

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

T. Y. Al-Naffouri^{a,*}, M. S. Sohail^b

^a*Electrical Engineering Department, University of Southern California, CA 90089*

^b*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 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:* tareq@stanfordalumni.org (T. Y. Al-Naffouri),
saqibsohail@gmail.com (M. S. Sohail)

1 Introduction

Orthogonal frequency division multiplexing (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], [2], [3], [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 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.

The aim of this work is to perform channel estimation in the frequency domain in multiple access OFDM. Before introducing this work, let us look at previous approaches to channel estimation. Numerous research contributions have appeared in literature on the topic of channel estimation in recent years. One way to classify these works is according to whether they performed estimation in the time domain or the frequency domain.

A lot of researchers have opted for 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 LS estimation for OFDM systems with timing offsets while [12] used implicit pilots for joint detection and channel estimation. 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

[13].

Various techniques for channel estimation in the frequency domain have also been explored in the past years. Authors of [14] apply phase shifted samples in the frequency-domain to an interpolated LS to estimate the channel while [15] proposed to include a phase rotation term in the frequency domain interpolation for better channel impulse response (CIR) window location. The authors of [16] proposed a channel estimation using polynomial cancelation coding (PCC) training symbols and frequency domain windowing. The authors presented a sub-band approach to channel estimation and channel equalization is proposed in [17] while a low-complexity iterative channel estimator is proposed in [18]. The minimum mean square error (MMSE) channel estimation in the frequency domain is considered in [19] while authors of [20] explored delay subspace-based channel estimation techniques for OFDM systems over fast-fading channels.

1.1 Disadvantage of performing channel estimation in time domain

Most channel estimation techniques estimate the channel in the time domain. The main reason is that channel length is usually (much) less than cyclic prefix length and so the number of parameters to estimate in the time domain is much less than the number of parameters in the frequency domain leading to higher estimation accuracy. Another advantage of time domain based estimation is that it allows us to make a collective use of underlying communication structure; a structure induced by the constraints on the channel (time and frequency correlation) and the data (cyclic prefix, finite alphabet constraint and coding). The collective use of these constraints in turn allows us to reduce the training overhead [8].

Apart from that, time domain based estimate estimate is plagued by several disadvantages. By performing the estimation in the time domain, we loose the diagonal structure of the channel. Thus, instead of the diagonal frequency domain relationship¹ (see equation (4) further ahead)

$$\mathbf{y}_i = \text{diag}(\mathbf{x}_i)\mathbf{h}_i + \mathcal{N}_i \tag{1}$$

¹The system model and input/output equations are developed further ahead in Section 2. We use some of these equations here to motivate our work.

in which the various equations are decoupled, we employ the time-frequency relationship

$$\mathbf{y}_i = \text{diag}(\mathbf{x}_i)\mathbf{Q}_{P+1}\mathbf{h}_i + \mathcal{N}_i = \mathbf{X}_i\mathbf{h}_i + \mathcal{N}_i \quad (2)$$

which is no more diagonal (decoupled). This loss in transparency in return complicates channel estimation and makes it more computationally complex. For example, while the estimation in (1) is performed on a bin by bin basis according to

$$\hat{\mathcal{H}}_i(l) = \frac{\mathcal{Y}_i(l)}{\mathcal{X}_i(l)} \quad l = 1, 2, \dots, N$$

channel estimation in (2) requires size $L + 1$ matrix inversion

$$\hat{\mathbf{h}}_i = (\mathbf{X}_i^* \mathbf{X}_i)^{-1} \mathbf{X}_i^* \mathbf{y}_i$$

Moreover, since data detection is best performed in the frequency domain, estimating the channel in the time domain makes it necessary to perform an extra IFFT operation (to obtain the frequency domain estimate $\hat{\mathcal{H}}_i$ from the time domain estimate $\hat{\mathbf{h}}_i$ and use it for data detection). Thus, for data-aided channel estimation techniques, each channel estimation step requires one such IFFT operation.

Apart from the computational complexity, performing channel estimation in the time domain might be oversolving a problem. For example, in multiple access OFDM systems, like WiMAX, users are not interested in the whole frequency spectrum, but only that part of the spectrum in which they are operating. Moreover, even if some users were interested in estimating the whole spectrum, many standards would not be able support that as there are not enough pilots to do so.

1.2 Can we perform channel estimation reliably in the frequency domain?

The only problem with channel estimation in the frequency domain is the increase in the number of parameter to be estimated [19]. If we can reduce the parameter estimation space, then we can avoid the one disadvantage of frequency domain estimation as compared to time domain

estimation. The frequency response of the channel is inherently limited by the degrees of freedom of the time domain impulse response. How does this limited degree of freedom manifests itself in the frequency domain? Figure 1 demonstrates the length 64 frequency response resulting from a 16 tap channel with exponential decay profile similar to the one we employ in our simulations. The figure shows that within a narrow enough band, the spectrum looks linear or quadratic.

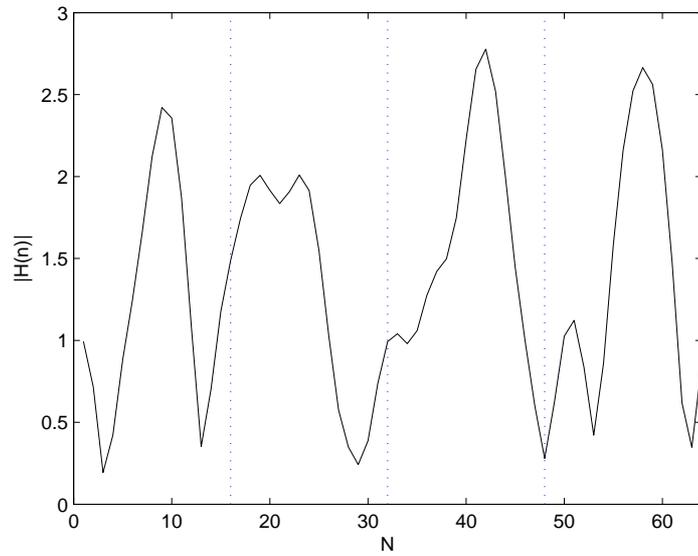


Figure 1: Channel Impulse Response in the Frequency Domain partitioned in four subchannels.

Thus the number of parameters needed to represent the CIR can be reduced. In Section 4, we will show how to utilize this property to represent the CIR using the dominant eigenvalues.

This paper is organized as follows. Section 2 describes the OFDM system model. The time domain channel estimator is reviewed in Section 3. Section 4 presents a new frequency domain data aided parameter reduction model for channel estimation. Section 5 introduces time correlation information to improve the receiver design. Section 7 discusses the simulation results and Section 8 provides the concluding remarks.

2 System Model

Figure 2 shows the block diagram of the system under consideration. The data bits, to be sent over the OFDM 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 [21] (further discussion on pilot design is provided in Section 7.4). We will use calligraphic notations (e.g., \mathcal{X}) for vectors in the frequency domain. Now consider a sequence of $T+1$ such OFDM symbols $\mathcal{X}_0, \mathcal{X}_1, \dots, \mathcal{X}_T$ to be transmitted. Each

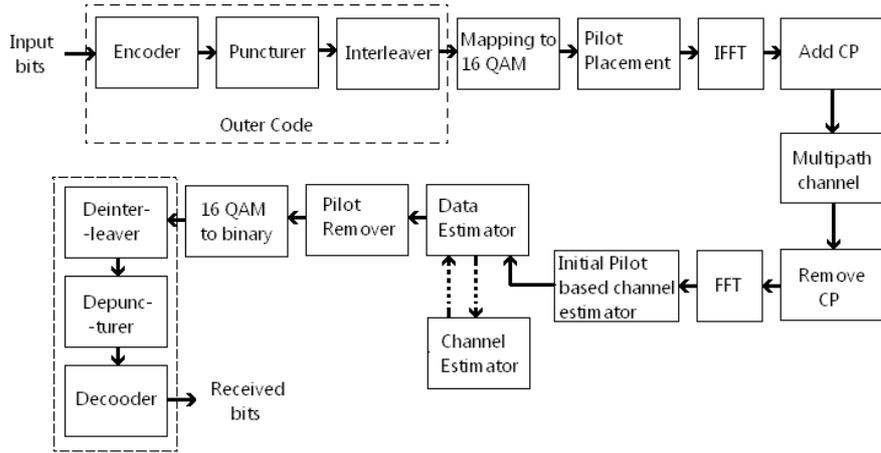


Figure 2: Block Diagram

symbol \mathcal{X}_i (of length N), undergoes an IFFT operation to produce the time domain symbol $\mathbf{x}_i = \sqrt{N}\mathbf{Q}^*\mathcal{X}_i$, where \mathbf{Q} is the $N \times N$ discrete Fourier transform (DFT) matrix given by $\mathbf{Q} = [e^{-j\frac{2\pi}{N}(l-1)(m-1)}]$ and the operator $*$ denotes conjugate transpose. The transmitter then appends a cyclic prefix (CP) of length P and transmits the resulting super symbol (the CP serves to mitigate the multi-path effect but the estimation of channel characteristics of fading channels

require densely spaced pilot tones specially for those channels with a small coherence bandwidth [14]). The channel \mathbf{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

$$\mathbf{h}_{i+1} = \mathbf{F}\mathbf{h}_i + \mathbf{G}\mathbf{u}_i \quad \mathbf{h}_0 \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_n), \quad \mathbf{u}_i \sim \mathcal{N}(\mathbf{0}, \sigma_u^2 \mathbf{I}), \quad (3)$$

The matrices \mathbf{F} and \mathbf{G} are a function of the doppler spread, the power delay profile and the transmit filter, and thus the model in (3) captures both frequency and time correlation (for details on the construction of these matrices, and a justification of this model, see [8]). The subscript i shows the time dependence of the variables. At the channel output and after stripping the CP, we obtain the time domain symbol \mathbf{y}_i . The input/output relationship of the OFDM system is best described in the frequency domain as

$$\mathbf{y}_i = \text{diag}(\boldsymbol{\mathcal{X}}_i)\boldsymbol{\mathcal{H}}_i + \mathcal{N}_i = \text{diag}(\boldsymbol{\mathcal{X}}_i)\mathbf{Q}_{P+1}\mathbf{h}_i + \mathcal{N}_i \quad (4)$$

where \mathbf{y}_i and $\boldsymbol{\mathcal{H}}_i$ are the length- N FFT's of \mathbf{y}_i and \mathbf{h}_i , respectively, and \mathcal{N}_i is the additive white Gaussian noise $\mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I})$. \mathbf{Q}_{P+1} is the matrix which contains the first $P + 1$ columns of \mathbf{Q} . The second line (4) follows from the FFT relationship

$$\boldsymbol{\mathcal{H}}_i = \mathbf{Q} \begin{bmatrix} \mathbf{h}_i \\ \mathbf{0} \end{bmatrix} = \mathbf{Q}_{P+1}\mathbf{h}_i \quad (5)$$

Alternatively, with $\mathbf{X}_i \triangleq \text{diag}(\boldsymbol{\mathcal{X}}_i)\mathbf{Q}_{P+1}$ we can write

$$\mathbf{y}_i = \mathbf{X}_i\mathbf{h}_i + \mathcal{N}_i \quad (6)$$

3 Channel Estimation in Time Domain: MMSE Estimator

When the channel obeys the I/O relationship (6), we estimate \mathbf{h}_i by maximizing the log-likelihood function

$$\hat{\mathbf{h}}_i^{\text{MAP}} = \arg \max_{\mathbf{h}_i} \{\ln p(\mathbf{y}_i | \mathbf{X}_i, \mathbf{h}_i) + \ln p(\mathbf{h}_i)\} \quad (7)$$

The first term on the right hand side in the above equation is given by $\ln p(\mathbf{y}_i | \mathbf{X}_i, \mathbf{h}_i) = -\|\mathbf{y}_i - \mathbf{X}_i \mathbf{h}_i\|_{\sigma^{-2}}^2$. Assuming \mathbf{h}_i to be $\mathcal{N}(\mathbf{0}, \mathbf{R}_h)$, the second term on the right side of equation (7) is given by $\ln p(\mathbf{h}_i) = -\|\mathbf{h}_i\|_{\mathbf{R}_h^{-1}}^2$. The MAP estimate is then given by²

$$\hat{\mathbf{h}}_i^{\text{MAP}} = \arg \min_{\mathbf{h}_i} \left\{ \|\mathbf{y}_i - \mathbf{X}_i \mathbf{h}_i\|_{\sigma^{-2}}^2 + \|\mathbf{h}_i\|_{\mathbf{R}_h^{-1}}^2 \right\} \quad (8)$$

As \mathbf{X}_i is not completely known at the receiver, we can use (6) to obtain a pilot/output equation that can be used for initial channel estimation. Let the index set $I_p = \{i_1, i_2, \dots, i_{L_p}\}$ denote the pilot locations within the OFDM symbol known *a priori* at the receiver. Also, let \mathbf{X}_{I_p} denote the matrix \mathbf{X} pruned of the rows that do not belong to I_p . Then, the pilot/output equation can be derived from the I/O relationship (6) as

$$\mathbf{y}_{I_p} = \mathbf{X}_{I_p} \mathbf{h} + \mathcal{N}_{I_p} \quad (9)$$

where we removed the time dependence for notational convenience. Deriving the MAP estimator for the above equation we get

$$\hat{\mathbf{h}} = \mathbf{R}_h \mathbf{X}_{I_p}^* [\sigma^2 \mathbf{I} + \mathbf{X}_{I_p} \mathbf{R}_h \mathbf{X}_{I_p}^*]^{-1} \mathbf{y}_{I_p} \quad (10)$$

where \mathbf{R}_h is the autocorrelation matrix of \mathbf{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.

4 Channel Estimation in Frequency Domain

4.1 A Parameter Reduction Approach

In this section we introduce the frequency domain based channel estimation algorithm. Our starting point is to partition the frequency response into a number of sections each of length L_f producing a total of $\lceil \frac{N}{L_f} \rceil$ sections³. Let the j^{th} section of the frequency response be denoted by

²We use the weighted norm $\|\mathbf{h}\|_{\Sigma}^2$ to denote $\mathbf{h}^* \Sigma \mathbf{h}$.

³In a multi-access OFDM system, we can choose the section length to be the number of carriers allocated to each user. However, the sections need not have equal length over the frequency response.

$\underline{\mathcal{H}}_i^{(j)}$. Then, from (4), the input/output equation that involves this section is given by

$$\underline{\mathbf{y}}_i^{(j)} = \text{diag}(\underline{\mathbf{x}}_i^{(j)})\underline{\mathcal{H}}_i^{(j)} + \underline{\mathcal{N}}_i^{(j)} \quad (11)$$

where $\underline{\mathbf{y}}_i^{(j)}$, $\underline{\mathbf{x}}_i^{(j)}$, $\underline{\mathcal{H}}_i^{(j)}$ and $\underline{\mathcal{N}}_i^{(j)}$ are the j^{th} sections of \mathbf{y}_i , \mathbf{x}_i , \mathcal{H}_i and \mathcal{N}_i , respectively. Now let $I_p^{(j)}$ denote the pilot locations within this section, then the pilot/output equations corresponding to (11) are given by

$$\underline{\mathbf{y}}_{I_p} = \text{diag}(\underline{\mathbf{x}}_{I_p})\underline{\mathcal{H}} + \underline{\mathcal{N}}_{I_p} \quad (12)$$

where in (12) and thereafter we suppress the dependence on the section index j and on the time index i for notational convenience. Here \mathbf{A}_{I_p} denotes the matrix \mathbf{A} pruned of the rows that don't belong to I_p .

Obviously, the pilots are not enough to estimate the elements of $\underline{\mathcal{H}}$. So we resort to model reduction starting from the autocorrelation function of $\underline{\mathcal{H}}$, $\mathbf{R}_{\underline{\mathcal{H}}}$ (where $\mathbf{R}_{\underline{\mathcal{H}}}$ is the FFT of \mathbf{R}_h). To this end, consider the eigenvalue decomposition of $\mathbf{R}_{\underline{\mathcal{H}}}$,

$$\mathbf{R}_{\underline{\mathcal{H}}} = \sum_{l=1}^{L_f} \lambda_l \mathbf{v}_l \mathbf{v}_l^T$$

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

$$\underline{\mathcal{H}} = \sum_{l=1}^{L_f} \alpha_l \mathbf{v}_l$$

where $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \dots, \alpha_{L_f}]^T$ is a parameter vector, to be estimated, with zero mean and autocorrelation matrix $\boldsymbol{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_{L_f})$. We now represent $\underline{\mathcal{H}}$ using the dominant eigenvalues and treat the rest as modeling noise⁴, i.e.

$$\underline{\mathcal{H}} = \mathbf{V}_d \boldsymbol{\alpha}_d + \mathbf{V}_n \boldsymbol{\alpha}_n \quad (13)$$

⁴The 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 λ_j 's. In our simulations, we use the condition $\frac{\lambda_j}{\lambda_{j+1}} > 5$ to place our cutoff.

Upon substituting (13) in (11), we obtain

$$\underline{\mathbf{y}} = \text{diag}(\underline{\mathbf{X}})\mathbf{V}_d\boldsymbol{\alpha}_d + \underline{\mathcal{N}} + \text{diag}(\underline{\mathbf{X}})\mathbf{V}_n\boldsymbol{\alpha}_n = \underline{\mathbf{X}}_d\boldsymbol{\alpha}_d + \underline{\mathcal{N}}' \quad (14)$$

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

$$\mathbf{R}_{\underline{\mathcal{N}}'} = \mathbf{R}_{\underline{\mathcal{N}}} + \text{diag}(\underline{\mathbf{X}})\mathbf{V}_n\text{diag}(\lambda_n)\mathbf{V}_n^*\text{diag}(\underline{\mathbf{X}})^* \quad (15)$$

Now equation (14) can be used to construct a pilot/output equation, similar to (12), as

$$\underline{\mathbf{y}}_{I_p} = \underline{\mathbf{X}}_{d,I_p}\boldsymbol{\alpha}_d + \underline{\mathcal{N}}'_{I_p} \quad (16)$$

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

$$\hat{\boldsymbol{\alpha}}_d^{\text{MAP}} = \arg \max_{\boldsymbol{\alpha}_d} \left\{ \ln p(\underline{\mathbf{y}}_{I_p} | \underline{\mathbf{X}}_{d,I_p}, \boldsymbol{\alpha}_d) + \ln p(\boldsymbol{\alpha}_d) \right\} \quad (17)$$

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

$$\hat{\boldsymbol{\alpha}}_d^{\text{MAP}} = \arg \min_{\boldsymbol{\alpha}_d} \left\{ \|\underline{\mathbf{y}}_{I_p} - \underline{\mathbf{X}}_{d,I_p}\boldsymbol{\alpha}_d\|_{\mathbf{R}_{\underline{\mathcal{N}}'}^{-1}}^2 + \|\boldsymbol{\alpha}_d\|_{\boldsymbol{\Lambda}_d^{-1}}^2 \right\} \quad (18)$$

which simplifies to

$$\hat{\boldsymbol{\alpha}}_d = \boldsymbol{\Lambda}_d \underline{\mathbf{X}}_{d,I_p}^* \left[\mathbf{R}_{\underline{\mathcal{N}}'} + \underline{\mathbf{X}}_{d,I_p} \boldsymbol{\Lambda}_d \underline{\mathbf{X}}_{d,I_p}^* \right]^{-1} \underline{\mathbf{y}}_{I_p} \quad (19)$$

The resulting mean square error is given by

$$\mathbf{R}_e = \left[\boldsymbol{\Lambda}_d^{-1} + \underline{\mathbf{X}}_{I_p}^* \mathbf{R}_{\underline{\mathcal{N}}'}^{-1} \underline{\mathbf{X}}_{I_p} \right]^{-1} \quad (20)$$

The estimate of the j^{th} section of the spectrum is then given by $\hat{\underline{\mathcal{H}}} = \mathbf{V}_d \hat{\boldsymbol{\alpha}}_d$. The concatenation of all $\lceil \frac{N}{L_f} \rceil$ sections produces the frequency domain based estimate of the frequency response $\hat{\underline{\mathcal{H}}}$.

4.2 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 which we discuss next.

4.2.1 The Maximization Step

In the previous subsection we find $\hat{\boldsymbol{\alpha}}_d$ by maximizing the log likelihood function given by equation (17). Since the input \mathcal{X} (and hence $\underline{\mathbf{X}}_d$) is not observable, we can employ the EM algorithm and instead of maximizing (17) 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 k^{th} iteration given by

$$\hat{\boldsymbol{\alpha}}_d^{(k)} = \arg \max_{\boldsymbol{\alpha}_d} \left\{ E_{\mathcal{X}_i | \mathcal{Y}_i, \hat{\boldsymbol{\alpha}}_d^{(k-1)}} \ln p(\underline{\mathbf{Y}}_{I_p} | \underline{\mathbf{X}}_{d, I_p}, \boldsymbol{\alpha}_d) + \ln p(\boldsymbol{\alpha}_d) \right\} \quad (21)$$

which simplifies to ⁵

$$\hat{\boldsymbol{\alpha}}_d^{\text{MAP}} = \arg \min_{\boldsymbol{\alpha}_d} \left\{ E \left\| \underline{\mathbf{Y}}_{I_p} - \underline{\mathbf{X}}_{d, I_p} \boldsymbol{\alpha}_d \right\|_{\mathbf{R}_{\mathcal{N}'}}^{-1}^2 + \|\boldsymbol{\alpha}_d\|_{\boldsymbol{\Lambda}_d^{-1}}^2 \right\} \quad (22)$$

Strictly speaking, the noise correlation $\mathbf{R}_{\mathcal{N}'}$ is itself dependent on the input due to the modeling noise (see equation (15)). Hence in performing the expectation in (22), we need to take this into account. Treating the general case is difficult, so we consider the following three cases for $\mathbf{R}_{\mathcal{N}'}$:

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

This happens when we ignore the modeling noise so that $\mathbf{R}_{\mathcal{N}'} = \sigma^2 \mathbf{I}$ where the expectation in (22) is taken with respect to $\underline{\mathbf{X}}_d$ given $\underline{\mathbf{Y}}$ and the most recent estimate $\boldsymbol{\alpha}_d$. In this case $\mathbf{R}_{\mathcal{N}'}$ becomes independent of $\underline{\mathbf{X}}_d$ and it would then be straight forward to carry the expectation in (22). Specifically, upon completing the squares, (22) can be equivalently written as

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

⁵the expectation is taken with respect to the input given the output and the most recent estimate $\hat{\boldsymbol{\alpha}}_d^{k-1}$. This information is understood and dropped for notational convenience.

which can be simplified to

$$\hat{\alpha}_d^{\text{MAP}} = \arg \min_{\alpha_d} \|\underline{\mathbf{y}} - E[\underline{\mathbf{X}}_d] \alpha_d\|_{\frac{1}{\sigma_n^2} \mathbf{I}}^2 + \|\alpha_d\|_{\frac{1}{\sigma_n^2} \text{Cov}[\underline{\mathbf{X}}_d^*]}^2 + \|\alpha_d\|_{\Lambda_d^{-1}}^2 \quad (23)$$

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

Instead of ignoring the modeling noise, we can split the expectation in (22) into an expectation over $\mathbf{R}_{\mathcal{N}'}$ and an *independent* expectation taken over the rest of the terms i.e., we can approximate (22) as

$$\hat{\alpha}_d^{\text{MAP}} = \arg \min_{\alpha_d} \left\{ E \left\| \underline{\mathbf{y}}_{I_p} - \underline{\mathbf{X}}_{d,I_p} \alpha_d \right\|_{E[\mathbf{R}_{\mathcal{N}'}]^{-1}}^2 + \|\alpha_d\|_{\Lambda_d^{-1}}^2 \right\} \quad (24)$$

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

$$E[\mathbf{R}_{\mathcal{N}'}] = \sigma^2 \mathbf{I} + E[\text{diag}(\underline{\mathbf{X}}) \mathbf{V}_n \mathbf{\Lambda}_n \mathbf{V}_n^* \text{diag}(\underline{\mathbf{X}}^*)] \quad (25)$$

We show in Appendix A that this expectation can be expressed as

$$E[\mathbf{R}_{\mathcal{N}'}] = \sigma^2 \mathbf{I} + E[\underline{\mathbf{D}}] \mathbf{V}_n \mathbf{\Lambda}_n \mathbf{V}_n^* E[\underline{\mathbf{D}}^*] + \text{Cov}[\underline{\mathbf{D}}] \text{diag}(\mathbf{V}_n \mathbf{\Lambda}_n \mathbf{V}_n^*)$$

where $\underline{\mathbf{D}} = \text{diag}(\underline{\mathbf{X}})$ and where $\text{diag}(\mathbf{V}_n \mathbf{\Lambda}_n \mathbf{V}_n^*)$ is a diagonal matrix whose diagonal coincides with the diagonal of the matrix $\mathbf{V}_n \mathbf{\Lambda}_n \mathbf{V}_n^*$. The now averaged $\mathbf{R}_{\mathcal{N}'}$ does not depend on $\underline{\mathbf{X}}$ any more. Replacing $\mathbf{R}_{\mathcal{N}'}$ by its expectation, it is then straight forward to carry the expectation in (24) which comes out to be

$$\hat{\alpha}_d^{\text{MAP}} = \arg \min_{\alpha_d} \|\underline{\mathbf{y}} - E[\underline{\mathbf{X}}_d] \alpha_d\|_{E[\mathbf{R}_{\mathcal{N}'}]^{-1}}^2 + \|\alpha_d\|_{\text{Cov}[\underline{\mathbf{D}}] \text{diag}(\mathbf{V}_n \mathbf{\Lambda}_n \mathbf{V}_n^*)}^2 + \|\alpha_d\|_{\Lambda_d^{-1}}^2 \quad (26)$$

Case 3: \mathbf{X} is constant modulus:

In the constant modulus case, it is possible to evaluate (22) exactly. Specifically, and starting from the expression for the autocorrelation as $\mathbf{R}_{\mathcal{N}'} = \sigma^2 \mathbf{I} + \underline{\mathbf{D}} \mathbf{V}_n \mathbf{\Lambda}_n \mathbf{V}_n^* \underline{\mathbf{D}}^*$ and we can write

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

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

$$\begin{aligned} \underline{\mathbf{X}}_d^* \mathbf{R}_{\underline{N}'}^{-1} &= \mathbf{V}_d^* \underline{\mathbf{D}}^* \mathbf{R}_{\underline{N}}^{-1} = \mathbf{V}_d^* \mathbf{R}_{\underline{N}''}^{-1} \underline{\mathbf{D}}^{-1} \\ \mathbf{R}_{\underline{N}'}^{-1} \underline{\mathbf{X}}_d &= \mathbf{D}^{-1*} \mathbf{R}_{\underline{N}''}^{-1} \mathbf{V}_d \end{aligned}$$

and

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

Thus, in the constant modulus case, (22) can be equivalently written as

$$\begin{aligned} \hat{\alpha}_d^{(j)} = \arg \min_{\alpha_d} & \mathbf{y}^* E[\mathbf{D}^{-1*}] \mathbf{R}_{\underline{N}''}^{-1} E[\mathbf{D}^{-1}] \mathbf{y} - \mathbf{y}^* E[\mathbf{D}^{-1*}] \mathbf{R}_{\underline{N}''}^{-1} \mathbf{V}_d \alpha_d \\ & - \alpha_d^* \mathbf{V}_d^* \mathbf{R}_{\underline{N}''}^{-1} E[\mathbf{D}^{-1}] \mathbf{y} + \alpha_d^* \mathbf{V}_d^* \mathbf{R}_{\underline{N}''}^{-1} \mathbf{V}_d \alpha_d + \alpha_d^* \mathbf{\Lambda}_d^{-1} \alpha_d \end{aligned}$$

which upon simplification becomes

$$\hat{\alpha}_d^{\text{MAP}} = \arg \min_{\alpha_d} \|E[\mathbf{D}^{-1}] \mathbf{y} - \mathbf{V}_d \alpha_d\|_{\mathbf{R}_{\underline{N}''}^{-1}}^2 + \|\alpha_d\|_{\mathbf{\Lambda}_d^{-1}}^2 \quad (27)$$

In the simulations further ahead, we compare the approximate solutions (23) and (26) with the exact EM solution (27) for a constant modulus input. Simulations show that replacing $\mathbf{R}_{\underline{N}'}$ with its expectation is almost as good as calculating the expectation exactly.

4.2.2 The Expectation Step

As we have seen above, the maximization step assumes the presence of some expectations. By inspecting subsection 4.2.1, we see we need to calculate the following moments.

$$E[\underline{\mathbf{X}}_d], \text{Cov}[\underline{\mathbf{X}}_d^*], E[\underline{\mathbf{D}}], E[\underline{\mathbf{D}} \underline{\mathbf{D}} \underline{\mathbf{D}}^*], \text{ and } E[\underline{\mathbf{D}}^{-1}] \quad (28)$$

Now as $\underline{\mathbf{X}}_d = \text{diag}(\underline{\mathbf{x}}) \mathbf{V}_d = \underline{\mathbf{D}} \mathbf{V}_d$ we can see that we can express the moments of $\underline{\mathbf{X}}_d$ in terms of moments of $\underline{\mathbf{D}}$. Specifically we have that

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

and

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

Moreover, we show in appendix A that

$$E[\underline{\mathbf{D}} \underline{\mathbf{B}} \underline{\mathbf{D}}^*] = E[\underline{\mathbf{D}}] \underline{\mathbf{B}} E[\underline{\mathbf{D}}^*] + Cov[\underline{\mathbf{D}}] \text{diag}(\underline{\mathbf{B}}) \quad (29)$$

From above it follows that in order to calculate the expectations in (28), it is enough to calculate the following three moments

$$E[\text{diag}(\underline{\mathcal{X}})], Cov[\text{diag}(\underline{\mathcal{X}})] \ \& \ E[\text{diag}(\underline{\mathcal{X}})^{-1}] \quad (30)$$

where the expectation is performed given the output $\underline{\mathcal{Y}}$ and the most recent channel estimate $\hat{\underline{\mathcal{H}}}$. In carrying out these expectations, we will assume that the elements of $\underline{\mathcal{X}}$ are independent.⁶ With this in mind, it is easy to see that we can evaluate the moments in (30) and hence in (28) by calculating $E[\underline{\mathcal{X}}(l)]$, $Cov[\underline{\mathcal{X}}(l)] = E[|\underline{\mathcal{X}}(l)|^2] - |E[\underline{\mathcal{X}}(l)]|^2$, $E[\frac{1}{\underline{\mathcal{X}}(l)}]$. Now assuming that $\underline{\mathcal{X}}(l)$ is drawn from the alphabet $A = \{A_1, \dots, A_M\}$ with equal probability, it can be shown that [8]

$$E[\underline{\mathcal{X}}(l) | \underline{\mathcal{Y}}(l), \underline{\mathcal{H}}(l)] = \frac{\sum_{j=1}^M A_j e^{-\frac{|\underline{\mathcal{Y}}(l) - \underline{\mathcal{H}}(l) A_j|^2}{\sigma^2}}}{\sum_{j=1}^M e^{-\frac{|\underline{\mathcal{Y}}(l) - \underline{\mathcal{H}}(l) A_j|^2}{\sigma^2}}} \quad (31)$$

$$E[|\underline{\mathcal{X}}(l)|^2 | \underline{\mathcal{Y}}(l), \underline{\mathcal{H}}(l)] = \frac{\sum_{j=1}^M |A_j|^2 e^{-\frac{|\underline{\mathcal{Y}}(l) - \underline{\mathcal{H}}(l) A_j|^2}{\sigma^2}}}{\sum_{j=1}^M e^{-\frac{|\underline{\mathcal{Y}}(l) - \underline{\mathcal{H}}(l) A_j|^2}{\sigma^2}}} \quad (32)$$

$$E[\frac{1}{\underline{\mathcal{X}}(l)} | \underline{\mathcal{Y}}(l), \underline{\mathcal{H}}(l)] = \frac{\sum_{j=1}^M \frac{1}{A_j} e^{-\frac{|\underline{\mathcal{Y}}(l) - \underline{\mathcal{H}}(l) A_j|^2}{\sigma^2}}}{\sum_{j=1}^M e^{-\frac{|\underline{\mathcal{Y}}(l) - \underline{\mathcal{H}}(l) A_j|^2}{\sigma^2}}} \quad (33)$$

⁶This is in general not true because the elements of $\underline{\mathcal{H}}$ are not independent (as the elements of $\underline{\mathbf{h}}$ are the Fourier transform of the impulse response \mathbf{h}). However, we continue to use this approximation as this maintains the transparency of element-by-element equalization in OFDM.

4.2.3 Summary of the EM Algorithm

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

1. Calculate the initial channel estimate $\hat{\mathcal{H}}_0$ using pilots (18).
2. Calculate the moments of the input given the current channel estimate $\hat{\mathcal{H}}_i$ and the output \mathcal{Y} using equations (31)-(33).
3. Calculate the channel estimate using either one of the methods (23), (26) or (27) outlined in Section 4.2.
4. Iterate between step 2 and 3.

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

5 Using Time-Correlation to Improve the Channel Estimate

The receiver developed in the previous section performs channel estimation symbol by symbol. In other words, the channel is block fading and hence is totally independent from symbol to symbol. 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 α_d . To this end, let's first develop a model for the time variation of the parameter α_d .

5.1 Developing a Frequency Domain Time-Variant Model

Consider the block fading model in (3) and let's assume for simplicity that the diagonal matrices \mathbf{F} and \mathbf{G} are actually scalar multiples of the identity, i.e. $\mathbf{F} = f\mathbf{I}$ and $\mathbf{G} = \sqrt{1 - f^2}\mathbf{I}$ where f is a function of Doppler frequency (see [8]). We will use the time domain model in (3) to derive a similar model for α . To this end, recall that $\mathcal{H}_i = \mathbf{Q}_{P+1}\mathbf{h}_i$. Thus, the j^{th} section of \mathcal{H}_i , $\mathcal{H}_i^{(j)}$,

is related to \mathbf{h}_i by

$$\underline{\mathbf{H}}_i^{(j)} = \mathbf{Q}_{P+1}^{(j)} \mathbf{h}_i \quad (34)$$

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

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

or

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

where \mathbf{V}_d^+ is the pseudo inverse of \mathbf{V}_d . Multiplying both sides of (3) by $\mathbf{V}_d^+ \mathbf{Q}_{P+1}^{(j)}$ yields a dynamical recursion for $\boldsymbol{\alpha}_d$

$$\boldsymbol{\alpha}_{d,i+1} = \mathbf{F}_\alpha \boldsymbol{\alpha}_{d,i} + \mathbf{G}_\alpha \mathbf{u}_i \quad (35)$$

where $\mathbf{F}_\alpha = f\mathbf{I}$ and $\mathbf{G}_\alpha = \sqrt{1-f^2} \mathbf{V}_d^+ \mathbf{Q}_{P+1}^{(j)}$ and where $E[\boldsymbol{\alpha}_{d,0} \boldsymbol{\alpha}_{d,0}^*] = \boldsymbol{\Lambda}_d$. Note that the dependence of \mathbf{G}_α and $\boldsymbol{\alpha}_d$ on j has been suppressed for notational convenience. We are now ready to implement the EM algorithm to the frequency domain system governed by the dynamical equation (35). As we have seen in section 4.2, the algorithm will consist of an initial estimation step, a maximization step, and an expectation step.

5.2 Initial (Pilot-Based) Channel Estimation

In the initial channel estimation step, the frequency domain system is described by equations (16) and (35), reproduced here for convenience.

$$\underline{\mathbf{y}}_{I_p,i} = \underline{\mathbf{X}}_{d,I_p,i} \boldsymbol{\alpha}_{d,i} + \underline{\mathcal{N}}'_{I_p,i} \quad (36)$$

$$\boldsymbol{\alpha}_{d,i+1} = \mathbf{F}_\alpha \boldsymbol{\alpha}_{d,i} + \mathbf{G}_\alpha \mathbf{u}_i \quad (37)$$

Now given a sequence $i = 0, 1, \dots, T$ 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 (36)-(37)(see [22]), i.e., by implementing the following equations

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

$$\mathbf{R}_{e,i} = \mathbf{R}_{\underline{\mathcal{N}}'} + \underline{\mathbf{X}}_{d,I_p,i} \mathbf{P}_{i|i-1} \underline{\mathbf{X}}_{d,I_p,i}^* \quad (38)$$

$$\mathbf{K}_{f,i} = \mathbf{P}_{i|i-1} \underline{\mathbf{X}}_{d,I_p,i}^* \mathbf{R}_{e,i}^{-1} \quad (39)$$

$$\hat{\boldsymbol{\alpha}}_{i|i} = \left(\mathbf{I} - \mathbf{K}_{f,i} \underline{\mathbf{X}}_{d,I_p,i} \right) \hat{\boldsymbol{\alpha}}_{i|i-1} + \mathbf{K}_{f,i} \boldsymbol{\mathcal{Y}}_i \quad (40)$$

$$\hat{\boldsymbol{\alpha}}_{i+1|i} = \mathbf{F}_\alpha \hat{\boldsymbol{\alpha}}_{i|i} \quad (41)$$

$$\mathbf{P}_{i+1|i} = \mathbf{F}_\alpha \left(\mathbf{P}_{i|i-1} - \mathbf{K}_{f,i} \mathbf{R}_{e,i} \mathbf{K}_{f,i}^* \right) \mathbf{F}_\alpha^* + \frac{1}{\sigma_n^2} \mathbf{G}_\alpha \mathbf{G}_\alpha^* \quad (42)$$

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

$$\boldsymbol{\lambda}_{i|T} = \left(\mathbf{I}_{P+N} - \underline{\mathbf{X}}_{d,I_p,i}^* \mathbf{K}_{f,i} \right) \mathbf{F}_i^* \boldsymbol{\lambda}_{i+1|T} + \underline{\mathbf{X}}_{d,I_p,i} \mathbf{R}_{e,i}^{-1} \left(\boldsymbol{\mathcal{Y}}_i - \underline{\mathbf{X}}_{d,I_p,i} \hat{\boldsymbol{\alpha}}_{i|i-1} \right) \quad (43)$$

$$\hat{\boldsymbol{\alpha}}_{i|T} = \hat{\boldsymbol{\alpha}}_{i|i-1} + \mathbf{P}_{i|i-1} \boldsymbol{\lambda}_{i|T} \quad (44)$$

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

5.3 Iterative (Data-Aided) Channel Estimation

For this part, we use the whole data symbol and not just the pilot part. Thus, in this case our system is described by equations (14) and (35) also reproduced here for convenience

$$\boldsymbol{\mathcal{Y}}_i = \underline{\mathbf{X}}_{d,i} \boldsymbol{\alpha}_{i,d} + \underline{\mathcal{N}}'_i \quad (45)$$

$$\boldsymbol{\alpha}_{d,i+1} = \mathbf{F}_\alpha \boldsymbol{\alpha}_{d,i} + \mathbf{G}_\alpha \mathbf{u}_i \quad (46)$$

If the data symbols $\underline{\mathbf{X}}_{d,i}$ were known, we would have employed the forward-backward Kalman-Filter (38)-(44) on the above state-space model. Since the input is not available, we replace it by its estimate along an expectation maximization algorithm. 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

$$\underline{\mathbf{y}}_i = \begin{bmatrix} E[\underline{\mathbf{X}}_{d,i}] \\ Cov[\underline{\mathbf{X}}_{d,i}^*]^{\frac{1}{2}} \end{bmatrix} \boldsymbol{\alpha}_{i,d} + \begin{bmatrix} \underline{\mathcal{N}}'_i \\ \mathbf{0} \end{bmatrix} \quad (47)$$

$$\boldsymbol{\alpha}_{d,i+1} = \mathbf{F}_\alpha \boldsymbol{\alpha}_{d,i} + \mathbf{G}_\alpha \mathbf{u}_i \quad (48)$$

where the expectations in (47) are taken given the output $\underline{\mathbf{y}}_i$ and most recent channel estimate $\boldsymbol{\alpha}_{d,i}$. The expectations that appears in (47) are calculated as we did in Section 4.2.2. In contrast to the symbol by symbol EM algorithm of section 4.2, there are several ways of implementing the EM iterations in the time-correlated multi-symbol case. In the symbol by symbol algorithm of Section 4, there was one dimension to iterate against (channel estimation vs data detection). When the channels are time correlated over several OFDM symbols as is the case here, there are two dimensions we can iterate against:

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. We discuss two such receivers here, the Cyclic and the Helix Kalman based receivers.

5.4 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.

5.5 Helix based FB Kalman

The Helix based FB Kalman is a more general 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 using the Kalman to the next OFDM symbol. The iterations in this case draw a helix shape, hence the name.

5.6 Using Code to Enhance the Estimate

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 simulation shows that invoking the code can have a profound effect on performance.

5.7 Forward Kalman Filter

One drawback of the FB Kalman implementation 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 data estimate.

6 Time Domain multiple access channel estimation

For fair comparison, we need to compare the frequency domain (LS and Kalman) receiver with the time domain counterpart. 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 j^{th} user is given by (see (11)) $\underline{\mathbf{y}}_i^{(j)} = \text{diag}(\underline{\mathbf{x}}_i^{(j)})\underline{\mathbf{h}}_i^{(j)} + \underline{\mathbf{N}}_i^{(j)}$. Now $\underline{\mathbf{h}}_i^{(j)}$ is related to the impulse response by (see (34)) $\underline{\mathbf{h}}_i^{(j)} = \mathbf{Q}_{P+1}^{(j)}\mathbf{h}_i$ where as described in Section 5.1, $\mathbf{Q}_{P+1}^{(j)}$ is \mathbf{Q}_{P+1} pruned of all rows that don't belong to the j^{th} section. So, we can write

$$\underline{\mathbf{y}}_i^{(j)} = \text{diag}(\underline{\mathbf{x}}_i^{(j)})\mathbf{Q}_{P+1}^{(j)}\mathbf{h}_i + \underline{\mathbf{N}}_i^{(j)} \quad (49)$$

Equation (49) 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 (3) $\mathbf{h}_{i+1} = \mathbf{F}\mathbf{h}_i + \mathbf{G}\mathbf{u}_i$ we can implement the various kind of Kalman filters discussed in the previous section for *time-domain* channel estimation. It is important to note that the computational complexity involved in the time domain case is much higher than in the Eigen estimate as the significant eigenvalues α_d are less than the channel length.

7 Simulation Results

We consider an OFDM system that transmits 6 symbols with 64 carriers and a cyclic prefix of length $P = 15$ each with a time variation of $f = 0.9$. The data bits are mapped to 16 QAM through Gray coding (except for Figures 3(a) and 3(b) which use a 4 QAM). The OFDM symbol serves 4 users each occupying 16 frequency bins. In addition, the OFDM symbol carries 16 or 24 pilots equally divided between the users. The channel impulse response consists of 15 complex taps (the maximum length possible). It has an exponential delay profile $E[|h_0(k)|^2] = e^{-0.2k}$ and remains fixed over any OFDM symbol. Where specified, an outer code is used to provide robustness. The outer code is 1/2 rate convolutional code. In what follows, we compare the performance of frequency domain based channel estimation using various techniques for the both the coded and uncoded cases. We also benchmark our method with the time domain method briefly described in Section 6 (see [8] also).

7.1 Effect of Modeling Noise

Figures 3(a) and 3(b) show the MSE and BER curves for the three cases considered in section 4.2 comparing the various treatment of the noise. We plot the Figures 3(a) and 3(b) for constant modulus using 16 pilots. As evident from the graphs, the inclusion of the modeling noise improves the result. We also note that the expectation of the noise and the exact solution have almost comparable results.

7.2 EM based Least Squares

In order to see a fair comparison between the time domain and the frequency domain techniques for a multiple access system, we compare the time domain LS estimate with the frequency domain LS and LS with EM estimate. Figures 4(a) show the MSE while Figures 4(b) show the BER performance for these methods for the uncoded case at 16 pilots.

7.3 Kalman Filter based Receivers

Figure 5(a) compares the BER performance of frequency domain Forward Kalman, Cyclic and Helical Kalman filters with the time domain LS method and Helix Kalman for the uncoded case at 16 pilots. 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 for the uncoded 16 pilot case, the frequency domain methods fairs better than the time domain methods.

Figure 5(b) shows the same comparison for 24 pilots uncoded case. For the case of 24 pilots, we note that though the time domain estimate methods perform better than frequency domain methods, the performance of the frequency domain Helix Kalman is comparable to the time domain Helix Kalman.

Figure 6(a) compares the BER performances of frequency domain channel estimation of various Kalman filters with the LS and LS EM estimate for the 16 pilot case. Here we utilize

the outercode to enhance the estimate. We see that the code enhancement technique is superior to the rest of the techniques.

Figure 6(b) shows the result of the comparison of frequency domain Helix Kalman and coded Kalman with the time domain Helix Kalman (16 pilots). We can see that for the multiple access case, the frequency domain technique fairs better than the time domain estimation method, while the coded Kalman outperforms all other techniques. In order to see a fair comparison between the time domain and the frequency domain techniques for a multiple access system, we compare the frequency domain Helix Kalman with the time domain Helix Kalman obtained from the procedure outlined in Section 6.

7.4 Pilot Design

From Figure 4(b), it can be established that pilot density has a profound effect on the channel estimation algorithm. It will be worthwhile to investigate the effect of pilot pattern on the channel estimation algorithm as well. Here we find the optimal pilot pattern that minimizes the MSE of the estimate for the pilot placement, given by (20), in the frequency domain. Consider the case of 16 pilots, with an OFDM symbol of length 64. Considering 4 users, each user will have access to 16 frequency bins. Assuming the pilots be equally divided among all the users, the spectrum available to every user will have 4 pilots each. This means there are a total of $C_4^{16} = 1820$ different combination of pilot patters that are possible. We perform an exhaustive search and find that the minimum MSE occurs at equispaced combinations (Figure 7(a) and Figure 7(b)).

Now we use either one of these two equispaced pilot patterns as the piloting scheme for all OFDM symbols (called the non alternating scheme) or we can use both of these patterns such that each alternate OFDM symbol has the same pilot pattern (called the alternating scheme). A comparison of BER of uncoded equidistance non alternating and alternating schemes is shown in Figure 7(c), while Figure 7(d) shows the same comparison for the coded case.

8 Conclusion

We present an OFDM receiver design based on a semi-blind low complexity frequency domain channel estimation algorithm for multi-access OFDM system. Opposed to the time domain case which estimates the whole spectrum, we propose a 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. Also, the user might not have access to the entire spectrum. We estimate the channel parameters based on the eigenvalue technique, greatly reducing the number of parameters to be estimated. The receiver uses the pilots to kick start the estimation process and then iterates between channel and data recovery. Our 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 the channel estimation boils down to solving a regularized least squares 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. The simulation results show the performance of our algorithm. Our results maybe extended to multiple antenna OFDM systems.

A Appendix

Now to calculate an expectation of the form $E[\underline{D}\underline{B}\underline{D}^*]$, which appears in (25), we note that by our assumption different elements of \underline{D} are independent making the expectation that involves them in $E[\underline{D}\underline{B}\underline{D}^*]$ separable, i.e. for these terms, we have

$$E[\underline{D}\underline{B}\underline{D}^*] = E[\underline{D}]\underline{B}E[\underline{D}^*] \quad (50)$$

The identical forms, however, interact according to

$$E[\underline{\mathbf{D}}\underline{\mathbf{B}}\underline{\mathbf{D}}] = E[\underline{\mathbf{D}}\text{diag}(\mathbf{B})\underline{\mathbf{D}}] = E[\underline{\mathbf{D}}\underline{\mathbf{D}}^*]E[\text{diag}(\mathbf{B})] \quad (51)$$

By combining (50) and (51), we see that

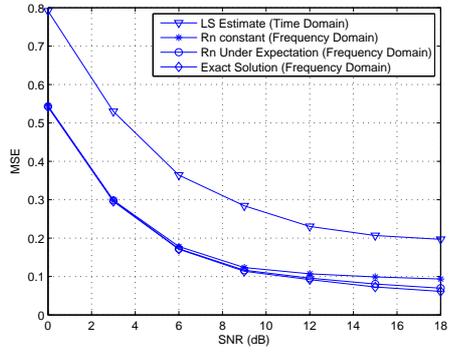
$$E[\underline{\mathbf{D}}\underline{\mathbf{B}}\underline{\mathbf{D}}^*] = E[\underline{\mathbf{D}}\mathbf{B}E[\underline{\mathbf{D}}^*] + \text{Cov}[\underline{\mathbf{D}}]\text{diag}(\mathbf{B})] \quad (52)$$

References

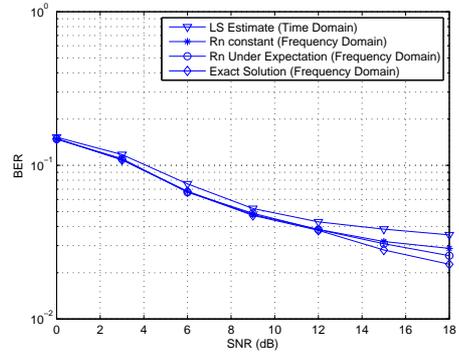
- [1] M. Speth, S. A. Fechtel, G. Fock, and H. Meyr, Optimum receiver design for wireless broadband systems using OFDM-Part 1, *IEEE Trans. Commun.*, vol. 47, no. 11, pp. 1668-1677, Nov. 1999.
- [2] European Telecommunications Standards Institute, ETS 300 744: Digital Video Broadcasting (DVB-T).
- [3] 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*, vol. 17, no. 3, pp. 29-48, May 2000.
- [6] Jong-Ho Lee, Jae Choong Han and Seong-Cheol Kim, Joint carrier frequency synchronization and channel estimation for OFDM systems via the EM algorithm, *IEEE Transactions on Vehicular Technology*, Volume 55, Issue 1, Jan. 2006 Page(s):167 - 172
- [7] Xiaolin Hou, Zhan Zhang, Kayama, H., Time-Domain MMSE Channel Estimation Based on Subspace Tracking for OFDM Systems, *Vehicular Technology Conference, 2006. VTC 2006-Spring. IEEE 63rd Volume 4*, 7-10 May 2006 Page(s):1565 - 1569

- [8] T. Y. Al-Naffouri, An EM-Based Forward-Backward Kalman Filter for the Estimation of Time-Variant Channels in OFDM, *IEEE Trans. Signal Processing*, vol. 55, Jul. 2007.
- [9] Xianbin Wang, Wu, Y., Chouinard, J.-Y. and Hsiao-Chun Wu, On the design and performance analysis of multisymbol encapsulated OFDM systems, *IEEE Transactions on Vehicular Technology*, Volume 55, Issue 3, May 2006 Page(s):990 - 1002.
- [10] Jianhua Liu, Jian Li, Parameter estimation and error reduction for OFDM-based WLANs, *Mobile Computing*, *IEEE Transactions on* Volume 3, Issue 2, April-June 2004 Page(s):152 - 163
- [11] Shengping Qin, Peide Liu, Xin Zhang, Laibo Zheng and Deqiang Wang, Channel estimation with timing offset based on PSD & LS estimation for wireless OFDM systems, *International Symposium on Intelligent Signal Processing and Communication Systems*, 2007. page(s): 248-251 *ISPACS 2007*.
- [12] 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, Vol. 1 pp. 1-5, July 2007.
- [13] Jeremic, A., Thomas, T.A., Nehorai, A., OFDM channel estimation in the presence of interference, *IEEE Transactions on Signal Processing*, Volume 52, Issue 12, Dec. 2004 Page(s):3429 - 3439
- [14] Chae-Hyun Lim, Dong Seog Han, Robust LS channel estimation with phase rotation for single frequency network in OFDM, *IEEE Transactions on Consumer Electronics*, Volume 52, Issue 4, Nov. 2006 Page(s):1173 - 1178
- [15] P.Yun Tsai, T. Dar Chiueh, Frequency-Domain Interpolation-Based Channel Estimation in Pilot Aided OFDM Systems, *IEEE 59th VTC*, vol 1, pp. 420-424, 2004.

- [16] J. Jiang, Tian Tang, Yongjin Zhang, Ping Zhang, A channel Estimation Algorithm for OFDM based on PCC training Symbols and Frequency Domian Windowing, IEEE ISGIT 2006, pp. 629-632, 2006.
- [17] Damini Marelli and Minyue Fu, A subband approach to channel estimation and equalization for DMT and OFDM systems, IEEE Transactions on Communications, Vol. 53, No. 11, November 2005.
- [18] Dowler, A., Nix, A., McGeehan, J., Data-derived iterative channel estimation with channel tracking for a mobile fourth generation wide area OFDM system, IEEE GLOBECOM 03, Global Telecommunications Conference, 2003. , Volume 2, 1-5 Dec. 2003 Page(s):804 - 808 Vol.2
- [19] O. Edfors, M. Sandell, J. van de Beek, K. S. Wilson, and P. O. Brjesson, OFDM channel estimation by singular value decomposition, IEEE Trans. Signal Proc., vol. 46, no. 7, pp. 931-939, Jul. 1998.
- [20] Simeone, O., Bar-Ness, Y., Spagnolini, U., Pilot-based channel estimation for OFDM systems by tracking the delay-subspace, IEEE Transactions on Wireless Communications, Volume 3, Issue 1, Jan. 2004 Page(s):315 - 325
- [21] R. V. Nee and R. Prasad, OFDM for wireless multimedia communications, Boston, MA: Artech house, 2000.
- [22] T. Kailath, A. H. Sayed and B. Hassibi, Linear estimation, Prentice Hall, 2000.

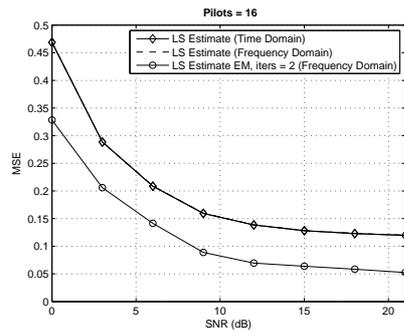


(a) MSE

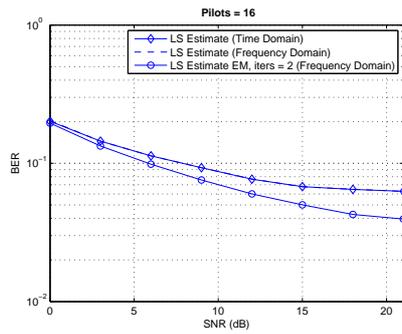


(b) BER

Figure 3: Effect of modeling noise.

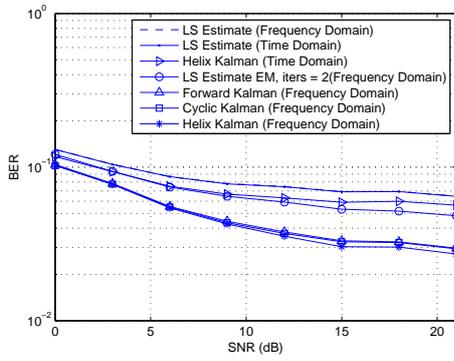


(a) MSE (16 pilots)

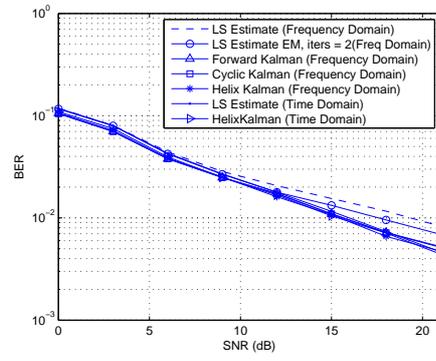


(b) BER (16 pilots)

Figure 4: EM based Least Squares comparison .

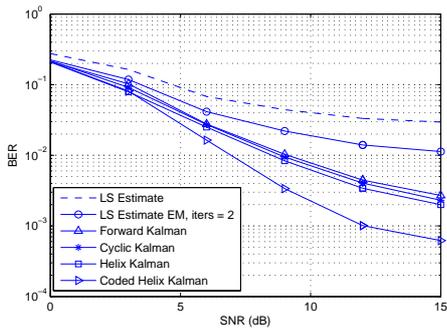


(a) 16 pilots

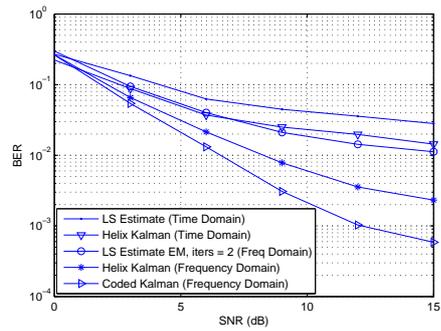


(b) 24 pilots

Figure 5: BER comparison for various uncoded frequency domain methods.



(a) Frequency domain coded methods

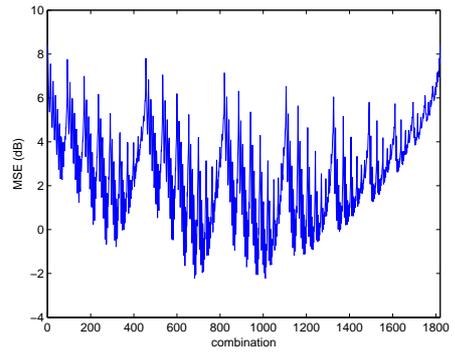


(b) Frequency and Time domain methods

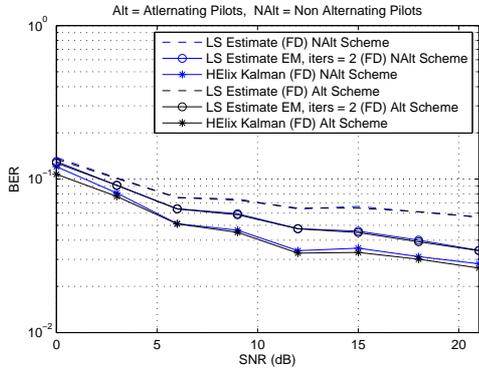
Figure 6: BER comparison for coded frequency and time domain methods (16 pilots).



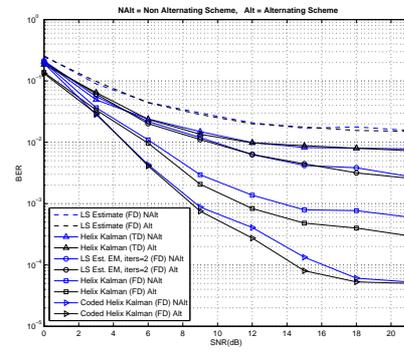
(a) Pilot patterns



(b) MSE for 16 pilots



(c) Uncoded 16 pilots



(d) Coded 16 pilots

Figure 7: (a) Pilot patterns with minimum MSE (b) MSE for all pilot patterns. Comparison of alternating and non alternating pilot schemes for (c) Uncoded case (d) Coded case.