Receiver Design for MIMO-OFDM Transmission over Time Variant Channels

T. Y. Al-Naffouri, O. Awoniyi, O. Oteri, A. Paulraj
Information Systems Laboratory
Dept. of Electrical Engineering, Stanford University
Email: {nafouri,lawoniyi,koteri,apaulraj}@stanford.edu

***Abstract*— This paper considers receiver design for Space Time Block Coded MIMO OFDM transmission over frequency selective time-variant channels. The receiver employs the expectation-maximization (EM) algorithm for joint channel and data recovery. It makes collective use of the data and channel constraints that characterize the communication problem. The data constraints include pilots, the finite alphabet constraint, and space-time block coding. The channel constraints include the finite delay spread and frequency and time correlation. The receiver employs an EM-based Kalman filter for channel estimation. The receiver is able to recover the channel (which varies from one space-time block to the next) and the data with no latency and to reduce the number of pilots needed. Simulations show that the receiver outperforms other least-squares based iterative receivers.**

I. Introduction

Orthogonal Frequency Division Multiplexing (OFDM) coupled with Multiple-Input-Multiple-Output techniques is an emerging technique for reliable, high speed wireless communications over frequency selective channels [3] [4].

The OFDM receiver needs Channel State Information (CSI) to recover the data. With no receiver CSI, both the channel and the data are unknowns that have to be estimated and decoded. Estimation and decoding can be carried out jointly or separately. Techniques for channel estimation fall into 3 distinct classes; training/pilot-symbol based, semi-blind and blind methods. Pilot/training based methods estimate the channel from a known preamble or pilot sequence sent at the transmitter and use the estimated channel to decode the data [4]. Blind methods do not use any preambles/pilots but rely instead on *a priori* constraints to recover the channel and data [5]. Semi-blind methods make use of both pilots and additional channel/input data constraints to perform channel identification and data decoding. The schemes use pilots to obtain an initial channel estimate and improve the estimate by using a variety of *priori* information [10][12].

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 Expectation Maximization (EM) algorithm for joint channel and data recovery.

The EM algorithm is used in estimating a desired parameter when some of the data required for the estimation is unobserved. The algorithm first performs an initial estimate of the unobserved data and uses the information to compute the ML estimate of the desired parameter. This is the maximization or M-step. The algorithm then uses the parameter estimate to update (compute the conditional expectation) of the unobserved data. This is the expectation or E-step. The process alternates between the M- and E-steps till a convergence criterion is satisfied [13].

EM-algorithm receivers have been applied to receiver design in [10] and [11]. These works consider the channel as the unobserved data and the transmitted signal as the desired parameter. The M-step is a maximum likelihood hard decision of the transmitted signal based on the previously calculated channel estimate and the E-step is based on an MMSE estimate of the channel.

In contrast to the prior approach, we take a channel estimation centric viewpoint and reverse the roles of the channel and the transmitted signal as in [9] and [12]. In addition, we make a collective use of the constraints induced by the data and channel that underly the communication problem.

The change in EM order and the general framework facilitate the use of a Kalman filter for estimation of the unknown channel using these constraints. The use of the Kalman filter in channel estimation has been proposed in [14] for SISO systems and [15] for MIMO systems.

Our contributions are as follows

- Incorporation of the channel and data constraints mentioned above in a general framework suitable for channel estimation.
- Use of the Kalman filter for the maximization step in the EM algorithm for joint channel and data recovery.
- Channel estimation in time-correlated, time-varying channel within an OFDM packet with no additional latency.

The paper is organized as follows; We give an overview of the transceiver in section II. In Section III, the input/output equations for the system are derived. We then describe the channel model in detail in section II-B. Channel estimation using Kalman filtering is described in section IV and a summary of the steps in the algorithm is discussed in section V. Our simulations are presented in section VI and finally we conclude in section VII.

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. A hat over a variable indicates an estimate of the variable. We use $\boldsymbol{h}^-$ and $\boldsymbol{h}^+$ to refer to the past and future value of $\boldsymbol{h}$, respectively, and use $\boldsymbol{h}^{(0)}$ to denote the initial value of $\boldsymbol{h}$. Given a sequence of vectors $\boldsymbol{h}_{r_x}^{t_x}$ for $r_x = 1 \cdots R_x$ and $t_x = 1 \cdots T_x$, we define the following stack variables

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

II. System Overview

In this section, we give an overview of the communications system, the transmitter, channel and receiver.

A. Transmitter

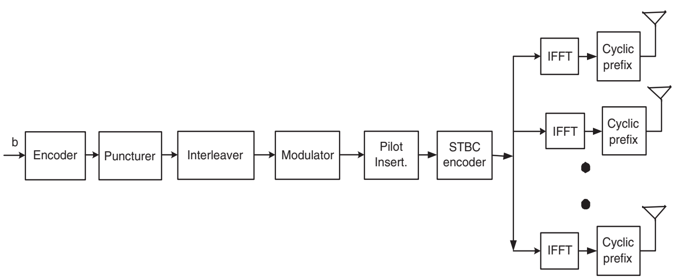

Fig. 1. Transmitter

A block diagram of the transmitter is shown in Figure 1. The bit sequence to be transmitted is passed to a convolutional encoder, punctured, interleaved then QAM-modulated. The QAM symbols are then mapped to the OFDM symbols with pilot insertion. The OSTBC encoder maps the OFDM symbols needed for one ST block to the various antennas based on the OSTBC used. The first ST block in the OFDM packet uses N_{p} pilots. Subsequent ST blocks can have no pilots (as is the case in [10] and [11]) or use a reduced number of them.

B. Channel Model

The input/output (I/O) relationship for a MIMO system with T_x transmit antennas and R_x receive antennas is give by

$$\boldsymbol{y}(m) \quad = \quad \sum_{p=0}^{P} \boldsymbol{H}(p)\boldsymbol{x}(m-p) + \boldsymbol{n}(m)$$

where $\boldsymbol{H}(p)$ is the $R_x \times T_x$ matrix representing the MIMO impulse response at tap p and m is the time. We will assume that $\boldsymbol{H}(p)$ is i.i.d. for all p and that it remains constant over a single STBC block. From one STBC block to the next, the MIMO taps change according to the dynamical equation

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

where $\alpha(p)$ is related to the Doppler frequency $f_D(p)$ by $\alpha(p) = J_0(2\pi f_D(p))$ and where $\boldsymbol{U}(p)$ is an i.i.d. matrix with entries that are $\mathcal{N}(0,1)$. The factor $\sqrt{(1-\alpha^2(p))e^{-\beta p}}$ ensures that each link maintains the same profile ($e^{-\beta p}$) for all time. This is similar to the MIMO model of [15]). Using this model, we can obtain the dynamical equation for the impulse response $\boldsymbol{h}_{r_x}^{t_x}$ between any pair of transmit and receive antennas

$$h_{r_x}^{t_x(+)}(p) = \alpha(p)h_{r_x}^{t_x}(p) + \sqrt{(1-\alpha^2(p))e^{-\beta p}}u_{r_x}^{t_x}(p) \tag{3}$$

By stacking (3) for $p = 0, 1, \ldots, P$, and further stacking the result over all transmit and receive antennas, we obtain

$$\boldsymbol{h}^{(+)} = (\boldsymbol{I}_{T_x R_x} \otimes \boldsymbol{F})\,\boldsymbol{h} + (\boldsymbol{I}_{T_x R_x} \otimes \boldsymbol{G})\,\boldsymbol{u} \tag{4}$$

where

$$\begin{aligned} \boldsymbol{F} &= \operatorname{diag}\left(\alpha(0), \quad \ldots, \quad \alpha(P) \right) \\ \boldsymbol{G} &= \left(\sqrt{1-\alpha^2(0)}, \quad \ldots, \quad \sqrt{(1-\alpha^2(P))e^{-\beta P}} \right) \end{aligned}$$

Here $\boldsymbol{h}$, $\boldsymbol{u}$, and $\boldsymbol{h}^{(+)}$ are vectors of size $T_x R_x(P+1) \times 1$ and $E\left[\boldsymbol{u}\boldsymbol{u}^*\right] = \boldsymbol{I}_{T_x R_x(P+1)}$. The channel covariance is given by

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

C. Receiver

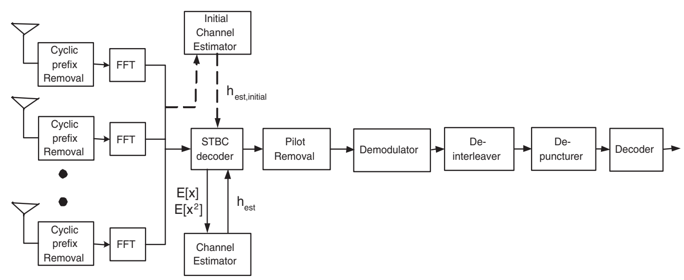

Fig. 2. OSTBC-OFDM Receiver

A block diagram of the receiver is shown in Figure 2. The pilots are first used to obtain an initial channel estimate. Subsequently, the Expectation Maximization (EM) algorithm alternates between data recovery (expectation) and channel recovery (maximization). This is done by the iterative module, made up of the Space-Time-Block decoder/data detector and the channel estimator. It is then followed by the familiar blocks of demodulation, deinterleaving, and decoding.

III. Input-output Equations for MIMO-OFDM

Consider the interaction between transmit antenna t_x and receive antenna r_x in a SISO link. The frequency domain OFDM symbol $\boldsymbol{\mathcal{X}}_{t_x}$ is related to the output symbol $\boldsymbol{\mathcal{Y}}_{t_x}$ by

$$\boldsymbol{\mathcal{Y}}_{r_x}^{t_x} = \operatorname{diag}\left(\boldsymbol{\mathcal{X}}_{t_x}\right)\tilde{\boldsymbol{Q}}_{P+1}^{*}\boldsymbol{h}_{r_x}^{t_x} + \boldsymbol{\mathcal{N}}_{r_x}^{t_x} \tag{5}$$

where $\tilde{\boldsymbol{Q}}_{P+1}$ represents the first $P+1$ rows of the size-N IDFT matrix $\mathbf{Q}$ and where $\boldsymbol{h}_{r_x}^{t_x}$ is the time-domain impulse response (which is a vector of length $P+1$). By superposition, we can express the I/O equation at receive antenna r_x of a MIMO system having T_x transmit antennas as

$$\boldsymbol{\mathcal{Y}}_{r_x} = \left[\ \operatorname{diag}(\boldsymbol{\mathcal{X}}_1) \quad \cdots \quad \operatorname{diag}(\boldsymbol{\mathcal{X}}_{T_x})\ \right]\left(\boldsymbol{I} \otimes \tilde{\boldsymbol{Q}}_{P+1}^{*}\right)\boldsymbol{h}_{r_x} + \boldsymbol{\mathcal{N}}_{r_x} \tag{6}$$

where $\otimes$ is the Kronecker product operator. This relationship can be used to derive I/O equation in the presence of space-time coding.

A. Input/Output Equations with STBC: Channel Estimation Version

Consider the set of N_s uncoded OFDM symbols $\{\boldsymbol{\mathcal{S}}(1), \ldots, \boldsymbol{\mathcal{S}}(N_s)\}$. Using ST coding, we wish to transmit these symbols in one OSTBC block using T_x antennas and T_b time slots. We achieve this using the set of $T_x \times T_b$ matrices $\{\boldsymbol{A}(1), \boldsymbol{B}(1), \ldots, \boldsymbol{A}(N_s), \boldsymbol{B}(N_s)\}$ which characterize the ST code used (see [17]). We can show that the I/O equation at receive antenna r_x for a MIMO system implementing such a code is given by

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

where $\boldsymbol{\mathcal{Y}}_{r_x} = \left[\begin{array}{ccc} \boldsymbol{\mathcal{Y}}^T_{r_x}(1) & \cdots & \boldsymbol{\mathcal{Y}}^T_{r_x}(T_b) \end{array}\right]$ and

$$\boldsymbol{X} = \sum_{n_s=1}^{N_s} \left[\begin{array}{cc} \operatorname{diag}\left(\boldsymbol{a}_1(n_s) \otimes \mathcal{R}\boldsymbol{\mathcal{S}}(n_s) + j\boldsymbol{b}_1(n_s) \otimes \mathcal{I}\boldsymbol{\mathcal{S}}(n_s)\right) & \cdots \end{array} \right.$$
$$\left. \operatorname{diag}\left(\boldsymbol{a}_{T_b}(n_s) \otimes \mathcal{R}\boldsymbol{\mathcal{S}}(n_s) + j\boldsymbol{b}_{T_b}(n_s) \otimes \mathcal{I}\boldsymbol{\mathcal{S}}(n_s)\right) \right] \left(\boldsymbol{I} \otimes \tilde{\boldsymbol{Q}}^*_{P+1}\right),$$

and where $\boldsymbol{a}_{t_x}(n_s)$ ($\boldsymbol{b}_{t_x}(n_s)$) is the t_xth row of $\boldsymbol{A}(n_s)$ ($\boldsymbol{B}(n_s)$). Collecting this relationship over all receive antennas yields

$$\boxed{\boldsymbol{\mathcal{Y}} = (\boldsymbol{I} \otimes \boldsymbol{X})\boldsymbol{h} + \boldsymbol{\mathcal{N}}} \tag{7}$$

For initial channel estimation, we construct the set of pilot/output equations given by

$$\boxed{\boldsymbol{\mathcal{Y}}_{I_\mathrm{p}} = (\boldsymbol{I} \otimes \boldsymbol{X}_{I_\mathrm{p}})\boldsymbol{h} + \boldsymbol{\mathcal{N}}_{I_\mathrm{p}}} \tag{8}$$

These equations are a pruned version of (7) determined by the index set I_p of the pilot locations.

B. Input/Output Equations with STBC: Data Detection Version

Signal detection in ST-coded OFDM is done on a tone-by-tone basis, except that the tones are collected for the whole ST block (for R_x receive antennas and over T_b time slots). From (6), we extract the following I/O equation for tone n belonging to the OFDM symbol t_b

$$\boldsymbol{\mathcal{Y}}(t_b, n) = \boldsymbol{\mathcal{H}}(n)\boldsymbol{\mathcal{X}}(t_b, n) + \boldsymbol{\mathcal{N}}(t_b, n) \tag{9}$$

where $\boldsymbol{\mathcal{Y}}(t_b, n) = \left[\begin{array}{ccc} \mathcal{Y}_1(t_b, n) & \ldots & \mathcal{Y}_{R_x}(t_b, n) \end{array}\right]^T$, $\big(\boldsymbol{\mathcal{X}}(t_b, n)$ and $\boldsymbol{\mathcal{N}}(t_b, n$ are defined similarly$\big)$, and where

$$\boldsymbol{\mathcal{H}}(n) = \left[\begin{array}{ccc} \mathcal{H}^1_1(n) & \cdots & \mathcal{H}^{T_x}_1(n) \\ \vdots & \cdots & \vdots \\ \mathcal{H}^1_{R_x}(n) & \cdots & \mathcal{H}^{T_x}_{R_x}(n) \end{array}\right] \tag{10}$$

By concatenating (9) for $t_b = 1, \cdots, T_b$, we can show that (see [17])

$$\boldsymbol{\mathcal{Y}}(n) = \boldsymbol{C}(n) \left[\begin{array}{c} \mathcal{R}\mathcal{S}_n \\ \mathcal{I}\mathcal{S}_n \end{array}\right] + \boldsymbol{\mathcal{N}}(n) \tag{11}$$

where

$$\boldsymbol{\mathcal{Y}}(n) = \left[\begin{array}{c} \boldsymbol{\mathcal{Y}}(1, n) \\ \vdots \\ \boldsymbol{\mathcal{Y}}(T_b, n) \end{array}\right] \quad \boldsymbol{\mathcal{S}}_n = \left[\begin{array}{c} \boldsymbol{\mathcal{S}}(1, n) \\ \vdots \\ \boldsymbol{\mathcal{S}}(N_s, n) \end{array}\right]$$

$$\begin{aligned} \boldsymbol{C}(n) &= \left[\begin{array}{cc} \boldsymbol{C}_a(n) & \boldsymbol{C}_b(n) \end{array}\right] \\ \boldsymbol{C}_a(n) &= \left[\begin{array}{ccc} \operatorname{vec}(\boldsymbol{\mathcal{H}}(n)\boldsymbol{A}(1)) & \cdots & \operatorname{vec}(\boldsymbol{\mathcal{H}}(n)\boldsymbol{A}(N_s)) \end{array}\right] \\ \boldsymbol{C}_b(n) &= \left[\begin{array}{ccc} \operatorname{vec}(\boldsymbol{\mathcal{H}}(n)\boldsymbol{B}(1)) & \cdots & \operatorname{vec}(\boldsymbol{\mathcal{H}}(n)\boldsymbol{B}(N_s)) \end{array}\right] \end{aligned}$$

For an OSTBC code, $\mathcal{R}\left[\boldsymbol{C}^*(n)\boldsymbol{C}(n)\right] = ||\boldsymbol{\mathcal{H}}(n)||^2\boldsymbol{I}$. Thus, on multiplying both sides of (11) by $\boldsymbol{C}^*(n)$, taking the real part, and rearranging terms, we can show that

$$\boxed{\widetilde{\boldsymbol{\mathcal{Y}}}(n) = \|\boldsymbol{\mathcal{H}}(n)\|^2\boldsymbol{\mathcal{S}}(n) + \widetilde{\boldsymbol{\mathcal{N}}}(n)} \tag{12}$$

where

$$\begin{aligned} \widetilde{\boldsymbol{\mathcal{Y}}}(n) &= \mathcal{R}\left[\boldsymbol{C}^*_a(n)\boldsymbol{C}(n)\boldsymbol{\mathcal{Y}}(n)\right] + j\mathcal{R}\left[\boldsymbol{C}^*_b(n)\boldsymbol{C}(n)\boldsymbol{\mathcal{Y}}(n)\right] \\ \widetilde{\boldsymbol{\mathcal{N}}}(n) &= \mathcal{R}\left[\boldsymbol{C}^*_a(n)\boldsymbol{C}(n)\boldsymbol{\mathcal{N}}(n)\right] + j\mathcal{R}\left[\boldsymbol{C}^*_b(n)\boldsymbol{C}(n)\boldsymbol{\mathcal{N}}(n)\right] \end{aligned}$$

Since $\boldsymbol{C}(n)$ is orthogonal, the noise $\widetilde{\boldsymbol{\mathcal{N}}}(n)$ remains white and $\boldsymbol{\mathcal{S}}(n)$ can be detected from (12) on an element-by-element basis.

IV. Channel Estimation

We start this section by explaining how to estimate the channel when the data is known and use that to treat the unobserved data case. Throughout this section, we will assume that the channel, $\boldsymbol{h}_d$, satisfies the generic state-space model

$$\begin{aligned} \boldsymbol{\mathcal{Y}}_d &= \boldsymbol{X}_d\boldsymbol{h}_d + \boldsymbol{\mathcal{N}}_d & (13) \\ \boldsymbol{h}^{(+)}_d &= \boldsymbol{F}_d\boldsymbol{h}_d + \boldsymbol{G}_d\boldsymbol{u}_d & (14) \end{aligned}$$

where the subscript d indicates dummy variables[1]

A. Known data case

When the input $\boldsymbol{X}_d$ is available, we perform channel estimation by maximizing the log-likelihood function

$$\begin{aligned} \hat{\boldsymbol{h}}_d &= \arg\max_{\boldsymbol{h}_d} p(\boldsymbol{h}_d|\boldsymbol{\mathcal{Y}}_d, \boldsymbol{X}_d) & (15) \\ &= \arg\max_{\boldsymbol{h}_d} p(\boldsymbol{h}_d)p(\boldsymbol{\mathcal{Y}}_d, \boldsymbol{X}_d|\boldsymbol{h}_d) & (16) \end{aligned}$$

Here $p(\boldsymbol{h}_d|\boldsymbol{\mathcal{Y}}_d, \boldsymbol{X}_d)$ is the pdf of the channel given the input and output data. More precisely, the dynamical dependence between the present and the past expressed by (14) allows us to use all *past* input and output data in addition to the present ones. In this case, the log-likelihood function (15) is maximized given all the past and present data and is achieved efficiently using the Kalman filter [18], described by the equations below ($\boldsymbol{\Pi}$ denotes the covariance of $\boldsymbol{h}^{(0)}$)

$$\begin{aligned} \boldsymbol{P}^{(+|-)} &= \begin{cases} \boldsymbol{\Pi} & \text{for first time instant} \\ \boldsymbol{F}_d\boldsymbol{P}^{(-)}\boldsymbol{F}^*_d + \boldsymbol{G}_d\boldsymbol{G}^*_d \end{cases} & (17) \\ \boldsymbol{R}_e &= \sigma^2_n\boldsymbol{I} + \boldsymbol{X}_d\boldsymbol{P}^{(+|-)}\boldsymbol{X}^*_d & (18) \\ \boldsymbol{K}_f &= \boldsymbol{P}^{(+|-)}\boldsymbol{X}^*_d\boldsymbol{R}^{-1}_e & (19) \\ \hat{\boldsymbol{h}}^{(+)}_d &= \begin{cases} \boldsymbol{0} & \text{for the first time instant} \\ \left(\boldsymbol{I} - \boldsymbol{K}_f\boldsymbol{X}_d\right)\boldsymbol{F}_d\hat{\boldsymbol{h}}_d + \boldsymbol{K}_f\boldsymbol{\mathcal{Y}}_d \end{cases} & (20) \\ \boldsymbol{P}^{(+)} &= \boldsymbol{P}^{(+|-)} - \boldsymbol{K}_f\boldsymbol{R}_e\boldsymbol{K}^*_f & (21) \end{aligned}$$

[1] In this section, we describe channel estimation in terms of a generic state-space model and dummy variables. This allows us to describe channel estimation in general and succinct terms and without having to carry complicated expressions around (involving the kronecker product, for example).

B. Unknown data case: The EM Algorithm

The challenge in our algorithm is that the input is not available. Hence, instead of maximizing the conditional distribution in (15), we maximize an *averaged* form of the distribution, i.e.

$$\hat{\boldsymbol{h}}_d^{\text{newiter}} = \arg\max_{h_d} E_{X|\mathcal{Y}_d, \hat{h}_d^{\text{olditer}}} \left[\ln p(\boldsymbol{h}_d|\boldsymbol{X}_d, \boldsymbol{\mathcal{Y}}_d)\right] \quad (22)$$

where averaging is performed over the unknown input given the output $\boldsymbol{\mathcal{Y}}_d$ and the channel estimate of the previous iteration. This represents the EM algorithm. Each iteration of the algorithm produces an estimate $\hat{\boldsymbol{h}}_d$ that monotonically increases the likelihood of the channel $\boldsymbol{h}_d$. When the data $\boldsymbol{X}_d$ is unobserved, we can not employ the Kalman filter (17)-(21) to estimate the channel. Instead, the EM-based channel estimate is obtained by employing the Kalman filter (17)-(21) to the following state-space model [8]

$$\boldsymbol{h}_d^{(+)} = \boldsymbol{F}_d\boldsymbol{h}_d + \boldsymbol{G}_d\boldsymbol{u}_d \quad (23)$$

$$\begin{bmatrix} \boldsymbol{\mathcal{Y}}_d \\ \boldsymbol{0}_{P\times 1} \end{bmatrix} = \begin{bmatrix} E\left[\boldsymbol{X}_d\right] \\ \text{Cov}\left[\boldsymbol{X}_d{}^*\right]^{1/2} \end{bmatrix} \boldsymbol{h}_d + \begin{bmatrix} \boldsymbol{\mathcal{N}}_d \\ \underline{\boldsymbol{z}}_d \end{bmatrix} \quad (24)$$

where $\underline{\boldsymbol{z}}_d$ is Gaussian $\mathcal{N}(\boldsymbol{0}_{P\times 1}, \sigma_n^2\boldsymbol{I})$ and independent from $\boldsymbol{\mathcal{N}}_d$. In other words, we employ the Kalman filter (17)-(21) with the following change of variables

$$\boldsymbol{X}_d \longrightarrow \begin{bmatrix} E\left[\boldsymbol{X}_d\right] \\ \text{Cov}\left[\boldsymbol{X}_d{}^*\right]^{1/2} \end{bmatrix}, \quad \boldsymbol{\mathcal{Y}}_d \longrightarrow \begin{bmatrix} \boldsymbol{\mathcal{Y}}_d \\ \boldsymbol{0}_{P\times 1} \end{bmatrix} \quad (25)$$

V. Algorithm Summary

In this section we summarize the steps taken in the algorithm

- **Initial Channel Estimation** The first step in the receiver operation is to obtain an initial estimate of the channel. We achieve this applying the Kalman filter to the dynamical channel model (4) together with the pilot/output equations (8)

$$\boldsymbol{h}^{(+)} = \left(\boldsymbol{I}_{T_xR_x} \otimes \boldsymbol{F}\right)\boldsymbol{h} + \left(\boldsymbol{I}_{T_xR_x} \otimes \boldsymbol{G}\right)\boldsymbol{u} \quad (26)$$

$$\boldsymbol{\mathcal{Y}}_{I_p} = (\boldsymbol{I} \otimes \boldsymbol{X}_{I_p})\boldsymbol{h} + \boldsymbol{\mathcal{N}}_{I_p} \quad (27)$$

  The Kalman filter (17)-(21) thus provides the initial channel estimate by performing the substitution

$$\begin{array}{ll} \boldsymbol{F}_d \longrightarrow \left(\boldsymbol{I}_{T_xR_x} \otimes \boldsymbol{F}\right) & \boldsymbol{G}_d \longrightarrow \left(\boldsymbol{I}_{T_xR_x} \otimes \boldsymbol{G}\right) \\ \boldsymbol{\mathcal{Y}}_d \longrightarrow \boldsymbol{\mathcal{Y}}_{I_p} & \boldsymbol{X}_d \longrightarrow \left(\boldsymbol{I}_{R_x} \otimes \boldsymbol{X}_{I_p}\right) \end{array} \quad (28)$$

  When time correlation information is not available to the receiver, the initial estimate can be obtained by setting $\boldsymbol{F} = \boldsymbol{0}$ in (28). In subsequent ST blocks, the final estimate calculated in the previous block is used to calculate the predicted portion of the channel estimate.

- **Expectation Step - Data** The receiver uses the latest channel estimate to perform the expectation step on the data. Let $\mathcal{S} = \left\{S_1, \ldots, S_{|\mathcal{S}|}\right\}$ where $|\mathcal{S}|$ is the size of the set $\mathcal{S}$, denote the alphabet set from which the elements of $\boldsymbol{\mathcal{S}}_n$ take their values. Based on the data detection relationship in (12), we can derive the conditional pdf $f(\mathcal{S}(n_s, n)|\widetilde{\mathcal{Y}}(n_s, n))$ and use this to calculate conditional expectation of $\mathcal{S}(n_s, n)$ and its second moment given the output $\widetilde{\mathcal{Y}}(n_s, n)$

$$E[\mathcal{S}(n_s, n)|\widetilde{\mathcal{Y}}(n_s, n)] = \frac{\sum_{i=1}^{|\mathcal{S}|} S_i e^{-\frac{\left|\tilde{y}(n_s,n) - \left\|\boldsymbol{\mathcal{H}}(n)\right\|^2 S_i\right|^2}{2\sigma_n^2}}}{\sum_{i=1}^{|\mathcal{S}|} e^{-\frac{\left|\tilde{y}(n_s,n) - \left\|\boldsymbol{\mathcal{H}}(n)\right\|^2 S_i\right|^2}{2\sigma_n^2}}}$$

$$E[|\mathcal{S}(n_s, n)|^2\,|\widetilde{\mathcal{Y}}(n_s, n)] = \frac{\sum_{i=1}^{|\mathcal{S}|} |S_i|^2 e^{-\frac{\left|\tilde{y}(n_s,n) - \left\|\boldsymbol{\mathcal{H}}(n)\right\|^2 S_i\right|^2}{2\sigma_n^2}}}{\sum_{i=1}^{|\mathcal{S}|} e^{-\frac{\left|\tilde{y}(n_s,n) - \left\|\boldsymbol{\mathcal{H}}(n)\right\|^2 S_i\right|^2}{2\sigma_n^2}}}$$

  These two moments are in turn used to calculate the first two moments of the ST coded input ($E[\boldsymbol{X}]$ and $\text{Cov}[\boldsymbol{X}]$).

- **Maximization Step - Channel Estimation** The receiver now uses the first two moments of the data to perform the maximization step on the channel. As we argued in subsection IV-B, the maximization step is carried out by running the Kalman filter (17)-(21) with the following change of variables

$$\boldsymbol{F}_d \longrightarrow \left(\boldsymbol{I}_{T_xR_x} \otimes \boldsymbol{F}\right), \quad \boldsymbol{G}_d \longrightarrow \left(\boldsymbol{I}_{T_xR_x} \otimes \boldsymbol{G}_d\right)$$

$$\boldsymbol{\mathcal{Y}}_d \longrightarrow \begin{bmatrix} \boldsymbol{\mathcal{Y}} \\ \boldsymbol{0}_{P\times 1} \end{bmatrix}, \quad \boldsymbol{X} \longrightarrow \begin{bmatrix} \boldsymbol{I}_{R_x} \otimes E\left[\boldsymbol{X}\right] \\ \boldsymbol{I}_{R_x} \otimes \text{Cov}\left[\boldsymbol{X}^*\right]^{1/2} \otimes \boldsymbol{I}_{R_x} \end{bmatrix}$$

- The expectation and maximization steps are alternated till a stopping criterion is satisfied. Once the stopping criterion is satisfied, the detected QAM-symbols are demodulated, de-punctured and de-interleaved. The resulting bits are then decoded by a Viterbi decoder.

VI. Simulations

The transmitter and receiver illustrated in Figure 1 and Figure 2 were implemented. The outer encoder is a rate 1/2 convolutional encoder and the coded bits are mapped to 16-QAM symbols using gray coding. We use the Alamouti code (so $N_s = T_b = 2$). Our MIMO channel model is simulated using the state-space model (4) with parameters, $\alpha(p) = \alpha = 0.985$, $\beta = 0.2$, $P = 7$ and $\boldsymbol{U}$ is $\mathcal{N}(0, I)$. The number of receive antennas, R_x, is set to 1 or 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. Sixteen pilots are used in the OFDM symbols making up the first ST block (25% of the data). The cyclic prefix is 16 samples long.

A. Bench Marking

We compare our algorithm with an EM-based iterative MMSE receiver such as those proposed in [10] and [11]. In contrast to our work, the authors in [10] and [11], take a data-centric approach, treating the transmitted signal as the desired parameter and the channel as the unobserved data. The algorithms confine their pilots to the first ST block which are used to produce an initial channel estimate. This initial estimate is used to predict the initial channel estimate for the subsequent ST blocks by employing a time correlation

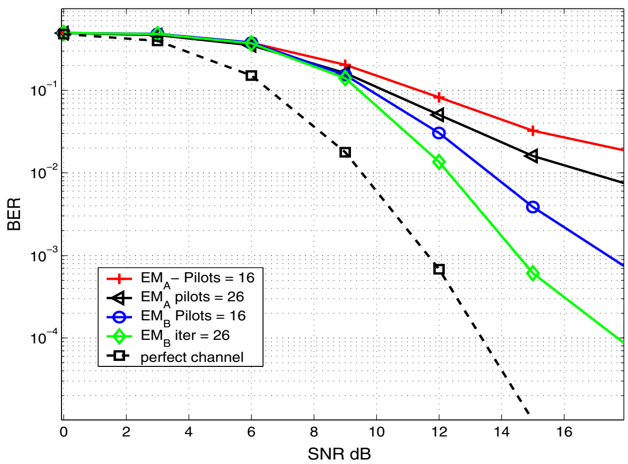

Fig. 3. Receiver Design Comparison

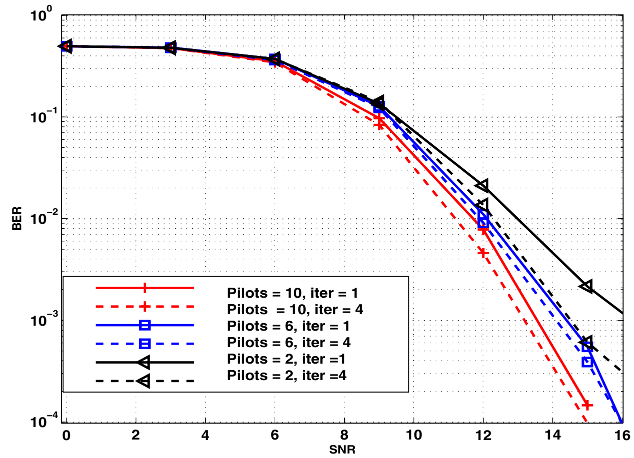

Fig. 4. BER Performance with Varied Number of Pilots

filter [10]. These initial estimates are used to kick-start the EM 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 (an MMSE estimation). The maximization step is simply the hard decision, ML decoding of the transmitted data with the previously obtained value of the channel.

In Figure 3, 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. We also implement both schemes with a total of 26 pilots as shown in Figure 3. The EM_A confines the pilots to the first ST block while in EM_B, we place 16 pilots in the first ST block and 2 pilots each in subsequent blocks. This ensures that both schemes incur the same pilot overhead.

With the parameters used in both scenarios, EM_B outperforms EM_A. One of the reasons for this performance improvement is the incorporation of correlation information and the most recent channel estimate of prior ST blocks at every iteration of the EM algorithm.

B. Sensitivity to Number of Iterations and Pilots

Here, we keep the number of pilots in the first ST block at 16, and we vary the number of pilot tones in the subsequent ST blocks (we use 10, 6 and 2 pilots). In Figure 4, the solid lines represent one EM iteration while the dashed lines represent four iterations.

The BER performance improves with increasing number of pilots. However, additional iterations give more significant gain when a small number of pilots are used. We also note that additional iterations can have substantial improvement for the low number of pilots case (e.g. the BER curve for the 2 pilots case is almost similar to that of the 6 pilots case).

C. Effect of Including Frequency and Time Correlation in the Channel Estimation

The impact of using both frequency and time correlation of the channel in the estimation is shown in Figure 5 for the 6-pilot scenario. In this figure, the solid lines are for $R_x = 1$ and the dashed lines for $R_x = 2$. $P_e = 1$ refers to channel estimation (for both pilot and data) using only frequency correlation while $P_e = 2$ implies the use of both frequency and time correlation in the channel estimation (see section V for details).

We observe an error floor when only the frequency correlation information is used in the channel estimation. This error floor remains regardless of the number of iterations. However, when we incorporate both frequency and time correlation information, we observe a significant improvement in BER (at a BER = 10^{-2}, the error floor drops by more than 10dB for $R_x = 1$ and $R_x = 2$). A single additional iteration shows a substantial improvement when compared to the pilot-based estimation case.

We conclude that including time correlation in the channel estimation process (especially for channels with high time correlation) increases the amount of information that can be harnessed by iterating.

D. Sensitivity to Time Variation

This is parameterized by $\alpha\,(0 \leq \alpha \leq 1)$ with lower values of α indicating a more time-variant channel. In Figure 6 we show the BER curves for a system with 6-pilots in subsequent symbols for $\alpha = 0.7, 0.8$ and 0.985. Results pertaining to channel estimates obtained for one iteration after the pilot based estimation are shown in solid lines while the dash lines represent the results with perfect channel knowledge.

We observe error-floors as the channel variation increases. It is obvious that as α decreases, the significance of including time correlation in the channel estimation process also decreases and more pilots are needed for better performance.

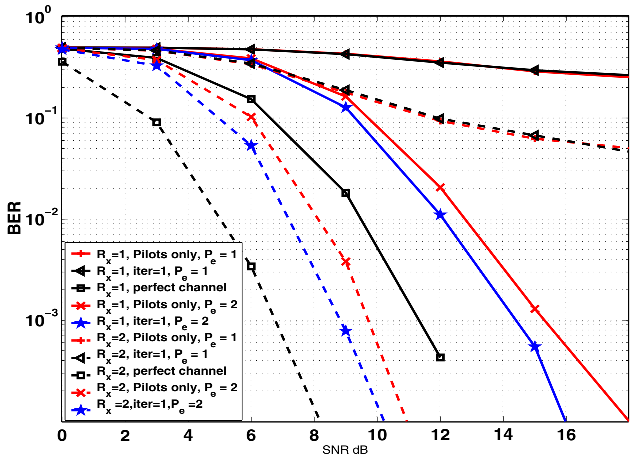

Fig. 5. BER Performance with Frequency and Time correlation

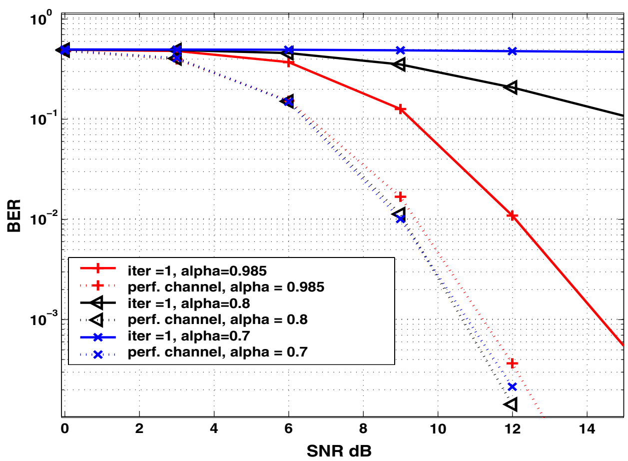

Fig. 6. BER Performance with Varying Time Correlation with 6 Pilots

VII. Conclusion

In this paper, we have proposed a receiver for MIMO-OFDM transmission over time-variant channels. The receiver makes full use of the data constraints (pilots, finite alphabet constraint and space-time code). It also exploits the the time and frequency correlation (channel constraints). The paper assumes the channel to be constant within the same space-time block but varying from one block to the next. This allows the receiver to operate in high speed environments. The receiver performs channel and data recovery within the same space-time block and hence avoids the need for data storage, making the receiver suitable for real-time applications. When compared with other MIMO receivers, our receiver makes the most use of the underlying data and channel constraints.

The receiver employs the EM algorithm to achieve channel and data recovery. Specifically, the data recovery (or the expectation step) is as simple as decoding a space-time block code. Channel recovery (or the maximization step) is performed using a Kalman filter. Simulation demonstrated the favorable behavior of our receiver as compared to other receivers.

We can generalize the algorithm presented in this paper to include the effects of the transmit filter and the channel transmit and receive spatial correlation. We can also modify the receiver to take care of (space-time) trellis as opposed to block codes. Neglecting storage and latency issues, we can modify the filter to perform estimation in the forward and backward direction resulting in better estimates. Reference [8] discusses the computational complexity of the algorithm, choosing the state-space parameters $\boldsymbol{A}$ and $\boldsymbol{B}$, and how to make the algorithm robust to uncertainties in these parameters.

References

[1] Paulraj A., Nabar R., Gore D. Introduction to Space Time Wireless Communications, Cambridge University Press. 2003

[2] Tujkovic D. , Juntti M. , Latva-Aho M. "Space-Frequency-Time Turbo Coded Modulation" ,IEEE Comm. Letters, , Vol. 5 Issue 12 , Dec. 2001 pp. 480 -482

[3] Gong Y., Letaief, K.B.,"Space-frequency-time coded OFDM for broadband wireless" IEEE Global TeleComm. Conference, 2001. Vol. 1 , 25-29 Nov. 2001 pp. 519 -523

[4] Li Y., Seshadri N., Ariyavisitakul S., "Channel Estimation For OFDM Systems With Transmitter Diversity In Mobile Wireless Channels" IEEE Journal on Sel. Areas in Comm. , Vol. 17 Issue 3, Mar 1999 pp. 461 -471

[5] Heath, R.W., Jr., Giannakis, G.B., "Exploiting Input Cyclostationarity For Blind Channel Identification In OFDM systems" IEEE Trans. Sig. Proc., Vol. 47 Issue 3 , Mar 1999 pp. 848 -856

[6] Edfors O., Sandell M., van de Beek J., WilsonS .K., Brjesson P.O., "OFDM Channel Estimation By Singular Value Decomposition". IEEE Trans. Comm., vol. 46, no. 7, July 1998, pp. 931-939

[7] Yang B., Letaief, K.B., Cheng, R.S., Zhigang C., "Channel Estimation For OFDM Transmission In Multipath Fading Channels Based On Parametric Channel Modeling" IEEE Trans. Comm., Vol.: 49 Issue 3 , March 2001 pp. 467 -479

[8] T.Y. Al-Naffouri, Adaptive Algorithms for Wireless Channel Estimation, PhD Dessertation, Under preparation, Stanford University

[9] Al-Naffouri, T.Y., Bahai, A., Paulraj, A., "Semi-Blind Channel Identification And Equalization In OFDM: An Expectation-Maximization Approach" IEEE 56th Vehicular Technology Conference, 2002., Vol. 1 2002 pp. 13 - 17

[10] Lu, B., Wang X., Li Y., "Iterative Receivers For Space-Time Block-Coded OFDM Systems In Dispersive Fading Channels" IEEE Trans. Wireless Comm., Vol. 1 Issue 2 , Apr 2002 pp. 213 -225

[11] Cozzo, C., Hughes, B.L., "Joint Channel Estimation And Data Detection In Space-Time Communications" IEEE Trans. Comm. , Vol.: 51 Issue 8 , Aug. 2003 pp. 1266 -1270

[12] Aldana C., de Carvalho E.,Cioffi J. M. "Channel Estimation for Multicarrier Multiple Input Single Output Systems using the EM Algorithm" In press, IEEE Trans. on Sig. Proc.

[13] Dempster A. P. , Laird N. M., Rubin D. B., "Maximum likelihood from incomplete data via the EM algorithm," *J. Royal Statistical Soc., Ser. B*, vol. 39, no. 1, pp. 1 - 38, 1977.

[14] Iltis, R.A., "Joint estimation of PN code delay and multipath using the extended Kalman filter" Comm., IEEE Trans. on , Vol. 38 Issue 10 , Oct. 1990 pp. 1677 -1685

[15] Komninakis, C.; Fragouli, C.; Sayed, A.H.; Wesel, R.D.; "Multi-Input Multi-Output Fading Channel Tracking And Equalization Using Kalman Estimation" IEEE Trans. Sig. Proc., Vol. 50 Issue 5 , May 2002 Pg, 1065 -1076

[16] Zhendao, W., G. B. Giannakis, "Wireless Multicarrier Communications", IEEE Sig. Proc. Mag., Vol. 17, Issue 3, May 2000 , pp. 29 -48,

[17] Larsson E. , Stoica P. Space-Time Block Coding for Wireless Comm., Cambridge University Press, 2003

[18] Kailath T., Sayed A. H., Hassibi B., Linear Estimation, Prentice Hall, NJ, 2000.