Channel Estimation and Equalization in Fading

Christos Komninakis, Christina Fragouli, Ali H. Sayed* and Richard D. Wesel†

Electrical Engineering Department
University of California at Los Angeles

{chkomn, christin, sayed, wesel}@ee.ucla.edu

Abstract

This paper addresses the problem of channel tracking and equalization in frequency-selective fading channels. Low-order autoregressive models approximate the channel taps, where each tap is a circular complex Gaussian random process with the typical U-shaped spectrum, and uncorrelated with each other. A Kalman filter tracks the time-varying channel, using the decisions of an adaptive minimum-mean-squared-error decision-feedback equalizer (DFE). The DFE is optimized for decision delay $\Delta > 0$, which exhibits performance advantages over decision delay $\Delta = 0$ for a wide range of channels. The DFE staggered decisions cause the Kalman filter to also track the channel with a delay. The receiver also uses a channel prediction module to bridge the time gap between the Kalman channel estimation and the channel estimates needed for the DFE adaptation. The proposed algorithm offers good tracking behavior, thus allowing for reduction in the amount of training symbols needed to effectively track a time-varying frequency selective channel.

1 Introduction

Time-variant frequency-selective fading channels present a severe challenge to the designer of a wireless communication system. They are usually considered to comprise a rather small number of taps $M + 1$ (M is of the order of 5 for GSM, and it rarely exceeds 2 for IS-136), each of which is modeled as a circular complex Gaussian random process, uncorrelated with the other taps and having a locally constant mean (giving rise to a wide-sense stationary and uncorrelated scattering channel, called "WSSUS", see [1]). If the tap mean is zero, the channel is said to introduce Rayleigh fading (worst case), while a non-zero mean tap corresponds to Ricean fading.

*Supported by NSF grants CCR-9732376 and ECS-9820765.
†Supported by NSF CAREER award #9733089, and Texas Instruments.

The receiver has a dual role, to estimate the time-varying channel tap coefficients (tracking) and to equalize the channel. Several choices are available for the implementation of the estimation and equalization tasks, depending on the modeling of the channel and the complexity invested in each task. For example, the tap amplitudes can be approximated by a finite-state Markov chain and then MAP channel estimation and/or equalization may be used. The number of states though for the decoding increases exponentially with the constellation size and the number of channel taps.

This paper uses a Kalman filter to track the channel and a Decision-Feedback Equalizer (DFE) for subsequent channel equalization, following the approach of [2]. The main difference is that we use an MMSE decision-feedback equalizer [3] optimized for decision delay $\Delta > 0$, which exhibits performance advantages over decision delay $\Delta = 0$ for a wide range of channels. Simulation results demonstrate that the performance of the system proposed in this paper (see Section 3) is better than employing a fast adaptive algorithm (e.g., plain RLS) to track the channel, albeit at a higher computational cost. We also examine the effect in the performance when the DFE is substituted with a linear equalizer optimized for delay Δ.

We may add that Kalman-based channel estimation methods are quite common in the literature, (e.g., [4] uses the extended Kalman filter to track a channel with delays unknown a priori). Also, in [5] 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.

The paper is organized as follows: Section 2 presents the channel model. Section 3 introduces the receiver block diagram, and discusses in respective subsections the Kalman-based tracking, the channel prediction, and the delay-optimized adaptive DFE design. Section 4 presents the simulation results and Section 5 concludes the paper.

0-7803-5700-0/99/$10.00©1999 IEEE

2 Channel Model

The uncorrelated scattering assumption leads to an FIR channel model, in which each of the $M+1$ taps varies independently of every other tap, and with a time autocorrelation function giving rise to the well-known U-shaped spectrum [6]. The physical situation underlying this model is the existence of a few large scatters far from the mobile receiver and the existence of a large number (virtually a continuum) of small scatterers in the vicinity of the mobile receiver. Thus the channel at time n is described by a vector $\mathbf{h}_n = [h(n;0)\ \ h(n;1)\ \ldots\ h(n;M)]^T$ and the received value is given by:

$$u(n) = \sum_{k=0}^{M} h(n;k)s(n-k) + v(n) = \mathbf{s}_{n-M}^{n}\mathbf{h}_n + v(n)$$

where $s(n)$ is the transmitted constellation point at time n, the vector $\mathbf{s}_{n-M}^{n} = [s(n)\ldots s(n-M)]$ contains the $M+1$ most recently transmitted symbols, and $v(n)$ is a zero-mean, complex Gaussian i.i.d. random process, with variance σ_v^2.

The real and imaginary parts of the channel tap coefficients are uncorrelated [6], and since Gaussian, they are also independent. A generally accepted mathematical expression for the time-correlation of the channel taps (their real parts (R), imaginary parts (I) and amplitudes (r) respectively) can be summarized by the following correlation coefficients:

$$\begin{aligned}
\rho_{RR}(h_R(n_1;k_1),h_R(n_2;k_2)) &= \delta(k_1-k_2)\mathcal{J}_o(2\pi f_d T|n_1-n_2|)\\
\rho_{II}(h_I(n_1;k_1),h_I(n_2;k_2)) &= \delta(k_1-k_2)\mathcal{J}_o(2\pi f_d T|n_1-n_2|)\\
\rho_{RI}(h_R(n_1;k_1),h_I(n_2;k_2)) &= 0,\ \forall n_1,n_2,k_1,k_2\\
\rho_{rr}(|h(n_1;k_1)|,|h(n_2;k_2)|) &= \delta(k_1-k_2)\mathcal{J}_o{}^2(2\pi f_d T|n_1-n_2|)
\end{aligned}$$

where f_d is the Doppler frequency, T is the baud duration, and $\mathcal{J}_o(\cdot)$ is the zero-order Bessel function of the first kind.

Notice that since the autocorrelation functions are nonrational, no ARMA model is an exact representation of the time evolution of the channel taps. However, since only the first few correlation terms (for small $|n_1-n_2|$) are important for the design of any receiver, 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.

Assuming that the channel vector process $\{\mathbf{h}_n\}$ is zero-mean (the effect of the mean is just a constant addition), the receiver can model this vector process as a multichannel AR process of order p, as in [2]:

$$\mathbf{h}_n = \sum_{l=1}^{p} \mathbf{A}(l)\mathbf{h}_{n-l} + \mathbf{G}\mathbf{q}_n \qquad (1)$$

where $\mathbf{q}_n$ is an i.i.d. circular complex Gaussian vector process with correlation matrix $\mathbf{R}_{qq}(j) = E\{\mathbf{q}_n\mathbf{q}_{n+j}^*\} = \mathbf{I}_{M+1}\delta(j)$. For $p=1$, the best fit of the above AR(1) model with the theoretical autocorrelation is achieved by choosing $\mathbf{A}(1) = \mathbf{F}$ to be a diagonal matrix with entries $f_k = \mathcal{J}_o(2\pi f_d^{(k)}T)$, $k=0,1,\ldots,M$ where f_d^k is the Doppler of the k^{th} tap, and $\mathbf{G}$ to be also diagonal with entries $\sqrt{1-f_k^2}$. 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 situation with Doppler frequency $f_d = 200$ Hz (corresponding to vehicular velocity of 90 Km/h or 56 mph) and a baud rate of 20 Ksps, $f_k = 0.999$.

A useful method to obtain the sequence of matrices $\mathbf{A}(l)$, $l=1,\ldots,p$ during a training mode is provided in [2], via higher-than-second-order statistics (HOS). Their method is effective and requires only reasonable assumptions about the transmitted sequence and the noise. In this paper we assume that $\mathbf{F}$ is known from a training phase, and focus on decision-aided tracking of the channel.

3 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. For a wide range of channels, it makes sense to use a DFE with delayed decisions, because it is a more powerful equalizer (see, e.g., Fig. 2). But then a time gap is created. At time n, when the last received value is $u(n)$, the DFE produces the hard-decision $\hat{s}(n-\Delta)$. The staggered decisions cause the Kalman filter to operate with delay, that is, operate at time $n-\Delta$, since it only has available hard decisions from the DFE up to then. However, the DFE still needs channel estimates up to time n. Thus the receiver needs to use a channel prediction module, 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. 1 is meant to show the time succession of steps (1) through (4) below. In Fig. 1 the flow of new information is clockwise, starting from top left, with each of the blocks corresponding to one of the following actions:

1. $\hat{\mathbf{h}}_{n-\Delta} = \mathcal{K}(\hat{\mathbf{h}}_{n-\Delta-1}, u(n-\Delta-1), \hat{\mathbf{s}}_{n-\Delta-M-1}^{n-\Delta-1})$
2. $\hat{\mathbf{h}}_{n-\Delta+1}^{n} = \mathcal{B}\left(\hat{\mathbf{h}}_{n-\Delta}, \mathbf{u}_{n-\Delta}^{n}\right)$
3. $[f_n^{opt}, b_n^{opt}] = \text{design DFE}(\hat{\mathbf{h}}_{n-N}^{n})$
4. $\hat{s}(n-\Delta) = \text{DFE}(f_n^{opt}, b_n^{opt})$

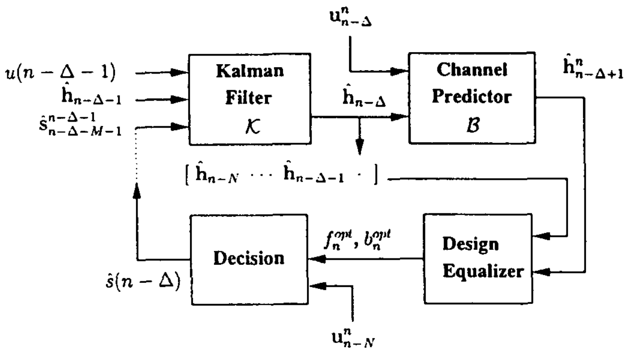

Figure 1: Receiver block diagram.

The iteration starts with the well-known Kalman filter recursions denoted by $\mathcal{K}(\cdot)$, which at time n yield the optimum linear estimator of the channel $\hat{\mathbf{h}}_{n-\Delta}$ as it was at time $n-\Delta$, because it is based on the (assumed reliable) DFE decisions vector $\hat{\mathbf{s}}^{n-\Delta-1}_{n-\Delta-M-1} = [\hat{s}(n-\Delta-1), \cdots, \hat{s}(n-\Delta-M-1)]$, the received $u(n-\Delta-1)$ and the previously estimated channel vector $\hat{\mathbf{h}}_{n-\Delta-1}$. In the second step, $\mathcal{B}(\cdot)$ denotes a predictor that may exploit the additional received values $\mathbf{u}^n_{n-\Delta} = [u(n), \cdots, u(n-\Delta)]$, along with the estimate $\hat{\mathbf{h}}_{n-\Delta}$ to compute the sequence of Δ new predicted channels $\hat{\mathbf{h}}^n_{n-\Delta+1}$.

Those Δ predicted channels, along with the $N+1-\Delta$ most recent channel estimates from the Kalman filter, are used by the DFE design module (see Sect. 3.3) to design the optimum feedforward filter, f^{opt}_n, and the feedback filter, b^{opt}_n of an MMSE DFE. Finally, the newly designed DFE decodes one more symbol $\hat{s}(n-\Delta)$, which is added to the vector of past (assumed reliable) decisions, which will help the Kalman filter make a new channel estimate $\hat{\mathbf{h}}_{n-\Delta+1}$ at the next iteration, taking place at time instant $n+1$. In the following subsections we look at the implementation of each receiver module in greater detail.

3.1 Kalman Filter Tracking

For simplicity, we limit our discussion to the AR(1) channel model, but the extension to higher order AR models is straightforward. Define the $(M+1)$-dimensional vector of transmitted points $\mathbf{s}_n = [s(n) \;\; s(n-1) \;\; \ldots \;\; s(n-M)]$. The channel at time n has a constant non-zero mean $\bar{\mathbf{h}}$ (Ricean fading), and a zero-mean time-varying part $\mathbf{h}_n$, which follows the AR(1) model:

$$\mathbf{h}_{n+1} = \mathbf{F}\mathbf{h}_n + \mathbf{G}\mathbf{q}_n. \qquad (2)$$

At time n, the (zero-mean) received value $u(n)$ is given by:

$$u(n) = \mathbf{s}_n \cdot (\bar{\mathbf{h}} + \mathbf{h}_n) + v(n). \qquad (3)$$

Assuming the matrices $\mathbf{F}$ and $\mathbf{G}$ and the mean channel vector $\bar{\mathbf{h}}$ are known from a preceding training phase, and assuming the vector of the most recent decisions $\hat{\mathbf{s}}_n = \hat{\mathbf{s}}^n_{n-M}$ to be equal to the true $\mathbf{s}_n$, the receiver can use the Kalman filter to track the channel variation $\mathbf{h}_n$, using as observables the $u(n) - \mathbf{s}_n\bar{\mathbf{h}}$. The Kalman filter, operating with a delay Δ is described at time n by the series of equations [7]:

$$\begin{aligned}
\hat{\mathbf{h}}_{n-\Delta} &= \mathbf{F}\hat{\mathbf{h}}_{n-\Delta-1} + \mathbf{K}_{n-\Delta-1}e_{n-\Delta-1} \\
e_{n-\Delta-1} &= u(n-\Delta-1) - \hat{\mathbf{s}}_{n-\Delta-1}(\hat{\mathbf{h}}_{n-\Delta-1} + \bar{\mathbf{h}}) \\
\mathbf{K}_{n-\Delta-1} &= (\mathbf{F}\mathbf{P}_{n-\Delta-1}\hat{\mathbf{s}}^*_{n-\Delta-1})/\mathbf{R}_{e,n-\Delta-1} \\
\mathbf{R}_{e,n-\Delta-1} &= \sigma_v^2 + \hat{\mathbf{s}}_{n-\Delta-1}\mathbf{P}_{n-\Delta-1}\hat{\mathbf{s}}^*_{n-\Delta-1} \\
\mathbf{P}_{n-\Delta} &= \mathbf{F}\mathbf{P}_{n-\Delta-1}\mathbf{F}^* + \mathbf{G}\mathbf{G}^* \\
&\quad -\mathbf{K}_{n-\Delta-1}\mathbf{R}_{e,n-\Delta-1}\mathbf{K}^*_{n-\Delta-1}
\end{aligned}$$

The above Kalman recursions implement the optimum linear estimator for the time-varying part of the channel $\mathbf{h}_{n-\Delta}$. The last reliable decision made by the DFE and used by the Kalman filter at this time is $\hat{s}(n-\Delta-1)$. For matrices $\mathbf{F}$ and $\mathbf{G}$ that are multiples of the identity (produced, for instance, by uncorrelated fading with the same Doppler for all taps) fast algorithms for the Kalman recursions can be pursued (see, e.g., [8]).

3.2 Channel Prediction

At time n the last channel estimate from the Kalman filter is $\hat{\mathbf{h}}_{n-\Delta}$, but the predicted channels $\hat{\mathbf{h}}_{n-\Delta+k}$, for $k = 1, \ldots, \Delta$ are needed for the DFE design. The channel prediction implementation 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:

$$\hat{\mathbf{h}}_n = \hat{\mathbf{h}}_{n-1} = \cdots = \hat{\mathbf{h}}_{n-\Delta} \qquad (4)$$

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

More generally, the optimal linear predictions, given that the channel follows the model of Eq. (2), but ignoring the additional received values $u(n), \ldots, u(n-\Delta+1)$ are:

$$\hat{\mathbf{h}}_n = \mathbf{F}^{\Delta}\hat{\mathbf{h}}_{n-\Delta}, \ldots, \hat{\mathbf{h}}_{n-\Delta+1} = \mathbf{F}\hat{\mathbf{h}}_{n-\Delta} \qquad (5)$$

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

The received values $u(n), \ldots, u(n-\Delta+1)$, which are also available, can be used to improve the prediction for a fast varying channel at high SNR. For example one could pick the channel estimates $\hat{\mathbf{h}}_k, k = 1, \ldots, \Delta$, as the arguments that minimize a weighted regularized least-squares cost $J(\mathbf{h})$, say of the form:

$$\begin{aligned}
J(\mathbf{h}) &= (\mathbf{h} - \mathbf{F}^k\hat{\mathbf{h}}_{n-\Delta})^*\pi_o(\mathbf{h} - \mathbf{F}^k\hat{\mathbf{h}}_{n-\Delta}) \\
&\quad + \left|\, \left|\|\mathbf{h}\|^2 - \alpha\right|^2 - \beta \,\right|^2, \;\; k = 1, \ldots, \Delta. \qquad (6)
\end{aligned}$$

where α and β are suitable constants, depending on the observables $u(n-\Delta+k)$, the mean channel $\bar{\mathbf{h}}$ and the noise variance σ_v^2, and π_o is a positive weight scalar, selected according to how fast the channel varies and the SNR.

3.3 DFE Design

The design of the optimum mmse feedforward and feedback filters f_N^{opt} and b_n^{opt} of lengths $N+1$ and Q respectively, uses the most recent $N+1$ channel tap coefficients (of which some are predicted and some are estimated). Under the assumption of no error propagation, the filter coefficients can be calculated as [3]:

$$b_n^{opt} = \frac{e_1^T R_\delta^{-1}}{e_1^T R_\delta^{-1} e_1} \quad (7)$$

$$f_n^{opt} = b_n^{opt} R_{sy} R_y^{-1} \quad (8)$$

where e_1 is the vector with first entry 1 and all others zero. To form the matrices R_δ^{-1}, R_y and R_{sy}, we use the $(N+1)\times(M+N+1)$ pre-windowed channel matrix C, containing the tap values:

$$C = \begin{bmatrix} (\bar{\mathbf{h}}+\hat{\mathbf{h}}_n)^T & 0 & \cdots & & 0 \\ & & \ddots & \ddots & \\ 0 & & \cdots & 0 & (\bar{\mathbf{h}}+\hat{\mathbf{h}}_{n-N})^T \end{bmatrix} \quad (9)$$

and form R_δ and R_y as:

$$R_\delta = R_s - R_{sy} R_y^{-1} Rys \quad (10)$$

$$R_y = C R_s C^* + \sigma_v^2 I_{N+1} \quad (11)$$

where R_s is the autocorrelation matrix of the input constellation points (typically $R_s = I_{N+1}$), and σ_v^2 is the variance of the zero-mean white noise. R_{sy} is a banded matrix of dimensions $(Q+1)\times(N+1)$ obtained from C^*, starting Δ rows after the first, where Δ is the decision delay of the DFE, and C^* denotes conjugate transpose of C in (9).

4 Simulation Results

The simulation results examine a 2-tap fading channel, with mean value $(1+j)\ [1\ \ 2.5]^T$ and Doppler rate $f_d T = 0.01$. The transmitted constellation is 4-PSK. We assumed that the channel variation follows an AR(1) model, designed to provide the best first-order AR approximation to a Ricean channel with this Doppler rate.

This is a maximum phase channel, and was chosen as an example of a short channel where decisions made with delay $\Delta > 0$ offer a significant performance improvement. Fig. 2 plots the performance of the MMSE DFE over this channel for different Δ as a function of SNR. The feedforward and feedback filters of the DFE have 7 and 1 taps, respectively. Also for this simulation we assumed that the DFE coefficients are calculated with perfect channel information.

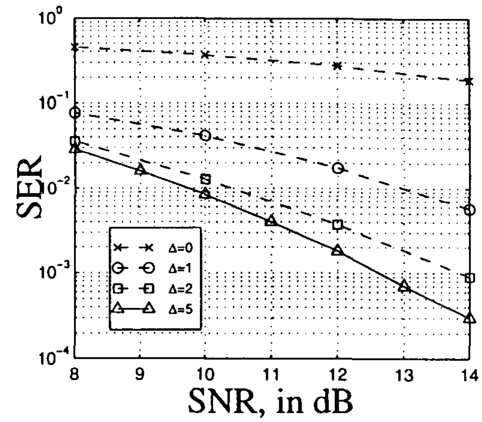

Figure 2: Performance of DFE for different values of Δ, assuming perfect channel information.

For known tap coefficients the decision delay that results in the minimum MMSE can be calculated analytically [3]. Generally though, using a decision delay larger than the optimum does not deteriorate the DFE performance given, of course, a sufficient feedforward filter length. For the rest of the simulations we used $\Delta = 5$, which is larger than required for this channel. The reason is that we wanted to have a large prediction interval to compare different prediction approaches. Also, we assume that we have perfect channel information at the beginning of operating intervals of 5000 symbols, during which the Kalman filter is used to track the channel.

Figures 3 and 4 show the simulation results for a baud-spaced DFE and a linear equalizer respectively, both using decision delay $\Delta = 5$. The solid line represents the proposed system performance, using Eq. (5) for prediction. The results indicate that our Symbol-Error Rate (SER) performance is better than the performance the LMS and RLS adaptive versions of the DFE or the linear equalizer can provide. Comparison of Fig. 3 and Fig. 4 shows that the DFE performs better than the linear equalizer, when both use decision delay $\Delta = 5$. However, a finite-length DFE does not always outperform a finite linear equalizer, as shown in [9].

Fig. 3 also plots the system performance assuming perfect channel information. Note that this curve is very close to the solid line that employs Kalman estimated tap values. Fig. 5 exhibits the tracking performance of the Kalman-based algorithm for this system at SNR=10 dB, which indeed is very good.

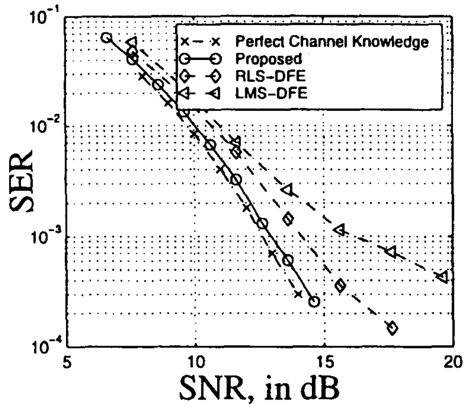

Figure 3: The DFE with Kalman-based tracking, and with RLS and LMS.

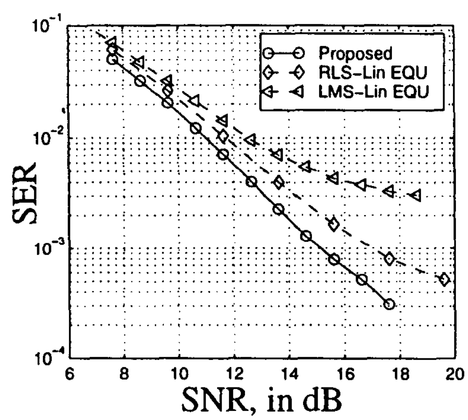

Figure 4: 8 tap Linear Equalizer with Kalman-based tracking, RLS and LMS.

5 Conclusions

This paper proposed a receiver structure to track and equalize a frequency selective fading channel. A Kalman filter was used for tracking the channel. We exploited the fact that allowing for a delay Δ in the DFE or linear equalizer before making a decision, achieves better equalization. The time gap between channel estimates produced by the Kalman filter and those needed for the DFE adaptation was bridged by using a 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.

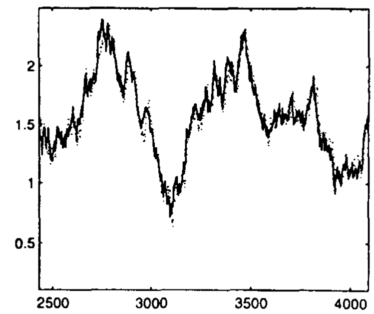

Figure 5: Tracking performance of the Kalman-based algorithm, plotted for the real part of the first tap. The solid line represents the true channel tap trajectory, and the dotted line the estimated trajectory. SNR=10 dB.

References

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

[2] 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.

[3] N. Al-Dhahir and J. Cioffi. MMSE Decision-Feedback Equalizers: Finite-Length Results. *IEEE Trans. on Inform. Theory*, 41(4):961–975, July 1995.

[4] 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.

[5] 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.

[6] W. C. Y. Lee. *Mobile Communications Engineering*. McGraw Hill, NY, 1982.

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

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

[9] W. Shi and R. D. Wesel. When the best decision-feedback equalizer is a linear equalizer. *In Thirty-Sixth Annual Allerton Conference*, pages 338–9, Sept. 1998.