Adaptive Multi-Input Multi-Output Fading Channel Equalization using Kalman Estimation

Christos Komninakis, Christina Fragouli, Ali H. Sayed, and Richard D. Wesel
Electrical Engineering Department
University of California at Los Angeles

Abstract—This paper addresses the problem of adaptive channel tracking and equalization for multi-input multi-output (MIMO) time-variant frequency-selective channels. A finite-length minimum-mean-squared-error decision-feedback equalizer (MMSE-DFE) performs the equalization task, while a Kalman filter tracks the MIMO channel, which models the corrupting effects of inter-symbol interference (ISI), inter-user interference (IUI), and noise. The Kalman tracking is aided by previous hard decisions produced by the DFE, with a decision delay $\Delta > 0$, which causes the Kalman filter to track the channel with a delay. A channel prediction module bridges the time gap between the channel estimates produced by the Kalman filter and those needed for the DFE adaptation. The proposed algorithm offers good tracking behavior for multi-user fading ISI channels at the expense of higher complexity.

I. INTRODUCTION

This paper generalizes the work of [1], where Kalman estimation and DFE equalization of a single user linear dispersive channel was presented. However, in modern wireless communication systems, apart from ISI and fading, a significant source of performance degradation is also inter-user interference (IUI). This effect arises in many scenarios such as wideband multiple access systems [2] or narrowband systems with many transmit/receive antennas [3]. In general, such systems can be modeled as multi-input multi-output (MIMO) time-variant frequency-selective channels with a few (matrix valued) taps, for which the receiver has a dual role, to track the channel tap coefficients and to equalize the channel, while suppressing IUI. Different choices are available to the designer for implementing channel estimation and equalization, depending on channel modeling and on the complexity invested in each task.

This paper uses a Kalman filter to track the MIMO channel taps. These taps are typically modeled as mutually uncorrelated circular complex Gaussian random processes, having locally constant means, due to large scale path loss, reflections and shadowing effects. The taps also have a time-varying part with autocorrelation properties corresponding to the wide-sense stationary and uncorrelated scattering "WSSUS" model of Bello [4]. If the tap means are zero, the channel is said to introduce Rayleigh fading (worst case), while a non-zero mean tap corresponds to Ricean fading. The Kalman channel estimator is aided by previous hard decisions about the transmitted symbols from all users, produced by the MIMO equalizer. It should be noted that, at least for the single-user channel, Kalman-based estimation methods are quite common

in the literature (e.g., [5] uses the extended Kalman filter to track a channel with unknown delays). Also, in [6] the Kalman approach is used to formulate extended forms of the RLS algorithm, and the tracking superiority of those is demonstrated compared to the standard RLS and LMS algorithms.

Assuming perfect knowledge of the MIMO channel, the optimum receiver is a maximum likelihood sequence estimator (MLSE), but its complexity is prohibitive, even for low order channels with a small number of inputs and outputs. Here we use the MIMO finite-length minimum-mean-squared-error decision-feedback equalizer (MMSE-DFE), derived in [7] and optimized for decision delay $\Delta \geq 0$. The choice of $\Delta > 0$ improves performance for a wide range of channels, as shown in [1] for the single-user channel. However, this delay poses the problem of channel prediction when combined with the Kalman tracking procedure mentioned above, because there is a time gap of Δ between channel estimates produced by the Kalman filter (aided by the delayed DFE decisions) and the channel estimates needed for the current DFE adaptation. We discuss simple methods to bridge this time gap, and show simulation results to demonstrate that the joint tracking and equalization algorithm proposed in this paper offers good performance. In fact, it outperforms conventional adaptive equalization algorithms, such as LMS or RLS, which do not have an explicit mechanism for incorporating the largely invariant channel statistics, such as the Doppler rate and the channel mean, in case they are known to the receiver from a previous training phase.

The paper is organized as follows. Section II presents the channel model. Section III introduces the receiver block diagram, and discusses the Kalman-based tracking, the channel prediction, and the delay-optimized adaptive DFE design. Section IV presents simulation results, and Section V concludes the paper.

II. CHANNEL MODEL

The MIMO channel with n_i inputs and n_o outputs can be modeled as a matrix FIR filter. Specifically, consider the received observable $u_t^{(j)}$ from receiver j (with $j = 1, \ldots, n_o$) at time t:

$$u_t^{(j)} = \sum_{i=1}^{n_i} \sum_{m=0}^{\nu^{(i,j)}} c_m^{(i,j)} x_{t-m}^{(i)} + v_t^{(j)} \tag{1}$$

where $c_m^{(i,j)}$ is the mth tap of the impulse response of order $\nu^{(i,j)}$ between the ith input $x^{(i)}$ and the jth output, $x_{t-m}^{(i)}$ is the complex baseband constellation point transmitted by the ith user at time $t - m$, and $v_t^{(j)}$ is the complex noise sample at the jth receiver. Setting $\nu = \max_{i,j} \nu^{(i,j)}$, we write the MIMO channel input-output relationship in vector form as follows:

$$\mathbf{u}_t = \sum_{m=0}^{\nu} \mathbf{C}_m(t)\mathbf{x}_{t-m} + \mathbf{v}_t \qquad (2)$$

where $\mathbf{u}_t$ and $\mathbf{v}_t$ are column vectors of length n_o, $\mathbf{C}_m(t)$ are $n_o \times n_i$ matrix channel taps, and $\mathbf{x}_t$ are column vectors of length n_i. In essence, there exist a total of $n_i n_o$ interfering ISI channels (see Fig. 1 for (2, 2) system):

$$\mathbf{c}^{(i,j)}(t) = \left[c_0^{(i,j)}(t) \quad c_1^{(i,j)}(t) \quad \cdots \quad c_{\nu^{(i,j)}}^{(i,j)}(t) \right] \qquad (3)$$

and each of their taps can be written as:

$$c_m^{(i,j)}(t) = \bar{c}_m^{(i,j)} + h_m^{(i,j)}(t), \quad m = 0, \ldots, \nu \qquad (4)$$

where we set the constant part $\bar{c}_m^{(i,j)} = 0$ for $m > \nu^{(i,j)}$. The time-varying part $h_m^{(i,j)}(t)$ is a zero-mean, wide-sense-stationary complex Gaussian process, uncorrelated with any other $h_{m'}^{(i',j')}(t)$, and has time-autocorrelation properties governed by the Doppler rate $f_d T$, (T is the baud duration) as in [8]:

$$E\{h_m^{(i,j)}(t_1)[h_m^{(i,j)}(t_2)]^*\} \sim \mathcal{J}_o\left(2\pi f_d^{(m,i,j)} T |t_1 - t_2|\right) \qquad (5)$$

where $\mathcal{J}_o(\cdot)$ is the zero-order Bessel function of the first kind.

$x^{(1)}(t)$ $\mathbf{c}^{(1,1)}(t)$ $u^{(1)}(t)$

$\mathbf{c}^{(2,1)}(t)$ $\mathbf{c}^{(1,2)}(t)$

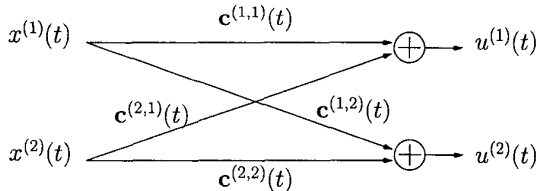$x^{(2)}(t)$ $\mathbf{c}^{(2,2)}(t)$ $u^{(2)}(t)$

Fig. 1. The interfering ISI channels in a (2, 2) MIMO system.

Notice that each of the $n_i n_o (\nu + 1)$ taps varies independently (although this will not be essential in our development) and can have a different Doppler rate. More importantly, the autocorrelation functions are non-rational, hence no ARMA model is an exact representation of the time evolution of the channel taps. However, only the first few correlation terms (for small $|t_1 - t_2|$) are important for the design of any receiver. Thus, even low order autoregressive models, or even a simple Markov model, can capture most of the channel tap dynamics and lead to effective tracking algorithms, as demonstrated below. For this purpose, we will need an alternative notation for the input-output relationship of the channel, of the form:

$$\mathbf{u}_t = \mathbf{X}_t \cdot \mathbf{c}_t + \mathbf{v}_t, \qquad (6)$$

where $\mathbf{c}_t$ is a long vector of length $n_i n_o (\nu + 1)$ containing all the channel taps at time t, $\mathbf{c}_t = \bar{\mathbf{c}} + \mathbf{h}_t$:

$$\mathbf{c}_t = \left[c_0^{(1,1)}(t) \ldots c_0^{(1,n_o)}(t) \ldots c_0^{(n_i,1)}(t) \ldots c_0^{(n_i,n_o)}(t) \ldots \right. $$
$$\left. c_\nu^{(1,1)}(t) \ldots c_\nu^{(1,n_o)}(t) \ldots c_\nu^{(n_i,1)}(t) \ldots c_\nu^{(n_i,n_o)}(t) \right]^T \quad (7)$$

and the $n_o \times n_o n_i (\nu + 1)$ data matrix $\mathbf{X}_t$ is:

$$\mathbf{X}_t = \left[x_t^{(1)}\mathbf{I}_{n_o} \ldots x_t^{(n_i)}\mathbf{I}_{n_o} \quad \cdots \quad x_{t-\nu}^{(1)}\mathbf{I}_{n_o} \ldots x_{t-\nu}^{(n_i)}\mathbf{I}_{n_o} \right] \quad (8)$$

The time-variant part of the channel is a vector process $\{\mathbf{h}_t\}$, which the receiver can model as a multichannel AR process of order p, as done in [9] for $n_i = n_o = 1$:

$$\mathbf{h}_t = \sum_{l=1}^{p} \mathbf{A}(l)\mathbf{h}_{t-l} + \mathbf{G}\mathbf{w}_t \qquad (9)$$

where $\mathbf{w}_t$ is a zero-mean i.i.d. circular complex Gaussian vector process with correlation matrix $\mathbf{R}_{ww}(j) = E\{\mathbf{w}_t \mathbf{w}_{t+j}^*\} = \mathbf{I}_{M+1}\delta(j)$. For $p = 1$, the best fit of the above AR(1) model with the theoretical autocorrelation given by (5) is achieved by choosing $\mathbf{A}(1) = \mathbf{F}$ to be a diagonal matrix with entries $f_k = \mathcal{J}_o(2\pi f_d^{(m,i,j)} T)$, where the index k enumerates all the taps $k = 1, \ldots, n_i n_o (\nu + 1)$, and $f_d^{(m,i,j)}$ is the Doppler of the mth tap of the channel from input i to output j. The matrix $\mathbf{G}$ is also diagonal with entries g_k. Each g_k is specified from the "specular-to-diffuse power ratio" K_k for each tap, as follows:

$$K_k \stackrel{def}{=} 10 \log\left(\frac{|\bar{c}_k|^2}{E\{|h_k|^2\}}\right) = 10 \log\left(\frac{|\bar{c}_k|^2}{g_k^2/(1 - f_k^2)}\right) \quad (10)$$

The ratio K_k indicates the ratio of power of the kth mean channel tap to the mean-squared power of the random, time-variant part of that tap, analogous to the Ricean factor defined in the wireless channel literature.

Although higher order models can be constructed for larger p, it turns out that this simple first-order approximation is enough to model the channel dynamics to the extent necessary for a receiver to operate. For perspective, in a 2.4 GHz transmission with baud rate of 20 Ksps and Doppler frequency $f_d = 200$ Hz (corresponding to vehicular velocity of 90 Km/h or 56 mph), then $f_d T = 0.01$ and $f_k = 0.999$.

A useful method to obtain the sequence of matrices $\mathbf{A}(l)$, $l = 1, \ldots, p$ during a training mode and for the SISO case (i.e. $n_i = 1, n_o = 1$) is provided in [9], via higher-than-second-order statistics (HOS). Their method is effective and requires only reasonable assumptions about the transmitted sequence and the noise. An analogous training method can be adopted for general MIMO channels. The information needed is only the channel mean $\bar{\mathbf{c}}$, the Doppler rates, the ratios K_k and the noise variance. In this paper we assume that these are known from a training phase, and focus on decision-aided tracking of the channel for relatively long time spans, without retraining.

III. Receiver Structure

The receiver uses a Kalman filter to track the channel and a DFE to equalize it. The Kalman filter assumes that the DFE hard decisions are correct and uses them to estimate the next channel value, while the DFE assumes correct Kalman filter channel estimates, and uses them in turn to equalize the channel. In general, the optimum decision delay $\Delta \geq 0$ can be determined analytically given a channel (see [7]). For a wide range of channels (including, but not limited to, non-minimum-phase channels), it turns out that a DFE producing decisions with $\Delta > 0$ is optimal. Even for the few channels where $\Delta = 0$ is best, it doesn't degrade performance to use a DFE with $\Delta > 0$, provided there are enough taps in the feedforward and feedback filters. Thus, it makes sense, particularly for time-varying channels like the ones treated here, to use decision delays $\Delta > 0$.

But when $\Delta > 0$, a time gap is created. At time t, when the last received vector is $\mathbf{u}_t$, the DFE produces the hard-decision $\hat{x}_{t-\Delta}$. The staggered decisions cause the Kalman filter to operate with delay, that is, operate at time $t - \Delta$, since it only has available hard decisions from the DFE up to then. However, the DFE design still needs channel estimates up to time t. Thus the receiver needs to use channel prediction to bridge the time gap between the Kalman channel estimation and the channel estimates needed for the current DFE adaptation.

The proposed system block diagram of Fig. 2 is meant to show the time succession of steps (1) through (4) below. The notation $\mathbf{y}_{n_1}^{n_2}$ (or $y_{n_1}^{n_2}$) means the collection of vector valued (or scalar) variables $\mathbf{y}_{n_1} \ldots \mathbf{y}_{n_2}$ or $y_{n_1} \ldots y_{n_2}$. In Fig. 2 the flow of new information is clockwise, starting from top left, with each of the blocks corresponding to one of the following actions:

1. $\hat{\mathbf{h}}_{t-\Delta} = \mathcal{K}\left(\hat{\mathbf{h}}_{t-\Delta-1}, \mathbf{u}_{t-\Delta-1}, \hat{x}_{t-\Delta-\nu-1}^{t-\Delta-1}\right)$
2. $\hat{\mathbf{h}}_{t-\Delta+1}^{t} = \mathcal{P}\left(\hat{\mathbf{h}}_{t-\Delta}, \mathbf{u}_{t-\Delta}^{t}\right)$
3. $\left[W_t^{opt}, B_t^{opt}\right] = \text{design DFE}(\hat{\mathbf{h}}_{t-N_f}^{t})$
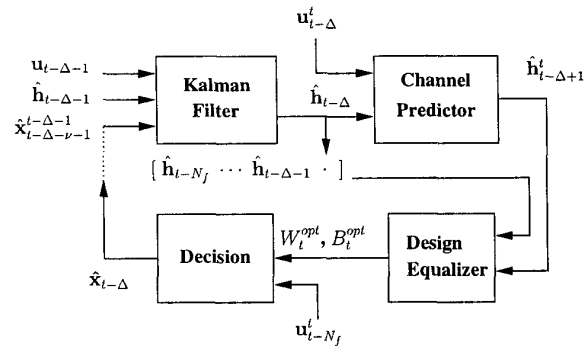4. $\hat{x}_{t-\Delta} = \text{DFE}(W_t^{opt}, B_t^{opt})$

Fig. 2. Receiver block diagram.

The iteration starts with the well-known Kalman filter recursions denoted by $\mathcal{K}(\cdot)$, which at time t yield the optimum

linear estimator of the time-varying part of the channel $\hat{\mathbf{h}}_{t-\Delta}$ as it was at time $t - \Delta$, because it is based on the (assumed reliable) DFE decisions $\hat{x}_{t-\Delta-1}, \ldots, \hat{x}_{t-\Delta-\nu-1}$, the received vector $\mathbf{u}_{t-\Delta-1}$ and the previously estimated channel vector $\hat{\mathbf{h}}_{t-\Delta-1}$. In the second step, $\mathcal{P}(\cdot)$ denotes a predictor that may exploit the additional received vectors $\mathbf{u}_t, \ldots \mathbf{u}_{t-\Delta}$, along with the last Kalman estimate $\hat{\mathbf{h}}_{t-\Delta}$ to compute the sequence of Δ predicted channels $\hat{\mathbf{h}}_t, \ldots, \hat{\mathbf{h}}_{t-\Delta+1}$.

Those Δ predicted channels, along with the $N_f - \Delta$ most recent channel estimates from the Kalman filter, are used by the DFE design module (see Sect. III-B) to design the optimum feedforward matrix filter, W_t^{opt}, and the feedback matrix filter, B_t^{opt} of an MMSE DFE. Finally, the newly designed DFE decodes one more n_i-dimensional symbol $\hat{x}_{t-\Delta}$, which is added to the collection of past (assumed reliable) decisions, which will help the Kalman filter make a new channel estimate $\hat{\mathbf{h}}_{t-\Delta+1}$ at the next iteration, taking place at time instant $t + 1$. In the following subsections we look at the implementation of the two main receiver modules in greater detail.

A. Kalman tracking and channel prediction

For simplicity, we limit our discussion to the AR(1) channel model, but the extension to higher order AR models is straightforward. The channel at time t has a constant (assumed known) mean $\bar{\mathbf{c}}$ and a zero-mean time-varying part $\mathbf{h}_t$, which follows the AR(1) model:

$$\mathbf{h}_{t+1} = \mathbf{F}\mathbf{h}_t + \mathbf{G}\mathbf{w}_t. \tag{11}$$

At time t, the (zero-mean) received vector $\mathbf{u}_t$ is given by:

$$\mathbf{u}_t = \mathbf{X}_t \cdot (\bar{\mathbf{c}} + \mathbf{h}_t) + \mathbf{v}_t. \tag{12}$$

Assuming the matrices $\mathbf{F}$ and $\mathbf{G}$ and the mean channel vector $\bar{\mathbf{c}}$ are known from a preceding training phase, and assuming the matrix of the most recent available decisions $\hat{\mathbf{X}}_{t-\Delta-1}$ to be equal to the true $\mathbf{X}_{t-\Delta-1}$ defined in (8), the receiver can use the Kalman filter to track the channel variation $\mathbf{h}_{t-\Delta}$, using as observable the vector $\mathbf{u}_{t-\Delta-1} - \hat{\mathbf{X}}_{t-\Delta-1}\bar{\mathbf{c}}$. The Kalman filter, operating with a delay Δ is described at time t by the series of equations [10]:

$$\hat{\mathbf{h}}_{t-\Delta} = \mathbf{F}\hat{\mathbf{h}}_{t-\Delta-1} + \mathbf{K}_{t-\Delta-1}\mathbf{e}_{t-\Delta-1}$$

$$\mathbf{e}_{t-\Delta-1} = \mathbf{u}_{t-\Delta-1} - \hat{\mathbf{X}}_{t-\Delta-1}(\hat{\mathbf{h}}_{t-\Delta-1} + \bar{\mathbf{c}})$$

$$\mathbf{K}_{t-\Delta-1} = (\mathbf{F}\mathbf{P}_{t-\Delta-1}\hat{\mathbf{X}}_{t-\Delta-1}^*)\mathbf{R}_{e,t-\Delta-1}^{-1}$$

$$\mathbf{R}_{e,n-\Delta-1} = \mathbf{R}_{vv} + \hat{\mathbf{X}}_{t-\Delta-1}\mathbf{P}_{t-\Delta-1}\hat{\mathbf{X}}_{t-\Delta-1}^*$$

$$\mathbf{P}_{t-\Delta} = \mathbf{F}\mathbf{P}_{t-\Delta-1}\mathbf{F}^* + \mathbf{G}\mathbf{G}^*$$
$$-\mathbf{K}_{t-\Delta-1}\mathbf{R}_{e,t-\Delta-1}\mathbf{K}_{t-\Delta-1}^*$$

The above Kalman recursions implement the optimum linear estimator for the time-varying part of the channel $\mathbf{h}_{t-\Delta}$. The last reliable decision made by the DFE and used by the Kalman filter at this time is $\hat{x}_{t-\Delta-1}$. For matrices $\mathbf{F}$ and $\mathbf{G}$ that are

multiples of the identity (produced, for instance, by uncorre-lated fading with the same Doppler and ratio K_k for all taps) fast algorithms for the above Kalman recursions can be pursued (see, e.g., [11]).

For designing the DFE at time t (step 3), the N_f most recent channel estimates are needed, where N_f is the order of the matrix feedforward filter W^{opt} of the DFE. Up to time $t - \Delta$ channel estimates are available from the Kalman filter. But the last Δ channel vectors $\hat{\mathbf{h}}_t, \ldots, \hat{\mathbf{h}}_{t-\Delta+1}$ have to be predicted. The implementation of the prediction depends upon the SNR of operation and how fast the channel varies. For a very slow varying channel, the simplest choice is to assume that the channel remains constant, that is:

$$\hat{\mathbf{h}}_t = \hat{\mathbf{h}}_{t-1} = \cdots = \hat{\mathbf{h}}_{t-\Delta} \qquad (13)$$

where $\hat{\mathbf{h}}_{t-\Delta}$ is already provided by the Kalman filter.

More generally, the optimal linear predictions, given that the channel follows the AR(1) model of (11), but ignoring the additional received vectors $\mathbf{u}_t, \ldots, \mathbf{u}_{t-\Delta}$ are:

$$\hat{\mathbf{h}}_t = \mathbf{F}^\Delta \hat{\mathbf{h}}_{t-\Delta}, \ldots, \hat{\mathbf{h}}_{t-\Delta+1} = \mathbf{F}\hat{\mathbf{h}}_{t-\Delta} \qquad (14)$$

where again $\hat{\mathbf{h}}_{t-\Delta}$ is the last Kalman channel estimate.

The received vectors $\mathbf{u}_t, \ldots, \mathbf{u}_{t-\Delta}$, which are also available, can be used to improve the prediction for a fast varying channel at high SNR. For example, one could formulate and optimize least-squares cost functions $J(\mathbf{h})$, forcing the predicted channel vectors $\hat{\mathbf{h}}_k, k = t, \ldots, t - \Delta + 1$ to the vicinity of the values of (14), weighted by how well a certain $\hat{\mathbf{h}}_k$ "justifies" the received vectors $\mathbf{u}_t, \ldots, \mathbf{u}_{t-\Delta}$ (see [1]). We do not pursue this prediction avenue in this paper, and all simulations are done using (14) for channel prediction.

B. DFE Design

The design of the optimum MMSE feedforward and feedback matrix filters W^{opt} and B^{opt} of lengths N_f and N_b respectively, as well as the optimum selection of the delay Δ for any (n_i, n_o) system is carefully solved in [7] and will not be repeated here. Suffice it to say that an essential part of the DFE design at every time instant t is the formulation of the $n_o N_f \times n_i (N_f + \nu)$ block pre-windowed channel matrix $\mathbf{H}$:

$$\mathbf{H} = \begin{bmatrix} \hat{\mathbf{C}}_0^t & \cdots & \hat{\mathbf{C}}_\nu^t & 0 & \cdots & 0 \\ 0 & \hat{\mathbf{C}}_0^{t-1} & \cdots & \hat{\mathbf{C}}_\nu^{t-1} & 0 & \cdots \\ \vdots & & & & & \vdots \\ 0 & \cdots & 0 & \hat{\mathbf{C}}_0^{t-N_f+1} & \cdots & \hat{\mathbf{C}}_\nu^{t-N_f+1} \end{bmatrix} \qquad (15)$$

where $\hat{\mathbf{C}}_m^k, m = 0, 1, \ldots, \nu, k = t, t - 1, \ldots, t - N_f + 1$ are the estimates of the $n_o \times n_i$ channel matrices $\mathbf{C}_m(k)$ in (2). It is obvious that constructing $\mathbf{H}$ of (15) merely involves adding the constant part $\bar{\mathbf{c}}$ to the long vector estimates $\hat{\mathbf{h}}_k, k = t, t - 1, \ldots, t - N_f + 1$ of the time-variant part of the channel (some of which are estimated via the Kalman filter and some are predicted), and then rearranging the resulting long vectors into the dimensions specified by Eq. (2).

IV. SIMULATION RESULTS

In all simulations presented in this section we implemented the receiver algorithm outlined above and compared the performance of the system, as measured by its symbol error rate (SER) when transmitting 4-PSK constellation points through $(2, 2)$ and single-user (i.e. $(1, 1)$) Ricean fading channels. We assumed the channel mean to be known at the beginning of each block of N symbols per user and tracked with the Kalman filter thereafter. For comparison purposes the plots include (dashed lines) the SER performance with the receiver having access to "genie-provided" perfect channel information, as well as that of conventional adaptive receiver algorithms, such as the LMS and RLS, admittedly less computationally intensive than the tracking algorithm proposed in this paper.

The SNR is set to be the same for both users (worst-case scenario, because having one user stronger than others facilitates decoding of *every* user). Also, all $n_i n_o$ mean interfering channels of (3) are normalized $||\bar{\mathbf{c}}^{(i,j)}||^2 = 1$, and all ratios K_k are chosen equal to a single K. Thus, with σ_v^2 being the noise variance at each receiver, the SNR plotted is the SNR of each interfering channel:

$$\text{SNR} = 10 \log \frac{1 + 10^{K/10}}{\sigma_v^2}, \qquad (16)$$

since the input 4-PSK points are normalized to unit power.

In Fig. 3 the mean channels were the normalized versions of $(1 + j)[1 \ 0.8]$ and $(1 + j)[1 \ 0.3]$ for the direct paths, and $(1 + j)[1 \ -0.8]$ and $(1 + j)[1 \ -0.5]$ for the interfering paths, the Doppler rate was $f_d T = 0.01$ (resulting in an AR(1) coefficient $f = 0.999$) and the specular-to-diffuse power ratio was $K = 6$ dB. The indicated SER performance represents unsupervised channel tracking for long blocks of $N = 5000$ symbols per user. The proposed Kalman-aided DFE performs less than 1 dB from perfect channel knowledge, while LMS adaptive DFE fails completely with such long blocklength. In Fig. 4 we kept the same *mean* $(2, 2)$ channel, but made it vary faster ($f = 0.99$, resulting from $f_d T = 0.032$), set $K = 10$ dB, and reduced N to 500. Performance is generally better (for LMS as well), and still the proposed algorithm outperforms LMS adaptation of the DFE. In both cases, the DFE has $N_f = 3$ and $N_b = 1$ matrix taps, and $\Delta = 2$.

Figures 5 and 6 show simulation results for a $(1, 1)$ system with mean channel $(1 + j)[1 \ 2.5]$, $K = 10$ dB for both taps, and a DFE with $\Delta = 5$, and $N_f = 7$ and $N_b = 1$ scalar taps. In both Figures the Doppler rate is $f_d T = 0.01$. However, in Fig. 5 the channel is indeed the AR(1) model with $f = 0.999$ (derived from $f_d T = 0.01$), while in Fig. 6 the channel is in fact generated according to Bello's model with the statistics described in (5), and is only approximated by the AR model of (11) with $f = 0.999$. We observe in both figures that the proposed Kalman-aided DFE outperforms conventional LMS/RLS adaptive algorithms. Note that this fact is more pronounced in the second case (Fig. 6), where the receiver exploits its knowledge about the channel statistics (i.e.
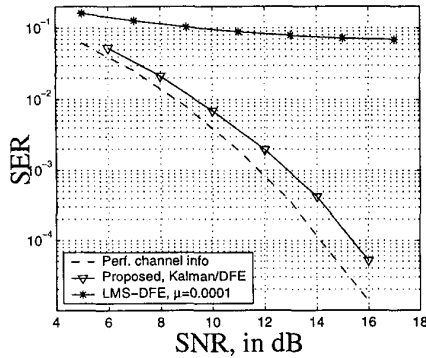
Fig. 3. Performance of $(2,2)$ system vs. SNR, with blocklength $N = 5000$, and channel derived from $f_d T = 0.01$. $K = 6$ dB.
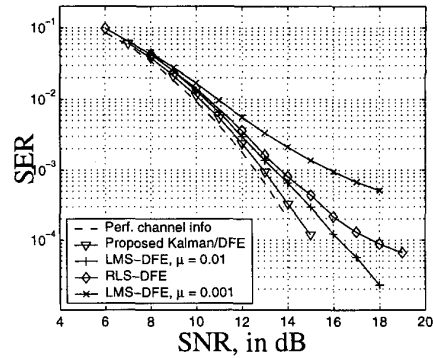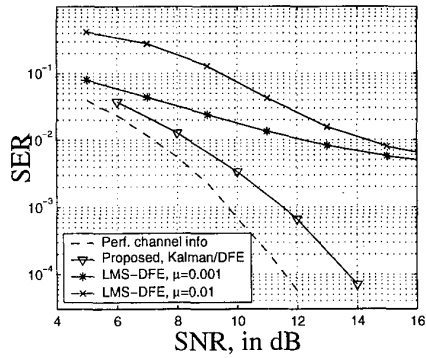

Fig. 4. Performance of $(2,2)$ system vs. SNR, with blocklength $N = 500$, and channel derived from $f_d T = 0.032$. $K = 10$ dB.

Fig. 5. Performance of $(1,1)$ system vs. SNR, with blocklength $N = 500$, and channel derived from $f_d T = 0.01$. $K = 10$ dB.
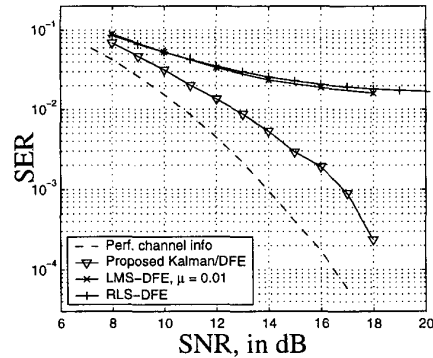

Fig. 6. Performance of $(1,1)$ system vs. SNR, with blocklength $N = 200$, and real Ricean channel with $f_d T = 0.01$ and $K = 10$ dB.

the Doppler) to form a first-order autoregressive model for the channel. This gives much better results than plain LMS/RLS adaptations, which do not have an explicit mechanism of incorporating known channel statistics, and only rely on careful selection of the parameters μ and λ to perform the best tracking they can.

V. CONCLUSIONS

This paper proposed a receiver structure to track and equalize a MIMO frequency selective fading channel. A Kalman filter was used for tracking the channel, and an MMSE DFE, optimized for decision delay $\Delta \geq 0$ to equalize the channel and supress IUI. The time gap between channel estimates produced by the Kalman filter and those needed for the DFE adaptation was bridged by using a simple prediction module. This algorithm, in exchange for larger complexity when compared to simple LMS/RLS updates of the DFE, offers improved performance and good tracking behavior for long unsupervised blocks, close to the performance of the idealistic case where perfect channel information is available at the receiver.

REFERENCES

[1] C. Komninakis, C. Fragouli, A.H. Sayed, and R.D. Wesel. Channel es-

timation and equalization in fading. *Proc. of 33rd Asilomar Conference on Signals, Systems and Computers*, pages 1159–1163, 1999.

[2] A. Duel-Hallen. A family of multiuser decision-feedback detectors for asynchronous code-division, multiple-access channels. *IEEE Trans. on Comm.*, pages 421–34, Feb/Mar/April 1995.

[3] A. Naguib, V. Tarokh, N. Seshadri, and A.R. Calderbank. A space-time coding modem for high-data-rate wireless communications. *IEEE Journal on Sel. Areas in Communications*, pages 1459–77, Oct. 1998.

[4] P. A. Bello. Characterization of randomly time-variant linear channels. *Transactions on Communication Systems*, CS(11):360–393, Dec. 1963.

[5] R. A. Iltis. Joint estimation of PN code delay and multipath using extended Kalman filter. *IEEE Trans. on Comm.*, 38(10):1677–1685, Oct. 1990.

[6] S. Haykin, A. H. Sayed, J. R. Zeidler, P. Yee, and P. C. Wei. Adaptive tracking of linear time-variant systems by extended RLS algorithms. *IEEE Transactions on Signal Processing*, 45(5):1118–1128, May 1997.

[7] N. Al-Dhahir and A.H. Sayed. The finite-length multi-input multi-output MMSE-DFE. *to appear in IEEE Trans. on Sign. Processing, 2000*, (see also Proc. of 33-rd Asilomar Conference on Signals, Systems and Computers, pp. 207–209, Oct. 1999).

[8] W. C. Jakes, Jr. *Microwave Mobile Communications*. John Wiley & Sons, NY, 1974.

[9] M. K. Tsatsanis, G. B. Giannakis, and G. Zhou. Estimation and equalization of fading channels with random coefficients. *Signal Processing*, 53(2-3):211–229, Sept. 1996.

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

[11] A. H. Sayed and T. Kailath. Extended Chandrasekhar recursions. *IEEE Trans. on Aut. Control*, 39(3):619–623, Mar. 1994.