A Forward-Backward Kalman for the Estimation of Time-Variant Channels in OFDM

Tareq Y. Al-Naffouri
Electrical Engineering Dept.
King Fahd University of Petroleum and Minerals
Dhahran, Saudi Arabia

Arogyswami Paulraj
Electrical Engineering Dept.
Stanford University
USA

***Abstract*— OFDM 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. In this paper, we propose an expectation-maximization (EM) algorithm for joint channel and data recovery. The algorithm makes use of the rich structure of the underlying communication problem– a structure induced by the data and channel constraints. These constraints include pilots, the cyclic prefix, and the finite alphabet constraints on the data, and sparsity, finite delay spread, and the statistical properties of the channel (frequency and time correlation). The algorithm boils down to a forward-backward (FB) Kalman filter. We also suggest a suboptimal modification that is able to track the channel and recover the data with no latency.**

I. Introduction

Orthogonal frequency division multiplexing (OFDM) is an effective technique for high bit-rate transmission. It has found widespread applications and is already part of many standards. For proper operation of an OFDM system, the receiver needs an accurate estimate of the channel state information. For rapidly time-variant channels, the receiver faces the additional challenge of performing channel (and data) recovery for each OFDM symbol. In performing these two operations, the receiver takes advantage of the rich structure of the underlying communication problem. This structure can be either traced back to some inherent constraints on the data or on the channel. Table I lists the most common of these constraints and the works that employed them.

TABLE I
Data and channel constraints used for channel estimation

Constraints	Assumptions	Reference
Data Constraints	Finite alphabet constraint	[1]
	Code	[2]
	Transmit precoding	[2]–[5]
	Pilots	[6],[7]
Channel Constraints	Finite delay spread	[2],[3],[6]
	Sparsity	[8]
	Frequency correlation	[2],[9],[10]
	Time correlation	[11]–[13]

A. Approaches to channel estimation in OFDM

Several algorithms were suggested in literature for channel estimation in OFDM transmission. Each of these algorithms makes use of a subset of the constraints in Table I. These algorithms can be classified into one of the following categories

1) Training-based estimation: Pilots are used to perform channel estimation as in [6] and [7].

2) Blind estimation: At the other extreme, blind algorithms rely completely on natural constraints underlying the communication problem to perform recovery (as in [2]–[4]).

3) Semi-blind estimation: Semi-blind techniques are a hybrid of blind and training based techniques, utilizing pilots and other natural constraints to perform channel estimation (as in [2], [10]–[15]).

4) Data-aided channel estimation: The receiver uses the channel estimate to detect the data, which in turn can be used to enhance the channel estimate, giving rise to an iterative technique for channel and data recovery [11], [12]. Other works, like [13], [16], and [17], arrived at iterative techniques more rigorously by employing the expectation-maximization (EM) algorithm. The data-aided approach seems the most sensible for channel estimation as it is more general.

The affortmentioned works utilize *only a subset* of the constraints on the channel and data. In this paper, however, we present a (data-aided EM) method that can make use of all the constraints in Table I[1]. The method boils down to a forward-backward (FB) Kalman filter. One disadvantage of our approach is the storage and latency requirements of the FB-Kalman as it has to process multiple OFDM symbols simultaneously. We thus suggest a suboptimal forward-only version (basically a Kalman filter) that is able to perform channel recovery with no latency.

A remark about notation: We use bold face letters (e.g., $\boldsymbol{y}$) to denote vectors and caligraphic notation to denote variables in the frequency domain (e.g., $\boldsymbol{\mathcal{Y}}$ is the DFT of $\boldsymbol{y}$).

II. System model

Consider a sequence of $T+1$ data symbols $\boldsymbol{\mathcal{X}}_0^T$ to be transmitted. In an OFDM system, each symbol $\boldsymbol{\mathcal{X}}_i$ (length-N) undergoes an IDFT operation to produce the time domain symbol

$$\boldsymbol{x}_i = \sqrt{N}\boldsymbol{Q}^*\boldsymbol{\mathcal{X}}_i \tag{1}$$

where $\boldsymbol{Q}$ is the DFT matrix. The transmitter then appends a cyclic prefix (CP) $\underline{\boldsymbol{x}}_i$ (of length P) to $\boldsymbol{x}_i$, resulting finally in a sequence of super-sybmols $\overline{\boldsymbol{x}}_0^T$.

We assume that the channel $\underline{\boldsymbol{h}}_i$ (of maximum length $P+1$) remains fixed over any one OFDM symbol (and associated CP) and varies from one symbol to the next according to a state-space model

$$\underline{\boldsymbol{h}}_{i+1} = \boldsymbol{F}\underline{\boldsymbol{h}}_i + \boldsymbol{G}\boldsymbol{u}_i \quad \underline{\boldsymbol{h}}_0 \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Pi}_0) \tag{2}$$

(In the Appendix, we show how we can construct such a model from the Doppler frequency (time-correlation), the power-delay profile (frequency-correlation), and the receive filter). At the channel output, we obtain a sequence of time-domain super-symbols $\overline{\boldsymbol{y}}_0^T$, which after stripping the cyclic prefix $\underline{\boldsymbol{y}}_i$, produces a sequence of time-domain symbols $\boldsymbol{y}_0^T$. The input/output (I/O) relationship of the OFDM system is best described in the frequency domain

$$\begin{aligned}\boldsymbol{\mathcal{Y}}_i &= \operatorname{diag}(\boldsymbol{\mathcal{X}}_i)\boldsymbol{\mathcal{H}}_i + \boldsymbol{\mathcal{N}}_i & (3)\\ &= \operatorname{diag}(\boldsymbol{\mathcal{X}}_i)\boldsymbol{Q}_{P+1}\underline{\boldsymbol{h}}_i + \boldsymbol{\mathcal{N}}_i & (4)\end{aligned}$$

[1]Due to space limitation, we don't elaborate on how the algorithm makes use of the code and sparsity. However, the algorithm can incorporate these constraints in a straightforward manner [18].

The second line (4) follows from the DFT relationship

$$\mathcal{H}_i = Q \begin{bmatrix} \underline{h}_i \\ \mathbf{0} \end{bmatrix} = Q_{P+1}\underline{h}_i \tag{5}$$

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

$$X_i \triangleq \mathrm{diag}(\mathcal{X}_i) Q_{P+1} \tag{6}$$

we can write

$$\mathcal{Y}_i = X_i \underline{h}_i + \mathcal{N}_i \tag{7}$$

We can also construct a similar I/O relationship that incorporates (7) as well as the effect of the cyclic prefix observation

$$\overline{\mathcal{Y}}_i = \overline{X}_i \underline{h}_i + \overline{\mathcal{N}}_i \tag{8}$$

A. Pilot/output relationships

The receiver needs pilots to initialize channel estimation. Let the index set $I_{\mathrm{p}} = \{i_1, i_2, \ldots, i_{L_{\mathrm{p}}}\}$ denote the pilot locations within the OFDM symbol. Also, let the notation $X_{I_{\mathrm{p}}}$ denote the matrix X pruned of the rows that don't belong to I_{p}. Then, the pilot/output equation can be derived from the I/O relationship (7) as

$$\mathcal{Y}_{i\,I_{\mathrm{p}}} = X_{i\,I_{\mathrm{p}}} \underline{h}_i + \mathcal{N}_{i\,I_{\mathrm{p}}} \tag{9}$$

III. The EM Algorithm for Joint Channel and Data Estimation

A. The EM algorithm

Ideally, we estimate $\underline{h}_i$ using some I/O relationship, e.g. (7), by maximizing the corresponding log-likelihood function

$$\hat{\underline{h}}_i^{\mathsf{MAP}} = \max_{\underline{h}_i} \ \ln p(\mathcal{Y}_i | X_i, \underline{h}_i) + \ln p(\underline{h}_i)$$

In our case, however, the input $\mathcal{X}_i$ (or X_i) [2] is not observable. Thus, we use the expectation-maximization algorithm and maximize instead an averaged form of the log-likelihood function. Specifically, starting from an initial estimate $\hat{\underline{h}}_i^{(0)}$, the estimate $\hat{\underline{h}}_i$ is calculated iteratively, with the estimate at the jth iteration given by

$$\hat{\underline{h}}_i^{(j)} = \arg\max_{\underline{h}_i} E_{X_i|\mathcal{Y}_i, \hat{\underline{h}}_i^{(j-1)}} \ln p(\mathcal{Y}_i | X_i, \underline{h}_i) + \ln p(\underline{h}_i)$$

For example, when the channel obeys the I/O relationship (7) and $\underline{h}_i$ is $\mathcal{N}(\mathbf{0}, \mathbf{\Pi})$, the EM-based estimate (at the jth iteration) is given by

$$\hat{\underline{h}}_i^{(j)} = \arg\min_{\underline{h}_i} \|\mathcal{Y}_i - E[X_i]\underline{h}_i\|^2_{\sigma_n^{-2}} + \|\underline{h}_i\|^2_{\mathrm{Cov}[X_i^*]} + \|\underline{h}_i\|^2_{\Pi^{-1}}$$

where the two moments of X_i are taken given the output $\mathcal{Y}_i$ and the most recent channel estimate $\underline{h}_i^{(j-1)}$. We now derive the EM algorithm for the time-variant case.

B. The EM-based forward-backward Kalman

Consider the OFDM system of section 2, essentially described by the state-space model

$$\underline{h}_{i+1} = F\underline{h}_i + G u_i \tag{10}$$

$$\mathcal{Y}_i = X_i \underline{h}_i + \mathcal{N}_i \tag{11}$$

with $\underline{h}_0 \sim \mathcal{N}(\mathbf{0}, \mathbf{\Pi}_0)$ and $u_i \sim \mathcal{N}(\mathbf{0}, \sigma_u^2 I)$. Given a sequence of $T+1$ input and output symbols X_0^T and $\mathcal{Y}_0^T$, we obtain the MAP estimate of the channel sequence $\underline{h}_0^T$ by the maximizing the log-likelihood

$$\mathcal{L} = \ln p(\mathcal{Y}_0^T | X_0^T, \underline{h}_0^T) + \ln p(\underline{h}_0^T)$$

Now, using (11), we can express the first term of the log-likelihood (up to some additive constants) as

$$\begin{aligned} \ln p(\mathcal{Y}_0^T | \mathcal{X}_0^T, \underline{h}_0^T) &= \sum_{i=0}^{T} \ln p(\mathcal{Y}_i | \mathcal{X}_i, \underline{h}_i) \\ &= -\sum_{i=0}^{T} \|\mathcal{Y}_i - X_i \underline{h}_i\|^2_{\frac{1}{\sigma_n^2}} \end{aligned}$$

Similarly, using (10), we can express the second term (again up to some additive constants) as

$$\ln p(\underline{h}_0^T) = \sum_{i=0}^{T} \ln p(\underline{h}_i | \underline{h}_{i-1}) + \ln p(\underline{h}_0) \tag{12}$$

$$= -\sum_{k=1}^{T} \|\underline{h}_k - F\underline{h}_{k-1}\|^2_{\frac{1}{\sigma_u^2} GG^*} - \|\underline{h}_0\|^2_{\Pi_0^{-1}} \tag{13}$$

Combining these two expressions yields

$$\mathcal{L} = -\sum_{i=0}^{T} \|\mathcal{Y}_i - X_i \underline{h}_i\|^2_{\frac{1}{\sigma_n^2}} - \sum_{i=1}^{T} \|\underline{h}_i - F\underline{h}_{i-1}\|^2_{\frac{1}{\sigma_u^2} GG^*} - \|\underline{h}_0\|^2_{\Pi_0^{-1}} \tag{14}$$

Since the channel sequence $\underline{h}_0^T$ is jointly Gaussian, the MAP estimate of the channel sequence given the input and output sequences X_0^T and $\mathcal{Y}_0^T$ is the same as the MMSE estimate given the same sequences. The MMSE estimate itself is obtained by the FB Kalman filter. This allows us to state the following theorem

Theorem 1: **Channel estimation–Known input case** Consider the state-space model (10)–(11). Given the input and output sequences X_0^T and $\mathcal{Y}_0^T$, the MAP (or equivalently MMSE) estimate of $\underline{h}_0^T$ is obtained by applying the following (forward-backward Kalman) filter to the state-space model (10)–(11)

Forward run: For $i = 1, \ldots, T$, calculate

$$R_{e,i} = \sigma_n^2 I_{N+P} + X_i P_{i|i-1} X_i^* \quad P_{0|-1} = \mathbf{\Pi}_0 \tag{15}$$

$$K_{f,i} = P_{i|i-1} X_i^* R_{e,i}^{-1} \tag{16}$$

$$\hat{\underline{h}}_{i|i} = (I_{N+P} - K_{f,i} X_i)\hat{\underline{h}}_{i|i-1} + K_{f,i}\mathcal{Y}_i, \tag{17}$$

$$\hat{\underline{h}}_{i+1|i} = F\hat{\underline{h}}_{i|i}, \quad \underline{h}_{0|-1} = \mathbf{0} \tag{18}$$

$$P_{i+1|i} = F_i\left(P_{i|i-1} - K_{f,i} R_{e,i} K_{f,i}^*\right) F^* + \frac{1}{\sigma_n^2} GG^* \tag{19}$$

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

$$\boldsymbol{\lambda}_{i|T} = \left(I_{P+N} - X_i^* K_{f,i}^*\right) F_i^* \boldsymbol{\lambda}_{i+1|T} + X_i R_{e,i}^{-1}\left(\mathcal{Y}_i - X_i \hat{\underline{h}}_{i|i-1}\right) \tag{20}$$

$$\hat{\underline{h}}_{i|T} = \hat{\underline{h}}_{i|i-1} + P_{i|i-1}\boldsymbol{\lambda}_{i|T} \tag{21}$$

The desired estimate is $\hat{\underline{h}}_{i|T}$. For a proof, see problem 10.9 in [19]. This theorem allows us to obtain the estimate of $\underline{h}_0^T$ when the input sequence X_0^T is not available. For in this case, we maximize the log-likelihood (14) *averaged* over the sequence X_0^T. Thus, the j-th iteration of the EM algorithm is now obtained by maximizing the averaged log-likelihood

$$\overline{\mathcal{L}} = E_{X_0^T | \underline{h}_0^T, \mathcal{Y}_0^T} \mathcal{L} \tag{22}$$

[2] Since $X_i = \mathrm{diag}(\mathcal{X}_i) Q_{P+1}$, conditioning on $\mathcal{X}_i$ can be replaced by conditioning on X_i.

By inspecting (14), we note that the only term that is modified under expectation is the first summand, and its expectation is given by

$$\begin{aligned} E\left\|\boldsymbol{\mathcal{Y}}_i-\boldsymbol{X}_i\underline{\boldsymbol{h}}_i\right\|^2_{\frac{1}{2\sigma_n^2}} &= \left\|\boldsymbol{\mathcal{Y}}_i-E\left[\boldsymbol{X}_i\right]\underline{\boldsymbol{h}}_i\right\|^2_{\frac{1}{2\sigma_n^2}}+\left\|\underline{\boldsymbol{h}}_i\right\|^2_{\frac{1}{2\sigma_n^2}\mathrm{Cov}[X_i^*]} \\ &= \left\|\begin{bmatrix}\boldsymbol{\mathcal{Y}}_i\\ \mathbf{0}_{P\times 1}\end{bmatrix}-\begin{bmatrix}E[\boldsymbol{X}_i]\\ \mathrm{Cov}[\boldsymbol{X}_i^*]^{1/2}\end{bmatrix}\underline{\boldsymbol{h}}_i\right\|^2_{\frac{1}{2\sigma_n^2}} \end{aligned}$$

where the expectations are taken given the previous estimate $\underline{\hat{\boldsymbol{h}}}_0^{(j-1)}$ and the output symbols $\boldsymbol{\mathcal{Y}}_0^T$. We thus have

$$\overline{\mathcal{L}}=-\sum_{i=0}^{T}\left\|\begin{bmatrix}\boldsymbol{\mathcal{Y}}_i\\ \mathbf{0}_{P\times 1}\end{bmatrix}-\begin{bmatrix}E[\boldsymbol{X}_i]\\ \mathrm{Cov}[\boldsymbol{X}_i^*]^{1/2}\end{bmatrix}\underline{\boldsymbol{h}}_i\right\|^2_{\frac{1}{2\sigma_n^2}}- \sum_{i=1}^{T}\left\|\underline{\boldsymbol{h}}_i-\boldsymbol{F}\underline{\boldsymbol{h}}_{i-1}\right\|^2_{\frac{1}{\sigma_u^2}GG^*}-\left\|\underline{\boldsymbol{h}}_0\right\|^2_{\Pi_0^{-1}} \qquad (23)$$

Note that we can obtain the averaged likelihood (23) from the original likelihood (14) by performing the substitution

$$\boldsymbol{X}_i\longrightarrow\begin{bmatrix}E[\boldsymbol{X}_i]\\ \mathrm{Cov}[\boldsymbol{X}_i^*]^{1/2}\end{bmatrix}\quad \boldsymbol{\mathcal{Y}}_i\longrightarrow\begin{bmatrix}\boldsymbol{\mathcal{Y}}_i\\ \mathbf{0}_{P\times 1}\end{bmatrix}$$

We can thus state the following theorem

Theorem 2: **Channel estimation–Unknown input case** Consider the state-space model (10)–(11) and assume that the receiver does not have access to the transmitted data $\boldsymbol{X}_0^T$. The channel estimate at the jth iteration $\underline{\boldsymbol{h}}_0^{T(j)}$ of the EM algorithm is obtained by applying the forward-backward Kalman (15)–(21) to the following state-space model

$$\begin{aligned} \underline{\boldsymbol{h}}_{i+1} &= \boldsymbol{F}\underline{\boldsymbol{h}}_i+\boldsymbol{G}\boldsymbol{u}_i && (24)\\ \begin{bmatrix}\boldsymbol{\mathcal{Y}}_i\\ \mathbf{0}_{P\times 1}\end{bmatrix} &= \begin{bmatrix}E[\boldsymbol{X}_i]\\ \mathrm{Cov}[\boldsymbol{X}_i^*]^{1/2}\end{bmatrix}\underline{\boldsymbol{h}}_i+\begin{bmatrix}\boldsymbol{\mathcal{N}}_i\\ \underline{\boldsymbol{n}}_i\end{bmatrix} && (25) \end{aligned}$$

where $\boldsymbol{n}_i$ is virtual noise that is independent of the physical noise $\boldsymbol{\mathcal{N}}_i$.

To fully implement the EM algorithm, we need to initialize the algorithm and calculate the first and second moments of the input, which we do next.

C. Calculating the input moments

Using the relationship $\boldsymbol{\mathcal{X}}_i=\mathrm{diag}(\boldsymbol{\mathcal{X}}_i)\boldsymbol{Q}_{P+1}$, we can write

$$\begin{aligned} E[\boldsymbol{X}_i] &= \mathrm{diag}(E[\boldsymbol{\mathcal{X}}_i])\boldsymbol{Q}_{P+1} && (26)\\ \mathrm{Cov}[\boldsymbol{X}_i^*] &= \boldsymbol{Q}_{P+1}^*\mathrm{Cov}[\boldsymbol{\mathcal{X}}_i^*]\boldsymbol{Q}_{P+1} && (27) \end{aligned}$$

Now we can calculated the mean and covariance of $\boldsymbol{\mathcal{X}}_i$ by calculating the first two moments of the individual elements $\mathcal{X}_i(l)\ \ l=1,\cdots N$ since these elements are independent. Assuming that $\mathcal{X}_i(l)$ takes its values from the alphabet $A=\{A_1,\ldots,A_{|A|}\}$ with equal probability, we can show that

$$E[\mathcal{X}_i(l)|\mathcal{Y}_i(l)]=\frac{\sum_{j=1}^{j=|A|}A_je^{-\frac{|\mathcal{Y}_i(l)-\mathcal{H}(l)A_j|^2}{\sigma_n^2}}}{\sum_{j=1}^{|A|}e^{-\frac{|\mathcal{Y}_i(l)-\mathcal{H}(l)A_j|^2}{\sigma_n^2}}} \qquad (28)$$

$$E[|\mathcal{X}_i(l)|^2|\mathcal{Y}_i(l)]=\frac{\sum_{j=1}^{j=|A|}|A_j|^2e^{-\frac{|\mathcal{Y}_i(l)-\mathcal{H}(l)A_j|^2}{\sigma_n^2}}}{\sum_{j=1}^{|A|}e^{-\frac{|\mathcal{Y}_i(l)-\mathcal{H}(l)A_j|^2}{\sigma_n^2}}} \qquad (29)$$

D. Initial channel estimation

We can obtain the initial channel estimate from the pilot/output equation (9). We do this by applying the FB Kalman to the state-space model

$$\begin{aligned} \underline{\boldsymbol{h}}_{i+1} &= \boldsymbol{F}\underline{\boldsymbol{h}}_i+\boldsymbol{G}\boldsymbol{u}_i && (30)\\ \boldsymbol{\mathcal{Y}}_{i\,I_{\mathrm{p}}} &= \boldsymbol{X}_{i\,I_{\mathrm{p}}}\underline{\boldsymbol{h}}_i+\boldsymbol{\mathcal{N}}_{i\,I_{\mathrm{p}}} && (31) \end{aligned}$$

i.e., by applying the FB Kalman (15)–(21) with the substitutions

$$\boldsymbol{X}_i\longrightarrow\boldsymbol{X}_{i\,I_{\mathrm{p}}},\quad \boldsymbol{\mathcal{Y}}_i\longrightarrow\boldsymbol{\mathcal{Y}}_{i\,I_{\mathrm{p}}}$$

E. Summary of the EM-based FB Kalman

In the following, we summarize the FB Kalman for channel and data recovery

1) Obtain the initial channel estimate $\underline{\boldsymbol{h}}_0^{T(0)}$ by applying the FB Kalman filter (15)–(21) to the state-space model (30)–(31)
2) Iterate between the expectation and maximization steps for $j=1,\ldots,N_{\mathrm{iter}}$:
   a) **Expectation:** Compute the first two moments of the input $\boldsymbol{X}_0^T$ given the output $\boldsymbol{\mathcal{Y}}_0^T$ and the previous estimate of the channel, $\underline{\boldsymbol{h}}_0^{T(j-1)}$, using (6), and (28)– (29).
   b) **Maximization:** Obtain the channel estimate $\underline{\boldsymbol{h}}_0^{T(j)}$ by employing the FB Kalman (15)–(21) to the state-space model (24)–(25).

The algorithm can be stopped when the difference between two consecutive estimates $\|\underline{\boldsymbol{h}}_0^{T(j)}-\underline{\boldsymbol{h}}_0^{T(j-1)}\|^2$ is below a certain threshold or when the maximum number of iterations N_{iter} is reached

IV. Two Extensions

A. Kalman- (forward-only) based estimation

One disadvantage of the FB Kalman is the storage and latency involved. The algorithm needs to wait for all $T+1$ symbols before it can execute the backward run and hence obtain the channel estimate. One way around this is to reduce the window size T. Alternatively, we can run the filter in the forward direction only (i.e., run (15)–(19)) for both the initial estimation and the EM iteration. The algorithm then collapses to the Kalman-based filter proposed in [20] where the data and channel are recovered within one OFDM symbol.

B. Using the cyclic prefix observation

The FB Kalman can also make use of the CP observation. Here pilot-based estimation remains the same while the EM algorithm is run on the I/O equation (8) which contains the effect of the cyclic prefix. Thus, in this case, we apply the FB Kalman (15)–(21) to the state-space model

$$\begin{aligned} \underline{\boldsymbol{h}}_{i+1} &= \boldsymbol{F}\underline{\boldsymbol{h}}_i+\boldsymbol{G}\boldsymbol{u}_i && (32)\\ \begin{bmatrix}\overline{\boldsymbol{\mathcal{Y}}}_i\\ \mathbf{0}_{P\times 1}\end{bmatrix} &= \begin{bmatrix}E[\overline{\boldsymbol{X}}_i]\\ \mathrm{Cov}[\overline{\boldsymbol{X}}_i^*]^{1/2}\end{bmatrix}\underline{\boldsymbol{h}}_i+\begin{bmatrix}\overline{\boldsymbol{\mathcal{N}}}_i\\ \underline{\boldsymbol{n}}_i\end{bmatrix} && (33) \end{aligned}$$

The two moments of $\overline{\boldsymbol{X}}_i$ can be obtained from (28)–(29) but the calculations become more cumbersome due to the presence of the CP (see [18]).

V. Simulations

We consider an OFDM system that transmits a sequence of 5 symbols each with 64 carriers and a cyclic prefix of length $P=15$. The input data is 16 QAM mapped from a binary bit stream through gray coding. We will use two pilot configurations. The first employs 16 pilots in the first symbol and x number of them in the subsequent four symbols with $x=4,8,12,16$. We denote this configuration by $16xxxx$. The second configuration, denoted $xx16xx$, is a cyclic rearrangement of the first with the 16 pilots in the middle (3rd) symbol and x pilots in the other symbols.

The channel IR consists of 16 complex taps (the maximum length possible). The initial IR $\underline{\boldsymbol{h}}_0$ has an exponential delay profile $E[|\underline{h}_0(k)|^2] = e^{-0.2k}$. For $i > 0$, $\underline{\boldsymbol{h}}_i$ is generated according to the dynamical model $\underline{\boldsymbol{h}}_{i+1} = \boldsymbol{F}\underline{\boldsymbol{h}}_i + \boldsymbol{G}\boldsymbol{u}_i$. Both $\boldsymbol{F}$ and $\boldsymbol{G}$ are diagonal matrices. Specifically, we set $\boldsymbol{F} = f\boldsymbol{I}$ with $0 < f < 1$ and set the diagonal entries of $\boldsymbol{G}$ as $G(k,k) = \sqrt{(1-f^2)E[|\underline{h}_0(k)|^2]}$. Throughout the simulations, we run the EM algorithm for 10 iterations.

A. Comparing the Kalman and the forward-backward Kalman

In Figure 1 we compare the performance of the Kalman receiver employing the $16xxxx$ pilot configuration with that of the FB-Kalman receiver employing the $xx16xx$ configuration [3]. We carry out this comparison for different levels of time variations ($f = .1, .3, .5, .7, .9$.) We note that the BER curves are quite comparable for the extreme cases of time variation (low and high values of f). However, for moderate levels of variation ($f = .7$), the FB-Kalman consistently outperforms the the forward Kalman. This is not unexpected for when the variation is too slow, the two filters are equally able to track the channel with only a few pilots. When the time variation is too high, time correlation information becomes of little use. It is only at a moderate level of time-variations that the additional signal processing of the FB-Kalman becomes valuable.

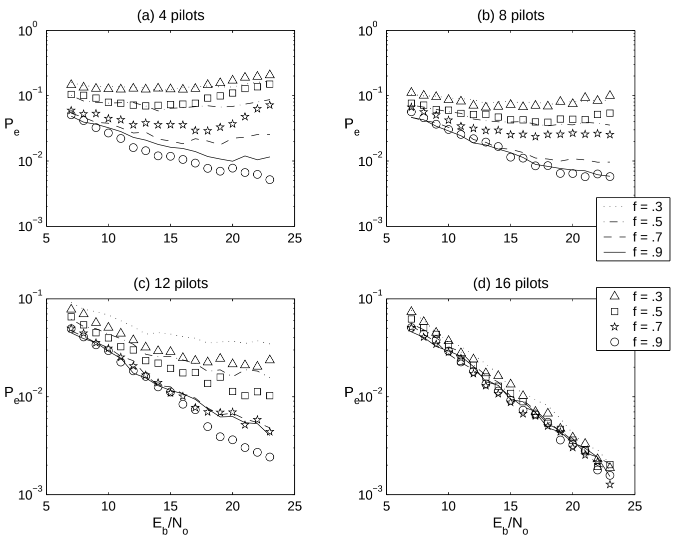

Fig. 1. Comparing the Kalman and FB-Kalman for various levels of time-variation (Solid and dotted lines are for the Kalman receiver).

B. Effect of increased signal processing

We next consider the effect of increased signal processing on the BER curves for the FB-Kalman receiver [4]. Specifically, we implement this receiver using 1) the CP observation and the soft estimate of the input, 2) the CP and the hard estimate of the input, and 3) no CP observation and the hard estimate of the input. Figure 2 shows that increasing the level of signal processing pays off producing better BER performance. This applies for different number of pilots and different degrees of time variation.

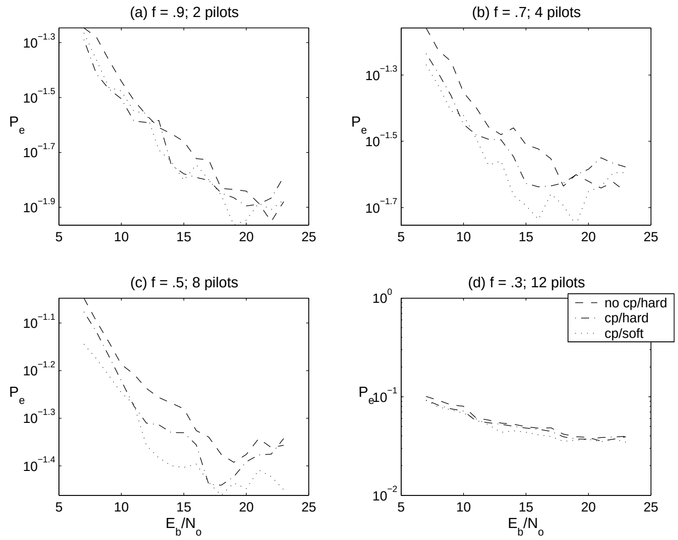

Fig. 2. The FB-Kalman based receiver demonstrates improved BER with increasing levels of signal processing

C. Bench marking

Finally, we bench mark the BER performance of the Kalman and FB-Kalman receivers against receivers that have been suggested in literature and also against the known-channel case. Specifically, Figure 3 compares the BER performance of the following five receivers: 1) EM-based least-squares (LS) receiver (i.e. a receiver employing frequency correlation only), 2) the EM-based receiver proposed by Lu, Wang, and Li in [13] [5], 3) the EM-based Kalman receiver 4) EM-based FB-Kalman receiver, and 5) a receiver with perfect channel knowledge. All receivers implement the $16xxxx$ pilot configuration except the FB Kalman which implements the $xx16xx$ configuration and the receiver with perfect channel knowledge which uses no pilots. We test these receivers against the $f = .7$ case.

Figure 3 demonstrates that the Kalman and FB-Kalman outperform the LS receiver and the receiver of [13], especially for low number of pilots. Moreover, for this case of moderate time variation, the FB-Kalman consistently outperforms the Kalman receiver.

VI. Conclusion

In this paper, we considered the problem of semi-blind channel and data recovery in OFDM transmission over time-variant channels. Motivated by the EM approach, the algorithm boils down to a FB Kalman filter. It makes a collective use of the channel and data constraints in Table I. Specifically, the algorithm makes use of the finite alphabet constraints (in (28–(29)), the data in its soft form (in (24)–(25)), pilots (in (30)–(31)), transmission precoding (in (32)–(33)), finite-delay spread (in that channel estimation is done in the time domain), and frequency- and time-correlation (in (2)). It is also straightforward to incorporate the effect of an outer code and sparsity (see [18]). We also suggested a relaxed version of the algorithm that is able to perform recovery with no latency and hence avoid the

[3]Both simulation and intuition suggest that the Kalman performs betters with the $16xxxx$ configuration while FB Kalman does better with the $xx16xx$ configuration.

[4]We omit the corresponding simulations for the Kalman filter due to the lack of space.

[5]This receiver is similar to our Kalman-based (forward-only) receiver in that it makes use of the time and frequency correlation.

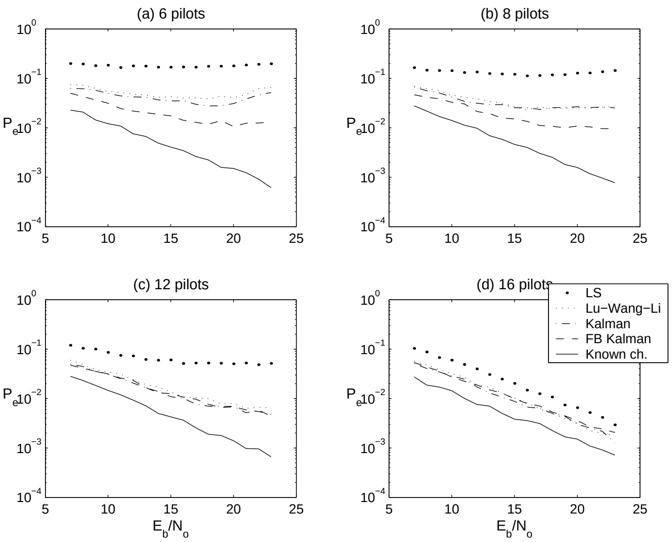

Fig. 3. Comparing the BER curves for various receivers.

delay and storage shortcomings of the FB-Kalman. Our simulation show the favorable behavior of the two Kalman filters. Specifically, simulations show that increased signal processing always results in better BER behavior.

VII. Appendix: Channel model

The channel $\underline{\boldsymbol{h}}_i$ is the convolution of the physical channel $\underline{\boldsymbol{c}}_i$ (which consists of $L+1$ paths arriving at instants $\tau_0, \tau_1 \ldots, \tau_L$) and the receive filter $\boldsymbol{r}$. Thus, we can write

$$\underline{\boldsymbol{h}}_i = \boldsymbol{R}_i \underline{\boldsymbol{c}}_i \tag{34}$$

where $\boldsymbol{R}_i$ is the receive filter matrix given by

$$\boldsymbol{R}_i = \begin{bmatrix} r(-\tau_0) & r(-\tau_1) & \cdots & r(-\tau_L) \\ r(T-\tau_0) & r(T-\tau_1) & \cdots & r(T-\tau_L) \\ \vdots & \vdots & \vdots & \vdots \\ r(PT-\tau_0) & r(PT-\tau_2) & \cdots & r(PT-\tau_L) \end{bmatrix}$$

Due to the mobile nature of the channel, the physical channel taps $c_i(k)$ are time-variant. According to the WSSUS model, the process $c_i(k)$ is zero-mean wide-sense stationary complex Gaussian process with autocorrelation

$$E\left[c_i(k) c_{i'}(k')\right] = \mathcal{J}_0\left(\alpha_k |i-i'|\right) \delta_{kk'} \quad \alpha_k = 2\pi f_c(k)(N+P)T$$

where T is the sampling (baud) rate, $f_c(k)$ is the Doppler frequency associated with the kth tap, and $\mathcal{J}_0$ denotes the zero-order Bessel function of the first kind. We can approximate the time-variant behavior of the tap $c_i(k)$ by a first-order AR model (see [11], [12])

$$\underline{c}_{i+1}(k) = \mathcal{J}_0\left(\alpha_k\right) \underline{c}_i(k) + \sqrt{(1-\mathcal{J}_0^2(\alpha_k))E[|\underline{c}_0(k)|^2]}\, u_i(k) \tag{35}$$

The factor $\sqrt{(1-\mathcal{J}_0^2(\alpha_k))E[|\underline{c}_0(k)|^2]}$ ensures that the tap $c_i(k)$ maintains the same power profile for all time. Collecting (35) for all taps yields

$$\boldsymbol{c}_{i+1} = \boldsymbol{F}_c \boldsymbol{c}_i + \boldsymbol{G}_c \boldsymbol{u}_i \tag{36}$$

where $\boldsymbol{F}_c = \text{diag}\left(\ \mathcal{J}_0(\alpha_1), \ \cdots, \ \mathcal{J}_0(\alpha_{L+1})\ \right)$ and $\boldsymbol{G}_c = \text{diag}\left(\ \sqrt{(1-\mathcal{J}_0^2(\alpha_1))E[|\underline{c}_i(1)|^2]}, \ldots, \sqrt{(1-\mathcal{J}_0^2(\alpha_{L+1}))E[|\underline{c}_i(L+1)|^2]}\ \right)$. We can use this dynamical relationship along with (34) to derive a dynamical relationship for the impulse response $\underline{\boldsymbol{h}}$. Specifically, multiplying both sides of (36) by $\boldsymbol{R}$ and noting that $\boldsymbol{R}^\dagger \boldsymbol{R} = \boldsymbol{I}$ [6], we obtain $\underline{\boldsymbol{h}}_{i+1} = \boldsymbol{F}\underline{\boldsymbol{h}}_i + \boldsymbol{G}\boldsymbol{u}_i$ where $\boldsymbol{F} = \boldsymbol{R}\boldsymbol{F}_c\boldsymbol{R}^\dagger$ and $\boldsymbol{G} = \boldsymbol{R}\boldsymbol{G}_c$.

Acknowledgement

The authors would like to acknowledge King Fahd University of Petroleum and Minerals, Saudi Arabia, supporting this work.

References

[1] 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.

[2] G. Alrawi, T. Y. Al-Naffouri, A. Bahai, and J. Cioffi, "Exploiting error-control coding and cyclic prefix in channel estimation for coded ofdm systems," *IEEE Commun. Lett.*, vol. 7, no. 7, pp. 388–390, Jul. 2003.

[3] H. Blcskei, R. W. Heath, and A. J. Paulraj, "Blind channel identification and equalization in ofdm-based multi-antenna systems," *IEEE Trans. Signal Proc.*, vol. 50, no. 1, pp. 96–109, Jan. 2002.

[4] R.W. Heath and G. B. Giannakis, "Exploiting input cyclostationarity for blind channel identification in ofdm systems," *IEEE Trans. Signal Proc.*, vol. 47, no. 3, pp. 848–856, Mar. 1999.

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

[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," in *Proc. IEEE Vehicular Tech. Conf.*, 1999, vol. 2, pp. 1131–1135.

[8] 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.

[9] 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.

[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.

[11] 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.

[12] 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.

[13] B. Lu, X. Wang, and Y. Li, "Iterative receivers for space-time block-coded ofdm systems in dispersive fading channels," vol. 1, no. 2, pp. 213–225, Apr. 2002.

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

[15] Xiao ming Chen and P. A. Hoeher, "Blind equalization with iterative joint channel and data estimation for wireless dpsk systems," in *Proc. IEEE GlobeCom*, 2001, pp. 274–279.

[16] T. Y. Al-Naffouri, A. Bahai, and A. Paulraj, "Semi-blind channel identification and equalization in ofdm: an expectation-maximization approach," in *Proc. IEEE Vehicular Tech. Conf.*, 2002, vol. 2, pp. 13–17.

[17] C. Cozzo and B. L. Hughes, "Joint channel estimation and data detection in space-time communications," *IEEE Trans. Commun.*, vol. 51, no. 8, pp. 1266–1270, Aug. 2003.

[18] T. Y. Al-Naffouri, *Adaptive algorithms for wireless channel estimation: Transient analysis and semi-blind design*, Ph.D Thesis, Stanford University, Jan. 2005.

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

[20] T. Y. Al-Naffouri, A. Bahai, and A. Paulraj, "Em algorithm for channel estimation and equalization in ofdm," in *Proc. IEEE GlobeCom*, 2002, pp. 589–593.

[6]For this to be true, the matrix $\boldsymbol{R}$ has to be tall which will be the case if the sampling rate is high enough so that the number of channel taps $P+1$ is larger than the number of physical paths.