Fast Block LMS Based Estimation of Angularly Sparse Channels for Single-Carrier Wideband Millimeter Wave Hybrid MIMO Systems

Suraj Srivastava, *Graduate Student Member, IEEE,* Pooja Sharma, Saumya Dwivedi,
Aditya K. Jagannatham, *Member, IEEE,* Lajos Hanzo, *Fellow, IEEE*

Abstract—Adaptive block-based least-mean squares (BLMS)-based techniques are conceived for channel estimation in single-carrier (SC) wideband millimeter wave (mmWave) hybrid MIMO systems. In this context, a frequency-domain channel estimation model is developed for SC wideband systems, followed by a novel fast BLMS (FBLMS) technique, which has a significantly lower computational complexity than the existing channel estimation schemes designed for mmWave hybrid MIMO systems. The proposed FBLMS technique is also robust, since it does not require any second-order statistical information, such as the cross-covariance vector and covariance matrix. Next a beamspace domain representation of the mmWave MIMO channel is obtained, followed by the development of the sparse-FBLMS (SF-BLMS) scheme for the estimation of the wideband mmWave MIMO channel that additionally exploits the angular-sparsity for improved channel estimation performance. Analytical expressions are derived for the mean squared estimation error (MSEE) and mean squared observation error (MSOE) of both the proposed FBLMS and SFBLMS techniques. Furthermore, a systematic procedure is developed for determining the beneficial range of the values of the regularization parameter, which ensures a high channel estimation accuracy of the SFBLMS over FBLMS. A hybrid precoder and combiner design is also proposed for SC wideband systems by employing the channel estimates obtained using the above techniques. Simulation results are presented to illustrate the performance of the proposed BLMS-based schemes in comparison to the existing schemes, which closely match the theoretical results derived.

Index Terms—mmWave communication, single-carrier (SC), frequency-domain equalization (FDE), hybrid MIMO, channel estimation, sparsity, least mean squares, sparse LMS, block LMS.

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

Pooja Sharma is with Qualcomm, Hyderabad-500081, India (e-mail: poosha@qti.qualcomm.com).

Saumya Dwivedi is with Radisys, Bangalore-560103, India (e-mail: saumya0389@gmail.com).

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).

I. INTRODUCTION

MILLIMETER wave (mmWave) technology, which relies on the wide bandwidth available in 30 GHz-300 GHz band for wireless communication, has emerged as one of the leading candidates for achieving ultra-high data rates in 5th generation (5G) wireless systems [1]–[3]. mmWave technology is envisioned to empower various cutting edge applications such as wireless high-definition (HD) [4], smart wearables [5], virtual reality (VR) [6], high-speed wireless local area networks (WLANs) [7] and most importantly, 5G new-radio (5G-NR) devices [3]. However, communication in the mmWave regime faces a number of challenges such as higher free-space losses, reduced scattering and diffraction of the transmitted radio signal [1], [2], [8]. The solution therefore is to leverage the smaller wavelength [9], which allows us to pack numerous antennas even into a headset to support multiple-input multiple-output (MIMO) transceivers to enhance the signal-to-noise-ratio (SNR). On the other hand, the implementation of large antenna arrays in mmWave MIMO systems is quite challenging in terms of the associated hardware complexity due to the high sampling rate of analog to digital/ digital to analog converters (ADC/ DACs) coupled with the high power consumption of the radio-frequency (RF) chains connected to each transmitter/ receiver. To overcome this impediment, without substantially compromising the gains promised by mmWave MIMO systems, a hybrid MIMO architecture [2], [10]–[13] has been proposed as a practical solution. In contrast to the conventional communication systems wherein a majority of the MIMO signal processing tasks are performed in the baseband, the hybrid mmWave MIMO architecture partitions the overall signal processing between the RF and baseband domains, which leads to a substantially reduced number of RF chains (RFCs) compared to the number of transmit antennas (TAs) and receive antennas (RAs). In hybrid mmWave MIMO systems, the analog beamforming is performed using a digitally controlled network of phase-shifters to achieve a beneficial beamforming gain, while the baseband precoding/ combining is done in the digital domain for attaining a substantial multiplexing gain. It is important to note that the design and analysis of this pioneering mmWave MIMO architecture is critically dependent on the accuracy of the available channel state information (CSI) of the mmWave

MIMO channel [14]. Hence, channel estimation plays a central role in such systems and therefore it has been explored in numerous research contributions, such as [11], [12], [15]. A brief review of the seminal treatises addressing this challenging problem is provided next.

A. Review of the Existing Contributions in mmWave MIMO Channel Estimation

Due to the associated large antenna array size and low SNR experienced prior to beamforming [2], [16], the traditional least squares (LS) and minimum mean squared error (MMSE)-based channel estimation schemes rely on a substantial pilot overhead in mmWave hybrid MIMO systems. Additionally, these traditional approaches can only provide a sub-optimal performance even at significantly higher pilot overheads, since they do not exploit the angular-domain sparsity of the mmWave MIMO channel [2] arising due to the reduced multipath components [8] coupled with the highly directional beamforming of large antenna arrays. The natural approach is to estimate the parameters of the mmWave MIMO channel, namely the angles of arrival (AoAs), angles of departure (AoDs) and the corresponding path-gains. Early contributions, such as [17]–[19] have proposed various beam-training methods for CSI acquisition, which rely on pilot-beams from a codebook to estimate the AoA/ AoD pairs of the available spatial multipaths. Although these schemes are straightforward to implement, they require substantial amount of feedback for finer angular resolutions. Moreover, the feedback increases linearly with the number of devices connected to the access-point. In [16], [20], the authors have proposed to invoke the classic multiple signal classification (MUSIC) algorithm for mmWave MIMO channel estimation. Although, the MUSIC algorithm is indeed capable of extracting multiple paths, it is quite sensitive both to the position of antenna elements, as well as to their gain and phase errors.

Several studies, such as [11], [15], [21]–[25] have exploited the sparse scattering nature of mmWave propagation for acquiring the CSI of mmWave hybrid MIMO systems. In these approaches, the AoA and AoD spaces are divided into a finite number of grid-points in order to obtain the *beamspace* representation of the mmWave MIMO channel, which is sparse. Subsequently, the sparse beamspace channel can be estimated using a variety of compressive sensing-based approaches, such as sparse Bayesian learning (SBL) [12], [22], [26], orthogonal matching pursuit (OMP) [15] and their variants, using a significantly reduced pilot overhead. However, most of the above contributions consider a narrowband frequency-flat mmWave MIMO channel. Recent studies, such as [10], [13], [27]–[32], have focused on the estimation of a practical wideband frequency-selective mmWave MIMO channel. In this context, the authors of [29] develop a distributed compressive sensing (DCS)-based sparse channel estimation scheme for a wideband multiuser uplink scenario. However, a key drawback of their scheme is that it only estimates the line-of-sight (LoS) channel. In [13], the authors represented the frequency-selective mmWave MIMO channel in the sparse beamspace domain by employing a 'sparsifying-dictionary' matrix comprising the array response vectors from a suitably chosen angular-grid for the AoA/ AoD spaces. The pertinent sparse signal recovery problem is then solved using the OMP technique, which selects the columns of the 'sparsifying-dictionary' matrix in a greedy manner to obtain the best approximation of the received pilot vector [13]. However, the performance of the OMP algorithm is heavily dependent on the selection of the stopping criterion and the dictionary-matrix employed, with minor variations leading to the potential error propagation and convergence errors that degrade the performance. The wideband mmWave MIMO channel also exhibits a *spatial-wideband* effect [32], which arises due to the non-negligible time delay of the electromagnetic waves across the large antenna array aperture. Moreover, it is sparse in the delay-domain [31], [32]. Most of the existing literature ignores this spatial-selectivity. The work in [32] considers the dual wideband effects, i.e. spatial- and frequency-wideband effects. Efficient uplink and downlink CSI estimation approaches are developed that exploit the beamspace domain sparsity together with the delay-domain sparsity, and require significantly reduced training overheads. On the other hand, most of the existing contributions in mmWave do not consider time-selectivity arising due to mobility. In this context, a recent contribution [31] exploits the above-mentioned doubly-sparse nature of the mmWave MIMO channel for time- and frequency-selective, i.e. doubly-selective CSI estimation, by suitably designing the training pattern. The proposed doubly-sparse doubly-selective (DSDS) CSI estimator proceeds along the following lines: 1) random probing followed by identifying dominant channel taps via the well-established energy detector, 2) employing an adaptive-block OMP (A-BOMP) to recover the dominant angular directions, 3) finally, the path gains and Doppler shifts are estimated using steering-probing. The authors of [30] also exploit the joint angular and delay domain sparsity of the wideband mmWave MIMO channel, considering a MIMO system having low-resolution ADCs. The channel estimation schemes proposed therein are based on the approximate message passing (AMP) algorithm. A key drawback of the many schemes proposed in the existing literature is that they estimate the mmWave MIMO channel only after receiving all the pilot beams, thereby increasing the processing delay. Moreover, the existing sparse channel estimation schemes such as [12], [13], [15] are computationally complex due to the requirement of several matrix inversions. A brief summary of the various existing contributions is presented in a tabular format in Table-I. Furthermore, single-carrier (SC) transmission is favored over multicarrier orthogonal frequency division multiplexing (OFDM) in uplink [33], [34], since the mmWave transmitters at the users are designed to transmit with low peak to average power ratio (PAPR). Against this background, we develop the novel wideband mmWave hybrid MIMO channel estimation techniques for overcoming the drawbacks of the existing schemes listed above. In this context, low complexity block least-mean square (BLMS)-based [35]–[37] online channel estimation techniques capable of accurate tracking and learning capabilities are eminently suitable for practical wideband mmWave hybrid MIMO systems. Moreover, these schemes only rely on the instantaneous estimates of the 2nd-order

statistics of the signal, namely the cross-covariance vector and covariance matrix, but do not require prior knowledge of these quantities. Thus, these benefits can be exploited for channel estimation in a wideband mmWave hybrid MIMO system, which is the focus of this study. The key contributions of this treatise are summarized next.

B. Contributions of the Paper

1) Motivated by the inherent tracking capability and low computational complexity, this paper develops novel BLMS-based adaptive channel estimation schemes for SC wideband mmWave hybrid MIMO systems. To this end, a computationally-efficient frequency-domain equalization (FDE)-based channel estimation model is developed for wideband mmWave hybrid MIMO systems, which substantially simplifies the channel estimation procedure and the subsequent data detection process, both of its time-domain counterpart.

2) A frequency-domain sparse channel estimation model is developed using the beamspace domain model of the mmWave MIMO channel, which utilizes a suitable sparsifying-dictionary comprised of the quantized receive and transmit array response vectors.

3) Next, a stochastic gradient based fast BLMS (FBLMS) framework is developed for adaptive online channel estimation in SC-FDE-based wideband mmWave hybrid MIMO systems, which minimizes the mean squared observation error (MSOE). The proposed FBLMS technique does not require knowledge of the first and second order statistical information of the wideband mmWave MIMO channel, and therefore it is eminently suited for both stationary and non-stationary environments, strengthening its practical importance. Analytical expressions are also derived for characterizing both the finite frame length and asymptotic mean squared estimation error (MSEE) as well as MSOE performance of the proposed FBLMS technique.

4) In order to additionally exploit the sparsity for attaining an improved performance, a novel sparse-FBLMS (SF-BLMS) framework is developed, which makes use of a regularized cost function by incorporating a sparsity-inducing penalty term along with the standard MSOE. The corresponding analytical MSEE and MSOE results are also developed. Furthermore, a systematic approach is conceived for determining the regularization parameter that ensures lower MSEE of SFBLMS in comparison to FBLMS.

5) A hybrid precoder and combiner design is also proposed for the SC-FDE based wideband mmWave MIMO systems by employing the channel estimates obtained using the above techniques.

6) Analytical bound and simulation results are presented for characterizing the performance of the proposed BLMS-based schemes, which are benchmarked against the existing techniques.

C. Organization of the Work

The organization of this paper is as follows. Section-II presents the wideband SC-FDE-based mmWave hybrid MIMO system model, followed by the sparse channel estimation model. Our FBLMS technique is developed in Section-III, while Section-IV develops SFBLMS technique, followed by the proposed design of hybrid precoder and combiner in Section-V. This is followed by our simulation results in Section-VI. Section-VII concludes the paper, which is followed by the proofs of the results in the Appendices.

D. Notation

The quantity $\mathrm{diag}(a_1, a_2, \cdots, a_N)$ represents a diagonal matrix having the principal diagonal elements given by $a_1, a_2, \cdots, a_N$ and $\mathbf{I}_N$ denotes the $N \times N$ identity matrix. Superscripts $\mathbf{A}^T$, $\mathbf{A}^H$, $\mathbf{A}^*$ and $\mathbf{A}^{-1}$ denote the transpose, Hermitian, conjugation and inverse respectively. The vector equivalent of the matrix $\mathbf{A}$ is denoted by $\mathrm{vec}(\mathbf{A})$, which is formed by stacking the columns to form a single column vector. Similarly, $\mathrm{vec}^{-1}(\mathbf{a})$ denotes the corresponding inverse vectorization operation to provide the original matrix. The maximum eigenvalue of a matrix is denoted by $\lambda_{max}(\cdot)$ and the trace operator is denoted by $\mathrm{Tr}(\cdot)$, while $\mathbb{E}\{\cdot\}$ denotes the statistical expectation. The probability density function (pdf) of a complex Gaussian random vector having a mean vector of $\boldsymbol{\mu}$ and covariance matrix of $\boldsymbol{\Sigma}$ is denoted by $\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. $\mathcal{U}[a, b]$ denotes uniform distribution between the range a and b. Furthermore, $\nabla(\cdot)$ represents the gradient of a function, while $\|\mathbf{x}\|_p$ denotes the l_p norm of vector $\mathbf{x}$. A well-known property of the $\mathrm{vec}(\cdot)$ operator, given by $\mathrm{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A})\mathrm{vec}(\mathbf{B})$, is used in the paper, where $\otimes$ denotes the Kronecker product of two matrices.

II. SC WIDEBAND MMWAVE HYBRID MIMO SYSTEM AND CHANNEL MODEL

Consider a SC wideband mmWave hybrid MIMO system having N_t TAs, N_r RAs and N_{RF} RFCs both at the transmitter and receiver to transmit N_s data streams, where $N_s \leq N_{RF} << \min(N_t, N_r)$. The mmWave hybrid MIMO transmitter comprises a baseband precoder $\mathbf{F}_{BB} \in \mathbb{C}^{N_{RF} \times N_s}$ and RF precoder $\mathbf{F}_{RF} \in \mathbb{C}^{N_t \times N_{RF}}$ in cascade. Similarly, the receiver consists of an RF combiner $\mathbf{W}_{RF} \in \mathbb{C}^{N_r \times N_{RF}}$ and baseband combiner $\mathbf{W}_{BB} \in \mathbb{C}^{N_{RF} \times N_s}$. It is important to note that the RF precoder and combiner matrices $\mathbf{F}_{RF}$ and $\mathbf{W}_{RF}$, respectively, constitute a network of digitally controlled phase-shifters, i.e., their elements are constrained to constant magnitudes [2]. Without loss of generality, we assume $|\mathbf{F}_{RF}(i, j)| = \frac{1}{\sqrt{N_t}}, |\mathbf{W}_{RF}(i, j)| = \frac{1}{\sqrt{N_r}}$. Let us consider a wideband mmWave MIMO channel, which is frequency-selective with L delay taps, where the lth, $\forall 0 \leq l \leq L - 1$, delay tap is represented by the complex channel matrix $\mathbf{H}_l$ of size $N_r \times N_t$. Let furthermore $\mathbf{x}(\tilde{n}) \in \mathbb{C}^{N_s \times 1}$ denote the transmit symbol vector at time instant $\tilde{n}$. The output signal vector $\widetilde{\mathbf{y}}(\tilde{n}) \in \mathbb{C}^{N_s \times 1}$, after the baseband combining, is given by

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

TABLE I: Summary of literature survey on mmWave hybrid MIMO channel estimation

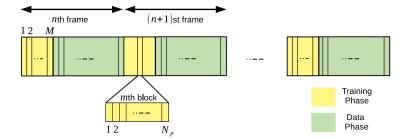|	[15]	[13]	[10]	[12]	[21]	[26]	[29]	[32]	[31]	Proposed
Spatial-sparsity	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Delay-domain sparsity	×	×	×	×	×	×	×	✓	✓	×
Time-selectivity	×	×	×	✓	×	✓	×	×	✓	✓
Frequency-selectivity	×	✓	✓	×	✓	✓	✓	✓	✓	✓
Cluster-based channel model	×	×	×	✓	×	✓	×	×	×	✓
Single-carrier (SC) system	×	✓	×	×	✓	✓	×	×	✓	✓
SC-FDE system	×	✓	×	×	×	×	×	×	×	✓
Low Complexity (Matrix-inversion not required)	×	×	×	×	×	×	×	×	×	✓
SC-FDE hybrid precoder/ combiner design	×	×	×	×	×	×	×	×	×	✓
Online and adaptive schemes	×	×	×	×	×	×	×	×	×	✓
Convergence analysis	×	×	×	×	×	×	×	×	×	✓

Fig. 1: Frame-structure for the proposed SC-FDE-based wideband mmWave hybrid MIMO channel estimation.
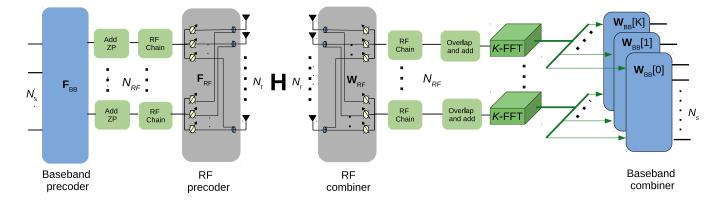

Fig. 2: Block diagram of frequency-domain signal processing in a SC-FDE-based wideband mmWave hybrid MIMO system.

where $\widetilde{\mathbf{x}}(\tilde{n}) = \mathbf{F}_{RF}\mathbf{F}_{BB}\mathbf{x}(\tilde{n}) \in \mathbb{C}^{N_t \times 1}$ and $\mathbf{v}(\tilde{n}) \in \mathbb{C}^{N_r \times 1}$ denotes the noise at the receiver that is distributed as $\mathcal{CN}(\mathbf{0}_{N_r \times 1}, \sigma^2\mathbf{I}_{N_r})$. The notation $*$ represents linear convolution between the mmWave MIMO channel taps $\mathbf{H}_l$ and the precoded symbol vector $\widetilde{\mathbf{x}}(\tilde{n})$. As observed in [13], owing to the higher complexity encountered in channel estimation and the subsequent data detection in the above time-domain (TD) model, one can employ the *overlap-add* method of [38] described next in order to obtain an equivalent frequency-domain (FD) representation.

A. SC-FD Channel Estimation Model

Consider the transmission frame-structure shown in Fig. 1. Each frame in this transmission scenario is comprised of training and data phases. The training phase in a frame is further divided into M blocks, each having N_p pilot vectors. The block diagram of the associated FD signal processing in a typical SC wideband mmWave hybrid MIMO system is illustrated in Fig. 2. Let $\mathbf{s}_{m,n}^{(p)} \in \mathbb{C}^{N_{RF} \times 1}$ denote the pth, $0 \leq p \leq N_p - 1$, complex pilot vector corresponding to the mth block in the nth frame. In each mth block, the Add ZP block appends $(L-1)$ zeros to the transmit pilot sequence of the individual RFC in order to generate a zero-padded (ZP) block of size $K = N_p + L - 1$. The sequence of channel taps $\mathbf{H}_l$ is also padded with $(N_p - 1)$ zero-matrices of size $N_r \times N_t$ in order to obtain a ZP block of size K. For compactness of notation, the ZP pilot vector and the mmWave MIMO channel tap sequences are described

as $\left\{\mathbf{s}_{m,n}^{(q)}\right\}_{q=0}^{K-1} = \{\mathbf{s}_{m,n}^{(0)}, \mathbf{s}_{m,n}^{(1)}, \cdots, \mathbf{s}_{m,n}^{(N_p-1)}, \underbrace{\mathbf{0}, \cdots, \mathbf{0}}_{L-1}\}$ and $\{\mathbf{H}_q\}_{q=0}^{K-1} = \{\mathbf{H}_0, \mathbf{H}_1, \cdots, \mathbf{H}_{L-1}, \underbrace{\mathbf{0}, \cdots, \mathbf{0}}_{N_p-1}\}$, respectively.

Let $\mathbf{F}_{RF,m} \in \mathbb{C}^{N_t \times N_{RF}}$ and $\mathbf{W}_{RF,m} \in \mathbb{C}^{N_r \times N_{RF}}$ represent the training RF precoder and combiner, respectively, corresponding to the mth block. The output signal vector $\mathbf{y}_{m,n}(q) \in \mathbb{C}^{N_{RF} \times 1}, 0 \le q \le K-1$, corresponding to the mth block in the nth frame is given as

$$\mathbf{y}_{m,n}(q) = \mathbf{W}_{RF,m}^H \left(\mathbf{H}_q \circledast_K \widetilde{\mathbf{s}}_{m,n}^{(q)} \right) + \mathbf{W}_{RF,m}^H \mathbf{v}_{m,n}(q), \quad (2)$$

where $\widetilde{\mathbf{s}}_{m,n}^{(q)} = \mathbf{F}_{RF,m} \mathbf{s}_{m,n}^{(q)}$ and $\mathbf{v}_{m,n}(q)$ denotes the noise vector distributed as $\mathcal{CN}(\mathbf{0}_{N_r \times 1}, \sigma^2 \mathbf{I}_{N_r})$. The notation $\circledast_K$ denotes the circular-convolution of the K-length sequences $\left\{\widetilde{\mathbf{s}}_{m,n}^{(q)}\right\}$ and $\{\mathbf{H}_q\}$. Let the mth TD output block be defined as $\{\mathbf{y}_{m,n}(q)\}_{q=0}^{K-1} = \{\mathbf{y}_{m,n}(0), \mathbf{y}_{m,n}(1), \cdots, \mathbf{y}_{m,n}(K-1)\}$. The equivalent FD representation of the wideband mmWave hybrid MIMO system is obtained by taking the K-point fast Fourier transform (FFT) of the above output block. Thus, the mth FD output block $\{\check{\mathbf{y}}_{m,n}[k]\}_{k=0}^{K-1}$ is obtained as

$$\begin{aligned} \{\check{\mathbf{y}}_{m,n}[k]\}_{k=0}^{K-1} &= \text{FFT}\left(\{\mathbf{y}_{m,n}(q)\}_{q=0}^{K-1} \right) \\ &= \{\check{\mathbf{y}}_{m,n}[0], \check{\mathbf{y}}_{m,n}[1], \cdots, \check{\mathbf{y}}_{m,n}[K-1]\}. \end{aligned} \quad (3)$$

Let $\mathbf{H}[k] \in \mathbb{C}^{N_r \times N_t}$ denote the mmWave MIMO channel frequency response (CFR) matrix at the kth subcarrier, which can be evaluated as

$$\mathbf{H}[k] = \sum_{l=0}^{L-1} \mathbf{H}_l e^{-j\frac{2\pi l}{K}k}. \quad (4)$$

The output vector $\check{\mathbf{y}}_{m,n}[k] \in \mathbb{C}^{N_{RF} \times 1}$ in (3) at the kth subcarrier index is given by

$$\check{\mathbf{y}}_{m,n}[k] = \mathbf{W}_{RF,m}^H \mathbf{H}[k] \mathbf{F}_{RF,m} \mathbf{s}_{m,n}[k] + \mathbf{W}_{RF,m}^H \mathbf{v}_{m,n}[k], \quad (5)$$

where $\mathbf{s}_{m,n}[k] \in \mathbb{C}^{N_{RF} \times 1}$ and $\mathbf{v}_{m,n}[k] \in \mathbb{C}^{N_r \times 1}$ represent the kth FFT-point of the transmit pilot sequence $\mathbf{s}_{m,n}^{(q)}$ and noise sequence $\mathbf{v}_{m,n}(q)$, respectively. Let $\mathbf{h}[k] = \text{vec}(\mathbf{H}[k]) \in \mathbb{C}^{N_r N_t \times 1}$ denote the vectorized mmWave MIMO CFR at the kth subcarrier. Upon exploiting properties of the $\text{vec}(\cdot)$ operator described in Section-I-D, Eq. (5) yields

$$\check{\mathbf{y}}_{m,n}[k] = \underbrace{\left(\mathbf{s}_{m,n}^T[k]\mathbf{F}_{RF,m}^T \otimes \mathbf{W}_{RF,m}^H\right)}_{\check{\boldsymbol{\Phi}}_{m,n}[k] \in \mathbb{C}^{N_{RF} \times N_r N_t}} \mathbf{h}[k] + \check{\mathbf{v}}_{m,n}[k], \quad (6)$$

where $\check{\mathbf{v}}_{m,n}[k] = \mathbf{W}_{RF,m}^H \mathbf{v}_{m,n}[k] \in \mathbb{C}^{N_{RF} \times 1}$ denotes the noise at the output of the combiner at the kth subcarrier, which is distributed as $\mathcal{CN}(\mathbf{0}, \sigma^2 \boldsymbol{\Gamma}_m)$ with $\boldsymbol{\Gamma}_m = K\mathbf{W}_{RF,m}^H \mathbf{W}_{RF,m}$. Finally, the whitened received signal $\mathbf{y}_{m,n}[k] = \boldsymbol{\Gamma}_m^{-1/2} \check{\mathbf{y}}_{m,n}[k]$ corresponding to the mth block is given by

$$\mathbf{y}_{m,n}[k] = \boldsymbol{\Phi}_{m,n}[k]\mathbf{h}[k] + \tilde{\mathbf{v}}_{m,n}[k], \quad (7)$$

where $\boldsymbol{\Phi}_{m,n}[k] = \boldsymbol{\Gamma}_m^{-1/2}\check{\boldsymbol{\Phi}}_{m,n}[k]$ denotes the sensing-matrix and $\tilde{\mathbf{v}}_{m,n}[k] = \boldsymbol{\Gamma}_m^{-1/2}\check{\mathbf{v}}_{m,n}[k]$. It can be readily verified that

the noise vector $\tilde{\mathbf{v}}_{m,n}[k]$ is distributed as $\mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_{N_{RF}})$. The next subsection details the mmWave MIMO specific clustered channel model that exploits the spatial-sparsity for channel recovery.

B. Sparse Channel Estimation Model for the Wideband mmWave Hybrid MIMO Systems

The lth delay tap $\mathbf{H}_l$ of a wideband frequency-selective mmWave MIMO channel can be modeled using the spatial channel model (SCM) formulated in [12], [14], [23] as

$$\mathbf{H}_l = \gamma \sum_{i=1}^{N_{cl}} \sum_{j=1}^{N_{ray,i}} \alpha_{ij} p\left(lT_s - \tau_{ij}\right) \mathbf{a}_R\left(\phi_{ij}\right) \mathbf{a}_T^H\left(\theta_{ij}\right), \quad (8)$$

where N_{cl} denotes the number of clusters, $N_{ray,i}$ represents the number of rays in the ith cluster and $\gamma = \sqrt{N_r N_t / N_{ray}}$ with $N_{ray} = \sum_{i=1}^{N_{cl}} N_{ray,i}$. The quantities T_s and $p(\cdot)$ represent the symbol duration and pulse shaping filter response, respectively. The 4-tuple $(\alpha_{ij}, \phi_{ij}, \theta_{ij}, \tau_{ij})$ denotes the complex channel gain, AoA, AoD and the delay corresponding to the jth spatial ray in the ith cluster. The vectors $\mathbf{a}_R(\phi_{ij}) \in \mathbb{C}^{N_r \times 1}$ and $\mathbf{a}_T(\theta_{ij}) \in \mathbb{C}^{N_t \times 1}$ denote the receive and transmit array response vectors corresponding to the AoA ϕ_{ij} and AoD θ_{ij}, respectively, and are expressed as

$$\begin{aligned} &\mathbf{a}_R\left(\phi_{ij}\right) \\ &= \frac{1}{\sqrt{N_R}} \left[1, e^{-j\frac{2\pi}{\lambda}d_R \cos\phi_{ij}}, \ldots, e^{-j\frac{2\pi}{\lambda}(N_R-1)d_R \cos\phi_{ij}}\right]^T, \end{aligned} \quad (9)$$

$$\begin{aligned} &\mathbf{a}_T\left(\theta_{ij}\right) \\ &= \frac{1}{\sqrt{N_T}} \left[1, e^{-j\frac{2\pi}{\lambda}d_T \cos\theta_{ij}}, \ldots, e^{-j\frac{2\pi}{\lambda}(N_T-1)d_T \cos\theta_{ij}}\right]^T, \end{aligned} \quad (10)$$

where λ denotes the carrier wavelength, and d_R and d_T represent the inter-antenna spacings of the uniform linear RA and TA array, respectively. Let $\mathbf{A}_R \in \mathbb{C}^{N_r \times N_{ray}}$ and $\mathbf{A}_T \in \mathbb{C}^{N_t \times N_{ray}}$ denote the receive and transmit array response matrices, respectively, defined as $\mathbf{A}_R = \left[\mathbf{a}_R\left(\phi_{11}\right), \mathbf{a}_R\left(\phi_{12}\right), \ldots, \mathbf{a}_R\left(\phi_{N_{cl}N_{ray,N_{cl}}}\right)\right]$ and $\mathbf{A}_T = \left[\mathbf{a}_T\left(\theta_{11}\right), \mathbf{a}_T\left(\theta_{12}\right), \ldots, \mathbf{a}_T\left(\theta_{N_{cl}N_{ray,N_{cl}}}\right)\right]$. The model in (8) can be compactly written as

$$\mathbf{H}_l = \mathbf{A}_R \mathbf{D}_l \mathbf{A}_T^H, \quad (11)$$

where $\mathbf{D}_l \in \mathbb{C}^{N_{ray} \times N_{ray}}$ is a diagonal matrix defined as

$$\mathbf{D}_l = \text{diag}\left(d_{11}, d_{12}, \cdots, d_{N_{cl}N_{ray,N_{cl}}}\right), \quad (12)$$

with $d_{ij} = \gamma \alpha_{ij} p(lT_s - \tau_{ij})$. As described in [2], [8], [16], [39], the number of clusters obeys $N_{cl} << \min(N_t, N_r)$, due to the reduced scattering, diffraction and highly directional signal propagation, coupled with the increased attenuation from blockage in the mmWave regime. This motivates the development of the sparse beamspace model for representing the wideband mmwave MIMO channel tap $\mathbf{H}_l$, which is described in the sequel.

Let the AoA and AoD spaces spanning the interval of $[0, \pi)$ be partitioned into grids of size G_r and G_t, respectively, where

we have $G_r, G_t \geq \max\{N_t, N_r\}$. Let the uniformly-spaced AoA-grid Φ_R and AoD-grid Θ_T be defined as

$$\Phi_R = \left\{ \phi_g : \phi_g = \frac{\pi(g-1)}{G_r}, \forall\, 1 \leq g \leq G_r \right\},$$

$$\Theta_T = \left\{ \theta_g : \theta_g = \frac{\pi(g-1)}{G_t}, \forall\, 1 \leq g \leq G_t \right\}.$$

Let furthermore $\mathbf{A}_T(\Theta_T) \in \mathbb{C}^{N_t \times G_t}$ and $\mathbf{A}_R(\Phi_R) \in \mathbb{C}^{N_r \times G_r}$ denote the dictionary matrices corresponding to the transmit and receive array responses, respectively, which are defined as $\mathbf{A}_R(\Phi_R) = [\mathbf{a}_R(\phi_1), \mathbf{a}_R(\phi_2), \ldots, \mathbf{a}_R(\phi_{G_r})]$ and $\mathbf{A}_T(\Theta_T) = [\mathbf{a}_T(\theta_1), \mathbf{a}_T(\theta_2), \ldots, \mathbf{a}_T(\theta_{G_t})]$. The beamspace domain representation of the wideband mmWave MIMO channel can thus be obtained as [2], [12], [15]

$$\mathbf{H}_l \approx \mathbf{A}_R(\Phi_R) \mathbf{H}_{b,l} \mathbf{A}_T^H(\Theta_T), \tag{13}$$

where $\mathbf{H}_{b,l} \in \mathbb{C}^{G_r \times G_t}$ represents the beamspace channel matrix corresponding to $\mathbf{H}_l$. Since, the number of clusters N_{cl} is much lower than the grid-size G_r and G_t, due to the reasons described earlier in this section, the beamspace channel matrix $\mathbf{H}_{b,l}$ is sparse, i.e., only a few of its entries are significant, with the rest being close to zero. Furthermore, when the quantization intervals π/G_r and π/G_t of the AoA and AoD spaces, respectively, are suitably fine, any column of the receive and transmit array response matrices $\mathbf{A}_R$ and $\mathbf{A}_T$ of the mmWave MIMO channel can be closely approximated by a column of the corresponding dictionary matrices $\mathbf{A}_R(\Phi_R)$ and $\mathbf{A}_T(\Theta_T)$, respectively. Thus, ignoring the effect of quantization errors in (13) and substituting it into (4), the mmWave MIMO matrix $\mathbf{H}[k]$ can be formulated as

$$\mathbf{H}[k] = \mathbf{A}_R(\Phi_R) \underbrace{\left(\sum_{l=0}^{L-1} \mathbf{H}_{b,l} e^{-j\frac{2\pi kl}{K}} \right)}_{\mathbf{H}_b[k]} \mathbf{A}_T^H(\Theta_T), \tag{14}$$

where $\mathbf{H}_b[k] \in \mathbb{C}^{G_r \times G_t}$ represents the beamspace CFR matrix corresponding to $\mathbf{H}[k]$. Since we know from (8) that the AoAs/AoDs corresponding to all the delay taps $\mathbf{H}_l$ are identical, this implies that the locations of the significant elements in $\mathbf{H}_{b,l}$ do not change across the delay taps. Furthermore, we also know from (14), the beamspace CFR matrix $\mathbf{H}_b[k], \forall\, 0 \leq k \leq K-1$, also exhibits an identical sparsity profile as that to $\mathbf{H}_{b,l}$. Note that although the channel impulse response (CIR) $\{\mathbf{H}_l\}_{l=0}^{L-1}$ is sparse in the delay domain [31], [32], the CFR $\{\mathbf{H}[k]\}_{k=0}^{K-1}$ obtained via FFT is not sparse in the subcarrier domain. Hence, in this work, the beamspace domain sparsity of the CFRs is exploited for CSI estimation, rather than capitalizing on the delay-domain sparsity. Once again, by exploiting the properties of the vec$(\cdot)$ operator, the vectorized mmWave MIMO channel $\mathbf{h}[k]$ can be expressed as

$$\mathbf{h}[k] = [\mathbf{A}_T^*(\Theta_T) \otimes \mathbf{A}_R(\Phi_R)] \mathbf{h}_b[k], \tag{15}$$

where $\mathbf{h}_b[k] = \text{vec}(\mathbf{H}_b[k]) \in \mathbb{C}^{G_r G_t \times 1}$ represents the vectorized beamspace CFR, which is sparse in nature. Substituting (15) in (7), the sparse channel estimation model of the system under consideration is given as

$$\mathbf{y}_{m,n}[k] = \mathbf{\Psi}_{m,n}[k] \mathbf{h}_b[k] + \tilde{\mathbf{v}}_{m,n}[k], \tag{16}$$

where $\mathbf{\Psi}_{m,n}[k] = \mathbf{\Phi}_{m,n}[k] [\mathbf{A}_T^*(\Theta_T) \otimes \mathbf{A}_R(\Phi_R)] \in \mathbb{C}^{N_{RF} \times G_r G_t}$ denotes the *equivalent sensing matrix*. The next section develops an *online* adaptive filtering based FBLMS scheme for estimating the wideband mmWave MIMO channel.

III. FBLMS-BASED WIDEBAND MMWAVE HYBRID MIMO CHANNEL ESTIMATION

Let $\widehat{\mathbf{h}}_{b,m,n}[k]$ denote estimate of the channel $\mathbf{h}_b[k]$ in the mth block of the nth frame. Let furthermore the corresponding MSOE cost function be defined as

$$\mathcal{J}_{m,n}[k] = \mathbb{E}\left\{ \left\| \mathbf{y}_{m,n}[k] - \mathbf{\Psi}_{m,n}[k]\widehat{\mathbf{h}}_{b,m,n}[k] \right\|^2 \right\}. \tag{17}$$

Employing the steepest-descent method [35], [36], the estimate $\widehat{\mathbf{h}}_{b,m+1,n}[k]$ is iteratively updated as

$$\widehat{\mathbf{h}}_{b,m+1,n}[k] = \widehat{\mathbf{h}}_{b,m,n}[k] - \frac{\mu}{2}\nabla_{\widehat{\mathbf{h}}_{b,m,n}[k]}(\mathcal{J}_{m,n}[k]), \tag{18}$$

where the quantity μ represents the step-size parameter. Upon defining $\mathbf{p} = \mathbb{E}\{\mathbf{\Psi}_{m,n}^H[k]\mathbf{y}_{m,n}[k]\} \in \mathbb{C}^{G_r G_t \times 1}$ and $\mathbf{R} = \mathbb{E}\{\mathbf{\Psi}_{m,n}^H[k]\mathbf{\Psi}_{m,n}[k]\} \in \mathbb{C}^{G_r G_t \times G_r G_t}$, the gradient of the cost function $\mathcal{J}_{m,n}[k]$ is given by

$$\nabla_{\widehat{\mathbf{h}}_{b,m,n}[k]}(\mathcal{J}_{m,n}[k]) = 2\mathbf{R}\widehat{\mathbf{h}}_{b,m,n}[k] - 2\mathbf{p}. \tag{19}$$

Upon substituting (19) in (18), the update equation is formulated as

$$\widehat{\mathbf{h}}_{b,m+1,n}[k] = \widehat{\mathbf{h}}_{b,m,n}[k] + \mu\left(\mathbf{p} - \mathbf{R}\widehat{\mathbf{h}}_{b,m,n}[k]\right). \tag{20}$$

It can be noted that in practical scenarios, it is rather challenging to obtain accurate estimates of the quantities $\mathbf{R}$ and $\mathbf{p}$, since this necessitates averaging over a large number of sample blocks. Thus, to avoid this difficulty, the proposed FBLMS technique employs a stochastic-gradient approach [35], [36], which replaces the quantities $\mathbf{R}$ and $\mathbf{p}$ by their instantaneous estimates $\widehat{\mathbf{R}}_{m,n} = \mathbf{\Psi}_{m,n}^H[k]\mathbf{\Psi}_{m,n}[k]$ and $\widehat{\mathbf{p}}_{m,n} = \mathbf{\Psi}_{m,n}^H[k]\mathbf{y}_{m,n}[k]$. Therefore, the FBLMS estimate $\widehat{\mathbf{h}}_{b,m+1,n}[k]$ of the wideband mmWave MIMO channel can be updated as

$$\widehat{\mathbf{h}}_{b,m+1,n}[k] = \widehat{\mathbf{h}}_{b,m,n}[k] + \mu\mathbf{\Psi}_{m,n}^H[k]\mathbf{e}_{m,n}[k], \tag{21}$$

where $\mathbf{e}_{m,n}[k] \in \mathbb{C}^{N_{RF} \times 1}$ denotes the instantaneous observation error vector obeying

$$\mathbf{e}_{m,n}[k] = \mathbf{y}_{m,n}[k] - \mathbf{\Psi}_{m,n}[k]\widehat{\mathbf{h}}_{b,m,n}[k]. \tag{22}$$

Finally, the mmWave MIMO CFR matrix estimate $\widehat{\mathbf{H}}_n^{\text{FBLMS}}[k]$ in the nth frame is expressed as

$$\widehat{\mathbf{H}}_n[k] = \mathbf{A}_R(\Phi_R) \text{vec}^{-1}\left(\widehat{\mathbf{h}}_{b,M,n}[k]\right) \mathbf{A}_T^H(\Theta_T), \tag{23}$$

where $\widehat{\mathbf{h}}_{b,M,n}[k]$ denotes the final estimate of the beamspace CFR vector in the nth frame. The proposed FBLMS technique is initialized as below:

$$\widehat{\mathbf{h}}_{b,1,0}[k] = \mathbf{0}_{G_r G_t \times 1}, \ \ \widehat{\mathbf{h}}_{b,1,n}[k] = \widehat{\mathbf{h}}_{b,M+1,n-1}[k]. \tag{24}$$

A compact representation of the steps involved in the FBLMS technique for wideband mmWave MIMO channel estimation is provided in Algorithm-1. The next subsection derives the analytical expressions of the resultant MSEE and MSOE of the proposed FBLMS technique for wideband mmWave MIMO CFR estimation.

Algorithm 1: FBLMS-based wideband mmWave hybrid MIMO channel estimation

Input: Observation $\mathbf{y}_{m,n}[k] \in \mathbb{C}^{N_{RF}\times 1}$, sensing matrix $\boldsymbol{\Psi}_{m,n}[k] \in \mathbb{C}^{N_{RF}\times G_r G_t}$

Output: $\widehat{\mathbf{H}}_n^{\text{FBLMS}}[k]$

1 **Initialization:** $\widehat{\mathbf{h}}_{b,M+1,-1}[k] = \mathbf{0}_{G_r G_t \times 1}$
2 **for** $n = 0,1,2,\cdots$ **do**
3    $\widehat{\mathbf{h}}_{b,1,n}[k] = \widehat{\mathbf{h}}_{b,M+1,n-1}[k]$
4   **for** $m = 1,2,\cdots,M$ **do**
5      $\mathbf{e}_{m,n}[k] = \mathbf{y}_{m,n}[k] - \boldsymbol{\Psi}_{m,n}[k]\widehat{\mathbf{h}}_{b,m,n}[k]$
6      $\widehat{\mathbf{h}}_{b,m+1,n}[k] = \widehat{\mathbf{h}}_{b,m,n}[k] + \mu\boldsymbol{\Psi}_{m,n}^H[k]\mathbf{e}_{m,n}[k]$
7   **end**
8   **return:** $\widehat{\mathbf{H}}_n^{\text{FBLMS}}[k] =$
     $\mathbf{A}_R(\Phi_R)\,\text{vec}^{-1}\left(\widehat{\mathbf{h}}_{b,M,n}[k]\right)\mathbf{A}_T^H(\Theta_T)$
9 **end**

A. Mean Squared Estimation Error (MSEE) for FBLMS

In order to simplify the analysis, the subcarrier index k is dropped, since the analysis is similar for each subcarrier. The result below, pertaining to the step-size parameter μ, ensures the convergence of the ensemble mean of the estimation error $\boldsymbol{\epsilon}_{m,n}$ that is defined as

$$\boldsymbol{\epsilon}_{m,n} = \widehat{\mathbf{h}}_{b,m,n} - \mathbf{h}_{b,\text{opt}}, \tag{25}$$

where $\mathbf{h}_{b,\text{opt}}$ and $\widehat{\mathbf{h}}_{b,m,n}$ denote the underlying beamspace channel and its FBLMS estimate, respectively.

Lemma 1. *A sufficient condition for the convergence of the ensemble mean of the estimation error $\boldsymbol{\epsilon}_{m,n}$ to zero, i.e, $\lim_{n\to\infty} \mathbb{E}[\boldsymbol{\epsilon}_{m,n}] = \mathbf{0}_{G_r G_t \times 1}$, is given by*

$$0 < \mu < \frac{2}{Tr(\mathbf{R})}. \tag{26}$$

Proof. Given in Appendix-A. $\square$

Let the MSEE of the proposed FBLMS technique be defined as

$$\mathcal{E}_{m,n}^{\text{FBLMS}} = \mathbb{E}\left\{\text{Tr}\left[\boldsymbol{\epsilon}_{m,n}\boldsymbol{\epsilon}_{m,n}^H\right]\right\}. \tag{27}$$

Let furthermore the eigenvalue decomposition of the matrix $\mathbf{R}$ be given by $\mathbf{R} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^H$, where the diagonal matrix $\boldsymbol{\Lambda}$ is comprised of the eigenvalues $\lambda_i, 1 \le i \le G_r G_t$, on its principal diagonal. Let $\tilde{\boldsymbol{\epsilon}}_{m,n} = \mathbf{U}^H\boldsymbol{\epsilon}_{m,n}$ denote the transformed estimation error. Furthermore, let $\boldsymbol{\Omega} = [\mathbf{A}_T^*(\Theta_T) \otimes \mathbf{A}_R(\Phi_R)]$ represent the sparsifying dictionary matrix, so that $\mathbf{h} = \boldsymbol{\Omega}\mathbf{h}_b$, with $\mathbf{K}_{m,n} = \mathbb{E}\left[\tilde{\boldsymbol{\epsilon}}_{m,n}\tilde{\boldsymbol{\epsilon}}_{m,n}^H\right]$ denoting the covariance matrix of the transformed estimation error vector in the beamspace domain, while the quantity $\tilde{\boldsymbol{\epsilon}}_{1,0} = \mathbf{U}^H\left(\widehat{\mathbf{h}}_{b,1,0} - \mathbf{h}_{b,\text{opt}}\right)$, where $\widehat{\mathbf{h}}_{b,1,0}$ represents the initial estimate for the 0th frame. The Lemma below determines the resultant MSEE for the proposed FBLMS technique.

Lemma 2. *The MSEE $\mathcal{E}_{m,n}^{FBLMS}$ for the FBLMS framework, corresponding to block m and frame n is given by*

$$\mathcal{E}_{m,n}^{FBLMS} = Tr\left[\boldsymbol{\Omega}\mathbf{U}\mathbf{K}_{m,n}\mathbf{U}^H\boldsymbol{\Omega}^H\right], \tag{28}$$

where $\mathbf{K}_{m,n}$ is given by the following expression

$$\mathbf{K}_{m,n} = \bar{\mathbf{I}}^{s-1}\mathbf{K}_{1,0}\bar{\mathbf{I}}^{s-1} + \boldsymbol{\Delta}_{m,n}. \tag{29}$$

The quantity $\bar{\mathbf{I}} = (\mathbf{I}_{G_r G_t} - \mu\boldsymbol{\Lambda})$ and $\boldsymbol{\Delta}_{m,n}$ in (29) is a diagonal matrix whose ith diagonal element is given by

$$\boldsymbol{\Delta}_{m,n}(i,i) = -\frac{\mu\sigma^2\left(1 - \mu\lambda_i\right)^{2(s-1)}}{2 - \mu\lambda_i} + \frac{\mu\sigma^2}{2 - \mu\lambda_i}, \tag{30}$$

where $s = Mn + m$. Furthermore, the asymptotic MSEE $\mathcal{E}_\infty^{FBLMS} \triangleq \lim_{n\to\infty} \mathbb{E}\left[\|\tilde{\boldsymbol{\epsilon}}_{m,n}\|^2\right]$ in the beamspace domain can be derived as

$$\mathcal{E}_\infty^{FBLMS} = Tr\left[\boldsymbol{\Delta}_\infty\right], \tag{31}$$

where the quantity $\boldsymbol{\Delta}_\infty$ is a diagonal matrix with its ith diagonal elements given by

$$\boldsymbol{\Delta}_\infty(i,i) = \frac{\mu\sigma^2}{2 - \mu\lambda_i}. \tag{32}$$

Proof. Appendix-A provides the detailed proof. $\square$

B. Mean Squared Observation Error (MSOE) for FBLMS

Let $\mathcal{J}_{\min}$ be defined as

$$\mathcal{J}_{\min} = \mathbb{E}\left\{\|\mathbf{y}_{m,n} - \boldsymbol{\Psi}_{m,n}\mathbf{h}_{b,\text{opt}}\|^2\right\}, \tag{33}$$

which denotes the minimum MSOE achieved by employing the true beamspace channel $\mathbf{h}_{b,\text{opt}}$. It follows from (16) that $\mathcal{J}_{\min} = \sigma^2 N_{RF}$. The Lemma below derives the MSOE for the proposed FBLMS technique.

Lemma 3. *The MSOE $\mathcal{J}_{m,n}^{FBLMS}$ for the FBLMS framework is given by*

$$\mathcal{J}_{m,n}^{FBLMS} = \mathcal{J}_{min} + \sum_{i=1}^{G_r G_t}\left[\lambda_i\mathbf{K}_{m,n}(i,i)\right]. \tag{34}$$

Furthermore, the asymptotic MSOE $\mathcal{J}_\infty^{FBLMS} \triangleq \lim_{n\to\infty}\mathcal{J}_{m,n}^{FBLMS}$ is given by

$$\mathcal{J}_\infty^{FBLMS} = \mathcal{J}_{min} + \sum_{i=1}^{G_r G_t}\left[\frac{\lambda_i\mu\sigma^2}{2 - \mu\lambda_i}\right]. \tag{35}$$

Proof. The proof is given in Appendix-B. $\square$

Furthermore, as described in [35] [Eq. (5.104)], the *average time constant* τ_{av} of the proposed FBLMS method is given by $\tau_{\text{av}} = \frac{G_r G_t}{2\mu Tr(\mathbf{R})}$. The *settling time* of the FBLMS technique, i.e. the time taken for the transients to die out, is proportional to τ_{av}. This implies that a higher value of the step-size parameter μ, while satisfying Lemma-1, yields a faster convergence. However, since the term $\frac{\mu}{2-\mu\lambda_i}$ appearing in Eq. (32) and (35) is an increasing function of μ, the asymptotic quantities MSEE $\mathcal{E}_\infty^{\text{FBLMS}}$ and MSOE $\mathcal{J}_\infty^{\text{FBLMS}}$, given by (31) and (35), respectively, increase for higher values of μ. Thus, there is a trade-off between the lower MSEE/ MSOE and faster rate of convergence for the proposed FBLMS algorithm, which depends on the choice of the step-size parameter μ. This is also illustrated by our simulations in Section-VI.

Note that the FBLMS technique estimates the wideband mmWave MIMO channel $\mathbf{h}[k]$ without exploiting the sparse nature of the multipath components in a typical mmWave MIMO channel, which is a unique signal propagation characteristic in the mmWave regime. Exploiting this aspect can further lead to a significant improvement in the quality of the estimated CSI. To this end, an online adaptive filtering based sparsity inducing SFBLMS framework is developed next for estimating the wideband mmWave MIMO channel.

IV. SFBLMS-based Sparse Channel Estimation in Wideband mmWave Hybrid MIMO Systems

The problem of estimating the sparse beamspace channel vector $\mathbf{h}_b[k]$ can now be formulated as

$$\underset{\mathbf{h}_b[k]}{\text{minimize}} \qquad \|\mathbf{h}_b[k]\|_0 \tag{36}$$

$$\text{subject to} \qquad \|\mathbf{y}_{m,n}[k] - \boldsymbol{\Psi}_{m,n}[k]\mathbf{h}_b[k]\|_2^2 \le \epsilon, \tag{37}$$

where ϵ is a configurable parameter. The above optimization problem can be readily observed to be non-convex due to the minimization of the l_0-norm. Although there exist a variety of sparse signal recovery algorithms, such as the SBL [12], [22] and OMP [15] etc., these are computationally complex and require multiple matrix inversions in each iteration. Additionally, they are *offline* in nature, which leads to a high processing delay. These challenges are addressed by the SFBLMS framework developed next.

The regularized cost function [40], [41] for the corresponding MSOE is defined as

$$\bar{\mathcal{J}}_{m,n}[k] = \mathbb{E}\left\{\left\|\mathbf{y}_{m,n}[k] - \boldsymbol{\Psi}_{m,n}[k]\widehat{\mathbf{h}}_{b,m,n}[k]\right\|^2\right\} + \delta' f\left(\widehat{\mathbf{h}}_{b,m,n}[k]\right), \tag{38}$$

where $f(\cdot)$ represents the sparsity-inducing penalty term and δ' denotes the regularization parameter. Upon employing the steepest descent algorithm [35], the estimate $\widehat{\mathbf{h}}_{b,m+1,n}[k]$ can be iteratively updated as

$$\widehat{\mathbf{h}}_{b,m+1,n}[k] = \widehat{\mathbf{h}}_{b,m,n}[k] - \frac{\mu}{2}\nabla_{\widehat{\mathbf{h}}_{b,m,n}[k]}\left(\bar{\mathcal{J}}_{m,n}[k]\right). \tag{39}$$

The gradient of the cost function above is formulated as

$$\nabla_{\widehat{\mathbf{h}}_{b,m,n}[k]}\left(\bar{\mathcal{J}}_{m,n}[k]\right) = 2\mathbf{R}\widehat{\mathbf{h}}_{b,m,n}[k] - 2\mathbf{p} - \delta\mathbf{g}\left(f(\widehat{\mathbf{h}}_{b,m,n}[k])\right), \tag{40}$$

where $\delta = \dfrac{\mu\delta'}{2}$ denotes the regularization step-size and $\mathbf{g}\left(f(\widehat{\mathbf{h}}_{b,m,n}[k])\right) = \nabla_{\widehat{\mathbf{h}}_{b,m,n}[k]}\left(f(\widehat{\mathbf{h}}_{b,m,n}[k])\right)$ represents the gradient of the sparsity-inducing penalty function $f(\cdot)$. Thus, the update equation in (39) can be simplified to

$$\widehat{\mathbf{h}}_{b,m+1,n}[k] = \widehat{\mathbf{h}}_{b,m,n}[k] + \mu\left(\mathbf{p} - \mathbf{R}\widehat{\mathbf{h}}_{b,m,n}[k]\right) - \delta\mathbf{g}\left(f(\widehat{\mathbf{h}}_{b,m,n}[k])\right). \tag{41}$$

Similar to the FBLMS framework developed Section-III, the quantities $\mathbf{R}$ and $\mathbf{p}$ can now be replaced by their instantaneous approximations $\boldsymbol{\Psi}_{m,n}^H[k]\boldsymbol{\Psi}_{m,n}[k]$ and $\boldsymbol{\Psi}_{m,n}^H[k]\mathbf{y}_{m,n}[k]$,

respectively. Following some rearrangement of the terms, it is simplified to

$$\widehat{\mathbf{h}}_{b,m+1,n}[k] = \widehat{\mathbf{h}}_{b,m,n}[k] + \mu\boldsymbol{\Psi}_{m,n}^H[k]\mathbf{e}_{m,n}[k] - \delta\mathbf{g}\left(f(\widehat{\mathbf{h}}_{b,m,n}[k])\right), \tag{42}$$

where $\mathbf{e}_{m,n}[k]$, similar to (22), denotes the instantaneous observation error. Lastly, the mmWave MIMO CFR matrix estimate $\widehat{\mathbf{H}}_n[k]$ for the nth frame can be recovered from $\widehat{\mathbf{h}}_{b,M,n}[k]$ using (23). In this contribution, we consider both l_0-norm approximation and l_1-norm based sparsity-inducing penalty functions. The analysis of each is described in separate subsections below.

A. SFBLMS using l_0-norm Approximation (SFBLMS-l_0)

The l_0-norm penalty function, denoted by $f_0(\cdot)$, is defined as

$$f_0\left(\widehat{\mathbf{h}}_{b,m,n}[k]\right) = \left\|\widehat{\mathbf{h}}_{b,m,n}[k]\right\|_0 = \sum_{i=1}^{G_r G_t} \mathcal{I}\left(\left|\widehat{\mathbf{h}}_{b,m,n}[k](i)\right| > 0\right), \tag{43}$$

where $\mathcal{I}(\cdot)$ denotes the indicator function. Since, the l_0-norm defined above is non-convex, we can employ a suitable approximation, similar to the one described in [40], [41], as

$$\sum_{i=1}^{G_r G_t} \mathcal{I}\left(\left|\widehat{\mathbf{h}}_{b,m,n}[k](i)\right| > 0\right) \approx \sum_{i=1}^{G_r G_t} \left(1 - e^{-\alpha\left(|\widehat{\mathbf{h}}_{b,m,n}[k](i)|\right)}\right), \tag{44}$$

where α is a parameter that defines the accuracy of the approximation above. Thus, the ith element of the gradient term $\mathbf{g}\left(f_0(\widehat{\mathbf{h}}_{b,m,n}[k])\right)$, denoted by $\left[\mathbf{g}\left(f_0(\widehat{\mathbf{h}}_{b,m,n}[k])\right)\right]_{(i)}$, is obtained as

$$\left[\mathbf{g}\left(f_0(\widehat{\mathbf{h}}_{b,m,n}[k])\right)\right]_{(i)} = \alpha e^{-\alpha\left(|\widehat{\mathbf{h}}_{b,m,n}[k](i)|\right)} \times \operatorname{sgn}\left(\widehat{\mathbf{h}}_{b,m,n}[k](i)\right). \tag{45}$$

Substituting (45) into (42) yields the update equation for SFBLMS-l_0.

B. SFBLMS using l_1-norm (SFBLMS-l_1)

The l_1-norm penalty function, denoted as $f_1(\cdot)$, is defined as [40], [41]

$$f_1\left(\widehat{\mathbf{h}}_{b,m,n}[k]\right) = \left\|\widehat{\mathbf{h}}_{b,m,n}[k]\right\|_1 = \sum_{i=1}^{G_r G_t} \left|\widehat{\mathbf{h}}_{b,m,n}[k](i)\right|. \tag{46}$$

The ith element of the gradient term $\mathbf{g}\left(f_1(\widehat{\mathbf{h}}_{b,m,n}[k])\right)$, denoted by $\left[\mathbf{g}\left(f_1(\widehat{\mathbf{h}}_{b,m,n}[k])\right)\right]_{(i)}$, is computed as

$$\left[\mathbf{g}\left(f_1(\widehat{\mathbf{h}}_{b,m,n}[k])\right)\right]_{(i)} = \operatorname{sgn}\left(\widehat{\mathbf{h}}_{b,m,n}[k](i)\right). \tag{47}$$

Once again, substituting (47) into (42), yields the update equation for SFBLMS-l_1. A step-by-step procedure along with the initialization required for the proposed SFBLMS technique is given in Algorithm-2. Although the conventional initialization

---

Algorithm 2: SFBLMS-based wideband mmWave hybrid MIMO channel estimation

---

Input: Observation $\mathbf{y}_{m,n}[k] \in \mathbb{C}^{N_{RF} \times 1}$, sensing matrix $\boldsymbol{\Psi}_{m,n}[k] \in \mathbb{C}^{N_{RF} \times G_r G_t}$

Output: $\widehat{\mathbf{H}}_n^{\text{SFBLMS}}[k]$

1 **Initialization:** $\widehat{\mathbf{h}}_{b,M+1,-1}[k] = \mathbf{0}_{G_r G_t \times 1}$

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

3     $\widehat{\mathbf{h}}_{b,1,n}[k] = \widehat{\mathbf{h}}_{b,M+1,n-1}[k]$

4   **for** $m = 1, 2, \cdots, M$ **do**

5        $\mathbf{e}_{m,n}[k] = \mathbf{y}_{m,n}[k] - \boldsymbol{\Psi}_{m,n}[k]\widehat{\mathbf{h}}_{b,m,n}[k]$

6   Update $\widehat{\mathbf{h}}_{b,m+1,n}[k]$ using (42)

7   **end**

8   **return:** $\widehat{\mathbf{H}}_n^{\text{SFBLMS}}[k] =$
       $\mathbf{A}_R(\Phi_R) \text{vec}^{-1}\left(\widehat{\mathbf{h}}_{b,M,n}[k]\right) \mathbf{A}_T^H(\Theta_T)$

9 **end**

$\widehat{\mathbf{h}}_{b,M+1,-1}[k] = \mathbf{0}_{G_r G_t \times 1}$ of the beamspace channel provides accurate estimates, as described in the simulation results of Section-VI, it typically requires a large number of frames for the convergence of the MSEE. To overcome this drawback, we can initialize the proposed SFBLMS framework with a suitable sparse estimate obtained using any of the existing sparse signal recovery techniques applied in the 0th frame. This technique is termed sparse-initialized (SI) SFBLMS (SI-SFBLMS) in the sequel and has been shown to have better convergence performance than the non-sparse initialization based SFBLMS.

C. MSEE for SFBLMS

For simplicity, once again the subcarrier index k is dropped. One can define the estimation error vector corresponding to the SFBLMS estimate $\widehat{\mathbf{h}}_{b,m,n}$ as

$$\check{\boldsymbol{\epsilon}}_{m,n} = \widehat{\mathbf{h}}_{b,m,n} - \mathbf{h}_{b,\text{opt}}. \qquad (48)$$

The condition for convergence of the ensemble mean of the estimation error $\check{\boldsymbol{\epsilon}}_{m,n}$ to zero, i.e $\lim_{n \to \infty} \mathbb{E}[\check{\boldsymbol{\epsilon}}_{m,n}] = \mathbf{0}_{G_r G_t \times 1}$, is once again given by Lemma-1. Hence, it is not repeated here. Let $\bar{\boldsymbol{\epsilon}}_{m,n} = \mathbf{U}^H \check{\boldsymbol{\epsilon}}_{m,n}$ denote the transformed estimation error vector. Furthermore, let the quantities $\mathbf{u}_s = \mathbf{u}_{m,n}$ and $\mathbf{t}_s = \mathbf{t}_{m,n}$ be defined as

$$\mathbf{u}_s = \mathbf{u}_{m,n} = \mu \mathbf{U}^H \boldsymbol{\Psi}_{m,n}^H \tilde{\mathbf{v}}_{m,n},$$
$$\mathbf{t}_s = \mathbf{t}_{m,n} = -\delta \mathbf{U}^H \mathbf{g}\left(f(\widehat{\mathbf{h}}_{b,m,n})\right), \qquad (49)$$

where the index $s = Mn + m$, or equivalently $n = \lfloor \frac{s-1}{M} \rfloor$ and $m = (s-1)_M + 1$. Here, $\lfloor \cdot \rfloor$ and $(\cdot)_M$ denote the floor and modulo-M operations, respectively. Let $\bar{\mathbf{K}}_{m,n} = \mathbb{E}[\bar{\boldsymbol{\epsilon}}_{m,n} \bar{\boldsymbol{\epsilon}}_{m,n}^H]$ denote the covariance matrix of the transformed estimation error vector $\bar{\boldsymbol{\epsilon}}_{m,n}$ and let $\bar{\boldsymbol{\epsilon}}_{1,0} = \mathbf{U}^H\left(\widehat{\mathbf{h}}_{b,1,0} - \mathbf{h}_{b,\text{opt}}\right)$, where $\widehat{\mathbf{h}}_{b,1,0}$ represents the initial estimate for the 0th frame. The Lemma below describes the MSEE for the proposed SFBLMS technique.

Lemma 4. *The MSEE $\mathcal{E}_{m,n}^{SFBLMS}$ for the SFBLMS framework is given as*

$$\mathcal{E}_{m,n}^{SFBLMS} = Tr\left[\boldsymbol{\Omega}\mathbf{U}\bar{\mathbf{K}}_{m,n}\mathbf{U}^H\boldsymbol{\Omega}^H\right], \qquad (50)$$

where $\bar{\mathbf{K}}_{m,n}$ is obtained by the following expression

$$\bar{\mathbf{K}}_{m,n} = \bar{\mathbf{I}}^{s-1}\bar{\mathbf{K}}_{1,0}\bar{\mathbf{I}}^{s-1} + \boldsymbol{\Delta}_{m,n} + \sum_{p=1}^{s-1}\bar{\mathbf{I}}^{s-1}\bar{\boldsymbol{\epsilon}}_{1,0}\mathbf{T}_p^H\bar{\mathbf{I}}^{s-1-p}$$
$$+ \sum_{p=1}^{s-1}\bar{\mathbf{I}}^{s-1-p}\mathbf{T}_p\bar{\boldsymbol{\epsilon}}_{1,0}^H\bar{\mathbf{I}}^{s-1} + \sum_{p,q=1}^{s-1}\bar{\mathbf{I}}^{s-1-p}\mathbf{K}_{u,t}^{p,q}\bar{\mathbf{I}}^{s-1-q}$$
$$+ \sum_{p,q=1}^{s-1}\bar{\mathbf{I}}^{s-1-q}\mathbf{K}_{t,u}^{q,p}\bar{\mathbf{I}}^{s-1-p} + \sum_{p,q=1}^{s-1}\bar{\mathbf{I}}^{s-1-p}\mathbf{K}_{t,t}^{p,q}\bar{\mathbf{I}}^{s-1-q},$$
$$(51)$$

relying on the quantities $\mathbf{K}_{t,t}^{p,q} = \mathbb{E}[\mathbf{t}_p\mathbf{t}_q^H]$, $\mathbf{K}_{u,t}^{p,q} = \mathbb{E}[\mathbf{u}_p\mathbf{t}_q^H]$, $\mathbf{K}_{t,u}^{p,q} = \mathbb{E}[\mathbf{t}_p\mathbf{u}_q^H]$ and $\mathbf{T}_p = \mathbb{E}[\mathbf{t}_p]$.

Proof. Given in Appendix-C $\qquad\square$

D. MSOE for SFBLMS

The following Lemma provides the MSOE corresponding to the proposed SFBLMS framework.

Lemma 5. *The MSOE $\mathcal{J}_{m,n}^{SFBLMS}$ for the mth block in the nth frame is given by*

$$\mathcal{J}_{m,n}^{SFBLMS} = \mathcal{J}_{min} + \sum_{i=1}^{G_r G_t}\left[\lambda_i \bar{\mathbf{K}}_{m,n}(i,i)\right]. \qquad (52)$$

Proof. The proof is similar to that of the FBLMS framework given in the Appendix-B. $\qquad\square$

E. Choice of the Regularization parameter δ

As seen in (42), the update equation of the SFBLMS framework employs a regularization parameter δ. The simulation results of Section-VI illustrate that the choice of the regularization parameter is critical and hence it should be appropriately chosen. The following Lemma provides a systematic bound on the choice of δ, which guarantees that the MSEE performance of SFBLMS is always better than that of the FBLMS.

Lemma 6. *The regularization parameter $\delta_{m,n}$ in the mth block of the nth frame, which ensures a lower MSEE of SFBLMS in comparison to FBLMS technique, is given by*

$$0 \leq \delta_{m,n} \leq \frac{2\Re\left\{\check{\boldsymbol{\epsilon}}_{m,n}^H(\mathbf{I}_{G_r G_t} - \mu\mathbf{R})\mathbf{g}\left(f(\widehat{\mathbf{h}}_{b,m,n}[k])\right)\right\}}{\left\|\mathbf{g}\left(f(\widehat{\mathbf{h}}_{b,m,n}[k])\right)\right\|^2}.$$
$$(53)$$

Proof. Given in our technical report in [42]. $\qquad\square$

The discussion related to the convergence behavior of the FBLMS technique, described after Lemma-3, also holds for the SFBLMS approach. The same is also illustrated by our simulations in Section-VI.

F. Complexity Analysis

The computational complexities incurred for each frame of the proposed FBLMS and SFBLMS schemes are evaluated next. The number of complex additions and multiplications required for estimating the beamspace channel vector $\mathbf{h}_b[k]$ using the proposed FBLMS, SFBLMS and the existing OMP [13] schemes are compared in Table-II. Observe from the table that the computational complexity of the channel estimation in each frame for both the FBLMS and SFBLMS schemes is in the order of $\mathcal{O}\left(MN_{RF}\right)$, while the same for the OMP is $\mathcal{O}\left[\left(MN_{RF}\right)^3\right]$, which is significantly higher in comparison to our BLMS-based schemes. In addition, as it will be demonstrated in the simulation results of Section-VI, the FBLMS and SFBLMS schemes provide an improved MSEE in comparison to the existing OMP technique, thus making them extremely well suited for practical wideband mmWave hybrid MIMO systems.

V. PRECODER AND COMBINER DESIGN FOR SC-FDE-BASED WIDEBAND MMWAVE HYBRID MIMO SYSTEMS

This section develops a comprehensive procedure to design the hybrid precoder and combiner at the transmitter and the receiver, respectively, of the SC-FDE-based wideband system using the CSI estimated through the approaches described in the previous sections. Let $\mathbf{s}[k] \in \mathbb{C}^{N_s \times 1}$ denote the baseband transmit symbol vector, which has been normalized to unit power so that $\mathbb{E}\left[\mathbf{s}[k]\mathbf{s}^H[k]\right] = \frac{1}{N_s}\mathbf{I}_{N_s}$. In the proposed SC-FDE-based mmWave hybrid MIMO system, the received symbol vector $\mathbf{y}[k] \in \mathbb{C}^{N_s \times 1}$, after the baseband combining, is obtained similar to (5) as

$$\mathbf{y}[k] = \mathbf{W}_{BB}^H[k]\mathbf{W}_{RF}^H\mathbf{H}[k]\mathbf{F}_{RF}\mathbf{F}_{BB}\mathbf{s}[k] + \mathbf{W}_{BB}^H[k]\mathbf{W}_{RF}^H\mathbf{v}[k],$$ (54)

where $\mathbf{v}[k] \in \mathbb{C}^{N_r \times 1} \sim \mathcal{CN}(\mathbf{0}, K\sigma^2\mathbf{I}_{N_r})$ is the receiver noise. Let the singular value decomposition (SVD) of the mmWave MIMO CFR matrix $\mathbf{H}[k]$ be given as

$$\mathbf{H}[k] = \mathbf{U}[k]\mathbf{\Sigma}[k]\mathbf{V}^H[k].$$ (55)

Note that for the mmWave MIMO channel $\mathbf{H}[k]$, the frequency-selective optimal digital precoder $\mathbf{F}^{\text{opt}}[k]$ is given as $\mathbf{F}^{\text{opt}}[k] = \mathbf{V}_1[k]$, where $\mathbf{V}_1[k]$ is comprised of the first N_s columns of the right singular matrix $\mathbf{V}[k]$. However, for the SC-FDE-based system model of (54), one has to design the subcarrier-independent hybrid precoders $\mathbf{F}_{RF}$ and $\mathbf{F}_{BB}$. The best approximation problem for the design of the optimal baseband and RF precoders $\mathbf{F}_{BB}^{\text{opt}}$ and $\mathbf{F}_{RF}^{\text{opt}}$, respectively, can be formulated as

$$\left(\mathbf{F}_{RF}^{\text{opt}}, \mathbf{F}_{BB}^{\text{opt}}\right) = \arg\min_{(\mathbf{F}_{RF}, \mathbf{F}_{BB})} \sum_{k=0}^{K-1}\left\|\mathbf{F}^{\text{opt}}[k] - \mathbf{F}_{RF}\mathbf{F}_{BB}\right\|_F^2.$$ (56)

As described in Appendix-D, we can reformulate the above design problem as

$$\left(\mathbf{F}_{RF}^{\text{opt}}, \mathbf{F}_{BB}^{\text{opt}}\right) = \arg\min_{(\mathbf{F}_{RF}, \mathbf{F}_{BB})}\left\|\mathbf{F}^{\text{opt}} - \mathbf{F}_{RF}\mathbf{F}_{BB}\right\|_F^2,$$ (57)

where $\mathbf{F}^{\text{opt}} = \frac{1}{K}\sum_{k=0}^{K-1}\mathbf{F}^{\text{opt}}[k]$. It can be readily observed that we have $\mathcal{C}\left(\mathbf{F}^{\text{opt}}\right) \subset \mathcal{R}\left(\mathbf{H}[k]\right)$, where $\mathcal{C}(\cdot)$ and $\mathcal{R}(\cdot)$ denote the column and row space of a matrix. This implies that the columns of the RF precoder $\mathbf{F}_{RF}$ can be suitably selected from the columns of the transmit array response dictionary matrix $\mathbf{A}_T(\Theta_T)$, which satisfy the constant magnitude constraints $|\mathbf{F}_{RF}(i,j)| = \frac{1}{\sqrt{N_t}}, |\mathbf{W}_{RF}(i,j)| = \frac{1}{\sqrt{N_r}}$. The popular simultaneous orthogonal matching pursuit (SOMP) technique [14], also described in our technical report [42], can now be employed to decompose the optimal frequency-flat precoder $\mathbf{F}^{\text{opt}}$ into the corresponding hybrid precoders $\mathbf{F}_{RF}$ and $\mathbf{F}_{BB}$. The design procedure of the frequency-flat RF combiner $\mathbf{W}_{RF}$ and of the frequency-selective baseband combiner $\mathbf{W}_{BB}[k]$ is presented next. In this context, the optimal digital MMSE combiner $\mathbf{W}_{\text{M}}[k]$ is given by

$$\mathbf{W}_{\text{M}}[k] = \mathbf{H}_e[k]\left(\mathbf{H}_e^H[k]\mathbf{H}_e[k] + \sigma^2 N_s\mathbf{I}_{N_s}\right)^{-1},$$ (58)

where $\mathbf{H}_e[k] = \mathbf{H}[k]\mathbf{F}_{RF}\mathbf{F}_{BB} \in \mathbb{C}^{N_r \times N_s}$. Let the concatenated MMSE combiner $\mathbf{W}_{\text{M}}$ and the baseband combiner $\mathbf{W}_{BB}$ be defined as

$$\mathbf{W}_{\text{M}} = [\mathbf{W}_{\text{M}}[0], \mathbf{W}_{\text{M}}[1], \cdots, \mathbf{W}_{\text{M}}[K-1]]$$
$$\mathbf{W}_{BB} = [\mathbf{W}_{BB}[0], \mathbf{W}_{BB}[1], \cdots, \mathbf{W}_{BB}[K-1]].$$ (59)

Since, the RF combiner $\mathbf{W}_{RF}$ is frequency-flat, the best approximation of the hybrid combiner is formulated as

$$\left(\mathbf{W}_{RF}^{\text{opt}}, \mathbf{W}_{BB}^{\text{opt}}\right) = \arg\min_{(\mathbf{W}_{RF}, \mathbf{W}_{BB})}\left\|\mathbf{W}_{\text{M}} - \mathbf{W}_{RF}\mathbf{W}_{BB}\right\|_F^2.$$ (60)

It can be once again verified that $\mathcal{C}\left(\mathbf{W}_{\text{M}}[k]\right) \subset \mathcal{C}\left(\mathbf{H}[k]\right), \forall 0 \leq k \leq K-1$, which implies that the columns of $\mathbf{W}_{RF}$ can be appropriately picked from the columns of $\mathbf{A}_R(\Phi_R)$. Thus, similar to (57), we can employ the SOMP technique for solving the hybrid combiner design problem stated above.

VI. SIMULATION RESULTS

A SC wideband mmWave hybrid MIMO system is considered with the number of TAs and RAs set to $N_t = N_r \in \{8, 16, 32\}$, the number of RFCs both at the transmitter and receiver side set to $N_{RF} \in \{4, 8\}$. The inter-antenna separation in the TA and RA arrays is $d_T = d_R = \frac{\lambda}{2}$. The wideband mmWave MIMO channel is assumed to be spatially-sparse having $N_{cl} \in \{4, 6\}$ clusters, $N_{ray,i} \in \{1, 4\}, \forall 1 \leq i \leq N_{cl}$, rays per cluster and the number of delay taps is set to $L = 4$. The complex path-gain α_{ij} and delay τ_{ij} corresponding to the jth spatial ray of the ith cluster is modeled as $\alpha_{ij} \sim \mathcal{CN}(0,1)$ and $\tau_{ij} \sim \mathcal{U}[0, (L-1)T_s]$. The raised-cosine pulse-shaping filter is considered with the roll-off factor set to 0.85. During channel estimation, the entries of the RF precoder and combiner are set as [13]

$$\mathbf{F}_{RF,m}(i,j) = \frac{1}{\sqrt{N_t}}e^{j\varphi_{i,j}} \text{ and } \mathbf{W}_{RF,m}(i,j) = \frac{1}{\sqrt{N_r}}e^{j\psi_{i,j}},$$

where the phases $\varphi_{i,j}$ and $\psi_{i,j}$ are drawn with uniform probability distribution from the set

$$\mathcal{A} = \left\{0, \frac{2\pi}{2^{N_Q}}, \cdots, \frac{\left(2^{N_Q}-1\right)2\pi}{2^{N_Q}}\right\},$$ (61)

TABLE II: Computational complexity of FBLMS, SFBLMS and OMP per frame

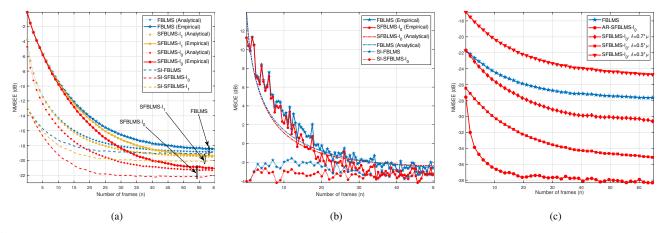| Algorithm	Multiplications	Additions
FBLMS	$M\left(2G_rG_tN_{RF} + N_{RF} + G_rG_t\right)$	$2MN_{RF}G_rG_t$
SFBLMS	$2MG_rG_t(N_{RF}+1)$	$MG_rG_t(2N_{RF}+1)$
OMP	$\frac{7}{3}(MN_{RF})^3 + \frac{7}{2}(MN_{RF})^2 + (G_rG_t+1)MN_{RF}$	$\frac{5}{2}(MN_{RF})^3 - (MN_{RF})^2 + (G_rG_t-1)MN_{RF} - G_rG_t$

Fig. 3: (a) NMSEE versus number of frames n performance for the mmWave hybrid MIMO system with simulation parameter $N_t = N_r = 16, N_{RF} = 4, G_r = G_t = 16, M = 20, N_{cl} = 4$, and $N_{ray,i} = 1$. The analytical NMSEEs for FBLMS and SFBLMS have been evaluated using Lemma-2 and Lemma-4, respectively. 'SI-' stands for sparse-initialized version of the proposed algorithms. (b) MSOE versus number of frames n performance for the mmWave hybrid MIMO system with simulation parameter $N_t = N_r = 16, N_{RF} = 4, G_r = G_t = 16, M = 20, N_{cl} = 4$, and $N_{ray,i} = 1$. The analytical MSOEs for FBLMS and SFBLMS have been evaluated using Lemma-3 and Lemma-5, respectively. (c) NMSEE versus number of frames n performance of the FBLMS and SFBLMS-l_0 scheme with varying regularizer δ for the mmWave hybrid MIMO system with simulation parameters $N_t = N_r = 32, N_{RF} = 8, G_r = G_t = 32, M = 20, N_{cl} = 4, N_{ray,i} = 1$ and $\mu = \frac{1}{10\lambda_{max}}$. For the adaptive regularization (AR)-based AR-SFBLMS-l_0, the regularizer δ is derived using Lemma-6.

with the angle quantization parameter of $N_Q = 3$. The number of pilot vectors in each block is set as $N_p = 13$, which implies that the FFT-block size is $K = N_p + L - 1 = 16$. The elements of the pilot symbol vector $\mathbf{s}_{m,n}^{(p)}$ are drawn from an 8-PSK (phase shift keying) constellation with an average power of unity. The AoA/ AoD space $[0,\pi)$ is quantized into $G_r = G_t \in \{16,32,64\}$ angular grid-points. The SNR is defined as $\frac{1}{\sigma^2}$. The parameter α in the l_0-approximation in (44) is set as $\alpha = 20$. The performance of the proposed and existing schemes are compared in terms of the normalized MSEE (NMSEE), which is defined as [10]

$$\text{NMSEE}_n = \frac{\sum_{k=0}^{K-1}\left\|\mathbf{H}[k] - \widehat{\mathbf{H}}_n[k]\right\|_F^2}{\sum_{k=0}^{K-1}\|\mathbf{H}[k]\|_F^2}. \quad (62)$$

Fig. 3(a) compares the NMSEE performance versus the number of frames (n) of the various BLMS-based schemes proposed in this paper. The simulation parameters are set as $N_t = N_r = 16, N_{RF} = 4, G_r = G_t = 16, M = 20$ and $N_{cl} = 4$. An *on-grid* scenario is considered, where the AoA/ AoDs corresponding to each spatial path are generated with uniform probability from the angular grid with $N_{ray,i} = 1$ ray per cluster. The SNR for this study is set to 15 dB. For the BLMS-based schemes, the step-size parameter μ is set to $\frac{1}{8\lambda_{max}}$, where λ_{max} denotes the maximum eigenvalue value of $\mathbf{R}$. The regularization parameter δ is set as $\delta = 0.08\mu$ for SFBLMS-l_0 and $\delta = 0.5\mu$ for SFBLMS-l_1. Observe from the figure

that the NMSEE performance of all the proposed BLMS-based schemes improve with the frame-index n. It can also be observed that the SFBLMS techniques outperform the FBLMS technique. This is due to the fact that the SFBLMS techniques exploit the spatial-sparsity inherent in the beamspace representation of the wideband mmWave MIMO channel, which the conventional FBLMS could not. Furthermore, the SFBLMS-l_0 variant of SFBLMS has a better NMSEE performance than SFBLMS-l_1. This is because the SFBLMS-l_0 employs the l_0-norm approximation given in (44), which yields a better fit to the original l_0-norm based optimization objective given in (36), in comparison to its l_1-norm approximation employed by SFBLMS-l_1. Additionally, it can also be observed that the theoretical NMSEE curves evaluated for the proposed FBLMS and SFBLMS techniques closely match their counterparts obtained via simulation. Finally, the proposed sparse initialized SI-FBLMS and SI-SFBLMS techniques, which utilize the initial estimate $\widehat{\mathbf{h}}_{b,1,0}[k]$ obtained from the OMP technique employed over the 0th frame, are seen to converge significantly faster in comparison to the zero-initialization given in (24), which is the conventional initialization employed for the BLMS-based schemes.

Fig. 3(b) presents the analytical and simulated MSOE versus the number of frames n for the proposed BLMS-based techniques for a mmWave setup with simulation parameters set as $N_t = N_r = 16, N_{RF} = 4, G_r = G_t = 16, M = 20$. The noise variance for this analysis is set to $\sigma^2 = -10$ dB, which

implies that the $\mathcal{J}_{\min} = -3.97$ dB. Once again, the MSOE for all the proposed schemes is seen to decrease upon increasing the frame-index n. Also, the theoretical MSOEs derived for all the proposed schemes coincide with their simulated MSOEs. Finally, the MSOE of the SI-SFBLMS technique is once again seen to converge significantly faster than the SFBLMS due to the sparse-initialization.

Fig. 3(c) compares the NMSEE of the proposed SFBLMS-l_0 technique, which employs a fixed regularization parameter δ, to that of the adaptive regularization (AR) based AR-SFBLMS-l_0 employing δ from Lemma-6. For this study, a mmWave setup is considered using simulation parameters of $N_t = N_r = 32, N_{RF} = 8, G_r = G_t = 32, M = 20$. From the plot, it can be readily seen that the SFBLMS-l_0 with an empirically tuned $\delta = 0.5\mu$ has the lowest NMSEE. Any other fixed value of δ results in a higher NMSEE. Furthermore, the AR-SFBLMS-l_0 uses the regularization parameter $\delta_{m,n}$ derived from Lemma-6, which ensures a lower instantaneous NMSEE in comparison to any other fixed value of δ.

Fig. 4(a) compares the NMSEE versus SNR performance of the proposed BLMS-based schemes upon convergence and that of the existing OMP technique for mmWave hybrid MIMO systems using the simulation parameters $N_t = N_r = 8, N_{RF} = 4, G_r = G_t = 16$ and $M = 20$. As the SNR increases, the NMSEE of all the schemes is seen to improve. The performance of the existing OMP scheme is seen to improve with increasing number of blocks M in each frame, thanks to the larger number of observations. Furthermore, it can be observed that the NMSEE performance of the BLMS-based techniques is notably better than that of the existing OMP scheme. Moreover, the latter has a significantly lower computational complexity, as derived in Section-IV-F. This is because the estimation accuracy of the existing OMP scheme is quite sensitive to the selection of the stopping threshold and to the equivalent sensing matrix $\boldsymbol{\Psi}_{m,n}[k]$, which lead to its poor overall performance. Thus, its performance is not robust, unlike that of the BLMS-based approaches. Any minor modification in the stopping criterion or the sensing matrix leads to structural and convergence errors [12]. It can also be observed that the SFBLMS based techniques, which also exploit the spatial-sparsity of the mmWave MIMO channel, outperform the FBLMS scheme. Fig. 4(a) also compares the on-grid and *off-grid* performance of the various techniques. The following simulation setup is considered for an off-grid scenario. The mean-angles of all the N_{cl} clusters are considered to be uniformly-distributed over the AoA/ AoD spaces. Furthermore, the AoA/ AoDs of the $N_{ray,i} = 4$ spatial-rays corresponding to a cluster are generated using Laplacian-distribution with standard deviation $\sigma_{as} = 0.05$ radian, around the mean-angle of that cluster. It can be readily observed in Fig. 4(a) that the performance of all the schemes slightly deteriorates in the off-grid scenario. This performance degradation arises because the actual AoAs/ AoD pairs for $N_{ray,i} = 4$ rays of the ith cluster differ from the grid-angles in the set Θ_T and Φ_R, which are used for constructing the equivalent sensing matrix $\boldsymbol{\Psi}_{m,n}[k]$. Finally, one can also observe an improved NMSEE performance for the BLMS-based techniques compared to the OMP in the off-grid scenario.

Fig. 4(b) demonstrates the tracking capabilities of the BLMS-based approaches in a non-stationary scenario for a wideband mmWave hybrid MIMO setup with simulation parameters of $N_t = N_r = 32, N_{RF} = 4, G_r = G_t = 64, M = 40$ at the SNR = 15 dB. The mmWave cellular system is assumed to operate in the Q-band at 28 GHz. At this carrier frequency, the user velocity of $v = 5$ km/h results in a Doppler shift of $f_D = 130$ Hz. Furthermore, the coherence-time T_c for the doubly-selective mmWave MIMO channel is set to $T_c = 1$ ms, while the frame-length T_F is set to $T_F = T_c/10$. These values can be substituted into Jake's well established model [43], to evaluate the temporal correlation coefficient ρ yielding $\rho = J_0(2\pi f_D T_F) \approx 0.9938$. For this mobility scenario, the time-selective complex path gains $\alpha_{ij,n}$ follow the first-order autoregressive (AR1) model given as

$$\alpha_{ij,n} = \rho\alpha_{ij,n-1} + \sqrt{1-\rho^2}u_{ij,n}, \tag{63}$$

where $u_{ij,n}$ represents the innovation noise that is assumed to be distributed as $\mathcal{CN}(0,1)$. Furthermore, the AoA/ AoDs corresponding to the spatial multipath components are changed at frame-indices of $n = 100, 200, 300$. It can be readily observed from the figure that the proposed BLMS-based techniques are able to track a time-varying channel, since when the AoA/ AoDs change, the algorithms converge within as few as $30 - 40$ frames. Hence they are eminently suited for the estimation of a practical non-stationary wideband mmWave MIMO channel.

Fig. 4(c) demonstrates the effect of step-size parameter μ on the converged NMSEE and on the rate of convergence for the proposed BLMS-based approaches. The simulation parameters are set to $N_t = N_r = 16, N_{RF} = 4, G_r = G_t = 16, M = 20$ and $N_{cl} = 4$. From the figure, it can be readily observed that the higher value of μ yields faster convergence, but it also leads higher NMSEE, as analytically described after Lemma-3. On the other hand, for a lower value of μ, the proposed BLMS-based algorithms converge to lower NMSEE, however their convergence rate slows down. Thus, the step-size parameter μ needs to be appropriately chosen to strike the best NMSEE versus rate of convergence trade-off.

Fig. 5 demonstrates the resultant bandwidth-efficiency (BE) with the estimated CSI from the various proposed and existing channel estimation schemes, which is defined as

$$\text{BE} = \sum_{k=0}^{K-1} \log_2\left(\left|\mathbf{I}_{N_s} + \mathbf{R}_n^{-1}\mathbf{H}_e[k]\mathbf{R}_x\mathbf{H}_e^H[k]\right|\right), \tag{64}$$

where $\mathbf{H}_e[k]$ is the equivalent channel for kth subcarrier given by $\mathbf{H}_e[k] = \mathbf{W}_{BB}^H[k]\mathbf{W}_{RF}^H \mathbf{H}[k]\mathbf{F}_{RF}\mathbf{F}_{BB}$. Furthermore, $\mathbf{R}_x$ is the transmit covariance matrix and $\mathbf{R}_n$ is the noise covariance matrix. The parameters of our mmWave system are set as $N_t = N_r = 8, N_s = N_{RF} = 4, G_r = G_t = 16, M = 20$ and $N_{cl} = 6$. The transmit covariance is set as $\frac{1}{N_s}\mathbf{I}_{N_s}$. For benchmarking the BE performance, the digital capacity of a hypothetical genie-receiver is considered, which uses the optimal digital precoder and MMSE combiner with the help of perfect knowledge of the mmWave MIMO CSI. The hybrid capacity obtained using the proposed hybrid precoder and combiner design discussed in Section-V using the perfect
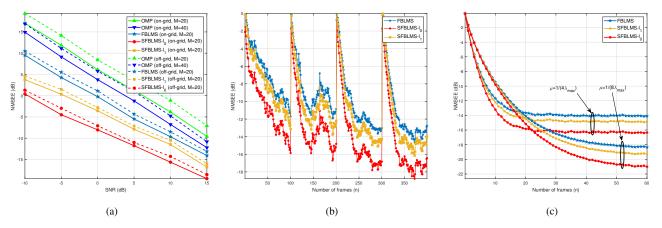
Fig. 4: (a) NMSEE versus SNR performance comparison of the BLMS-based schemes with the existing OMP scheme for the mmWave hybrid MIMO system with simulation parameters $N_t = N_r = 8, N_{RF} = 4, G_r = G_t = 16, N_{cl} = 4, N_{ray,i} \in \{1, 4\}$ and $\mu = \frac{1}{7\lambda_{\max}}$. (b) Tracking performance comparison of the BLMS-based schemes for the mmWave hybrid MIMO system with simulation parameters $N_t = N_r = 32, N_{RF} = 8, G_r = G_t = 64, M = 40, N_{cl} = 4, N_{ray,i} = 1$ and $\rho = 0.9983$. (c) NMSEE versus number of frames n performance with two different values of the step-size parameter μ for the mmWave hybrid MIMO system with simulation parameter $N_t = N_r = 16, N_{RF} = 4, G_r = G_t = 16, M = 20, N_{cl} = 4,$ and $N_{ray,i} = 1$.
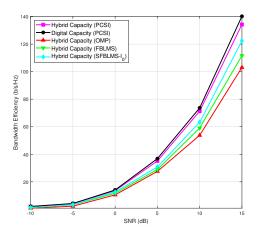

Fig. 5: BE versus SNR performance for the mmWave hybrid MIMO setup with $N_t = N_r = 8, N_{RF} = 4, G_r = G_t = 16, N_{cl} = 6, N_{ray,i} = 1, M = 20$.

CSI and CSI estimated by the various techniques are also shown therein. It can be verified that the enhanced mmWave MIMO CSI estimation accuracy of the proposed FBLMS and SFBLMS schemes is reflected in its improved BE. Moreover, the BE of the SFBLMS technique is close to that of various benchmarks.

VII. CONCLUSIONS

We developed an FBLMS-based low-complexity, adaptive online channel estimation procedure for SC-FDE-based wideband millimeter wave hybrid MIMO systems. Subsequently, the SFBLMS technique was presented, which also exploits the sparsity inherent in the beamspace domain model of the wideband frequency-selective mmWave MIMO channel, in addition to FBLMS, through a regularized cost function. This was seen to lead to a substantial improvement in the channel

estimation performance. The theoretical MSEE, MSOE and the corresponding expressions for their asymptotic values are also derived for both the schemes. A systematic procedure was also derived for determining the range of the feasible values of the regularization parameter that can be employed for SFBLMS. Finally, a hybrid precoder and combiner design is proposed for SC-FDE-based wideband mmWave MIMO system by employing the channel estimates obtained using the above techniques. Our simulation results illustrated the enhanced CSI estimation and tracking performance of the proposed BLMS-based techniques compared to the existing OMP-based technique. More importantly, the proposed techniques do not require additional information, such as the prior knowledge of AoA/AoDs. Future research may consider frequency-domain correlation across CFR [44], [45], and develop suitable schemes to exploit the delay-domain sparsity for estimating the wideband mmWave MIMO channel.

APPENDIX A

Let the estimation error corresponding to $\mathbf{h}$ be given by $\boldsymbol{\epsilon}'_{m,n} = \widehat{\mathbf{h}}_{m,n} - \mathbf{h}$, where $\widehat{\mathbf{h}}_{m,n} = \boldsymbol{\Omega}\widehat{\mathbf{h}}_{b,m,n}$. The relationship between $\boldsymbol{\epsilon}'_{m,n}$ and the estimation error in the beamspace domain $\boldsymbol{\epsilon}_{m,n}$ is given as $\boldsymbol{\epsilon}'_{m,n} = \boldsymbol{\Omega}\boldsymbol{\epsilon}_{m,n}$. The corresponding MSEE for FBLMS is given by

$$
\begin{aligned}
\mathcal{E}_{m,n}^{\text{FBLMS}} &= \text{Tr}\left[\mathbb{E}\left(\boldsymbol{\epsilon}'_{m,n}\boldsymbol{\epsilon}'^{H}_{m,n}\right)\right] \\
&= \text{Tr}\left[\boldsymbol{\Omega}\mathbf{U}\mathbb{E}\left(\tilde{\boldsymbol{\epsilon}}_{m,n}\tilde{\boldsymbol{\epsilon}}^{H}_{m,n}\right)\mathbf{U}^{H}\boldsymbol{\Omega}^{H}\right] \\
&= \text{Tr}\left[\boldsymbol{\Omega}\mathbf{U}\mathbf{K}_{m,n}\mathbf{U}^{H}\boldsymbol{\Omega}^{H}\right].
\end{aligned}
\tag{65}
$$

Subtracting $\mathbf{h}_{b,\text{opt}}$ from both sides in (21), the update equation for the estimation error $\boldsymbol{\epsilon}_{m+1,n}$ can be obtained as

$$
\boldsymbol{\epsilon}_{m+1,n} = \left(\mathbf{I}_{G_r G_t} - \mu\boldsymbol{\Psi}^{H}_{m,n}\boldsymbol{\Psi}_{m,n}\right)\boldsymbol{\epsilon}_{m,n} + \mu\boldsymbol{\Psi}^{H}_{m,n}\tilde{\mathbf{v}}_{m,n}.
\tag{66}
$$

Using the small step-size theory, the solution of the stochastic difference equation in (66) can be approximated by the solution of the following differential equation [35]

$$\boldsymbol{\epsilon}_{m+1,n} = (\mathbf{I} - \mu\mathbf{R})\,\boldsymbol{\epsilon}_{m,n} + \mu\boldsymbol{\Psi}_{m,n}^H \tilde{\mathbf{v}}_{m,n}. \tag{67}$$

Employing the eigenvalue decomposition $\mathbf{R} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^H$, the differential equation above can be transformed as

$$\tilde{\boldsymbol{\epsilon}}_{m+1,n} = (\mathbf{I} - \mu\boldsymbol{\Lambda})\,\tilde{\boldsymbol{\epsilon}}_{m,n} + \mu\mathbf{U}^H \boldsymbol{\Psi}_{m,n}^H \tilde{\mathbf{v}}_{m,n}. \tag{68}$$

Now, taking the expectation $\mathbb{E}\{\cdot\}$ on both the sides in (68), the ith element, $\forall\,1 \leq i \leq G_r G_t$, of $\tilde{\boldsymbol{\epsilon}}_{m,n}$ is obtained as

$$\mathbb{E}\left[\tilde{\boldsymbol{\epsilon}}_{m+1,n}(i)\right] = (1 - \mu\lambda_i)\,\mathbb{E}\left[\tilde{\boldsymbol{\epsilon}}_{m,n}(i)\right], \tag{69}$$

which follows from the fact that the noise vector $\tilde{\mathbf{v}}_{m,n}$ and the sensing matrix $\boldsymbol{\Psi}_{m,n}$ are statistically independent, thereby leading to

$$\mathbb{E}\left\{\boldsymbol{\Psi}_{m,n}^H \tilde{\mathbf{v}}_{m,n}\right\} = \mathbb{E}\left\{\boldsymbol{\Psi}_{m,n}^H\right\} \mathbb{E}\left\{\tilde{\mathbf{v}}_{m,n}\right\} = \mathbf{0}_{G_r G_t \times 1}. \tag{70}$$

The result in (69) can be rewritten as

$$\mathbb{E}\left[\tilde{\boldsymbol{\epsilon}}_{m+1,n}(i)\right] = (1 - \mu\lambda_i)^s\,\tilde{\boldsymbol{\epsilon}}_{1,0}(i), \tag{71}$$

where $s = Mn + m$. It can be readily observed that the condition $\lim_{n\to\infty}\mathbb{E}\left[\tilde{\boldsymbol{\epsilon}}_{m,n}\right] = \mathbf{0}_{G_r G_t \times 1}$ is satisfied if and only if $|1 - \mu\lambda_i| < 1, \forall\,i$. This imposes the following constraint on the value of the step-size parameter μ for ensuring convergence in the ensemble mean: $0 < \mu < \frac{2}{\lambda_{\max}}$, where $\lambda_{\max}$ is the maximum eigenvalue of $\mathbf{R}$. Furthermore, by exploiting the property $\lambda_{\max} \leq \mathrm{Tr}(\mathbf{R}) = \sum_{i=1}^{G_r G_t}\lambda_i$, a tighter bound on the upper limit of μ can be obtained as

$$0 < \mu < \frac{2}{\mathrm{Tr}(\mathbf{R})},$$

which completes the proof of Lemma-1.

Employing the definition of $\mathbf{u}_s$ from (49), Eq. (68) can now be rewritten as

$$\tilde{\boldsymbol{\epsilon}}_{m,n} = \bar{\mathbf{I}}^{s-1}\tilde{\boldsymbol{\epsilon}}_{1,0} + \sum_{p=1}^{s-1}\bar{\mathbf{I}}^{s-1-p}\mathbf{u}_p. \tag{72}$$

Furthermore, the quantity $\tilde{\boldsymbol{\epsilon}}_{m,n}\tilde{\boldsymbol{\epsilon}}_{m,n}^H$ can be derived as

$$\tilde{\boldsymbol{\epsilon}}_{m,n}\tilde{\boldsymbol{\epsilon}}_{m,n}^H = \bar{\mathbf{I}}^{s-1}\tilde{\boldsymbol{\epsilon}}_{1,0}\tilde{\boldsymbol{\epsilon}}_{1,0}^H\bar{\mathbf{I}}^{s-1} + \sum_{p=1}^{s-1}\bar{\mathbf{I}}^{s-1}\tilde{\boldsymbol{\epsilon}}_{1,0}\mathbf{u}_p^H\bar{\mathbf{I}}^{s-1-p}$$
$$+ \sum_{p=1}^{s-1}\bar{\mathbf{I}}^{s-1-p}\mathbf{u}_p\tilde{\boldsymbol{\epsilon}}_{1,0}^H\bar{\mathbf{I}}^{s-1}$$
$$+ \sum_{p,q=1}^{s-1}\bar{\mathbf{I}}^{s-1-p}\mathbf{u}_p\mathbf{u}_q^H\bar{\mathbf{I}}^{s-1-q}. \tag{73}$$

Now, applying the expectation operator $\mathbb{E}\{\cdot\}$ on the both sides of the above equation, the matrix $\mathbf{K}_{m,n}$ is obtained as

$$\mathbf{K}_{m,n} = \mathbb{E}\left[\tilde{\boldsymbol{\epsilon}}_{m,n}\tilde{\boldsymbol{\epsilon}}_{m,n}^H\right] = \bar{\mathbf{I}}^{s-1}\mathbf{K}_{1,0}\bar{\mathbf{I}}^{s-1} + \boldsymbol{\Delta}_{m,n}, \tag{74}$$

where the following simplifications have been employed. The 2nd and 3rd terms of (73) become zero due to the independence assumption. Furthermore, as described next, the expectation of the last term of (73) simplifies to $\boldsymbol{\Delta}_{m,n}$. Note

that $\boldsymbol{\Delta}_{m,n} = \sum_{p,q=1}^{s-1}\bar{\mathbf{I}}^{s-1-p}\mathbb{E}\left[\mathbf{u}_p\mathbf{u}_q^H\right]\bar{\mathbf{I}}^{s-1-q}$. It follows from (49) that $\mathbb{E}\left[\mathbf{u}_p\mathbf{u}_q^H\right] = \mathbf{0}$, for $p \neq q$. Whereas, for $p = q$,

$$\mathbb{E}\left[\mathbf{u}_p\mathbf{u}_p^H\right] = \mu^2\mathbf{U}^H\mathbb{E}\left[\boldsymbol{\Psi}_{m,n}^H\tilde{\mathbf{v}}_{m,n}\tilde{\mathbf{v}}_{m,n}^H\boldsymbol{\Psi}_{m,n}\right]\mathbf{U}$$
$$= \mu^2\sigma^2\mathbf{U}^H\underbrace{\mathbb{E}\left[\boldsymbol{\Psi}_{m,n}^H\boldsymbol{\Psi}_{m,n}\right]}_{\mathbf{R}}\mathbf{U}$$
$$= \mu^2\sigma^2\boldsymbol{\Lambda}. \tag{75}$$

The simplifications in the above employ $\mathbb{E}\left[\tilde{\mathbf{v}}_{m,n}\tilde{\mathbf{v}}_{m,n}^H\right] = \sigma^2\mathbf{I}_{N_{RF}}$ and the eigenvalue decomposition of $\mathbf{R}$ as $\mathbf{R} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^H$. Substituting $\mathbb{E}\left[\mathbf{u}_p\mathbf{u}_q^H\right] = \mu^2\sigma^2\boldsymbol{\Lambda}\delta_{\mathrm{K}}(p-q)$, where $\delta_{\mathrm{K}}(\cdot)$ denotes the Kronecker-delta function, into the expression of $\boldsymbol{\Delta}_{m,n}$ yields

$$\boldsymbol{\Delta}_{m,n} = \mu^2\sigma^2\sum_{p=1}^{s-1}\bar{\mathbf{I}}^{s-1-p}\boldsymbol{\Lambda}\bar{\mathbf{I}}^{s-1-p}. \tag{76}$$

From the above, it can be readily observed that the matrix $\boldsymbol{\Delta}_{m,n}$ is a diagonal matrix, whose ith diagonal entry $\boldsymbol{\Delta}_{m,n}(i,i)$ is simplified as follows:

$$\boldsymbol{\Delta}_{m,n}(i,i) = \mu^2\sigma^2\lambda_i\sum_{p=1}^{s-1}(1 - \mu\lambda_i)^{2(s-1-p)}$$
$$= \mu^2\sigma^2\lambda_i(1 - \mu\lambda_i)^{2(s-1)}\sum_{p=1}^{s-1}(1 - \mu\lambda_i)^{-2p}$$
$$= \frac{\mu\sigma^2}{(2 - \mu\lambda_i)} - \frac{\mu\sigma^2(1 - \mu\lambda_i)^{2(s-1)}}{(2 - \mu\lambda_i)}.$$

This completes the proof of Lemma 2.

APPENDIX B

The MSOE defined in (17) can be modified for the FBLMS framework as

$$\mathcal{J}_{m,n}^{\mathrm{FBLMS}} = \mathbb{E}\left\{\left\|\mathbf{y}_{m,n} - \boldsymbol{\Psi}_{m,n}\mathbf{h}_{b,\mathrm{opt}} - \boldsymbol{\Psi}_{m,n}\boldsymbol{\epsilon}_{m,n}\right\|^2\right\}$$
$$= \mathbb{E}\left\{\left\|\tilde{\mathbf{v}}_{m,n} - \boldsymbol{\Psi}_{m,n}\boldsymbol{\epsilon}_{m,n}\right\|^2\right\}. \tag{77}$$

Also, it follows from (33) that $\mathcal{J}_{\min} = \mathbb{E}\left\{\|\tilde{\mathbf{v}}_{m,n}\|^2\right\}$. Therefore (77) can be further simplified using the independence assumption in (70) as

$$\mathcal{J}_{m,n}^{\mathrm{FBLMS}} = \mathcal{J}_{\min} + \mathbb{E}\left\{\boldsymbol{\epsilon}_{m,n}^H\boldsymbol{\Psi}_{m,n}^H\boldsymbol{\Psi}_{m,n}\boldsymbol{\epsilon}_{m,n}\right\}$$
$$= \mathcal{J}_{\min} + \mathbb{E}\left\{\mathrm{Tr}\left(\boldsymbol{\epsilon}_{m,n}^H\mathbf{R}\boldsymbol{\epsilon}_{m,n}\right)\right\}$$
$$= \mathcal{J}_{\min} + \mathbb{E}\left\{\mathrm{Tr}\left(\tilde{\boldsymbol{\epsilon}}_{m,n}^H\boldsymbol{\Lambda}\tilde{\boldsymbol{\epsilon}}_{m,n}\right)\right\}, \tag{78}$$

which simplifies to the desired result in (34). Finally, the asymptotic MSOE of the FBLMS framework is obtained as

$$\mathcal{J}_{\infty}^{\mathrm{FBLMS}} = \mathcal{J}_{\min} + \sum_{i=1}^{G_r G_t}\left[\lambda_i\left(\lim_{n\to\infty}\mathbb{E}\left[|\tilde{\boldsymbol{\epsilon}}_{m,n}(i)|^2\right]\right)\right]$$
$$= \mathcal{J}_{\min} + \sum_{i=1}^{G_r G_t}\left[\lambda_i\boldsymbol{\Delta}_{\infty}(i,i)\right], \tag{79}$$

which gives the desired result of (35).

APPENDIX C

Subtracting $\mathbf{h}_{b,\text{opt}}$ from both sides in Eq. (42) and rearranging the terms, the estimation error $\check{\boldsymbol{\epsilon}}_{m+1,n}$ for the proposed SFBLMS framework is given by

$$\check{\boldsymbol{\epsilon}}_{m+1,n} = \left(\mathbf{I} - \mu\boldsymbol{\Psi}_{m,n}^H\boldsymbol{\Psi}_{m,n}\right)\check{\boldsymbol{\epsilon}}_{m,n} + \mu\boldsymbol{\Psi}_{m,n}^H\tilde{\mathbf{v}}_{m,n}$$
$$-\delta\mathbf{g}\left(f\left(\widehat{\mathbf{h}}_{b,m,n}\right)\right). \quad (80)$$

Employing (49), together with the small step-size theory similar to our MSEE analysis of FBLMS in Appendix-A, the transformed estimation error $\bar{\boldsymbol{\epsilon}}_{m,n}$ can be written as

$$\bar{\boldsymbol{\epsilon}}_{m,n} = \bar{\mathbf{I}}^{s-1}\bar{\boldsymbol{\epsilon}}_{1,0} + \sum_{p=1}^{s-1}\bar{\mathbf{I}}^{s-1-p}\mathbf{u}_p + \sum_{p=1}^{s-1}\bar{\mathbf{I}}^{s-1-p}\mathbf{t}_p. \quad (81)$$

Furthermore, the quantity $\bar{\boldsymbol{\epsilon}}_{m,n}\bar{\boldsymbol{\epsilon}}_{m,n}^H$ for the SFBLMS can be formulated as

$$
\begin{aligned}
\bar{\boldsymbol{\epsilon}}_{m,n}\bar{\boldsymbol{\epsilon}}_{m,n}^H &= \bar{\mathbf{I}}^{s-1}\bar{\boldsymbol{\epsilon}}_{1,0}\bar{\boldsymbol{\epsilon}}_{1,0}^H\bar{\mathbf{I}}^{s-1} \\
&+ \sum_{p=1}^{s-1}\bar{\mathbf{I}}^{s-1}\bar{\boldsymbol{\epsilon}}_{1,0}\mathbf{u}_p^H\bar{\mathbf{I}}^{s-1-p} + \sum_{p=1}^{s-1}\bar{\mathbf{I}}^{s-1-p}\mathbf{u}_p\bar{\boldsymbol{\epsilon}}_{1,0}^H\bar{\mathbf{I}}^{s-1} \\
&+ \sum_{p=1}^{s-1}\bar{\mathbf{I}}^{s-1}\bar{\boldsymbol{\epsilon}}_{1,0}\mathbf{t}_p^H\bar{\mathbf{I}}^{s-1-p} + \sum_{p,q=1}^{s-1}\bar{\mathbf{I}}^{s-1-q}\mathbf{t}_q\mathbf{u}_p^H\bar{\mathbf{I}}^{s-1-p} \\
&+ \sum_{p=1}^{s-1}\bar{\mathbf{I}}^{s-1-p}\mathbf{t}_p\bar{\boldsymbol{\epsilon}}_{1,0}^H\bar{\mathbf{I}}^{s-1} + \sum_{p,q=1}^{s-1}\bar{\mathbf{I}}^{s-1-p}\mathbf{u}_p\mathbf{t}_q^H\bar{\mathbf{I}}^{s-1-q} \\
&+ \sum_{p,q=1}^{s-1}\bar{\mathbf{I}}^{s-1-p}\mathbf{u}_p\mathbf{u}_q^H\bar{\mathbf{I}}^{s-1-q} + \sum_{p,q=1}^{s-1}\bar{\mathbf{I}}^{s-1-p}\mathbf{t}_p\mathbf{t}_q^H\bar{\mathbf{I}}^{s-1-q}.
\end{aligned}
$$
$$(82)$$

Now, taking expectation $\mathbb{E}\{\cdot\}$ on both sides of the above equation, the expression for the matrix $\bar{\mathbf{K}}_{m,n}$ is obtained as given in Lemma-4, where the following simplifications have been employed. The 2nd and 3rd terms of (82) become zero due to the independence assumption. The 2nd-last term of (82) simplifies to a diagonal matrix $\boldsymbol{\Delta}_{m,n}$, whose diagonal elements can be derived similar to the simplification of the last term of Eq. (73). The simplification of the remaining terms in (82) can be readily carried out using the definitions of $\mathbf{K}_{t,t}^{p,q}, \mathbf{K}_{t,u}^{p,q}, \mathbf{K}_{u,t}^{p,q}, \mathbf{T}_p$ in Lemma-4. This completes the proof.

APPENDIX D

$$
\left(\mathbf{F}_{RF}^{\text{opt}}, \mathbf{F}_{BB}^{\text{opt}}\right) = \underset{(\mathbf{F}_{RF},\mathbf{F}_{BB})}{\arg\min} \sum_{k=0}^{K-1}\left\|\mathbf{F}^{\text{opt}}[k] - \mathbf{F}_{RF}\mathbf{F}_{BB}\right\|_F^2
$$

$$
\equiv \underset{(\mathbf{F}_{RF},\mathbf{F}_{BB})}{\arg\min} \sum_{k=0}^{K-1}\text{Tr}\left[-\mathbf{F}^{\text{opt}}[k]\mathbf{F}_{BB}^H\mathbf{F}_{RF}^H - \mathbf{F}_{RF}\mathbf{F}_{BB}\left(\mathbf{F}^{\text{opt}}[k]\right)^H\right.
$$
$$
\left. + \mathbf{F}_{RF}\mathbf{F}_{BB}\mathbf{F}_{BB}^H\mathbf{F}_{RF}^H\right] \quad (83)
$$

$$
\equiv \underset{(\mathbf{F}_{RF},\mathbf{F}_{BB})}{\arg\min} \text{Tr}\left[-\left(\sum_{k=0}^{K-1}\mathbf{F}^{\text{opt}}[k]\right)\mathbf{F}_{BB}^H\mathbf{F}_{RF}^H\right.
$$
$$
\left. - \mathbf{F}_{RF}\mathbf{F}_{BB}\left(\sum_{k=0}^{K-1}\mathbf{F}^{\text{opt}}[k]\right)^H + K\mathbf{F}_{RF}\mathbf{F}_{BB}\mathbf{F}_{BB}^H\mathbf{F}_{RF}^H\right]
$$
$$(84)$$

$$
\equiv \underset{(\mathbf{F}_{RF},\mathbf{F}_{BB})}{\arg\min} \text{Tr}\left[-\mathbf{F}^{\text{opt}}\mathbf{F}_{BB}^H\mathbf{F}_{RF}^H - \mathbf{F}_{RF}\mathbf{F}_{BB}\left(\mathbf{F}^{\text{opt}}\right)^H\right.
$$
$$
\left. + \mathbf{F}_{RF}\mathbf{F}_{BB}\mathbf{F}_{BB}^H\mathbf{F}_{RF}^H\right], \quad (85)
$$

where $\mathbf{F}^{\text{opt}} = \frac{1}{K}\sum_{k=0}^{K-1}\mathbf{F}^{\text{opt}}[k]$. Adding the term $\mathbf{F}^{\text{opt}}\left(\mathbf{F}^{\text{opt}}\right)^H$ that is independent of the optimization variable $\mathbf{F}_{RF}$ and $\mathbf{F}_{BB}$, we can reformulate the above design problem as

$$
\left(\mathbf{F}_{RF}^{\text{opt}}, \mathbf{F}_{BB}^{\text{opt}}\right) = \underset{(\mathbf{F}_{RF},\mathbf{F}_{BB})}{\arg\min} \text{Tr}\left[\mathbf{F}^{\text{opt}}\left(\mathbf{F}^{\text{opt}}\right)^H - \mathbf{F}^{\text{opt}}\mathbf{F}_{BB}^H\mathbf{F}_{RF}^H\right.
$$
$$
\left. - \mathbf{F}_{RF}\mathbf{F}_{BB}\left(\mathbf{F}^{\text{opt}}\right)^H + \mathbf{F}_{RF}\mathbf{F}_{BB}\mathbf{F}_{BB}^H\mathbf{F}_{RF}^H\right]
$$
$$(86)$$

$$
= \underset{(\mathbf{F}_{RF},\mathbf{F}_{BB})}{\arg\min} \left\|\mathbf{F}^{\text{opt}} - \mathbf{F}_{RF}\mathbf{F}_{BB}\right\|_F^2. \quad (87)
$$

REFERENCES

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

[2] R. W. Heath, N. Gonzalez-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 3, pp. 436–453, 2016.

[3] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE access*, vol. 1, pp. 335–349, 2013.

[4] "Wireless HD specification version 1.1 overview," *Wireless HD, Morgan Hill, CA, USA*, 2010.

[5] K. Venugopal, M. C. Valenti, and R. W. Heath, "Analysis of millimeter wave networked wearables in crowded environments," in *2015 49th Asilomar Conference on Signals, Systems and Computers*. IEEE, 2015, pp. 872–876.

[6] M. S. Elbamby, C. Perfecto, M. Bennis, and K. Doppler, "Edge computing meets millimeter-wave enabled VR: Paving the way to cutting the cord," in *2018 IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE, 2018, pp. 1–6.

[7] P. Kumari, N. Gonzalez-Prelcic, and R. W. Heath, "Investigating the IEEE 802.11 ad standard for millimeter wave automotive radar," in *2015 IEEE 82nd Vehicular Technology Conference (VTC2015-Fall)*. IEEE, 2015, pp. 1–5.

[8] S. Geng, J. Kivinen, X. Zhao, and P. Vainikainen, "Millimeter-wave propagation channel characterization for short-range wireless communications," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 1, pp. 3–13, 2008.

[9] T. Bai, A. Alkhateeb, and R. W. Heath, "Coverage and capacity of millimeter-wave cellular networks," *IEEE Communications Magazine*, vol. 52, no. 9, pp. 70–77, 2014.

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

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

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

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

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

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

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

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

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

[19] Y. M. Tsang, A. S. Poon, and S. Addepalli, "Coding the beams: Improving beamforming training in mmwave communication system," in *Global Telecommunications Conference (GLOBECOM 2011), 2011 IEEE*. IEEE, 2011, pp. 1–6.

[20] Z. Guo, X. Wang, and W. Heng, "Millimeter-wave channel estimation based on 2-D beamspace MUSIC method," *IEEE Transactions on Wireless Communications*, vol. 16, no. 8, pp. 5384–5394, 2017.

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

[22] X. Cheng, C. Tang, and Z. Zhang, "Accurate channel estimation for millimeter-wave MIMO systems," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 5, pp. 5159–5163, 2019.

[23] Q. Qin, L. Gui, P. Cheng, and B. Gong, "Time-varying channel estimation for millimeter wave multiuser MIMO systems," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 10, pp. 9435–9448, 2018.

[24] Z. Wan, Z. Gao, B. Shim, K. Yang, G. Mao, and M.-S. Alouini, "Compressive sensing based channel estimation for millimeter-wave full-dimensional MIMO with lens-array," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 2, pp. 2337–2342, 2019.

[25] X. Cheng, Y. Yang, B. Xia, N. Wei, and S. Li, "Sparse channel estimation for millimeter wave massive MIMO systems with lens antenna array," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 11, pp. 11 348–11 352, 2019.

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

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

[28] W. Shen, X. Bu, X. Gao, C. Xing, and L. Hanzo, "Beamspace precoding and beam selection for wideband millimeter-wave MIMO relying on lens antenna arrays," *IEEE Transactions on Signal Processing*, vol. 67, no. 24, pp. 6301–6313, 2019.

[29] Z. Gao, C. Hu, L. Dai, and Z. Wang, "Channel estimation for millimeter-wave massive MIMO with hybrid precoding over frequency-selective fading channels," *IEEE Communications Letters*, vol. 20, no. 6, pp. 1259–1262, 2016.

[30] J. Mo, P. Schniter, and R. W. Heath, "Channel estimation in broadband millimeter wave MIMO systems with few-bit ADCs," *IEEE Transactions on Signal Processing*, vol. 66, no. 5, pp. 1141–1154, 2017.

[31] S. Gao, X. Cheng, and L. Yang, "Estimating doubly-selective channels for hybrid mmwave massive MIMO systems: A doubly-sparse approach," *IEEE Transactions on Wireless Communications*, 2020.

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

[33] W. Huang, Y. Huang, R. Zhao, S. He, and L. Yang, "Wideband millimeter wave communication: Single carrier based hybrid precoding with sparse optimization," *IEEE Transactions on Vehicular Technology*, 2018.

[34] S. Buzzi, C. D Andrea, T. Foggi, A. Ugolini, and G. Colavolpe, "Single-carrier modulation versus OFDM for millimeter-wave wireless MIMO," *IEEE Transactions on Communications*, vol. 66, no. 3, pp. 1335–1348, 2018.

[35] S. S. Haykin, *Adaptive filter theory*. Pearson Education India, 2005.

[36] B. Farhang-Boroujeny, *Adaptive filters: theory and applications*. John Wiley & Sons, 2013.

[37] A. H. Sayed, *Adaptive filters*. John Wiley & Sons, 2011.

[38] A. V. Oppenheim and R. W. Schafer, *Digital signal processing*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1975.

[39] S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter wave cellular wireless networks: Potentials and challenges," *arXiv preprint arXiv:1401.2560*, 2014.

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

[41] J. Jin, Y. Gu, and S. Mei, "A stochastic gradient approach on compressive sensing signal reconstruction based on adaptive filtering framework," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 409–420, 2010.

[42] S. Srivastava, P. Sharma, S. Dwivedi, A. K. Jagannatham, and L. Hanzo, "Technical Report: Fast Block LMS Based Estimation of Angularly Sparse Channels for Hybrid Wideband Millimeter Wave MIMO Systems," IIT Kanpur, Tech. Rep., 2020, [Online]. Available: http://www.iitk.ac.in/mwn/documents/MWNLab_TR_SFBLMS_mmWave_2020.pdf.

[43] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge university press, 2005.

[44] A. Alkhateeb and R. W. Heath, "Frequency selective hybrid precoding for limited feedback millimeter wave systems," *IEEE Transactions on Communications*, vol. 64, no. 5, pp. 1801–1818, 2016.

[45] J. Li, S. Gao, X. Cheng, and L. Yang, "Hybrid precoded spatial modulation (hPSM) for mmwave massive MIMO systems over frequency-selective channels," *IEEE Wireless Communications Letters*, 2020.