Subspace-Based Estimation of Rapidly Varying Mobile Channels for OFDM Systems

Habib Şenol, *Senior Member, IEEE*, and Cihan Tepedelenlioğlu, *Senior Member, IEEE*

***Abstract*—It is well-known that time-varying channels can provide time diversity and improve error rate performance compared to time-invariant fading channels. However, exploiting time diversity requires very accurate channel estimates at the receiver. In order to reduce the number of unknown channel coefficients while estimating the time-varying channel, basis expansion models can be used along with long transmission frames that contain multiple orthogonal frequency division multiplexing (OFDM) symbols that experience the channel variation. The design of these OFDM frames need to judiciously incorporate training and data insertions in the transmitted signal while maintaining orthogonality. In this work, we propose an inter channel interference (ICI)-free training model depending on pilot symbols only and provide a corresponding time-varying channel estimation method. This scheme relies on an algorithm to determine the number of OFDM symbols per frame and the number of basis functions per path with minimal information about the Doppler bandwidth. As a performance benchmark, Bayesian Cramér Rao lower bound (CRLB) and the corresponding MSE bound are derived analytically for the proposed training model. Theoretical MSE expressions of the proposed estimation scheme are also derived as well as the MSE expressions in the presence of Doppler frequency mismatch. Simulations exhibit substantial MSE improvement and the corresponding Symbol Error Rate (SER) performances of the low complexity estimation scheme. They also corroborate that, unlike the common results in the literature, an OFDM system can perform better as the Doppler frequency increases with judicious design of training and channel estimation schemes.**

***Index Terms*—Basis expansion model (BEM), channel estimation, null space, orthogonal frequency-division multiplexing (OFDM), rapidly varying multipath channel.**

I. Introduction

ORTHOGONAL frequency division multiplexing (OFDM), due to its robustness against frequency selectivity in wireless channels and ability to mitigate multipath fading causing inter symbol interference (ISI), has been widely adopted as a standard modulation technique in modern wireless communication technologies operating over frequency selective fading channels. Consequently, OFDM is a backbone physical layer technology that is employed in the wireless cellular standards such as mobile Worldwide Interoperability for Microwave Access (WiMAX) systems, Third-Generation Partnership (3GPP) Long Term Evolution (LTE) project, and mobile Wireless Local Area Network (WLAN) standards. Additionally, in order to enable mobile broadband communications at high speeds beyond 120 km/h, the LTE standards are promoted by introducing a high mobility feature of LTE Advanced (LTE-A).

The Doppler frequency emerging because of relative motion between transceivers makes the multipath channel time-variant, destroys the subcarrier orthogonality, and leads to inter channel interference (ICI) causing performance degradation of the OFDM system due to power leakage among subchannels. On the other hand, since developing modern wireless communication systems should support high mobility, estimation of the channel and relevant channel parameters of an OFDM system operating effectively under such conditions is a crucial problem. In the literature, while early studies on OFDM channel estimation considered quasi-static channels [1]–[10], the subsequent studies focused on OFDM receiver design for the transmission over block-wise [11], [12] and very rapidly time varying multipath fading channels [13]–[21] that are mentioned briefly in the next paragraph.

In order to mitigate the ICI, authors in [13] proposed a time-domain estimation method for a rapidly time-varying channel. In the proposed method in [13], reducing the computational load of the linear minimum mean squared error (MMSE) channel estimator by using the singular value decomposition (SVD) based low rank approximation, the fading channel was estimated. However, in the proposed method the linear MMSE equalization with a successive detection performed along with the channel estimation in which detection ordering was determined with respect to signal to interference-and-noise power ratio (SINR) criteria required a large number of subcarriers and requires high computational load. In [14], exploiting the channel interpolation to handle rapid variation within an OFDM transmission block, authors proposed a pilot-assisted channel estimation method which has to be performed in two estimation steps. However the computational load of the proposed method is very high since it requires large size matrix inversion. In order to reduce the computational load of the proposed method, they also proposed a simplified scheme incurring a substantial performance loss.

Manuscript received April 29, 2020; revised September 11, 2020 and December 4, 2020; accepted December 13, 2020. Date of publication December 17, 2020; date of current version January 11, 2021. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Sangarapillai Lambotharan. The work of Habib Şenol was supported by The Scientific and Technological Research Council of Turkey (TÜBITAK) between September 2018 and June 2019. *(Corresponding author: Habib Şenol.)*

Habib Şenol is with the Department of Computer Engineering, Kadir Has University, 34 083 Istanbul, Turkey (e-mail: hsenol@khas.edu.tr).

Cihan Tepedelenlioğlu is with the Department of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, AZ 85 287 USA (e-mail: cihan@asu.edu).

Digital Object Identifier 10.1109/TSP.2020.3045562

In [15], assuming that each path of the time-varying multipath channel varies linearly within one OFDM symbol duration, authors proposed a time-domain channel estimator. While this assumption holds for reasonable small values of the Doppler frequency, it no longer holds as the Doppler frequency increases. In [16], approximating channel time-variations by a piece-wise linear model, authors proposed two methods to mitigate ICI in an OFDM system with a coherent channel estimation technique. In the first method [16], while information of channel time-variations was extracted from the cyclic prefix, in the second method, these variations were estimated using the next symbol. However, as mentioned earlier for [15], piece-wise linear model assumption in [16] no longer holds for higher values of Doppler frequency causing rapidly time-varying multipath channels. In [17], assuming a banded channel transfer matrix, an equalizer with low computational cost was designed first, then a pilot aided MMSE channel estimation algorithm for the time-varying wide-sense stationary uncorrelated scattering (WSSUS) channel model was proposed. However, the proposed method suffers from an irreducible error floor when ICI dominates due to the time-varying channels and requires higher pilot overhead to decrease this error floor. Reference [18] considers a nonlinear optimization problem of pilot sequence design to minimize the MSE of the channel estimator and proposes a path-following procedure with sequential convex programming with the sub-optimal solution. However, while the proposed pilot design algorithm is employed during channel estimation where the ICI effect is not considered, the referred successive interference cancellation method has been used to eliminate the ICI effect during data recovery. The work in [19] focused on the estimation problem of time-varying multipath channels for a time-domain synchronous OFDM (TDS-OFDM). The proposed channel estimation scheme in [19] divides one OFDM symbol into partitions and inserts known pseudo-noise (PN) random sequences for training in between these partitions and the guard interval. The PN sequence inserted to the guard interval is exploited also for synchronization. In addition to increasing the length of one OFDM symbol duration due to these added time-domain PN sequences, also the proposed method suffers from the mutual interference introduced between the PN sequence in the guard interval and the OFDM symbol which destroys the cyclic property of the received OFDM symbol. Therefore, before channel estimation in [19], the mutual interference between the PN sequence and the OFDM symbol must be eliminated in iterative fashion causing additional computational load. In [20] and [21], expectation maximization (EM) variant techniques were used to estimate the time-varying tap coefficients of multipath OFDM channels. However, since EM variant OFDM channel estimation algorithms are iterative, they have much higher computational load as compared with batch channel estimation algorithms and are not suitable candidates in search of efficient time-varying channel estimator providing minimum latency between OFDM transceivers.

In the literature, the Gauss-Markov model [22] and the basis expansion model (BEM) [23]–[26] are common models for time-varying channels. In the BEM, each path of the time-varying multipath channel is expressed efficiently as a linear combination of a few basis functions and corresponding basis coefficients while path variations are tracked through symbol-by-symbol updates in Gauss-Markov models. Therefore, in BEM based channel estimation, the estimation problem of the time-varying multipath channel turns to estimation of a few unknown BEM coefficients rather than much higher number of unknown channel coefficients. The proposed BEMs in [19], [20], [23]–[29] are discrete Karhunen-Loeve (DKL), discrete-prolate spheroidal, complex exponential, discrete cosine and discrete Legendre polynomial (DLP) BEMs, respectively. Among these BEMs, the channel modeling error of DKL BEM is the minimum as compared with that of other BEMs and its expansion coefficients are statistically uncorrelated [30]. On the other hand, the DKL BEM based algorithm requires the prior knowledge of channel statistics that depend on Doppler frequency as well, and therefore, is not robust against maximum Doppler frequency mismatch. As proved in [20], the modeling error performance of the DLP BEM approaches very rapidly to the performance of the DKL BEM as the resolution increases. Moreover, as shown by performance plots in [20], DLP BEM models the time-varying channel very accurately without the knowledge of channel statistics.

In this paper, we propose efficient subspace-based batch estimation of rapidly varying mobile multipath channels for OFDM systems by avoiding computationally complex iterative approaches of [20], while retaining the orthogonality, unlike [19]. Doppler shift that changes the position of the subcarriers randomly, causes the loss of the orthogonality of the subcarriers which leads to ICI due to power leakage among neighboring OFDM subcarriers. Therefore, data ICI effect due to unknown neighboring data symbols of pilots symbols degrades the performance of the channel estimator. So, defining a suitable frequency-to-time mapping, a higher performance channel estimator may be possible with a so-called ICI-free time-domain training signal model which does not include data symbols. In the proposed scheme without destroying the orthogonality of the subcarriers, in order to obtain an ICI-free time-domain training signal model depending on pilot symbols only, a frequency-to-time mapping vector is added to frequency domain OFDM symbol such that the sum of the OFDM data (data vector is expressed by substituting zero into pilot symbols) and mapping vectors remains in the null space of the inverse Fast Fourier Transform (FFT) fat submatrix obtained by removing row vectors with data indices. Similarly, the sum of the OFDM pilot (pilot vector is expressed by substituting zero into data symbols) and mapping vectors remains in the null space of the inverse FFT fat submatrix obtained by removing row vectors with training indices. This way, choosing time-domain consecutive training indices, one OFDM symbol duration is divided in to a single training interval and two data intervals. In the OFDM system setup, discrete time training interval is selected so as to be in the middle of a cyclic prefixed OFDM (CP-OFDM) duration. We employed the DLP BEMs which capture the time-domain rapid variations of the channel paths by a few parameters to be estimated. We highlight the main contributions and the gained advantages by using our proposed estimation scheme as follows:

- For a fixed modelling MSE, using so-called the Padé method that is a rational fraction approximation, we analytically expressed the number of BEM coefficients per

path as a function of maximum Doppler frequency and the length of OFDM frame.
- We provide an algorithm that can determine the number of OFDM symbols per frame and the number of BEM coefficients per path for given number of subcarriers, number of paths, pilot spacing and number of observation per one BEM coefficient.
- We propose a new training signal model where the received signal corresponding to the pilot symbols are free of interference even in the presence of Doppler.
- In order to present performance benchmark for the estimator of the complex valued DLP BEM channel coefficients, the analytical expressions of the Bayesian Cramér Rao lower bound (CRLB) and theoretical MSEs considering mismatch in the maximum Doppler frequency scenario as well.
- We show by performance plots that our training scheme proposed for OFDM systems provides the time diversity property resulting in better channel estimation performance for higher Doppler frequency (due to ICI-free training model and setting-up the system so as to exploit higher Doppler frequencies)

The rest of the paper is organized as follows: Section II describes the signal and the time-varying channel models for an OFDM frame with cyclic prefixed OFDM (CP-OFDM) symbols. Section III provides an ICI-free new training signal model depending on pilot symbols only, and employing DLP BEM, presents estimator for rapidly time-varying channels as well as the calculation of the computational complexity of the estimator. Section IV considers the calculation of the Doppler related parameters such as the number of OFDM symbols per frame and the number of basis functions per single channel path. Section V evaluates the performances of the system via computer simulations. Finally, Section VI summarizes the main conclusions of the paper.

II. System Model

A. Signal and Channel Models

We consider an OFDM frame with M OFDM symbols and each OFDM symbol in the frame has N subcarriers. Employing the Inverse Fast Fourier Transform (IFFT), the time-domain signal of mth OFDM symbol in the frame is obtained at the output of the IFFT as follows

$$x[m,n]=\frac{1}{N}\sum_{k=-N/2}^{N/2-1}X[m,k]\,e^{j2\pi n\,k/N},\quad n=0,1,\ldots,N-1,\ m=0,1,\ldots,M-1, \tag{1}$$

where n is the discrete-time index within one OFDM symbol duration, k denotes for the discrete-frequency index and $X[m,k]$ stands for the frequency domain digitally modulated symbol transmitted over the kth subcarrier of the mth OFDM symbol. We assume an L-path time-varying multipath mobile radio channel with discrete-time impulse response $g[n,l]=\sum_{\ell=0}^{L-1}h_\ell[n]\,\delta[l-\ell]$ between transceivers where zero-mean complex Gaussian $h_\ell[n]$ represents time-varying multipath coefficient of the ℓth path and $\delta[\cdot]$ stands for the Kronecker delta function. Assuming Jakes' model [31], the discrete autocorrelation function of $h_\ell[n]$ is given by

$$\rho_h[n-n,'\,\ell-\ell']=E\left\{h_\ell[n]\,h^*_{\ell'}[n']\right\}=\Omega_\ell J_0\left(2\pi f_D T_s(n-n')\right)\,\delta[\ell-\ell'], \tag{2}$$

where $(\cdot)^*$ denotes the complex conjugate operator, $\Omega_\ell=E\{|h_\ell[t]|^2\}$ is the normalized power of the ℓth path such that $\sum_{\ell=0}^{L-1}\Omega_\ell=1$, $J_0(\cdot)$ is the zeroth-order Bessel function of the first kind, f_D is the maximum (one-sided) Doppler shift and T_s represents the sampling period. L-path multipath channel causes an interference between two consecutive OFDM symbols at receiver side. This is called as inter symbol interference (ISI). In order to combat with ISI, a cyclic prefix (CP) with length $L_c\geq L-1$ is added to each time-domain OFDM signal. At the receiver side, after matched filtering, symbol-rate sampling, and discarding the time domain observation samples falling in the CP, received signal at the input of the Fast Fourier transform (FFT) is obtained as

$$y[mN_g+L_c+n]=\sum_{\ell=0}^{L-1}h_\ell[mN_g+L_c+n]\,x[m,n-\ell]+w[mN_g+L_c+n],\ n=0,1,\ldots,N-1,\ m=0,1,\ldots,M-1, \tag{3}$$

where $N_g=N+L_c$ is the length of a CP-OFDM signal and $w[\cdot]$ stands for zero-mean complex additive Gaussian noise with variance σ_w^2. Using (3), the observation vector for the mth OFDM symbol can be given by following two forms

$$\mathbf{y}[m]=\sum_{\ell=0}^{L-1}\mathrm{diag}(\mathbf{x}_\ell[m])\,\mathbf{h}_\ell[m]+\mathbf{w}[m] \tag{4}$$

$$\text{or}\ \ \mathbf{y}[m]=\mathbf{H}[m]\mathbf{x}[m]+\mathbf{w}[m],\ m=0,1,\ldots,M-1, \tag{5}$$

where

$$\begin{aligned}\mathbf{y}[m]&=[y[mN_g+L_c+n]\,;\,n=0,1,\ldots,N-1]\in\mathbb{C}^{N\times1}\\ \mathbf{h}_\ell[m]&=[h_\ell[mN_g+L_c+n]\,;\,n=0,1,\ldots,N-1]\in\mathbb{C}^{N\times1}\\ \mathbf{w}[m]&=[w[mN_g+L_c+n]\,;\,n=0,1,\ldots,N-1]\in\mathbb{C}^{N\times1}\\ \mathbf{x}[m]&=[x[m,n]\,;\,n=0,1,\ldots,N-1]\in\mathbb{C}^{N\times1}\end{aligned} \tag{6}$$

and recalling $x[m,-\ell]=x[m,N-\ell]$, the ℓ-step circular shifted version of $\mathbf{x}_\ell[m]$ is given by

$$\mathbf{x}_\ell[m]=[x[m,n-\ell]\,;\,n=0,1,\ldots,N-1]\in\mathbb{C}^{N\times1} \tag{7}$$

In (5), the channel convolution matrix $\mathbf{H}[m]$ is defined as

$$\mathbf{H}[m]=\sum_{\ell=0}^{L-1}\mathrm{colshift}\left(\mathrm{diag}(\mathbf{h}_\ell[m]),-\ell\right)\ \in\mathbb{C}^{N\times N}, \tag{8}$$

where $\mathrm{diag}(\mathbf{h}_\ell)$ creates a diagonal matrix whose main diagonal is $\mathbf{h}_\ell$ and $\mathrm{colshift}(\mathbf{A},q)$ represents a column-wise q-step circular shift of matrix $\mathbf{A}$.

III. Training Signal Model and Channel Estimation

Doppler frequency leading to ICI causes performance degradation of the OFDM receiver due to power leakage among subchannels. In other words, an observation received over the kth subcarrier corresponding to a pilot symbol transmitted over the kth subcarrier may depend also on unknown data symbols in neighboring subcarriers. In addition, as the Doppler frequency increases, the interference begins to depend on neighboring symbols further away. So, in order to present a channel estimator that will not be affected by the negative impact of ICI, it is necessary to obtain a training signal model that can be decoupled from the data symbols in the presence of Doppler.

A. Training Signal Model

In this section, preserving the frequency domain signal structure of OFDM, the ICI-free time domain training signal model is obtained over an OFDM symbol duration. This is achieved with a frequency-to time mapping to ensure data and training samples are not overlapping in time, and also training samples remain within a single discrete-time training interval, which is different than conventional OFDM. We express the frequency domain transmit signal $X[m,k]$ in (1) as follows

$$\begin{aligned} X[m,k] &= S[m,k]+V[m,k], k=-\frac{N}{2},-\frac{N}{2}+1,\ldots,\frac{N}{2}-1 \\ &= \begin{cases} S^{(tr)}[m,k]+V^{(tr)}[m,k]\ , & k\in\mathcal{I}^{(tr)} \\ S^{(dt)}[m,k]+V^{(dt)}[m,k]\ , & k\in\mathcal{I}^{(dt)} \end{cases} \end{aligned} \tag{9}$$

where $S^{(tr)}[m,k]$ and $S^{(dt)}[m,k]$ are digitally modulated training and data symbols, respectively, and $V[m,k]$ is the frequency-to-time mapping signal. The index set $\mathcal{I}^{(tr)}$ with cardinality $N^{(tr)}$ represents the index set of subcariers on transmitted training symbols whereas the index set $\mathcal{I}^{(dt)}$ with cardinality $N^{(dt)}=N-N^{(tr)}$ stands for the index set of subcariers on transmitted data symbols. Recalling that $\mathcal{I}^{(tr)}$ and $\mathcal{I}^{(dt)}$ in (9) represent the frequency-domain pilot and data index sets, respectively, in the frequency domain comb-type pilot structure, we want to determine $V^{(tr)}[m,k]$ and $V^{(dt)}[m,k]$ such that the time-domain training signal for channel estimation stays in the interval of $n_0\le n<(n_0+N^{(tr)})$ and time-domain data signal remains within the rest of the OFDM symbol duration $0\le n<N$, where n_0 is the location of the first training sample and it will be calculated such that the interval of the path samples involving in the training observations is in the middle of one CP-OFDM symbol duration. This can be written for mth time-domain OFDM signal $x[m,n]$ in (1) as follows

$$x[m,n]=\begin{cases} s^{(dt)}[m,n], & 0\le n<n_0 \\ s^{(tr)}[m,n-n_0], & n_0\le n<(n_0+N^{(tr)}) \\ s^{(dt)}[m,n-N^{(tr)}], & (n_0+N^{(tr)})\le n<N \end{cases} \tag{10}$$

where samples of $s^{(tr)}[m,n]$ and $s^{(dt)}[m,n]$ are not overlapping in time and we want to obtain $s^{(tr)}[m,n]$ and $s^{(dt)}[m,n]$ signals by only the frequency domain training and data symbols, respectively, such that

$$\begin{aligned} s^{(tr)}[m,n-n_0] &= \frac{1}{N}\sum_{k\in\mathcal{I}^{(tr)}} S^{(tr)}[m,k]\, e^{j2\pi nk/N}, \\ &\quad n_0\le n<(n_0+N^{(tr)}), \\ s^{(dt)}[m,n-aN^{(tr)}] &= \frac{1}{N}\sum_{k\in\mathcal{I}^{(dt)}} S^{(dt)}[m,k]\, e^{j2\pi nk/N}, \\ &\quad a(n_0+N^{(tr)})\le n<(1-a)n_0+aN, \\ &\quad a\in\{0,1\}. \end{aligned} \tag{11}$$

We will calculate the estimate of the multipath channels within the data intervals by basis expansion model (BEM) based interpolation method. As seen from the convolution operation in (3), training signal defined in the interval of $n_0\le n<(n_0+N^{(tr)})$ is convolved by the path coefficients within the interval of $n_0+(L-1)\le n<(n_0+N^{(tr)})$. In order to minimize the interpolation error symmetrically on both sides of this interval of $n_0+(L-1)\le n<(n_0+N^{(tr)})$ within one CP-OFDM symbol duration N_g, we calculate the location of the first training sample n_0 in (10) such that this interval is in the middle of one CP-OFDM symbol duration as follows

$$\begin{aligned} n_0 &= \left\lfloor \frac{N_g-(N^{(tr)}-L+1)}{2}-(L-1)\right\rceil \\ &= \left\lfloor \frac{N-N^{(tr)}+L_c-(L-1)}{2}\right\rceil, \end{aligned} \tag{12}$$

where $\lfloor\cdot\rceil$ rounds to the nearest integer. Recalling (9), (10) and $\mathbf{F}^{-1}=\frac{1}{N}\mathbf{F}^{\dagger}$ where $(\cdot)^{\dagger}$ denotes the Hermitian transpose, the Fourier Transform relationship between training and data parts of $X[m,k]$ and $x[m,n]$ in (1) can be given for mth OFDM symbol as follows

$$\begin{bmatrix} \mathbf{S}^{(tr)}[m]+\mathbf{V}^{(tr)}[m] \\ \mathbf{S}^{(dt)}[m]+\mathbf{V}^{(dt)}[m] \end{bmatrix} = \begin{bmatrix} \mathbf{F}_0 & \mathbf{F}_1 \\ \mathbf{F}_2 & \mathbf{F}_3 \end{bmatrix} \begin{bmatrix} \mathbf{s}^{(tr)}[m] \\ \mathbf{s}^{(dt)}[m] \end{bmatrix} \tag{13}$$

and

$$\begin{bmatrix} \mathbf{s}^{(tr)}[m] \\ \mathbf{s}^{(dt)}[m] \end{bmatrix} = \frac{1}{N}\begin{bmatrix} \mathbf{F}_0^{\dagger} & \mathbf{F}_2^{\dagger} \\ \mathbf{F}_1^{\dagger} & \mathbf{F}_3^{\dagger} \end{bmatrix} \begin{bmatrix} \mathbf{S}^{(tr)}[m]+\mathbf{V}^{(tr)}[m] \\ \mathbf{S}^{(dt)}[m]+\mathbf{V}^{(dt)}[m] \end{bmatrix}, \tag{14}$$

where

$$\begin{aligned} \mathbf{S}^{(tr)}[m] &= \left[S[m,k]\,;\, k\in\mathcal{I}^{(tr)}\right]\in\mathbb{C}^{N^{(tr)}\times 1} \\ \mathbf{S}^{(dt)}[m] &= \left[S[m,k]\,;\, k\in\mathcal{I}^{(dt)}\right]\in\mathbb{C}^{N^{(dt)}\times 1} \\ \mathbf{V}^{(tr)}[m] &= \left[V[m,k]\,;\, k\in\mathcal{I}^{(tr)}\right]\in\mathbb{C}^{N^{(tr)}\times 1} \\ \mathbf{V}^{(dt)}[m] &= \left[V[m,k]\,;\, k\in\mathcal{I}^{(dt)}\right]\in\mathbb{C}^{N^{(dt)}\times 1} \\ \mathbf{s}^{(tr)}[m] &= [x[m,n]\,;\, n=n_0,n_0+1,\ldots, \\ &\qquad n_0+N^{(tr)}-1\Big]\in\mathbb{C}^{N^{(tr)}\times 1} \\ \mathbf{s}^{(dt)}[m] &= \Big[x[m,n]\,;\, n=0,1,\ldots,n_0-1,n_0+N^{(tr)},\ldots, \\ &\qquad N-1]\in\mathbb{C}^{N^{(dt)}\times 1}. \end{aligned} \tag{15}$$

In (13) and (14), the submatrices of the FFT matrix are defines as follows

$\mathbf{F}_0$: an $N^{(tr)} \times N^{(tr)}$ submatrix of the FFT matrix with column training and row training index sets.

$\mathbf{F}_1$: an $N^{(tr)} \times N^{(dt)}$ submatrix of the FFT matrix with column data and row training index sets.

$\mathbf{F}_2$: an $N^{(dt)} \times N^{(tr)}$ submatrix of the FFT matrix with column training and row data index sets.

$\mathbf{F}_3$: an $N^{(dt)} \times N^{(dt)}$ submatrix of the FFT matrix with column data and row data index sets.

Writing the training and data equation set of (14) separately and composing these equations so as to satisfy the equations in (11), we obtain

$$\mathbf{s}^{(tr)}[m] = \underbrace{\frac{1}{N}\left[\mathbf{F}_0^{\dagger}\mathbf{F}_2^{\dagger}\right]\begin{bmatrix}\mathbf{V}^{(tr)}[m]\\ \mathbf{S}^{(dt)}[m]+\mathbf{V}^{(dt)}[m]\end{bmatrix}}_{0} + \frac{1}{N}\mathbf{F}_0^{\dagger}\,\mathbf{S}^{(tr)}[m], \tag{16}$$

and

$$\mathbf{s}^{(dt)}[m] = \underbrace{\frac{1}{N}\left[\mathbf{F}_1^{\dagger} \quad \mathbf{F}_3^{\dagger}\right]\begin{bmatrix}\mathbf{S}^{(tr)}[m]+\mathbf{V}^{(tr)}[m]\\ \mathbf{V}^{(dt)}[m]\end{bmatrix}}_{0} + \frac{1}{N}\mathbf{F}_3^{\dagger}\,\mathbf{S}^{(dt)}[m], \tag{17}$$

where note that the vectors $[\mathbf{V}^{(tr)}[m]\,;\,\mathbf{S}^{(dt)}[m]+\mathbf{V}^{(dt)}[m]]$ and $[\mathbf{S}^{(tr)}[m]+\mathbf{V}^{(tr)}[m]\,;\,\mathbf{V}^{(dt)}[m]]$ inside zero terms on the right hand side of (16) and (17) are chosen to be null vectors of the fat matrices $[\mathbf{F}_0^{\dagger}\ \mathbf{F}_2^{\dagger}]$ and $[\mathbf{F}_1^{\dagger}\ \mathbf{F}_3^{\dagger}]$, respectively. Thus, after this choice, the vectors $\mathbf{s}^{(tr)}[m]$ and $\mathbf{s}^{(dt)}[m]$ remain in the range spaces of the matrices $\mathbf{F}_0^{\dagger}$ and $\mathbf{F}_3^{\dagger}$, respectively. From the joint solution of these null vectors, we have

$$\frac{1}{N}\begin{bmatrix}\mathbf{F}_0^{\dagger} & \mathbf{F}_2^{\dagger}\\ \mathbf{F}_1^{\dagger} & \mathbf{F}_3^{\dagger}\end{bmatrix}\begin{bmatrix}\mathbf{V}^{(tr)}[m]\\ \mathbf{V}^{(dt)}[m]\end{bmatrix} = -\frac{1}{N}\begin{bmatrix}0 & \mathbf{F}_2^{\dagger}\\ \mathbf{F}_1^{\dagger} & 0\end{bmatrix}\begin{bmatrix}\mathbf{S}^{(tr)}[m]\\ \mathbf{S}^{(dt)}[m]\end{bmatrix} \tag{18}$$

and thus, respecting the frequency domain training and data locations, (18) can be rewritten as follows

$$\frac{1}{N}\mathbf{F}^{\dagger}\mathbf{V}[m] = -\frac{1}{N}\left(\mathbf{F} - \mathbf{F}_{\mathbf{z}}\right)^{\dagger}\mathbf{S}[m], \tag{19}$$

where

$$\mathbf{S}[m] = \left[S[m,k]; k = -\frac{N}{2}, -\frac{N}{2}+1, \ldots, \frac{N}{2}-1\right] \in \mathbb{C}^{N\times 1}$$
$$\mathbf{V}[m] = \left[V[m,k]; k = -\frac{N}{2}, -\frac{N}{2}+1, \ldots, \frac{N}{2}-1\right] \in \mathbb{C}^{N\times 1}, \tag{20}$$

and the FFT submatrix $\mathbf{F}_{\mathbf{z}} \in \mathbb{C}^{N\times N}$ is obtained by setting to zero all the entries of FFT matrix with column-training/row-data and column-data/row-training indices. The solution of the mapping vector $\mathbf{V}[m]$ in (19) is obtained with respect to digitally modulated OFDM symbol vector $\mathbf{S}[m]$ as follows

$$\mathbf{V}[m] = -\left(\mathbf{I}_N - \frac{1}{N}\mathbf{F}\mathbf{F}_{\mathbf{z}}^{\dagger}\right)\mathbf{S}[m]. \tag{21}$$

Eventually, recalling (9) and substituting (21) into the following equation, frequency domain transmit signal vector $\mathbf{X}[m]$ with respect to digitally modulated OFDM symbol vector $\mathbf{S}[m]$ as follows

$$\begin{aligned}\mathbf{X}[m] &= \mathbf{S}[m] + \mathbf{V}[m] \ \in \mathbb{C}^{N\times 1}\\ &= \frac{1}{N}\mathbf{F}\mathbf{F}_{\mathbf{z}}^{\dagger}\,\mathbf{S}[m],\end{aligned} \tag{22}$$

where note that $\frac{1}{N}\mathbf{F}\mathbf{F}_{\mathbf{z}}^{\dagger}$ is a precomputed matrix and $\mathbf{X}[m]$ is defined as

$$\mathbf{X}[m] = \left[X[m,k]\,;\, k = -\frac{N}{2}, -\frac{N}{2}+1, \ldots, \frac{N}{2}-1\right]. \tag{23}$$

As a result of (23), we can drop the zero terms on the right hand sides of (16) and (17) to obtain

$$\begin{aligned}\mathbf{s}^{(tr)}[m] &= \frac{1}{N}\mathbf{F}_0^{\dagger}\,\mathbf{S}^{(tr)}[m] \ \in \mathbb{C}^{N^{(tr)}\times 1}\\ \mathbf{s}^{(dt)}[m] &= \frac{1}{N}\mathbf{F}_3^{\dagger}\,\mathbf{S}^{(dt)}[m] \ \in \mathbb{C}^{N^{(dt)}\times 1}.\end{aligned} \tag{24}$$

Recalling the definition in (10), as a result of (24), the time domain OFDM transmit signal vector is given by

$$\mathbf{x}[m] = \frac{1}{N}\mathbf{F}_{\mathbf{z}}^{\dagger}\,\mathbf{S}[m]. \tag{25}$$

Using the observation model in (4), the vector definitions in (6), (7) and (10), the training observation model for an OFDM frame can be given by

$$\mathbf{y}^{(tr)}[m] = \sum_{\ell=0}^{L-1}\mathrm{diag}\left(\mathbf{s}_{\ell}^{(tr)}[m]\right)\,\mathbf{h}_{\ell}^{(tr)}[m] + \mathbf{w}^{(tr)}[m], \tag{26}$$

where complex valued $(N^{(tr)} - L + 1) \times 1$ vectors $\mathbf{y}^{(tr)}[m]$, $\mathbf{h}_{\ell}^{(tr)}[m]$ and $\mathbf{w}^{(tr)}[m]$ are defined as

$$\begin{aligned}\mathbf{y}^{(tr)}[m] &= [y[mN_g + L_c + n]; n = n_0 + L - 1, n_0 + L, \ldots, n_0 + N^{(tr)} - 1]\\ \mathbf{h}_{\ell}^{(tr)}[m] &= [h_{\ell}[mN_g + L_c + n]; n = n_0 + L - 1, n_0 + L, \ldots, n_0 + N^{(tr)} - 1]\\ \mathbf{w}^{(tr)}[m] &= [w[mN_g + L_c + n]; n = n_0 + L - 1, n_0 + L, \ldots, n_0 + N^{(tr)} - 1]\end{aligned} \tag{27}$$

and the ℓ-step circular shifted version of $\mathbf{s}^{(tr)}[m]$ is

$$\mathbf{s}_{\ell}^{(tr)}[m] = [x[m, n-\ell]; n = n_0 + L - 1, n_0 + L, \ldots, n_0 + N^{(tr)} - 1]. \tag{28}$$

Note that, if we substitute $\ell = 0, 1, \ldots, (L-1)$ in (28), the time range of $\mathbf{s}_\ell^{(tr)}[m]$ satisfies the whole training interval of $n_0 \leq n < (n_0 + N^{(tr)})$ in (10). In (26), defining

$$\begin{aligned}&\boldsymbol{\mathcal{S}}^{(tr)}[m]\\&= \left[\text{diag}\left(\mathbf{s}_0^{(tr)}[m]\right), \text{diag}\left(\mathbf{s}_1^{(tr)}[m]\right), \ldots, \text{diag}\left(\mathbf{s}_{L-1}^{(tr)}[m]\right)\right]\\&\in \mathbb{C}^{(N^{(tr)}-L+1)\times L(N^{(tr)}-L+1)}\end{aligned} \tag{29}$$

and

$$\begin{aligned}\mathbf{h}^{(tr)}[m] &= \left[\mathbf{h}_0^{(tr)}[m]\,;\, \mathbf{h}_1^{(tr)}[m]\,;\, \cdots\,;\, \mathbf{h}_{L-1}^{(tr)}[m]\right]\\&\in \mathbb{C}^{L(N^{(tr)}-L+1)\times 1},\end{aligned} \tag{30}$$

the training signal model in (26) can be expressed as

$$\mathbf{y}^{(tr)}[m] = \boldsymbol{\mathcal{S}}^{(tr)}[m]\,\mathbf{h}^{(tr)}[m] + \mathbf{w}^{(tr)}[m]. \tag{31}$$

However in (31), we can not still estimate unknown channel vector $\mathbf{h}^{(tr)}[m]$ since it's length is L times the length of observation vector $\mathbf{y}^{(tr)}[m]$ which means that there are more unknowns than observation equations. Therefore, in order to reduce the number of unknown channel coefficients, and make channel estimation possible, we introduce the basis expansion model (BEM) of the time-varying multipath channel in the following section.

B. Basis Expansion Model of Time-Varying Channel

The performance of the receiver critically depends on the estimate of the time-varying channel vectors $\{\mathbf{h}^{(tr)}[m]\}_{m=0}^{M-1}$. In (31), we have $L(N^{(tr)} - L + 1) \times 1$ unknown channel vector $\mathbf{h}^{(tr)}[m]$ to be estimated and a $(N^{(tr)} - L + 1) \times 1$ training observation vector $\mathbf{y}^{(tr)}[m]$. As we consider the whole OFDM frame with M-OFDM symbols, it seems the estimation of unknowns channel coefficient vectors $\{\mathbf{h}^{(tr)}[m]\}_{m=0}^{M-1}$ is impossible by means of observation equations $\{\mathbf{y}^{(tr)}[m]\}_{m=0}^{M-1}$ since there are more unknowns to be determined than known equations. In order to express the time variations of the channel by a finite number of parameters, BEM can be applied to approximate the time-varying channel vector $\mathbf{h}^{(tr)}[m]$ in (31). After applying the BEM, channel estimation problem turns to the estimation of the BEM coefficients, and using the estimated BEM coefficients, it will be possible to obtain the estimate of the time-varying multipath coefficient throughout the OFDM frame with M-OFDM symbols. As the time-varying multipath channel coefficients $\{\mathbf{h}_\ell[t],\, t = 0, 1 \cdots, (MN_g - 1)\}_{\ell=0}^{L-1}$ with discrete-time index t throughout the OFDM frame are essentially low-pass processes whose bandwidths are determined by the maximum Doppler frequency, they can be well-approximated by the weighted sum of substantially fewer number of orthonormal basis functions $\{\psi_q[t]\}$ in the discrete time interval $[0, MN_g - 1]$ as follows

$$h_\ell[t] = \sum_{q=0}^{Q-1} \psi_q[t]\, c_\ell[q], \quad t = 0, 1, \ldots, MN_g - 1, \tag{32}$$

where $c_\ell[q]$ is the qth BEM coefficient of the ℓth path. Similarly, using the orthogonality property of the basis functions, the expansion coefficients can be evaluated by the inverse transformation as

$$c_\ell[q] \;=\; \sum_{t=0}^{MN_g-1} \psi_q^*[t]\, h_\ell[t], \quad q = 0, 1, \ldots, Q-1. \tag{33}$$

Thus, for each channel path ℓ ($\ell = 0, 1, \ldots, L-1$), the channel and the BEM coefficients can be expressed in matrix form:

$$\mathbf{h}_\ell = \Psi \mathbf{c}_\ell, \tag{34}$$

$$\mathbf{c}_\ell = \Psi^\dagger \mathbf{h}_\ell, \tag{35}$$

where

$$\begin{aligned}\mathbf{h}_\ell &= \left[h_\ell[t]\,;\, t = 0, 1, \ldots, MN_g - 1\right] \in \mathbb{C}^{MN_g \times 1}\\ \mathbf{c}_\ell &= \left[c_\ell[q]\,;\, q = 0, 1, \ldots, Q-1\right] \in \mathbb{C}^{Q\times 1},\end{aligned} \tag{36}$$

and the matrix Ψ contains the orthonormal basis vectors as

$$\Psi = \left[\psi_0, \psi_1, \ldots, \psi_{(Q-1)}\right] \in \mathbb{R}^{MN_g \times Q} \tag{37}$$

together with the following column vector

$$\psi_q = \left[\psi_q[t]\,;\, t = 0, 1, \ldots, MN_g - 1\right],\; q = 0, 1, \ldots, Q-1\,. \tag{38}$$

In (32), the number of basis functions Q is a function of $f_D T_{\max}$ where f_D as given in (2) is the maximum Doppler frequency in Hz of the time-varying channel coefficients $\{h_\ell[t]\}_{\ell=0}^{L-1}$ of the multipath channel and $T_{\max}$ is the observation length in seconds. In this work, we take $T_{\max} = MN_g T_s$ that is the length of the frame with M-OFDM symbol where T_s is the sampling period. We will give the detailed calculations of how to choose the number of basis functions Q and the number of OFDM symbols M in a frame in Section IV-A.

In our work, we make use of a BEM, based on the orthonormal discrete Legendre polynomial (DLP) basis expansion model (DLP-BEM), to represent the time variations of the channel in an observation interval. DLP-BEM is well suited to represent the low-pass equivalent of the Doppler channel by means of a small number of basis functions [20]. Also, the DLP basis functions have the advantages of being independent of the channel statistics and having expansion coefficients that become uncorrelated as the number of observations MN_g gets larger, as proven in [20]. The Legendre polynomials are generated by carrying out Gram-Schmidt orthogonalization on the polynomials $\{1, t, t^2, \cdots\}$ with respect to the time-varying channels in a neighborhood of the middle point of the considered interval. The orthonormal Legendre polynomials are real valued and defined as [20]

$$\psi_q[t] = \frac{\nu_q[t]}{\sqrt{\sum_{t=0}^{MN_g-1} \nu_q^2[t]}}\,, \tag{39}$$

where $\{\nu_q[t]\}_{q=2}^{(Q-1)}$ denote discrete orthogonal Legendre polynomials that can be computed recursively for $t = 0, 1, \ldots, (MN_g - 1)$ as follows

$$\begin{aligned}\nu_q[t] = {}& \frac{(2q-1)(MN_g - 1 - 2t)}{q(MN_g - q)}\, \nu_{(q-1)}[t]\\ &- \frac{(q-1)(MN_g + q - 1)}{q(MN_g - q)}\, \nu_{(q-2)}[t],\end{aligned} \tag{40}$$

with the following initial polynomials and coefficients

$$\nu_0[t] = 1,\ \nu_1[t] = 1 - \frac{2t}{MN_g - 1}\,.$$

It follows from (34) that

$$\mathbf{h}_\ell^{(tr)}[m] = \Psi^{(tr)}[m]\,\mathbf{c}_\ell, \tag{41}$$

where

$$\begin{aligned}\Psi^{(tr)}[m] =&\ \left[\psi_0^{(tr)}[m], \psi_1^{(tr)}[m], \ldots, \psi_{(Q-1)}^{(tr)}[m]\right] \\ &\in \mathbb{R}^{(N^{(tr)}-L+1)\times Q} \\ \psi_q^{(tr)}[m] =&\ [\psi_q[mN_g + L_c + n]; n = n_0 + L - 1, n_0 \\ &+ L, \ldots, n_0 + N^{(tr)} - 1\Big]\,.\end{aligned} \tag{42}$$

Accordingly, the channel vector $\mathbf{h}^{(tr)}[m]$ in (31) can be expressed with respect to BEM coefficients as follows

$$\mathbf{h}^{(tr)}[m] = \mathbf{\Phi}^{(tr)}[m]\,\mathbf{c}\,, \tag{43}$$

where

$$\mathbf{c} = \left[\mathbf{c}_\ell\,;\, \ell = 0, 1, \ldots, L-1\right] \in \mathbb{C}^{QL\times 1} \tag{44}$$

$$\mathbf{\Phi}^{(tr)}[m] = \mathbf{I}_L \otimes \mathbf{\Psi}^{(tr)}[m] \in \mathbb{R}^{L(N^{(tr)}-L+1)\times QL} \tag{45}$$

and $\otimes$ denotes the Kronecker product. After substituting (43) into (31), the training signal model for mth OFDM symbol can be rewritten with respect to the unknown BEM coefficient vector $\mathbf{c}$ as follows

$$\mathbf{y}^{(tr)}[m] = \mathbf{Z}[m]\,\mathbf{c} + \mathbf{w}^{(tr)}[m]\,, \tag{46}$$

where $\mathbf{Z}[m] = \boldsymbol{\mathcal{S}}^{(tr)}[m]\,\mathbf{\Phi}^{(tr)}[m] \in \mathbb{C}^{(N^{(tr)}-L+1)\times QL}$ and $\boldsymbol{\mathcal{S}}^{(tr)}[m]$ is defined in (29). Stacking all the training observation vectors throughout the OFDM frame, $\{\mathbf{y}^{(tr)}[m]\}_{m=0}^{(M-1)}$, the overall training signal model for the estimation of the BEM coefficient vector is given by

$$\mathbf{y}^{(tr)} = \mathbf{Z}\,\mathbf{c} + \mathbf{w}^{(tr)}\,, \tag{47}$$

where

$$\begin{aligned}\mathbf{y}^{(tr)} =&\ \left[\mathbf{y}^{(tr)}[m]\,;\, m = 0, 1, \ldots, M-1\right] \in \mathbb{C}^{M(N^{(tr)}-L+1)\times 1} \\ \mathbf{Z} =&\ [\mathbf{Z}[m]\,;\, m = 0, 1, \ldots, M-1] \in \mathbb{C}^{M(N^{(tr)}-L+1)\times QL} \\ \mathbf{w}^{(tr)} =&\ \left[\mathbf{w}^{(tr)}[m]; m = 0, 1, \ldots, M-1\right] \in \mathbb{C}^{M(N^{(tr)}-L+1)\times 1}\end{aligned} \tag{48}$$

C. Estimation of the BEM and the Multipath Coefficients

Training signal model in (47) is linear with respect to $\mathbf{c}$. So, using (47), we obtain the linear minimum mean square error (MMSE) estimate of $\mathbf{c}$ as

$$\widehat{\mathbf{c}} = \Gamma_{\mathbf{c}}^{\dagger}\,\mathbf{y}^{(tr)}, \tag{49}$$

where the coefficient matrix $\Gamma_{\mathbf{c}}^{\dagger} \in \mathbb{C}^{QL\times M(N^{(tr)}-L+1)}$ that minimizes the average MSE, $\frac{1}{QL}E\{(\mathbf{c}-\widehat{\mathbf{c}})^\dagger(\mathbf{c}-\widehat{\mathbf{c}})\}$, is obtained as

$$\begin{aligned}\Gamma_{\mathbf{c}}^{\dagger} &= E\left\{\mathbf{c}\,\mathbf{y}^{(tr)\dagger}\right\}\left(E\left\{\mathbf{y}^{(tr)}\mathbf{y}^{(tr)\dagger}\right\}\right)^{-1} \\ &= \mathbf{R}_{\mathbf{c}}\mathbf{Z}^\dagger\left(\mathbf{Z}\mathbf{R}_{\mathbf{c}}\mathbf{Z}^\dagger + \sigma_w^2\,\mathbf{I}_{M(N^{(tr)}-L+1)}\right)^{-1}\end{aligned} \tag{50}$$

and applying the matrix inversion lemma, it can be expressed also as

$$\Gamma_{\mathbf{c}}^{\dagger} = \left(\mathbf{Z}^\dagger\,\mathbf{Z} + \sigma_w^2\mathbf{R}_{\mathbf{c}}^{-1}\right)^{-1}\mathbf{Z}^\dagger. \tag{51}$$

Since the BEM coefficient vector $\mathbf{c}_\ell$ is transformed by normally distributed ℓth path vector $\mathbf{h}_\ell$ as seen in (35), $\mathbf{c}_\ell$ is also normally distributed with $\mathbf{c}_\ell \sim \mathcal{N}(0, \mathbf{R}_{\mathbf{c}_\ell})$. Note that the covariance matrix of $\mathbf{c}$ in (44) can be determined from (35) and (2) as

$$\mathbf{R}_{\mathbf{c}} = \text{blkdiag}\left(\mathbf{R}_{\mathbf{c}_0}, \mathbf{R}_{\mathbf{c}_1}, \ldots, \mathbf{R}_{\mathbf{c}_{(L-1)}}\right), \tag{52}$$

where

$$\mathbf{R}_{\mathbf{c}_\ell} = \Psi^\dagger\mathbf{R}_{\mathbf{h}_\ell}\Psi. \tag{53}$$

Here, $\mathbf{R}_{\mathbf{h}_\ell}$ is the autocorrelation matrix of $\mathbf{h}_\ell$ and the entries of $\mathbf{R}_{\mathbf{h}_\ell}$ is obtained from (2). Both calculations of the coefficient matrix $\Gamma^\dagger$ in (50) and (51) can be pre-computed and do not contribute to the computational complexity, as will be also discussed in Section III-E. After estimating $\mathbf{c}$ using (49), and recalling the definition in (44), we can obtain the estimate of the subvectors,$\{\widehat{\mathbf{c}}_\ell\}_{\ell=0}^{L-1}$, by dividing $\widehat{\mathbf{c}}$ into subvectors with Q entries. Subsequently, similar to (41), the estimate of the $\mathbf{h}_\ell[m]$ defined in (6) is calculated as follows

$$\widehat{\mathbf{h}}_\ell[m] = \Psi[m]\widehat{\mathbf{c}}_\ell \in \mathbb{C}^{N\times 1}\,, \tag{54}$$

where

$$\begin{aligned}\Psi[m] =&\ \left[\psi_0[m], \psi_1[m], \ldots, \psi_{(Q-1)}[m]\right] \in \mathbb{R}^{N\times Q} \\ \psi_q[m] =&\ \left[\psi_q[mN_g + L_c + n]\,;\, n = 0, 1, \ldots, N-1\right].\end{aligned} \tag{55}$$

Similar to (45), it is straightforward from (54) that

$$\widehat{\mathbf{h}}[m] = \mathbf{\Phi}[m]\,\widehat{\mathbf{c}} \tag{56}$$

with the following definitions

$$\widehat{\mathbf{h}}[m] = \left[\widehat{\mathbf{h}}_0[m]; \widehat{\mathbf{h}}_1[m]; \cdots; \widehat{\mathbf{h}}_{L-1}[m]\right] \in \mathbb{C}^{NL\times 1} \tag{57}$$

$$\mathbf{\Phi}[m] = \mathbf{I}_L \otimes \Psi[m] \in \mathbb{R}^{NL\times QL}\,. \tag{58}$$

Eventually, using (8) and (57), the estimate of the channel convolution matrix $\mathbf{H}[m]$ in (5) is determined as

$$\widehat{\mathbf{H}}[m] = \sum_{\ell=0}^{L-1}\text{colshift}\left(\text{diag}(\widehat{\mathbf{h}}_\ell[m]), -\ell\right) \in \mathbb{C}^{N\times N} \tag{59}$$

that is required for data detection.

D. Symbol Detection

Denoting the BEM coefficient estimation error vector by ϵ such that $\mathbf{c} = \widehat{\mathbf{c}} + \epsilon$ because of orthogonality of zero-mean normally distributed vectors $\widehat{\mathbf{c}}$ and ϵ, and following the similar steps between (54) and (59), we obtain the channel convolution matrix in (5) as $\mathbf{H}[m] = \widehat{\mathbf{H}}[m] + \mathbf{\Xi}[m]$ where $\mathbf{\Xi}[m]$ represents

the estimation error matrix. Incorporating the channel estimation error into the receive signal model in (5), the data detection performance can be marginally improved. However, since our focus in this paper is pilot design and channel estimation, we do not consider channel estimation error in data detection calculations, and keep the computational complexity of the receiver low as well. Thus, in order to obtain observation model to detect the digitally modulated symbols transmitted within the mth OFDM signal, we substitute the estimate of the channel convolution matrix in (59) into the observation equation in (5) and we have

$$\mathbf{y}[m] = \widehat{\mathbf{H}}[m]\mathbf{x}[m] + \mathbf{w}[m]\,, \tag{60}$$

where, taking into account the column indices of the matrix $\widehat{\mathbf{H}}[m]$ corresponding to the time domain training and data indices, and recalling (10) and (15), the term $\widehat{\mathbf{H}}[m]\mathbf{x}[m]$ in (60) can be divided into the following two parts

$$\widehat{\mathbf{H}}[m]\mathbf{x}[m] = \widehat{\mathbf{H}}^{(tr)}[m]\mathbf{s}^{(tr)}[m] + \widehat{\mathbf{H}}^{(dt)}[m]\mathbf{s}^{(dt)}[m]. \tag{61}$$

Substituting (61) into (60) and keeping in mind that $\mathbf{s}^{(dt)}[m] = \frac{1}{N}\mathbf{F}_3^{\dagger}\,\mathbf{S}^{(dt)}[m]$ given in (17) , we can recompose the observation equation as follows

$$\mathbf{z}[m] = \Upsilon[m]\,\mathbf{S}^{(dt)}[m] + \mathbf{w}[m] \tag{62}$$

with

$$\mathbf{z}[m] = \mathbf{y}[m] - \widehat{\mathbf{H}}^{(tr)}[m]\mathbf{s}^{(tr)}[m] \;\in \mathbb{C}^{N\times 1} \tag{63}$$

$$\Upsilon[m] = \frac{1}{N}\widehat{\mathbf{H}}^{(dt)}[m]\mathbf{F}_3^{\dagger} \;\in \mathbb{C}^{N\times N^{(dt)}} \tag{64}$$

and also note that the term $\widehat{\mathbf{H}}^{(tr)}[m]\mathbf{s}^{(tr)}[m]$ in (63) is known. Using (62), the linear MMSE estimate of the data vector of the mth OFDM symbol is given by

$$\widehat{\mathbf{S}}^{(dt)}[m] = \Gamma_{\mathbf{d}}^{\dagger}[m]\,\mathbf{z}[m], \tag{65}$$

where the coefficient matrix $\Gamma_{\mathbf{d}}^{\dagger}[m] \;\in \mathbb{C}^{N^{(dt)}\times N}$ for the mth OFDM symbol is obtained as follows

$$\begin{aligned}\Gamma_{\mathbf{d}}^{\dagger}[m] &= E\left\{\mathbf{S}^{(dt)}[m]\,\mathbf{z}^{\dagger}[m]\right\}\left(E\left\{\mathbf{z}[m]\,\mathbf{z}^{\dagger}[m]\right\}\right)^{-1}\\ &= \Upsilon^{\dagger}[m]\left(\Upsilon[m]\Upsilon^{\dagger}[m] + \sigma_w^2\,\mathbf{I}_N\right)^{-1}\end{aligned} \tag{66}$$

and applying the matrix inversion lemma, it can be expressed also by

$$\Gamma_{\mathbf{d}}^{\dagger}[m] = \left(\Upsilon^{\dagger}[m]\,\Upsilon[m] + \sigma_w^2\,\mathbf{I}_{N^{(dt)}}\right)^{-1}\Upsilon^{\dagger}[m]. \tag{67}$$

We prefer the expression in (67) while calculating $\widehat{\mathbf{S}}^{(dt)}[m]$ in (65) since matrix inverse operation in (67) has less computational complexity than that of (66). After estimating the unknown data vector $\mathbf{S}^{(dt)}[m]$ of the mth OFDM symbol by employing the linear MMSE estimator in (65) together with (63), (64) and (67), digitally modulated data symbols are detected by rounding to the nearest constellation point.

E. Computational Complexity of the Channel Estimator

In this subsection, we compare our estimator with [20] and [11]. In order to evaluate the computational complexity of the channel estimators, the computational loads of our proposed estimator, and that of [20] and [11] are calculated in terms of complex multiplications (CMs) and complex additions (CAs). The computational complexities of the estimators are determined by the number of BEM coefficients per path Q, the number of paths L, the number of OFDM symbols within one OFDM frame M, the number of pilot symbols $N^{(tr)}$ and the pilot spacing Δ. Table I lists the overall computational complexities of the methods with respect to CMs and CAs for estimation of the unknown parameters $\mathbf{c}$. Since it is stated in [20] that the iterative SAGE algorithm converges in 4 iterations, the computational complexity of the SAGE algorithm is calculated under this assumption. On the other hand, since the MMSE-PAT based channel estimator in [11] is a batch estimator, it is a linear transformation of observation vector such that

$$\widehat{\mathbf{c}} = \mathbf{T}\mathbf{y}, \tag{68}$$

where $\mathbf{T} \in \mathbb{C}^{QL\times MN}$ is a precomputed transformation matrix that depends on the training signal, data and noise statistics, and recalling (6), the observation vector $\mathbf{y} \in \mathbb{C}^{MN\times 1}$ is defined as

$$\mathbf{y} = \left[\mathbf{y}^T[0], \mathbf{y}^T[1], \ldots, \mathbf{y}^T[M-1]\right]^T. \tag{69}$$

From the matrix and vector multiplication on the right hand side of (68), the estimate of $\mathbf{c}$ in [11] needs a total of $QLMN$ CMs and $QL(M(N-1)$ CAs as seen in Table I. On the other hand, in our case, the coefficient matrix $\Gamma_{\mathbf{c}}^{\dagger} \in \mathbb{C}^{NL\times M(N^{(tr)}-L+1)}$ is pre-computed matrix since it is completely dependent on the training signal, and thus, the our proposed estimator of $\mathbf{c}$ in (49) requires total $QLM(N^{(tr)}-L+1)$ CMs and $QL(M(N^{(tr)}-L+1)-1)$ CAs as given in the last row of Table I. From the computational complexities given in the Table I, recalling that $N^{(tr)} < N$, it is clear that the computational complexity of our proposed method is considerably less than that of [20] and [11]. Finally, since $N^{(tr)} = N/\Delta = N/8$ considering the parameters in Table III, the total computational load per OFDM symbol to implement the estimation algorithm employing the proposed training scheme is approximately $\mathcal{O}(QLN/8)$ which means that the computational load of our proposed scheme is significantly lower.

IV. Calculation of Doppler Related OFDM Parameters

Doppler frequency due to the relative motion between transceivers causes time variations of the paths of the channel. The polynomial degree of the channel paths increases with the increasing the maximum Doppler frequency. On the other hand, recalling that the qth Legendre polynomial is actually a $(q+1)$th degree time polynomial, it can be concluded that the number of Legendre polynomials (Q) per path increases as the maximum Doppler frequency increases. Thus, the number of OFDM symbols (M) within one OFDM frame also increases with increasing maximum Doppler frequency for a fixed number of observation equation (K) per unknown BEM coefficients.

A. Calculation of Q and M

In this section, we calculate Q and M for a given maximum Doppler frequency f_D, the number of the subcarriers N, the number of paths L, the pilot spacing Δ, the cyclic prefix length L_c and the average number of equations per unkown BEM

TABLE I
COMPUTATIONAL COMPLEXITY COMPARISON OF [20], [11] AND THE PROPOSED METHOD FOR $\widehat{\mathbf{c}}$

SAGE Method in [20]		
Operation	CMs	CAs
Initialization	$MNQL$	$MN(Q-1)L$
E-Step	$4MNQL^2+4MNL(\Delta+L)+2MNL(N+1)$	$4MN(Q-1)L^2+4MNL(\Delta+L)+2MNL(N+1)$
M-Step	$2MNL(Q+1)(QL+2)$	$4M(N-1)QL+2MNL^2Q(Q+1)$
MMSE-PAT Method in [11]		
Operation	CMs	CAs
$\widehat{\mathbf{c}}=\mathbf{Ty}$	$QLMN$	$QL(MN-1)$
Proposed Method		
Operation	CMs	CAs
$\widehat{\mathbf{c}}$ in (49)	$QLM(N^{(tr)}-L+1)$	$QL\big(M(N^{(tr)}-L+1)-1\big)$

TABLE II
f_DT_{MAX} AND $\mathcal{Q}$ VALUES AT $\text{MSE}_{mod}=10^{-8}$

f_DT_{max}	$\mathcal{Q}$	f_DT_{max}	$\mathcal{Q}$	f_DT_{max}	$\mathcal{Q}$
0	1.000	0.08	3.228	0.7	7.325
0.01	1.979	0.09	3.339	0.8	7.729
0.02	2.276	0.1	3.446	0.9	8.314
0.03	2.495	0.2	4.332	1	8.789
0.04	2.676	0.3	5.052	2	13.120
0.05	2.837	0.4	5.677	3	17.090
0.06	2.983	0.5	6.258	4	20.880
0.07	3.110	0.6	6.806		

TABLE III
DOPPLER RELATED OFDM PARAMETERS FOR $\text{MSE}_{Mod}=10^{-8}$, $N=1024$, $\Delta=8$ AND $L=5$

f_{Dnorm}	Case	M	Q	K
0.02	I	1	2	12.40
	II	3	3	24.80
	III	4	3	33.06
0.08	I	2	4	12.40
	II	6	6	24.80
	III	12	9	33.06

coefficient K. In [20], it is pointed out that Discrete Legendre Polynomials are well suited to represent the low-pass equivalent representation of the time varying multipath channels. In order to determine the Q and M, we first consider as a metric the modeling MSE expression given in (80) in Appendix-VII-A in which we derive the expression of the modeling MSE. From (80), the modeling MSE for a single path can be given by

$$\text{MSE}_{mod}=\frac{1}{N_{tot}}\,\text{tr}\left[(\mathbf{I}_{N_{tot}}-\Psi\Psi^{\dagger})\mathbf{R}\right]\,, \tag{70}$$

where tth row t'th column entry of the matrix $\mathbf{R}$ is given by $[\mathbf{R}]_{t,t'}=J_0(2\pi f_DT_{\max}(t-t')/N_{tot})$, $T_{\max}=N_{tot}T_s$ is the CP-OFDM frame length in continuous time-domain and $N_{tot}=MN_g$ total number of time samples within CP-OFDM frame duration. In (70), the modeling MSE depends on the parameters $f_DT_{\max}$, Q (the number of Legendre polynomials, in other words, the number of columns of Ψ) and N_{tot} only. In (70), the average value of main diagonal entries of the matrix inside the trace operator is calculated by summing N_{tot} entries first by trace operator and then dividing by N_{tot}. In other words, this operation is a form of averaging by discrete integration. So, we can interpret N_{tot} as the resolution of this discrete integration which means that, for fixed $f_DT_{\max}$ and Q, as N_{tot} increases precision of MSE_{mod} increases, and vice versa. In order to observe how the modeling MSE changes with the parameters Q and $f_DT_{\max}$, we plot MSE_{mod} vs Q curves for different $f_DT_{\max}$

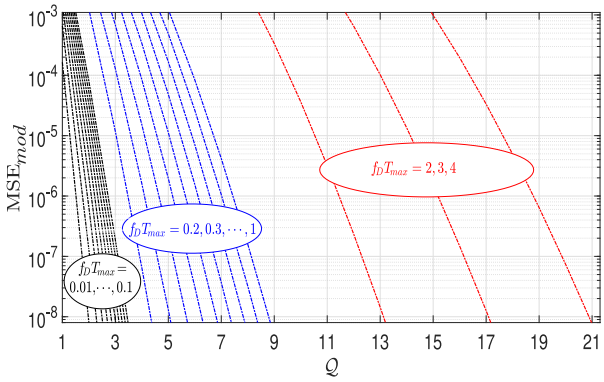


Fig. 1. Modeling MSE vs Number of Legendre Polynomials (Q) for a single path.

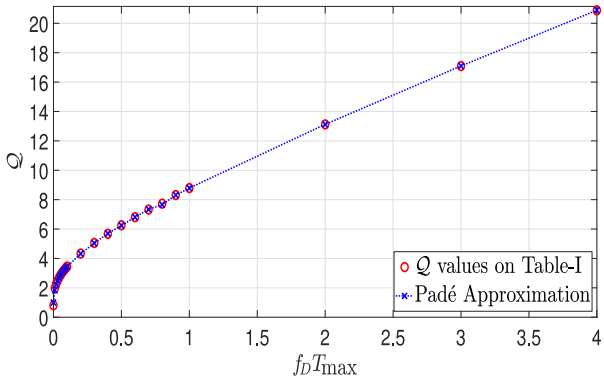


Fig. 2. Number of BEM Coefficients per Path ($\mathcal{Q}$) vs $f_DT_{\max}$.

values as seen Fig. 1. Considering the plots in Fig. 1, we want to express the relationship between Q and $f_DT_{\max}$ for a desired modeling MSE level. In [20], the minimum value of MSE is arround 10^{-4}, and in order not to dominate the estimation MSE, we choose modeling MSE level at 10^{-8}. Even though the number of Legendre polynomials Q is an integer, in order to minimize the approximation error, we estimate the number of Legendre polynomials as a real number while reading from Fig. 1. This estimate is a function of $f_DT_{\max}$ and denoted by $\mathcal{Q}(f_DT_{\max})$. In Table II, $f_DT_{\max}$ versus $\mathcal{Q}$ values are listed for $\text{MSE}_{mod}=10^{-8}$. Using $f_DT_{\max}$ versus $\mathcal{Q}$ values for $\text{MSE}_{mod}=10^{-8}$ in Table II, the function $\mathcal{Q}(f_DT_{\max})$ and its Padé approximation [32] are plotted in Fig. 2. From the curve plotted for the true values of $\mathcal{Q}(f_DT_{\max})$, one can easily note that two different behaviours in terms of the degrees of the function for high and low $f_DT_{\max}$

values. In order to model these two behaviours, we approximate the function as the ratio of two polynomials of $f_D T_{\max}$ by employing the Padé method. The approximate function $\mathcal{Q}(f_D T_{\max})$ is expressed as follows

$$\mathcal{Q}(\xi) = \frac{15.432\xi^5+11.979\xi^4-27.115\xi^3+3.653\xi^2+2.071\xi+0.010}{0.048\xi^5+3.738\xi^4-2.593\xi^3-1.222\xi^2+0.705\xi+0.010}, \tag{71}$$

where $\xi = f_D T_{\max}$. Although small fitting errors can be tolerated for the values of $\mathcal{Q}(f_D T_{\max})$ around an integer since $\mathcal{Q}(f_D T_{\max})$ is rounded to a nearest integer Q, as seen from Fig. 1, the approximate function $\mathcal{Q}(f_D T_{\max})$ in (71) fits the values in Table II well. Recalling (73), $N_g = N + L_c$ and assuming $N \gg L_c$, the parameter ξ of the $\mathcal{Q}$ function in (71) can be expressed in terms of normalized maximum Doppler frequency $f_{D_{norm}}$ and the OFDM frame length M as follows

$$\begin{aligned} \xi &= f_D T_{\max} \\ &= f_D N_{tot} T_s \\ &= f_D M N_g T_s \\ &= f_{D_{norm}} M \frac{N_g}{N} \end{aligned} \tag{72}$$

where the normalized maximum Doppler frequency $f_{D_{norm}}$ is defined as the ratio of the maximum Doppler frequency to one of the OFDM subband width and given by

$$f_{D_{norm}} = \frac{f_D}{F_s/N} = f_D T_s N, \tag{73}$$

with $F_s = 1/T_s$. As we mention in Section III-B, by applying the BEM, we reduce the total number unknown channel parameters to the total number of unknown BEM coefficients, $Q(\xi)L$. As seen from (71) and (72), the total number of unknown BEM coefficients is a function of maximum Doppler frequency and the number OFDM symbols per frame, $\mathcal{Q}(f_{D_{norm}} M N_g/N)L$. So, the number of the unknown BEM coefficients per OFDM symbol is calculated as $\mathcal{Q}(f_{D_{norm}} M N_g/N)L/M$. We assume comb-type training (pilot) structure in which equally spaced pilots are carried by each OFDM symbol with N subcarriers and denote the pilot spacing by Δ. Recalling the number of observations per OFDM symbol is $N^{(tr)} - L + 1$ as seen from the dimension of $\mathbf{y}^{(tr)}[m]$ in (27) and the number time domain training samples is equal to the number of pilots that means $N^{(tr)} = \lfloor N/\Delta \rfloor$ where $\lfloor \cdot \rfloor$ represents the floor operator, we can express the number of observations per OFDM symbol in two ways as follows

$$K \frac{\mathcal{Q}\left(f_{D_{norm}} M N_g/N\right) L}{M} = \frac{N}{\Delta} - (L-1), \tag{74}$$

where we denote the desired average number observations per unknown BEM coefficient by K. However, the equality in (74) can not hold exactly always since the parameters involving in it are integers, but we can choose these integer parameters so as to minimize the difference between both sides of (74). Accordingly, for given K, $f_{D_{norm}}$, L, N_g, N and Δ, the parameter M that is the number of OFDM symbols per frame is obtained with the help of (71) by performing the following simple search algorithm

$$M = \arg\min_{M'} \left| \frac{\mathcal{Q}\left(f_{D_{norm}} M' N_g/N\right)}{M'} - \frac{N-(L-1)\Delta}{KL\Delta} \right|, \quad M' = 1, 2, \cdots \tag{75}$$

So, substituting M as the parameter of (71) and rounding the value of the function to the nearest integer we determine the number of Legendre polynomials as

$$Q = \lfloor \mathcal{Q}(f_{D_{norm}} M N_g/N) \rceil, \tag{76}$$

where $\lfloor \cdot \rceil$ represents the operator for rounding to nearest integer. Increasing the value of Q regardless of the solution of $\mathcal{Q}(f_D T_{\max})$ and M employing the algorithm in Eq. (75) causes an increasing number of unnecessary unknown channel parameters to be estimated. Accordingly, estimator performance degrades as the number of unnecessary channel parameters increases. This causes a performance loss of the receiver and also additional computational complexity. Therefore, using the ceiling operator rather than the nearest integer in (76) might not be reasonable. This is why we chose an operator for rounding to the nearest integer in (76).

V. Simulation Results

In this section, computer simulation results to assess the performance of the rapidly time-varying OFDM system operating with the proposed training model and the channel estimation algorithm are presented. Considering given parameters $N = 1024, \Delta = 8$ and $L = 5, \text{MSE}_{mod} = 10^{-8}$, using (71), (75) and (76), Doppler related parameters M and Q are calculated to provide fairness between different Doppler scenarios achieved by keeping the average number of equations per unknown, K, fixed. These calculated parameters for each of the three K values are listed in Table III as Case-I, Case-II and Case-III, respectively.

In simulations, we assume that the normalized power of the ℓth path obeys the following exponential power decaying profile

$$\Omega_\ell = \frac{e^{-\ell/L}}{\sum_{l=0}^{L-1} e^{-l/L}}. \tag{77}$$

Based on the proposed ICI-free training signal model, the estimates of the time-varying multipath channel is performed by the low complexity MMSE estimation technique. In order to investigate the performance of the OFDM system operating with the proposed scheme, we plot MSE and SER performances for $f_{D_{norm}} = 0.02, 0.08$ and $K = 12.40, 24.80, 33.06$. In addition, we also consider Doppler mismatch scenarios where the Doppler frequency assumed on the receiver side is different from its true value. In these figures, we compare MSE and SER performance results with the results of CRLB and the perfectly known channel state information (CSI) cases, respectively. Theoretical calculations of modeling and estimation MSEs are given in Appendix A, Appendix B, respectively. In Appendix C, the expression of the overall MSE is given by showing that it can be calculated as a sum of modeling and estimation MSEs. In addition, in Appendix D, we derive the overall MSE expression for Doppler frequency mismatch scenarios. Finally in Appendix E, in order

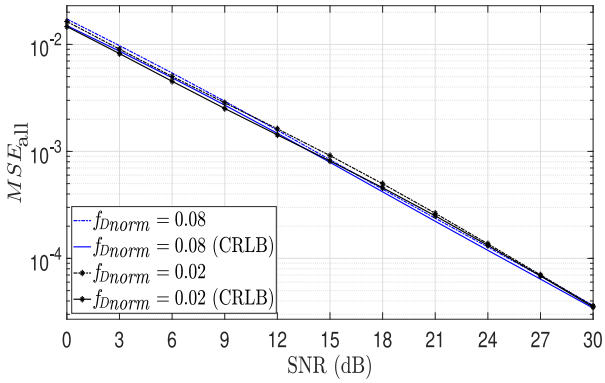


Fig. 3. Overall MSE vs. SNR performances (using Case-II parameters).

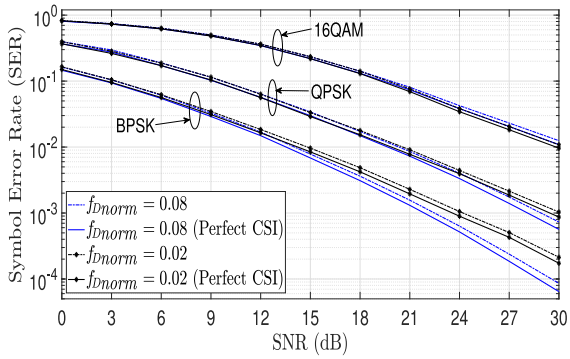


Fig. 4. SER vs. SNR performances (using Case-II parameters).

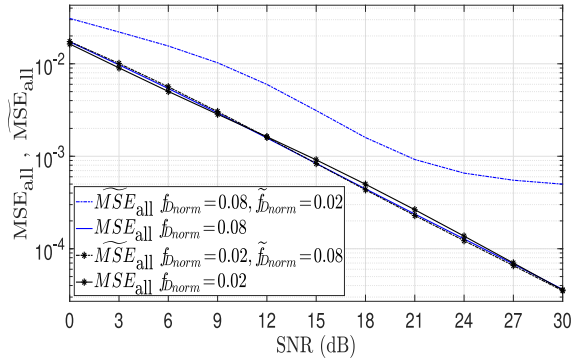


Fig. 5. Effect of Doppler mismatch on the MSE performance (using Case-II parameters).

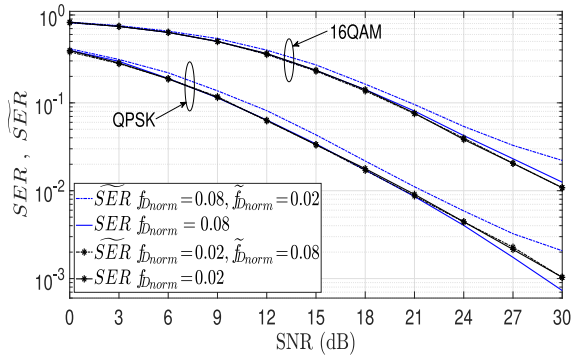


Fig. 6. Effect of Doppler mismatch on the SER performance (using Case-II parameters).

to present performance benchmark for the estimator of the complex valued multipath channel coefficients, we also present an analytical expression of the Bayesian CRLB.

The plots in Figs. 3 and 4 show the overall MSE and the SER performances of the system for fair comparison parameters of Case-II scenarios in Table III. It is widely reported in the literature that the higher the Doppler frequency, the lower the performance of the channel estimation and therefore the receiver. Unlike the commonly reported in the literature, MSE curves of the proposed channel estimator plotted for lower and higher Doppler scenarios in Fig. 3 are almost the same which confirms that the MSE performance of the estimator for high Doppler scenario is as good as that of small Doppler scenario, and this verifies our Doppler related calculations in Section IV-A for fairness under different Doppler scenarios. On the other hand, we also observe that the MSE performances almost attain the corresponding CRLBs as the contribution of our proposed ICI-free training signal model. In Fig. 4, SER performance curves plotted for the channel estimation are very close to that of perfectly known channel state information (CSI) cases. It is also seen that a better SER performance is obtained for the higher Doppler frequency, which means that our proposed training scheme and Doppler related calculations in Section IV-A for fairness provide time diversity as well.

Figs. 5 and 6 exhibit the effects of maximum Doppler frequency mismatch on the overall MSE and the SER performances for true ($f_{D\,norm}$) and mismatch ($\widetilde{f_{D\,norm}}$) values of the normalized maximum Doppler frequencies at 0.02 and 0.08 assuming Case-II parameters in Table III. In these figures, Doppler mismatch scenarios are addressed in two ways; 1.) the maximum Doppler frequency on the receiver side is considered at a much higher than its true value, 2.) the maximum Doppler frequency on the receiver side is considered at a much lower than its true value. In the first scenario, because the maximum Doppler frequency is considered to be higher than its true value, an unnecessarily higher number of unknown BEM coefficients per path are used, and therefore a slight loss in the MSE performance of the channel estimator is observed. On the other hand, with an increasing number of some unnecessarily used unknown BEM coefficients, the computational complexity of the estimator also begins to increase. In the second scenario which is the opposite of the first scenario, because the maximum Doppler frequency is considered at a much lower than its true value, fewer number of BEM coefficients per path than the required is used. This means to try to represent the time-varying channel with a polynomial degree less than a degree that fits channel variations, and therefore due to a considerable interpolation error, a substantial loss in the MSE performance of the estimator is observed. The MSE performance plots for those two Doppler mismatch scenarios are given in Fig. 5. In Fig. 6, as compared to the SER curves plotted for the cases without Doppler mismatch, the loss in SER performance observed in the second Doppler mismatch scenario is much higher than that of the first Doppler mismatch scenario. Although SER performance loss seems to be almost absent in the first scenario, the loss for this scenario can be interpreted as

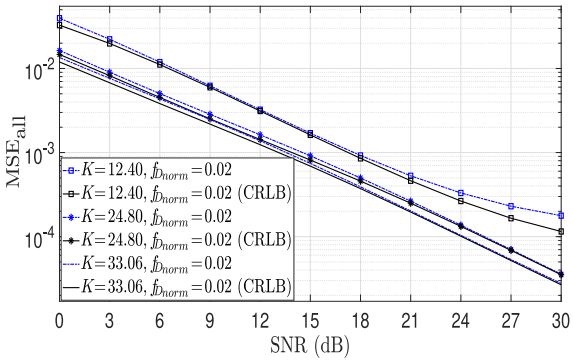


Fig. 7. Effect of K on the MSE performance ($f_{D\,norm} = 0.02$).

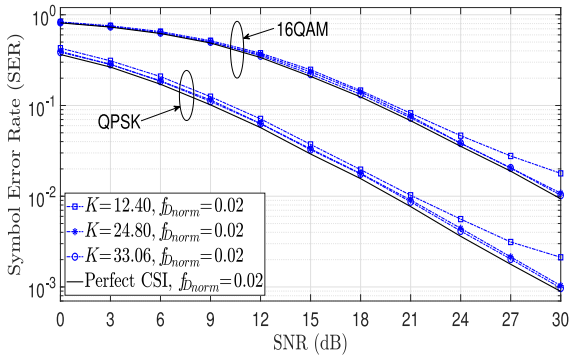


Fig. 8. Effect of K on the SER performance ($f_{D\,norm} = 0.02$).

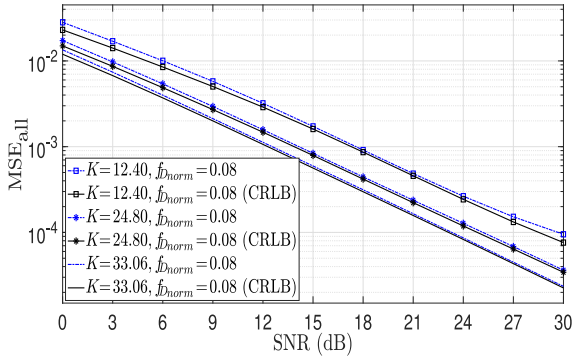


Fig. 9. Effect of K on the MSE performance ($f_{D\,norm} = 0.08$).

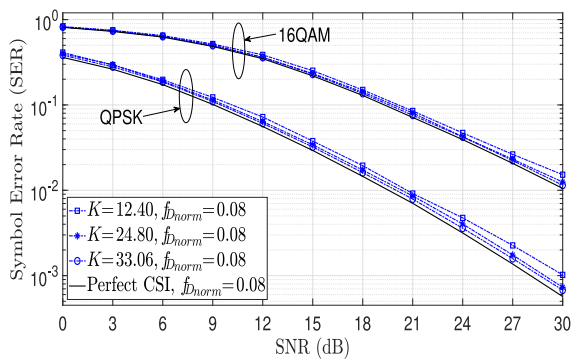


Fig. 10. Effect of K on the SER performance ($f_{D\,norm} = 0.08$).

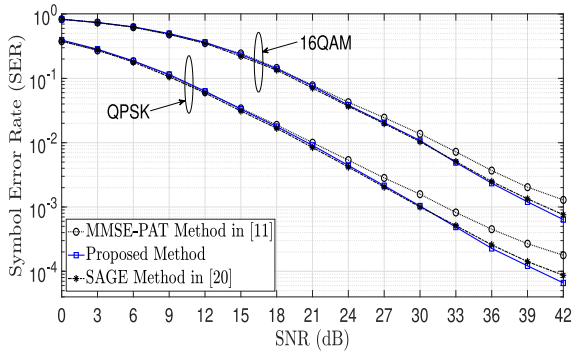


Fig. 11. Performance comparisons of the proposed scheme with [11] and [20] for $f_{D\,norm} = 0.02$.

an increased computational load because of some unnecessarily used unknown BEM coefficients per path.

Figs. 7 and 8 depict the effect of number of observations per unknown DLP-BEM coefficients on the overall MSE and SER performances for the normalized maximum Doppler frequency of 0.02 assuming Case-I, Case-II and Case-III parameters in Table III. In Fig. 7, although the MSE performances of the channel estimator for three cases are quite close to the corresponding CRLBs at SNR levels below 20 dB, the MSE performance for $K = 12.40$ is limited by an error floor at SNR levels beyond 20 dB and the MSE performances for $K = 24.80$ and $K = 33.06$ asymptotically attain the bounds. In addition, as seen from the SER curves in Fig. 8, SER performance for perfect CSI case is almost achieved for $K = 24.80$ and $K = 33.06$, while a loss in the SER performance is observed for $K = 12.40$ at SNR levels beyond 20 dB which means there is almost no SER performance gain between the cases of $K = 24.80$ and $K = 33.06$. As a result, considering the trade-off between computational load due to unnecessary BEM coefficients and MSE/SER performances, we can conclude that there is diminishing returns when K is increased for this system scenario.

The MSE and the SER performance curves in Figs. 9 and 10 are plotted by repeating the simulation scenario in Figs. 7 and 8, respectively, for the normalized maximum Doppler frequency of 0.08 and the corresponding parameters in Table III. As compared the MSE curves in Fig. 9 with that of Fig. 7, although the normalized maximum Doppler frequency is higher in Fig. 9, in the similar MSE manners as in Fig. 7, better MSE performances in all three cases are obtained as well as the MSE performance loss compared to CRLB is also reduced in Case-I as seen from Fig. 9. In Fig. 10, as a corroboration of the fact that the proposed training scheme and the Doppler related calculations in Section IV-A required to obtain time diversity at higher Doppler levels, better SER performances compared to the SER curves in Fig. 8 are obtained, which are more clearly seen especially for QPSK constellation. In addition, the loss of SER performance for Case-I in Fig. 8 is decreased as seen in Fig. 10.

In Figs. 11 and 12, we compare the SER performance of our proposed scheme with that of [20] and [11] for the normalized maximum Doppler frequencies at 0.02 and 0.08, respectively. In [20], an iterative space alternating generalized expectation-maximization (SAGE) algorithm is proposed for channel estimation in OFDM systems operating over frequency selective and mobile wireless channels. On the other hand, [11] considers a design problem of pilot-aided cyclic-prefixed block

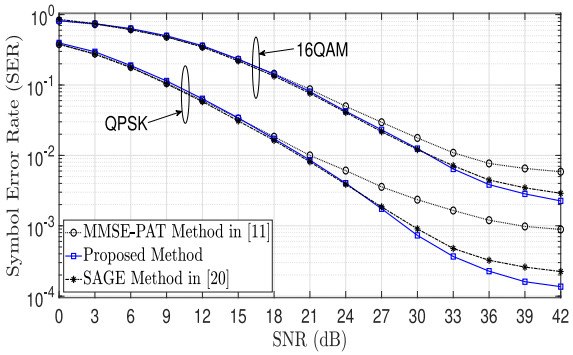


Fig. 12. Performance comparisons of the proposed scheme with [11] and [20] for $f_{Dnorm} = 0.08$.

transmissions for doubly selective channels which disperse the transmitted signal in both the time and frequency domains and assumes that the multipath channel varies independently from block to block. In the proposed scheme in [11], pilot symbols as many as the number of paths are placed in the frequency domain, and the neighboring subcarriers as the number of complex exponential BEM coefficients to the right and left of each pilot symbol are zero padded. The pilot-aided transmission (PAT) design criterion in [11] relies on the minimum MSE (MMSE) of the channel estimator. According to this MMSE-PAT criterion, the pilot symbols, which are equal to the number of paths, are much more strengthened than data symbols with respect to the minimum MSE performance limit of the channel estimator. In our simulations in order to provide fair performance comparison between [20], [11] and our proposed scheme, we consider the same total transmit power and the same training power for each scheme in the same scenario. As seen SER performance curves in Fig. 11, while our proposed outperforms the MMSE-PAT method proposed in [11] especially at high SNR levels beyond 21 dB, it is almost the same as the performance of the proposed SAGE algorithm in [20] at low and mid-range SNR values but it starts to perform slightly better than [20] at SNR levels beyond 33 dB and 36 dB for QPSK and 16QAM modulation types, respectively. The primary reason our proposed scheme outperforms the scheme in [11] at high SNR levels is that the method in [11] assumes that the multipath channel is changing from block to block, which is not always suitable for the rapidly varying multipath channel scenario. The secondary reason is that [11] uses fewer pilot symbols and therefore more data symbols as compared to our scheme, and therefore [11] has to allocate the same total data power to a larger number of data symbols resulting weak data symbols. In other words, the scheme in [11] is not power efficient as compared our proposed scheme. In addition, the fact that it starts to be better than the SAGE algorithm at high SNR levels can be evaluated as the beginning of the time diversity gain of our proposed method at the normalized the Doppler frequency of 0.02. However, we can not observe this at low and mid-range SNR levels because AWGN dominates the performance of the estimators. As seen from the SER curves in Fig. 12, we observe that the performance gap between our proposed method and the MMSE-PAT method proposed in [11] is further increased at SNR levels beyond 21 dB and 24 dB for QPSK and 16QAM modulation types, respectively, and the performance gap in the case of QPSK is much larger than the case of 16QAM. In addition, it is also seen in Fig. 12 that the SER performance of our proposed method starts to get much better than the performance of the iterative SAGE method proposed in [20] at SNR levels beyond 27 dB and 30 dB for QPSK and 16QAM modulations, respectively. Similar to Fig. 11, the performance gap in the case of QPSK is much larger than the case of 16QAM. This is due to the time diversity gain of our proposed method being much higher than that of iterative SAGE algorithm in [20] at the higher normalized Doppler frequency of 0.08.

VI. Conclusion

In this work, based on the sub-spaces of the training and data signal spaces, we propose an ICI-free time domain training signal model for a computationally and performance efficient estimation of frequency selective and rapidly time-varying multipath channels in OFDM systems. To reduce the number of unknown channel parameters, DLP-BEM have been employed to represent the rapidly time-varying fading channel. As another contribution of this work, we have provided also an algorithm that can determine the number of OFDM symbols per frame and the number of BEM coefficients per path for a given number of subcarriers, number of paths, pilot spacing and the average number of observations per BEM coefficient. This is done while maintaining fairness between lower and higher Doppler scenarios achieved by keeping the average number of observations per unknown BEM coefficient fixed. While doing this, using Padé rational fraction approximation method, we analytically expressed the number of BEM coefficients per path as a function of maximum Doppler frequency and the length of OFDM frame for a fixed modeling error. In order to present performance benchmark for the estimator of the complex valued DLP-BEM channel coefficients, the analytical expressions of the exact Bayesian CRLB and theoretical MSEs have been derived considering maximum Doppler frequency mismatch scenario as well.

As shown in computer simulations, with the help of proposed scheme and the Doppler related parameter calculations required for fairness at different Doppler levels, for a reasonable number of observation equations per BEM coefficient, an MSE performance that is very close to the corresponding Bayesian CRLB, and a resulting SER performance that almost attains to the SER performance of the perfect CSI case are obtained with a smaller computational load during the estimation of the rapidly varying multipath channel in an OFDM system. We conclude that, unlike the common results in the literature, an OFDM system can perform better as the Doppler frequency increases as long as the system setup for exploiting higher Doppler frequency is provided.

Appendix

A. Calculation of Modeling MSE

In (34), $MN_g \times 1$ multipath channel vector $\mathbf{h}_\ell$ is approximated by $\Psi \mathbf{c}_\ell$ using a lower dimension $Q \times 1$ BEM coefficient vector $\mathbf{c}_\ell$. Considering also (35), modeling error vector for ℓth

path is obtained as follows

$$\begin{aligned}\epsilon_\ell &= \mathbf{h}_\ell - \Psi \mathbf{c}_\ell \\ &= (\mathbf{I}_{MN_g} - \Psi\Psi^\dagger)\mathbf{h}_\ell. \end{aligned} \tag{78}$$

Noting that $\Psi^\dagger\Psi = \mathbf{I}_Q$, the average MSE for modeling the long multipath channel vector $\mathbf{h} = [\mathbf{h}_\ell\,;\, \ell = 0, 1, \ldots, (L-1)]$ is given by

$$\begin{aligned}\text{MSE}_{mod} &= \frac{1}{MN_gL}\sum_{\ell=0}^{L-1} E\{\epsilon_\ell^\dagger \epsilon_\ell\} \\ &= \frac{1}{MN_gL}\sum_{\ell=0}^{L-1} E\left\{\mathbf{h}_\ell^\dagger (\mathbf{I}_{MN_g} - \Psi\Psi^\dagger)^\dagger (\mathbf{I}_{N_{tot}} - \Psi\Psi^\dagger)\mathbf{h}_\ell\right\} \\ &= \frac{1}{MN_gL}\sum_{\ell=0}^{L-1} E\left\{\mathbf{h}_\ell^\dagger (\mathbf{I}_{MN_g} - \Psi\Psi^\dagger)\mathbf{h}_\ell\right\} \\ &= \frac{1}{MN_gL}\sum_{\ell=0}^{L-1} \text{tr}\left[(\mathbf{I}_{MN_g} - \Psi\Psi^\dagger)\mathbf{R}_{h_\ell}\right], \end{aligned} \tag{79}$$

where $\mathbf{R}_{h_\ell}$ is the autocorrelation matrix of the ℓth multipath and its entries are obtained from (2). As seen from (2), $\mathbf{R}_{h_\ell}$ can be written as $\mathbf{R}_{h_\ell} = \Omega_\ell \mathbf{R}$. Substituting $\mathbf{R}_{h_\ell} = \Omega_\ell \mathbf{R}$ into (79) and noting that $\sum_{\ell=0}^{L-1}\Omega_\ell = 1$, we arrive the following general expression of MSE_{mod}

$$\text{MSE}_{mod} = \frac{1}{MN_gL}\,\text{tr}\left[(\mathbf{I}_{MN_g} - \Psi\Psi^\dagger)\mathbf{R}\right], \tag{80}$$

where tth row t'th column entry of the matrix $\mathbf{R}$ is given by $[\mathbf{R}]_{t,t'} = J_0(2\pi f_D T_s(t-t'))$ using (2).

B. Calculation of Estimation MSE

Recalling (34), the average MSE for the Bayesian estimator of the $MN_gL \times 1$ multipath vector $\mathbf{h} = [\mathbf{h}_\ell; \ell = 0, 1, \ldots, (L-1)]$ is calculated as follows

$$\begin{aligned}\text{MSE}_{est} &= \frac{1}{MN_gL} E_{\mathbf{h},\hat{\mathbf{h}}}\left\{(\mathbf{h}-\hat{\mathbf{h}})^\dagger(\mathbf{h}-\hat{\mathbf{h}})\right\} \\ &= \frac{1}{MN_gL}\sum_{\ell=0}^{L-1} E_{\mathbf{h}_\ell,\hat{\mathbf{h}}_\ell}\left\{(\mathbf{h}_\ell-\hat{\mathbf{h}}_\ell)^\dagger(\mathbf{h}_\ell-\hat{\mathbf{h}}_\ell)\right\} \\ &= \frac{1}{MN_gL}\sum_{\ell=0}^{L-1} E_{\mathbf{c}_\ell,\hat{\mathbf{c}}_\ell}\left\{(\mathbf{c}_\ell-\hat{\mathbf{c}}_\ell)^\dagger \underbrace{\Psi^\dagger\Psi}_{\mathbf{I}_Q}(\mathbf{c}_\ell-\hat{\mathbf{c}}_\ell)\right\} \\ &= \frac{1}{MN_gL} E_{\mathbf{c},\hat{\mathbf{c}}}\left\{(\mathbf{c}-\hat{\mathbf{c}})^\dagger(\mathbf{c}-\hat{\mathbf{c}})\right\} \\ &= \frac{1}{MN_gL}\,\text{tr}\left[E_{\mathbf{c},\mathbf{y}^{(tr)}}\left\{(\mathbf{c}-\hat{\mathbf{c}})(\mathbf{c}-\hat{\mathbf{c}})^\dagger\right\}\right], \end{aligned} \tag{81}$$

where $E_{\mathbf{c},\mathbf{y}^{(tr)}}\{(\mathbf{c}-\hat{\mathbf{c}})(\mathbf{c}-\hat{\mathbf{c}})^\dagger\}$ is the estimation error covariance matrix for BEM coefficient vector $\mathbf{c}$. Recalling that $E_{\mathbf{c},\mathbf{y}^{(tr)}}\{(\mathbf{c}-\hat{\mathbf{c}})\hat{\mathbf{c}}^\dagger\} = 0$ since $E_{\mathbf{c}|\mathbf{y}^{(tr)}}\{\mathbf{c}\} = \hat{\mathbf{c}}$ and substituting (49) into (81), the estimation error covariance matrix is derived as

$$\begin{aligned}E_{\mathbf{c},\mathbf{y}^{(tr)}}\left\{(\mathbf{c}-\hat{\mathbf{c}})(\mathbf{c}-\hat{\mathbf{c}})^\dagger\right\} &= E_{\mathbf{c},\mathbf{y}^{(tr)}}\left\{(\mathbf{c}-\hat{\mathbf{c}})\mathbf{c}^\dagger\right\} \\ &= \mathbf{R}_{\mathbf{c}} - \Gamma_{\mathbf{c}}^\dagger E_{\mathbf{c},\mathbf{y}^{(tr)}}\left\{\mathbf{y}^{(tr)}\mathbf{c}^\dagger\right\} \\ &= \mathbf{R}_{\mathbf{c}} - \Gamma_{\mathbf{c}}^\dagger \mathbf{Z}\mathbf{R}_{\mathbf{c}}\,, \end{aligned} \tag{82}$$

where $\Gamma_{\mathbf{c}}^\dagger$ can be replaced by one of the expressions in (50) and (51), and consequently applying the matrix inversion lemma, we obtain the estimation error covariance of the BEM coefficient vector $\mathbf{c}$ as follows

$$E_{\mathbf{c},\mathbf{y}^{(tr)}}\left\{(\mathbf{c}-\hat{\mathbf{c}})(\mathbf{c}-\hat{\mathbf{c}})^\dagger\right\} = \left(\frac{1}{\sigma_w^2}\mathbf{Z}^\dagger\mathbf{Z} + \mathbf{R}_{\mathbf{c}}^{-1}\right)^{-1}. \tag{83}$$

Eventually, using (83) in (81), average MSE for the estimation of multipath channels is given by

$$\text{MSE}_{est} = \frac{1}{MN_gL}\,\text{tr}\left[\left(\frac{1}{\sigma_w^2}\mathbf{Z}^\dagger\mathbf{Z} + \mathbf{R}_{\mathbf{c}}^{-1}\right)^{-1}\right]. \tag{84}$$

C. Calculation of Overall MSE

Recalling (34) and (35), the total error between true and the estimate values of the ℓth multipath channel vector is defined as follows

$$\begin{aligned}\mathbf{h}_\ell - \hat{\mathbf{h}}_\ell &= (\mathbf{h}_\ell - \Psi\mathbf{c}_\ell) + \left(\Psi\mathbf{c}_\ell - \hat{\mathbf{h}}_\ell\right) \\ &= (\mathbf{h}_\ell - \Psi\mathbf{c}_\ell) + \Psi\left(\mathbf{c}_\ell - \hat{\mathbf{c}}_\ell\right) \\ &= \underbrace{(\mathbf{I}_{MN_g} - \Psi\Psi^\dagger)\mathbf{h}_\ell}_{\text{modeling error}} + \underbrace{\Psi\left(\mathbf{c}_\ell - \hat{\mathbf{c}}_\ell\right)}_{\text{estimation error}}. \end{aligned} \tag{85}$$

Using the overall error definition in (85) and noting that $\Psi^\dagger(\mathbf{I}_{MN_g} - \Psi\Psi^\dagger) = 0$, it is straightforward that the overall MSE is obtained as the sum of *modeling* and *estimation* MSEs as follows

$$\begin{aligned}\text{MSE}_{all} &= \frac{1}{MN_gL} E_{\mathbf{h},\hat{\mathbf{h}}}\left\{(\mathbf{h}-\hat{\mathbf{h}})^\dagger(\mathbf{h}-\hat{\mathbf{h}})\right\} \\ &= \frac{1}{MN_gL}\sum_{\ell=0}^{L-1} E_{\mathbf{h}_\ell,\hat{\mathbf{h}}_\ell}\left\{(\mathbf{h}_\ell-\hat{\mathbf{h}}_\ell)^\dagger(\mathbf{h}_\ell-\hat{\mathbf{h}}_\ell)\right\} \\ &= \text{MSE}_{mod} + \text{MSE}_{est}\,, \end{aligned} \tag{86}$$

where the MSE expressions MSE_{mod} and MSE_{est} are obtained from (80) and (84), respectively.

D. Calculation of Overall MSE for Doppler Mismatch Case

In the Doppler mismatch case, we show the maximum (one-sided) true and mismatch values of the maximum Doppler frequencies by f_D and $\widetilde{f}_D$, respectively. At the receiver side, the number OFDM symbols per frame, number of Legendre polynomials in BEM and the autocorrelation matrix of the the BEM coefficient vector $\mathbf{c}$ are determined with respect to $\widetilde{f}_D$ and we denote these mismatch parameters by $\widetilde{M}$, $\widetilde{Q}$, and $\mathbf{R}_{\widetilde{\mathbf{c}}}$. In the maximum Doppler frequency mismatch scenario, we assume we have L multipath channels varying in discrete time interval $[0, \widetilde{M}N_g)$ due to Doppler frequencies having maximum value

f_D and estimate these channels using the mismatch parameters by $\widetilde{M}$, $\widetilde{Q}$, and $\mathbf{R}_{\widetilde{\mathbf{c}}}$ at the receiver side. Following the method given in Section IV-A, $\widetilde{M}$ and $\widetilde{Q}$ are determined according to $\widetilde{f}_D$. Moreover, we show the MSE and SER metrics associated with maximum Doppler frequency mismatch by $\widetilde{\text{MSE}}$ and $\widetilde{\text{SER}}$, respectively. In the Doppler mismatch scenario, since we estimate the mismatch channel $\widetilde{\mathbf{h}} = \widetilde{\mathbf{\Phi}}\widetilde{\mathbf{c}}$ instead of the true channel $\mathbf{h}$, the overall average MSE can be found as follows

$$\begin{aligned}
\widetilde{\text{MSE}}_{all} &= \frac{1}{\widetilde{M}N_gL} E_{\mathbf{h},\mathbf{y}^{(tr)}}\left\{(\mathbf{h}-\widetilde{\mathbf{\Phi}}\widehat{\widetilde{\mathbf{c}}})^\dagger(\mathbf{h}-\widetilde{\mathbf{\Phi}}\widehat{\widetilde{\mathbf{c}}})\right\} \\
&= \frac{1}{\widetilde{M}N_gL}\text{tr}\left[E_{\mathbf{h},\mathbf{y}^{(tr)}}\left\{(\mathbf{h}-\widetilde{\mathbf{\Phi}}\widehat{\widetilde{\mathbf{c}}})(\mathbf{h}-\widetilde{\mathbf{\Phi}}\widehat{\widetilde{\mathbf{c}}})^\dagger\right\}\right] \\
&= \frac{1}{\widetilde{M}N_gL}\text{tr}\left[E_{\mathbf{h},\mathbf{y}^{(tr)}}\left\{(\mathbf{h}-\widetilde{\mathbf{\Phi}}\Gamma_{\widetilde{\mathbf{c}}}^\dagger\mathbf{y}^{(tr)})(\mathbf{h}-\widetilde{\mathbf{\Phi}}\Gamma_{\widetilde{\mathbf{c}}}^\dagger\mathbf{y}^{(tr)})^\dagger\right\}\right] \\
&= \frac{1}{\widetilde{M}N_gL}\text{tr}\left[\mathbf{R}_{\mathbf{h}}\right] + \frac{1}{\widetilde{M}N_gL}\text{tr}\left[\Gamma_{\widetilde{\mathbf{c}}}^\dagger\left(\mathbf{Z}\mathbf{R}_{\mathbf{c}}\mathbf{Z}^\dagger+\sigma_w^2\mathbf{I}\right)\Gamma_{\widetilde{\mathbf{c}}}\right] \\
&- \frac{2}{\widetilde{M}N_gL}\Re\left\{\text{tr}\left[\Gamma_{\widetilde{\mathbf{c}}}^\dagger\mathbf{Z}\mathbf{\Phi}^T\mathbf{R}_{\mathbf{h}}\widetilde{\mathbf{\Phi}}\right]\right\} \\
&= 1 + \frac{1}{\widetilde{M}N_gL}\text{tr}\left[\Gamma_{\widetilde{\mathbf{c}}}^\dagger\left(\mathbf{Z}\mathbf{R}_{\mathbf{c}}\mathbf{Z}^\dagger+\sigma_w^2\mathbf{I}\right)\Gamma_{\widetilde{\mathbf{c}}}\right] \\
&- \frac{2}{\widetilde{M}N_gL}\Re\left\{\text{tr}\left[\Gamma_{\widetilde{\mathbf{c}}}^\dagger\mathbf{Z}\mathbf{\Phi}^T\mathbf{R}_{\mathbf{h}}\widetilde{\mathbf{\Phi}}\right]\right\} \qquad (87)
\end{aligned}$$

where, $\Re\{\cdot\}$ denotes the real part of complex valued parameter, and assuming mistakenly $\widetilde{\mathbf{Z}} \in \mathbb{C}^{\widetilde{M}(N^{(tr)}-L+1)\times\widetilde{Q}L}$ and $\widetilde{\mathbf{c}} \in \mathbb{C}^{\widetilde{Q}L\times 1}$ in (47) instead of $\mathbf{Z}$ and $\mathbf{c}$, the MMSE estimation coefficient matrix for the Doppler mismatch case is determined as

$$\Gamma_{\widetilde{\mathbf{c}}}^\dagger = (\mathbf{R}_{\widetilde{\mathbf{c}}}\widetilde{\mathbf{Z}}^\dagger\widetilde{\mathbf{Z}}+\sigma_w^2\mathbf{I}_{\widetilde{Q}L})^{-1}\mathbf{R}_{\widetilde{\mathbf{c}}}\widetilde{\mathbf{Z}}^\dagger \in \mathbb{C}^{\widetilde{Q}L\times\widetilde{M}(N^{(tr)}-L+1)}. \qquad (88)$$

We can also list the matrices in (87) and (88) as follows

$$\begin{aligned}
\mathbf{R}_{\widetilde{\mathbf{c}}} &= \widetilde{\mathbf{\Phi}}^T\mathbf{R}_{\widetilde{\mathbf{h}}}\widetilde{\mathbf{\Phi}} \in \mathbb{C}^{\widetilde{Q}L\times\widetilde{Q}L} \\
\widetilde{\mathbf{\Phi}} &= \mathbf{I}_L\otimes\widetilde{\Psi} \in \mathbb{R}^{\widetilde{M}N_gL\times\widetilde{Q}L} \\
\mathbf{R}_{\widetilde{\mathbf{h}}} &= \Omega\otimes\widetilde{\mathbf{R}} \in \mathbb{C}^{\widetilde{M}N_gL\times\widetilde{M}N_gL} \\
\Omega &= \text{diag}\{\Omega_0,\Omega_1,\ldots,\Omega_{L-1}\} \in \mathbb{R}^{L\times L}, \qquad (89)
\end{aligned}$$

where, it is straightforward to obtain $\widetilde{\mathbf{Z}}$ by following the similar steps taken for $\mathbf{Z}$ in between (46) and (48), and tth row t'th column entry of the matrix $\widetilde{\mathbf{R}} \in \mathbb{C}^{\widetilde{M}N_g\times\widetilde{M}N_g}$ is obtained according to mismatch value of the maximum Doppler frequency $\widetilde{f}_D$ by $[\widetilde{\mathbf{R}}]_{t,t'} = J_0(2\pi\widetilde{f}_DT_s(t-t'))$ using (2).

E. Evaluation of Bayesian Crámer Rao Lower Bound

In this part of the appendix, we give the analytical expression of the Bayesian CRLB for the overal MSE. The error coveriance matrix of the BEM coefficient vector $\mathbf{c}$ is defined by

$$\begin{aligned}
E_{\mathbf{c},\hat{\mathbf{c}}}\left\{(\mathbf{c}-\hat{\mathbf{c}})(\mathbf{c}-\hat{\mathbf{c}})^\dagger\right\} &= E_{\mathbf{y}^{(tr)},\mathbf{c}}\left\{(\hat{\mathbf{c}}-\mathbf{c})(\hat{\mathbf{c}}-\mathbf{c})^\dagger\right\} \\
&= E_{\mathbf{c}}E_{\mathbf{y}^{(tr)}|\mathbf{c}}\left\{\left(\hat{\mathbf{c}}-E_{\mathbf{y}|\mathbf{c}}\{\hat{\mathbf{c}}\}\right)\left(\hat{\mathbf{c}}-E_{\mathbf{y}^{(tr)}|\mathbf{c}}\{\hat{\mathbf{c}}\}\right)^\dagger\right\} \\
&+ E_{\mathbf{c}}E_{\mathbf{y}^{(tr)}|\mathbf{c}}\left\{\left(E_{\mathbf{y}^{(tr)}|\mathbf{c}}\{\hat{\mathbf{c}}\}-\mathbf{c}\right)\left(E_{\mathbf{y}^{(tr)}|\mathbf{c}}\{\hat{\mathbf{c}}\}-\mathbf{c}\right)^\dagger\right\} \\
&= \Sigma_{\hat{\mathbf{c}}} + E_{\mathbf{c}}\left\{\mathbf{b}(\mathbf{c})\,\mathbf{b}^\dagger(\mathbf{c})\right\}, \qquad (90)
\end{aligned}$$

where $\Sigma_{\hat{\mathbf{c}}} = E_{\mathbf{y}^{(tr)},\mathbf{c}}\{(\hat{\mathbf{c}}-E_{\mathbf{y}^{(tr)}|\mathbf{c}}\{\hat{\mathbf{c}}\})(\hat{\mathbf{c}}-E_{\mathbf{y}^{(tr)}|\mathbf{c}}\{\hat{\mathbf{c}}\})^\dagger\}$ is the covariance matrix of the estimate vector $\hat{\mathbf{c}}$ and $\mathbf{b}(\mathbf{c})$ is the bias function of the vector $\mathbf{c}$ such that is $E_{\mathbf{y}^{(tr)}|\mathbf{c}}\{\hat{\mathbf{c}}\} = \mathbf{c}+\mathbf{b}(\mathbf{c})$.

The conventional CRLB derived for unbiased estimators is not directly applicable since the bias term in (90) is considered as zero by assuming $\mathbf{b}(\mathbf{c}) = 0_{QL\times 1}$. However, the bias term should be taken into account, but in general, the bias function is not known for any estimator. In [20], to find a feasible solution under these circumstances, the bias function is assumed as a linear function of $\mathbf{c}$. Consequently in [20], obtaining the lower bound of $\Sigma_{\hat{\mathbf{c}}}$ by taking into account the bias function, substituting it into (90) and solving the coefficients of the linear bias function so as to minimize the estimation error covariance in (90), the CRLB for a linearly biased Bayesian estimator of the expansion coefficients is derived as follows

$$E_{\mathbf{y}^{(tr)},\mathbf{c}}\left\{(\hat{\mathbf{c}}-\mathbf{c})(\hat{\mathbf{c}}-\mathbf{c})^\dagger\right\} \geq (\mathbf{J}+\mathbf{R}_{\mathbf{c}}^{-1})^{-1}. \qquad (91)$$

where $\mathbf{J}$ is the Bayesian Fisher information matrix (FIM) defined as

$$\begin{aligned}
\mathbf{J} &\triangleq E_{\mathbf{y}^{(tr)},\mathbf{c}}\left\{\frac{\partial\log p(\mathbf{y}^{(tr)},\mathbf{c})}{\partial\mathbf{c}^*}\;\frac{\partial\log p(\mathbf{y}^{(tr)},\mathbf{c})}{\partial\mathbf{c}^T}\right\} \\
&= -\;E_{\mathbf{y}^{(tr)},\mathbf{c}}\left\{\frac{\partial^2\log p(\mathbf{y}^{(tr)},\mathbf{c})}{\partial\mathbf{c}^*\,\partial\mathbf{c}^T}\right\}. \qquad (92)
\end{aligned}$$

$\mathbf{J}$ in (92) can be expressed as the sum of $\mathbf{J}^{\text{data}}$ and $\mathbf{J}^{\text{prior}}$ that are the FIMs evaluated from data and prior information, respectively:

$$\mathbf{J} = \mathbf{J}^{\text{data}} + \mathbf{J}^{\text{prior}}, \qquad (93)$$

where

$$\mathbf{J}^{\text{data}} = -E_{\mathbf{c}}\,E_{\mathbf{y}^{(tr)}|\mathbf{c}}\left\{\frac{\partial^2\log\,p\left(\mathbf{y}^{(tr)}\middle|\mathbf{c}\right)}{\partial\mathbf{c}^*\partial\mathbf{c}^T}\right\} \qquad (94)$$

and

$$\mathbf{J}^{\text{prior}} = -E_{\mathbf{c}}\left\{\frac{\partial^2\log\,p(\mathbf{c})}{\partial\mathbf{c}^*\partial\mathbf{c}^T}\right\}. \qquad (95)$$

Since $(\mathbf{y}^{(tr)}|\mathbf{c}) \sim \mathcal{CN}(\mathbf{Z}\mathbf{c},\sigma_w^2\mathbf{I})$ and $\mathbf{c} \sim \mathcal{CN}(0_{QL},\mathbf{R}_{\mathbf{c}})$, the FIMs $\mathbf{J}^{\text{data}}$ and $\mathbf{J}^{\text{prior}}$ are obtained as

$$\mathbf{J}^{\text{data}} = \frac{1}{\sigma_w^2}\mathbf{Z}^\dagger\mathbf{Z} \qquad (96)$$

and

$$\mathbf{J}^{\text{prior}} = \mathbf{R}_{\mathbf{c}}^{-1}, \qquad (97)$$

respectively. Eventually, substituting the CRLB in (91) into (81), and then using (93), (96) and (97), the CRLB bound for the average MSE for the Bayesian estimation of $\mathbf{h}$ is found as

$$\text{MSE}_{est} \geq \frac{1}{MN_gL}\text{tr}\left[\left(\frac{1}{\sigma_w^2}\mathbf{Z}^\dagger\mathbf{Z}+2\mathbf{R}_{\mathbf{c}}^{-1}\right)^{-1}\right]. \qquad (98)$$

Finally, substituting the CRLB in (98) into (86) and using (80) as well, the Bayesian CRLB for MSE_{all} can be given by

$$\mathrm{MSE}_{all} \geq \frac{1}{MN_gL}\,\mathrm{tr}\left[(\mathbf{I}_{MN_g} - \Psi\Psi^{\dagger})\mathbf{R}\right] + \frac{1}{MN_gL}\,\mathrm{tr}\left[\left(\frac{1}{\sigma_w^2}\mathbf{Z}^{\dagger}\mathbf{Z} + 2\mathbf{R}_{\mathbf{c}}^{-1}\right)^{-1}\right]. \quad (99)$$

REFERENCES

[1] Y. (G.) Li, L. J. Cimini Jr., and N. R. Sollenberger, "Robust channel estimation for OFDM systems with rapid dispersive fading channels," *IEEE Trans. Commun.*, vol. 46, no. 4, pp. 902–914, Jul. 1998.

[2] O. Edfors, M. Sandell, J. van de Beek, S. Wilson, and P. Borjesson, "OFDM channel estimation by singular value decomposition," *IEEE Trans. Commun.*, vol. 46, no. 7, pp. 931–939, Jul. 1998.

[3] Y. (G.) Li, L. J. Cimini Jr. and N. R. Sollenberger, "Pilot-symbol-aided channel estimation for OFDM in wireless systems," *IEEE Trans. Veh. Technol.*, vol. 49, no. 4, pp. 1207–1215, Jul. 2000.

[4] M. Morelli and U. Mengali, "A comparison of pilot-aided channel estimation methods for OFDM systems," *IEEE Trans. Signal Process.*, vol. 49, no. 12, pp. 3065–3073, Dec. 2001.

[5] M.-H. Ng and S.-W. Cheung, "Bandwidth-efficient pilot-symbol-aided technique," *Electron. Lett.*, vol. 34, no. 16, pp. 1548–1550, 1998.

[6] S. Coleri, M. Ergen, A. Puri, and A. Bahai, "Channel estimation techniques based on pilot arrangement in OFDM systems," *IEEE Trans. Broadcast.*, vol. 48, no. 3, pp. 223–229, Sep. 2002.

[7] H. Senol, H. A. Cirpan, E. Panayirci, and M. Cevik, "A low complexity time-domain MMSE channel estimator for space-time/frequency block coded OFDM systems," *EURASIP J. Appl. Signal Process.*, vol. 2006, no. 19, 2006, Art. no. 39026.

[8] H. A. Cirpan, E. Panayirci, and H. Dogan, "Non-data-aided channel estimation for OFDM systems with space-frequency transmit diversity," *IEEE Trans. Veh. Technol.*, vol. 55, no. 2, pp. 449–457, Mar. 2006.

[9] H. Dogan, H.A. Cirpan, and E. Panayirci, "An efficient joint channel estimation and decoding algorithm for turbo-coded space-time orthogonal frequency division multiplexing receivers," *IET Commun.*, vol. 2, no. 7, pp. 886–894, Aug. 2008.

[10] H. Dogan, H. A. Cirpan, and E. Panayirci, "Iterative channel estimation and decoding of turbo coded SFBC-OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 6, no. 8, pp. 3090–3101, Aug. 2007.

[11] A. P. Kannu and P. Schniter, "Design and analysis of MMSE pilot-aided cyclic-prefixed block transmissions for doubly selective channels," *IEEE Trans. Signal Process.*, vol. 56, no. 3, pp. 1148–1160, Mar. 2008.

[12] M. Almoneer, C. Rohde, K. Hassan, and W. H. Gerstacker, "Intercarrier interference-aware pilot-aided channel estimation in OFDM systems," *IEEE Trans. Broadcast.*, vol. 63, no. 3, pp. 449–462, Sep. 2017.

[13] Y.-S. Choi, P. J. Voltz, and F. A. Cassara, "On channel estimation and detection for multicarrier signals in fast and selective rayleigh fading channels," *IEEE Trans. Commun.*, vol. 49, no. 8, pp. 1375–1387, Aug. 2001.

[14] W. Song and J. Lim, "Pilot-symbol aided channel estimation for OFDM with fast fading channels," *IEEE Trans. Broadcast.*, vol. 49, no. 4, pp. 398–402, Dec. 2003.

[15] A. Gorokhov and J. P. Linnartz, "Iterative interference cancellation and channel estimation for mobile OFDM," *IEEE Trans. Wireless Commun.*, vol. 4, no. 1, pp. 238–245, Jan. 2005.

[16] Y. Mostofi and D. Cox, "ICI mitigation for pilot-aided OFDM mobile systems," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 765–774, Mar. 2005.

[17] X. Huang and H.-C. Wu, "Robust and efficient intercarrier interference mitigation for OFDM systems in time-varying fading channels," *IEEE Trans. Veh. Technol.*, vol. 56, no. 5, pp. 2517–2528, Sep. 2007.

[18] Z. Sheng, H. D. Tuan, H. H. Nguyen, and Y. Fang, "Pilot optimization for estimation of high-mobility OFDM channels," *IEEE Trans. Veh. Technol.*, vol. 66, no. 10, pp. 8795–8806, Oct. 2017.

[19] M. Basaran, H. Senol, S. Erkucuk, and H. Cirpan, "Channel estimation for TDS-OFDM systems in rapidly time-varying mobile channels," *IEEE Trans. Wireless Commun.*, vol. 17, no. 12, pp. 8123–8135, Dec. 2018.

[20] H. Senol, E. Panayirci, and H. V. Poor, "Non-data-aided joint channel estimation and equalization for OFDM systems in very rapidly varying mobile channels," *IEEE Trans. Signal Process.*, vol. 60, no. 8, pp. 4236–4253, Aug. 2012.

[21] S. B. Amor, S. Affes, and F. Bellili, "ML EM estimation of fast time-varying OFDM-type channels," in *Proc. IEE 15th Int. Wireless Commun. Mobile Comput. Conf.*, Tangier, Morocco, Jun. 24–28 2019, pp. 1–6.

[22] A. Gaston, W. Chriss, and E. Walker, "A multipath fading simulator for radio," *IEEE Trans. Veh. Technol.*, vol. 22, no. 4, pp. 241–244, Nov. 1973.

[23] K. A. D. Teo and S. Ohno, "Optimal MMSE finite parameter model for doubly-selective channels," in *Proc. IEEE Global Telecommun. Conf.*, St. Louis, MO, USA, Nov. 2, 2005, pp. 3503–3507.

[24] T. Zemen and C. F. Mecklenbrauker, "Time-variant channel estimation using discrete prolate spheroidal sequences," *IEEE Trans. Signal Process.*, vol. 53, no. 9, pp. 3597–3607, Sep. 2005.

[25] G. B. Giannakis and C. Tepedelenlioglu, "Basis expansion models and diversity techniques for blind identification and equalization of time varying channels," *Proc. IEEE*, vol. 86, no. 10, pp. 1969–1986, Oct. 1998.

[26] E. Panayirci, H. Senol and H. V. Poor, "Joint channel estimation, equalization and data detection for OFDM systems in the presence of very high mobility," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4225–4238, Aug. 2010.

[27] S. Tomasin, A. Gorokhov, H. Yang, and J.-P. Linnartz, "Iterative interference cancellation and channel estimation for mobile OFDM," *IEEE Trans. Wireless Commun.*, vol. 4, no. 1, pp. 238–245, Jan. 2005.

[28] H. Senol, "Joint channel estimation and symbol detection for OFDM systems in rapidly time-varying sparse multipath channels," *Wireless Pers. Commun.*, vol. 82, no. 3, pp. 1161–1178, Jun. 2015.

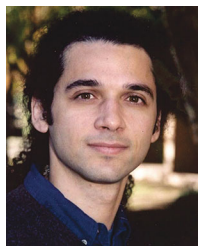
[29] H. Senol, X. Li, and C. Tepedelenlioglu, "Rapidly time-varying channel estimation for full-duplex amplify-and-forward one-way relay networks," *IEEE Trans. Signal Process.*, vol. 66, no. 11, pp. 3056–3069, Jun. 2018.

[30] M. Visintin, "Karhunen-loéve expansion of a fast rayleigh fading process," *IEEE Electron. Lett.*, vol. 32, no. 8, pp. 1712–1713, Aug. 1996.

[31] W. C. Jakes and D. C. Cox, *Microwave Mobile Communications*. New York, NY, USA: Wiley-IEEE, 1994.

[32] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. New York, NY, USA: Cambridge University Press, Sep. 2007.

Habib Şenol (Senior Member, IEEE) was born in Nazilli, Turkey, in 1971. He received the B.S. and M.S. degrees from Istanbul University, Istanbul, Turkey, in 1993 and 1999, respectively, both in electronics engineering. He received the Ph.D. degree in electronics engineering from Işık University, Istanbul, Turkey, in 2006. Dr. Şenol spent the academic year 2007–2008 with the Department of Electrical Engineering, Arizona State University, Tempe, USA, working on channel estimation and power optimization algorithms for Wireless Sensor Networks. He is currently an Associate Professor of Computer Engineering with Kadir Has University, Istanbul, Turkey. He was recently on sabbatical leave from Kadir Has University as a Visiting Research Professor with Arizona State University during 2018–2019 academic year. His research interests include statistical signal processing techniques and their applications to wireless electrical/underwater acoustic/optical communication systems, estimation and equalization algorithms for wireless communications, orthogonal frequency division multiplexing (OFDM), multi-carrier communications for 5G and beyond wireless networks, and statistical machine learning.



Cihan Tepedelenlioğlu (Senior Member, IEEE) was born in Ankara, Turkey, in 1973. He received the B.S. degree with highest honors from the Florida Institute of Technology in 1995, and the M.S. degree from the University of Virginia in 1998, both in electrical engineering. From January 1999 to May 2001, he was a Research Assistant with the University of Minnesota, where he completed the Ph.D. degree in electrical and computer engineering. He is currently an Associate Professor of Electrical Engineering with Arizona State University. He was awarded the NSF (early) Career grant in 2001, and has served as an Associate Editor for several IEEE TRANSACTIONS including IEEE TRANSACTIONS ON COMMUNICATIONS, IEEE SIGNAL PROCESSING LETTERS, and IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY. His research interests include statistical signal processing, system identification, wireless communications, estimation and equalization algorithms for wireless systems, multi-antenna communications, OFDM, ultra-wideband systems, distributed detection and estimation, and data mining for PV systems.