An EM Based Frequency Domain Channel Estimation Algorithm for Multi-Access OFDM Systems

M. S. Sohail, T. Y. Al-Naffouri*

Electrical Engineering Department, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia

Abstract

Channel estimation is an important prerequisite for receiver design. In this paper we present a semi-blind low complexity frequency domain based channel estimation algorithm for multi-access Orthogonal Frequency Division Multiplexing (OFDM) systems. Our algorithm is based on eigenvalues interpolation and makes a collective use of data and channel constraints. We exploit these constraints to derive a frequency domain Maximum A Posteriori (MAP) channel estimator. Furthermore, we develop a data aided (expectation maximization based) estimator incorporating frequency correlation information. The estimator is further enhanced by utilizing the time correlation information through a Forward Backward (FB) Kalman filter. We also explore various implementation for the FB Kalman filter. The simulation results are provided validating the applicability of the proposed algorithm.

Key words: OFDM, channel estimation, model reduction, Multi-access Systems, Kalman Filtering.

*corresponding author *E-mail addresses:* saqib@kfupm.edu.sa (M. S. Sohail), naffouri@kfupm.edu.sa (T. Y. Al-Naffouri)

1 Introduction

OFDM is a technology that promises to meet the high transmission demands of modern times. Since the last decade, OFDM has attracted considerable attention and has been selected as the physical layer of choice for broadband wireless communications systems [1]-[4]. The main reason for this interest is the substantial advantage it offers in high rate transmissions over frequency selective fading channels like robustness to multi-path fading and capability to control the data rate according to the transmission channel [5]. The other main advantage of OFDM is its simple receiver structure utilizing a Frequency-domain Equalizer (FEQ) with only one complex multiplication per subcarrier to mitigate frequency selectivity. As such, OFDM has found wide acceptance and application.

Channel estimation is an integral part of any receiver design. Numerous research contributions have appeared in literature on the topic of channel estimation in recent years. Most works have investigated channel estimation for single user OFDM systems. One way to classify these works is according to whether they perform estimation in the time domain or the frequency domain.

Many researchers have pursued channel estimation in the time domain. A joint carrier frequency synchronization and channel estimation scheme using the Expectation-Maximization (EM) approach is presented in [6] while [7] used subspace tracking. In [8], a joint channel and data estimation algorithm is presented which makes a collective use of data and channel constraints. A joint frequency-offset and channel estimation technique for Multi-Symbol Encapsulated MSE OFDM system is proposed in [9], while the authors of [10] presented a sequential method based on carrier frequency offset and symbol timing estimation. The authors of [11] estimated the channel based on Power Spectral Density (PSD) and Least Squares (LS) estimation for OFDM systems with timing offsets. A pilot aided channel estimation algorithm in the presence of synchronous noise by exploiting the a priori available information about the interference structure was presented in [12] while [13] used implicit pilots for joint detection and

channel estimation. A joint time domain tracking of channel frequency offset for OFDM systems is suggested in [14] while a time domain Carrier Frequency Offset (CFO) tracking method based on Particle filtering is presented in [15].

Various techniques for channel estimation in the frequency domain have also been explored in the past years. Authors of [16] applied phase shifted samples in the frequency-domain to an LS interpolator to estimate the channel while [17] proposed to include a phase rotation term in the frequency domain interpolation for better Channel Impulse Response (CIR) window location. The authors of [18] proposed channel estimation using polynomial cancellation coding, training symbols and frequency domain windowing. The authors in [19] presented a sub-band approach to channel estimation and channel equalization. The Minimum Mean Square Error (MMSE) channel estimation in the frequency domain is considered in [20] while authors of [21] explored delay subspace-based channel estimation techniques for OFDM systems over fast-fading channels.

Due to the complex nature of the problem, various iterative methods have also been explored [22, 23]. Such techniques iterate between data detection step and channel estimation step, using the result of one to improve the other. In [24], two EM based channel estimation algorithm for OFDM systems using transmit diversity are compared, [25] considered a Space Time Block Coded (STBC) EM based approach for OFDM systems while an iterative method for detection and frequency synchronization in multi access OFDM systems is presented in [26]. All three of these works perform channel estimation in time domain. The data detection step is always performed in frequency domain. Thus for iterative methods, performing channel estimation in time domain would not only incur more computations but introduce some latency as well.

In multiuser OFDM scenario, channel estimation has several dissimilarities from the single user case, and poses several additional challenges, e.g., in a multiuser OFDM system, the user has access to a part of the frequency band and not the whole band. Also the number of pilots available per user are much less in the multiuser case as compared to the single user OFDM

case. In practical systems like the long term evolution (LTE), the cell specific reference signal (CRS) in the downlink transmission is accessible to all users. Although this CRS can be used for channel estimation purposes, this information might not be frequent enough if a user is moving at a higher speed relative to other users. All of these hurdles are elegantly handled in case of frequency domain channel estimation. The correlation among the carriers in the frequency domain is usually higher than the correlation among the time domain taps. The reason being the time domain taps represent the whole OFDM bandwidth while the frequency domain "bins" correspond to only a narrow band of the spectrum. This correlation can be exploited to drastically reduce the training overhead in the frequency domain and estimation can be performed using $L_P < L + 1$ pilots (where L_P is the number of pilots used and $L + 1$ is the number of time domain channel taps). Of course one cannot reduce training indefinitely as this will result in an error floor.

Particle filtering based solutions (like in [15]) have limited applications as Particle filtering has been shown to require up to 2 to 4 times more computations than the extended Kalman filter to achieve the same accuracy [27]. As such, Kalman filtering based solutions are more practical. Furthermore, using system constraints (frequency and time correlation) to aid the estimation process is bound to give better results. It is, therefore, worthwhile to investigate a frequency domain approach for channel estimation and the aim of this work is to present a low complexity frequency domain channel estimation algorithms for multi-access OFDM systems.

A problem associated with frequency domain channel estimation is the increase in the parameter estimation space [20]. Let $L+1$ be the number of time domain channel taps and N be the length of frequency domain channel response with $N > L + 1$. The frequency response of the channel is inherently limited by the degrees of freedom of the time domain impulse response. As a user is only assigned a small band of the spectrum in multi access OFDM system, the number of parameters needed to represent the CIR can be reduced. In Section 3, we will show how to utilize this property to represent the CIR using the dominant eigenvalues.

The paper is organized as follows. Section 2 describes the OFDM system model. Section 3 presents a new frequency domain data aided parameter reduction model for channel estimation. Section 4 introduces time correlation information to improve the receiver design. Section 5 gives a quick overview of the equivalent time domain channel estimation method. Section 6 compares the computational complexity of the presented algorithms in time and frequency domain. Section 7 discusses the simulation results and Section 8 provides the concluding remarks.

2 System Model

We consider the downlink of a multi-access OFDM system in which K users are active simultaneously. The total number of subcarriers of the system is denoted by N, out of which each user has access to $N_K = \frac{N}{K}$ subcarriers[1]. Figure 1 shows the block diagram of the system under consideration. The data bits, to be sent over the system, are first fed to a convolutional encoder, punctured and then passed through a random interleaver. The bit sequence thus obtained, is mapped to QAM symbols using Gray code. The QAM symbols are then mapped to OFDM symbols at the data tones and pilots are inserted at the pilot tones. Here we consider comb-type pilots as they are more robust in fast fading channels than block-type pilots [28]. We will use calligraphic notations (e.g., $\boldsymbol{\mathcal{X}}$) for vectors in the frequency domain. Now consider a sequence of $T+1$ such OFDM symbols $\{\boldsymbol{\mathcal{X}}_0, \boldsymbol{\mathcal{X}}_1, \ldots, \boldsymbol{\mathcal{X}}_T\}$to be transmitted. Each OFDM symbol $\boldsymbol{\mathcal{X}}_i$ (of length N), undergoes an IFFT operation to produce the time domain symbol $\boldsymbol{x}_i = \sqrt{N}\boldsymbol{Q}^*\boldsymbol{\mathcal{X}}_i$, where $\boldsymbol{Q}$ is the $N \times N$ Discrete Fourier Transform (DFT) matrix given by $\boldsymbol{Q} = [e_{m,n}^{-j\frac{2\pi}{N}(m-1)(n-1)}]$ with $m = 1, 2, \ldots, N$, $n = 1, 2, \ldots, N$ and where $*$ denotes conjugate transpose. The transmitter then appends a Cyclic Prefix (CP) of length P and transmits the resulting super symbol. The CP is discarded at the receiver thereby mitigating the effect of Inter Symbol Interference (ISI).

[1]For simplicity we assume that the available subcarriers are distributed equally among all users. However, this need not be the case and our development applies even when the number of carriers per user is different from one user to another.

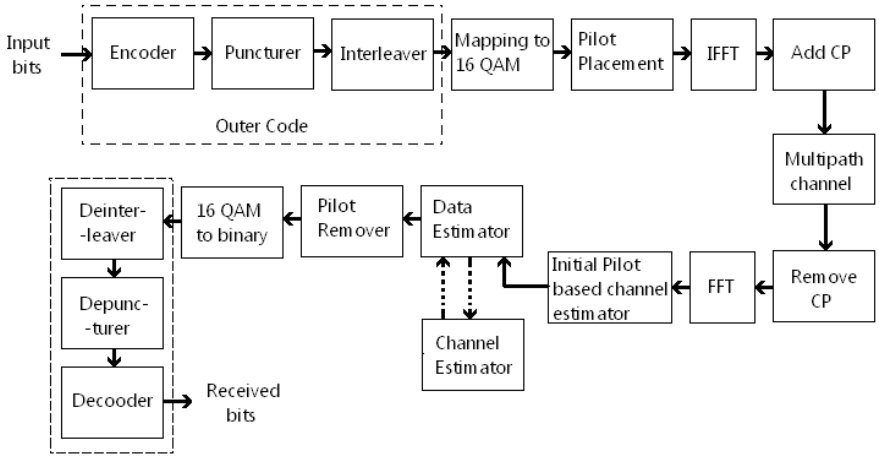


Figure 1: Block diagram.

The channel $\boldsymbol{h}_i$, of length $L+1$ ($\leq P+1$), remains fixed over any OFDM symbol and associated cyclic prefix and varies from one symbol to the next according to a state-space model[2]

$$\boldsymbol{h}_{i+1} = \boldsymbol{F}\boldsymbol{h}_i + \boldsymbol{G}\boldsymbol{u}_i \quad \boldsymbol{h}_0 \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_h), \quad \boldsymbol{u}_i \sim \mathcal{N}(\boldsymbol{0}, \sigma_u^2 \boldsymbol{I}), \tag{1}$$

where the subscript i shows the time dependence of the variables. $\boldsymbol{R}_h$ is the channel correlation assumed to be known a priori at the receiver (in our simulations we take it as the power delay profile) and $\boldsymbol{u}_i$ is the additive noise[3]. The matrices $\boldsymbol{F}$ and $\boldsymbol{G}$ are a function of the doppler spread, the power delay profile and the transmit filter, and thus the model in (1) captures both frequency and time correlation (for details on the construction of these matrices, and a justification of this model, see [8]). Now $\boldsymbol{\mathcal{H}}_i$ is the length-N FFT of $\boldsymbol{h}_i$ given as

$$\boldsymbol{\mathcal{H}}_i = \boldsymbol{Q} \begin{bmatrix} \boldsymbol{h}_i \\ \boldsymbol{0} \end{bmatrix} = \boldsymbol{Q}_{P+1} \boldsymbol{h}_i, \tag{2}$$

[2]The noise and channel statistics are assumed to be stationary with respect to the channel coherence time and generally estimated by sending a training sequence for this purpose.

[3]$\boldsymbol{u}_i$ models the randomness in $\boldsymbol{h}_{i+1}$. So part of $\boldsymbol{h}_{i+1}$ is a function of $\boldsymbol{h}_i$, i.e., $\boldsymbol{F}\boldsymbol{h}_i$ and the other part is random modeled by $\boldsymbol{G}\boldsymbol{u}_i$.

where $\boldsymbol{Q}_{P+1}$ is the matrix which contains the first $P+1$ columns of $\boldsymbol{Q}$. Let $\underline{\boldsymbol{\mathcal{H}}}_i^{(k)}$ corresponding to the k^{th} section of the frequency response of the channel in which the k^{th} user operates. The input/output relationship for this user is best described in the frequency domain and is given by

$$\underline{\boldsymbol{\mathcal{Y}}}_i^{(k)} = \text{diag}(\underline{\boldsymbol{\mathcal{X}}}_i^{(k)})\underline{\boldsymbol{\mathcal{H}}}_i^{(k)} + \underline{\boldsymbol{\mathcal{N}}}_i^{(k)}, \tag{3}$$

where $\underline{\boldsymbol{\mathcal{Y}}}_i^{(k)}$, $\underline{\boldsymbol{\mathcal{X}}}_i^{(k)}$, $\underline{\boldsymbol{\mathcal{H}}}_i^{(k)}$ and $\underline{\boldsymbol{\mathcal{N}}}_i^{(k)}$ are the k^{th} sections of $\boldsymbol{\mathcal{Y}}_i$ (the received signal), $\boldsymbol{\mathcal{X}}_i$, $\boldsymbol{\mathcal{H}}_i$ and $\boldsymbol{\mathcal{N}}_i$ (the additive white Gaussian noise $\mathcal{N}(\boldsymbol{0}, \sigma^2\boldsymbol{I})$), respectively. Let the index set $I_p^{(k)} = \{i_1, i_2, \ldots, i_{L_p}\}$ denote the pilot locations within the k^{th} section of the transmitted data known *a priori* at the receiver. Also, let $\boldsymbol{A}_{I_p}$ denote the matrix $\boldsymbol{A}$ pruned of the rows that do not belong to $I_p^{(k)}$, then the pilot/output equation corresponding to (3) is given by

$$\underline{\boldsymbol{\mathcal{Y}}}_{I_p,i} = \text{diag}(\underline{\boldsymbol{\mathcal{X}}}_{I_p,i})\underline{\boldsymbol{\mathcal{H}}}_{I_p,i} + \underline{\boldsymbol{\mathcal{N}}}_{I_p,i}, \tag{4}$$

where in (4) and thereafter we suppress the dependence on the user index k for notational convenience. This suppression is possible because the users are independent from each other in the frequency domain.

3 Frequency Domain Channel Estimation

Consider the input/output equation in 3, of the k^{th} user. It consists of N_K equations in N_K unknowns. So to estimate the frequency response $\underline{\boldsymbol{\mathcal{H}}}$, all carriers should be used for training. In the following subsection , we pursue a parameter reduction approach for estimating $\underline{\boldsymbol{\mathcal{H}}}$.

3.1 A parameter reduction approach

We begin by presenting a frequency domain based LS algorithm for channel estimation and later extend the idea to incorporate time correlation over several symbols in Section 4. The algorithms presented in this section process on a symbol by symbol basis, therefore we will drop the dependence of variables on time index i for clarity. In order to reduce the parameter space

to be estimated we resort to model reduction, starting from the autocorrelation function of $\underline{\boldsymbol{\mathcal{H}}}$, $\boldsymbol{R}_{\underline{\boldsymbol{\mathcal{H}}}}$ (where $\boldsymbol{R}_{\underline{\boldsymbol{\mathcal{H}}}}$ is the FFT of $\boldsymbol{R}_h$ which itself is the autocorrelation function of $\boldsymbol{h}$)[4]. To this end, consider the eigenvalue decomposition of $\boldsymbol{R}_{\underline{\boldsymbol{\mathcal{H}}}}$

$$\boldsymbol{R}_{\underline{\boldsymbol{\mathcal{H}}}} = \sum_{n_k=1}^{N_K} \lambda_{n_k} \mathbf{v}_{n_k} \mathbf{v}_{n_k}^T$$

where $\lambda_1 \geq \lambda_2 \ldots \geq \lambda_{N_K}$ are the *ordered* eigenvalues of $\boldsymbol{R}_{\underline{\boldsymbol{\mathcal{H}}}}$ and $\mathbf{v}_1, \ldots, \mathbf{v}_{N_K}$ are the corresponding eigenvectors. We can use this decomposition to represent $\underline{\boldsymbol{\mathcal{H}}}$ as

$$\underline{\boldsymbol{\mathcal{H}}} = \sum_{n_k=1}^{N_K} \alpha_{n_k} \mathbf{v}_{n_k},$$

where $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \ldots, \alpha_{N_K}]^T$ is a parameter vector, to be estimated, with zero mean and autocorrelation matrix $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_{N_K})$. We now represent $\underline{\boldsymbol{\mathcal{H}}}$ using the dominant eigenvalues and treat the rest as modeling noise [5], i.e.

$$\underline{\boldsymbol{\mathcal{H}}} = \boldsymbol{V}_d \boldsymbol{\alpha}_d + \boldsymbol{V}_n \boldsymbol{\alpha}_n. \tag{5}$$

Upon substituting (5) in (3), we obtain

$$\underline{\boldsymbol{\mathcal{Y}}} = \mathrm{diag}(\underline{\boldsymbol{\mathcal{X}}}) \boldsymbol{V}_d \boldsymbol{\alpha}_d + \underline{\boldsymbol{\mathcal{N}}} + \mathrm{diag}(\underline{\boldsymbol{\mathcal{X}}}) \boldsymbol{V}_n \boldsymbol{\alpha}_n = \underline{\boldsymbol{X}}_d \boldsymbol{\alpha}_d + \underline{\boldsymbol{\mathcal{N}}}^{'}, \tag{6}$$

where $\underline{\boldsymbol{X}}_d = \mathrm{diag}(\underline{\boldsymbol{\mathcal{X}}}) \boldsymbol{V}_d$ and $\underline{\boldsymbol{\mathcal{N}}}^{'} = \underline{\boldsymbol{\mathcal{N}}} + \underline{\boldsymbol{X}}_n \boldsymbol{\alpha}_n$ with $\underline{\boldsymbol{X}}_n = \mathrm{diag}(\underline{\boldsymbol{\mathcal{X}}}) \boldsymbol{V}_n$. The noise $\underline{\boldsymbol{\mathcal{N}}}'$ includes both the additive and *modeling* noise. We consider it to be zero mean white gaussian noise with autocorrelation

$$\boldsymbol{R}_{\underline{\mathcal{N}}'} = \boldsymbol{R}_{\underline{\mathcal{N}}} + \mathrm{diag}(\underline{\boldsymbol{\mathcal{X}}}) \boldsymbol{V}_n \mathrm{diag}(\boldsymbol{\Lambda}_n) \boldsymbol{V}_n^* \mathrm{diag}(\underline{\boldsymbol{\mathcal{X}}})^*, \tag{7}$$

where $\boldsymbol{\Lambda}_n$ is diagonal matrix constituting of the non-dominant ordered eigenvalues of $\boldsymbol{R}_{\underline{\boldsymbol{\mathcal{H}}}}$. Now equation (6) can be used to construct a pilot/output equation similar to (4), as

$$\underline{\boldsymbol{\mathcal{Y}}}_{I_p} = \underline{\boldsymbol{X}}_{d,I_p} \boldsymbol{\alpha}_d + \underline{\boldsymbol{\mathcal{N}}}^{'}_{I_p}, \tag{8}$$

[4] Although channel taps are usually assumed to be uncorrelated, the presence of the receive filter and the sampling process introduce correlation in the channel taps.

[5] The cutoff between the parameters that are considered dominant and the ones that are considered as part of the modeling noise depends on the relative values of the $\lambda_j^{'}s$. In our simulations, we use the condition $\frac{\lambda_j}{\lambda_{j+1}} > 5$ to place our cutoff.

which can be used to estimate $\boldsymbol{\alpha}_d$ by maximizing the log likelihood function

$$\hat{\boldsymbol{\alpha}}_d^{\mathsf{MAP}} = \arg\max_{\alpha_d} \left\{ \ln\, p(\underline{\boldsymbol{\mathcal{Y}}}_{I_p} | \underline{\boldsymbol{X}}_{d,I_p}, \boldsymbol{\alpha}_d) + \ln\, p(\boldsymbol{\alpha}_d) \right\}. \tag{9}$$

The MAP estimate of parameter $\boldsymbol{\alpha}$ is thus given by[6]

$$\hat{\boldsymbol{\alpha}}_d^{\mathsf{MAP}} = \arg\min_{\alpha_d} \left\{ \|\underline{\boldsymbol{\mathcal{Y}}}_{I_p} - \underline{\boldsymbol{X}}_{d,I_p}\boldsymbol{\alpha}_d\|^2_{\boldsymbol{R}^{-1}_{\underline{\mathcal{N}}'_{I_p}}} + \|\boldsymbol{\alpha}_d\|^2_{\boldsymbol{\Lambda}_d^{-1}} \right\}, \tag{10}$$

where $\boldsymbol{\Lambda}_d$ is a diagonal matrix constituting of the dominant ordered eigenvalues of $\boldsymbol{R}_{\underline{\boldsymbol{\mathcal{H}}}}$. As the channel is assumed to be jointly gaussian, so the MAP estimator is the same as the MMSE estimator for the same input and output sequence (for proof see Problem 10.9 in [29]). The above optimization simplifies to

$$\hat{\boldsymbol{\alpha}}_d = \boldsymbol{\Lambda}_d \underline{\boldsymbol{X}}^*_{d,I_p} \left[\boldsymbol{R}_{\underline{\mathcal{N}}'_{I_p}} + \underline{\boldsymbol{X}}_{d,I_p} \boldsymbol{\Lambda}_d \underline{\boldsymbol{X}}^*_{d,I_p} \right]^{-1} \underline{\boldsymbol{\mathcal{Y}}}_{I_p}, \tag{11}$$

and the resulting mean square error is given by

$$\boldsymbol{R}_e = \left[\boldsymbol{\Lambda}_d^{-1} + \underline{\boldsymbol{X}}^*_{I_p} \boldsymbol{R}^{-1}_{\underline{\mathcal{N}}'_{I_p}} \underline{\boldsymbol{X}}_{I_p} \right]^{-1}. \tag{12}$$

The estimate of the k^{th} section of the spectrum is then given by $\hat{\underline{\boldsymbol{\mathcal{H}}}}^{(k)} = \boldsymbol{V}_d^{(k)} \hat{\boldsymbol{\alpha}}_d^{(k)}$.

3.2 Iterative channel estimation using the expectation maximization approach

Pilot based channel estimation, whether in the time domain or frequency domain, does not make full use of the constraints on the data. One can thus implement iterative (data-aided) techniques for channel estimation [8]. A formal way to do so is by implementing the expectation maximization (EM) algorithm.As its name suggests, the EM algorithm consists of two steps: a maximization step and an expectation step. We discuss them in the following subsections.

3.2.1 The maximization step

In the previous subsection we found $\hat{\boldsymbol{\alpha}}_d$ by maximizing the log likelihood function given by equation (9). Since the input $\boldsymbol{\mathcal{X}}$ (and hence $\underline{\boldsymbol{X}}_d$) is not observable, we can employ the EM

[6] We use the weighted norm $\|\boldsymbol{\alpha}_d\|^2_{\Sigma}$ to denote $\boldsymbol{\alpha}_d^* \boldsymbol{\Sigma} \boldsymbol{\alpha}_d$.

algorithm and instead of maximizing (9) we can maximize an averaged form of the log likelihood function. Specifically, starting from an initial estimate $\hat{\boldsymbol{\alpha}}_d^{(0)}$, calculated say using pilots, the estimate $\hat{\boldsymbol{\alpha}}_d$ is calculated iteratively with the estimate at the j^{th} iteration given by

$$\hat{\boldsymbol{\alpha}}_d^{(j)} = \arg\max_{\alpha_d} \left\{ E_{\mathcal{X}_i|\mathcal{Y}_i,\hat{\boldsymbol{\alpha}}_d{}^{(j-1)}} \ln\, p(\underline{\boldsymbol{\mathcal{Y}}}|\underline{\boldsymbol{X}}_d, \boldsymbol{\alpha}_d) + \ln\, p(\boldsymbol{\alpha}_d) \right\}, \tag{13}$$

which simplifies to [7]

$$\hat{\boldsymbol{\alpha}}_d^{\mathsf{MAP}} = \arg\min_{\alpha_d} \left\{ E\|[\underline{\boldsymbol{\mathcal{Y}}} - \underline{\boldsymbol{X}}_d\boldsymbol{\alpha}_d\|^2_{\boldsymbol{R}^{-1}_{\underline{\mathcal{N}'}}} + \|\boldsymbol{\alpha}_d\|^2_{\boldsymbol{\Lambda}_d^{-1}} \right\}. \tag{14}$$

Strictly speaking, the noise correlation $\boldsymbol{R}_{\underline{\mathcal{N}'}}$ is itself dependent on the input due to the modeling noise (see equation (7)). Hence in performing the expectation in (14), we need to take this into account. Treating the general case is difficult, so we consider the following three cases for $\boldsymbol{R}^{-1}_{\mathcal{N}'}$:

Case 1: $\boldsymbol{\mathcal{R}}_{\mathcal{N}'}$ is a constant:

This happens when we ignore the modeling noise so that $\boldsymbol{\mathcal{R}}_{\underline{\mathcal{N}'}} = \sigma^2\boldsymbol{I}$ where the expectation in (14) is taken with respect to $\underline{\boldsymbol{X}}_d$ given $\underline{\boldsymbol{\mathcal{Y}}}$ and the most recent estimate $\boldsymbol{\alpha}_d$. In this case $\boldsymbol{R}_{\boldsymbol{\mathcal{N}}'}$ becomes independent of $\underline{\boldsymbol{X}}_d$ and it would then be straight forward to carry the expectation in (14). Specifically, upon completing the squares, (14) can be equivalently written as

$$\begin{aligned}\min_{\alpha_d} \boldsymbol{\mathcal{Y}}_i^* \boldsymbol{R}^{-1}_{\underline{\mathcal{N}'}} \boldsymbol{\mathcal{Y}}_i - \boldsymbol{\alpha}_d^* E[\underline{\boldsymbol{X}}_d^*] \boldsymbol{R}^{-1}_{\underline{\mathcal{N}'}} \boldsymbol{\mathcal{Y}}_i - \boldsymbol{\mathcal{Y}}_i^* \boldsymbol{R}^{-1}_{\underline{\mathcal{N}'}} E[\underline{\boldsymbol{X}}_d]\boldsymbol{\alpha}_d \\ + \boldsymbol{\alpha}_d^* E[\underline{\boldsymbol{X}}_d^*] \boldsymbol{R}^{-1}_{\underline{\mathcal{N}'}} E[\underline{\boldsymbol{X}}_d]\boldsymbol{\alpha}_d - \boldsymbol{\alpha}_d^* E[\underline{\boldsymbol{X}}_d^*] \boldsymbol{R}^{-1}_{\underline{\mathcal{N}'}} E[\underline{\boldsymbol{X}}_d]\boldsymbol{\alpha}_d \\ + \boldsymbol{\alpha}_d^* E[\underline{\boldsymbol{X}}_d^* \boldsymbol{R}^{-1}_{\underline{\mathcal{N}'}} \underline{\boldsymbol{X}}_d]\boldsymbol{\alpha}_d + \boldsymbol{\alpha}_d^* \boldsymbol{\Lambda}_d^{-1} \boldsymbol{\alpha}_d,\end{aligned}$$

which can be simplified to

$$\hat{\boldsymbol{\alpha}}_d^{\mathsf{MAP}} = \arg\min_{\alpha_d} \ \|\underline{\boldsymbol{\mathcal{Y}}} - E[\underline{\boldsymbol{X}}_d]\boldsymbol{\alpha}_d\|^2_{\frac{1}{\sigma_n^2} I} + \|\boldsymbol{\alpha}_d\|^2_{\frac{1}{\sigma_n^2} Cov[\underline{\boldsymbol{X}}_d^*]} + \|\boldsymbol{\alpha}_d\|^2_{\Lambda_d^{-1}}. \tag{15}$$

Case 2: Taking Expectation of $\boldsymbol{\mathcal{R}}_{\mathcal{N}'}$:

Instead of ignoring the modeling noise, we can split the expectation in (14) into an expectation over $\boldsymbol{\mathcal{R}}_{\underline{\mathcal{N}'}}$ and an *independent* expectation taken over the rest of the terms i.e., we can

[7] The expectation is taken with respect to the input given the output and the most recent estimate $\hat{\boldsymbol{\alpha}}_d^{j-1}$. This information is understood and dropped for notational convenience.

approximate (14) as

$$\hat{\boldsymbol{\alpha}}_d^{\mathsf{MAP}} = \arg\min_{\alpha_d} \left\{ E\|[\underline{\boldsymbol{\mathcal{Y}}} - \underline{\boldsymbol{X}}_d \boldsymbol{\alpha}_d\|^2_{E[R_{\underline{\mathcal{N}}'}]^{-1}} + \|\boldsymbol{\alpha}_d\|^2_{\boldsymbol{\Lambda}_d^{-1}} \right\}. \tag{16}$$

Now the expectation of $\boldsymbol{\mathcal{R}}_{\underline{\mathcal{N}}'}$ is given by

$$\begin{aligned} E[\boldsymbol{R}_{\underline{\boldsymbol{\mathcal{N}}}'}] &= \sigma^2 \boldsymbol{I} + E[\mathrm{diag}(\underline{\boldsymbol{\mathcal{X}}}) \boldsymbol{V}_n \boldsymbol{\Lambda}_n \boldsymbol{V}_n^* \mathrm{diag}(\underline{\boldsymbol{\mathcal{X}}}^*)] & (17) \\ &= \sigma^2 \boldsymbol{I} + E[\underline{\boldsymbol{D}}] \boldsymbol{V}_n \boldsymbol{\Lambda}_n \boldsymbol{V}_n^* E[\underline{\boldsymbol{D}}^*] + \mathrm{Cov}[\underline{\boldsymbol{D}}] \mathrm{diag}(\boldsymbol{V}_n \boldsymbol{\Lambda}_n \boldsymbol{V}_n^*), & (18) \end{aligned}$$

where $\underline{\boldsymbol{D}} = \mathrm{diag}(\underline{\boldsymbol{\mathcal{X}}})$ and where $\mathrm{diag}(\boldsymbol{V}_n \boldsymbol{\Lambda}_n \boldsymbol{V}_n^*)$ is a diagonal matrix whose diagonal coincides with the diagonal of the matrix $\boldsymbol{V}_n \boldsymbol{\Lambda}_n \boldsymbol{V}_n^*$ (see Appendix A for details). The now averaged $\boldsymbol{R}_{\underline{\boldsymbol{\mathcal{N}}}'}$ does not depend on $\underline{\boldsymbol{\mathcal{X}}}$ any more. Replacing $\boldsymbol{R}_{\underline{\boldsymbol{\mathcal{N}}}'}$ by its expectation, it is then straight forward to carry the expectation in (16) which comes out to be

$$\hat{\boldsymbol{\alpha}}_d^{\mathsf{MAP}} = \arg\min_{\alpha_d} \ \|\underline{\boldsymbol{\mathcal{Y}}} - E[\underline{\boldsymbol{X}}_d]\boldsymbol{\alpha}_d\|^2_{E[R_{\underline{\mathcal{N}}'}]^{-1}} + \|\boldsymbol{\alpha}_d\|^2_{\mathrm{Cov}[\underline{\boldsymbol{D}}]\mathrm{diag}(\boldsymbol{V}_n \boldsymbol{\Lambda}_n \boldsymbol{V}_n^*)} + \|\boldsymbol{\alpha}_d\|^2_{\Lambda_d^{-1}}. \tag{19}$$

Case 3: $\underline{\boldsymbol{\mathcal{X}}}$ is constant modulus:

In the constant modulus case, it is possible to evaluate (14) exactly. Specifically, and starting from the expression for the autocorrelation as $\boldsymbol{R}_{\underline{\mathcal{N}}'} = \sigma^2 \boldsymbol{I} + \underline{\boldsymbol{D}} \boldsymbol{V}_n \boldsymbol{\Lambda}_n \boldsymbol{V}_n^* \underline{\boldsymbol{D}}^*$ and we can write

$$\begin{aligned} \boldsymbol{R}_{\underline{\mathcal{N}}'}^{-1} &= (\sigma^2 \boldsymbol{I} + \underline{\boldsymbol{D}} \boldsymbol{V}_n \boldsymbol{\Lambda}_n \boldsymbol{V}_n^* \underline{\boldsymbol{D}}^*)^{-1} \\ &= \underline{\boldsymbol{D}}^{-*} (\frac{\sigma^2}{\mathcal{E}} \boldsymbol{I} + \boldsymbol{V}_n \boldsymbol{\Lambda}_n \boldsymbol{V}_n^*)^{-1} \underline{\boldsymbol{D}}^{-1} \\ &= \underline{\boldsymbol{D}}^{-*} \boldsymbol{R}_{\underline{\mathcal{N}}''}^{-1} \underline{\boldsymbol{D}}^{-1}, \end{aligned}$$

where $\boldsymbol{R}_{\underline{\mathcal{N}}''} \triangleq \frac{\sigma^2}{\mathcal{E}} \boldsymbol{I} + \boldsymbol{V}_n \boldsymbol{\Lambda}_n \boldsymbol{V}_n^*$ and where we used the fact that $\underline{\boldsymbol{D}}\underline{\boldsymbol{D}}^* = \mathcal{E}\boldsymbol{I}$ since the input is constant modulus. With this in mind, we conclude that

$$\underline{\boldsymbol{X}}_d^* \boldsymbol{R}_{\underline{\mathcal{N}}'}^{-1} = \boldsymbol{V}_d^* \underline{\boldsymbol{D}}^* \boldsymbol{R}_{\underline{\mathcal{N}}}^{-1} = \boldsymbol{V}_d^* \boldsymbol{R}_{\underline{\mathcal{N}}''}^{-1} \underline{\boldsymbol{D}}^{-1}$$

$$\boldsymbol{R}_{\underline{\mathcal{N}}'}^{-1} \underline{\boldsymbol{X}}_d = \underline{\boldsymbol{D}}^{-1*} \boldsymbol{R}_{\underline{\mathcal{N}}''}^{-1} \boldsymbol{V}_d$$

and

$$\underline{\boldsymbol{X}}_d^* \boldsymbol{R}_{\underline{\mathcal{N}}'}^{-1} \underline{\boldsymbol{X}}_d = \boldsymbol{V}_d^* \boldsymbol{R}_{\underline{\mathcal{N}}''}^{-1} \boldsymbol{V}_d.$$

Thus, in the constant modulus case, (14) can be equivalently written as

$$\begin{aligned}\hat{\boldsymbol{\alpha}}_d^{\mathsf{MAP}} = \arg\min_{\alpha_d} \boldsymbol{\mathcal{Y}}^* E[\underline{\boldsymbol{D}}^{-1*}]\boldsymbol{R}_{\underline{\mathcal{N}}''}^{-1} E[\underline{\boldsymbol{D}}^{-1}]\boldsymbol{\mathcal{Y}} - \boldsymbol{\mathcal{Y}}^* E[\underline{\boldsymbol{D}}^{-1*}]\boldsymbol{R}_{\underline{\mathcal{N}}''}^{-1}\boldsymbol{V}_d\boldsymbol{\alpha}_d \\ -\boldsymbol{\alpha}_d^*\boldsymbol{V}_d^*\boldsymbol{R}_{\underline{\mathcal{N}}''}^{-1} E[\underline{\boldsymbol{D}}^{-1}]\boldsymbol{\mathcal{Y}} + \boldsymbol{\alpha}_d^*\boldsymbol{V}_d^*\boldsymbol{R}_{\underline{\mathcal{N}}''}^{-1}\boldsymbol{V}_d\boldsymbol{\alpha}_d + \boldsymbol{\alpha}_d^*\boldsymbol{\Lambda}_d^{-1}\boldsymbol{\alpha}_d,\end{aligned}$$

which upon simplification becomes

$$\hat{\boldsymbol{\alpha}}_d^{\mathsf{MAP}} = \arg\min_{\alpha_d} \|E[\underline{\boldsymbol{D}}^{-1}]\boldsymbol{\mathcal{Y}} - \boldsymbol{V}_d\boldsymbol{\alpha}_d\|^2_{R^{-1}_{\underline{\mathcal{N}}''}} + \|\boldsymbol{\alpha}_d\|^2_{\Lambda_d^{-1}}. \tag{20}$$

In the simulations further ahead, we compare the approximate solutions (15) and (19) with the exact EM solution (20) for a constant modulus input. Simulations show that replacing $\boldsymbol{R}'_{\mathcal{N}}$ with its expectation is almost as good as calculating the expectation exactly.

3.2.2 The expectation step

By inspecting subsection 3.2.1, we see that to perform the maximization step, we need to calculate the following moments

$$E[\underline{\boldsymbol{X}}_d],\ \ Cov[\underline{\boldsymbol{X}}_d^*],\ \ E[\underline{\boldsymbol{D}}],\ \ E[\underline{\boldsymbol{D}}\boldsymbol{B}\underline{\boldsymbol{D}}^*],\ \text{ and }\ E[\underline{\boldsymbol{D}}^{-1}]. \tag{21}$$

Now as $\underline{\boldsymbol{X}}_d = \mathrm{diag}(\underline{\boldsymbol{\mathcal{X}}})\boldsymbol{V}_d = \underline{\boldsymbol{D}}\boldsymbol{V}_d$ we can see that the moments of $\underline{\boldsymbol{X}}_d$ can be expressed in terms of moments of $\underline{\boldsymbol{D}}$. Specifically we have that

$$E[\underline{\boldsymbol{X}}_d] = E[\underline{\boldsymbol{D}}]\boldsymbol{V}_d,$$

and

$$\begin{aligned} Cov[\underline{\boldsymbol{X}}_d^*] &= E[\underline{\boldsymbol{X}}_d\underline{\boldsymbol{X}}_d^*] - E[\underline{\boldsymbol{X}}_d]E[\underline{\boldsymbol{X}}_d^*] \\ &= E[\underline{\boldsymbol{D}}]\boldsymbol{V}_d\boldsymbol{V}_d^*E[\underline{\boldsymbol{D}}^*] + Cov[\underline{\boldsymbol{D}}]\mathrm{diag}(\boldsymbol{V}_d\boldsymbol{V}_d^*) - E[\underline{\boldsymbol{D}}]\boldsymbol{V}_d\boldsymbol{V}_d^*E[\underline{\boldsymbol{D}}^*] \\ &= Cov[\underline{\boldsymbol{D}}]\mathrm{diag}(\boldsymbol{V}_d\boldsymbol{V}_d^*). \end{aligned}$$

Moreover, we show in Appendix A that

$$E[\underline{\boldsymbol{D}}\boldsymbol{B}\underline{\boldsymbol{D}}^*] = E[\underline{\boldsymbol{D}}]\boldsymbol{B}E[\underline{\boldsymbol{D}}^*] + \mathrm{Cov}[\underline{\boldsymbol{D}}]\mathrm{diag}(\boldsymbol{B}). \tag{22}$$

From above, it follows that in order to calculate the expectations in (21), it is enough to calculate the following three moments

$$E[\text{diag}(\underline{\boldsymbol{\mathcal{X}}})],\ \ Cov[\text{diag}(\underline{\boldsymbol{\mathcal{X}}})]\ \text{ and }\ E[\text{diag}(\underline{\boldsymbol{\mathcal{X}}})^{-1}], \tag{23}$$

where the expectation is performed given the output $\underline{\boldsymbol{\mathcal{Y}}}$ and the most recent channel estimate $\underline{\hat{\boldsymbol{\mathcal{H}}}}$. In carrying out these expectations, we will assume that the elements of $\underline{\boldsymbol{\mathcal{X}}}$ are independent[8]. With this in mind, it is easy to see that we can evaluate the moments in (23) and hence in (21) by calculating $E[\underline{\mathcal{X}}(l)],\ \ \text{Cov}[\underline{\mathcal{X}}(l)] = E[|\underline{\mathcal{X}}(l)|^2] - |E[\underline{\mathcal{X}}(l)]|^2,\ \ E[\frac{1}{\underline{\mathcal{X}}(l)}]$. Now assuming that $\underline{\mathcal{X}}(l)$ is drawn from the alphabet $A = \{A_1, \ldots, A_M\}$ with equal probability, it can be shown that [8]

$$E[\underline{\mathcal{X}}(l)|\underline{\mathcal{Y}}(l), \underline{\mathcal{H}}(l)] = \frac{\sum_{j=1}^{M} A_j e^{-\frac{|\underline{\mathcal{Y}}(l) - \underline{\mathcal{H}}(l) A_j|^2}{\sigma^2}}}{\sum_{j=1}^{M} e^{-\frac{|\underline{\mathcal{Y}}(l) - \underline{\mathcal{H}}(l) A_j|^2}{\sigma^2}}} \tag{24}$$

$$E[|\underline{\mathcal{X}}(l)|^2|\underline{\mathcal{Y}}(l), \underline{\mathcal{H}}(l)] = \frac{\sum_{j=1}^{M} |A_j|^2 e^{-\frac{|\underline{\mathcal{Y}}(l) - \underline{\mathcal{H}}(l) A_j|^2}{\sigma^2}}}{\sum_{j=1}^{M} e^{-\frac{|\underline{\mathcal{Y}}(l) - \underline{\mathcal{H}}(l) A_j|^2}{\sigma^2}}} \tag{25}$$

$$E[\frac{1}{\underline{\mathcal{X}}(l)}|\underline{\mathcal{Y}}(l), \underline{\mathcal{H}}(l)] = \frac{\sum_{j=1}^{M} \frac{1}{A_j} e^{-\frac{|\underline{\mathcal{Y}}(l) - \underline{\mathcal{H}}(l) A_j|^2}{\sigma^2}}}{\sum_{j=1}^{M} e^{-\frac{|\underline{\mathcal{Y}}(l) - \underline{\mathcal{H}}(l) A_j|^2}{\sigma^2}}}. \tag{26}$$

3.2.3 Summary of the EM algorithm

Now let us summarize the EM based estimation algorithm developed so far.

1. Calculate the initial channel estimate $\underline{\hat{\boldsymbol{\mathcal{H}}}}^{(0)}$ using pilots (10).

2. Calculate the moments of the input given the current (j^{th}) channel estimate $\underline{\hat{\boldsymbol{\mathcal{H}}}}^{(j)}$ and the output $\underline{\boldsymbol{\mathcal{Y}}}$ using equations (24)-(26).

3. Calculate the channel estimate using either one of the methods (15), (19) or (20) outlined in Section 3.2.

[8]This is in general not true because the elements of $\underline{\boldsymbol{\mathcal{H}}}$ are not independent (as the elements of $\underline{\boldsymbol{\mathcal{H}}}$ are the Fourier transform of the impulse response $\boldsymbol{h}$). However, we continue to use this approximation as this maintains the transparency of element-by-element equalization in OFDM.

4. Iterate between steps 2 and 3.

The algorithm can be run for a specific number of times or until some predefined minimum error threshold is reached.

4 Using Time-Correlation to Improve the Channel Estimate

The receiver developed in the previous section performs channel estimation on a symbol by symbol basis. In a practical scenario the channel impulse responses are correlated over time. In this section, we will show how to use time correlation to enhance the estimate of $\boldsymbol{\alpha}_{d,i}$ (here we reintroduce the dependence of variables on time index i). To this end, let us first develop a model for the time variation of the parameter $\boldsymbol{\alpha}_{d,i}$.

4.1 Developing a frequency domain time-variant model

Consider the block fading model in (1) and lets assume for simplicity that the diagonal matrices $\boldsymbol{F}$ and $\boldsymbol{G}$ are actually scalar multiples of the identity, i.e. $\boldsymbol{F} = f\boldsymbol{I}$ and $\boldsymbol{G} = \sqrt{1-f^2}\boldsymbol{I}$ where f is a function of Doppler frequency (see [8]). We will use the time domain model in (1) to derive a similar model for $\boldsymbol{\alpha}$. To this end, recall that $\boldsymbol{\mathcal{H}}_i = \boldsymbol{Q}_{P+1}\boldsymbol{h}_i$. Thus, the k^{th} section of $\boldsymbol{\mathcal{H}}_i$, $\underline{\boldsymbol{\mathcal{H}}}_i^{(k)}$, is related to $\boldsymbol{h}_i$ by

$$\underline{\boldsymbol{\mathcal{H}}}_i^{(k)} = \boldsymbol{Q}_{P+1}^{(k)}\boldsymbol{h}_i, \tag{27}$$

where $\boldsymbol{Q}_{P+1}^{(k)}$ corresponds to the k^{th} section of $\boldsymbol{Q}_{P+1}$, i.e., $\boldsymbol{Q}_{P+1}$ pruned of all its rows except those of the k^{th} section. Now, we can replace $\underline{\boldsymbol{\mathcal{H}}}_i^{(k)}$ by its representation using the dominant parameters $\boldsymbol{\alpha}_d$, to get

$$\boldsymbol{V}_d\boldsymbol{\alpha}_{d,i} = \boldsymbol{Q}_{P+1}^{(k)}\boldsymbol{h}_i,$$

or

$$\boldsymbol{\alpha}_{d,i} = \boldsymbol{V}_d^{+}\boldsymbol{Q}_{P+1}^{(k)}\boldsymbol{h}_i,$$

where $\boldsymbol{V}_d^+$ is the pseudo inverse of $\boldsymbol{V}_d$. Multiplying both sides of (1) by $\boldsymbol{V}_d^+ \boldsymbol{Q}_{P+1}^{(k)}$ yields a dynamical recursion for $\boldsymbol{\alpha}_d$

$$\boldsymbol{\alpha}_{d,i+1} = \boldsymbol{F}_\alpha \boldsymbol{\alpha}_{d,i} + \boldsymbol{G}_\alpha \boldsymbol{u}_i, \tag{28}$$

where $\boldsymbol{F}_\alpha = f\boldsymbol{I}$ and $\boldsymbol{G}_\alpha = \sqrt{1-f^2}\boldsymbol{V}_d^+ \boldsymbol{Q}_{P+1}^{(k)}$ and where $E[\boldsymbol{\alpha}_{d,0}\boldsymbol{\alpha}_{d,0}^*] = \boldsymbol{\Lambda}_d$. Note that the dependence of $\boldsymbol{G}_\alpha$ and $\boldsymbol{\alpha}_d$ on k has again been suppressed for notational convenience. We are now ready to implement the EM algorithm to the frequency domain system governed by the dynamical equation (28). As we have seen in Section 3, the algorithm will consist of an initial estimation step, a maximization step, and an expectation step.

4.2 Initial (Pilot-Based) Channel Estimation

In the initial channel estimation step, the frequency domain system is described by equations (8) and (28), reproduced here for convenience.

$$\underline{\boldsymbol{\mathcal{Y}}}_{I_p,i} = \underline{\boldsymbol{X}}_{d,I_p,i}\boldsymbol{\alpha}_{d,i} + \underline{\boldsymbol{\mathcal{N}}}'_{I_p,i} \tag{29}$$

$$\boldsymbol{\alpha}_{d,i+1} = \boldsymbol{F}_\alpha \boldsymbol{\alpha}_{d,i} + \boldsymbol{G}_\alpha \boldsymbol{u}_i. \tag{30}$$

Now given a sequence $i = 0, 1, \ldots, T$ of pilot bearing symbols, we can obtain the optimum estimate of $\{\boldsymbol{\alpha}_{i,d}\}_{i=0}^T$ by applying a forward-backward Kalman to (29)-(30) (see[8], [29]), i.e., by implementing the following equations.

Forward run: Starting from the initial conditions $\boldsymbol{P}_{0|-1} = \boldsymbol{\Pi}_0$ and $\boldsymbol{\alpha}_{0|-1} = \boldsymbol{0}$ and for $i = 0, \ldots, T$, calculate

$$\boldsymbol{R}_{e,i} = \boldsymbol{R}_{\underline{\mathcal{N}}'} + \underline{\boldsymbol{X}}_{d,I_p,i}\boldsymbol{P}_{i|i-1}\underline{\boldsymbol{X}}_{d,I_p,i}^* \tag{31}$$

$$\boldsymbol{K}_{f,i} = \boldsymbol{P}_{i|i-1}\underline{\boldsymbol{X}}_{d,I_p,i}^*\boldsymbol{R}_{e,i}^{-1} \tag{32}$$

$$\hat{\boldsymbol{\alpha}}_{i|i} = \left(\boldsymbol{I} - \boldsymbol{K}_{f,i}\underline{\boldsymbol{X}}_{d,I_p,i}\right)\hat{\boldsymbol{\alpha}}_{i|i-1} + \boldsymbol{K}_{f,i}\boldsymbol{\mathcal{Y}}_i \tag{33}$$

$$\hat{\boldsymbol{\alpha}}_{i+1|i} = \boldsymbol{F}_\alpha \hat{\boldsymbol{\alpha}}_{i|i} \tag{34}$$

$$\boldsymbol{P}_{i+1|i} = \boldsymbol{F}_\alpha\left(\boldsymbol{P}_{i|i-1} - \boldsymbol{K}_{f,i}\boldsymbol{R}_{e,i}\boldsymbol{K}_{f,i}^*\right)\boldsymbol{F}_\alpha^* + \frac{1}{\sigma_n^2}\boldsymbol{G}_\alpha\boldsymbol{G}_\alpha^*. \tag{35}$$

Backward run: Starting from $\boldsymbol{\lambda}_{T+1|T} = \mathbf{0}$ and for $i = T, T-1, \ldots, 0$, calculate

$$\boldsymbol{\lambda}_{i|T} = \left(\boldsymbol{I}_{P+N} - \underline{\boldsymbol{X}}^*_{d,I_p,i}\boldsymbol{K}^*_{f,i}\right)\boldsymbol{F}^*_i\boldsymbol{\lambda}_{i+1|T} + \underline{\boldsymbol{X}}^*_{d,I_p,i}\boldsymbol{R}^{-1}_{e,i}\left(\boldsymbol{\mathcal{Y}}_i - \underline{\boldsymbol{X}}_{d,I_p,i}\hat{\boldsymbol{\alpha}}_{i|i-1}\right) \tag{36}$$

$$\hat{\boldsymbol{\alpha}}_{i|T} = \hat{\boldsymbol{\alpha}}_{i|i-1} + \boldsymbol{P}_{i|i-1}\boldsymbol{\lambda}_{i|T}. \tag{37}$$

The desired estimate is $\hat{\boldsymbol{\alpha}}_{i|T}$. This gives us an initial estimate to run the data-aided part of the algorithm.

4.3 Iterative (data-aided) channel estimation

In the data aided channel estimation problem, we use the whole data symbol and not just the pilot part (i.e. we use the whole k^{th} section of the spectrum in which the user is operating). Thus, in this case our system is described by equations (6) and (28) also reproduced here for convenience

$$\underline{\boldsymbol{\mathcal{Y}}}_i = \underline{\boldsymbol{X}}_{d,i}\boldsymbol{\alpha}_{i,d} + \underline{\boldsymbol{\mathcal{N}}}'_i \tag{38}$$

$$\boldsymbol{\alpha}_{d,i+1} = \boldsymbol{F}_\alpha\boldsymbol{\alpha}_{d,i} + \boldsymbol{G}_\alpha\boldsymbol{u}_i. \tag{39}$$

If the data symbols $\underline{\boldsymbol{X}}_{d,i}$ were known, we would have employed the forward-backward Kalman-Filter (31)-(37) on the above state-space model. Since the input is not available, we replace it by its estimate. Specifically, along the lines developed in [8] we can show that the FB Kalman filter needs to be applied to the following state space model

$$\underline{\boldsymbol{\mathcal{Y}}}_i = \begin{bmatrix} E[\underline{\boldsymbol{X}}_{d,i}] \\ Cov[\underline{\boldsymbol{X}}^*_{d,i}]^{\frac{1}{2}} \end{bmatrix}\boldsymbol{\alpha}_{i,d} + \begin{bmatrix} \underline{\boldsymbol{\mathcal{N}}}'_i \\ \mathbf{0} \end{bmatrix} \tag{40}$$

$$\boldsymbol{\alpha}_{d,i+1} = \boldsymbol{F}_\alpha\boldsymbol{\alpha}_{d,i} + \boldsymbol{G}_\alpha\boldsymbol{u}_i, \tag{41}$$

where the expectations in (40) are taken given the output $\underline{\boldsymbol{\mathcal{Y}}}_i$ and most recent channel estimate $\boldsymbol{\alpha}_{d,i}$. The expectations that appear in (40) are calculated as we did in Subsection 3.2.2.

In contrast to the symbol by symbol EM algorithm of Section 3 where there was one dimension to iterate against (channel estimation vs data detection), there are two dimensions that we can iterate against when the channel is correlated over several symbols:

1. We can iterate between channel estimation and data detection.

2. We could also iterate against time using the Kalman filter where the previous channel estimate informs the subsequent channel estimate.

Depending on how we schedule iterations across these two dimensions, we get different receivers. In the following, we discuss the implementation structure of these variations.

4.4 Cyclic FB Kalman

In the cyclic based Kalman, we initialize the algorithm using the FB Kalman implemented over the pilot symbols. This is then used to initialize the data aided version, where the channel estimate is used to obtain the data estimate, and that allows us to propagate the estimate to the next symbol. The process is continued until the forward steps are completed followed by the backward run. The EM steps are repeated again (2^{nd} forward run followed by 2^{nd} backward run and so on). In other words, we iterate only *once* between channel estimation and data detection before invoking the Kalman to move to the next symbol and so on. The iterations thus trace circles over the OFDM symbols which motivates the name Cyclic Kalman.

4.5 Helix based FB Kalman

The Helix based FB Kalman is a more generalized version of the Cyclic Kalman. The two filters are initialized in the same way. However at each symbol, we iterate *several* times between channel estimation and data detection before moving on the next symbol (whereas the cyclic Kalman iterates once between the channel estimate and data estimate at each step). This allows us to refine the channel estimate as much as possible before propagating it to the next OFDM symbol using the Kalman filter. The iterations in this case draw a helix shape, hence the name.

4.6 Code enhanced Kalman

In any practical system, an outer code is usually implemented that extends over several OFDM symbols. The outer code can be used to enhance the data aided channel estimate. Specifically, following data detection, the code can be invoked to enhance the data estimate (through error correction). Now the (hard) data obtained is more refined and hence can be used to enhance the channel estimate by employing the FB Kalman again. Our simulations show that invoking the code can have a profound effect on performance.

4.7 Forward Kalman filter

One drawback of the three FB Kalman implementations detailed above is the latency and memory involved as one needs to store *all* symbols to perform the backward run. One way around that is to implement the forward only Kalman which avoids the latency problem. The forward only Kalman thus suffers as a result in performance and is not able to make use of the code to enhance the channel estimate.

5 Benchmarking

For benchmarking purposes, we compare the frequency domain (LS and Kalman) receivers of Sections 3 and 4 with a frequency domain estimator (LMMSE). We also compare with a time domain estimation method to signify the advantage offered by the proposed methods over time domain estimation techniques under same constraints.

5.1 Frequency domain LMMSE estimator

To gauge the performance of the proposed frequency domain methods, we compare them with the standard LMMSE algorithm. We compare the results with two implementations of the LMMSE estimator, one using only frequency correlation and second using both time and frequency correlation information for estimating the channel (for further detail see [30]).

5.2 Time domain multiple access channel estimation

We also compare the performance of our proposed algorithms against time domain estimation method. How do users estimate the channel in the time domain given their limited share of the spectrum? To describe this, we just need to write the input/output equations seen by each user. The input/output equation for the k^{th} user is given by (3) as $\underline{\boldsymbol{\mathcal{Y}}}_i^{(k)} = \text{diag}(\underline{\boldsymbol{\mathcal{X}}}_i^{(k)})\underline{\boldsymbol{\mathcal{H}}}_i^{(k)} + \underline{\boldsymbol{\mathcal{N}}}_i^{(k)}$. Recall that the k^{th} section of frequency response $\underline{\boldsymbol{\mathcal{H}}}_i^{(k)}$ can be written as $\underline{\boldsymbol{\mathcal{H}}}_i^{(k)} = \boldsymbol{Q}_{P+1}^{(k)}\boldsymbol{h}_i$ using (27). So, we can write

$$\underline{\boldsymbol{\mathcal{Y}}}_i^{(k)} = \text{diag}(\underline{\boldsymbol{\mathcal{X}}}_i^{(k)})\boldsymbol{Q}_{P+1}^{(k)}\boldsymbol{h}_i + \underline{\boldsymbol{\mathcal{N}}}_i^{(k)}. \tag{42}$$

Equation (42) can be used for initial *time domain* estimate using pilots and for symbol by symbol EM based estimation. If we use in addition the dynamic recursion of (1) we can implement the various kind of Kalman filters discussed in the previous section for *time domain* channel estimation (see also [8]).

6 Computational Complexity

We compare the computational complexity of the various algorithms presented and compare it with the standard Linear Minimum Mean Square Error (LMMSE) algorithm. We represent the subcarriers assigned to the k^{th} user by $N_K = \frac{N}{K}$, d is the number of dominant eigenvalues, n is the number of non-dominant eigenvalues, l is the OFDM frame length, j is the number of EM iterations, M is the size of the alphabet set and P is the number of time domain channel taps. Results are given in Table 1 in approximate number of complex multiplications (CM). For a fair comparison between algorithms that perform estimation on a symbol by symbol basis (Section 3) and those that perform estimation frame by frame basis (Section 4), we compare the approximate number of CM each algorithm will perform for l symbols.

Table 1 lists the various algorithms compared in the paper. The estimation domain of the algorithms, frequency domain (FD) or time domain (TD), is also indicated. We see that when

A PRIORI INFO. USED	ALGORITHM	COMPLEXITY
frequency correlation	LMMSE (FD)	$ljN_K^3 + lj(3M+1)N_K^2 + 2ljN_K$
	$\boldsymbol{\mathcal{R}}_{\mathcal{N}'}$ constant (FD)	$ljN_K^3 + lj(4M+d+1)N_K^2 + 4ljdN_K$
	Constant modulus (FD)	$ljN_K^3 + lj(3M+d+n+1)N_K^2 + lj(3d+n+1)N_K$
	Expection of $\boldsymbol{\mathcal{R}}_{\mathcal{N}'}$ (FD)	$ljN_K^3 + lj(4M+d+n+3)N_K^2 + lj(4d+n)N_K$
frequency and time correlation	Helix Kalman (FD)	$l(j+1)N_K^3 + lj(3d+n+3M+3)N_K^2$ $+l(4jd^2+jd+jn^2+d^2+2d)N_K$
	Helix Kalman (TD)	$l(j+1)N_K^3 + l(3jM+4P)N_K^2 + l(4P^2+6P+2jP)N_K$
	Forward Kalman (FD)	$ljN_K^3 + lj(3d+n+3M+3)N_K^2 + lj(4d^2+d+n^2)N_K$
	Cyclic Kalman (FD)	$2ljN_K^3 + lj(3d+n+3M+3)N_K^2 + lj(5d^2+3d+n^2)N_K$
	Code Enhanced Kalman (FD)	$2l(j+1)N_K^3 + 2lj(3d+n+3M+3)N_K^2$ $+2l(4jd^2+jd+jn^2+d^2+2d)N_K$
	LMMSE (FD)	$l^3jN_K^3 + l^2j(3M+2)N_K^2 + ljN_K$

Table 1: Comparison of computational complexities.

only frequency correlation information is used, the algorithms presented here are of the same complexity order as the LMMSE. However, when both frequency and time domain correlation are used, the Kalman filters (Forward/Helix/Cyclic/Code Enhanced) are of lower complexity than the LMMSE. Furthermore, the computational complexity of the frequency domain Helix Kalman is lower than the time domain Helix Kalman as the number of significant eigenvalues $\boldsymbol{\alpha}_d$ are less than the channel length P.

7 Simulation Results

We consider a multi-access OFDM system which has 4 active users and frame size of 6 symbols ($l = 6$). Each OFDM symbol has 64 carriers, of which 16 are (equispaced) pilot tones distributed equally among the 4 users. The CP used is of length 15. Each user has access to 16 subcarriers. The data bits are mapped to 16 QAM through Gray coding (except for Figure 2 which uses

4 QAM). Where specified, an outer code is used to provide robustness. The outer code is 1/2 rate convolutional code. The time domain channel impulse response consists of 15 complex taps (worst case scenario while avoiding ICI) and has an exponential delay profile $E[|\underline{h}_0(t)|^2] = e^{-0.2t}$ and remains fixed over any OFDM symbol. We also test our methods on the ITU vehicular B channel model (Figures 5(a) and 5(b)). In what follows, we compare the performance of various frequency domain based channel estimation methods for the both the coded and uncoded cases. We benchmark our proposed methods with the standard LMMSE algorithm with frequency and time correlation and the time domain method briefly described in Section 5.

7.1 Effect of modeling noise

Figure 2 shows the MSE for the three cases considered in Subsection 3.2. The figure is plotted for constant modulus data using 16 pilots. The figure shows that the MSE of the channel estimate decreases when modeling noise is taken into account. It also shows that for the constant modulus case, splitting the expectation in (14) or calculating the expectation exactly (exact solution), have almost comparable results.

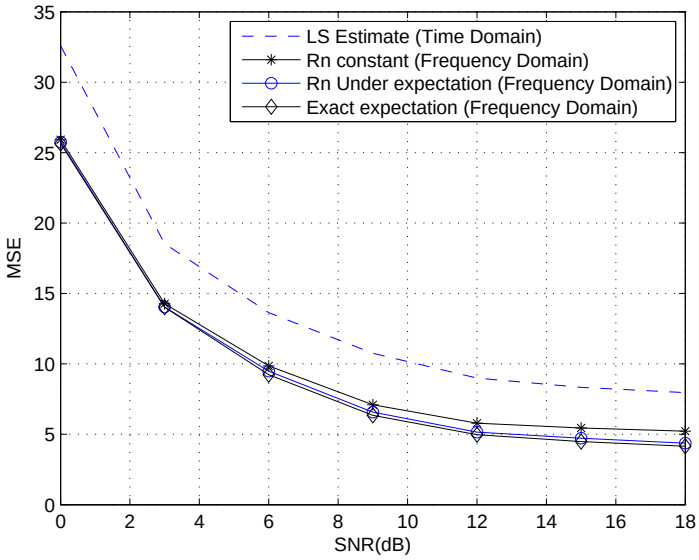


Figure 2: Effect of modeling noise.

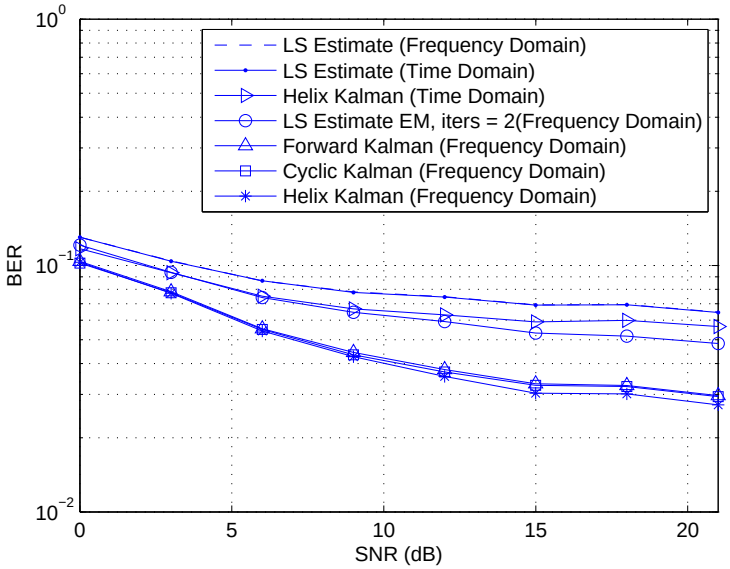


Figure 3: BER comparison for various uncoded freq. domain methods.

7.2 Kalman filter based receivers

Figure 3 compares the BER performance of frequency domain Forward Kalman, Cyclic and Helix Kalman filters with the time domain LS method and Helix Kalman for uncoded data. As expected, we see that using Kalman filter improves the EM based estimate in the frequency domain. We also see that Helix based Kalman performs better than other frequency domain based techniques and that the frequency domain methods fair better than the time domain methods.

Figure 4(a) compares the BER performance, for the coded case, of various frequency domain Kalman based algorithms presented in this paper with frequency domain variants of LMMSE estimator, EM based LS and time domain Helix Kalman estimator. When only the frequency correlation information is used at the receiver, the EM based LS algorithm gives better results than the standard LMMSE estimator. When both the frequency and time correlation information is utilized at the receiver for computing the estimate, the Kalman based receivers

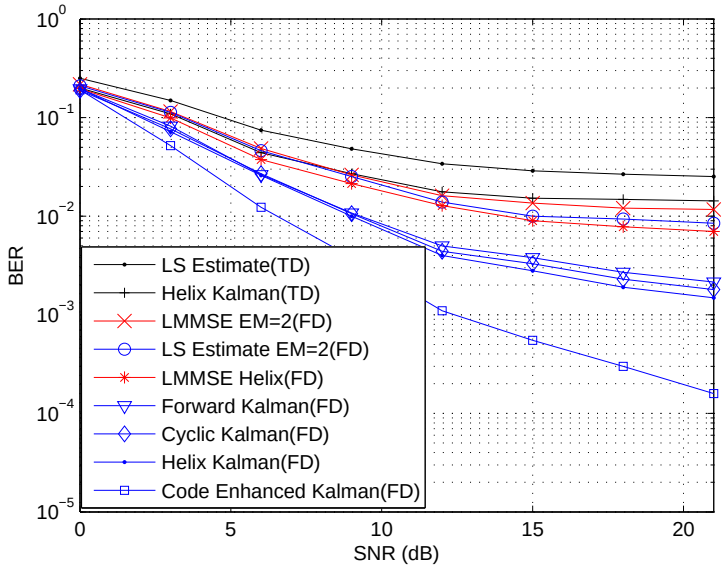
outperform the LMMSE based receiver. The code enhanced Kalman proves to be the best of the lot.

Figure 4(b) compares the presented methods when the channel taps are not located on the FFT grid. When the delay profile of the channel is not uniform, the taps will not lie on the FFT grid and the channel taps can now be considered to be correlated. This correlation will be known at the receiver as the delay profile of the channel and the receive filter response are assumed to be known. Figure 4(b) shows the error floor is much lower specially in the case of the Code Enhanced Kalman.

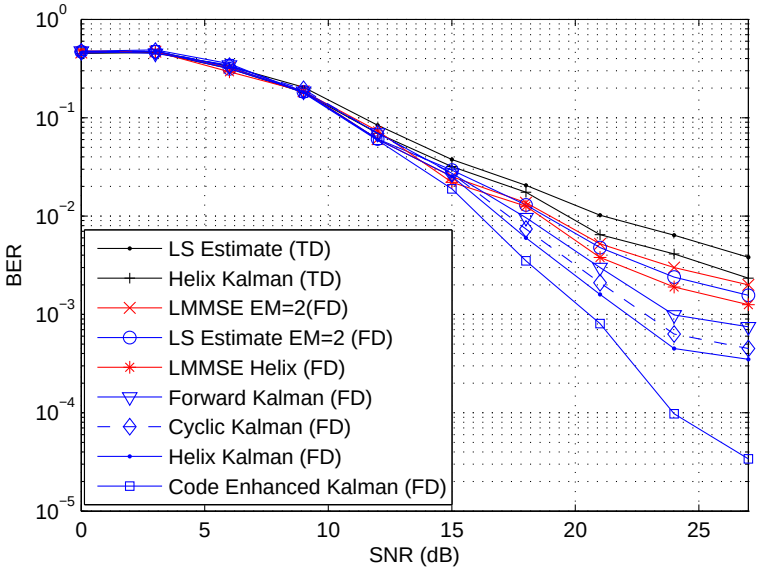
Figures 5(a) and 5(b) compare the proposed methods for a practical channel model. We test our algorithms for the ITU vehB channel model (this model is recommended by WiMAX forum). The rest of the system parameters are the same as specified above. Figure 5(a) shows the BER comparison of uncoded system, while Figure 5(b) compares the same methods with $\frac{1}{2}$ rate convolutional coding. The proposed method offers up to 2.5 dB gain for the uncoded case and 1 dB in the coded case over the time domain methods.

8 Conclusion

We present a receiver design based on a semi-blind low complexity frequency domain channel estimation algorithm for multi-access OFDM systems. Opposed to the time domain case which estimates the whole spectrum, we propose an eigenvalue based reduced parameter frequency domain approach in which the user estimates the part of the spectrum in which he operates. The advantage of this is reduction in computational cost incurred by each user. The receiver uses pilots to kick start the estimation process and then iterates between channel and data recovery. The proposed receiver utilizes data (finite alphabet set, code, transmit precoding, pilots) and channel (finite delay spread, frequency correlation, time correlation) constraints. Thanks to the decoupled relation in the frequency domain, data recovery is done on an element by element basis while channel estimation boils down to solving a regularized least squares

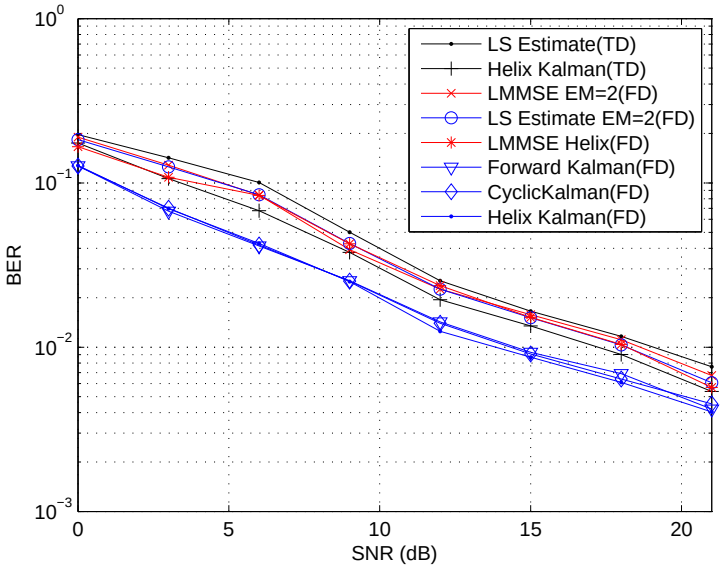


(a) BER comparison when channel taps lie on the FFT grid (no correlation).

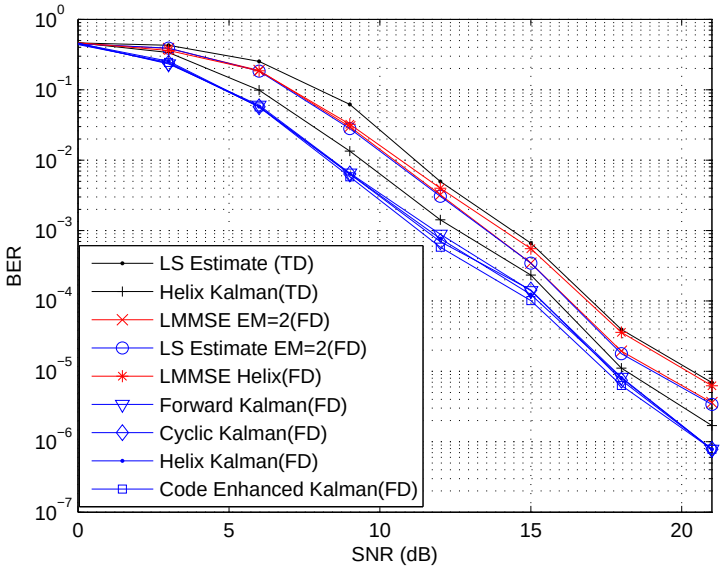


(b) BER comparison with non integer delay profile.

Figure 4: BER comparison with outer code (FD = Frequency Domain, TD = Time Domain).



(a) Uncoded case.



(b) Coded case.

Figure 5: BER comparison of proposed methods in ITU vehB channel (FD = Frequency Domain, TD = Time Domain).

problem. We propose to improve the estimate making use of the time correlation information of the channel by relaxing the latency requirement. For this purpose, we employ Cyclic and Helix based FB Kalman filters and use the outer code to enhance the channel estimate. We make use of both the frequency and time correlation which results in a relatively low training overhead and comparatively low computational cost than the standard LMMSE estimator. The simulation results show the performance of our algorithm is better than the standard LMMSE estimator and equivalent time domain technique.

A Appendix

To calculate an expectation of the form $E[\underline{\boldsymbol{D}}\boldsymbol{B}\underline{\boldsymbol{D}}]$, which appears in (17), we note that by our assumption different elements of $\underline{\boldsymbol{D}}$ are independent, making the expectation that involves them in $E[\underline{\boldsymbol{D}}\boldsymbol{B}\underline{\boldsymbol{D}}]$ separable, i.e. for these terms, we have

$$E[\underline{\boldsymbol{D}}\boldsymbol{B}\underline{\boldsymbol{D}}^*] = E[\underline{\boldsymbol{D}}]\boldsymbol{B}E[\underline{\boldsymbol{D}}^*]. \tag{43}$$

The identical forms, however, interact according to

$$E[\underline{\boldsymbol{D}}\boldsymbol{B}\underline{\boldsymbol{D}}] = E[\underline{\boldsymbol{D}}\mathrm{diag}(\boldsymbol{B})\underline{\boldsymbol{D}}] = E[\underline{\boldsymbol{D}}\boldsymbol{D}^*]E[\mathrm{diag}(\boldsymbol{B})]. \tag{44}$$

By combining (43) and (44), we see that

$$E[\underline{\boldsymbol{D}}\boldsymbol{B}\underline{\boldsymbol{D}}^*] = E[\underline{\boldsymbol{D}}]\boldsymbol{B}E[\underline{\boldsymbol{D}}^*] + \mathrm{Cov}[\underline{\boldsymbol{D}}]\mathrm{diag}(\boldsymbol{B}). \tag{45}$$

References

[1] M. Speth, S. A. Fechtel, G. Fock, and H. Meyr, Optimum receiver design for wireless broadband systems using OFDM-Part 1, IEEE Transaction on Communications, 47 (11) (1999) 1668-1677.

[2] European Telecommunications Standards Institute, ETS 300 744: Digital Video Broadcasting (DVB-T).

[3] IEEE Std 802.11a/D7.0-1999, Part11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: High Speed Physical Layer in the 5GHz Band.

[4] ETSI, Broadband Radio Access Networks (BRAN): HIPERLAN type 2 technical specification Physical (PHY) layer, August 1999. DTS/BRAN-0023003 V0.k.

[5] W. Zhendao and G. B. Giannakis, Wireless multicarrier communications, IEEE Signal Processing Magazine, 17 (3) (2000) 29-48.

[6] J. H. Lee, J. C. Han and S. C. Kim,Joint carrier frequency synchronization and channel estimation for OFDM systems via the EM algorithm, IEEE Transactions on Vehicular Technology, 55 (1) (2006) 167-172.

[7] X. Hou, Z. Zhang, H. Kayama, Time-Domain MMSE Channel Estimation Based on Subspace Tracking for OFDM Systems, IEEE 63rd Vehicular Technology Conference, Melbourne, Vic, 7-10 May 2006, pp. 1565-1569.

[8] T. Y. Al-Naffouri, An EM-Based Forward-Backward Kalman Filter for the Estimation of Time-Variant Channels in OFDM, IEEE Transaction on Signal Processing, 1 (11) (2007) 1-7.

[9] X. Wang, Y. Wu, J.Y. Chouinard, and H. C. Wu,On the design and performance analysis of multisymbol encapsulated OFDM systems, IEEE Transactions on Vehicular Technology, 55 (3) (2006) 990-1002.

[10] J. Liu and J. Li,Parameter estimation and error reduction for OFDM-based WLANs, IEEE Transactions on Mobile Computing, 3 (2) (2004) 152-163.

[11] S. Qin, P. Liu, X. Zhang, L. Zheng and D. Wang, Channel estimation with timing offset based on PSD & LS estimation for wireless OFDM systems, International Symposium on Intelligent Signal Processing and Communication Systems, 28 Nov. - 1 Dec. 2007, pp 248-251.

[12] A. Jeremic, T.A. Thomas, A. Nehorai,OFDM channel estimation in the presence of interference, IEEE Transactions on Signal Processing, 52 (12) (2004) 3429-3439.

[13] R. Dinis, N. Souto, J. Silva, A. Kumar and A. Correia,Joint Detection and Channel Estimation for OFDM Signals with Implicit Pilots, IEEE Mobile and Wireless Communications Summit, Hungary, July 2007, pp. 1-5.

[14] T. Roman, M. Enescu, and V. Koivunen, Joint time-domain tracking of channel and frequency offsets for MIMO OFDM systems, Wireless Personal Communications, 31 (3-4) (2004) 181-200.

[15] T. Nyblom, T. Roman, M. Enescu, and V. Koivunen, Time-varying carrier offset tracking in OFDM systems using particle filtering, Proc. of the Fourth IEEE International Symposium on Signal Processing and Information Technology, Rome, Italy, December 2004, pp. 217-220.

[16] C.H. Lim and D. S. Han, Robust LS channel estimation with phase rotation for single frequency network in OFDM, IEEE Transaction on Consumer Electronics, 52 (4) (2006) 1173-1178.

[17] P.Y. Tsai and T. D. Chiueh, Frequency-Domain Interpolation-Based Channel Estimation in Pilot Aided OFDM Systems, IEEE 59th Vehicular Technology Conference, 17-19 May 2004, pp. 420-424.

[18] J. Jiang, T. Tang, Y. Zhang and P. Zhang, A channel Estimation Algorithm for OFDM based on PCC training Symbols and Frequency Domian Windowing, International Symposium on Communications and Information Technologies, Bangkok, Thailand, Oct. 18 2006-Sept. 20 2006, pp. 629-632.

[19] D. Marelli and M. Fu, A subband approach to channel estimation and equalization for DMT and OFDM systems, IEEE Transaction on Communications, 53 (11) (2005) 1850-1858.

[20] O. Edfors, M. Sandell, J. van de Beek, K. S. Wilson, and P. O. Börjesson, OFDM channel estimation by singular value decomposition, IEEE Transaction on Signal Processing, 46 (7) (1998) 931-939.

[21] O. Simeone, Y. Bar-Ness, U. Spagnolini, Pilot-based channel estimation for OFDM systems by tracking the delay-subspace, IEEE Transaction on Wireless Communications, 3 (1) (2004) 315-325.

[22] A. Dowler, A. Nix, J. McGeehan, Data-derived iterative channel estimation with channel tracking for a mobile fourth generation wide area OFDM system, Procedings of IEEE Global Telecommunications Conference, 1-5 Dec. 2003 pp. 804-808.

[23] Y. Xie and C.N. Georghiades, Two EM-type channel estimation algorithm for OFDM with transmitter diversity, IEEE Transaction on Communications, 51 (1) (2003) 106 - 115.

[24] Y. Xie and C.N. Georghiades, Two EM-type channel estimation algorithm for OFDM with transmitter diversity, IEEE Transaction on Communications, 51 (1) (2003) 106 - 115.

[25] B. Lu, X. Wang and Y. Li, Iterative receivers for space-time block-coded OFDM systems in dispersive fading channels, IEEE Transacion on Wireless Communications, 1 (2) (2002) 213 - 225.

[26] M.O. Pun, M. Morelli and C.-C. Kuo, Iterative Detection and Frequency Synchronization for Generalized OFDMA Uplink Transmissions, IEEE Transaction on Wireless Communications, 5 (12) (2007) 629-639.

[27] Daum Fred, Jim Huang, Mysterious computational complexity of particle filter, SPIE proceedings series, Signal and data processing of small targets 2002. Orlando FL, ETATS-UNIS 2002, vol. 4728, pp. 418-426.

[28] R. V. Nee and R. Prasad, OFDM for wireless multimedia communications, Artech house, Boston, MA, 2000.

[29] T. Kailath, A. H. Sayed and B. Hassibi, Linear estimation, Prentice Hall, 2000.

[30] Y. Li, L.J. Cimini and N.R. Sollenberger, Robust channel estimation for OFDM systems with rapid dispersive fading channels, IEEE Transaction on Communications, 46 (1) (1998) 902-915.