Bayesian Learning-Based Doubly-Selective Sparse Channel Estimation for Millimeter Wave Hybrid MIMO-FBMC-OQAM Systems

Suraj Srivastava, *Graduate Student Member, IEEE*, Prem Singh, *Student Member, IEEE*, Aditya K. Jagannatham, *Member, IEEE*, Abhay Karandikar, *Member, IEEE* and Lajos Hanzo, *Fellow, IEEE*

***Abstract*—We design and analyse filter bank multicarrier (FBMC) offset quadrature amplitude modulation (OQAM)-based millimeter wave (mmWave) hybrid multiple-input multiple-output (MIMO) systems. Furthermore, a novel channel estimation model is conceived for quasi-static mmWave hybrid MIMO-FBMC-OQAM (mmH-MFO) systems that reconfigures the radio-frequency (RF) circuitry during the transmission of zero symbols. Subsequently, a Bayesian learning (BL) technique is proposed for sparse channel estimation, which relies on multiple measurement vectors combined with selective subcarrier grouping for enhanced estimation. Additionally, an online BL based Kalman filter (OBL-KF) is designed for sparse channel tracking in doubly-selective mmH-MFO systems. Then the Bayesian Cramér-Rao lower bounds (BCRLBs) are derived for characterizing the performance of the proposed frequency-selective and doubly-selective channel estimation techniques. Finally, a limited feedback based algorithm relying on beamspace channel estimates is proposed for hybrid precoder/combiner design. The accuracy of our analytical results is confirmed by our simulation results.**

***Index Terms*—mmWave communication, hybrid MIMO architecture, filter bank multicarrier, Bayesian learning (BL), expectation maximization, Bayesian Cramér-Rao bound.**

I. Introduction

MILLIMETER wave (mmWave) technology that enables leveraging the huge bandwidth available in the 30 to 300 GHz band has emerged as a promising solution for next generation wireless systems. It has been recently employed in various standards such as the 5th generation New Radio (5G-NR), IEEE 802.11ad for wireless local area networks (WLAN) [1] and WirelessHD for personal area networks (WPAN) [2]. To overcome the hardware complexity of the conventional fully-digital mmWave multiple-input multiple-output (MIMO) implementation, the hybrid MIMO architecture has been shown to be eminently suitable for communication at mmWave frequencies [3], [4]. In such a mmWave hybrid MIMO system, digital precoding is performed in the baseband at the transmitter, followed by analog precoding for directional beamforming using a network of digitally controlled phase-shifters in the radio-frequency (RF) front-end. Recently, orthogonal frequency division multiplexing (OFDM) has also been proposed for hybrid mmWave MIMO systems in order to combat multipath fading in wideband mmWave wireless channels [5], [6]. However, the rectangular time-domain window used in OFDM for each subcarrier leads to significant out-of-band (OOB) radiation. Furthermore, OFDM systems are also susceptible to synchronization errors [7] due to the limited frequency localization of the rectangular time domain pulse, especially in high-velocity mobile scenarios where it is difficult to track the Doppler shifts [8]. Therefore, OFDM may not necessarily be well suited for all the use cases in future wireless communication networks. As an appealing design alternative, filter bank multicarrier (FBMC)-offset quadrature amplitude modulation (OQAM) employing banks of frequency-time (FT) localized filters for modulation and demodulation has emerged as a potential waveform candidate for future systems [9]. Its pulse shaping filters significantly reduce the OOB emission and readily meets the stringent synchronization requirement of these systems [10]. Thus, FBMC-OQAM is a viable competitor of OFDM in future mmWave hybrid MIMO systems. The next subsection presents a brief review of the recent contributions in this context.

A. Review of Existing Works

Pérez-Neira *et al.* [11] described the signal processing solutions incorporating the FBMC-OQAM waveform in MIMO schemes. Singh *et al.* [12], [13] derived an MMSE receiver for MIMO-FBMC-OQAM (MFO) systems and characterized them in the presence of perfect and imperfect channel state information (CSI), respectively. Li *et al.* [14] have harnessed Alamouti's encoder and decoder for highly frequency-selective FBMC-OQAM systems. Liu *et al.* [15] designed an optimal frequency division multiplexing (FDM)-structured preamble sequence for enhancing the channel estimation accuracy of MFO systems. Singh *et al.* [16] presented a two-step training

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

Abhay Karandikar is currently Director, Indian Institute of Technology Kanpur, Kanpur, 208016, India, as well as Professor, Department of Electrical Engineering, IIT Bombay, Mumbai, 400076, India (e-mail: karandi@iitk.ac.in).

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

L. Hanzo would like to acknowledge the financial support of the Engineering and Physical Sciences Research Council projects EP/N004558/1, EP/P034284/1, EP/P034284/1, EP/P003990/1 (COALESCE), of the Royal Society's Global Challenges Research Fund Grant as well as of the European Research Council's Advanced Fellow Grant QuantCom. A. K. Jagannatham would like to acknowledge the research supported by the Science and Engineering Research Board (SERB), Department of Science and Technology, Government of India, Space Technology Cell, IIT Kanpur, IIMA IDEA Telecom Centre of Excellence, Qualcomm Innovation Fellowship and Arun Kumar Chair Professorship.

approach for joint frequency-selective channel and carrier frequency offset (CFO) estimation of MFO systems. Singh *et al.* [17] have also investigated semi-blind, data-aided and training based channel estimation schemes designed for MFO systems. Wang [18] designed a sparse channel estimation scheme for smart city applications using compressive sensing for MFO systems. Chen *et al.* [19] analysed the performance of FBMC-OQAM based systems in the context of Internet of Things. As a further advance, the authors of [20]–[22] have described the benefits of FBMC-OQAM signaling over its OFDM counterpart in massive MIMO systems.

Let us now focus our attention on mmWave hybrid MIMO systems. The authors of [23], [24] have proposed various beam training strategies for CSI acquisition in such systems. The papers [25]–[27] described techniques exploiting the angular sparsity of the mmWave channel for CSI estimation. The authors of [5], [6] have addressed the problem of frequency-selective mmWave MIMO channel estimation. The orthogonal matching pursuit (OMP) and its variant termed as simultaneous OMP (SOMP) have been employed for sparse channel recovery in [5], [6] for mmWave MIMO-OFDM systems. However, to the best of our knowledge, and also seen from Table-I, none of the existing contributions have described the complete signal processing framework of mmWave hybrid MIMO FBMC-OQAM (mmH-MFO) systems. The focus of this paper is therefore to develop efficient channel estimation schemes for mmH-MFO systems followed by the associated precoder/combiner design. The main contributions of this treatise are presented next.

B. Contributions of the Present Work

- A novel transceiver architecture is developed for mmH-MFO systems, which intrinsically amalgamates the hybrid RF/ baseband precoding and combining operations with filter banks for modulation and demodulation. Initially, conventional LS and MMSE schemes are developed for quasi-static channel frequency response (CFR) estimation.
- A beamspace channel model is presented next that also reflects the sparse nature of the multipath scattering at mmWave frequencies, and thereby can be characterized only by a few set of parameters [28]–[31]. Subsequently, an interference approximation method (IAM) based CFR estimation model is developed that exploits the transmission of zero symbols used for ISI reduction during channel estimation in FBMC systems, for the sake of reconfiguring the analog RF circuitry of the mmH-MFO system.
- Next, sparse channel estimation techniques are conceived for mmH-MFO systems based on the OMP and Bayesian learning (BL) paradigms. Additionally, the *simultaneous-sparsity* across the subcarriers is also exploited via selective subcarrier grouping (SSG) in the context of a multiple measurement vector (MMV)-based BL (MBL-SSG) framework. This leads to a further improvement in the channel estimation performance. A doubly-selective sparse channel estimation model and the pertinent online BL-based Kalman filtering (OBL-KF) framework is invoked for channel tracking.
- The Bayesian Cramér-Rao lower bounds (BCRLBs) are derived for all the proposed estimation schemes. Finally, a limited CSI based scheme is designed for our RF precoders/combiners using the estimated beamspace channel. Our simulation results validate the various findings, and compare the performance of mmH-MFO and the conventional OFDM based mmWave hybrid MIMO systems.

C. Notation

The operations $\Re\{\cdot\}$ and $\Im\{\cdot\}$ represent real and imaginary parts, and $[\cdot]_N$ denotes the modulo-N operation. The operations $\mathbf{A}(m,:)$ and $\mathbf{A}(:,m)$ access the mth row and column of the matrix $\mathbf{A}$, respectively. The floor function $\lfloor u \rfloor$ denotes the largest integer that is no larger than u, and $\lceil u \rceil$ denotes the smallest integer that is no smaller than u. The operation $\mathbf{A}\otimes\mathbf{B}$ denotes Kronecker product of the matrices $\mathbf{A}$ and $\mathbf{B}$. For $p \geq 1$, $\|\cdot\|_p$ denotes the ℓ_p-norm of a vector [32], whereas $\|\mathbf{a}\|_0$ denotes the ℓ_0-norm of a vector $\mathbf{a}$ of size $N \times 1$, defined as $\|\mathbf{a}\|_0 = \sum_{i=1}^{N} \mathcal{I}\left(a_i \neq 0\right)$, where $\mathcal{I}(\cdot)$ represents the indicator function [33]. Note that $\|\mathbf{a}\|_0$ denotes the number of non-zero entries in the vector $\mathbf{a}$ [34].

II. mmWave Hybrid MIMO-FBMC System Model

Consider a single-user mmH-MFO baseband system employing N subcarriers for transmitting N_s data streams using N_t transmit antennas (TAs) and N_r receive antennas (RAs). As shown in Fig. 1, both the transmitter and receiver are assumed to have a hybrid MIMO architecture using N_{RF} RF chains (RFCs), where $N_{RF} < \min(N_t, N_r)$ [4]. The matrices $\mathbf{F}_{\text{BB},m} \in \mathbb{C}^{N_{RF}\times N_s}$ at the transmitter and $\mathbf{W}_{\text{BB},m} \in \mathbb{C}^{N_s \times N_{RF}}$ at the receiver for the mth subcarrier, $0 \leq m \leq N-1$, represent a bank of N baseband precoders and combiners, respectively. Note that similar to [5], [6], the baseband precoders and combiners are different for each subcarrier, whereas the RF precoders and combiners, represented by the matrices $\mathbf{F}_{\text{RF}} \in \mathbb{C}^{N_t \times N_{RF}}$ and $\mathbf{W}_{\text{RF}} \in \mathbb{C}^{N_{RF}\times N_r}$ respectively, are identical across all the subcarriers, since they are applied to the time-domain FBMC signal. As described in [5], the RF precoder and combiner are implemented using a digitally controlled network of unit-magnitude phase shifters.

Let $c^i_{m,n}$ denote the QAM symbol in the ith data stream on subcarrier m at time instant n. The function $\Phi(c^i_{m,n})$ separates the real and imaginary parts of the complex QAM symbol $c^i_{m,n}$ to extract real OQAM symbols $d^i_{m,2n}$ and $d^i_{m,2n+1}$ as per the procedure described in [35, Eq. (2), (3)]. Let T_s represent the duration of the QAM symbol $c^i_{m,n}$ with $\frac{T_s}{2}$ denoting the duration of each of the component OQAM symbols $d^i_{m,2n}$ and $d^i_{m,2n+1}$. The real and imaginary parts of the QAM symbol are assumed to be spatially and temporally independent and identically distributed (i.i.d.) with power P_d, i.e., $\mathbb{E}\left[d^i_{m,n}\left(d^i_{m,n}\right)^*\right] = P_d$, from which it follows that $\mathbb{E}\left[c^i_{m,n}\left(c^i_{m,n}\right)^*\right] = 2P_d$. Let the symbol vector $\mathbf{d}_{m,n} \in \mathbb{C}^{N_s\times 1}$ be defined as

$$\mathbf{d}_{m,n} = \left[d^1_{m,n}, d^2_{m,n}, \cdots, d^{N_s}_{m,n}\right]^T. \tag{1}$$

TABLE I: Summary of literature survey on mmWave hybrid MIMO and MIMO-FBMC-OQAM systems

	[16], [17], [36] [19]	[18]	[39]	[26]	[5]	[6]	**Proposed**
OFDM	×	×	×	×	✓	✓	×
FBMC	✓	✓	×	×	×	×	✓
mmWave hybrid MIMO	×	×	✓	✓	✓	✓	✓
sub-6 GHz MIMO	✓	✓	×	×	×	×	×
Spatial channel model	×	×	✓	✓	✓	✓	✓
Sparsity	×	✓	✓	✓	✓	✓	✓
Simultaneous-sparsity	×	×	×	✓	✓	×	✓
Doubly-selective channel	×	×	×	×	×	×	✓
Online channel estimation	×	×	×	✓	×	×	✓
CRLB bounds	×	×	×	✓	✓	×	✓
Limited CSI feedback	×	×	×	×	×	×	✓

The vector $\tilde{\mathbf{d}}_{m,n} \in \mathbb{C}^{N_{RF}\times 1}$ denotes the baseband precoded vector generated as $\tilde{\mathbf{d}}_{m,n} = \mathbf{F}_{\text{BB},m}\mathbf{d}_{m,n}$. Let $\tilde{d}^q_{m,n}$ denote the qth element of $\tilde{\mathbf{d}}_{m,n}$, which is fed into the qth transmit FBMC processing block, $1 \le q \le N_{RF}$, as shown in Fig. 2(a). Next, the baseband signal $s^q[k]$ at the output of qth transmit RFC can be obtained as [36]

$$s^q[k] = \sum_{m=0}^{N-1}\sum_{n\in\mathbb{Z}} \tilde{d}^q_{m,n}\chi_{m,n}[k], \tag{2}$$

where k denotes the sample index corresponding to the sampling interval T_s/N and the FBMC basis function $\chi_{m,n}[k]$ is defined as

$$\chi_{m,n}[k] \quad = \quad p[k-nN/2]e^{j2\pi mk/N}e^{j\phi_{m,n}}. \tag{3}$$

The symmetric real-valued pulse $p[k]$ of length L_p represents the impulse response of the prototype filter of the FBMC system. The phase factor $\phi_{m,n}$ is defined as $\frac{\pi}{2}(m+n)-\pi mn$. The basis functions $\chi_{m,n}[k]$ are required to satisfy the following real field orthogonality condition

$$\Re\Bigg\{\underbrace{\sum_{k=-\infty}^{+\infty}\chi_{m,n}[k]\chi^*_{\bar{m},\bar{n}}[k]}_{\xi^{\bar{m},\bar{n}}_{m,n}}\Bigg\} = \delta_{m,\bar{m}}\delta_{n,\bar{n}}, \tag{4}$$

where $\delta_{m,\bar{m}}$ denotes the Kronecker delta with $\delta_{m,\bar{m}} = 1$ if $m = \bar{m}$ and zero otherwise. Thus, we have $\xi^{\bar{m},\bar{n}}_{m,n} = 1$ if $(m,n) = (\bar{m},\bar{n})$, and $\xi^{\bar{m},\bar{n}}_{m,n} = j\langle\xi\rangle^{\bar{m},\bar{n}}_{m,n}$ if $(m,n) \neq (\bar{m},\bar{n})$, where $\langle\xi\rangle^{\bar{m},\bar{n}}_{m,n}$ denotes the imaginary part of $\xi^{\bar{m},\bar{n}}_{m,n}$ [37]. Let $\mathbf{F}_{\text{RF}}(t,q)$ be the (t,q)th element of the RF precoder matrix $\mathbf{F}_{\text{RF}}$. Then, the signal $\tilde{s}^t[k]$ destined for the tth TA is formulated as:

$$\tilde{s}^t[k] = \sum_{q=1}^{N_{RF}} \mathbf{F}_{\text{RF}}(t,q)s^q[k]. \tag{5}$$

Let $h^{r,t}[l]$, $0 \le l \le L_h-1$, denote the lth tap of the channel filter between the tth TA and the rth RA. The signal received by the mmH-MFO system is given as

$$y^r[k] = \sum_{t=1}^{N_t}\Big(\tilde{s}^t[k] * h^{r,t}[k]\Big) + \eta^r[k], \tag{6}$$

where the noise samples $\eta^r[k]$ are assumed to be i.i.d. as $\mathcal{CN}(0,\sigma^2_\eta)$. Let $\mathbf{W}_{\text{RF}}(j,r)$ denote the (j,r)th element of the RF combiner matrix $\mathbf{W}_{\text{RF}}$. The received signal $\tilde{y}^j[k]$ at the output of the jth RFC can be written as

$$\tilde{y}^j[k] = \sum_{r=1}^{N_r}\mathbf{W}_{\text{RF}}(j,r)y^r[k]. \tag{7}$$

Substituting $y^r[k]$ from (6) and in turn using $\tilde{s}^t[k]$ from (5), the expression for $\tilde{y}^j[k]$ above can be expanded as

$$\tilde{y}^j[k] = \sum_{r=1}^{N_r}\sum_{t=1}^{N_t}\sum_{q=1}^{N_{RF}}\mathbf{W}_{\text{RF}}(j,r)\left(\sum_{l=0}^{L_h-1}h^{r,t}[l]s^q[k-l]\right) \times \mathbf{F}_{\text{RF}}(t,q) + \tilde{\eta}^j[k], \tag{8}$$

where the noise at the output of the jth RFC is $\tilde{\eta}^j[k] = \sum_{r=1}^{N_r}\eta^r[k]\mathbf{W}_{\text{RF}}(j,r)$. Since the RF combiner matrix is comprised of unit-magnitude phase elements, the noise $\tilde{\eta}^j[k]$ is distributed as $\mathcal{CN}(0,N_r\sigma^2_\eta)$. Note that due to the RF combining, $\tilde{\eta}^j[k]$ is spatially-correlated across the RFC index j. The signal $\tilde{y}^j[k]$ is subsequently processed by the jth receive FBMC block as shown in Fig. 2(b), wherein $\tilde{y}^j[k]$ is matched with the FBMC basis function $\chi_{m,n}[k]$ to obtain the demodulated signal $\tilde{y}^j_{\bar{m},\bar{n}}$ as [36]

$$\tilde{y}^j_{\bar{m},\bar{n}} = \sum_{k=-\infty}^{+\infty}\tilde{y}^j[k]\chi^*_{\bar{m},\bar{n}}[k]. \tag{9}$$

Substituting the expressions for $s^q[k]$, $\chi_{\bar{m},\bar{n}}[k]$ and $\tilde{y}^j[k]$ from (2), (3) and (8) respectively in (9), the expression for $\tilde{y}^j_{\bar{m},\bar{n}}$ above can be expanded as

$$\tilde{y}^j_{\bar{m},\bar{n}} = \sum_{r=1}^{N_r}\sum_{t=1}^{N_t}\sum_{q=1}^{N_{RF}}\mathbf{W}_{\text{RF}}(j,r)\mathbf{F}_{\text{RF}}(t,q)\sum_{m=0}^{N-1}\sum_{n\in\mathbb{Z}}\tilde{d}^q_{m,n} \times \sum_{l=0}^{L_h-1}h^{r,t}[l]e^{-j2\pi ml/N}\sum_{k}p[k-l-nN/2]p[k-\bar{n}N/2] \times e^{j2\pi(m-\bar{m})k/N}e^{j(\phi_{m,n}-\phi_{\bar{m},\bar{n}})} + \tilde{\eta}^j_{\bar{m},\bar{n}}, \tag{10}$$

where the demodulated noise $\tilde{\eta}^j_{\bar{m},\bar{n}} = \sum_{k=-\infty}^{+\infty}\tilde{\eta}^j[k]\ \chi^*_{\bar{m},\bar{n}}[k]$ is distributed as $\mathcal{CN}(0,N_r\sigma^2_\eta)$ owing to the fact that the FBMC basis function $\chi_{m,n}[k]$ has unit-energy. It is worth noting that

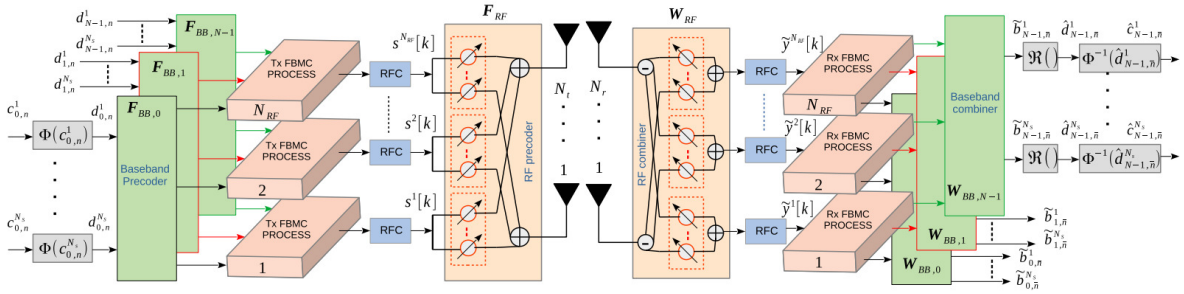


Fig. 1: Architecture of mmWave hybrid MIMO-FBMC transmitter and receiver.

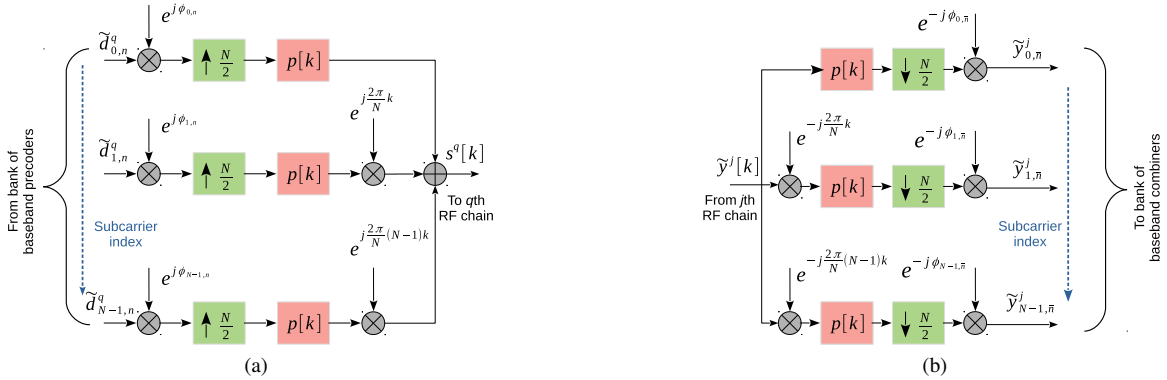


Fig. 2: Block diagram of FBMC (a) transmit processing for input of the qth RFC; (b) receiver processing for output of the jth RFC.

due to the real field orthogonality in FBMC systems, the noise $\tilde{\eta}^j_{\bar{m},\bar{n}}$ across the FT indices $(\bar{m},\bar{n})$, strictly speaking, is correlated. However, this correlation is insignificant due to the sharp nature of the pulse shaping filters in these systems [37], [38] and is neglected to render the analysis tractable. Since the channel's delay spread at mmWave frequencies is typically small in comparison to the symbol duration [29], which also implies that $p[k-l-nN/2] \approx p[k-nN/2]$ for $l \in [0, L_h]$, the expression above in (10) can be simplified to:

$$\tilde{y}^j_{\bar{m},\bar{n}} = \sum_{r=1}^{N_r}\sum_{t=1}^{N_t}\sum_{q=1}^{N_{RF}} \mathbf{W}_{\text{RF}}(j,r)\mathbf{F}_{\text{RF}}(t,q) \times \sum_{m=0}^{N-1}\sum_{n\in\mathbb{Z}} \tilde{d}^q_{m,n} H^{r,t}_m \xi^{\bar{m},\bar{n}}_{m,n} + \tilde{\eta}^j_{\bar{m},\bar{n}}, \tag{11}$$

where the quantity $H^{r,t}_m = \sum_{l=0}^{L_h-1} h^{r,t}[l] e^{-j2\pi ml/N}$ denotes the CFR for the mth subcarrier between the tth TA and the rth RA. One can now separate the term $\sum_{m=0}^{N-1}\sum_{n\in\mathbb{Z}} \tilde{d}^q_{m,n} H^{r,t}_m \xi^{\bar{m},\bar{n}}_{m,n}$ into desired and interfering components as

$$\sum_{m=0}^{N-1}\sum_{n\in\mathbb{Z}} \tilde{d}^q_{m,n} H^{r,t}_m \xi^{\bar{m},\bar{n}}_{m,n} = H^{r,t}_{\bar{m}} \tilde{d}^q_{\bar{m},\bar{n}} + j \sum_{(m,n)\neq(\bar{m},\bar{n})} \tilde{d}^q_{m,n} H^{r,t}_m \langle\xi\rangle^{\bar{m},\bar{n}}_{m,n}, \tag{12}$$

where the fact that $\xi^{\bar{m},\bar{n}}_{\bar{m},\bar{n}} = 1$ has been exploited in the above simplification. The latter term denotes the intrinsic interference that is characteristic of FBMC systems. For a FT localized filter $p[k]$, a significant portion of the interference can be attributed to the first-order neighborhood of the FT point $(\bar{m},\bar{n})$, denoted by $\Omega_{\bar{m},\bar{n}} = \{(\bar{m}\pm1,\bar{n}\pm1),(\bar{m},\bar{n}\pm1),(\bar{m}\pm1,\bar{n})\}$. For the subcarriers m in this neighbourhood, the CFRs $H^{r,t}_m$ can be well approximated by $H^{r,t}_{\bar{m}}$. Thus, the expression in (12) above can be simplified as $H^{r,t}_{\bar{m}} \tilde{b}^q_{\bar{m},\bar{n}}$, where the virtual symbol $\tilde{b}^q_{\bar{m},\bar{n}}$ obeys $\tilde{b}^q_{\bar{m},\bar{n}} = \tilde{d}^q_{\bar{m},\bar{n}} + j\tilde{I}^q_{\bar{m},\bar{n}}$. The intrinsic interference component $\tilde{I}^q_{\bar{m},\bar{n}}$ is expressed as

$$\tilde{I}^q_{\bar{m},\bar{n}} = \sum_{(m,n)\in\Omega_{\bar{m},\bar{n}}} \tilde{d}^q_{m,n} \langle\xi\rangle^{\bar{m},\bar{n}}_{m,n}. \tag{13}$$

The term $\tilde{I}^q_{\bar{m},\bar{n}}$ consists of the ISI and inter-carrier interference (ICI) imposed by the adjacent FT symbols around the desired symbol $\tilde{d}^q_{\bar{m},\bar{n}}$. This is in contrast to classic OFDM systems, where the ISI and ICI are suppressed using the cyclic-prefix (CP) and the orthogonality among the subcarriers, respectively. Using the above simplification, the expression for $\tilde{y}^j_{\bar{m},\bar{n}}$ in (11) can be further reduced to

$$\tilde{y}^j_{\bar{m},\bar{n}} = \sum_{r=1}^{N_r}\sum_{t=1}^{N_t}\sum_{q=1}^{N_{RF}} \mathbf{W}_{\text{RF}}(j,r) H^{r,t}_{\bar{m}} \mathbf{F}_{\text{RF}}(t,q) \tilde{b}^q_{\bar{m},\bar{n}} + \tilde{\eta}^j_{\bar{m},\bar{n}}.$$

Let $\tilde{\mathbf{y}}_{\bar{m},\bar{n}} = [\tilde{y}^1_{\bar{m},\bar{n}}, \tilde{y}^2_{\bar{m},\bar{n}}, \ldots, \tilde{y}^{N_{RF}}_{\bar{m},\bar{n}}]^T \in \mathbb{C}^{N_{RF}\times 1}$ denote the vector of received symbols at the output of the N_{RF} receive RF chains at the FT index $(\bar{m},\bar{n})$. This can be determined as

$$\tilde{\mathbf{y}}_{\bar{m},\bar{n}} = \mathbf{W}_{\text{RF}} \mathbf{H}_{\bar{m}} \mathbf{F}_{\text{RF}} \tilde{\mathbf{b}}_{\bar{m},\bar{n}} + \tilde{\boldsymbol{\eta}}_{\bar{m},\bar{n}}, \tag{14}$$

where $\tilde{\boldsymbol{\eta}}_{\bar{m},\bar{n}} = [\tilde{\eta}^1_{\bar{m},\bar{n}}, \tilde{\eta}^2_{\bar{m},\bar{n}}, \ldots, \tilde{\eta}^{N_{RF}}_{\bar{m},\bar{n}}]^T \in \mathbb{C}^{N_{RF}\times 1}$ is the corresponding noise vector so that $\mathbb{E}[\tilde{\boldsymbol{\eta}}_{\bar{m},\bar{n}} \tilde{\boldsymbol{\eta}}^H_{\bar{m},\bar{n}}] = \sigma^2_\eta \mathbf{W}_{\text{RF}} \mathbf{W}^H_{\text{RF}}$.

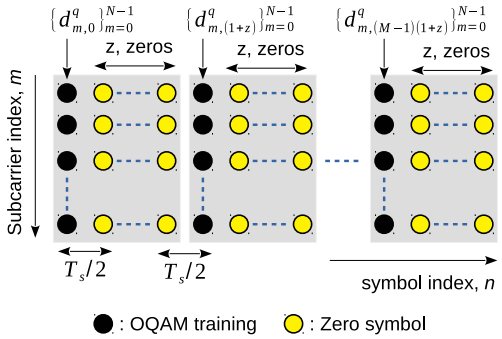


Fig. 3: Training frame structure for the qth RFC.

The vector $\tilde{\mathbf{b}}_{\bar{m},\bar{n}} = [\tilde{b}^1_{\bar{m},\bar{n}}, \tilde{b}^2_{\bar{m},\bar{n}}, \ldots, \tilde{b}^{N_{RF}}_{\bar{m},\bar{n}}]^T \in \mathbb{C}^{N_{RF}\times 1}$ is comprised of the precoded virtual symbols. Finally, the matrix $\mathbf{H}_{\bar{m}} \in \mathbb{C}^{N_r\times N_t}$ denotes the mmWave MIMO CFR matrix for the $\bar{m}$th subcarrier with its (r,t)th element given by $H^{r,t}_{\bar{m}}$.

III. Conventional Channel Estimation for mmH-MFO Systems

Let us now arrange for the qth RFC to transmit OQAM training symbols on each subcarrier according to the frame structure shown in Fig. 3, for the purpose of channel estimation. Note that similar to the data symbols, the training symbols are also zero-mean i.i.d. random variables, but known at the receiver. Each OQAM training symbol is followed by z zero symbols that are inserted for meeting the following requirements. Firstly, the zero symbols curb the ISI arising due to the overlapping nature of the FBMC pulse-shaping filters [37], [38]. Secondly, this allows reconfiguration of the RF precoder and combiner matrices $\mathbf{F}_{\text{RF}}$ and $\mathbf{W}_{\text{RF}}$ without encountering any loss of training data [6]. In contrast to OFDM, the adjacent time-domain FBMC symbols interfere due to the overlapping nature of the associated prototype filter [37]. To suppress this interference, similar to several authoritative references, such as [17], [37], [38], zero symbols have been inserted between the adjacent training symbols, as shown in Fig. 3. However, it is important to note that for $z = 1$, the duration of an OQAM training symbol followed by a zero, prior to FBMC modulation, is equal to T_s, which equals the duration of a QAM symbol. Thus, for $z = 1$, there is no additional overhead in an mmH-MFO system and it is typically sufficient to mitigate the ISI to an acceptable level [38]. The performance significantly improves further for $z = 2$ and achieves the ideal BCRLB performance, as seen in the various results in Section VII. Thus, although strictly speaking z can be arbitrary, in practice, typically $z = 1$ (no additional overhead) or $z = 2$ (marginal additional overhead) is sufficient to achieve the best performance. As shown in Fig. 3, the training symbols are located at the indices $n = i(1 + z)$, for $0 \le i \le M-1$. Let the matrices $\mathbf{F}_{\text{TR},i(1+z)}$ and $\mathbf{W}_{\text{TR},i(1+z)}$ represent the training RF precoder and combiner matrices, respectively, during the transmission of the $i(1+z)$th training symbol. The training vector $\mathbf{y}_{\bar{m},i(1+z)}$ received at these indices is obtained using (14) as

$$\mathbf{y}_{\bar{m},i(1+z)} = \mathbf{W}_{\text{TR},i(1+z)}\mathbf{H}_{\bar{m}}\mathbf{F}_{\text{TR},i(1+z)}\mathbf{b}_{\bar{m},i(1+z)} + \tilde{\boldsymbol{\eta}}_{\bar{m},i(1+z)}, \tag{15}$$

where $\mathbf{b}_{\bar{m},i(1+z)} = [b^1_{\bar{m},i(1+z)}, \ldots, b^{N_{RF}}_{\bar{m},i(1+z)}]^T \in \mathbb{C}^{N_{RF}\times 1}$ is the virtual training vector at the FT index $(\bar{m}, i(1+z))$ and $\tilde{\boldsymbol{\eta}}_{\bar{m},i(1+z)}$ is the noise vector, whose covariance matrix $\mathbf{R}_{\eta,i}$ obeys $\mathbf{R}_{\eta,i} = \sigma^2_\eta\left[\mathbf{W}_{\text{TR},i(1+z)}\mathbf{W}^H_{\text{TR},i(1+z)}\right]$. The qth element of $\mathbf{b}_{\bar{m},i(1+z)}$ is given by $b^q_{\bar{m},i(1+z)} = d^q_{\bar{m},i(1+z)} + jI^q_{\bar{m},i(1+z)}$, where the intrinsic interference $I^q_{\bar{m},i(1+z)}$ is [17]

$$I^q_{\bar{m},i(1+z)} = \sum_{m\neq\bar{m}} d^q_{m,i(1+z)}\Im\Bigg\{\sum_{l=-\infty}^{+\infty} p^2[l]e^{j(\phi_{m,0}-\phi_{\bar{m},0})} \times e^{j2\pi(m-\bar{m})l/N}\Bigg\} = \sum_{m\neq\bar{m}} d^q_{m,i(1+z)}\langle\xi\rangle^{\bar{m},0}_{m,0}. \tag{16}$$

Since the training symbols $d^q_{\bar{m},i(1+z)}$ are zero mean i.i.d. random variables, it follows that the interference $I^q_{\bar{m},i(1+z)}$ is spatially uncorrelated across the RF chains q. Furthermore, assuming z being sufficiently large, they are also temporally uncorrelated across the training symbol index i. Hence, the variance of the virtual training symbol $b^q_{\bar{m},i(1+z)}$ can be evaluated as

$$\mathbb{E}[|b^q_{\bar{m},i(1+z)}|^2] = \mathbb{E}[|d^q_{\bar{m},i(1+z)}|^2] + \mathbb{E}[|I^q_{\bar{m},i(1+z)}|^2] = \sigma^2_b. \tag{17}$$

Using the properties of the $\text{vec}(\cdot)$ operator and matrix Kronecker product $\otimes$, (15) can be recast as

$$\mathbf{y}_{\bar{m},i(1+z)} = \boldsymbol{\Psi}_{\bar{m},i(1+z)}\mathbf{h}_{\bar{m}} + \tilde{\boldsymbol{\eta}}_{\bar{m},i(1+z)}, \tag{18}$$

where $\boldsymbol{\Psi}_{\bar{m},i(1+z)} = \left[\left(\mathbf{b}^T_{\bar{m},i(1+z)}\mathbf{F}^T_{\text{TR},i(1+z)}\right)\otimes\left(\mathbf{W}_{\text{TR},i(1+z)}\right)\right] \in \mathbb{C}^{N_{RF}\times N_rN_t}$ and $\mathbf{h}_{\bar{m}} = \text{vec}\left(\mathbf{H}_{\bar{m}}\right)$ is the mmWave CFR vector of size $N_rN_t \times 1$. Upon concatenating the observations $\mathbf{y}_{\bar{m},i(1+z)}$ for all the training vectors with indices, $0 \le i \le M-1$, as $\mathbf{y}_{\bar{m}} = [\mathbf{y}^T_{\bar{m},0}, \mathbf{y}^T_{\bar{m},(1+z)}, \ldots, \mathbf{y}^T_{\bar{m},(M-1)(1+z)}]^T \in \mathbb{C}^{N_{RF}M\times 1}$, the equivalent channel estimation model can be expressed as

$$\mathbf{y}_{\bar{m}} = \boldsymbol{\Psi}_{\bar{m}}\mathbf{h}_{\bar{m}} + \tilde{\boldsymbol{\eta}}_{\bar{m}}, \tag{19}$$

where $\boldsymbol{\Psi}_{\bar{m}} = [\boldsymbol{\Psi}^T_{\bar{m},0}, \boldsymbol{\Psi}^T_{\bar{m},(1+z)}, \ldots, \boldsymbol{\Psi}^T_{\bar{m},(M-1)(1+z)}]^T \in \mathbb{C}^{N_{RF}M\times N_rN_t}$. The noise vector $\tilde{\boldsymbol{\eta}}_{\bar{m}} \in \mathbb{C}^{N_{RF}M\times 1}$ obeys $\tilde{\boldsymbol{\eta}}_{\bar{m}} = [\tilde{\boldsymbol{\eta}}^T_{\bar{m},0}, \tilde{\boldsymbol{\eta}}^T_{\bar{m},(1+z)}, \ldots, \tilde{\boldsymbol{\eta}}^T_{\bar{m},(M-1)(1+z)}]^T$, with its covariance matrix given by $\mathbf{R}_\eta \triangleq \mathbb{E}\left[\tilde{\boldsymbol{\eta}}_{\bar{m}}\tilde{\boldsymbol{\eta}}^H_{\bar{m}}\right] = \text{blkdiag}\left(\mathbf{R}_{\eta,0}, \mathbf{R}_{\eta,1}, \cdots, \mathbf{R}_{\eta,M-1}\right)$, which is a block-diagonal matrix with $\mathbf{R}_{\eta,i}, 0 \le i \le M-1$, on its principal diagonal. The conventional LS estimate of the CFR vector $\mathbf{h}_{\bar{m}}$ at the $\bar{m}$th subcarrier can now be determined in a straightforward fashion as [39]

$$\hat{\mathbf{h}}^{\text{LS}}_{\bar{m}} = \boldsymbol{\Psi}^\dagger_{\bar{m}}\mathbf{y}_{\bar{m}}, \tag{20}$$

where $\boldsymbol{\Psi}^\dagger_{\bar{m}} = (\boldsymbol{\Psi}^H_{\bar{m}}\boldsymbol{\Psi}_{\bar{m}})^{-1}\boldsymbol{\Psi}^H_{\bar{m}}$ represents the pseudo-inverse of the matrix $\boldsymbol{\Psi}_{\bar{m}}$. Next, the conventional MMSE estimate, which minimizes the MSE, can be determined as

$$\hat{\mathbf{h}}^{\text{MMSE}}_{\bar{m}} = \left(\mathbf{R}^{-1}_{\mathbf{h}_{\bar{m}}} + \boldsymbol{\Psi}^H_{\bar{m}}\mathbf{R}^{-1}_\eta\boldsymbol{\Psi}_{\bar{m}}\right)^{-1}\boldsymbol{\Psi}^H_{\bar{m}}\mathbf{R}^{-1}_\eta\mathbf{y}_{\bar{m}}, \tag{21}$$

where $\mathbf{R}_{\mathbf{h}_{\bar{m}}} = \mathbb{E}\left[\mathbf{h}_{\bar{m}}\mathbf{h}_{\bar{m}}^H\right] \in \mathbb{C}^{N_rN_t \times N_rN_t}$. The conventional LS and LMMSE schemes described above suffer from the drawback that they require an over-determined system for providing reliable estimates, which necessitates the transmission of at least $M = \lceil \frac{N_tN_r}{N_{RF}} \rceil$ training symbols on each of the subcarriers. Given that $N_{RF} < \min(N_t, N_r)$ for a typical mmWave hybrid MIMO system, this leads to a potentially excessive training overhead, which implies that the conventional channel estimators are spectrally inefficient. However, by exploiting the frequency-domain correlation among subcarriers [40], [41], together with the delay-domain sparsity of the mmWave MIMO channel [42], [43], the pilot overhead may be further reduced. Future work may extend the proposed frameworks for incorporating these aspects. On the other hand, it is important to note that these conventional schemes fail to exploit the spatially-sparse nature of the multipath components in a typical mmWave MIMO channel, which is a unique characteristic of the wireless channel in the mmWave regime [28]–[31]. Leveraging this aspect can lead to a significant improvement in the quality of the channel estimate obtained. Hence, the next section develops a mmWave-specific clustered MIMO channel model followed by novel schemes that also exploit the sparsity for channel recovery.

IV. Sparse mmWave MIMO FBMC Channel Estimation Schemes

Let $\mathbf{H}[l] \in \mathbb{C}^{N_r \times N_t}$ denote the lth tap of the frequency-selective mmWave MIMO channel where the (r,t)th element is given as $h^{r,t}[l]$. Using the narrowband clustered channel model described in [3], [4], [44] for a mmWave MIMO system, the matrix $\mathbf{H}[l]$ can be expressed as

$$\mathbf{H}[l] = \sqrt{\frac{N_tN_r}{N_{cl}N_{ray}}} \sum_{i=1}^{N_{cl}} \sum_{j=1}^{N_{ray}} \alpha_{ij}[l]\mathbf{a}_R(\theta_{ij}^a)\mathbf{a}_T^H(\theta_{ij}^d), \quad (22)$$

where the quantity N_{cl} denotes the total number of clusters, with each cluster contributing N_{ray} spatial multipath components. The angles θ_{ij}^a and θ_{ij}^d respectively represent the angles-of-arrival (AoAs) and angles-of-departure (AoDs) associated with the jth ray in the ith cluster. The quantity $\alpha_{ij}[l]$ represents the complex gain of the lth channel tap. The vectors $\mathbf{a}_R(\theta_{ij}^a) \in \mathbb{C}^{N_r \times 1}$ and $\mathbf{a}_T(\theta_{ij}^d) \in \mathbb{C}^{N_t \times 1}$, which respectively denote the receive and transmit array response vectors of the jth ray in the ith cluster, are defined as

$$\mathbf{a}_R(\theta_{ij}^a) = \frac{1}{\sqrt{N_r}}\left[1, e^{-j\tilde{d}_r}, \ldots, e^{-j\tilde{d}_r(N_r-1)}\right]^T,$$
$$\mathbf{a}_T(\theta_{ij}^d) = \frac{1}{\sqrt{N_t}}\left[1, e^{-j\tilde{d}_t}, \ldots, e^{-j\tilde{d}_t(N_t-1)}\right]^T,$$

where $\tilde{d}_r = \frac{2\pi}{\lambda} d_r \cos(\theta_{ij}^a)$ and $\tilde{d}_t = \frac{2\pi}{\lambda} d_t \cos(\theta_{ij}^d)$. The quantities λ, d_r and d_t denote the carrier wavelength, antenna spacings of the receive and transmit arrays, respectively. Consider a partition of the feasible AoD and AoA spaces using grids comprised of G_t, G_r angles, respectively, where $G_t, G_r \geq \max\{N_t, N_r\}$ angles. The set of spatial angles in the grids Θ_t and Θ_r at the transmitter and receiver, respectively, is chosen for ensuring that the following conditions are satisfied [45]

$$\Theta_t = \left\{\phi_g^t : \cos(\phi_g^t) = \frac{2}{G_t}(g-1) - 1,\ 1 \leq g \leq G_t\right\},$$
$$\Theta_r = \left\{\phi_g^r : \cos(\phi_g^r) = \frac{2}{G_r}(g-1) - 1,\ 1 \leq g \leq G_r\right\}.$$

The quantized transmit and receive array response matrices $\mathbf{A}_T(\Theta_t) \in \mathbb{C}^{N_t \times G_t}$ and $\mathbf{A}_R(\Theta_r) \in \mathbb{C}^{N_r \times G_r}$ corresponding to the grids Θ_t and Θ_r, respectively, are constructed as $\mathbf{A}_T(\Theta_t) = [\mathbf{a}_T(\phi_1^t), \cdots, \mathbf{a}_T(\phi_{G_t}^t)]$ and $\mathbf{A}_R(\Theta^r) = [\mathbf{a}_R(\phi_1^r), \cdots, \mathbf{a}_R(\phi_{G_r}^r)]$. When the quantization intervals of the AoA/ AoD spaces are suitably fine, the beamspace representation of the channel matrix $\mathbf{H}[l]$ can be obtained as

$$\mathbf{H}[l] = \mathbf{A}_R(\Theta_r)\mathbf{H}_b[l]\mathbf{A}_T^H(\Theta_t), \quad (23)$$

where $\mathbf{H}_b[l] \in \mathbb{C}^{G_r \times G_t}$ denotes the equivalent beamspace channel matrix corresponding to $\mathbf{H}[l]$. Owing to the highly directional nature of signal propagation at mmWave frequencies coupled with the reduced scattering effects, typically only a few active spatial multipath components exist [28]–[31]. Thus, the resultant beamspace channel matrix $\mathbf{H}_b[l]$ is *sparse* in nature. Note that the different rays in a given cluster associated with closely spaced AoAs/ AoDs can be assumed to be mapped to a single beamspace component. Furthermore, the non-zero coefficients corresponding to different clusters can be assumed to be independent due to the wide angular spread. The beamspace representation of the mmWave MIMO-FBMC CFR matrix can now be obtained as

$$\mathbf{H}_{\bar{m}} = \sum_{l=0}^{L_h-1} \mathbf{H}[l]e^{-j\frac{2\pi\bar{m}}{N}l}$$
$$= \mathbf{A}_R(\Theta_r) \underbrace{\left(\sum_{l=0}^{L_h-1} \mathbf{H}_b[l]e^{-j\frac{2\pi\bar{m}}{N}l}\right)}_{\mathbf{H}_{b,\bar{m}}} \mathbf{A}_T^H(\Theta_t), \quad (24)$$

where $\mathbf{H}_{b,\bar{m}} \in \mathbb{C}^{G_r \times G_t}$ represents the beamspace channel matrix corresponding to $\mathbf{H}_{\bar{m}}$. It follows from (22) and (24) that the beamspace channel matrices $\mathbf{H}_b[l],\ 0 \leq l \leq L_h - 1$, and $\mathbf{H}_{b,\bar{m}},\ 0 \leq \bar{m} \leq N-1$, share a common sparsity profile, i.e., the locations of their non-zero entries coincide. One can now express the mmWave CFR vector $\mathbf{h}_{\bar{m}}$ as

$$\mathbf{h}_{\bar{m}} \triangleq \mathrm{vec}(\mathbf{H}_{\bar{m}}) = \left(\mathbf{A}_T^*(\Theta_t) \otimes \mathbf{A}_R(\Theta_r)\right)\mathbf{h}_{b,\bar{m}}, \quad (25)$$

where $\mathbf{h}_{b,\bar{m}} \triangleq \mathrm{vec}(\mathbf{H}_{b,\bar{m}}) \in \mathbb{C}^{G_rG_t \times 1}$ represents the equivalent beamspace CFR vector that is sparse. Substituting this into (19), the sparse channel estimation model of mmH-MFO system is

$$\mathbf{y}_{\bar{m}} = \boldsymbol{\Upsilon}_{\bar{m}}\mathbf{h}_{b,\bar{m}} + \tilde{\boldsymbol{\eta}}_{\bar{m}}, \quad (26)$$

where $\boldsymbol{\Upsilon}_{\bar{m}} = \boldsymbol{\Psi}_{\bar{m}}\left[\mathbf{A}_T^*(\Theta_t) \otimes \mathbf{A}_R(\Theta_r)\right]$ denotes the dictionary matrix. Sparse channel estimation techniques designed for mmWave MIMO FBMC channel estimation are presented next.

Algorithm 1: OMP-based CFR estimation for mmWave hybrid MIMO-FBMC systems

Input: Dictionary matrix $\boldsymbol{\Upsilon}_{\bar{m}}$, Observation vector $\mathbf{y}_{\bar{m}}$, Stopping parameter ϵ, Array response matrices $\mathbf{A}_T(\Theta_t)$ and $\mathbf{A}_R(\Theta_r)$

Output: Estimate $\widehat{\mathbf{H}}_{\bar{m}}^{\text{OMP}}$ of the mmWave MIMO CFR matrix $\mathbf{H}_{\bar{m}}$

1 **Initialization:** $\mathcal{I}_0 = [\,]$, residue $\mathbf{r}_{-1} = \mathbf{0}$, $\mathbf{r}_0 = \mathbf{y}_{\bar{m}}$, $\hat{\mathbf{h}}_{b,\bar{m}}^{\text{OMP}} = \mathbf{0}, \boldsymbol{\Upsilon}_{\bar{m}}^{\mathcal{I}} = [\,]$, $i = 1$

2 **while** $\left(\parallel \mathbf{r}_{i-1} \parallel_2^2 - \parallel \mathbf{r}_{i-2} \parallel_2^2 \geq \epsilon\right)$ **do**

3 $\quad j = \underset{k=1,2,\ldots,G_rG_t}{\text{argmax}} \left|\boldsymbol{\Upsilon}_{\bar{m}}^H(:,k)\mathbf{r}_{i-1}\right|$

4 $\quad \mathcal{I}_i = \mathcal{I}_{i-1} \cup j$

5 $\quad \boldsymbol{\Upsilon}_{\bar{m}}^{\mathcal{I}} = \left[\boldsymbol{\Upsilon}_{\bar{m}}^{\mathcal{I}} \;\; \boldsymbol{\Upsilon}_{\bar{m}}(:,j)\right]$

6 $\quad \mathbf{h}_{b,\bar{m}}^i = \left(\boldsymbol{\Upsilon}_{\bar{m}}^{\mathcal{I}}\right)^{\dagger}\mathbf{y}_{\bar{m}}$

7 $\quad \mathbf{r}_i = \mathbf{y}_{\bar{m}} - \boldsymbol{\Upsilon}_{\bar{m}}^{\mathcal{I}}\mathbf{h}_{b,\bar{m}}^i$

8 $\quad i = i + 1$

9 **end**

10 $\hat{\mathbf{h}}_{b,\bar{m}}^{\text{OMP}}(\mathcal{I}_i) = \mathbf{h}_{b,\bar{m}}^i$

11 **return**: $\widehat{\mathbf{H}}_{\bar{m}}^{\text{OMP}} = \mathbf{A}_R(\Theta_r)\text{vec}^{-1}\left(\hat{\mathbf{h}}_{b,\bar{m}}^{\text{OMP}}\right)\mathbf{A}_T^H(\Theta_t)$

A. OMP-based mmWave hybrid MIMO-FBMC channel estimation

Given the model in (26), the estimate of the mmWave beamspace CFR vector $\mathbf{h}_{b,\bar{m}}$ can be obtained by solving the following sparse signal recovery problem

$$\min_{\mathbf{h}_{b,\bar{m}}} \parallel \mathbf{h}_{b,\bar{m}} \parallel_0, \text{ s.t. } \parallel \mathbf{y}_{\bar{m}} - \boldsymbol{\Upsilon}_{\bar{m}}\mathbf{h}_{b,\bar{m}} \parallel_2^2 \leq \epsilon_t, \tag{27}$$

where ϵ_t is a tunable parameter that depends on the noise variance σ_η^2. The popular orthogonal matching pursuit (OMP) [45] adopted for sparse channel recovery is described in Algorithm-1. The key steps of this algorithm can be briefly described as follows. In each iteration i, step-3 computes the index j of the column of the dictionary matrix $\boldsymbol{\Upsilon}_{\bar{m}}$ that is maximally correlated with the residue vector $\mathbf{r}_{i-1}$. Step-4 and step-5 respectively update the index set $\mathcal{I}_i$ and the corresponding matrix $\boldsymbol{\Upsilon}_{\bar{m}}^{\mathcal{I}}$ by including the index j and the jth column of the dictionary matrix $\boldsymbol{\Upsilon}_{\bar{m}}$. The LS solution $\mathbf{h}_{b,\bar{m}}^i$ and the corresponding residue $\mathbf{r}_i$ are obtained in step-7 and step-8, respectively. This procedure is concluded when the difference between the l_2-norm of the consecutive residues becomes lower than a predetermined threshold ϵ. As shown in step-3, the OMP procedure picks the column of the dictionary matrix $\boldsymbol{\Upsilon}_{\bar{m}}$ in a greedy fashion for minimizing the residue $\mathbf{r}_i$. Therefore, its convergence and in turn the resultant performance is sensitive both to the choice of the dictionary matrix $\boldsymbol{\Upsilon}_{\bar{m}}$ as well as to the threshold parameter ϵ. Thus, OMP results in structural errors arising due to the convergence to sub-optimal solutions [33]. The next subsection develops BL-based techniques for sparse channel estimation in mmH-MFO systems.

Algorithm 2: BL-based CFR estimation for mmWave hybrid MIMO-FBMC systems

Input: Dictionary matrix $\boldsymbol{\Upsilon}_{\bar{m}}$, Observation vector $\mathbf{y}_{\bar{m}}$, Stopping Parameters ϵ_0 and N_{max}, Array response dictionary matrices $\mathbf{A}_T(\Theta_t)$ and $\mathbf{A}_R(\Theta_r)$

Output: Estimate $\widehat{\mathbf{H}}_{\bar{m}}^{\text{BL}}$ of the mmWave MIMO CFR matrix $\mathbf{H}_{\bar{m}}$

1 **Initialization:** $\widehat{\boldsymbol{\Gamma}}^0 = \mathbf{I}_{G_rG_t}$, $k = 0$ and $\widehat{\boldsymbol{\Gamma}}^{(-1)} = \mathbf{0}$

2 **while** $\left(\parallel \widehat{\boldsymbol{\Gamma}}^{(k)} - \widehat{\boldsymbol{\Gamma}}^{(k-1)} \parallel_F^2 > \epsilon_0 \;\; \&\& \;\; k < N_{max}\right)$ **do**

3 $\quad k = k + 1$

4 $\quad \boldsymbol{\Sigma}_{\bar{m}}^{(k)} = \left(\boldsymbol{\Upsilon}_{\bar{m}}^H\mathbf{R}_\eta^{-1}\boldsymbol{\Upsilon}_{\bar{m}} + (\widehat{\boldsymbol{\Gamma}}^{(k-1)})^{-1}\right)^{-1}$;
$\quad\quad \boldsymbol{\mu}_{\bar{m}}^{(k)} = \boldsymbol{\Sigma}_{\bar{m}}^{(k)}\boldsymbol{\Upsilon}_{\bar{m}}^H\mathbf{R}_\eta^{-1}\mathbf{y}_{\bar{m}}$

5 $\quad$ Update hyperparameters $\widehat{\gamma}_i^{(k)}$ using (31)

6 **end**

7 $\widehat{\mathbf{h}}_{b,\bar{m}} = \boldsymbol{\mu}_{\bar{m}}^{(k)}$

8 **return**: $\widehat{\mathbf{H}}_{\bar{m}}^{\text{BL}} = \mathbf{A}_R(\Theta_r)\text{vec}^{-1}\left(\widehat{\mathbf{h}}_{b,\bar{m}}\right)\mathbf{A}_T^H(\Theta_t)$

B. BL-based sparse channel Estimation for mmWave Hybrid MIMO-FBMC Systems

The BL framework of sparse signal recovery begins by assigning the parameterized Gaussian prior seen below to the beamspace channel vector $\mathbf{h}_{b,\bar{m}}$ [33]

$$f(\mathbf{h}_{b,\bar{m}};\boldsymbol{\Gamma}) = \prod_{i=1}^{G_rG_t}(\pi\gamma_i)^{-1}\exp\left(-\frac{|\mathbf{h}_{b,\bar{m}}(i)|^2}{\gamma_i}\right), \tag{28}$$

where γ_i denotes the hyperparameter corresponding to the ith element $\mathbf{h}_{b,\bar{m}}(i)$ of the beamspace channel vector $\mathbf{h}_{b,\bar{m}}$ and the diagonal matrix $\boldsymbol{\Gamma} = \text{diag}(\gamma_1, \gamma_2, \cdots, \gamma_{G_rG_t}) \in \mathbb{R}^{G_rG_t \times G_rG_t}$ is comprised of the hyperparameters. Corresponding to the prior above, the *a posteriori* pdf of the beamspace channel vector $\mathbf{h}_{b,\bar{m}}$ is determined as $\mathcal{CN}(\boldsymbol{\mu}_{\bar{m}}, \boldsymbol{\Sigma}_{\bar{m}})$, where the mean $\boldsymbol{\mu}_{\bar{m}} \in \mathbb{C}^{G_rG_t \times 1}$ and covariance matrix $\boldsymbol{\Sigma}_{\bar{m}} \in \mathbb{C}^{G_rG_t \times G_rG_t}$ are obtained as [39]

$$\boldsymbol{\mu}_{\bar{m}} = \boldsymbol{\Sigma}_{\bar{m}}\boldsymbol{\Upsilon}_{\bar{m}}^H\mathbf{R}_\eta^{-1}\mathbf{y}_{\bar{m}}, \quad \boldsymbol{\Sigma}_{\bar{m}} = \left(\boldsymbol{\Upsilon}_{\bar{m}}^H\mathbf{R}_\eta^{-1}\boldsymbol{\Upsilon}_{\bar{m}}^H + \boldsymbol{\Gamma}^{-1}\right)^{-1}. \tag{29}$$

Note that the MMSE estimate $\boldsymbol{\mu}_{\bar{m}}$ of the beamspace channel vector $\mathbf{h}_{b,\bar{m}}$ reduces to the estimation of the hyperparameter matrix $\boldsymbol{\Gamma}$. The BL approach chooses the hyperparameter matrix $\widehat{\boldsymbol{\Gamma}}$ that maximizes the Bayesian evidence $\log\, f(\mathbf{y}_{\bar{m}};\boldsymbol{\Gamma}) = c_0 - \log\left(\det\left(\boldsymbol{\Sigma}_{\mathbf{y}_{\bar{m}}}\right)\right) - \mathbf{y}_{\bar{m}}^H\boldsymbol{\Sigma}_{\mathbf{y}_{\bar{m}}}^{-1}\mathbf{y}_{\bar{m}}$, where $c_0 = -MN_{RF}\log(\pi)$ and $\boldsymbol{\Sigma}_{\mathbf{y}_{\bar{m}}} = \mathbf{R}_\eta + \boldsymbol{\Upsilon}_{\bar{m}}\boldsymbol{\Gamma}\boldsymbol{\Upsilon}_{\bar{m}}^H$. As it can be readily seen, the optimization objective above is non-concave. Hence, direct maximization of the above cost-function for the estimation of the hyperparameters γ_i becomes intractable [33]. Therefore, the expectation maximization (EM) framework provides an ideal tool for iterative estimation of the sparse beamspace vector due to its low complexity and guaranteed convergence to a local optima [39]. The key steps in the EM procedure are described below.

Let $\widehat{\boldsymbol{\Gamma}}^{(k-1)}$ denote the estimate of $\boldsymbol{\Gamma}$ in the $(k-1)$st

iteration. The expectation-step (E-step) in the kth iteration determines the expected log-likelihood of the complete data set $\{\mathbf{y}_{\bar{m}}, \mathbf{h}_{b,\bar{m}}\}$ as

$$\begin{aligned}\mathcal{L}(\mathbf{\Gamma}|\widehat{\mathbf{\Gamma}}^{(k-1)}) &= \mathbb{E}_{\mathbf{h}_{b,\bar{m}}|\mathbf{y}_{\bar{m}};\widehat{\mathbf{\Gamma}}^{(k-1)}}\Big\{\log f(\mathbf{y}_{\bar{m}}, \mathbf{h}_{b,\bar{m}}; \mathbf{\Gamma})\Big\} \\ &= \mathbb{E}_{\mathbf{h}_{b,\bar{m}}|\mathbf{y}_{\bar{m}};\widehat{\mathbf{\Gamma}}^{(k-1)}}\Big\{\log f(\mathbf{y}_{\bar{m}}|\mathbf{h}_{b,\bar{m}}) + \log f(\mathbf{h}_{b,\bar{m}}; \mathbf{\Gamma})\Big\}. \end{aligned} \tag{30}$$

Since the first term inside the $\mathbb{E}\{\cdot\}$ operator of (30) does not depend on the matrix $\mathbf{\Gamma}$, the estimate $\widehat{\mathbf{\Gamma}}^{(k)}$ in the kth iteration is obtained via maximization of the second term in the M-step as

$$\begin{aligned}\widehat{\mathbf{\Gamma}}^{(k)} &= \arg\max_{\mathbf{\Gamma}} \mathbb{E}\big\{\log f(\mathbf{h}_{b,\bar{m}}; \mathbf{\Gamma})\big\} \\ &= -\arg\max_{\mathbf{\Gamma}} \sum_{i=1}^{G_r G_t} \left(\log(\gamma_i) + \frac{\mathbb{E}_{\mathbf{h}_{b,\bar{m}}|\mathbf{y}_{\bar{m}};\widehat{\mathbf{\Gamma}}^{(k)}}\big\{|\mathbf{h}_{b,\bar{m}}(i)|^2\big\}}{\gamma_i}\right).\end{aligned}$$

Differentiating the objective function above with respect to γ_i and setting the derivative equal to zero, the estimate of the hyperparameter $\widehat{\gamma}_i^{(k)}$ can be obtained as

$$\begin{aligned}\widehat{\gamma}_i^{(k)} &= \mathbb{E}_{\mathbf{h}_{b,\bar{m}}|\mathbf{y}_{\bar{m}};\widehat{\mathbf{\Gamma}}^{(k-1)}}\Big\{|\mathbf{h}_{b,\bar{m}}(i)|^2\Big\} \\ &= \mathbf{\Sigma}_{\bar{m}}^{(k)}(i,i) + |\boldsymbol{\mu}_{\bar{m}}^{(k)}(i)|^2, \end{aligned} \tag{31}$$

where the mean $\boldsymbol{\mu}_{\bar{m}}^{(k)}$ and the covariance $\mathbf{\Sigma}_{\bar{m}}^{(k)}$ are obtained using (29) by substituting $\mathbf{\Gamma} = \widehat{\mathbf{\Gamma}}^{(k-1)}$. The E-step and M-step above are repeated in succession, and the algorithm terminates either when $\| \widehat{\mathbf{\Gamma}}^{(k)} - \widehat{\mathbf{\Gamma}}^{(k-1)} \|_F^2 \leq \epsilon_0$ or after $N_{\max}$ EM iterations, whichever is achieved earlier, where the stopping parameters ϵ_0 and $N_{\max}$ have to be suitably chosen. Upon convergence, the BL-based sparse estimate of the mmH-MFO beamspace channel vector is determined as $\widehat{\mathbf{h}}_{b,\bar{m}} = \boldsymbol{\mu}_{\bar{m}}$. Subsequently, the estimate of the mmWave MIMO CFR matrix estimate $\mathbf{H}_{\bar{m}}$ can be obtained as

$$\widehat{\mathbf{H}}_{\bar{m}}^{\mathrm{BL}} = \mathbf{A}_R(\Theta_r)\mathrm{vec}^{-1}\left(\widehat{\mathbf{h}}_{b,\bar{m}}\right)\mathbf{A}_T^H(\Theta_t). \tag{32}$$

Algorithm-2 briefly summarizes the various steps of the proposed BL scheme. Let us now consider the matrix $\mathbf{H}_b = [\mathbf{h}_{b,0}, \mathbf{h}_{b,1}, \ldots, \mathbf{h}_{b,N-1}] \in \mathbb{C}^{G_r G_t \times N}$, which denotes the concatenated beamspace mmWave MIMO CFR across all the subcarriers. As described previously, since the sparsity profile of the beamspace CFR vector $\mathbf{h}_{b,\bar{m}}$ does not change with the subcarrier index $\bar{m}$, the beamspace CFR matrix $\mathbf{H}_b$ has entries in a particular row which are simultaneously zero or non-zero. Therefore, the next subsection develops an MMV based MBL-SSG approach, which exploits this 'simultaneous-sparsity' with the aid of special training sequences and selective subcarrier grouping, i.e., SSG, for achieving an improved estimation performance.

C. MMV-based MBL-SSG Channel Estimation Scheme for mmH-MFO Systems

It can be observed from (15) that the performance of channel estimation scheme in mmH-MFO systems depends on the power of the virtual training symbols $b_{\bar{m},i(1+z)}^q = d_{\bar{m},i(1+z)}^q + jI_{\bar{m},i(1+z)}^q$. As shown in Appendix-A, the intrinsic interference term can be computed as

$$I_{\bar{m},i(1+z)}^q \approx d_{\bar{m}+1,i(1+z)}^q \alpha_0 - d_{\bar{m}-1,i(1+z)}^q \alpha_0, \tag{33}$$

where the real constant α_0 obeys $\alpha_0 = \langle\xi\rangle_{\bar{m}+1,0}^{\bar{m},0} = -\langle\xi\rangle_{\bar{m}-1,0}^{\bar{m},0} > 0$. For improving the channel estimation performance, one can employ the IAM-real (IAMR) and IAM-complex (IAMC) training sequences of [38], which boost the power of the virtual training symbols without increasing the training overhead and/ or transmit power. Fig. 4 pictorially represents these sequences for $N_{RF} = 2$ RF chains, $z = 1$ zero and $N = 8$ subcarriers. For the IAMR training sequence, the OQAM training symbols adjacent to the $\bar{m}$th subcarrier are identical with opposite sign, i.e., we have $d_{\bar{m}+1,i(1+z)}^q = -d_{\bar{m}-1,i(1+z)}^q$. Thus, the power of the virtual symbols for this training sequence is $\sigma_{b,\mathrm{R}}^2 = P_d(1+4\alpha_0^2)$. Similarly, one can readily compute the power for the virtual training symbol with IAMC training sequence as $\sigma_{b,\mathrm{C}}^2 = P_d(1+4\alpha_0+4\alpha_0^2)$. Note that $\sigma_{b,\mathrm{C}}^2 > \sigma_{b,\mathrm{R}}^2$. It follows from (33) that the intrinsic interference at the FT index $(\bar{m}, i(1+z))$ depends only on the adjacent $(\bar{m}+1)$st and $(\bar{m}-1)$st subcarriers. Due to the construction of both the IAMR and IAMC training sequences in Fig. 4, it follows that the virtual training symbols $b_{\bar{m},i(1+z)}^q$ at the subcarrier indices $[\bar{m} \pm 4m_0]_N$ are identical, where $m_0 \in \mathbb{Z}$. Furthermore, they are opposite in sign to those at the subcarrier indices $[(\bar{m} \pm 2) \pm 4m_0]_N$. Thus, the dictionary matrix $\mathbf{\Upsilon}_{\bar{m}}$ in (26) obeys $\mathbf{\Upsilon}_0 = -\mathbf{\Upsilon}_2 = \mathbf{\Upsilon}_4 = -\mathbf{\Upsilon}_6 = \ldots = \mathbf{\Upsilon}_{N-4} = -\mathbf{\Upsilon}_{N-2}$. Consequently, by concatenating $\mathbf{Y}^e = [\mathbf{y}_0, -\mathbf{y}_2, \mathbf{y}_4, \ldots, \mathbf{y}_{N-4}, -\mathbf{y}_{N-2}] \in \mathbb{C}^{MN_{RF} \times (N/2)}$ across the even-indexed subcarriers, one can formulate the MMV-SSG channel estimation model as

$$\mathbf{Y}^e = \mathbf{\Upsilon}_0 \mathbf{H}_b^e + \mathbf{\Xi}^e,$$

where $\mathbf{H}_b^e \in \mathbb{C}^{G_r G_t \times (N/2)}$ is defined as $\mathbf{H}_b^e = [\mathbf{h}_{b,0}, -\mathbf{h}_{b,2}, \mathbf{h}_{b,4}, \ldots, \mathbf{h}_{b,N-4}, -\mathbf{h}_{b,N-2}]$ and $\mathbf{\Upsilon}_0 = \mathbf{\Psi}_0\left(\mathbf{A}_T^*(\Theta_t) \otimes \mathbf{A}_R(\Theta_r)\right)$. The corresponding noise matrix $\mathbf{\Xi}^e \in \mathbb{C}^{MN_{RF} \times (N/2)}$ obeys $\mathbf{\Xi}^e = \left[\tilde{\boldsymbol{\eta}}_0, -\tilde{\boldsymbol{\eta}}_2, \ldots, \tilde{\boldsymbol{\eta}}_{N-4}, -\tilde{\boldsymbol{\eta}}_{N-2}\right]$. Similarly, for the odd case, the MMV-SSG channel estimation model is obtained as

$$\mathbf{Y}^o = \mathbf{\Upsilon}_1 \mathbf{H}_b^o + \mathbf{\Xi}^o,$$

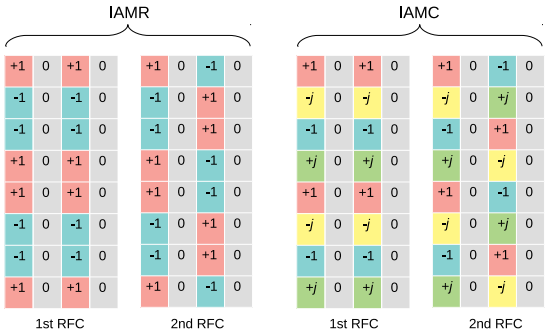


Fig. 4: IAMR and IAMC training sequences for $N_{RF} = 2$, $N = 8$, $z = 1$.

where $\mathbf{Y}^o = [\mathbf{y}_1, -\mathbf{y}_3, \ldots, \mathbf{y}_{N-3}, -\mathbf{y}_{N-1}] \in \mathbb{C}^{MN_{RF}\times(N/2)}$, $\mathbf{H}_b^o = [\mathbf{h}_{b,1}, -\mathbf{h}_{b,3}, \ldots, \mathbf{h}_{b,N-3}, -\mathbf{h}_{b,N-1}] \in \mathbb{C}^{G_rG_t\times(N/2)}$ and $\boldsymbol{\Upsilon}_1 = \boldsymbol{\Psi}_1\left(\mathbf{A}_T^*(\Theta_t) \otimes \mathbf{A}_R(\Theta_r)\right)$. The corresponding noise matrix is $\boldsymbol{\Xi}^o = \left[\tilde{\boldsymbol{\eta}}_1, -\tilde{\boldsymbol{\eta}}_3, \ldots, \tilde{\boldsymbol{\eta}}_{N-3}, -\tilde{\boldsymbol{\eta}}_{N-1}\right] \in \mathbb{C}^{MN_{RF}\times(N/2)}$. The proposed MBL-SSG scheme is now described below for the estimation of the beamspace matrix $\mathbf{H}_b^e$ for the even subcarriers. Estimation of $\mathbf{H}_b^o$ can be carried out similarly.

The parameterized prior assigned to the simultaneous sparse beamspace CFR $\mathbf{H}_b^e$ is

$$f(\mathbf{H}_b^e;\boldsymbol{\Gamma}) = \prod_{i=1}^{G_rG_t} f\Big(\mathbf{H}_b^e(i,:);\gamma_i\Big), \tag{34}$$

where the prior $f\big(\mathbf{H}_b^e(i,:);\gamma_i\big)$ assigned to the ith row is given as

$$f\big(\mathbf{H}_b^e(i,:);\gamma_i\big) = \prod_{j=1}^{N/2}(\pi\gamma_i)^{-1}\exp\left(-\frac{|\mathbf{H}_b^e(i,j)|^2}{\gamma_i}\right). \tag{35}$$

The *a posteriori* density function for the $\bar{m}$th column of the matrix $\mathbf{H}_b^e$ is given as $\mathbf{H}_b^e(:,\bar{m}) \sim \mathcal{CN}\Big(\boldsymbol{\mathcal{M}}(:,\bar{m}),\boldsymbol{\Sigma}\Big)$ [39], where

$$\boldsymbol{\mathcal{M}} = \boldsymbol{\Sigma}\boldsymbol{\Upsilon}_0^H\mathbf{R}_\eta^{-1}\mathbf{Y}^e \text{ and } \boldsymbol{\Sigma} = \left(\boldsymbol{\Upsilon}_0^H\mathbf{R}_\eta^{-1}\boldsymbol{\Upsilon}_0 + \boldsymbol{\Gamma}^{-1}\right)^{-1}. \tag{36}$$

Similar to the BL framework described in the previous subsection, the EM algorithm can now be employed for estimating the hyperparameters. Let $\widehat{\boldsymbol{\Gamma}}^{(k-1)}$ denote the estimate of hyperparameter matrix $\boldsymbol{\Gamma}$ in the $(k-1)$st iteration. The update equation for γ_i in the kth iteration is

$$\widehat{\gamma}_i^{(k)} = \frac{2}{N}\left\|\boldsymbol{\mathcal{M}}^{(k)}(i,:)\right\|_2^2 + \boldsymbol{\Sigma}^{(k)}(i,i), \tag{37}$$

where $\boldsymbol{\mathcal{M}}^{(k)}$ and $\boldsymbol{\Sigma}^{(k)}$ are obtained from (36) by setting $\boldsymbol{\Gamma} = \widehat{\boldsymbol{\Gamma}}^{(k-1)}$. Upon convergence, the MBL-SSG based estimate of the matrix $\mathbf{H}_b^e$ is obtained as $\widehat{\mathbf{H}}_b^e = \boldsymbol{\mathcal{M}}$. The estimate $\widehat{\mathbf{H}}_{\bar{m}}^{\text{SSG}}$ of the mmWave MIMO channel at the $\bar{m}$th subcarrier can subsequently be obtained as

$$\widehat{\mathbf{H}}_{\bar{m}}^{\text{SSG}} = \mathbf{A}_R(\Theta_r)\text{vec}^{-1}\left(\widehat{\mathbf{H}}_b^e(:,\bar{m})\right)\mathbf{A}_T^H(\Theta_t). \tag{38}$$

The algorithmic form of the MBL-SSG technique is provided in our technical report in [46].

V. Doubly-selective channel estimation for mmH-MFO systems

In practice, a typical mmWave MIMO channel is also temporally correlated, due to mobility of user-equipment. This section presents an online OBL-KF scheme for time and frequency-selective, i.e., doubly-selective, mmH-MFO systems that exploits both the sparsity as well as the temporal correlation. Since the data rate of mmWave systems is typically of the order of Gbps, offline channel estimation schemes such as that of [47], require multiple blocks of training symbols to be processed simultaneously. This leads to a significant increase in the processing delay. By contrast, the proposed OBL-KF technique does not have to store multiple blocks due to its online nature. It is therefore ideally suited for employment in mmH-MFO systems.

The lth MIMO delay tap in the uth block of a doubly-selective mmWave channel $\mathbf{H}_u[l]$ is modeled as [26], [48]

$$\mathbf{H}_u[l] = \sqrt{\frac{N_TN_R}{N_{cl}N_{ray}}}\sum_{i=1}^{N_{cl}}\sum_{j=1}^{N_{ray}}\alpha_{ij,u}[l]\mathbf{a}_R(\theta_{ij}^a)\mathbf{a}_T^H(\theta_{ij}^d), \tag{39}$$

where $\alpha_{ij,u}[l]$ represents the complex channel gain in the uth block for the jth ray in the ith cluster, which can be modeled for the time-selective channel as [48]

$$\alpha_{ij,u}[l] = \rho\ \alpha_{ij,u-1}[l] + \sqrt{1-\rho^2}\ w_{ij,u}[l]. \tag{40}$$

The quantity ρ represents the temporal correlation coefficient, and can be computed employing Jake's model formulated as $\rho = J_0(2\pi f_DT_B)$, where $J_0(\cdot)$ is the zeroth order Bessel function of first kind. The quantities f_D and T_B denote the maximum Doppler frequency and block duration, respectively. This is derived by modelling the time-varying coefficient $\alpha_{ij,u}[l]$ similar to a time-varying SISO wireless channel, as described in standard references such as [49], and used in other works, such as [26], [47], [48]. The complex model noise $w_{ij,u}[l]$ is represented by $\mathcal{CN}(0,\sigma_w^2)$, and it is assumed to be independent of $\alpha_{ij,u}[l],\ \forall\ u$. Following the procedure described in Section-IV, the beamspace channel vector $\mathbf{h}_{b,\bar{m},u} \in \mathbb{C}^{G_rG_t\times 1}$ in the uth block can be modeled as

$$\mathbf{h}_{b,\bar{m},u} = \rho\mathbf{h}_{b,\bar{m},u-1} + \sqrt{1-\rho^2}\ \mathbf{w}_{\bar{m},u}, \tag{41}$$

where both the beamspace channel vector $\mathbf{h}_{b,\bar{m},u}$ and the model noise vector $\mathbf{w}_{\bar{m},u} \in \mathbb{C}^{G_rG_t\times 1}$ are *sparse* and the locations of their nonzero elements in the successive blocks remain the same. For a doubly-selective mmWave MIMO channel, the measurement model of (26) can readily be modified for the uth block as

$$\mathbf{y}_{\bar{m},u} = \boldsymbol{\Upsilon}_{\bar{m}}\mathbf{h}_{b,\bar{m},u} + \tilde{\boldsymbol{\eta}}_{\bar{m},u}, \tag{42}$$

where the noise vector obeys $\tilde{\boldsymbol{\eta}}_{\bar{m},u} \in \mathbb{C}^{MN_{RF}\times 1} \sim \mathcal{CN}(\mathbf{0},\mathbf{R}_\eta)$. The OBL-KF technique once again assigns a Gaussian prior to the beamspace channel vector $\mathbf{h}_{b,\bar{m},u}$ with hyperparameter $\gamma_{i,u},\ 1 \le i \le G_rG_t$, representing the prior variance of the ith element of $\mathbf{h}_{b,\bar{m},u}$. Let $\boldsymbol{\Gamma}_u \in \mathbb{R}^{G_rG_t\times G_rG_t}$ denote the diagonal matrix of these hyperparameters and $\widehat{\boldsymbol{\Gamma}}_u^{(k-1)}$ symbolize the estimate of the hyperparameter matrix obtained in the $(k-1)$st EM iteration of the uth block. The *a posteriori* pdf of the beamspace channel vector $\mathbf{h}_{b,\bar{m},u}$ in the kth EM iteration can be evaluated similar to (29) as $\mathcal{CN}\big(\boldsymbol{\mu}_{\bar{m},u}^{(k)},\boldsymbol{\Sigma}_{\bar{m},u}^{(k)}\big)$, where

$$\begin{aligned}\boldsymbol{\mu}_{\bar{m},u}^{(k)} &= \boldsymbol{\Sigma}_{\bar{m},u}^{(k)}\boldsymbol{\Upsilon}_{\bar{m}}^H\mathbf{R}_\eta^{-1}\mathbf{y}_{\bar{m},u},\\ \boldsymbol{\Sigma}_{\bar{m},u}^{(k)} &= \left(\boldsymbol{\Upsilon}_{\bar{m}}^H\mathbf{R}_\eta^{-1}\boldsymbol{\Upsilon}_{\bar{m}} + \left(\widehat{\boldsymbol{\Gamma}}_u^{(k-1)}\right)^{-1}\right)^{-1}.\end{aligned} \tag{43}$$

Along similar lines to Section-IV-C, the estimates $\widehat{\gamma}_{i,u}^{(k)}$ in the kth iteration can be obtained as

$$\widehat{\gamma}_{i,u}^{(k)} = \boldsymbol{\Sigma}_{\bar{m},u}^{(k)}(i,i) + |\boldsymbol{\mu}_{\bar{m},u}^{(k)}(i)|^2, \tag{44}$$

with $\widehat{\boldsymbol{\Gamma}}_u^{(k)}$ denoting the diagonal matrix of hyperparameter estimates. Let $\widehat{\mathbf{h}}_{b,\bar{m},u-1|u-1}$ and $\boldsymbol{\Sigma}_{\bar{m},u-1|u-1}$ represent the filtered estimate and the associated error covariance matrix,

respectively, of the sparse beamspace channel vector $\mathbf{h}_{b,\bar{m},u}$ in the $(u-1)$st block. The prediction $\widehat{\mathbf{h}}_{b,\bar{m},u|u-1}$ and the corresponding error covariance matrix $\mathbf{\Sigma}_{\bar{m},u|u-1}$ of the OBL-KF scheme are given by similar to Kalman filter (KF) prediction equations [39], [47]:

$$\begin{aligned}\widehat{\mathbf{h}}_{b,\bar{m},u|u-1} &= \rho\ \widehat{\mathbf{h}}_{b,\bar{m},u-1|u-1},\\ \mathbf{\Sigma}_{\bar{m},u|u-1} &= \rho^2\ \mathbf{\Sigma}_{\bar{m},u-1|u-1} + (1-\rho^2)\ \widehat{\mathbf{\Gamma}}_u^{(k)}.\end{aligned} \tag{45}$$

It can be observed that the covariance matrix of the driving noise $\mathbf{w}_{\bar{m},u}$ has been set to the estimated hyperparameter matrix $\widehat{\mathbf{\Gamma}}_u^{(k)}$, which exploits the fact that the beamspace channel vector $\mathbf{h}_{b,\bar{m},u}$ and the model noise $\mathbf{w}_{\bar{m},u}$ share a common sparsity profile. The filtered estimate in the uth block and the associated error covariance is obtained using KF filtering as [39], [47]

$$\begin{aligned}\widehat{\mathbf{h}}_{b,\bar{m},u|u} &= \widehat{\mathbf{h}}_{b,\bar{m},u|u-1} + \mathbf{K}_{\bar{m},u}(\mathbf{y}_{\bar{m},u} - \mathbf{\Upsilon}_{\bar{m}}\widehat{\mathbf{h}}_{b,\bar{m},u|u-1}),\\ \mathbf{\Sigma}_{\bar{m},u|u} &= (\mathbf{I} - \mathbf{K}_{\bar{m},u}\mathbf{\Upsilon}_{\bar{m}})\mathbf{\Sigma}_{\bar{m},u|u-1},\end{aligned} \tag{46}$$

where the Kalman gain matrix $\mathbf{K}_{\bar{m},u}$ obeys

$$\mathbf{K}_{\bar{m},u} = \mathbf{\Sigma}_{\bar{m},u|u-1}\mathbf{\Upsilon}_{\bar{m}}^H(\mathbf{R}_\eta + \mathbf{\Upsilon}_{\bar{m}}\mathbf{\Sigma}_{\bar{m},u|u-1}\mathbf{\Upsilon}_{\bar{m}}^H)^{-1}.$$

The estimate $\widehat{\mathbf{H}}_{\bar{m},u}^{\text{OBL-KF}}$ of the doubly-selective mmWave MIMO CFR at the $\bar{m}$th subcarrier in the uth block can now be expressed as

$$\widehat{\mathbf{H}}_{\bar{m},u}^{\text{OBL-KF}} = \mathbf{A}_R(\Theta_r)\text{vec}^{-1}\big(\widehat{\mathbf{h}}_{b,\bar{m},u|u}\big)\mathbf{A}_T^H(\Theta_t). \tag{47}$$

The OBL-KF procedure is initialized as

$$\widehat{\mathbf{h}}_{b,\bar{m},-1|-1} = \mathbf{0}_{G_rG_t\times 1}, \mathbf{\Sigma}_{\bar{m},-1|-1} = \widehat{\mathbf{\Gamma}}_0^{(k)}, \widehat{\mathbf{\Gamma}}_0^{(0)} = \mathbf{I}_{G_rG_t}, \tag{48}$$

with the hyperparameter matrix $\widehat{\mathbf{\Gamma}}_u^{(0)}$ for the uth block initialized as $\widehat{\mathbf{\Gamma}}_u^{(0)} = \widehat{\mathbf{\Gamma}}_{u-1}^{(k)}$. Algorithm-3 briefly summarizes the various steps of the proposed OBL-KF scheme.

Algorithm 3: OBL-KF based sparse doubly-selective CFR estimation for mmWave hybrid MIMO-FBMC systems

Input: Observation $\mathbf{y}_{\bar{m},u}$, dictionary matrix $\mathbf{\Upsilon}_{\bar{m}}$, correlation coefficient ρ, noise covariance $\mathbf{R}_\eta$, stopping parameters ϵ_0 and $N_{\max}$

Output: Estimate $\widehat{\mathbf{H}}_{\bar{m},u}^{\text{OBL-KF}}$ of the doubly-selective mmWave MIMO CFR matrix $\mathbf{H}_{\bar{m},u}$

Initialization: Use (48)

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

  Update hyperparameters $\widehat{\gamma}_{i,u}^{(k)}$ using (43) and (44)

  Obtain $\widehat{\mathbf{h}}_{b,\bar{m},u|u-1}$ and $\mathbf{\Sigma}_{\bar{m},u|u-1}$ using (45)

  Obtain $\widehat{\mathbf{h}}_{b,\bar{m},u|u}$ and $\mathbf{\Sigma}_{\bar{m},u|u}$ using (46)

  return:$\widehat{\mathbf{H}}_{\bar{m},u}^{\text{OBL-KF}} = \mathbf{A}_R(\Theta_r)\text{vec}^{-1}\big(\widehat{\mathbf{h}}_{b,\bar{m},u|u}\big)\mathbf{A}_T^H(\Theta_t)$

end

VI. Bayesian Cramér-Rao Lower Bounds

This section determines the BCRLBs that serve as our performance benchmarks for the various channel estimation schemes developed above for the beamspace channel model of mmH-MFO systems described in Eq. (23) with the assumptions justified therein.

A. BCRLB for quasi-static mmH-MFO channel estimation

As shown in our technical report in [46], the Bayesian Fisher information matrix (BFIM) $\mathbf{J}_B \in \mathbb{C}^{G_rG_t\times G_rG_t}$ can be derived as

$$\mathbf{J}_B = \mathbf{\Upsilon}_{\bar{m}}^H\mathbf{R}_\eta^{-1}\mathbf{\Upsilon}_{\bar{m}} + \mathbf{\Gamma}^{-1}. \tag{49}$$

The BCRLB for the MSE of the estimate $\widehat{\mathbf{h}}_{b,\bar{m}}$ of the beamspace channel vector is given by $\text{MSE}\left(\widehat{\mathbf{h}}_{b,\bar{m}}\right) \geq \text{Tr}\left\{\mathbf{J}_B^{-1}\right\}$. The corresponding BCRLB for the estimate $\widehat{\mathbf{H}}_{\bar{m}}$ of the mmWave MIMO CFR $\mathbf{H}_{\bar{m}}$ can now be obtained as

$$\text{MSE}\big(\widehat{\mathbf{H}}_{\bar{m}}\big) \geq \text{Tr}\Big\{\mathbf{\Psi}_\Theta\mathbf{J}_B^{-1}\mathbf{\Psi}_\Theta^H\Big\},$$

where $\mathbf{\Psi}_\Theta = \mathbf{A}_T^*(\Theta_t)\otimes\mathbf{A}_R(\Theta_r)$. The BCRLB corresponding to the MMV-based MBL-SSG technique used for the estimation of the concatenated CFR $\mathbf{H}^e = [\mathbf{H}_0, \mathbf{H}_2, \cdots, \mathbf{H}_{N-2}] \in \mathbb{C}^{N_r\times N_tN/2}$ across the even subcarriers can be derived along similar lines as

$$\text{MSE}\big(\widehat{\mathbf{H}}^e\big) \geq \text{Tr}\left\{\big(\mathbf{I}_{N/2}\otimes\mathbf{\Psi}_\Theta\big)\mathbf{J}_{B,e}^{-1}\Big(\mathbf{I}_{N/2}\otimes\mathbf{\Psi}_\Theta^H\Big)\right\},$$

where the corresponding BFIM obeys $\mathbf{J}_{B,e} = \mathbf{I}_{N/2}\otimes\Big(\mathbf{\Upsilon}_0^H\mathbf{R}_\eta^{-1}\mathbf{\Upsilon}_0 + \mathbf{\Gamma}^{-1}\Big) \in \mathbb{C}^{G_rG_tN/2\times G_rG_tN/2}$.

B. BCRLB for doubly-selective mmH-MFO channel estimation

Let $\mathbf{J}_{B,\bar{m},u} \in \mathbb{C}^{G_rG_t\times G_rG_t}$ be the BFIM for the estimate of the beamspace channel $\mathbf{h}_{b,\bar{m},u}$ in (42). As shown in our technical report [46], the matrix $\mathbf{J}_{B,\bar{m},u+1}$ can be recursively derived as

$$\mathbf{J}_{B,\bar{m},u+1} = \big(\rho^2\mathbf{J}_{B,\bar{m},u}^{-1} + (1-\rho^2)\mathbf{\Gamma}_u\big)^{-1} + \mathbf{\Upsilon}_{\bar{m}}^H\mathbf{R}_\eta^{-1}\mathbf{\Upsilon}_{\bar{m}}. \tag{50}$$

Employing the above result, the MSE of the estimated doubly-selective mmWave MIMO CFR $\widehat{\mathbf{H}}_{\bar{m},u}$ in (47) is lower bounded by

$$\text{MSE}\big(\widehat{\mathbf{H}}_{\bar{m},u}^{\text{OBL-KF}}\big) \geq \text{Tr}\Big\{\mathbf{\Psi}_\Theta\mathbf{J}_{B,\bar{m},u}^{-1}\mathbf{\Psi}_\Theta^H\Big\}.$$

VII. Simulation Results

For this study, the number of TAs and RAs for the mmH-MFO system is set as $N_t = N_r \in \{8, 32\}$ with $N = 64$ subcarriers, and the number of RF chains is set as $N_{RF} \in \{4, 8\}$, unless stated otherwise. The data and training symbols are drawn from the real and imaginary parts of the 4-QAM symbols, unless stated otherwise. The spacing between adjacent antennas both at the transmitter and receiver is fixed as $d_t = d_r = \lambda/2$. The AoA and AoD range $(0, \pi)$ is divided into $G_r, G_t \in \{10, 16, 32, 48, 64\}$ angular grid points. An $L_h = 4$ tap clustered mmWave MIMO channel is considered with $N_{cl} = 4$ clusters. The AoA/ AoD associated with $N_{ray} = 1$ ray per cluster, for an *on-grid* scenario, are

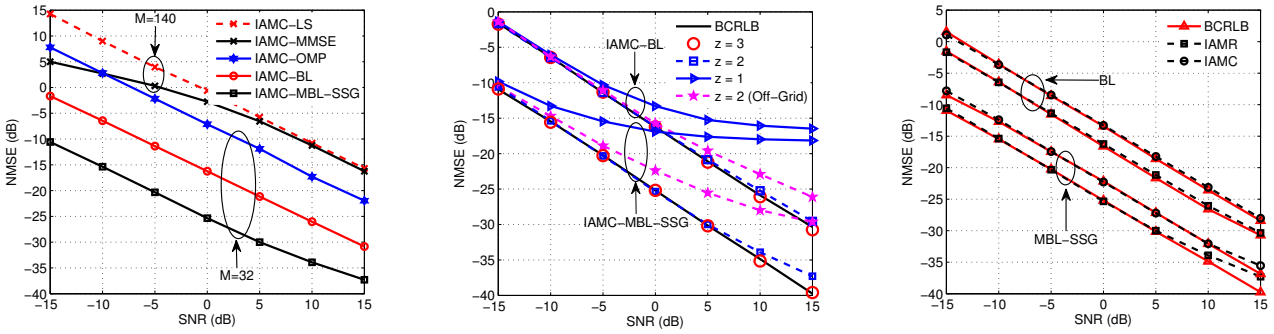


Fig. 5: Common simulation parameters: $N_t = N_r = 32, N_{RF} = 8, G_r = G_t = 32$. NMSE versus SNR comparison of a) the non-sparse LS and MMSE schemes with $M = 140$, and the sparse schemes with $M = 32$, $z = 2$ and *on-grid* setting; b) the proposed BL and MBL-SSG schemes with $M = 32$ and varying z for both *on-grid* and *off-grid* ($\sigma_{as} = 10^{-1}$) settings; and c) the proposed BL and MBL-SSG schemes relying on IAMR and IAMC training sequences for an *on-grid* setting with $z = 2$.

assumed to be uniformly distributed over $G_r/\ G_t$ angular grid points and the corresponding channel gains are modeled as i.i.d. $\mathcal{CN}(0, 1)$. This setting represents a mmWave MIMO channel without any grid-mismatch. For an *off-grid* scenario, $N_{ray} = 4$ rays associated with a cluster are assumed to have a Laplacian distribution with an angular spread σ_{as} around the mean angle of the cluster. The corresponding path gains are modeled as i.i.d. $\mathcal{CN}(0, 1/N_{ray})$. Furthermore, the mean angles of the clusters are assumed to be uniformly distributed over the angular grid points. An isotropic orthogonal transform algorithm (IOTA) [36] pulse shaping filter of duration $4T_s$ is used as the prototype filter for the mmH-MFO system. The SNR of operation on each subcarrier is defined as $2P_d/\sigma_\eta^2$. The stopping parameters ϵ_0 and $N_{\max}$ for the BL and MBL-SSG schemes are set as $\epsilon_0 = 10^{-5}$ and $N_{\max} = 100$, whereas the stopping criterion for OMP is set as $\epsilon = 0.1$. The normalized mean square error (NMSE) is defined as $\frac{1}{N}\sum_{\bar{m}=0}^{N-1} \left|\left|\widehat{\mathbf{H}}_{\bar{m}} - \mathbf{H}_{\bar{m}}\right|\right|_F^2 / \left|\left|\mathbf{H}_{\bar{m}}\right|\right|_F^2$. The corresponding BCRLBs have also been similarly normalized. The legend entries in the various plots are marked by the acronyms, $\mathtt{A} - \{\mathtt{LS}, \mathtt{MMSE}, \mathtt{OMP}, \mathtt{BL}, \mathtt{MBL-SSG}, \mathtt{OBL-KF}\}$, where A identifies the name of the associated training sequence employed for the channel estimation with $\mathtt{A} \in \{\mathtt{IAMR}, \mathtt{IAMC}\}$. For the conventional sparsity-agnostic MMSE estimation in Eq. (21), the covariance matrix $\mathbf{R}_{\mathbf{h}_{\bar{m}}}$ is set to an identity matrix.

A. Quasi-Static Channel Estimation

Fig. 5(a) compares the NMSE versus SNR performance of the non-sparse (LS and MMSE) and the sparse channel estimation schemes proposed for mmH-MFO systems relying on the IAMC training sequence using the simulation parameters of $N_t = N_r = 32, N_{RF} = 8$, $G_r = G_t = 32$ and $z = 2$. The sparse channel estimation schemes (OMP, BL and MBL-SSG) that exploit the sparsity of the mmWave MIMO channel are seen to perform significantly better than the conventional non-sparse LS and MMSE schemes. Furthermore, the sparse schemes require only $M = 32$ training vectors while, as shown in the paragraph below (21), the non-sparse schemes need at least $\lceil \frac{N_t N_r}{N_{RF}} \rceil = 128$ training vectors. Therefore, for $M = 32$, the conventional non-sparse schemes cannot provide reliable solutions, since their estimation model becomes highly underdetermined. It can also be observed that the BL and MBL-SSG schemes significantly outperform the OMP scheme, since the latter is sensitive both to the choice of the dictionary matrix $\mathbf{\Upsilon}_{\bar{m}}$ and to the stopping parameter ϵ. The MMV based MBL-SSG scheme that exploits simultaneous-sparsity leads to a further performance improvement.

Fig. 5(b) demonstrates the effect of ISI between adjacent training symbols on the NMSE versus SNR performance of the BL and MBL-SSG schemes using $N_t = N_r = 32, N_{RF} = 8, M = 32$ and $G_r = G_t = 32$. To benchmark the performance of the proposed schemes, their corresponding BCRLBs have also been plotted. It is observed that the performance of both the schemes associated with $z = 1$ exhibit floors at high SNR owing to the ISI between the adjacent training symbols. Both the BL and MBL-SSG schemes are seen to achieve their respective BCRLBs for $z \geq 2$. Note that the BCRLB is derived for an ideal scenario with the known zero and non-zero locations of the beamspace channel vector, while the BL and MBL-SSG schemes do not require this knowledge. This demonstrates the efficacy of the proposed schemes. Thus, $z = 2$ strikes an appealing spectral efficiency versus NMSE trade-off. One can also observe that the performance does not degrade significantly for this *off-grid* scenario in which the AoAs and AoDs deviate from the set of feasible AoA and AoD space Θ_r and Θ_t, respectively, employed for constructing the dictionary matrix $\mathbf{\Upsilon}_{\bar{m}}$.

Fig. 5(c) compares the NMSE versus SNR performance of the proposed BL and MBL-SSG schemes with IAMR and IAMC training sequences of Fig. 4 and the various parameters set as $N_t = N_r = 32, N_{RF} = 8, M = 32$, $G_r = G_t = 32$ and $z = 2$. Their corresponding BCRLBs have also been plotted for reference. As shown in Section-IV-C, the power of the resultant virtual symbols for the IAMR training sequence is less that that of the IAMC, i.e., $\sigma_{b,\mathrm{R}}^2 < \sigma_{b,\mathrm{C}}^2$. Therefore, the NMSE of both the BL and MBL-SSG schemes improves as the training sequence changes from IAMR to IAMC. Furthermore, both the schemes with IAMR and IAMC sequences achieve their

respective BCRLBs for $z = 2$.

B. Doubly-selective Channel Estimation

This section evaluates the performance of the proposed $\texttt{OBL}-\texttt{KF}$ technique for channel tracking in doubly-selective mmH-MFO systems. For this study, the simulations are conducted with $N_t = N_r = 8, N_{RF} = 4, M = 8$ and $G_r = G_t = 10$. An mmH-MFO wireless system having a carrier frequency of $f_c = 28$ GHz in the Q-band is considered at a user velocity of 5 Km/h, which leads to a Doppler shift of $f_D = 130$ Hz. The block length T_B is assumed to be 0.1 ms. Given these parameters, the temporal correlation coefficient becomes $\rho = J_0(2\pi f_D T_B) \approx 0.9983$.

Fig. 6(a) illustrates the NMSE versus the number of blocks u for the proposed OBL-KF technique, relying on the IAMR and IAMC training sequences of Fig. 4. The performance is also compared to that of the BL scheme that exploits sparsity. However, it does not exploit the temporal correlation of the beamspace channel vector across the blocks. The proposed $\texttt{OBL}-\texttt{KF}$ technique specifically developed for the doubly-selective channel estimation exploits both the sparsity as well as the temporal correlation, and therefore significantly outperforms the BL scheme. Furthermore, the IAMC training sequence based BL and $\texttt{OBL}-\texttt{KF}$ techniques perform better than their IAMR counterparts, since the former training sequence is more efficient in boosting the power of the virtual training symbols. Fig. 6(b) shows the NMSE versus SNR performance of the conventional KF, BL and $\texttt{OBL}-\texttt{KF}$ techniques with the IAMR training sequence for different number of zeros z. It can be observed that the conventional KF yields a poor performance, since it is sparsity-agnostic. The BL and $\texttt{OBL}-\texttt{KF}$ schemes, which leverage sparsity, lead to a considerable performance improvement over the conventional KF. Furthermore, the $\texttt{OBL}-\texttt{KF}$ technique significantly outperforms the BL scheme, because the former also exploits the temporal correlation among the blocks. Moreover, both the BL and $\texttt{OBL}-\texttt{KF}$ schemes show a similar NMSE trend upon varying the number of zeros z, as observed in Fig. 5(b), and achieve the respective BCRLBs for $z \geq 2$.

Fig. 6(c) evaluates the effect of the temporal correlation coefficient ρ over the NMSE performance for the IAMR and IAMC training sequences. Since the BL technique does not exploit the temporal correlation among the blocks, its NMSE remains unchanged with respect to the parameter ρ. However, the NMSE of the $\texttt{OBL}-\texttt{KF}$ scheme improves progressively as ρ is increased from 0.6 to 0.9. Furthermore, the NMSE improvement of this scheme is significant, when $\rho \geq 0.9$.

C. Data Detection

For the data detection, equal power allocation is considered for all the N_s data streams. Hence, the baseband precoder matrix is set as $\mathbf{F}_{\text{BB},\bar{m}} = \mathbf{I}_{N_{RF}},\ 0 \leq \bar{m} \leq N-1$, which also implies $N_s = N_{RF}$. The RF precoder and combiner matrices are designed using the AoD and AoA estimates obtained from the various channel estimation techniques presented in this work. The detailed procedure is given in Algorithm-4. It is important to note that the proposed Algorithm-4 employs only a limited information concerning the CSI, i.e., the indices of the quantized AoAs/ AoDs for the RF precoder/combiner design and does not require the full CSI to either be stored at the receiver or be fed back to the transmitter. Hence, the proposed CSI estimation and precoder/combiner design algorithms are eminently suitable for practical mmH-MFO systems. Finally, the baseband combiner matrix $\mathbf{W}_{\text{BB},\bar{m}}$ is set as $\mathbf{W}_{\text{BB},\bar{m}} = \left(\mathbf{W}_{\text{RF}}\widehat{\mathbf{H}}^{\text{A}}_{\bar{m}}\mathbf{F}_{\text{RF}}\right)^{\dagger}$, where $\text{A} \in \{\text{BL}, \text{MBL-SSG}\}$.

Algorithm 4: Design of directional RF precoder and combiner

Input: The estimated beamspace mmWave MIMO channel $\widehat{\mathbf{H}}_b \in \mathbb{C}^{G_r G_t \times N}$, number of RF chains N_{RF}, Array response dictionary matrices $\mathbf{A}_T(\Theta_t)$ and $\mathbf{A}_R(\Theta_r)$

Output: RF precoder $\mathbf{F}_{\text{RF}}$ and $\mathbf{W}_{\text{RF}}$

Initialization: $\mathbf{F}_{\text{RF}} = [\,]$, $\mathbf{W}_{\text{RF}} = [\,]$

$\mathbf{h}_b = |\widehat{\mathbf{H}}^{\text{A}}_b| * \text{diag}(\mathbf{I}_N)$

$[\mathbf{h}^{\text{sort}}_b\ \tilde{\mathbf{v}}_{\text{index}}]$ = sort($\mathbf{h}_b$,'descend')

$\mathbf{v}_{\text{index}} = \tilde{\mathbf{v}}_{\text{index}}(1 : N_{RF})$

for $i = 1 : N_{RF}$ **do**

  AoD = $\lfloor(\mathbf{v}_{\text{index}}(i) - 1)/G_t\rfloor + 1$; AoA = $\text{rem}(\mathbf{v}_{\text{index}}(i) - 1), G_r) + 1$

  $\mathbf{F}_{\text{RF}} = [\mathbf{F}_{\text{RF}}\ \mathbf{A}_T(\Theta_t)(:, \text{AoD})]$;
  $\mathbf{W}_{\text{RF}} = [\mathbf{W}_{\text{RF}}\ \mathbf{A}_R(\Theta_r)(:, \text{AoA})]$

end

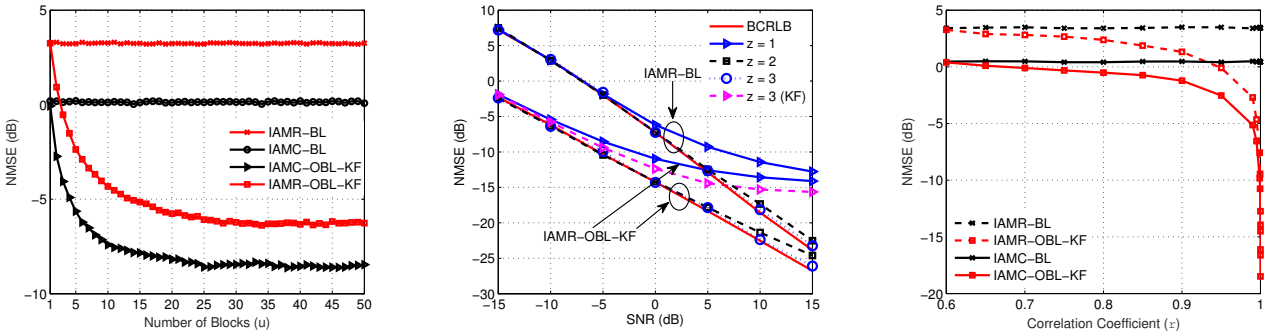
return: $\mathbf{F}_{\text{RF}}$ and $\mathbf{W}_{\text{RF}}$

Fig. 7(a) presents the BER versus SNR performance of the various competing schemes for a quasi-static scenario. A mmH-MFO system is considered using the simulation parameters of $N_t = N_r = 32, N_{RF} = 8, M = 32$, $G_r = G_t = 32$ and $z = 2$. For this study, the hybrid precoder/combiner are designed using Algorithm-4. The NMSE performance trend of the various schemes is also reflected in their respective BER performances.

In order to study the performance of the proposed algorithms for practical mmWave channels, Fig. 7(b) and 7(c) show the results for channel realizations obtained using the NYUSIM wireless channel simulator [50], which has been developed based on the statistical spatial channel models (SSCM) derived as per the 3GPP recommendations [51]. The other simulation parameters for this study are set to $N_t = N_r = 8, N_{RF} = 4, M = 16$ and $z = 2$ for Fig. 7(b), and $N_t = N_r = 32, N_{RF} = 8, M = 128$ and $z = 2$ for Fig. 7(c). It is observed that the NMSEs of the OMP, BL and MBL-SSG schemes, for suitable choice of grid-sizes G_r and G_t, show a trend similar to Fig. 5(a) upon varying the SNR. This shows that the proposed schemes are well-suited for mmWave MIMO channel estimation in practical implementations.

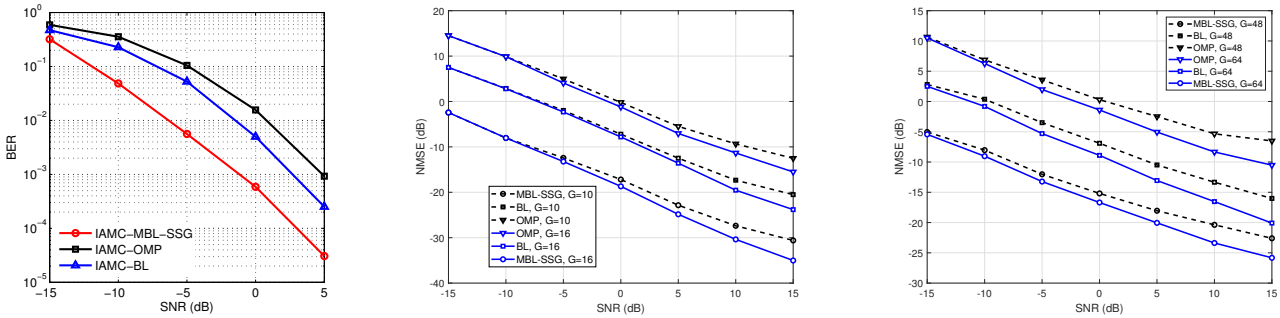
VIII. Conclusions

This work investigated the performance of FBMC-OQAM signalling conceived for next-generation mmWave communication. An architecture was developed to incorporate FBMC-OQAM filter banks in mmWave hybrid MIMO systems. Bayesian learning based sparse channel estimation techniques



(a) (b) (c)

Fig. 6: Performance comparison of the conventional KF, BL and OBL − KF schemes for doubly-selective mmWave channel estimation with $N_t = N_r = 8, N_{RF} = 4, M = 8$, $G_r = G_t = 10$ and an *off-grid* setting ($\sigma_{as} = 10^{-1}$): a) NMSE versus number of blocks (u); b) NMSE versus SNR; and c) NMSE versus temporal correlation ρ.



(a) (b) (c)

Fig. 7: (a) BER versus SNR performance of the OMP, BL and MBL − SSG based quasi-static channel estimation schemes for mmWave hybrid MIMO-FBMC systems with simulation parameters $N_t = N_r = 32, N_{RF} = 8, M = 32, G_r = G_t = 32$ and $z = 2$. (b) NMSE versus SNR performance comparison of the OMP, BL and MBL − SSG schemes for mmH-MFO systems with mmWave MIMO channel generated using NYUSIM [50] with simulation parameters set to $N_t = N_r = 8, N_{RF} = 4$ and $M = 16$. (c) NMSE versus SNR performance comparison of the OMP, BL and MBL − SSG schemes for mmH-MFO systems with mmWave MIMO channel generated using NYUSIM [50] with simulation parameters set to $N_t = N_r = 32, N_{RF} = 8$ and $M = 128$.

were proposed both for quasi-static as well as for doubly-selective mmWave hybrid MIMO-FBMC systems, which also incorporated SSG for attaining an improved performance. The Bayesian Cramér-Rao lower bounds were derived for analytically benchmarking the NMSE performance of the proposed schemes. The proposed BL and MBL techniques were seen to significantly outperform the conventional OMP, LS and MMSE schemes in terms of both the NMSE as well as BER. A Kalman filtering based BL technique has also been developed for online tracking of a doubly-selective mmWave MIMO channel. Future extensions of this work may explore the approximate message passing (AMP)-based implementation of BL [52] and sparse adaptive channel estimation schemes, similar to [53], in order to limit the complexity of FBMC-OQAM-based mmWave hybrid MIMO systems. The system model and the optimal MMSE techniques developed in our paper can potentially serve as starting points and performance benchmarks for these schemes.

Appendix A
Analysis of Intrinsic Interference

The intrinsic interference in (16) can be rewritten as

$$I^t_{\bar{m},i(1+z)} = \sum_{m\neq 0} d^q_{\bar{m}+m,i(1+z)} \langle \xi \rangle^{\bar{m},0}_{\bar{m}+m,0}, \tag{51}$$

with the quantity $\langle \xi \rangle^{\bar{m},0}_{\bar{m}+m,0} = \Im\{\xi^{\bar{m},0}_{\bar{m}+m,0}\}$. Employing (3) and the relationship of $\phi_{m,n} = \frac{\pi}{2}(m+n) - \pi mn$, the term $\xi^{\bar{m},0}_{\bar{m}+m,0}$ can be expressed as $\xi^{\bar{m},0}_{\bar{m}+m,0} = e^{j\pi m/2} \sum_{k=0}^{L_p-1} p^2[k] e^{j2\pi mk/N}$. Since the dominant contribution to the intrinsic interference arises from the symbols in the neighborhood $\Omega_{\bar{m},i(1+z)}$, it is desirable to compute the term $\xi^{\bar{m},0}_{\bar{m}+m,0}$ for $m \in \{\pm 2, \pm 1\}$. For $m = \pm 1$, we have: $\xi^{\bar{m},0}_{\bar{m}\pm 1,0} = \pm j \sum_{k=0}^{L_p-1} p^2[k] e^{\pm j2\pi k/N}$, which also implies that $\langle \xi \rangle^{\bar{m},0}_{\bar{m}\pm 1,0} = \pm \sum_{k=0}^{L_p-1} p^2[k] e^{\pm j2\pi k/N}$. Exploiting the fact that the quantity $\sum_{k=0}^{L_p-1} p^2[k] e^{\pm j2\pi k/N}$ is real and thus equal to its complex conjugate, one obtains $\langle \xi \rangle^{\bar{m},0}_{\bar{m}+1,0} = -\langle \xi \rangle^{\bar{m},0}_{\bar{m}-1,0} = \alpha_0$. Next, for $m = \pm 2$, it can be seen that $\xi^{\bar{m},0}_{\bar{m}\pm 2,0} = -\sum_{k=0}^{L_p-1} p^2[k] e^{\pm j2\pi 2k/N}$. Since the right hand side of the this expression coincides with the ambiguity

function $A_p(n,m)$ computed at $(n,m)=(0,\pm 2)$, it follows from [54, Eq. (19)] that $\xi^{\bar{m},0}_{\bar{m}\pm 2,0}=0$. Thus, $\langle\xi\rangle^{\bar{m},0}_{\bar{m}\pm 2,0}=0$. Exploiting the above analysis in (51), the intrinsic interference at the FT index $(\bar{m},i(1+z))$ can be evaluated as

$$I^q_{\bar{m},i(1+z)} \approx d^q_{\bar{m}+1,i(1+z)}\alpha_0 - d^q_{\bar{m}-1,i(1+z)}\alpha_0. \tag{52}$$

REFERENCES

[1] C. Cordeiro, D. Akhmetov, and M. Park, "IEEE 802.11 ad: Introduction and performance evaluation of the first multi-Gbps WiFi technology," in *Proceedings of the 2010 ACM international workshop on mmWave communications: from circuits to networks*. ACM, 2010, pp. 3–8.

[2] F. Giannetti, M. Luise, and R. Reggiannini, "Mobile and personal communications in the 60 GHz band: A survey," *Wireless Personal Communications*, vol. 10, no. 2, pp. 207–243, 1999. [Online]. Available: https://doi.org/10.1023/A:1018308429332

[3] 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.

[4] 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.

[5] 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.

[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] T. Pollet, M. V. Bladel, and M. Moeneclaey, "BER sensitivity of OFDM systems to carrier frequency offset and Wiener phase noise," *IEEE Trans. Communications*, vol. 43, no. 234, pp. 191–193, 1995.

[8] M. Morelli, C. J. Kuo, and M. Pun, "Synchronization techniques for orthogonal frequency division multiple access (OFDMA): A tutorial review," *Proceedings of the IEEE*, vol. 95, no. 7, pp. 1394–1427, 2007.

[9] B. Farhang-Boroujeny, "OFDM versus filter bank multicarrier," *IEEE Signal Process. Mag.*, vol. 28, no. 3, pp. 92–112, 2011.

[10] A. Aminjavaheri, A. Farhang, A. R. Reyhani, and B. Farhang-Boroujeny, "Impact of timing and frequency offsets on multicarrier waveform candidates for 5G," *CoRR*, vol. abs/1505.00800, 2015.

[11] A. I. Pérez-Neira, M. Caus, R. Zakaria, D. L. Ruyet, E. Kofidis, M. Haardt, X. Mestre, and Y. Cheng, "MIMO signal processing in Offset-QAM based filter bank multicarrier systems," *IEEE Trans. Signal Processing*, vol. 64, no. 21, pp. 5733–5762, 2016.

[12] P. Singh, R. Budhiraja, and K. Vasudevan, "Probability of error in MMSE detection for MIMO-FBMC-OQAM systems," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 8, pp. 8196–8200, Aug 2019.

[13] P. Singh, R. Budhiraja, and K. Vasudevan, "SER analysis of MMSE combining for MIMO FBMC-OQAM systems with imperfect CSI," *IEEE Communications Letters*, vol. 23, no. 2, pp. 226–229, 2019.

[14] J. Li, D. Chen, D. Qu, Y. Zhang, and T. Jiang, "Receiver design for Alamouti coded FBMC system in highly frequency selective channels," *IEEE Transactions on Broadcasting*, vol. 65, no. 3, pp. 601–608, 2018.

[15] W. Liu, D. Chen, K. Luo, T. Jiang, and D. Qu, "FDM-structured preamble optimization for channel estimation in MIMO-OQAM/FBMC systems," *IEEE Trans. Wireless Communications*, vol. 17, no. 12, pp. 8433–8443, 2018.

[16] P. Singh, E. Sharma, K. Vasudevan, and R. Budhiraja, "CFO and channel estimation for frequency selective MIMO-FBMC/OQAM systems," *IEEE Wireless Commun. Letters*, vol. 7, no. 5, pp. 844–847, 2018.

[17] P. Singh, H. B. Mishra, A. K. Jagannatham, and K. Vasudevan, "Semi-blind, training, and data-aided channel estimation schemes for MIMO-FBMC-OQAM systems," *IEEE Transactions on Signal Processing*, vol. 67, no. 18, pp. 4668–4682, Sep. 2019.

[18] H. Wang, "Sparse channel estimation for MIMO-FBMC/OQAM wireless communications in smart city applications," *IEEE Access*, vol. 6, pp. 60 666–60 672, 2018.

[19] D. Chen, Y. Tian, D. Qu, and T. Jiang, "OQAM-OFDM for wireless communications in future internet of things: A survey on key technologies and challenges," *IEEE Internet of Things Journal*, vol. 5, no. 5, pp. 3788–3809, 2018.

[20] A. Farhang, N. Marchetti, L. E. Doyle, and B. Farhang-Boroujeny, "Filter bank multicarrier for massive MIMO," in *IEEE 80th Vehicular Technology Conference, VTC Fall, Vancouver, BC, Canada, September 14-17*, 2014, pp. 1–7.

[21] A. Aminjavaheri, A. Farhang, N. Marchetti, L. E. Doyle, and B. Farhang-Boroujeny, "Frequency spreading equalization in multicarrier massive MIMO," in *IEEE International Conference on Communication, ICC, United Kingdom, June 8-12, Workshop Proceedings*, 2015, pp. 1292–1297.

[22] P. Singh, H. B. Mishra, A. K. Jagannatham, K. Vasudevan, and L. Hanzo, "Uplink sum-rate and power scaling laws for multi-user massive MIMO-FBMC systems," *IEEE Transactions on Communications*, pp. 1–1, 2019.

[23] 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.

[24] 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.

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

[26] 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.

[27] X. Gao, L. Dai, S. Zhou, A. M. Sayeed, and L. Hanzo, "Wideband beamspace channel estimation for millimeter-wave MIMO systems relying on lens antenna arrays," *IEEE Transactions on Signal Processing*, vol. 67, no. 18, pp. 4809–4824, 2019.

[28] S. Pratschner, M. Lerch, D. Schutzenhofer, M. Hofer, J. Blumenstein, S. Sangodoyin, T. Zemen, A. Prokes, A. F. Molisch, S. Caban *et al.*, "Sparsity in the delay-Doppler domain for measured 60 GHz vehicle-to-infrastructure communication channels," in *2019 IEEE International Conference on Communications Workshops (ICC Workshops)*, pp. 1–6.

[29] V. Raghavan, A. Partyka, L. Akhoondzadeh-Asl, M. A. Tassoudji, O. H. Koymen, and J. Sanelli, "Millimeter wave channel measurements and implications for PHY layer design," *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 12, pp. 6521–6533, 2017.

[30] H. Deng and A. Sayeed, "Mm-wave MIMO channel modeling and user localization using sparse beamspace signatures," in *2014 IEEE 15th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. IEEE, 2014, pp. 130–134.

[31] S. Hur, S. Baek, B. Kim, Y. Chang, A. F. Molisch, T. S. Rappaport, K. Haneda, and J. Park, "Proposal on millimeter-wave channel modeling for 5G cellular system," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 3, pp. 454–469, 2016.

[32] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[33] 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.

[34] D. L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ_1 minimization," *Proceedings of the National Academy of Sciences*, vol. 100, no. 5, pp. 2197–2202, 2003.

[35] A. Viholainen, T. Ihalainen, T. H. Stitz, M. Renfors, and M. G. Bellanger, "Prototype filter design for filter bank based multicarrier transmission," in *17th European Signal Processing Conference, EUSIPCO 2009, Glasgow, Scotland, UK, August 24-28*, 2009, pp. 1359–1363.

[36] P. Siohan, C. Siclet, and N. Lacaille, "Analysis and design of OFDM/OQAM systems based on filterbank theory," *IEEE Trans. Signal Processing*, vol. 50, no. 5, pp. 1170–1183, 2002.

[37] C. Lélé, J. Javaudin, R. Legouable, A. Skrzypczak, and P. Siohan, "Channel estimation methods for preamble-based OFDM/OQAM modulations," *European Transactions on Telecommunications*, vol. 19, no. 7, pp. 741–750, 2008.

[38] E. Kofidis, D. Katselis, A. A. Rontogiannis, and S. Theodoridis, "Preamble-based channel estimation in OFDM/OQAM systems: A review," *Signal Processing*, vol. 93, no. 7, pp. 2038–2054, 2013.

[39] S. M. Kay, "Fundamentals of statistical signal processing, Volume 1," *PTR Prentice-Hall, Englewood Cliffs, NJ*, 1993.

[40] S. Srivastava, A. K. Jagannatham, and L. Hanzo, "Sparse doubly-selective channel estimation techniques for OSTBC MIMO-OFDM systems: A hierarchical Bayesian Kalman filter based approach." IIT

Kanpur, Kanpur, India, Tech. Rep. TR 1.2020, 2020. [Online]. Available: http://www.iitk.ac.in/mwn/documents/MWNLab_TR_HBKF_MIMO_OFDM
[41] W. G. Jeon, K. H. Paik, and Y. S. Cho, "Two-dimensional MMSE channel estimation for OFDM systems with transmitter diversity," in *IEEE 54th Vehicular Technology Conference. VTC Fall 2001. Proceedings (Cat. No. 01CH37211)*, vol. 3. IEEE, 2001, pp. 1682–1685.
[42] S. Gao, X. Cheng, and L. Yang, "Estimating doubly-selective channels for hybrid mmwave massive MIMO systems: A doubly-sparse approach," *IEEE Transactions on Wireless Communications*, 2020.
[43] B. Wang, F. Gao, S. Jin, H. Lin, and G. Y. Li, "Spatial-and frequency-wideband effects in millimeter-wave massive MIMO systems," *IEEE Transactions on Signal Processing*, vol. 66, no. 13, pp. 3393–3406, 2018.
[44] T. S. Rappaport, R. W. Heath Jr, R. C. Daniels, and J. N. Murdock, *Millimeter Wave Wireless Communications*. Pearson Education, 2014.
[45] 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.
[46] S. Srivastava, P. Singh, A. K. Jagannatham, A. Karandikar, and L. Hanzo, "Technical report: Channel estimation schemes for millimeter wave hybrid MIMO-FBMC-OQAM systems." IIT Kanpur, Kanpur, India, Tech. Rep. TR 2.2020, 2020. [Online]. Available: http://www.iitk.ac.in/mwn/documents/MWNLab_TR_mmWave_MIMO_FBM
[47] R. Prasad, C. R. Murthy, and B. D. Rao, "Joint channel estimation and data detection in MIMO-OFDM systems: A sparse Bayesian learning approach," *IEEE Trans. Signal Processing*, vol. 63, no. 20, pp. 5369–5382, 2015.
[48] J. He, T. Kim, H. Ghauch, K. Liu, and G. Wang, "Millimeter wave MIMO channel tracking systems," in *2014 IEEE Globecom Workshops (GC Wkshps)*. IEEE, 2014, pp. 416–421.
[49] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge university press, 2005.
[50] S. Sun, G. R. MacCartney, and T. S. Rappaport, "A novel millimeter-wave channel simulator and applications for 5G wireless communications," in *2017 IEEE International Conference on Communications (ICC)*. IEEE, 2017, pp. 1–7.
[51] 3GPP, "Study on channel model for frequencies from 0.5 to 100 GHz (release 15)," *3rd Generation Partnership Project (3GPP), TR 38.901 V14.1.1, Jul. 2017.*
[52] J. P. Vila and P. Schniter, "Expectation-maximization Gaussian-mixture approximate message passing," *IEEE Transactions on Signal Processing*, vol. 61, no. 19, pp. 4658–4672, 2013.
[53] Y. Chen, Y. Gu, and A. O. Hero, "Regularized least-mean-square algorithms," *arXiv preprint arXiv:1012.5066*, 2010.
[54] J. Du and S. Signell, "Classic OFDM systems and pulse shaping OFDM/OQAM systems," 2007.

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.

Prem Singh received the M.Tech and PhD degrees in Electrical Engineering from the Indian Institute of Technology Kanpur, India, in 2011 and 2020 respectively. He is currently employed as a Project Executive Officer on the Indigenous 5G TestBed project, where he is developing software and hardware algorithms for an end-to-end 3GPP compliant 5G-NR Testbed. His current research interest includes parameter estimation and receiver design for 5G and next-generation technologies including Filter Bank Multicarrier (FBMC), Orthogonal Time-Frequency space (OTFS), massive MIMO and Millimeter Wave, and designing end-to-end practical 4G/5G wireless systems using 3GPP standards.

Aditya K. Jagannatham (S'04-M'05) received his Bachelors degree from Indian Institute of Technology, Bombay and M.S. and Ph.D. degrees from University of California, San Diego, U.S.A. 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 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. He was awarded the CAL(IT)2 fellowship at the University of California San Diego and the Upendra Patel Achievement Award at Qualcomm.

Abhay Karandikar is currently the Director of Indian Institute of Technology Kanpur (on leave from IIT Bombay). He is also a member (part-time) of Telecom Regulatory Authority of India. At IIT Bombay, he served as Institute Chair Professor in the Electrical Engineering Department, Dean (Faculty Affairs) from 2017 to 2018, and Head of the Electrical Engineering Department from 2012 to 2015. He is a founding member of Telecom Standards Development Society, India (TSDSI), India's standards body for telecommunications. He was the Chairman of TSDSI from 2016 to 2018. His research interests include resource allocation in wireless networks, software defined networking, frugal 5G, and rural broadband. A detailed biography can be found at https://www.ee.iitb.ac.in/ karandi/.

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. Over 40 of them are Professors at various stages of their careers in academia and many of them are leading scientists in the wireless industry.