Chapter 1

Iterative Forward-Backward Kalman Filtering for Data Recovery in (Multiuser) OFDM Communications[1]

Tareq Y. Al-Naffouri, Muhammad S. Sohail and Ahmed A. Quadeer

King Fahd University of Petroleum & Minerals

Saudi Arabia

1. Introduction

The current era of connectivity and information transfer demands high data rates from broadband wireless systems. The main problem faced by such high data rate systems is multipath fading as it causes inter symbol interference (ISI). Thus, there is a need for a technique that avoids ISI while still providing high speed data. Orthogonal Frequency Division Multiplexing (OFDM) emerged as a technique that combines the advantages of high achievable data rates, relatively easy implementation, high spectral efficiency and robustness to multipath fading. This is reflected by many standards that considered and adopted OFDM including digital audio and video broadcasting (DAB and DVB), WIMAX (Worldwide Inter-operability for Microwave Access), high speed modems over digital subscriber lines, and local area wireless broadband standards such as the HIPERLAN/2 and IEEE 802.11a (Stuber et al., 2004).

In a wireless communication system, data is sent over a channel. This channel is time variant and undergoes fading. The exact state of the channel is unknown at the receiver. The transmitted signal is received at the receiver convolved with the channel and corrupted with noise. The main interest at the receiver is to recover the transmitted data. If we consider that the channel state information is known at the receiver, we can faithfully extract the transmitted data from the received signal with this knowledge (through equalization). In practice however, we do not have prior knowledge of the channel state information and hence we have to employ some estimation technique to obtain an estimate of the channel. Channel estimation is thus an important step in receiver design.

We will start this chapter by reviewing the various approaches to channel estimation available in literature. We will then present an iterative channel estimation technique based on the Forward-Backward Kalman filter for simple OFDM systems and later extend the concept for multi-access OFDM and Multi Input Multi Output (MIMO) OFDM systems.

[1]This work was partly supported by King Abdul Aziz City for Science & Technology, Project No. AR-27-98 and partly by a Junior Faculty Project JF060003, King Fahd University of Petroleum and Minerals, Saudi Arabia.

2. Various Approaches to Channel Estimation

An OFDM receiver needs to be able to accurately estimate the transmitted data and for that, it needs accurate channel state information. For time variant channel, an additional problem is that channel state keeps changing. One way to deal with such channels is to periodically send a training (known) sequence which can be used at the receiver to estimate the channel. However this would severely strain the bandwidth efficiency of the system. An alternative approach is to use a priori constraints on the communication system to reduce the training overhead. The most popular of these constraints are summarized in Table 1.1 together with the works that employed them. All channel estimation techniques use all or a subset of these constraints. For example, to reduce the training overhead, we could assume that the channel is sparse or that it exhibits strong time correlation from one instant into another. Similarly, we could use some a priori information about the transmitted data such as the fact that it is drawn from some finite alphabet (see Table 1.1).

TYPE	CONSTRAINTS	REFERENCE
Data Constraints	Finite alphabet constraint	(Al-Naffouri et al., 2001); (Shengli and Giannakis, 2001); (Al-Naffouri, 2007); (Yang and Ser, 2004)
	Code	(Al-Rawi et al., 2003); (Zhang and Chen, 2004); (Petropulu et al., 2004); (Gao and Nallanathan, 2007)
	Transmit precoding (e.g., cyclic prefix, zero-padding, virtual carriers)	(Al-Naffouri, 2007); (Bölcskei et al., 2002); (Al-Rawi et al., 2003); (Kim and Eo, 2006); (Muquet et al., 2000); (Shin et al., 2007); (Kunji et al., 2006)
	Pilots	(Edfors et al., 1998); (Negi and Cioffi, 1998); (Biguesh and Gershman, 2004); (Li et al., 1998); (Minn and Al-Dhahir, 2006)
Channel Constraints	Finite delay spread	(Bölcskei et al., 2002); (Negi and Cioffi, 1998)
	Sparsity	(Yang et al., 2001); (Kang et al., 1999)
	Frequency correlation	(Al-Naffouri, 2007); (Edfors et al., 1998); (Al-Rawi et al., 2003); (Chang and Su, 2004); (Cui and Tellambura, 2006)
	Time correlation	(Muquet et al., 2000); (Al-Naffouri et al., 2004); (Necker and Stuber, 2004); (Zhang and Chen, 2004); (Al-Naffouri, 2007)
	Uncertainty information	(Sayed, 2001); (Li et al., 1998)

Table 1.1: Constraints used for channel estimation

2.1. Iterative Receivers for Channel Estimation & Data Recovery

There are several methods for channel estimation. However, the state of the art receiver is iterative in nature in that it iterates between finding a channel estimate which is used for data detection and between finding a data estimate which is in turn used to enhance channel estimation. Training data kick starts the iterative process by providing an initial channel estimate. Moreover, the a priori information that we have about the channel and data enhance the channel estimation and data detection steps which in turn reduces the required transmission overhead (Al-Naffouri et al., 2002); (Aldana et al., 2003); (Cozzo and Hughes, 2003).

In this chapter, we use the Expectation-Maximization (EM) algorithm to design an iterative receiver for the estimation and equalization of time variant channels. We will show that the receiver boils down to a forward-backward Kalman filter. We will discuss the use of Kalman filter for channel and data recovery in single user as well as multiuser OFDM systems. To setup the stage, we introduce our notation in the following section followed by the system model.

3. Notation

In this chapter, scalars are denoted by small-case letters (e.g., y), vectors by small-case boldface letters (e.g., $\boldsymbol{y}$), and matrices by uppercase boldface letters (e.g., $\boldsymbol{X}$). For vectors in frequency domain, calligraphic notation (e.g. $\boldsymbol{\mathcal{Y}}$) is used. An estimate of a variable is indicated by a hat on that variable (e.g., $\hat{\boldsymbol{h}}$ is an estimate of $\boldsymbol{h}$). Also, conjugate transpose is represented by $*$, Kronecker product by $\otimes$, size $N \times N$ identity matrix by $\boldsymbol{I}_N$, and the all zero $M \times N$ matrix by $\mathbf{0}_{M\times N}$. For a vector $\boldsymbol{a}$, $\text{diag}(\boldsymbol{a})$ represents a diagonal matrix with the elements of $\boldsymbol{a}$ on its diagonal. Finally, we use $\boldsymbol{h}_0^T$ to denote the sequence $\boldsymbol{h}_0, \boldsymbol{h}_1, ..., \boldsymbol{h}_T$.

4. System Model

Consider an OFDM system where a sequence of $T+1$ symbols $\boldsymbol{\mathcal{X}}_0, \boldsymbol{\mathcal{X}}_1, \cdots, \boldsymbol{\mathcal{X}}_T$, each of length N, are to be transmitted. The data bits to be transmitted are first passed through a convolutional encoder, punctured and interleaved. The resulting bit sequence is mapped to QAM symbols using Gray code. The OFDM symbol is then constructed by inserting these QAM symbols at data positions and pilot symbols at training positions. Here we consider comb type pilots as they are more robust in fading channels than block type pilots (Nee and Prasad, 2000). Each symbol $\boldsymbol{\mathcal{X}}_i$ undergoes an IDFT operation to produce the time domain symbol $\boldsymbol{x}_i = \sqrt{N}\boldsymbol{Q}^*\boldsymbol{\mathcal{X}}_i$ where $\boldsymbol{Q}$ is the $N \times N$ DFT matrix. A length P cyclic prefix $\underline{\boldsymbol{x}}_i$ is appended to the symbol $\boldsymbol{x}_i$ to counter the effect of inter symbol interference. This transforms the complex equalization problem into parallel single tap equalizers. The transmitter then transmits the resulting super symbol $\overline{\boldsymbol{x}}_i$ of length $N+P$ as shown in figure (1.1)

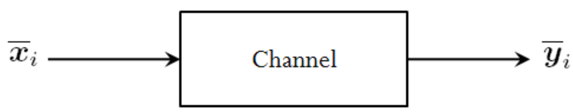

Figure 1.1: Simple OFDM System

We consider a block fading channel model, meaning the channel $\boldsymbol{h}_i$ (length $< P+1$) remains unchanged for each super symbol but varies from one super symbol to the next according to a state space model

$$\boxed{\boldsymbol{h}_{i+1} = \boldsymbol{F}\boldsymbol{h}_i + \boldsymbol{G}\boldsymbol{u}_i \qquad \boldsymbol{h}_0 \sim \mathcal{N}(\mathbf{0}, \boldsymbol{R}_n)} \tag{1}$$

where $\boldsymbol{h}_o \sim \boldsymbol{\mathcal{N}}(0, \Pi_0)$, and $\mathbf{u}_o \sim \boldsymbol{\mathcal{N}}(0, \sigma_u^2)$. The matrices F and G are square matrices of size $P \times P$ and are a function of Doppler spread (time correlation), power delay profile (frequency correlation) and the transmit filter. These matrices are assumed to be known at the receiver and are given as

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

where $\alpha(p)$ 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. The constraints captured by the state space model (1) include frequency correlation, time correlation, finite delay spread, and sparsity.

The passage of $\bar{\boldsymbol{x}}_i$ symbols through the channel $\boldsymbol{h}$ produces the received sequence $\bar{\boldsymbol{y}}_i$ at the receiver. The received packet of length $N+P$ is split into a length N packet $\boldsymbol{y}_i$ and a length P prefix $\underline{\boldsymbol{y}}_i$. The prefix absorbs all the inter symbol interference between $\bar{\boldsymbol{x}}_{(i-1)}$ and $\bar{\boldsymbol{x}}_i$ and is hence discarded. The time domain relation of the input/output equation can thus be expressed as

$$\boldsymbol{y}_i = \boldsymbol{x}_i \circledast \boldsymbol{h}_i + \boldsymbol{n}_i \tag{3}$$

where $\circledast$ represents circular convolution. This relation takes a more transparent form in the frequency domain as

$$\begin{aligned} \boldsymbol{\mathcal{Y}}_i &= \mathrm{diag}(\boldsymbol{\mathcal{X}}_\mathrm{i})\boldsymbol{\mathcal{H}}_\mathrm{i} + \boldsymbol{\mathcal{N}}_\mathrm{i} & (4)\\ &= \mathrm{diag}(\boldsymbol{\mathcal{X}}_\mathrm{i})\boldsymbol{Q}_\mathrm{P}\underline{\boldsymbol{h}}_\mathrm{i} + \boldsymbol{\mathcal{N}}_\mathrm{i} & (5)\end{aligned}$$

where $\boldsymbol{Q}_P$ is the FFT matrix comprised of the first P columns of $\boldsymbol{Q}$, $\boldsymbol{\mathcal{H}}_i$ is related to $\underline{\boldsymbol{h}}_i$ by the relation

$$\boldsymbol{\mathcal{H}}_i = \boldsymbol{Q}\begin{bmatrix} \underline{\boldsymbol{h}}_i \\ 0 \end{bmatrix} = \boldsymbol{Q}_P\underline{\boldsymbol{h}}_i$$

and $\boldsymbol{\mathcal{N}}_i$ is additive white Gaussian noise $\mathcal{N}(0, \sigma_n^2\boldsymbol{I})$. For ease of notation, let us define $\boldsymbol{X}_i$ as $\mathrm{diag}(\boldsymbol{\mathcal{X}}_\mathrm{i})\boldsymbol{Q}_\mathrm{P}$. Thus we can rewrite equation (5) as

$$\boxed{\boldsymbol{\mathcal{Y}}_i = \boldsymbol{X}_i\underline{\boldsymbol{h}}_i + \boldsymbol{\mathcal{N}}_i} \tag{6}$$

This equation gives the input/output relationship of the OFDM system. Now the OFDM symbol $\boldsymbol{X}_i$ consists of both data and pilots. It will be useful for our discussion in the next sections to define a pilot/output relation. Let $I_p = \{i_1, i_2, \cdots, i_{L_p}\}$ be the set of pilot indices within the OFDM symbol known *a priori* at the receiver and let $\boldsymbol{X}_{I_\mathrm{p}}$ be the matrix pruned of rows that do not belong to I_p, i.e. $\boldsymbol{X}_{I_\mathrm{p}}$ is comprised of the pilot rows only of $\boldsymbol{X}$. The pilot/output relationship will thus be a pruned version of (6) and is given as

$$\boxed{\boldsymbol{\mathcal{Y}}_{i,I_p} = \boldsymbol{X}_{i,I_p}\underline{\boldsymbol{h}}_i + \boldsymbol{\mathcal{N}}_{i,I_p}} \tag{7}$$

5. Channel Estimation

In the following, we consider the channel estimation problem when the data is (i) completely known and when it is (ii) partially known (i.e. some training data is available).

5.1. Data Completely Known

Our goal here is to estimate $\underline{\boldsymbol{h}}_i$. We start by performing channel estimation using only one OFDM symbol. We start with the assumption that the transmitted OFDM symbols $\boldsymbol{X}_i$ are known completely at the receiver. This enables us to use input/output equation (6) to obtain the channel estimate $\hat{\underline{\boldsymbol{h}}}_i$ by maximizing the following log likelihood function

$$\hat{\underline{\boldsymbol{h}}}_i^{\mathsf{MAP}} = \arg\max_{\underline{\boldsymbol{h}}_i}\left\{\ln\ p\left(\boldsymbol{\mathcal{Y}}_i|\boldsymbol{X}_i, \underline{\boldsymbol{h}}_i\right) + \ln\ p\left(\underline{\boldsymbol{h}}_i\right)\right\} \tag{8}$$

where $p(z)$ stands for probability density function of z. As the noise is white gaussian, the maximization reduces to

$$\underline{\hat{\boldsymbol{h}}}_i^{\mathsf{MAP}} = \arg\min_{\underline{h}_i} \left\{ ||\boldsymbol{\mathcal{Y}}_i - \boldsymbol{X}_i \underline{\boldsymbol{h}}_i||^2_{\boldsymbol{R}_{\mathcal{N}}^{-1}} + ||\underline{\boldsymbol{h}}_i||^2_{\boldsymbol{\Pi}^{-1}} \right\} \quad (9)$$

$$= \boldsymbol{\Pi} \boldsymbol{X}_i^* \left[\boldsymbol{R}_{\mathcal{N}} + \boldsymbol{X}_i \boldsymbol{\Pi} \boldsymbol{X}_i^* \right]^{-1} \boldsymbol{\mathcal{Y}}_i \quad (10)$$

where $||\boldsymbol{a}||^2{}_{\boldsymbol{A}} \triangleq \boldsymbol{a}^* \boldsymbol{A} \boldsymbol{a}$ The Maximum A Posteriori (MAP) solution (10) makes use of the information at the i^{th} symbol $\boldsymbol{\mathcal{Y}}_i$ only to estimate $\underline{\boldsymbol{h}}_i$. This is clearly suboptimal as $\underline{\boldsymbol{h}}_i$ is correlated with $\underline{\boldsymbol{h}}_j$ for $j = 0, 1, ..., T$ and hence it is correlated with all the output symbols $\boldsymbol{\mathcal{Y}}_0, \boldsymbol{\mathcal{Y}}_1, ..., \boldsymbol{\mathcal{Y}}_T$ (and not just $\boldsymbol{\mathcal{Y}}_i$). To make full use of this correlation, we maximize the log likelihood function of the whole sequence $\boldsymbol{h}_0^T = \boldsymbol{h}_0, \boldsymbol{h}_1, ..., \boldsymbol{h}_T$ given the whole output sequence $\boldsymbol{\mathcal{Y}}_0^T$, i.e. we maximize the MAP estimate of channel tap sequence $\underline{\boldsymbol{h}}_0^T$ by maximizing the following log likelihood function

$$\mathcal{L} = \ln\ p(\boldsymbol{\mathcal{Y}}_0^T | \boldsymbol{X}_0^T, \underline{\boldsymbol{h}}_0^T) + \ln\ p(\underline{\boldsymbol{h}}_0^T) \quad (11)$$

It can be shown that since the noise and channel IR sequence are Gaussian, the likelihood function takes the quadratic form (Al-Naffouri, 2007)

$$\mathcal{L} = \sum_{i=0}^{T} \ln\ p(\boldsymbol{\mathcal{Y}}_i | \boldsymbol{X}_i, \underline{\boldsymbol{h}}_i) + \sum_{i=1}^{T} \ln\ p(\underline{\boldsymbol{h}}_i | \underline{\boldsymbol{h}}_{i-1}) + \ln\ p(\underline{\boldsymbol{h}}_0) \quad (12)$$

$$= -\sum_{i=0}^{T} ||\boldsymbol{\mathcal{Y}}_i - \boldsymbol{X}_i \underline{\boldsymbol{h}}_i||^2_{\boldsymbol{R}_{\mathcal{N}}^{-1}} - \sum_{i=1}^{T} ||\underline{\boldsymbol{h}}_i - \boldsymbol{F} \underline{\boldsymbol{h}}_{i-1}||^2_{\frac{1}{\sigma_u^2} GG^*} - ||\underline{\boldsymbol{h}}_0||^2_{\Pi_0^{-1}} \quad (13)$$

up to some additive constant.

This can be formulated as a least square problem in the sequence $\underline{\boldsymbol{h}}_0^T$. Alternatively, since the MAP estimate coincide with the Minimum Mean Square Error (MMSE) estimate for Gaussian data, we can obtain the sequence $\underline{\boldsymbol{h}}_0^T$ by applying the Forward Backward (FB) Kalman filter to the state space model

$$\underline{\boldsymbol{h}}_{i+1} = \boldsymbol{F} \underline{\boldsymbol{h}}_i + \boldsymbol{G} \underline{\boldsymbol{u}}_i \quad (14)$$

$$\boldsymbol{\mathcal{Y}}_i = \boldsymbol{X}_i \underline{\boldsymbol{h}}_i + \boldsymbol{\mathcal{N}}_i \quad (15)$$

The Forward backward Kalman filter which provides the MAP estimate (and also the MMSE estimate) is described by the following set of equations (Kailath et al., 2000):

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

$$\boldsymbol{R}_{e,i} = \sigma_n^2 \boldsymbol{I}_{N+P} + \boldsymbol{X}_i \boldsymbol{P}_{i|i-1} \boldsymbol{X}_i^* \quad (16)$$

$$\boldsymbol{K}_{f,i} = \boldsymbol{P}_{i|i-1} \boldsymbol{X}_i^* \boldsymbol{R}_{e,i}^{-1} \quad (17)$$

$$\underline{\hat{\boldsymbol{h}}}_{i|i} = (\boldsymbol{I}_{N+P} - \boldsymbol{K}_{f,i} \boldsymbol{X}_i)\, \underline{\hat{\boldsymbol{h}}}_{i|i-1} + \boldsymbol{K}_{f,i} \boldsymbol{\mathcal{Y}}_i \quad (18)$$

$$\underline{\hat{\boldsymbol{h}}}_{i+1|i} = \boldsymbol{F} \underline{\hat{\boldsymbol{h}}}_{i|i} \quad (19)$$

$$\boldsymbol{P}_{i+1|i} = \boldsymbol{F} \left(\boldsymbol{P}_{i|i-1} - \boldsymbol{K}_{f,i} \boldsymbol{R}_{e,i} \boldsymbol{K}_{f,i}^* \right) \boldsymbol{F}^* + \frac{1}{\sigma_n^2} \boldsymbol{G}_\alpha \boldsymbol{G}_\alpha^* \quad (20)$$

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

$$\boldsymbol{\lambda}_{i|T} = \left(\boldsymbol{I}_{P+N} - \boldsymbol{X}_i^* \boldsymbol{K}_{f,i}^* \right) \boldsymbol{F}^* \boldsymbol{\lambda}_{i+1|T} + \boldsymbol{X}_i \boldsymbol{R}_{e,i}^{-1} \left(\boldsymbol{\mathcal{Y}}_i - \boldsymbol{X}_i \underline{\hat{\boldsymbol{h}}}_{i|i-1} \right) \quad (21)$$

$$\underline{\hat{\boldsymbol{h}}}_{i|T} = \underline{\hat{\boldsymbol{h}}}_{i|i-1} + \boldsymbol{P}_{i|i-1} \boldsymbol{\lambda}_{i|T} \quad (22)$$

where $\hat{\underline{\boldsymbol{h}}}_{i|T}$ is the desired estimate of the channel taps.

5.2. Data Partially Known

When only the training part of the data is known, we estimate the channel from the state space model

$$\underline{\boldsymbol{h}}_{i+1} = \boldsymbol{F}\underline{\boldsymbol{h}}_i + \boldsymbol{G}\underline{\boldsymbol{u}}_i \tag{23}$$

$$\boldsymbol{\mathcal{Y}}_{i,I_p} = \boldsymbol{X}_{i,I_p}\underline{\boldsymbol{h}}_i + \boldsymbol{\mathcal{N}}_{i,I_p} \tag{24}$$

This dynamical model is different from the model (14) - (15) through the input equation (specifically $\boldsymbol{X}_i$ is replaced by $\boldsymbol{X}_{i,I_p}$ and $\boldsymbol{\mathcal{Y}}_i$ by $\boldsymbol{\mathcal{Y}}_{i,I_p}$). Thus the training based MAP estimate of the channel sequence $\underline{\boldsymbol{h}}_0^T$ is obtained by applying the Forward Backward Kalman filter to the dynamical model (14) - (15), i.e. by applying (16) - (22) with the substitution $\boldsymbol{X}_i \rightarrow \boldsymbol{X}_{i,I_p}$ and $\boldsymbol{\mathcal{Y}}_i \rightarrow \boldsymbol{\mathcal{Y}}_{i,I_p}$.

5.3. Iterative Channel Estimation

Using a sufficiently large number of pilots would yield a good estimate of channel but it will consume the valuable bandwidth of the system. For this reason, it is desirable to keep the number of pilots in the system to a minimum. The data, which constitutes the larger part of the received symbol, is not being used and hence we are not fully using the constraints on the data. This provides the motivation for using some data aided technique. Since the data is unknown at the receiver, we make use of the expectation maximization (EM) algorithm. The EM algorithm is used when some of the data needed for the estimation process is unavailable. To motivate the EM algorithm, consider the MAP estimate of the channel $\underline{\boldsymbol{h}}_i$ described by (9). Since the data $\boldsymbol{X}_i$ is unknown, we can not minimize (9) for $\underline{\boldsymbol{h}}_i$. To go around this, we minimize (9) averaged over the data $\boldsymbol{X}_i$ i.e.

$$\hat{\underline{\boldsymbol{h}}}_i^{(j)} = \arg\min_{\underline{h}_i} \left\{ E_{X_i|\mathcal{Y}_i,\underline{h}_i^{(j-1)}} \left\{ \ln\ p\left(\boldsymbol{\mathcal{Y}}_i|\boldsymbol{X}_i,\underline{\boldsymbol{h}}_i\right) + \ln\ p\left(\underline{\boldsymbol{h}}_i\right)\right\}\right\} \tag{25}$$

As indicated in (25), this is an iterative procedure as it gives the estimate of $\underline{\boldsymbol{h}}_i$ at the j^{th} iteration by utilizing the estimate at the $(j-1)^{th}$ iteration. Thus, the expectation with respect to $\boldsymbol{X}_i$ is taken given the output data $\boldsymbol{\mathcal{Y}}_i$ and given the channel estimate at the $(j-1)^{th}$ step, $\hat{\underline{\boldsymbol{h}}}_i^{(j-1)}$. By evaluating the expectation and completing the squares, we can rewrite (25) as

$$\hat{\underline{\boldsymbol{h}}}_i^{(j)} = \arg\min_{\underline{h}_i} \left\{ ||\boldsymbol{\mathcal{Y}}_i - E[\boldsymbol{X}_i]\underline{\boldsymbol{h}}_i||^2_{\frac{1}{\sigma_n^2}} + ||\underline{\boldsymbol{h}}_i||^2_{\frac{1}{\sigma_n^2}Cov[X_i^*]} + ||\underline{\boldsymbol{h}}_i||^2_{\Pi^{-1}} \right\} \tag{26}$$

where $E[\boldsymbol{X}_i]$ is the expected value of $\boldsymbol{X}_i$ and $Cov[\boldsymbol{X}_i^*]$ is its covariance. Combining the first two terms of the above equation, we can get a concise form of the log likelihood function as

$$\hat{\underline{\boldsymbol{h}}}_i^{(j)} = \arg\min_{\underline{h}_i} \left\{ \left\| \begin{bmatrix} \boldsymbol{\mathcal{Y}}_i \\ 0_{P\times 1} \end{bmatrix} - \begin{bmatrix} E[\boldsymbol{X}_i] \\ Cov[\boldsymbol{X}_i^*]^{1/2} \end{bmatrix} \underline{\boldsymbol{h}}_i \right\|^2_{\frac{1}{\sigma_n^2}} + ||\underline{\boldsymbol{h}}_i||^2_{\Pi^{-1}} \right\} \tag{27}$$

Comparing the quadratic form (9) for the known data case with the form (27) for the unknown (or partially known) data case, we note that to move from the former to the latter, we need to perform the substitution

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

This step is the maximization step of the EM algorithm. It remains to perform the expectation step which will be considered below.

As $\boldsymbol{X}_i = \text{diag}(\boldsymbol{\mathcal{X}}_i)\boldsymbol{Q}_P$, we can express the mean and the covariance of $\boldsymbol{X}_i$ in terms of $\boldsymbol{\mathcal{X}}_i$ as

$$E[\boldsymbol{X}_i|\boldsymbol{\mathcal{Y}}_i, \hat{h}_i^{(j-1)}] = \text{diag}(E[\boldsymbol{\mathcal{X}}_i|\boldsymbol{\mathcal{Y}}_i, \hat{h}_i^{(j-1)}])\boldsymbol{Q}_P \tag{28}$$

$$Cov[\boldsymbol{X}_i^*|\boldsymbol{\mathcal{Y}}_i, \boldsymbol{\mathcal{H}}_i] = \boldsymbol{Q}_P^* Cov[\boldsymbol{\mathcal{X}}_i^*|\boldsymbol{\mathcal{Y}}_i, \boldsymbol{\mathcal{H}}_i]\boldsymbol{Q}_P \tag{29}$$

This decoupled form enables us to calculate the above expectations on an element by element basis $\mathcal{X}_i(l)$ $l = 1, 2, \ldots, N$. For the case when $\mathcal{X}_i(l)$ is drawn from a finite alphabet set $A = \{A_1, A_2, \ldots, A_{|A|}\}$ with equal probability, we can show that

$$f(\mathcal{X}_i(l)|\mathcal{Y}_i(l), \mathcal{H}_i(l)) = \frac{e^{-\frac{|\underline{\mathcal{Y}}(l)-\underline{\mathcal{H}}(l)A_j|^2}{\sigma^2}}}{\sum_{j=1}^{M} e^{-\frac{|\underline{\mathcal{Y}}(l)-\underline{\mathcal{H}}(l)A_j|^2}{\sigma^2}}} \tag{30}$$

using Bayes rule. This result leads to the expectations

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

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

which can in turn be used to evaluate the expectation and covariance of $\boldsymbol{X}_i$.

The initial estimate $\underline{\hat{\boldsymbol{h}}}_i^{(0)}$ is obtained using pilots. From this channel estimate, we make an estimate of data and use (26) to estimate the channel again. This process is iterated a number of times to get the final estimate $\underline{\hat{\boldsymbol{h}}}_i$.

6. The EM-based Forward Backward Kalman Filter

The algorithm above performs iterative channel estimation (of $\underline{\boldsymbol{h}}_i$) and data detection (of $\boldsymbol{\mathcal{X}}_i$) using the i^{th} OFDM symbol $\boldsymbol{\mathcal{Y}}_i$ only. However, as we mentioned earlier, this is suboptimal as the channel $\underline{\boldsymbol{h}}_i$ is correlated with the symbols $\boldsymbol{\mathcal{Y}}_0$, $\boldsymbol{\mathcal{Y}}_1, \ldots, \boldsymbol{\mathcal{Y}}_T$. Thus, in the partially known data case, we can design an EM based FB Kalman filter as follows

Maximization Step: The maximization step obtains estimate of $\underline{\boldsymbol{h}}_i$ at the j^{th} iteration by applying the FB Kalman to the state space model

$$\underline{\boldsymbol{h}}_{i+1} = \boldsymbol{F}\underline{\boldsymbol{h}}_i + \boldsymbol{G}\underline{\boldsymbol{u}}_i \tag{33}$$

$$\begin{bmatrix} \boldsymbol{\mathcal{Y}}_i \\ 0_{P\times 1} \end{bmatrix} = \begin{bmatrix} E[\boldsymbol{X}_i] \\ Cov[\boldsymbol{X}_i^*]^{1/2} \end{bmatrix} \underline{\boldsymbol{h}}_i + \begin{bmatrix} \boldsymbol{\mathcal{N}}_i \\ \underline{\boldsymbol{n}}_i \end{bmatrix} \tag{34}$$

i.e. by applying the FB Kalman (16) - (22) with the change of variables

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

Expectation Step: The expectation step in the FB Kalman case is exactly the same as described by equations (28) - (32).

Initialization Step: The initial channel estimate $\underline{\hat{\boldsymbol{h}}}_i^{(0)}$ is obtained by applying the FB Kalman filter to the dynamical model (23) - (24).

There are several possible implementations of incorporating the EM algorithm in the FB Kalman filter. For time correlated channels, there are two dimensions we can iterate along (i) between channel estimation & data detection or (ii) against time. We get different receiver structures depending upon the scheduling of these iterations.

6.1. Cyclic FB Kalman

If we iterate between channel estimation & data detection after the entire forward-backward run of the Kalman, the filter formed is called a cyclic FB Kalman. The pilot based estimate serves as the initial estimate to jump start the data aided version. The channel estimate is then given by the FB Kalman. This estimate is again used to refine the data and the entire process is iterated. The iterations process the OFDM symbols in a circular manner motivating the name of this filter structure as shown in figure 1.2.

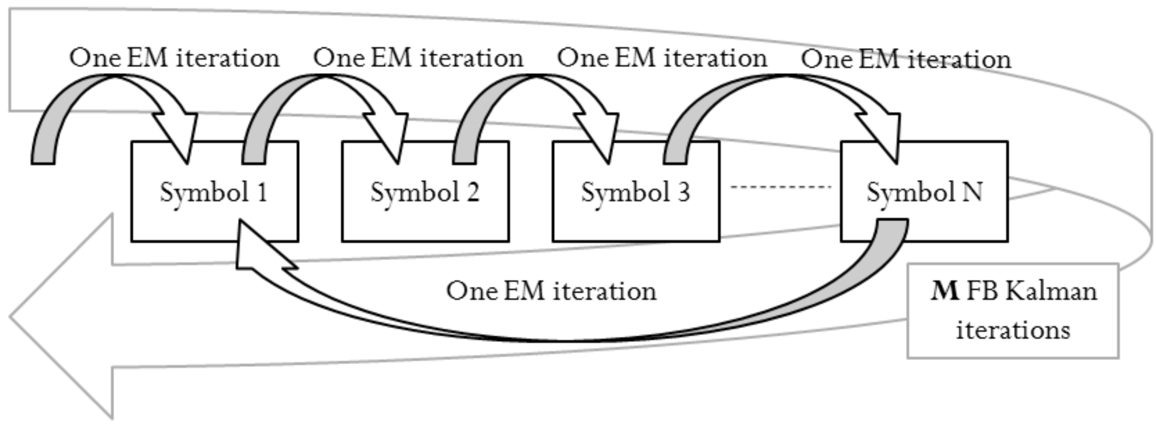

Figure 1.2: Cyclic FB Kalman

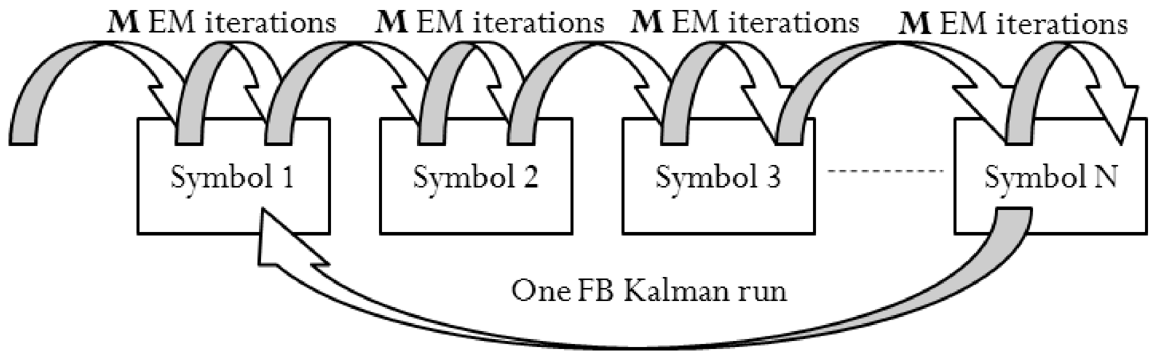

Figure 1.3: Helix FB Kalman

6.2. Helix FB Kalman

If we structure the filter to iterate between channel estimation and data detection at each step of the Kalman, the filter formed is called a helix FB Kalman. Here, the initial estimate is again provided by the pilots. Now, at each step of the Kalman, the filter iterates several times between channel estimation and data detection at each OFDM symbol, using one to enhance the other and so on, before moving to the next symbol using the Kalman filter. Figure 1.3 shows that the iterations trace a helix which explains the name we chose for this filter.

6.3. Forward only Kalman

Latency and memory is an issue with the FB Kalman filter as the receiver needs all the symbols within an OFDM block to estimate the channel. By implementing the forward part of the Kalman filter only, we can get rid of the storage and latency issues at the cost of decreased performance.

7. The FB-Kalman based Receivers

The FB Kalman filter described in the previous section can be applied in several other receiver scenarios. To do so, we need to provide the following

1. A dynamical equation that describes the evolution of the channel impulse response (similar to (1)). The FB Kalman can be applied then to obtain the iterative channel estimate (maximization step).
2. Input/output equation for data recovery. This is used to evaluate the first and second moments of the data (expectation step).
3. A dynamical equation that can be used for initial channel estimation via the FB Kalman filter.

In the following, we shall apply this procedure to two communication scenarios namely channel estimation in multiuser OFDM and channel estimation in MIMO OFDM.

8. Channel Estimation in Multi-Access OFDM Systems

In a multi-access OFDM system, the available bandwidth of the system is shared between a number of users. Each user thus has access to only a specific portion of the OFDM spectrum. This has some important ramifications as far as channel estimation is concerned. Specifically, each user is only interested in the estimate of the particular band in which it is operating. In time domain based channel estimation techniques the user is required to estimate the entire spectrum and hence this proves to be computationally expansive in the multi-user case. As such, frequency domain channel estimation methods for the multi-access scenario would make more sense as this would allow each user to estimate the part of spectrum in which it is operating and not the entire spectrum.

The only problem with frequency domain channel estimation is the increased number of parameters to be estimated in this case (from P to N, where P is the number of time domain channel taps and N is the number of frequency bins used). As such, we can use some parameter reduction model to reduce the number of parameters to be estimated.

8.1. A Parameter Reduction Approach

Let k users be operating in a multi-access OFDM system. The frequency response of the entire spectrum is of length N. For simplicity, we will assume that all users share the bandwidth equally. Each user thus operates in a band (or section) of length $L_f = \lceil \frac{N}{k} \rceil$. Let $\underline{\boldsymbol{\mathcal{H}}}_i^{(j)}$ be the j^{th} section of the frequency response, then from (1), the input/output equation of the j^{th} user is given by

$$\underline{\boldsymbol{\mathcal{Y}}}_i^{(j)} = \operatorname{diag}(\underline{\boldsymbol{\mathcal{X}}}_i^{(j)})\underline{\boldsymbol{\mathcal{H}}}_i^{(j)} + \underline{\boldsymbol{\mathcal{N}}}_i^{(j)} \tag{35}$$

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. We will suppress the dependence on the user index j and time index i for notational convenience. Denoting the pilot locations by the subscript $I_p^{(j)}$, we can write the pilot/output equation as

$$\underline{\boldsymbol{\mathcal{Y}}}_{I_p} = \operatorname{diag}(\underline{\boldsymbol{\mathcal{X}}}_{I_p})\underline{\boldsymbol{\mathcal{H}}}_{I_p} + \underline{\boldsymbol{\mathcal{N}}}_{I_p} \tag{36}$$

There are not enough pilots to estimate $\underline{\boldsymbol{\mathcal{H}}}$. So, we resort to model reduction starting from the autocorrelation function of $\underline{\boldsymbol{\mathcal{H}}}$, $\boldsymbol{R}_{\underline{\boldsymbol{\mathcal{H}}}}$. The eigenvalue decomposition of $\boldsymbol{R}_{\underline{\boldsymbol{\mathcal{H}}}}$ is given by

$$\boldsymbol{R}_{\underline{\boldsymbol{\mathcal{H}}}} = \sum_{l=1}^{L_f} \lambda_l \mathbf{v}_l \mathbf{v}_l^T \tag{37}$$

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. Using this decomposition, $\underline{\boldsymbol{\mathcal{H}}}$ an be represented as

$$\underline{\boldsymbol{\mathcal{H}}} = \sum_{l=1}^{L_f} \alpha_l \mathbf{v_l} \tag{38}$$

where $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \ldots, \alpha_{L_f}]^T$ is the parameter vector to be estimated, with zero mean and autocorrelation matrix $\boldsymbol{\Lambda} = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_{L_f})$. Using (37) and (38), we can represent $\underline{\boldsymbol{\mathcal{H}}}$ in terms of its dominant eigenvalues and treat the insignificant eigenvalues as modeling noise, i.e.

$$\underline{\boldsymbol{\mathcal{H}}} = \boldsymbol{V}_d \boldsymbol{\alpha}_d + \boldsymbol{V}_n \boldsymbol{\alpha}_n \tag{39}$$

Upon substituting (39) in (35), we get

$$\underline{\boldsymbol{\mathcal{Y}}} = \operatorname{diag}(\underline{\boldsymbol{\mathcal{X}}})\boldsymbol{V}_d \boldsymbol{\alpha}_d + \underline{\boldsymbol{\mathcal{N}}} + \operatorname{diag}(\underline{\boldsymbol{\mathcal{X}}})\boldsymbol{V}_n \boldsymbol{\alpha}_n \tag{40}$$

$$= \underline{\boldsymbol{X}}_d \boldsymbol{\alpha}_d + \underline{\boldsymbol{\mathcal{N}}}' \tag{41}$$

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

$$\boldsymbol{R}_{\underline{\mathcal{N}}'} = \boldsymbol{R}_{\underline{\mathcal{N}}} + \operatorname{diag}(\underline{\boldsymbol{\mathcal{X}}})\boldsymbol{V}_n \operatorname{diag}(\lambda_n) \boldsymbol{V}_n^* \operatorname{diag}(\underline{\boldsymbol{\mathcal{X}}})^* \tag{42}$$

Based on (41) we can construct a pilot/output equation similar to (36), as

$$\boxed{\underline{\boldsymbol{\mathcal{Y}}}_{I_p} = \underline{\boldsymbol{X}}_{d,I_p} \boldsymbol{\alpha}_d + \underline{\boldsymbol{\mathcal{N}}}'_{I_p}} \tag{43}$$

Equation (41) is the equation we need for data recovery. Equation (43) can be used for initial channel estimation (See (Al-Naffouri and Sohail, 2008) for further details). All that is needed now is to develop a dynamical model for the interpolation parameter.

8.2. Developing a Frequency Domain Time-Variant Model

In this subsection, we develop a state space model for the parameter $\boldsymbol{\alpha}_d$. To this end, let us consider the channel model of (1) and 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 (Al-Naffouri, 2007)). Now as $\boldsymbol{\mathcal{H}}_i$ is just the channel response $\boldsymbol{h}_i$ in frequency domain ($\boldsymbol{\mathcal{H}}_i = \boldsymbol{Q}_P \boldsymbol{h}_i$), the j^{th} section of $\boldsymbol{\mathcal{H}}_i$(i.e. $\boldsymbol{\mathcal{H}}_i^{(j)}$) is related to $\boldsymbol{h}_i$ by

$$\underline{\boldsymbol{\mathcal{H}}}_i^{(j)} = \boldsymbol{Q}_P^{(j)} \boldsymbol{h}_i \tag{44}$$

where $\boldsymbol{Q}_P^{(j)}$ is the j^{th} section of $\boldsymbol{Q}_P$. Replacing $\underline{\boldsymbol{\mathcal{H}}}_i^{(j)}$ in (44) by its representation, we get a relation between the time domain channel response and the dominant parameters $\boldsymbol{\alpha}_d$ as

$$\boldsymbol{V}_d\boldsymbol{\alpha}_{d,i} = \boldsymbol{Q}_P^{(j)}\boldsymbol{h}_i \tag{45}$$

$$\boldsymbol{\alpha}_{d,i} = \boldsymbol{V}_d^{+}\boldsymbol{Q}_P^{(j)}\boldsymbol{h}_i \tag{46}$$

where $\boldsymbol{V}_d^{+}$ is the pseudo inverse of $\boldsymbol{V}_d$. Combining (1) and (46) yields a dynamical recursion for $\boldsymbol{\alpha}_d$

$$\boxed{\boldsymbol{\alpha}_{d,i+1} = \boldsymbol{F}_\alpha\boldsymbol{\alpha}_{d,i} + \boldsymbol{G}_\alpha\boldsymbol{u}_i} \tag{47}$$

where $\boldsymbol{F}_\alpha = f\boldsymbol{I}$ and $\boldsymbol{G}_\alpha = \sqrt{1-f^2}\boldsymbol{V}_d^{+}\boldsymbol{Q}_P^{(j)}$ and where

$$E[\boldsymbol{\alpha}_{d,0}\boldsymbol{\alpha}_{d,0}^{*}] = \boldsymbol{\Lambda}_d \tag{48}$$

here we suppress the dependence of $\boldsymbol{G}_\alpha$ and $\boldsymbol{\alpha}_d$ on j for notational convenience.

We are now ready to implement our FB Kalman based receiver, which consists of an initial channel estimation step and an iterative channel estimation step.

8.3. Initial (Pilot-Based) Channel Estimation

In multi-access OFDM, the initial estimate is given by applying the FB Kalman filter (16) - (22) to the following state space model:

$$\underline{\boldsymbol{\mathcal{Y}}}_{I_p,i} = \underline{\boldsymbol{X}}_{d,I_p,i}\boldsymbol{\alpha}_{d,i} + \underline{\boldsymbol{\mathcal{N}}}'_{I_p,i} \tag{49}$$

$$\boldsymbol{\alpha}_{d,i+1} = \boldsymbol{F}_\alpha\boldsymbol{\alpha}_{d,i} + \boldsymbol{G}_\alpha\boldsymbol{u}_i \tag{50}$$

8.4. Iterative (Data-Aided) Channel Estimation

The iterative channel estimation step is obtained by applying FB Kalman to the state space model

$$\underline{\boldsymbol{\mathcal{Y}}}_i = \begin{bmatrix} E[\underline{\boldsymbol{X}}_{d,i}] \\ Cov[\underline{\boldsymbol{X}}_{d,i}^{*}]^{\frac{1}{2}} \end{bmatrix}\boldsymbol{\alpha}_{i,d} + \begin{bmatrix} \underline{\boldsymbol{\mathcal{N}}}'_i \\ \boldsymbol{0} \end{bmatrix} \tag{51}$$

$$\boldsymbol{\alpha}_{d,i+1} = \boldsymbol{F}_\alpha\boldsymbol{\alpha}_{d,i} + \boldsymbol{G}_\alpha\boldsymbol{u}_i \tag{52}$$

The data expectations in (51) are obtained from the input/output equation (41). Consequently, the FB Kalman is applied to the above set. As mentioned in Section 5, the FB Kalman can be implemented as Cyclic Kalman, Helix Kalman or forward only Kalman.

Figure 1.4 compares the Bit Error Rate (BER) performance of three implementations of Kalman filter with the simple EM based least square estimation method. 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. The OFDM symbol serves 4 users each occupying 16 frequency bins. In addition, the OFDM symbol carries 16 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. As expected, the estimate improves when we use time correlation information (by using a Kalman filter). Figure 1.4 is plotted for the uncoded case. Since we are relying on the data to improve the channel estimate, an outer code can improve the data quality and hence the quality of the channel estimate. Figure 1.6 shows the advantage of using coding (1/2 rate convolutional code) during the iterative process over the performance of iterative process that only makes use of the code to correct the data at the last step of the algorithm.

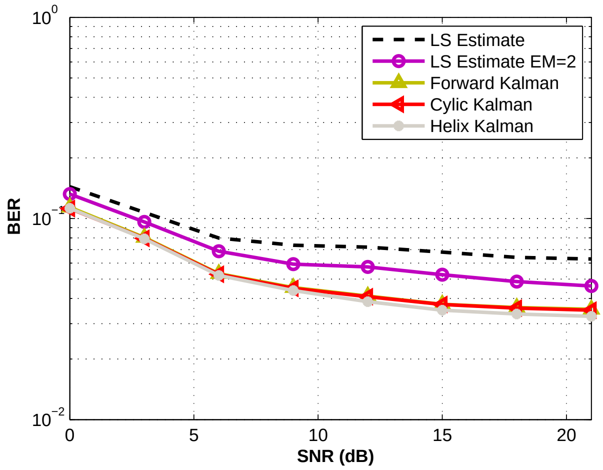

Figure 1.4: Comparison of various uncoded Kalman implementations.

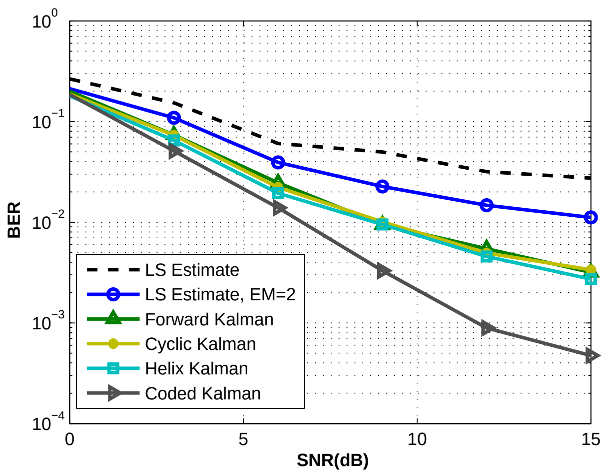

Figure 1.5: Comparison of Kalman implementations using code.

9. Channel Estimation in MIMO OFDM Systems

As wireless communication is becoming more and more integrated in everyday life, service operators are offering new and advanced features like video streaming and broadband connections. These services demand high spectral performance. By using Multi Input Multi Output (MIMO) OFDM techniques, we can increase the spectral efficiency and throughput of wireless systems. As we have seen so far, the problem of channel estimation in a wireless system is a challenging one. In the case of MIMO OFDM systems, it becomes compounded by the fact that transmitters and receivers employ multiple antennas which substantially increase the number of parameters to be estimated. In this section, we will extend the FB Kalman receiver to a Space Time Block Coded (STBC) MIMO OFDM system.

9.1. MIMO Channel Model

We start by defining the MIMO channel model. Thus consider a MIMO system with T_x transmission antennas and R_x receiver antennas. The time domain input/output relationship for a general MIMO system can be described by

$$\boldsymbol{y}(m) = \sum_{p=0}^{P} \boldsymbol{H}(p)\boldsymbol{x}(m-p) \tag{53}$$

where $\boldsymbol{H}(p)$ is the $R_x \times T_x$ MIMO impulse response at tap p and where m represents the sample time index. The effect of the transmit filter and transmit and receive correlation are incorporated in $\boldsymbol{H}(p)$ making it correlated across space and taps. For simplicity, $\boldsymbol{H}(p)$ is assumed to be iid. Again, a block fading channel model is considered where the changes from the current block ($\boldsymbol{H}_t(p)$) to the next block ($\boldsymbol{H}_{t+1}(p)$) take place according to the dynamical equation

$$\boldsymbol{H}_{t+1}(p) = \alpha(p)\boldsymbol{H}_t(p) + \sqrt{(1-\alpha^2(p))e^{-\beta p}}\boldsymbol{U}_t(p) \tag{54}$$

where $\alpha(p)$ is related to the Doppler frequency $f_D(p)$ by $\alpha(p) = J_0(2\pi f_D(p)T)$ (T being the time duration of one ST block), β is the exponent of the channel decay profile while the factor $\sqrt{(1-\alpha^2(p))e^{-\beta p}}$ represents the exponential decay profile ($e^{-\beta p}$) for all time and $\boldsymbol{U}_t(p) \sim \mathcal{N}(0,1)$ is an iid matrix. The model approximates the non rational Jakes model by a 1^{st} order AR model. A higher order AR model would give a better approximation but at the expense of increased latency at the receiver. From (54), we can obtain the impulse response $\underline{\boldsymbol{h}}_{r_x}^{t_x}$ between transmit antenna t_x and receive antenna r_x.

$$h_{r_x t+1}^{t_x}(p) = \alpha(p)h_{r_x t}^{t_x}(p) + \sqrt{(1-\alpha^2(p))e^{-\beta p}}u_{r_x t}^{t_x}(p) \tag{55}$$

Stacking (55) over the taps $p = 0, 1, \ldots, P$, we get the dynamical model

$$\underline{\boldsymbol{h}}_{r_x t+1}^{t_x} = \boldsymbol{F}\underline{\boldsymbol{h}}_{r_x}^{t_x} + \boldsymbol{G}\boldsymbol{u}_{r_x}^{t_x} \tag{56}$$

where $\underline{\boldsymbol{h}}_{r_x}^{t_x}$ is the channel IR at $r_x = 1 \cdots R_x$ and $t_x = 1 \cdots T_x$, and $\boldsymbol{F}$ and $\boldsymbol{G}$ are the same as given by (2). Stacking (56) for all transmit and receive antennas, we obtain

$$\boxed{\underline{\boldsymbol{h}}_{t+1} = (\boldsymbol{I}_{T_x R_x} \otimes \boldsymbol{F})\,\underline{\boldsymbol{h}}_t + (\boldsymbol{I}_{T_x R_x} \otimes \boldsymbol{G})\,\boldsymbol{u}_t} \tag{57}$$

where

$$\underline{\boldsymbol{h}}_t = \begin{bmatrix} \underline{\boldsymbol{h}}_1^1 & \cdots & \underline{\boldsymbol{h}}_1^{T_x} & \underline{\boldsymbol{h}}_2^1 & \cdots & \underline{\boldsymbol{h}}_2^{T_x} & \cdots & \underline{\boldsymbol{h}}_{R_x}^1 & \cdots & \underline{\boldsymbol{h}}_{R_x}^{T_x} \end{bmatrix}^T$$

and where the vectors $\underline{\boldsymbol{h}}_{t+1}$, $\underline{\boldsymbol{h}}_t$, and $\boldsymbol{u}_t$, are of size $T_x R_x(P) \times 1$. To employ the Kalman filter, we still need to characterize the covariance information of the dynamical model, specifically we need the covariance of $\boldsymbol{u}_t$ and also the covariance of the channel at $t = 0$. It can be shown that

$$E\left[\boldsymbol{u}_t \boldsymbol{u}_t^*\right] = \boldsymbol{I}_{R_x} \otimes E\left[\boldsymbol{u}_{r_x} \boldsymbol{u}_{r_x}^*\right] \tag{58}$$

$$= \boldsymbol{I}_{R_x} \otimes \left(\boldsymbol{I}_{T_x} \otimes E\left[u_{r_x}^{t_x} u_{r_x}^{t_x *}\right]\right) \tag{59}$$

$$= \boldsymbol{I}_{R_x} \otimes \boldsymbol{I}_{T_x} \otimes \boldsymbol{I}_P = \boldsymbol{I}_{T_x R_x(P)} \tag{60}$$

and

$$E\left[\underline{\boldsymbol{h}}_0 \underline{\boldsymbol{h}}_0^*\right] = \boldsymbol{I}_{T_x R_x} \otimes \boldsymbol{G}\boldsymbol{G}^*$$

It is worthwhile to note that while (54) and (57) are equivalent, the latter model is in vector form and will be useful for Kalman filter operations.

In the channel model described above, we did not consider the transmit/receive correlation between the antennas. When both transmit/receive correlation are incorporated in the channel model, the dynamical equation remains same as in (57) but the covariance of $\boldsymbol{u}_t$ is given by (see (Al-Naffouri and Quadeer, 2008) for further details)

$$E[\boldsymbol{u}\boldsymbol{u}^*] = \sum_{p=0}^{P} \boldsymbol{R}(p) \otimes \boldsymbol{T}(p) \otimes \left(\underline{\boldsymbol{I}}^p \boldsymbol{B} \overline{\boldsymbol{I}}^p\right)$$

where $\boldsymbol{T}(p)$ and $\boldsymbol{R}(p)$ are the transmit and receive correlation matrix (of size T_x and R_x) respectively.

The receiver will perform two functions namely channel estimation and data detection. So we need to derive two forms of the input/output equation. The first is a *channel estimation* form, which treats the channel impulse response as the unknown and which together with the dynamical model (57) forms the state space model that is used by the FB Kalman filter. The second is a *data detection* form, which treats the input in its uncoded form as the unknown. In the Single Input Single Output (SISO) case, we defined the frequency domain input/output relationship by (4). For the MIMO OFDM case, the input/output relation between transmit antenna t_x and receive antenna r_x is given by

$$\boldsymbol{\mathcal{Y}}_{r_x}^{t_x} = \text{diag}\left(\boldsymbol{\mathcal{X}}_{t_x}\right) \boldsymbol{Q}_P^* \underline{\boldsymbol{h}}_{r_x}^{t_x} + \boldsymbol{\mathcal{N}}_{r_x} \tag{61}$$

By stacking, the input/output equation at receive antenna r_x can be expressed as

$$\boldsymbol{\mathcal{Y}}_{r_x} = \left[\text{diag}(\boldsymbol{\mathcal{X}}_1) \ \cdots \ \text{diag}(\boldsymbol{\mathcal{X}}_{T_x})\right] (\boldsymbol{I}_{T_x} \otimes \boldsymbol{Q}_P^*) \underline{\boldsymbol{h}}_{r_x} + \boldsymbol{\mathcal{N}}_{r_x} \tag{62}$$

In what follows, we describe the channel estimation version of the input/output equation for STBC MIMO OFDM transmission over block fading channels. We omit the data detection version as it is similar to the SISO case and as it does not directly relate to the operation of the FB Kalman, which is the center of attention of this chapter. The reader can find more information in (Al-Naffouri and Quadeer, 2008).

9.2. Input/Output Equation

The input/output equation in STBC MIMO OFDM transmission can be divided into the following two categories:

1. Channel Estimation Version

2. Data Detection Version

Channel Estimation Version

Consider a MIMO OFDM system which has N_c time slots, T_x transmit antennas and R_x receive antennas. Let N_u be the set of uncoded OFDM symbols $\{\boldsymbol{\mathcal{S}}(1), \ldots, \boldsymbol{\mathcal{S}}(N_u)\}$ which are to be transmitted. We can implement the Alamouti (ST) code using the set of $T_x \times N_c$ matrices $\{\boldsymbol{A}(1), \boldsymbol{B}(1), \ldots, \boldsymbol{A}(N_u), \boldsymbol{B}(N_u)\}$ following the procedure in (Larsson and Stoica, 2003). Thus, the OFDM symbol transmitted from antenna t_x at time n_c is given by

$$\boldsymbol{\mathcal{X}}_{t_x}(n_c) = \sum_{n_u=1}^{N_u} a_{t_x,n_c}(n_u)\text{Re}\ \boldsymbol{\mathcal{S}}(n_u) + jb_{t_x,n_c}(n_u)\text{Im}\ \boldsymbol{\mathcal{S}}(n_u) \tag{63}$$

where $a_{t_x,n_c}(n_u)$ is the (t_x, n_c) element of $\boldsymbol{X}(n_u)$ and $b_{t_x,n_c}(n_u)$ is the (t_x, n_c) element of $\boldsymbol{B}(n_u)$. So, instead of (62), the input/output equation at antenna r_x at OFDM symbol n_c of a ST block takes the form

$$\boldsymbol{\mathcal{Y}}_{r_x}(n_c) = [\text{diag}(\boldsymbol{\mathcal{X}}_1(n_c)) \cdots \text{diag}(\boldsymbol{\mathcal{X}}_{T_x}(n_c))](\boldsymbol{I}_{T_x} \otimes \boldsymbol{Q}_P^*)\underline{\boldsymbol{h}}_{r_x} + \boldsymbol{\mathcal{N}}_{r_x}(n_c)$$

Stacking over all symbols yields

$$\boldsymbol{\mathcal{Y}}_{r_x} = \boldsymbol{X}\underline{\boldsymbol{h}}_{r_x} + \boldsymbol{\mathcal{N}}_{r_x} \tag{64}$$

where

$$\boldsymbol{\mathcal{Y}}_{r_x} = \begin{bmatrix} \boldsymbol{\mathcal{Y}}_{r_x}(1) \\ \vdots \\ \boldsymbol{\mathcal{Y}}_{r_x}(N_c) \end{bmatrix}, \ \boldsymbol{X} = \begin{bmatrix} \{\text{diag}(\boldsymbol{\mathcal{X}}_1(1)) \cdots \text{diag}(\boldsymbol{\mathcal{X}}_{T_x}(1))\}(\boldsymbol{I}_{T_x} \otimes \boldsymbol{Q}_{P+1}^*) \\ \{\text{diag}(\boldsymbol{\mathcal{X}}_1(2)) \cdots \text{diag}(\boldsymbol{\mathcal{X}}_{T_x}(2))\}(\boldsymbol{I}_{T_x} \otimes \boldsymbol{Q}_{P+1}^*) \\ \vdots \\ \{\text{diag}(\boldsymbol{\mathcal{X}}_1(N_c)) \cdots \text{diag}(\boldsymbol{\mathcal{X}}_{T_x}(N_c))\}(\boldsymbol{I}_{T_x} \otimes \boldsymbol{Q}_{P+1}^*) \end{bmatrix} \tag{65}$$

Further stacking this relation for all receive antennas, we get an input/output relationship at *all* frequency bins, for *all* input and output antennas, and for *all* OFDM symbols of the t^{th} ST block

$$\boxed{\boldsymbol{\mathcal{Y}}_t = (\boldsymbol{I}_{R_x} \otimes \boldsymbol{X}_t)\underline{\boldsymbol{h}}_t + \boldsymbol{\mathcal{N}}_t} \tag{66}$$

Pruning (66), we get the set of those equations where the pilots are present and the pilot/output equation takes the form

$$\boxed{\boldsymbol{\mathcal{Y}}_{t_{I_p}} = (\boldsymbol{I}_{R_x} \otimes \boldsymbol{X}_{t_{I_p}})\underline{\boldsymbol{h}}_t + \boldsymbol{\mathcal{N}}_{t_{I_p}}} \tag{67}$$

Data Detection Version

Signal detection MIMO case is done in the same fashion as the SISO case, i.e. on a tone-by-tone basis except that the tones are collected for R_x receive antennas and over N_c time slots (the whole ST block). As mentioned above, we will omit the details of the data detection version as it is similar to the SISO case discussed in Section 5.3. The reader is also invited to check (Al-Naffouri and Quadeer, 2008) for more details.

9.3. Channel Estimation using EM based FB Kalman

As we showed in Section 5, the MAP estimate for a sequence of $T+1$ input and output symbols $\boldsymbol{X}_0^T$ and $\boldsymbol{\mathcal{Y}}_0^T$ is given by (11). For the MIMO case, we can use (57) and (66) to express the channel log likelihood as

$$\begin{aligned}\mathcal{L} &= -\sum_{t=1}^{T}\|\boldsymbol{\mathcal{Y}}_t-(\boldsymbol{I}_{R_x}\otimes\boldsymbol{X}_t)\underline{\boldsymbol{h}}_t\|^2_{\frac{1}{\sigma_n^2}}\\ &\quad -\sum_{t=1}^{T}\|\underline{\boldsymbol{h}}_t-(\boldsymbol{I}_{T_xR_x}\otimes\boldsymbol{F})\underline{\boldsymbol{h}}_{t-1}\|^2_{(GR_uG^*)^{-1}}-\|\underline{\boldsymbol{h}}_0\|^2_{\Pi_0^{-1}}\end{aligned} \tag{68}$$

In what follows, we present the known and the unknown input data cases for channel estimation.

Known Input Case (Initial Channel Estimation)

Let us start with the simple case of known input i.e. when pilots are used to estimate the channel. The MAP estimate of $\underline{\boldsymbol{h}}_0^T$ for the input and output sequences $\boldsymbol{X}_0^T$ and $\boldsymbol{\mathcal{Y}}_0^T$ respectively can be obtained by applying the FB Kalman (16) - (22) on the following state-space model

$$\begin{aligned}\underline{\boldsymbol{h}}_{t+1} &= (\boldsymbol{I}_{T_xR_x}\otimes\boldsymbol{F})\underline{\boldsymbol{h}}_t+(\boldsymbol{I}_{T_xR_x}\otimes\boldsymbol{G})\boldsymbol{u}_t\\ \boldsymbol{\mathcal{Y}}_{t_{I_\mathrm{p}}} &= (\boldsymbol{I}_{R_x}\otimes\boldsymbol{X}_{t_{I_\mathrm{p}}})\underline{\boldsymbol{h}}_t+\boldsymbol{\mathcal{N}}_{t_{I_\mathrm{p}}}\end{aligned}$$

where $\underline{\boldsymbol{h}}_0\sim\mathcal{N}(\mathbf{0},\mathbf{\Pi})$ and $\boldsymbol{u}_t\sim\mathcal{N}(\mathbf{0},\boldsymbol{R}_u)$.

Unknown Input Case

When input is unknown, the EM algorithm can be used to estimate the channel similar to what we described in Section 6. In this case, we maximize the averaged form of log-likelihood (68). Thus, the j^{th} iteration of the EM algorithm is now obtained by maximizing the averaged log-likelihood

$$\overline{\mathcal{L}}=E_{X_0^T|\mathcal{Y}_0^T,h_0^{T(j-1)}}[\mathcal{L}] \tag{69}$$

As shown in Section 5.3, this is done by representing the input by its first and second moments and applying the FB Kalman (16) - (22) to the following state space model

$$\underline{\boldsymbol{h}}_{t+1} = (\boldsymbol{I}_{T_xR_x}\otimes\boldsymbol{F})\underline{\boldsymbol{h}}_t+(\boldsymbol{I}_{T_xR_x}\otimes\boldsymbol{G})\boldsymbol{u}_t \tag{70}$$

$$\begin{bmatrix}\boldsymbol{\mathcal{Y}}_t\\ \mathbf{0}_{T_xR_x(P)\times 1}\end{bmatrix} = \begin{bmatrix}\boldsymbol{I}_{R_x}\otimes E[\boldsymbol{X}_t]\\ \boldsymbol{I}_{R_x}\otimes \mathrm{Cov}[\boldsymbol{X}_t^*]^{1/2}\end{bmatrix}\underline{\boldsymbol{h}}_t+\begin{bmatrix}\boldsymbol{\mathcal{N}}_t\\ \underline{\boldsymbol{n}}_t\end{bmatrix} \tag{71}$$

where $\underline{\boldsymbol{n}}_t$ is virtual noise and is independent of the physical noise $\boldsymbol{\mathcal{N}}_t$.

9.4. Data Detection

The data is detected by using the data detection version of the input/output equation which in turn is used to obtain the first and second moments of the inputs needed in (70) - (71). This is similar to the SISO case described in Section 5.3. The reader can also refer to (Al-Naffouri and Quadeer, 2008) for more details.

To show the favorable behavior of the algorithm, we simulate a MIMO OFDM system in which a 1/2 rate convolutional encoder is used as an outer encoder. 16-QAM with Gray coding is used as the modulation scheme. Orthogonal space time block coding (OSTBC) commonly known as Alamouti code (number of time slots, N_s = 2 and number of transmitters, T_x = 2) is used (Alamouti, 1998). Other parameters used are $\alpha = 0.985$, $\beta = 0.2$, and $P = 7$. Each packet consists of 12 OFDM symbols transmitted over six ST blocks. Each OFDM symbol consists of 64 frequency tones and a cyclic prefix of length 8. The first ST block is comprised of 16 pilots while the number of pilots in subsequent blocks can be varied from zero to 16.

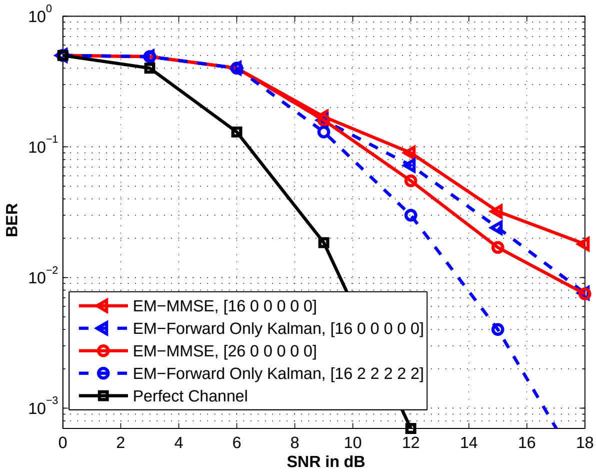

Figure 1.6: Comparison of EM-MMSE and EM-FB Kalman algorithms

To benchmark the proposed algorithm, it is compared with an EM-based iterative MMSE receiver (proposed in (Li et al., 2001) and (Cozzo and Hughes, 2003)) in Figure 1.6 over a spatially white channel. The Expectation step in this algorithm is calculated through MMSE estimation i.e. by a conditional expectation of the channel given the received symbol and the current estimate of the transmitted data. The Maximization step is simply the maximum likelihood estimate of the transmitted data. In this algorithm, the pilots are confined only to the first space time (ST) block which is used to produce the initial channel estimate. In Figure 1.6, the two algorithms are compared for two scenarios with respect to pilots. In first scenario, 16 pilots are used in the first ST block and zero pilots in the subsequent ST blocks. In the second scenario, the EM-MMSE algorithm has 26 pilots in the first ST block and zero pilots in the subsequent ones while the proposed algorithm (EM-FB Kalman) has 16 pilots in the first ST block and 2 pilots in the subsequent blocks, ensuring the same pilot overhead. It can be easily observed that our algorithm outperforms the EM-MMSE algorithm in both scenarios.

Figure 1.7 describes the effect of spatial correlation over performance of the algorithm for a MIMO OFDM transmission. The parameters used here are same except that channel length, $P = 15$, $\alpha = 0.8$, and transmit and receive correlation matrices are given by

$$\boldsymbol{T}(p) = \begin{bmatrix} 1 & \zeta \\ \zeta & 1 \end{bmatrix} \quad \text{and} \quad \boldsymbol{R}(p) = I$$

where $\zeta = 0.8$. The number of pilots in subsequent blocks is fixed at 12 and two EM iterations are used. It can be observed that the performance of both Forward only and FB Kalman is better over spatially correlated channel (practical scenario) as compared to their performance over spatially white channel.

Figure 1.8 compares the performance of the different implementations of the Kalman filter (Forward Only Kalman, Cyclic FB Kalman and Helix based FB Kalman discussed in Section 6) over spatially correlated channel. It can be seen that the Helix based FB Kalman filter outperforms the other two implementations.

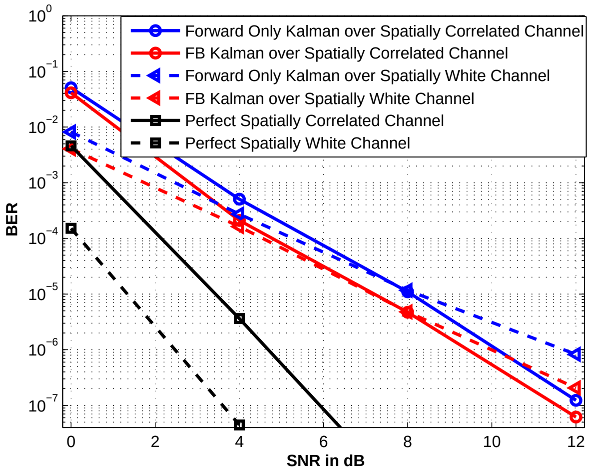

Figure 1.7: BER performance of Forward only Kalman and FB Kalman using Soft data over spatially white and correlated channel models

10. Conclusion

In this chapter, we presented an iterative receiver for data transmission over time variant channels. Such a receiver needs to perform the two tasks of channel estimation and data detection. Moreover, since these two tasks can enhance each other, they were run iteratively. The focus in this chapter was on the channel estimation part, whose maximum likelihood estimate boils down to a forward backward Kalman filter. To run this filter we just need to construct an input/output equation that describes the operation of the channel and a dynamical model that describes the time evolution of the channel. In addition we need an input/output equation that can be used for data detection.

We demonstrated the receiver design for 3 OFDM systems: single user single antenna (SISO) OFDM, multi-access OFDM, and multiple antenna (MIMO) OFDM. Moreover, the three different implementations of the Kalman filter (Cyclic FB Kalman, Helix FB Kalman and Forward only Kalman) were presented along with their comparative performance. The simulation results demonstrated in this chapter indicate that Kalman filter based receivers perform quite well in wireless environment and thus are potential contenders for practical receivers.

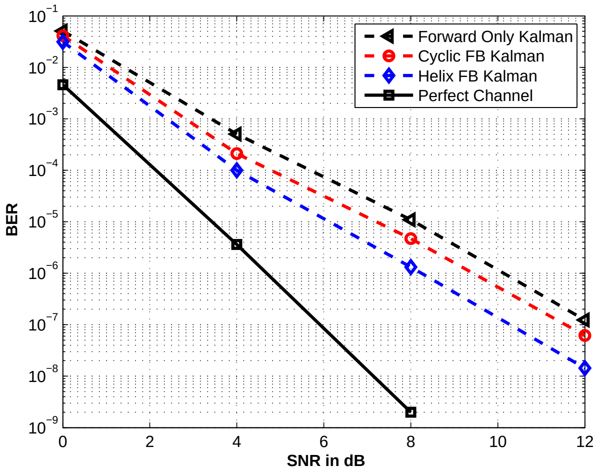

Figure 1.8: Comparison of Forward only Kalman, Cyclic FB Kalman, and Helix FB Kalman with 12 pilots over spatially correlated channel

Bibliography

T. Y. Al-Naffouri. An EM-based forward-backward Kalman filter for the estimation of time-variant channels in OFDM. *IEEE Trans. Signal Proc.*, 55(7):3924–3930, Jul. 2007.

T. Y. Al-Naffouri and A. Quadeer. A Forward-Backward Kalman Filter based STBC MIMO OFDM Receiver. submitted to EURASIP Journal Adv. Signal Proc., May 2008.

T. Y. Al-Naffouri and M. S. Sohail. An EM based Frequency Domain Channel Estimation Algorithm for Multi-Access OFDM Systems. submitted to Elsavier Signal Proc. Journal, Jul. 2008.

T. Y. Al-Naffouri, D. Toumpakaris, A. Bahai, and A. Paulraj. An adaptive semi-blind algorithm for channel identification in OFDM. *Asilomar Conf. on Signals, Syst., and Computers*, 2:921–925, Nov. 2001.

T. Y. Al-Naffouri, A. Bahai, and A. Paulraj. Semi-blind channel identification and equalization in OFDM: an expectation-maximization approach. *IEEE Vehicular Tech. Conf.*, 1:13–17, Sep. 2002.

T. Y. Al-Naffouri, O. Awoniyi, O. Oteri, and A. Paulraj. Receiver design for MIMO-OFDM transmission over time variant channels. *IEEE GLOBECOM*, 4:2487–2492, Dec. 2004.

G. A. Al-Rawi, 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.*, 7(8):388–390, Aug. 2003.

S. M. Alamouti. A simple transmit diversity technique for wireless communications. *IEEE J. Select. Areas Commun.*, 16: 1451–1458, Oct. 1998.

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.*, 51(12):3280–3292, Dec. 2003.

M. Biguesh and A. B. Gershman. MIMO channel estimation: optimal training and tradeoffs between estimation techniques. *IEEE Int. Conf. Commun.*, 5:2658–2662, Jun. 2004.

H. Bölcskei, R. W. Heath, and A. J. Paulraj. Blind channel identification and equalization in OFDM-based multi-antenna systems. *IEEE Trans. Signal Proc.*, 50(1):96–109, Jan. 2002.

M. Chang and Y. T. Su. Blind and semiblind detections of OFDM signals in fading channels. *IEEE Trans. Commun.*, 52(5): 744–754, May 2004.

C. Cozzo and B. L. Hughes. Joint channel estimation and data detection in space-time communications. *IEEE Trans. Commun.*, 51(8):1266–1270, Aug. 2003.

T. Cui and C. Tellambura. Joint data detection and channel estimation for OFDM systems. *IEEE Trans. Commun.*, 54(4): 670–679, Apr. 2006.

O. Edfors, M. Sandell, J. van de Beek, K. S. Wilson, and P. O. Börjesson. OFDM channel estimation by singular value decomposition. *IEEE Trans. Signal Proc.*, 46(7):931–939, Jul. 1998.

F. Gao and A. Nallanathan. Blind channel estimation for MIMO OFDM systems via nonredundant linear precoding. *IEEE Trans. Signal Proc.*, 55(2):784–789, Jan. 2007.

T. Kailath, A. H. Sayed, and B. Hassibi. *Linear Estimation*. Prentice Hall, 2000.

I. Kang, M. P. Fitz, and S. B. Gelfand. Blind estimation of multipath channel parameters: a modal analysis approach. *IEEE Trans. Commun.*, 47(8):1140–1150, Aug. 1999.

T. Kim and I. Eo. Reliable blind channel estimation scheme based on cross-correlated cyclic prefix for OFDM system. *Int. Conf. Adv. Commun. Technol.*, 1, Feb. 2006.

W. Kunji, Z. Jianhua, L. Chaojun, and H. Chen. Semi-blind OFDM channel estimation using receiver diversity in the presence of virtual carriers. *Int. Conf. Commun. and Networking*, 1:1–4, Oct. 2006.

E. Larsson and P. Stoica. *Space-Time Block Coding for Wireless Communications*. Cambridge University Press, 2003.

Y. Li, L. J. Cimini, and N. R. Sollenberger. Robust channel estimation for OFDM systems with rapid dispersive fading channels. *IEEE Trans. Commun.*, 46(7):902–915, Jul. 1998.

Y. Li, C. N. Georghiades, and G. Huang. Iterative maximum-likelihood sequence estimation for space-time coded systems. *IEEE Trans. Commun.*, 49(6):948–951, Jun. 2001.

H. Minn and N. Al-Dhahir. Optimal training signals for MIMO OFDM channel estimation. *IEEE Trans. Wireless Commun.*, 5(5):1158–1168, May 2006.

B. Muquet, M. de Courville, G. B. Giannakis, Z. Wang, and P. Duhamel. Reduced complexity equalizers for zero-padded OFDM transmissions. *IEEE Int. Conf. Acoust., Speech, and Signal Proc.*, 5:2973–2976, Jun. 2000.

M. C. Necker and G. L. Stuber. Totally blind channel estimation for OFDM on fast varying mobile radio channels. *IEEE Trans. Wireless Commun.*, 3(5):1514–1525, Sep. 2004.

R. V. Nee and R. Prasad. *OFDM for wireless multimedia communications*. MA: Artech house, Boston, 2000.

R. Negi and J. Cioffi. Pilot tone selection for channel estimation in a mobile OFDM system. *IEEE Trans. Consumer Electr.*, 44 (3):1122–1128, Aug. 1998.

A. P. Petropulu, R. Zhang, and R. Lin. Blind OFDM channel estimation through simple linear precoding. *IEEE Trans. Wireless Commun.*, 3(4):647–655, Mar. 2004.

A. H. Sayed. A framework for state-space estimation with uncertain models. *IEEE Trans. Auto. Control.*, 46(7):998–1013, Jul. 2001.

Z. Shengli and G. B. Giannakis. Finite-alphabet based channel estimation for OFDM and related multicarrier systems. *IEEE Trans. Commun.*, 49(8):1402–1414, Aug. 2001.

C. Shin, R. W. Heath, and E. J. Powers. Blind channel estimation for MIMO-OFDM systems. *IEEE Trans. Vehicular Tech.*, 56 (2):670–685, Mar. 2007.

G. Stuber, J. Barry, S. Mclaughlin, Y. Li, M. Ingram, and T. Pratt. Broadband MIMO-OFDM wireless communications. *Proc. IEEE*, 92(2):271–294, Feb. 2004.

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.*, 49(3):467–479, Mar. 2001.

F. Yang and W. Ser. Adaptive semi-blind channel estimation for OFDM systems. *IEEE Veh. Technol. Conf.*, 3:1773–1776, May 2004.

R. Zhang and W. Chen. A mixture Kalman filter approach for blind OFDM channel estimation. *Asilomar Conf. on Signals, Syst., and Computers*, 1:350–354, Nov. 2004.