

Receiver Design for MIMO OFDM Transmission over Time Variant Channels

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

Abstract—This paper considers receiver design for space time block coded MIMO OFDM transmission over frequency and time selective channels. The receiver employs the expectation-maximization algorithm for joint channel and data recovery. It makes collective use of the data and channel constraints of the communication problem. The data constraints include pilots, the cyclic prefix, the finite alphabet constraint, and space-time block coding. The channel constraints include the finite delay spread and frequency and time correlation as well as spatial correlation. The channel estimation part of the receiver boils down to an EM-based forward-backward Kalman filter. To avoid the latency and storage associated with smoothing, we introduce a forward-only Kalman that performs channel (and data) recovery with no latency. Simulations show that the receiver outperforms other least-squares based iterative receivers.

I. INTRODUCTION

Orthogonal frequency division multiplexing (OFDM) is a technique that enables high speed transmission over frequency selective channels with simple equalizers. When MIMO is combined with OFDM, space-time codes (STC) can be used per tone, providing the benefit of multiple antennas (diversity and coding gain) with simple channel equalization. Channel state information (CSI) is needed at the receiver. With no receiver CSI, both the channel and the data have to be recovered. Channel estimation techniques rely on some training data as well as *a priori* constraints on the channel and data. The data constraints include pilots, redundancy due to the cyclic prefix [1], [2], the code, and the finite alphabet constraint [3]. The channel constraints include the finite impulse response and time and frequency correlation [4]. The channel estimate can also be improved iteratively in a data-aided fashion [5] or more rigorously by the expectation maximization (EM) approach [6], [7], [8], [9]. This paper considers receiver design for OSTBC-OFDM transmission over a frequency selective, time-variant channel. We propose a semi-blind iterative receiver using the EM algorithm for joint channel and data recovery. Our contributions are as follows:

- 1) We make a collective and optimal use of the structure of the communication problem, i.e. constraints on the data (pilots, space-time code, and soft estimate of the data) and on the channel (finite delay spread, time and frequency correlation, and spatial correlation).
- 2) The channel estimation and data detection as well as the exploitation of the system constraints through the expectation and maximization algorithm which guarantees a relatively simple receiver structure.
- 3) In spite of the complexity of the problem we work on and the many constraints we incorporate, our algorithm maintains its transparency: a) the maximization step is used for channel estimation and makes use of the channel constraints by employing a forward-backward Kalman filter and b) the expectation step is used for data detection and makes use of the data constraints.

A similar algorithm was proposed by the author in [9] for the SISO case. Extending this algorithm to the MIMO case is non-trivial. For in addition to the scale up in the number of transmit and receive antennas, the MIMO case incorporates ST coding and makes a full use of transmit and receive correlation.

A. Notation

We denote scalars with small-case letters, vectors with small-case boldface letters, and matrices with uppercase boldface letters. Calligraphic notation (e.g. \mathcal{X}) is reserved for vectors in the frequency domain. Given a sequence of vectors $\mathbf{h}_{r_x}^{t_x}$ for $r_x = 1 \cdots R_x$ and $t_x = 1 \cdots T_x$, we define the following stack variables

$$\mathbf{h}_{r_x} = \begin{bmatrix} \mathbf{h}_{r_x}^1 \\ \vdots \\ \mathbf{h}_{r_x}^{T_x} \end{bmatrix} \quad \text{and} \quad \mathbf{h} = \begin{bmatrix} \mathbf{h}_1 \\ \vdots \\ \mathbf{h}_{R_x} \end{bmatrix} \quad (1)$$

II. CHANNEL MODEL

The I/O equation for a MIMO system is given by

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

where $\mathbf{H}(p)$ is the $R_x \times T_x$ MIMO impulse response at tap p . The tap $\mathbf{H}(p)$ incorporates the effect of the transmit and receive correlation

$$\mathbf{H}(p) = \mathbf{R}^{1/2}(p) \mathbf{W}(p) \mathbf{T}^{1/2}(p) \quad (2)$$

where $\mathbf{T}(p)$ ($\mathbf{R}(p)$) is the transmit (receive) correlation matrix at tap p and is of size T_x (R_x). The matrix $\mathbf{W}(p)$ consists of iid elements and varies from a ST block to the next according to

$$\mathbf{W}_{t+1}(p) = \alpha(p) \mathbf{W}_t(p) + \sqrt{(1 - \alpha^2(p))e^{-\beta p}} \mathbf{U}_t(p) \quad (3)$$

Here, $\mathbf{U}_t(p)$ is an iid matrix with $\mathcal{N}(0, 1)$ entries and $\alpha(p) = J_0(2\pi f_D T(p))$, where T is the time duration of a ST block and $f_D(p)$ the Doppler frequency. The variable β corresponds to the exponent of the channel decay profile while the factor $\sqrt{(1 - \alpha^2(p))e^{-\beta p}}$ ensures that each link maintains the decay profile ($e^{-\beta p}$) for all time.

This channel model pushes the time variation to the limit (as the channel can change arbitrarily from one ST block to the next) while avoiding intercarrier interference and ensuring the proper operation of STC. This model was adopted in [10] and [11] in a SISO context. Here, we scale it up to the MIMO case and show how to incorporate the effect of spatial correlation. We can use (2) and (3) to obtain a dynamical equation for $\mathbf{h}_{r_x}^{t_x}$, the IR between antenna t_x and antenna r_x

$$\mathbf{h}_{r_x}^{t_x(+)} = \mathbf{F} \mathbf{h}_{r_x}^{t_x} + \mathbf{G} \mathbf{u}_{r_x}^{t_x} \quad (4)$$

where \mathbf{F} and \mathbf{G} are diagonal matrices with $f_{ii} = \alpha(i-1)$ and $g_{ii} = \sqrt{1 - \alpha^2(i-1)}$. Subsequently, we can obtain a dynamical relation for the stack variables \mathbf{h}_{r_x} and \mathbf{h} (see the stacking convention (1))

$$\mathbf{h}_{t+1} = (\mathbf{I}_{T_x R_x} \otimes \mathbf{F}) \mathbf{h}_t + (\mathbf{I}_{T_x R_x} \otimes \mathbf{G}) \mathbf{u}_t \quad (5)$$

where \mathbf{h}_{t+1} , \mathbf{h}_t , and \mathbf{u}_t , are vectors of size $T_x R_x (P+1) \times 1$. We also need to characterize the covariance of \mathbf{h}_0 and \mathbf{u}_t which we can show to be [12]

$$E[\mathbf{h}_0 \mathbf{h}_0^*] = \mathbf{I}_{T_x R_x} \otimes \mathbf{G} \mathbf{G}^* \quad (6)$$

$$E[\mathbf{u}_t \mathbf{u}_t^*] = \sum_{p=0}^P \mathbf{I}_{R_x} \otimes \mathbf{T}(p) \otimes (\mathbf{I}^P \mathbf{B} \bar{\mathbf{T}}^P) \quad (7)$$

where

$$\mathbf{B} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}, \quad \mathbf{I} = \begin{bmatrix} 0 & & & & & \\ 1 & 0 & & & & \\ & 1 & \ddots & & & \\ & & \ddots & \ddots & & \\ & & & & 0 & \\ & & & & 1 & 0 \end{bmatrix}$$

$$\bar{\mathbf{I}} = \begin{bmatrix} 0 & 1 & & & & \\ & 0 & \ddots & & & \\ & & \ddots & \ddots & & \\ & & & & 1 & \\ & & & & 0 & 1 \\ & & & & & 0 \end{bmatrix}$$

A. Receiver

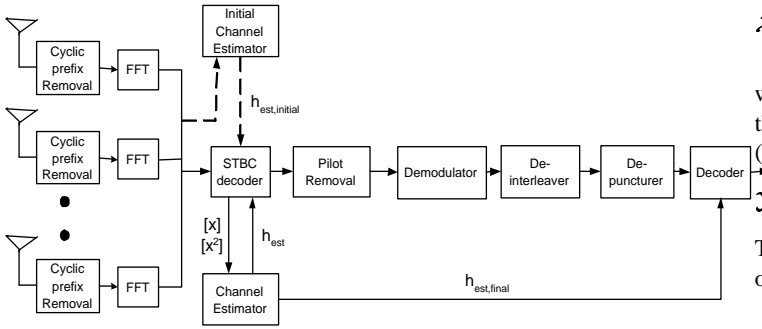


Fig. 1. OSTBC OFDM receiver

A block diagram of the receiver is shown in Figure 1. The receiver's core operation is based on the EM algorithm which performs joint channel and data recovery:

1) *STBC Decoder/Data Detector (Estimation Step)*: The STBC decoder/data detector calculates the conditional first and second moments of the transmitted data (soft estimate) to be used by the channel estimator.

2) *Channel Estimator (Maximization Step)*: Pilots are used to initialize channel estimation. The channel estimator then uses the soft data estimates together with the data and channel constraints to improve the channel estimate. These two processes (channel estimation and data detection) go on iteratively until a stopping criterion is satisfied.

III. I/O EQUATIONS FOR MIMO-OFDM

We derive two forms of the (I/O) equations: one that lends itself to *channel estimation* and a dual version that lends itself to *data detection*. To this end, Let \mathcal{X}_{t_x} be the OFDM symbol transmitted through antenna t_x which first undergoes an IDFT $\mathbf{x}_{t_x} = 1/N\mathbf{Q}\mathcal{X}_{t_x}$ where \mathbf{Q} is the $N \times N$ IDFT matrix. The system then appends a cyclic prefix before transmission. At the receiver end, the receiver strips the cyclic prefix to obtain the time domain symbol $\mathbf{y}_{r_x}^{t_x}$. The I/O equation of the OFDM system between transmit antenna t_x and receive antenna r_x is best described in the frequency domain

$$\mathbf{y}_{r_x}^{t_x} = \text{diag}(\mathcal{X}_{t_x}) \mathbf{Q}_{P+1}^* \mathbf{h}_{r_x}^{t_x} + \mathcal{N}_{r_x} \quad (8)$$

where $\mathbf{y}_{r_x}^{t_x}$, \mathcal{X}_{t_x} , $\mathcal{H}_{r_x}^{t_x}$, and $\mathcal{N}_{r_x}^{t_x}$ are the (length- N) DFT's of $\mathbf{y}_{r_x}^{t_x}$, \mathbf{x}_{t_x} , $\mathbf{h}_{r_x}^{t_x}$, \mathbf{n}_{r_x} , respectively, and where (8) follows from the fact that

$$\mathcal{H}_{r_x}^{t_x} = \mathbf{Q}^* \begin{bmatrix} \mathbf{h}_{r_x}^{t_x} \\ \mathbf{0}_{(N-P-1) \times 1} \end{bmatrix} = \mathbf{Q}_{P+1}^* \mathbf{h}_{r_x}^{t_x} \quad (9)$$

Here \mathbf{Q}_{P+1} represents the first $P+1$ rows of \mathbf{Q} . By superposition and using the stacking notation (1), we can express the I/O equation at receive antenna r_x as

$$\mathbf{y}_{r_x} = [\text{diag}(\mathcal{X}_1) \cdots \text{diag}(\mathcal{X}_{T_x})] (\mathbf{I}_{T_x} \otimes \mathbf{Q}_{P+1}^* \mathbf{S}) \mathbf{h}_{r_x} + \mathcal{N}_{r_x} \quad (10)$$

A. I/O Equations with Space Time Coding: Channel Estimation Version

Let $\{\mathcal{S}(1), \dots, \mathcal{S}(N_u)\}$ denote the set of uncoded OFDM symbols which we transmit over T_x antennas and N_c time slots. Following [13], we can perform ST coding using the set of $T_x \times N_c$ matrices $\{\mathbf{A}(1), \mathbf{B}(1), \dots, \mathbf{A}(N_u), \mathbf{B}(N_u)\}$ which characterizes the ST code. We can now show that the OFDM symbol transmitted from antenna t_x at time n_c is given by

$$\mathcal{X}_{t_x}(n_c) = \sum_{n_u=1}^{N_u} a_{t_x, n_c}(n_u) \text{Re} \mathcal{S}(n_u) + j b_{t_x, n_c}(n_u) \text{Im} \mathcal{S}(n_u) \quad (11)$$

where $a_{t_x, n_c}(n_u)$ is the (t_x, n_c) element of $\mathbf{A}(n_u)$ and $b_{t_x, n_c}(n_u)$ is the (t_x, n_c) element of $\mathbf{B}(n_u)$. Thus, in the presence of ST coding, (10) reads

$$\mathbf{y}_{r_x}(n_c) = [\text{diag}(\mathcal{X}_1(n_c)) \cdots \text{diag}(\mathcal{X}_{T_x}(n_c))] (\mathbf{I}_{T_x} \otimes \mathbf{Q}_{P+1}^*) \mathbf{h}_{r_x} + \mathcal{N}_{r_x}(n_c)$$

This represents the I/O equation at antenna r_x at OFDM symbol n_c of a ST block. Collecting this equation for all such symbols yields

$$\mathbf{y}_{r_x} = \mathbf{X} \mathbf{h}_{r_x} + \mathcal{N}_{r_x} \quad (12)$$

where

$$\mathbf{y}_{r_x} = \begin{bmatrix} \mathbf{y}_{r_x}(1) \\ \vdots \\ \mathbf{y}_{r_x}(N_c) \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} \text{diag}(\mathcal{X}_1(1)) \cdots \text{diag}(\mathcal{X}_{T_x}(1)) \\ \text{diag}(\mathcal{X}_1(2)) \cdots \text{diag}(\mathcal{X}_{T_x}(2)) \\ \vdots \\ \text{diag}(\mathcal{X}_1(N_c)) \cdots \text{diag}(\mathcal{X}_{T_x}(N_c)) \end{bmatrix} \quad (13)$$

Now, by further collecting this relationship over all receive antennas, we obtain

$$\mathbf{y}_t = (\mathbf{I}_{R_x} \otimes \mathbf{X}_t) \mathbf{h}_t + \mathcal{N}_t \quad (14)$$

This equation captures the I/O relationship at *all* frequency bins, for *all* input and output antennas, and for *all* OFDM symbols of the t th ST block. To perform initial channel estimation, we select those equations where the pilots are present. Let I_p denote the index set of the pilots bins. Then, the pilot/output equation takes the form

$$\mathbf{y}_{t_{I_p}} = (\mathbf{I}_{R_x} \otimes \mathbf{X}_{t_{I_p}}) \mathbf{h}_t + \mathcal{N}_{t_{I_p}} \quad (15)$$

B. I/O Equations with Space Time Coding: Data Detection Version

Signal detection in ST-coded OFDM is done on a tone-by-tone basis (i.e., as in SISO OFDM), except that the tones are collected for the whole ST block (i.e., for R_x receive antennas and over N_c time slots). From (8), we can construct the following I/O equation at *any tone* n belonging to the OFDM symbol n_c

$$\mathbf{y}_{r_x}(n_c) = [\mathcal{H}_{r_x}^1 \cdots \mathcal{H}_{r_x}^{T_x}] \begin{bmatrix} \mathcal{X}_1(n_c) \\ \vdots \\ \mathcal{X}_{T_x}(n_c) \end{bmatrix} + \mathcal{N}_{r_x}(n_c) \quad (16)$$

We suppress the dependence on n for notational convenience. Collecting this relationship for all receive antennas yields

$$\begin{bmatrix} \mathbf{y}_1(n_c) \\ \vdots \\ \mathbf{y}_{R_x}(n_c) \end{bmatrix} = \begin{bmatrix} \kappa_1^1 & \cdots & \kappa_1^{T_x} \\ \vdots & \ddots & \vdots \\ \kappa_{R_x}^1 & \cdots & \kappa_{R_x}^{T_x} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1(n_c) \\ \vdots \\ \mathbf{x}_{T_x}(n_c) \end{bmatrix} + \begin{bmatrix} \mathcal{N}_1(n_c) \\ \vdots \\ \mathcal{N}_{R_x}(n_c) \end{bmatrix}$$

Or, more succinctly,

$$\mathbf{y}(n_c) = \mathbf{H}\mathbf{x}(n_c) + \mathcal{N}(n_c)$$

By further concatenating this relationship for $n_c = 1, \dots, N_c$, we can show that the following relationship holds (see [13])

$$\mathbf{y} = \mathbf{C} \begin{bmatrix} \text{Re } \mathbf{S} \\ \text{Im } \mathbf{S} \end{bmatrix} + \mathcal{N} \quad (17)$$

where

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}(1) \\ \vdots \\ \mathbf{y}(N_c) \end{bmatrix}, \quad \mathbf{S} = \begin{bmatrix} \mathbf{S}(1) \\ \vdots \\ \mathbf{S}(N_u) \end{bmatrix}, \quad \text{and } \mathbf{C} = [\mathbf{C}_a \quad \mathbf{C}_b]$$

with $\mathbf{C}_a = [\text{vec}(\mathbf{H}\mathbf{A}(1)) \quad \cdots \quad \text{vec}(\mathbf{H}\mathbf{A}(N_u))]$ and $\mathbf{C}_b [\text{vec}(\mathbf{H}\mathbf{B}(1)) \quad \cdots \quad \text{vec}(\mathbf{H}\mathbf{B}(N_u))]$. We finally note that the STBC code is orthogonal if and only if the matrix \mathbf{C} satisfies [13]

$$\text{Re}[\mathbf{C}^* \mathbf{C}] = \|\mathbf{H}\|^2 \mathbf{I}_{2N_u} \quad \forall \mathbf{H} \quad (18)$$

This property is essential to perform data detection. We stress that the relationships (16) through (18) apply at a particular tone n and that this dependence has been omitted for notational convenience.

IV. THE EM ALGORITHM FOR JOINT CHANNEL AND DATA ESTIMATION

A. The EM-Based Forward-Backward Kalman

Consider the OFDM system of this paper, essentially described by

$$\mathbf{h}_{t+1} = (\mathbf{I}_{T_x R_x} \otimes \mathbf{F})\mathbf{h}_t + (\mathbf{I}_{T_x R_x} \otimes \mathbf{G})\mathbf{u}_t \quad (19)$$

$$\mathbf{y}_t = (\mathbf{I}_{R_x} \otimes \mathbf{X}_t)\mathbf{h}_t + \mathcal{N}_t \quad (20)$$

with $\mathbf{h}_0 \sim \mathcal{N}(\mathbf{0}, \mathbf{\Pi})$ and $\mathbf{u}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_u)$. Given a sequence of $T+1$ input and output symbols \mathbf{X}_0^T and \mathbf{y}_0^T ,¹ we obtain the MAP estimate of the sequence \mathbf{h}_0^T by the maximizing the likelihood

$$\mathcal{L} = \ln p(\mathbf{y}_0^T | \mathbf{X}_0^T, \mathbf{h}_0^T) + \ln p(\mathbf{h}_0^T) \quad (21)$$

Using (20), we can express the first term of the log-likelihood as

$$\begin{aligned} \ln p(\mathbf{y}_0^T | \mathbf{X}_0^T, \mathbf{h}_0^T) &= \sum_{t=0}^T \ln p(\mathbf{y}_t | \mathbf{X}_t, \mathbf{h}_t) \\ &= - \sum_{t=0}^T \|\mathbf{y}_t - (\mathbf{I}_{R_x} \otimes \mathbf{X}_t)\mathbf{h}_t\|_{\frac{1}{2\sigma_n^2}}^2 \end{aligned}$$

Using (19), we can express the second term of (21) as

$$\begin{aligned} \ln p(\mathbf{h}_0^T) &= \sum_{t=1}^T \ln p(\mathbf{h}_t | \mathbf{h}_{t-1}) + \ln p(\mathbf{h}_0) \\ &= - \sum_{t=1}^T \|\mathbf{h}_t - (\mathbf{I}_{T_x R_x} \otimes \mathbf{F})\mathbf{h}_{t-1}\|_{(G\mathbf{R}_u G^*)^{-1}}^2 - \|\mathbf{h}_0\|_{\mathbf{\Pi}_0^{-1}}^2 \end{aligned}$$

¹We use \mathbf{X}_0^T to denote the sequence $\mathbf{X}_0, \mathbf{X}_1, \dots, \mathbf{X}_T$.

Combining these two expressions yields

$$\begin{aligned} \mathcal{L} &= - \sum_{t=1}^T \|\mathbf{y}_t - (\mathbf{I}_{R_x} \otimes \mathbf{X}_t)\mathbf{h}_t\|_{\frac{1}{2\sigma_n^2}}^2 \\ &\quad - \sum_{t=1}^T \|\mathbf{h}_t - (\mathbf{I}_{T_x R_x} \otimes \mathbf{F})\mathbf{h}_{t-1}\|_{(G\mathbf{R}_u G^*)^{-1}}^2 - \|\mathbf{h}_0\|_{\mathbf{\Pi}_0^{-1}}^2 \end{aligned} \quad (22)$$

Since the channel sequence \mathbf{h}_0^T is jointly Gaussian, the MAP estimate of the channel sequence given the input and output sequences \mathbf{X}_0^T and \mathbf{y}_0^T is the same as the MMSE estimate given the same sequences. The MMSE estimate itself is obtained by the forward-backward (FB) Kalman filter. This allows us to state the following theorem.

Theorem 1: Channel estimation–Known input case Consider the state-space model (19)–(20). Given the input and output sequences \mathbf{X}_0^T and \mathbf{y}_0^T , the MAP (or equivalently MMSE) estimate of \mathbf{h}_0^T is obtained by applying the following (forward-backward Kalman) filter to the state-space model (19)–(20)

Forward run: Start with $\mathbf{P}_{0|-1} = \mathbf{\Pi}_0$ and for $i = 1, \dots, T$, calculate

$$\mathbf{R}_{e,t} = \sigma_n^2 \mathbf{I}_{T_x R_x N} + (\mathbf{I}_{R_x} \otimes \mathbf{X}_t) \mathbf{P}_{t|t-1} (\mathbf{I}_{R_x} \otimes \mathbf{X}_t^*) \quad (23)$$

$$\mathbf{K}_t = \mathbf{P}_{t|t-1} (\mathbf{I}_{R_x} \otimes \mathbf{X}_t^*) \mathbf{R}_{e,t}^{-1} \quad (24)$$

$$\hat{\mathbf{h}}_{t|t} = (\mathbf{I}_{T_x R_x (P+1)} - \mathbf{K}_t (\mathbf{I}_{R_x} \otimes \mathbf{X}_t)) \hat{\mathbf{h}}_{t|t-1} + \mathbf{K}_t \mathbf{y}_t \quad (25)$$

$$\hat{\mathbf{h}}_{t+1|t} = (\mathbf{I}_{T_x R_x} \otimes \mathbf{F}) \hat{\mathbf{h}}_{t|t}, \quad \mathbf{h}_{0|-1} = \mathbf{0} \quad (26)$$

$$\begin{aligned} \mathbf{P}_{t+1|t} &= (\mathbf{I}_{T_x R_x} \otimes \mathbf{F}) (\mathbf{P}_{t|t-1} - \mathbf{K}_t \mathbf{R}_{e,t} \mathbf{K}_t^*) (\mathbf{I}_{T_x R_x} \otimes \mathbf{F}^*) \\ &\quad + \mathbf{G} \mathbf{R}_u \mathbf{G}^* \end{aligned} \quad (27)$$

Backward run: Starting from $\lambda_{T+1|T} = \mathbf{0}$ and for $t = T, T-1, \dots, 0$, calculate

$$\lambda_{t|T} = (\mathbf{I}_{P+N} - (\mathbf{I}_{R_x} \otimes \mathbf{X}_t^*) \mathbf{K}_t^*) (\mathbf{I} \otimes \mathbf{F}^*) \lambda_{t+1|T} \quad (28)$$

$$+ (\mathbf{I} \otimes \mathbf{X}_t) \mathbf{R}_{e,t}^{-1} (\mathbf{y}_t - (\mathbf{I} \otimes \mathbf{X}_t) \hat{\mathbf{h}}_{t|t-1})$$

$$\hat{\mathbf{h}}_{t|T} = \hat{\mathbf{h}}_{t|t-1} + \mathbf{P}_{t|t-1} \lambda_{t|T} \quad (29)$$

The desired estimate is $\hat{\mathbf{h}}_{t|T}$. For a proof, see problem 10.9 in [14]. This theorem allows us to obtain the estimate of \mathbf{h}_0^T when the input sequence \mathbf{X}_0^T is not available. For in this case, we maximize the likelihood (22) averaged over the sequence \mathbf{X}_0^T . Thus, the j -th iteration of the EM algorithm is now obtained by maximizing the averaged log-likelihood $\bar{\mathcal{L}} = E_{\mathbf{X}_0^T | \mathbf{y}_0^T, \mathbf{h}_0^{T(j-1)}} [\mathcal{L}]$. By inspecting (22), we note that the only term that is modified under expectation is the first summand, and its expectation is given by

$$\left\| \begin{bmatrix} \mathbf{y}_t \\ \mathbf{0}_{T_x R_x (P+1) \times 1} \end{bmatrix} - \begin{bmatrix} \mathbf{I}_{R_x} \otimes E[\mathbf{X}_t] \\ \mathbf{I}_{R_x} \otimes \text{Cov}[\mathbf{X}_t^*]^{1/2} \end{bmatrix} \mathbf{h}_t \right\|_{\frac{1}{2\sigma_n^2}}^2$$

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

$$\begin{aligned} \bar{\mathcal{L}} &= - \sum_{t=0}^T \left\| \begin{bmatrix} \mathbf{y}_t \\ \mathbf{0}_{T_x R_x (P+1) \times 1} \end{bmatrix} - \begin{bmatrix} \mathbf{I}_{R_x} \otimes E[\mathbf{X}_t] \\ \mathbf{I}_{R_x} \otimes \text{Cov}[\mathbf{X}_t^*]^{1/2} \end{bmatrix} \mathbf{h}_t \right\|_{\frac{1}{2\sigma_n^2}}^2 - \\ &\quad \sum_{t=1}^T \|\mathbf{h}_t - \mathbf{F}\mathbf{h}_{t-1}\|_{(G\mathbf{R}_u G^*)^{-1}}^2 - \|\mathbf{h}_0\|_{\mathbf{\Pi}_0^{-1}}^2 \end{aligned}$$

Note that we can obtain the averaged likelihood (30) from the original likelihood (22) by performing the substitution

$$\mathbf{I}_{R_x} \otimes \mathbf{X}_t \longrightarrow \begin{bmatrix} \mathbf{I}_{R_x} \otimes E[\mathbf{X}_t] \\ \mathbf{I}_{R_x} \otimes \text{Cov}[\mathbf{X}_t^*]^{1/2} \end{bmatrix}, \quad \mathbf{y}_t \longrightarrow \begin{bmatrix} \mathbf{y}_t \\ \mathbf{0}_{T_x R_x (P+1) \times 1} \end{bmatrix} \quad (30)$$

We can thus state the following theorem

Theorem 2: Channel estimation–Unknown input case Consider the state-space model (19)–(20) and assume that the receiver does not have access to the transmitted data \mathbf{X}_0^T . The channel estimate at the j th iteration $\mathbf{h}_0^{T(j)}$ of the EM algorithm is obtained by applying the forward-backward Kalman (23)–(29) to the following state-space model

$$\mathbf{h}_{t+1} = (\mathbf{I}_{T_x R_x} \otimes \mathbf{F}) \mathbf{h}_t + (\mathbf{I}_{T_x R_x} \otimes \mathbf{G}) \mathbf{u}_t \quad (31)$$

$$\begin{bmatrix} \mathbf{y}_t \\ \mathbf{0}_{T_x R_x (P+1) \times 1} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{R_x} \otimes E[\mathbf{X}_t] \\ \mathbf{I}_{R_x} \otimes \text{Cov}[\mathbf{X}_t^*]^{1/2} \end{bmatrix} \mathbf{h}_t + \begin{bmatrix} \mathcal{N}_t \\ \mathbf{n}_t \end{bmatrix} \quad (32)$$

where \mathbf{n}_t is virtual noise that is independent of \mathcal{N}_t .

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.

B. Initial Channel Estimation

We obtain the initial channel estimate from the pilot/output equation (15) together with the dynamical channel model (5). Specifically, we do this by applying the FB Kalman to the following state-space model

$$\mathbf{h}_{t+1} = (\mathbf{I}_{T_x R_x} \otimes \mathbf{F}) \mathbf{h}_t + (\mathbf{I}_{T_x R_x} \otimes \mathbf{G}) \mathbf{u}_t \quad (33)$$

$$\mathbf{y}_{t_{I_p}} = (\mathbf{I}_{R_x} \otimes \mathbf{X}_{t_{I_p}}) \mathbf{h}_t + \mathcal{N}_{t_{I_p}} \quad (34)$$

i.e., by applying The FB Kalman (23)–(29) with the substitution

$$\mathbf{y}_t \longrightarrow \mathbf{y}_{t_{I_p}}, \quad \mathbf{X}_t \longrightarrow \mathbf{X}_{t_{I_p}}, \quad \text{and} \quad \mathbf{I}_{T_x R_x N} \longrightarrow \mathbf{I}_{T_x R_x |I_p|} \quad (35)$$

C. Data Detection

To detect the data, we use the data detection version of the I/O equation (17). Upon multiplying both sides by \mathbf{C}^* and taking the real part, we obtain

$$\tilde{\mathbf{y}} = \|\mathcal{H}\|^2 \begin{bmatrix} \text{Re } \mathcal{S} \\ \text{Im } \mathcal{S} \end{bmatrix} + \tilde{\mathcal{N}} \quad (36)$$

where $\tilde{\mathbf{y}}$ and $\tilde{\mathcal{N}}$ are $2N_u \times 1$ vectors defined by

$$\tilde{\mathbf{y}} = \text{Re } \mathbf{C}^* \mathbf{y} \quad \text{and} \quad \tilde{\mathcal{N}} = \text{Re } \mathbf{C}^* \mathcal{N}$$

Since \mathbf{C} is orthogonal, the noise $\tilde{\mathcal{N}}$ remains white, and the input can be detected on an element-by-element basis. We will now demonstrate how to detect the elements of $\text{Re } \mathcal{S}$ (the imaginary part can be treated similarly). So let $\mathcal{R} = \{r_1, \dots, r_{|\mathcal{R}|}\}$ denote the alphabet set from which the elements of $\text{Re } \mathcal{S}$ take their values. We can show that the conditional pdf $f(\text{Re } \mathcal{S}(n_u) | \tilde{\mathbf{y}}(n_u))$ is given by

$$f(r_i | \tilde{\mathbf{y}}(n_u)) = \frac{e^{-\frac{|\tilde{\mathbf{y}}(n_u) - \|\mathcal{H}\|^2 r_i|^2}{2\sigma_n^2}}}{\sum_{i=1}^{|\mathcal{R}|} e^{-\frac{|\tilde{\mathbf{y}}(n_u) - \|\mathcal{H}\|^2 r_i|^2}{2\sigma_n^2}}} \quad (37)$$

We can use this pdf to two moments of $\text{Re } \mathcal{S}(n_u)$

$$E[\text{Re } \mathcal{S}(n_u) | \tilde{\mathbf{y}}(n_u)] = \frac{\sum_{i=1}^{|\mathcal{R}|} r_i e^{-\frac{|\tilde{\mathbf{y}}(n_u) - \|\mathcal{H}\|^2 r_i|^2}{2\sigma_n^2}}}{\sum_{i=1}^{|\mathcal{R}|} e^{-\frac{|\tilde{\mathbf{y}}(n_u) - \|\mathcal{H}\|^2 r_i|^2}{2\sigma_n^2}}} \quad (38)$$

$$E[|\text{Re } \mathcal{S}(n_u)|^2 | \tilde{\mathbf{y}}(n_u)] = \frac{\sum_{i=1}^{|\mathcal{R}|} r_i^2 e^{-\frac{|\tilde{\mathbf{y}}(n_u) - \|\mathcal{H}\|^2 r_i|^2}{2\sigma_n^2}}}{\sum_{i=1}^{|\mathcal{R}|} e^{-\frac{|\tilde{\mathbf{y}}(n_u) - \|\mathcal{H}\|^2 r_i|^2}{2\sigma_n^2}}} \quad (39)$$

We can similarly calculate the two moments of the imaginary part. Now equations (38)–(39), just like (16)–(18), apply at a certain frequency tone n . So collecting (38) and (39)

for all tones ($n = 1, \dots, N$) produces the two moments of the uncoded OFDM symbols. Specifically, we can calculate $E[\text{Re } \mathcal{S}(n_u)]$, $E[\text{Im } \mathcal{S}(n_u)]$, $E[\text{diag}(\text{Re } \mathcal{S}(n_u))^2]$, and $E[\text{diag}(\text{Im } \mathcal{S}(n_u))^2]$.

D. Summary of the EM-Based FB Kalman

Given a sequence of input and output symbols \mathbf{X}_0^T and \mathbf{y}_0^T perform the following operations:

- 1) calculate the initial channel estimate $\mathbf{h}_0^{T(0)}$ by applying the FB Kalman filter to the state-space model (33)–(34), i.e. by applying (23)–(29) the substitutions (35).
- 2) Iterate between the expectation and maximization steps for $j = 1, \dots, N_{\text{iter}}$:
 - a) **Expectation:** Compute the first two moments of the uncoded OFDM symbols $\mathcal{S}(1), \dots, \mathcal{S}(n_u)$, given the output \mathbf{y}_0^T and the most recent estimate of the channel, $\mathbf{h}_0^{T(j-1)}$. Use these moments to calculate the moments of \mathbf{X} through the relationships (11) and (13).
 - b) **Maximization:** Obtain the channel estimate $\mathbf{h}_0^{T(j)}$ by employing the FB Kalman to the state-space model (31)–(34), i.e. by applying (23)–(29) the substitutions (30).

E. A Simplification: 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 (23)–(27)) for both the initial estimation and the EM iteration. The algorithm then collapses to the Kalman-based filter proposed in [15] where the data and channel are recovered within one ST symbol.

V. SIMULATION RESULTS

The two Kalman filters (forward only and forward-backward) derived in this paper apply to OFDM transmission over both SISO and MIMO transmission. We demonstrate the behavior of the forward only Kalman for MIMO-OFDM transmission and bench mark the FB Kalman with SISO-OFDM (see [9]). The conclusions that we draw from our simulations apply equally well to both algorithms.

A. MIMO-OFDM: Testing the Forward-Only Kalman

Throughout the simulation, we use a rate 1/2 convolutional encoder. The coded bits are mapped to 16-QAM symbols. We use the Alamouti STC code for which $N_s = 2$ and $T_b = 2$. We use the channel model (5) with parameters, $\alpha = 0.985$, $\beta = 0.2$, and $P = 7$ and set the number of antennas to $T_x = 2$ $R_x = 1$ or $R_x = 2$. Three thousand packets were simulated per SNR value. Each packet is comprised of 12 OFDM symbols transmitted over 6 ST blocks. Each OFDM symbol consists of 64 frequency tones and a cyclic prefix of length 16. We employ 16 pilots in the OFDM symbols making up the first ST block, while the number of pilots in subsequent symbols vary between 2, 6, and 10. In the following, we discuss the effect of various parameters on the BER performance of the receiver design.

1) **Bench Marking:** We compare our algorithm with an EM-based iterative MMSE receiver such as the one proposed in [7] and [16]. In contrast to our work, the authors in [7] and [16] take a data-centric approach, treating the transmitted signal as the desired parameter and the channel as the unobserved data. This algorithm further confines its pilots to the first ST block. The pilots are used to produce an initial channel estimate for the first ST block. This estimate is in turn used to predict the initial channel estimate for the subsequent

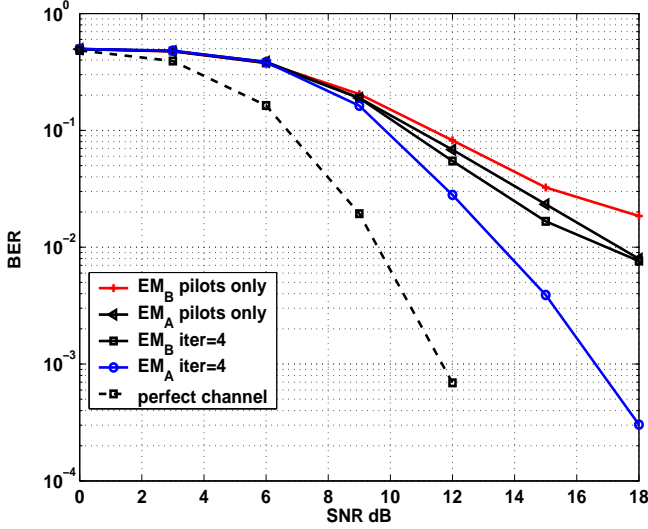


Fig. 2. Receiver design comparison

ST blocks by employing a time correlation filter [7]. These initial estimates are used to kick-start the EM algorithm.

In this algorithm, the E-step is calculated by a conditional expectation of the channel given the received symbol and the current estimate of the transmitted data (i.e., through MMSE estimation). The maximization step is simply the hard decision, i.e. the ML estimate of the transmitted data.

In Figure 2, we compare both schemes with 16 pilots in the initial ST block and zero pilots in the subsequent blocks. EM_A refers to the iterative MMSE scheme while EM_B refers to the Kalman filter based scheme proposed in this chapter. Our algorithm (EM_B) outperforms EM_A of [7] in both pilot scenarios. One reason for this performance improvement is that our algorithm incorporates the time correlation information and the most recent channel estimate in every iteration of the EM algorithm.

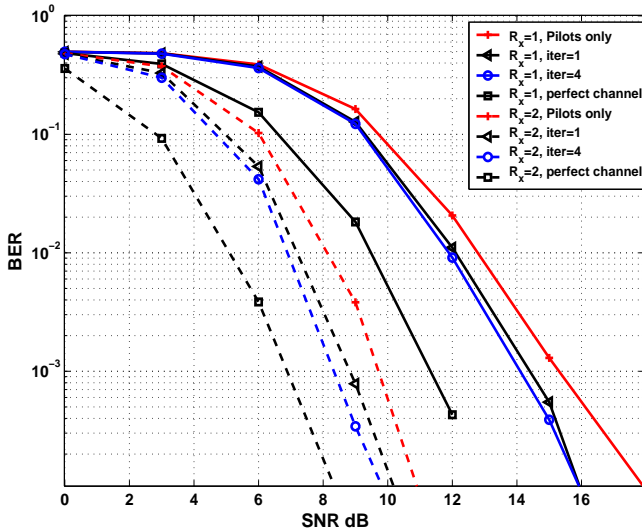


Fig. 3. BER performance with iterations and spatial diversity

2) *Effect of Number of Iterations and Spatial Diversity:* In this subsection, we test the sensitivity of our algorithm to the number of EM iterations used. We demonstrate this for one and two receive antennas. Here we employ 6 pilots per OFDM symbol (in addition to the 16 pilots per symbol employed in the first ST block). From Figure 3, we see that the first iteration yields substantial improvement over the pilot-based estimation. Iterating beyond that yields diminishing returns. The results apply to $R_x = 1, 2$.

3) *Sensitivity to Time Variation:* In this subsection, we test the performance of our receiver against different degrees of time variation. This is parameterized by α ($0 \leq \alpha \leq 1$) with lower values of α indicating a more time-variant channel. In Figure 4 we show the BER curves for a system that employs 10-pilots per OFDM symbol.

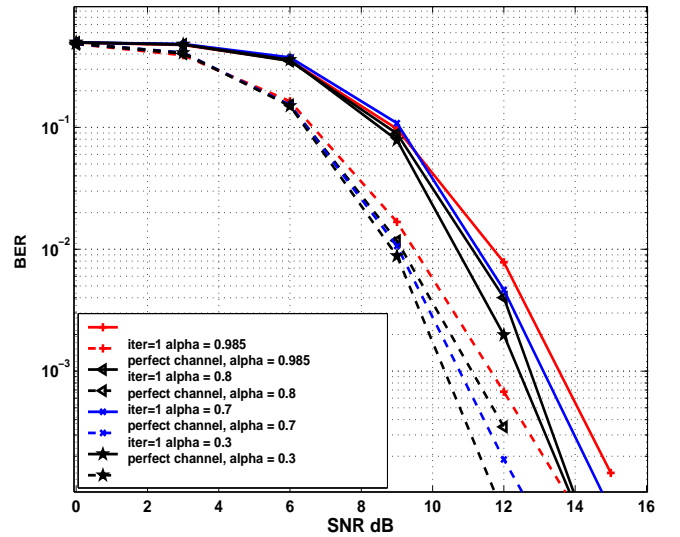


Fig. 4. BER performance with varying channel correlation with 10 pilots

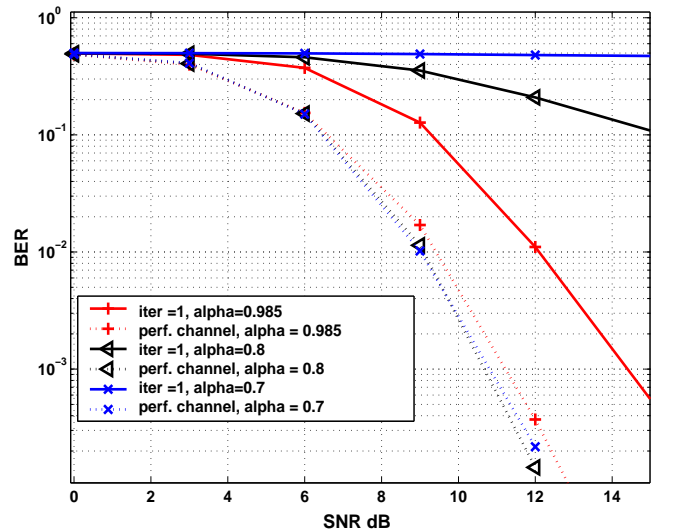


Fig. 5. BER performance with varying channel correlation with 6 pilots

From this figure, we observe that as α decreases (indicating more channel variation), the BER improves. This comes from increasing

time diversity in the channel. Therefore, with enough number of pilots, we are able to track the channel and capture time diversity.

For comparison, in Figure 5, we show the BER curves for a system with fewer pilots (6-pilots per OFDM symbol) for $\alpha = 0.7, 0.8$ and 0.985 . We observe an error-floor as the channel variation increases. So, in this case, we are unable to capture the time diversity. More pilots are thus needed to capture diversity and improve performance.

B. SISO-OFDM: Testing the FB Kalman

We use an input similar to the one employed in the MIMO case; a sequence of 5 OFDM symbols each with 64 carriers and CP of length 15. The OFDM symbols are constructed from an uncoded bit stream mapped to 16 QAM symbols through gray coding. We use a SISO channel model similar to the one employed for the MIMO simulation except that the number of transmit and receive antennas is set to 1 and the number of channel taps is about doubled to 16.

We bench mark the BER performance of the FB-Kalman against two receivers that have been suggested in literature 1) EM-based least-squares (LS) receiver (i.e. a receiver employing frequency correlation only), 2) the EM-based receiver proposed in [7] in addition to the forward-only Kalman that is suggested in this paper. These receivers employ the same number of pilots (16 in the first OFDM symbol and x number of them in the subsequent 4 (with $0 \leq x \leq 16$)).

Figure 6 demonstrates that the Kalman and FB-Kalman outperform the LS receiver and the receiver of [7], especially for low number of pilots. Moreover, for this case of moderate time variation, the FB-Kalman consistently outperforms the Kalman receiver.

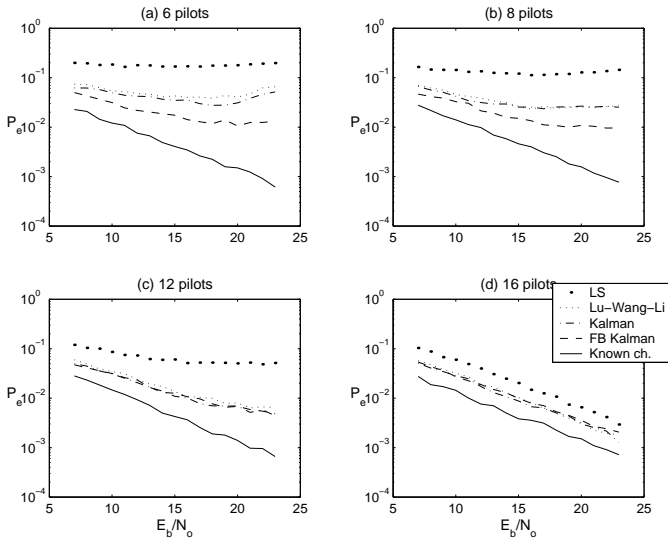


Fig. 6. Comparing the BER curves for various receivers.

VI. CONCLUSION

This paper proposed a receiver for MIMO-OFDM transmission over time-variant channels. While the paper assumed the channel to be constant within any ST block, it is allowed to vary from one block to the next. This makes the receiver suitable for operation in high-speed environments.

The receiver employs the EM algorithm to achieve channel and data recovery. Specifically, data recovery (or the E-step) is as simple as decoding a space-time block code. Channel recovery (or the M-step) is performed using a forward-backward Kalman filter. We also

suggested a relaxed (forward-only) version of the algorithm that is able to perform recovery with no latency and hence avoid the delay and storage shortcomings of the FB-Kalman.

When compared with other MIMO receivers, our receiver makes the most use of the underlying structure. Specifically, the algorithm makes use of the finite alphabet constraints (in (38)–(39)), the data in its soft form (in (31)–(32)), pilots (in (33)–(34)), finite-delay spread (in that channel estimation is done in the time domain), frequency- and time-correlation (in (5)), spatial correlation in (7), and space-time coding. It is also straightforward to incorporate the effect of an outer code [17], sparsity [18] and the CP [9]. Our simulations show the favorable behavior of the two Kalman filters as compared to other receivers.

Acknowledgement: This work was supported by King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia.

REFERENCES

- [1] 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.
- [2] B. Muquet, M. de Courville, P. Duhamel, and V. Buzenac, "A subspace based blind and semi-blind channel identification method for OFDM systems," in *Proc. IEEE Workshop on Signal Proc. Advances in Wireless Comm.*, 1999, pp. 170–173.
- [3] 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.
- [4] 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.
- [5] S. Zhou, B. Muquet, and G. B. Giannakis, "Subspace-based (semi-) blind channel estimation for block precoded space-time OFDM," *IEEE Trans. Signal Proc.*, vol. 50, no. 5, pp. 1215–1228, May 2002.
- [6] 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.*, Nov. 2002, vol. 2, pp. 13–17.
- [7] B. Lu, X. Wang, and Y. Li, "Iterative receivers for space-time block-coded OFDM systems in dispersive fading channels," *IEEE Trans. Wireless Commun.*, vol. 1, no. 2, pp. 213–225, Apr. 2002.
- [8] Y. Li, C. N. Georghiades, and H. Garnig, "Iterative maximum-likelihood sequence estimation for space-time coded systems," *IEEE Trans. Commun.*, vol. 49, no. 6, pp. 948–951, Jun. 2001.
- [9] T. Y. Al-Naffouri, "An EM-based forward-backward Kalman filter for the estimation of time-variant channels in OFDM," *IEEE Trans. Signal Proc.*, 2006.
- [10] 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.
- [11] 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.
- [12] O. Oteri T. Y. Al-Naffouri, O. Awoniyi and A. Paulraj, "Receiver design for MIMO OFDM transmission over time variant channels,"
- [13] E. Larsson and P. Stoica, *Space-Time Block Coding for Wireless Communications*, Cambridge University Press, 2003.
- [14] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear estimation*, Prentice Hall, 2000.
- [15] 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.
- [16] 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.
- [17] 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.
- [18] T. Y. Al-Naffouri, Adaptive algorithms for wireless channel estimation, Department of Electrical Engineering, Stanford University, Jan. 2005.