FREQUENCY DOMAIN ESTIMATION OF TIME VARYING CHANNELS IN OFDMA SYSTEMS: AN EM APPROACH

Tareq Y. Al-Naffouri

Electrical Engineering Department,
King Fahd University of Petroleum and Minerals,
Dhahran 31261, Saudi Arabia
naffouri@kfupm.edu.sa

ABSTRACT

OFDM modulation combines advantages of high achievable data rates and relatively easy implementation. However, for proper recovery of input, the OFDM receiver needs accurate channel information. Most algorithms proposed in literature perform channel estimation in time domain which increases computational complexity in multi-access situations where the user is only interested in part of the spectrum. In this paper, we propose a frequency domain algorithm for channel estimation in OFDMA systems. The algorithm performs eigenvalue decomposition of channel autocorrelation matrix and approximates channel frequency response seen by each user using the first few dominant eigenvectors. In a time variant environment, we derive a state space model for the evolution of the eigenmodes that help us to track them. This is done using a forward backward Kalman filter. The performance of the algorithm is further improved by employing a data-aided approach (based on expectation maximization).

Index Terms**—** Channel estimation, Kalman filtering, reduced order systems, OFDMA, iterative methods

1. INTRODUCTION

There has been increasing interest in OFDM as it combines the advantages of high achievable rates and easy implementation. This is reflected by the many standards that considered and adopted OFDM as the modulation scheme of choice for the physical layer. For proper operation of OFDM receiver, it needs accurate estimate of frequency domain carrier gains. This becomes especially challenging when the channel is time variant. In OFDMA systems, like WiMAX, each user is assigned a part of the spectrum and only needs to estimate its particular band. Estimating the entire spectrum would be over solving the problem. Many standards would not be able to support that due to limited number of pilots.

Historically, most channel estimation techniques estimate channel in time domain. A joint carrier frequency synchronization and channel estimation approach using expectation maximization (EM) approach is proposed in [1]. An EM based Kalman filtering approach is proposed in [2]. In [3], the authors have presented channel parameter estimation and error reduction algorithms to improve overall performance of OFDM system using Maximum Likelihood (ML) phase tracking and Least Squares (LS) phase fitting approaches while [4] proposed the use of implicit pilots for joint detection and channel estimation. In recent years, however, frequency domain channel estimation has attracted attention. In [5], a priori available information about interference structure is used to reduce the number of covariance parameters to be determined in the synchronous case using basis functions while in [6], an interpolated LS estimator is proposed by applying phase shifted samples in the frequency domain. Iterative solutions have also been explored due to the complexity of the problem despite the fact that iterative solutions are computationally intense. Researchers have explored various techniques for iterative channel estimation in OFDMA scenario i.e., Genetic algorithms [7], extrapolation [8] and Wiener filter [9] based channel estimation techniques.

A major problem associated with frequency domain channel estimation, as compared to time domain estimation techniques, is the increase in the number of parameters to be estimated. Iterative techniques iterate between data detection and channel estimation, using one to improve the estimate of the other. As data detection step is always performed in frequency domain, it would be convenient if channel estimation is also performed in frequency domain provided the parameter estimation space does not increase.

In this paper we introduce a class of iterative algorithms for channel and data recovery in frequency domain for OFDMA systems using the EM approach. An important aspect of the present work is that we employ a parameter reduction model to reduce the parameters to be estimated. The approach we pursue in this paper builds on the approach of [2]. A major difference from [2] is that this work performs both channel estimation and data detection in the frequency domain and specifically considers the OFDMA scenario.

The paper is organized as follows. In the following section, we introduce the system model including input/output equations in frequency domain. In Section 3, we introduce the parameter reduction model for frequency domain channel estimation and present an expectation maximization based iterative least square algorithm for channel and data recovery. In Section 4, we incorporate time correlation and use an EM forward-backward Kalman filter for channel estimation in frequency domain. Simulation results are presented in Section 6 and Section 7 gives concluding remarks.

2. SYSTEM MODEL

Consider the downlink of a narrowband OFDMA system with J users. Let the OFDM frame consist of K OFDM symbols where each OFDM symbol is of length N and consists of data and pilot tones. We consider comb-type pilots as they are more robust in fast fading channels than block-type pilots. The data bits to be transmitted are first encoded, punctured, interleaved, then mapped to QAM symbols using Gray code and then inserted at data tones in the OFDM symbols. A cyclic prefix (CP) is also appended to counter the intersymbol interference (ISI). At the receiver, we discard the CP and get the ISI free symbol. We consider a block fading channel $\boldsymbol{h}_i$ of length L. The channel varies from one OFDM symbol to the next according to the state space model

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

where i is the time index, $\boldsymbol{h}_0 \sim \mathcal{N}(\mathbf{0}, \mathbf{\Pi}_0)$ and $\boldsymbol{u}_i \sim \mathcal{N}(\mathbf{0}, \sigma_u^2\boldsymbol{I})$. The matrices $\boldsymbol{F}$ and $\boldsymbol{G}$ in (1) are square matrices of size L and are function of Doppler spread, power delay profile and transmit filter. The matrices are assumed to be known at the receiver and given as

$$\boldsymbol{F} = \begin{bmatrix} \alpha(0) & & \\ & \ddots & \\ & & \alpha(L-1) \end{bmatrix} \tag{2}$$

and

$$\boldsymbol{G} = \begin{bmatrix} \sqrt{1-\alpha^2(0)} & & \\ & \ddots & \\ & & \sqrt{(1-\alpha^2(L-1))e^{-\beta(L-1)}} \end{bmatrix} \tag{3}$$

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

The input/output relationship of the OFDM system is best described in the frequency domain as

$$\begin{aligned} \boldsymbol{\mathcal{Y}}_i &= \operatorname{diag}(\boldsymbol{\mathcal{X}}_i)\boldsymbol{\mathcal{H}}_i + \boldsymbol{\mathcal{N}}_i & (4) \\ &= \operatorname{diag}(\boldsymbol{\mathcal{X}}_i)\boldsymbol{Q}_L\boldsymbol{h}_i + \boldsymbol{\mathcal{N}}_i & (5) \end{aligned}$$

where $\boldsymbol{\mathcal{X}}_i$ is the transmitted signal, $\boldsymbol{\mathcal{Y}}_i$ is the received signal, $\boldsymbol{\mathcal{H}}_i$ is the FFT of $\boldsymbol{h}_i$ and $\boldsymbol{\mathcal{N}}_i$ is the additive white Gaussian noise $\mathcal{N}(\mathbf{0}, \sigma^2\boldsymbol{I})$, all of length N. Equation (5) follows from the FFT relationship

$$\boldsymbol{\mathcal{H}}_i = \boldsymbol{Q}\begin{bmatrix} \boldsymbol{h}_i \\ \mathbf{0} \end{bmatrix} = \boldsymbol{Q}_L\boldsymbol{h}_i \tag{6}$$

where $\boldsymbol{Q}$ is the FFT matrix and $\boldsymbol{Q}_L$ consists of the first L columns of $\boldsymbol{Q}$.

3. FREQUENCY DOMAIN CHANNEL ESTIMATION: A PARAMETER REDUCTION APPROACH

In this section we introduce a frequency domain based channel estimation algorithm. In an OFDMA system, the available spectrum is shared among a number of users. The number of subcarriers assigned may vary from one user to another. For simplicity, we will consider that the available bandwidth is shared equally among all active users. However, all the derivations and results presented here are also applicable to the case where different users are allocated different number of subcarriers. Consider such an OFDMA system in which J users are active simultaneously. From (4), the input/output equation for each user is given by

$$\underline{\boldsymbol{\mathcal{Y}}}_i^{(j)} = \operatorname{diag}(\underline{\boldsymbol{\mathcal{X}}}_i^{(j)})\underline{\boldsymbol{\mathcal{H}}}_i^{(j)} + \underline{\boldsymbol{\mathcal{N}}}_i^{(j)} \tag{7}$$

where $\underline{\boldsymbol{\mathcal{Y}}}_i^{(j)}$, $\underline{\boldsymbol{\mathcal{X}}}_i^{(j)}$, $\underline{\boldsymbol{\mathcal{H}}}_i^{(j)}$ and $\underline{\boldsymbol{\mathcal{N}}}_i^{(j)}$ are the j^{th} sections of $\boldsymbol{\mathcal{Y}}_i$, $\boldsymbol{\mathcal{X}}_i$, $\boldsymbol{\mathcal{H}}_i$ and $\boldsymbol{\mathcal{N}}_i$ respectively corresponding to the j^{th} user. Let the pilot locations within each section be denoted by $I_p^{(j)}$ then the pilot/output equation corresponding to (7) is given as

$$\underline{\boldsymbol{\mathcal{Y}}}_{I_p} = \operatorname{diag}(\underline{\boldsymbol{\mathcal{X}}}_{I_p})\underline{\boldsymbol{\mathcal{H}}}_{I_p} + \underline{\boldsymbol{\mathcal{N}}}_{I_p} \tag{8}$$

We suppress the dependence of (8) on the section (user) index j and on the time index i for notational convenience. Here $\boldsymbol{\mathcal{X}}_{I_p}$ denotes the matrix $\boldsymbol{\mathcal{X}}$ pruned of the rows that don't belong to I_p. Henceforth, we will use this simplified notation. It is desirable to keep the number of pilots to a minimum for high bandwidth efficiency.

Transforming the channel estimation step from time domain to frequency domain increases the estimation parameter space. A remedy to this problem is the singular value decomposition approach suggested in [10]. Based on this we resort to a model reduction approach starting from the autocorrelation function of $\underline{\boldsymbol{\mathcal{H}}}$, i.e. $\boldsymbol{R}_{\underline{\boldsymbol{\mathcal{H}}}}$. We find its eigenvalue decomposition as $\boldsymbol{R}_{\underline{\boldsymbol{\mathcal{H}}}} = \sum_{l=1}^{L_f} \lambda_l \mathbf{v}_l \mathbf{v}_l^T$ 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. The eigenvalue decomposition is performed only once hence the overhead incurred by it does not burden the receiver. Using this decomposition, $\underline{\boldsymbol{\mathcal{H}}}$ can be represented 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 $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_{L_f})$. Using the dominant eigenvalues to represent $\underline{\boldsymbol{\mathcal{H}}}$ and treating the rest as modeling noise[1], we get

$$\underline{\boldsymbol{\mathcal{H}}} = \mathbf{V}_d \boldsymbol{\alpha}_d + \mathbf{V}_n \boldsymbol{\alpha}_n \tag{9}$$

The first term here is the dominant component while the second term is considered as modeling noise. Substituting (9) in (7) results in

$$\begin{aligned} \underline{\boldsymbol{\mathcal{Y}}} &= \mathrm{diag}(\underline{\boldsymbol{\mathcal{X}}}) \boldsymbol{V}_d \boldsymbol{\alpha}_d + \underline{\boldsymbol{\mathcal{N}}} + \mathrm{diag}(\underline{\boldsymbol{\mathcal{X}}}) \boldsymbol{V}_n \boldsymbol{\alpha}_n & (10) \\ &= \underline{\boldsymbol{X}}_d \boldsymbol{\alpha}_d + \underline{\boldsymbol{\mathcal{N}}}' & (11) \end{aligned}$$

where $\underline{\boldsymbol{X}}_d = \mathrm{diag}(\underline{\boldsymbol{\mathcal{X}}}) \boldsymbol{V}_d$ and $\underline{\boldsymbol{\mathcal{N}}}' = \underline{\boldsymbol{\mathcal{N}}} + \underline{\boldsymbol{X}}_n \boldsymbol{\alpha}_n$ with $\underline{\boldsymbol{X}}_n = \mathrm{diag}(\underline{\boldsymbol{\mathcal{X}}}) \boldsymbol{V}_n$. We consider $\underline{\boldsymbol{\mathcal{N}}}'$ to be zero mean white Gaussian noise with autocorrelation

$$\boldsymbol{R}_{\underline{\mathcal{N}}'} = \boldsymbol{R}_{\underline{\mathcal{N}}} + \mathrm{diag}(\underline{\boldsymbol{\mathcal{X}}}) \boldsymbol{V}_n \mathrm{diag}(\boldsymbol{\Lambda}_n) \boldsymbol{V}_n^* \mathrm{diag}(\underline{\boldsymbol{\mathcal{X}}})^* \tag{12}$$

where $\boldsymbol{\Lambda}_n$ is a diagonal matrix comprising of the non-dominant ordered eigenvalues of $\boldsymbol{R}_{\underline{\boldsymbol{\mathcal{H}}}}$. We can use (11) to construct a pilot/output equation (similar to (8)) as

$$\underline{\boldsymbol{\mathcal{Y}}}_{I_p} = \underline{\boldsymbol{X}}_{I_p} \boldsymbol{\alpha}_d + \underline{\boldsymbol{\mathcal{N}}}'_{I_p} \tag{13}$$

Using equation(13), the parameter $\boldsymbol{\alpha}_d$ can be estimated from the log likelihood function

$$\hat{\boldsymbol{\alpha}}_d^{\mathsf{MAP}} = \arg\max_{\alpha_d} \left\{ \ln\, p(\underline{\boldsymbol{\mathcal{Y}}}_{I_p} | \underline{\boldsymbol{X}}_{d,I_p}, \boldsymbol{\alpha}_d) + \ln\, p(\boldsymbol{\alpha}_d) \right\} \tag{14}$$

As the noise is white Gaussian, the MAP estimate of $\hat{\boldsymbol{\alpha}_d}$ is given by

$$\hat{\boldsymbol{\alpha}}_d = \arg\min_{\alpha_d} \left\{ \| \underline{\boldsymbol{\mathcal{Y}}}_{I_p} - \underline{\boldsymbol{X}}_{d,I_p} \boldsymbol{\alpha}_d \|^2_{\boldsymbol{R}^{-1}_{\underline{\mathcal{N}}'_{I_p}}} + \|\boldsymbol{\alpha}_d\|^2_{\boldsymbol{\Lambda}_d^{-1}} \right\} \tag{15}$$

which simplifies to

$$\hat{\boldsymbol{\alpha}}_d = \boldsymbol{\Lambda}_d \underline{\boldsymbol{X}}^*_{d,I_p} \left[\boldsymbol{R}_{\underline{\mathcal{N}}'_{I_p}} + \underline{\boldsymbol{X}}_{d,I_p} \boldsymbol{\Lambda}_d \underline{\boldsymbol{X}}^*_{d,I_p} \right]^{-1} \underline{\boldsymbol{\mathcal{Y}}}_{I_p} \tag{16}$$

The estimate of the channel of the j^{th} user is then obtained by $\hat{\underline{\boldsymbol{\mathcal{H}}}} = \mathbf{V}_d \hat{\boldsymbol{\alpha}}_d$. In order to fully exploit the data constraints of the OFDM system, we will use the EM approach. So instead of maximizing (14), we maximize an average form of the log likelihood function. Starting from an initial estimate calculated using only pilots, the estimate is improved iteratively with the k^{th} estimate given by

$$\hat{\boldsymbol{\alpha}}_d^{(k)} = \arg\max_{\alpha_d} \left\{ E_{\mathcal{X}_i | \mathcal{Y}_i, \hat{\boldsymbol{\alpha}}_d{}^{(k-1)}} \ln\, p(\underline{\boldsymbol{\mathcal{Y}}} | \underline{\boldsymbol{X}}_d, \boldsymbol{\alpha}_d) + \ln\, p(\boldsymbol{\alpha}_d) \right\} \tag{17}$$

The expectation is taken with respect to the input given the output and the most recent parameter estimate $\hat{\boldsymbol{\alpha}}_d{}^{(k-1)}$. The iterations maybe performed for a specified number of times or until some predetermined condition is fulfilled. We drop the iteration index k from the equations henceforth for notational convenience. For the data aided case, the estimate of $\boldsymbol{\alpha}_d$ is given as

$$\hat{\alpha}_d = \arg\min_{\alpha_d} \left\{ E\left[\| \underline{\boldsymbol{\mathcal{Y}}} - \underline{\boldsymbol{X}}_d \boldsymbol{\alpha}_d \|^2_{\boldsymbol{R}^{-1}_{\underline{\mathcal{N}}'}} \right] + \|\boldsymbol{\alpha}_d\|^2_{\boldsymbol{\Lambda}_d^{-1}} \right\} \tag{18}$$

It turns out that solving this minimization is tricky as the noise $\underline{\boldsymbol{\mathcal{N}}}'$ itself is dependent on the input. We can proceed in one of the following three ways: 1) neglect the modeling noise altogether, in this case $\underline{\boldsymbol{\mathcal{N}}}'$ will no longer be dependent on the input and solving the minimization in (18) would be straight forward. A consequence of this approach would be some loss in accuracy. 2) split the expectation into two parts, one taken over the noise $\underline{\boldsymbol{\mathcal{N}}}'$ and the second taken *independently* over the data. This would yield a good approximation of (18). 3) if the signal is constant modulus, the expectation can be calculated exactly. It turns out that splitting the expectation yields almost comparable results as calculating the exact solution. We compare the MSE for these three method for the (constant modulus) 4 QAM case (see Section 6 for the simulation setup).

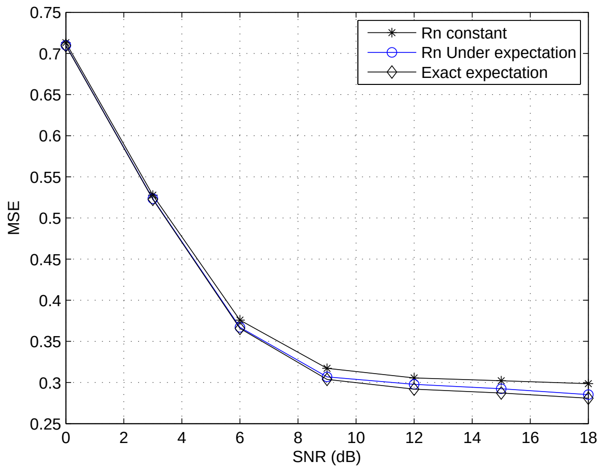

Fig. 1. Comparison of the three approaches to the EM based algorithm.

4. USING TIME-CORRELATION TO IMPROVE THE CHANNEL ESTIMATE

The above iterative technique for channel estimation does not use any time-correlation information. As the practical channel will exhibit some time correlation, this information can be used to enhance the estimate of $\boldsymbol{\alpha}_d$ further. Let us first develop a model for the variation of the parameter $\boldsymbol{\alpha}_d$.

[1]We use the condition $\frac{\lambda_j}{\lambda_{j+1}} > 5$ to determine the dominant eigenvalues.

4.1. Developing a Frequency Domain Time-Variant Model

For simplicity, let us assume that the diagonal matrices $\boldsymbol{F}$ and $\boldsymbol{G}$ in (1) are actually scalar multiples of the identity matrix, i.e. $\boldsymbol{F} = f\mathbf{I}$ and $\boldsymbol{G} = \sqrt{1-f^2}\mathbf{I}$. We will derive a model similar to the time domain model in (1). From (6), the j^{th} section of $\boldsymbol{\mathcal{H}}_i$ (corresponding to the spectrum of the j^{th} user), $\underline{\boldsymbol{\mathcal{H}}}_i^{(j)}$, is related to $\boldsymbol{h}_i$ by

$$\underline{\boldsymbol{\mathcal{H}}}_i^{(j)} = \boldsymbol{Q}_P^{(j)}\boldsymbol{h}_i \tag{19}$$

where $\boldsymbol{Q}_P^{(j)}$ is the j^{th} section of $\boldsymbol{Q}_P$ obtained by pruning of all its rows except those of the j^{th} section. Replacing $\underline{\boldsymbol{\mathcal{H}}}_i^{(j)}$ by its representation using the dominant parameters $\boldsymbol{\alpha}_d$, we get

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

or

$$\boldsymbol{\alpha}_{d,i} = \boldsymbol{V}_d^{+}\boldsymbol{Q}_P^{(j)}\boldsymbol{h}_i \tag{20}$$

where $\boldsymbol{V}_d^{+}$ is the pseudo inverse of $\boldsymbol{V}_d$. From (1) and (20) we get the dynamical recursion

$$\boldsymbol{\alpha}_{d,i+1} = \boldsymbol{F}_\alpha\boldsymbol{\alpha}_{d,i} + \boldsymbol{G}_\alpha\boldsymbol{u}_i \tag{21}$$

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

$$E[\boldsymbol{\alpha}_{d,0}\boldsymbol{\alpha}_{d,0}^*] = \boldsymbol{\Lambda}_d$$

Now that we have a model for time variation for $\boldsymbol{\alpha}_d$ we can incorporate the time correlation information given by the dynamical equation (21) to improve the estimate of $\boldsymbol{\alpha}_d$. The algorithm will again comprise of two steps. An initial pilot based estimation step and then an iterative data aided step to improve the original estimate.

4.2. Initial (Pilot-Based) Channel Estimation

In this step, we use only the pilots to estimate the parameter $\boldsymbol{\alpha}_d$. This initial pilot based estimate is computed from the state space system described by equations (13) and (21). For a sequence of $T+1$ pilot bearing symbols, the optimum estimate of $\{\boldsymbol{\alpha}_{i,d}\}_{i=0}^{T}$ is given by applying a forward-backward Kalman to equations (13) and (21).

Forward run: The initial conditions, for $i = 1,\ldots,T$, are $\boldsymbol{P}_{0|-1} = \boldsymbol{\Pi}_0$ and $\boldsymbol{\alpha}_{0|-1} = \mathbf{0}$. The Forward run of the Kalman filter is described by the following set of equations

$$\boldsymbol{R}_{e,i} = \boldsymbol{R}_{\underline{\mathcal{N}}} + \underline{\boldsymbol{X}}_{iI_p}\boldsymbol{P}_{i|i-1}\underline{\boldsymbol{X}}_{iI_p}^* \tag{22}$$

$$\boldsymbol{K}_{f,i} = \boldsymbol{P}_{i|i-1}\underline{\boldsymbol{X}}_{iI_p}^*\boldsymbol{R}_{e,i}^{-1} \tag{23}$$

$$\hat{\boldsymbol{\alpha}}_{i|i} = \left(\boldsymbol{I} - \boldsymbol{K}_{f,i}\underline{\boldsymbol{X}}_{iI_p}\right)\hat{\boldsymbol{\alpha}}_{i|i-1} + \boldsymbol{K}_{f,i}\boldsymbol{\mathcal{Y}}_i \tag{24}$$

$$\hat{\boldsymbol{\alpha}}_{i+1|i} = \boldsymbol{F}_\alpha\hat{\boldsymbol{\alpha}}_{i|i} \tag{25}$$

$$\boldsymbol{P}_{i+1|i} = \boldsymbol{F}_\alpha\left(\boldsymbol{P}_{i|i-1} - \boldsymbol{K}_{f,i}\boldsymbol{R}_{e,i}\boldsymbol{K}_{f,i}^*\right)\boldsymbol{F}_\alpha^* + \frac{1}{\sigma_n^2}\boldsymbol{G}_\alpha\boldsymbol{G}_\alpha^* \tag{26}$$

Backward run: The backward run ($i = T, T-1, \ldots, 0$) starts from $\boldsymbol{\lambda}_{T+1|T} = \mathbf{0}$ and is described by

$$\boldsymbol{\lambda}_{i|T} = \left(\boldsymbol{I}_{P+N} - \underline{\boldsymbol{X}}_{iI_p}^*\boldsymbol{K}_{f,i}^*\right)\boldsymbol{F}_i^*\boldsymbol{\lambda}_{i+1|T} + \underline{\boldsymbol{X}}_{iI_p}\boldsymbol{R}_{e,i}^{-1}\left(\boldsymbol{\mathcal{Y}}_i - \boldsymbol{X}_i\hat{\boldsymbol{\alpha}}_{i|i-1}\right) \tag{27}$$

$$\hat{\boldsymbol{\alpha}}_{i|T} = \hat{\boldsymbol{\alpha}}_{i|i-1} + \boldsymbol{P}_{i|i-1}\boldsymbol{\lambda}_{i|T} \tag{28}$$

where $\hat{\boldsymbol{\alpha}}_{i|T}$ gives us an initial pilot based estimate used to boot start the data-aided part of the algorithm.

4.3. Iterative (Data-Aided) Channel Estimation

In the data-aided step, we use data tones of the user to improve the pilot based estimate of $\boldsymbol{\alpha}_d$ in an iterative manner. The system is described by the state space equations (11) and (21). However, the receiver has no prior knowledge of the actual data symbols $\underline{\boldsymbol{X}}_i$ hence we use instead an estimate of the transmitted data. The Forward Backward Kalman is therefore applied to the following state space model(see [2] for a proof)

$$\begin{bmatrix} \underline{\boldsymbol{\mathcal{Y}}}_i \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} E[\underline{\boldsymbol{X}}_d] \\ Cov[\underline{\boldsymbol{X}}_d^*]^{\frac{1}{2}} \end{bmatrix}\boldsymbol{\alpha}_{i,d} + \begin{bmatrix} \underline{\boldsymbol{\mathcal{N}}}_i' \\ \mathbf{0} \end{bmatrix} \tag{29}$$

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

Recall that $\underline{\boldsymbol{X}}_d = \text{diag}(\underline{\boldsymbol{\mathcal{X}}})\boldsymbol{V}_d$. Thus,

$$E[\underline{\boldsymbol{X}}_d] = E[\text{diag}(\underline{\boldsymbol{\mathcal{X}}})]\boldsymbol{V}_d$$
$$Cov[\underline{\boldsymbol{X}}_d^*] = \boldsymbol{V}_d^*Cov[\text{diag}(\underline{\boldsymbol{\mathcal{X}}}^*)]\boldsymbol{V}_d$$

i.e., we need to evaluate the first two moments of $\boldsymbol{\mathcal{X}}$ (given $\underline{\boldsymbol{\mathcal{Y}}}$ and the most recent estimate $\widehat{\alpha}_{\boldsymbol{d}}$). The first moment is evaluated element-by-element. Moreover, in carrying out the second expectation, we will assume that the elements of $\underline{\boldsymbol{\mathcal{X}}}$ are independent.[2] With this assumption, it is easy to see that we can calculate the first two moments by calculating

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

Given that $\underline{\mathcal{X}}(l)$ is drawn from the alphabet set $A = \{A_1, \ldots, A_M\}$ with equal probability, the first and second moments are calculated along the same lines as in [2] and given as

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

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

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

While there was just a single implementation of the symbol by symbol algorithm of Section 3, several implementations are possible for the time-correlation case depending on the structure of the data-aided iterations in the FB Kalman filter. Next we briefly described the implementation structure of these variations.

4.4. Cyclic FB Kalman

In the cyclic FB Kalman filter, the algorithm starts with calculating the initial pilot based estimate. This initial estimate is used to kick start the data-aided step where the algorithm obtains the channel estimate (forward run) traversing over the entire OFDM block and then performs the back ward run of the Kalman filter. This channel estimate is then used to obtain a more refined data estimate and the entire cycle is repeated.

4.5. Helix FB Kalman

In the Helix FB Kalman filter, the initial estimate is obtained using the pilots. Using this initial estimate, the algorithm then iterates between channel and data estimation within one symbol and after completing the designated number of iterations, moves to the next symbol where the same iterations are repeated again (Forward run). Once the algorithm has iterated over all the symbols within the OFDM block, it performs the back ward run of the Kalman filter.

4.6. Using Code to Enhance the Estimate

When an outer code extending over the entire OFDM block is implemented, the estimate can be further enhanced. As can be seen, the channel estimate obtained by the above two implementation depends on the data estimate. Therefore, improving the data estimate will improve the channel estimate. Using the code, we can reduce the number of errors in the received data and use this corrected estimate of data to enhance the channel estimate. This results in marked improvement in the bit error rate (BER) as corroborated by our simulation results.

5. COMPUTATIONAL COMPLEXITY

The data detection step is always performed in frequency domain. For iterative methods, like the ones proposed here, performing channel estimation in time domain will require 2 extra FFT operations, one to transform the time domain channel estimate to frequency domain (to aid in the next data detection step) and the second to transform the resulting data estimate back to time domain (for further iterations). Thus not only does the time domain estimation require more computations, it also introduces latency in the process. Frequency domain channel estimation methods avoid both of these issues. Table 1 shows the computational complexities of the various algorithms presented in the paper in terms of approximate number of complex multiplications. In the table, $N_J = \frac{N}{J}$ represents the subcarriers assigned to the j^{th} user, d and n represent the dominant and non-dominant eigenvalues, k is the number of EM iterations, M is the size of the alphabet set and P represents the number of time domain channel taps.

6. SIMULATION RESULTS

We consider an OFDMA system with 4 users. The OFDM block length is 6. Each OFDM symbol has 64 carriers out of which 16 are reserved for pilots. A cyclic prefix of length 15 is used. The outer code is a 1/2 rate convolutional code. The channel model is considered to be block fading described by state space equation (1) with a time variation of $f = 0.9$. It consists of 15 complex tabs (worst case scenario) and has an exponential delay profile $E[|\underline{h}_0(k)|^2] = e^{-0.2k}$.

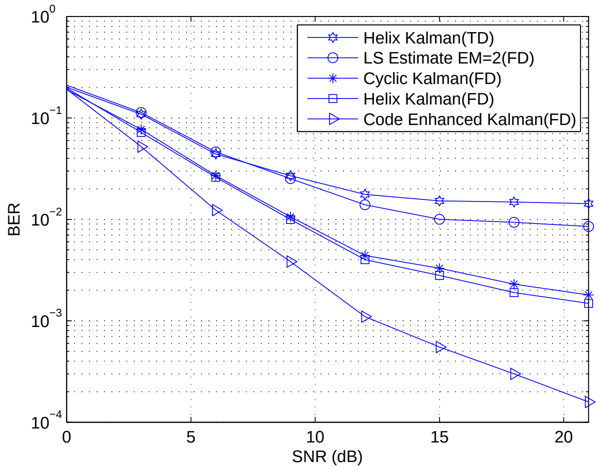

Fig. 2. BER comparison of entire system (TD: Time Domain, FD: Frequency Domain).

Figure 1 compares the three approaches to solve (18) for the case when the input signal is of constant modulus. We can see that neglecting the modeling noise results in a loss of accuracy while splitting the expectation into two independent expectations gives us almost comparable results as calculating the expectation exactly.

Figure 2 compares BER performances of frequency domain channel estimation techniques described in Section 4 with the equivalent time domain FB Helix Kalman based on the technique presented in [2]. FD and TD indicate whether the method works in frequency or time domain. The graph shows the overall BER performance of the system for all k users. Figure 3 gives the comparison of these techniques when the channel is correlated. We can see that the frequency domain Kalman based methods perform better than the time domain methods with the coded Kalman providing the best performance.

Table 1. COMPUTATIONAL COMPLEXITIES.

Algorithm	Complexity
Helix Kalman (FD)	$2N_J^3 + (3kM + 3d)N_J^2 + (5d^2 + 5d + kd + Pd)N_J$
Helix Kalman (TD)	$2N_J^3 + (3kM + 4P)N_J^2 + (4P^2 + 6P + 2kP)N_J$
Forward Kalman (FD)	$2kN_J^3 + k(3M + 5d)N_J^2 + k(5d^2 + 4d + Pd)N_J$
Cyclic Kalman (FD)	$2kN_J^3 + k(3M + 5d)N_J^2 + k(5d^2 + 4d + Pd)N_J$
Code Enhanced Kalman (FD)	$4N_J^3 + 2(3kM + 5d)N_J^2 + 2(5d^2 + 3d + kd + Pd)N_J$

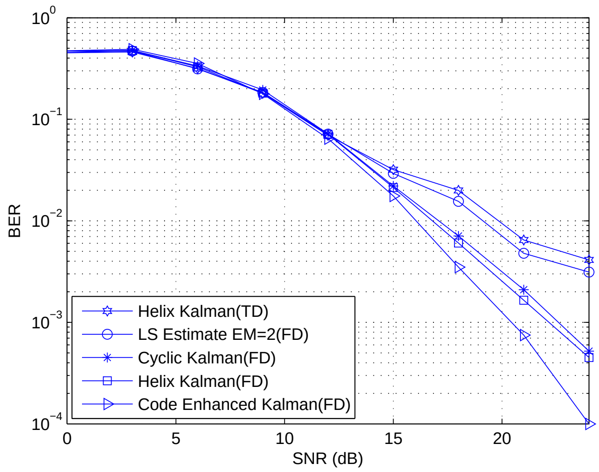

Fig. 3. BER comparison for correlated case (TD: Time Domain, FD: Frequency Domain).

7. CONCLUSION

A frequency domain data aided algorithm for channel estimation in OFDMA systems is proposed. In OFDMA systems, users are only interest in the band allocated to them and not the entire spectrum so frequency domain approach to channel estimation reduces the computational cost incurred by each user. A parameter reduction model is employed to counter the increased number of estimation parameters in the frequency domain. An EM based forward backward Kalman filter is applied to exploit the time correlation of the channel. The outercode's ability to purge the received data packet from error s is used to further improve the estimate. The simulation results indicate the viability of the proposed methods for the OFDMA scenario.

8. ACKNOWLEDGEMENT

This work was supported by King Abdul Aziz City for Science and Technology (KACST), Saudi Arabia, Project no. AR 27-98.

9. REFERENCES

[1] J. H. Lee, J. C. Han and S. C. Kim, "Joint carrier frequency synchronization and channel estimation for OFDM systems via the EM algorithm", *IEEE Transactions on Vehicular Technology*, 55 (1) pp167-172 2006.

[2] 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, no. 7, Jul. 2007.

[3] J. Liu, J. Li, "Parameter estimation and error reduction for OFDM-based WLANs ", *IEEE Transactions on Mobile Computing*, Volume 3, pp.152 - 163, Issue 2, April-June 2004.

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

[5] A. Jeremic, T. A. Thomas, A. Nehorai, "OFDM channel estimation in the presence of interference", *IEEE Transactions on Signal Processing*, Volume 52, Issue 12, pp. 3429 - 3439, Dec. 2004.

[6] C. H. Lim, D. S. Han, "Robust LS channel estimation with phase rotation for single frequency network in OFDM", *IEEE Transactions on Consumer Electronics*, Volume 52, Issue 4, pp. 1173 - 1178, Nov. 2006.

[7] M. Jiang, J. Akhtman, L. Hanzo, "Iterative Joint Channel Estimation and Multi-User Detection for Multiple-Antenna Aided OFDM Systems", *IEEE Transactions on Wireless Communications*, Vol. 6(8), pp. 2904 - 2914, August 2007.

[8] A. A. Tahat, "Multi-user channel estimation in a 4G OFDM system", *IEEE 18th International Symposium on Personal, Indoor and Mobile Radio Communications*, 3-7 Sep 2007, pp. 1-5.

[9] J. Bonnet, G. Auer,"Chunk-based Channel Estimation for Uplink OFDM", *IEEE Vehicular Technology Conference*, Melbourne, Vic. 2006., vol. 4, pp. 1555-1559.

[10] 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.