ESTIMATION AND EQUALIZATION OF TIME-SELECTIVE FADING CHANNELS

G. Tong Zhou[1], *Yongsub Kim*[2], *Georgios B. Giannakis*[3]

[1,2]School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0250, USA.
[3] Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455, USA.

ABSTRACT

Time-selective fading effects can occur in communication systems due to channel variations, carrier frequency offset, and oscillator's phase noise. In this paper, we model the multiplicative channel fading effect as an autoregressive process and the objective is to recover the original transmitted symbols. Filterbanks are used here to precode the transmitted symbols and second-order statistics of the output data are computed which in turn, yield the correlation information of the time-varying channel. Next, we employ the Kalman filter to track channel variations and recover the transmitted symbols in a block-by-block fashion. For the Rayleigh fading case, we also take advantage of the additional information provided by pilot symbol assisted modulation (PSAM) to resolve channel ambiguities. Computer simulations are provided to illustrate the concepts.

1. INTRODUCTION

Fading is a major impairment in communication systems because it can cause the instantaneous channel signal-to-noise ratio (SNR) to suddenly drop. Fading occurs due to channel variations, carrier frequency offset, oscillator's phase noise, etc. An example is given in [8] which illustrates motion induced fading effects: consider a 900 MHz carrier in an IS-54 system and a mobile moving at 120 km/h. The minimum time between the two fading nulls is 5 ms, which is even shorter than the proposed burst length of 6.7 ms. Changing characteristics of the propagation media can also cause fading. For example, in the HF band (3-30 MHz), the radio channel is randomly time varying because the ions in the ionospheric layers are in constant motion [10, p. 758].

Denote by $s(n)$ the i.i.d. information sequence, $v(n)$ the additive noise, and $x(n)$ the channel output. A linear, discrete-time, baseband equivalent channel model can be described by

$$x(n) = \sum_{l=0}^{L} h(n;l)\ s(n-l) + v(n), \tag{1}$$

where $h(n;l)$ is the channel impulse response at time n and tap delay l and its frequency domain counterpart is the scattering function $\mathcal{H}(n;\omega)$. When the transmitted signal bandwidth is much smaller than the coherence bandwidth of the channel, the channel is frequency-nonselective (flat); i.e., $\mathcal{H}(n;\omega) = \mathcal{H}(n;0)$, $\forall\omega$ [10, p. 770]. In the time domain this implies $h(n;l) = h(n)\,\delta(l)$ and hence (1) becomes

$$x(n) = h(n)s(n) + v(n). \tag{2}$$

The time-selective nature is clearly evident from (2) where the channel $h(n)$ acts as multiplicative distortion to the information sequence $s(n)$.

Given the received data $x(n)$, our goal is to recover $s(n)$. Without imposing structure on the channel variations, the problem is ill-posed because with every incoming data $x(n)$, there is one more unknown $h(n)$ in addition to $s(n)$. When $h(n)$ (or $h(n;l)$ in general) is viewed as deterministic, basis function expansion techniques have been explored (e.g., [14]). When $h(n)$ (or $h(n;l)$) appears random, first order autoregressive (AR(1)) or the Gauss-Markov model has demonstrated great potential [4, 13]. For the remainder of the paper, we will focus our attention on frequency-flat, random time-selective fading channels.

In many applications such as scattering function characterization [1], network performance evaluation [3], adaptive signaling [5], and vehicle speed estimation [9], it is useful to know the channel correlation information. For the model in (2), we will show in Section 2 that the correlation function of $h(n)$ cannot be recovered from the output only data. In [13], channel correlation function is estimated from training data by exploring a joint fourth-order statistic between the input and the output. We will show in Section 3 that by employing a simple precoding scheme, the channel correlation function can be recovered from the output second-order statistic and training is not necessary.

We model the random channel as AR here because it leads to a state-space formulation and hence is amenable to the Kalman filter analysis. With precoding, it is feasible to update the channel correlation coefficients and hence the AR parameters throughout the operation mode, which is advantageous as compared to [13].

When the channel mean $m_h = E[h(n)] > 0$, we have the Rician fading model, which accounts for the presence of a line-of-sight (LOS) component. When the LOS component is absent or has negligible power, the Rician model reduces into the Rayleigh one [15]. We will see in Section 4 that for the Rayleigh fading channel, a sign ambiguity (or in general, phase rotation) problem exists and hence reliable decoding cannot be achieved without additional information. For the Rayleigh fading channel, we propose to periodically insert pilot symbols into the transmitted data stream – an approach called pilot symbol assisted modulation (PSAM) (see e.g., [2], [6]). With PSAM, a loss of information rate is traded for improved probability of error. Simulation examples are presented in Section 5 to illustrate the performance of the algorithms.

2. CHANNEL IDENTIFIABILITY

We make the following assumptions regarding the model in (2): the input $s(n)$ is zero-mean, i.i.d., symmetrically distributed, and comes from a known alphabet; the channel $h(n)$ is stationary; and the additive noise $v(n)$ is zero-mean with a known covariance function $c_{2v}(k) = E[v(n)v^*(n+k)]$. We further assume that $s(n)$, $h(n)$, and $v(n)$ are mutually independent.

0-7803-5700-0/99/$10.00©1999 IEEE

2.1. Identification of the channel mean

We infer from (2) that the mean of $x(n)$ is zero regardless of whether $h(n)$ has zero mean or not. As we will see in the next subsection, m_h cannot be found from the output second-order statistic either. However, knowing m_h is key to classifying the channel as Rayleigh or Rician. It is therefore worthwhile to utilize a small amount of training data to gain information about m_h. We obtain the following relationship based on (2):

$$\begin{aligned} E[x(n)s^*(n)] &= E[h(n)]\ E[|s(n)|^2] + E[v(n)]\ E[s^*(n)] \\ &= m_h\ \sigma_s^2, \end{aligned}$$

where σ_s^2 is the variance of $s(n)$. Therefore, relying on the input and output data, an estimate of the channel mean can be obtained as

$$\hat{m}_h = \frac{\sum_{n=0}^{N-1} x(n)s^*(n)}{\sum_{n=0}^{N-1} |s(n)|^2}. \tag{3}$$

Since this is a first-order sample statistic, a small amount of data points (usually < 100) can yield good results.

2.2. Identification of the channel correlation

Because we allow $h(n)$ to have non-zero mean, the autocovariance function $c_{2h}(k)$ and the autocorrelation function $m_{2h}(k)$ are generally different:

$$\begin{aligned} c_{2h}(k) &= E\{[h(n) - m_h][h^*(n+k) - m_h^*]\}, \\ m_{2h}(k) &= E[h(n)h^*(n+k)] = c_{2h}(k) + |m_h|^2. \end{aligned}$$

If the training data $\{x(n), s(n)\}_{n=0}^{N-1}$ are used to estimate the mean m_h as in (3), the same data can be used to estimate the channel autocorrelation function as well:

$$\begin{aligned} \hat{h}(n) &= \frac{x(n)}{s(n)}, \\ \hat{m}_{2h}(k) &= \frac{1}{N}\sum_{n=0}^{N-1} \hat{h}(n)\hat{h}^*(n+k). \end{aligned}$$

Afterwards, the covariance function can be found:

$$\hat{c}_{2h}(k) = \hat{m}_{2h}(k) - |\hat{m}_h|^2. \tag{4}$$

If $h(n)$ is modeled as AR, the AR parameters can be solved from $\hat{c}_{2h}(k)$ using the Yule-Walker equations. For example, for an AR(1) channel model

$$h(n) = a_1 h(n-1) + w(n), \tag{5}$$

the AR(1) coefficient can be found from

$$a_1 = \frac{c_{2h}^*(1)}{c_{2h}(0)}. \tag{6}$$

However, it usually takes an order of magnitude more data for $\hat{c}_{2h}(k)$ to have comparable performance as $\hat{m}_h$. This also motivates us to search for methods which can estimate the channel autocorrelation function from output data only so the $\hat{c}_{2h}(k)$ estimate can be updated throughout the operation mode.

Let us start with the second-order statistic of the output:

$$c_{2x}(k) = E[x(n)x^*(n+k)] = c_{2s}(k)\ m_{2h}(k) + c_{2v}(k). \tag{7}$$

Since $s(n)$ is i.i.d., we have $c_{2s}(k) = \sigma_s^2\ \delta(k)$. Therefore, only $m_{2h}(0) = |m_h|^2 + \sigma_s^2$ can be estimated from (7).

Recall that all odd-order statistics of the symmetrically distributed $s(n)$ vanish. We consider next the fourth-order cumulant of $x(n)$. For simplicity, let us consider the case where $h(n)$ is zero-mean, and $h(n)$ and $v(n)$ are both Gaussian (so their fourth-order cumulants are zero). When $x(n)$ is real valued, it is shown in [12] that only the magnitude of the channel covariance function can be recovered as

$$|c_{2h}(k)| = \begin{cases} \sqrt{c_{4x}(0,0,0)/[3(\gamma_{4s} + 2\sigma_s^4)]}, & k = 0, \\ \sqrt{c_{4x}(k,0,k)/(2\sigma_s^4)}, & k \neq 0, \end{cases} \tag{8}$$

where γ_{4s} is the kurtosis of $s(n)$, and the fourth-order cumulant of the zero-mean $x(n)$ is defined as

$$\begin{aligned} & c_{4x}(k_1,k_2,k_3) \\ =\ & \text{cum}\{x(n), x(n+k_1), x(n+k_3), x(n+k_4)\} \\ =\ & m_{4x}(k_1,k_2,k_3) - c_{2x}(k_1)\ c_{2x}(k_2-k_3) \\ -\ & c_{2x}(k_2)\ c_{2x}(k_3-k_1) - c_{2x}(k_3)\ c_{2x}(k_1-k_2). \end{aligned}$$

Note that even though $x(n)$ has zero-mean, the fourth-order cumulant $c_{4x}(k_1,k_2,k_3)$ and the fourth-order moment $m_{4x}(k_1,k_2,k_3) = E[x(n)x(n+k_1)x(n+k_3)x(n+k_4)]$ generally differ.

When $x(n)$ is complex valued, we define its fourth-order cumulant as

$$\begin{aligned} & c_{4x}(k_1,k_2,k_3) \\ =\ & \text{cum}\{x(n), x(n+k_1), x^*(n+k_3), x^*(n+k_4)\}. \end{aligned}$$

When $s(n)$ and $h(n)$ are zero-mean, circular symmetrically distributed (i.e., real and imaginary parts are mutually independent but share the same distribution), and $h(n)$ and $v(n)$ are Gaussian, the formula in (8) needs to be modified to

$$|c_{2h}(k)| = \begin{cases} \sqrt{c_{4x}(0,0,0)/[2(\gamma_{4s} + \sigma_s^4)]}, & k = 0, \\ \sqrt{c_{4x}(k,0,k)/\sigma_s^4}, & k \neq 0, \end{cases} \tag{9}$$

In both the real and complex cases, $c_{2h}(k)$ itself cannot be recovered from the 1st- through 4th-order statistics of the data. This fact motivates us to seek out diversity techniques which allow blind estimation of the channel correlation function.

3. CHANNEL ESTIMATION AND EQUALIZATION WITH PRECODING

A general framework is developed in [11] which introduces redundancy into the data by placing filterbank precoders at the transmitter. For our problem here, let us first define $\mathbf{s}(n)$ as a length-P block of $s(n)$:

$$\mathbf{s}(n) = \begin{bmatrix} s(nP) \\ s(nP+1) \\ \vdots \\ s(nP+P-1) \end{bmatrix} \tag{10}$$

Pre-multiply $\mathbf{s}(n)$ by a $P \times P$ matrix $\mathbf{F}$ to obtain

$$\mathbf{u}(n) = \mathbf{F}\ \mathbf{s}(n). \tag{11}$$

We therefore map a length-P block of $s(n)$ into a length-P block of $u(n)$ and for the mapping to be unique, $\mathbf{F}$ has to be non-singular. It is not difficult to verify that $u(n)$ is now cyclostationary whose time-varying statistics are periodic in n with period P.

3.1. Recovering channel correlation information

Since $s(n)$ is i.i.d., the covariance matrix of $\mathbf{u}(n)$ is

$$\mathbf{C}_u = E[\mathbf{u}(n)\ \mathbf{u}^H(n)] = \sigma_s^2\ \mathbf{F}\mathbf{F}^H \tag{12}$$

The purpose of precoding is to introduce color into $u(n)$ which subsequently enables the estimation of $m_{2h}(k)$. It is evident from (12) that $\mathbf{F}$ cannot be unitary for $u(n)$ to be correlated.

Similar to (10), we define length-P vectors $\mathbf{x}(n)$, $\mathbf{h}(n)$ and $\mathbf{v}(n)$. Using these vector notations, equation (2) can be rewritten as

$$\mathbf{x}(n) = \mathbf{U}(n)\ \mathbf{h}(n) + \mathbf{v}(n), \tag{13}$$

where $\mathbf{U}(n) = \text{diag}(\mathbf{u}(n))$.

Denote by $\mathbf{C}_x$, $\mathbf{C}_v$, and $\mathbf{M}_h$ the covariance matrices of $\mathbf{x}(n)$, $\mathbf{v}(n)$, and the correlation matrix of $\mathbf{h}(n)$, respectively. They are related through the following equation:

$$\mathbf{C}_x = \mathbf{C}_u \ . \ \mathbf{M}_h + \mathbf{C}_v, \tag{14}$$

where the dot between $\mathbf{C}_u$ and $\mathbf{C}_h$ indicates element-wise multiplication of the two matrices. $\mathbf{C}_u$ can be pre-calculated from (12), and $\mathbf{C}_v$ is assumed known. From the data $\mathbf{x}(n)$, we first estimate $\mathbf{C}_x$ by averaging the instantaneous $\mathbf{x}(n)\mathbf{x}^H(n)$ products over blocks, and then recover $\mathbf{M}_h$ from (14). Since $h(n)$ is stationary, $\mathbf{M}_h$ is Hermitian and Toeplitz whose first row is $\{m_{2h}(k)\}_{k=0}^{P-1}$.

When modeling $h(n)$ as an AR(p) process, we need $\{c_{2h}(k)\}_{k=0}^{p}$ to solve for the AR parameters via the Yule-Walker equations. Therefore we must have the block length P greater than the channel AR order p; i.e., $P > p$. It has been shown by many authors (e.g., [4, 13]) that an AR(1) model is sufficient to capture slow channel variations. Therefore for many channels, the block size can be as small as 2 and is quite manageable.

Because $x(n)$ is cyclostationary, the covariance matrix $\mathbf{C}_x$ is no longer Toeplitz. For $m_{2h}(k)$ to be identifiable, we only need any one of the elements along the kth diagonal of $\mathbf{F}\mathbf{F}^H$ to be non-zero, and this has to be true for all $0 \le k \le p$.

Summarizing, our requirements for $\mathbf{F}$ are: (i) $\mathbf{F}$ is non-singular; (ii) $\mathbf{F}$ is non-unitary. If $h(n)$ is AR(p), then (iii) the dimension of $\mathbf{F}$ must satisfy $P > p$; and (iv) $\mathbf{F}\mathbf{F}^H$ cannot have an all-zero kth diagonal for $0 \le k \le p$.

Assumptions (i)-(iv) are usually met, even by a randomly generated $P \times P$ matrix so they are fairly relaxed conditions. If certain optimality criteria are imposed, then $\mathbf{F}$ has to be specially designed – a topic of our future research.

3.2. Channel tracking and symbol decoding

The vector version of the AR(1) recursion in (5) is

$$\mathbf{h}(n) = \mathbf{A}\mathbf{h}(n-1) + B\mathbf{w}(n), \tag{15}$$

where $\mathbf{A} = a_1^P \mathbf{I}$ is $P \times P$,

$$\mathbf{B} = \begin{bmatrix} a_1^{P-1} & \dots & a_1 & 1 & 0 & \dots & 0 \\ 0 & a_1^{P-1} & \dots & a_1 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & \dots & a_1^2 & a_1 & 1 \end{bmatrix},$$

is $P \times (2P-1)$, and $\mathbf{w}(n)$ is a $(2P-1) \times 1$ vector consisting of $w(nP+k)$ for $-(P-1) \le k \le (P-1)$.

Combining (15) and (13), we see that $\mathbf{h}(n)$ is the state, (15) is the state equation and (13) is the measurement equation. We then employ the Kalman filter to track the variations in $\mathbf{h}(n)$. The challenge though, lies in that $\mathbf{U}(n)$ is unknown and needs to be estimated from the data as well.

Formulas for the vector Kalman filtering can be found e.g., in [7, pp. 446-447]. For every block of $\hat{\mathbf{u}}(n)$ estimates, we first obtain symbol estimates (c.f. (11)):

$$\hat{\mathbf{s}}(n) = \mathbf{F}^{-1}\ \hat{\mathbf{u}}(n). \tag{16}$$

We then take advantage of the finite alphabet information and obtain quantized symbol estimates $\bar{\mathbf{s}}(n)$.

We now summarize our block-based Kalman filtering algorithm. First, we initialize $\hat{\mathbf{h}}(0)$ by the channel mean estimate $\hat{m}_h$, and then use the Kalman filter to predict the next value $\hat{\mathbf{h}}(1)$. For $n \ge 1$, we follow these steps:

Step 1: Form the tentative $\mathbf{u}(n)$ estimate:

$$\hat{\mathbf{u}}(n) = \hat{\mathbf{H}}^{-1}(n)\ \mathbf{x}(n), \qquad \hat{\mathbf{H}}(n) = \text{diag}(\hat{\mathbf{h}}(n))$$

Step 2: Form $\hat{\mathbf{s}}(n) = \mathbf{F}^{-1}\ \hat{\mathbf{u}}(n)$.
Step 3: Quantize $\hat{\mathbf{s}}(n)$ to obtain $\bar{\mathbf{s}}(n)$.
Step 4: Update $\bar{\mathbf{u}}(n) = \mathbf{F}\ \bar{\mathbf{s}}(n)$.
Step 5: Run the Kalman filter with $\mathbf{U}(n) = \text{diag}(\bar{\mathbf{u}}(n))$ to predict $\hat{\mathbf{h}}(n+1)$.

Our simulations have shown that this is a fairly robust routine. Once the $\hat{\mathbf{h}}(n)$ estimate is "on-track", the algorithm runs well without retraining.

Instead of the $P \times P$ matrix $\mathbf{F}$, we can introduce correlation into $u(n)$ by pre-multiplying $\mathbf{s}(n)$ with a vector $\mathbf{f}^T = [f_0\ f_1\ \dots\ f_{P-1}]$ which corresponds to $u(n) = f(n) * s(n)$. In this case, we run the scalar version of the Kalman filter on the state equation (5) and measurement equation (2). A disadvantage is that decoding is not as easy as (16) and a decision directed equalizer is employed.

4. PILOT SYMBOL ASSISTED MODULATION FOR RAYLEIGH CHANNELS

In wireless applications when no LOS is present, the channel $h(n)$ has a zero (or close to zero) mean and is called Rayleigh fading. In this case, the tracking algorithm proposed in Section 3.2 fails even with sufficient training. This is because given any channel-input product $y(n) = h(n)s(n)$, $y(n)$ can be written as $y(n) = [-h(n)][-s(n)]$, or in general, $y(n) = [h(n)\ e^{j\theta}][s(n)\ e^{-j\theta}]$ as well. Therefore a sign ambiguity (or in general, phase rotation) problem exists. A remedy is to employ pilot symbol assisted modulation (PSAM) where a known pilot symbol is periodically inserted into the data stream [2], [6]. The channel value at the time of pilot symbol transmission can be fairly accurately estimated (any error is only due to the additive noise). If the channel is sufficiently lowpass (i.e., slowly time-varying), then the next few channel values are also likely to be tracked accurately. Obviously, more frequent pilot symbol transmission results in better tracking capability of the Kalman filter and lower error rates in the decoded symbols, but the price paid is some loss in information rate. The frequency of pilot symbol transmission is directly influenced by the bandwidth of the channel, which again shows the importance of the channel correlation (and hence power spectrum) estimation problem addressed in subsections 2.2 and 3.1.

5. SIMULATIONS

Extensive computer simulations were carried out to investigate the performance of the proposed algorithms. For illustration purposes, we only show results on BPSK signals and $\mathbf{F} = \mathbf{I}$ (no precoding). We model the channel as AR(1) which has been shown by many authors as suitable for slowly time varying channels. Both the additive noise $v(n)$ in (2) and the driving noise $w(n)$ in (5) were i.i.d. Gaussian with variance σ_v^2 and σ_w^2, respectively.

The initial training period used 200 samples of input and output data to estimate the mean m_h according to (3) and

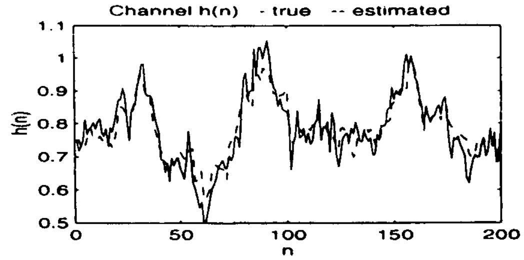

Figure 1. Example 1 (Rician fading). The true channel $h(n)$ and its estimate $\hat{h}(n)$, at SNR = 20 dB.

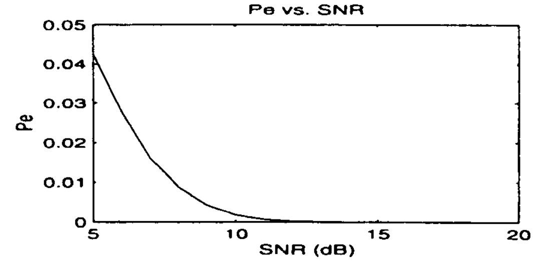

Figure 2. Example 1 (Rician fading). Probability of error in $\hat{s}(n)$.

the AR(1) parameter a_1 as in (6). Afterwards, we measured the probability of error using 10,000 samples without retraining. For a given SNR, 100 independent realizations were used to obtain the average probability of error (P_e). In this paper, SNR is defined as

$$\text{SNR} = \frac{\sigma_s^2 m_{2h}(0)}{\sigma_v^2} = \frac{\sigma_s^2 \left(\sigma_h^2 + |m_h|^2\right)}{\sigma_v^2}.$$

Since the Rician and Rayleigh fading channels require different treatments, we give one example for each.

Example 1: Rician ($m_h \neq 0$) fading channel. The channel $h(n)$ is modeled as AR(1) (c.f. equation (5)) with $a_1 = 0.9$. Its driving noise $w(n)$ has variance $\sigma_w^2 = 0.002$, and the mean of $h(n)$ is $m_h = 0.8$. The input $s(n)$ is BPSK with $\sigma_s^2 = 1$. Figure 1 shows the true $h(n)$ (solid line) and its estimate $\hat{h}(n)$ (dashed line) for the first 200 symbols in the operation mode at SNR = 20 dB. The Kalman filter is seen to yield excellent tracking results. Next, we varied SNR between 5 and 20 dB, and obtained the probability of error curve as shown in Figure 2.

Example 2: Rayleigh ($m_h = 0$) fading channel. The parameters are the same as in the previous example except that now $m_h = 0$, $\sigma_w^2 = 0.1$, and SNR is 30 dB.

As we explained in Section 4, without additional information, a sign ambiguity exists for the Rayleigh channel. Figure 3 illustrates this phenomenon. The top plot shows the true channel $h(n)$ (solid line) and its estimate $\hat{h}(n)$ (dashed line). We see that sometimes $\hat{h}(n)$ follows $h(n)$ quite well but some other times, $\hat{h}(n)$ is almost a mirror image of $h(n)$. Figure 3b shows the locations where errors had occurred in the decoded symbols. Without PSAM, the probability of error is very high, close to the guessing probability of 0.5.

PSAM offers a remedy. In our example, a known pilot symbol was transmitted once every 12 samples, resulting in approximately 8% loss of information rate. But as we see in Figure 4, both the channel tracking capability and the probability of error were improved.

Acknowledgment: This work was supported in part by the National Science Foundation grant MIP 9703312.

REFERENCES

[1] H. Artes, G. Matz, and F. Hlawatsch, "An unbiased scattering function estimator for fast time-varying channels," *Proc. IEEE Signal Processing Workshop on Advances in Wireless Communications*, pp. 411-414, Annapolis, MD, May 1999.

[2] J. K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels," *IEEE Trans. on Vehicular Technology*, vol. 40, no. 4, pp. 686-693, 1991.

[3] A. Chockalingam, M. Zorzi, L. B. Milstein, and P. Venkataram, "Performance of a wireless protocol on correlated Rayleigh-fading channels with capture," *IEEE Trans. on Communications*, vol. 46, no. 5, pp. 644-655, 1998.

[4] A. W. Fuxjaeger and R. A. Iltis, "Adaptive parameter estimation using parallel Kalman filtering for spread spectrum code and Doppler tracking," *IEEE Trans. on Communications*, vol. 42, no 6, pp. 2227-2230, June 1994.

[5] D. L. Goeckel, "Adaptive coding for time-varying channels using outdated fading estimates," *IEEE Trans. on Communications*, vol. 47, no. 6, pp. 844-855, 1999.

[6] J.-C. Guey, M. P. Fitz, M. R. Bell, and W.-Yi. Kuo, "Signal design for transmitter diversity wireless communication systems over Rayleigh fading channels," *IEEE Trans. on Communications*, vol. 47, no. 4, pp. 527-537, 1999.

[7] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice Hall, 1993.

[8] H.-N. Lee and G. J. Pottie, "Fast adaptive equalization/diversity combining for time-varying dispersive channels," *IEEE Trans. on Communications*, vol. 46, no. 9, pp. 1146-1162, 1998.

[9] R. Narasimhan and D. C. Cox, "Speed estimation in wireless systems using wavelets," *IEEE Trans. on Communications*, vol. 47, no. 9, pp. 1357-1364, 1999.

[10] J. G. Proakis, *Digital Communications*, 3rd edition, McGraw-Hill, 1995.

[11] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant filterbank precoders and equalizers. Part I: unification and optimal designs," *IEEE Trans. on Signal Processing*, vol. 47, no. 7, pp. 1988-2005, 1999.

[12] A. Swami, "Multiplicative noise models: parameter estimation using cumulants," *Signal Processing*, vol. 36, pp. 355-373, 1994.

[13] M. K. Tsatsanis, G. B. Giannakis, and G. Zhou, "Estimation and equalization of fading channels with random coefficients," *Signal Processing*, vol. 53, no. 2/3, pp. 211-229, 1996.

[14] M. K. Tsatsanis and G. B. Giannakis, "Equalization of rapidly fading channels: self-recovering methods," *IEEE Trans. on Communications*, vol. 44, no. 5, pp. 619-630, 1996.

[15] M. Zorzi, R. Rao, and L. B. Milstein, "Error statistics in data transmission over fading channels," *IEEE Trans. on Communications*, vol. 46, no. 11, pp. 1468-1477, 1998.

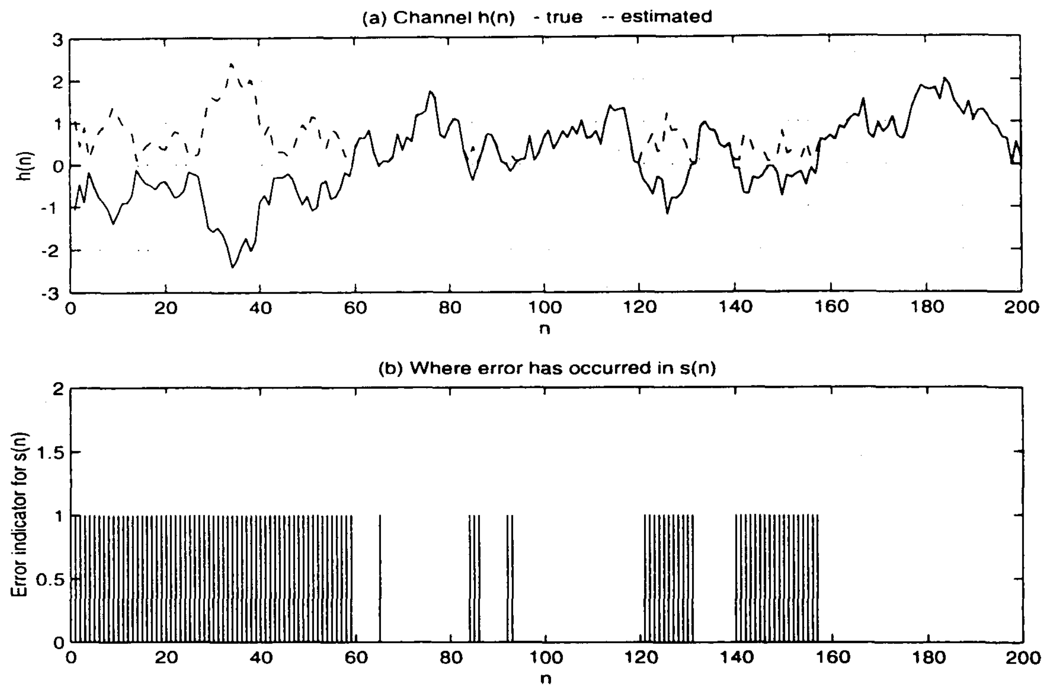

Figure 3. Example 2 (Rayleigh fading). Without PSAM.

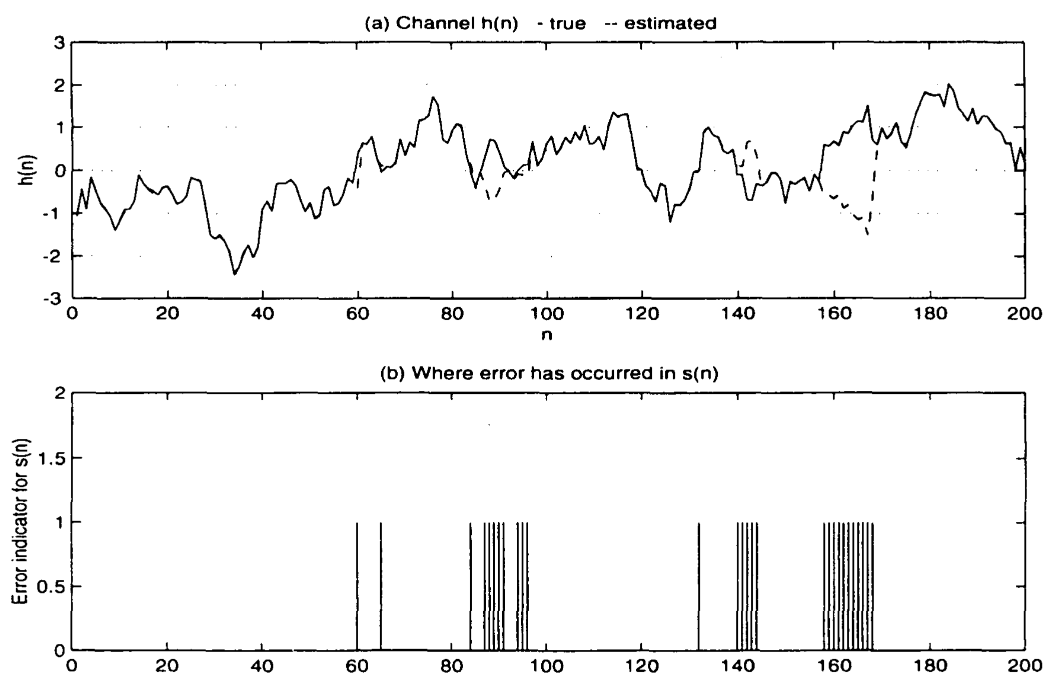

Figure 4. Example 2 (Rayleigh fading). PSAM is employed at the rate of 1/12.