Channel Estimation and Equalization for Spread-Response Precoding Systems in Fading Environments

by

J. Nicholas Laneman

B.S.E.E., Washington University (1995)
B.S.C.S., Washington University (1995)

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Master of Science

at the

Massachusetts Institute of Technology

June 1997

Author: _____

Department of Electrical Engineering and Computer Science
May 23, 1997

Certified by: _____

Gregory W. Wornell
Associate Professor of Electrical Engineering
Thesis Supervisor

Accepted by: _____

Arthur C. Smith
Chairman, Departmental Committee on Graduate Students

Channel Estimation and Equalization for Spread-Response Precoding Systems in Fading Environments

by

J. Nicholas Laneman

Submitted to the Department of Electrical Engineering and Computer Science
on May 23, 1997, in partial fulfillment of the
requirements for the degree of
Master of Science

Abstract

In the wireless communications setting, channel coding schemes must combat not only additive noise, but also distortions such as intersymbol interference and signal fading caused by multipath propagation. Spread-response precoding is a newly proposed method for communicating over such channels, which is attractive because it offers good performance while having efficient implementations and requiring no additional average power or bandwidth. Furthermore, these systems essentially transform the intersymbol interference into a more benign form of uncorrelated, additive noise at the receiver. In this thesis, we examine several theoretical and practical issues relating to fading channel estimation and equalization for use in these spread-response precoding systems.

We first compare linear fading channel equalizers designed according to two seemingly different criteria of optimality, namely, the maximum signal-to-noise ratio (SNR) criterion and the minimum mean-square error (MMSE) criterion. Initially, we assume the fading channel impulse response is known at the receiver, and we show that, while the maximum SNR equalizer corresponds exactly to the MMSE equalizer for the case of a frequency nonselective channel, for the case of a slowly-varying, frequency selective channel this result holds only approximately. We examine the more general frequency selective case in some detail by comparing the frequency responses of the two equalizers and the bit-error probability of systems employing them.

We then consider the case in which the fading channel coefficients are unknown at the receiver, but some form of direct channel measurement is obtained through the use of a pilot-tone. From these observations, we develop techniques based on the Kalman filter and the Expectation-Maximization algorithm for estimating and tracking the fading channel parameters. We examine how these estimates alter the form of the optimal equalizers, and assess their impact on the performance of spread-response precoding systems through analysis and simulation.

Finally, we give some preliminary results from an implementation of these ideas within an indoor wireless testbed. We demonstrate the use of our channel estimation algorithms in characterizing the wireless channel found in the laboratory, and we compare transmissions over this channel with and without spread-response precoding.

Thesis Supervisor: Gregory W. Wornell

Title: Associate Professor of Electrical Engineering

Acknowledgments

I wish to express my sincere appreciation to my advisor, Prof. Gregory Wornell, for his guidance, encouragement, and support throughout my graduate studies in general and this thesis research in particular. For creating a productive environment in which to work, and nurturing me as a beginning researcher, I am truly grateful to him, and I look forward to continuing my studies under his supervision. I am also very indebted to two senior members of the Digital Signal Processing Group, Haralabos Papadopoulos and Dr. Steven Isabelle, for their assistance in developing some of the ideas in this thesis and their many thought provoking discussions about related research topics.

Thanks go also to Alan Seefeldt, Richard Barron, and Matthew Secor of the Digital Signal Processing Group for sharing their random ideas, research-related or otherwise; to Walter Sun, Andrew Kim, and Dewey Tucker for our late-night study sessions and periodic lunch breaks; and finally to Jeffrey Karam, Brian Hearing, Santanu Das, and David Cuneo for joining me in stress-relieving athletic pursuits and other escapes from work.

Lastly, I wish to offer a very special word of tribute to my family, Mom, Dad, Jenny, and Jill, for their continuous love and caring. Mom and Dad, you are my oldest teachers, always sharing your wisdom and supporting my endeavors unconditionally. Your sacrifices through the years have given me every opportunity to succeed, and as but a small token of my gratitude, I dedicate this work to you.

Contents

List of Figures

Chapter 1

Introduction

A central issue in the wireless communications setting is the signal fading and intersymbol interference (ISI) introduced by the channel [1]. Due to multipath propagation, many copies of the transmitted signal arrive at the receiver antenna, each with a given attenuation and delay. When these delays are separated by more than the symbol duration, intersymbol interference (ISI) results. Furthermore, when the receiver antenna is set in motion, as is usually the case in such applications such as cellular telephony, signal fading results as the received power level fluctuates because the multipath components add constructively or destructively. As a result of these distortions, wireless fading channels exhibit dramatically poorer bit-error performance than traditional additive white Gaussian noise channels when using uncoded transmissions.

Recent work [2], [3] has suggested a technique called *spread-response precoding* for combating ISI and signal fading found in wireless links. The idea behind this sort of precoding is to temporally distribute the energy of each symbol, and effectively recombine these transmissions, to achieve the average effect of the channel rather than the instantaneous fade. A key element of the receiver is an equalizer that partially compensates for the instantaneous ISI and fading effects of the channel, and constructing this equalizer requires some knowledge of the fading channel impulse response. In a real wireless system, especially when the receiver is mobile, knowledge of the fading characteristics must be obtained through adaptive channel estimation, either blind or trained, based on a model for the channel. The impact of estimation errors must then be incorporated into the performance analysis of the communication system. Different techniques for estimating the channel must be compared

based on their bit-error probability, robustness, and computational complexity.

In this thesis, we address several theoretical and practical topics relating to fading channel estimation and equalization for use in spread-response precoding systems. After summarizing the relevant background material on fading channels and spread-response precoding systems in Chapter 2, we examine in Chapter 3 two different criteria for obtaining optimal linear equalizers that rely on complete knowledge of the fading channel impulse response. Chapter 4 develops a particular method for obtaining fading channel estimates based on observations of the fading in white noise (a pilot-tone), and Chapter 5 integrates these channel estimates into the equalizers found in Chapter 3 and examines the bit-error probability of the complete system. In Chapter 6, we present the results of transmitting data over a real wireless channel using a preliminary implementation of the ideas outlined in this thesis. Finally, in Chapter 7, we make some concluding remarks and suggests areas of further research.

Chapter 2

Background

For the purposes of this research, we consider point-to-point communication using the baseband equivalent system shown in Figure 2-1. In this figure, $x[n]$ represents generally the complex-valued sequence of (coded) information-bearing symbols. For simplicity of exposition, we assume that $x[n]$ is a zero-mean, white QPSK sequence with energy $\mathcal{E}_s$ per symbol. Additional channel coding is performed by the spread-response precoder $h[k;n]$, which generates the transmit sequence $y[n]$ from $x[n]$. The fading channel corrupts $y[n]$ with complex-valued fading and additive noise, to yield the received sequence $r[n]$. These effects are modeled by the time-varying linear filter $a[k;n]$ and additive term $w[n]$, respectively. Finally, the receiver processes $r[n]$ to obtain symbol estimates $\hat{x}[n]$, which may be further decoded.

In this chapter, we shed some light on the model for the fading channel and the structure of the spread-response precoder. For a more thorough introduction to fading channels and diversity techniques in general, the interested reader is referred to Proakis [1]. Spread-

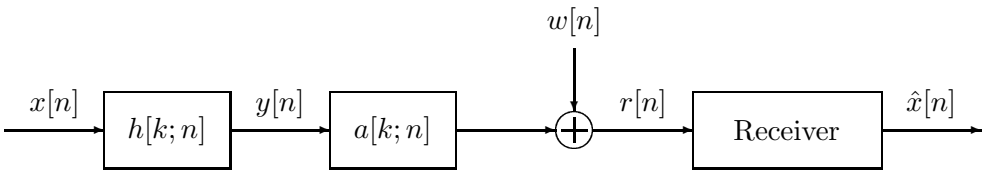


Figure 2-1: System model for spread-response precoding systems in fading environments.

response precoding is a particular example of a time diversity scheme, and was recently explored by Wornell [2], [3].

2.1 Rayleigh Fading Channels

The channel in Figure 2-1 represents a fairly general, stationary Rayleigh fading channel with uncorrelated scattering and additive noise. Specifically, we regard the kernel $a[k;n]$ as the fading response at time n to a unit-sample applied at time $n-k$, so that the received sequence may be written as the convolution

$$r[n] = \sum_k a[k;n]\, y[n-k] + w[n] \tag{2.1}$$

where $w[n]$ captures the effects of receiver thermal noise, and more importantly, co-channel interference. We model $w[n]$ as a zero-mean, white complex Gaussian sequence with variance

$$\sigma_w^2 = \mathrm{E}\left[|w[n]|^2\right] = \mathcal{N}_0 \tag{2.2}$$

Also, we note that $x[n]$, $a[k;n]$, and $w[n]$ are mutually independent.

2.1.1 Statistical Characterization

Statistically, we model the fading response $a[k;n]$ as a collection of complex jointly Gaussian sequences in n indexed by k. When these sequences have zero-mean, as we assume, the magnitude $|a[k;n]|$ is Rayleigh distributed, hence the name *Rayleigh fading channel*. Furthermore, when the sequences are jointly stationary, we may define the variance of each sequence as

$$\sigma_{a_k}^2 = \mathrm{E}\left[|a[k;n]|^2\right] \tag{2.3}$$

Finally, we assume that the sequences corresponding to distinct values of k are independent, which corresponds to what is referred to as the *uncorrelated scattering* assumption.

The above assumptions imply that the time-varying frequency response[1]

$$A(\omega; n] = \sum_k a[k; n] e^{-j\omega k} \tag{2.4}$$

is also Gaussian and stationary in ω and n. Moreover, its first and second moments are given by

$$\mu_A = \mathrm{E}\left[A(\omega; n]\right] = 0 \tag{2.5a}$$

$$\sigma_A^2 = \mathrm{E}\left[\left|A(\omega; n]\right|^2\right] = \sum_k \sigma_{a_k}^2 \tag{2.5b}$$

When we wish to classify the fading channel broadly in terms of its frequency and time selectivity, we examine the spaced-frequency spaced-time correlation function

$$\Psi_A(\omega; m] = \mathrm{E}\left[A(\theta; n]\, A^*(\theta - \omega; n - m]\right] \tag{2.6}$$

We note that the function $\Psi_A(\omega; 0]$ is a correlation function for the frequency response at frequencies separated by ω; hence, we expect the support of this function to correspond to the *coherence bandwidth* of the channel, normalized by the sampling frequency. When the transmission bandwidth is much less than the coherence bandwidth of the channel, the signal is affected uniformly in frequency, and the channel is termed *frequency nonselective*; otherwise, the channel is termed *frequency selective*. We point out that it will often be convenient to simplify our notation for the case of a frequency nonselective fading channel, for which

$$a[k; n] = A(0; n]\, \delta[k] = a[n]\, \delta[k] \tag{2.7}$$

where $a[n]$ is a zero-mean, complex Gaussian sequence with variance σ_A^2. In this case, the received sequence is given more simply as

$$r[n] = a[n]\, y[n] + w[n] \tag{2.8}$$

[1]We adopt the useful notation "$F(\cdot, \cdot]$" to indicate that the first argument to F is continuous-valued, while the second argument to F is discrete-valued.

Similarly, $\Psi_A(0;m]$ is a correlation function for the frequency response at sample times separated by m; hence, we expect the support of this function to correspond to the *coherence time* of the channel, normalized by the sample period. Specifically, we define the coherence time of the channel as

$$\tau_A = \frac{1}{\sigma_A^2} \sum_{m=0}^{\infty} \mathrm{Re}\left\{\Psi_A(0;m]\right\} \tag{2.9}$$

When the symbol duration is much less than the coherence time of the channel, the channel is termed *slowly-varying*.

2.1.2 State-Space Evolution Model

In addition to the system-level descriptions above, we need to specify an evolution model for the fading that captures some of the correlations of these sequences. Such a model allows us to simulate these characteristics as well as develop our estimation techniques in Chapter 4. We begin with a simple construction for the frequency nonselective channel, and extend it to the frequency selective channel.

Frequency Nonselective, Time Selective Channel

When the fading is frequency nonselective, but time selective, the impulse response satisfies (2.7), and the fading can be thought of as a time-varying gain applied to the transmitted signal. In order to simulate or form a model-based estimator of the channel, we need to augment (2.7) with a model for the time evolution of the sequence $a[n]$. For our purposes, we utilize a first-order autoregressive model of the form

$$a[n+1] = \rho\, a[n] + v[n+1] \tag{2.10}$$

where ρ is a real number satisfying $0 < \rho < 1$, and $v[n]$ is another zero-mean, stationary, complex Gaussian random sequence, with variance $\sigma_v^2 = \sigma_A^2\left(1-\rho^2\right)$, and which is independent of $a[k]$ for all $k < n$. The model (2.10) has been proposed in the literature for modeling time-selective fading channels, and is adequate for capturing, at least at a high level, the effects of correlation in the process. Furthermore, its simplicity helps in making analytical results more tractable.

The time constant ρ is determined by the coherence time of the fading channel. Specif-

ically, when we examine the spaced-time correlation function for this case, we find

$$\Psi_A(0;m] = \sigma_A^2 \ \rho^{|m|} \tag{2.11}$$

The coherence time, in samples, is then given by

$$\tau_A = \frac{1}{1-\rho} \tag{2.12}$$

For example, with $\rho = 0.99$, $\tau_A = 100$ samples, and the channel is highly correlated, or slowly-varying.

Frequency Selective Channel

For the frequency selective case, the channel response $a[k;n]$ consists of several taps, each of which is modeled as a zero-mean, stationary, complex Gaussian random sequence with variance $\sigma_{a_k}^2$. Furthermore, we restrict our attention to the case when the channel is causal and has impulse response of finite duration K, *i.e.*,

$$a[k;n] = 0 \quad \text{for } k < 0 \text{ and } k \geq K \tag{2.13}$$

Because we assume that the channel exhibits uncorrelated scattering, we can model each tap according to (2.10), *i.e.*,

$$a[k;n+1] = \rho_k \ a[k;n] + v[k;n+1] \tag{2.14}$$

where each ρ_k is a real number satisfying $0 < \rho_k < 1$, and the kernel $v[k;n]$ is a collection of independent, zero-mean, stationary complex Gaussian sequences with variances $\sigma_{v_k}^2 = \sigma_{a_k}^2(1-\rho_k^2)$, respectively.

We find it useful to collect the taps into the vectors

$$\mathbf{a}[n] = [\, a[0;n] \ \cdots \ a[K-1;n] \,]^T \tag{2.15a}$$

$$\mathbf{v}[n] = [\, v[0;n] \ \cdots \ v[K-1;n] \,]^T \tag{2.15b}$$

where the length K is defined by (2.13). Then the evolution model becomes

$$\mathbf{a}[n+1] = \mathbf{P}\ \mathbf{a}[n] + \mathbf{v}[n+1] \tag{2.16}$$

in which the state transition matrix $\mathbf{P} = \text{diag}\,(\rho_0, \ldots, \rho_{K-1})$.

Again, we may examine the spaced-time correlation function to find

$$\Psi_A(0; m] = \sum_k \sigma_{a_k}^2\, \rho_k^{|m|} \tag{2.17}$$

so that the coherence time of the frequency selective channel is the weighted sum

$$\tau_A = \frac{1}{\sigma_A^2} \sum_k \sigma_{a_k}^2 \left(\frac{1}{1-\rho_k}\right) \tag{2.18}$$

2.2 Spread-Response Precoding Systems

As a result of intersymbol interference and signal fading, uncoded transmissions over fading channels often exhibit dramatically poorer bit-error performance than uncoded transmissions over traditional additive white Gaussian noise (AWGN) channels. To combat these effects, systems often employ diversity schemes, including time, frequency, and spatial diversity methods [1]. The spread-response precoder $h[k; n]$ in Figure 2-1 represents a recently proposed technique for obtaining time diversity in fading environments [2], [3]. We focus on this technique throughout the thesis.

The idea behind time diversity schemes in general is to temporally distribute the energy of each symbol, and effectively recombine these transmissions, to achieve closer to the average effect of the channel rather than the instantaneous fade. Spread-response precoding systems in particular achieve time diversity without requiring additional power or bandwidth, by passing the symbol sequence through a dispersive, invertible, linear transformation before transmission. Accordingly, spread-response precoding represents a computationally efficient alternative to error-control coding for achieving time diversity. Furthermore, spread-response precoding effectively transforms the fading channel into an additive, white marginally Gaussian noise channel.

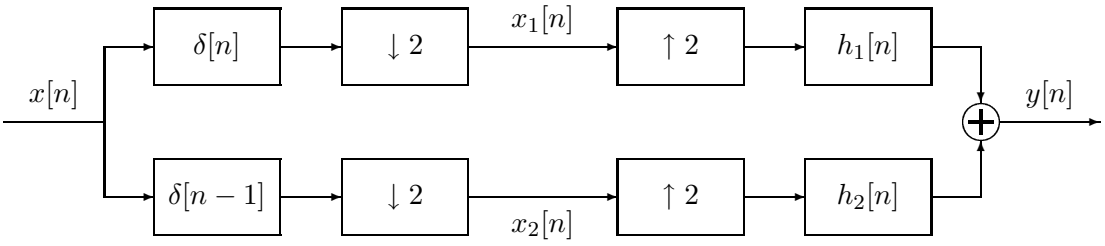


Figure 2-2: Linear, periodically time-varying spread-response precoder structure.

2.2.1 Transmitter: Linear, Periodically Time-Varying Precoder

The spread-response precoder $h[k;n]$ produces as its output the sequence

$$y[n] = \sum_k h[k;n]\, x[n-k] \tag{2.19}$$

For the purposes of this research, we restrict our attention to the special case in which the impulse response $h[k;n]$ corresponds to a linear, periodically time-varying (LPTV) system of the form shown in Figure 2-2. We view the precoding structure of Figure 2-2 in the following way. The first two stages demultiplex the symbol sequence $x[n]$ into its even samples, $x_1[n]$, and odd samples, $x_2[n]$, respectively. Note that we can easily recover $x[n]$ at this point by appropriately interleaving the two sequences. Another interpretation of this operation is that of converting the original user of the channel into two "virtual" users. In fact, the last two stages correspond to a spread-signature filter bank of order $L = 2$ and finite duration N developed in [3] for multiuser communication over fading channels. We now describe some of the important properties of these spread-signature sets, which we denote by the vector

$$\mathbf{h}[n] = [\, h_1[n] \;\cdots\; h_L[n] \,]^T \tag{2.20}$$

First, we note that the signature set, together with its translates by integer multiples of L, constitute a complete orthonormal set, so that the transformation introduced by precoding is simply invertible. Note that this is not a necessary condition for recovery of $x[n]$ from $y[n]$, but the orthogonal construction is convenient in terms of both analysis and implementation. Furthermore, for practical reasons, we restrict out attention to real-valued

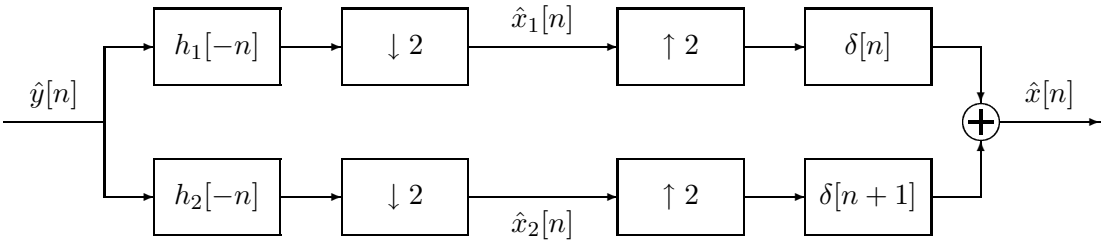


Figure 2-3: Linear, periodically time-varying postcoder structure for inverting the effects of the precoder of Figure 2-2.

signature sets. Consequently, the inversion is simply the transpose of the system given in Figure 2-2, which we give in block diagram form in Figure 2-3 for completeness. The notation "$\hat{(\cdot)}$" in Figure 2-3 is necessary to indicate that the sequences are not equivalent to their counterparts in Figure 2-2 when channel distortions are introduced between the precoder and postcoder.

Secondly, the sequences take values

$$h_l[n] = \pm 1/\sqrt{N} \tag{2.21}$$

for each $1 \leq l \leq L$ and $0 \leq n < N$. This property allows for efficient implementations of these systems via simple additions and subtractions. Furthermore, (2.21) implies that these sequences are the most effective, among sequences of length N, at temporally spreading each symbol. More specifically, we define the dispersion factor $\mathcal{D}_{\mathbf{h}}$ of the spread-signature set as

$$\frac{1}{\mathcal{D}_{\mathbf{h}}} = \frac{1}{L} \sum_l \frac{1}{\mathcal{D}_{h_l}} \tag{2.22}$$

where $\mathcal{D}_{h_l}$ represents the dispersion of the sequence $h_l[n]$, namely,

$$\mathcal{D}_{h_l} = \left(\sum_n h_l^4[n] \right)^{-1} \tag{2.23}$$

Then these signature sets achieve the bound

$$\mathcal{D}_{\mathbf{h}} \leq N \tag{2.24}$$

with equality. This property is important because the effective time diversity offered by the spread-response precoder increases with N. Furthermore, we see that since the bound is achieved for any $L \geq 2$, there is no need to consider precoders with $L > 2$ for the single-user scenario.

However, increasing the length N of the precoder to obtain higher diversity benefit precludes the use of optimal maximum-likelihood sequence detection (MLSD) due to its computation complexity. To make this statement more precise, consider the channel formed by the cascade of the precoder and fading channel response

$$\tilde{a}[k;n] = \sum_{l} a[l;n]\, h[k-l;n-l] \tag{2.25}$$

Let us assume that the fading channel impulse response is known at the receiver, so the composite channel (2.25) is also known. Recall that the precoder at any time n is causal and has finite duration N, and the fading channel at any time n is also causal and has finite duration K. Hence we expect the cascaded channel $\tilde{a}[k;n]$ to satisfy

$$\tilde{a}[k;n] = 0 \quad \text{for } k < 0 \text{ and } k \geq (N+K-1)$$

for any time n. When each symbol in the sequence $x[n]$ takes value in the finite set $\mathcal{X}$, we may model the cascade channel as a $|\mathcal{X}|^{\nu}$ state finite-state machine, where $\nu = (N+K-2)$ under our assumptions. Then implementing the MLSD via the Viterbi algorithm requires

$$|\mathcal{X}|^{\nu+1} = |\mathcal{X}|^{(N+K-1)} \geq |\mathcal{X}|^{N} \tag{2.26}$$

computations for each received sample [4]. Consequently, we see that increasing the length of the precoder leads to an exponential increase in the amount of computation required for the MLSD.

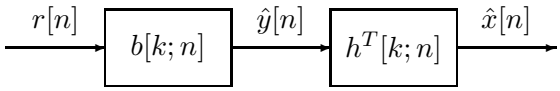


Figure 2-4: Linear receiver consisting of an equalizer and postcoder.

2.2.2 Receiver: Linear Equalizer–Postcoder

Instead of the MLSD, the receiver block of Figure 2-4 consists of a linear equalizer $b[k;n]$, which partially compensates for the effects of fading, followed by the postcoder[2] $h^T[k;n]$, which inverts the precoding process.

We define the cascade of the fading channel and equalizer as the kernel

$$c[k;n] = \sum_l b[l;n]\, a[k-l;n-l] \tag{2.27}$$

and assume that all three kernels are wide-sense stationary, uncorrelated scattering channels that satisfy certain ergodicity constraints. Under these assumptions, Wornell [2], [3] has shown the following result: As the dispersion $\mathcal{D}_{\mathbf{h}} \to \infty$, the symbol estimates $\hat{x}[n]$ converge to

$$\hat{x}[n] \xrightarrow{\text{m.s.}} \mu_C\, x[n] + z[n] \tag{2.28}$$

where we use the notation "m.s." to indicate mean-square convergence. The sequence $z[n]$ in (2.28) is a zero-mean, complex-valued, marginally Gaussian white-noise sequence with variance

$$\sigma_z^2 = \mathcal{E}_s \sigma_C^2 + \mathcal{N}_0 \left(\sigma_B^2 + |\mu_B|^2 \right) \tag{2.29}$$

and μ_B, μ_C and σ_B^2, σ_C^2 are the means and variances of the kernels $B(\omega;n]$ and $C(\omega;n]$, respectively, found from (2.5).

Accordingly, we may compute the signal-to-noise ratio (SNR) in the symbol estimates

[2] We use a special notation for the postcoding filter to simplify the receiver block diagram. Specifically, if $h[k;n]$ is the spread-response precoder of Figure 2-2, the the notation "$h^T[k;n]$" indicates the corresponding postcoder of Figure 2-3.

as

$$\gamma(b) = \frac{|\mu_C|^2}{\sigma_C^2 + \xi_0 \left(\sigma_B^2 + |\mu_B|^2\right)} \tag{2.30}$$

where

$$\xi_0 = \frac{\mathcal{N}_0}{\mathcal{E}_s} \tag{2.31}$$

We may then approximate the bit-error probability for symbol-by-symbol detection when using QPSK signaling over this composite channel as the corresponding bit-error probability over an AWGN channel with SNR γ. This probability of error, given by [1], is

$$\Pr(e) \approx \mathcal{Q}(\sqrt{\gamma}) \tag{2.32}$$

where

$$\mathcal{Q}(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-x^2/2} dx \tag{2.33}$$

This approximation is simple, and turns out to be a good one as the length N, and therefore dispersion $\mathcal{D}_{\mathbf{h}}$, of the precoder becomes large [2].

2.3 Summary

In this chapter, we have presented a context for the thesis and introduced important notation and terminology for fading channels and spread-response precoding systems. We have indicated the structure of the spread-response precoder and briefly stated some of its properties. Most importantly, we have repeated Wornell's theorem on the ability of spread-response precoding systems to transform the arbitrary Rayleigh fading channel into an additive marginally Gaussian white-noise channel, and given a useful approximation for the bit-error probability of QPSK signaling over the composite channel. Finally, we have set the stage for Chapter 3, in which we address the problem of choosing the equalizer $b[k;n]$ to obtain optimal performance from the class of linear equalizers.

Chapter 3

Optimal Linear Equalizers

In Chapter 2, we introduced the concept of spread-response precoding for communication over fading channels. Our main result was that under certain ergodicity constraints on the channel, long precoding essentially transforms the fading channel into an additive marginally Gaussian white-noise channel.

In this chapter, we examine two criteria for deriving optimal equalizers. First, we develop equalizers which maximize the output SNR of the composite channel, following the development of [2]. Because the precoder is an invertible, all-pass transformation of the transmitted sequence, we will see that the resulting receivers correspond to linear, minimum mean-square equalizers for two special cases of interest. Due to this overlap for certain cases of the channel, we examine the minimum mean-square error criterion in more detail and develop single user versions of the recursive equalizers found in [5]. Comparisons between the equalizers are made in terms of their time-varying frequency responses and bit-error probabilities.

3.1 Maximum SNR Criterion

Our first design criterion for obtaining optimal equalizers is to maximize the composite channel SNR given by

$$\gamma(b) = \frac{|\mu_C|^2}{\sigma_C^2 + \xi_0 \left(\sigma_B^2 + |\mu_B|^2\right)} \tag{3.1}$$

with $\xi_0 = \mathcal{N}_0/\mathcal{E}_s$ as in (2.31). Such an approach seems reasonable because we approximate the composite channel as Gaussian, and the bit-error probability for Gaussian channels (2.32) decreases monotonically with increasing SNR. While the development in this section mirrors that of [2], the individual steps will prove useful at a later point in the thesis. We examine two specific cases of the channel, namely, frequency nonselective fading and frequency selective, slowly-varying fading, respectively.

3.1.1 Frequency Nonselective Channel

For the case of a frequency nonselective channel, the fading response satisfies

$$a[k;n] = a[n]\,\delta[k]$$

in which case the equalizer is of the form

$$b[k;n] = b[n]\,\delta[k]$$

and

$$c[k;n] = a[n]\,b[n]\,\delta[k]$$

Consequently, we may rewrite (3.1) as

$$\gamma(b) = \frac{\left|\mathrm{E}\,[ab]^2\right|}{\mathrm{Var}\,[ab] + \xi_0\,\mathrm{E}\left[|b|^2\right]} \tag{3.2}$$

where we have dropped the dependence on n due to stationarity. To derive the optimal equalizer, we first rewrite (3.2) as [2]

$$\gamma(b) = \frac{1}{1/\phi(b) - 1} \tag{3.3}$$

where

$$\phi(b) = \frac{|\mathrm{E}\,[ab]|^2}{\mathrm{E}\left[\left(|a|^2 + \xi_0\right)|b|^2\right]} \tag{3.4}$$

and we note that maximizing (3.4) is equivalent to maximizing (3.3). Now by the Schwartz inequality, we have

$$\begin{aligned} |\mathrm{E}\,[ab]|^2 &= \left| \mathrm{E}\left[\frac{a}{\sqrt{|a|^2 + \xi_0}} \cdot b\sqrt{|a|^2 + \xi_0} \right] \right|^2 \\ &\leq \mathrm{E}\left[\frac{|a|^2}{|a|^2 + \xi_0} \right] \cdot \mathrm{E}\left[\left(|a|^2 + \xi_0 \right) |b|^2 \right] \end{aligned}$$

with equality just in case

$$b[n] \propto \frac{a^*[n]}{|a[n]|^2 + \xi_0} \tag{3.5}$$

Thus we maximize (3.4) and therefore (3.3) through this choice of $b[n]$, which is specified to within a (complex) scale factor.

3.1.2 Frequency Selective, Slowly-Varying Channel

In this section we consider the more general, frequency selective fading channel model in which the response has more than a single tap. We note that as we increase the transmission bandwidth relative to the coherence bandwidth of the channel, we also decrease the symbol period relative to the coherence time of the channel. Hence, for very short symbol times, the channel is very slowly-varying, and the cascade of the fading and equalizer can be approximated via

$$C(\omega; n] \approx A(\omega; n] \cdot B(\omega; n] \tag{3.6}$$

Essentially this approximation requires that the fading channel is time-invariant over a block of length N, with $K < N < \tau_A$ and where we have assumed $\tau_A \gg K$. Consequently, the composite channel frequency response should be well approximated by the product of the two constituent frequency responses, as in (3.6).

Using (3.6) in (3.1), we find that the output SNR for the frequency selective channel is given approximately by

$$\gamma(b) \approx \frac{|\mathrm{E}\,[AB]|^2}{\mathrm{Var}\,[AB] + \xi_0\,\mathrm{E}\left[|B|^2\right]} \tag{3.7}$$

where we have dropped the dependence of these statistics on ω and n due to stationarity.

Following an argument similar to that of the previous section, we find that the equalizer

that approximately maximizes the composite channel SNR has frequency response [2]

$$B(\omega; n] \propto \frac{A^*(\omega; n]}{|A(\omega; n]|^2 + \xi_0} \tag{3.8}$$

where again the equalizer is specified to within an arbitrary (complex) scale factor.

3.1.3 Average Output SNR and Bit-Error Probability

With the above choices for the equalizer $b[k; n]$, we may compute the output SNR directly. In fact, we see by inspection of the results (3.5) and (3.8) that the maximum output SNR are identical in both cases. This is due to the fact that for the nonselective channel, the processes $a[n]$ and $A(\omega; n] = A(0; n]$ are equivalent, with mean-zero and variance σ_A^2.

We solve for the maximum SNR by substituting (3.5) into (3.4), and computing the required expectations. The details of these calculations are a special case of those given in Appendix A, with $\eta = \eta' = \xi_0$, resulting in

$$\phi_{\max} = 1 - \zeta_0 e^{\zeta_0} E_1(\zeta_0) \tag{3.9}$$

where the receiver input SNR is given by

$$\text{SNR} = 1/\zeta_0 = \frac{\mathcal{E}_s \, \sigma_A^2}{\mathcal{N}_0} \tag{3.10}$$

and $E_1(\cdot)$ denotes the exponential integral

$$E_1(\nu) = \int_\nu^\infty \frac{e^{-t}}{t} dt \tag{3.11}$$

Hence, by applying (3.3), we obtain

$$\gamma_{\max} = \frac{1}{\zeta_0 e^{\zeta_0} E_1(\zeta_0)} - 1 \tag{3.12}$$

Equation (3.12) is useful for our purposes because it gives us an approximate upper bound on the output SNR of spread-response precoding systems for the two cases we have addressed. Similarly, for uncoded transmissions, it gives, in conjunction with (2.32), at least an approximate lower bound on the bit-error probability of symbol-by-symbol detection.

Figure 3-1 shows the bit-error probability as a function of the receiver input signal-to-noise ratio, defined in (3.10), for various lengths of the precoding filters. As the figure indicates, for practical systems with finite length precoding, the performance approaches the bound as the length of the precoder increases. In particular, for the input SNR values of interest, the asymptotic result is virtually achieved by precoding of length $N = 256$ times the coherence time of the channel in samples.

3.2 Linear MMSE Criterion

As Wornell [2] also points out for the frequency nonselective case, the output $\hat{y}[n]$ from the equalizer (3.5) corresponds to a linear, MMSE estimate of the transmitted sequence $y[n]$. Hence, since the postcoder $h^T[k;n]$ inverts the precoder, its output $\hat{x}[n]$ corresponds to a linear, MMSE estimate of $x[n]$. Similarly, under the assumption (3.6) and, with $h^T[k;n]$ the precoder inverse, $\hat{x}[n]$ corresponds, at least approximately, to linear MMSE estimates of the symbol sequence $x[n]$.

These observations suggest that a reasonable alternative and sometimes equivalent criterion for the receiver structure is to determine the linear receiver that has minimal mean-square error. In fact, this criterion was used to develop the recursive multiuser receiver of [5]. In this section, we examine the structure of this equalizer, which is based on the Kalman filter [6], for the single user case.

3.2.1 Channel State-Space Model

To derive the Kalman filter, we must cast the channel of Figure 2-1 into state-space form. The approach taken in [5] is to used the cascade of the precoder and fading channel (2.25), so that the received sequence is given simply as

$$r[n] = \sum_k \tilde{a}[k;n]\, x[n-k] + w[n] \tag{3.13}$$

Recall that the precoder is causal and has finite length N, so that $h[n] = 0$ for $n < 0$ and $n \geq N$. With the additional assumption that $a[k;n]$ is of finite length K, the combined

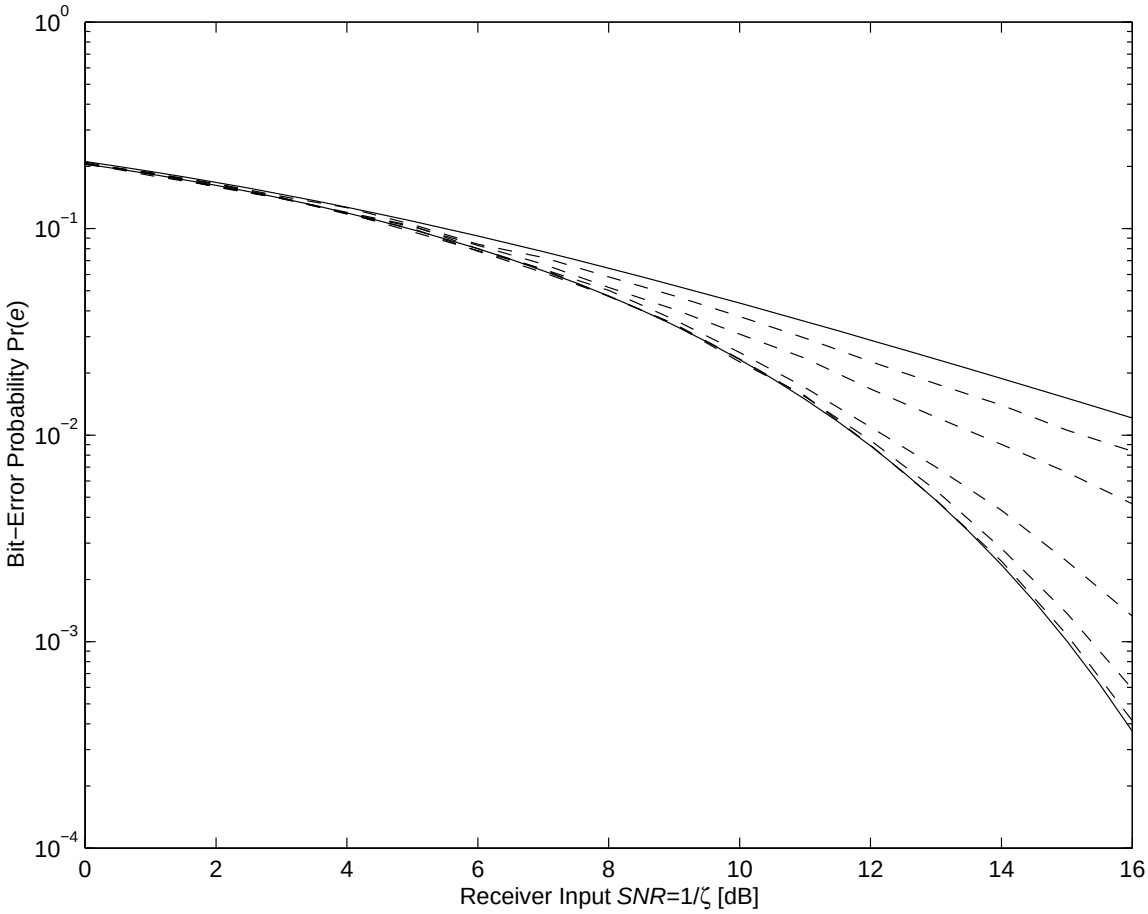


Figure 3-1: Bit-error probability for QPSK signaling using spread-response precoding of various lengths over known fading channels. The successively lower solid curves represent the analytical bit-error performance over the fading channel without precoding and for infinite-length precoding, respectively. The successively lower dashed curves are the performance of spread-response precoders of lengths $N = 2, 4, 16, 64$ and 256, respectively.

kernel $\tilde{a}[k;n]$ also has finite duration, namely, $(N+K-1)$. Hence, with the definitions

$$\mathbf{s}[n] = [\, x[n] \; x[n-1] \; \cdots \; x[n-(N+K-2)] \,]^T \tag{3.14a}$$

$$\tilde{\mathbf{a}}[n] = [\, \tilde{a}[0;n] \; \tilde{a}[1;n] \; \cdots \; \tilde{a}[N+K-2;n] \,] \tag{3.14b}$$

a state-space model for the channel takes the form

$$\mathbf{s}[n+1] = \mathbf{F}\,\mathbf{s}[n] + \mathbf{G}\,x[n+1] \tag{3.15a}$$

$$r[n] = \tilde{\mathbf{a}}[n]\,\mathbf{s}[n] + w[n] \tag{3.15b}$$

where

$$\mathbf{G} = [\, 1 \; 0 \; \cdots \; 0 \,]^T$$

and $\mathbf{F}$ is the $(N+K-1)$ dimensional shift matrix

$$\mathbf{F} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix}$$

3.2.2 Recursive Equalizer Equations

Given the state-space model (3.15), it is straightforward to derive the Kalman filter, consisting of two sets of equations, namely, the update equations, and the prediction equations. For ease of presentation, we make the definitions

$$\hat{\mathbf{s}}[n|m] = \mathrm{E}\,[\, \mathbf{s}[n] \;\; | \; r[l],\; l \leq m \;]\, \tag{3.16a}$$

$$\mathbf{\Sigma}_{\mathbf{s}}[n|m] = \mathrm{E}\left[(\mathbf{s}[n] - \hat{\mathbf{s}}[n|m])\,(\mathbf{s}[n] - \hat{\mathbf{s}}[n|m])^{\dagger} \;\; | \; r[l],\; l \leq m \;\right] \tag{3.16b}$$

to denote the state estimate and its corresponding conditional mean-square error, respectively. We also fix the initial conditions of the Kalman filter to $\hat{\mathbf{s}}[0|-1] = \mathbf{0}$ and

$\mathbf{\Sigma_s}[0|-1] = \mathcal{E}_s\mathbf{I}$. Finally, the symbol estimates may be taken as

$$\hat{x}[n] = \mathbf{B}_j\,\hat{\mathbf{s}}[n|n] \tag{3.17}$$

where $\mathbf{B}_j$ is the unit vector whose j^{th} entry is 1, and j is chosen to achieve a specified degree of smoothing in the estimates.

Symbol Estimate Update Equations

Given the estimate $\hat{\mathbf{s}}[n|n-1]$ and its corresponding conditional error variance $\mathbf{\Sigma_s}[n|n-1]$, as well as the observation $r[n]$, we compute the updated estimate and its conditional error covariance according to

$$\hat{\mathbf{s}}[n|n] = \hat{\mathbf{s}}[n|n-1] + \mathbf{K}[n]\,(r[n] - \tilde{\mathbf{a}}[n]\,\hat{\mathbf{s}}[n|n-1]) \tag{3.18a}$$

$$\mathbf{\Sigma_s}[n|n] = (\mathbf{I} - \mathbf{K}[n]\,\tilde{\mathbf{a}}[n])\,\mathbf{\Sigma_s}[n|n-1] \tag{3.18b}$$

with

$$\mathbf{K}[n] = \mathbf{\Sigma_s}[n|n-1]\,\tilde{\mathbf{a}}^\dagger[n]\left(\tilde{\mathbf{a}}[n]\,\mathbf{\Sigma_s}[n|n-1]\,\tilde{\mathbf{a}}^\dagger[n] + \mathcal{N}_0\right)^{-1} \tag{3.18c}$$

Symbol Estimate Prediction Equations

Next, given the estimate $\hat{\mathbf{s}}[n|n]$ and $\mathbf{\Sigma_s}[n|n]$, we compute the predicted symbol estimates according to

$$\hat{\mathbf{s}}[n+1|n] = \mathbf{F}\,\hat{\mathbf{s}}[n|n] \tag{3.19a}$$

$$\mathbf{\Sigma_s}[n+1|n] = \mathbf{F}\,\mathbf{\Sigma_s}[n|n]\,\mathbf{F}^T + \mathcal{E}_s\,\mathbf{G}\,\mathbf{G}^T \tag{3.19b}$$

Using these predicted estimates, the algorithm can continue with (3.18) at time $(n+1)$ as soon as the observation $r[n+1]$ is available.

Equations (3.18) and (3.19) yield efficient, recursive implementations of the linear, MMSE equalizer based on observations from the present and into the infinite past. Additional smoothing of the symbol estimates may be obtained by augmenting the state vector

$\mathbf{s}[n]$ and state-space matrices $\mathbf{F}$, $\mathbf{G}$, and $\tilde{\mathbf{a}}[n]$ accordingly. There are more efficient methods for obtaining smoothed estimates [6], but we restrict our attention to a small amount of smoothing to simplify the exposition.

3.3 Comparison of the Equalizers for the Two Criteria

The equations of the previous section specify exactly how to implement a linear MMSE equalizer in recursive form in the time domain, regardless of whether or not the precoding filter is lossless, or whether or not the channel is slowly-varying. In this sense, the linear MMSE equalizer appears more appropriate for implementation purposes than the equalizer of (3.8), whose implementation generally requires truncation of the impulse response. However, we do not want to sacrifice performance for ease of implementation; therefore, we need a framework for comparing the Kalman filter with the equalizers of (3.5) and (3.8), which we have seen maximize the output signal-to-noise ratio in the frequency nonselective and slowly-varying, frequency selective cases, respectively.

Such a comparison is very complicated to formalize based on the recursive equations given (3.18) and (3.19), because in general the state-space model is time-varying. Instead, we consider the effects of the Kalman filter/smoother in terms of non-recursive, linear MMSE estimation equations, and find that we may compute the time-varying frequency response of the Kalman filter/smoother directly from this point of view.

More specifically, we treat a random sequence $f[n]$ as the infinite random vector

$$\mathbf{f} = [\cdots \; f[-1] \; f[0] \; f[1] \; \cdots]^T$$

Hence, the signals $x[n]$, $y[n]$, $w[n]$, and $r[n]$ become the vectors $\mathbf{x}$, $\mathbf{y}$, $\mathbf{w}$, and $\mathbf{r}$, respectively. From our system model, we have that $\mathrm{E}[\mathbf{x}] = \mathbf{0}$ and $\mathrm{E}\left[\mathbf{x}\mathbf{x}^\dagger\right] = \mathcal{E}_s \mathbf{I}$. Similarly, $\mathrm{E}[\mathbf{w}] = \mathbf{0}$, $\mathrm{E}\left[\mathbf{w}\mathbf{w}^\dagger\right] = \mathcal{N}_0 \mathbf{I}$, and $\mathrm{E}\left[\mathbf{x}\mathbf{w}^\dagger\right] = \mathbf{0}$.

To specify the second-order statistics of the vectors $\mathbf{y}$ and $\mathbf{r}$, we must first determine matrix representations of the precoder and fading channel by examining the convolution relationships (2.1) and (2.19) for these two systems. The received sequence may be written as

$$\mathbf{r} = \mathbf{A}\mathbf{y} + \mathbf{w} \tag{3.20}$$

where the matrix $\mathbf{A}$ has the form

$$\mathbf{A} = \begin{bmatrix} & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & & \\ \cdots & 0 & a[K-1;-1] & \cdots & a[0;-1] & 0 & 0 & 0 & \cdots \\ \cdots & 0 & 0 & a[K-1;0] & \cdots & a[0;0] & 0 & 0 & \cdots \\ \cdots & 0 & 0 & 0 & a[K-1;1] & \cdots & a[0;1] & 0 & \cdots \\ & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \end{bmatrix}$$

We note that $a[0;0]$ in the middle row aligns with $y[0]$ in the vector $\mathbf{y}$. Similarly, it is straightforward to write the precoder convolution as

$$\mathbf{y} = \mathbf{H}\mathbf{x} \tag{3.21}$$

where the $\mathbf{H}$ matrix has the form

$$\mathbf{H} = \begin{bmatrix} & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & & \\ \cdots & 0 & h[N-1;-1] & \cdots & h[0;-1] & 0 & 0 & 0 & \cdots \\ \cdots & 0 & 0 & h[N-1;0] & \cdots & h[0;0] & 0 & 0 & \cdots \\ \cdots & 0 & 0 & 0 & h[N-1;1] & \cdots & h[0;1] & 0 & \cdots \\ & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \end{bmatrix}$$

Again, we note that $h[0;0]$ in the middle row aligns with $x[0]$ in the vector $\mathbf{x}$. Furthermore, we could group these two matrices into the combined matrix $\tilde{\mathbf{A}} = \mathbf{A}\mathbf{H}$, but we have reasons for keeping them separate, as we will soon point out.

Using (3.20) as well as (3.21), we find that

$$\mathrm{E}\left[\mathbf{r}\right] = \mathbf{0} \tag{3.22a}$$

$$\mathrm{E}\left[\mathbf{x}\mathbf{r}^{\dagger}\right] = \mathcal{E}_s \mathbf{H}^{\dagger}\mathbf{A}^{\dagger} \tag{3.22b}$$

$$\mathrm{E}\left[\mathbf{r}\mathbf{r}^{\dagger}\right] = \mathcal{E}_s \mathbf{A}\mathbf{H}\mathbf{H}^{\dagger}\mathbf{A}^{\dagger} + \mathcal{N}_0\mathbf{I} \tag{3.22c}$$

From these statistics, we can write the Kalman smoother equations in non-recursive vector form as

$$\hat{\mathbf{x}} = \mathbf{H}^{\dagger}\mathbf{A}^{\dagger}\left(\mathbf{A}\mathbf{H}\mathbf{H}^{\dagger}\mathbf{A}^{\dagger} + \xi_0\mathbf{I}\right)^{-1}\mathbf{r} \tag{3.23}$$

with $\xi_0 = \mathcal{N}_0/\mathcal{E}_s$ as defined in (2.31). As with the other convolutions, (3.23) is an infinite set of equations for computing all the symbol estimates $\hat{\mathbf{x}}$ based on the entire observed sequence $\mathbf{r}$. We note that this formulation is only of conceptual value; to implement the linear MMSE equalizer, we employ the recursive equations of the previous section.

For the special case in which the precoder inverse is the transpose, $\mathbf{H}$ is a unitary matrix, so the Kalman smoother simplifies to

$$\hat{\mathbf{x}} = \mathbf{H}^\dagger \mathbf{A}^\dagger \left(\mathbf{A}\mathbf{A}^\dagger + \xi_0 \mathbf{I} \right)^{-1} \mathbf{r} \tag{3.24}$$

which is similar in form to (3.8) followed by the postcoder. This result confirms the partition of the receiver into an equalizer and postcoder: if the precoding transformation was not orthogonal, then we could not partition the linear MMSE equalizer in this manner. Because of this partitioning, we can simply compare the filter

$$\hat{\mathbf{B}} = \mathbf{A}^\dagger \left(\mathbf{A}\mathbf{A}^\dagger + \xi_0 \mathbf{I} \right)^{-1} \tag{3.25}$$

to the equalizers of (3.5) and (3.8).

Just as the matrix $\mathbf{A}$ has a special structure, so does the filter matrix $\hat{\mathbf{B}}$. Specifically, each row of the matrix contains the time-reversed version of the equalizer impulse response $\hat{b}[k;n]$ for a particular time n. Hence to compare the two equalizers at time n_0, we simply pick the corresponding row of $\hat{\mathbf{B}}$, reverse this row in time, and compute its frequency response. Although we cannot compute the doubly-infinite matrix $\hat{\mathbf{B}}$ for the general case, we can make several observations in the two cases of immediate interest.

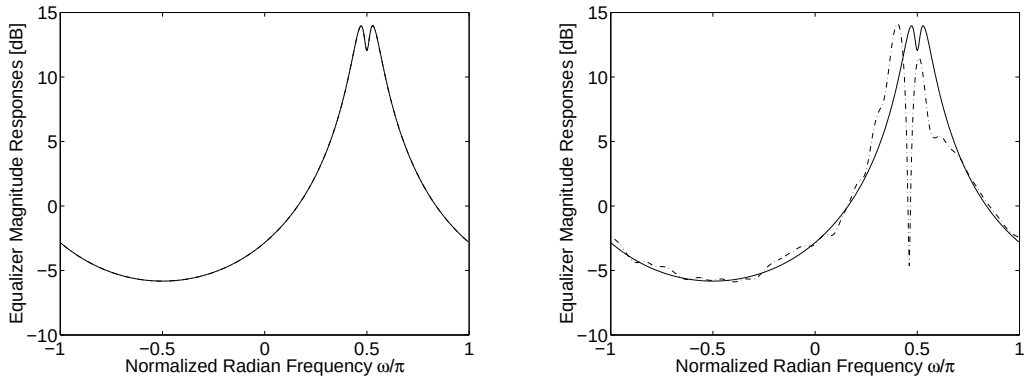
For the case of the frequency nonselective channel $a[k;n] = a[n]\delta[k]$, and the fading matrix is

$$\mathbf{A} = \mathrm{diag}(\ldots, a[-1], a[0], a[1], \ldots)$$

so that $\hat{\mathbf{B}}$ is also a diagonal matrix. Picking the elements along the diagonal, we find

$$\hat{b}[n] = \frac{a^*[n]}{|a[n]|^2 + \xi_0} \tag{3.26}$$

which is identical to (3.5). Hence, the two optimization problems, based on seemingly different criteria, give the same result for this case.



(a) (b)

Figure 3-2: Frequency response comparisons between the (approximate) maximum SNR equalizer and the MMSE equalizer, for (a) a slowly-varying channel, and (b) a more quickly-varying channel. The solid lines corresponds to the (approximate) maximum SNR equalizer $B(\omega; 100]$, while the dash-dotted lines correspond to the MMSE equalizer $\hat{B}(\omega; 100]$.

For an arbitrary frequency selective channel, $\hat{\mathbf{B}}$ is impossible to compute. However, we might examine the finite-length version of the equation (3.20) for a particular realization of the fading. For example, if we assume $K = 2$ and write the equation (3.20) for a segment of the data of length M, then we can compute the truncated approximation $\hat{\mathbf{B}}_M$ to $\hat{\mathbf{B}}$.[1] We may then compute the frequency responses (3.8) and compare to the frequency responses of the corresponding rows of $\hat{\mathbf{B}}_M$.

Illustrative results of these computations are shown in Figure 3-2 for a slowly-varying channel and a more quickly-varying channel. In these figures, we compare the same fading channel frequency response surrounded by different contexts. That is, we fix the frequency response at time $n = 100$ and use forward and backward recursions of the model (2.16) to generate a block of length $M = 200$. In both cases, we arbitrarily set $\xi_0 = 0.01$.

For the slowly-varying case, we use the parameters $\rho_0 = \rho_1 = 0.9999$ and $\sigma_{a_0}^2 = \sigma_{a_1}^2 = 1/2$, corresponding to a channel with coherence time $\tau_A = 10000$. Figure 3-2(a) shows the frequency responses of the two equalizers in the middle of this data segment, *i.e.*, the (approximate) maximum SNR equalizer $B(\omega; 100]$ computed from (3.8), and the linear MMSE equalizer $\hat{B}(\omega; 100]$ computed by taking the discrete-time Fourier transform of the

[1]When M is large, we expect the middle rows of $\hat{\mathbf{B}}_M$ to be good approximations of the corresponding rows of $\hat{\mathbf{B}}$. In fact, computer experiments have suggested that the results for $M \geq 50$ are largely similar.

middle row of $\hat{\mathbf{B}}_{200}$. We see from the figure that the approximation (3.6) does result in a good match between (3.8) and the Kalman filter/smoother in terms of their frequency responses. Hence we expect similar performance for both equalizers in terms of output signal-to-noise ratio and mean-square error. Furthermore, we find it appropriate to utilize the equalizer of (3.8) for analysis of both approaches; however, we find the Kalman filter/smoother more suitable for implementation purposes because it is constrained to be realizable.

A similar comparison is offered by Figure 3-2(b), in which the same frequency response is fixed at time $n = 100$, but for this example, the parameters of the model (2.16) were $\rho_0 = \rho_1 = 0.99$, and $\sigma_{a_0}^2 = \sigma_{a_1}^2 = 1/2$, corresponding to a channel with coherence time $\tau_A = 100$. For this channel, the two equalizers are very different in terms of their frequency responses; however, this figure does not offer the whole picture. Ultimately, we are interested in the bit-error probability of the systems employing these equalizers. While Figure 3-2(b) suggests that the bit-error probabilities are different, it does not suggest which of the two is smaller.

Figure 3-3 shows empirical bit-error probability for the case of a white channel, namely, $\rho_0 = \rho_1 = 0$, and $\sigma_{a_0}^2 = \sigma_{a_1}^2 = 1/2$ using the two equalizers with length $N = 256$ precoding. The successively lower curves correspond to the (approximate) maximum SNR equalizer (3.8), the Kalman filter (without smoothing), and the Kalman filter with a smoothing factor of 100, respectively. It is interesting to note that the performance of the equalizer (3.8) does not improve with increasing SNR. This result suggests that, for the quickly-varying, frequency selective channel, this equalizer is incapable of reducing the time-varying intersymbol interference introduced by the channel. On the other hand, the Kalman filter equalizer's performance does improve with increasing SNR, which suggests that it is in fact capable of reducing the ISI. Moreover, it seems plausible that the performance improves with additional smoothing, although more extensive investigation of this result is required.

3.4 Summary

In this chapter, we have examined two criteria for obtaining optimal linear equalizers for use in spread-response precoding systems. Specifically, we have demonstrated that the maximum output SNR and MMSE criteria yield exactly the same equalizer in the frequency nonselective case. For the frequency selective channel, our numerical results suggest that

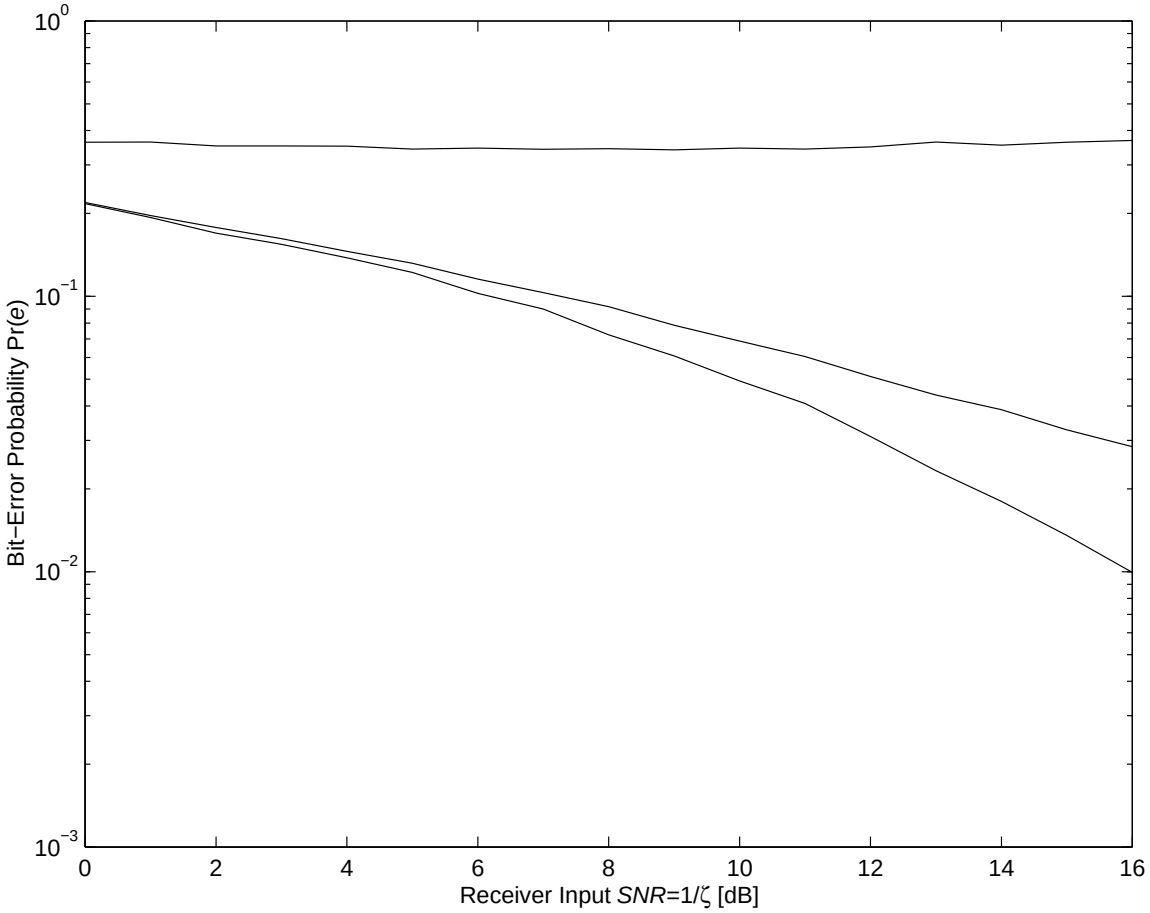


Figure 3-3: Bit-error probability over a very fast-fading channel using the two equalizers with length $N = 256$ precoding. The successively lower solid curves correspond to the (approximate) maximum SNR equalizer, the Kalman filter (without smoothing), and the Kalman filter with a degree of smoothing 100.

the (approximate) maximum SNR equalizer and the linear MMSE equalizer are similar when the channel is slowly-varying. However, we have shown through simulation that the (approximate) maximum SNR equalizer is severely interference limited when the channel exhibits fast fading, while the linear MMSE equalizer compensates for some of the ISI effect in this case. Whether the true maximum SNR equalizer corresponds to the linear MMSE equalizer remains as an interesting question for future research.

Chapter 4

Fading Channel Estimation

As derived in Chapter 3, the fading channel equalizers of Section 3.1 that maximize the output SNR of the composite channel, and the recursive equalizer of Section 3.2 that minimizes the output mean-square error both require exact knowledge of the fading impulse response $a[k;n]$, as well as the additive noise intensity $\mathcal{N}_0$. However, in practice, only estimates of these quantities may be available at the receiver. For spread-response precoding systems over the frequency nonselective channel, Wittneben [7] demonstrated SNR loss due to noisy channel estimates and only moderate sensitivity of the output SNR to estimates of $\mathcal{N}_0$. These preliminary results motivate us to look at the effects of channel estimation in more detail. Our approach is to formulate a channel estimator independently, and later examine how estimation errors impact the performance of spread-response precoding systems.

In this chapter we develop a fading channel estimator based on the Kalman filter, while in Chapter 5 we analyze the performance of spread-response precoding systems when these estimates are used in the equalizer. First, we develop a state-space model for the channel by augmenting the evolution model of (2.10) with a channel observation model in the form of a pilot tone. We then determine the Kalman filtering equations in order to estimate the channel, and compute the mean-square estimation error as a function of the known channel parameters. Next, we discuss an approach for obtaining estimates of the fading process when the model parameters are unknown. This approach includes a sequential parameter estimator similar in principle to the Expectation-Maximization (EM) algorithm. Finally, we comment on how these ideas might be extended in order to estimate the frequency selective channel.

4.1 State-Space Model for the Fading Channel

In addition to the autoregressive evolution model

$$a[n+1] = \rho\, a[n] + v[n+1] \tag{4.1}$$

we need to define an observation model for the fading channel. In other words, we have to specify how the receiver obtains information as a channel measurement from which to estimate the fading. At a high level, we assume that the channel information we obtain is in some sense "orthogonal" to the data transmissions. For example, in the frequency nonselective, time selective channel case, a pilot tone for measuring the channel can be transmitted outside the transmission bandwidth but within the coherence bandwidth of the channel. The data and pilot tone "see" the same fading, but the additive noises in the two problems are independent because they do not overlap in frequency. This orthogonality is important because it allows us to examine the problem of estimating the channel independently of the problem of decoding the symbols.

In particular, suppose the pilot-tone is a sinusoid with center frequency just outside the combined transmission bandwidth and the Doppler spread $B_d \propto 1/\tau_A$ of the channel. In baseband equivalent form, using the frequency nonselective channel response $a[k;n] = a[n]\,\delta[k]$ of (2.7), the observation sequence for this model is

$$q[n] = \sqrt{\mathcal{E}_p}\, a[n] + \tilde{w}[n] \tag{4.2}$$

where $\mathcal{E}_p$ is the energy of the pilot-tone, and $\tilde{w}[n]$ is yet another zero-mean, stationary, complex Gaussian sequence with variance $\sigma^2_{\tilde{w}} = \mathcal{N}_0$ representing the receiver thermal noise and co-channel interference in the pilot-tone bandwidth.

Equations (4.1) and (4.2) together form a minimal state-space model for the fading process $a[n]$. However, we will find it helpful in later sections to utilize a state-space model based on the augmented state vector

$$\mathbf{s}[n] = \begin{bmatrix} a[n] \\ a[n-1] \end{bmatrix} \tag{4.3}$$

Combining (4.1) and (4.2), we obtain the state-space model

$$\mathbf{s}[n+1] = \mathbf{F}_{\boldsymbol{\theta}}\,\mathbf{s}[n] + \mathbf{G}\,v[n] \tag{4.4a}$$

$$q[n] = \mathbf{C}\,\mathbf{s}[n] + \tilde{w}[n] \tag{4.4b}$$

where

$$\mathbf{F}_{\boldsymbol{\theta}} = \begin{bmatrix} \rho & 0 \\ 1 & 0 \end{bmatrix}, \quad \mathbf{G} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} \sqrt{\mathcal{E}_p} & 0 \end{bmatrix} \tag{4.5}$$

and where $v[n]$ and $\tilde{w}[n]$ are mutually independent, zero-mean complex Gaussian white-noise sequences with variances σ_v^2 and $\sigma_{\tilde{w}}^2$, respectively. This model captures the actual evolution of the fading process $a[n]$ as well as the generation of measurements $q[n]$. We note that the state-space model is parameterized by three parameters, namely, the time constant ρ, the process noise variance σ_v^2, and the additive noise variance $\sigma_{\tilde{w}}^2$, which we collect into the vector

$$\boldsymbol{\theta} = \begin{bmatrix} \rho \\ \sigma_v^2 \\ \sigma_{\tilde{w}}^2 \end{bmatrix} \tag{4.6}$$

4.2 Recursive MMSE Channel Estimator

To estimate the fading process state $\mathbf{s}[n]$ from the received pilot-tone sequence $q[n]$, we utilize the Kalman filter [6] based on the state-space model with parameters $\boldsymbol{\theta}$, *i.e.*, the state-space model of (4.4). By filtering the observed sequence $q[k]$ for $k \leq n$, the Kalman filter produces the estimate[1] $\hat{\mathbf{s}}_{\boldsymbol{\theta}}[n|n]$ which corresponds to the MMSE estimate of the state $\mathbf{s}[n]$.

Before presenting the Kalman filtering equations, we first define two useful shorthands for the conditional mean and covariance of the fading state vector $\mathbf{s}[n]$ based on the state-

[1] We utilize the subscript notation $\hat{\mathbf{s}}_{\boldsymbol{\theta}}[n|n]$ to indicate that the estimates are based on the model with the true parameters $\boldsymbol{\theta}$. In a later section, this notation will allow us to discuss the estimates $\hat{\mathbf{s}}_{\hat{\boldsymbol{\theta}}}[n|n]$ based on the model with estimated parameters $\hat{\boldsymbol{\theta}}$.

space model with parameters $\boldsymbol{\theta}$, namely,

$$\hat{\mathbf{s}}_{\boldsymbol{\theta}}[n|m] = \mathrm{E}\left[\, \mathbf{s}[n] \;\middle|\; q[l],\ l \leq m \right] \tag{4.7a}$$

$$\boldsymbol{\Sigma}_{\boldsymbol{\theta}}[n|m] = \mathrm{E}\left[(\mathbf{s}[n] - \hat{\mathbf{s}}_{\boldsymbol{\theta}}[n|m])\,(\mathbf{s}[n] - \hat{\mathbf{s}}_{\boldsymbol{\theta}}[n|m])^{\dagger} \;\middle|\; q[l],\ l \leq m \right] \tag{4.7b}$$

We initialize the filter by setting

$$\hat{\mathbf{s}}_{\boldsymbol{\theta}}[0|-1] = \mathbf{0} \tag{4.8a}$$

$$\boldsymbol{\Sigma}_{\boldsymbol{\theta}}[0|-1] = \frac{\sigma_v^2}{1-\rho^2} \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \tag{4.8b}$$

the *a priori* mean and covariance of the fading process, respectively. The Kalman filter is described by two sets of equations, namely, the update equations and the prediction equations, which we give next.

4.2.1 Update Equations

Given $\hat{\mathbf{s}}_{\boldsymbol{\theta}}[n|n-1]$, $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}[n|n-1]$, and the observation $q[n]$, we form the updated estimate $\hat{\mathbf{s}}_{\boldsymbol{\theta}}[n|n]$ and its corresponding covariance $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}[n|n]$ via

$$\hat{\mathbf{s}}_{\boldsymbol{\theta}}[n|n] = \hat{\mathbf{s}}_{\boldsymbol{\theta}}[n|n-1] + \mathbf{K}_{\boldsymbol{\theta}}[n]\,(q[n] - \mathbf{C}\,\hat{\mathbf{s}}_{\boldsymbol{\theta}}[n|n-1]) \tag{4.9a}$$

$$\boldsymbol{\Sigma}_{\boldsymbol{\theta}}[n|n] = (\mathbf{I} - \mathbf{K}_{\boldsymbol{\theta}}[n]\,\mathbf{C})\,\boldsymbol{\Sigma}_{\boldsymbol{\theta}}[n|n-1] \tag{4.9b}$$

with

$$\mathbf{K}_{\boldsymbol{\theta}}[n] = \boldsymbol{\Sigma}_{\boldsymbol{\theta}}[n|n-1]\,\mathbf{C}^{\dagger}\left(\mathbf{C}\,\boldsymbol{\Sigma}_{\boldsymbol{\theta}}[n|n-1]\,\mathbf{C}^{\dagger} + \sigma_{\tilde{w}}^2\right)^{-1} \tag{4.9c}$$

4.2.2 Prediction Equations

Next, given $\hat{\mathbf{s}}_{\boldsymbol{\theta}}[n|n]$ and $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}[n|n]$ as computed via (4.9), we compute the predicted estimate $\hat{\mathbf{s}}_{\boldsymbol{\theta}}[n+1|n]$ and its corresponding covariance $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}[n+1|n]$ via

$$\hat{\mathbf{s}}_{\boldsymbol{\theta}}[n+1|n] = \mathbf{F}_{\boldsymbol{\theta}}\, \hat{\mathbf{s}}_{\boldsymbol{\theta}}[n|n] \tag{4.10a}$$

$$\boldsymbol{\Sigma}_{\boldsymbol{\theta}}[n+1|n] = \mathbf{F}_{\boldsymbol{\theta}}\, \boldsymbol{\Sigma}_{\boldsymbol{\theta}}[n|n]\, \mathbf{F}_{\boldsymbol{\theta}}^{\dagger} + \sigma_v^2 \mathbf{G}\mathbf{G}^{\dagger} \tag{4.10b}$$

Using these predicted estimates, the algorithm can continue with (4.9) at time $(n+1)$, as soon as the observation $q[n+1]$ is available.

We obtain an estimate of the fading at time n via

$$\hat{a}_{\boldsymbol{\theta}}[n] = \begin{bmatrix} 1 & 0 \end{bmatrix} \hat{\mathbf{s}}_{\boldsymbol{\theta}}[n|n] \tag{4.11}$$

and the corresponding conditional variance is the element in the first row and column of $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}[n|n]$, *i.e.*,

$$\Sigma_a[n] = [\ \boldsymbol{\Sigma}_{\boldsymbol{\theta}}[n|n]\ ]_{1,1} \tag{4.12}$$

The conditional mean (4.11) is the unbiased estimator which minimizes the mean-square error defined as $\sigma^2_{e[n]} = \mathrm{E}\left[|e[n]|^2\right]$, where

$$e[n] = a[n] - \hat{a}_{\boldsymbol{\theta}}[n] \tag{4.13}$$

A key result of Gaussian estimation theory is that the conditional covariance (4.7b), and therefore, (4.12), is independent of the observations $q[k]$; hence, (4.12) also represents the unconditional mean-square error $\sigma^2_{e[n]}$.

4.3 Channel Model Parameter Estimation

The Kalman filter channel estimator developed in the previous section requires the state-space model of (4.4) along with the set of parameters $\boldsymbol{\theta}$. In this section, we consider the situation in which the form of the model (4.4) is known correctly, *i.e.*, the order of the state-space model is known at the receiver; however, the parameter vector $\boldsymbol{\theta}$ is not known at the receiver, and must be estimated in order to apply the Kalman filter equations of the

previous section.

4.3.1 Expectation-Maximization Algorithm

Our approach to finding parameter estimates $\hat{\boldsymbol{\theta}}$ is similar in principle to the Expectation-Maximization (EM) algorithm [8], which iteratively computes Maximum-Likelihood (ML)[2] estimates of the parameters $\boldsymbol{\theta}$ based on the observations of the random vector

$$\mathbf{q} = [\ q[0]\ \cdots\ q[N-1]\ ]^T \tag{4.14}$$

Direct computation of the ML estimates based on the observation $\mathbf{q} = \mathbf{q}_0$, which is referred to as the *incomplete data*, often requires multidimensional optimization algorithms. On the other hand, computations of the ML estimates based on observations of the random vector

$$\begin{aligned}\mathbf{d} &= [\ \mathbf{a}^\dagger\ \mathbf{q}^\dagger\ ]^\dagger \\ &= [\ a[-1]\ a[0]\ \cdots\ a[N-1]\ q[0]\ \cdots\ q[N-1]\ ]^T\end{aligned} \tag{4.15}$$

are more straightforward, because, as we will see, estimation of $\sigma_{\tilde{w}}^2$ can be performed separately. In the nomenclature of the EM algorithm, the observation $\mathbf{d} = \mathbf{d}_0$ is called the *complete data*.

The EM algorithm forms parameter estimates in the following iterative manner given the observations $\mathbf{q} = \mathbf{q}_0$. First, assuming a previous parameter estimate, we compute the expected value of the log-likelihood function of the complete data $\mathbf{d}$. Next, we maximize over this expected log-likelihood to obtain a new parameter estimate. Mathematically, the steps of the EM algorithm are:

1. Pick an initial parameter estimate $\hat{\boldsymbol{\theta}}_0$.

2. For each iteration $k \geq 1$,

[2]We note that for our particular problem, the ML estimator is biased, and consequently not an efficient estimator. (In fact, an unbiased efficient estimator does not exist from the complete data, so it seems plausible that none exist from the incomplete data.) However, we explore ML estimation for several reasons. First, as we will see, maximization over the log-likelihoods of Gaussian random variables lead to closed-form expressions for the parameter estimates. Secondly, although a proof of these properties is beyond the scope of this thesis, ML estimators are often asymptotically efficient and consistent estimators.

Expectation Step Compute the function $U(\boldsymbol{\theta}', \hat{\boldsymbol{\theta}}_{k-1})$ defined as

$$U(\boldsymbol{\theta}', \hat{\boldsymbol{\theta}}_{k-1}) = \mathrm{E}\left[\log p_{\mathbf{d}}(\mathbf{d}; \boldsymbol{\theta}') \;\middle|\; \mathbf{q} = \mathbf{q}_0,\ \boldsymbol{\theta} = \hat{\boldsymbol{\theta}}_{k-1}\right] \tag{4.16}$$

Maximization Step Compute the parameter estimates $\hat{\boldsymbol{\theta}}_k$ for the next iteration via

$$\hat{\boldsymbol{\theta}}_k = \arg\max_{\boldsymbol{\theta}'} U(\boldsymbol{\theta}', \hat{\boldsymbol{\theta}}_{k-1}) \tag{4.17}$$

By maximizing over (4.16), the expected value of the log-likelihood of the complete data, the EM algorithm increases the log-likelihood of the incomplete data with each iteration. In fact, the EM algorithm is guaranteed to converge to a local minimum of the log-likelihood of the incomplete data [8]. Consequently, it may be necessary to use several initial conditions for the algorithm, with the hope that one of the trajectories reaches the global maximum corresponding to the ML estimates.

In the following two sections, we compute (4.16) and (4.17) for our problem by following the development of [8], and we point out how the batch structure and resulting computational complexity of the exact EM algorithm are impractical for our intended application.

E-Step Computations

In this section, we compute (4.16) for our fading channel model. First, noting that

$$p_{\mathbf{d}}(\mathbf{d}_0; \boldsymbol{\theta}') = p_{\mathbf{a}}(\mathbf{a}_0; \boldsymbol{\theta}')\ p_{\mathbf{q}|\mathbf{a}}(\mathbf{q}_0|\mathbf{a}_0; \boldsymbol{\theta}') \tag{4.18}$$

we obtain

$$\log\left[p_{\mathbf{d}}(\mathbf{d}; \boldsymbol{\theta}')\right] = \log\left[p_{\mathbf{a}}(\mathbf{a}; \boldsymbol{\theta}')\right] + \log\left[p_{\mathbf{q}|\mathbf{a}}(\mathbf{q}|\mathbf{a}; \boldsymbol{\theta}')\right] \tag{4.19}$$

Because both the fading and the additive noise are complex Gaussian, we can easily compute the log-likelihoods in (4.19). From the state evolution model (4.1), the fading process is first-order Markov, *i.e.*, given only $a[n-1]$, $a[n]$ is complex Gaussian with mean $\rho'\, a[n-1]$

and variance $(\sigma_v^2)'$. Hence[3]

$$p_{\mathbf{a}}(\mathbf{a};\boldsymbol{\theta}') = p_{a[-1]}(a_{-1}) \prod_{n=0}^{N-1} \frac{1}{\pi(\sigma_v^2)'} \; e^{-|a[n]-\rho' \, a[n-1]|^2/(\sigma_v^2)'} \tag{4.20}$$

Also, because the additive noise $\tilde{w}[n]$ is zero-mean, white and complex Gaussian, we have that given $a[n]$, $q[n]$ is complex Gaussian with mean $\sqrt{\mathcal{E}_p}\, a[n]$ and variance $(\sigma_{\tilde{w}}^2)'$. Consequently,

$$p_{\mathbf{q}|\mathbf{a}}(\mathbf{q}|\mathbf{a};\boldsymbol{\theta}') = \prod_{n=0}^{N-1} \frac{1}{\pi(\sigma_{\tilde{w}}^2)'} \; e^{-\left|q[n]-\sqrt{\mathcal{E}_p}\; a[n]\right|^2/(\sigma_{\tilde{w}}^2)'} \tag{4.21}$$

Combining (4.20) with (4.21) into (4.19), we find

$$\begin{aligned}
&\log\left[p_{\mathbf{d}}(\mathbf{d};\boldsymbol{\theta}')\right] = \\
&C - N\log(\sigma_v^2)' - \frac{1}{(\sigma_v^2)'} \sum_{n=0}^{N-1} \left[|a[n]|^2 - 2\rho' \mathrm{Re}\left\{a[n]a^*[n-1]\right\} + (\rho')^2 \, |a[n-1]|^2\right] \\
&\quad - N\log(\sigma_{\tilde{w}}^2)' - \frac{1}{(\sigma_{\tilde{w}}^2)'} \sum_{n=0}^{N-1} \left[|q[n]|^2 - 2\sqrt{\mathcal{E}_p}\; \mathrm{Re}\left\{q[n]a^*[n]\right\} + \mathcal{E}_p \;\; |a[n]|^2\right]
\end{aligned} \tag{4.22}$$

where C is a constant, independent of $\boldsymbol{\theta}'$. Now we take the expected value of both sides of (4.22), conditional on the observation $\mathbf{q} = \mathbf{q}_0$ and assuming that the parameters are $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}_{k-1}$. Using the shorthand notation

$$\overline{(\cdot)} = \mathrm{E}\left[\; \cdot \; \middle| \; \mathbf{q} = \mathbf{q}_0, \;\; \boldsymbol{\theta} = \hat{\boldsymbol{\theta}}_{k-1} \; \right] \tag{4.23}$$

we find that

$$\begin{aligned}
&U(\boldsymbol{\theta}', \hat{\boldsymbol{\theta}}_{k-1}) = \\
&C - N\log(\sigma_v^2)' - \frac{1}{(\sigma_v^2)'} \sum_{n=0}^{N-1} \left[\overline{|a[n]|^2} - 2\rho' \mathrm{Re}\left\{\overline{a[n]a^*[n-1]}\right\} + (\rho')^2 \overline{|a[n-1]|^2}\right] \\
&\quad - N\log(\sigma_{\tilde{w}}^2)' - \frac{1}{(\sigma_{\tilde{w}}^2)'} \sum_{n=0}^{N-1} \left[|q[n]|^2 - 2\sqrt{\mathcal{E}_p}\, \mathrm{Re}\left\{q[n]\overline{a^*[n]}\right\} + \mathcal{E}_p \, \overline{|a[n]|^2}\right]
\end{aligned} \tag{4.24}$$

Equation (4.24) requires computation of the conditional mean vector and (certain ele-

[3]We assume that $p_{a[-1]}(a_{-1})$ is independent of $\boldsymbol{\theta}$ to simplify the M-step computations. This assumption might create an edge effect, but it dies away for N large.

ments of the) correlation matrix of $\mathbf{a}$ given observation $\mathbf{q} = \mathbf{q}_0$, assuming the parameters are $\hat{\boldsymbol{\theta}}_{k-1}$. In fact, if we are willing to accept noisier estimates due to lack of smoothing, the Kalman filter based on the state-space model with parameters $\hat{\boldsymbol{\theta}}_{k-1}$ yields exactly the required estimates. Hence, given the parameter estimate, we can compute (approximately) the quantities $\overline{a[n]}$, $\overline{a[n]a^*[n-1]}$, and $\overline{|a[n]|^2}$ needed above to compute (4.24) by filtering the incomplete data vector $\mathbf{q}_0$ with the Kalman filter of the previous section.

Since this Kalman filter gives the estimates $\hat{\mathbf{s}}_{\hat{\boldsymbol{\theta}}_{k-1}}[n|n]$ with corresponding covariances $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}_{k-1}}[n|n]$, the required correlation matrix is

$$\mathbf{R}_{\hat{\boldsymbol{\theta}}_{k-1}}[n|n] = \boldsymbol{\Sigma}_{\hat{\boldsymbol{\theta}}_{k-1}}[n|n] + \hat{\mathbf{s}}_{\hat{\boldsymbol{\theta}}_{k-1}}[n|n]\,\hat{\mathbf{s}}^{\dagger}_{\hat{\boldsymbol{\theta}}_{k-1}}[n|n] \tag{4.25}$$

Then we have the following relationships, for $0 \le n \le N$,

$$\overline{a[n]} = \left[\hat{\mathbf{s}}_{\hat{\boldsymbol{\theta}}_{k-1}}[n|n]\right]_1 \tag{4.26a}$$

$$\overline{a[n]a^*[n-1]} = \left[\mathbf{R}_{\hat{\boldsymbol{\theta}}_{k-1}}[n|n]\right]_{1,2} \tag{4.26b}$$

$$\overline{|a[n]|^2} = \left[\mathbf{R}_{\hat{\boldsymbol{\theta}}_{k-1}}[n|n]\right]_{1,1} \tag{4.26c}$$

$$\overline{|a[n-1]|^2} = \left[\mathbf{R}_{\hat{\boldsymbol{\theta}}_{k-1}}[n|n]\right]_{2,2} \tag{4.26d}$$

More importantly, we see that estimates of the fading itself are required in order to estimate the parameters, and vise versa. Hence, the EM algorithm really gives us a joint parameter and state estimator for the fading channel. More generally, the EM algorithm can be viewed as a parameter estimation algorithm that produces estimates of the underlying symbols as side information. However, it is worth emphasizing that in our application, it is these signal estimates in which we are ultimately interested.

M-Step Computations

Now that we have determined $U(\boldsymbol{\theta}', \hat{\boldsymbol{\theta}}_{k-1})$ (the E-step) for our problem, we must compute $\hat{\boldsymbol{\theta}}_k$ according to (4.17). We solve for these estimates by setting

$$\left[\frac{d}{d\boldsymbol{\theta}'}\,U(\boldsymbol{\theta}', \hat{\boldsymbol{\theta}}_{k-1})\right]\Bigg|_{\boldsymbol{\theta}'=\hat{\boldsymbol{\theta}}_k} = \mathbf{0} \tag{4.27}$$

and solving. Substitution of (4.24) into (4.27) yields, after a bit of algebra,

$$\hat{\rho}_k = \left(\sum_{n=0}^{N-1} \mathrm{Re}\left\{ \overline{a[n]a^*[n-1]} \right\} \right) \Bigg/ \left(\sum_{n=0}^{N-1} \overline{|a[n-1]|^2} \right) \tag{4.28a}$$

$$\widehat{(\sigma_v^2)}_k = \frac{1}{N} \sum_{n=0}^{N-1} \left[\; \overline{|a[n]|^2} - 2\hat{\rho}_k \mathrm{Re}\left\{ \overline{a[n]a^*[n-1]} \right\} + \hat{\rho}_k^2 \overline{|a[n-1]|^2} \; \right] \tag{4.28b}$$

$$\widehat{(\sigma_{\tilde{w}}^2)}_k = \frac{1}{N} \sum_{n=0}^{N-1} \left[\; |q[n]|^2 - 2\sqrt{E_p}\ \mathrm{Re}\left\{ q[n]\overline{a^*[n]} \right\} + E_p\ \overline{|a[n]|^2} \; \right] \tag{4.28c}$$

4.3.2 Sequential EM Approximation

As we see from the structure of (4.24) and (4.28), the EM algorithm operates on a block of data in an iterative fashion. Using a previous parameter estimate $\hat{\boldsymbol{\theta}}_{k-1}$, (4.24) must be implemented, requiring estimates of the entire fading vector $\mathbf{a}$ and parts of its correlation matrix. A new parameter estimate $\hat{\boldsymbol{\theta}}_k$ is obtained via (4.28), and the process may be continued until some suitable level of convergence is obtained. Presumably, both the parameter and state estimates become more accurate with more and more iterations of the algorithm.

Unfortunately, there are two practical problems with the EM approach for our intended application. First and foremost, computer experiments have demonstrated slow convergence of this algorithm requiring many iterations, in which we must compute the entire set of signal estimates. Secondly, sequential implementations of the above algorithm are not known. In other words, when we obtain a new piece of data, say $q[N+1]$, it is not known how one could incorporate this data with previous parameter estimates to directly form new parameter estimates. Instead, we must perform a number of iterations, re-estimating the entire length $(N+1)$ sequence, which is impractical.

As an alternative, we combine the notions of iterations k and time n, to form a sequential algorithm such as that developed in [9]. In particular, upon the observation $q[n]$, we use the previous parameter estimate $\hat{\boldsymbol{\theta}}_{n-1}$ to iterate the Kalman filter equations (4.9) and (4.10) (with $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}_{n-1}$), thereby obtaining the estimates (4.26) (with $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}_{n-1}$). We then form

the updated parameter estimates according to a sequential version of (4.28), namely,

$$\hat{\rho}_n = R_2[n]/R_3[n] \tag{4.29a}$$

$$\widehat{(\sigma_v^2)}_n = \frac{1}{N}\left[R_1[n] - 2\hat{\rho}_n R_2[n] + \hat{\rho}_n^2 R_3[n]\right] \tag{4.29b}$$

$$\widehat{\left(\sigma_{\tilde{w}}^2\right)} = \frac{1}{N} R_4[n] \tag{4.29c}$$

in which

$$R_1[n] = R_1[n-1] + \overline{|a[n]|^2} \tag{4.30a}$$

$$R_2[n] = R_2[n-1] + \mathrm{Re}\left\{\overline{a[n]a^*[n-1]}\right\} \tag{4.30b}$$

$$R_3[n] = R_3[n-1] + \overline{|a[n-1]|^2} \tag{4.30c}$$

$$R_4[n] = R_4[n-1] + |q[n]|^2 - 2\sqrt{\mathcal{E}_p}\mathrm{Re}\left\{q[n]\overline{a^*[n]}\right\} + \mathcal{E}_p\overline{|a[n]|^2} \tag{4.30d}$$

Of course, the above algorithm is not an EM algorithm, so we cannot expect that $\hat{\boldsymbol{\theta}}_n$ converges over time to the ML estimates. However, this algorithm is more suitable for our application because of its sequential nature and dramatic computational reduction (per sample) compared the EM algorithm of the previous section.

4.4 Estimator Mean-Square Error Performance

As is usually done, we characterize the performance of the fading channel estimator in terms of its mean-square error. That is, defining the error sequence as

$$e[n] = a[n] - \hat{a}_{\boldsymbol{\theta}}[n] \tag{4.31}$$

the mean-square error is the quantity

$$\mathrm{MSE}[n] = \mathrm{E}\left[|e[n]|^2\right] \tag{4.32}$$

Even when the model parameters are known at the receiver, the error sequence $e[n]$ at the output of the Kalman filter is nonstationary because we fix a time to begin estimating the stationary process $a[n]$. However, we can expect the filter to reach steady-state behavior

(and $e[n]$ to become stationary) as the length of our observation time becomes large. Similarly, we hope that the parameter estimates converge, so that the error sequence becomes at least stationary, if not unbiased as well. In this case, we can characterize the performance of the Kalman filter estimator in terms of its steady-state mean-square error, defined as

$$\mathrm{MSE}_\infty = \lim_{n\to\infty} \mathrm{MSE}[n] \tag{4.33}$$

when the limit exists. We may then examine the signal-to-estimation-noise ratio (SENR) for the estimator defined as

$$\mathrm{SENR} = \frac{\sigma_A^2}{\mathrm{MSE}_\infty} \tag{4.34}$$

and investigate how this normalized quantity behaves as a function of n, τ_A, and the received pilot-channel SNR defined as

$$\mathrm{SNR_p} = \frac{\mathcal{E}_p\,\sigma_A^2}{\mathcal{N}_0} \tag{4.35}$$

4.4.1 Performance with Known Parameters

As we mentioned previously, when the model parameters are known, the channel estimator is unbiased and the mean-square error corresponds to the error variance. In this case, we use the typical approach of iterating the Kalman filter covariance equations (4.9b) and (4.10b), beginning with the initial condition (4.8b), to solve for the steady-state mean-square error. This approach also allows us to characterize how the steady-state mean-square error decays with time.

Inspection of the results of SENR versus n indicates the following. First, the rate of convergence to the steady-state solution goes up with increasing $\mathrm{SNR_p}$, for any choice of the time constant ρ. Thus we plot in Figure 4-1 the time evolution for $\mathrm{SNR_p} = 0$ dB, since these should be worst case results in the range of receiver input $\mathrm{SNR_p}$ we consider. Secondly, we see from the figure that the convergence rate goes down as ρ becomes closer to one. Nevertheless, we expect the Kalman filter to reach steady-state in roughly 25 samples in the worst case. Fast convergence results such as these allow us to treat the Kalman filter as an LTI processor, corresponding to the causal Wiener filter, in its steady-state.

Finally, Figure 4-2 displays the results for SENR versus $\mathrm{SNR_p}$. Again, several choices

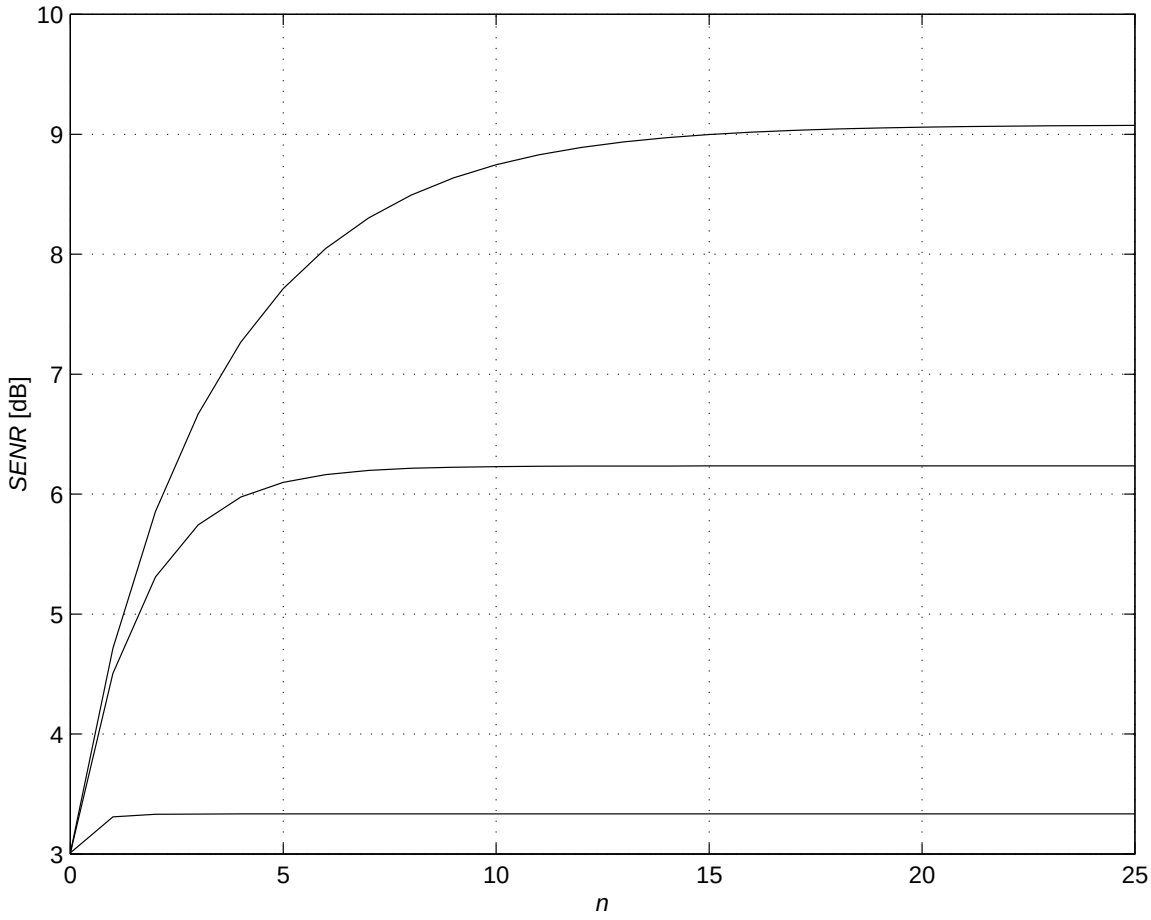


Figure 4-1: Evolution of the channel estimator SENR over time for the case in which the model parameters are known and for $\mathrm{SNR_p} = 0$ dB. The successively lower curves correspond to $\rho = 0.99$, $\rho = 0.95$, and $\rho = 0.5$, respectively.

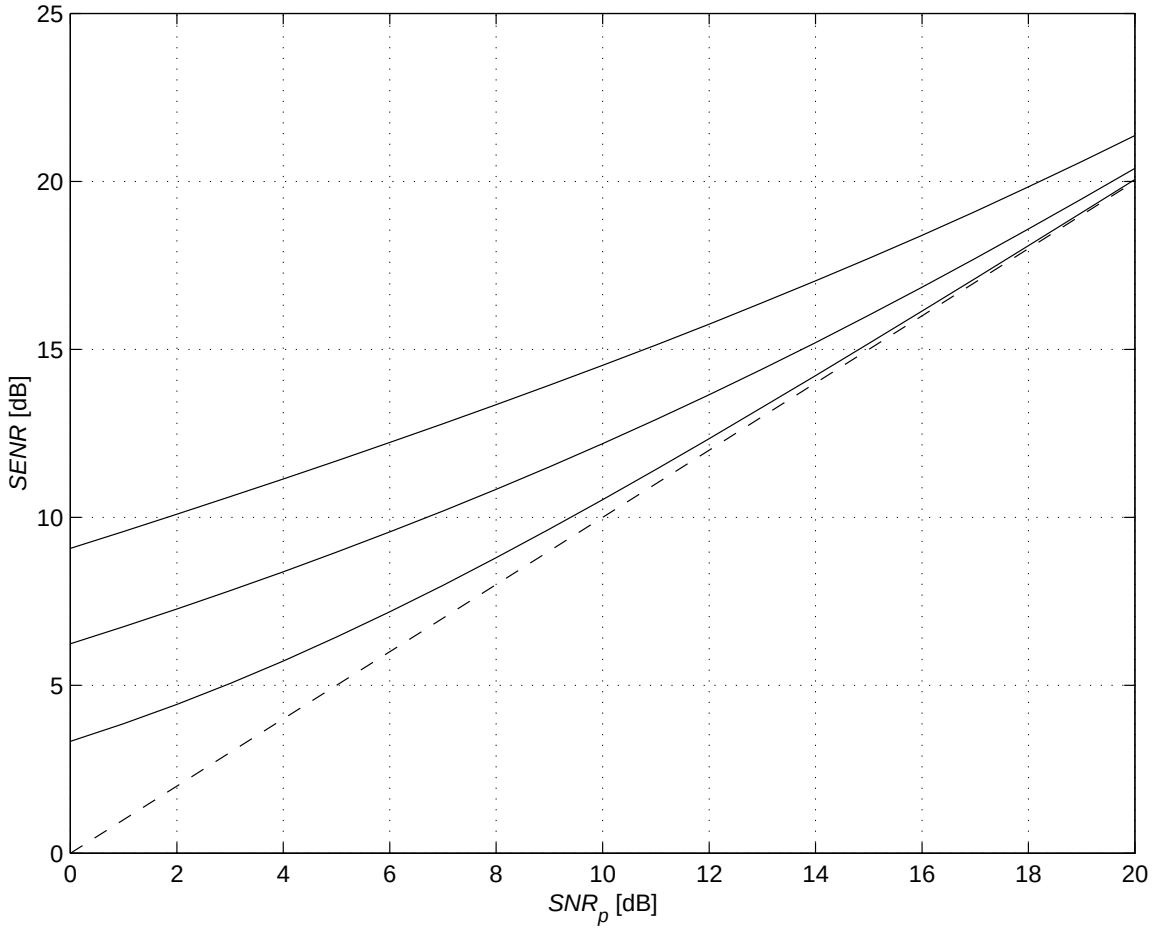


Figure 4-2: Channel estimator SENR versus $\mathrm{SNR_p}$ for the case in which the model parameters are known. The successively lower solid curves correspond to $\rho = 0.99$, $\rho = 0.95$, and $\rho = 0.5$, respectively. The dashed curve is the curve $\mathrm{SENR} = \mathrm{SNR_p}$, which we see gives a lower bound for the estimator performance.

of ρ are displayed on the plot. As these results reflect,

$$\text{SENR} \geq \text{SNR}_\text{p} \tag{4.36}$$

for all choices of ρ, and this lower bound becomes tight for high SNR_p.

4.4.2 Performance with Unknown Parameters

In this section, we give results similar to those of the previous section but for the case in which the model parameters are unknown at the receiver. Because we expect the channel estimator to be biased in this case, we examine both error bias and error variance of the estimation algorithm through Monte Carlo simulations.

Empirical averages for the error sequence $e[n]$ indicate that the estimator remains approximately unbiased when the parameter estimates are unknown at the receiver. We make this conclusion based on the observation that the numerical error variance and mean-square error were very similar, differing by no more than 1 dB.

Figure 4-3 illustrates the results of our empirical averages for the output SENR versus the number of observations n. Again, we see that the convergence rate decreases as ρ becomes closer to one. We also note a loss in the output SENR due to parameter estimation.

Finally, we present in Figure 4-4 preliminary results of the mean-square estimation error for the joint parameter and state estimator versus SNR_p.

4.5 Frequency Selective Fading Channel Estimation

In this section we comment on how to extend the algorithms of this chapter to the case of a frequency selective channel. Following the frequency nonselective case, we give a state-space characterization of the frequency selective channel and then discuss a model-based channel estimation scheme.

4.5.1 State-Space Model

Similar to the frequency nonselective case, we wish to obtain channel measurements which are orthogonal to the transmit symbol sequence. For the frequency selective channel, a pilot-tone is not suitable because only a certain frequency can be measured. Instead, we

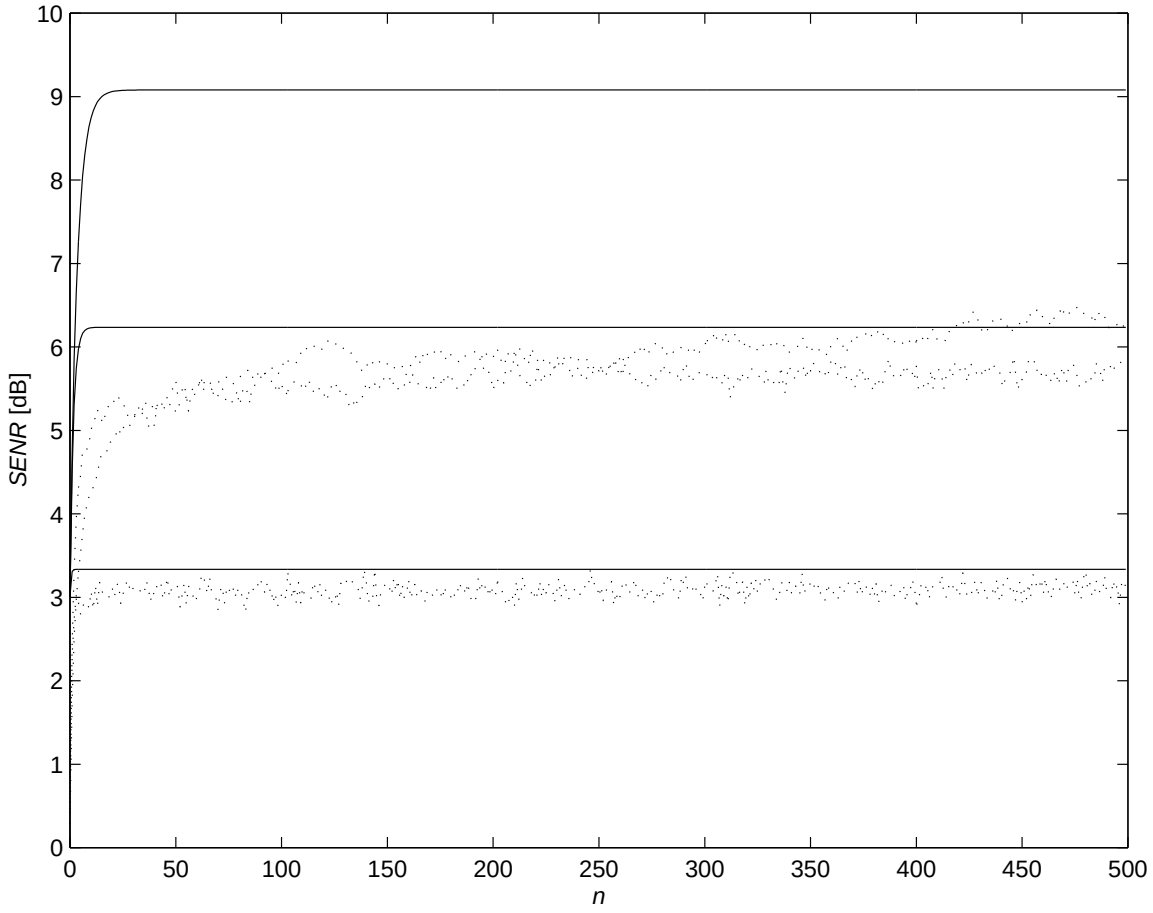


Figure 4-3: Evolution of the channel estimator SENR over time for the case in which the model parameters are estimated and for $\mathrm{SNR}_{\mathrm{p}} = 0$ dB. The successively lower dotted curves corresponds to empirical averages of the mean-square error for the values $\rho = 0.99$, $\rho = 0.95$, and $\rho = 0.5$, respectively. For comparison, the successively lower solid curves are copied from Figure 4-1, where the parameter are assumed known at the receiver.

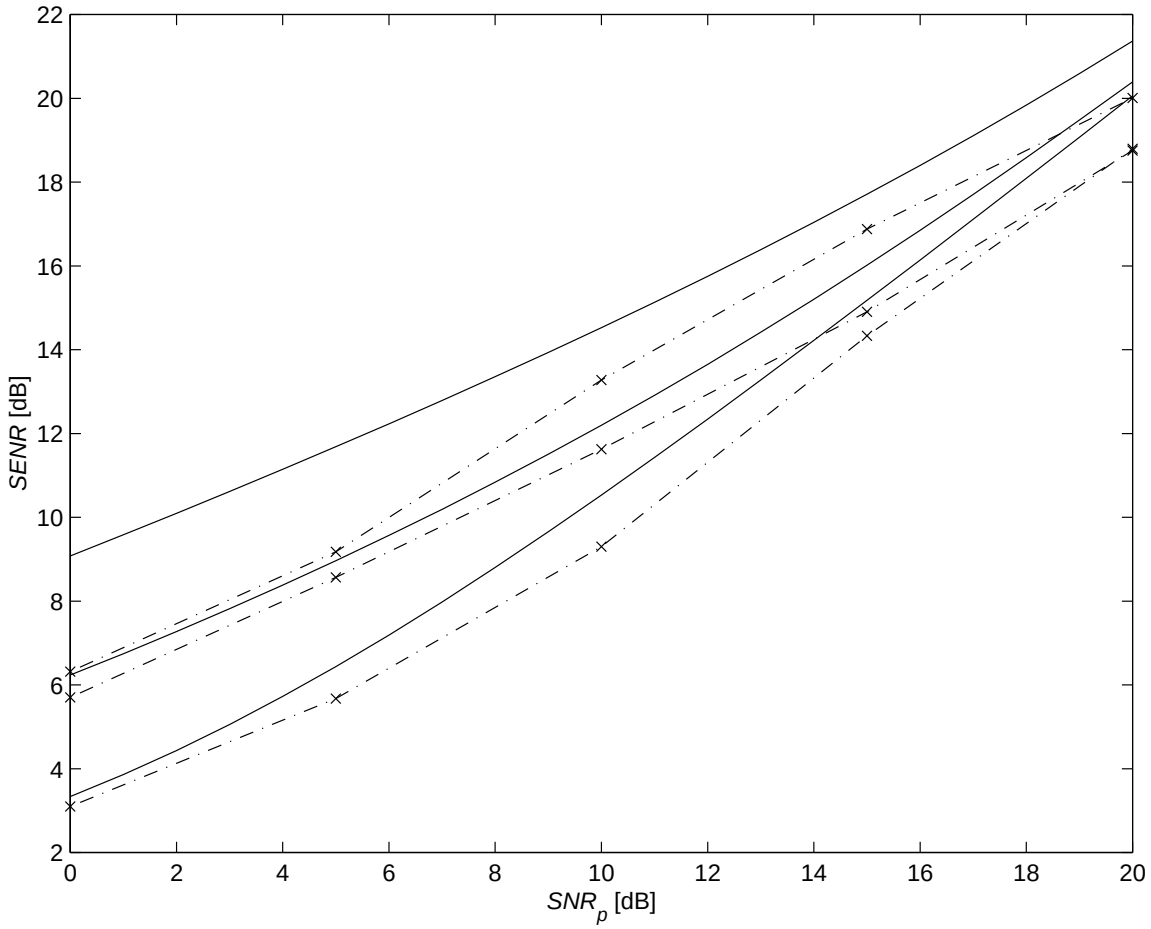


Figure 4-4: Channel estimator SENR versus $\mathrm{SNR_p}$ for the case in which the model parameters are estimated. The successively lower dash-dotted curves correspond to the mean-square error performance of the joint parameter and state estimator for values $\rho = 0.99$, $\rho = 0.95$, and $\rho = 0.5$, respectively. The successively lower solid curves corresponding to known model parameters are copied from Figure 4-2 for comparison.

separate the channel measurements and transmit sequence in time by using training periods with pilot symbols. Specifically, over a block of length M, the first T symbols transmitted over the channel are training symbols, while the remaining $M - T$ symbols are data. We note that such a scheme should be useful when the channel has coherence time $\tau_A > M$. We also note that this scheme, as with a pilot-tone, increases the bandwidth required for transmissions at a fixed rate. In both cases, it is necessary to properly divide available power resources between the data and pilot channels.

Mathematically, the above properties imply that the training sequence $t[n]$ satisfies

$$t[n] = 0 \quad \text{for } lT \leq n < lM,\ l = \ldots, -1, 0, 1, \ldots \tag{4.37}$$

Collecting this sequence into the vector

$$\mathbf{t}[n] = [\, t[n]\ t[n-1]\ \cdots\ t[n-K+1] \,]^T$$

whose length K corresponds to the length of the fading channel impulse response, our observation model for the channel becomes

$$q[n] = \mathbf{a}^T[n]\, \mathbf{t}[n] + \tilde{w}[n] \tag{4.38}$$

where $\mathbf{a}[n]$ is defined as in (2.15a). Coupled with the evolution model of (2.16), we obtain the complete state-space model

$$\mathbf{a}[n+1] = \mathbf{P}\ \mathbf{a}[n] + \mathbf{v}[n+1] \tag{4.39a}$$

$$q[n] = \mathbf{t}^T[n]\, \mathbf{a}[n] + \tilde{w}[n] \tag{4.39b}$$

4.5.2 Channel Estimator

Based on the model (4.39) and the observations $q[n]$, it is straightforward to write down the Kalman filter for estimating the state vector $\mathbf{a}[n]$ based on a model with parameters

$$\boldsymbol{\theta} = \begin{bmatrix} \rho_0 \\ \vdots \\ \rho_{K-1} \\ \sigma^2_{v_0} \\ \vdots \\ \sigma^2_{v_{K-1}} \\ \sigma^2_{\tilde{w}} \end{bmatrix}$$

When the parameter vector $\boldsymbol{\theta}$ is unknown at the receiver, we may utilize the results of the previous section to obtain these parameter estimates. Specifically, we first estimate the fading tap vector $\mathbf{a}[n]$. Then since each of the taps are uncorrelated, we may apply the estimation equations of the previous section to each of them. A detailed study of this technique is beyond the scope of this thesis, but it represents an interesting topic for further research.

4.6 Summary

In this chapter, we have developed a technique for obtaining estimates of the fading channel impulse response. While we have focused exclusively on the frequency nonselective channel for brevity, we have commented on how these ideas might be extended to the frequency selective case. To recap, the channel estimator consists of a Kalman filter for estimating the fading state vector. When the parameters of the state-space model are unknown, they too must be estimated to apply the recursive estimator equations. Because of the difficulties inherent in obtaining true ML parameter estimates, we utilize a sequential algorithm which performs satisfactorily and with significantly reduced complexity. We have summarized the estimator performance in terms of SENR versus both the number of observations n and the signal-to-noise ratio SNR_p.

Chapter 5

System Performance Using Channel Estimates

In this chapter, we examine the performance of spread-response precoding systems when only channel estimates of the form developed in Chapter 4 are available at the receiver. We begin by reviewing some of the properties of the (approximate) MMSE channel estimates, and we use these properties in the derivation of the optimal equalizer along the lines of Section 3.1. We then compute the average output SNR and predict the bit-error performance using the additive Gaussian white-noise approximation for the composite channel. Finally, we confirm these predicted results through simulation. As has usually been the case, we consider only the frequency nonselective fading channel in detail; consequently, the equalizers for the two criteria are equivalent. We also comment on how to extend the results to the frequency selective channel.

5.1 Optimal Linear Equalizers Revisited

In this section, we develop the optimal linear equalizer (both in term of maximum SNR and linear, MMSE) based on MMSE channel estimates. We begin our discussion by reviewing several properties of the channel estimate, and then determine the optimal equalizer.

5.1.1 MMSE Channel Estimator Properties

The channel estimator developed in Chapter 4 has several important properties that we may utilize to simplify the calculation of the maximum SNR equalizer based on channel estimates. Specifically, when the model parameters are known at the receiver, the channel estimate $\hat{a}_{\boldsymbol{\theta}}[n]$ is unbiased and has minimum mean-square error. Furthermore, the results of Chapter 4 indicate that these properties hold approximately for the channel estimate $\hat{a}_{\hat{\boldsymbol{\theta}}}[n]$ based on estimates of the model parameters. In the remainder of the thesis, we develop our analytical results for the channel estimate based on the true model parameters, with the understanding that these results hold approximately for the latter case. Consequently, we eliminate the subscript $\boldsymbol{\theta}$ notation without risk of confusion, and make it clear from the context which estimation technique is employed.

We can express the properties of the channel estimate more concisely in terms of the error sequence

$$e[n] = a[n] - \hat{a}[n] \tag{5.1}$$

Mathematically, we have that the error sequence (5.1) is zero-mean

$$\mathrm{E}\left[e[n]\right] = 0 \tag{5.2}$$

for all n. Furthermore, because the estimator is linear and has minimum mean-square error, $e[n]$ is orthogonal to the estimate $\hat{a}[n]$, *i.e.*,

$$\mathrm{E}\left[\hat{a}[n]e^*[n]\right] = 0 \tag{5.3}$$

for each n. Since $a[n]$ is Gaussian, the condition (5.3) implies that $\hat{a}[n]$ and $e[n]$ are independent. This result is useful since we constrain our equalizer $b[n]$ to rely only on the estimate $\hat{a}[n]$, so that we also have

$$\mathrm{E}\left[b[n]e^*[n]\right] = 0 \tag{5.4}$$

for each n. Also by independence of $\hat{a}[n]$ and $e[n]$, we see that

$$\sigma_{\hat{A}}^2 = \sigma_A^2 - \sigma_E^2 \tag{5.5}$$

5.1.2 Derivation of the Optimal Equalizers

As we mentioned in Chapter 2, the composite channel consisting of the precoder, fading channel, equalizer, and postcoder is well approximated as an additive, marginally Gaussian white-noise channel when the length of the precoding filter becomes large. Asymptotically, the output SNR for this channel is

$$\gamma(b) = \frac{|\mathrm{E}\,[ab]|^2}{\mathrm{Var}\,[ab] + \xi_0\,\mathrm{E}\left[|b|^2\right]} \tag{5.6}$$

where $\xi_0 = \mathcal{N}_0/\mathcal{E}_s$ as in (2.31). Note that we have dropped the dependence of these statistics on n due to stationarity. As in Chapter 3, we wish to find the linear equalizer $b[n]$ which maximizes this quantity; however, in this section, we consider the effects of the unknown error sequence (5.1) in the analysis.

Again following the development of [2], we first rewrite $\gamma(b)$ as

$$\gamma(b) = \frac{1}{1/\phi(b) - 1} \tag{5.7}$$

where

$$\phi(b) = \frac{|\mathrm{E}\,[ab]|^2}{\mathrm{E}\left[\left(|a|^2 + \xi_0\right)|b|^2\right]} \tag{5.8}$$

and in which we have exploited the identity

$$\mathrm{Var}\,[ab] = \mathrm{E}\left[|ab|^2\right] - |\mathrm{E}\,[ab]|^2 \tag{5.9}$$

We note that maximizing $\phi(b)$ is equivalent to maximizing $\gamma(b)$.

Next we make use of our relationship between the true and estimated channel responses, by substituting

$$a[n] = \hat{a}[n] + e[n]$$

into (5.8) to obtain

$$\phi(b) = \frac{|\mathrm{E}\,[\hat{a}b] + \mathrm{E}\,[eb]|^2}{\mathrm{E}\left[|\hat{a}b|^2\right] + 2\mathrm{Re}\left\{\mathrm{E}\left[\hat{a}e^*\,|b|^2\right]\right\} + \mathrm{E}\left[|eb|^2\right] + \xi_0\mathrm{E}\left[|b|^2\right]} \tag{5.10}$$

When we apply the properties (5.3) and (5.4), we find that (5.10) simplifies to

$$\phi(b) = \frac{\left|\mathrm{E}\left[\hat{a}b\right]\right|^2}{\mathrm{E}\left[\left(\left|\hat{a}\right|^2 + \xi'\right)\left|b\right|^2\right]} \tag{5.11}$$

with $\xi_0' = \xi_0 + \sigma_E^2$. The form (5.11) was maximized in Section 3.1. Consequently, we may use our previous result to obtain the maximum SNR equalizer

$$b[n] \propto \frac{\hat{a}^*[n]}{\left|\hat{a}[n]\right|^2 + \xi_0'} \tag{5.12}$$

We note that this equalizer is optimal in the class of equalizers constrained to depend on only the channel estimate $\hat{a}[n]$ and only the statistics of the error sequence $e[n]$.

Several observations are important at this point in the discussion. First, we see by inspection of the result (5.12) that the optimal equalizer in terms of SNR must take into account the mean-square estimation error of the channel estimator. Similarly, since the two equalizers are equivalent in the frequency nonselective case, the linear MMSE equalizer must also take into account the mean-square estimation error. Thus, as we would expect, the two problems of channel estimation and equalization are coupled. Secondly, we note that this result is consistent with the known channel case in (3.5): as $\sigma_E^2 \to 0$, we have $\hat{a}[n] \to a[n]$ and $\xi_0' \to \xi_0$, which implies (5.12) becomes (3.5). Thirdly, note that (5.12) corresponds to the optimal equalizer for the "known channel" $\hat{a}[n]$ with additive noise of intensity $\mathcal{E}_s\sigma_E^2 + \mathcal{N}_0$. This interpretation becomes very intuitive when we examine the channel as shown in Figure 5-1. Let us examine the sequence

$$\hat{w}[n] = e[n]\, y[n] \tag{5.13}$$

Because $y[n]$ is white and zero-mean, and $e[n]$ is zero-mean, we have that $\hat{w}[n]$ is a white sequence with intensity $\mathcal{E}_s\sigma_E^2$ that is uncorrelated with the input $y[n]$. Thus, the total additive noise has variance $\mathcal{E}_s\sigma_E^2 + \mathcal{N}_0$, since $\hat{w}[n]$ and $w[n]$ are independent. We also point out that even though $\hat{w}[n]$ is not a Gaussian sequence, treating it as another additive noise term is valid in our analysis because we only examine second-order statistics in computing and maximizing the composite channel SNR. Furthermore, as Wornell points out in [2], the fact that $\hat{w}[n]$ is not Gaussian does not change the fact that the composite channel can still

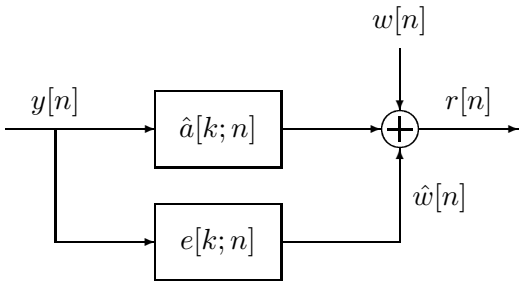


Figure 5-1: Interpretation of the fading channel as a "known" estimated channel $\hat{a}[k;n]$ in parallel with an"unknown" estimation-error channel $e[k;n]$.

be well approximated as an additive marginally Gaussian white-noise channel. We shall make use of this result in the next section.

As a final comment, we note that similar results can in principle be obtained for the frequency selective channel if we assume that the error kernel $e[k;n]$ corresponds to a wide-sense stationary, uncorrelated scattering channel. Then again, the sequence $\hat{w}[n]$ is a white sequence that is uncorrelated with the input sequence $y[n]$, and all of the results cited in this chapter apply. However, the validity of this uncorrelated scattering assumption for the error kernel warrants further investigation.

5.2 Average Output SNR and Bit-Error Probability

The last section demonstrated that the optimal equalizer (5.12) must take into account the mean-square estimation error of the channel estimator. In this section, we compute the average output SNR for this optimal equalizer, and compare it to the equalizer which uses only the channel estimate $\hat{a}[n]$ and does not account for mean-square estimation error, *i.e.*, the equalizer

$$b[n] = \frac{\hat{a}^*[n]}{|\hat{a}[n]|^2 + \xi_0} \tag{5.14}$$

First, when the optimal equalizer (5.12) is used, we may utilize the results of Appendix A with $a = \hat{a}$, b given by (5.12), and $\eta = \eta' = \xi_0'$. The maximum output SNR is given by

$$\gamma_{\max} = \frac{1}{\zeta_0' e^{\zeta_0'} E_1(\zeta_0')} - 1 \tag{5.15}$$

where $\zeta_0' = \xi_0'/\sigma_{\hat{A}}^2$. Again, we note that as $\sigma_E^2 \to 0$, (5.15) reduces to the previous result found in (3.12).

For the mismatched equalizer (5.14), we obtain the output SNR using the computations of Appendix A with $a = \hat{a}$, b given by (5.14), $\eta = \xi_0'$, and $\eta' = \xi_0''$. The result is

$$\gamma_{\mathrm{m}} = \frac{1}{1/\phi_{\mathrm{m}} - 1} \tag{5.16}$$

in which

$$\phi_{\mathrm{m}} = \frac{\left|1 - \zeta_0'' e^{\zeta_0''} E_1(\zeta_0'')\right|^2}{1 - \zeta_0'' e^{\zeta_0''} E_1(\zeta_0'') + (\zeta_0' - \zeta_0'')\left[e^{\zeta_0''} E_1(\zeta_0'')(1 + \zeta_0'') - 1\right]} \tag{5.17}$$

and where $\zeta_0'' = \xi_0/\sigma_{\hat{A}}^2$.

We are ultimately interested in the bit-error performance of the spread-response precoding systems which employ the equalizers (5.12) and (5.14). To compare their bit-error performance, we make use of the fact that the composite channels can be well approximated as additive marginally Gaussian white-noise channels, with average output SNRs given by (5.15) and (5.16), respectively. Then we approximate the of a bit-error probability using (2.32).

Implicitly, the performance of these systems depends on the effectiveness of the channel estimation algorithm. Recall that in Chapter 4 we characterized the performance of the channel estimation algorithm in terms of the normalized quantity

$$\mathrm{SENR} = \frac{\sigma_A^2}{\mathrm{MSE}_\infty} \tag{5.18}$$

representing the signal-to-estimation-noise ratio in the channel estimates. We solved for SENR through iteration, and examined this quantity as a function of several parameters, including the autoregressive time constant ρ, the observation length n, and the input pilot-tone signal-to-noise ratio

$$\mathrm{SNR_p} = \frac{\mathcal{E}_p\, \sigma_A^2}{\mathcal{N}_0} \tag{5.19}$$

To combine all of our results, we assume that the transmit and pilot-tone energies are related according to

$$\mathcal{E}_p = D \cdot \mathcal{E}_s \tag{5.20}$$

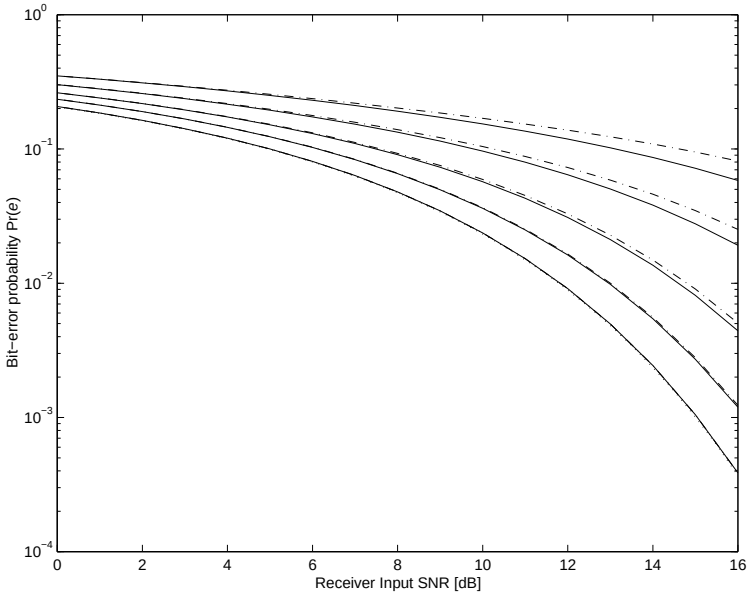
where $D > 0$ is a constant. Under this assumption, the transmit and pilot-tone signal-to-noise ratios scale proportionally, so that the performance of the channel estimation algorithm improves with increasing SNR. We plot the bit-error performance of the two equalizers using (2.32) as a function of the receiver input signal-to-noise ratio

$$\text{SNR} = 1/\zeta_0 = \frac{\mathcal{E}_s \, \sigma_A^2}{\mathcal{N}_0} \tag{5.21}$$

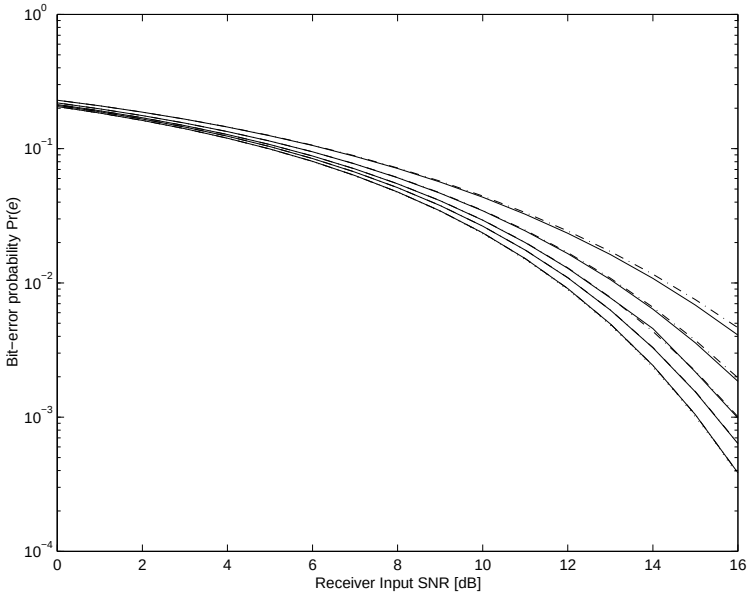
for various values of ρ and D.

We show in Figure 5-2 analytical results for two different fading channels with time constants $\rho = 0.9$ and $\rho = 0.999$, respectively. In each plot, we compute the bit-error performance for pilot-tone power factors of $D = -10, -5, 0, 5$ and 20 dB for the two equalizers. We see from the figure that the more slowly-varying the channel, the more critical it becomes to account for the mean-square estimation error in the equalizer. Furthermore, we see that as the channel becomes more slowly-varying, less energy is required in the pilot-tone to achieve good performance, as we would expect. In particular, we point out that the curve for $D = 0$ dB with $\rho = 0.999$ is lower than the curve for $D = 5$ dB with $\rho = 0.9$. This suggests that the total energy per bit required to achieve these performance curves goes down as the channel becomes more slowly varying.

The plots in Figure 5-2 assume that the fading channel model parameters are known at the receiver. To characterize the performance when the parameters are unknown and estimated using the algorithms of the previous section, we turn to a simulation of these systems. In Figure 5-3, we show the result of simulations for the frequency nonselective fading channel with $\rho = 0.9$ and $D = 0, 5$ and 20 dB. First, in Figure 5-3(a), we demonstrate the consistency of our analytical and simulated results for the case when the model parameters are known perfectly at the receiver. We note that because the difference in performance between the optimal and mismatched equalizers is relatively small for the large pilot-tone power factors considered, we examine only the optimal equalizer in Figure 5-3. Secondly, we compare in Figure 5-3(b) the analytical results based on knowledge of the parameters with simulated results based on parameter estimates. We see that a slight loss in SNR is incurred due to the parameter estimation algorithm; however, the exponential decay of the probability of error with increasing input SNR remains. Consequently, these results suggest that the composite channel is still well modeled as an additive marginally Gaussian

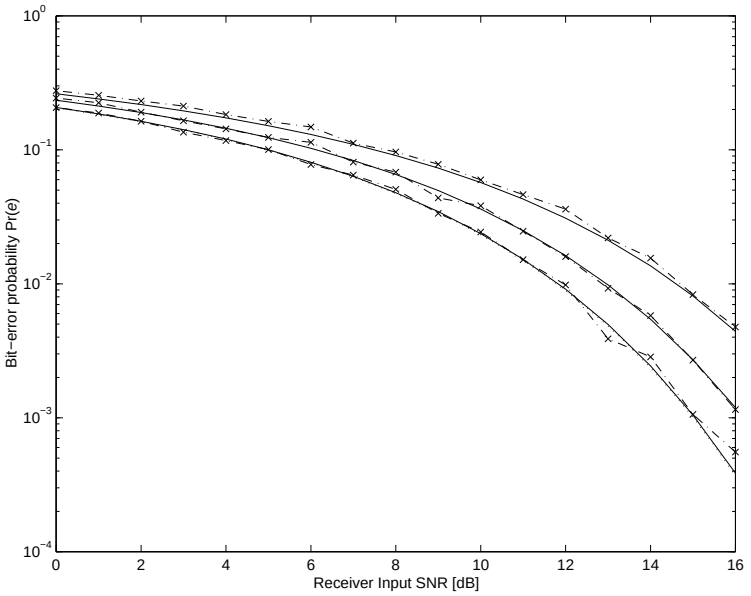


(a)

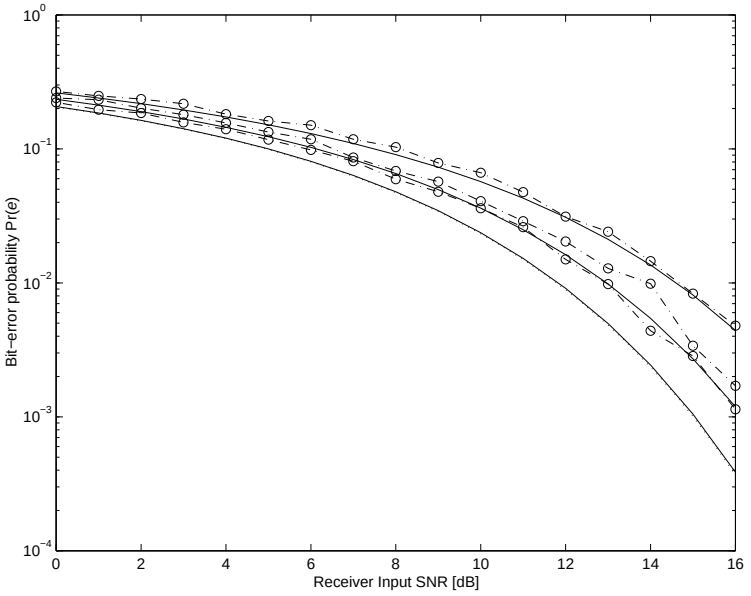


(b)

Figure 5-2: Analytical bit-error probability for QPSK signaling using spread-response precoding systems based on channel estimates. The two plots correspond to the frequency nonselective channel model (4.1) with (a) $\rho = 0.9$ and (b) $\rho = 0.999$. The successively lower curves correspond to the pilot-tone power factors of $D = -10, -5, 0, 5$ and 20 dB, respectively. The solid curves correspond to the optimal equalizer (5.12), while the dash-dotted curves correspond to the mismatched equalizer (5.14).



(a)



(b)

Figure 5-3: Simulated bit-error probability for QPSK signaling using spread-response precoding systems based on channel estimates. The two plots correspond to using the frequency nonselective channel model (4.1) with (a) known model parameters and (b) estimated model parameters. The successively lower curves correspond to pilot-tone power factors of $D = 0, 5$ and 20 dB, respectively. The solid curves are the results from Figure 5-2(a), while the curves marked with "$\times$" and "$\circ$" correspond to simulated performance values.

white-noise channel.

5.3 Summary

In this chapter, we have examined the impact of channel estimation errors on the performance of spread-response precoding systems. We began by re-deriving the optimal fading channel equalizers based on (approximate) MMSE estimates of the fading, and showed that the optimal equalizer must take into account a measure of the estimation noise, in particular, the mean-square estimation error. We then computed the output SNR for the optimal equalizer as well as for an equalizer which does not take into account a measure of the estimation noise. Through empirical investigations, we discovered that the optimal equalizer does not improve significantly on the performance of the mismatched equalizer when the pilot-tone SNR is at least as large as the transmit channel SNR. Furthermore, we found the difference to be even less significant as the channel became more correlated. Finally, we confirmed these results through simulation for a particular channel, and gave simulated results when the channel model parameters are estimated as well.

Chapter 6

Laboratory Experiments

All of the results of Chapters 3–5 were obtained through analysis and simulation. To generate each of the curves, we assumed a particular set of model parameters and evaluated how the various estimation and equalization techniques performed under this assumption. Although such an approach yields understanding of how the algorithms behave under various conditions, our ultimate interest lies in the performance of real systems employing spread-response precoding. To begin developing this picture, we have performed several experiments within an indoor wireless communication laboratory. We first summarize the laboratory setup and point out the additional channel distortions we have suppressed in our previous development. We then present the results of two sets of experiments, namely, measurement and estimation of the fading channel, and an implementation of spread-response precoding systems over this channel.

6.1 Laboratory Setup

Our experimental wireless system can be summarized according to the block diagram of Figure 6-1. We process the complex discrete-time, baseband transmit and receive signals $s[n]$ and $r[n]$, respectively, using a workstation. Another program handles data transfer between the workstation and the analog-to-digital (A/D) and digital-to-analog (D/A) converters. These converters can provide a sampling frequency of at least $f_s = 10$ kHz. At the transmitter, a set of analog hardware implements the frequency up-conversion to a carrier of $f_c = 915$ MHz. The up-conversion process is shown in the figure as a multiplication

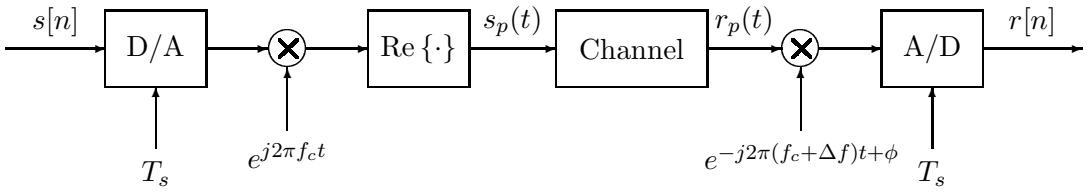


Figure 6-1: Passband block diagram summarizing the indoor wireless laboratory environment.

by the complex exponential $e^{j2\pi f_c t}$ followed by taking the real part of the resulting signal, yielding the passband signal $s_p(t)$. The channel consists of indoor wireless transmission of $s_p(t)$ between two antennae, with the receive antenna potentially in motion. At the receiver, the received passband signal $r_p(t)$ is down-converted using another set of analog components; however, due to offsets in instantaneous frequency and phase in the two oscillators, we model the down-conversion process as multiplication by the complex exponential $e^{-j2\pi(f_c+\Delta f)t+\phi}$. Appropriate anti-aliasing filters are present in the A/D converter.

Several distortions are introduced by the analog hardware, only some of which are shown in Figure 6-1. To reduce their effects, several restrictions must be placed on the transmitted signal. For example, harmonic distortions are introduced when the input signal power is too large. Similarly, when the input signal power is too small, the demodulated carrier introduces DC bias in the received signal. Both of these distortions restrict the effective transmit signal energy which can be employed. To reduce the effect of carrier offsets Δf, we employ a tunable oscillator (frequency generator) at the receiver, allowing us to reduce Δf to under 10 Hz which should be small relative to our transmission bandwidth for signaling experiments. Finally, for the case of QPSK symbols, we find it appropriate to employ differential encoding and detection to eliminate the effects of the constant phase offset ϕ. All of these restrictions are implementation specific and were not required in our general analysis of the previous chapters. However, to perform various experiments with a real wireless system, these effects must be considered.

6.2 Fading Measurements

In our first set of experiments, we explore the properties of the particular fading channel found within our wireless system of Figure 6-1. It seems plausible to model this channel as being frequency nonselective, due to the large coherence bandwidth, or small multipath delay spread, typically present in such indoor channels. For example, using our maximum sampling rate of $f_s = 10$ kHz, corresponding to a sample period of $T_s = 10^{-4}$ seconds, we see that for two multipath components to be resolve more than one tenth of a sample apart, the paths would need to have lengths differing by

$$3 \times 10^8 \,\mathrm{m/s} \cdot 10^{-5} \,\mathrm{seconds} = 3 \,\mathrm{km}$$

Because differences in path lengths on this order are not possible in our small indoor environment, we may faithfully regard the channel as having a small delay spread relative to the sample period, or alternatively, a large coherence bandwidth relative to the transmission bandwidth.

Since the experimental wireless channel is frequency nonselective, we are interested in assessing the applicability of the assumed first-order autoregressive model

$$a[n+1] = \rho\, a[n] + v[n+1] \tag{6.1}$$

where the (possibly complex[1]) time constant $|\rho| < 1$ and $v[n]$ is a zero-mean, complex Gaussian white-noise sequence with variance $\mathrm{E}\left[|v[n]|^2\right] = \sigma_v^2$.

To measure the channel, we transmit a complex exponential with continuous-time frequency f_p as a pilot-tone, *i.e.*,

$$s[n] = \sqrt{\mathcal{E}_p}\, e^{j2\pi(f_p/f_s)n} \tag{6.2}$$

through the system shown in Figure 6-1, while walking around with the receiver antenna. To process the received signal $r[n]$, we utilize the system shown in Figure 6-2. First, we demodulate the pilot-tone to baseband and then low-pass filter to eliminate out-of-band

[1] We find it convenient to incorporate the effects of the small frequency offset Δf into the model (6.1) by allowing the time constant ρ to be complex. Hence, $\rho = |\rho|\, e^{j \angle\rho}$, where $\angle\rho \propto \Delta f / f_s$.

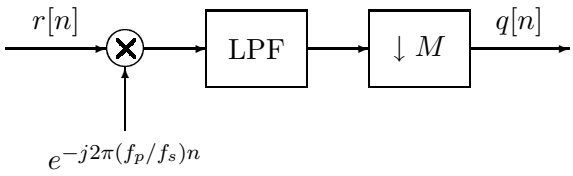


Figure 6-2: Baseband system for processing the received pilot-tone signal.

noise. Finally, we downsample to obtain the reduced rate sequence

$$q[n] = \sqrt{\mathcal{E}_p}\, a[n] + \tilde{w}[n] \tag{6.3}$$

where $\tilde{w}[n]$ is a zero-mean, complex Gaussian white-noise sequence with variance

$$\mathrm{E}\left[|\tilde{w}[n]|^2\right] = \mathcal{N}_0/M \tag{6.4}$$

In this setting, we assume $a[n]$ evolves according to the model (6.1).

Figure 6-3 shows the received power level of $q[n]$ in decibels for a particular experiment in which we walked around the room with the receiver antenna to create fading effects in the channel. The setup of this experiment was as follows: $f_s = 2$ kHz, $f_p = 500$ Hz, and $\mathcal{E}_p = 0.25$ V (-12 dB). Examination of the received sequence $r[n]$ revealed that a downsampling factor of $M = 5$ was suitable for eliminating all but the highest level noise, while maintaining the entire fading spectrum. From the results in Figure 6-3, we see two effects often found in wireless links, large-scale path loss of roughly 3 dB, and, more significantly, small-scale fading by as much as 30 dB due to receiver motion. We also note that the channel appears to be slowly-varying, relative to our sample time. For example, this particular realization remains essentially constant for half of a second in several places, which means we should be able to estimate the fading accurately.

We consider applying our parameter estimation algorithms to the sequence $q[n]$, since it is of the form of a channel measurement on which these algorithms were based. We apply two different parameter estimation algorithms, namely, the complete EM algorithm described in Section 4.3.1, and the sequential approximation to the EM algorithm discussed in Section 4.3.2. These algorithms yield practical values of the model parameters ρ, σ_v^2, and

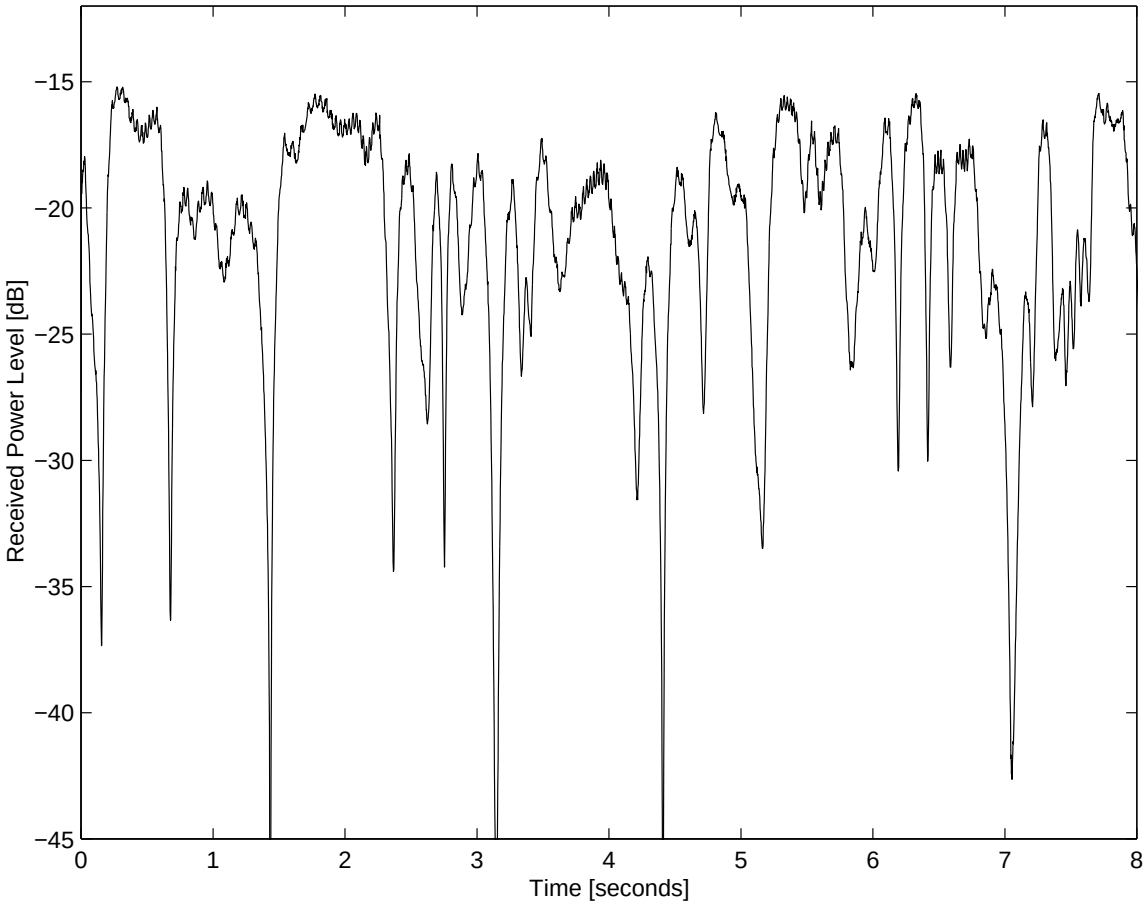


Figure 6-3: Measured received power level $|q[n]|$ in dB with time.

$\sigma_{\tilde{w}}^2$.

A typical set of results for iterations of the complete EM algorithm are given in Figure 6-4. Due to computational complexity of this algorithm, only the last half second, corresponding to two hundred samples, were used when applying the complete EM algorithm. Our convergence criterion was that the absolute difference in the estimates from one iteration to the next be no more than 1%. As we see from this set of results, 100 iterations were required to meet this level of convergence, which is a large computational burden, even with the reduced data set.

In Figure 6-5, we show the results of the sequential approximation to the EM algorithm for estimating the parameters. Recall that for the sequential EM approximation, we perform only one iteration for each new data point; hence, the horizontal axis in Figure 6-5 represents the number of observations, n, instead of the number of iterations, k. We see from these results that the two algorithms seem to converge to similar solutions; therefore, based on empirical results, we regard our sequential algorithm as a good approximation to the complete EM algorithm when given enough data. More importantly, we have received a large computational benefit in the sequential algorithm. It is important to point out, however, that direct comparisons of the results in Figure 6-4 and Figure 6-5 are difficult because the parameter estimates are formed based on different sets of data.

In addition, we show in Figure 6-6 the power density spectrum for the received signal $q[n]$ compared with that predicted by the first-order model (6.1) with the final parameter estimates from Figure 6-4 and Figure 6-5, respectively. Specifically, the model power density spectra are given by

$$S_{qq}(e^{j\omega}) = \left| \frac{\sigma_v}{1 - \rho\, e^{j\omega}} \right|^2 + \sigma_{\tilde{w}}^2 \tag{6.5}$$

with the parameter values obtained from applying the two parameter estimation algorithms to the sequence $q[n]$. We see from Figure 6-6 that the first-order model seems to captures some, but not all, of the dynamics of the fading process. In particular, we note the inability of the first-order model to decay quickly enough in frequency, which might explain why the complete EM algorithm repeatedly underestimated the additive receiver thermal noise. Fortunately, the sequential algorithm seems to measure this quantity more accurately. To capture all of these dynamics more precisely, higher-order models, potentially with moving average components, may be necessary.

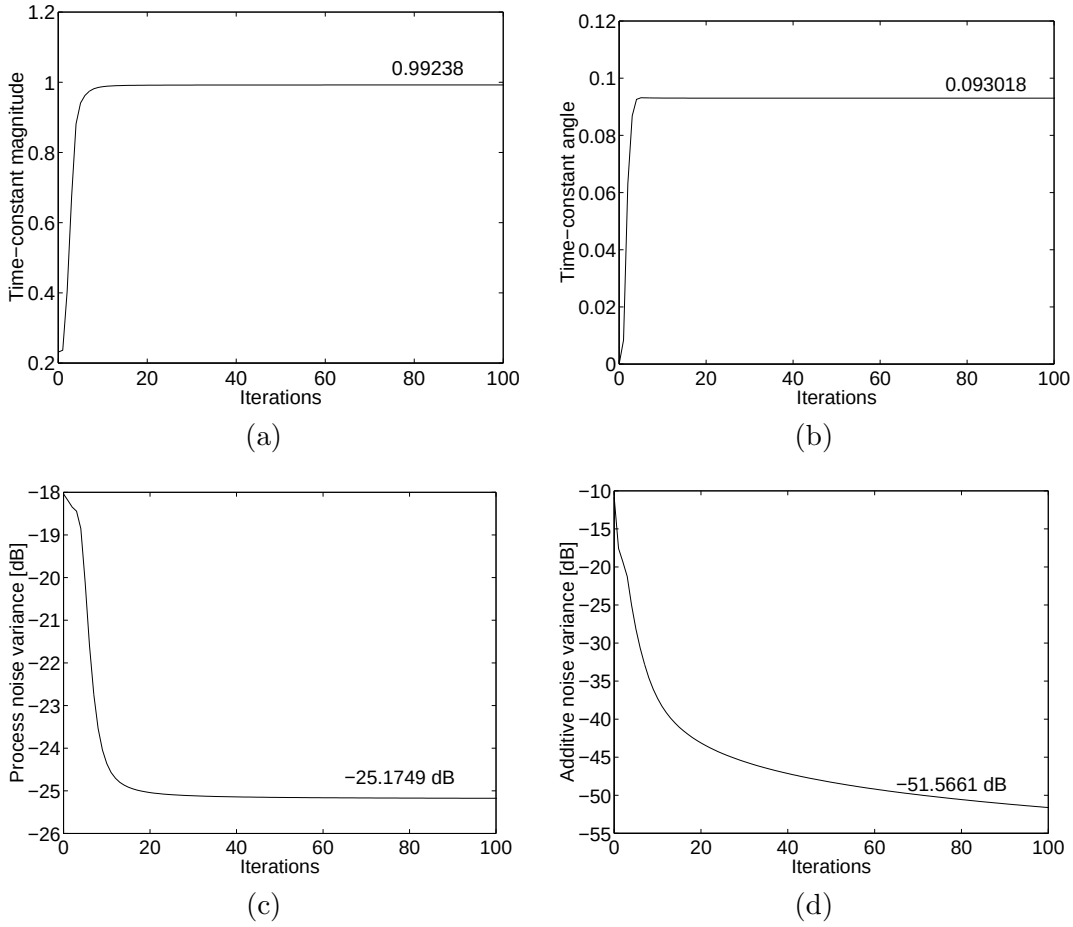


Figure 6-4: Parameter estimation results of the complete EM algorithm using the experimental pilot-tone data. The four plots show the evolution of the estimated (a) time-constant magnitude $|\hat{\rho}_k|$, (b) time-constant angle $\angle\hat{\rho}_k$, (c) process noise variance $\widehat{(\sigma_v^2)}_k$, and (d) additive noise variance $\widehat{(\sigma_{\tilde{w}}^2)}_k$ versus iterations k, respectively.

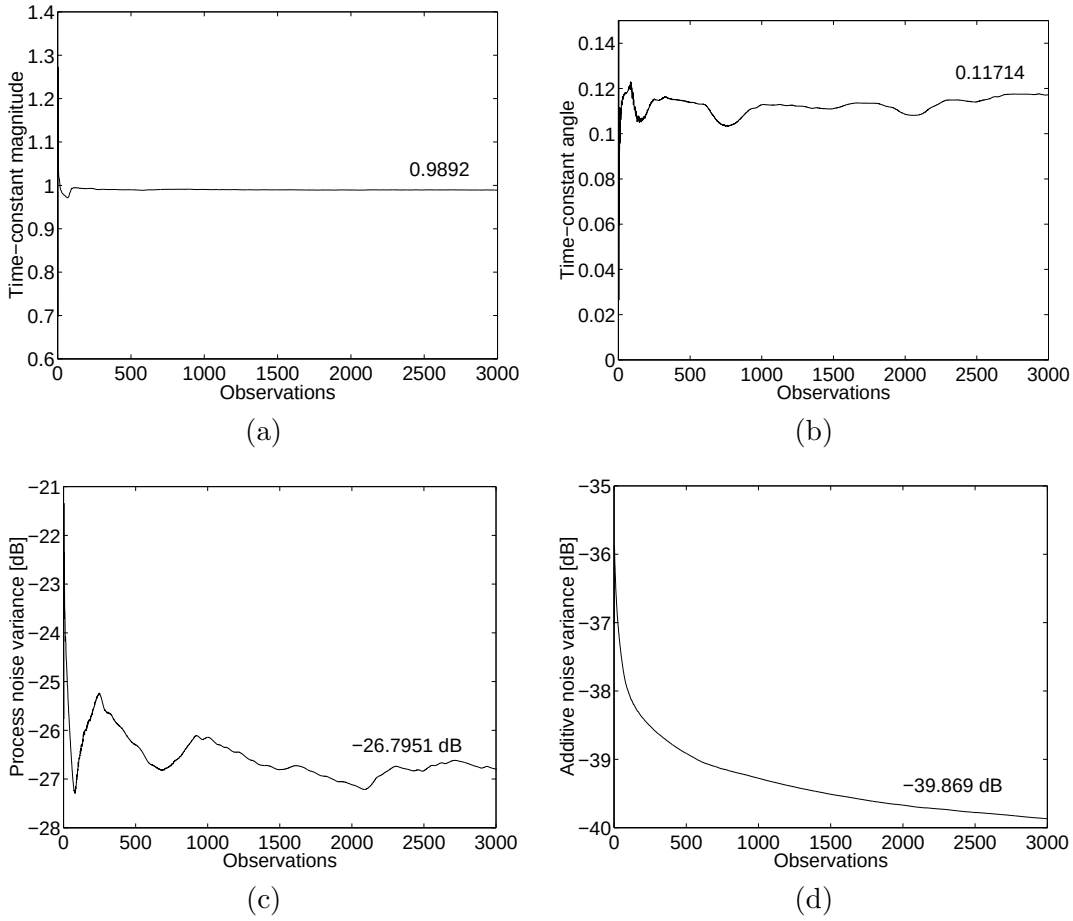


Figure 6-5: Parameter estimation results of the sequential approximation to the EM algorithm using the experimental pilot-tone data. The four plots show the evolution of the estimated (a) time-constant magnitude $|\hat{\rho}_n|$, (b) time-constant angle $\angle\hat{\rho}_n$, (c) process noise variance $\widehat{(\sigma_v^2)}_n$, and (d) additive noise variance $\widehat{(\sigma_{\tilde{w}}^2)}_n$ versus observations n, respectively.

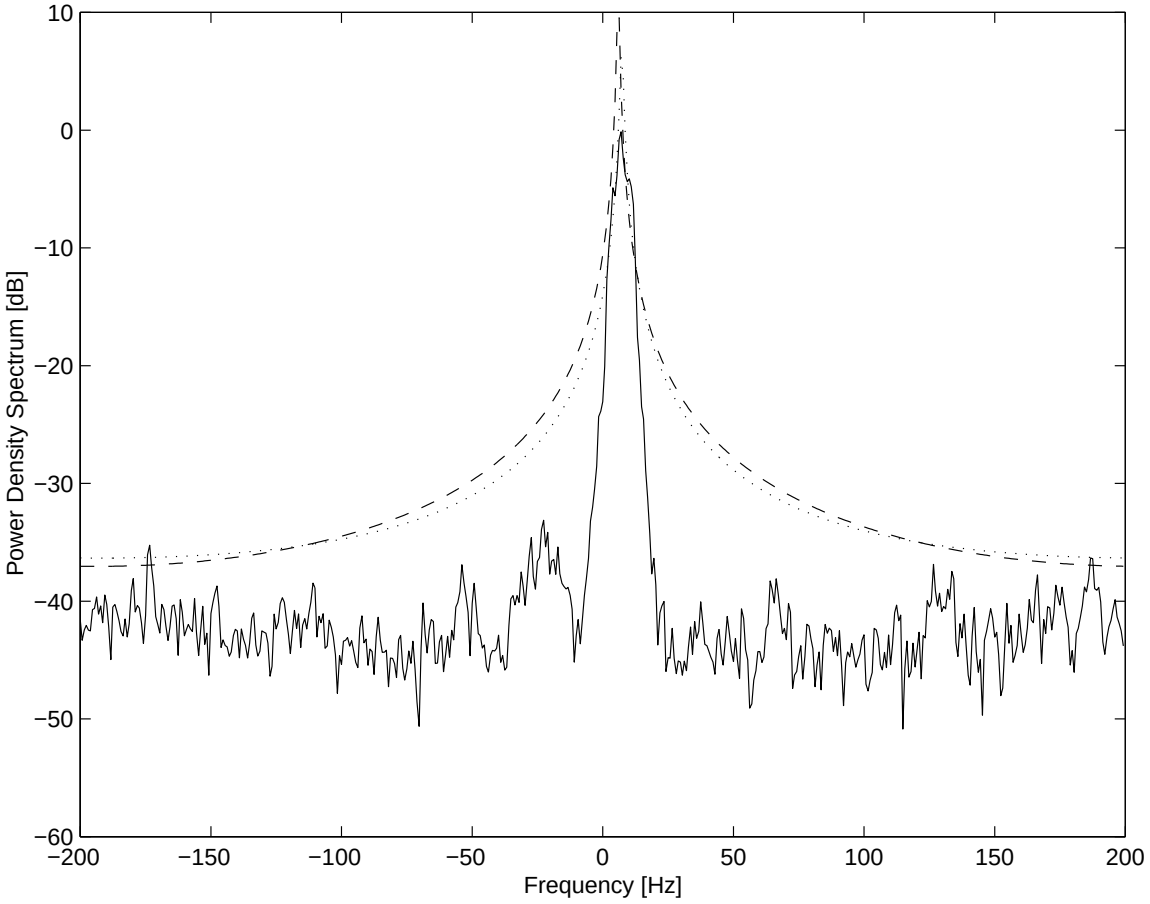


Figure 6-6: Power density spectrum of the received pilot-tone $q[n]$. The solid line corresponds to the measured power density spectrum obtained through the Welch periodogram method. The dashed line corresponds to the fading channel mode with parameters obtained by applying the complete EM algorithm to $q[n]$, while the dotted line corresponds to the fading channel mode with parameters obtained by applying the sequential algorithm to $q[n]$.

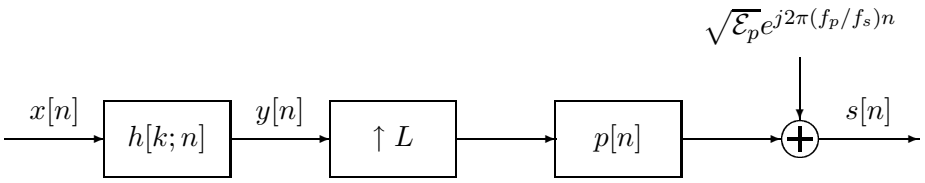


Figure 6-7: Block diagram of the transmitter for the spread-response precoding implementation.

Finally, we note the similarity of the two models based on parameter estimates from the different algorithms. These results suggest that our indoor wireless channel can be reasonably modeled using the autoregressive model, with a coherence time of approximately 0.25 seconds ($\rho \approx 0.99$ with $f_s = 400$ Hz), average fading power of 0.15 V^2 ($\rho \approx 0.99$, $\sigma_v^2 \approx 0.003$), and additive noise intensity of less than -30 dB.

6.3 Spread-Response Precoding Implementation

The results of the previous section suggest that the first-order model is a reasonable if simple one for our wireless channel. We next consider an implementation of spread-response precoding systems for use over this channel, based on the estimation and equalization schemes from the previous chapters. While this implementation is preliminary, and limited in its ability to verify the performance curves obtained through analysis and simulation, it does shed some light on the assumptions used throughout the thesis. Our approach in this section is to first discuss the implementation details, and then present several results obtained in the laboratory.

To obtain the transmit sequence $s[n]$ which we pass through the channel of Figure 6-1, we process the QPSK symbol stream $x[n]$ using the system shown in Figure 6-7. As we mentioned in Section 6.1, due to the phase offset ϕ between the transmitter and receiver carriers, it is necessary to differentially encode $x[n]$. This sequence is further encoded using the spread-response precoder $h[k;n]$, which we implement according to the block diagram of Figure 2-2. We upsample the precoded sequence $y[n]$ by the factor L, and then pass the sequence through the square-root, raised-cosine filter $p[n]$ for pulse-shaping. Finally, the pilot-tone of the previous section is added to the pulse-shaped sequence. Using the sampling

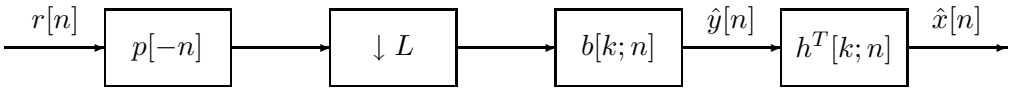
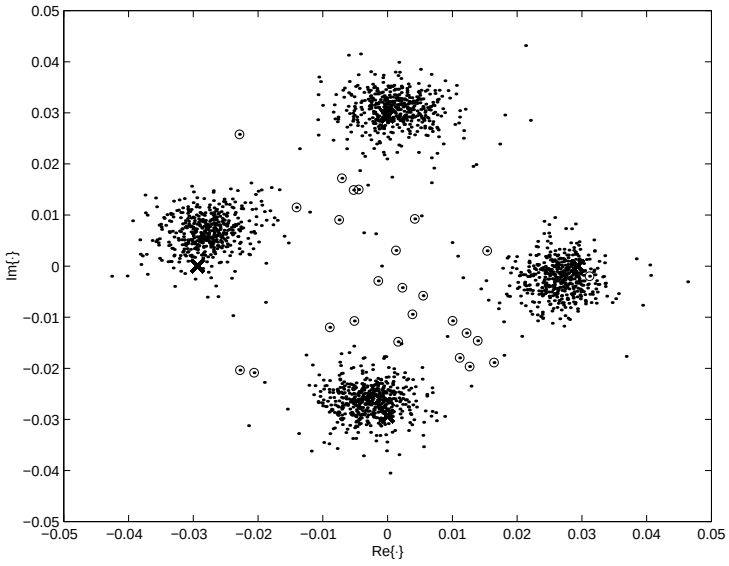


Figure 6-8: Block diagram of the receiver for the spread-response precoding implementation.

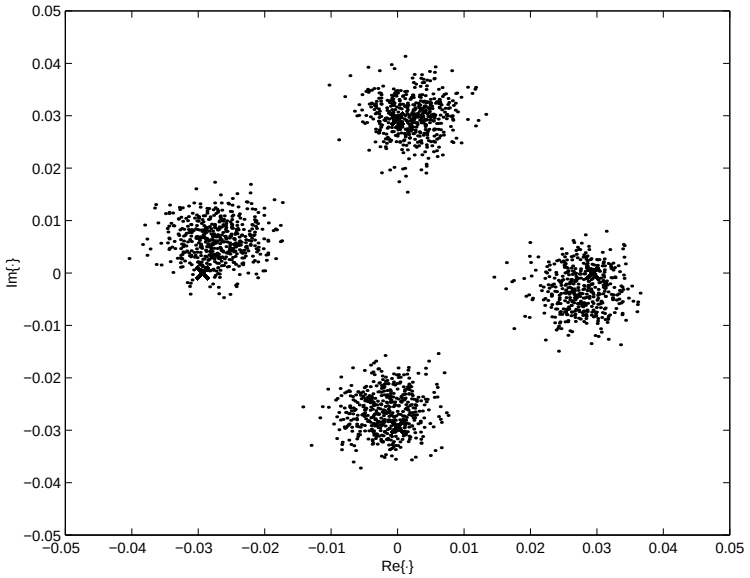
frequency $f_s = 5$ kHz with $L = 8$, our symbol rate is $R = 625$ samples per second. Since we use 50% excess bandwidth in the raised-cosine filter to provide some immunity to timing errors, we use the pilot-tone frequency $f_p = 1$ kHz.

At the receiver, we have two separate processing steps. First is the pilot-tone processing shown in Figure 6-2, with $M = L = 8$, yielding the channel observation sequence $q[n]$. Secondly, we process the data sequence using the system in Figure 6-8. First, we matched-filter the received sequence using $p[-n]$, and then downsample by the factor L. We equalize the resulting sequence using the equalizer $b[k;n]$, and finally "undo" the precoding using the postcoder $h^T[k;n]$, whose implementation is shown in Figure 2-3. Finally, the estimated symbols $\hat{x}[n]$ are differentially detected.

Using the systems as described above, we transmit a block of 2000 symbols over the channel for two cases, without spread-response precoding, and with spread-response precoding of length $N = 128$ times the coherence time of the channel in samples. In both cases, channel estimates are derived from pilot-tone observations and used to equalize the received sequence. In Figure 6-9 we show scatter plots of the receiver symbol constellations (before hard-limiting) for the two cases. We see from inspection of this figure that finite-length spread-response precoding dramatically changes the interference characteristics at the output of the receiver. These results seem reasonable based on simulations using the same about of data bits and similar fading rates. In fact, if we subtract the original symbols $x[n]$ from symbol estimates $\hat{x}[n]$ generated through simulation of these laboratory conditions, we may then compute histograms for the output noise $z[n]$ as shown in Figure 6-10. From these observations we see that the marginally Gaussian model seems to be a good approximation. Furthermore, estimates of the power density spectrum of $z[n]$ indicate that the white approximation is also a reasonable one.



(a)



(b)

Figure 6-9: Output constellations (a) without spread-response precoding precoding, and (b) with length $N = 128$ spread-response precoding. The transmitted symbols $x[n]$ are indicated by "$\times$", while the symbol estimates (before hard-limiting) $\hat{x}[n]$ are indicated by "$\cdot$". Those symbol estimates which are in error after hard limiting are indicated by "$\odot$".

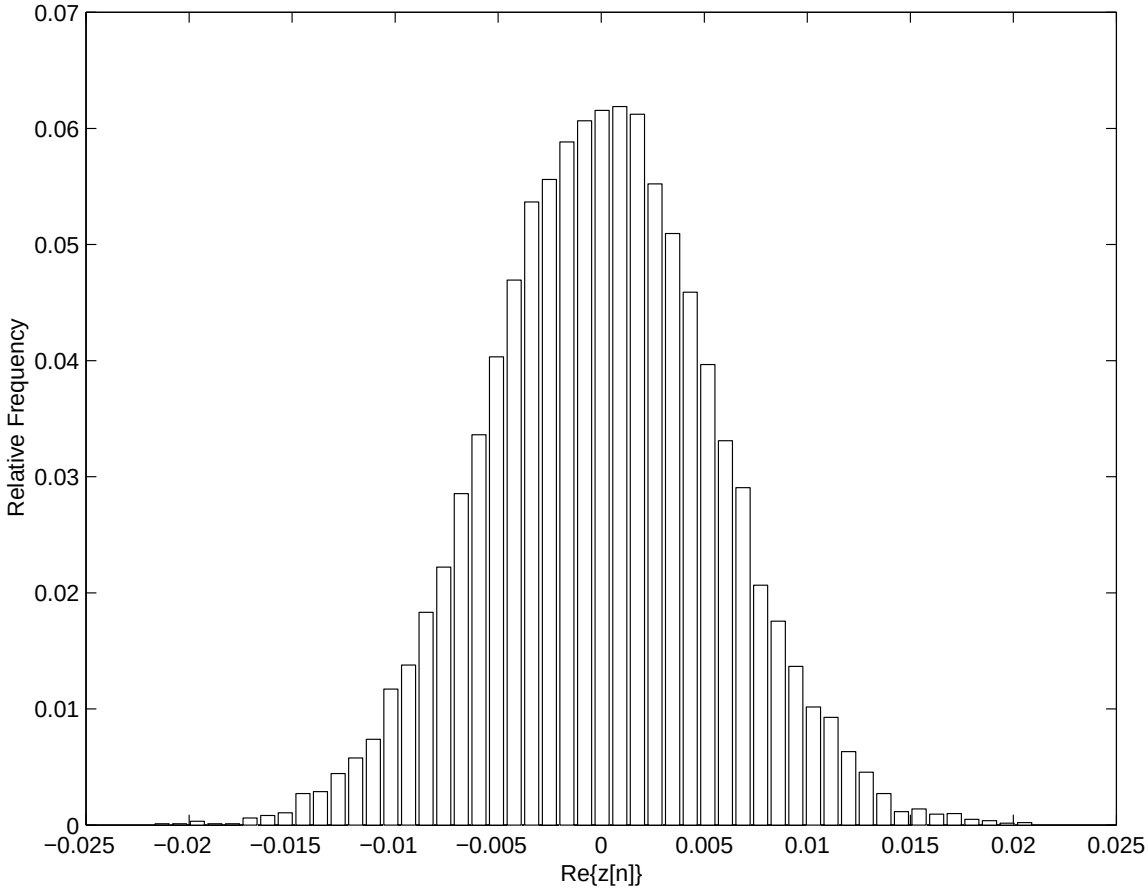


Figure 6-10: Histogram for the output noise $\mathrm{Re}\{z[n]\}$ of a spread-response precoding system.

6.4 Summary

In this chapter we have presented the results of several laboratory experiments with spread-response precoding systems over an indoor wireless fading channel. We first identified the fading channel as being frequency nonselective, and found useful values for the parameters of the first-order autoregressive evolution model of the channel. In obtaining these values, we confirmed the notion that our sequential approximation to the EM algorithm for estimating these parameters performs well and gives a large computational benefit over the EM algorithm. We also described an implementation of spread-response precoding systems within our wireless laboratory, and presented several results which seem to confirm the theoretical results (and the assumptions on which they are based) concerning spread-response precoding systems. In particular, we showed empirical results which suggest that the remaining additive interference at the output of the receiver appears marginally Gaussian and white.

Chapter 7

Conclusions

The goal of this research has been to explore several of the many issues involved with implementing spread-response precoding systems for use in fading environments. Specifically, we have addressed several questions related to channel estimation and equalization in these systems, and we have demonstrated our results through analysis, simulation, and a preliminary implementation within an indoor wireless communication system. We have approached these problems somewhat independently in that we first form a channel estimate and then use this information in a optimal way in the equalizer.

In Chapter 3, we explored two equalizers obtained through seemingly different criteria, namely, the maximum output SNR criterion and the linear MMSE criterion. We demonstrated that these two equalizers are equivalent for the frequency nonselective channel, and very similar in terms of their frequency responses for the slowly-varying, frequency selective channel. Finally, we gave an example of a quickly-varying fading channel for which the two equalizers are vastly different. From the results of Chapter 3, we find that the Kalman filter equalizer based on the linear MMSE criterion has a more straightforward implementation and gives performance which is equal to or better than that obtained using the the (approximate) maximum SNR equalizer. However, solution of the true maximum output SNR equalizer for the quickly-varying, frequency selective channel remains an open problem.

In Chapter 4, we described a pilot-tone channel observation model and channel estimation techniques based on the Kalman filter and a sequential approximation to the EM algorithm. Our empirical results suggest that these estimators remain unbiased and approximate MMSE channel estimators even when the model parameters are unknown *a priori*

at the receiver. Using these two important properties of the channel estimates, we were able in Chapter 5 to find the optimal equalizers based on channel estimates, and characterize the degradation in performance due to mean-square estimation error of the channel. We also gave a useful interpretation of the equalizer as being based on the known channel corresponding to the channel estimate, with additive noise depending on the mean-square estimation error as well as the original additive noise.

Finally, in Chapter 6, we presented the results of several laboratory experiments with the algorithms given in Chapters 2–5. We found characteristic values of the fading channel model parameters by applying the parameter estimation algorithms of Sections 4.3.1 and 4.3.2 to received pilot-tone data obtained experimentally in the laboratory. Furthermore, in the context of implementing spread-response precoding systems, we provided experimental evidence that spread-response precoding changes the characteristics of the output noise to be well approximated as marginally Gaussian and white. This property is very important from the point of view of implementing spread-response precoding systems, because it allows us to employ symbol-by-symbol detection at the receiver.

Several other important issues have been left for future work. Many of the results of this thesis could be extended formally to the frequency selective channel. As we mentioned in Chapter 4, a different channel observation model based on the insertion of pilot-symbols would be preferable for this case. Moreover, the pilot-symbol observation model would include the frequency nonselective channel as a special case, and would eliminate the peak-power issues associated with a pilot-tone. We feel that such a framework could be extended to employ decision feedback equalization, which might lead to an increase in throughput as well as in performance. We also note that the area of uncoordinated multiuser communication over fading channels presents another natural area of extension of this research. Finally, as with any idea, demonstration of these schemes within our wireless testbed would provide a useful tool for confirming our analytical results and the assumptions on which they are based.

Appendix A

Computations

A.1 Average Output SNR Computations

In this section, we give details on some of the results utilized in Chapters 3 and 5. Specifically, we compute the quantity

$$\phi = \frac{|\mathrm{E}\,[ab]|^2}{\mathrm{E}\left[\left(|a|^2 + \eta\right)|b|^2\right]} \tag{A.1}$$

where a is a zero-mean complex Gaussian random variable with variance σ_a^2, the random variable b depends on a according to the relationship

$$b = \frac{a^*}{|a|^2 + \eta'} \tag{A.2}$$

and η, η' are two constants. Our first step is to rewrite (A.1) as

$$\phi = \frac{|\mathrm{E}\,[ab]|^2}{\mathrm{E}\left[\left(|a|^2 + \eta' + (\eta - \eta')\right)|b|^2\right]} \tag{A.3}$$

Then when we substitute (A.2) into (A.3), we obtain

$$\phi = \frac{|\mathrm{E}\,[ab]|^2}{\mathrm{E}\,[ab] + (\eta - \eta')\mathrm{E}\left[|b|^2\right]} \tag{A.4}$$

We see from (A.4) that we must compute two expectations, namely, $\mathrm{E}\,[ab]$, and $\mathrm{E}\left[|b|^2\right]$.

We begin by noting that a has the probability density function

$$p_a(a_0) = \frac{1}{\pi\sigma_a^2} e^{-|a_0|^2/\sigma_a^2} \tag{A.5}$$

so that its squared amplitude $s = |a|^2$ has the exponential density function

$$p_s(s_0) = \frac{1}{\sigma_a^2} e^{-s_0/\sigma_a^2} \tag{A.6}$$

for $s_0 \geq 0$, and 0 otherwise.

A.1.1 Computation of $\mathrm{E}[ab]$

By substituting (A.2), we obtain

$$\mathrm{E}[ab] = \mathrm{E}\left[\frac{|a|^2}{|a|^2 + \eta'}\right] = \mathrm{E}\left[\frac{s}{s + \eta'}\right] \tag{A.7}$$

which we solve for using the density function (A.6) via

$$\mathrm{E}\left[\frac{s}{s+\eta'}\right] = \frac{1}{\sigma_a^2}\int_0^\infty \frac{s_0}{s_0+\eta'} e^{-s_0/\sigma_a^2} ds_0 \tag{A.8a}$$

$$= \frac{1}{\sigma_a^2}\int_{\eta'}^\infty \frac{t-\eta'}{t} e^{-(t-\eta')/\sigma_a^2} dt \tag{A.8b}$$

$$= \frac{1}{\sigma_a^2}\int_0^\infty e^{-u/\sigma_a^2} du - \frac{\eta'}{\sigma_a^2} e^{\eta'/\sigma_a^2} \int_{\eta'}^\infty \frac{e^{-t/\sigma_a^2}}{t} dt \tag{A.8c}$$

$$= 1 - \zeta' e^{\zeta'} E_1(\zeta') \tag{A.8d}$$

where $\zeta' = \eta'/\sigma_a^2$ and $E_1(\cdot)$ is the exponential integral

$$E_1(\nu) = \int_\nu^\infty \frac{e^{-t}}{t} dt \tag{A.9}$$

We note that (A.8b) is obtained through the change of variable $t = s_0 + \eta$, (A.8c) results from the change of variable $u = t - \eta$, and (A.8d) results from the change of variable $v = t/\sigma_a^2$.

We also note that when $\eta = \eta'$, we need only solve for $\mathrm{E}[ab]$, in which case we find

$$\phi = 1 - \zeta e^{\zeta} E_1(\zeta) \tag{A.10}$$

in which $\zeta = \eta/\sigma_a^2$.

A.1.2 Computation of $\mathrm{E}\left[|b|^2\right]$

When $\eta \neq \eta'$, we must also compute the expectation $\mathrm{E}\left[|b|^2\right]$. Substituting (A.2) as before,

$$\mathrm{E}\left[|b|^2\right] = \frac{|a|^2}{(|a| + \eta')^2} = \mathrm{E}\left[\frac{s}{(s+\eta')^2}\right] \tag{A.11}$$

Using the density function (A.6), the steps in the computation of this expectation are as follows.

$$\mathrm{E}\left[\frac{s}{(s+\eta')^2}\right] = \frac{1}{\sigma_a^2}\int_0^\infty \frac{s_0}{(s_0+\eta')^2} e^{-s_0/\sigma_a^2} ds_0 \tag{A.12a}$$

$$= \frac{1}{\sigma_a^2}\int_{\eta'}^\infty \frac{t-\eta'}{t^2} e^{-(t-\eta')/\sigma_a^2} dt \tag{A.12b}$$

$$= \frac{1}{\sigma_a^2} e^{\eta'/\sigma_a^2}\int_{\eta'}^\infty \frac{e^{-t/\sigma_a^2}}{t} dt - \frac{\eta'}{\sigma_a^2} e^{\eta'/\sigma_a^2}\int_{\eta'}^\infty \frac{e^{-t/\sigma_a^2}}{t^2} dt \tag{A.12c}$$

$$= \frac{1}{\sigma_a^2} e^{\zeta'} E_1(\zeta') - \frac{\eta'}{\sigma_a^2} e^{\eta'/\sigma_a^2}\left[\frac{1}{\eta'} e^{-\eta'/\sigma_a^2} - \frac{1}{\sigma_a^2} E_1(\zeta')\right] \tag{A.12d}$$

$$= \frac{1}{\sigma_a^2}\left[e^{\zeta'} E_1(\zeta')(1+\zeta') - 1\right] \tag{A.12e}$$

Again we note that (A.12b) results from the change of variable $t = s_0 + \eta'$, while (A.12d) results from integrating the second term of (A.12c) by parts.

Combining the results (A.8) and (A.12), we obtain, after some simplification

$$\phi = \frac{\left|1 - \zeta' e^{\zeta'} E_1(\zeta')\right|^2}{1 - \zeta' e^{\zeta'} E_1(\zeta') + (\zeta - \zeta')\left[e^{\zeta'} E_1(\zeta')(1+\zeta') - 1\right]} \tag{A.13}$$

Bibliography

[1] John G. Proakis. *Digital Communications*. McGraw-Hill Book Company, Third edition, 1995.

[2] Gregory W. Wornell. Spread-response precoding for communicating over fading channels. *IEEE Transactions on Information Theory*, 42(2), March 1996.

[3] Gregory W. Wornell. Spread-signature CDMA: Efficient multiuser communication in the presence of fading. *IEEE Transactions on Information Theory*, 41(5), September 1995.

[4] G. David Forney, Jr. The Viterbi algorithm. *Proceedings of the IEEE*, 61(3), March 1973.

[5] Steven H. Isabelle and Gregory W. Wornell. Recursive multiuser equalization for CDMA systems in fading environments. In *Proceedings of the Allerton Conference on Communications, Control, and Signal Processing*, October 1996.

[6] Brian O. Anderson and John B. Moore. *Optimal Filtering*. Prentice-Hall, Inc., 1979.

[7] Armin Wittneben. An energy- and bandwidth efficient data transmission system for time-selective fading channels. In *IEEE GLOBCOMM*, volume 3, 1993.

[8] Meir Feder and Ehud Weinstein. Parameter estimation of superimposed signals using the EM algorithm. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 36(4), April 1988.

[9] E. Weinstein, A.V. Oppenheim, and M. Feder. Signal enhancement using single and multi-sensor measurements. Technical Report 560, Massachusetts Institute of Technology, Research Laboratory of Electronics, November 1990.