Nondata-Aided Joint Channel Estimation and Equalization for OFDM Systems in Very Rapidly Varying Mobile Channels

Habib Şenol, *Member, IEEE*, Erdal Panayırcı, *Fellow, IEEE*, and H. Vincent Poor, *Fellow, IEEE*

***Abstract*—This paper is concerned with the challenging and timely problem of joint channel estimation and equalization for orthogonal frequency division multiplexing (OFDM) systems in the presence of frequency selective and very rapidly time varying channels. The resulting algorithm is based on the space alternating generalized expectation maximization—maximum *a posteriori* probability (SAGE-MAP) technique which is particularly well suited to multicarrier signal formats. The algorithm is implemented in the time-domain which enables one to use the Gaussian approximation of the transmitted OFDM samples. Consequently, the averaging process of the nonpilot data symbols becomes analytically possible resulting in a feasible and computationally efficient channel estimation algorithm leading to a receiver structure that yields also an equalized output from which the data symbols are detected with excellent symbol error rate (SER) performance. Based on this Gaussian approximation the exact Bayesian Cramér Rao lower bound (CRLB) as well as the convergence rate of the algorithm are derived analytically. To reduce the computational complexity of the algorithm, discrete Legendre orthogonal basis functions are employed to represent the rapidly time-varying fading channel. It is shown that, depending on the normalized Doppler frequency, only a small number of expansion coefficients is sufficient to approximate the channel very well and there is no need to know the correlation function of the input signal. The computational complexity of the algorithm is shown to be $\sim O(NL)$ per detected data symbol and per SAGE-MAP algorithm cycle where N is the number of OFDM subcarriers and L is the number of multipath components.**

***Index Terms*—Basis expansion model (BEM), joint channel estimation and equalization, orthogonal frequency-division multiplexing (OFDM), rapidly varying wireless channels, SAGE-MAP algorithm.**

I. Introduction

INCREASING demand on data rates to support broadband high speed communication systems operating over frequency selective fading channels and with very high mobilities has resulted in research on designing fast, computationally efficient algorithms for channel estimation, equalization and data detection. Orthogonal frequency-division multiplexing (OFDM) is becoming a backbone structure of such systems, which are being standardized as the IEEE's 802.16 family—better known as Mobile Worldwide Interoperability Microwave Systems for Next-Generation Wireless Communication Systems (WiMAX)—and by the Third-Generation Partnership Project (3GPP) in the form of its Long-Term Evolution (LTE) project. Both systems employ orthogonal frequency division multiplexing/multiple access (OFDMA) as well as a new single-carrier frequency-division multiple access (SC-FDMA) format. To promote the IEEE 802.16 and LTE standards, recently, a high mobility feature has been introduced [IEEE 802.16 m, LTE Advanced (LTE-A)] to enable mobile broadband services at vehicular speeds beyond 120 km/h.

In fading channels with very high mobilities, the time variation of the channel over an OFDM symbol period results in a loss of subchannel orthogonality which leads to interchannel interference (ICI) due to power leakage among OFDM subcarriers. Consequently, estimation of the channel and channel parameters, as well as channel equalization become critical issues in the design of wireless systems operating under such conditions. Since mobility support is widely considered to be one of the key features in next generation wireless communication systems, OFDM transmission over very rapidly time varying multipath fading channels has been considered in a number of recent papers. They can be summarized as follows.

For a rapidly time-varying channel, the time-domain channel estimation method proposed in [1] is a potential candidate for the channel estimator, in order to mitigate the ICI. This technique estimates the fading channel by exploiting the time-varying nature of the channel as a provider of time diversity and reduces the computational complexity using the singular value decomposition (SVD) method. However, the linear minimum mean-square error (LMMSE) equalizer and a successive interference cancelation (SIC) scheme with optimal ordering proposed in [1] along with channel estimation demand very high computation, since the number of subcarriers is usually very large; thus it may not be feasible in practical systems. In [2], to handle rapid variation within an OFDM symbol, a pilot-based estimation scheme using channel interpolation was proposed. However, since the proposed estimation scheme has two estimation steps and requires large size matrix inversion, its computational complexity is very high. The authors of [2] also proposed a simplified scheme to reduce the complexity at the expense of substantial performance loss. A time-domain

Manuscript received July 21, 2011; revised January 09, 2012 and March 24, 2012; accepted April 04, 2012. Date of publication April 20, 2012; date of current version July 10, 2012. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Sofiene Affes. This work was conducted within the NEWCOM++ Network of Excellence in Wireless Communications and WIMAGIC Strep projects funded through the EC 7th Framework Programs, and by the U.S. Air Force Office of Scientific Research under MURI Grant FA9550-09-1-0643.

H. Şenol is with the Department of Computer Engineering, Faculty of Engineering and Natural Sciences, Kadir Has University, 34083, Istanbul, Turkey (e-mail: hsenol@khas.edu.tr).

E. Panayırcı is with the Department of Electrical and Electronics Engineering, Faculty of Engineering and Natural Sciences, Kadir Has University, 34083, Istanbul, Turkey (e-mail: eepanay@khas.edu.tr).

H. V. Poor is with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail poor@princeton.edu).

Digital Object Identifier 10.1109/TSP.2012.2195657

channel estimator was proposed in [3] which assumes that the channel impulse response (CIR) varies in a linear fashion within the symbol duration. However, this assumption no longer holds when the normalized Doppler frequency takes substantially higher values. In [4], two methods to mitigate ICI in an OFDM system with coherent channel estimation were proposed. Both methods employed a piece-wise linear approximation to estimate channel time-variations in each OFDM symbol. In [5], least-squares (LS) channel estimation has been proposed to minimize the squared differences between the received signal and estimated signal. However, the LS algorithm is not very suitable for the rapidly varying channel models. Also the inversion of the large dimensional square matrix turns out to be ill conditioned and limits the direct application of the LS estimator. In [6], a low-complexity equalizer is designed first, assuming the channel is banded. Then, employing this equalizer, a pilot-aided MMSE channel estimation scheme for a time-varying wide-sense stationary uncorrelated scatters channel model is proposed.

Recently, several channel estimation and equalization techniques have been proposed based on the basis expansion method (BEM) which can approximate the time-varying channel efficiently. Consequently, we only need to estimate a much smaller number of unknown channel coefficients than the total number of coefficients. The optimal BEM in terms of the mean square error is the so-called discrete Karhunen-Loeve BEM (DKL-BEM). However it is quite sensitive to variations in the channel statistics [7]. Therefore, as a compromise, several BEMs that are independent of the channel statistics have been derived. The most popular of them are the so-called discrete-prolate spheroidal BEM [8], and the complex-exponential BEM [9], [10]. Focusing on the estimation of channels for OFDM systems, the works [11] and [12] are concerned with channel estimation and equalization based on different BEM assumptions. In [13], the channel is modeled by combining linear and exponential basis functions and a two dimensional interpolation approach is used to track the flat fading channel in the frequency domain. A channel estimation technique based on a frequency-domain, pilot-aided modulation has been investigated in [14] where BEM is applied to approximate the channel. The LMMSE channel estimation algorithm has been proposed as well. In [15] the channel state information (CSI) is estimated iteratively exploiting previously estimated CSI in an iterative LMMSE algorithm. However, all of these channel estimation techniques suffer from an error floor because of ICI in the high Doppler spread scenario. Therefore in pilot-aided channel estimation approaches in OFDM systems, it is not clear how to place the pilots optimally since the pilot symbols are corrupted by the unknown data symbols. Taking this fact into account, many existing works view the frequency-domain channel matrix either as being diagonal, thus ignoring the ICI completely, or strictly banded as in [8], [16], and [17]. These approaches suffer from large estimation errors for channels having large Doppler spreads. On the other hand, the works in [1] and [3] view the frequency domain channel matrix as a full matrix, which reflects the true situation but generally requires the pilots to occupy the entire OFDM symbol. Along this line, an algorithm has been proposed in [18] for estimating the Rayleigh complex channel gains and detecting the data jointly for OFDM systems in fast fading channels, assuming the channel delays are known *a priori*. However the complexity of the algorithms was shown to be $O(N^3)$ where N is the number of OFDM subcarriers, which is quite high for WiMAX and LTE applications where N can be as large as 1024. Recently in [19], a joint data detection, channel estimation and equalization algorithm has been presented for OFDM systems operating over high mobility channels based on the SAGE technique. To compare our current work with that of [19], we can highlight the following differences: The main objective of the work [19] was the detection of data symbols rather than estimating the channel coefficients directly. Thus, the parameters of interest to be estimated in [19] are discrete-valued whereas in the present work, a continuous-valued parameter estimation problem is considered as the unknown data symbols, treated as nuisance parameters, are averaged out by the SAGE algorithm. Consequently, derivations and structures as well as the properties of the two algorithms are completely different. For example, the convergence of the SAGE algorithm for estimation of discrete parameters is not guaranteed. Therefore the initialization of the SAGE algorithm becomes very critical for the algorithm to converge. On the other hand, in the current work, while deriving the SAGE based continuous-valued channel estimate in the time-domain, the continuity of the multipath channels makes it possible to derive the convergence rate of the SAGE algorithm and the Cramér Rao lower bound (CRLB) analytically. Also, unlike that of [19], in this work we propose a discrete Legendre polynomial (DLP) basis expansion model (DLP-BEM) to express the rapidly time-varying wireless channels. It is proved that the performance of the DLP-BEM based channel estimation algorithm approaches very rapidly to the performance of the one employing the optimal BEM as the resolution increases. While the block-type pilot structure was used in our earlier work, in this paper we consider LTE and WiMAX types of more realistic pilot structures and scenarios.

In this paper, a new nondata-aided space alternating generalized expectation maximization—maximum *a posteriori* probability (SAGE-MAP) algorithm is proposed for joint channel estimation and equalization in OFDMA systems operating over frequency selective and highly mobile wireless channels. The main contributions of the work and advantages gained by using this proposed method are as follows:

- As opposed to many existing algorithms in the literature that are implemented in the frequency-domain, the algorithm proposed is implemented in the time-domain which enables one to use the Gaussian approximation of the transmitted OFDM samples resulting in a feasible and computationally efficient channel estimation and equalization algorithm with excellent symbol error rate (SER) performance.
- An important feature of the proposed technique is the capability to exploit the time-varying channel as a provider of time diversity. The resulting detection step makes good use of it. Consequently, the time-domain approach achieves performance superior to other techniques without increasing bandwidth or incorporating redundancy.
- The orthonormal DLP-BEM is employed to represent the rapidly time-varying fading channel. We show in

the paper that the mean-square error performance of DLP-BEM is closest to the optimal BEM, the so-called discrete Karhunen-Loeve BEM, and it is independent of the channel statistics.
- The exact Bayesian Cramér-Rao lower bound (CRLB) is derived for the estimator of the complex valued DLP-BEM channel coefficient vector regarding the Gaussian approximated transmit signal vector as a nuisance parameter. We propose a Gibbs sampling technique to compute the Bayesian Fisher information matrix (FIM) in a computationally feasible way.
- The convergence rate of the SAGE-MAP based channel estimator is derived analytically regarding again the transmit signal vector as a nuisance parameter.

II. System Model

A. Signal and Channel Models

We consider an OFDM system with N subcarriers. At the transmitter, K out of N subcarriers are actively employed to transmit data symbols and nothing is transmitted from the remaining $N-K$ carriers. The time-domain transmitted signal is denoted as

$$s(m,n)=\frac{1}{N}\sum_{k=0}^{K-1}d(m,k)e^{j2\pi n\frac{k}{N}} \tag{1}$$

where n and k are the discrete-time and the discrete-frequency indices during the mth OFDM symbol, respectively. $d(m,k)$ stands for the frequency domain data symbol transmitted at discrete time m over the kth OFDM subchannel. By the central limit theorem, the transmitted signal $s(m,n)$ can be modeled as a zero-mean complex Gaussian sequence provided K is sufficiently large. A cyclic prefix (CP) of length L_c is then added. We assume a time-varying multipath mobile radio channel with discrete-time impulse response $h(n,\ell),\ell=0,1,\ldots,L-1$ where L is the maximum channel length and that $L\le L_c$. At the receiver, after matched filtering, symbol-rate sampling and discarding the symbols falling in the cyclic prefix, the received signal at the input of the discrete Fourier transform (DFT) can be expressed as

$$y(mN_g+n)=\sum_{\ell=0}^{L-1}h(mN_g+n,\ell)s(m,n-\ell)+w(mN_g+n) \tag{2}$$

for $n=0,1,\ldots,N-1$ and $m=0,1,\ldots,M-1$, where M represents one OFDM frame length consisting of M consecutive OFDM symbols. $N_g\triangleq N+L_c$ and $w(\cdot)$ is zero-mean complex additive Gaussian noise with variance N_0. Note that when the normalized Doppler frequency is sufficiently small, the time varying-channel impulse response can be assumed to be constant over the duration of one OFDM symbol; that is $h(mN_g+n,\ell)\approx h(m,\ell)$ for $n=0,1,\ldots N-1$. Then it can be easily shown from (2) that the received signal at the output of DFT takes the known form

$$Y(m,k)=d(m,k)H(m,k)+W(m,k)\,,\quad k=0,1,\ldots,K-1$$

where $Y(m,k)$, $W(m,k)$, and $H(m,k)$ are the received signal, noise and channel coefficients, respectively, all represented in the frequency-domain, corresponding to the mth OFDM symbol and kth subchannel.

By collecting received signal samples in a vector, the above model can be expressed in vector form as follows:

$$\mathbf{y}(m)=\sum_{\ell=0}^{L-1}\operatorname{diag}\big(\mathbf{s}_\ell(m)\big)\mathbf{h}_\ell(m)+\mathbf{w}(m)\in\mathcal{C}^N \tag{3}$$

where

$$\begin{aligned}\mathbf{y}(m)&=[y(mN_g),y(mN_g+1),\ldots,y(mN_g+N-1)]^T\\&\in\mathcal{C}^N\\\mathbf{w}(m)&=[w(mN_g),w(mN_g+1),\ldots,w(mN_g+N-1)]^T\\&\in\mathcal{C}^N\end{aligned}$$

and $\mathbf{h}_\ell(m)=[h(mN_g,\ell),h(mN_g+1,\ell),\ldots,h(mN_g+N-1,\ell)]^T$, $\ell=0,1,\ldots,L-1$ represents L—path wide-sense stationary uncorrelated scattering (WSSUS) Rayleigh fading coefficients at the (mN_g+n)th discrete-times. Assuming the Jakes model, the autocorrelation function of the channel is

$$\begin{aligned}&E\{h(mN_g+n,\ell)h^*(m'N_g+n',\ell')\}\\&\quad=\sigma_\ell^2J_0\Big(2\pi f_DT_s((m-m')N_g\\&\qquad+(n-n'))\Big)\delta(\ell-\ell')\end{aligned} \tag{4}$$

where $\sigma_\ell^2,\ell=0,1,\ldots,L-1$, represents the normalized power of the lth path of the channel satisfying $\sum_\ell\sigma_\ell^2=1$. $J_0(\cdot)$ is the zeroth-order Bessel function of the first kind, f_D is the Doppler shift due to the vehicle motion and $\delta(\cdot)$ is the Kronecker delta. $T_s=\frac{T}{(N+L_c)}$, T being the OFDM symbol duration. $\mathbf{w}(m)$ is a complex white Gaussian noise vector with zero-mean and $E[\mathbf{w}(m)\mathbf{w}^\dagger(m)]=N_0\mathbf{I}_N$. Finally, $\mathbf{s}_\ell(m)$ in (3) is defined as

$$\begin{aligned}\mathbf{s}_\ell(m)&\triangleq[s(m,-\ell),s(m,-(\ell-1)),\ldots,s(m,N-(\ell+1))]^T\\&\in\mathcal{C}^N.\end{aligned} \tag{5}$$

Note that due to the cyclic prefix employed at the transmitter, $s(m,-\ell)=s(m,N-l)$ for $\ell=0,1,\ldots,L-1$. Consequently, $\mathbf{s}_\ell(m)=\operatorname{vshift}(\mathbf{s}(m),\ell)$, where $\operatorname{vshift}(\mathbf{s}(m),\ell)$ represents the ℓ-step circular shift operator for a column vector $\mathbf{s}(m)=[s(m,0),s(m,1),\ldots,s(m,N-1)]^T$. Defining

$$\begin{aligned}\mathbf{S}_\ell(m)&=\operatorname{diag}(\mathbf{s}_\ell(m)),\\\mathbf{S}(m)&\triangleq[\mathbf{S}_0(m),\mathbf{S}_1(m),\ldots,\mathbf{S}_{L-1}(m)]\in\mathcal{C}^{N\times LN}\end{aligned} \tag{6}$$

and

$$\mathbf{h}(m)\triangleq[\mathbf{h}_0^T(m),\mathbf{h}_1^T(m),\ldots,\mathbf{h}_{L-1}^T(m)]^T\in\mathcal{C}^{LN}$$

the received signal model in (3) can be expressed as

$$\mathbf{y}(m)=\mathbf{S}(m)\mathbf{h}(m)+\mathbf{w}(m). \tag{7}$$

B. Selection of Channel Basis Expansion Model (BEM)

The performance of the receiver depends critically on the estimate of the time-varying channel impulse response $\mathbf{h} = [\mathbf{h}^T(0), \mathbf{h}^T(1) \ldots, \mathbf{h}^T(M-1)]^T \in \mathcal{C}^{MNL}$ from the $MN(MN < MNL)$ dimensional received vector $\mathbf{y} = [\mathbf{y}^T(0), \mathbf{y}^T(1) \ldots, \mathbf{y}^T(M-1)]^T$. It seems the estimation of the $MNL \times 1$ channel vector $\mathbf{h}$ is impossible by means of $\mathbf{y}$ since there are more unknowns to be determined than known equations. However, the banded property of the channel matrix [20], [21] enables us to reduce the number of parameters needed for channel estimation and consequently reduce the computational complexity of the channel estimation step substantially.

We first apply a suitable BEM which describes the time variations of the discrete-time channel impulse response $h(mN_g + n, \ell)$ over a data block consisting of M OFDM symbols. We do not make any assumption regarding the amount of time-variation (equivalently, Doppler frequency) in the channel. For notational simplicity, let $t \triangleq mN_g + n$. Then

$$\begin{pmatrix} m = 0, 1, \ldots, M-1 \\ \text{and} \\ n = 0, 1, \ldots, N_g - 1 \end{pmatrix} \Leftrightarrow \quad t = 0, 1, \ldots, MN_g - 1.$$

For each channel path $\ell = 0, 1, \ldots, L-1$, the channel coefficients, $h(t, \ell)$, can be represented as weighted sums of $N_g M$ orthogonal basis functions $\{\psi_q(t)\}$ in the interval $[0, N_g M T_s]$

$$h(t, \ell) = \sum_{q=0}^{MN_g - 1} \psi_q(t) c(q, \ell), \quad t = 0, 1, \ldots, MN_g - 1 \quad (8)$$

where $\{c(q, \ell)\}$ are the expansion coefficients. As $h(\cdot, \ell)$ is essentially a lowpass process whose bandwidth is determined by the Doppler frequency, it can be well approximated by the weighted sum of a substantially fewer number D ($\ll MN_g$) of suitable orthonormal basis functions:

$$\widetilde{h}(t, \ell) = \sum_{q=0}^{D-1} \psi_q(t) c(q, \ell), \;\; t = 0, 1, \ldots, MN_g - 1. \quad (9)$$

Similarly, using the orthogonality property of the basis functions, the expansion coefficients can be evaluated by the inverse transformation as

$$c(q, \ell) = \sum_{t=0}^{MN_g - 1} \psi_q(t) h(t, \ell), \;\; q = 0, 1, \ldots, D-1. \quad (10)$$

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

$$\widetilde{\mathbf{h}}_\ell = \mathbf{\Psi} \mathbf{c}_\ell \quad (11)$$

$$\text{and} \quad \mathbf{c}_\ell = \mathbf{\Psi}^\dagger \mathbf{h}_\ell \quad (12)$$

where

$$\widetilde{\mathbf{h}}_\ell = [\widetilde{h}(0, \ell), \widetilde{h}(1, \ell), \ldots, \widetilde{h}(MN_g - 1, \ell)]^T \in \mathcal{C}^{MN_g}$$
$$\mathbf{c}_\ell = [c(0, \ell), c(1, \ell), \ldots, c(D-1, \ell)]^T \in \mathcal{C}^D$$

and the matrix $\mathbf{\Psi}$ contains the orthogonal basis vectors as

$$\mathbf{\Psi} = [\boldsymbol{\psi}(0), \boldsymbol{\psi}(1), \ldots, \boldsymbol{\psi}(MN_g - 1)]^T \in \mathcal{R}^{MN_g \times D} \quad (13)$$

with

$$\boldsymbol{\psi}(t) = [\psi_0(t), \psi_1(t), \ldots, \psi_{D-1}(t)]^T$$
$$t = 0, 1, \ldots, MN_g - 1.$$

The dimension D of the basis expansion satisfies $D_{lower} \leq D \leq MN_g$. The lower bound D_{lower} is given by $D_{lower} = \lceil 2(f_D)_{\max} TM + 1 \rceil$, where $(f_D)_{\max}$ is the maximum (one-sided) Doppler bandwidth defined by $(f_D)_{\max} = \frac{v_{\max} f_c}{c}$ with $v_{\max}$, f_c, and c is the maximum supported speed, the carrier frequency, and the speed of light, respectively, and T the OFDM symbol duration. By choosing D, we control the channel modeling mean square error (MSE)

$$\text{MSE}_{trun} \triangleq \frac{1}{L} \sum_{\ell=0}^{L-1} \text{MSE}_\ell. \quad (14)$$

where $\text{MSE}_\ell \triangleq \left(\frac{1}{MN_g}\right) E\{(\mathbf{h}_\ell - \widetilde{\mathbf{h}}_\ell)^\dagger (\mathbf{h}_\ell - \widetilde{\mathbf{h}}_\ell)\}$. From (11) and (12) using the relationship $\widetilde{\mathbf{h}}_\ell = \mathbf{\Psi}\mathbf{\Psi}^\dagger \mathbf{h}_\ell$, and taking into account that the path powers are normalized to unity, it can be easily shown that

$$\text{MSE}_{trun} = \frac{1}{MN_g L} \text{tr}\{(\mathbf{I}_{MN_g} - \mathbf{\Psi}\mathbf{\Psi}^\dagger)\mathbf{R}\} \quad (15)$$

where $\mathbf{R}$ represents the correlation matrix of the channel normalized to the path powers and for the WSSUS fading channel model its (t, t')th entry can be expressed from (4) as $\mathbf{R}[t, t'] = J_0(2\pi f_D T_s (t - t'))$.

Employing complex exponentials as basis functions that have a period equal to the length of the considered interval has been widely considered in the literature due to its algebraically simple implementation and the orthogonality among columns of the basis-expansion matrix. However, this induces a larger modeling error [12]. Similarly, the discrete cosine transformation (DCT) was considered as a BEM in [19]. However, since the orthogonal discrete cosine basis functions are perfectly bandlimited and lowpass, they are not well suited to represent rapidly time varying channels. Recently, polynomial BEMs [10], [22] have been also employed for channel estimation in which each channel coefficient is modeled as a linear combination of a set of polynomials. However, the modeling performance of this technique is rather sensitive to the Doppler spread. Note that the BEM examples mentioned so far do not require statistical channel knowledge. On the other hand, the so-called DKL-BEM is optimal in terms of the mean square channel modeling error and the the expansion coefficients are uncorrelated [23]. However, the implementation of the DKL-BEM based algorithm is computationally expensive and requires knowledge of the channel statistics. Also, it has been observed that if the assumed channel statistics deviate from the true ones, the DKL-BEM performs suboptimally. In our work, we make use of a BEM, based on the orthonormal DLP-BEM, to represent the time variations of the channel in an observation interval. As can be seen from Fig. 1, the DLP-BEM is well suited to represent the low-pass equivalent of the channel by

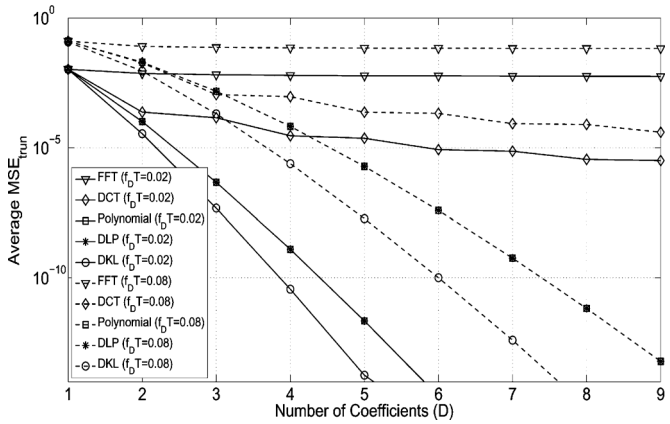


Fig. 1. Average truncation MSE versus number of coefficients $M = 7$.

means of a small number of basis functions. 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 Appendix A. The Legendre polynomials are generated by carrying out Gram-Schmidt orthogonalization on the polynomials $\{1, t, t^2, \ldots\}$ with respect to the time-varying channel coefficients in a neighborhood of the middle point of the considered interval. They can be defined as

$$\psi_q(t) = \frac{\nu_q(t)}{\xi_q} \tag{16}$$

where ξ_q denotes the normalization coefficient. The discrete orthogonal Legendre polynomials $\nu_q(t)$ and their corresponding normalization coefficients can be computed recursively as

$$\begin{aligned}\nu_q(t) &= \frac{(2q-1)(MN_g-1-2t)}{q(MN_g-q)}\nu_{q-1}(t) \\ &\quad - \frac{(q-1)(MN_g+q-1)}{q(MN_g-q)}\nu_{q-2}(t) \\ \xi_q &= \sqrt{\frac{(2q-1)(MN_g+q)}{(2q+1)(MN_g-q)}}\xi_{q-1} \\ t &= 0, 1, \ldots, MN_g-1,\ q = 2, 3, \ldots, D-1\end{aligned} \tag{17}$$

with the following initial polynomials and coefficients:

$$\begin{aligned}\nu_0(t) &= 1,\ \nu_1(t) = 1 - \frac{2t}{MN_g-1} \\ \xi_0 &= \sqrt{MN_g},\ \xi_1 = \sqrt{\frac{MN_g(MN_g+1)}{3(MN_g-1)}}.\end{aligned} \tag{18}$$

After removing the cyclic prefix, it can be shown from (8) and (2) that the dimension of the channel vectors $\widetilde{\mathbf{h}}_\ell$, $\ell = 0, 1, \ldots, L-1$, in (11) reduces from MN_g to MN. Then it can be easily seen that the channel vector $\mathbf{h}_\ell(m)$ in (3) is related to $\widetilde{\mathbf{h}}_\ell$ in (11) as

$$\widetilde{\mathbf{h}}_\ell(m) = \widetilde{\mathbf{h}}_\ell[mN_g : mN_g + N - 1]$$

where we have used the notation $\mathbf{x}[i : j]$ to denote entry i to entry j of a vector $\mathbf{x}$. Finally from (11) it follows that:

$$\widetilde{\mathbf{h}}(m) = \mathbf{\Phi}(m)\mathbf{c} \tag{19}$$

where

$$\begin{aligned}\widetilde{\mathbf{h}}(m) &= [\widetilde{\mathbf{h}}_0^T(m), \widetilde{\mathbf{h}}_1^T(m), \ldots, \widetilde{\mathbf{h}}_{L-1}^T(m)]^T \in \mathcal{C}^{NL} \\ \mathbf{c} &= [\mathbf{c}_0^T, \mathbf{c}_1^T, \ldots, \mathbf{c}_{L-1}^T]^T \in \mathcal{C}^{DL}\end{aligned} \tag{20}$$

and

$$\begin{aligned}\mathbf{\Phi}(m) &= \mathbf{I}_L \otimes \mathbf{\Psi}(m), \\ \mathbf{\Psi}(m) &= [\boldsymbol{\psi}(mN_g), \boldsymbol{\psi}(mN_g+1), \ldots, \boldsymbol{\psi}(mN_g+N-1)]^T \\ &\quad \in \mathcal{R}^{N\times D}.\end{aligned}$$

Note that $\otimes$ denotes the Kronecker product. Finally, substituting (19) into (7), the received signal is expressed in terms of the reduced dimensional channel vector $\mathbf{c}$ as follows:

$$\mathbf{y}(m) = \mathbf{Z}(m)\mathbf{c} + \mathbf{w}(m) \tag{21}$$

where $\mathbf{Z}(m) \triangleq \mathbf{S}(m)\mathbf{\Phi}(m)$ and $\mathbf{S}(m)$ is defined in (6).

C. Pilot Symbol Selection

For channel estimation and especially for initialization of the SAGE-MAP algorithm as explained in Section III, insertion of pilot symbols is necessary. Although several pilot patterns are possible, in our simulations, we employ more practical pilot distributions as defined in 4th generation (4G) wireless mobile communication systems. We consider both downlink LTE and WiMAX pilot patterns as shown in Fig. 2(a). Note that the number of pilots in LTE is much smaller than in other existing systems, including WiMAX. In the LTE context, the resource blocks (RBs) (in both time and frequency dimensions) are associated to users based on subframes of two RBs as depicted in Fig. 2(a). A simple RB may consist of 6 or 7 OFDM symbols (depending on the CP length) and 12 subcarriers. As for reference signals, a sparse pilot distribution in both frequency and time dimensions has been specified. In particular, the first and the fifth OFDM symbols of each RB are defined to contain pilots, and in the frequency dimension, the subcarriers with indices that are multiples of 6 are used for carrying pilots. This pilot distribution is shown in Fig. 2(a) for an RB having 7 OFDM symbols. On the other hand in WiMAX systems, multiple subchannels (slots) can be allocated to each user. As seen in Fig. 2(b), each slot consists of 2 OFDM blocks such that each one is defined to contain pilots with a pilot spacing of 6.

As notation, we assume that in each time-slot $m = 0, 1, \ldots, M-1$, there are P_m frequency domain pilot symbols located at OFDM subcarriers indexed by $i_1(m), i_2(m), \ldots, i_{P_m}(m)$ where $i_p(m) \in \{0, 1, \ldots, K-1\}$. They can be arranged in a way so that pilot spacings can be adjusted with respect to the time variations of the channel. When the channel is slowly varying, the time-domain correlation decays at a slower rate and consequently pilot spacing can be chosen wider.

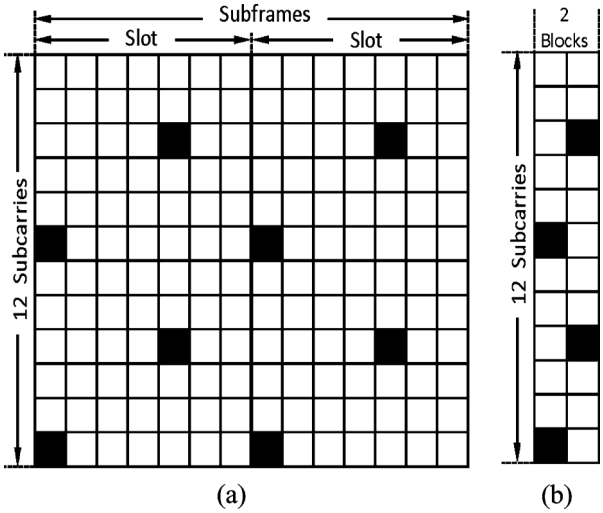


Fig. 2. Pilot Schemes of LTE and WiMAX OFDM Systems. (a) LTE. (b) Wimax.

III. Channel Estimation Using the Sage-Map Technique

The problem of interest is to derive an iterative algorithm based on the SAGE-MAP technique for channel estimation in a nondata-aided fashion (excepting the pilot symbols), employing the signal model given by (21). Since the SAGE method has been studied and applied to a number of problems in communications over the years, the details of the algorithm will not be presented in this paper. The reader is referred to [24] for a general exposition of the SAGE algorithm and to [25] for its application to an estimation problem related to the work herein. A suitable approach for applying the SAGE-MAP algorithm for estimation of the reduced dimensional channel vector $\mathbf{c}$, is to decompose the received signal in (21) into the sum [26]

$$\mathbf{y}(m) = \mathbf{x}_\ell(m) + \overline{\mathbf{x}}_\ell(m) \tag{22}$$

where

$$\mathbf{x}_\ell(m) = \mathbf{S}_\ell(m)\mathbf{\Psi}(m)\mathbf{c}_\ell + \mathbf{w}(m), \tag{23}$$

$$\text{and} \quad \overline{\mathbf{x}}_\ell(m) = \sum_{\ell'=0,\ell'\neq\ell}^{L-1} \mathbf{S}_{\ell'}(m)\mathbf{\Psi}(m)\mathbf{c}_{\ell'}. \tag{24}$$

We now derive the SAGE-MAP algorithm for estimating the channel vectors in the set $\mathbf{c} \triangleq \{\mathbf{c}_0, \mathbf{c}_1, \ldots, \mathbf{c}_{L-1}\}$, where $\mathbf{c}_\ell = [c(0,\ell), c(1,\ell), \ldots, c(D-1,\ell)]^T$, is the reduced dimensional channel vector of the ℓth path of the channel, based on the received vector $\mathbf{y} = [\mathbf{y}^T(0), \mathbf{y}^T(1), \ldots, \mathbf{y}^T(M-1)]^T$. To obtain a receiver architecture that iterates between soft-data and channel estimation, one might choose the parameter vector to be $\{\mathbf{c}\}$. At iteration (i), only the ℓth path channel vector $\{\mathbf{c}_\ell\}$ is updated, while the channel vectors of other channel paths $\overline{\mathbf{c}}_\ell = \mathbf{c}\backslash\mathbf{c}_\ell$ are kept fixed, where the notation "\" denotes the set exclusion operator. In the SAGE-MAP algorithm, we view the observed data $\mathbf{y}$ as the *incomplete* data and since $\mathbf{s} \triangleq \{\mathbf{s}(0), \mathbf{s}(1), \ldots, \mathbf{s}(M-1)\}$ is unknown, except the pilots, we incorporate $\mathbf{s}$ into the *admissible hidden* data set as $\chi_\ell = \{\mathbf{x}_\ell, \mathbf{s}\}$, where $\mathbf{x}_\ell \triangleq \{\mathbf{x}_\ell(0), \mathbf{x}_\ell(1), \ldots, \mathbf{x}_\ell(M-1)\}$, to which the incomplete data $\mathbf{y}$ are related through a possibly nondeterministic mapping [24]. In a first step of the SAGE algorithm called the *expectation step*, the algorithm computes an estimate of the log-likelihood function of the hidden data based on $\mathbf{y}$ and a current estimate of $\mathbf{c}$. Then, in a second step called the *maximization step*, $\mathbf{c}_\ell$ is updated as the value that maximizes this estimate of the log-likelihood function. The two steps are repeated in the subsequent iterations until convergence is achieved.

A. Expectation-Step (E-Step)

To perform *expectation step (E-Step)* of the SAGE-MAP algorithm, the conditional expectation is taken over $\boldsymbol{\chi}_\ell$ given the observation $\mathbf{y}$ and given that $\mathbf{c}$ equals its estimate calculated at ith iteration:

$$Q_\ell\big(\mathbf{c}_\ell|\mathbf{c}^{(i)}\big) = E\big\{\log p(\mathbf{x}_\ell|\mathbf{c}_\ell, \overline{\mathbf{c}}_\ell^{(i)}, \mathbf{s})\big|\mathbf{y}, \mathbf{c}^{(i)}\big\} + \log p(\mathbf{c}_\ell). \tag{25}$$

The prior probability density function (pdf) $p(\mathbf{c}_\ell)$ of $\mathbf{c}_\ell$ in (20) is chosen as $\mathbf{c}_\ell \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_{\mathbf{c}_\ell})$. Note that the covariance matrix of $\mathbf{c} = [\mathbf{c}_0^T, \mathbf{c}_1^T, \ldots, \mathbf{c}_{L-1}^T]^T$ can be determined from (12) and (4) as

$$\mathbf{\Sigma}_\mathbf{c} = \text{diag}\big(\mathbf{\Sigma}_{\mathbf{c}_0}, \mathbf{\Sigma}_{\mathbf{c}_1}, \ldots, \mathbf{\Sigma}_{\mathbf{c}_{L-1}}\big) \in \mathcal{C}^{LD\times LD}$$

where

$$\mathbf{\Sigma}_{\mathbf{c}_\ell} = \mathbf{\Psi}^\dagger \mathbf{\Sigma}_{\mathbf{h}_\ell} \mathbf{\Psi} \in \mathcal{C}^{D\times D}. \tag{26}$$

$\mathbf{\Sigma}_{\mathbf{h}_\ell}$, the covariance matrix of $\mathbf{h}_\ell$, can be obtained from (4) as

$$\mathbf{\Sigma}_{\mathbf{h}_\ell} = \sigma_\ell^2 \begin{bmatrix} r(0) & r(1) & \cdots & r(MN_g-1) \\ r(1) & r(0) & \cdots & r(MN_g-2) \\ \vdots & \vdots & \ddots & \vdots \\ r(MN_g-1) & r(MN_g-2) & \cdots & r(0) \end{bmatrix} \tag{27}$$

with $r(k) = J_0(2\pi f_D k T_s)$. However, since $D \ll MN_g$, it can be shown (see Appendix A) that for each channel path $\ell = 0, 1, \ldots, L-1$, the reduced dimensional channel coefficients $c(d,\ell)$, $d = 0, 1, \ldots, D-1$, become uncorrelated and consequently $\mathbf{\Sigma}_{\mathbf{c}_\ell}$, the covariance matrix of $\mathbf{c}_\ell$, turns out to be diagonal as

$$\mathbf{\Sigma}_{\mathbf{c}_\ell} = \sigma_\ell^2 \text{diag}\big(\lambda(0,\ell), \lambda(1,\ell), \ldots, \lambda(D-1,\ell)\big) \tag{28}$$

where $\lambda(d,\ell) = \mathcal{S}_h\big(\frac{\pi d}{N_g M}\big)$, $d = 0, 1, \ldots, D-1$,, and $\mathcal{S}_h(\cdot,\cdot)$ is the channel's scattering function defined by the Fourier transform of $r(k)$. For Jakes' Doppler profile the scattering function is given by $\mathcal{S}_h(\Omega,\ell) = \frac{2\sigma_\ell^2 T_s}{\sqrt{\Omega_{\text{dopp}}^2 - \Omega^2}}$ for $|\Omega| < \Omega_{\text{dopp}} < \pi$, [27], where $\Omega_{\text{dopp}} = \omega_{\text{dopp}} T_s$ is the discrete normalized angular Doppler frequency, and $\omega_{\text{dopp}} = 2\pi f_D T$. By neglecting the terms independent of $\mathbf{c}$, $\log p(\mathbf{x}_\ell|\mathbf{c}_\ell, \overline{\mathbf{c}}_\ell^{(i)}, \mathbf{s})$ can be calculated from (23) as

$$\log p(\mathbf{x}_\ell|\mathbf{c}_\ell, \overline{\mathbf{c}}_\ell^{(i)}, \mathbf{s}) \sim \frac{1}{N_0} \sum_{m=0}^{M-1} \Big(2\Re\big\{\mathbf{c}_\ell^\dagger \mathbf{\Psi}^\dagger(m) \mathbf{S}_\ell^\dagger(m) \mathbf{x}_\ell(m)\big\} - \mathbf{c}_\ell^\dagger \mathbf{\Psi}^\dagger(m) \mathbf{S}_\ell^\dagger(m) \mathbf{S}_\ell(m) \mathbf{\Psi}(m) \mathbf{c}_\ell\Big) \tag{29}$$

where $\Re\{\cdot\}$ denotes the real part of its argument. Inserting (29) into (25), we have for $Q_\ell(\mathbf{c}_\ell|\mathbf{c}^{(i)})$

$$Q_\ell(\mathbf{c}_\ell|\mathbf{c}^{(i)}) = -\mathbf{c}_\ell^\dagger \boldsymbol{\Sigma}_{\mathbf{c}_\ell}^{-1}\mathbf{c}_\ell + \frac{1}{N_0}\sum_{m=0}^{M-1}\Big(2\Re\Big\{\mathbf{c}_\ell^\dagger\boldsymbol{\Psi}^\dagger(m)\big[\mathbf{S}_\ell^\dagger(m)\mathbf{x}_\ell(m)\big]^{(i)}\Big\} - \mathbf{c}_\ell^\dagger\boldsymbol{\Psi}^\dagger(m)\big[\mathbf{S}_\ell^\dagger(m)\mathbf{S}_\ell(m)\big]^{(i)}\boldsymbol{\Psi}(m)\mathbf{c}_\ell\Big) \quad (30)$$

where

$$\big[\mathbf{S}_\ell^\dagger(m)\mathbf{x}_\ell(m)\big]^{(i)} \triangleq E\Big\{\mathbf{S}_\ell^\dagger(m)\mathbf{x}_\ell(m)\big|\mathbf{y},\mathbf{c}^{(i)}\Big\} \quad (31)$$

and

$$\big[\mathbf{S}_\ell^\dagger(m)\mathbf{S}_\ell(m)\big]^{(i)} \triangleq E\Big\{\mathbf{S}_\ell^\dagger(m)\mathbf{S}_\ell(m)\big|\mathbf{y},\mathbf{c}^{(i)}\Big\}. \quad (32)$$

Equation (31) can be calculated by applying the conditional expectation rule as

$$\big[\mathbf{S}_\ell^\dagger(m)\mathbf{x}_\ell(m)\big]^{(i)} = E\Big\{\mathbf{S}_\ell^\dagger(m)E\big\{\mathbf{x}_\ell(m)|\mathbf{s},\mathbf{y},\mathbf{c}^{(i)}\big\}\big|\mathbf{y},\mathbf{c}^{(i)}\Big\}. \quad (33)$$

The conditional distribution of $\mathbf{x}_\ell(m)$ given $\mathbf{y}$, $\mathbf{s}$ and $\mathbf{c}=\mathbf{c}^{(i)}$ is Gaussian with mean

$$E(\mathbf{x}_\ell(m)|\mathbf{y},\mathbf{s},\mathbf{c}^{(i)}) = \mathbf{S}_\ell(m)\boldsymbol{\Psi}(m)\mathbf{c}_\ell^{(i)} + \Big(\mathbf{y}(m) - \sum_{\ell'=0}^{L-1}\mathbf{S}_{\ell'}(m)\boldsymbol{\Psi}(m)\mathbf{c}_{\ell'}^{(i)}\Big) \quad (34)$$

where $\mathbf{c}^{(i)}$ is the estimated value of the channel $\mathbf{c}$ at the ith iteration step. Inserting (34) in (33), and subsequently substituting (33) and (32) in (30), we can rewrite (30) as

$$Q_\ell(\mathbf{c}_\ell|\mathbf{c}^{(i)}) = 2\Re\big\{\mathbf{c}_\ell^\dagger\mathbf{F}_\ell^{(i)}\big\} - \mathbf{c}_\ell^\dagger\mathbf{G}_\ell^{(i)}\mathbf{c}_\ell \quad (35)$$

where

$$\mathbf{F}_\ell^{(i)} = \frac{1}{N_0}\sum_{m=0}^{M-1}\boldsymbol{\Psi}^\dagger(m)\Big(\big[\mathbf{S}_\ell^\dagger(m)\big]^{(i)}\mathbf{y}(m) - \sum_{\ell'=0,\ell'\neq\ell}^{L-1}\big[\mathbf{S}_\ell^\dagger(m)\mathbf{S}_{\ell'}(m)\big]^{(i)}\boldsymbol{\Psi}(m)\mathbf{c}_{\ell'}^{(i)}\Big) \quad (36)$$

$$\mathbf{G}_\ell^{(i)} = \boldsymbol{\Sigma}_{\mathbf{c}_\ell}^{-1} + \frac{1}{N_0}\sum_{m=0}^{M-1}\boldsymbol{\Psi}^\dagger(m) \times [\mathbf{S}_\ell^\dagger(m)\mathbf{S}_\ell(m)]^{(i)}\boldsymbol{\Psi}(m). \quad (37)$$

After some algebra, we obtain the following:

$$\big[\mathbf{S}_\ell(m)\big]^{(i)} = E\big\{\mathbf{S}_\ell(m)\big|\mathbf{y},\mathbf{c}^{(i)}\big\} \in \mathcal{C}^{N\times LN} = \mathrm{diag}\Big(\mathrm{vshift}\big(\boldsymbol{\mu}_{\mathbf{s}}^{(i)}(m),\ell\big)\Big) \quad (38)$$

where $\boldsymbol{\mu}_{\mathbf{s}}^{(i)}(m) = E_{\mathbf{s}}\big[\mathbf{s}(m)\big|\mathbf{y},\mathbf{c}^{(i)}\big] \in \mathcal{C}^N$ is the posterior mean of $\mathbf{s}(m)$, the transmitted signal in the time-domain, given $\mathbf{y}$ and $\mathbf{c}^{(i)}$. $\boldsymbol{\mu}_{\mathbf{s}}^{(i)}(m)$ represents the soft-statistics generated by the SAGE-MAP algorithm at the ith iteration step. It can be easily shown that after taking the discrete Fourier transform of $\boldsymbol{\mu}_{\mathbf{s}}^{(i)}(m)$, an MMSE equalizer output is generated in the frequency-domain from which the transmitted data can be demodulated effectively.

In order to obtain $\boldsymbol{\mu}_{\mathbf{s}}^{(i)}(m)$, we consider the following alternative form of the observation equation in (7):

$$\mathbf{y}(m) = \boldsymbol{\Gamma}(m)\mathbf{s}(m) + \mathbf{w}(m). \quad (39)$$

It is straightforward to see that $\boldsymbol{\Gamma}(m)$ stands for the convolution matrix and can be expressed as

$$\boldsymbol{\Gamma}(m) \triangleq \sum_{\ell=0}^{L-1}\mathrm{mshift}\Big(\mathrm{diag}\big(\mathbf{h}_\ell(m)\big),0,-\ell\Big) \in \mathcal{C}^{N\times N} \quad (40)$$

where $\mathrm{mshift}(\mathbf{A},q,p)$ represents a row-wise q-step and column-wise p-step circular shift of matrix $\mathbf{A}$. Consequently, the posterior mean of $\mathbf{s}(m)$ given $\mathbf{c}^{(i)}$ at the ith step can be obtained by using the formula [28, 10.24] along with the matrix identity $\mathbf{A}(a\mathbf{I}+\mathbf{B}\mathbf{A})^{-1} = (a\mathbf{I}+\mathbf{A}\mathbf{B})^{-1}\mathbf{A}$ as follows:

$$\boldsymbol{\mu}_{\mathbf{s}}^{(i)}(m) = \mathbf{s}_P(m) + \frac{1}{N_0}\boldsymbol{\Xi}_{\mathbf{s}}^{(i)}(m){\boldsymbol{\Gamma}^{(i)}}^\dagger(m)\Big(\mathbf{y}(m) - \boldsymbol{\Gamma}^{(i)}(m)\mathbf{s}_P(m)\Big) \quad (41)$$

together with the following posterior covariance matrix of $\mathbf{s}(m)$:

$$\boldsymbol{\Xi}_{\mathbf{s}}^{(i)}(m) = \boldsymbol{\Sigma}_{\mathbf{s}}(m)\Big(\mathbf{I}_N + \frac{1}{N_0}{\boldsymbol{\Gamma}^{(i)}}^\dagger(m)\boldsymbol{\Gamma}^{(i)}(m)\boldsymbol{\Sigma}_{\mathbf{s}}(m)\Big)^{-1} \quad (42)$$

where $\boldsymbol{\Sigma}_{\mathbf{s}}(m) = E\Big\{\big(\mathbf{s}(m)-\mathbf{s}_P(m)\big)\big(\mathbf{s}(m)-\mathbf{s}_P(m)\big)^\dagger\Big\}$ and $\boldsymbol{\Gamma}^{(i)}(m)$ is given by (40). Note that, in (41), $\mathbf{s}_P(m) = \mathbb{F}^{-1}\mathbf{d}_P(m)$, $\mathbb{F}$ denotes the DFT matrix and $\mathbf{d}_P(m) = [d_P(m,0), d_P(m,1),\ldots,d_P(m,K-1)]^T$ is the frequency domain pilot symbol vector with the following entries:

$$d_P(m,k) = \begin{cases}\pi(m,k), & k\in\{i_1(m),i_2(m),\ldots,i_{P_m}(m)\}\\ 0, & \text{otherwise}\end{cases} \quad k=0,1,\ldots,K-1 \text{ and } m=0,1,\ldots,M-1 \quad (43)$$

where $\pi(m,i_p(m))$ is the pilot symbol at the pilot location $i_p(m)$, $p=1,2,\ldots,P_m$ and P_m is the number of pilots of the OFDM symbol transmitted at the mth time-slot. Subsequently, in (36), we obtain

$$\big[\mathbf{S}_\ell^\dagger(m)\mathbf{S}_{\ell'}(m)\big]^{(i)} = \mathrm{diag}\Big(\mathrm{dg}\Big(\mathrm{mshift}\big([\mathbf{s}(m)\mathbf{s}^\dagger(m)]^{(i)},\ell',\ell\big)\Big)\Big) \quad (44)$$

where the operator $\mathrm{dg}(\cdot)$ denotes the main diagonal vector of a matrix, and $[\mathbf{s}(m)\mathbf{s}^\dagger(m)]^{(i)}$ represents the posterior autocorrelation matrix of $\mathbf{s}(m)$ given $\mathbf{c}^{(i)}$. The latter quantity is given by

$$[\mathbf{s}(m)\mathbf{s}^\dagger(m)]^{(i)} = \boldsymbol{\mu}_{\mathbf{s}}^{(i)}(m){\boldsymbol{\mu}_{\mathbf{s}}^{(i)}}^\dagger(m) + \boldsymbol{\Xi}_{\mathbf{s}}^{(i)}(m). \quad (45)$$

B. Maximization-Step (M-Step)

In the M-step of the SAGE-MAP algorithm, the estimates of the channel coefficients are updated at the $(i+1)$th iteration according to

$$\begin{aligned}\mathbf{c}_\ell^{(i+1)} &= \arg\max_{\mathbf{c}_\ell} Q_\ell(\mathbf{c}_\ell|\mathbf{c}^{(i)}),\ \text{and}\\ \overline{\mathbf{c}}_\ell^{(i+1)} &= \overline{\mathbf{c}}_\ell^{(i)}\end{aligned} \tag{46}$$

where $Q_\ell(\mathbf{c}_\ell|\mathbf{c}^{(i)})$ is given by (35). Substituting (35) into the above equation yields the following for $\mathbf{c}_\ell^{(i+1)}$:

$$\begin{aligned}\mathbf{c}_\ell^{(i+1)} &= {\mathbf{G}_\ell^{(i)}}^{-1}\mathbf{F}_\ell^{(i)},\ \text{and}\\ \overline{\mathbf{c}}_\ell^{(i+1)} &= \overline{\mathbf{c}}_\ell^{(i)}.\end{aligned} \tag{47}$$

C. Initialization of Channel Coefficients

The initial channel estimate $\mathbf{c}^{(0)}$ can be determined with the aid of the pilot symbols which is computed as follows. From (21), the received vector corresponding to one of the OFDM frame of length M can be expressed as

$$\mathbf{y} = \mathbf{Z}\mathbf{c} + \mathbf{w} \tag{48}$$

where

$$\begin{aligned}\mathbf{y} &= [\mathbf{y}^T(0), \mathbf{y}^T(1), \ldots, \mathbf{y}^T(M-1)]^T\\ \mathbf{Z} &= [\mathbf{Z}^T(0), \mathbf{Z}^T(1), \ldots, \mathbf{Z}^T(M-1)]^T\end{aligned} \tag{49}$$

and

$$\begin{aligned}\mathbf{Z}(m) &= \mathbf{S}(m)\mathbf{\Phi}(m)\\ &= [\mathbf{S}_0(m)\mathbf{\Psi}(m), \mathbf{S}_1(m)\mathbf{\Psi}(m), \ldots, \mathbf{S}_{L-1}(m)\mathbf{\Psi}(m)].\end{aligned} \tag{50}$$

Note that, using the inverse Fourier transform, the following equality is obtained for the diagonal matrix $\mathbf{S}_\ell(m)$:

$$\begin{aligned}\mathbf{S}_\ell(m) &= \operatorname{diag}\Big(\operatorname{vshift}\big(\mathbf{s}(m), \ell\big)\Big)\\ &= \frac{1}{N}\sum_{k=0}^{K-1} d(m,k) e^{\frac{-j2\pi\ell k}{N}} \operatorname{diag}\big(\mathbb{F}_N^*(k)\big)\end{aligned} \tag{51}$$

where $\mathbb{F}_N(k)$ denotes the kth column of the DFT matrix. Using (51), it is straightforward to show that

$$\mathbf{Z}(m) = \sum_{k=0}^{K-1} d(m,k)\mathbf{U}_k(m) \tag{52}$$

with

$$\mathbf{U}_k(m) = \mathbb{F}_L^T(k) \otimes \Big(\big(\mathbf{1}_D^T \otimes \frac{1}{N}\mathbb{F}_N^*(k)\big) \odot \mathbf{\Psi}(m)\Big) \in \mathcal{C}^{N\times DL} \tag{53}$$

where $\mathbb{F}_L(k)$ represents the first L terms of the kth column of the DFT matrix $\mathbb{F}$, $\odot$ denotes the element by element product and $\mathbf{1}_D$ stands for the all-one column vector with length D. In (52), we consider $\mathbf{Z}(m) = \mathbf{Z}_P(m) + \mathbf{Z}_D(m)$, where $\mathbf{Z}_P(m) = \sum_{k\in\mathcal{I}_P}^{K-1} d(m,k)\mathbf{U}_k(m)$ and $\mathbf{Z}_D(m) = \sum_{k\in\mathcal{I}_D}^{K-1} d(m,k)\mathbf{U}_k(m)$ are the matrices obtained from pilot and data symbols, respectively. So, the initial value of the reduced dimensional channel vector $\mathbf{c}$ can be determined from the received signal model (48) by an LMMSE estimation technique as follows:

$$\mathbf{c}^{(0)} = \mathbf{\Sigma}_\mathbf{c}\mathbf{Z}_P^\dagger\big(\mathbf{Z}_P\mathbf{\Sigma}_\mathbf{c}\mathbf{Z}_P^\dagger + \mathbf{V}_D + N_0\mathbf{I}_{MN}\big)^{-1}\mathbf{y} \tag{54}$$

where $\mathbf{Z}_P = [\mathbf{Z}_P^T(0), \mathbf{Z}_P^T(1), \ldots, \mathbf{Z}_P^T(M-1)]^T \in \mathcal{C}^{MN\times DL}$, $\mathbf{V}_D \triangleq \operatorname{diag}\big\{\mathbf{V}_D(0), \mathbf{V}_D(1), \ldots, \mathbf{V}_D(M-1)\big\} \in \mathcal{C}^{MN\times MN}$, $\mathbf{V}_D(m) \triangleq \sum_{k\in\mathcal{I}_D(m)} \mathbf{U}_k(m)\mathbf{\Sigma}_\mathbf{c}\mathbf{U}_k^\dagger(m) \in \mathcal{C}^{N\times N}$. To simplify the matrix inversion above we apply the matrix inversion lemma above to obtain

$$\begin{aligned}\mathbf{c}^{(0)} = \Big(&\mathbf{Z}_P^\dagger\big(\mathbf{V}_D + N_0\mathbf{I}_{MN}\big)^{-1}\mathbf{Z}_P\\ &+\mathbf{\Sigma}_\mathbf{c}^{-1}\Big)^{-1}\mathbf{Z}_P^\dagger\big(\mathbf{V}_D + N_0\mathbf{I}_{MN}\big)^{-1}\mathbf{y}.\end{aligned} \tag{55}$$

Note that by this transformation, we need to take a matrix inversion of only size $DL \times DL$ rather than $MN \times MN$ since $MN \gg DL$ and $\big(\mathbf{V}_D + N_0\mathbf{I}_N\big)^{-1}$ is precomputed.

A complete block diagram of the SAGE-MAP algorithm including the initialization step is given in Fig. 3. Note that, as shown in Fig. 3, (47), (36)–(38) can be interpreted as joint channel estimation and equalization implemented in the time-domain, immediately following the analog-to-digital conversion and cyclic prefix deletion processes at the OFDM receiver. Consequently, we can think of (41) as the output of an LMMSE equalizer, generated at the ith iteration step of the SAGE-MAP algorithm. After the algorithm converges at some iteration step, say the Ith step, the original data is detected by conventional coherent detection in the frequency-domain by taking the DFT of the time-domain equalized signal vector; that is, DFT $\{\boldsymbol{\mu}_\mathbf{s}^{(I)}\}$.

IV. Performance Limits of the SAGE-MAP Channel Estimation Algorithm

A. Bayesian MSE Lower Bound

Since the SAGE-MAP channel estimation algorithm cannot remove all the ICI in data and pilot symbols, the residual ICI results in an estimation bias in estimating the channel coefficients. To serve as a benchmark, we now derive an overall Bayesian MSE bound for the channel estimator proposed in this paper. The results are very general and can be applied for any parameter estimation problem having any additive and multiplicative bias terms. The overall Bayesian MSE for the channel impulse response vector is defined as follows:

$$\mathrm{MSE}_{\mathrm{all}} \triangleq \frac{1}{MN_gL}E_{\mathbf{h},\hat{\mathbf{h}}}\big\{(\mathbf{h}-\hat{\mathbf{h}})^\dagger(\mathbf{h}-\hat{\mathbf{h}})\big\} \tag{56}$$

Theorem 1: The overall Bayesian MSE in (56), is bounded by

$$\begin{aligned}\mathrm{MSE}_{\mathrm{all}} \geq \frac{1}{MN_gL}\Big(&\operatorname{tr}\big\{(\mathbf{J} + \mathbf{\Sigma}_\mathbf{c}^{-1})^{-1}\big\}\\ &+\operatorname{tr}\big\{(\mathbf{I}_{MN_g} - \mathbf{\Psi}\mathbf{\Psi}^\dagger)\mathbf{R}\big\}\Big)\end{aligned} \tag{57}$$

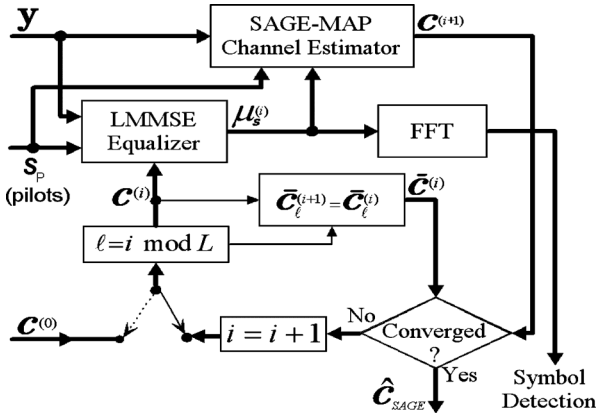


Fig. 3. Block diagram of the SAGE-MAP channel estimation algorithm.

where $\mathbf{\Psi}$, $\mathbf{R}$ and $\mathbf{\Sigma_c}$ are defined in (13), (15), and (26), respectively.

Proof: Proof of the Theorem 1 is given in Appendix B. ■

B. Convergence Rate of the SAGE-MAP Algorithm

In this section, we derive the convergence rate of our SAGE-MAP algorithm. Assuming that the algorithm is initialized in a region suitably close to a local maximum $\mathbf{c}^\star$ in the interior of the parameter space, the sequence of estimates converges monotonically in norm to it. Using the Taylor expansion in the neighborhood of $\mathbf{c}^\star$, the relationship between the estimates $\mathbf{c}^{(i)}$ and $\mathbf{c}^{(i+1)}$ can be expressed as

$$\mathbf{c}^{(i+1)} - \mathbf{c}^\star = \mathcal{M}(\mathbf{c}^{(i)} - \mathbf{c}^\star) \tag{58}$$

where $\mathcal{M}$ denotes the rate matrix of the SAGE-MAP algorithm, and is given by [26]

$$\mathcal{M} = \mathbf{I}_{DL} - (\mathcal{D}_{\bar{\mathcal{H}}} + \mathcal{D}_{\bar{\mathcal{I}}} + \mathcal{L}_{\bar{\mathcal{H}}})^{-1}\bar{\mathcal{H}} \tag{59}$$

$\mathcal{D}_{\bar{\mathcal{H}}}$ and $\mathcal{L}_{\bar{\mathcal{H}}}$ in (59) represent the block diagonal and strictly lower triangular parts of the average Hessian matrix $\bar{\mathcal{H}}$, respectively. The ℓth block of $\mathcal{D}_{\bar{\mathcal{H}}}$ corresponds to $\mathbf{c}_\ell$. On the other hand, $\mathcal{D}_{\bar{\mathcal{I}}}$ is also a block diagonal matrix with $\bar{\mathcal{I}}_\ell$ in the ℓth block. $\bar{\mathcal{I}}_\ell$ is the average value of the expected augmented information matrix, defined as follows [24]:

$$\bar{\mathcal{I}}_\ell = E_{\mathbf{y},\mathbf{c}}\{\mathcal{I}_\ell\} \tag{60}$$

where

$$\mathcal{I}_\ell \triangleq -E\Big\{\frac{\partial^2}{\partial \mathbf{c}_\ell^* \partial \mathbf{c}_\ell^T}\big(\log p(\chi_\ell|\mathbf{y},\mathbf{c}_\ell,\bar{\mathbf{c}}_\ell)\big)\Big|\mathbf{y},\mathbf{c}\Big\}. \tag{61}$$

Recalling that the *admissible hidden* data set $\chi_\ell = \{\mathbf{x}_\ell, \mathbf{s}\}$ and bearing in mind (35), it follows from (61) that:

$$\begin{aligned}\mathcal{I}_\ell &= \frac{\partial^2}{\partial \mathbf{c}_\ell^* \partial \mathbf{c}_\ell^T}\big(\log p(\mathbf{y},\mathbf{c}_\ell,\bar{\mathbf{c}}_\ell)\big) \\ &\quad - E\Big\{\frac{\partial^2}{\partial \mathbf{c}_\ell^* \partial \mathbf{c}_\ell^T}\big(\log p(\mathbf{c}_\ell|\chi_\ell,\bar{\mathbf{c}}_\ell)\big)\Big|\mathbf{y},\mathbf{c}\Big\} \\ &= \frac{\partial^2}{\partial \mathbf{c}_\ell^* \partial \mathbf{c}_\ell^T}\big(\log p(\mathbf{y},\mathbf{c}_\ell,\bar{\mathbf{c}}_\ell)\big) \\ &\quad - \Big(\frac{\partial^2}{\partial \mathbf{c}_\ell^* \partial \mathbf{c}_\ell^T} Q_\ell\big(\mathbf{c}_\ell|\mathbf{c}^{(i)}\big)\Big)\Big|_{\mathbf{c}^{(i)}=\mathbf{c}} \\ &= \frac{\partial^2}{\partial \mathbf{c}_\ell^* \partial \mathbf{c}_\ell^T}\big(\log p(\mathbf{y},\mathbf{c}_\ell,\bar{\mathbf{c}}_\ell)\big) + G_\ell^{(i)}\big|_{\mathbf{c}^{(i)}=\mathbf{c}}. \end{aligned} \tag{62}$$

Substituting (62) into (60), we obtain

$$\bar{\mathcal{I}}_\ell = -\mathbf{J}_\ell + \bar{G}_\ell \tag{63}$$

where $\bar{G}_\ell \triangleq E_{\mathbf{c}}\{G_\ell(\mathbf{c})\}$, $G_\ell(\mathbf{c}) \triangleq E_{\mathbf{y}|\mathbf{c}}\{G_\ell^{(i)}\big|_{\mathbf{c}^{(i)}=\mathbf{c}}\}$, and $\mathbf{J}_\ell$ is ℓth subblock matrix on the main diagonal of the Bayesian FIM $\mathbf{J}$ (see Appendix B). In calculating $G_\ell(\mathbf{c})$, we use (37), (44), (45), (42), and (41), respectively, and note that

$$\begin{aligned}\mathbf{\Xi_s}(m) &\triangleq E_{\mathbf{y}|\mathbf{c}}\Big\{\mathbf{\Xi}_\mathbf{s}^{(i)}(m)\big|_{\mathbf{c}^{(i)}=\mathbf{c}}\Big\} \\ &= \mathbf{\Sigma_s}(m)\Big(\mathbf{I}_N + \frac{1}{N_0}\mathbf{\Gamma}^\dagger(m)\mathbf{\Gamma}(m)\mathbf{\Sigma_s}(m)\Big)^{-1}\end{aligned} \tag{64}$$

and

$$\begin{aligned}E_{\mathbf{y}|\mathbf{c}}\Big\{\boldsymbol{\mu}_\mathbf{s}^{(i)}(m)\boldsymbol{\mu}_\mathbf{s}^{(i)^\dagger}(m)\big|_{\mathbf{c}^{(i)}=\mathbf{c}}\Big\} &= \mathbf{s}_P(m)\mathbf{s}_P{}^\dagger(m) \\ + \frac{1}{N_0^2}\mathbf{\Xi_s}(m)\mathbf{\Gamma}(m)^\dagger\mathbf{\Sigma}(m)\mathbf{\Gamma}(m)\mathbf{\Xi_s}{}^\dagger(m)\end{aligned} \tag{65}$$

where $\mathbf{\Sigma}(m) \triangleq \sum_{k\in\mathbf{I}_D}\mathbf{U}_k(m)\mathbf{c}\mathbf{c}^\dagger\mathbf{U}_k^\dagger(m) + N_0\mathbf{I}_N$ (see Appendix B). To compute the expectation $\bar{G}_\ell \triangleq E_{\mathbf{c}}\{G_\ell(\mathbf{c})\}$, we simply use the Gibbs sampling technique and generate samples $\mathbf{c}[1], \mathbf{c}[2], \ldots, \mathbf{c}[N_s]$ from the pdf $p(\mathbf{c}) = \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma_c})$. Then

$$\bar{G}_\ell \approx \frac{1}{N_s}\sum_{n=1}^{N_s} G_\ell\big(\mathbf{c}[n]\big). \tag{66}$$

In (59), the average Hessian matrix $\bar{\mathcal{H}}$ is defined as

$$\bar{\mathcal{H}} = E_{\mathbf{y},\mathbf{c}}\{\mathcal{H}\} \tag{67}$$

where $\mathcal{H}$ is the Hessian matrix and given as follows:

$$\begin{aligned}\mathcal{H} &= -\frac{\partial^2 \log p(\mathbf{y},\mathbf{c})}{\partial \mathbf{c}^* \partial \mathbf{c}^T} \\ &= -\frac{\partial^2 \log p(\mathbf{c})}{\partial \mathbf{c}^* \partial \mathbf{c}^T} - \frac{\partial^2 \log p(\mathbf{y}|\mathbf{c})}{\partial \mathbf{c}^* \partial \mathbf{c}^T}.\end{aligned} \tag{68}$$

Substituting (68) into (67), we obtain the average Hessian matrix as

$$\bar{\mathcal{H}} = \mathbf{J}. \tag{69}$$

The convergence rate is obtained by the spectral radius, that is the maximum magnitude eigenvalue, of the rate matrix $\mathcal{M}$ in (59). The convergence rate of the SAGE-MAP algorithm can be improved if one chooses a less informative hidden-data

TABLE I
COMPUTATIONAL COMPLEXITY DETAILS

Eq. No	Variable	Complexity	
		CMs	CAs
(54)	$\mathbf{c}^{(0)}$	$MNDL$	$MN(D-1)L$
(40) (using (19))	$\mathbf{\Gamma}^{(i)}(m)$	NDL^2	$N(D-1)L^2$
(42)	$\mathbf{\Xi}_{\mathbf{s}}^{(i)}(m)$	$2\Delta^2+\Delta(L+1)+L(L+1)/2$	$2\Delta^2-\Delta+1+L(L+1)/2$
(41)	$\boldsymbol{\mu}_{\mathbf{s}}^{(i)}(m)$	$N(\Delta+L)+\Delta(L+1)$	$N(\Delta+L)+L$
(45)	$\left[\mathbf{s}(m)\mathbf{s}^{\dagger}(m)\right]^{(i)}$	$N(N+1)/2$	$N(N+1)/2$
(38)	$\left[\mathbf{S}^{\dagger}(m)\right]^{(i)}$	0	0
(44)	$\left[\mathbf{S}_{\ell}^{\dagger}(m)\mathbf{S}_{\ell'}(m)\right]^{(i)}$	0	0
(36)	$\mathbf{F}_{\ell}^{(i)}$	$M\big[N(D+1)+$ $D(L-1)(D+N(D+1)/2)\big]$	$MD\big[N-1+$ $(L-1)(D-1+N(D+1)/2)\big]$
(37)	$\mathbf{G}_{\ell}^{(i)}$	$MND(D+1)/2$	$(M+1)ND(D+1)/2$
(47)	$\mathbf{c}_{\ell}^{(i+1)}$	$2D^2$	$D(2D-1)$

space. Less informative hidden-data spaces lead to shorter convergence time and hence a faster converging SAGE-MAP algorithm.

C. Complexity Analysis

The computational complexity of the algorithm is presented in Table I under the assumption that $K = N$. Note that, in the initialization step of the algorithm in (54), the term $\mathbf{\Sigma}_{\mathbf{c}}\mathbf{Z}_P^{\dagger}\left(\mathbf{Z}_P\mathbf{\Sigma}_{\mathbf{c}}\mathbf{Z}_P^{\dagger}+\mathbf{V}_D+N_0\mathbf{I}_{MN}\right)^{-1} \in \mathcal{C}^{DL\times MN}$ is a precomputed matrix. Therefore, the initialization step requires only a multiplication of this precomputed matrix with the vector $\mathbf{y} \in \mathcal{C}^{MN}$ resulting in $MNDL$ complex multiplications (CMs) and $MN(D-1)L$ complex additions (CAs). On the other hand, the covariance matrix $\mathbf{\Sigma}_{\mathbf{s}}(m)$, necessary for computation of (41), is a block matrix whose submatrices are diagonal with constant entries. Also, the convolution matrix $\mathbf{\Gamma}^{(i)}(m)$ in (41) and (42) is a sparse matrix whose columns have only L nonzero entries. Consequently, in the computation of $\boldsymbol{\mu}_{\mathbf{s}}^{(i)}(m)$ and $\mathbf{\Xi}_{\mathbf{s}}^{(i)}(m)$ in (41) and (42), the terms $\mathbf{\Gamma}^{(i)}(m)\mathbf{\Sigma}_{\mathbf{s}}(m)$, ${\mathbf{\Gamma}^{(i)}}^{\dagger}(m)\mathbf{\Gamma}^{(i)}(m)\mathbf{\Sigma}_{\mathbf{s}}(m)$, $\left(\mathbf{I}_N+\frac{1}{N_0}{\mathbf{\Gamma}^{(i)}}^{\dagger}(m)\mathbf{\Gamma}^{(i)}(m)\mathbf{\Sigma}_{\mathbf{s}}(m)\right)^{-1}$ and $\mathbf{\Xi}_{\mathbf{s}}^{(i)}(m){\mathbf{\Gamma}^{(i)}(m)}^{\dagger}$ are Hermitian Toeplitz matrices and can be approximated by block matrices whose submatrices are diagonal with constant entries resulting in a reduction in the complexity of the algorithm. By means of the above results and from Table I, it follows that the total computational complexity per detected symbol and per cycle, needed to estimate the channel coefficient vector $\mathbf{c}_{\ell}^{(i+1)}$ in (47) is approximately $\frac{L(N+1)}{2}+DL^2+DL+L(\Delta+L)+L^2D(D+1)$ CMs and $\frac{L(N+1)}{2}+2DL^2+DL+L(\Delta-1)+L^2\frac{D(D+1)}{2}$CAs $\sim \mathcal{O}(NL)$.

We now compare the computational complexity of our algorithm with some channel estimation/equalization algorithms which have been proposed in the literature recently. For fast time-varying channels, many existing works, including ours, resort to estimating the complex time-varying channel coefficients either in the time-domain or in the frequency domain, which are modeled by the BEM. In practice, these coefficients are estimated by inserting pilot tones and conventional methods generally consist of estimating the channel at pilot tones and then interpolating the channel frequency response. The estimation of the channel at pilots tones can be based on LMMSE, LS, best linear unbiased estimation (BLUE) or Kalman type of estimation techniques. To reduce the computational complexity of the algorithm, the banded structure of the channel matrix is taken into account resulting in fewer BEM coefficients to be estimated depending on the Doppler spread. The paper [29] proposes an iterative algorithm with Kalman filtering for estimating the time-varying, multipath Rayleigh channel gains and detecting data jointly for OFDM systems, assuming the knowledge of the delay-related information. It was reported in [29] that the total computational complexity of the proposed algorithm is $O(N^2)$ per detected symbols where N is the number of OFDM subcarriers. Note that this complexity is substantially greater than that of our algorithm (i.e., $O(N^2)$ versus $O(NL)$ with $L \ll N$). In [30], an iterative LS algorithm is presented for estimating the multipath complex channel gains, assuming again the knowledge of the delay-related information. The overall complexity of that algorithm was found to be $O(N_cN)$ where N_c denotes the reduced number of BEM coefficients determined by the Doppler spread. Although the complexity of this algorithms is almost the same as our algorithm, the LS estimator performs substantially worse when ICI is prominent, especially for high mobilities, and suffers from a large performance gap as compared to the CRLB. In [31], BEM-based channel estimation schemes are developed for high mobility OFDM uplink systems. Specifically, the authors expand the time varying channel into a small number of complex exponential basis functions spanning the given Doppler range and then formulate the LS and LMMSE algorithms to estimate those basis coefficients for some pilot patterns. The computational complexity of the LS algorithm and the LMMSE algorithms implemented in the time domain were found to be ρ^2 and ρ^2N CM per detected data symbol, respectively, where $\rho = D_{lower}L$ is the number of BEM coefficients to be estimated. In comparison with our algorithm, note that the algorithm in [31] is only concerned with the channel estimation and the complexity due to equalization is not taken into account. Consequently, the complexity of our algorithm which achieves channel estimation and equalization jointly, is much lower than that of those presented in [31]. In addition to this, note that the LMMSE-based channel estimator

TABLE II
SIMULATION PARAMETERS

Channel Bandwidth (BW)	$10\ MHz$
Number of Subcarriers (N)	1024
Number of Occupied Subcarriers (K)	720
Subcarrier Spacing	$15\ KHz$
Sampling Frequency (f_s)	$15.36\ MHz$
Carrier Frequency (f_c)	$2.5\ GHz$
Resource Block Length (M)	2 (WiMAX) , 7 (LTE)
Number of Multipaths (L)	3
Number of SAGE-MAP iterations (i_{max})	3×Number of Multipath
Modulation Formats	BPSK, QPSK, 16-QAM, 64-QAM

is suboptimal [12] due to the fact that it has the extra task to process the interference term which contains the information of the channel coefficients being estimated. Consequently, the SER performance of our SAGE-MAP based algorithm is always better than that of the LMMSE-based algorithms when operating over rapidly time varying channels.

V. SIMULATION RESULTS

In this section, we present computer simulation results to assess the performance of OFDM systems operating with the proposed channel estimation algorithm. Simulation parameters are chosen as in Table II. The initial estimate of the channel is performed by the reduced-complexity LMMSE estimation technique based on the pilot symbols.

In our simulations, we employ more practical pilot distributions as defined in 4th generation wireless mobile communication systems. We consider both downlink LTE and WiMAX pilot patterns as shown in Fig. 2.

The main motivation for choosing the practical pilot schemes in our computer simulations is to test the maximum SER and MSE performance capabilities of our channel estimation algorithm with respect to the channel mobility as well as the signaling format. The performance results presented shortly will give us some indications on how far we can increase the normalized Doppler frequency with suitably chosen modulation type and the pilot schemes in Fig. 2 to reach an outage limit of the system.

Figs. 4 and 5 show SER and MSE performance curves of the SAGE-MAP algorithm corresponding to the LTE and WiMAX pilot patterns with mobilities $f_DT = 0.02(v = 120\ \frac{\mathrm{km}}{\mathrm{h}})$ and $f_DT = 0.08(v = 480\ \frac{\mathrm{km}}{\mathrm{h}})$ and with binary phase shift keying (BPSK), quadrature phase shift keying (QPSK), 16 quadrature amplitude modulation (16-QAM) and 64-QAM signaling formats. The performance curves corresponding to perfect CSI are also included in Figs. 4 and 5 for comparison purposes. The multipath wireless channel having an exponentially decaying power delay profile with the normalized powers, $\sigma_0^2 = 0.448$, $\sigma_1^2 = 0.321$, and $\sigma_2^2 = 0.230$, is chosen. The number of Legendre polynomial coefficients, corresponding to the channel modeling $\mathrm{MSE} \approx 10^{-6}$, are chosen according to the Table III for LTE and WiMAX standards. We conclude from the curves in Figs. 4 and 5 that even when the number of Legendre coefficients is chosen to be fairly small as compared to the total number of coefficients, the performance loss in SER is not significant when CSI is not available. Our extensive computer simulations have shown that the SER and MSE performance of the algorithm employing the LTE pilot pattern is quite sensitive to the Doppler spread and degrades rapidly for mobilities beyond 120 km/h and for signaling formats beyond BPSK as compared to the WiMAX pilot pattern; this is mainly due to the fact that the number of pilots in LTE is much smaller and sparser than that for WiMAX. Note that as the speed increases, the rapidly varying channel not only destroys the orthogonality but also provides the receiver with time diversity. Especially when the channel is perfectly known (CSI is available) the system is able to make good use of the time diversity. Consequently, especially in Figs. 4(b) and 5(b), the SER performance for the case with $v = 480\ \frac{\mathrm{km}}{\mathrm{h}}$ is better than that of the case with $v = 120\ \frac{\mathrm{km}}{\mathrm{h}}$. However, when the CSI is not available, the gain from diversity may be overwhelmed by the channel estimation errors as well as by the complexity of the modulation format. This effect is clearly seen from the performance curves in Fig. 4(b) in which the SER performance of the SAGE-MAP algorithm for $v = 480\ \frac{\mathrm{km}}{\mathrm{h}}$ is slightly better than the case of $v = 120\ \frac{\mathrm{km}}{\mathrm{h}}$ when the CSI is obtained through channel estimation and a BPSK signaling format is employed. However, the time diversity vanishes as we move to higher dimensional modulations. A similar effect is also observed in the SER performance curves in Fig. 5(b), although it is less pronounced. In Figs. 4(a) and 5(a), the effects of channel estimation on the average overall MSE performance are investigated for different values of the Doppler spread and for the pilot structures shown in Fig. 2. The Bayesian MSE lower bounds computed from (57) are also given. We observe that the MSE performance is very close to the lower bound especially for small Doppler shifts and for WiMAX, which employs denser pilot patterns. However, the MSE performance gap between the MSE and the lower bound increases as the channel changes faster. As can be noticed from Fig. 5(a) this gap becomes more significant mainly because a sparser pilot pattern is employed in LTE systems.

In Fig. 6, we show that a maximum of three iterations are sufficient for the SAGE-MAP algorithm to converge.

We also study the average overall MSE versus the normalized Doppler spread (f_DT) for different SNR values for WiMAX pilot structure. From Fig. 7, it is observed that the MSE performance of the algorithm is quite robust against Doppler spread especially for lower SNR values. For SNR greater that 20 dB, we observe a gradual increase in MSE values as the normalized Doppler spread approaches 0.1, which is an extreme case in real applications.

In Figs. 8 and 9, the MSE and SER performance of our algorithm is compared to that of the DKL-BEM LMMSE algorithm for two different mobilities: $f_DT = 0.08\left(480\ \frac{\mathrm{km}}{\mathrm{h}}\right)$, $f_DT = 0.1\left(600\ \frac{\mathrm{km}}{\mathrm{h}}\right)$, and for the system employing the QPSK signaling format and the WiMAX pilot structure as shown in Fig. 2. The performance curves shown in Figs. 8 and 9 indicate that the SAGE-MAP algorithm clearly outperforms the LMMSE technique.

In Fig. 10, we investigate the MSE performance of our algorithm for the WiMAX-OFDM system in the presence of mismatch due to an overestimated Doppler frequency. We allow for mismatch by assuming $f_DT = 0.01$, 0.02 and 0.04 at the operational Doppler frequency $f_DT = 0.03$ chosen in the simulations. We note from Fig. 10 that the SAGE-MAP channel estimator, which is suboptimal due to Doppler frequency mismatch,

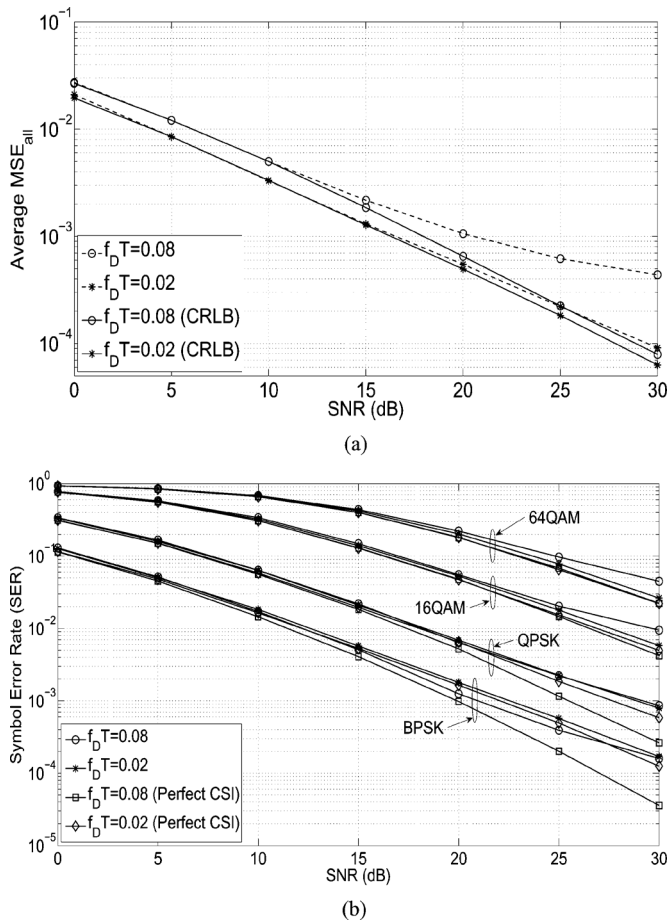


Fig. 4. SER and MSE performances of the SAGE-MAP algorithm for WiMAX Systems. (a) Average overall MSE versus SNR for WiMAX Systems. (b) SER versus SNR for WiMAX Systems.

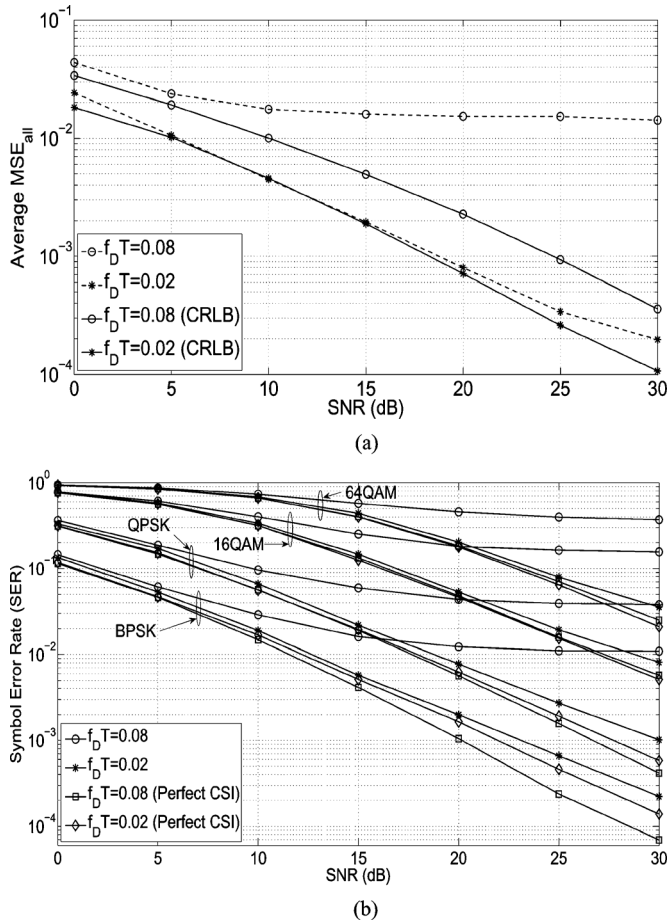


Fig. 5. SER and MSE performances of the SAGE-MAP algorithm for LTE Systems. (a) Average overall MSE versus SNR for LTE systems. (b) SER versus SNR for LTE Systems.

performs much better under the channels with $f_DT = 0.02$, 0.04 than $f_DT = 0.01$. Our computer simulations also indicate that underestimating the Doppler frequency is more harmful than overestimating it.

Fig. 11 shows the convergence rate of the SAGE-MAP algorithm versus the SNR, corresponding to the normalized Doppler shifts $f_DT = 0.02$ and $f_DT = 0.08$ for WiMAX and LTE pilot patterns. It can be seen that convergence rate increases monotonically and approaches a constant value, for SNR values greater than approximately 15 dB. A lower convergence rate is achieved at low SNR values and consequently the algorithm converges more often to a global maximum with fewer number of iterations. We also notice that the mobility does not affect the convergence rate significantly. However, depending on the pilot schemes employed during estimation, the convergence rate may change significantly as seen from Fig. 11.

Finally, note that, in a mobile network, although the performance depends on the location of the mobile terminal and distance from the base station, a general rule of thumb is that the SER values should be lower than 10^{-4} if a hybrid automatic repeat request (HARQ) is employed and 10^{-3} without HARQ, within the range of the acceptable operational SNR values. From the SER performance presented in Figs. 4(a) and 5(b), we can determine how much coding gain is needed and consequently what kind of coding technique should be chosen to reach an acceptable level of performance for different mobilities and modulation formats.

TABLE III
NUMBER OF DLP COEFFICIENTS FOR TRUNCATION MSE $\approx 10^{-6}$

Normalized Doppler Frequency	WiMAX	LTE
$f_DT = 0.02\,(v = 120\,km/h)$	D=2	D=3
$f_DT = 0.08\,(v = 480\,km/h)$	D=3	D=5

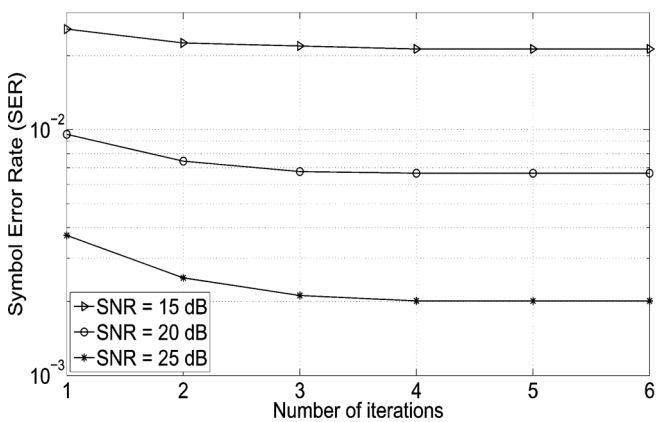


Fig. 6. SER versus Number of Iterations.

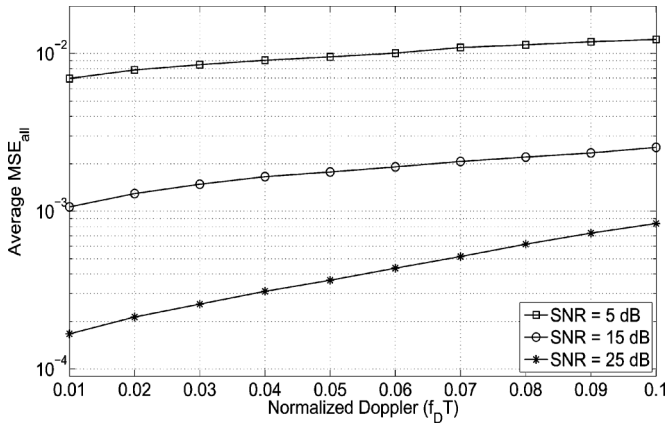


Fig. 7. Overall MSE versus normalized Doppler performance.

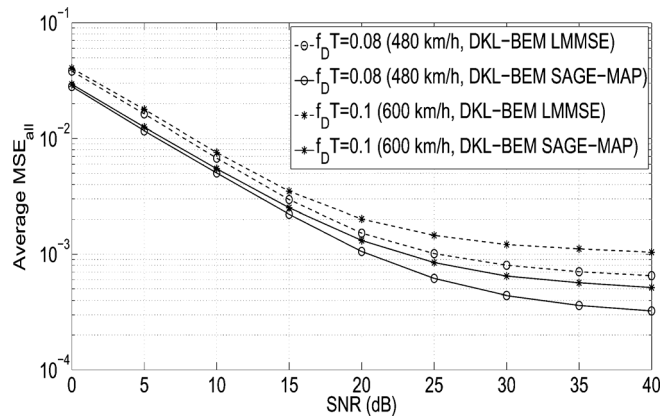


Fig. 8. Overall MSE performance comparisons of the DKL-BEM LMMSE and DKL-BEM SAGE-MAP channel estimation techniques.

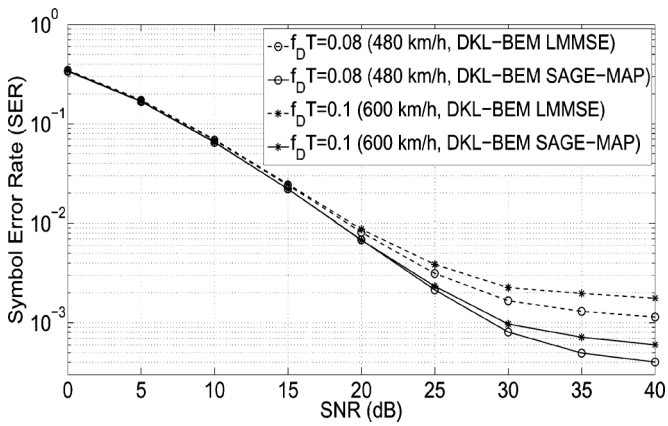


Fig. 9. SER performance comparisons of the DKL-BEM LMMSE and DKL-BEM SAGE-MAP channel estimation techniques.

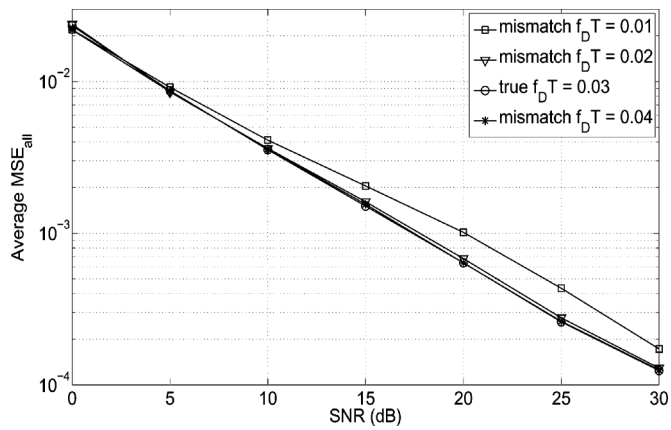


Fig. 10. Effects of Doppler mismatch on the overall MSE performance.

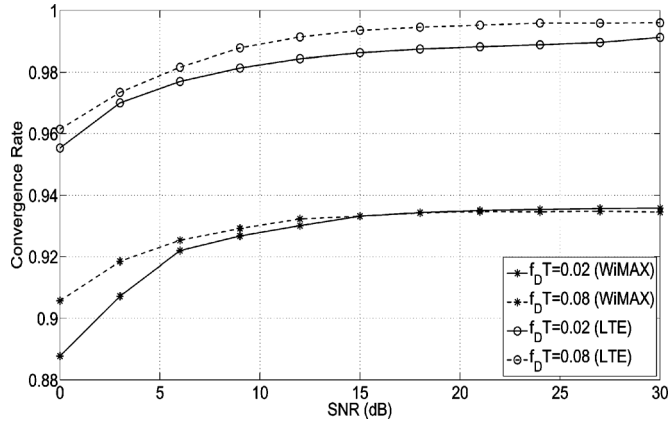


Fig. 11. Convergence rate versus SNR.

VI. Conclusion

In this paper, a new channel estimation algorithm has been proposed for OFDM systems operating over frequency selective and very rapidly time-varying channels, based on the SAGE-MAP technique which incorporates also channel equalization. This algorithm is implemented in the time-domain which enables one to use a Gaussian approximation for the transmitted OFDM samples. Consequently, the averaging process of the nonpilot data symbols in the expectation step of the algorithm becomes analytically possible resulting in a feasible and computationally efficient algorithm. To reduce the computational complexity of the algorithm, discrete Legendre orthogonal basis functions have been employed to represent the rapidly time-varying fading channel. It has been shown that, depending on the normalized Doppler frequency, only a small number of expansion coefficients is sufficient to approximate the channel perfectly and there is no need to know the correlation function of the input signal. Initial channel coefficients are effectively obtained by the pilot aided LMMSE estimator and unknown data symbols are averaged out in the algorithm.

The exact Bayesian CRLB, as well as the convergence rate of the SAGE-MAP based channel estimator have been derived while regarding the transmit signal vector as a nuisance parameter. This is analytically intractable in general. However, thanks to the Gaussian approximation assumed in the time-domain on the OFDM transmitted samples, it has been possible to derive those averaging processes analytically.

It has been shown via computer simulation that the proposed algorithm has excellent symbol error rate and channel estimation performance even with a very small number of channel expansion coefficients, resulting in reduction of the computational complexity to as low as $\sim O(NL)$ per detected data symbol and per SAGE-MAP algorithm cycle. Finally, computer simulations performed with OFDM based systems using the pilot structures

described in WiMAX and LTE standards also enabled us to determine the tradeoff between the SER performance of the overall system and the mobility of the channel.

APPENDIX A
PROOF OF ASYMPTOTIC CONVERGENCE OF THE CHANNEL COVARIANCE MATRIX

The correlation between $c(d,\ell)$ and $c(d',\ell)$ can be computed from (10) as

$$E\{c(d,\ell)c^*(d',\ell)\} = \sum_{t=0}^{MN_g-1} \psi_d(t) \sum_{t'=0}^{MN_g-1} E\{h(t,\ell)h^*(t',\ell)\}\psi_{d'}(t') \quad (70)$$

where $E\{h(t,\ell)h^*(t',\ell)\} = R_h(t-t',\ell)$.

We now show that for $d' = 0,1,\ldots,D-1$ and $D \ll MN_g$

$$\sum_{t'=0}^{MN_g-1} R_h(t-t',\ell)\psi_{d'}(t') \approx \lambda(d',\ell)\psi_{d'}(t). \quad (71)$$

Consequently, since the $\psi_q(\cdot)$'s are orthonormal, it follows from (70) that:

$$E\{c(d,\ell)c^*(d',\ell)\} = \begin{cases} \lambda(d,\ell), & \text{if } d = d' \\ 0, & \text{if } d \neq d' \end{cases}. \quad (72)$$

Let

$$g_d(t) \triangleq \sum_{t'=0}^{N_g M-1} R_h(t-t',\ell)\psi_{d'}(t'). \quad (73)$$

With $\mathcal{G}_d(\Omega)$, $\mathcal{S}_h(\Omega,\ell)$, and $\Psi_d(\Omega)$ denoting the DFTs of $g_d(t)$, $R_h(t,\ell)$, and $\psi_d(t)$, respectively, we have from (73)

$$\mathcal{G}_d(\Omega) \triangleq \mathcal{S}_h(\Omega,\ell)\Psi_d(\Omega) \quad (74)$$

where $\Omega = 2\pi f T_s$. Now consider the channel's scattering function which is the DFT of $R_h(t,\ell)$:

$$\mathcal{S}_h(\Omega,\ell) = \sum_{t=0}^{MN_g-1} R_h(t,\ell)e^{-j\Omega t}. \quad (75)$$

Expanding $e^{-j\Omega t}$ around the frequency Ω_d, at which $\Psi_d(\Omega)$ is peaked, and substituting the expansion in (75) we have

$$\mathcal{S}_h(\Omega,\ell) = \sum_{k=0}^{\infty} \mathfrak{m}_{k,d} \frac{(-j)^k(\Omega-\Omega_d)^k}{k!} \quad (76)$$

where

$$\mathfrak{m}_{k,d} \triangleq \sum_{t=0}^{MN_g-1} t^k R_h(t,\ell)e^{-j\Omega_d t}.$$

It can be easily shown from (76) that

$$\mathfrak{m}_{k,d} = (-j)^k \left.\frac{\partial^k \mathcal{S}_h(\Omega,\ell)}{\partial\Omega^k}\right|_{\Omega=\Omega_d}. \quad (77)$$

TABLE IV
THE VALUES OF $\eta_{k,0}$ AS A FUNCTION OF k

k	0	2	4	6
$\eta_{k,0}$	1	2.5×10^{-5}	5.6×10^{-9}	3.5×10^{-12}

Consequently, it follows from (74) that

$$\mathcal{G}_d(\Omega) \triangleq \mathfrak{m}_{0,d}\Psi_d(\Omega) + \sum_{k=1}^{\infty} \mathfrak{m}_{k,d}\Psi_d(\Omega)\frac{(-j)^k(\Omega-\Omega_d)^k}{k!}. \quad (78)$$

Finally, taking the inverse DFT (IDFT) of (78) we arrive at

$$g_d(t) \triangleq \mathfrak{m}_{0,d}\psi_d(t) + \sum_{k=1}^{\infty} \mathfrak{m}_{k,d}\ \mathrm{IDFT}\Big\{\frac{(-j)^k(\Omega-\Omega_d)^k}{k!}\Psi_d(\Omega)\Big\}. \quad (79)$$

Note that for WSSUS fading channels and for the Jakes' Doppler profile, the channel's scattering function for the ℓth path is given by

$$\mathcal{S}_h(\Omega,\ell) = \frac{2\sigma_\ell^2 T_s}{\sqrt{\Omega_{\mathrm{dopp}}^2 - \Omega^2}} \text{ for } |\,\Omega\,| < \Omega_{\mathrm{dopp}} < \pi \quad (80)$$

where Ω_{dopp} denotes the discrete normalized angular Doppler frequency. For $d = 0,1,\ldots,D-1 \ll MN_g$, the frequency spectra of the Legendre polynomials are concentrated in a narrow band $2\Delta\Omega$ around $\Omega_d = \frac{\pi d}{N_g M}$ and $\mathcal{S}_h(\Omega,\ell)$ is sufficiently smooth within this interval. Therefore, except for $\mathfrak{m}_{0,d}$, all other coefficients $\mathfrak{m}_{k,d}, k = 1,2,\ldots$ can be neglected. Let us explain this for the case $d = 0$ ($\Omega_d = 0$). By taking the kth derivative of (80) and replacing $\Omega = \Omega_d$ in the resulting equation, we have

$$\mathfrak{m}_{k,0} = \begin{cases} \zeta_k \left(\frac{T_s}{\Omega_{\mathrm{dopp}}}\right)^{k+1}, & \text{for } k = 0,2,\ldots \\ 0, & \text{for } k = 1,3,\ldots \end{cases} \quad (81)$$

where $\zeta_k \triangleq 2(2\pi)^k\ 1^2.3^2.\ldots.(k-1)^2$. As can be seen from Table IV, for a typical normalized Doppler angular frequency $\Omega_{\mathrm{dopp}} = 0.015$, the ratio $\eta_{k,0} = \frac{\mathfrak{m}_{k,0}}{\mathfrak{m}_{0,0}}$ vanishes very rapidly as k gets larger.

Consequently, the claim in (28) is supported.

APPENDIX B
PROOF OF THEOREM 1

In this appendix, a lower bound on the overall MSE, $\mathrm{MSE}_{\mathrm{all}}$, is derived. From (56), $\mathrm{MSE}_{\mathrm{all}}$ can be expressed as

$$\mathrm{MSE}_{\mathrm{all}} = \mathrm{MSE}_{\mathrm{trun}} + \mathrm{MSE}_{\mathrm{est}} + 2\Re\{\mathrm{MSE}_{\mathrm{cross}}\} \quad (82)$$

where

$$\mathrm{MSE}_{\mathrm{trun}} = \frac{1}{MN_gL}E_{\mathbf{h},\tilde{\mathbf{h}}}\{(\mathbf{h}-\tilde{\mathbf{h}})^\dagger(\mathbf{h}-\tilde{\mathbf{h}})\}$$
$$\mathrm{MSE}_{\mathrm{est}} = \frac{1}{MN_gL}E_{\tilde{\mathbf{h}},\hat{\tilde{\mathbf{h}}}}\{(\tilde{\mathbf{h}}-\hat{\tilde{\mathbf{h}}})^\dagger(\tilde{\mathbf{h}}-\hat{\tilde{\mathbf{h}}})\}$$
$$\mathrm{MSE}_{\mathrm{cross}} = \frac{1}{MN_gL}E_{\mathbf{h},\tilde{\mathbf{h}},\hat{\tilde{\mathbf{h}}}}\{(\mathbf{h}-\tilde{\mathbf{h}})^\dagger(\tilde{\mathbf{h}}-\hat{\tilde{\mathbf{h}}})\}. \quad (83)$$

$\mathrm{MSE}_{\mathrm{trun}}$, $\mathrm{MSE}_{\mathrm{est}}$, and $\mathrm{MSE}_{\mathrm{cross}}$, denote the truncation MSE, the estimation MSE and the cross correlation between truncation and estimation errors, respectively. Bearing in mind that $\tilde{\mathbf{h}} = \mathbf{\Phi}\mathbf{c}$, $\mathbf{c} = \mathbf{\Phi}^{\dagger}\mathbf{h}$, i.e $\tilde{\mathbf{h}} = \mathbf{\Phi}\mathbf{\Phi}^{\dagger}\mathbf{h}$, $\hat{\tilde{\mathbf{h}}} = \mathbf{\Phi}\hat{\mathbf{c}}$, and $(\mathbf{I}_{MN_gL} - \mathbf{\Phi}\mathbf{\Phi}^{\dagger})\mathbf{\Phi} = \mathbf{0}_{MN_gL\times DL}$, it can be shown from (83) that $\mathrm{MSE}_{\mathrm{cross}} = 0$. Now, noting that $\mathbf{c}_\ell = \mathbf{\Psi}^{\dagger}\mathbf{h}_\ell$ and $\tilde{\mathbf{h}}_\ell = \mathbf{\Psi}\mathbf{c}_\ell$, i.e $\tilde{\mathbf{h}}_\ell = \mathbf{\Psi}\mathbf{\Psi}^{\dagger}\mathbf{h}_\ell$, and after some algebra we have the following for $\mathrm{MSE}_{\mathrm{trun}}$:

$$\begin{aligned}\mathrm{MSE}_{\mathrm{trun}} &= \frac{1}{MN_gL}E_{\mathbf{h},\tilde{\mathbf{h}}}\{(\mathbf{h}-\tilde{\mathbf{h}})^{\dagger}(\mathbf{h}-\tilde{\mathbf{h}})\}\\ &= \frac{1}{MN_gL}\sum_{\ell=0}^{L-1}E_{\mathbf{h}_\ell,\tilde{\mathbf{h}}_\ell}\{(\mathbf{h}_\ell-\tilde{\mathbf{h}}_\ell)^{\dagger}(\mathbf{h}_\ell-\tilde{\mathbf{h}}_\ell)\}\\ &= \frac{1}{MN_gL}\mathrm{tr}\{(\mathbf{I}_{MN_g}-\mathbf{\Psi}\mathbf{\Psi}^{\dagger})\mathbf{R}\} \qquad (84)\end{aligned}$$

where $\mathbf{R}$ represents the correlation matrix of the channel normalized to the path powers and is given in (15). Finally, recalling that $\tilde{\mathbf{h}} = \mathbf{\Phi}\mathbf{c}$, $\hat{\tilde{\mathbf{h}}} = \mathbf{\Phi}\hat{\mathbf{c}}$, and $\mathbf{\Phi}^{\dagger}\mathbf{\Phi} = \mathbf{I}_{DL}$, we have

$$\begin{aligned}MSE_{\mathrm{est}} &\triangleq \frac{1}{MN_gL}E_{\tilde{\mathbf{h}},\hat{\tilde{\mathbf{h}}}}\{(\tilde{\mathbf{h}}-\hat{\tilde{\mathbf{h}}})^{\dagger}(\tilde{\mathbf{h}}-\hat{\tilde{\mathbf{h}}})\}\\ &= \frac{1}{MN_gL}\mathrm{tr}\Big\{E_{\mathbf{c},\hat{\mathbf{c}}}\{(\mathbf{c}-\hat{\mathbf{c}})(\mathbf{c}-\hat{\mathbf{c}})^{\dagger}\}\Big\} \qquad (85)\end{aligned}$$

Note that the error covariance matrix $E_{\mathbf{c},\hat{\mathbf{c}}}\Big\{(\mathbf{c}-\hat{\mathbf{c}})(\mathbf{c}-\hat{\mathbf{c}})^{\dagger}\Big\}$ is bounded as, $E_{\mathbf{c},\hat{\mathbf{c}}}\Big\{(\mathbf{c}-\hat{\mathbf{c}})(\mathbf{c}-\hat{\mathbf{c}})^{\dagger}\Big\} \geq (\mathbf{J}+\mathbf{\Sigma}_{\mathbf{c}}^{-1})^{-1}$, $\mathbf{J}$ being the Bayesian FIM. This bound is based on the Crámer Rao inequality. However, the conventional CRLB is not directly applicable, because the bias is not known for any estimator in general. To find a feasible solution under these circumstances, we assume that the bias is a linear function of $\mathbf{c}$, i.e., $E\{\hat{\mathbf{c}}|\mathbf{c}\} = \mathbf{A}\mathbf{c}+\mathbf{b}$, and that the bias terms $\mathbf{A}$ and $\mathbf{b}$ are determined by minimizing $E_{\mathbf{y},\mathbf{c}}\{(\hat{\mathbf{c}}-\mathbf{c})(\hat{\mathbf{c}}-\mathbf{c})^{\dagger}\}$. In the most general setting, let the observation equation be defined as $\mathbf{y} = \mathbf{f}(\mathbf{c})+\mathbf{w}$ and the Bayesian estimation of $\hat{\mathbf{c}} = \hat{\mathbf{c}}(\mathbf{y})$, where $\mathbf{f}(\cdot)$ is a function of $\mathbf{c}$. Note that $\mathbf{f}(\mathbf{c}) = \mathbf{Z}\mathbf{c}$ in our problem, as seen in (21). The exact derivations for the CRLB of $E_{\mathbf{c},\hat{\mathbf{c}}}\{(\mathbf{c}-\hat{\mathbf{c}})(\mathbf{c}-\hat{\mathbf{c}})^{\dagger}\}$ are given as follows:

$$\begin{aligned}&E_{\mathbf{c},\hat{\mathbf{c}}}\{(\mathbf{c}-\hat{\mathbf{c}})(\mathbf{c}-\hat{\mathbf{c}})^{\dagger}\}\\ &\quad= E_{\mathbf{y},\mathbf{c}}\{(\hat{\mathbf{c}}-\mathbf{c})(\hat{\mathbf{c}}-\mathbf{c})^{\dagger}\}\\ &\quad= E_{\mathbf{c}}E_{\mathbf{y}|\mathbf{c}}\Big\{\big(\hat{\mathbf{c}}-E_{\mathbf{y}|\mathbf{c}}\{\hat{\mathbf{c}}\}\big)\big(\hat{\mathbf{c}}-E_{\mathbf{y}|\mathbf{c}}\{\hat{\mathbf{c}}\}\big)^{\dagger}\Big\}\\ &\qquad+E_{\mathbf{c}}E_{\mathbf{y}|\mathbf{c}}\Big\{\big(E_{\mathbf{y}|\mathbf{c}}\{\hat{\mathbf{c}}\}-\mathbf{c}\big)\big(E_{\mathbf{y}|\mathbf{c}}\{\hat{\mathbf{c}}\}-\mathbf{c}\big)^{\dagger}\Big\}\\ &\quad= \mathbf{\Sigma}_{\hat{\mathbf{c}}}+E_{\mathbf{c}}\{\boldsymbol{\beta}(\mathbf{c})\boldsymbol{\beta}(\mathbf{c})^{\dagger}\} \qquad (86)\end{aligned}$$

where $\mathbf{\Sigma}_{\hat{\mathbf{c}}} = E_{\mathbf{y},\mathbf{c}}\Big\{\big(\hat{\mathbf{c}}-E_{\mathbf{y}|\mathbf{c}}\{\hat{\mathbf{c}}\}\big)\big(\hat{\mathbf{c}}-E_{\mathbf{y}|\mathbf{c}}\{\hat{\mathbf{c}}\}\big)^{\dagger}\Big\}$ and $\boldsymbol{\beta}(\mathbf{c}) \triangleq E_{\mathbf{y}|\mathbf{c}}\{\hat{\mathbf{c}}\}-\mathbf{c}$. We now derive the CRLB for $\mathbf{\Sigma}_{\hat{\mathbf{c}}}$ in (86). We first define $\boldsymbol{\gamma}(\mathbf{c}) \triangleq E_{\mathbf{y}|\mathbf{c}}\{\hat{\mathbf{c}}\}$. Recalling $E_{\mathbf{y}|\mathbf{c}}\Big\{\frac{\partial \log p(\mathbf{y}|\mathbf{c})}{\partial \mathbf{c}^T}\Big\} = \mathbf{0}_{DL\times 1}^T$ and after some algebra, it can be then shown that

$$E_{\mathbf{c}}\Big\{\frac{\partial\boldsymbol{\gamma}(\mathbf{c})}{\partial\mathbf{c}^T}\Big\} = E_{\mathbf{y},\mathbf{c}}\Big\{\big(\hat{\mathbf{c}}-\boldsymbol{\gamma}(\mathbf{c})\big)\frac{\partial\log p(\mathbf{y},\mathbf{c})}{\partial\mathbf{c}^T}\Big\}. \qquad (87)$$

Consequently, using arbitrary complex valued $\mathbf{a}_1$ and $\mathbf{a}_2$ vectors and applying the Cauchy-Schwarz inequality

$$\begin{aligned}&\Big|\mathbf{a}_1^{\dagger}E_{\mathbf{c}}\Big\{\frac{\partial\boldsymbol{\gamma}(\mathbf{c})}{\partial\mathbf{c}^T}\Big\}\mathbf{a}_2\Big|^2\\ &\quad\leq\int_{\mathbf{c}}\int_{\mathbf{y}}\mathbf{a}_1^{\dagger}\big(\hat{\mathbf{c}}-\boldsymbol{\gamma}(\mathbf{c})\big)\big(\hat{\mathbf{c}}-\boldsymbol{\gamma}(\mathbf{c})\big)^{\dagger}\mathbf{a}_1p(\mathbf{y},\mathbf{c})d\mathbf{y}d\mathbf{c}\\ &\qquad\times\int_{\mathbf{c}}\int_{\mathbf{y}}\mathbf{a}_2^{\dagger}\frac{\partial\log p(\mathbf{y},\mathbf{c})}{\partial\mathbf{c}^*}\frac{\partial\log p(\mathbf{y},\mathbf{c})}{\partial\mathbf{c}^T}\mathbf{a}_2p(\mathbf{y},\mathbf{c})d\mathbf{y}d\mathbf{c}\\ &\quad=\big(\mathbf{a}_1^{\dagger}\mathbf{\Sigma}_{\hat{\mathbf{c}}}\mathbf{a}_1\big)\big(\mathbf{a}_2^{\dagger}\mathbf{J}\mathbf{a}_2\big) \qquad (88)\end{aligned}$$

where $\mathbf{J}$ is the Bayesian FIM defined as

$$\begin{aligned}\mathbf{J} &\triangleq E_{\mathbf{y},\mathbf{c}}\Big\{\frac{\partial\log p(\mathbf{y},\mathbf{c})}{\partial\mathbf{c}^*}\frac{\partial\log p(\mathbf{y},\mathbf{c})}{\partial\mathbf{c}^T}\Big\}\\ &= -E_{\mathbf{y},\mathbf{c}}\Big\{\frac{\partial^2\log p(\mathbf{y},\mathbf{c})}{\partial\mathbf{c}^*\partial\mathbf{c}^T}\Big\}. \qquad (89)\end{aligned}$$

Since $\mathbf{a}_2$ is arbitrary, let

$$\mathbf{a}_2 = \mathbf{J}^{-1}E_{\mathbf{c}}\Big\{\frac{\partial\boldsymbol{\gamma}(\mathbf{c})}{\partial\mathbf{c}^T}\Big\}^{\dagger}\mathbf{a}_1. \qquad (90)$$

Substituting in (88), and then canceling positive semidefinite terms $\mathbf{a}_1^{\dagger}E_{\mathbf{c}}\Big\{\frac{\partial\boldsymbol{\gamma}(\mathbf{c})}{\partial\mathbf{c}^T}\Big\}\mathbf{J}^{-1}E_{\mathbf{c}}\Big\{\frac{\partial\boldsymbol{\gamma}(\mathbf{c})}{\partial\mathbf{c}^T}\Big\}^{\dagger}\mathbf{a}_1$ on both sides, we have

$$\mathbf{a}_1^{\dagger}\Big(\mathbf{\Sigma}_{\hat{\mathbf{c}}}-E_{\mathbf{c}}\Big\{\frac{\partial\boldsymbol{\gamma}(\mathbf{c})}{\partial\mathbf{c}^T}\Big\}\mathbf{J}^{-1}E_{\mathbf{c}}\Big\{\frac{\partial\boldsymbol{\gamma}(\mathbf{c})}{\partial\mathbf{c}^T}\Big\}^{\dagger}\Big)\mathbf{a}_1\geq 0. \qquad (91)$$

From (91), it follows that

$$\mathbf{\Sigma}_{\hat{\mathbf{c}}}\geq E_{\mathbf{c}}\Big\{\frac{\partial\boldsymbol{\gamma}(\mathbf{c})}{\partial\mathbf{c}^T}\Big\}\mathbf{J}^{-1}E_{\mathbf{c}}\Big\{\frac{\partial\boldsymbol{\gamma}(\mathbf{c})}{\partial\mathbf{c}^T}\Big\}^{\dagger}. \qquad (92)$$

Finally, inserting (92) into (86), the lower bound in (86) is obtained as

$$\begin{aligned}E_{\mathbf{y},\mathbf{c}}\{(\hat{\mathbf{c}}-\mathbf{c})(\hat{\mathbf{c}}-\mathbf{c})^{\dagger}\} &\geq \mathbf{B}_{\mathbf{c}}\\ &\triangleq E_{\mathbf{c}}\Big\{\frac{\partial\boldsymbol{\gamma}(\mathbf{c})}{\partial\mathbf{c}^T}\Big\}\mathbf{J}^{-1}E_{\mathbf{c}}\Big\{\frac{\partial\boldsymbol{\gamma}(\mathbf{c})}{\partial\mathbf{c}^T}\Big\}^{\dagger}\\ &\quad+E_{\mathbf{c}}\{\boldsymbol{\beta}(\mathbf{c})\boldsymbol{\beta}(\mathbf{c})^{\dagger}\}. \qquad (93)\end{aligned}$$

Now, assuming $\boldsymbol{\gamma}(\mathbf{c}) = \boldsymbol{\beta}(\mathbf{c})+\mathbf{c} = \mathbf{A}\mathbf{c}+\mathbf{b}$, we want to find a matrix $\mathbf{A}\in\mathcal{C}^{DL\times DL}$ and vector $\mathbf{b}\in\mathcal{C}^{DL\times 1}$ that minimize the lower bound in (93). So the lower bound in (93) turns into

$$\mathbf{B}_{\mathbf{c}} = \mathbf{A}(\mathbf{J}^{-1}+\mathbf{\Sigma}_{\mathbf{c}})\mathbf{A}^{\dagger}-\mathbf{A}\mathbf{\Sigma}_{\mathbf{c}}-\mathbf{\Sigma}_{\mathbf{c}}\mathbf{A}^{\dagger}+\mathbf{\Sigma}_{\mathbf{c}}+\mathbf{b}\mathbf{b}^{\dagger}. \qquad (94)$$

Note that $\mathbf{B}_{\mathbf{c}}$ is a positive definite matrix and we can minimize $\mathbf{a}_3^{\dagger}\mathbf{B}_{\mathbf{c}}\mathbf{a}_3$ instead of just minimizing $\mathbf{B}_{\mathbf{c}}$ itself using an arbitrary complex valued nonzero vector $\mathbf{a}_3$:

$$\begin{aligned}\frac{\partial\big(\mathbf{a}_3^{\dagger}\mathbf{B}_{\mathbf{c}}\mathbf{a}_3\big)}{\partial\mathbf{b}^{\dagger}} &= \mathbf{b}^T\mathbf{a}_3^*\mathbf{a}_3^T = \mathbf{0}_{DL\times 1}^T\\ \frac{\partial\big(\mathbf{a}_3^{\dagger}\mathbf{B}_{\mathbf{c}}\mathbf{a}_3\big)}{\partial\mathbf{A}^{\dagger}} &= \Big(\mathbf{A}(\mathbf{J}^{-1}+\mathbf{\Sigma}_{\mathbf{c}})-\mathbf{\Sigma}_{\mathbf{c}}\Big)^T\\ \mathbf{a}_3^*\mathbf{a}_3^T &= \mathbf{0}_{DL\times DL}. \qquad (95)\end{aligned}$$

From the solution of this equation set, we obtain the bias function coefficients that minimize the lower bound $\mathbf{B_c}$ in (94) as

$$\mathbf{b} = \mathbf{0}_{DL\times 1}, \quad \mathbf{A} = \mathbf{\Sigma_c}(\mathbf{J}^{-1} + \mathbf{\Sigma_c})^{-1}. \tag{96}$$

If we substitute (96) into (94) and apply the matrix inversion lemma, we obtain

$$\mathbf{B_c} = \mathbf{\Sigma_c} - \mathbf{\Sigma_c}(\mathbf{J}^{-1} + \mathbf{\Sigma_c})^{-1}\mathbf{\Sigma_c} = (\mathbf{J} + \mathbf{\Sigma_c}^{-1})^{-1}. \tag{97}$$

So, inserting (97) into (93), the minimum value of the CRLB for a linearly biased Bayesian estimator of the expansion coefficients is

$$E_{\mathbf{y},\mathbf{c}}\{(\hat{\mathbf{c}} - \mathbf{c})(\hat{\mathbf{c}} - \mathbf{c})^\dagger\} \geq (\mathbf{J} + \mathbf{\Sigma_c}^{-1})^{-1}. \tag{98}$$

Evaluation of Bayesian Fisher Information Matrix $\mathbf{J}$*:* $\mathbf{J}$ in (89) can be expressed as the sum of $\{\mathbf{J}_D(m)\}$ and $\mathbf{J}_P$ that are the FIMs evaluated from data and prior information, respectively

$$\mathbf{J} = \sum_{m=0}^{M-1} \mathbf{J}_D(m) + \mathbf{J}_P \tag{99}$$

where

$$\mathbf{J}_D(m) \triangleq -E_{\mathbf{c}}E_{\mathbf{y}(m)|\mathbf{c}}\Big\{\frac{\partial^2 \log p\big(\mathbf{y}(m)|\mathbf{c}\big)}{\partial \mathbf{c}^* \partial \mathbf{c}^T}\Big\} \tag{100}$$

and

$$\mathbf{J}_P \triangleq -E_{\mathbf{c}}\Big\{\frac{\partial^2 \log p(\mathbf{c})}{\partial \mathbf{c}^* \partial \mathbf{c}^T}\Big\}. \tag{101}$$

Since $\mathbf{c} \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma_c})$ from (101), it follows that:

$$\mathbf{J}_P = \mathbf{\Sigma_c}^{-1}. \tag{102}$$

We now compute the (i,j)th component of the matrix $\mathbf{J}_D(m)$ as follows. Consider the receive signal model given by (21). For simplicity of notation we use $\mathbf{\Sigma}(m) \triangleq \mathbf{\Sigma}_{\mathbf{y}(m)|\mathbf{c}}$. Noting that $\mathbf{y}(m)|\mathbf{c} \sim \mathcal{CN}\big(\mathbf{Z}_p(m)\mathbf{c}, \mathbf{\Sigma}(m)\big)$ and $\mathbf{\Sigma}(m) = \sum_{k\in\mathbf{I}_D} \mathbf{U}_k(m)\mathbf{c}\mathbf{c}^\dagger\mathbf{U}_k^\dagger(m) + N_0\mathbf{I}_N$, the $\log p\big(\mathbf{y}(m)|\mathbf{c}\big)$ in (100) can be expressed as

$$\begin{aligned}\log p\big(\mathbf{y}(m)|c\big) \sim & -\log\det\big(\mathbf{\Sigma}(m)\big) \\ & - \big(\mathbf{y}(m) - \mathbf{Z}_p(m)\mathbf{c}\big)^\dagger \mathbf{\Sigma}^{-1}(m) \\ & \times \big(\mathbf{y}(m) - \mathbf{Z}_p(m)\mathbf{c}\big).\end{aligned} \tag{103}$$

Using the identities

$$\begin{aligned}\frac{\partial \mathbf{\Sigma}^{-1}(m)}{\partial c(j)} &= -\mathbf{\Sigma}^{-1}(m)\frac{\partial \mathbf{\Sigma}(m)}{\partial c(j)}\mathbf{\Sigma}^{-1}(m) \\ \frac{\partial \log\det\big(\mathbf{\Sigma}(m)\big)}{\partial c(j)} &= \mathrm{tr}\Big\{\frac{\partial \mathbf{\Sigma}(m)}{\partial c(j)}\mathbf{\Sigma}^{-1}(m)\Big\}\end{aligned} \tag{104}$$

we first take partial derivatives with respect to $c^*(i)$ and $c(j)$, the ith and jth components of $\mathbf{c}^*$ and $\mathbf{c}$, respectively. Then the conditional expectation is taken with respect to $\mathbf{y}(m)|\mathbf{c}$ of (103) resulting in

$$\mathbf{J}_D(m) = E_c\big\{\Upsilon_m(\mathbf{c})\big\} \tag{105}$$

where

$$\begin{aligned}\Upsilon_m(\mathbf{c}) = & \big([\mathbf{Z}_p(m)]_{(:,i)}\big)^\dagger \mathbf{\Sigma}^{-1}(m)[\mathbf{Z}_p(m)]_{(:,j)} \\ & + \mathrm{tr}\Big\{\frac{\partial \mathbf{\Sigma}(m)}{\partial c^*(i)}\mathbf{\Sigma}^{-1}(m)\frac{\partial \mathbf{\Sigma}(m)}{\partial c(j)}\mathbf{\Sigma}^{-1}(m)\Big\}\end{aligned} \tag{106}$$

and

$$\begin{aligned}\mathbf{\Sigma}(m) &= \sum_{k\in\mathbf{I}_D} \mathbf{U}_k(m)\mathbf{c}\mathbf{c}^\dagger\mathbf{U}_k^\dagger(m) + N_o\mathbf{I}_N \\ \frac{\partial \mathbf{\Sigma}(m)}{\partial c(j)} &= \sum_{k\in\mathbf{I}_D} [\mathbf{U}_k(m)]_{(:,j)}\mathbf{c}^\dagger\mathbf{U}_k^\dagger(m) \\ \frac{\partial \mathbf{\Sigma}(m)}{\partial c^*(i)} &= \sum_{k\in\mathbf{I}_D} \mathbf{U}_k(m)\mathbf{c}\big([\mathbf{U}_k(m)]_{(:,i)}\big)^\dagger\end{aligned} \tag{107}$$

where the notation $[\mathbf{A}]_{(:,j)}$ denotes the jth column vector of the matrix $\mathbf{A}$. To compute the expectation in (105), we simply use the Gibbs sampling technique and generate samples $\mathbf{c}[1], \mathbf{c}[2], \ldots, \mathbf{c}[N_s]$ from the pdf $p(\mathbf{c}) = \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma_c})$. Then

$$\mathbf{J}_D(m) \approx \frac{1}{N_s}\sum_{n=1}^{N_s} \Upsilon_m\big(\mathbf{c}[n]\big). \tag{108}$$

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers whose insightful and constructive comments significantly improved the quality of this paper.

REFERENCES

[1] 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.

[2] 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, 2003.

[3] 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.

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

[5] A. Ancora, C. Bona, and D. T. M. Slock, "A down-sampled impulse response least-squares channel estimation for LTE OFDMA," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Honolulu, HI, Apr. 15–20, 2007, pp. 293–296.

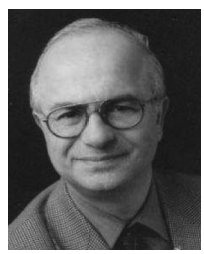
[6] X. Huang and H.-C. Wu, "Robust and efficient intercarrier interference mitigation for OFDM systems in time-varying fading channels," *IEEE Trans. Veh. Technol.*, vol. 56, no. 5, pp. 2517–2528, Sep. 2007.

[7] K. D. Teo and S. Ohno, "Optimal MMSE finite parameter model for doubly-selective channels," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, St. Louis, MO, Nov.–Dec. 28–2, 2005, pp. 3503–3507.
[8] T. Zemen and C. F. Mecklenbruker, "Time-variant channel estimation using discrete prolate spheroidal sequences," *IEEE Trans. Signal Process.*, vol. 53, no. 9, pp. 3597–3607, Sep. 2005.
[9] G. B. Giannakis and C. Tepedelenlioglu, "Basis expansion models and diversity techniques for blind equalization of time-varying channels," *Proc. IEEE*, vol. 86, pp. 1969–1986, Oct. 1998.
[10] 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.
[11] I. Barhumi, G. Leus, and M. Moonen, "Low-complexity pilot-aided channel estimation for OFDM systems over doubly-selective channels," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Seoul, Korea, May 16–20, 2005, pp. 1980–1984.
[12] Z. Tang, R. C. Cannizzaro, G. Leus, and P. Banelli, "Pilot-assisted time-varying channel estimation for OFDM systems," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 2226–2238, 2007.
[13] Y. Ma and R. Tafazolli, "Channel estimation for OFDMA uplink: A hybrid of linear and BEM interpolation approach," *IEEE Trans. Signal Process.*, vol. 55, no. 4, pp. 1568–1573, Apr. 2007.
[14] I. Barhumi, G. Leus, and M. Moonen, "Equalization for OFDM over doubly selective channels," *IEEE Trans. Signal Process.*, vol. 54, no. 4, pp. 1145–1443, 2008.
[15] J. Jhang, X. Mu, E. Chen, and S. Yang, "Decision-directed channel estimation based on iterative linear minimum mean square error for orthogonal frequency division multiplexing systems," *IET Commun.*, vol. 3, no. 7, pp. 1136–1143, 2008.
[16] R. Negi and J. Cioffi, "Pilot tone selection for channel estimation in a mobile OFDM system," *IEEE Trans. Consum. Electron.*, vol. 44, no. 8, pp. 1122–1128, Aug. 1998.
[17] C. R. N. Athaudage and A. D. S. Jayalath, "Enhanced MMSE channel estimation using timing error statistics for wireless OFDM system," *IEEE Trans. Broadcast.*, vol. 50, no. 4, pp. 369–376, Dec. 2004.
[18] H. Hijazi and L. Ros, "Joint data QR-detection and Kalman estimation for OFDM time-varying Rayleigh channel complex gains," *IEEE Trans. Commun.*, vol. 50, no. 1, pp. 170–178, Jan. 2010.
[19] 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.
[20] W. G. Jeon, K. H. Chang, and Y. S. Cho, "An equalization technique for orthogonal frequency-division multiplexing systems in time-variant multipath channels," *IEEE Trans. Commun.*, vol. 47, no. 1, pp. 27–32, Jan. 1999.
[21] P. Schniter, "Low-complexity equalization of OFDM in doubly-selective channels," *IEEE Trans. Signal Process.*, vol. 52, no. 4, pp. 1002–1011, Apr. 2004.
[22] D. K. Borah and B. D. Hart, "Frequency-selective fading channel estimation with a polynomial time-varying channel model," *IEEE Trans. Commun.*, vol. 47, no. 6, pp. 862–873, Jun. 1999.
[23] M. Visintin, "Karhunen-Loéve expansion of a fast Rayleigh fading process," *IEEE Electron. Lett.*, vol. 32, no. 8, pp. 1712–1713, Aug. 1996.
[24] J. A. Fessler and A. O. Hero, "Space-alternating generalized expectation-maximization algorithm," *IEEE Trans. Signal Process.*, vol. 42, no. 10, pp. 2664–2677, 1994.
[25] 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.
[26] M. Feder and E. Weinstein, "Parameter estimation of superimposed signals using the EM algorithm," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, no. 4, pp. 477–489, Apr. 1988.
[27] W. C. Jakes and D. C. Cox, *Microwave Mobile Communications*. New York: Wiley-IEEE, 1994.
[28] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
[29] H. Hijazi and L. Ros, "Joint data QR-detection and Kalman estimation for OFDM time-varying Rayleigh channel complex gains," *IEEE Trans. Commun.*, vol. 58, no. 1, pp. 170–177, Jan. 2010.
[30] H. Hijazi and L. Ros, "Polynomial estimation of time-varying multipath gains with intercarrier interference mitigation in OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 58, no. 1, pp. 140–151, Jan. 2009.
[31] M. F. Rabbi, S.-W. Hou, and C. C. Ko, "High mobility orthogonal frequency division multiple access channel estimation using basis expansion model," *IET Commun.*, vol. 4, no. 3, pp. 353–367, Jan. 2010.

Habib Şenol (S'04–M'07) 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.

From 1996 to 1999, he was a Research Assistant at Istanbul University. He is currently a faculty member of Computer Engineering at Kadir Has University, Istanbul. His research interests cover statistical signal processing, estimation and equalization algorithms for wireless communications, multicarrier (OFDM) systems, distributed detection, and estimation. He was a Postdoctoral Researcher with the Department of Electrical Engineering, Arizona State University, Tempe, in 2007. He has been a researcher of a 6th and 7th Frame European projects called Network of Excellent on Wireless Communications (NEWCOM) and WIMAGIC STREP.

Dr. Şenol served as a reviewer for the IEEE TRANSACTIONS including the IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, and IEEE TRANSACTIONS ON COMMUNICATIONS.



Erdal Panayırcı (S'73–M'80–SM'91–F'03) received the Diploma Engineering degree in electrical engineering from Istanbul Technical University, Istanbul, Turkey, and the Ph.D. degree in electrical engineering and system science from Michigan State University, East Lansing.

Until 1998, he was with the Faculty of Electrical and Electronics Engineering, Istanbul Technical University, where he was a Professor and Head of the Telecommunications Chair. Currently, he is a Professor of Electrical Engineering and Head of the Electrical and Electronics Engineering Department, Kadir Has University, Istanbul. His recent research interests include communication theory, synchronization, advanced signal processing techniques and their applications to wireless communications, coded modulation and interference cancelation with array processing. He published extensively in leading scientific journals and international conferences. He has coauthored the book *Principles of Integrated Maritime Surveillance Systems* (Boston, MA: Kluwer Academic, 2000). He spent two years (1980–1981), with the Department of Computer Science, Michigan State University, as a Fulbright-Hays Fellow and a NATO Senior Scientist. Between 1990–1991, he was with the Center for Communications and Signal Processing, New Jersey Institute of Technology, Hoboken, as a Visiting Professor. In 1998–2000, he was Visiting Professor with the Department of Electrical Engineering, Texas A&M University, College Station. In 2008–2009, he was a Research Scholar with the Department of Electrical Engineering, Princeton University, Princeton, NJ, working on new channel estimation and equalization algorithms for high mobility WiMAX and LTE systems. He has been the principal coordinator of a 6th and 7th Frame European project called Network of Excellent on Wireless Communications (NEWCOM) representing Kadir Has University for five years and WIMAGIC Strep project for two years.

Dr. Panayırcı was an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS in the areas of synchronizations and equalizations in 1995–1999. He served as a Member of the IEEE Fellow Committee in 2005–2008. He was the Technical Program Chair of the IEEE International Conference on Communications (ICC-2006) and Technical Program Chair of the IEEE PIMRC held in Istanbul in 2006 and 2010, respectively. He is the Vice Chair of the upcoming IEEE Wireless Communications and Networking (WCNC 2014) to be held in Istanbul, in April 2014. Presently, he is head of the Turkish Scientific Commission on Signals and Systems of International Union of Radio Science (URSI).

H. Vincent Poor (S'72–M'77–SM'82–F'87) received the Ph.D. degree in electrical engineering and computer science from Princeton University, Princeton, NJ, in 1977.

From 1977 until 1990, he was on the faculty of the University of Illinois at Urbana-Champaign. Since 1990, he has been on the faculty at Princeton University, where he is the Michael Henry Strater University Professor of Electrical Engineering and Dean of the School of Engineering and Applied Science. His research interests are in the areas of stochastic analysis, statistical signal processing, and information theory, and their applications in wireless networks and related fields such as social networks and smart grid. Among his publications in these areas are the recent books *Classical, Semi-classical and Quantum Noise* (New York: Springer, 2012) and *Smart Grid Communications and Networking* (Cambridge, U.K.: Cambridge University Press, 2012).

Dr. Poor is a member of the National Academy of Engineering and the National Academy of Sciences, a Fellow of the American Academy of Arts and Sciences, and an International Fellow of the Royal Academy of Engineering (U.K.). He is also a Fellow of the Institute of Mathematical Statistics, the Acoustical Society of America, and other organizations. In 1990, he served as President of the IEEE Information Theory Society, and in 2004–2007, he served as the Editor-in-Chief of the IEEE TRANSACTIONS ON INFORMATION THEORY. He received a Guggenheim Fellowship in 2002 and the IEEE Education Medal in 2005. Recent recognition of his work includes the 2010 IET Ambrose Fleming Medal, the 2011 IEEE Eric E. Sumner Award, the 2011 Society Award of the IEEE Signal Processing Society, and honorary doctorates from the University of Edinburgh and Aalborg University, conferred in 2011 and 2012, respectively.