Frequency Domain Estimation of Time Variant Channels in OFDM

Principal Investigator: Dr. Tareq Y. Al-Naffouri
Department of Electrical Engineering
KFUPM, Dhahran 31261
Saudi Arabia
e-mail: naffouri@kfupm.edu.sa

Junior Project JF060003
Final Report

Contents

Abstract

OFDM modulation combines the advantages of high achievable rates and relatively easy implementation. However, for proper recovery of the input, the OFDM receiver needs accurate channel information. Most algorithms proposed in literature perform channel estimation in the time domain. This increases the computational complexity especially for data aided algorithms and does not lend itself to multiaccess situations where the user is only interested in part of the spectrum. In this final project report, we propose a frequency domain algorithm for channel estimation and tracking in OFDM. The algorithm is based on interpolating the spectrum and tracking the interpolation parameters instead of tracking the actual response. In this project we propose two interpolation techniques. The first is a polynomial interpolation method (e.g., using piece-wise linear or quadratic interpolation) which proved to be quite inadequate for channel tracking in a wireless environment. The second tracking method is based on eigen-decomposition of the channel correlation and proves to be superior to former tracking technique and to channel estimation techniques. A data-aided forward-backward Kalman filter is also implemented to to enhance the performance of the channel estimation algorithm.

Chapter 1

Introduction to Channel Estimation

With the advent of the modern digital communication age, demands on the data transmission rates have exceeded several Mbps and will continue to grow in the foreseeable future as the telecommunication industry continues to offer more sophisticated and advanced services. Orthogonal frequency division multiplexing (OFDM) is a technology that promises to meet these transmission demands. Since the last decade, OFDM has attracted considerable attention. 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 [1]. OFDM effectively divides a wide band frequency selective fading channel into a large number of narrow band flat fading channels over which parallel data streams are transmitted thereby increasing the symbol duration. The insertion of a cyclic prefix (CP), of adequate length, in the transmission symbol reduces the inter symbol interference (ISI). The CP, which is a cyclic extension of the IFFT output, has to be at least as long as the cannel impulse response (CIR) in order to avoid ISI. This also enables the OFDM system to have 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. It is already a part of many digital communication standards and is being used the world over. OFDM has been selected as the physical layer of choice for broadband wireless communications systems ([1], [2], [3], [4], [5]).

The aim of this project is to explore techniques frequency-domain techniques for channel estimation and equalization techniques in multiple access OFDM. To pave the way for this, we introduce in this first chapter the OFDM system model in a time-variant environment and perform extensive literature review of available techniques (blind, semi-blind, training-based, and data-aided) for channel estimation. We also survey techniques used for channel estimation in the time and frequency domains. We then contrast the two method of time and frequency domain channel estimation and argue that multiple access OFDM lends itself better to frequency domain channel estimation.

In Chapter 2 we introduce an interpolation method for channel estimation in the frequency domain and use least-squares to perform channel estimation. Our simulations show that this technique is not suitable for recovering the channel in highly time-variant environments. In Chapter 3, we introduce a new method based on retaining the dominant eigenvectors of the channel covariance matrix. We also use a data-aided forward-backward Kalman filter to track the dominant eigenvalues of the channel. This techniques shows favorable behavior as compared with channel estimation

1.1 System Model

Consider a sequence of $T+1$ data symbols $\boldsymbol{\mathcal{X}}_0, \boldsymbol{\mathcal{X}}_1, \cdots, \boldsymbol{\mathcal{X}}_T$, each of length N, to be transmitted in an OFDM system. Every symbol $\boldsymbol{\mathcal{X}}_i$, undergoes an IFFT operation to produce the time domain symbol

$$\boldsymbol{x}_i = \sqrt{N} Q^* \boldsymbol{\mathcal{X}}_i \tag{1.1}$$

where Q is the $N \times N$ FFT matrix. In order to counter the effect of ISI, a length P CP $\underline{\boldsymbol{x}}_i$ is appended to the symbol $\boldsymbol{x}_i$, which results in the super symbol $\overline{\boldsymbol{x}}_i$, each of length $N+P$. 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 [31]. Figure 1.1 shows the basic elements of an OFDM transmitter.

Let $\underline{\boldsymbol{h}}_i$ be the channel of maximum length $P+1$. We consider a block fading model and assume that the channel remains unchanged for each super-symbol but varies from one super-symbol to the

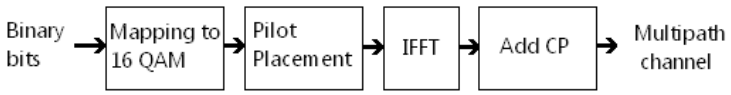


Figure 1.1: OFDM Transmitter Block Diagram.

next according to the following state space model.

$$\underline{\boldsymbol{h}}_{i+1} = \boldsymbol{F}\underline{\boldsymbol{h}}_{\boldsymbol{i}} + \boldsymbol{G}\underline{\boldsymbol{u}}_{\boldsymbol{i}} \tag{1.2}$$

where $\underline{\boldsymbol{h}}_o \sim \boldsymbol{\mathcal{N}}(0, \Pi_o)$ and $\underline{u}_o \sim \boldsymbol{\mathcal{N}}(0, \sigma_u^2)$. the matrices $\boldsymbol{F}$ and $\boldsymbol{G}$ are a function of the doppler spread, the power delay profile (frequency correlation), and the transmit filter. The matrices are given as

$$\boldsymbol{F} = \begin{bmatrix} \alpha(0) & & \\ & \ddots & \\ & & \alpha(P) \end{bmatrix} \text{ and } \boldsymbol{G} = \begin{bmatrix} \sqrt{1-\alpha^2(0)} & & \\ & \ddots & \\ & & \sqrt{(1-\alpha^2(P))e^{-\beta P}} \end{bmatrix}$$

$\alpha(p)$ is related to the Doppler frequency $f_D(p)$ by $\alpha(p) = J_0(2\pi f_D T(p))$. The variable β corresponds to the exponent of the channel decay profile while the factor $\sqrt{(1-\alpha^2(p))e^{-\beta p}}$ ensures that each link maintains the exponential decay profile ($e^{-\beta p}$) for all time. We assume this information in known at the receiver. The model thus captures both frequency and time correlation.

The passage of $\overline{\boldsymbol{x}}_k$ symbols through the channel $\underline{\boldsymbol{h}}$, produces the received sequence $\overline{\boldsymbol{y}}_k$ at the receiver. The received packet (of length $N+P$) is split into a length N packet $\boldsymbol{y}_k$ and a length P prefix $\underline{\boldsymbol{y}}_k$. The prefix absorbs all the ISI present between the $\overline{\boldsymbol{x}}_{k-1}$ and $\overline{\boldsymbol{x}}_k$ packets and is hence discarded. The time domain relation of the input and the output can be expressed as

$$\boldsymbol{y}_i = \boldsymbol{x}_i \otimes \boldsymbol{h}_i + \boldsymbol{n}_i \tag{1.3}$$

Equation (1.3) takes a more transparent form in the frequency domian as

$$\boldsymbol{\mathcal{Y}}_i = \mathrm{diag}(\boldsymbol{\mathcal{X}}_{\mathrm{i}})\boldsymbol{\mathcal{H}}_{\mathrm{i}} + \boldsymbol{\mathcal{N}}_{\mathrm{i}} \tag{1.4}$$

or

$$\boldsymbol{\mathcal{Y}}_i = \mathrm{diag}(\boldsymbol{\mathcal{X}}_{\mathrm{i}})\boldsymbol{Q}_{\mathrm{P+1}}\underline{\boldsymbol{h}}_{\mathrm{i}} + \boldsymbol{\mathcal{N}}_{\mathrm{i}} \tag{1.5}$$

The relationship in (1.5) follows from the FFT relationship

$$\boldsymbol{\mathcal{H}}_i = \boldsymbol{Q} \begin{bmatrix} \underline{\boldsymbol{h}}_i \\ 0 \end{bmatrix} = \boldsymbol{Q}_{P+1} \underline{\boldsymbol{h}}_i \tag{1.6}$$

where $\boldsymbol{Q}_{P+1}$ consists of the first $P+1$ columns of $\boldsymbol{Q}$. Alternatively, with

$$\boldsymbol{X}_i \triangleq \text{diag}(\boldsymbol{\mathcal{X}}_\text{i}) \boldsymbol{Q}_{\text{P+1}} \tag{1.7}$$

we can write

$$\boldsymbol{\mathcal{Y}}_i = \boldsymbol{X}_i \underline{\boldsymbol{h}}_i + \boldsymbol{\mathcal{N}}_i \tag{1.8}$$

which is no longer diagonal. We will discuss the disadvantage of this decoupled relationship in this part.

1.2 Literature Review

As mentioned in the introduction, our aim is to design an algorithm for channel estimation in OFDM. In this section, we will take a look at the literature relating to channel estimation in OFDM systems. We will provide an overview of the various approaches to channel estimation and the different constraints assumed on channel and data.

The availability of an accurate channel transfer function estimate is one of the prerequisites for coherent symbol detection in an OFDM receiver. Numerous research contributions have appeared in literature on the topic of channel estimation, in recent years. One way to classify these works is as according to the method used for channel estimation (training based, semi blind, blind and data aided). Another approach to classify these algorithms is based on the constraints used for channel and data recovery.

1.2.1 Channel Estimation using Pilots

One technique for channel estimation is to use pilots. As equalization requires channel state information (CSI), pilots on predetermined subcarriers are sent as training signals in OFDM systems, and

the channels for pilot subcarriers are directly estimated, while those for non pilot subcarriers need to be estimated through interpolation with the channel estimates from adjacent pilot subcarriers [6], [7], [8], [9], [10], [11]. This in turn, is achieved at the cost of a reduction in the number of useful subcarriers available for data transmission. In [12], the authors have developed a channel estimator by introducing an extended channel and its finite impulse response approximation.

1.2.2 Blind Channel Estimation

Since the number of pilots must be greater than the number of channel taps, the use of cyclic prefix (CP) and pilot symbols entails a significant bandwidth loss, motivating blind methods. Several works have attempted to perform blind channel estimation in OFDM. The authors in [13] explored transmitter redundancy for blind channel estimation while in [14], a blind identification exploiting receiver diversity which can get CSI during one OFDM symbol was investigated. In [16] the authors present a fast converging blind channel estimator for OFDM-systems based on the Maximum Likelihood principle. A non redundant precoding along with cyclic prefix was explored in [17]. In [18], second-order cyclostationary statistics and antenna precoding are used while [19] employs finite-alphabet constraint for blind channel estimation. The authors in [20] suggest an approach which relies on the i.i.d. assumption of the data sequence and uses the cyclic prefix redundancy present in OFDM systems and [22] developed a posteriori probability based two dimensional channel estimation algorithm.

1.2.3 Semi Blind Channel Estimation

In semi blind methods, both the pilots and natural constraints are used for channel estimation([23], [24]). In [25] a semi-blind channel estimation using receiver diversity is proposed for OFDM systems in the presence of virtual carriers. The authors in [26] employed a semiblind channel estimation method using selected channel parameter estimation and error reduction algorithms. The work presented in [45] proposes a pilot aided algorithm for frequency domain channel estimation for a single-user and multiple receiving antennas system in the presence of synchronous interference while

the authors in [27] used delay sub-space based approach for channel estimation. In [28], coding along with pilots was used for channel estimation. Similarly other papers have explored various other semi-blind techniques for channel estimation. Coding and cyclic prefix was investigated for channel estimation in [30]. Authors of [31] used interpolated LS by applying phase shifted samples while [32] proposed to include a phase rotation term in the frequency domain interpolation.

1.2.4 Data Aided Channel Estimation

The motivation behind estimating the channel response is to recover the data being transmitted. The detected data can be, in turn, used to improve the channel estimate, thus giving rise to an iterative method for channel and data recovery. Several works have explored this idea of joint data and channel recovery ([29], [47] , [48], [49], [51], [52], [38], [42], [53], [44], [57], [58]). A data aided approach seems most appropriate for channel estimation as it makes a collective use of data and channel constraints for estimation.

1.2.5 Constraints Used in Channel Estimation/Data Detection

All the works mentioned above, use a subset of the following constraints on the channel estimate or data, regardless of the estimation technique used. Following is a survey of these constraints and the work that employs them.

Data Constraints:

Finite alphabet constraint: Data is usually drawn from a finite alphabet set. The authors in [19], [35] and [38] make use of this constraint.

Code: Data usually exhibits some form of redundancy like a code that helps reduce the row probability or err [22], [30], [49].

Transmit precoding: The data might also contain some form of precoding (to facilitate equalization at the receiver) such as a cyclic prefix, silent guard bands [46], [50] and known symbol precoding [61].

Pilots: Pilots represent the most primitive form of redundancy and are usually inserted to perform channel estimation or simply to initialize the estimation process [11], [7], [8], [28], [44], [12], [27].

Channel Constraints:

Finite delay spread: The channel is usually of finite impulse response with a maximum delay spread that is assumed available to the receiver.

Sparsity: the sparsity of a multipath fading channel is defined as the ratio of the time duration spanned by the multipaths to their number [15], [34], [39]. the number of paths and their delays are usually stationary. However, their amplitudes and relative phases usually very much more rapidly with time. this essentially reduces the number of parameters to be estimated to that of the number of multipaths in the channel.

Frequency correlation: In addition to information about which of the channel taps are inactive, we usually have additional statistical information about the active ones. Thus, it is usually assumed that the taps are Gaussian (zero mean or not depending on whether the channel exhibits Rayleigh or Rician fading) with a certain covariance matrix. this matrix is a measure of the frequency correlation among the taps [37], [54].

Time correlation: As channels vary with time, they exhibit some form of time correlation. time-variant behavior could also be more structured, e.g., following a state-space model [47], [48], [56].

Uncertainty information: Channel also suffers from non ideal effects such as nonlinearities and rapid time-variations that are difficult to model. The aggregate effect of this non ideal behavior could be represented as uncertainty information that can be used to build robust receivers [21].

Regardless of the approach used for channel estimation or the constraints employed, estimation can be carried out in any of the two domains (time and frequency). Below, we classify the approaches that are used in either of these two domians. We also discuss the advantages and disadvantages of estimation in these domains. All these methods for channel estimation are either in the frequency domain or in the time domain. Below is a survey of various works in the two domains.

1.2.6 Time Domain Channel Estimation

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 proposed in [40]. A time domain minimum mean square error (MMSE) channel estimation technique based on subspace tracking for OFDM system is put forward in [41]. In [42], a joint channel and data estimation algorithm is presented which makes a collective use of data and channel constraints. A simplified joint frequency-offsett and channel estimation technique for Multi-Symbol Encapsulated MSE OFDM system is proposed in [23], while authors in [26] present a sequential method for channel response estimation based on Carrier Frequency Offset and symbol timing estimation by exploiting the structure of the packet preamble of IEEE 802.11a standard. The authors in [43] take a statistical approach and estimate the channel based on Power Spectral Density (PSD) and LS estimation for OFDM systems with timing offsets. An iterative receiver structure with joint detection and channel estimation based on implicit pilots is proposed in [44] and [45] presents a pilot aided channel estimation algorithm in the presence of synchronous noise by exploiting the a priori available information about the interference structure.

1.2.7 Frequency Domain Channel Estimation

In the past years, various techniques for channel estimation in the frequency domain have also been explored. Researchers in [31] apply phase shifted samples in the frequency-domain to an interpolated LS to estimate the channel while in [32], the authors propose to include a phase rotation term in the frequency domain interpolation for better CIR window location. Channel estimation using polynomial cancelation coding (PCC) training symbols and frequency domain windowing is proposed in [33]. A sub-band approach to channel estimation and channel equalization is proposed in [36]. A low-complexity iterative channel estimator is proposed in [29]. The minimum mean square error (MMSE) channel estimation in the frequency domain is considered in [37] while researchers in [27] present delay subspace-based channel estimation techniques for OFDM systems over fast-fading channels.

1.3 Disadvantage of Performing Channel Estimation in Time Domain

Most channel estimation algorithms for OFDM presented in literature perform estimation in the time domain (instead of the frequency domain) [12], [40], [42], [23], [26]. By performing estimation in the time domain, one can decrease the degrees of freedom from N, the number of frequency bins, to $P+1$, the number of (time domain) channel taps. This is a drastic reduction since the number of channel taps is usually less than the cyclic prefix which is usually designed to be less than $\frac{N}{4}$. The reduction in the parameter estimation space in turn results in improved estimation accuracy.

There is a certain price that we have to pay, however, for this gain. We loose the diagonal structure of the channel by performing the estimation in the time domain. Thus, instead of frequency domain relationship (1.4) in which the various equations are decoupled, we employ the time-frequency relationship (1.8) 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 (1.4) is performed on a bin by bin basis according to

$$\widehat{\mathcal{H}}(l) = \frac{\mathcal{Y}(l)}{\mathcal{X}(l)} l = 1, 2, \cdots, N \tag{1.9}$$

channel estimation in (1.8) requires size P +1 matrix inversion

$$\underline{\widehat{\boldsymbol{h}}} = (\boldsymbol{X}^* \boldsymbol{X})^{-1} \boldsymbol{X}^* \boldsymbol{\mathcal{Y}} \tag{1.10}$$

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 $\mathcal{H}$ from the time domain estimate $\hat{\boldsymbol{h}}$ and use it for data detection). Thus, for data-aided channel estimation techniques, each channel estimation step would require one such IFFT operation.

Apart from the computational complexity, performing channel estimation in the time domain might be over solving a problem. For example, in multiple access OFDM systems, like WiMAX, users are not interested in the whole frequency spectrum, but only in that part of the spectrum in

which they are operating. In fact these users don't have access to the whole spectrum but only a part of it is available to them. Moreover, even if some users were interested in estimating the whole spectrum, many standards would not be able to support that as there are not enough pilots to do so.

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

Channel estimation in the frequency domain avoids the above disadvantages. Moreover, the structure that characterizes the estimation problem in the time domain continues to characterize the estimation problem in the frequency domain. Specifically, the time and frequency correlation exhibited by the time domain channel maps in to corresponding correlation of the channel frequency response.

The only problem with channel estimation in the frequency domain is the increase in the number of parameter to be estimated [37]. 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 1.2 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. Note that within a narrow enough band of spectrum, the spectrum looks linear or quadratic. As such, we employ model reduction in this paper to estimate the spectrum, thereby reducing the number of parameters to be estimated.

1.5 Input/Output Relationship in the Frequency Domain

The input/output relationship of the OFDM system is best described in the frequency domain. A frequency domain channel response of length N is shown in figure 1.2. We start by partitioning the

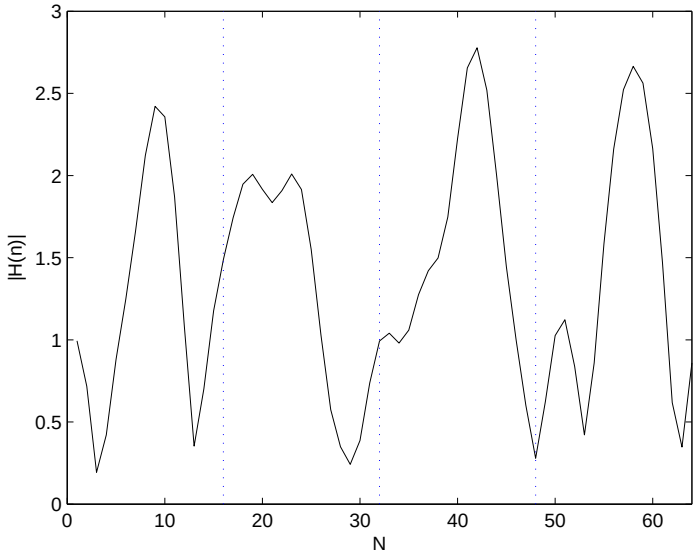


Figure 1.2: Channel Impulse Response in the Frequency Domain partitioned in four subchannels.

channel response into a number of sections each of length L_f producing a total of $\lceil \frac{N}{L_f} \rceil$ sections[1]. Let us denote the j^{th} section of the frequency response by $\underline{\boldsymbol{\mathcal{H}}}^{(j)}$. The input/output equation for this section is given by

$$\underline{\boldsymbol{\mathcal{Y}}}_i^{(j)} = \mathrm{diag}(\underline{\boldsymbol{\mathcal{X}}}_{\mathrm{i}}^{(\mathrm{j})})\underline{\boldsymbol{\mathcal{H}}}_{\mathrm{i}}^{(\mathrm{j})} + \underline{\boldsymbol{\mathcal{N}}}_{\mathrm{i}}^{(\mathrm{j})} \tag{1.11}$$

Where $\underline{\boldsymbol{\mathcal{Y}}}^{(j)}$, $\underline{\boldsymbol{\mathcal{X}}}^{(j)}$, $\underline{\boldsymbol{\mathcal{H}}}^{(j)}$ and $\underline{\boldsymbol{\mathcal{N}}}^{(j)}$ are the j^{th} sections of $\boldsymbol{\mathcal{Y}}$, $\boldsymbol{\mathcal{X}}$, $\boldsymbol{\mathcal{H}}$ and $\boldsymbol{\mathcal{N}}$ respectively. From now onwards, we will drop the dependence on j for notational convenience. Equation (1.11) can now be written as

$$\underline{\boldsymbol{\mathcal{Y}}}_i = \mathrm{diag}(\underline{\boldsymbol{\mathcal{X}}}_{\mathrm{i}})\underline{\boldsymbol{\mathcal{H}}}_{\mathrm{i}} + \underline{\boldsymbol{\mathcal{N}}}_{\mathrm{i}} \tag{1.12}$$

Where $\boldsymbol{\mathcal{N}}_i \sim \mathcal{N}(0, \sigma_i^2 I)$ is the additive white gaussian noise.

1.6 Pilot/Output Relationship in Frequency Domain

In general, the receiver needs pilots to obtain a channel estimate. The pilot locations within the OFDM symbol are denoted by the index set $I_p = i_1, i_2, \cdots, i_{L_p}$. Also, let $\mathrm{diag}(\boldsymbol{\mathcal{X}}_{\mathrm{I_p}})$ denote the

[1]In 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.

matrix diag($\boldsymbol{\mathcal{X}}$) pruned of the rows that do not belong to I_p. Then, the pilot/output equation can be derived from the input/output relation (1.12) as

$$\underline{\boldsymbol{\mathcal{Y}}}_{I_p} = \text{diag}(\underline{\boldsymbol{\mathcal{X}}}_{\text{I}_\text{p}})\underline{\boldsymbol{\mathcal{H}}} + \underline{\boldsymbol{\mathcal{N}}}_{\text{I}_\text{p}} \tag{1.13}$$

1.7 A Parameter Reduction Approach

The main hindrance in performing channel estimation in the frequency domain, as opposed to the time domain estimation, is the increased number of parameters to be estimated. Our goal here is to apply some parameter reduction technique to reduce the number of frequency domain parameters to be estimated. Dropping the dependence on time index $\boldsymbol{i}$ for notational convenience, we consider that $\underline{\boldsymbol{\mathcal{H}}}$ can be expressed as

$$\underline{\boldsymbol{\mathcal{H}}} = \boldsymbol{V}_p \boldsymbol{\alpha}_d \tag{1.14}$$

where $\boldsymbol{V}_p$ is a known matrix and $\boldsymbol{\alpha}_d$ is the vector of parameters to be determined. We will consider two different approaches for estimating $\underline{\boldsymbol{\mathcal{H}}}$. One approach is to consider a linear or quadratic approximation which we will discuss in chapter 2. Another way to go about solving for $\underline{\boldsymbol{\mathcal{H}}}$ is based on Eigenvalue decomposition and is discussed in chapter 3.

Chapter 2

Interpolation Based Frequency Domain Estimation

The problem that we encounter when performing channel estimation in frequency domain is the increased number of parameters to be estimated. For frequency domain estimation, we require to estimate N parameters while in the case of time domain estimate, we only need to estimate $P + 1$ parameters. We can eliminate this disadvantage if we can find a way to decrease the parameter estimation space for the frequency domain estimation, such that it is comparable to the number of parameters needed for time domain estimation. The frequency response of the channel is inherently limited by the degrees of freedom of the time domain impulse response. Figure 2.1 shows a length 64 frequency response of a 16 tap channel with an exponential delay profile similar to the one that will be use in simulations. We can see that with in a narrow enough band width, the spectrum can be approximated as linear or quadratic. Mathematically speaking, let $\boldsymbol{\mathcal{H}}$(k) be a sub band of the frequency spectrum of width L_f(where k = 1, 2, $\cdots$, $\lfloor \frac{N}{Lf} \rfloor$). If the frequency spectrum is linear in

this sub band, then we can write

$$\underline{\boldsymbol{\mathcal{H}}}(k) = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ \vdots & \vdots \\ 1 & L_f \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \tag{2.1}$$

If the spectrum is quadratic, we can write:

$$\underline{\boldsymbol{\mathcal{H}}}(k) = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2^2 \\ \vdots & \vdots & \vdots \\ 1 & L_f & L_f^2 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} \tag{2.2}$$

In general, we can write

$$\underline{\boldsymbol{\mathcal{H}}}(k) = \boldsymbol{V}_p \boldsymbol{\alpha}_d \tag{2.3}$$

where $\boldsymbol{V}_p$ is the interpolation matrix and $\boldsymbol{\alpha}_d$ is the vector of interpolation parameters.

The input/output relation is given by equation(1.12). Replacing $\underline{\boldsymbol{\mathcal{H}}}$ from equation (2.3) results in

$$\begin{aligned} \underline{\boldsymbol{\mathcal{Y}}} &= \mathrm{diag}(\underline{\boldsymbol{\mathcal{X}}}) \boldsymbol{V}_p \boldsymbol{\alpha}_d + \underline{\boldsymbol{\mathcal{N}}} \\ &= \underline{\boldsymbol{\Omega}} \boldsymbol{\alpha}_d + \underline{\boldsymbol{\mathcal{N}}} \end{aligned} \tag{2.4}$$

where $\underline{\boldsymbol{\Omega}} = \mathrm{diag}(\underline{\boldsymbol{\mathcal{X}}}) \boldsymbol{V}_p$ and $\boldsymbol{\mathcal{N}}$ is zero mean white gaussian noise. Pruning the above equation yields

$$\underline{\boldsymbol{\mathcal{Y}}}_{I_p} = \underline{\boldsymbol{\Omega}}_{I_p} \boldsymbol{\alpha}_d + \underline{\boldsymbol{\mathcal{N}}}_{I_p} \tag{2.5}$$

2.1 Least Squares

The solution of equation (2.5) is based on minimizing

$$\widehat{\boldsymbol{\alpha}}_d = \arg\min_{\alpha_d} \{ || \underline{\boldsymbol{\mathcal{Y}}}_{I_p} - \boldsymbol{\Omega}_{I_p} \boldsymbol{\alpha}_d ||^2 \} \tag{2.6}$$

Solving it as a least square problem [59], yields

$$\widehat{\boldsymbol{\alpha}}_d = (\boldsymbol{\Omega}_{I_p}^* \boldsymbol{\Omega}_{I_p})^{-1} \boldsymbol{\Omega}_{I_p}^* \underline{\boldsymbol{\mathcal{Y}}}_{I_p} \tag{2.7}$$

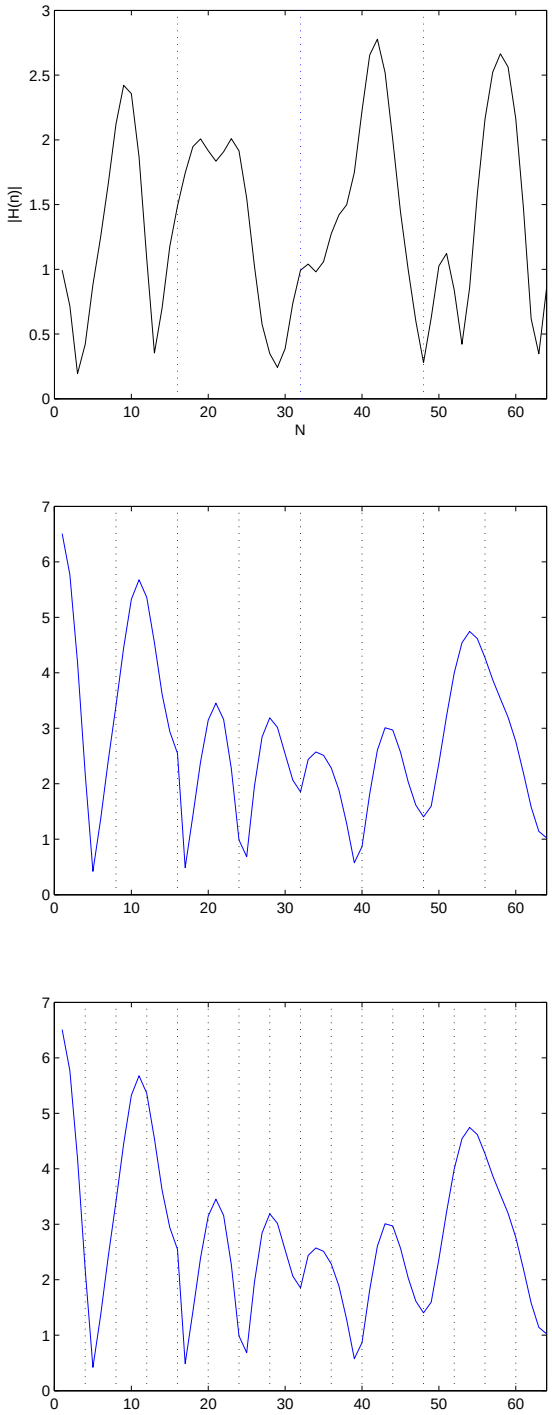


Figure 2.1: Channel Impulse Response in the Frequency Domain divided into 4, 8 and 16 parts.

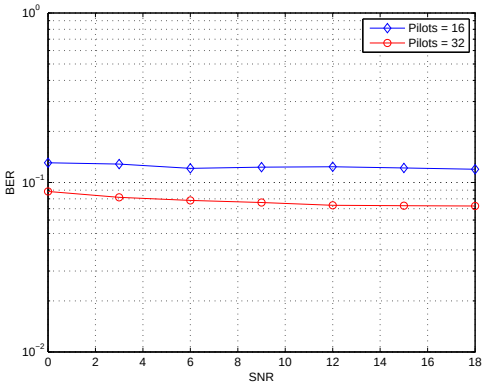


Figure 2.2: BER curve for 8 sections 2 parameters

The estimate of the j^{th} section of the channel is thus given by

$$\underline{\boldsymbol{\mathcal{H}}} = \boldsymbol{V}_p \widehat{\boldsymbol{\alpha}}_d \tag{2.8}$$

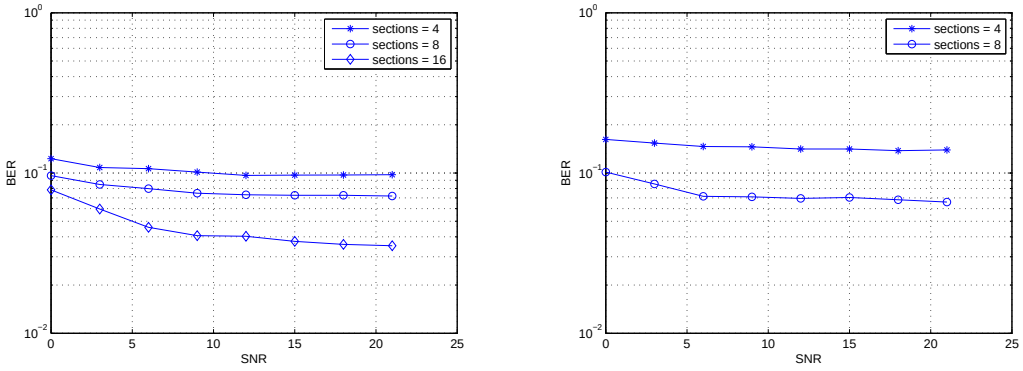
Concatenation of all M such section will give us the complete channel response $\boldsymbol{\mathcal{H}}$.

2.2 Simulation Parameters

Consider an OFDM system where an iid sequence of $T+1$ data symbols $\boldsymbol{\mathcal{X}}_o^T$ are to be transmitted. The length of each symbol N, is 64. We use a CP of length 15. The modulation scheme used is 16 QAM with grey coding. The channel impulse response(CIR) is considered to consist of 16 complex taps(maximum length allowable for the channel with a CP length of 15). The exponential decay profile $E[|\underline{h}_0(k)|^2]$ of the channel remains fixed over any OFDM symbol and is taken to be $e^{-0.2k}$. These parameters are used throughout the simulations.

2.3 Effect of number of pilots

By intuition, we know that increasing the number of pilots should yield a better channel estimate and hence better BER performance. Figure 2.2 is plotted for 16 and 32 pilots. In both cases, we use 2 interpolation parameters and 8 sections. As evident from the figure, increasing the number of pilots will lead to a better channel estimate.



(a) (b)

Figure 2.3: (a) BER curve for 32 pilots and 2 parameters (b) BER curve for 32 pilots and 3 parameters.

2.4 Effect of section length

Another way to improve the channel estimate is would be to divide the channel into a larger number of sections. We employ 32 pilots, 2 interpolation parameters and divide the channel into 4, 8 and 16 sections respectively. The BER curves for these three cases are shown in figure 2.3(a). We can see that decreasing the section length, i.e. increasing the number of section per channel, results in a better BER performance. Figure 2.3(b) shows BER performance for 32 pilots and 3 parameters and shows the same trend. So for a better channel estimate, we should use a larger number of sections.

2.5 Effect of varying the number of parameters

The channel estimate is also affected by varying the number of estimation parameters. We plot the BER of the system using 2 and 3 interpolation parameters. In both cases, we use 32 pilots and dived the channel into 8 sections. Figure 2.4 shows the effect of changing the number of parameter on the BER performance. As we can see, increasing the number of parameters from 2 to 3, results in improved BER performance specially at high SNR. So increasing the number of interpolation parameters improves the channel estimate. For figures 2.2-2.4 above, we conclude that

- Increasing the number of pilots improves the channel estimate.

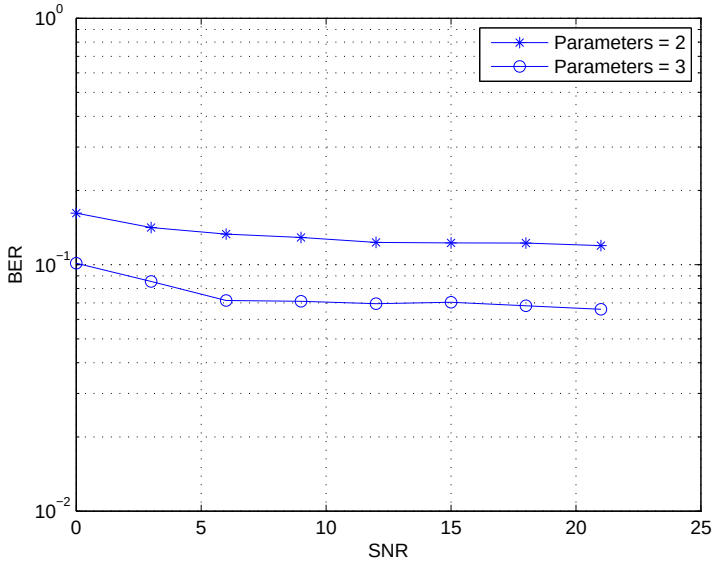


Figure 2.4: BER curve using 32 pilots and 8 sections

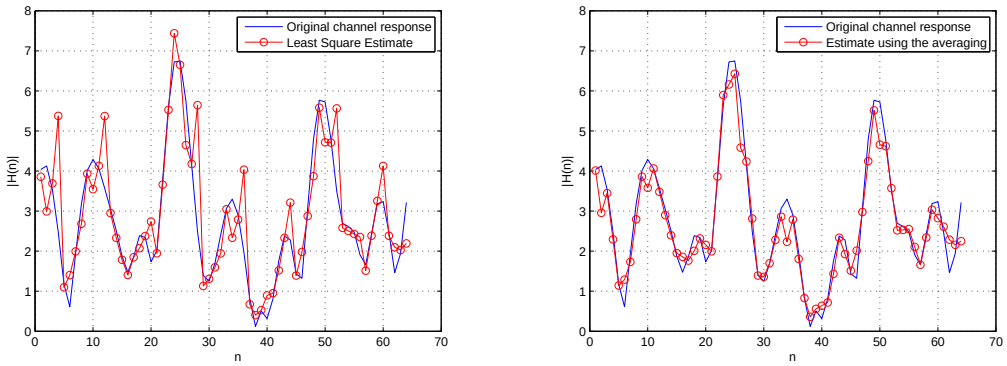
- Increasing the number of sections in which the frequency domain channel response is divided, improves the channel estimate.
- Increasing the number of interpolation parameters improves the channel estimate.

However, there is a limit to the extent we can increase these parameters. Increasing the number of pilots means fewer carriers are available for data transmission purpose. The number of sections and the number of interpolation parameters are in turn both limited by the number of pilots we use. For the Least Square solution of equation (2.5), requires the following condition to be fulfilled

$$\text{number of pilots in each section} \geq \text{number of interpolation parameters} \tag{2.9}$$

So if use 32 pilots and 2 interpolation parameters, then every section of the channel response must have at least 2 pilots. That means that at most we can divide the channel response into 16 sections. If we increase the number of interpolation parameters to 3, than each section must have at least 3 pilots. In this case channel response can be divided into a maximum of 8 sections.

This limitation can be avoided if we use a regularized Least Square solution for equation (2.5). In that case we can have as many sections and interpolation parameters as we want as long as there



(a) (b)

Figure 2.5: (a) CIR inflated at the edges(32/2/16) (b) Removing inflation using averaging scheme.

is at least one pilot per section. i.e.

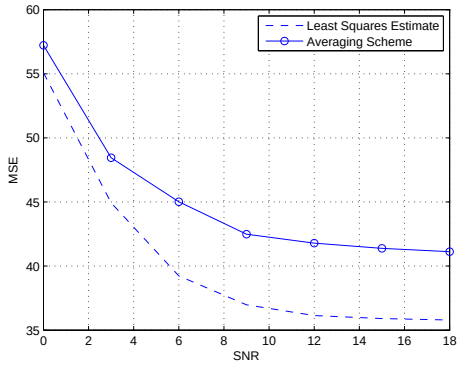
$$\text{number of pilots in each section} \geq 1 \tag{2.10}$$

2.6 Schemes to improve the channel estimate

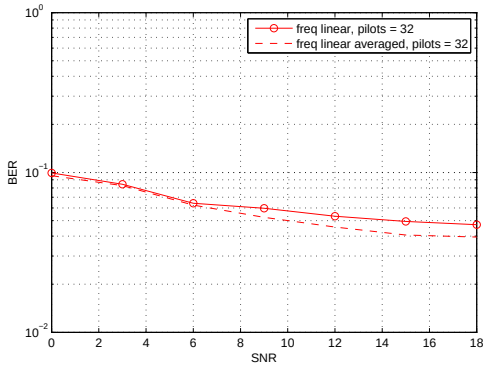
The interpolation method we use in this chapter is polynomial based. As such, we expect the point at the edge of each section to be inflated as shown in figure 2.5(a). The first figure is for 2 interpolation parameters and the second is for 3 parameters. both have 32 pilots and 16 sections. If we can somehow correct these inflated points, our estimate is bound to improve. In order to reduce this error, we use an **Averaging Scheme**. This scheme sets the estimate of the edge point of each section to be the average of the second last point of the current section and the first point of the next section.

Figures 2.5(a) and 2.5(b) show the original LS based channel estimate(32 pilots/2 parameters/16 sections) with inflated points and compares it with the averaging scheme. We can see that using the averaging scheme improves the channel estimate. Let uss compare the performance of the two methods to get a better idea of the advantage offered by the averaging scheme. We consider the case of 32 pilots. Figures 2.6(a) and 2.6(b) show the Error and BER plots for the two schemes using 3 interpolation parameters, 8 sections per channel and figures 2.6(c) and 2.6(d) show the same plots

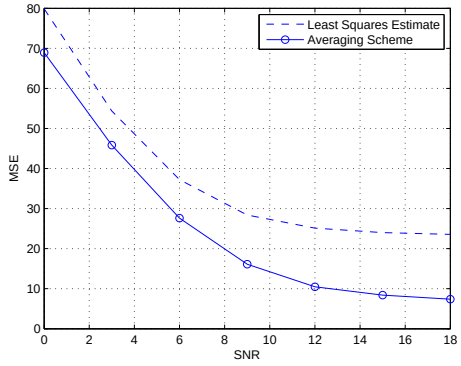
for 2 interpolation parameters, 16 sections per channel. It is evident that the averaging scheme outperforms the LS solution. Also an interesting observation is that the averaging scheme performs better in the case of 16 sections. This is logical as the later case has more number of sections and more edge points will be corrected by the averaging scheme than in the former case.



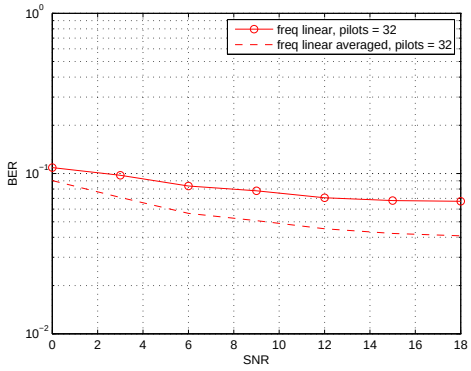
(a)



(b)



(c)



(d)

Figure 2.6: (a) Error plot for LS and Averaging Scheme for (32/3/8) (b) BER plots for LS and Averaged Scheme for (32/3/8) (c) Error plot for LS and Averaging Scheme for (32/2/16) (d) BER plots for LS and Averaged Scheme for (32/2/16).

2.7 Least Squares with regularization

Ideally, we estimate $\boldsymbol{\alpha}_d$ using some I/O relationship by maximizing the corresponding log-likelihood function

$$\hat{\boldsymbol{\alpha}}_d = \arg\max_{\alpha_d} \{\ln p(\underline{\boldsymbol{\mathcal{Y}}}|\underline{\boldsymbol{\Omega}}, \boldsymbol{\alpha}_d) + \ln p(\boldsymbol{\alpha}_d)\}$$

When the channel obeys the I/O relationship (2.4) (so that $\ln p(\underline{\boldsymbol{\mathcal{Y}}}|\underline{\boldsymbol{\Omega}}, \boldsymbol{\alpha}_d) = -\|\underline{\boldsymbol{\mathcal{Y}}} - \underline{\boldsymbol{\Omega}}\boldsymbol{\alpha}_d\|^2_{\sigma^{-2}}$ up to some additive constant $\ln p(\boldsymbol{\alpha}_d) = -\|\boldsymbol{\alpha}_d\|^2_{R^{-1}_{\alpha_d}}$ up to some additive constant), then the LS estimate is given by

$$\hat{\boldsymbol{\alpha}}_d = \arg\max_{\alpha_d} \left\{ ||\underline{\boldsymbol{\mathcal{Y}}}_{I_p} - \underline{\boldsymbol{\Omega}}_{I_p}\boldsymbol{\alpha}_d||^2_{\sigma_n^{-2}I} + ||\boldsymbol{\alpha}_d||^2_{R^{-1}_{\alpha_d}} \right\} \tag{2.11}$$

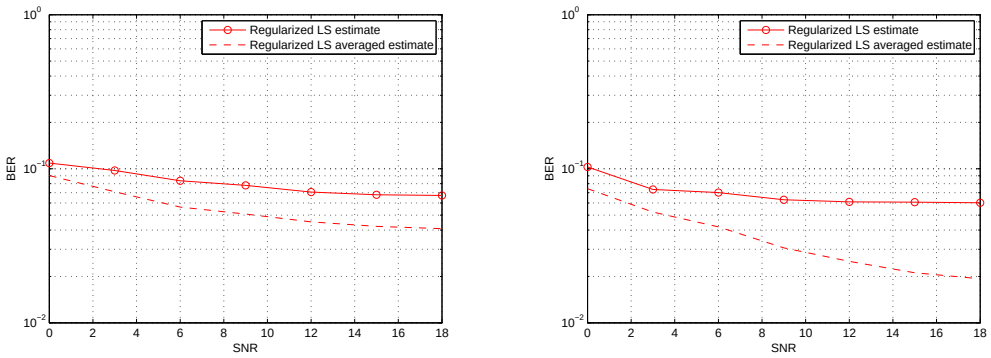
where σ_n^2 is the noise variance. The estimate of $\boldsymbol{\alpha}_d$ that minimizes the MSE is given by

$$\widehat{\boldsymbol{\alpha}}_{\boldsymbol{d}} = \boldsymbol{R}_{\boldsymbol{\alpha_d}}\underline{\boldsymbol{\Omega}}^{*}_{\boldsymbol{I_p}}[\boldsymbol{\sigma_n^{-2}I} + \underline{\boldsymbol{\Omega}}_{\boldsymbol{I_p}}\boldsymbol{R}_{\boldsymbol{\alpha_d}}\underline{\boldsymbol{\Omega}}^{*}_{\boldsymbol{I_p}}]^{-1}\underline{\boldsymbol{\mathcal{Y}}}_{\boldsymbol{I_p}} \tag{2.12}$$

We will assume that $\boldsymbol{R}_{\alpha_d}$ is the identity matrix, that is we assume no correlation between parameters. The advantage of regularized LS solution is that there is no restriction on the number of pilots per section and they can be less than the number of parameters. This allows us to try those combinations of interpolation parameters, number of pilots and number of sections which are not possible in the non-regularized case. Following are the BER curves for the regularized case with and without averaging. Figure 2.7(a) and 2.7(b) are both plotted for 32 pilots and 16 sections. The number of interpolation parameters used is 2 and 3 respectively.

2.8 Conclusion

The interpolation techniques based on simple linearization and quadratic approximation investigated in this chapter require very dense pilot placement and this increased number of frequency domain parameters to be estimated. The inherent limit on the number of interpolation parameters per section can be removed by considering a regularized LS solution for the estimation problem. Still, we find that this method requires a high number of estimation parameter for channel estimation. As such, this method proves to be too expansive. Hence the need to explore some other method



(a) (b)

Figure 2.7: (a) BER curve for regularized solution with 32 pilots/2 parameters/16 sections (b) BER curve for regularized solution with 32 pilots/3 parameters/16 sections.

to represent the channel in frequency domain. In the next chapter, we will consider an alternate method for this purpose.

Chapter 3

Eigenvalue Approach to Frequency Domain Channel Estimation

Another approach for reducing the parameters in frequency domain channel estimation is the eigenvalue approach. Assuming that the second order statistics of the channel is available at the receiver, we can find its Eigenvalue decomposition. Using model reduction, we can represent $\underline{\boldsymbol{\mathcal{H}}}$ using dominant eigenvalue and treat the rest as modeling noise.

The input/output equation that involves the j^{th} section is given by equation (1.12) while equation (1.13) gives its pruned form. We reproduce them here for easy reference

$$\underline{\boldsymbol{\mathcal{Y}}}_i^{(j)} = \text{diag}(\underline{\boldsymbol{\mathcal{X}}}_i^{(j)})\underline{\boldsymbol{\mathcal{H}}}_i^{(j)} + \underline{\boldsymbol{\mathcal{N}}}_i^{(j)} \tag{3.1}$$

Dropping the dependence on j and i for notational convenience and pruning

$$\underline{\boldsymbol{\mathcal{Y}}}_{I_p} = \text{diag}(\underline{\boldsymbol{\mathcal{X}}}_{I_p})\underline{\boldsymbol{\mathcal{H}}} + \underline{\boldsymbol{\mathcal{N}}}_{I_p} \tag{3.2}$$

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

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

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

$$\underline{\boldsymbol{\mathcal{H}}} = \sum_{l=1}^{L_f} \alpha_l \mathbf{v_l}$$

where $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \ldots, \alpha_{L_f}]^T$ is a parameter vector, to be estimated, with zero mean and autocorrelation matrix $\boldsymbol{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_{L_f})$. We now represent $\underline{\boldsymbol{\mathcal{H}}}$ using the dominant eigenvalues and treat the rest as modeling noise [1], i.e.

$$\underline{\boldsymbol{\mathcal{H}}} = \boldsymbol{V}_d \boldsymbol{\alpha}_d + \boldsymbol{V}_n \boldsymbol{\alpha}_n \tag{3.3}$$

Upon substituting (3.3) in (3.1), we obtain

$$\begin{aligned} \underline{\boldsymbol{\mathcal{Y}}} &= \text{diag}(\underline{\boldsymbol{\mathcal{X}}}) \boldsymbol{V}_d \boldsymbol{\alpha}_d + \underline{\boldsymbol{\mathcal{N}}} + \text{diag}(\underline{\boldsymbol{\mathcal{X}}}) \boldsymbol{V}_n \boldsymbol{\alpha}_n & (3.4) \\ &= \underline{\boldsymbol{X}}_d \boldsymbol{\alpha}_d + \underline{\boldsymbol{\mathcal{N}}}' & (3.5) \end{aligned}$$

where $\underline{\boldsymbol{X}}_d = \text{diag}(\underline{\boldsymbol{\mathcal{X}}}) \boldsymbol{V}_d$ and $\underline{\boldsymbol{\mathcal{N}}}' = \underline{\boldsymbol{\mathcal{N}}} + \underline{\boldsymbol{X}}_n \boldsymbol{\alpha}_n$ with $\underline{\boldsymbol{X}}_n = \text{diag}(\underline{\boldsymbol{\mathcal{X}}}) \boldsymbol{V}_n$. 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

$$\boldsymbol{R}_{\underline{\mathcal{N}}'} = \boldsymbol{R}_{\underline{\mathcal{N}}} + \text{diag}(\underline{\boldsymbol{\mathcal{X}}}) \boldsymbol{V}_n \text{diag}(\lambda_n) \boldsymbol{V}_n^* \text{diag}(\underline{\boldsymbol{\mathcal{X}}})^* \tag{3.6}$$

Now equation (3.5) can be used to construct a pilot/output equation, similar to (3.2), as

$$\underline{\boldsymbol{\mathcal{Y}}}_{I_p} = \underline{\boldsymbol{X}}_{d,I_p} \boldsymbol{\alpha}_d + \underline{\boldsymbol{\mathcal{N}}}'_{I_p} \tag{3.7}$$

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

$$\hat{\boldsymbol{\alpha}}_d^{\mathsf{MAP}} = \arg\max_{\alpha_d} \left\{ \ln\, p(\underline{\boldsymbol{\mathcal{Y}}}_{I_p} | \underline{\boldsymbol{X}}_{d,I_p}, \boldsymbol{\alpha}_d) + \ln\, p(\boldsymbol{\alpha}_d) \right\} \tag{3.8}$$

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

$$\hat{\boldsymbol{\alpha}}_d^{\mathsf{MAP}} = \arg\min_{\alpha_d} \left\{ \|\underline{\boldsymbol{\mathcal{Y}}}_{I_p} - \underline{\boldsymbol{X}}_{d,I_p} \boldsymbol{\alpha}_d\|^2_{\boldsymbol{R}^{-1}_{\underline{\mathcal{N}}'}} + \|\boldsymbol{\alpha}_d\|^2_{\boldsymbol{\Lambda}_d^{-1}} \right\} \tag{3.9}$$

[1] 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 $\frac{\lambda_{j+1}}{\lambda_j} > 5$ to place our cutoff.

which simplifies to

$$\hat{\boldsymbol{\alpha}}_d = \boldsymbol{\Lambda}_d \underline{\boldsymbol{X}}^*_{d,I_p} \left[\boldsymbol{R}_{\underline{\mathcal{N}}'} + \underline{\boldsymbol{X}}_{d,I_p} \boldsymbol{\Lambda}_d \underline{\boldsymbol{X}}^*_{d,I_p} \right]^{-1} \underline{\boldsymbol{\mathcal{Y}}}_{I_p} \tag{3.10}$$

The resulting mean square error is given by

$$\boldsymbol{R}_e = \left[\boldsymbol{\Lambda}_d^{-1} + \underline{\boldsymbol{X}}^*_{I_p} \boldsymbol{R}^{-1}_{\underline{\mathcal{N}}'} \underline{\boldsymbol{X}}_{I_p} \right]^{-1} \tag{3.11}$$

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

3.1 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 [38], [42]. Using the data aided approach, we can improve the channel estimate [29], [42]. Thus providing the motivation to use the expectation-maximization (EM) algorithm. The EM algorithm is used to estimate a parameter in the case when some of the date required for estimation is unobserved. The algorithm first performs an initial estimate of the unobserved data and uses this estimate to compute the maximum-likelihood (ML) estimate of the parameter to be estimated. This is the maximization step. Next, the algorithm uses the estimated parameter to update the estimate of the unobserved data. this is the expectation step. These steps are repeated iteratively until a convergent solution is reached [60]. Next, we will discuss the EM algorithm in detail.

3.1.1 The Maximization Step

In the previous subsection we find $\hat{\boldsymbol{\alpha}}_d$ by maximizing the log likelihood function given by equation (3.8). Since the input $\boldsymbol{\mathcal{X}}$ (and hence $\underline{\boldsymbol{X}}_d$) is not observable, we can employ the EM algorithm and instead of maximizing (3.8) we can maximize an averaged from 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

$$\hat{\boldsymbol{\alpha}}_d^{(k)} = \arg\max_{\alpha_d} \left\{ E_{\mathcal{X}_i|\mathcal{Y}_i,\hat{\boldsymbol{\alpha}}_d{}^{(k-1)}} \ln\, p(\underline{\boldsymbol{\mathcal{Y}}}_{I_p}|\underline{\boldsymbol{X}}_{d,I_p}, \boldsymbol{\alpha}_d) + \ln\, p(\boldsymbol{\alpha}_d) \right\} \tag{3.12}$$

which simplifies to [2]

$$\hat{\boldsymbol{\alpha}}_d^{\mathsf{MAP}} = \arg\min_{\alpha_d} \left\{ E\|[\underline{\boldsymbol{\mathcal{Y}}}_{I_p} - \underline{\boldsymbol{X}}_{d,I_p}\boldsymbol{\alpha}_d\|^2_{\boldsymbol{R}^{-1}_{\underline{\mathcal{N}'}}} + \|\boldsymbol{\alpha}_d\|^2_{\boldsymbol{\Lambda}_d^{-1}} \right\} \tag{3.13}$$

Strictly speaking, the noise correlation $\boldsymbol{R}_{\underline{\mathcal{N}'}}$ is itself dependent on the input due to the modeling noise (see equation (3.6)). Hence in performing the expectation in (3.13), 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: $\boldsymbol{\mathcal{R}}_{\underline{\mathcal{N}'}}$ 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 (3.13) 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 (3.13). Specifically, upon completing the squares, (3.13) 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

$$\hat{\boldsymbol{\alpha}}_d^{\mathsf{MAP}} = \arg\min_{\alpha_d} \ \|\underline{\boldsymbol{\mathcal{Y}}} - E[\underline{\boldsymbol{X}}_d]\boldsymbol{\alpha}_d\|^2_{\frac{1}{\sigma_n^2} I} + \|\boldsymbol{\alpha}_d\|^2_{\frac{1}{\sigma_n^2} Cov[\underline{\boldsymbol{X}}_d^*]} + \|\boldsymbol{\alpha}_d\|^2_{\Lambda_d^{-1}} \tag{3.14}$$

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

Instead of ignoring the modeling noise, we can split the expectation in (3.13) into an expectation

[2] the 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 & dropped for notational convenience.

over $\boldsymbol{\mathcal{R}}_{\underline{\mathcal{N}}'}$ and an *independent* expectation taken over the rest of the terms i.e., we can approximate (3.13) as

$$\hat{\boldsymbol{\alpha}}_d^{\mathsf{MAP}} = \arg\min_{\alpha_d} \left\{ E\|[\underline{\boldsymbol{\mathcal{Y}}}_{I_p} - \underline{\boldsymbol{X}}_{d,I_p}\boldsymbol{\alpha}_d\|^2_{E[R_{\underline{\mathcal{N}}'}]^{-1}} + \|\boldsymbol{\alpha}_d\|^2_{\boldsymbol{\Lambda}_d^{-1}} \right\} \tag{3.15}$$

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

$$E[\boldsymbol{R}_{\underline{\boldsymbol{\mathcal{N}}}'}] = \sigma^2\boldsymbol{I} + E[\mathrm{diag}(\underline{\boldsymbol{\mathcal{X}}})\boldsymbol{V}_n\boldsymbol{\Lambda}_n\boldsymbol{V}_n^*\mathrm{diag}(\underline{\boldsymbol{\mathcal{X}}}^*)] \tag{3.16}$$

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

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

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 $\boldsymbol{V}_n\boldsymbol{\Lambda}_n\boldsymbol{V}_n^*$. 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 (3.15) which comes out to be

$$\hat{\boldsymbol{\alpha}}_d^{\mathsf{MAP}} = \arg\min_{\alpha_d} \ \|\underline{\boldsymbol{\mathcal{Y}}} - E[\underline{\boldsymbol{X}}_d]\boldsymbol{\alpha}_d\|^2_{E[R_{\underline{\mathcal{N}}'}]^{-1}} + \|\boldsymbol{\alpha}_d\|^2_{\mathrm{Cov}[\underline{\boldsymbol{D}}]\mathrm{diag}(\boldsymbol{V}_n\boldsymbol{\Lambda}_n\boldsymbol{V}_n^*)} + \|\boldsymbol{\alpha}_d\|^2_{\Lambda_d^{-1}} \tag{3.17}$$

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

In the constant modulus case, it is possible to evaluate (3.13) exactly. Specifically, and starting from the expression for the autocorrelation $\boldsymbol{R}_{\underline{\mathcal{N}}'}$

$$\boldsymbol{R}_{\underline{\mathcal{N}}'} = \sigma^2\boldsymbol{I} + \underline{\boldsymbol{D}}\boldsymbol{V}_n\boldsymbol{\Lambda}_n\boldsymbol{V}_n^*\underline{\boldsymbol{D}}^*$$

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 $\boldsymbol{R}_{\underline{\mathcal{N}}''} \stackrel{\Delta}{=} \frac{\sigma^2}{\mathcal{E}}\boldsymbol{I} + \boldsymbol{V}_n\boldsymbol{\Lambda}_n\boldsymbol{V}_n^*$ 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

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

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

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, (3.13) can be equivalently written as

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

which upon simplification becomes

$$\hat{\boldsymbol{\alpha}}_d^{\mathsf{MAP}} = \arg\min_{\alpha_d} \|E[\mathbf{D}^{-1}]\boldsymbol{\mathcal{Y}} - \boldsymbol{V}_d\boldsymbol{\alpha}_d\|_{R_{\mathcal{N}''}^{-1}}^2 + \|\boldsymbol{\alpha}_d\|_{\Lambda_d^{-1}}^2 \tag{3.18}$$

In the simulations further ahead, we compare the approximate solutions (3.14) & (3.17) with the exact EM solution (3.18) 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.

3.1.2 The Expectation Step

As we have seen above, the maximization step assumes the presence of some expectations. By inspecting subsection 3.1.1, we see we need to calculate the following moments.

$$E[\underline{\boldsymbol{X}}_d],\ \ Cov[\underline{\boldsymbol{X}}_d^*],\ \ E[\underline{\boldsymbol{D}}],\ \ E[\underline{\boldsymbol{D}}\boldsymbol{B}\underline{\boldsymbol{D}}^*],\ \text{ and }\ E[\underline{\boldsymbol{D}}^{-1}] \tag{3.19}$$

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 we can express the moments of $\boldsymbol{X}_d$ in terms of moments of $\boldsymbol{D}$. Specifically we have that

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

and

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

Moreover, we show in appendix A that

$$E[\underline{\boldsymbol{D}}\boldsymbol{B}\underline{\boldsymbol{D}}^*] = E[\underline{\boldsymbol{D}}]\boldsymbol{B}E[\underline{\boldsymbol{D}}^*] + \mathrm{Cov}[\underline{\boldsymbol{D}}]\mathrm{diag}(\boldsymbol{B}) \tag{3.20}$$

From above it follows that in order to calculate the expectations in (3.19), it is enough to calculate the following three moments

$$E[\mathrm{diag}(\boldsymbol{\mathcal{X}})],\ \ Cov[\mathrm{diag}(\boldsymbol{\mathcal{X}})]\ \ \&\ \ E[\mathrm{diag}(\boldsymbol{\mathcal{X}})^{-}1] \tag{3.21}$$

where the expectation is performed given the output $\underline{\boldsymbol{\mathcal{Y}}}$ and the most recent channel estimate $\hat{\underline{\boldsymbol{\mathcal{H}}}}$. In carrying out these expectations, we will assume that the elements of $\underline{\boldsymbol{\mathcal{X}}}$ are independent.[3]. With this in mind, it is easy to see that we can evaluate the moments in (3.21) and hence in (3.19) by calculating

$$E[\underline{\mathcal{X}}(l)],\ \ \mathrm{Cov}[\underline{\mathcal{X}}(l)] = E[|\underline{\mathcal{X}}(l)|^2] - |E[\underline{\mathcal{X}}(l)]|^2,\ \ E[\frac{1}{\underline{\mathcal{X}}(l)}]$$

Now assuming that $\underline{\mathcal{X}}(l)$ is drawn from the alphabet $A = \{A_1, \ldots, A_M\}$ with equal probability, it is can be shown that [42]

$$E[\underline{\mathcal{X}}(l)|\underline{\mathcal{Y}}(l), \underline{\mathcal{H}}(l)] = \frac{\sum_{j=1}^{M} A_j e^{-\frac{|\underline{\mathcal{Y}}(l)-\underline{\mathcal{H}}(l)A_j|^2}{\sigma^2}}}{\sum_{j=1}^{M} e^{-\frac{|\underline{\mathcal{Y}}(l)-\underline{\mathcal{H}}(l)A_j|^2}{\sigma^2}}} \tag{3.22}$$

$$E[|\underline{\mathcal{X}}(l)|^2|\underline{\mathcal{Y}}(l), \underline{\mathcal{H}}(l)] = \frac{\sum_{j=1}^{M} |A_j|^2 e^{-\frac{|\underline{\mathcal{Y}}(l)-\underline{\mathcal{H}}(l)A_j|^2}{\sigma^2}}}{\sum_{j=1}^{M} e^{-\frac{|\underline{\mathcal{Y}}(l)-\underline{\mathcal{H}}(l)A_j|^2}{\sigma^2}}} \tag{3.23}$$

$$E[\frac{1}{\underline{\mathcal{X}}(l)}|\underline{\mathcal{Y}}(l), \underline{\mathcal{H}}(l)] = \frac{\sum_{j=1}^{M} \frac{1}{A_j} e^{-\frac{|\underline{\mathcal{Y}}(l)-\underline{\mathcal{H}}(l)A_j|^2}{\sigma^2}}}{\sum_{j=1}^{M} e^{-\frac{|\underline{\mathcal{Y}}(l)-\underline{\mathcal{H}}(l)A_j|^2}{\sigma^2}}} \tag{3.24}$$

3.1.3 Summary of the EM Algorithm

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

1. Calculate the initial channel estimate $\hat{\boldsymbol{\mathcal{H}}}_0$ using pilots (3.9).

[3]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.

2. Calculate the moments of the input given the current channel estimate $\hat{\boldsymbol{\mathcal{H}}}_i$ and the output $\boldsymbol{\mathcal{Y}}$ using equations (3.22)-(3.24).

3. Calculate the channel estimate using either one of the methods (3.14), (3.17) or (3.18) outlined in Section 3.1.

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.

3.2 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 & hence is totaly 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 $\boldsymbol{\alpha}_d$. To this end, let's first develop a model for the time variation of the parameter $\boldsymbol{\alpha}_d$.

3.2.1 Developing a Frequency Domain Time-Variant Model

Consider the block fading model in (1.2) and lets assume for simplicity that the diagonal matrices $\boldsymbol{F}$ and $\boldsymbol{G}$ are actually scalar multiples of the identity, i.e.

$$\boldsymbol{F} = f\boldsymbol{I} \qquad \boldsymbol{G} = \sqrt{1-f^2}\boldsymbol{I}$$

where f is a function of Doppler frequency (see [42]). We will use the time domain model in (1.2) 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 j^{th} section of $\boldsymbol{\mathcal{H}}_i$, $\boldsymbol{\mathcal{H}}_i^{(j)}$, is related to $\boldsymbol{h}_i$ by

$$\underline{\boldsymbol{\mathcal{H}}}_i^{(j)} = \boldsymbol{Q}_{P+1}^{(j)}\boldsymbol{h}_i \tag{3.25}$$

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

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

or

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

where $\boldsymbol{V}_d^{+}$ is the pseudo inverse of $\boldsymbol{V}_d$. Multiplying both sides of (1.2) by $\boldsymbol{V}_d^{+}\boldsymbol{Q}_{P+1}^{(j)}$ yields a dynamical recursion for $\boldsymbol{\alpha}_d$

$$\boldsymbol{\alpha}_{d,i+1} = \boldsymbol{F}_\alpha \boldsymbol{\alpha}_{d,i} + \boldsymbol{G}_\alpha \boldsymbol{u}_i \tag{3.26}$$

where $\boldsymbol{F}_\alpha = f\boldsymbol{I}$ and $\boldsymbol{G}_\alpha = \sqrt{1-f^2}\boldsymbol{V}_d^{+}\boldsymbol{Q}_{P+1}^{(j)}$ 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 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 (3.26). As we have seen in section 3.1, the algorithm will consist of an initial estimation step, a maximization step, and an expectation step.

3.2.2 Initial (Pilot-Based) Channel Estimation

In the initial channel estimation step, the frequency domain system is described by equations (3.7) and (3.26), reproduced here for convenience.

$$\underline{\boldsymbol{\mathcal{Y}}}_{I_p,i} = \underline{\boldsymbol{X}}_{d,I_p,i}\boldsymbol{\alpha}_{d,i} + \underline{\boldsymbol{\mathcal{N}}}'_{I_p,i} \tag{3.27}$$

$$\boldsymbol{\alpha}_{d,i+1} = \boldsymbol{F}_\alpha \boldsymbol{\alpha}_{d,i} + \boldsymbol{G}_\alpha \boldsymbol{u}_i \tag{3.28}$$

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 (3.27)-(3.28)(see [55]), i.e., by implementing the following equations

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

$$\begin{aligned}
\boldsymbol{R}_{e,i} &= \boldsymbol{R}_{\underline{\mathcal{N}}'} + \underline{\boldsymbol{X}}_{d,I_p,i} \boldsymbol{P}_{i|i-1} \underline{\boldsymbol{X}}^*_{d,I_p,i} && (3.29)\\
\boldsymbol{K}_{f,i} &= \boldsymbol{P}_{i|i-1} \underline{\boldsymbol{X}}^*_{d,I_p,i} \boldsymbol{R}^{-1}_{e,i} && (3.30)\\
\hat{\boldsymbol{\alpha}}_{i|i} &= \left(\boldsymbol{I} - \boldsymbol{K}_{f,i} \underline{\boldsymbol{X}}_{d,I_p,i} \right) \hat{\boldsymbol{\alpha}}_{i|i-1} + \boldsymbol{K}_{f,i} \boldsymbol{\mathcal{Y}}_i && (3.31)\\
\hat{\boldsymbol{\alpha}}_{i+1|i} &= \boldsymbol{F}_\alpha \hat{\boldsymbol{\alpha}}_{i|i} && (3.32)\\
\boldsymbol{P}_{i+1|i} &= \boldsymbol{F}_\alpha \left(\boldsymbol{P}_{i|i-1} - \boldsymbol{K}_{f,i} \boldsymbol{R}_{e,i} \boldsymbol{K}^*_{f,i} \right) \boldsymbol{F}^*_\alpha + \frac{1}{\sigma_n^2} \boldsymbol{G}_\alpha \boldsymbol{G}^*_\alpha && (3.33)
\end{aligned}$$

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

$$\begin{aligned}
\boldsymbol{\lambda}_{i|T} &= \left(\boldsymbol{I}_{P+N} - \underline{\boldsymbol{X}}^*_{d,I_p,i} \boldsymbol{K}^*_{f,i} \right) \boldsymbol{F}^*_i \boldsymbol{\lambda}_{i+1|T} + \underline{\boldsymbol{X}}_{d,I_p,i} \boldsymbol{R}^{-1}_{e,i} \left(\boldsymbol{\mathcal{Y}}_i - \underline{\boldsymbol{X}}_{d,I_p,i} \hat{\boldsymbol{\alpha}}_{i|i-1} \right) && (3.34)\\
\hat{\boldsymbol{\alpha}}_{i|T} &= \hat{\boldsymbol{\alpha}}_{i|i-1} + \boldsymbol{P}_{i|i-1} \boldsymbol{\lambda}_{i|T} && (3.35)
\end{aligned}$$

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.

3.2.3 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 (3.5) and (3.26) also reproduced here for convenience

$$\begin{aligned}
\underline{\boldsymbol{\mathcal{Y}}}_i &= \underline{\boldsymbol{X}}_{d,i} \boldsymbol{\alpha}_{i,d} + \underline{\boldsymbol{\mathcal{N}}}'_i && (3.36)\\
\boldsymbol{\alpha}_{d,i+1} &= \boldsymbol{F}_\alpha \boldsymbol{\alpha}_{d,i} + \boldsymbol{G}_\alpha \boldsymbol{u}_i && (3.37)
\end{aligned}$$

If the data symbols $\underline{\boldsymbol{X}}_{d,i}$ were known, we would have employed the forward-backward Kalman-Filter (3.29)-(3.35) 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 [42] we can show that the FB Kalman filter needs to be applied to the following state space model

$$\begin{bmatrix} \underline{\boldsymbol{\mathcal{Y}}}_i \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} E[\underline{\boldsymbol{X}}_{d,i}] \\ Cov[\underline{\boldsymbol{X}}^*_{d,i}]^{\frac{1}{2}} \end{bmatrix} \boldsymbol{\alpha}_{i,d} + \begin{bmatrix} \underline{\boldsymbol{\mathcal{N}}}'_i \\ \mathbf{0} \end{bmatrix} \qquad (3.38)$$

$$\boldsymbol{\alpha}_{d,i+1} = \boldsymbol{F}_{\alpha}\boldsymbol{\alpha}_{d,i} + \boldsymbol{G}_{\alpha}\boldsymbol{u}_i \tag{3.39}$$

where the expectations in (3.38) are taken given the output $\underline{\boldsymbol{\mathcal{Y}}}_i$ and most recent channel estimate $\boldsymbol{\alpha}_{d,i}$. The expectations that appears in (3.38) are calculated as we did in Section 3.1.2. In contrast to the symbol by symbol EM algorithm of section 3.1, there are several ways of implementing the EM iterations in the time-correlated multi-symbol case. In the symbol by symbol algorithm of Section 3, there was one dimension to iterate against (channel estimation vs data detection). When the channels are time correlated over symbols as is the case here, there are several dimensions we can iterate against:

1. We can iterate between channel estimation & data detection.
2. And 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 filters here, the Cyclic and the Helix Kalman filters.

3.2.4 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 & data detection before invoking the Kalman to move to the next symbol. The iterations thus trace circles over the OFDM symbols which motivates the name Cyclic Kalman.

3.2.5 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

estimate and data detection before moving on the next symbol (whereas the cyclic Kalman iterates once between channel estimate and data estimate at each step). This allows 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.

3.2.6 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). Our simulation shows that invoking the code can have a profound effect on performance. Now the (hard) data obtained is more refined and hence can be used enhance the channel estimate by employing the FB Kalman again.

3.2.7 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 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.

3.2.8 Estimating the channel in the time domain

For a fair comparison, we need to compare the estimate the channel $\hat{\boldsymbol{\mathcal{H}}}^{(FD)}$ obtained in the previous section with the time domain based estimate. This is done by following the similar course for the time domain. Pruning the time domain input/output relation is given by equation (1.8) results in

$$\boldsymbol{\mathcal{Y}}_{I_p} = \boldsymbol{X}_{I_p}\boldsymbol{h} + \boldsymbol{\mathcal{N}}_{I_p} \tag{3.40}$$

Deriving the MMSE estimator for the above equation we get

$$\hat{\boldsymbol{h}} = \mathbf{R}_h \boldsymbol{X}_{I_p}^* [\sigma^2 \boldsymbol{I} + \boldsymbol{X}_{I_p} \boldsymbol{R}_h \boldsymbol{X}_{I_p}^*]^{-1} \underline{\boldsymbol{\mathcal{Y}}}_{I_p} \tag{3.41}$$

where R_h is the autocorrelation matrix of $\boldsymbol{h}$. This estimate is then improved by applying the EM algorithm to it, similar to the approach we have shown previously for frequency domain. The minimization in this case reduces to

$$\hat{\boldsymbol{h}}^{(TD)} = \arg\min_{h} \left\{ \|\boldsymbol{\mathcal{Y}} - E[\boldsymbol{X}]\boldsymbol{h}\|^2_{\boldsymbol{R}_{\mathcal{N}}^{-1}} + \|\boldsymbol{h}\|^2_{\mathbf{B}} + \|\boldsymbol{h}\|^2_{\boldsymbol{R}_h^{-1}} \right\}$$

where $\mathbf{B} = \boldsymbol{Q}_P^* Cov[\mathrm{diag}(\boldsymbol{\mathcal{X}}^*)]\mathbf{R}_{\mathcal{N}}^{-1}\boldsymbol{Q}_P$. The frequency domain based estimate is obtained as

$$\hat{\boldsymbol{\mathcal{H}}}^{(TD)} = \boldsymbol{Q}_P \underline{\hat{\boldsymbol{h}}}$$

3.3 Time Domain multiple access channel estimation

For fair comparison, we need to compare the frequency domain (LS and Kalman) receiver withe time domain counter part. 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 imput/output eqation for the j^{th} user is given by (see (3.1))

$$\underline{\boldsymbol{\mathcal{Y}}}_i^{(j)} = \mathrm{diag}(\underline{\boldsymbol{\mathcal{X}}}_i^{(j)})\underline{\boldsymbol{\mathcal{H}}}_i^{(j)} + \underline{\boldsymbol{\mathcal{N}}}_i^{(j)}$$

Now $\underline{\boldsymbol{\mathcal{H}}}_i^{(j)}$ is related to the impulse response by (see (3.25))

$$\underline{\boldsymbol{\mathcal{H}}}_i^{(j)} = \boldsymbol{Q}_{P+1}^{(j)}\boldsymbol{h}_i$$

where as described in Section 3.2.1, $\boldsymbol{Q}_{P+1}^{(j)}$ is $\boldsymbol{Q}_{P+1}$ pruned of all rows that dont belong to the j^{th} section. So, we can write

$$\underline{\boldsymbol{\mathcal{Y}}}_i^{(j)} = \mathrm{diag}(\underline{\boldsymbol{\mathcal{X}}}_i^{(j)})\boldsymbol{Q}_{P+1}^{(j)}\boldsymbol{h}_i + \underline{\boldsymbol{\mathcal{N}}}_i^{(j)} \tag{3.42}$$

Equation (3.42) 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 (1.2)

$$\boldsymbol{h}_{i+1} = \boldsymbol{F}\boldsymbol{h}_i + \boldsymbol{G}\boldsymbol{u}_i$$

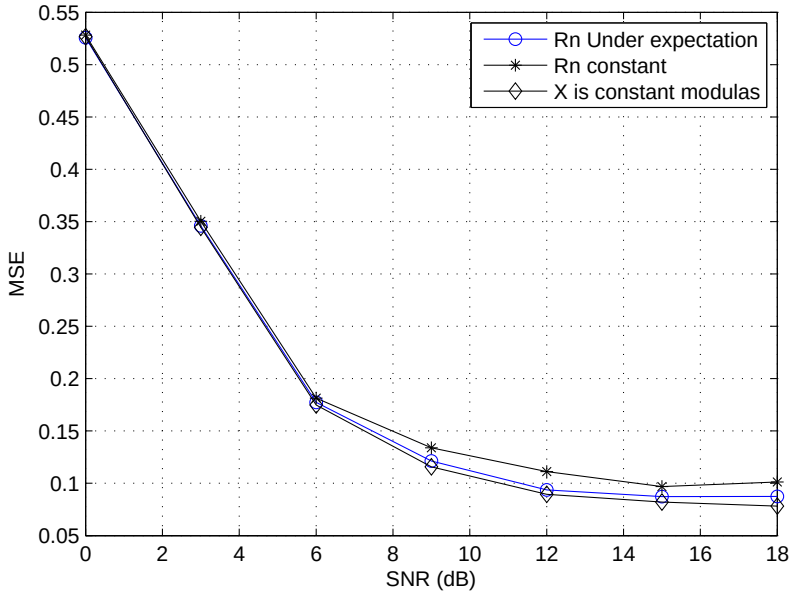
we can implement the various kind of Kalman filters discussed in the previous section for *time-domain* channel estimation.

3.4 Simulation Results

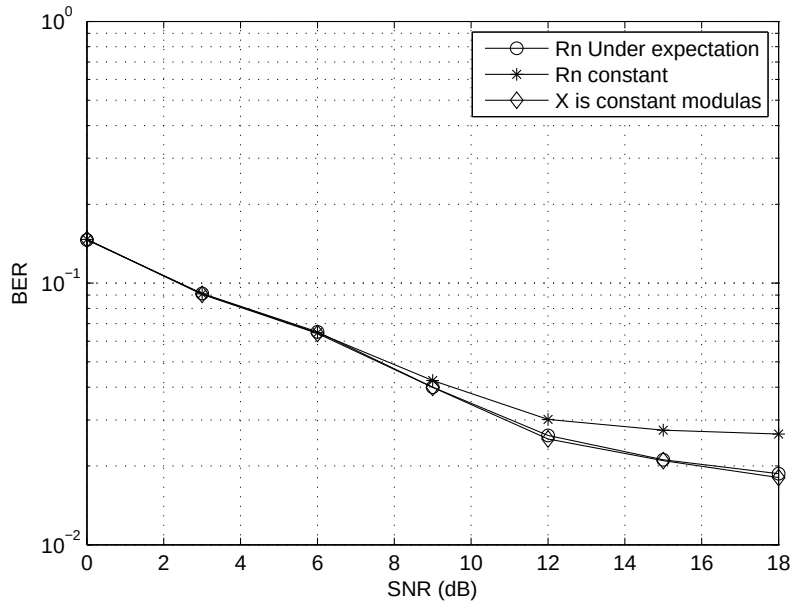
We consider an OFDM system that transmits 6 symbols with 64 carriers and a cyclic prefix of length $P = 15$ each with a time variation of $f = 0.9$. The data bits are mapped to 16 QAM through Gray coding (except for figures 3.1(a) and 3.1(b) which use a 4 QAM). The OFDM symbol serves 4 users each occupying 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 $E[|\underline{h}_0(k)|^2] = e^{-0.2k}$ 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 3.3 (see [42] also).

3.4.1 Effect of Modeling Noise

Figures 3.1(a) and 3.1(b) show the MSE and BER curves for the three cases considered in section 3.1 comparing the various treatment of the noise. We plot the figures 3.1(a) and 3.1(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.



(a)



(b)

Figure 3.1: (a) MSE (b) BER.

3.4.2 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 3.2-3.5 show the MSE of the channel estimate and the BER performance for these methods for the uncoded case. Figures 3.3 and 3.5 show the BER performance while figures 3.2 and 3.4 show the MSE at 16 and 24 pilots respectively. Comparing them, we see that increasing the number of pilots improves the LS estimate, with the frequency domain method wading better than the time domain method.

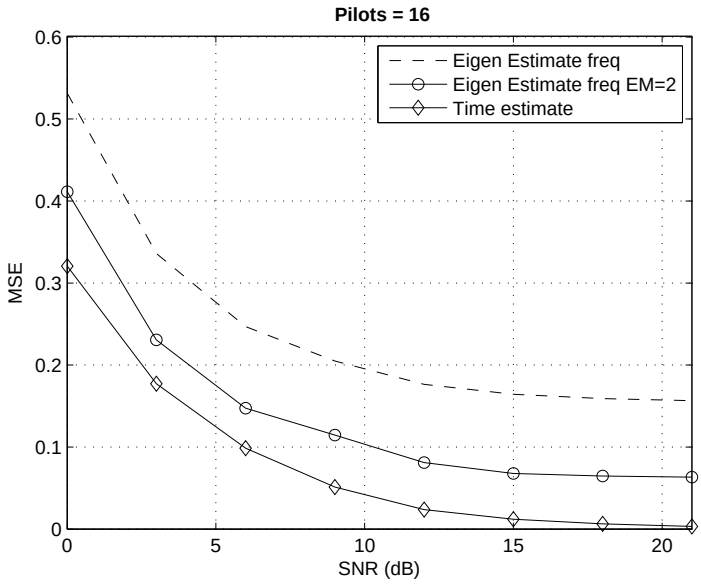


Figure 3.2: MSE comparison EM based Least Squares (16 pilots).

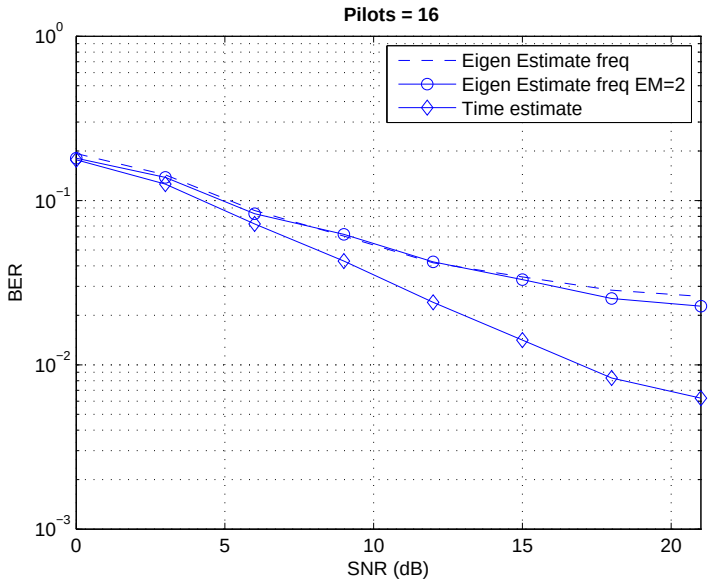


Figure 3.3: BER comparison EM based Least Squares (16 pilots).

3.4.3 Kalman Filter based Receivers

Figure 3.6(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 3.6(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 3.7 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 3.8 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

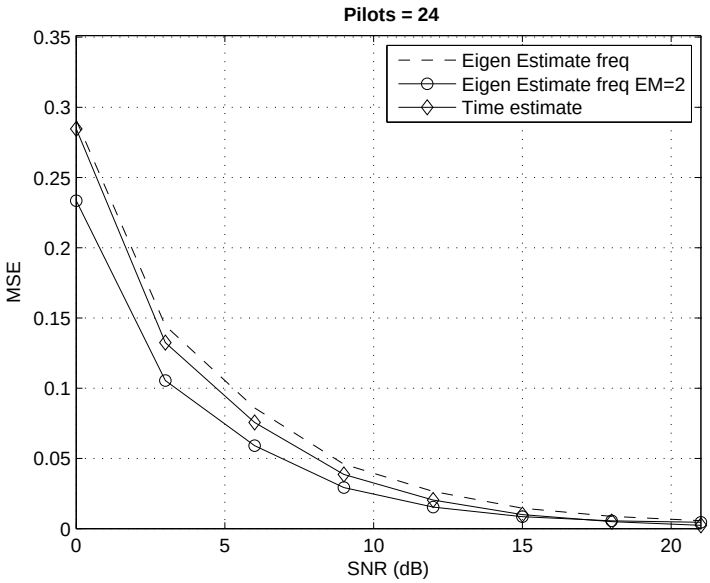


Figure 3.4: MSE comparison EM based Least Squares (24 pilots).

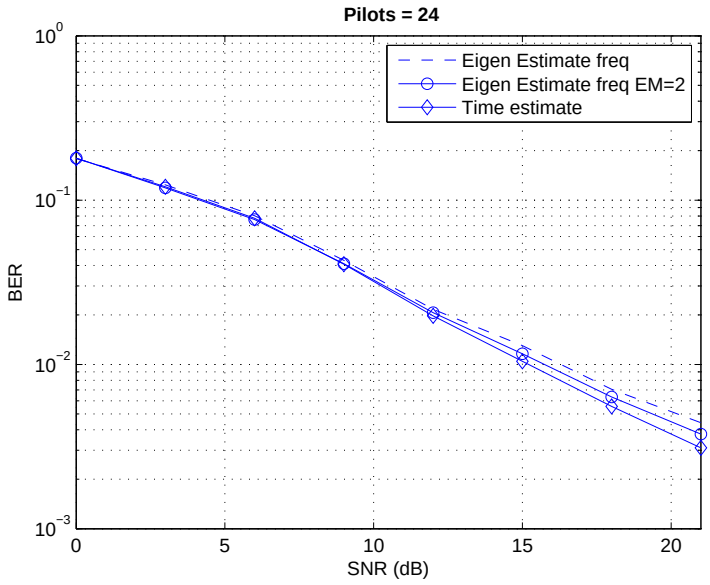
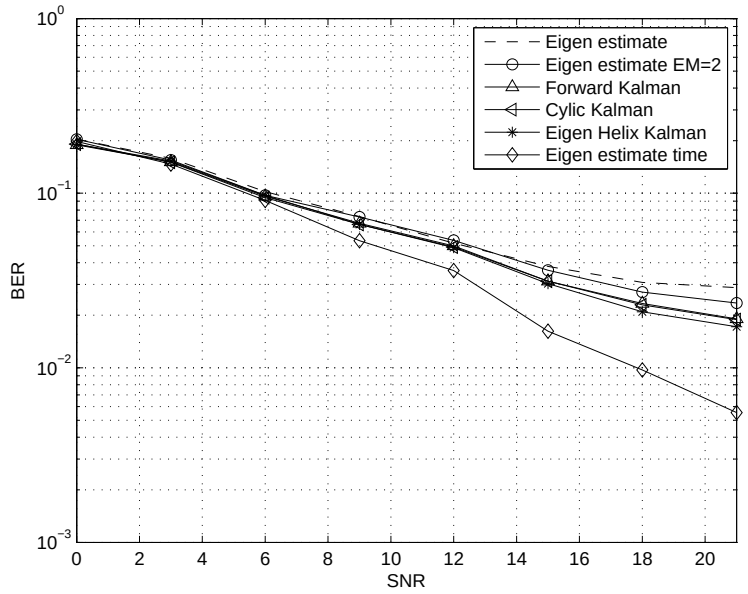
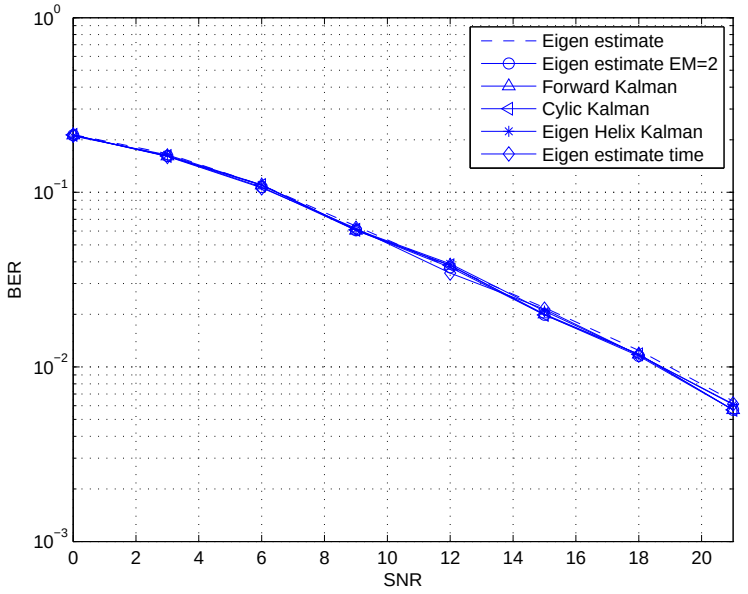


Figure 3.5: BER comparison EM based Least Squares (24 pilots).



(a)



(b)

Figure 3.6: BER comparison for various uncoded freq. domain methods (a) using 16 pilots (b) using 24 pilots.

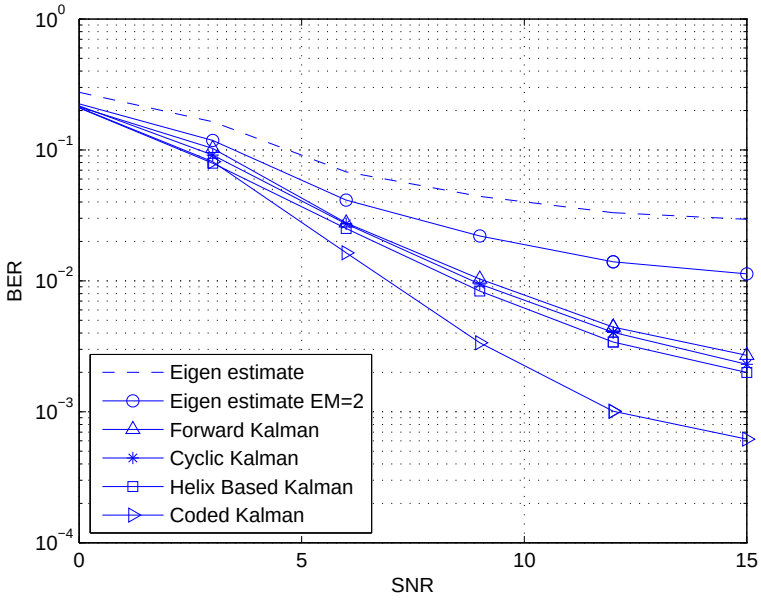


Figure 3.7: BER comparison for frequency domain coded methods (16 pilots).

multiple access case, the frequency domain technique fairs better than the time domain estimation method, while the coded Kalman outperforms all other techniques.

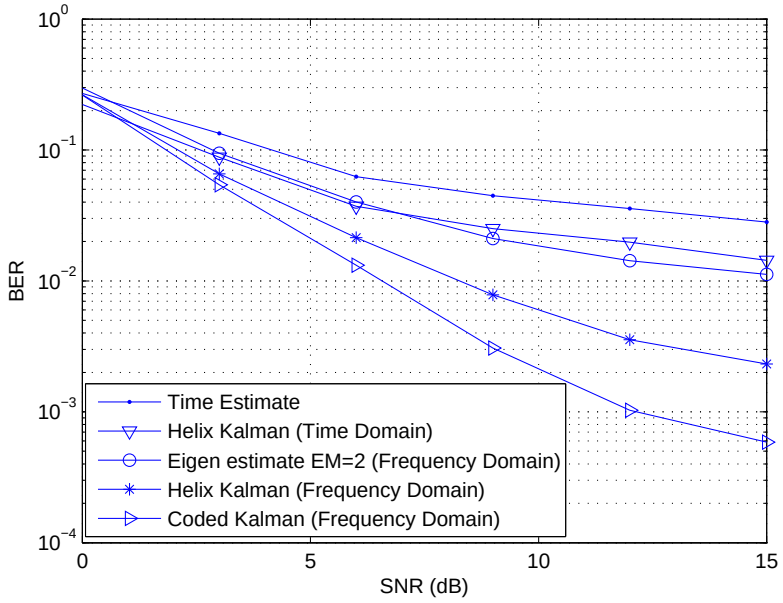


Figure 3.8: BER comparison for the coded case (16 pilots).

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 3.3. We plot figure 3.9 for 24 pilots with using the outercode and employing 6 eigenvalues per section to estimate the channel in the frequency domain. As we can see from the figure, the frequency domain Helix Kalman outperforms the time domain Helix Kalman.

3.5 Conclusion

We present an OFDM receiver design based on a semi-blind low complexity frequency domain channel estimation algorithm for multi-access OFDM system. As 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

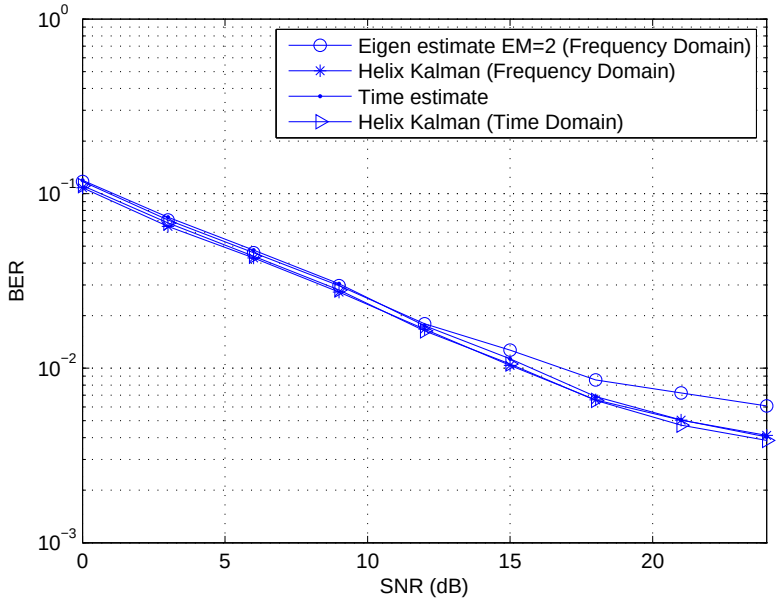


Figure 3.9: BER comparison of time and frequency domain uncoded techniques (24 pilots).

between channel and data recovery. Our receiver utilizes data constraints (finite alphabet set, code, and pilots) and channel constraints (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.

3.6 Appendix A

Now to calculate an expectation of the form $E[\underline{\boldsymbol{D}}\boldsymbol{B}\underline{\boldsymbol{D}}]$, which appears in (3.16), 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

$$E[\underline{\boldsymbol{D}}\boldsymbol{B}\underline{\boldsymbol{D}}^*] = E[\underline{\boldsymbol{D}}]\boldsymbol{B}E[\underline{\boldsymbol{D}}^*] \tag{3.43}$$

The identical forms, however, interact according to

$$\begin{aligned} E[\underline{\boldsymbol{D}}\boldsymbol{B}\underline{\boldsymbol{D}}] &= E[\underline{\boldsymbol{D}}\mathrm{diag}(\boldsymbol{B})\underline{\boldsymbol{D}}] && (3.44)\\ &= E[\underline{\boldsymbol{D}}\underline{\boldsymbol{D}}^*]E[\mathrm{diag}(\boldsymbol{B})] && (3.45) \end{aligned}$$

By combining (3.43) and (3.45), we see that

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

Chapter 4

Conclusions, Recommendations, Outcomes, and Publications

4.1 Conclusions

In this report, we studied the problem of channel estimation and data detection for OFDM transmission over block fading time-variant channels. Specifically, we pursued a frequency domain approach as opposed to the time-domain approach that is heavily pursued in literature. A frequency domain approach (for channel estimation and data detection) makes much more sense than a time-domain especially in multiple access OFDM.

We pursued two approaches for channel estimation, namely, an interpolation approach and an eigenvalue approach. The interpolation approach which was the basis of our original project proposal turned out to be ill-suited and produced unsatisfactory BER behavior. The eignenvalue approach proposed in Chapter 3 turned out to be the more natural approach to pursue for the problem at hand.

The eigenvalue approach was extended in several ways. It was applied to time-variant case where the channel changes from one symbol to the next according to some Doppler Frequency. Also, the performance of the channel estimation algorithm was enhanced by applying the expectation-maximization algorithm for joint data-detection and data-aided channel estimation.

The expectation maximization algorithm for channel estimation and data-detection boils is nothing but a Forward-Backward Kalman filter. The EM based Kalman implements two types of iterations (channel estimation and data detection). The performance of the receiver is affected by how these two types of iterations are scheduled. We particularly focused on two type of implementations called the cyclic and helix forward-backward Kalman. We showed that the helix Kalman has a superior performance. We also implemented the forward-only Kalman which is much less complex and suffers from no latency at the cost of inferior performance to the FB-Kalman.

4.2 Recommendations

Frequency domain receivers for multiple access OFDM are very promising. However, more research needs to be done to improve their performance. Specifically, the receiver assumed the knowledge of frequency correlation at the receiver. If we further assume that this correlation is available at the transmitter, one can further improve the performance by using that for optimal placement of pilots.

We have also assumed that the receiver has knowledge of such a priori information as the frequency correlation and time-correlation. This is not expected to be naturally available to the receiver, i.e., it need to be estimated. Also, one needs to investigate the effect of errors in the estimation of these parameters on the performance of the algorithms proposed in this project.

4.3 Summary of the Outcome of the Project Results

The project resulted in the following outcomes

1. Detailed literature review of available algorithms for channel estimation in the time and frequency domain.

2. Careful study of the advantages and challenges of channel estimation in the frequency and time domains

3. An interpolation-based algorithm for channel estimation in the frequency domain.

4. An eigenvalue-based approach for channel estimation in the frequency domain.

5. A forward-backward Kalman receiver for channel estimation in the frequency domain.

4.4 Publications that Resulted from the Project

Here is a summary of the publications that resulted from this work

1. 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.

2. T. Y. Al-Naffouri and A. Mukaddam "Frequency Domain Estimation of Multiple Access OFDM" IEEE International Conference on Signal Processing and Communication, Nov. 2007, Dubai, UAE.

3. T. Y. Al-Naffouri and Muhammad Saqib Sohail, "An EM based frequency domain channel estimation algorithm for multi access ODFM systems" *submitted to Elsavier Digital Signal Processing Journal.*

4. T. Y. Al-Naffouri and Muhammad Saqib Sohail "Frequency domain estimation of time varying multiple access OFDM channels: An EM approach", *to be submitted to ICASSP 2009.*

5. T. Y. Al-Naffouri, Ahmed Quadeer, and Muhammad Saqib Sohail "Iterative Forward-Backward Kalman Filtering for Data Recovery in (Multiuser) OFDM Communications" Book Chapter *Under preparation for submission to*

The project also resulted in the following Master thesis:

Mr. Muhammad Saqib Sohail, "Adaptive Algorithm for Channel Estimation: Using a Priori Information for Optimal Design," Electrical Engineering Department, King Fahd University of Petroleum and Minerals, June, 2008.

Bibliography

[1] W. Zhendao and G. B. Giannakis, "Wireless multicarrier communications," *IEEE Signal Processing Magazine*, vol. 17, no. 3, pp. 2948, May 2000.

[2] M. Speth, S. A. Fechtel, G. Fock, and H. Meyr, "Optimum receiver design for wireless broad-band systems using OFDM-Part 1,"*IEEE Trans. Commun.*, vol. 47, no. 11, pp. 1668-1677, Nov. 1999.

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

[4] 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.

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

[6] R. Negi and J. Cioffi, "Pilot tone selection for channel estimation in a mobile OFDM system," *IEEE Trans. Consumer Electr.*, vol. 44, no. 3, pp. 1122-1128, Aug. 1998.

[7] Y. Li, "Pilot-symbol-aided channel estimation for OFDM in wireless systems," *Proc. IEEE Vehicular Tech. Conf.*, 1999, vol. 2, pp. 1131-1135.

[8] J.K. Cavers, "An analysis of pilot symbol assisted modulation for rayleigh fading channels (mobile radio)," *IEEE Trans. Vehicular Tech.*, vol. 40, no. 4, pp. 686-693, Nov. 1991.

[9] F. Tufvesson and T. Maseng, "Pilot assisted channel estimation for OFDM in mobile cellular systems," *Proc. IEEE Vehicular Tech. Conf.*, May 1997, vol. 3, pp. 1693-1643.

[10] S. Ohno and G. B. Giannakis, "Optimal training and redundant precoding for block transmissions with application to wireless OFDM," *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Proc.*, 2001, pp. 2389-2392.

[11] J. W. Choi and Y. H. Lee, "Optimum Pilot Pattern for Channel Estimation in OFDM Systems", *IEEE Transactions on Wireless Communications*, Vol. 4, No. 5, September 2005.

[12] Seongwook Song and Andrew C. Singer, "Pilot-Aided OFDM Channel Estimation in the Presence of the Guard Band", *IEEE Transactions on Communications*, Vol. 55, No. 8, August 2007

[13] C. Tepedelenlioglu and G. B. Giannakis,"Transmitter redundancy for blind estimation and equalization of time- and frequency-selective channels," *IEEE Trans. Signal Processing*, vol. 48, pp. 2029-2043, July 2000.

[14] Chengyang Li, S.Roy,"Subspace-Based Blind Channel Estimation for OFDM by Exploiting Virtual Carriers," *IEEE Trans. on Wireless Comm.*, vol. 2, pp.141-150, Jan. 2003.

[15] B. Yang, K. Ben Letaief, R. Cheng, and Z. Cao, "Channel estimation for OFDM transmission in multipath fading channels based on parametric channel modeling," *IEEE Trans. Commun.*, vol. 49, no. 3, pp. 467-479, Mar. 2001.

[16] M. C. Necker and G. L. Stüber,"Totally blind channel estimation fpr OFDM over fast varying mobile channels," *Proc. IEEE Int. Conf. on Communications (ICC)*, New York, Apr. 2002, pp. 421-425.

[17] C. Shin, R. W. Heath, Jr. and E. J. Powers,"Blind Channel Estimation for MIMO OFDM Systems", *IEEE Transactions On Vehicular Technology*, Vol. 56, No. 2, March 2007

[18] H. Bölcskei, R.W. Heath, Jr., and A. J. Paulraj, "Blind channel identification and equalization in OFDM-based multi-antenna systems," I*EEE Trans. Signal Processing*, vol. 50, pp. 96-109, Jan. 2002.

[19] S. Zhou and G. B. Giannakis, "Finite-alphabet based channel estimation for OFDM and related multicarrier systems,"*IEEE Trans. Commun.*, vol. 49, pp. 1402-1414, Aug. 2001.

[20] X. Zhuang, Z. Ding, and A. L. Swindlehurst, "A statistical subspace method for blind channel identification in OFDM communications," *Proc. IEEE ICASSP*, Istanbul, Turkey, 2000.

[21] Y. Li, L. J. Cimini, and N. R. Sollenberger, "Robust channel estimation for OFDM systems with rapid dispersive fading channels", *IEEE Trans. Commun.*, vol. 46, no. 7, pp. 902-915, Jul. 1998.

[22] F. Sanzi and M. C. Necker, "Totally Blind APP Channel Estimation for Mobile OFDM Systems", *IEEE Communications Letters*, Vol. 7, No. 11, November 2003

[23] Xianbin Wang, Wu, Y., Chouinard, J.-Y., 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.

[24] Cirpan, H.A., Panayirci, E., Dogan, H., "Nondata-aided channel estimation for OFDM systems with space-frequency transmit diversity", *IEEE Transactions on Vehicular Technology*, Volume 55, Issue 2, March 2006 Page(s):449 - 457

[25] Wang Kunji, Zhang Jianhua, Li Chaojun, Huang Chen, "Semi-Blind OFDM Channel Estimation Using Reciever Diversity in the presence of virtual carriers",*IEEE First International Conference on Communications and Networking in China, ChinaCom '06*. Oct. 2006. Pages(s):1-4.

[26] Jianhua Liu, Jian Li,"Parameter estimation and error reduction for OFDM-based WLANs ", *IEEE Transactions on Mobile Computing*, Volume 3, Issue 2, April-June 2004 Page(s):152 - 163

[27] 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

[28] B. Song, L. Gui, and W. Zhang,"Comb Type Pilot Aided Channel Estimation in OFDM Systems with transmit diversity", *IEEE Transactions on Broadcasting*, Vol. 52, No. 1, March 2006

[29] 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

[30] G. Alrawi, T. Y. Al-Naffouri, A. Bahai, and J. Cioffi,"Exploiting errorcontrol coding and cyclic prefix in channel estimation for coded OFDM systems," *Proc. IEEE GlobeCom*, Nov. 2002, pp. 1152-1156.

[31] 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

[32] 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.

[33] J. Jiang, Tian Tang, Yongjin Zhang, Ping Zhang,"A channel Estimation Algorithm for OFDM based on PCC training Symbols and Frequency Domian Windowing", *IEEE ISCIT 2006*, pp. 629-632, 2006.

[34] I. Kang, M. P. Fitz, and S. B. Gelfand, "Blind estimation of multipath channel parameters: a modal analysis approach,"*IEEE Trans. Commun.*, vol. 47, no. 8, pp. 1140-1150, Aug. 1999.

[35] Z. Shengli and G. B. Giannakis, "Finite-alphabet based channel estimation for OFDM and related multicarrier systems," IEEE Trans. Commun., vol. 49, no. 8, pp. 1402-1414, Aug. 2001.

[36] Damián 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.

[37] 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.

[38] C. Aldana, E. de Carvalho, and J. M. Cioffi, "Channel estimation for multicarrier multiple input single output systems using the EM algorithm," *IEEE Trans. Signal Proc.*, vol. 51, no. 12, pp. 3280-3292, Dec. 2003.

[39] M.C. Vanderveen, A.-J. Van der Veen, and A. Paulraj, "Estimation of multipath parameters in wireless communications," *IEEE Trans. Signal Proc.*, vol. 46, no. 3, pp. 682-690, Mar. 1998.

[40] Jong-Ho Lee, Jae Choong Han, 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

[41] 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

[42] 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.

[43] Shengping Qin, Peide Liu, Xin Zhang, Laibo Zheng, 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.

[44] 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.

[45] 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

[46] G.B. Giannakis, "Filter banks for blind channel identification and equalization," *IEEE Signal Proc. Lett.*, vol. 4, no. 6, pp. 184-187, Jun. 1997.

[47] C. Komninakis, C. Fragouli, A. Sayed, and R. Wesel, "Multi-input multi-output fading channel tracking and equalization using Kalman estimation," *IEEE Trans. Signal Proc.*, vol. 50, no. 5, pp. 1065-1076, May 2002.

[48] R.A. Iltis, "Joint estimation of PN code delay and multipath using the extended Kalman filter," *IEEE Trans. Commun.*, vol. 38, no. 10, pp. 1677-1685, Oct. 1990.

[49] B. Lu, X. Wang, and Y. Li, "Iterative receivers for space-time blockcoded OFDM systems in dispersive fading channels," *IEEE Trans. Wireless Commun.*, vol. 1, no. 2, pp. 213-225, Apr. 2002.

[50] B. Muquet, M. de Courville, G. B. Giannakis, Z. Wang, and P. Duhamel, "Reduced complexity equalizers for zero-padded OFDM transmissions," *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Proc.*, Istanbul, Turkey, Jun. 2000, pp. 2973-2976

[51] A. Ghorokhov and J. Linnartz,"Robust OFDM receivers for dispersive time varying channels: equalization and channel acquisition," *Proc. IEEE Int. Conf. Commun.*, 2002, pp. 470-474.

[52] Y. Xie and C. N. Georghiades, "Two EM-type channel estimation algorithms for OFDM with transmitter diversity," *IEEE Trans. Commun.*, vol. 51, no. 1, pp. 106-115, Jan. 2003.

[53] M. Jiang, J. Akhtman, F. Guo, and L. Hanzo."Iterative joint channel estimation and multi-user detection for high-throughput multiple-antenna OFDM systems". *In Proceedings of the Spring'06 IEEE Vehicular Technology Conference*, Melbourne, Australia, May 2006.

[54] G. Sütber, *Principles of mobile communication*, Kluwer Academic, 2001.

[55] T. Kailath, A. H. Sayed and B. Hassibi, *Linear estimstion*, Prentice Hall, 2000.

[56] M. K. Tsatsanis, G. B. Giannakis, and G. Zhou, "Estimation and equalization of fading channels with random coefficients," *Signal Proc.*, vol. 53, no. 2-3, pp. 211-229, Sept. 1996.

[57] M. D. Estarki and E. Karami,"Joint Blind and Semi-Blind MIMO Channel Tracking and Decoding Using CMA Algorithm", *IEEE Vehicular Technology Conference*, Dublin, vol. 1, pp 2223-2227, April 2007.

[58] T. Cui and C. Tellambura,"Joint data detection and channel estimation for OFDM systems". *IEEE Transactions on Communications*, 54 (4). pp. 670-679, 2006.

[59] Ali H. Syed, *Fundamentals of Adaptive Filtering*, John Wiley & Sons, Inc. 2003.

[60] T. K. Moon, "The expectation-maximization algorithm", *IEEE Signal Proc. Mag.*, 13(6):47-60, Nov. 1996.

[61] G. Leus and M. Moonen, "Semi-blind channel estimation for block transmissions with non-zero padding",*Proc. Asilomar Conf. on Signals, Syst. and Computers*, Nov. 2001, pp. 762-766.