criterion for the design of optimal precoders and combiners, at the transmitter and receiver, respectively, is of significant importance. Let the transmitted signal vector, denoted by $\mathbf{t}[k] \in \mathbb{C}^{N_T \times 1}$, be generated as $\mathbf{t}[k] = \mathbf{F}_{RF}\mathbf{F}_{BB}[k]\,\mathbf{x}[k]$, with the covariance matrix of the baseband symbol vector $\mathbf{x}[k]$ normalized as $\mathbb{E}\left[\mathbf{x}[k]\,\mathbf{x}^H[k]\right] = \frac{1}{N_s}\mathbf{I}_{N_s}$, i.e., $\mathbf{x}[k]$ is comprised of independent and identically distributed (i.i.d.) symbols with power $\frac{1}{N_s}$. Let the maximum transmit power be denoted by P_t. It can be shown that, in order to restrict the total transmit power after precoding to P_t, i.e., $\sum_{k=1}^{K} \mathrm{Tr}\{\mathbb{E}\left[\mathbf{t}[k]\,\mathbf{t}^H[k]\right]\} \leq P_t N_s$, we employ the equivalent constraint $\sum_{k=1}^{K} \|\mathbf{F}_{RF}\mathbf{F}_{BB}[k]\|_F^2 \leq P_t N_s$. The optimal transmit precoders $\mathbf{F}_{BB}^{\mathrm{opt}}[k] \in \mathbb{C}^{N_{RF}^T \times N_s}\ \forall k$, and $\mathbf{F}_{RF}^{\mathrm{opt}} \in \mathbb{C}^{N_R \times N_{RF}^T}$ can be designed via the maximization of the sum of the mutual information, which is formulated as

$$\left(\left\{\mathbf{F}_{BB}^{\mathrm{opt}}[k]\right\}_{k=1}^{K}, \mathbf{F}_{RF}^{\mathrm{opt}}\right) = \underset{\left(\left\{\mathbf{F}_{BB}[k]\right\}_{k=1}^{K}, \mathbf{F}_{RF}\right)}{\arg\max} \sum_{k=1}^{K} \log_2\left|\mathbf{I}_{N_R} + \widetilde{\mathbf{H}}[k]\,\widetilde{\mathbf{H}}^H[k]\right| \quad \text{subject to} \sum_{k=1}^{K} \|\mathbf{F}_{RF}\mathbf{F}_{BB}[k]\|_F^2 \leq P_t N_s, \tag{49}$$

where $\widetilde{\mathbf{H}}[k] = \mathbf{H}[k]\,\mathbf{F}_{RF}\mathbf{F}_{BB}[k] \in \mathbb{C}^{N_R \times N_s}$ denotes the equivalent baseband channel. However, the optimization problem stated above is non-convex, hence it is challenging to solve as the elements of $\mathbf{F}_{RF}$ are constrained to have constant magnitude [4], [10]. In order to obtain a computationally tractable solution for the precoders, we may proceed as follows. Let $\mathbf{F}[k] = \mathbf{F}_{RF}\mathbf{F}_{BB}[k] \in \mathbb{C}^{N_T \times N_s}$ denote the equivalent digital precoder, with the constant magnitude constraint above relaxed. The optimal equivalent digital precoder $\mathbf{F}^{\mathrm{opt}}[k]$ can now be obtained as the solution of the modified optimization problem expressed as

$$\left\{\mathbf{F}^{\mathrm{opt}}[k]\right\}_{k=1}^{K} = \underset{\left\{\mathbf{F}[k]\right\}_{k=1}^{K}}{\arg\max} \sum_{k=1}^{K} \log_2\left|\mathbf{I}_{N_R} + \widetilde{\mathbf{H}}[k]\,\widetilde{\mathbf{H}}^H[k]\right| \quad \text{subject to} \sum_{k=1}^{K} \|\mathbf{F}[k]\|_F^2 \leq P_t N_s. \tag{50}$$

The closed form solution for the above problem can be obtained using the popular water-filling procedure as follows. Let the singular value decomposition (SVD) of $\mathbf{H}[k]$ be given as $\mathbf{H}[k] = \mathbf{U}[k]\,\mathbf{\Sigma}[k]\,\mathbf{V}^H[k]$. The optimal equivalent digital precoder $\mathbf{F}^{\mathrm{opt}}[k]$ is given as

$$\mathbf{F}^{\mathrm{opt}}[k] = \mathbf{V}_1[k]\,\mathbf{P}^{1/2}[k], \tag{51}$$

where the matrix $\mathbf{V}_1[k] \in \mathbb{C}^{N_T \times N_s}$ is the submatrix comprised of the first N_s columns of $\mathbf{V}[k]$ and $\mathbf{P}[k] \in \mathbb{R}_{+}^{N_s \times N_s}$ represents a diagonal power allocation matrix with the diagonal elements $p_{k,s}, \forall 1 \leq s \leq N_s$, given by: $p_{k,s} = \max\left\{0, \left(\lambda - \frac{\sigma_n^2}{(\mathbf{\Sigma}[k](s,s))^2}\right)\right\}$. The constant λ is chosen to satisfy the constraint $\sum_{k=1}^{K}\sum_{s=1}^{N_s} p_{k,s} \leq P_t N_s$. Similar to [21] that considers a frequency-flat mmWave MIMO channel, the hybrid precoders $\mathbf{F}_{BB}^{\mathrm{opt}}[k]$ and $\mathbf{F}_{RF}^{\mathrm{opt}}$ for the frequency-selective mmWave MIMO OFDM system can now be chosen to obtain the best possible approximation to the ideal precoder

Algorithm 3: Design of hybrid precoders and combiners from the estimated beamspace domain CSI

Input: Estimated beamspace channel matrix $\widehat{\mathbf{H}}_b \in \mathbb{C}^{G_R G_T \times K_p}$, optimal equivalent digital precoder $\mathbf{F}^{\text{opt}}[k] \in \mathbb{C}^{N_T \times N_s}$ and combiner $\mathbf{W}^{\text{opt}}[k] \in \mathbb{C}^{N_R \times N_s}$, number of receive and transmit RF chains N_{RF}^R and N_{RF}^T

Output: $\mathbf{F}_{BB}[k]$, $\mathbf{W}_{BB}[k], \forall 1 \leq k \leq K$, and $\mathbf{F}_{RF}$, $\mathbf{W}_{RF}$

Initialization: $\mathbf{F}_{RF} = [\,]$, $\mathbf{W}_{RF} = [\,]$, set $\mathbf{h}_{\text{temp}} = \text{abs}\left(\widehat{\mathbf{H}}_b\right) \times \text{diag}\left(\mathbf{I}_{K_p}\right)$

Obtain the set $\mathcal{S}$, which contains the indices of the elements of $\mathbf{h}_{\text{temp}}$, when magnitude of elements of $\mathbf{h}_{\text{temp}}$ are sorted in descending order

for $i = 1, 2, \ldots, N_{RF}^T$ **do**

  $g = \lfloor (\mathcal{S}(i) - 1)/G_R \rfloor + 1$; $\mathbf{F}_{RF} = [\mathbf{F}_{RF}\ \mathbf{a}_T(\theta_g)]$

end

for $i = 1, 2, \ldots, N_{RF}^R$ **do**

  $g = \text{rem}(\mathcal{S}(i) - 1, G_R) + 1$;
  $\mathbf{W}_{RF} = [\mathbf{W}_{RF}\ \mathbf{a}_R(\phi_g)]$

end

for $k = 1, 2, \ldots, K$ **do**

  $\mathbf{F}_{BB}[k] = (\mathbf{F}_{RF})^{\dagger}\mathbf{F}^{\text{opt}}[k]$;
  $\mathbf{W}_{BB}[k] = (\mathbf{W}_{RF})^{\dagger}\mathbf{W}^{\text{opt}}[k]$

end

return: $\mathbf{F}_{BB}[k]$, $\mathbf{W}_{BB}[k], \forall 1 \leq k \leq K$, and $\mathbf{F}_{RF}$, $\mathbf{W}_{RF}$

$\mathbf{F}^{\text{opt}}[k]$ via the optimization problem below

$$\left(\left\{\mathbf{F}_{BB}^{\text{opt}}[k]\right\}_{k=1}^{K}, \mathbf{F}_{RF}^{\text{opt}}\right) = \underset{\left(\left\{\mathbf{F}_{BB}[k]\right\}_{k=1}^{K}, \mathbf{F}_{RF}\right)}{\arg\min} \sum_{k=1}^{K} \|\mathbf{F}^{\text{opt}}[k] - \mathbf{F}_{RF}\mathbf{F}_{BB}[k]\|_F$$
$$\text{subject to} \sum_{k=1}^{K} \|\mathbf{F}_{RF}\mathbf{F}_{BB}[k]\|_F^2 \leq P_t N_s. \quad (52)$$

Note that, since the locations of the dominant components of the mmWave MIMO OFDM beamspace CFR $\mathbf{H}_b[k]$ may vary across the subcarriers, a straightforward subcarrier-wise implementation of the precoder design procedures, such as those described in [21], [22], may not result in frequency-flat RF precoders $\mathbf{F}_{RF}$. Therefore, the RF precoder $\mathbf{F}_{RF}$ has to be jointly optimized across all the subcarriers of the mmWave MIMO-OFDM system, as shown in the optimization problem (52). One can now use the following theorem to simplify the joint hybrid baseband-RF precoder design problem detailed above.

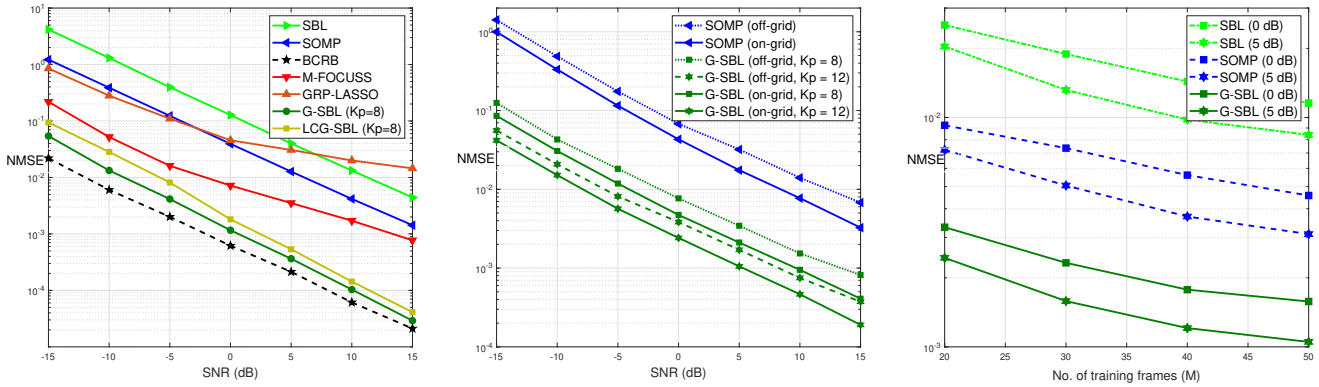
Theorem 1: $\mathcal{C}(\mathbf{F}^{\text{opt}}[k]) \subseteq \mathcal{C}(\mathbf{A}_T), \forall k$, where $\mathcal{C}(\cdot)$ denotes the column-space of a matrix.

Proof : Given in Appendix B.

The theorem above provides an important insight into the choice of the ideal frequency-flat RF precoder $\mathbf{F}_{RF}$ that contains only N_{RF} columns. It can be noted that employing $N_{RF}^T = N_{\text{ray}}$ number of RF chains and setting $\mathbf{F}_{RF} = \mathbf{A}_T$ as the RF precoder, the approximation error $\sum_{k=1}^{K} \|\mathbf{F}^{\text{opt}}[k] - \mathbf{F}_{RF}\mathbf{F}_{BB}[k]\|_F$ can be made zero. However, in practical scenarios wherein the transmit array response matrix $\mathbf{A}_T$ is frequently unknown and the number of RF chains obeys $N_{RF}^T < N_{\text{ray}}$, one can construct the RF precoder $\mathbf{F}_{RF}$ by choosing the N_{RF} dominant array response vectors obtained during the estimation of the beamspace channel $\mathbf{h}_b$. Moreover, the transmit array response vectors are comprised of constant-magnitude elements, as seen from (8). Thus, choosing the columns of $\mathbf{F}_{RF}$ as the transmit array response vectors also satisfies the implicit non-convex constraint in the optimization problem (52). Finally, the baseband precoder for the kth subcarrier can be obtained using the LS solution of $\mathbf{F}_{BB}[k] = (\mathbf{F}_{RF})^{\dagger}\mathbf{F}^{\text{opt}}[k]$. A concise step-by-step description of the proposed hybrid precoder design procedure for the mmWave MIMO OFDM system is presented in Algorithm-3. A similar approach can also be employed for deriving the hybrid combiners $\mathbf{W}_{RF}$ and $\mathbf{W}_{BB}[k]$. Note that the SOMP technique, as described in [21], requires N_{RF} iterations for selecting the N_{RF} dominant array response vectors via a computationally intensive correlation method (Step-4 and 5 of Algorithm-1 in [21]), followed by an intermediate LS solution in each iteration. By contrast, the proposed hybrid precoder design framework is directly able to compute the final baseband precoder of each subcarrier using the LS solution, once the RF precoder is derived using the estimated beamspace domain CSI. Thus, the proposed hybrid precoder design has a significantly lower computational cost, while performing very close to the SOMP, as demonstrated in our simulation results of Fig. 3(c). Furthermore, the CSI has to be fed back to the transmitter in a mmWave MIMO system. The framework for beamspace domain CSI estimation, followed by our hybrid precoder design developed requires significantly lower feedback, since the receiver only has to feed back the N_{RF} indices of the dominant beamspace components together with their quantized gains in order to construct the hybrid precoder of the transmitter.

VII. Simulation results

This section presents simulation results for characterizing the performance of the offline and online channel estimation schemes proposed for mmWave hybrid MIMO OFDM systems and compare them to that of the existing schemes. A mmWave MIMO OFDM system is considered where the number of TAs and RAs are set as $N_T = N_R \in \{8, 16, 32\}$, the number of RF chains at transmitter and receiver set as $N_{RF}^T = N_{RF}^R \in \{4, 8\}$ and the number of subcarriers is $K \in \{16, 128, 256\}$. The antenna spacings between the TA and RA arrays are set as $d_T = d_R = \frac{\lambda}{2}$. The AoA/ AoD space is partitioned into $G_R = G_T \in \{16, 32, 48\}$ angular grid points. The frequency-selective mmWave MIMO channel is assumed to be spatially sparse with $N_{\text{cl}} = 4$ clusters, $N_{\text{ray},i} \in \{1, 4\}$ rays per cluster and $L = 4$ delay taps. The complex channel gain α_{ij} corresponding to the jth ray in the ith cluster is modeled as $\alpha_{ij} \sim \mathcal{CN}(0, 1)$ and the raised-cosine filter is set to have a roll-off factor of 0.85 [4]. The SNR is defined as $\frac{1}{\sigma^2}$. For



(a) (b) (c)

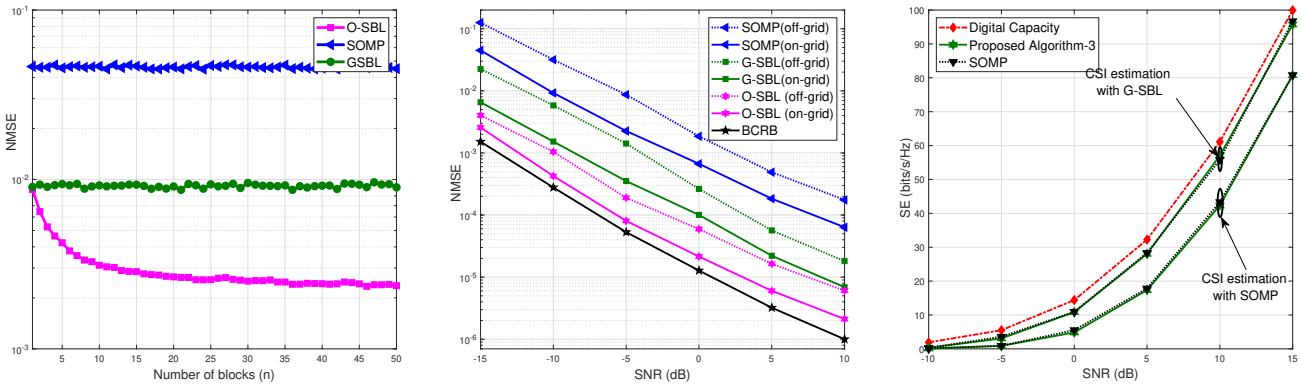
Fig. 2: (a) NMSE vs SNR comparison in an *on-grid* scenario for $N_R = N_T = 8$, $N_{RF}^R = N_{RF}^T = 4$, $M = 32$, $K = 256$, $K_p = 8$, $G_R = G_T = 16$, along with BCRB using Eq. (36). (b) NMSE vs SNR comparison in *on-grid* and *off-grid* scenarios for $N_R = N_T = 32$, $N_{RF}^R = N_{RF}^T = 8$, $M = 64$, $K = 64$ and $G_R = G_T = 48$. (c) NMSE vs the number of training frames M for $N_R = N_T = 16$, $N_{RF}^R = N_{RF}^T = 4$, $K = 128$, $K_p = 8$ and $G_R = G_T = 16$.

SBL based approaches, the stopping parameters are set as $\epsilon = 10^{-6}$ and $p_{\max} = 100$.

A. Quasi-static Offline mmWave MIMO OFDM Channel Estimation

Fig. 2(a) depicts the NMSE performance comparison of the proposed SBL-based schemes with the existing multiple-measurement-vectors (MMV)-based SOMP scheme of [4], in terms of the normalized mean squared error (NMSE) defined as NMSE $\triangleq \frac{\sum_{k=1}^{K} \|\widehat{\mathbf{H}}[k]-\mathbf{H}[k]\|_F^2}{KN_RN_T}$, where $\widehat{\mathbf{H}}[k]$ denotes the estimate of the channel. The performance is also benchmarked against the BCRB derived in (36) for quasi-static scenarios. The mmWave MIMO OFDM setup has the parameter values of $N_R = N_T = 8$, $N_{RF}^R = N_{RF}^T = 4$ and $G_R = G_T = 16$. The number of training frames and total number of subcarriers are set to $M = 32$ and $K = 256$, respectively. The AoA/ AoDs corresponding to each spatial path for this scenario, also termed *on-grid*, is obtained uniformly from the angular grid by setting $N_{\text{ray},i} = 1$. The number of pilot subcarriers for the SOMP-based scheme is set to $K_p = K$, while it is set to $K_p = 8$ for the proposed G-SBL technique. Observe from the figure that the proposed G-SBL scheme yields the best NMSE performance in comparison to the SOMP and to the single-measurement-vector (SMV)-based SBL scheme that does not exploit the group-sparsity. This clearly demonstrates the performance benefits achieved by exploiting the group-sparsity. The performance of SOMP is poor in comparison to the proposed G-SBL scheme due to the fact that it is sensitive to both the choice of the dictionary matrix $\bar{\mathbf{\Phi}}$ and to the stopping criterion employed. On the other hand, as descried in Section-V-C, the computational cost of SOMP is lower than that of the SBL-based approaches. Thus, there is a trade-off between the NMSE performance and the computational cost. However, it can be readily observed that the proposed LCG-SBL scheme is also as efficient as G-SBL, since its NMSE performance is very close to that of the G-SBL, but with a significantly lower computational cost. The performance is also compared to that of the popular MMV-based group least absolute shrinkage and selection operator (GRP-Lasso) [30] and to the multiple focal underdetermined system solver (M-FOCUSS) [31], which are based on l_1 and $l_p, p < 1$ norm minimization. The regularization parameter for the M-FOCUSS is set to the value of the noise variance σ^2, whereas for the GRP-Lasso, it is empirically tuned for minimizing the NMSE. Furthermore, for M-FOCUSS, the norm parameter is set to $p = 0.8$, whereas the stopping threshold and the maximum number of iterations are set to 10^{-5} and 800, respectively. Since the performance of the GRP-Lasso critically depends on the user-defined regularization parameter and that of the M-FOCUSS suffers from convergence problems, their NMSE is also inferior in comparison to that of the proposed G-SBL. It is also worth noting that the proposed G-SBL based scheme employs $K_p = 8$ pilot subcarriers, which is significantly lower than that of the SOMP that employs all the $K = 256$ subcarriers for pilot transmission. Thus, the proposed G-SBL technique is also bandwidth efficient, since it can employ the remaining $K - K_p$ subcarriers for data transmission. This is of significant importance in 5G new radio (5G-NR) systems for supporting ultra-high data rates. Finally, the NMSEs of the G-SBL and LCG-SBL techniques are also seen to be close to the corresponding BCRB, which strongly advocates the case for G-SBL and LCG-SBL in practical scenarios.

Fig. 2(b) illustrates the NMSE comparison of various schemes for the mmWave MIMO OFDM setup with parameters $N_R = N_T = 32$, $N_{RF}^R = N_{RF}^T = 8$ and $G_R = G_T = 48$. A trend similar to that of Fig. 2(a) can be discerned, wherein the LCG-SBL was seen to achieve significantly better NMSE performance in comparison to SOMP. The performance of the LCG-SBL scheme upon increasing the number of pilot subcarriers K_p is also shown progressively improving the estimation accuracy, as expected. Another interesting aspect of this plot is that it also shows the NMSE for an *off-grid* scenario, in which the true AoAs/ AoDs for the N_{ray} rays



(a) (b) (c)

Fig. 3: (a) NMSE vs number of blocks n comparison for the doubly-selective scenario with $N_R = N_T = 8$, $N_{RF}^R = N_{RF}^T = 4$, $M = 32$, $G_R = G_T = 16$. (b) NMSE vs SNR comparison for $N_R = N_T = 16$, $N_{RF}^R = N_{RF}^T = 8$, $M = 48$, $G_R = G_T = 24$ along with BCRB using Eq. (48). (c) SE vs SNR comparison for $N_R = N_T = 16$, $N_{RF}^T = N_{RF}^R = 6$, $N_s = 4$, $M = 32$, $G_R = G_T = 16$.

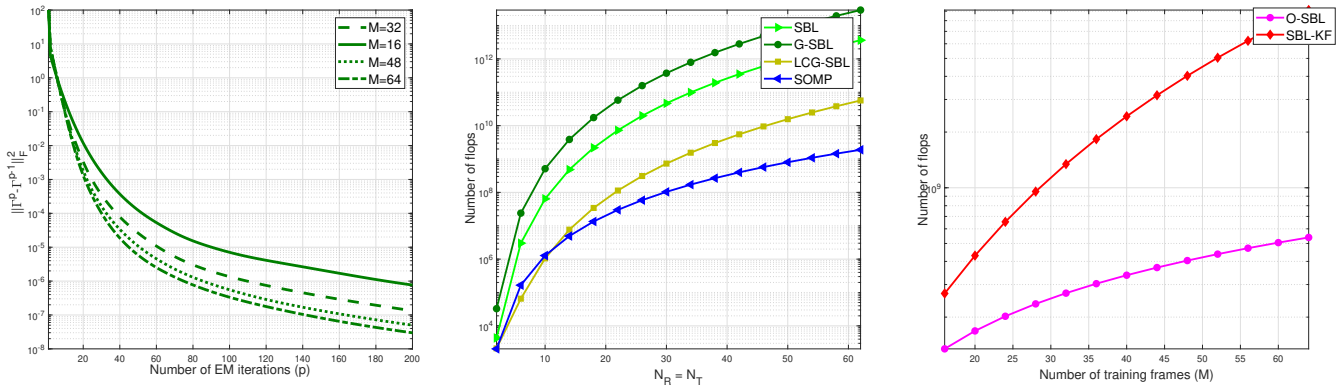
differ from the quantized angles in the set of feasible AoA/ AoD space Φ_R, Θ_T that is used to construct the sparsifying dictionary matrix $\mathbf{\Psi}$. For this scenario, the AoA/ AoDs of all the $N_{\text{ray},i} = 4$ rays associated with a cluster are assumed to have a Laplacian distribution around the mean-angle of the cluster with the standard deviation of $\sigma_{AS} = 1/10$ radian. Furthermore, the mean-angles of all the clusters are assumed to be uniformly distributed over the angular grid. We can readily note the performance degradation of various schemes for the off-grid scenario. However, even with this marginal degradation, the performance of the LCG-SBL scheme with a reduced number of pilot subcarriers is improved in comparison to the existing SOMP. Fig. 2(c) compares the NMSE of the G-SBL and SOMP techniques with varying the number of training frames M. For this, a mmWave MIMO OFDM setup is considered with parameters $N_R = N_T = 16$, $N_{RF}^R = N_{RF}^T = 4$, $G_R = G_T = 16$, $K = 128$ and $K_p = 8$. The NMSE performance of both the techniques is naturally seen to improve upon increasing the number of training frames, which can be attributed to the larger number of measurements. However, it can also be observed that the performance of the G-SBL scheme using $M = 20$ training frames is better than that of its SOMP counterpart employing $M = 50$. Thus, G-SBL achieves the desired level of estimation accuracy at significantly lower training overheads.

B. Doubly-selective Online mmWave MIMO OFDM Channel Estimation

To characterize the proposed estimation scheme in Section-V for a doubly-selective mmWave MIMO OFDM channel, a Q-band system is considered having a carrier frequency of 28 GHz. The user velocity is set to 5 km/h, which leads to a sizable Doppler shift of $f_D = 130$ Hz at this high carrier frequency. The coherence time is set to $T_c = 5$ ms with the block length of $T_B = T_c/10$. Substituting these values, the temporal correlation coefficient ρ of the doubly-selective mmWave MIMO OFDM channel is obtained as $\rho = J_0\left(2\pi f_D T_B\right) \approx 0.9983$. Fig. 3(a) shows the NMSE performance of the proposed online O-SBL scheme with respect to the number of blocks n for the parameters $N_R = N_T = 8$ and $N_{RF}^R = N_{RF}^T = 4$. The SNR for the scenario is set to 0 dB, while the number of training frames is set to $M = 32$. It can be observed that the NMSE of the O-SBL algorithm that beneficially exploits the temporal correlation is significantly better than that of the SOMP and G-SBL. Moreover, the performance gap is seen to increase with the number of blocks n, thanks to the improved estimation accuracy of the hyperparameter matrix $\mathbf{\Gamma}$. Finally, it is also worth noting that while the G-SBL performs several EM iterations per block, the online O-SBL procedure performs only a single iteration. Thus, it has a low complexity coupled with the ability to constantly track the channel, which renders it eminently suitable for practical implementation. Fig. 3(b) depicts our NMSE versus SNR performance comparison of a 16×16 system having 8 RF chains at both ends and $M = 48$ training frames. This lends additional credence to the trend observed previously, wherein the O-SBL scheme was seen to lead in performance in comparison to its quasi-static competitors, namely to G-SBL and SOMP. Furthermore, the NMSE performance of O-SBL is seen to be close to that of the recursive BCRB for the doubly-selective scenario, as determined in (48).

C. Spectral-Efficiency (SE) performance

Fig. 3(c) presents the SE performance of the system with the hybrid precoder/ combiner designed using the proposed Algorithm-3 and existing SOMP [21] with the mmWave MIMO channel estimates obtained from the various sparse channel estimation schemes. The various parameters of the system are set as $N_R = N_T = 16$, $N_{RF}^T = N_{RF}^R = 6$, $N_{\text{cl}} = 6$, $N_s = 4$ and $G_R = G_T = 16$. The expression used for evaluating the resultant SE is $\text{SE} = \sum_{k=1}^{K} \log_2 \left| \mathbf{I}_{N_s} + \frac{1}{N_s} \mathbf{R}_n^{-1} \mathbf{H}_{\text{eq}}[k] \mathbf{H}_{\text{eq}}^H[k] \right|$, where $\mathbf{H}_{\text{eq}}[k] = \bar{\mathbf{W}}_{BB}^H[k] \bar{\mathbf{W}}_{RF}^H \mathbf{H}[k] \bar{\mathbf{F}}_{RF} \bar{\mathbf{F}}_{BB}[k]$ denotes the equivalent baseband channel and



(a) (b) (c)

Fig. 4: (a) $\left\|\widehat{\mathbf{\Gamma}}^{(p)}-\widehat{\mathbf{\Gamma}}^{(p-1)}\right\|_F^2$ vs number of EM iterations p for $N_R=N_T=8$, $N_{RF}^R=N_{RF}^T=4$, $G_R=G_T=16$ and $K_p=8$. (b) Complexity comparison for $N_R=N_T=G_R=G_T=2N_{RF}^R=2N_{RF}^T=M/2, K_p=8, K=64$ in quasi-static scenario. (c) Complexity comparison for $N_R=N_T=8, G_R=G_T=10, N_{RF}^R=N_{RF}^T=4, K_p=8, K=64$ in doubly-selective scenario.

$\mathbf{R}_n[k] = \sigma^2\bar{\mathbf{W}}_{BB}^H[k]\,\bar{\mathbf{W}}_{RF}^H\bar{\mathbf{W}}_{RF}\bar{\mathbf{W}}_{BB}[k]$ is the noise covariance. The hybrid precoders $\bar{\mathbf{F}}_{BB}[k], \bar{\mathbf{F}}_{RF}$ and combiners $\bar{\mathbf{W}}_{BB}[k], \bar{\mathbf{W}}_{RF}$ are obtained using the estimated CSI $\widehat{\mathbf{H}}[k]$. The SE is benchmarked with respect to that of an ideal genie receiver, wherein the optimal digital precoder/ combiner matrices are designed assuming perfect knowledge of the mmWave MIMO channel matrix $\mathbf{H}[k]$. It can be readily observed that the SE evaluated using the hybrid precoders and combiners obtained from the proposed Algorithm-3 is very close to that of the SOMP technique of [21] suitably extended for mmWave MIMO OFDM systems. This shows the efficacy of our proposed limited CSI based low complexity hybrid transceiver design that employs the estimated channel of the beamspace domain for designing the RF precoder, followed by a simple LS solution for the baseband precoders. On the other hand, the SOMP requires several iterations to yield these quantities. The figure also reflects the fact that the improved channel estimation accuracy of the proposed G-SBL technique successfully translates into a corresponding gain in the end-to-end SE.

D. Complexity and Convergence Rate

Fig. 4(a) plots the convergence criterion of the proposed G-SBL based approach with respect to the number of EM iterations for different values of the training frames M. This signifies the number of EM iterations required for the convergence of hyperparameters, i.e., $\left\|\widehat{\mathbf{\Gamma}}^{(p)}-\widehat{\mathbf{\Gamma}}^{(p-1)}\right\|_F^2 < \epsilon$. For a fixed $\epsilon = 10^{-6}$, it can be readily observed that the number of iterations required for convergence decreases with increasing number of training frames M. Fig. 4(b) and Fig. 4(c) demonstrate a numerical comparison of the complexities in number of floating point operations (flops). As seen in Fig. 4(b) for the quasi-static scenario, the complexities of the various schemes can be ordered as $\mathcal{O}(\text{G-SBL}) > \mathcal{O}(\text{SBL}) > \mathcal{O}(\text{LCG-SBL}) > \mathcal{O}(\text{SOMP})$. Interesting, it can be seen that the LCG-SBL scheme has a complexity even lower than that of the SBL, as described in Section-V-C. Similarly, for the doubly-selective scenario of Fig. 4(c), $\mathcal{O}(\text{SBL-KF}) > \mathcal{O}(\text{O-SBL})$. This validates the results derived in Section-V-C.

VIII. Conclusions

We initially presented an SBL approach for channel estimation in mmWave hybrid MIMO OFDM systems that exploits the sparsity of the beamspace mmWave channel for each individual subcarrier. Subsequently, a G-SBL procedure has also been derived for the joint estimation of the beamspace channel vectors across all the pilot subcarriers, thus exploiting the group-sparsity and the associated frequency-domain correlation, which leads to a substantial performance improvement. An online procedure, termed O-SBL, has also been conceived for the estimation of a doubly-selective mmWave MIMO OFDM channel, which is particularly appealing for practical implementation due to its low processing delay and complexity. The BCRBs that determine the lower bounds for the achievable MSE were also derived for both the quasi-static and doubly-selective channel estimation techniques. This was followed by the development of a novel scheme for hybrid precoder/ combiner design that successfully exploited the knowledge of the beamspace domain limited CSI estimated using the proposed techniques. Simulation results were presented for demonstrating the improved performance of the proposed techniques in comparison to similar schemes in the existing literature, as well as the close agreement with various performance bounds. Future extensions of this work may explore the approximate message passing (AMP) [32] based implementation of SBL and sparse adaptive channel estimation schemes [33] to limit the complexity of channel estimation in mmWave hybrid MIMO OFDM systems. Additionally, one can also incorporate the effect of timing and synchronization errors in the system model followed by suitable updates of the SBL-based approaches.

APPENDIX A
UPDATE OF CORRELATION MATRIX $\mathbf{G}_c$ [17]

Ignoring terms that do not depend on the correlation matrix $\mathbf{G}_c$, the quantity $\mathbb{E}\{\log\, p\big(\mathbf{h}_b;\mathbf{\Gamma},\mathbf{G}_c\big)\}$ can be expressed as $\mathbb{E}\{\log\, p\big(\mathbf{h}_b;\mathbf{\Gamma},\mathbf{G}_c\big)\} \propto -G_R G_T \log\det(\mathbf{G}_c) - \sum_{i=1}^{G_R G_T} \frac{\mathbb{E}\left\{\left(\mathbf{h}_b^i\right)^H \mathbf{G}_c^{-1}\mathbf{h}_b^i\right\}}{\gamma_i}$. Gradient of the above term with respect to $\mathbf{G}_c^*$ yields $-G_R G_T \mathbf{G}_c^{-1} + \sum_{i=1}^{G_R G_T} \frac{1}{\gamma_i}\mathbf{G}_c^{-1}\mathbb{E}\left\{\mathbf{h}_b^i\left(\mathbf{h}_b^i\right)^H\right\}\mathbf{G}_c^{-1}$. Next, evaluating the gradient at $\gamma_i = \widehat{\gamma}_i^{(p)}$ and setting the resultant to zero gives the update of the correlation matrix $\mathbf{G}_c$ in the pth iteration as $\widehat{\mathbf{G}}_c^{(p)} = \frac{1}{G_R G_T}\sum_{i=1}^{G_R G_T} \frac{\mathbb{E}\left\{\mathbf{h}_b^i\left(\mathbf{h}_b^i\right)^H\right\}}{\widehat{\gamma}_i^{(p)}}$. Employing the *a posteriori* pdf of $\mathbf{h}_b$ parameterized by (28) for computing the expectation in the above expression yields the desired result.

APPENDIX B
PROOF OF *Theorem 1*

Using (9), the channel matrix $\mathbf{H}[k]$ for the kth subcarrier can be determined as $\mathbf{H}[k] = \sum_{l=0}^{L-1}\mathbf{H}_l\; e^{-j\frac{2\pi k}{K}l} = \mathbf{A}_R\mathbf{H}_D[k]\,\mathbf{A}_T^H$, where $\mathbf{H}_D[k] = \sum_{l=0}^{L-1}\mathbf{H}_{D,l}\; e^{-j\frac{2\pi k}{K}l}$. This implies that $\mathcal{R}(\mathbf{H}[k]) = \mathcal{C}(\mathbf{A}_T),\ \forall\ 0 \le k \le K-1$, where $\mathcal{R}(\cdot)$ denotes the row-space of a matrix. Furthermore, from (51), it can be seen that $\mathcal{C}(\mathbf{F}^{\rm opt}[k]) \subseteq \mathcal{R}(\mathbf{H}[k])$ when $N_s \le \mathrm{rank}(\mathbf{H}[k])$. Thus it can be readily seen that $\mathcal{C}(\mathbf{F}^{\rm opt}[k]) \subseteq \mathcal{C}(\mathbf{A}_T), \forall\ 0 \le k \le K-1$.

REFERENCES

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

[2] T. S. Rappaport, R. W. Heath Jr, R. C. Daniels, and J. N. Murdock, *Millimeter Wave Wireless Communications*. Pearson Education, 2014.

[3] L. Hanzo, M. Münster, B. Choi, and T. Keller, *OFDM and MC-CDMA for broadband multi-user communications, WLANs and broadcasting*. John Wiley & Sons, 2005.

[4] J. R. Fernández, N. G. 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.

[5] X. Zhang, A. F. Molisch, and S.-Y. Kung, "Variable-phase-shift-based RF-baseband codesign for MIMO antenna selection," *IEEE Transactions on Signal Processing*, vol. 53, no. 11, p. 4091, 2005.

[6] 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.

[7] J. Wang, "Beam codebook based beamforming protocol for multi-Gbps millimeter-wave WPAN systems," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 8, 2009.

[8] S. Hur, T. Kim, D. J. Love, J. V. Krogmeier, T. A. Thomas, A. Ghosh *et al.*, "Millimeter wave beamforming for wireless backhaul and access in small cell networks." *IEEE Trans. Communications*, vol. 61, no. 10, pp. 4391–4403, 2013.

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

[10] S. Srivastava, A. Mishra, A. Rajoriya, A. K. Jagannatham, and G. Ascheid, "Quasi-static and time-selective channel estimation for block-sparse millimeter wave hybrid MIMO systems: Sparse Bayesian learning (SBL) based approaches," *IEEE Transactions on Signal Processing*, vol. 67, no. 5, pp. 1251–1266, 2019.

[11] X. Lin, S. Wu, C. Jiang, L. Kuang, J. Yan, and L. Hanzo, "Estimation of broadband multiuser millimeter wave massive MIMO-OFDM channels by exploiting their sparse structure," *IEEE Transactions on Wireless Communications*, vol. 17, no. 6, pp. 3959–3973, 2018.

[12] 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.

[13] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on information theory*, vol. 53, no. 12, pp. 4655–4666, 2007.

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

[15] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM review*, vol. 43, no. 1, pp. 129–159, 2001.

[16] D. P. Wipf and B. D. Rao, "An empirical Bayesian strategy for solving the simultaneous sparse approximation problem," *IEEE Transactions on Signal Processing*, vol. 55, no. 7, pp. 3704–3716, 2007.

[17] Z. Zhang and B. D. Rao, "Sparse signal recovery with temporally correlated source vectors using sparse Bayesian learning," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 5, pp. 912–926, 2011.

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

[19] S. Srivastava, C. S. K. Patro, A. K. Jagannatham, and G. Sharma, "Sparse Bayesian learning (SBL)-based frequency-selective channel estimation for millimeter wave hybrid MIMO systems," in *2019 National Conference on Communications (NCC)*. IEEE, 2019, pp. 1–6.

[20] S. Srivastava and A. K. Jagannatham, "Sparse Bayesian learning-based Kalman filtering (SBL-KF) for group-sparse channel estimation in doubly selective mmWave hybrid MIMO systems," in *2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. IEEE, 2019, pp. 1–5.

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

[22] S. Srivastava, A. Mishra, A. K. Jagannatham, and G. Ascheid, "SBL-based hybrid precoder/combiner design for power and spectrally efficient millimeter wave MIMO systems," in *2020 International Conference on Signal Processing and Communications (SPCOM)*. IEEE, 2020, pp. 1–5.

[23] H. Zhang and F. Ding, "On the kronecker products and their applications," *Journal of Applied Mathematics*, vol. 2013, 2013.

[24] A. V. Oppenheim and R. W. Schafer, *Discrete-time signal processing*. Pearson Education, 2014.

[25] S. M. Kay, *Fundamentals of statistical signal processing*. Prentice Hall PTR, 1993.

[26] S. Srivastava, C. S. K. Patro, A. K. Jagannatham, and L. Hanzo, "Technical report: SBL-based group-sparse and online channel estimation for doubly-selective mmWave hybrid MIMO OFDM systems." IIT Kanpur, Tech. Rep. TR 1.2021, 2021. [Online]. Available: http://www.iitk.ac.in/mwn/documents/MWNLab_TR_SBL_MIMO_OFDM.pdf.

[27] H. L. Van Trees, *Optimum array processing: Part IV of detection, estimation, and modulation theory*. John Wiley & Sons, 2004.

[28] G. Li, Z. Wei, X. Zhang, and D. Yang, "Low overhead channel tracking strategy for millimetre wave multiple-input multiple-output system," *IET Communications*, vol. 11, no. 16, pp. 2544–2551, 2017.

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

[30] L. Meier, S. Van De Geer, and P. Bühlmann, "The group Lasso for logistic regression," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 70, no. 1, pp. 53–71, 2008.

[31] 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.

[32] A. Maleki, "Approximate message passing algorithms for compressed sensing," *A degree of Doctor of Philosophy, Stanford University*, 2011.

[33] Y. Chen, Y. Gu, and A. O. Hero, "Regularized least-mean-square algorithms," *arXiv preprint arXiv:1012.5066*, 2010.

Suraj Srivastava received the B.Tech. degree in electronics and communication engineering from Uttar Pradesh Technical University, India, in 2010, and M.Tech. degree in communication systems from Indian Institute of Technology (IIT) Roorkee, Roorkee, India, in 2012. 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 is currently working toward the Ph.D. degree with the Department of Electrical Engineering, IIT Kanpur, Kanpur, India. His research interests include applications of 5G mmWave MIMO wireless technology, Sparse Signal Processing, Distributed Signal Processing for Wireless Sensor Networks (WSNs) for IoT. He was awarded Qualcomm Innovation Fellowship (QInf)-2018 from Qualcomm.

Ch Suraj Kumar Patro received the B.Tech. degree in electronics and communication engineering from Biju Patnaik University of Technology, India, in 2015, and M.Tech. degree in electrical engineering from Indian Institute of Technology (IIT) Kanpur, Kanpur, India, in 2019. Since August 2019, he has been working in LTE modem systems team at Qualcomm India Private Limited, Hyderabad, India. He was awarded with academic excellence award at IIT Kanpur for the academic session 2017-2018. His research interests include wireless communication, signal processing, machine learning methods for channel estimation in 5G systems.

Aditya K. Jagannatham (S'04-M'05) received the bachelorâĂŹs degree from IIT Bombay and the M.S. and Ph.D. degrees from the University of California San Diego. From April '07 to May '09 he was employed as a senior wireless systems engineer at Qualcomm Inc., San Diego, California, where he was a part of the Qualcomm CDMA technologies (QCT) division. His research interests are in the area of next-generation wireless cellular and WiFi networks, with special emphasis on various 5G technologies such as massive MIMO, mmWave MIMO, FBMC, NOMA, Full Duplex and others. He was awarded the CAL(IT)2 fellowship at the University of California San Diego and the Upendra Patel Achievement Award at Qualcomm. He is currently a Professor in the Electrical Engineering department at IIT Kanpur, where he holds the Arun Kumar Chair Professorship, and is also associated with the BSNL-IITK Telecom Center of Excellence (BITCOE). He has been twice awarded the P.K. Kelkar Young Faculty Research Fellowship for excellence in research, the Qualcomm Innovation Fellowship (QInF) and the IIT Kanpur Excellence in Teaching Award.

Lajos Hanzo (http://www-mobile.ecs.soton.ac.uk, https://en.wikipedia.org/wiki/Lajos_Hanzo) (FIEEE'04, Fellow of the Royal Academy of Engineering F(REng), of the IET and of EURASIP), received his Master degree and Doctorate in 1976 and 1983, respectively from the Technical University (TU) of Budapest. He was also awarded the Doctor of Sciences (DSc) degree by the University of Southampton (2004) and Honorary Doctorates by the TU of Budapest (2009) and by the University of Edinburgh (2015). He is a Foreign Member of the Hungarian Academy of Sciences and a former Editor-in-Chief of the IEEE Press. He has served several terms as Governor of both IEEE ComSoc and of VTS. He has published 1900+ contributions at IEEE Xplore, 19 Wiley-IEEE Press books and has helped the fast-track career of 123 PhD students. He holds the Chair of Telecommunications and directs the research of Next-Generation Wireless at the University of Southampton, UK.