Sparse Bayesian Learning Aided Estimation of Doubly-Selective MIMO Channels for Filter Bank Multicarrier Systems

Prem Singh, Suraj Srivastava, Amrita Mishra, Aditya K. Jagannatham, and Lajos Hanzo, *Life Fellow IEEE//*

***Abstract*—Sparse Bayesian learning (SBL)-based channel state information (CSI) estimation schemes are developed for filter bank multicarrier (FBMC) systems using offset quadrature amplitude modulation (OQAM). Initially, an SBL-based channel estimation scheme is designed for a frequency-selective quasi-static single-input single-output (SISO)-FBMC system, relying on the interference approximation method (IAM). The IAM technique, although has low complexity, is only suitable for channels exhibiting mild frequency-selectivity. Hence, an alternative time-domain (TD) model based sparse channel estimation framework is developed for highly frequency-selective channels. Subsequently, the Kalman filtering (KF)-based IAM and its TD counterpart are developed for sparse doubly-selective CSI estimation in SISO-FBMC systems. These schemes are also extended to FBMC-based multiple-input multiple-output (MIMO) systems, for both quasi-static and doubly-selective channels, after demonstrating the special block and group-sparse structures of the IAM and TD-based models respectively, which are the characteristic features of such channels. The Bayesian Cramér-Rao lower bounds (BCRLBs) and the time-recursive BCRLBs are derived for the proposed quasi-static as well as doubly-selective sparse CSI estimation models, respectively. Our numerical results closely match the analytical findings, demonstrating the enhanced performance of the proposed schemes over the existing techniques.**

***Index Terms*—Filter bank multicarrier, sparse Bayesian learning, channel estimation, Kalman filtering.**

I. Introduction

Orthogonal frequency division multiplexing (OFDM) signaling has gained prominence for broadband transmission in both wired and wireless systems. The subcarriers in an OFDM system are modulated and demodulated using the computationally efficient IFFT and FFT operations, respectively, which leads to low complexity, thus rendering it well-suited for practical implementation [1], [2]. However, the rectangular prototype filter of the IFFT/ FFT filter bank in an OFDM system has a sinc-shaped spectrum in the frequency-domain that suffers from relatively high out-of-band (OOB) emission, which makes the performance of these systems sensitive to practical imperfections, such as timing- and carrier frequency-offset (CFO). Therefore, OFDM may not be ideally suited for all the use cases in future mobile communication systems [3], [4]. In order to address the above shortcomings of OFDM, alternative multicarrier waveforms based on filter bank processing at the transmitter and receiver relying on different prototype filters have attracted significant research interest. In particular, offset quadrature amplitude modulation (OQAM)-aided filter bank multicarrier (FBMC) transmission has emerged as one of the potential waveform candidates to replace OFDM in next generation wireless systems [2], [4], [5]. This is because of the fact that the sharp cut-off prototype filter in FBMC-OQAM can significantly lower the OOB emission, and also obviate the use of cyclic-prefix (CP). The key differences between OFDM and FBMC-OQAM systems lie i) in the fact that the latter adopts OQAM symbols rather than QAM symbols; and ii) in the specific choice of the time-domain prototype windowing. In contrast to OFDM, FBMC uses a non-rectangular pulse, e.g. the isotropic orthogonal transform algorithm (IOTA) [6], Phydyas [7], root raised cosine (RRC), whose duration is much greater than the FBMC symbol duration. Therefore, in order to attain an identical spectral efficiency as that of OFDM, the adjacent time-domain FBMC symbols overlap with each other [8]. Furthermore, it must be noted that orthogonality in an FBMC system holds only in the real field. The accuracy of channel estimate available at the receiver has a key role in the reliable detection of data in FBMC-OQAM systems and hence, channel state information (CSI) estimation for these systems has been at the center of various contributions, as described in the next subsection. In the sequel, FBMC-OQAM is referred to as FBMC, for brevity.

A. Literature Review

Based on the FBMC system models, training-based channel estimation schemes can be broadly classified as follows. The first category is based on the interference approximation method (IAM), which relies on the assumption that the channel is approximately frequency-flat at the subcarrier level [9]. The IAM-based schemes exploit the property that the desired FBMC symbol interferes only with its neighborhood symbols, and approximate the channel's frequency response

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

Prem Singh and Amrita Mishra are with Networking, Communication and Signal Processing group at IIIT Bangalore, Karnataka-560100, India (e-mail: {prem.singh, amrita.mishra}@iiitb.ac.in). Suraj Srivastava and Aditya K. Jagannatham are with the Department of Electrical Engineering, IIT Kanpur, Uttar Pradesh-208016, India (e-mail: {ssrivast, adityaj}@iitk.ac.in). Lajos 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). The first two authors contributed equally.

(CFR) by a constant envelope over this neighborhood [9]. The second category is based on the time-domain (TD) model, which is used for the estimation of highly frequency selective channels, since unlike the IAM approach, it does not require the channel to be frequency-flat over each subcarrier [10]. In the context of IAM-based channel estimation, Lele *et al.* in [9] have proposed a least-squares (LS) CFR estimation technique for FBMC-aided single-input single-output (SISO) systems. As a further advance, Katselis *et al.* [6] presented a comparative study of channel estimation in SISO-FBMC and CP-OFDM systems, while Choi *et al.* [7] investigated a pilot-aided LS channel estimation scheme that exploits the intrinsic interference for enhancing its accuracy. Savaux *et al.* [11] studied joint minimum mean square error (MMSE)-based channel and noise variance estimation for SISO-FBMC systems, while the authors of [12]–[14] developed IAM-based channel estimation schemes for SISO-FBMC systems using machine learning based approaches. Cui *et al.* in [15] proposed a novel scattered IAM-based pilot-aided channel estimation scheme for FBMC systems. One can find a comprehensive literature on IAM-aided CSI estimation schemes conceived for SISO/multiple-input multiple-output (MIMO)-FBMC systems in [16]. Reference [17] investigated the impact of scattered silot-based channel estimation at the receiver on the performance of channel adaptive modulation (CAM) in terms of channel estimation errors and bit error rate (BER).

When considering TD model based schemes, Kong *et al.* [10] have investigated the weighted LS (WLS) and MMSE schemes designed for highly frequency-selective CSI estimation in SISO-FBMC systems. As a further development, [18], [19] designed channel estimation schemes for highly frequency-selective MIMO-FBMC systems. Perez Neira *et al.* [20] reviewed the family of channel estimation techniques conceived for scenarios having both high and low frequency-selectivity. Zhang *et al.* in [21] analyzed the impact of a doubly-selective non-sparse channel on the performance of FBMC systems in terms of the mean square error (MSE) of the received signal. Ihalainen *et al.* in [22] extended the frequency sampling based equalization technique in MIMO-FBMC receivers for equalization at each subcarrier.

It has been shown in [23], [24], and in the references therein that due to the wide bandwidth and high sampling frequency, a typical wireless channel is sparse in the delay-domain and has a high delay spread. However, the training based schemes reviewed above do not exploit the sparsity of the multipath wireless channel, which tends to have only a few dominant channel taps. Exploiting the inherent sparsity can lead to a significant reduction in the required number of pilot transmissions, hence improving the bandwidth efficiency. Recently, He *et al.* [25] have proposed orthogonal matching pursuit (OMP)-based schemes for doubly-selective sparse channel estimation in SISO-FBMC systems, while Wang [26] designed OMP-based techniques for sparse CSI estimation in MIMO-FBMC systems. The OMP-based techniques developed in [25], [26], however rely on the IAM model of FBMC systems. Thus, their estimation performance degrades for highly frequency-selective sparse wireless channels. Furthermore, He *et al.* [25] do not consider the time-correlation of a doubly-selective transmission medium. Also, the performance of their solution is critically dependent on the selection of both the stopping criterion and of the sensing matrix. The performance of other popular sparse signal recovery schemes found in the literature, such as the least absolute shrinkage and selection operator (Lasso) [27], is sensitive to the regularization parameter, whereas the focal underdetermined system solver (FOCUSS) [28] suffers from convergence deficiencies [29]. The authors of [30] develops sparse Bayesian learning (SBL)-aided channel estimation schemes for millimeter (mmWave) hybrid MIMO-FBMC systems considering both quasi-static and doubly-selective scenarios. Their work is based on the assumption that the symbol duration is sufficiently longer than the maximum delay spread of the channel, hence making it feasible to employ only IAM-based estimation. Furthermore, since the mmWave MIMO channel is sparse in the angular domain, a frame-wise channel estimation model has been developed therein in order to excite all the spatial modes of the channel. Moreover, the work in [30] only considers forward filtering and derives its Bayesian Cramer-Rao lower bound (BCRLB). Donoho *et al.* developed a set of iterative algorithms in [31], popularly known as approximate message passing (AMP), for sparse signal estimation. The AMP-based basis pursuit (BP) and Lasso-techniques were derived in [32]. Subsequently, the AMP algorithm was also extended to a Bayesian setting in [33], where it has been demonstrated that in comparison to the original expectation-maximization (EM) based SBL framework, which necessitates a matrix inversion to compute the *a posteriori* covariance matrix, the AMP-based SBL algorithms have a low-complexity, since they only track the *a posteriori* mean and variance of each element of the sparse vector. The performance of the AMP-based algorithms is typically seen to be comparable, but sub-optimal, in comparison to the SBL algorithm, since the messages are approximated by the Gaussian distribution [33]. Furthermore, the AMP-based techniques are not suitable for applications wherein the entire *a posteriori* covariance matrix is required, such as for Kalman filtering. The authors of [34] developed SBL-based sparse channel estimation for OFDM systems considering both quasi-static and time-varying scenarios for a sparse multipath power delay profile (PDP). Another work in [35] designed CSI acquisition and tracking schemes for a massive MIMO system in a time-varying scenario considering the angular domain sparsity. However, these contributions considered batch processing of the pilot observations for CSI tracking and updating the hyperparameters, which make them block-based processing in nature. Another similar treatise [36] focuses on downlink channel estimation for a time-varying massive MIMO channel. Here, a coordinate-wise EM-based SBL framework is developed for estimating the sparse virtual channel and the noise covariance, followed by tracking via an optimal Bayesian Kalman filtering (KF). However, these contributions do not derive any performance bounds for their algorithms. In order to overcome the above drawbacks, SBL [29] and its multiple signal based extension M-SBL [37] have been demonstrated in [23] to attain superior sparse signal recovery. Thus, motivated by the success of SBL, this treatise develops SBL-based sparse channel estimation schemes for both

TABLE I: A tabular form of literature review on CSI estimation for SISO/MIMO-FBMC systems.

	[6], [8], [11]-[14]	[9]-[10]	[7]	[15]	[17]-[18]	[19]	[24]	[25]	**Proposed**
SISO-FBMC	✓	✓	✓	✓	✓	✓	✓	×	✓
MIMO-FBMC	×	×	×	✓	×	✓	×	✓	✓
IAM model	✓	✓	✓	✓	✓	✓	✓	✓	✓
TD model	×	×	✓	×	✓	✓	×	×	✓
Sparsity	×	×	×	×	×	×	✓	✓	✓
Simultaneous-sparsity	×	×	×	×	×	×	×	×	✓
Quasi-static channel	✓	✓	✓	✓	✓	✓	✓	✓	✓
Doubly-selective channel	×	×	×	×	×	×	×	×	✓
Online estimation	×	×	×	×	×	×	×	×	✓
BCRLB	×	✓	✓	×	×	×	×	×	✓

SISO and MIMO-FBMC systems, operating both in quasi-static and doubly-selective scenarios, while considering both low as well as high grade of frequency-selectivity. In addition, the designs herein are online in nature, which renders them ideally suited for tracking doubly-selective CSI associated with a sparse PDP. The present treatise also derives the BCRLBs for benchmarking the performance of the proposed algorithms. A tabular representation of the above literature review depicting the seminal contributions and explicitly contrasting our new contributions is given in Table I. Our key contributions are given next.

B. Contributions

- We commence with the development of sparse channel estimation schemes for SISO-FBMC systems. By considering a quasi-static scenario, initially, an SBL scheme is developed for IAM-based ill-posed sparse channel estimation. The proposed IAM-SBL technique employs a parameterized Gaussian prior and estimates the associated hyperparameters using the EM framework.
- In order to overcome the performance degradation of IAM-SBL for the estimation of a channel with high frequency-selectivity, an alternative TD model based scheme, termed as TD-SBL, is developed. In contrast to [30], the present work focuses on SBL-based channel estimation for sub-6 GHz. The proposed TD-SBL scheme herein, unlike the IAM model, is also capable of successfully estimating strongly frequency-selective channels. Furthermore, this work additionally develops the pertinent analytical results for backward smoothing along with forward filtering to enhance the estimation accuracy along with deriving their respective BCRLBs.
- For the estimation of a doubly-selective wireless channel, KF-based online schemes, referred to as IAM-SBL-KF and TD-SBL-KF, are developed that exploit both the temporal-correlation as well as the inherent sparsity.
- Subsequently, a novel IAM-based block-sparse model is developed for MIMO-FBMC systems. The IAM based block-sparse SBL (IAM-BSBL) technique is proposed for exploiting the block-sparsity for CSI estimation, which demonstrably leads to improved performance. Alternatively, for a MIMO channel that is highly frequency selective, an analogous TD-based estimation scheme is developed after proving that the model exhibits group-sparsity.
- The KF counterparts of the IAM-BSBL and TD-group sparse SBL (GSBL) schemes, termed as IAM-BSBL-KF and TD-GSBL-KF respectively, are also conceived for doubly-selective MIMO channels.
- The Bayesian Cramer-Rao lower bounds (BCRLBs) are derived for benchmarking the performance of the proposed CSI estimation techniques. Furthermore, detailed analysis is presented to determine and compare the computational complexities of the proposed and existing schemes.

C. Notations

The quantity $j \triangleq \sqrt{-1}$ and the operators $\Im\{\cdot\}$ and $\Re\{\cdot\}$ symbolize the imaginary and real parts. An $L \times L$ diagonal matrix is denoted by $\mathrm{diag}\big[\gamma_0,\ \gamma_1, \cdots, \gamma_{L-1}\big]$, while $\mathrm{Tr}(\cdot)$ represents the trace of a square matrix. The notation $\otimes$ symbolizes the Kronecker product, while the operation $\mathrm{vec}(\cdot)$ column-wise vectorizes a matrix. The operator $\mathbb{E}[\cdot]$ denotes the statistical expectation.

II. SISO/MIMO-FBMC System and Channel Estimation Models

1) SISO-FBMC System: In a SISO-FBMC system with N subcarriers, the transmitted baseband signal is given as [38]

$$s[k] \;=\; \sum_{m=0}^{N-1} \sum_{n\in\mathbb{Z}} d_{m,n}\chi_{m,n}[k]. \tag{1}$$

Here $\mathbb{Z}$ denotes set of integers and k symbolizes the sample-index associated with the sampling time T/N, where T denotes the duration of a QAM symbol $c_{m,n}$ at the symbol instant n and subcarrier m. The QAM symbols are drawn from a symmetric QAM constellation, such as $\{4, 16, 64, 256\}$-QAM. The $T/2$-duration OQAM symbol $d_{m,n}$ is extracted from the real and imaginary parts of the T-duration QAM symbol [16]. The real and imaginary parts of a QAM symbol are independent and identically distributed (i.i.d.) with power P_d such that $\mathbb{E}\big[d_{m,n}d^*_{m,n}\big] = P_d$, which implies that $\mathbb{E}\big[c_{m,n}c^*_{m,n}\big] = 2P_d$. The quantity $\chi_{m,n}[k]$ that symbolizes the FBMC basis function at the frequency-time (FT) index (m, n) is $\chi_{m,n}[k] = p[k - nN/2]e^{j2\pi mk/N}e^{j\phi_{m,n}}$, where $p[k]$ denotes the impulse response of the real valued symmetrical prototype filter of length L_p. The phase factor $\phi_{m,n}$ obeys $\phi_{m,n} = (\pi/2)(m + n) - \pi mn$ [38]. Let the quantity $\xi^{\bar{m},\bar{n}}_{m,n}$ be defined as $\xi^{\bar{m},\bar{n}}_{m,n} = \sum_{k=-\infty}^{+\infty} \chi_{m,n}[k]\chi^*_{\bar{m},\bar{n}}[k]$. In order

to be able to perfectly reconstruct $d_{m,n}$ in a distortion-free transmission medium, $\xi_{m,n}^{\bar{m},\bar{n}}$ must obey $\Re\{\xi_{m,n}^{\bar{m},\bar{n}}\} = \delta_{m,\bar{m}}\delta_{n,\bar{n}}$, which implies that $\xi_{m,n}^{\bar{m},\bar{n}} = 1$, for $(m,n) = (\bar{m},\bar{n})$, and for $(m,n) \neq (\bar{m},\bar{n})$, $\xi_{m,n}^{\bar{m},\bar{n}} = j\langle\xi\rangle_{m,n}^{\bar{m},\bar{n}} = \Im\{\xi_{m,n}^{\bar{m},\bar{n}}\}$. This is popularly known as the real field orthogonality condition. Let $h[l], 0 \leq l \leq L_h - 1$, denote an L_h-tap channel impulse response (CIR) of fading channel. The received FBMC signal $r[k]$ is

$$r[k] = s[k] * h[k] + \eta[k], \tag{2}$$

where $\eta[k]$ is the i.i.d. additive white noise distributed as $\mathcal{CN}\left(0, \sigma_\eta^2\right)$ and the operation '$*$' denotes the convolution. The demodulated signal at the FT index $(\bar{m},\bar{n})$ is obtained as $y_{\bar{m},\bar{n}} = \sum_{k=-\infty}^{+\infty} r[k]\chi_{\bar{m},\bar{n}}^*[k]$. Using (1) and (2), we arrive at

$$y_{\bar{m},\bar{n}} = \sum_{l=0}^{L_h-1} h[l] \sum_{k=-\infty}^{+\infty}\sum_{m=0}^{N-1}\sum_{n\in\mathbb{Z}} d_{m,n}e^{-j2\pi ml/N} p\left[k - \bar{n}\frac{N}{2}\right] p\left[k - l - n\frac{N}{2}\right] e^{j(\phi_{m,n}-\phi_{\bar{m},\bar{n}})} e^{j2\pi(m-\bar{m})k/N} + \eta_{\bar{m},\bar{n}}, \tag{3}$$

where $\eta_{\bar{m},\bar{n}} = \sum_{k=-\infty}^{+\infty} \eta[k]\chi_{\bar{m},\bar{n}}^*[k]$ represents noise at the output of the demodulator. It can be readily shown that $\eta_{\bar{m},\bar{n}} \sim \mathcal{CN}(0,\sigma_\eta^2)$. If the maximum channel delay spread is sufficiently shorter than the symbol duration T, one can exploit the property [9], [16]

$$p[k - l - nN/2] \approx p[k - nN/2] \text{ for } l \in [1, L_h]. \tag{4}$$

Hence, employing (4), and thanks to the fact that the FBMC prototype filter is both time and frequency localized, the demodulated signal $y_{\bar{m},\bar{n}}$ in (3) can be simplified to:

$$y_{\bar{m},\bar{n}} \approx H_{\bar{m}} b_{\bar{m},\bar{n}} + \eta_{\bar{m},\bar{n}}, \tag{5}$$

where the $\bar{m}$th subcarrier's CFR $H_{\bar{m}}$ is defined as $H_{\bar{m}} = \sum_{l=0}^{L_h-1} h[l]e^{-j2\pi\bar{m}l/N}$. The quantity $b_{\bar{m},\bar{n}}$, popularly known as the virtual symbol [16], obeys $b_{\bar{m},\bar{n}} = d_{\bar{m},\bar{n}} + jI_{\bar{m},\bar{n}}$. The quantity $I_{\bar{m},\bar{n}}$ represents the intrinsic interference, which is expressed as $I_{\bar{m},\bar{n}} = \sum_{(m,n)\in\Omega_{\bar{m},\bar{n}}} d_{m,n}\langle\xi\rangle_{m,n}^{\bar{m},\bar{n}}$. The set $\Omega_{\bar{m},\bar{n}}$ above is comprised of the symbols in the neighborhood of FT point $(\bar{m},\bar{n})$. Since the OQAM symbols are i.i.d. with power P_d, from [9], we obtain $\mathbb{E}\left[|b_{\bar{m},\bar{n}}^t|^2\right] = \mathbb{E}[|d_{\bar{m},\bar{n}}^t|^2] + \mathbb{E}[|I_{\bar{m},\bar{n}}^t|^2] \approx 2P_d$. The FBMC models in (3) and (5) are popularly refereed as the TD and IAM models, respectively. It can be readily seen that data detection in the IAM model in (5) can be performed using a simple one-tap equalizer. One the other hand, data detection in the TD in (3) requires multi-tap equalization, which makes it computationally intensive [39]. Thus, when the channel has relatively modest-frequency selectivity, one can use IAM-based sparse channel estimation followed by one-tap channel equalization for data detection. When the channel is highly frequency-selective, it is advisable to use TD based sparse channel estimation followed by multi-tap channel equalization for data detection.

2) MIMO-FBMC System: We consider a MIMO-FBMC system communicating using N_r receive antennas (RAs), N_t transmit antennas (TAs) and N subcarriers. For the tth TA, let a complex QAM symbol at the FT index (m,n) be denoted by $c_{m,n}^t$. The OQAM symbols $d_{m,2n}^t$ and $d_{m,2n+1}^t$ are drawn from the real and imaginary parts of the QAM symbols $c_{m,n}^t$, and are assumed to be i.i.d. with power P_d such that $\mathbb{E}\left[d_{m,n}^t \left(d_{m,n}^t\right)^*\right] = P_d$. This implies that $E\left[c_{m,n}^t \left(c_{m,n}^t\right)^*\right] = 2P_d$. Following the procedure described in the previous section, the demodulated signal of the rth RA at FT index $(\bar{m},\bar{n})$, similar to (3), can be expressed as

$$y_{\bar{m},\bar{n}}^r = \sum_{t=1}^{N_t}\sum_{l=0}^{L_h-1} h^{r,t}[l] \sum_{k=-\infty}^{+\infty}\sum_{m=0}^{N-1}\sum_{n\in\mathbb{Z}} d_{m,n}^t p[k - l - nN/2] \times p[k - \bar{n}N/2]e^{-j2\pi ml/N}e^{j(\phi_{m,n}-\phi_{\bar{m},\bar{n}})} \times e^{j2\pi(m-\bar{m})k/N} + \eta_{\bar{m},\bar{n}}^r, \tag{6}$$

where $\eta_{\bar{m},\bar{n}}^r \sim \mathcal{CN}(0,\sigma_\eta^2)$ represents the demodulated noise at the rth RA. Employing the approximation $p[k-l-kM/2] \approx p[k-nM/2]$ for $l \in [0, L_h]$, the IAM model of MIMO-FBMC systems can be derived similar to (5) as

$$y_{\bar{m},\bar{n}}^r \approx \sum_{t=1}^{N_t} H_{\bar{m}}^{r,t} b_{\bar{m},\bar{n}}^t + \eta_{\bar{m},\bar{n}}^r, \tag{7}$$

where $H_{\bar{m}}^{r,t} = \sum_{l=0}^{L_h-1} h^{r,t}[l]e^{-j2\pi\bar{m}l/N}$ symbolizes the CFR at subcarrier index $\bar{m}$ between the tth TA and rth RA. The complex quantity $b_{\bar{m},\bar{n}}^t$, popularly termed as the virtual symbol of the tth TA, obeys $b_{\bar{m},\bar{n}}^t = d_{\bar{m},\bar{n}}^t + jI_{\bar{m},\bar{n}}^t$. The associated intrinsic interference $I_{\bar{m},\bar{n}}^t$ is expressed as $I_{\bar{m},\bar{n}}^t = \sum_{(m,n)\in\Omega_{\bar{m},\bar{n}}} d_{m,n}^t\langle\xi\rangle_{m,n}^{\bar{m},\bar{n}}$. The relationship $\mathbb{E}\left[|b_{\bar{m},\bar{n}}^t|^2\right] = \mathbb{E}[|d_{\bar{m},\bar{n}}^t|^2] + \mathbb{E}[|I_{\bar{m},\bar{n}}^t|^2] \approx 2P_d$ follows from [9]. Let $\mathbf{y}_{\bar{m},\bar{n}} = [y_{\bar{m},\bar{n}}^1, y_{\bar{m},\bar{n}}^2, \ldots, y_{\bar{m},\bar{n}}^{N_r}]^T \in \mathbb{C}^{N_r\times 1}$ denote the vector of received symbols stacked across all the RAs. Eq. (7) can be written in vectorized form as

$$\mathbf{y}_{\bar{m},\bar{n}} \approx \mathbf{H}_{\bar{m}}\mathbf{b}_{\bar{m},\bar{n}} + \boldsymbol{\eta}_{\bar{m},\bar{n}}, \tag{8}$$

where the matrix $\mathbf{H}_{\bar{m}} \in \mathbb{C}^{N_r\times N_t}$ represents the MIMO-FBMC CFR, whose (r,t)th element is $H_{\bar{m}}^{r,t}$. The quantity $\boldsymbol{\eta}_{\bar{m},\bar{n}} = [\eta_{\bar{m},\bar{n}}^1, \eta_{\bar{m},\bar{n}}^2, \ldots, \eta_{\bar{m},\bar{n}}^{N_r}]^T \in \mathbb{C}^{N_r\times 1}$ symbolizes the noise vector having the covariance matrix $\mathbb{E}\left[\boldsymbol{\eta}_{\bar{m},\bar{n}}\boldsymbol{\eta}_{\bar{m},\bar{n}}^H\right] = \sigma_\eta^2\mathbf{I}_{N_r}$ while $\mathbf{b}_{\bar{m},\bar{n}} = [b_{\bar{m},\bar{n}}^1, b_{\bar{m},\bar{n}}^2, \ldots, b_{\bar{m},\bar{n}}^{N_t}]^T \in \mathbb{C}^{N_t\times 1}$ stacks the virtual symbols and satisfies $\mathbb{E}\left[\mathbf{b}_{\bar{m},\bar{n}}\mathbf{b}_{\bar{m},\bar{n}}^H\right] \approx 2P_d\mathbf{I}_{N_t}$.

3) IAM-based SISO-FBMC Sparse Channel Estimation Model: Fig. 1(a) shows the FBMC frame for sparse CSI estimation in SISO systems. It is comprised of L_0 OQAM symbols per subcarrier. In each frame, the first N_p subcarriers at index $n = 0$ carry the training symbols. Due to the overlap of adjacent time- and frequency-domain FBMC symbols, we insert z zero symbols across both the dimensions to curb interference between the training and data symbols [9], [16]. Typically, $z = 1$ suffices to suppress the interference to a tolerable level, since the pulse shaping filters in FBMC have a sharp cut-off. Evaluating (5) at $n = 0$, we get:

$$y_{\bar{m},0}^{\text{iam}} \approx H_{\bar{m}} b_{\bar{m},0} + \eta_{\bar{m},0}, \tag{9}$$

where $b_{\bar{m},0} = d_{\bar{m},0} + jI_{\bar{m},0}$ is the virtual training symbol at the $\bar{m}$th subcarrier. By utilizing zero symbols across the time and frequency axes, it has been shown in [16] that $I_{\bar{m},0} = \sum_{m=1,m\neq\bar{m}}^{N_p} d_{m,0}\langle\xi\rangle_{m,0}^{\bar{m},0}$. Let $\mathbf{h} \in \mathbb{C}^{L_h\times 1}$ denote

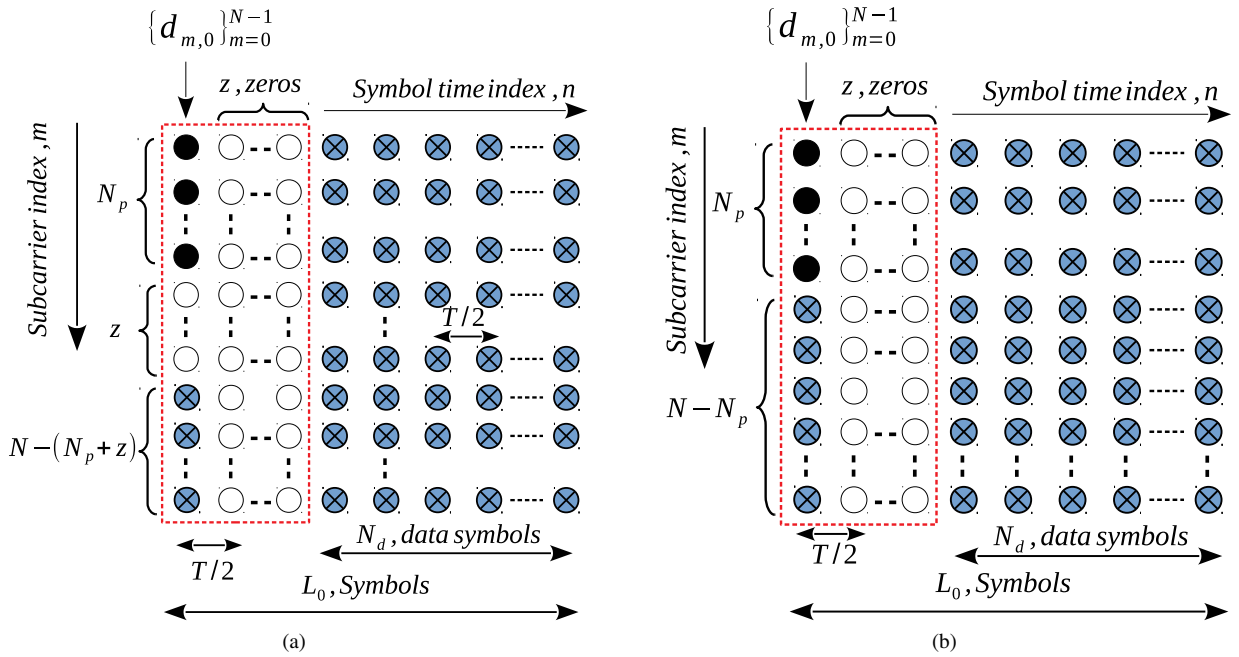


Fig. 1: FBMC frame for sparse CSI estimation in SISO systems: (a) IAM model; and (b) TD model. The notations ⊗, ◯ and ● represent data, zero and training symbols, respectively.

the CIR vector as $\mathbf{h} = [h[0], h[1], \ldots, h[L_h - 1]]^T$ and $\mathbf{F}_{\text{DFT}} \in \mathbb{C}^{N\times N}$ represent the discrete Fourier transform (DFT) matrix and $\mathbf{F}^{\text{tr}}_{\text{DFT}} \in \mathbb{C}^{N\times L_h}$ denote the truncated-DFT matrix defined as $\mathbf{F}^{\text{tr}}_{\text{DFT}} = \mathbf{F}_{\text{DFT}}(:, 1 : L_h)$. Let $\mathbf{F} \in \mathbb{C}^{N_p\times L_h}$ denote the matrix obtained by extracting the first N_p rows of $\mathbf{F}^{\text{tr}}_{\text{DFT}}$. For convenience, the model in (9) can be written in compact form by stacking the outputs across the N_p pilot subcarriers as

$$\mathbf{y}_0^{\text{iam}} = \mathbf{B}_0\mathbf{F}\mathbf{h} + \boldsymbol{\eta}_0, \tag{10}$$

where $\mathbf{y}_0^{\text{iam}} = [y_{0,0}^{\text{iam}}, y_{1,0}^{\text{iam}}, \ldots, y_{N_p-1,0}^{\text{iam}}]^T \in \mathbb{C}^{N_p\times 1}$ and $\mathbf{B}_0 = \text{diag}(b_{0,0}, b_{1,0}, \ldots, b_{N_p-1,0}) \in \mathbb{C}^{N_p\times N_p}$ are the received training vector and diagonal virtual training matrix, respectively. Finally, $\boldsymbol{\eta}_0 = [\eta_{1,0}, \eta_{2,0}, \ldots, \eta_{N_p-1,0}]^T \in \mathbb{C}^{N_p\times 1}$ is the stacked noise vector.

4) IAM-based MIMO-FBMC Sparse Channel Estimation Model: Let each TA transmit M training frames on each subcarrier, where each training frame comprises symbols as shown with dotted lines in Fig. 1(a). In each frame, the first N_p subcarriers at time instants $0 \le n \le (M-1)(1+z)$, carry the pilot symbols used for CSI estimation. This is followed by the transmission of N_d data symbols on each subcarrier. Evaluating (8) at the pilot positions $n = i(1+z)$, for $i = 0, \ldots, M-1$, and over the N_p pilot subcarriers, we get

$$\mathbf{Y}_{\bar{m}}^{\text{iam}} = \mathbf{H}_{\bar{m}}\mathbf{B}_{\bar{m}} + \boldsymbol{\eta}_{\bar{m}}, \tag{11}$$

where $\mathbf{Y}_{\bar{m}}^{\text{iam}} = [\mathbf{y}_{\bar{m},0}, \mathbf{y}_{\bar{m},(1+z)}, \ldots, \mathbf{y}_{\bar{m},(M-1)(1+z)}] \in \mathbb{C}^{N_r\times M}$ represents the observation matrix corresponding to all the pilot positions and the corresponding noise is given by $\boldsymbol{\eta}_{\bar{m}} = [\boldsymbol{\eta}_{\bar{m},0}, \boldsymbol{\eta}_{\bar{m},(1+z)}, \ldots, \boldsymbol{\eta}_{\bar{m},(M-1)(1+z)}] \in \mathbb{C}^{N_r\times M}$. The matrix $\mathbf{B}_{\bar{m}}$, comprising the virtual training symbols at the $\bar{m}$th pilot subcarrier, is obtained as $\mathbf{B}_{\bar{m}} = [\mathbf{b}_{\bar{m},0}, \mathbf{b}_{\bar{m},(1+z)}, \ldots, \mathbf{b}_{\bar{m},(M-1)(1+z)}] \in \mathbb{C}^{N_t\times M}$, where the tth element of $\mathbf{b}_{\bar{m},i(1+z)} \in \mathbb{C}^{N_t\times 1}$ is expressed as $b^t_{\bar{m},i(1+z)} = d^t_{\bar{m},i(1+z)} + jI^t_{\bar{m},i(1+z)}$. It follows from [16] that the interference term $I^t_{\bar{m},i(1+z)}$ can be formulated as

$$I^t_{\bar{m},i(1+z)} = \sum_{m\neq\bar{m}} d^t_{m,i(1+z)}\langle\xi\rangle^{\bar{m},0}_{m,0}. \tag{12}$$

Let $\mathbf{H}[l] \in \mathbb{C}^{N_r\times N_t}$ denote the lth MIMO channel tap with its (r,t)th element given by $h^{r,t}[l]$, and the concatenated channel matrix $\mathbf{H} \in \mathbb{C}^{N_r\times N_tL_h}$ be defined as $\mathbf{H} = \big[\mathbf{H}[0], \mathbf{H}[1], \ldots, \mathbf{H}[L_h-1]\big]$. The MIMO CFR matrix $\mathbf{H}_{\bar{m}}$ corresponding to the $\bar{m}$th subcarrier can be recast as $\mathbf{H}_{\bar{m}} = \mathbf{H}\,(\mathbf{f}_{\bar{m}} \otimes \mathbf{I}_{N_t})$, where the vector $\mathbf{f}_{\bar{m}} = [1, e^{-j2\pi\bar{m}/N}, \ldots, e^{-j2\pi\bar{m}(L_h-1)/N}]^T \in \mathbb{C}^{L_h\times 1}$ is obtained by extracting the $\bar{m}$th row of the matrix $\mathbf{F}$. With the above identity, the IAM model of CSI estimation of a MIMO-FBMC system in (11) can be expressed as

$$\mathbf{Y}_{\bar{m}}^{\text{iam}} = \mathbf{H}\big(\mathbf{f}_{\bar{m}} \otimes \mathbf{I}_{N_t}\big)\mathbf{B}_{\bar{m}} + \boldsymbol{\eta}_{\bar{m}}. \tag{13}$$

Vectorizing (13) and invoking the properties of the matrix Kronecker product [40], one obtains

$$\underbrace{\text{vec}(\mathbf{Y}_{\bar{m}}^{\text{iam}})}_{\tilde{\mathbf{y}}_{\bar{m}}^{\text{iam}}} = \underbrace{\big(\tilde{\mathbf{B}}_{\bar{m}}^T \otimes \mathbf{I}_{N_r}\big)}_{\boldsymbol{\Phi}_{\bar{m}}}\text{vec}(\mathbf{H}) + \underbrace{\text{vec}(\boldsymbol{\eta}_{\bar{m}})}_{\tilde{\boldsymbol{\eta}}_{\bar{m}}}, \tag{14}$$

where the matrix $\tilde{\mathbf{B}}_{\bar{m}}$ obeys $\tilde{\mathbf{B}}_{\bar{m}} = \big(\mathbf{f}_{\bar{m}} \otimes \mathbf{I}_{N_t}\big)\mathbf{B}_{\bar{m}}$. Upon concatenating $\tilde{\mathbf{y}}_{\bar{m}}^{\text{iam}}$ across the pilot subcarriers $0 \le \bar{m} \le N_p - 1$, the IAM-based channel estimation model is obtained as

$$\mathbf{y}^{\text{iam}} = \boldsymbol{\Phi}\bar{\mathbf{h}} + \boldsymbol{\eta}, \tag{15}$$

where the output vector is $\mathbf{y}^{\text{iam}} = \left[(\tilde{\mathbf{y}}_0^{\text{iam}})^T, (\tilde{\mathbf{y}}_1^{\text{iam}})^T, \ldots, (\tilde{\mathbf{y}}_{N_p-1}^{\text{iam}})^T\right]^T \in \mathbb{C}^{N_p N_r M \times 1}$, the MIMO CIR vector is $\bar{\mathbf{h}} = \text{vec}(\mathbf{H}) \in \mathbb{C}^{N_r N_t L_h \times 1}$ and the noise vector is $\boldsymbol{\eta} = \left[\tilde{\boldsymbol{\eta}}_0^T, \tilde{\boldsymbol{\eta}}_1^T, \ldots, \tilde{\boldsymbol{\eta}}_{N_p-1}^T\right]^T \in \mathbb{C}^{N_p N_r M \times 1}$. The dictionary matrix $\boldsymbol{\Phi} = \left[\boldsymbol{\Phi}_0^T, \boldsymbol{\Phi}_1^T, \ldots, \boldsymbol{\Phi}_{N_p-1}^T\right]^T \in \mathbb{C}^{N_p N_r M \times N_r N_t L_h}$. Note that one of the key differences between OFDM and FBMC systems is that the former does not experience intrinsic interference, because unlike FBMC, the time domain symbols in OFDM do not overlap with each other and the subcarriers are orthogonal. Therefore, in OFDM systems the matrix $\mathbf{B}_0$ in (10) and the above matrix $\boldsymbol{\Phi}$, depend on the training symbols only. Consequently, the resultant intrinsic interference in $\mathbf{B}_0$ and $\boldsymbol{\Phi}$ in FBMC play a crucial role, and also make channel estimation significantly different from that of its OFDM counterpart. For calculating the intrinsic interference, FBMC system, as shown in Fig. 1, requires the placement of z zero symbols between the adjacent training symbols to avoid ISI between them. One also has to find the optimal number of zero symbols in terms of the overhead required for channel estimation while avoiding any undue BER degradation. This work addresses this issue with the aid of the BCRLBs. It is observed in the simulations section that the proposed IAM and TD schemes achieve their respective BCRLBs using $z = 3$. Thus, $z = 3$ is ideally suited to avoid under-design/ over-design of an FBMC system. The approximation in (4), which leads to frequency flat models across the subcarriers for a mildly frequency-selective channel, plays a key role in the development of the IAM-based channel estimation model above. However, for a highly frequency-selective channel, where the assumption in (4) is not satisfied, the performance of channel estimation schemes based on the IAM model exhibits severe degradation. The TD-based channel estimation model developed next does not require a frequency-flat channel at the subcarrier level, and thus overcomes this limitation.

5) TD-based SISO-FBMC Sparse Channel Estimation Model: The frame structure of a TD-based model is shown in Fig. 1(b), which is similar to that of the IAM model of Fig. 1(a), except for the following differences. In contrast to the latter, the former does not require the insertion of zero symbols across the subcarrier axis at the time instant $n = 0$. This is due to the fact that unlike the IAM model, the operation of the TD-based channel estimation scheme does not require computation of the intrinsic interference. For the training subcarriers at time instant $n = 0$, the output training symbol at the FT index $(\bar{m}, 0)$ can be expressed with the aid of (3) as $y_{\bar{m},0} = \sum_{l=0}^{L_h-1} D_{\bar{m},l} h[l] + \eta_{\bar{m},0}$, where the quantity $D_{\bar{m},l}$ is formulated as

$$D_{\bar{m},l} = \sum_{k=l}^{L_p-1} \sum_{m=0}^{N_p-1} d_{m,0} p[k] p[k-l] e^{j(\phi_{m,0} - \phi_{\bar{m},0})} \times e^{j2\pi(m-\bar{m})k/N} e^{-j2\pi ml/N}. \quad (16)$$

For mathematical ease, the expression for $y_{\bar{m},0}$ can be succinctly written in vector form as [18]

$$\mathbf{y}_0^{\text{td}} = \mathbf{D}\mathbf{h} + \boldsymbol{\eta}_0, \quad (17)$$

where $\mathbf{y}_0^{\text{td}} = [y_{0,0}, y_{1,0}, \ldots, y_{N_p-1,0}]^T \in \mathbb{C}^{N_p \times 1}$ is once again the stacked vector of received outputs across the N_p training subcarriers and $\mathbf{D} \in \mathbb{C}^{N_p \times L_h}$ is a matrix whose $(\bar{m}, l)$th element is given by $D_{\bar{m},l}$ in (16). The main difference between FBMC and OFDM with respect to the TD model is that the calculation of the matrix $\mathbf{D}$ for the former requires the insertion of z zeros, as shown in Fig. 1. These z zeros reduce the overlap between the training and data symbols. Therefore, one also has to find the optimal number of zero symbols inserted between pilot and data symbols. As explained below (15), this issue is addressed in the simulation section with the help of BCRLBs derived in this work.

6) TD-based MIMO-FBMC Sparse Channel Estimation Model: Each TA employs N_p training subcarriers in the dotted box of the frame in Fig. 1(b) for CSI estimation. From (6), the pilot output at FT index $(\bar{m}, 0)$ of the rth RA is obtained as $y_{\bar{m},0}^r = \sum_{t=1}^{N_t} \sum_{l=0}^{L_h-1} h^{r,t}[l] D_{\bar{m},l}^t + \eta_{\bar{m},0}^r$, where the quantity $D_{\bar{m},l}^t$ is given by

$$D_{\bar{m},l}^t = \sum_{m=0}^{N-1} d_{m,0}^t e^{j(\phi_{m,0} - \phi_{\bar{m},0})} e^{-j2\pi ml/N} \times \sum_{k=l}^{L_p-1} p[k] p[k-l] e^{j2\pi(m-\bar{m})k/N}. \quad (18)$$

Let $\mathbf{h}^{r,t} = [h^{r,t}[0], h^{r,t}[1], \ldots, h^{r,t}[L_h - 1]]^T \in \mathbb{C}^{L_h \times 1}$ denote the CIR vector between the tth TA and rth RA. The TD-based MIMO-FBMC model can be succinctly cast in vectorized form as

$$\mathbf{y}_0^r = \sum_{t=1}^{N_t} \mathbf{D}^t \mathbf{h}^{r,t} + \boldsymbol{\eta}_0^r, \quad (19)$$

where $\mathbf{y}_0^r = [y_{0,0}^r, y_{1,0}^r, \ldots, y_{N_p-1,0}^r]^T \in \mathbb{C}^{N_p \times 1}$ denotes the vector of training symbols received by the rth RA and $\boldsymbol{\eta}_0^r = [\eta_{0,0}^r, \eta_{1,0}^r, \ldots, \eta_{N_p-1,0}^r]^T \in \mathbb{C}^{N_p \times 1}$ is the associated noise. The matrix $\mathbf{D}^t \in \mathbb{C}^{N_p \times L_h}$ is computed at the receiver using the training symbols $\{d_{m,0}^t\}_{m=0}^{N_p-1}$ and the pulse shaping filter $p[k]$, and its $(\bar{m}, l)$th element is given by $D_{\bar{m},l}^t$. Then we can rewrite (19) as

$$\mathbf{y}_0^r = \boldsymbol{\mathcal{D}}\mathbf{h}^r + \boldsymbol{\eta}_0^r, \quad (20)$$

where the matrix $\boldsymbol{\mathcal{D}} \in \mathbb{C}^{N_p \times N_t L_h} = [\mathbf{D}^1, \mathbf{D}^2, \ldots, \mathbf{D}^{N_t}]$, the channel vector $\mathbf{h}^r \in \mathbb{C}^{N_t L_h \times 1} = [(\mathbf{h}^{r,1})^T, (\mathbf{h}^{r,2})^T, \ldots, (\mathbf{h}^{r,N_t})^T]^T$ comprises the CIRs from all the TAs to the rth RA. Stacking the vectors $\mathbf{y}_0^r, 1 \le r \le N_r$, the TD CSI estimation model is formulated as

$$\bar{\mathbf{y}}_0^{\text{td}} = \bar{\boldsymbol{\mathcal{D}}}\tilde{\mathbf{h}} + \bar{\boldsymbol{\eta}}_0, \quad (21)$$

where $\bar{\mathbf{y}}_0^{\text{td}} \in \mathbb{C}^{N_p N_r \times 1} = [(\mathbf{y}_0^1)^T, (\mathbf{y}_0^2)^T, \ldots, (\mathbf{y}_0^{N_r})^T]^T$ consists of the received training vectors and $\boldsymbol{\eta}_0 \in \mathbb{C}^{N_p N_r \times 1}$ is the corresponding noise vector, and the block diagonal matrix $\bar{\boldsymbol{\mathcal{D}}} \in \mathbb{C}^{N_p N_r \times N_t N_r L_h} = \text{blkdiag}[\boldsymbol{\mathcal{D}}, \boldsymbol{\mathcal{D}}, \ldots, \boldsymbol{\mathcal{D}}]$ is the dictionary matrix. The quantity $\tilde{\mathbf{h}} \in \mathbb{C}^{N_r N_t L_h \times 1} = [(\mathbf{h}^1)^T, (\mathbf{h}^2)^T, \ldots, (\mathbf{h}^{N_r})^T]^T$ denotes the stacked CIR vector.

Commencing with the model in (17), the conventional LS and MMSE estimates of the CIR vector $\mathbf{h}$ for the TD model

can be obtained in a straightforward fashion as

$$\widehat{\mathbf{h}}^{\text{LS}} = \left(\mathbf{D}^H\mathbf{D}\right)^{-1}\mathbf{D}^H\mathbf{y}_0^{\text{td}} \text{ and}$$
$$\widehat{\mathbf{h}}^{\text{MMSE}} = \left(\mathbf{R}_h^{-1} + \mathbf{D}^H\mathbf{R}_\eta^{-1}\mathbf{D}\right)^{-1}\mathbf{D}^H\mathbf{R}_\eta^{-1}\mathbf{y}_0^{\text{td}},$$

where the matrices $\mathbf{R}_h = \mathbb{E}\left[\mathbf{h}\mathbf{h}^H\right] \in \mathbb{C}^{L_h\times L_h}$ and $\mathbf{R}_\eta = \mathbb{E}\left[\boldsymbol{\eta}_0\boldsymbol{\eta}_0^H\right] \in \mathbb{C}^{N_p\times N_p}$ denote the covariance matrices of the CIR vector $\mathbf{h}$ and the noise $\boldsymbol{\eta}_0$, respectively. More details and in-depth analysis of such l_2-norm minimization based conventional estimators can be found in standard texts such as [41]. A significant drawback of the LS approach is that it does not exploit the sparsity of the CIR vector $\mathbf{h}$ arising due to the sparse multipath delay profile of a typical wireless channel [23], [24], which implies that the number of non-zero taps L_s is in practice much smaller than the total number L_h of the CIR taps. Naturally, exploiting this sparsity can lead to performance benefits together with a reduced pilot overhead. As a result, the LS approach requires an overdetermined system [41], i.e., $N_p \geq L_h$, which leads to higher pilot overheads and reduced spectral efficiency. On the other hand, the MMSE technique can exploit this sparsity via the prior covariance matrix $\mathbf{R}_h$ of the CIR vector $\mathbf{h}$. However, $\mathbf{R}_h$ is unknown in practice. This motivates the development of sparse channel estimation techniques, which iteratively estimate the parameters corresponding to the prior distribution of the CIR vector $\mathbf{h}$ from the pilot output $\mathbf{y}_0^{\text{td}}$. This forms the focus of this paper.

III. Quasi-Static Sparse CSI Estimation in FBMC Systems

Channel estimation in FBMC systems, as shown in Fig. 1, requires the placement of zero symbols between adjacent training symbols to avoid inter-symbol-interference (ISI) due to the overlapping nature of the time-domain FBMC symbols. This, in turn, requires careful examination of the intrinsic interference at the receiver to evaluate the resultant virtual training symbols. Moreover, one also has to address the optimal number of zero symbols required to avoid system underdesign/ overdesign of FBMC system in terms of overhead requirement for channel estimation. Thus, while addressing the channel estimation problem in FBMC systems, it is not always possible to extend the existing schemes or corresponding analysis for OFDM systems.

1) TD-based SISO-FBMC Sparse Channel Estimation: The SBL, which has enjoyed popularity for sparse CSI estimation owing to its superior sparse signal recovery, employs a Bayesian philosophy for sparse estimation. We start with assigning the following parameterized Gaussian prior to the sparse CIR vector $\mathbf{h}$ [29]

$$f\left(\mathbf{h};\boldsymbol{\Gamma}\right) = \prod_{i=0}^{L_h-1}\left(\pi\gamma_i\right)^{-1}\exp\left(-\frac{\left|\mathbf{h}\left(i\right)\right|^2}{\gamma_i}\right), \qquad (22)$$

where $\gamma_i \geq 0, 0 \leq i \leq L_h-1$, denotes the hyperparameter that signifies the prior variance of the ith element of the CIR vector $\mathbf{h}$, while $\boldsymbol{\Gamma} \in \mathbb{R}_+^{L_h\times L_h}$ is a diagonal matrix comprising the hyperparameters γ_i on its principal diagonal. Upon employing the parameterized Gaussian prior above, the MMSE estimate $\boldsymbol{\mu} \in \mathbb{C}^{L_h\times 1}$ of the channel tap vector $\mathbf{h}$ and the corresponding error covariance matrix $\boldsymbol{\Sigma} \in \mathbb{C}^{L_h\times L_h}$ can be expressed as [41]

$$\boldsymbol{\mu} = \boldsymbol{\Sigma}\mathbf{D}^H\mathbf{R}_\eta^{-1}\mathbf{y}_0^{\text{td}} \text{ and } \boldsymbol{\Sigma} = \left(\boldsymbol{\Gamma}^{-1} + \mathbf{D}^H\mathbf{R}_\eta^{-1}\mathbf{D}\right)^{-1}. \quad (23)$$

It can be readily observed that the MMSE estimate $\boldsymbol{\mu}$ depends on the hyperparameter matrix $\boldsymbol{\Gamma}$. Therefore, the estimation of the CIR vector in a SISO-FBMC system reduces to that of the hyperparameters γ_i. We can readily estimate these hyperparameters by maximizing the Bayesian evidence $f\left(\mathbf{y}_0^{\text{td}};\boldsymbol{\Gamma}\right)$ expressed as $f\left(\mathbf{y}_0^{\text{td}};\boldsymbol{\Gamma}\right) = \left(1/(\pi)^{N_p}\det\left(\boldsymbol{\Sigma}_y\right)\right)\exp\left(-\left(\mathbf{y}_0^{\text{td}}\right)^H\boldsymbol{\Sigma}_y^{-1}\mathbf{y}_0^{\text{td}}\right)$ [41], where the matrix $\boldsymbol{\Sigma}_y = \left(\mathbf{D}\boldsymbol{\Gamma}\mathbf{D}^H + \mathbf{R}_\eta\right) \in \mathbb{C}^{N_p\times N_p}$ denotes the covariance of the received pilot vector $\mathbf{y}_0^{\text{td}}$. As described in [29], the log-likelihood $\log\left[f\left(\mathbf{y}_0^{\text{td}};\boldsymbol{\Gamma}\right)\right]$ of the optimization objective is non-concave and therefore its direct maximization with respect to the hyperparameters becomes intractable. Hence, the SBL-based channel estimation technique proposed in our treatise employs an expectation-maximization (EM) framework as follows.

Consider the received pilot output $\mathbf{y}_0^{\text{td}}$ as the known variable and the CIR vector $\mathbf{h}$ as the latent/ unknown variable. Therefore, the complete data set employed in the EM framework [41] is defined as $\left\{\mathbf{y}_0^{\text{td}},\mathbf{h}\right\}$. Let $\widehat{\boldsymbol{\Gamma}}^{(p-1)}$ represent the estimate of the matrix $\boldsymbol{\Gamma}$ in the $(p-1)$st iteration of the EM framework. In the pth iteration, the expectation step (E-step) computes the log likelihood $\mathcal{L}\left(\boldsymbol{\Gamma} \mid \widehat{\boldsymbol{\Gamma}}^{(p-1)}\right)$ of the complete information set as

$$\begin{aligned}\mathcal{L}\left(\boldsymbol{\Gamma} \mid \widehat{\boldsymbol{\Gamma}}^{(p-1)}\right) &= \mathbb{E}_{\mathbf{h}|\mathbf{y}_0^{\text{td}};\widehat{\boldsymbol{\Gamma}}^{(p-1)}}\left\{\log\left[f\left(\mathbf{y}_0^{\text{td}},\mathbf{h};\boldsymbol{\Gamma}\right)\right]\right\}\\ &= \mathbb{E}\left\{\log\left[f\left(\mathbf{y}_0^{\text{td}} \mid \mathbf{h}\right)\right] + \log\left[f\left(\mathbf{h};\boldsymbol{\Gamma}\right)\right]\right\}.\end{aligned}$$

Subsequently, the associated maximization step (M-step) of the EM procedure maximizes the log likelihood $\mathcal{L}\left(\boldsymbol{\Gamma} \mid \widehat{\boldsymbol{\Gamma}}^{(p-1)}\right)$ to estimate $\widehat{\boldsymbol{\Gamma}}^{(p)}$ as

$$\widehat{\boldsymbol{\Gamma}}^{(p)} = \arg\max_{\boldsymbol{\Gamma}}\ \mathcal{L}\left(\boldsymbol{\Gamma} \mid \widehat{\boldsymbol{\Gamma}}^{(p-1)}\right). \qquad (24)$$

The conditional log-likelihood $\log\left[f\left(\mathbf{y}_0^{\text{td}} \mid \mathbf{h}\right)\right] = \mathcal{C}_1 - \left(\mathbf{y}_0^{\text{td}} - \mathbf{D}\mathbf{h}\right)^H\mathbf{R}_\eta^{-1}\left(\mathbf{y}_0^{\text{td}} - \mathbf{D}\mathbf{h}\right)$ with the constant $\mathcal{C}_1 = -N_p\log(\pi) - \log\left[\det\left(\mathbf{R}_\eta\right)\right]$ can be seen to be independent of the hyperparameter matrix $\boldsymbol{\Gamma}$. Therefore, the M-step of hyperparameter estimation in (24) can be simplified to

$$\begin{aligned}\widehat{\boldsymbol{\Gamma}}^{(p)} &= \arg\max_{\boldsymbol{\Gamma}}\ \mathbb{E}\left\{\log\left[f\left(\mathbf{h};\boldsymbol{\Gamma}\right)\right]\right\}\\ &\equiv \arg\max_{\boldsymbol{\Gamma}}\ \mathbb{E}\left\{\sum_{i=1}^{L_h} -\log\left(\gamma_i\right) - \frac{1}{\gamma_i}\left|\mathbf{h}\left(i\right)\right|^2\right\}.\end{aligned}$$

By differentiating the above expression with respect to the hyperparameter γ_i and setting it equal to zero, we obtain the hyperparameter update equation for the pth EM iteration as $\widehat{\gamma}_i^{(p)} = \mathbb{E}_{\mathbf{h}|\mathbf{y}_0^{\text{td}};\widehat{\boldsymbol{\Gamma}}^{(p-1)}}\left\{\left|\mathbf{h}\left(i\right)\right|^2\right\}$. Next, to evaluate $\mathbb{E}_{\mathbf{h}|\mathbf{y}_0^{\text{td}};\widehat{\boldsymbol{\Gamma}}^{(p-1)}}\{\cdot\}$, we have to employ the *a posteriori* probability density function (pdf) $f\left(\mathbf{h} \mid \mathbf{y}_0^{\text{td}};\widehat{\boldsymbol{\Gamma}}^{(p-1)}\right) =$

$\mathcal{CN}\left(\boldsymbol{\mu}^{(p)}, \boldsymbol{\Sigma}^{(p)}\right)$, where the *a posteriori* mean $\boldsymbol{\mu}^{(p)}$ and the covariance matrix $\boldsymbol{\Sigma}^{(p)}$ can be determined by substituting $\boldsymbol{\Gamma} = \widehat{\boldsymbol{\Gamma}}^{(p-1)}$ into Eq. (23). Using this, the update equation of $\widehat{\gamma}_i^{(p)}$ for the pth EM iteration is expressed as

$$\widehat{\gamma}_i^{(p)} = \left|\boldsymbol{\mu}^{(p)}(i)\right|^2 + \boldsymbol{\Sigma}^{(p)}(i,i). \tag{25}$$

Next, the estimate of the hyperparameter matrix is obtained as $\widehat{\boldsymbol{\Gamma}}^{(p)} = \mathrm{diag}\big(\widehat{\gamma}_0^{(p)}, \widehat{\gamma}_1^{(p)}, \ldots, \widehat{\gamma}_{L_h-1}^{(p)}\big)$. The E-step and M-step derived above are repeated until the Frobenius norm of the difference of the estimated hyperparameter matrix falls below a threshold ϵ_1, i.e., $\|\widehat{\boldsymbol{\Gamma}}^{(p)} - \widehat{\boldsymbol{\Gamma}}^{(p-1)}\|_F < \epsilon_1$ or $N_{\max}$ EM-iterations are completed. Upon convergence, the SBL-based estimate of the CIR vector $\widehat{\mathbf{h}}^{\text{TD-SBL}}$ for the TD SISO-FBMC channel estimation model of (17) is given by $\widehat{\mathbf{h}}^{\text{TD-SBL}} = \boldsymbol{\mu}$. The concise step-by-step procedure of the proposed TD-SBL CSI estimation of SISO-FBMC systems is given in Algorithm-2 of our technical report [42]. Furthermore, the IAM-based SBL estimate $\widehat{\mathbf{h}}^{\text{IAM-SBL}}$ of the IAM model in (10) can be derived along similar lines.

2) IAM-based MIMO-FBMC Sparse Channel Estimation: It is interesting to note that the CIR vector $\bar{\mathbf{h}}$, as shown in Fig. 1(b) of the technical report [42], is *block-sparse* in nature, where each block of $N_t N_r$ consecutive entries has an identical sparsity profile. This block sparsity is exploited by the IAM-Block SBL (IAM-BSBL) technique described next for channel estimation.

The IAM-BSBL technique assigns the following parameterized Gaussian prior to the lth, $\forall\ 0 \le l \le L_h - 1$, vectorized channel tap denoted by $\mathbf{h}_l \triangleq \mathrm{vec}(\mathbf{H}[l])$:

$$f(\mathbf{h}_l; \gamma_l) = \prod_{i=1}^{N_r N_t} (\pi\gamma_l)^{-1} \exp\left(-\frac{\left|\mathbf{h}_l(i)\right|^2}{\gamma_l}\right). \tag{26}$$

Hence, the prior assigned to the vectorized channel $\bar{\mathbf{h}}$ is given by $f\left(\bar{\mathbf{h}}; \boldsymbol{\Gamma}\right) = \prod_{l=0}^{L_h-1} f(\mathbf{h}_l; \gamma_l)$. Employing the EM procedure described in Section III, the *a posteriori* pdf $f\left(\bar{\mathbf{h}} \mid \mathbf{y}^{\text{iam}}; \widehat{\boldsymbol{\Gamma}}^{(p-1)}\right)$ in the pth EM iteration can be formulated as $f\left(\bar{\mathbf{h}} \mid \mathbf{y}^{\text{iam}}; \widehat{\boldsymbol{\Gamma}}^{(p-1)}\right) = \mathcal{CN}\left(\bar{\boldsymbol{\mu}}^{(p)}, \bar{\boldsymbol{\Sigma}}^{(p)}\right)$, where the *a posteriori* mean $\bar{\boldsymbol{\mu}}^{(p)}$ and covariance matrix $\bar{\boldsymbol{\Sigma}}^{(p)}$ can be expressed as

$$\bar{\boldsymbol{\mu}}^{(p)} = \bar{\boldsymbol{\Sigma}}^{(p)} \boldsymbol{\Phi}^H \bar{\mathbf{R}}_\eta^{-1} \mathbf{y}^{\text{iam}} \ \text{ and}$$
$$\bar{\boldsymbol{\Sigma}}^{(p)} = \left(\boldsymbol{\Phi}^H \bar{\mathbf{R}}_\eta^{-1} \boldsymbol{\Phi} + \left(\widehat{\boldsymbol{\Gamma}}^{(p-1)} \otimes \mathbf{I}_{N_r N_t}\right)^{-1}\right)^{-1},$$

where $\bar{\mathbf{R}}_\eta = \mathbb{E}\left[\boldsymbol{\eta}\boldsymbol{\eta}^H\right] \in \mathbb{C}^{N_p N_r M \times N_p N_r M}$ denotes the noise covariance. Furthermore, the update equation of $\widehat{\gamma}_i^{(p)}$ in the pth EM iteration can be expressed as

$$\widehat{\gamma}_i^{(p)} = \frac{1}{N_t N_r} \sum_{\tilde{l}=iN_r N_t+1}^{iN_r N_t} \left|\bar{\boldsymbol{\mu}}^{(p)}(\tilde{l})\right|^2 + \bar{\boldsymbol{\Sigma}}^{(p)}(\tilde{l},\tilde{l}). \tag{27}$$

Upon convergence, the IAM-BSBL based estimate $\bar{\mathbf{h}}^{\text{IAM-BSBL}}$ is given by $\bar{\mathbf{h}}^{\text{IAM-BSBL}} = \bar{\boldsymbol{\mu}}$.

3) TD-Based MIMO-FBMC Sparse Channel Estimation: It is interesting to note that the CIR vector $\tilde{\mathbf{h}}$ of the TD channel estimation model formulated above exhibits a *group-sparse* structure. This is justified as follows. Let $\tilde{\mathbf{h}}^i \in \mathbb{C}^{N_r N_t \times 1}$ denote the ith group, $0 \le i \le L_h - 1$, of the CIR vector $\tilde{\mathbf{h}}$, which is defined as

$$\tilde{\mathbf{h}}^i = [\tilde{\mathbf{h}}(i+1), \tilde{\mathbf{h}}(L_h+i+1), \ldots, \tilde{\mathbf{h}}((N_r N_t - 1)L_h + i + 1)]^T. \tag{28}$$

As demonstrated in Fig. 1(b) of our technical report [42], the $N_r N_t$ elements within the group $\tilde{\mathbf{h}}^i$ of vector $\tilde{\mathbf{h}}$ are expected to be either all-zero or non-zero. The TD-GSBL framework is developed next, which exploits this *group-sparsity* for improved estimation of the CIR vector $\tilde{\mathbf{h}}$. This begins by assigning the Gaussian prior parameterized by γ_i to the ith group $\tilde{\mathbf{h}}^i$ as

$$f\left(\tilde{\mathbf{h}}^i; \gamma_i\right) = (\pi\gamma_i)^{-N_r N_t} \exp\left(-\gamma_i^{-1} \left\|\tilde{\mathbf{h}}^i\right\|_2^2\right). \tag{29}$$

The prior corresponding to the CIR vector $\tilde{\mathbf{h}}$ can be obtained as $f\left(\tilde{\mathbf{h}}; \boldsymbol{\Gamma}\right) = \prod_{i=0}^{L_h-1} f\left(\tilde{\mathbf{h}}^i; \gamma_i\right)$. Employing the EM procedure described in Section III, the *a posteriori* pdf $f\left(\tilde{\mathbf{h}} \mid \bar{\mathbf{y}}_0^{\text{td}}; \widehat{\boldsymbol{\Gamma}}^{(p-1)}\right)$ in the pth EM iteration can be calculated as $f\left(\tilde{\mathbf{h}} \mid \bar{\mathbf{y}}_0^{\text{td}}; \widehat{\boldsymbol{\Gamma}}^{(p-1)}\right) = \mathcal{CN}\left(\tilde{\boldsymbol{\mu}}^{(p)}, \tilde{\boldsymbol{\Sigma}}^{(p)}\right)$, where the *a posteriori* mean $\tilde{\boldsymbol{\mu}}^{(p)}$ and the covariance matrix $\tilde{\boldsymbol{\Sigma}}^{(p)}$ are formulated as

$$\tilde{\boldsymbol{\mu}}^{(p)} = \tilde{\boldsymbol{\Sigma}}^{(p)} \bar{\boldsymbol{\mathcal{D}}}^H \tilde{\mathbf{R}}_\eta^{-1} \bar{\mathbf{y}}_0^{\text{td}} \ \text{ and}$$
$$\tilde{\boldsymbol{\Sigma}}^{(p)} = \left(\bar{\boldsymbol{\mathcal{D}}}^H \tilde{\mathbf{R}}_\eta^{-1} \bar{\boldsymbol{\mathcal{D}}} + \left(\mathbf{I}_{N_r N_t} \otimes \widehat{\boldsymbol{\Gamma}}^{(p-1)}\right)^{-1}\right)^{-1}. \tag{30}$$

Employing the *a posteriori* pdf above, the update of the hyperparameters is expressed as

$$\widehat{\gamma}_i^{(p)} = \frac{1}{N_r N_t} \sum_{d=1}^{N_r N_t} \left|\tilde{\boldsymbol{\mu}}^{(p)}\left(\tilde{d}\right)\right|_2^2 + \tilde{\boldsymbol{\Sigma}}^{(p)}\left(\tilde{d}, \tilde{d}\right), \tag{31}$$

where $\tilde{d} = (i + 1 + (d-1)L_h)$. Upon convergence, the TD-GSBL based estimate of the CIR vector $\tilde{\mathbf{h}}$ is given by $\tilde{\mathbf{h}}^{\text{TD-GSBL}} = \tilde{\boldsymbol{\mu}}$.

A. Other Sparse Channel Estimation Techniques: OMP, FOCUSS and Lasso

As it transpires from the previous section, the SBL framework employs a Bayesian philosophy for sparse estimation. The orthogonal matching pursuit (OMP) is another popular greedy-search based sparse channel estimation technique [43]. This is also used for performance comparison in Section-VI. The operation of OMP-aided SISO-FBMC systems is described in a compact fashion in Algorithm-1 of the technical report in [42]. As described in Step-3 therein, the OMP algorithm chooses the column of the matrix $\mathbf{D}$ at each iteration in a greedy fashion. Its convergence and the resultant performance therefore is sensitive to the selection of the sensing matrix $\mathbf{D}$ and to the stopping-threshold ϵ_0, which gives rise to structural and convergence errors [29].

The least absolute shrinkage and selection operator (Lasso) [27] and focal underdetermined system solver (FOCUSS) [28] are other attractive convex-relaxation based techniques that can be employed for sparse channel estimation in FBMC systems. However, the performance of the Lasso technique is sensitive to the regularization parameter, whereas FOCUSS suffers from convergence deficiencies [29]. On the other hand, the SBL framework is free from any regularization/ stopping parameters and leads to fewer convergence errors, which can be attributed to the well-established properties of the EM method. Due to lack of space, these are not described here.

IV. Doubly-Selective Sparse CSI Estimation in FBMC Systems

In practice, the CIR vector $\mathbf{h}$ is a time- as well as frequency-selective [44], [45], i.e. doubly-selective sparse CIR vector $\mathbf{h}_u \in \mathbb{C}^{L_h \times 1}$, where u denotes the block-index. The time-evolution of the CIR vector $\mathbf{h}_u$ can be described by a first-order Gauss-Markov model [23], [46]

$$\mathbf{h}_u = \rho \mathbf{h}_{u-1} + \sqrt{1-\rho^2}\ \mathbf{w}_u, \tag{32}$$

where the correlation ρ symbolizes the time-domain correlation of the vector $\mathbf{h}_u$. Similar to [23], we can employ Jake's model for calculating the coefficient ρ as $\rho = J_0\left(2\pi f_D T_B\right)$, where J_0 is the zeroth-order Bessel function of first kind, while T_B and f_D denote the block-length and the Doppler spread, respectively. The complex vector $\mathbf{w}_u \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{\Gamma}\right) \in \mathbb{C}^{L_h \times 1}$ models the driving noise process and shares a common sparsity profile with $\mathbf{h}_u$, an assumption that is justified by the fact that the locations of the dominant components of the channel vector $\mathbf{h}_u$ do not change for several OFDM frames [23], [46], since the power delay profile is constant. This also implies that the process $\mathbf{h}_u$ considered in the state model (32) is wide-sense stationary (WSS). Furthermore, upon considering the block duration to be well within the coherence time interval, we can assume the channel $\mathbf{h}_u$ to be constant over each block. By exploiting the properties of a Gauss-Markov random process, the components of the sparse innovation vector $\mathbf{w}_u$ are assumed to be independent of $\mathbf{h}_v$, for $v < u$.

1) Doubly-Selective Sparse Channel Estimation in SISO-FBMC: Using (17), the doubly-selective TD model of sparse CSI estimation can be described as

$$\mathbf{y}_{0,u}^{\text{td}} = \mathbf{D}\mathbf{h}_u + \boldsymbol{\eta}_{0,u}, \tag{33}$$

where $\mathbf{y}_{0,u}^{\text{td}} \in \mathbb{C}^{N_p \times 1}$ denotes the received pilot vector corresponding to the block-index u. The TD-SBL-KF framework is now developed for the estimation of the doubly-selective sparse CSI vector $\mathbf{h}_u$ (33). To begin with, the following parameterized Gausssian prior is assigned to the sparse doubly-selective CIR vector $\mathbf{h}_u$

$$f(\mathbf{h}_u; \mathbf{\Gamma}_u) = \prod_{i=1}^{L_h} (\pi \gamma_{i,u})^{-1} \exp\left(-\frac{|\mathbf{h}_u(i)|^2}{\gamma_{i,u}} \right), \tag{34}$$

where $\gamma_{i,u},\ 1 \le i \le L_h$, represents the hyperparameter associated with the ith channel tap in the uth training block, and $\mathbf{\Gamma}_u \in \mathbb{R}_+^{L_h \times L_h}$ is the diagonal matrix comprising these hyperparameters. Let $\widehat{\mathbf{h}}_{u-1|u-1} \in \mathbb{C}^{L_h \times 1}$ and $\mathbf{\Sigma}_{u-1|u-1} \in \mathbb{C}^{L_h \times L_h}$ represent the filtered estimate and the corresponding estimation error covariance matrix of the CSI vector $\mathbf{h}_{u-1}$, respectively. Let $\widehat{\mathbf{\Gamma}}_u$ denote the estimate of the matrix $\mathbf{\Gamma}_u$ in the uth block. The MMSE prediction $\widehat{\mathbf{h}}_{u|u-1}$ of the CIR vector $\mathbf{h}_u$ and the corresponding error covariance matrix $\mathbf{\Sigma}_{u|u-1}$ are obtained as [41]

$$\begin{aligned} \widehat{\mathbf{h}}_{u|u-1} &= \rho\ \widehat{\mathbf{h}}_{u-1|u-1}, \quad \text{and} \\ \mathbf{\Sigma}_{u|u-1} &= \rho^2\ \mathbf{\Sigma}_{u-1|u-1} + (1-\rho^2)\ \widehat{\mathbf{\Gamma}}_u, \end{aligned} \tag{35}$$

where the covariance matrix of the driving noise vector $\mathbf{w}_u$ has been set to $\widehat{\mathbf{\Gamma}}_u$ in (35). Furthermore, upon employing the predicted quantities $\widehat{\mathbf{h}}_{u|u-1}$ and $\mathbf{\Sigma}_{u|u-1}$ above, the filtered estimate $\widehat{\mathbf{h}}_{u|u}$ and the associated error covariance $\mathbf{\Sigma}_{u|u}$ can be updated as

$$\begin{aligned} \widehat{\mathbf{h}}_{u|u} &= \widehat{\mathbf{h}}_{u|u-1} + \mathbf{K}_u(\mathbf{y}_{0,u}^{\text{td}} - \mathbf{D}\widehat{\mathbf{h}}_{u|u-1}) \quad \text{and} \\ \mathbf{\Sigma}_{u|u} &= (\mathbf{I}_{L_h} - \mathbf{K}_u\mathbf{D})\mathbf{\Sigma}_{u|u-1}, \end{aligned} \tag{36}$$

where $\mathbf{K}_u \in \mathbb{C}^{L_h \times N_p}$ is the Kalman-gain matrix given by $\mathbf{K}_u = \mathbf{\Sigma}_{u|u-1}\mathbf{D}^H(\mathbf{R}_\eta + \mathbf{D}\mathbf{\Sigma}_{u|u-1}\mathbf{D}^H)^{-1}$. Interestingly, since the matrix $\mathbf{K}_u$ depends on the matrix $\mathbf{\Sigma}_{u|u-1}$, which in turn only depends on $\widehat{\mathbf{\Gamma}}_u$, the filtered estimate $\widehat{\mathbf{h}}_{u|u}$ of the sparse CSI vector once again boils down to the estimation of the matrix $\mathbf{\Gamma}_u$. To achieve this, we can once again maximize the Bayesian evidence $f(\mathbf{y}_{0,u}^{\text{td}}; \mathbf{\Gamma}_u)$ using the EM iterations described next.

Let $\widehat{\mathbf{\Gamma}}_u^{(p-1)}$ be the hyperparameter matrix estimate obtained in the EM iteration $(p-1)$ corresponding to the uth block. The *a posteriori* pdf of the sparse doubly-selective CIR vector $\mathbf{h}_u$ in the pth EM iteration can be evaluated as $f(\mathbf{h}_u|\mathbf{y}_{0,u}^{\text{td}}; \mathbf{\Gamma}_u^{(p-1)}) = \mathcal{CN}\left(\boldsymbol{\mu}_u^{(p)}, \mathbf{\Sigma}_u^{(p)}\right)$, where the *a posteriori* mean $\boldsymbol{\mu}_u^{(p)} \in \mathbb{C}^{L_h \times 1}$ and the covariance matrix $\mathbf{\Sigma}_u^{(p)} \in \mathbb{C}^{L_h \times L_h}$ are determined as

$$\begin{aligned} \boldsymbol{\mu}_u^{(p)} &= \mathbf{\Sigma}_u^{(p)}\mathbf{D}^H\mathbf{R}_\eta^{-1}\mathbf{y}_{0,u}^{\text{td}} \quad \text{and} \\ \mathbf{\Sigma}_u^{(p)} &= \left(\mathbf{D}^H\mathbf{R}_\eta^{-1}\mathbf{D} + \left(\widehat{\mathbf{\Gamma}}_u^{(p-1)}\right)^{-1}\right)^{-1}. \end{aligned} \tag{37}$$

As seen in (25), the estimates $\widehat{\gamma}_{i,u}^{(p)}$ in the pth EM iteration can be obtained as

$$\widehat{\gamma}_{i,u}^{(p)} = \mathbf{\Sigma}_u^{(p)}(i,i) + |\boldsymbol{\mu}_u^{(p)}(i)|^2. \tag{38}$$

Upon convergence of the EM procedure, the estimate $\widehat{\mathbf{\Gamma}}_u$ in the uth block is obtained as $\widehat{\mathbf{\Gamma}}_u = \text{diag}\big(\gamma_{1,u}^{(p)}, \gamma_{2,u}^{(p)}, \cdots, \gamma_{L_h,u}^{(p)}\big)$. The various quantities of the TD-SBL-KF are initialized as $\widehat{\mathbf{h}}_{-1|-1} = \mathbf{0}_{L_h \times 1}, \mathbf{\Sigma}_{-1|-1} = \widehat{\mathbf{\Gamma}}_0^{(p)}, \widehat{\mathbf{\Gamma}}_0^{(0)} = \mathbf{I}_{L_h}$. The proposed SBL-KF framework initializes the hyperparameter matrix $\widehat{\mathbf{\Gamma}}_u^{(0)}$ for block u as $\widehat{\mathbf{\Gamma}}_u^{(0)} = \widehat{\mathbf{\Gamma}}_{u-1}^{(p)}$, i.e., to the converged estimate of the hyperparameter matrix obtained from the previous block. The advantage of this initialization procedure is two-fold: when the sparsity profile of the CIR does not change, the convergence is faster. On the other hand, when it changes suddenly, the proposed SBL-KF becomes capable of detecting the change and adapts to the new sparsity profile in a few iterations. This has been illustrated via a simulation result in Fig. 2 of our technical report in [42]. The schematic of the proposed TD-SBL-KF scheme is given in Fig. 1(a) of

the technical report [42]. The hyperparameter update block therein evaluates the hyperparameter matrix $\mathbf{\Gamma}_u$, which ensures sparsity in the estimate of the CIR vector $\mathbf{h}_u$ through the Prediction block. The Filtering block therein provides the final estimate $\widehat{\mathbf{h}}_u^{\text{TD-SBL-KF}} = \widehat{\mathbf{h}}_{u|u}$ and the error covariance matrix $\mathbf{\Sigma}_{u|u}$ of the sparse doubly-selective CIR vector $\mathbf{h}_u$. An algorithmic description of the proposed TD-SBL-KF technique is given in Algorithm-3 of the technical report [42]. A very important property of the proposed TD-SBL-KF technique is that it is *online* of nature, since it only employs the output vector $\mathbf{y}_{0,u}^{\text{td}}$ for updating $\mathbf{h}_u$. This significantly reduces both the estimation delay as well as the complexity. The IAM-based doubly-selective channel estimation can be developed following similar lines, commencing with the observation model of a quasi-static channel in (10).

Finally, in order to further improve the CSI estimation accuracy, one can employ a KF-smoother (KFS) framework, which employs all the available measurements to refine the filtered estimates. Toward this, let U represent the number of blocks. Using the standard Kalman filter (KF) notations, the backward smoothing steps are given as follows.

for $u = U, U-1, \cdots, 2$

$$\widetilde{\mathbf{\Sigma}}_{u-1} = \rho \mathbf{\Sigma}_{u-1|u-1} \mathbf{\Sigma}_{u|u-1}^{-1},$$
$$\widehat{\mathbf{h}}_{u-1|U} = \widehat{\mathbf{h}}_{u-1|u-1} + \widetilde{\mathbf{\Sigma}}_{u-1} \left(\widehat{\mathbf{h}}_{u|U} - \widehat{\mathbf{h}}_{u|u-1} \right),$$
$$\mathbf{\Sigma}_{u-1|U} = \mathbf{\Sigma}_{u-1|u-1} + \widetilde{\mathbf{\Sigma}}_{u-1} \left(\mathbf{\Sigma}_{u|U} - \mathbf{\Sigma}_{u|u-1} \right) \widetilde{\mathbf{\Sigma}}_{u-1}^{H},$$

where the quantity $\widehat{\mathbf{h}}_{u-1|U}$ represents the final estimate of the KFS for the $(u-1)$st block.

2) Doubly-Selective Sparse Channel Estimation in MIMO-FBMC: Let the doubly-selective CIR $\mathbf{h}_u^{r,t} \in \mathbb{C}^{L_h \times 1}$ corresponding to rth RA and tth TA be defined as

$$\mathbf{h}_u^{r,t} = \left[h_u^{r,t}[0], h_u^{r,t}[1], \ldots, h_u^{r,t}[L_h - 1] \right]^T, \tag{39}$$

where $h_u^{r,t}[l], 0 \leq l \leq L_h - 1$, denotes the lth CIR tap in the uth block. Similar to the doubly-selective channel model of (32), the time-evolution of the CIR vector $\mathbf{h}_u^{r,t}$ can be modeled as

$$\mathbf{h}_u^{r,t} = \rho \mathbf{h}_{u-1}^{r,t} + \sqrt{1-\rho^2}\ \mathbf{w}_u^{r,t}, \tag{40}$$

where the complex vector $\mathbf{w}_u^{r,t} \in \mathbb{C}^{L_h \times 1}$ denotes the driving noise process, and shares a common sparsity profile with $\mathbf{h}_u^{r,t}$. To avoid repetition, the notations defined for the quasi-static MIMO-FBMC system are extended to the doubly-selective scenario with the addition of the subscript u for the uth block-index. Let $\bar{\mathbf{h}}_u$ denote the doubly-selective block-sparse CIR vector corresponding to the uth block. Along similar lines to Section III-2, the IAM-model of doubly selective sparse CSI estimation in MIMO-FBMC systems can be constructed as

$$\mathbf{y}_u^{\text{iam}} = \mathbf{\Phi} \bar{\mathbf{h}}_u + \boldsymbol{\eta}_u, \tag{41}$$

where $\mathbf{y}_u^{\text{iam}}$ and $\boldsymbol{\eta}_u$ represent the received pilot and the corresponding noise vectors, similar to (15), for the block-index u. Similarly, let $\tilde{\mathbf{h}}_u$ denote the doubly-selective group-sparse CIR vector corresponding to the uth block. Using (21), the TD-based model of doubly-selective sparse CSI estimation for our MIMO-FBMC system can be expressed as

$$\bar{\mathbf{y}}_{0,u}^{\text{td}} = \bar{\boldsymbol{\mathcal{D}}} \tilde{\mathbf{h}}_u + \bar{\boldsymbol{\eta}}_{0,u}, \tag{42}$$

where $\bar{\mathbf{y}}_{0,u}^{\text{td}}$ and $\bar{\boldsymbol{\eta}}_{0,u}$ denote the received pilot and noise vectors in the block-index u. The doubly-selective sparse channel estimation technique TD-SBL-KF developed for SISO-FBMC systems can now be readily extended to the TD-GSBL-KF model of MIMO-FBMC systems by modifying the hyperparameter update equation derived in (31). The TD-GSBL-KF algorithm proposed for the doubly-selective channel estimation model of (42) is presented in Algorithm-4 of our technical report [42]. The equivalent procedure of the IAM-based doubly selective channel estimation model of (41) can be derived similarly.

V. Complexity Analysis

This subsection briefly describes the computational complexity of both the proposed and exiting SISO- and MIMO-FBMC sparse channel estimation schemes. A detailed complexity analysis is provided in the technical report in [42]. As shown in [42], the complexity order of the proposed IAM-SBL, TD-SBL, IAM-SBL-KF and TD-SBL-KF channel estimation schemes for SISO-FBMC is $\mathcal{O}(N_p^3 + L_h^3)$. On the other hand, the proposed IAM-BSBL/IAM-BSBL-KF and TD-GSBL/TD-GSBL-KF schemes conceived for MIMO-FBMC channel estimation have the computational complexity of $\mathcal{O}(N_p^3 N_r^3 M^3 + N_t^3 N_r^3 L_h^3)$ and $\mathcal{O}(N_p^3 N_r^3 + N_t^3 N_r^3 L_h^3)$, respectively. The OMP scheme, as described in [42] has the computational complexity of $\mathcal{O}(i^3)$ in the ith iteration. This leads to $\mathcal{O}(N_p^3)$ complexity for SISO-FBMC sparse channel estimation, whereas it has the complexities of $\mathcal{O}(N_p^3 N_r^3 M^3)$ and $\mathcal{O}(N_p^3 N_r^3)$ for MIMO-FBMC sparse channels using the IAM and TD models, respectively. It follows from (10) and (17) that the conventional LS estimator designed for SISO-FBMC has the complexity of $\mathcal{O}(L_h^3)$, while its MIMO counterpart incurs a complexity order of $\mathcal{O}(N_t^3 N_r^3 L_h^3)$, as inferred from (15) and (21). It can be observed that the proposed SBL-based schemes have complexity orders similar to their OMP counterparts, but they have slightly higher complexities than their LS counterparts. However, as shown in the next section, the proposed SBL-based schemes significantly outperform both the exiting OMP and LS-based schemes for sparse channel estimation in both SISO- and MIMO-FBMC systems.

VI. Simulation Results

We now demonstrate the sparse CSI estimation performance of the proposed schemes and compare them to that of a variety of other schemes such as OMP, FOCUSS [28], Lasso [27] and AMP [31]. The number of subcarriers N is chosen from the set $\{64, 256\}$. The IOTA based discrete time prototype filter $p[k]$ of length $L_p = 4N$ is utilized for FBMC modulation. The data and training symbols use QPSK (4-QAM). The SNR on each subcarrier is computed as $2P_d/\sigma_\eta^2$. The standard pedestrian-B and Vehicular-A channel models of the international telecommunication union (ITU) [47], which has $L_h = 32$ taps in conjunction with 6 dominant taps, model the wireless channel for quasi-static and doubly-selective scenarios, respectively,

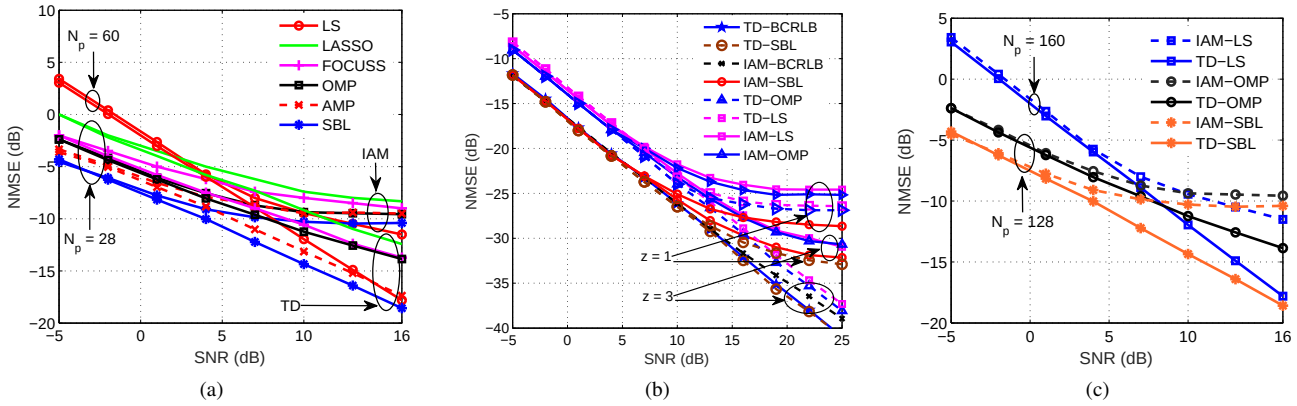


Fig. 2: SISO-FBMC quasi-static channel estimation performance comparison of the proposed schemes with the corresponding BCRLBs derived in our technical report in [42] and the existing schemes: (a) NMSE versus SNR with $N = 64$ and $z = 3$; (b) NMSE versus SNR with $N = N_p = 256$ and $z = 3$; and (c) NMSE versus SNR over 5G-NR CDL-B channel with $N = 256$, $N_p = 256$ and $z = 3$.

between each pair of TA and RA, unless stated otherwise. The number of TAs and RAs for the MIMO scenario is set to $N_r = N_t = 2$. However, note that the proposed designs are applicable for estimating quasi-static and doubly selective channels of any MIMO configuration. The hyperparameters γ_i, for $\gamma_i = 1, 0 \leq i \leq L_h - 1$, are initialized to unity. The stopping parameters ϵ_1 and $N_{\max}$ are set to 10^{-5} and 50, respectively, for all the proposed sparse schemes. As shown in (10), (17) and (21) for quasi-static channel estimation, and in (33) and (42) for doubly selective channel estimation, the proposed designs estimate the channel in the time-domain, in order to exploit the sparse nature of the respective channels. Hence the NMSE is calculated as $||\hat{\mathbf{h}} - \mathbf{h}||^2/||\mathbf{h}||^2$, where $\hat{\mathbf{h}}$ denotes the estimate of the CIR vector $\mathbf{h}$.

A. *Sparse CSI Estimation in SISO-FBMC Systems*

Fig. 2(a) depicts the NMSE of the SBL, OMP, AMP, Lasso and FOCUSS estimators for sparse CSI estimation in SISO-FBMC systems, using both the IAM and TD models in Section-II-3 and Section-II-5. The NMSE of the conventional LS estimator is also shown therein. However, while an ill-posed channel estimation model using $N_p = 28$ training subcarriers is considered for the sparse estimators, $N_p = 60$ pilot subcarriers are used for the LS estimator, in order to obtain a comparable NMSE. The proposed SBL-based techniques have significantly lower NMSE than the OMP, Lasso, FOCUSS and LS benchmark techniques. The poor performance of the LS approaches, despite employing more than twice the number of pilots, is due to the fact that they do not exploit the sparse nature of the wireless channel. This clearly demonstrates the importance and significant benefits that can be realized via exploiting sparsity. The OMP algorithm, which is fragile to the selection of both the sensing matrix and of the stopping threshold, also results in poor performance. The NMSE of the AMP-based algorithms is slightly inferior to its SBL counterpart, since the former approximates the messages by the Gaussian distribution [33]. Furthermore, it is observed that the TD schemes yield significantly lower NMSE than the IAM-based techniques in the high SNR regime, because, unlike the latter, the former does not necessitate a frequency-flat channel at the subcarrier level. The intrinsic interference, which dominates in the high-SNR regime, results in NMSE floors for the IAM-based channel estimation schemes at higher SNR.

Fig. 2(b) compares the NMSE versus SNR results for both the proposed and existing schemes using $N = N_p = 256$ and $z \in \{1, 3\}$. This renders the channel estimation models in (10) and (17) well-posed. The BCRLBs of the proposed IAM-SBL and TD-SBL schemes, as determined in Section-IV of our technical report [42], have also been plotted for reference. It may be observed that for $z = 1$, the NMSE of all the schemes saturates at high SNRs due to the ISI between the training symbols and the rest of the frame. It can be readily seen that the OMP techniques perform similarly to their LS counterparts for both the IAM and TD systems. The proposed SBL schemes once again outperform both the OMP and the conventional LS techniques. Interestingly, it is also observed that the proposed IAM-SBL scheme performs similarly to its TD counterpart and achieves its BCRLB at low SNR for $z = 3$. By contrast, the proposed TD-SBL approach achieves its BCRLB across the entire SNR range for $z = 3$. It is important to realize that the BCRLB corresponds to the ideal scenario of having a known channel profile and zero ISI between the training symbols. The fact that the SBL algorithm is capable of achieving this in a realistic system having an unknown channel support demonstrates its prowess and suitability for practical implementation.

The NMSE performance of Fig. 2(a) and Fig. 2(b) underlines an interesting aspect of the IAM-SBL scheme. It is observed that as the number of subcarriers N increases, the performance of IAM-SBL approaches that of the TD-SBL scheme. This can be explained by the observation that the IAM-based schemes rely on the approximation in (4). The corresponding approximation error diminishes progressively as the number of subcarriers N increases. Since the proposed TD-SBL approach does not rely on the approximation in (4), its performance does not exhibit an NMSE floor at high SNRs, regardless of the number of subcarriers. However, for data detection under the TD-based model, one has to use a

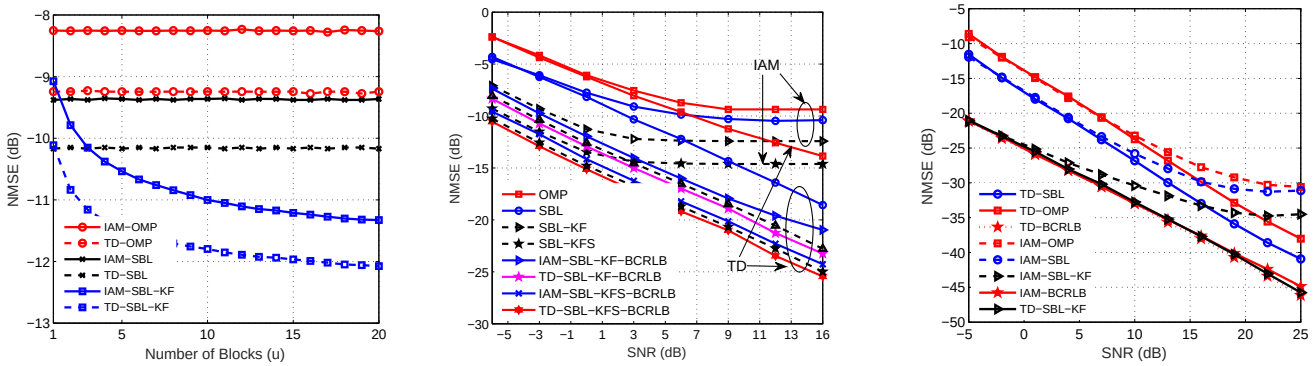


Fig. 3: SISO-FBMC doubly-selective channel estimation: (a) NMSE versus number of blocks (u) with $N = 64$, $N_p = 28$, $z = 3$, $\rho = 0.9981$ and SNR $= -5$ dB; (b) NMSE versus SNR with corresponding BCRLBs derived in [42] with $N = 64$, $N_p = 28$, $\rho = 0.9981$ and $z = 3$; and (b) NMSE versus SNR with $N = N_p = 256$, $\rho = 0.9981$ and $z = 3$.

multi-tap time-domain equalizer [20], [39], which suffers from a high computational complexity. On the other hand, data detection in the IAM-based model can be performed using the conventional single-tap frequency-domain equalizer [20] that has a low complexity. In light of the above observations, one can deploy the IAM-SBL scheme, when the number of subcarriers N is significantly larger than the number of channel taps L_h. Otherwise, the TD-SBL scheme is best suited for such systems.

Fig. 2(c) displays NMSE versus SNR performance of the proposed schemes for estimating CDL-B channel model of 5G-new radio (NR). A 50 MHz system is considered over CDL-B channel, which has a maximum normalized delay spread of 4.7834 [48, Table 7.7.1-2]. It follows from Eq. (7.7-1) of [48] that with the desired channel delay spread of 500 ns, the maximum delay spread of the CDL-B channel is $4.7834 \times 500 = 2391.7$ ns. Thus, with a sampling interval of $1/(50 \times 10^6) = 20$ ns, the CDL-B channel has $2391.7/20 \approx 120$ channel taps. Out of these 120 taps, 24 taps, as per Table 7.7.1-2 of [48], are the dominant taps, which leads to an approximately sparse channel. For estimating this channel, an FBMC system with $N = 256$ subcarriers and $N_p = 128$ pilot symbols is considered. It can seen from Fig. 2(c) that the proposed TD-SBL and IAM-SBL schemes significantly outperform their conventional counterparts. The NMSE of the IAM-based schemes exhibit a floor at high SNR due to the dominance of the intrinsic interference. This clearly demonstrates the accuracy of the proposed schemes for the CDL-B channel estimation.

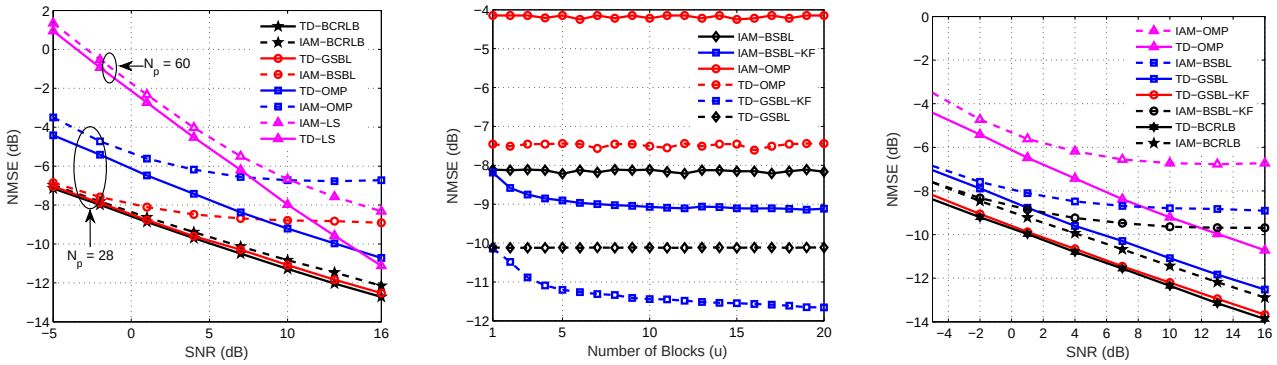
Fig. 3(a) plots the NMSE versus the number of blocks (u) both for the OMP and for the proposed SBL suite of schemes, which includes SBL and SBL-KF that incorporates the Kalman filter, for the CSI estimation of a doubly-selective scenario at SNR $= -5$dB. A SISO-FBMC-aided mobile user operating in the 2.4 GHz band and roaming at a velocity of $v = 30$ km/h, which results in a Doppler shift of $f_D = 69$ Hz, has been considered. The block-length T_B is set to 0.2 ms. With these settings in Jake's model, the time-correlation ρ of the doubly-selective CSI model of (32) becomes $\rho = J_0(2\pi f_d T_B) \approx 0.9981$. Interestingly, the NMSE of the OMP and of the general SBL schemes that exploit only sparsity is seen to remain unaffected by the number of blocks u. By contrast, the performance of the SBL-KF techniques that leverage the sparsity in addition to the temporal correlation across blocks, significantly improves upon increasing the number of blocks, thus outperforming the quasi-static approaches.

Fig. 3(b) displays the NMSE versus SNR of the OMP and SBL techniques for doubly-selective channel estimation. This study is performed using $N = 64$ and $N_p = 28$ pilot subcarriers, leading to ill-posed channel estimation models in (33) and (41). The SBL-KF schemes of both the IAM and TD systems, which were specifically designed for doubly-selective channel estimation, result in a significantly lower NMSE than the other schemes. The SBL-KFS further enhances the estimation performance in comparison to the SBL-KF, since it employs all the available measurements to further refine the filtered estimates. Once again, the NMSE of the IAM-based schemes exhibit floor at high SNR due to the high level of intrinsic interference. The SBL-KF and SBL-KFS techniques of the TD system achieve the time-recursive BCRLB derived in [42].

Fig. 3(c) displays the NMSE of the OMP and SBL class of schemes for doubly-selective channel estimation with pilots transmitted over all the subcarriers, i.e. for $N = N_p = 256$. Similar to our observation in Fig. 2(b) for a quasi-static channel, due to the increased number of subcarriers, the performance of the IAM algorithms exhibits flooring above SNR $= 10$dB.

B. Sparse CSI Estimation in MIMO-FBMC Systems

Fig. 4(a) provides our NMSE comparison between the SBL-based sparse CSI estimation schemes as well as the LS and OMP techniques. The number of pilot subcarriers of the LS-based schemes is set to $N_p = 60$, which is significantly higher than that of both the OMP and SBL-based estimators with $N_p = 28$. The IAM model based BSBL and TD model based GSBL schemes, which exploit the special block and



(a) (b) (c)

Fig. 4: Performance comparison of the proposed techniques with the corresponding BCRLBs derived in the technical report in [42] and the existing schemes for sparse CSI estimation in 2×2 MIMO-FBMC systems: (a) quasi-static channel estimation: NMSE versus SNR with $N = 64$ and $z = 3$; (b) doubly-selective channel estimation: NMSE versus number of blocks (u) with $N = 64$, $N_p = 28$, $z = 3$, $\rho = 0.9981$ and SNR $= 5$ dB (c) doubly-selective channel estimation: NMSE versus SNR with $N = 64$, $N_p = 28$, $\rho = 0.9981$ and $z = 3$.

group-sparse structures inherent in the sparse MIMO channel, significantly outperform their OMP counterparts. Once again, we have an NMSE floor for the IAM-based schemes in the high-SNR regime due to the approximation error arising from (4). The TD model based proposed GSBL is seen to achieve its BCRLB.

Fig. 4(b) considers a doubly-selective MIMO-FBMC system and demonstrates the NMSE performance versus the number of blocks (u) for the various proposed and existing schemes. The SNR is set to 5 dB for this study with $N_p = 28$ subcarriers loaded with pilots. It can seen that the NMSE of the BSBL-KF and GSBL-KF techniques proposed for doubly-selective sparse channel estimation is better than that of their quasi-static counterparts, namely of the BSBL, OMP and GSBL based on either the IAM or TD models. The NMSE improvement can be expected due to the reason that the BSBL-KF and GSBL-KF schemes, unlike their quasi-static counterparts, additionally exploit the temporal correlation across training blocks. Fig. 4(c) displays the NMSE of the various schemes for doubly-selective sparse channel estimation, which shows a similar trend as observed in Fig. 4(b). Furthermore, the TD model based GSBL-KF is seen to achieve the recursive BCRLB determined in [42].

Fig. 3 of our technical report in [42] shows the coded BER performance of MIMO-FBMC systems using channel estimates obtained from the proposed and existing schemes. The TD system using GSBL can be seen therein to achieve the lowest BER, outperforming, viz. IAM with BSBL and both TD as well as IAM-based OMP and LS estimators. The NMSE floor of channel estimation in the IAM-based techniques naturally leads to the BER floor at high SNRs.

VII. Conclusions

We developed a suite of SBL-based approaches for sparse CSI estimation in FBMC systems using the IAM and TD models that are well-suited for channels with low and high levels of frequency selectivity, respectively. Initially, SBL-based sparse channel estimators were conceived for CSI estimation in quasi-static SISO-FBMC systems. The SBL-KF framework was subsequently exploited to track the CSI of doubly-selective SISO-FBMC systems. Next, novel IAM and TD-based channel estimation models were developed for MIMO-FBMC systems that demonstrate block and group sparsity, respectively. These properties were successfully exploited by our IAM-based BSBL and TD-based GSBL schemes for sparse estimation, which were shown to lead to improved estimation performance arising from leveraging the simultaneous sparsity encountered. The BCRLBs were derived for all the estimation scenarios. The performance of the proposed SBL-based schemes were shown to be significantly improved in comparison to the various sparse channel estimators such as the OMP, Lasso, FOCUSS as well as the conventional LS technique, for both the ill-posed and well-posed observation models. Interestingly, the TD-based estimators are seen to achieve the respective BCRLBs, while the performance of IAM-based schemes floors at high SNRs.

References

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

[2] M. Renfors, X. Mestre, E. Kofidis, and F. Bader, *Orthogonal Waveforms and Filter Banks for Future Communication Systems*. Academic Press, 2017.

[3] L. Hanzo, Y. Akhtman, J. Akhtman, L. Wang, and M. Jiang, *MIMO-OFDM for LTE, WiFi and WiMAX: Coherent versus non-coherent and cooperative turbo transceivers*. John Wiley & Sons, 2011.

[4] R. Nissel, S. Schwarz, and M. Rupp, "Filter bank multicarrier modulation schemes for future mobile communications," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 8, pp. 1768–1782, 2017.

[5] B. Farhang-Boroujeny and H. Moradi, "OFDM inspired waveforms for 5G," *IEEE Communications Surveys and Tutorials*, vol. 18, no. 4, pp. 2474–2492, 2016.

[6] D. Katselis, E. Kofidis, A. A. Rontogiannis, and S. Theodoridis, "Preamble-based channel estimation for CP-OFDM and OFDM/OQAM systems: a comparative study," *IEEE Trans. Signal Processing*, vol. 58, no. 5, pp. 2911–2916, 2010.

[7] J. M. Choi, Y. Oh, H. Lee, and J. Seo, "Pilot-aided channel estimation utilizing intrinsic interference for FBMC/OQAM systems," *TBC*, vol. 63, no. 4, pp. 644–655, 2017.

[8] 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 Trans. Signal Processing*, vol. 67, no. 18, pp. 4668–4682, 2019.

[9] 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.

[10] D. Kong, D. Qu, and T. Jiang, "Time domain channel estimation for OQAM-OFDM systems: Algorithms and performance bounds," *IEEE Trans. Signal Processing*, vol. 62, no. 2, pp. 322–330, 2014.

[11] V. Savaux, F. Bader, and Y. Louët, "A joint MMSE channel and noise variance estimation for OFDM/OQAM modulation," *IEEE Trans. Communications*, vol. 63, no. 11, pp. 4254–4266, 2015.

[12] Z. He, L. Zhou, Y. Chen, and X. Ling, "Nonlinear complex support vector regression for fading channel estimation in FBMC-OQAM system," *IEEE Wireless Commun. Letters*, vol. 8, no. 3, pp. 753–756, 2019.

[13] X. Cheng, D. Liu, C. Wang, S. Yan, and Z. Zhu, "Deep learning-based channel estimation and equalization scheme for FBMC/OQAM systems," *IEEE Wireless Commun. Letters*, vol. 8, no. 3, pp. 881–884, 2019.

[14] X. Cheng, D. Liu, S. Yan, W. Shi, and Y. Zhao, "Channel estimation and equalization based on deep BLSTM for FBMC-OQAM, systems," in *2019 IEEE International Conference on Communications, ICC 2019, Shanghai, China, May 20-24, 2019*, 2019, pp. 1–6.

[15] W. Cui, D. Qu, T. Jiang, and B. Farhang-Boroujeny, "Coded auxiliary pilots for channel estimation in FBMC-OQAM systems," *IEEE Trans. Veh. Technol.*, vol. 65, no. 5, pp. 2936–2946, 2016.

[16] 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.

[17] M. Fuhrwerk, S. Moghaddamnia, and J. Peissig, "Scattered pilot-based channel estimation for channel adaptive FBMC-OQAM systems," *IEEE Trans. Wirel. Commun.*, vol. 16, no. 3, pp. 1687–1702, 2017.

[18] E. Kofidis, "Preamble-based estimation of highly frequency selective channels in FBMC/OQAM systems," *IEEE Trans. Signal Processing*, vol. 65, no. 7, pp. 1855–1868, 2017.

[19] E. Kofidis, "Preamble-based estimation of highly frequency selective channels in MIMO-FBMC/OQAM systems," in *21th European Wireless Conference; Proceedings of*. VDE, 2015, pp. 1–6.

[20] 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.

[21] L. Zhang, P. Xiao, A. Zafar, A. ul Quddus, and R. Tafazolli, "FBMC system: An insight into doubly dispersive channel impact," *IEEE Trans. Veh. Technol.*, vol. 66, no. 5, pp. 3942–3956, 2017.

[22] T. Ihalainen, A. Ikhlef, J. Louveaux, and M. Renfors, "Channel equalization for multi-antenna FBMC/OQAM receivers," *IEEE Trans. Veh. Technol.*, vol. 60, no. 5, pp. 2070–2085, 2011.

[23] R. Prasad, C. R. Murthy, and B. D. Rao, "Joint approximately sparse channel estimation and data detection in OFDM systems using sparse Bayesian learning." *IEEE Trans. Signal Processing*, vol. 62, no. 14, pp. 3591–3603, 2014.

[24] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak, "Compressed channel sensing: A new approach to estimating sparse multipath channels," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1058–1076, 2010.

[25] Z. He, L. Zhou, Y. Yang, Y. Chen, X. Ling, and C. Liu, "Compressive sensing-based channel estimation for FBMC-OQAM system under doubly selective channels," *IEEE Access*, vol. 7, pp. 51 150–51 158, 2019.

[26] 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.

[27] R. Tibshirani, "Regression shrinkage and selection via the Lasso," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, 1996.

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

[29] 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.

[30] S. Srivastava, P. Singh, A. K. Jagannatham, A. Karandikar, and L. Hanzo, "Bayesian learning-based doubly-selective sparse channel estimation for millimeter wave hybrid MIMO-FBMC-OQAM systems," *IEEE Trans. Commun.*, vol. 69, no. 1, pp. 529–543, 2021.

[31] D. L. Donoho, A. Maleki, and A. Montanari, "Message-passing algorithms for compressed sensing," *Proceedings of the National Academy of Sciences*, vol. 106, no. 45, pp. 18 914–18 919, 2009.

[32] D. L. Donoho, A. Maleki, and A. Montanari, "Message passing algorithms for compressed sensing: I. motivation and construction," in *2010 IEEE information theory workshop on information theory (ITW 2010, Cairo)*. IEEE, 2010, pp. 1–5.

[33] M. Al-Shoukairi, P. Schniter, and B. D. Rao, "A GAMP-based low complexity sparse Bayesian learning algorithm," *IEEE Transactions on Signal Processing*, vol. 66, no. 2, pp. 294–308, 2017.

[34] R. Prasad, C. R. Murthy, and B. D. Rao, "Joint approximately sparse channel estimation and data detection in OFDM systems using sparse bayesian learning," *IEEE Trans. Signal Processing*, vol. 62, no. 14, pp. 3591–3603, 2014.

[35] J. Ma, S. Zhang, H. Li, F. Gao, and S. Jin, "Sparse bayesian learning for the time-varying massive MIMO channels: Acquisition and tracking," *IEEE Trans. Commun.*, vol. 67, no. 3, pp. 1925–1938, 2019.

[36] M. Li, S. Zhang, N. Zhao, W. Zhang, and X. Wang, "Time-varying massive MIMO channel estimation: Capturing, reconstruction, and restoration," *IEEE Trans. Commun.*, vol. 67, no. 11, pp. 7558–7572, 2019.

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

[38] 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.

[39] A. Ikhlef and J. Louveaux, "Per subchannel equalization for MIMO FBMC/OQAM systems," in *2009 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing*. IEEE, 2009, pp. 559–564.

[40] K. B. Petersen, M. S. Pedersen *et al.*, "The matrix cookbook," *Technical University of Denmark*, vol. 7, p. 15, 2008.

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

[42] P. Singh, S. Srivastava, A. K. Jagannatham, and L. Hanzo, "Technical Report on SBL-Aided Approaches for Quasi-Static and Doubly-Selective Sparse CSI Estimation for FBMC-based SISO/ MIMO Systems," IIT Kanpur, Tech. Rep., 2020, [Online]. Available: http://www.iitk.ac.in/mwn/documents/MWNLab_TR_SBL_FBMC_2020.pdf.

[43] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *Conference Record of The Twenty-Seventh Asilomar Conference on Signals, Systems and Computers, 1993. 1993*. IEEE, 1993, pp. 40–44.

[44] A. Goldsmith, *Wireless communications*. Cambridge university press, 2005.

[45] 3GPP and R. Team, "Study on channel model for frequencies from 0.5 to 100 Ghz for 5G," *3GPP TR Technical Report 38.901*, vol. Release 15, no. version 15.0.0, pp. 48–59, 2018.

[46] T. Feng, T. R. Field, and S. Haykin, "Stochastic differential equation theory applied to wireless channels," *IEEE Transactions on Communications*, vol. 55, no. 8, pp. 1478–1483, 2007.

[47] I. Recommendation, "Guidelines for evaluation of radio transmission technologies for IMT-2000," *International Telecommunication Union*, no. M.1225, 1997.

[48] J. Meredith, "Study on channel model for frequency spectrum above 6 ghz," *3GPP TR 38.900, Jun, Tech. Rep.*, 2016.