ABSTRACT

COMMUNICATIONS OVER FADING CHANNELS WITH PARTIAL CHANNEL INFORMATION: PERFORMANCE AND DESIGN CRITERIA

by
Xinmin Deng

The effects of system parameters upon the performance are quantified under the assumption that some statistical information of the wireless fading channels is available. These results are useful in determining the optimal design of system parameters. Suboptimal receivers are designed for systems that are constrained in terms of implementation complexity.

The achievable rates are investigated for a wireless communication system when neither the transmitter nor the receiver has prior knowledge of the channel state information (CSI). Quantitative results are provided for independent and identically distributed (i.i.d.) Gaussian signals. A simple, low-duty-cycle signaling scheme is proposed to improve the information rates for low signal-to-noise ratio (SNR), and the optimal duty cycle is expressed as a function of the fading rate and SNR. It is demonstrated that the resource allocations and duty cycles developed for Gaussian signals can also be applied to systems using other signaling formats.

The average SNR and outage probabilities are examined for amplify-and-forward cooperative relaying schemes in Rayleigh fading channels. Simple power allocation strategies are determined by using knowledge of the mean strengths of the channels.

Suboptimal algorithms are proposed for cases that optimal receivers are difficult to implement. For systems with multiple transmit antennas, an iterative method is used to avoid the inversion of a data-dependent matrix in decision-directed channel estimation. When CSI is not available, two noncoherent detection algorithms are formulated based on the generalized likelihood ratio test (GLRT). Numerical results are presented to demonstrate the use of GLRT-based detectors in systems with cooperative diversity.

COMMUNICATIONS OVER FADING CHANNELS WITH PARTIAL CHANNEL INFORMATION: PERFORMANCE AND DESIGN CRITERIA

by
Xinmin Deng

A Dissertation
Submitted to the Faculty of
New Jersey Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Electrical Engineering

Department of Electrical and Computer Engineering

May 2005

APPROVAL PAGE

COMMUNICATIONS OVER FADING CHANNELS WITH PARTIAL CHANNEL INFORMATION: PERFORMANCE AND DESIGN CRITERIA

Xinmin Deng

5/2/2005

Dr. Alexander M. Haimovich, Dissertation Advisor — Date
Professor, Department of Electrical and Computer Engineering, NJIT

5/2/05

Dr. Ali Abdi, Committee Member — Date
Assistant Professor, Department of Electrical and Computer Engineering, NJIT

5/2/05

Dr. Yeheskel Bar-Ness, Committee Member — Date
Distinguished Professor, Department of Electrical and Computer Engineering, NJIT

5/2/05

Dr. Leonard J. Cimini, Jr., Committee Member — Date
Professor, Department of Electrical and Computer Engineering, University of Delaware

5/2/05

Dr. Roy You, Committee Member — Date
Assistant Professor, Department of Electrical and Computer Engineering, NJIT

BIOGRAPHICAL SKETCH

Author: Xinmin Deng

Degree: Doctor of Philosophy

Date: May 2005

Undergraduate and Graduate Education:

- Doctor of Philosophy in Electrical Engineering,
  New Jersey Institute of Technology, Newark, NJ, USA, 2005
- Master of Science in Computer Engineering,
  Southwest Jiaotong University, Chengdu, Sichuan, PRC, 1997
- Bachelor of Science in Information Engineering,
  Southwest Jiaotong University, Chengdu, Sichuan, PRC, 1994

Major: Electrical Engineering

Presentations and Publications:

Xinmin Deng and Alexander Haimovich, "Multichannel noncoherent detection with applications to cooperative diversity," to be submitted, 2005.

Xinmin Deng and Alexander Haimovich, "Cooperative relaying in wireless networks with local channel state information," to be presented at *IEEE Vehicular Technology Conference (VTC)*, Dallas, TX, USA, September 2005.

Xinmin Deng and Alexander Haimovich, "Power allocation for cooperative relaying in wireless networks," submitted for publication, 2005.

Xinmin Deng and Alexander Haimovich, "Achievable rates over time-varying Rayleigh fading channels," submitted for publication, 2004.

Xinmin Deng and Alexander Haimovich, "Information rates of time varying Rayleigh fading channels," *IEEE International Conference on Communications (ICC)*, vol. 1, pp. 573-577 Paris, France, June 2004.

Xinmin Deng and Alexander Haimovich, "On pilot symbol aided channel estimation for time varying Rayleigh fading channels," *The 38th Annual Conference on Information Sciences and Systems (CISS)*, Princeton, NJ, March 2004.

Hangjun Chen, Xinmin Deng, and Alexander Haimovich, "Layered turbo space-time coded MIMO-OFDM systems for time varying channels," *IEEE Global Telecommunications Conference (GLOBECOM)*, vol. 4, pp. 1831 - 1836, San Francisco, CA, December 2003.

Xinmin Deng, Alexander Haimovich, and Javier Garcia-Frias, "Decision directed iterative channel estimation for MIMO systems," *IEEE International Conference on Communications (ICC)*, vol. 4, pp. 2326 - 2329, Anchorage, AK, May 2003.

ACKNOWLEDGMENT

First I would like to thank my advisor, Prof. Alexander Haimovich, for his well-founded advice and constant encouragement. The assistance of the other committee members, Profs. Ali Abid, Yeheskel Bar-Ness, Leonard Cimini, and Roy You, is greatly appreciated. I would also like to thank Profs. Rick Blum, Javier Garcia-Frias, Larry Greenstein, and Dr. Jianhong Luo, for their helpful interactions.

The financial support of the National Science Foundation and the Air Force Office of Scientific Research is gratefully acknowledged.

Gratitude goes to Ms. Marlene Toeroek, and my fellow students within the Center for Communications and Signal Processing Research (CCSPR), whose warmth and friendship has provided a most pleasant environment during my years of study at NJIT.

Lastly, loving thanks go to my family for their support.

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

Chapter **Page**

TABLE OF CONTENTS
(Continued)

Chapter **Page**

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES
(Continued)

Figure **Page**

CHAPTER 1

INTRODUCTION

This opening chapter outlines the scope of this work, and establishes the notation that will be used throughout the dissertation.

The transmission of information over wireless channels is subject to fading, i.e., random variation of amplitude and phase of the received signal. Fading is a challenging obstacle to the practical implementation of reliable communications, due to the physical limitations on power and bandwidth resources of wireless systems.

The performance of wireless systems depends on channel models and on the availability of channel state information (CSI) at the transmitter and receiver. In practice, less is known about the channel at the transmitter than the receiver. Throughout this dissertation, attention is confined to the cases in which the transmitter does not have *full* knowledge of the channel. Specifically, it is supposed that the transmitter does not have any information on the *instantaneous* realization of the channel (*short-term* knowledge), except some statistical information of the channel (*long-term* knowledge). The channel statistical information includes the probability density function (pdf), moments of channel gain for a block-fading channel, and the Doppler spectrum for a time-varying channel.

At the receiver, a typical approach to deal with the fading channel is to estimate the channel gains by training with known pilot symbols inserted in the transmitted data sequence. In a fast-fading environment, the CSI is generally *imperfect*, and the quality of CSI is determined by the resources devoted for channel estimation and the rate of channel variations. For low-mobility applications, where the channel varies slowly enough that the receiver can estimate the channel with very high accuracy, it is reasonable to assume that the receiver has *perfect* knowledge of the CSI.

In a wireless network, information may be transmitted from the source to the destination with the assistance of relay nodes, and therefore multiple wireless links are involved. Although for simplicity it is usually assumed that the destination knows all the channels in the cooperative transmission (*global* CSI), this is not practically true. For example, the source-relay channels cannot be explicitly measured at the destination. Therefore, it is practically appealing to consider systems with more realistic knowledge of CSI. For example, it is of interest to consider the case where only *local* CSI is available, i.e., the destination only has access to the CSI of the channels from its immediate neighbors, or the case that no CSI is available at all.

1.1 Overview of the Dissertation

The purpose of this dissertation is to optimize performance over wireless channels by exploiting channel information. With statistical information of the channels known at the transmitters, optimal system parameters will be determined from quantitative results on the effects of design parameters upon achievable performance. These system parameters include the power and bandwidth allocation to pilot symbols, duty cycle, and power allocation among nodes in a wireless network. Although an ideal receiver should exploit the channel information whenever available, the optimal receiver may be too complicated to implement. Design criteria for low-complexity suboptimal receivers will be developed under these situations.

Detailed background for this study is presented in Chapter 2. The main results are contained in Chapters 3-6.

Chapter 3 investigates the achievable rates over time-varying fading channels for a point-to-point wireless communication system, under the assumption that neither the transmitter nor the receiver have prior knowledge of the instantaneous CSI. The dynamics of the time-varying channel are characterized by a Doppler spectrum with a known fading rate. The penalty due to unknown CSI is expressed in terms of the fading rate, the SNR, and the

block length. For a pilot-aided system, expressions for the achievable rates are developed, and are used to determine the optimal power and bandwidth allocations for pilot symbols. A low-duty-cycle signaling scheme is proposed to improve the performance in the low SNR regime.

Chapter 4 proposes an iterative algorithm for decision-directed channel estimation in systems with multiple transmit antennas to avoid the inversion of a data-dependent matrix. This estimator is used in joint channel estimation and data detection for systems space-time coded modulation.

Chapter 5 examines the performance of amplify-and-forward cooperative relaying schemes in a wireless network with Rayleigh fading. Both average SNR and outage performance are considered, for various cases of available CSI at the relay and destination nodes. Power allocation strategies are developed that optimize these performance metrics. Numerical results are prensented for optimal receivers that utilize the global knowledge of CSI, and for suboptimal receivers when only local knowledge of CSI is available.

Chapter 6 investigates the noncoherent detection with diversity reception. Two detection algorithms are formulated based on the generalized likelihood ratio test (GLRT) rules. These algorithms are used for noncoherent detection in systems with amplify-and-forward cooperative diversity, for which the optimal Bayesian-based detector is difficult to implement.

Conclusions and some thoughts on possible future work are presented in Chapter 7. The appendices contain the detailed derivations of some results obtained in this dissertation.

1.2 Notation

The following notation is generally used throughout the dissertation. More specialized notation will be established in subsequent sections when needed.

Lower- and upper-case boldface letters are used to denote column vectors and matrices, respectively. The symbol $\mathrm{diag}(s_1, s_2, \cdots, s_n)$ denotes a diagonal matrix whose main diagonal elements are $s_1, s_2, \cdots, s_n$. The all-zero vector is denoted by $\mathbf{0}$. Matrix $\mathbf{I}$ stands for the

identity matrix. Superscripts $*$, T and H denote the complex conjugation, transposition and Hermitian transposition (complex conjugation and transposition combined), respectively. The statistical expectation operator is denoted by $\mathbb{E}$. The variance of a random variable is denoted by $\mathrm{var}[.]$, where the quantity enclosed is the random variable of interest. The symbol $\mathcal{CN}(\mu, \mathbf{R})$ denotes the circularly-symmetric, complex Gaussian distribution with mean vector μ and covariance matrix $\mathbf{R}$.

CHAPTER 2

BACKGROUND

This chapter reviews the literature related to the topics studied for this dissertation.

2.1 Fading Channels and Estimations

For simplicity only frequency-flat fading channels are considered in this dissertation. This is because dispersive multipath channels can be converted into multiple flat fading channels by using orthogonal frequency division multiplexing (OFDM).

For a flat fading channel, the discrete-time received signal at time k can be formulated as

$$r_k = h_k s_k + n_k, \tag{2.1}$$

where s_k is the transmitted symbol, h_k and n_k are samples of the fading and noise processes, respectively. Without loss of generality the channel gain is normalized such that $\mathbb{E}[|h_k|^2] = 1$. The noise power is $\mathbb{E}[|n_k|^2] = N_0$. For Rayleigh fading and Gaussian noise, the channel gains and noise samples are distributed as $h_k \sim \mathcal{CN}(0,1)$ and $n_k \sim \mathcal{CN}(0, N_0)$, respectively. The noise process is assumed to be white.

It is assumed that data are transmitted over the channel in blocks of symbols of length L, and that the fading between different blocks is independent. This model is similar to the one used in [1] and it might apply to a TDMA system, where the user does not have access to contiguous blocks.

In order for the receiver to obtain an estimate of the channel gains, pilot symbols are inserted into the data sequences. In each L-symbol block, the pilot symbols occupy L_P symbols. The power of each pilot symbol is E_P. In the sequel, the pilot symbols

and the corresponding received signals are represented by the $L_P \times 1$ vectors $\mathbf{s}_P$ and $\mathbf{r}_P$, respectively.

Block-Fading Channels

For a block-fading channel, the channel gain remains constant over the transmission block. Let h be the channel gain. Then the input-output relationship can be written

$$\mathbf{r}_P = h\mathbf{s}_P + \mathbf{n}_P, \tag{2.2}$$

where $\mathbf{n}_P$ is the vector of noise samples with the same time indices as $\mathbf{r}_P$.

The optimal maximum-likelihood estimate for the channel gain is

$$\hat{h} = \frac{1}{L_P E_P}\mathbf{s}_P^H \mathbf{r}_P, \tag{2.3}$$

The resulting mean-squared error is

$$\sigma_{\hat{h}}^2 = \mathbb{E}[(h-\hat{h})(h-\hat{h})^*] = \frac{N_0}{L_P E_P}. \tag{2.4}$$

It is noted that for a block fading channel, the channel estimation error is inversely proportional to the SNR E_P/N_0 and the number of pilot symbols L_P. Thus, for systems with high SNR, or slow fading (large L), it is possible to obtain a very accurate estimate of the channel with relatively small amount of training resource. The assumption of perfect CSI is reasonable for these cases.

Time-Varying Channels

For a time-varying channel, the fading process h_k is assumed to be stationary and correlated in time. The dynamics of the time-varying fading process are characterized by the power spectral density function (or equivalently, the autocorrelation function) with specified normalized fading rate, f_D, defined as the maximum Doppler spread normalized by the symbol rate. For mobile communications, Doppler spread is usually determined by motion of the terminals

and the carrier frequency. For example, at carrier frequency 1800 MHz, a mobile terminal moving 10 m/s imparts a maximum Doppler spread of 60 Hz. If the symbol rate is 30×10^3 symbols/sec, then the normalized fading rate is $f_D = 0.002$.

For Clarke's two-dimensional isotropic scattering model [2], the power spectral density function of the continuous fading process is

$$H_{\mathrm{C}}(\omega) = \begin{cases} \frac{1}{\pi f_D \sqrt{1-(\omega/2\pi f_D)^2}}, & |\omega| < 2\pi f_D \\ 0, & \text{otherwise,} \end{cases} \tag{2.5}$$

and the corresponding autocorrelation function is given by

$$\mathbb{E}[h_i h_j^*] = J_0(2\pi f_D(i-j)), \tag{2.6}$$

where J_0 is the zero-order Bessel function of the first kind. For simpler results and more insight, it is useful to also consider a fading process with an ideal lowpass equivalent spectrum and with the same cutoff frequency as (2.5):

$$H_{\mathrm{L}}(\omega) = \begin{cases} \frac{1}{2f_D}, & |\omega| < 2\pi f_D \\ 0, & \text{otherwise.} \end{cases} \tag{2.7}$$

For the time-varying channel, the input-output relationship can be written

$$\mathbf{r}_P = \mathbf{S}_P \mathbf{h}_P + \mathbf{n}_P, \tag{2.8}$$

where $\mathbf{S}_P = diag(\mathbf{s}_P)$, $\mathbf{h}_P$ and $\mathbf{n}_P$ are respectively the vectors of channel gains and noise samples with the same time indices as $\mathbf{r}_P$.

The optimal linear minimum mean squared error estimate for the channel gain h_k is

$$\hat{h}_k = \sigma_{h_i \mathbf{h}_P} \mathbf{S}_P^H \Sigma_{\mathbf{r}_P \mathbf{r}_P}^{-1} \mathbf{r}_P, \tag{2.9}$$

where $\sigma_{h_k\mathbf{h}_P} = \mathbb{E}[h_k\mathbf{h}_P^H]$, $\Sigma_{\mathbf{r}_P\mathbf{r}_P} = \mathbf{S}_P\Sigma_{\mathbf{h}_P\mathbf{h}_P}\mathbf{S}_P^H + N_0\mathbf{I}$, with $\Sigma_{\mathbf{h}_P\mathbf{h}_P} = \mathbb{E}[\mathbf{h}_P\mathbf{h}_P^H]$. The resulting mean-squared error is

$$\sigma_{\tilde{h}_k}^2 = \mathbb{E}[(h_k - \hat{h}_k)(h_k - \hat{h}_k)^*] = 1 - \sigma_{h_k\mathbf{h}_P}\mathbf{S}_P^H\Sigma_{\mathbf{r}_P\mathbf{r}_P}^{-1}\mathbf{S}_P\sigma_{h_k\mathbf{h}_P}^H. \tag{2.10}$$

In general, the channel estimation error is dependent on the time index k, and through $\sigma_{h_k\mathbf{h}_P}$ and $\Sigma_{\mathbf{r}_P\mathbf{r}_P}$, is dependent on the positions of the pilot symbols. For simplicity, it is assumed that at the transmitter of a pilot-aided system, a pilot symbol is inserted after every block of $(T_P - 1)$ data symbols. The pilot symbol insertion frequency $1/T_P$ satisfies the Nyquist criterion $1/T_P \geq 2f_D$ so that no aliasing is incurred.

Estimation errors are mainly due to the fact that the estimation process utilizes noisy observations. In addition to that, for a strictly band-limited fading process, the correlation function is theoretically unlimited in time. The block length L provides the time support for observations of the fading process. According to [3], for a process with time-bandwidth product $\simeq 1$, the block length has to be $L \geq f_D^{-1}$ for adequate time support. A finite observation time, i.e., $L < \infty$, will serve as an additional source of estimation error. To simplify the analysis, the latter error can be eliminated by assuming that $L \to \infty$. In this case, the sequence of pilot symbols is infinite (although the allocation of power and bandwidth is finite), the channel estimation error is independent of the position of the symbol in the block, and therefore time indices are suppressed in the sequel. The actual channel coefficient h can be expressed as

$$h = \hat{h} + \tilde{h}, \tag{2.11}$$

where $\hat{h}$ denotes the estimate of h, and $\tilde{h}$ denotes the estimation error. For infinite block length, the channel estimation error is given by (see, e.g., [4])

$$\sigma_{\tilde{h}}^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} H(\omega)\left[1 - W(\omega)\right] d\omega, \tag{2.12}$$

where $H(\omega)$ is the power spectral density function of the fading process, and

$$W(\omega) = \left[1 + T_P \left(\frac{E_P}{N_0} H(\omega)\right)^{-1}\right]^{-1} \tag{2.13}$$

is the transfer function of the optimal Wiener filter used for channel interpolation. For a lowpass Doppler spectrum, such as given by (2.7), the MMSE of the channel estimate calculated from (2.12) and (2.13), is given by

$$\sigma_{\tilde{h}}^2 = \left(1 + \frac{1}{2 f_D T_P} \frac{E_P}{N_0}\right)^{-1}. \tag{2.14}$$

2.2 Performance Measures

In this work, the average SNR and average mutual information are used to measure the average system performance. The outage probability characterizes the worst-case performance. The average performance measures are useful when the system experiences the ergodic fluctuations of the channels over a sufficiently long time. The worst-case performance measure is more important when the delay and complexity constraints of the system prevent the inherent time diversity of the channels from being exploited.

Average SNR

Let S denote the random variable representing the SNR of the combined received signal. The statistical average of S over the fading distributions, $\mathbb{E}[S]$, provides a simple measure of average system performance.

Average Mutual Information

The average mutual information measures the maximum achievable rate for reliable communication. For systems with additive white Gaussian noise (AWGN), Gaussian signaling, and perfect CSI at the receiver, the mutual information becomes the ergodic capacity of the

channel

$$C = \mathbb{E}\left[\log_2\left(1+S\right)\right].$$

The average mutual information and the average SNR are closely related by Jensen's inequality. The average SNR provides an upper bound on the ergodic capacity

$$\mathbb{E}\left[\log_2\left(1+S\right)\right] \leq \log_2\left(1+\mathbb{E}\left[S\right]\right).$$

Outage Probability

The outage probability, which characterizes the worst-case performance, is defined as the probability that the SNR of the combined signal falls below some pre-specified threshold t, i.e.,

$$P_{\text{out}} \triangleq \Pr\left[S < t\right].$$

The threshold SNR t is the required SNR for the system to maintain a certain level of performance. [1]

Clearly for a given SNR threshold, the outage probability is a decreasing function of ρ. When SNR is high, the relation $P_{\text{out}} \propto \rho^{-n}$ holds for most diversity schemes in the literature, where n is usually referred to as the diversity order, and in most cases n is an integer.

2.3 Related Literature

A growing body of literature is available that explores the ultimate limits of achievable rates of wireless communications systems over fading channels. Results have been published for a large variety of cases, depending on channel models and on the availability of channel state information (CSI) at the transmitter and receiver [5].

[1] For example, this threshold can be directly tied to the transmission rate R (in bits/symbol), through $R = \log_2\left(1+t\right)$, or equivalently, $t = 2^R - 1$.

Achievable Rates without Perfect CSI

It is well known that when the CSI is perfectly known to the receiver, the channel capacity is achieved by independent and identically distributed (i.i.d.) Gaussian signals. If, in addition, the transmitter also has the CSI, then Gaussian signaling is still optimal, and the capacity can be achieved by "waterfilling" the transmit power. However, in scenarios where neither the transmitter nor the receiver have CSI, the channel and capacity achieving signaling are not as well understood. Considerable attention has been paid to wideband channels with large number of dimensions in either spectral, temporal, or spatial domains. In the limit, when the number of channel dimensions is infinite, this capacity is the same as that of the additive white Gaussian noise (AWGN) channel: the capacity can be attained by any zero-mean signaling for known channel, and by "flash signaling" for unknown channels. Stemming from that, for vanishing capacity (in bits per dimension), knowledge of CSI is useless. Significant differences between known and unknown channels, however, start to emerge when the number of transmitted bits per channel dimension is not zero [6].

Over unknown wideband channels that decorrelate in frequency and time, the mutual information for signals with bounded peakiness tends to zero as the number of dimensions increases to infinity [7,8,9]. The interpretation provided in the references is that each degree of freedom of the channel needs to be measured, and as the number of degrees of freedom increases there is less and less energy for each to be estimated accurately. The solution suggested to this problem is to concentrate the energy in fewer degrees of freedom, which can then be estimated with high accuracy. This leads to peakiness in time or frequency. The capacity of the multiple-input-multiple-output (MIMO) channel with block Rayleigh fading was studied in [10, 11]. It is shown that when the number of transmitter antennas is large and the channel coherence time is limited, the optimal signaling is peaky in space as well: the number of transmit antennas used should be no more than half of the coherence time (in symbols).

The time-varying nature of the channel is challenging since perfect knowledge of

the state information is not possible regardless of how much effort goes into estimating the channel. Early on, Viterbi has shown that for M-ary orthogonal signaling, in the limit when the cardinality of the source alphabet increases to infinity, the mutual information representing the channel randomness subtracts directly from the channel capacity with known CSI [12]. The capacity problem without CSI has been studied for some simplified channel models, e.g., memoryless Rayleigh-fading channels [13, 14], block-fading channels [10, 1, 11, 15, 16, 17], channels with unknown phases [18, 19], and finite-state Markov channels [20, 21]. Alternatively, in [22], the channel capacity is studied as a function of channel measurement errors, rather than being tied directly to a channel model. The unknown channel problem can be solved by joint decoding and channel estimation [4]. However, joint solutions are not affordable for most systems of practical interest, since the receiver complexity grows exponentially with the block length. A suboptimal, but practical, approach is to estimate the CSI by training with known pilot symbols inserted in the transmitted data sequence. Due to the presence of noise and the time variation of the channel, the receiver is provided with imperfect CSI and therefore the system performance depends on the quality of the CSI [22, 23]. Capacities of pilot-aided systems have been studied in [24, 25, 26, 27].

Publications referenced above, such as [13, 10], are part of an increasing body of literature that states that when CSI is not available, the capacity achieving distribution is discrete with one of the signals located at zero. This is in marked contrast to channels with perfect CSI for which continuous Gaussian signaling is optimal. However, the exact characterization of the capacity achieving signaling is a difficult problem that has been addressed mostly numerically and only for a few cases. For example, for fast fading (memoryless) channels, it is found empirically that binary signaling is optimal for low SNR and that the required number of signal levels increases monotonically with SNR [13]. Full characterization of the capacity achieving signaling is a daunting task, which consists of determining the number of signals in the source alphabet, their levels and their probabilities. Practical

implementation seems far fetched.

In Chapter 3 the achievable rates are evaluated for Gaussian signals. Optimal system parameters are determined by maximizing the achievable rates.

Joint Channel Estimation and Detection for MIMO Systems

The use of multiple antennas at both the transmitter and receiver in wireless communications has been shown to provide significant increase in capacity [28, 29]. In particular, if the fading between pairs of transmit and receive antenna are independently Rayleigh distributed, it is known that for high enough signal-to-noise ratio (SNR), the average capacity grows linearly with the smaller of the number of transmit or receive antennas, provided that perfect channel state information (CSI) is available at the receiver. However in the real world perfect CSI is never known *a priori*. In practice an estimate of CSI is obtained from known pilot symbols and subsequently used for decoding as if it were exact. Therefore the performance depends on the quality of the channel estimate and hence the number of pilot symbols. Estimation of the multiple-input-multiple-output (MIMO) fading channel is a major challenge for multiple antenna systems. When the number of antennas increases, accurate channel estimation becomes more difficult because of the increase in the number of parameters to be estimated. In [26, 24], the effect of the availability of pilot symbols on MIMO fading channels is determined by information theoretic approaches.

Recently there is increasing interest in joint channel and data decoding [30, 31, 32, 33, 21, 34, 35, 36], where data decision obtained from the decoding, either hard or soft, is used as additional training to refine the channel estimate. For single transmit antenna systems the optimal maximum *a posteriori* (MAP) channel estimate can be obtained by applying a data-independent Wiener filter to the maximum likelihood (ML) estimate. For multiple transmit antenna systems, however, the ML channel estimate does not exist, and it can be shown the computation of the optimal channel estimate involves the inversion of a data-dependent matrix [30, 32, 34]. Due to the high computation complexity of matrix inversion

and the fact that the matrix inversion has to be calculated in each iteration, optimal decision directed channel estimation for MIMO channels is difficult to implement in practice.

In [37, 38] an iterative method is proposed to reduce the complexity by avioding the data-depedent matrix inversion. This method is presented in Chapter 4.

Noncoherent Detection

For noncoherent modulation and detection, no attempt is made to estimate the channels parameters. Noncoherent detection is an attractive solution particularly when the phase of the channel cannot be tracked accurately. There are two major approaches to detect signals with unknown parameters. The first is the Bayesian approach, which considers the unknown parameters as realizations of random variables with known statistics. The second is the generalized likelihood ratio test (GLRT), which uses the maximum likelihood estimate (MLE) of the unknown parameter in a likelihood ratio test (LRT) [39]. The Bayesian-based noncoherent detector receives much of the attention in the literature because it exploits the prior information of the channel and is optimal in the maximum-likelihood (ML) sense. It is also noted that for the case of unitary signals and independent and identically distributed Rayleigh fading, the Bayesian-based detector and the GLRT-based detector are equivalent. With the Bayesian approach the unknown channel parameters are averaged out with respect to the known statistics, leading to the optimal maximum-likelihood detection techniques such as square-law detection of orthogonal signals or ML sequence detection of differential phase-shift keying (DPSK) signals [40,41,42]. However, for general channels, Bayesian-based detector may be difficult to implement since multiple integration is involved. In Chapter 6 two noncoherent detection algorithms are derived for multichannel receivers using the GLRT approach, which do not require knowledge of the channel statistics at the receiver.

Cooperative Relaying

Cooperative relaying has been introduced in recent literature as an efficient means for spatially dispersed nodes in wireless networks to relay signals for one another in order to exploit spatial diversity (referred to as *cooperative diversity* [43,44,45,46,47]) in fading channels, and to transmit information with minimum power [48,49].

For a more efficient use of the power resources, the problems of optimally distributing power among nodes have been considered in [50, 51, 52, 53]. When instantaneous channel gains are known at the transmitters, adaptive allocations are used to optimize the performance for amplify-and-forward (AF) diversity in [50, 51]. In [52] the optimal allocation are developed that minimizes the outage probability for decode-and-forward (DF) diversity and used as an approximate solution for AF diversity. Simple power allocation strategies with mean channel gain information are presented for DF and AF diversity in [53,54]. In Chapter 5 power allocations are developed to optimize both average and outage performance.

Among various cooperative diversity schemes, fixed AF relaying [44] has been shown to be promising. Besides its simplicity, it achieves full second-order diversity, provided that *global* knowledge of the channel state information (CSI) is available and the optimal maximal-ratio combining (MRC) is employed at the destination. By contrast, fixed DF relaying does not achieve full diversity. More complicated protocols, which exploit feedback from the destination, are required for the DF relaying to achieve full diversity [44,47].

Most work in the literature on AF diversity focuses on the case that global CSI is known [44, 55, 56, 57, 58, 59]. While the availability of perfect knowledge of global CSI simplifies the performance analysis, this may not be true in practice. For example, the source-relay channel cannot be explicitly measured at the destination, and it may take too much overhead to obtain this information. Therefore it is appealing to consider systems without CSI or with more realistic knowledge of CSI [60, 61, 62]. In Chapter 6 GLRT-based noncoherent detectors are considered for AF diversity, for which the optimal ML noncoherent detector is difficult to implement.

CHAPTER 3

ACHIEVABLE RATES OVER TIME-VARYING CHANNELS

This chapter considers a channel subject to a continuous flat fading process, i.e., the fading changes during a block of symbols rather than remaining constant. This channel model has received less attention in the literature (exceptions to that are recent work [26, 63, 64]). In particular, information rates of time-varying, Rayleigh fading channels are studied under the assumption that no *a priori* knowledge of the CSI is available at either the receiver or the transmitter. The dynamics of the fading channel are characterized by its Doppler spectrum with specified normalized fading rate f_D, known to the receiver [65]. Since little is known about the exact structure of optimal signaling under those conditions, and due to the difficulty of implementing signals for which the alphabet size, levels and probabilities depend on channel conditions, it is of interest to evaluate what information rates are achievable by applying, and possibly modifying, familiar modulation formats. In particular Gaussian signaling is focused since it is optimal when the CSI knowledge is perfect, and is asymptotically optimal even in the absence of CSI, when the fading is very slow. Gaussian signaling also makes it possible to obtain closed-form expressions that afford insight and intuition into the problem of information rates over channels without CSI. Finally, it is noted that it is not difficult to carry out the analysis presented in the chapter for other signaling formats (on-off keying, binary antipodal, orthogonal). However in those cases it does not seem possible to obtain closed-form solutions.

This chapter is organized as follows. The system and channel models are briefly described in Section 3.1. Section 3.2 contains the derivation of the penalty on information rate and a lower bound on capacity. The performance of pilot-aided systems and the optimal resource allocation are presented in Section 3.3. The asymptotic behavior in the high and

low SNR regimes are analyzed in Section 3.4. Section 3.5 provides the treatment on the low-duty-cycle signaling in the low SNR regime. Conclusions are given in Section 3.6.

In the sequel, $\mathcal{H}$ and $\mathcal{I}$ are used to denote the differential entropy and mutual information, respectively. The function ln is the natural logarithm so that information rates and entropies are expressed in natural units (nats, 1 nat = $1/(\ln 2)$ bits).

3.1 System and Channel Model

The channel between the transmitter and receiver is modeled as a flat fading process. The discrete time received signal is given by (2.1). The average transmitted signal power is $E_s = \mathbb{E}[|s_k|^2]$. Let ρ denote the average signal-to-noise ratio (SNR) per symbol at the receiver, i.e., $\rho = E_s/N_0$.

The fading process h_k is stationary, with normalized fading rate f_D. The effects on achievable information rates are analyzed in the next section, for the fading processes with Clarke's spectrum (2.5) and the low pass sprectrum (2.7).

3.2 A Lower Bound on Mutual Information

Let the transmitted block consist of L symbols. Let $\mathbf{s} = [s_1\ s_2\ \cdots\ s_L]^T$ be the sequence of transmitted symbols. The received sequence, denoted by an $L \times 1$ vector $\mathbf{r} = [r_1\ r_2\ \cdots\ r_L]^T$, is of the form

$$\mathbf{r} = \mathbf{S}\mathbf{h} + \mathbf{n}, \tag{3.1}$$

where $\mathbf{S} = \mathrm{diag}(s_1, s_2, \cdots, s_L)$, $\mathbf{h} = [h_1\ h_2\ \cdots\ h_L]^T$, and $\mathbf{n} = [n_1\ n_2\ \cdots\ n_L]^T$. The noise sample vector $\mathbf{n}$ is complex Gaussian, distributed in L dimensions with zero mean vector $\mathbf{0}$, and covariance matrix $N_0\mathbf{I}$. The channel parameter vector $\mathbf{h}$ is complex Gaussian with zero mean vector and Toeplitz positive semidefinite covariance matrix $\mathbf{R_h}$, whose elements are determined by the normalized fading rate. Under these assumptions, one can readily

show that given $\mathbf{S}$, $\mathbf{r}$ has an L-dimensional complex Gaussian distribution with mean vector 0 and covariance matrix $\mathbf{S}\mathbf{R}_\mathbf{h}\mathbf{S}^H + N_0\mathbf{I}$ (see, e.g., [66]).

The mutual information between $\mathbf{r}$ and $\{\mathbf{s}, \mathbf{h}\}$ can be expanded by using the chain rule as follows (see [67]):

$$\begin{aligned} \mathcal{I}(\mathbf{r}; \mathbf{s}, \mathbf{h}) &= \mathcal{I}(\mathbf{r}; \mathbf{s}) + \mathcal{I}(\mathbf{r}; \mathbf{h}|\mathbf{s}) \\ &= \mathcal{I}(\mathbf{r}; \mathbf{h}) + \mathcal{I}(\mathbf{r}; \mathbf{s}|\mathbf{h}). \end{aligned} \tag{3.2}$$

Thus, $\mathcal{I}(\mathbf{r}; \mathbf{s})$, the mutual information between channel input $\mathbf{s}$ and channel output $\mathbf{r}$, without CSI, can be expressed

$$\begin{aligned} \mathcal{I}(\mathbf{r}; \mathbf{s}) &= \mathcal{I}(\mathbf{r}; \mathbf{s}, \mathbf{h}) - \mathcal{I}(\mathbf{r}; \mathbf{h}|\mathbf{s}) \\ &= \mathcal{I}(\mathbf{r}; \mathbf{h}) + \mathcal{I}(\mathbf{r}; \mathbf{s}|\mathbf{h}) - \mathcal{I}(\mathbf{r}; \mathbf{h}|\mathbf{s}). \end{aligned} \tag{3.3}$$

Note that $\mathcal{I}(\mathbf{r}; \mathbf{s}|\mathbf{h})$ in (3.3) is the mutual information with perfect CSI. Therefore, $(\mathcal{I}(\mathbf{r}; \mathbf{h}|\mathbf{s}) - \mathcal{I}(\mathbf{r}; \mathbf{h}))$ is the penalty in information rate due to unknown CSI. Since $\mathcal{I}(\mathbf{r}; \mathbf{h})$ is nonnegative, $\mathcal{I}(\mathbf{r}; \mathbf{h}|\mathbf{s})$ can be interpreted as an upper bound on the penalty due to unknown CSI and since $0 \leq \mathcal{I}(\mathbf{r}; \mathbf{h}) \leq \mathcal{I}(\mathbf{r}; \mathbf{h}|\mathbf{s})$, there is always a cost due to unknown CSI. The cost vanishes if and only if $\mathcal{I}(\mathbf{r}; \mathbf{h}|\mathbf{s}) = \mathcal{I}(\mathbf{r}; \mathbf{h}) = 0$, i.e., the CSI is known.

In (3.3), the information rate without CSI, $\mathcal{I}(\mathbf{r}; \mathbf{s})$, is dependent on the distribution of the input $\mathbf{s}$. In the sequel, first an upper bound on the penalty $\mathcal{I}(\mathbf{r}; \mathbf{h}|\mathbf{s})$ is derived. This penalty is a function of the symbols $\mathbf{s}$. For any specific signaling format a lower bound on $\mathcal{I}(\mathbf{r}; \mathbf{s})$ can be obtained by starting from (3.3), neglecting the nonnegative term $\mathcal{I}(\mathbf{r}; \mathbf{h})$, evaluating $\mathcal{I}(\mathbf{r}; \mathbf{s}|\mathbf{h})$ and using the upper bound on $\mathcal{I}(\mathbf{r}; \mathbf{h}|\mathbf{s})$. Asymptotic results will be provided for long block lengths. It is noted that it is also possible to derive *upper* bounds on the achievable information rates by evaluating $\mathcal{I}(\mathbf{r}; \mathbf{h}|\mathbf{s})$, upper bounding $\mathcal{I}(\mathbf{r}; \mathbf{s}, \mathbf{h})$, and using the first relation in (3.3) [68].

3.2.1 An Upper Bound on Information Rate Penalty

The mutual information term $\mathcal{I}(\mathbf{r};\mathbf{h}|\mathbf{s})$ was previously introduced as an upper bound on the penalty due to lack of CSI. Expressing $\mathcal{I}(\mathbf{r};\mathbf{h}|\mathbf{s})$ as a difference of conditional differential entropies of the L dimensional vector $\mathbf{r}$,

$$\begin{aligned}\mathcal{I}(\mathbf{r};\mathbf{h}|\mathbf{s}) &= \mathcal{H}(\mathbf{r}|\mathbf{s}) - \mathcal{H}(\mathbf{r}|\mathbf{h},\mathbf{s}) \\ &= \mathcal{H}(\mathbf{r}|\mathbf{s}) - \mathcal{H}(\mathbf{n}),\end{aligned} \tag{3.4}$$

where for white Gaussian noise

$$\mathcal{H}(\mathbf{n}) = L\ln(\pi e N_0), \tag{3.5}$$

and

$$\begin{aligned}\mathcal{H}(\mathbf{r}|\mathbf{s}) &= \mathbb{E}\left[\ln\det(\pi e(\mathbf{S}\mathbf{R}_\mathbf{h}\mathbf{S}^H + N_0\mathbf{I}))\right] \\ &= L\ln(\pi e N_0) + \mathbb{E}\left[\ln\det(\mathbf{I} + \frac{1}{N_0}\mathbf{R}_\mathbf{h}\mathbf{S}^H\mathbf{S})\right],\end{aligned} \tag{3.6}$$

where the expectation is with respect over $\mathbf{S}$.

For general input distributions, the computation of $\mathcal{H}(\mathbf{r}|\mathbf{s})$ (and hence $\mathcal{I}(\mathbf{r};\mathbf{h}|\mathbf{s})$) is intractable. Instead, one can develop an upper bound on $\mathcal{I}(\mathbf{r};\mathbf{h}|\mathbf{s})$.

Application of Jensen's inequality to (3.6) gives

$$\mathcal{H}(\mathbf{r}|\mathbf{s}) \leq L\ln(\pi e N_0) + \ln\det\left(\mathbf{I} + \frac{1}{N_0}\mathbf{R}_\mathbf{h}\mathbb{E}\left[\mathbf{S}^H\mathbf{S}\right]\right), \tag{3.7}$$

where equality holds if and only if the input symbols have constant power, i.e., $|s_l|^2 = E_s$, $l = 1, 2, ..., L$, (e.g., M-ary phase-shift keying). Continuing the upper bounding of $\mathcal{H}(\mathbf{r}|\mathbf{s})$,

$$\begin{aligned}\mathcal{H}(\mathbf{r}|\mathbf{s}) &\leq L\ln(\pi e N_0) + \ln\det\left(\mathbf{I} + \rho\mathbf{R}_\mathbf{h}\right) \\ &= L\ln(\pi e N_0) + \sum_{l=1}^{L}\ln\left(1 + \rho\lambda_l\right),\end{aligned} \tag{3.8}$$

where λ_l, $l = 1, 2, ..., L$, are the eigenvalues of $\mathbf{R_h}$. Substituting (3.5) and (3.8) into (3.4) yields

$$\mathcal{I}(\mathbf{r}; \mathbf{h}|\mathbf{s}) \leq \sum_{l=1}^{L} \ln(1 + \rho\lambda_l). \tag{3.9}$$

As mentioned earlier, $\mathcal{I}(\mathbf{r}; \mathbf{h}|\mathbf{s})$ is interpreted as the upper bound on the penalty due to the lack of CSI at either the transmitter or receiver. The penalty attains its maximum value given by the right-hand side (RHS) of (3.9) for constant power signals. For future use, this value of the penalty normalized to the number of symbols is denoted by

$$P_\Delta = \frac{1}{L} \sum_{l=1}^{L} \ln(1 + \rho\lambda_l) \quad \text{nats/symbol.} \tag{3.10}$$

Notice that P_Δ is a function of the SNR ρ, the frame length L, and, through the assumed known channel covariance matrix $\mathbf{R_h}$, of the channel normalized fading rate f_D. Moreover, through the eigenvalues of channel covariance matrix, the penalty function is also dependent on the spectrum of the channel fading process.

It is noted that for fixed f_D and ρ, P_Δ is a nonincreasing sequence in L. This is a consequence of the fact that the matrix $\mathbf{I} + \rho\mathbf{R_h}$ is Toeplitz positive definite and Theorem 16.8.6 in [67]. This observation confirms the intuition that the channel dynamics can be learned better with an increasing sequence of observations, leading to a reduced loss in information rate.

3.2.2 Large Block Length

Since P_Δ is a nonincreasing sequence in L and is bounded below ($P_\Delta \geq 0$), it has a limit as block length L goes to infinity. In the sequel, this limit is denoted by $\widetilde{P}_\Delta$, i.e.,

$$\widetilde{P}_\Delta = \lim_{L \to \infty} P_\Delta.$$

By Szegö's theorem ([69], pp. 64-65), one can obtain the following asymptotic relation between the eigenvalues of the correlation matrix of the fading process and its

power spectral density:

$$\begin{aligned}\widetilde{P}_{\Delta} &= \lim_{L\to\infty}\frac{1}{L}\sum_{l=1}^{L}\ln(1+\rho\lambda_l)\\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\ln\left(1+\rho H(\omega)\right)d\omega,\end{aligned} \tag{3.11}$$

where $H(\omega)$ is the power spectral density function of the fading process h_k.

For Clarke's model with autocorrelation function (2.6), and power spectral density (2.5), the asymptotic value $\widetilde{P}_{\Delta,\mathrm{C}}$ can be evaluated analytically as shown in Appendix A.1, and is given by

$$\widetilde{P}_{\Delta,\mathrm{C}} = 2f_D\phi\left(\frac{\rho}{\pi f_D}\right), \tag{3.12}$$

where the function ϕ is defined as

$$\phi(x) = \begin{cases} 0, & x = 0 \\ \frac{\pi}{2}x + \ln\frac{x}{2} + \sqrt{1-x^2}\ln(\frac{1+\sqrt{1-x^2}}{x}), & 0 < x \leq 1 \\ \frac{\pi}{2}x + \ln\frac{x}{2} - 2\sqrt{x^2-1}\tan^{-1}(\sqrt{\frac{x-1}{x+1}}), & x > 1. \end{cases} \tag{3.13}$$

Expression (3.13) looks fairly complicated. For a simpler expression that yields better insight, one can also consider a fading process with an ideal lowpass equivalent spectrum with the same cutoff frequency as (2.7). Substitution of (2.7) into (3.11) leads to

$$\widetilde{P}_{\Delta,\mathrm{L}} = 2f_D\ln\left(1+\frac{\rho}{2f_D}\right). \tag{3.14}$$

This result indicates that the penalty for not knowing the channel is upper-bounded by the information contained in the fading process.

Fig. 3.1 shows (3.10), the upper bound on the information rate penalty due to unknown CSI, versus block length L for an SNR per symbol $\rho = 10$ dB, and for normalized fading rates $f_D = 0.01$ and 0.05. As expected, the penalty decreases in opposite direction

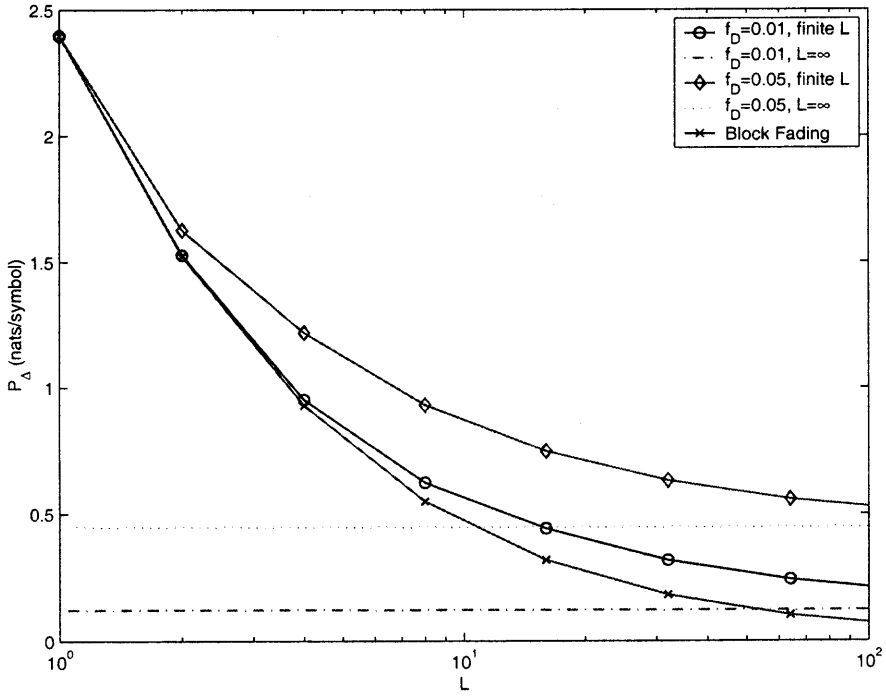

Figure 3.1 Upper bound on information rate penalty versus block length for Clarke's Doppler spectrum ($\rho = 10$ dB).

of the block length L, and it approaches the asymptotic bound in (3.12) as L goes to infinity. Conversely, short blocks do not allow observation of the channel over a period of time sufficient for learning the channel dynamics, and hence carry a higher penalty due to unknown CSI. The information rate penalty for block fading is also shown in the figure. It can be observed that to the extent that the block fading model is used as an approximation to continuous fading, it underestimates the loss in information rate.

Fig. 3.2 is a graph of the rate penalties for long blocks ($L \to \infty$) for the two fading spectra discussed above, (3.12) and (3.14), as a function of the SNR ρ, for various fading rates. It can be seen that the penalties for the ideal lowpass fading spectrum are only slightly higher than those for Clarke's spectrum with the same fading rate. With information rates seeming to be insensitive to the actual shape of the Doppler spectrum, and due to the simpler expressions, in the sequel, the ideal lowpass spectrum is used for performance

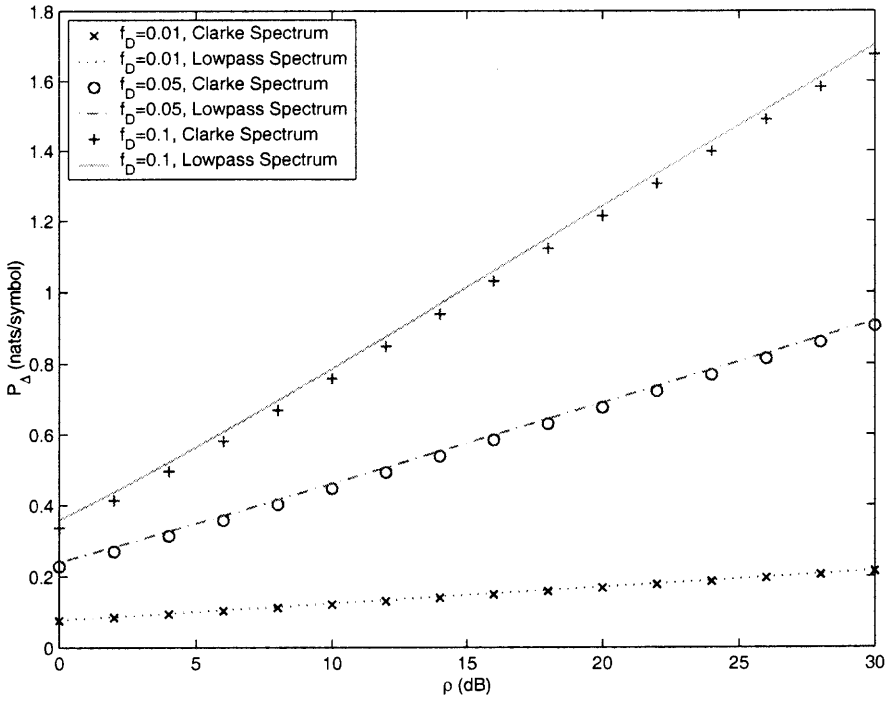

Figure 3.2 Asymptotic (large block length) information rate penalty versus SNR.

analysis. Observe that, as expected, the penalty increases with the fading rate and with SNR. The effect of SNR will be revisited in the next section.

3.2.3 A Lower Bound on Achievable Rate

Since the term $\mathcal{I}(\mathbf{r};\mathbf{h})$ in (3.3) is nonnegative, the mutual information for unknown CSI is lower-bounded by

$$\mathcal{I}(\mathbf{r};\mathbf{s}) \geqslant \mathcal{I}(\mathbf{r};\mathbf{s}|\mathbf{h}) - \mathcal{I}(\mathbf{r};\mathbf{h}|\mathbf{s}). \tag{3.15}$$

The second term on the right-hand side (RHS) of (3.15) is upper bounded in (3.9). The first term on the RHS of (3.15) is just the mutual information between the input and the output of the fading known channel and it is maximized by input symbols i.i.d. with Gaussian

density [67]

$$\begin{aligned} \mathcal{I}(\mathbf{r};\mathbf{s}|\mathbf{h}) &= \sum_{l=1}^{L} \mathcal{I}(r_l; s_l|h_l) \\ &= LC_{\text{Rayleigh}}(\rho), \end{aligned} \tag{3.16}$$

where $C_{\text{Rayleigh}}(x)$ is the ergodic capacity per symbol of a Rayleigh fading channel with perfect CSI as a function of average SNR. The capacity of the fading channel is given by (see [70, 71]):

$$\begin{aligned} C_{\text{Rayleigh}}(x) &\triangleq \mathbb{E}[\ln(1+x\lambda)] \\ &= -\exp\left(x^{-1}\right) \text{Ei}\left(-x^{-1}\right) \quad \text{nats/symbol}, \end{aligned} \tag{3.17}$$

where the expectation is with respect to an exponentially distributed random variable λ with unity mean, and Ei is the exponential-integral function

$$\text{Ei}(x) = \int_{-\infty}^{x} \frac{\exp(-t)}{t} dt. \tag{3.18}$$

The lower bound on the achievable information rate without CSI is formulated as the following proposition:

Proposition 1 *The capacity per symbol of a time-varying channel without CSI is lower-bounded by*

$$C \geq C_{Rayleigh}(\rho) - \frac{1}{L}\sum_{l=1}^{L} \ln(1+\rho\lambda_l).$$

As L goes to infinity, the following information rate can be achieved by i.i.d. Gaussian signaling:

$$R_{LB} = C_{Rayleigh}(\rho) - 2f_D \ln\left(1+\frac{\rho}{2f_D}\right) \quad \textit{nats/symbol}. \tag{3.19}$$

The proof of the proposition follows from (3.9), (3.15) and (3.16).

The block fading channel is a special case of the time-varying channel addressed here. The following corollary specializes the proposition to the block fading channel:

Corollary 1 *Let L be the number of symbols in the coherence interval of a block fading channel. Then the capacity per symbol of this channel without CSI is lower-bounded by*

$$C \geq C_{Rayleigh}(\rho) - \frac{1}{L}\ln(1+\rho L) \quad \textit{nats/symbol.} \tag{3.20}$$

The corollary is obtained directly from the observation that for block fading the only nonzero eigenvalue equals L. As L goes to infinity, this lower bound approaches the capacity with perfect CSI and that can be achieved with i.i.d. Gaussian signaling.

It is noted that for quasi-static channels, similar results have been derived for a variety of cases; see, e.g., [18, 19, 10, 15, 16].

The task of achieving the channel capacity for unknown channels is difficult even for channels that can be represented by a finite number of states such as the discrete memoryless channel [72]. It is more so for the continuous fading channel considered here. A suboptimal approach often adopted in practice is to estimate the channel from a training sequence referred to as pilot, and to follow that by conventional maximum-likelihood detection using the estimated channel values as a known channel. The design of such systems is considered in the next section.

3.3 Pilot-Aided Systems

This section studies achievable information rates for a system using pilot-aided channel estimation, and determines the optimal resource allocation for the pilot and data symbols that maximizes the information rate.

Recall that for a time-varying fading with lowpass Doppler spectrum given by (2.7), the MMSE of the channel estimate is given by (2.14), where E_P is the power of each pilot symbol, and $1/T_P$ is the pilot insertion frequency.

Next, the channel estimation error is regarded as an additive noise term with the variance given in (2.14), and it is used to evaluate system performance.

Let α and β denote the fractions of bandwidth and power devoted to pilot symbols, respectively. Using E_D to denote the average data symbol power, then for the pilot

$$\alpha = \frac{1}{T_P}, \tag{3.21}$$

and

$$E_P = \beta T_P E_s = \frac{\beta}{\alpha} E_s, \tag{3.22}$$

and for the data

$$E_D = \frac{T_P}{T_P - 1}(1 - \beta) E_s = \frac{1 - \beta}{1 - \alpha} E_s. \tag{3.23}$$

In the special case when the pilot symbol and data symbols are constrained to have the same power, i.e., $E_P = E_D = E_s$, the fractions of bandwidth and power that are allocated to the pilot are equal, $\alpha = \beta = 1/T_P$.

3.3.1 Effective SNR

By substituting (2.11) in (2.1), the signal model is rewritten

$$\begin{aligned} r &= s\hat{h} + s\tilde{h} + n \\ &= s\hat{h} + z. \end{aligned} \tag{3.24}$$

In this equation, $s\hat{h}$ is the signal term, and $z = s\tilde{h} + n$ as an additive noise term. The effective SNR can be expressed

$$\begin{aligned} \rho_{\text{eff}} &= \frac{\mathbb{E}\left[\left|s\hat{h}\right|^2\right]}{\mathbb{E}\left[|z|^2\right]} \\ &= \frac{\mathcal{E}_D\left(1 - \sigma_{\tilde{h}}^2\right)}{\mathcal{E}_D \sigma_{\tilde{h}}^2 + N_0}. \end{aligned} \tag{3.25}$$

Upon substitution of (3.22), (3.23) and (2.14), and after some manipulations, (3.25) becomes

$$\rho_{\text{eff}} = \eta\rho, \tag{3.26}$$

where η is the SNR efficiency due to imperfect CSI:

$$\eta = \frac{(1-\beta)\,\beta\rho}{(1-\alpha)\,(2f_D + \beta\rho) + 2f_D\,(1-\beta)\,\rho}. \tag{3.27}$$

In (3.27), the SNR efficiency due to the use of pilot symbols is expressed in terms of the power and bandwidth allocations to the pilot. The expression suggests that suitable choices of those allocations will optimize performance.

3.3.2 Rates with Optimal Resource Allocation

This section deals with systems with imperfect CSI. As is the case with systems with unknown CSI, little is known about the capacity and the capacity achieving signaling for imperfect CSI. Taking the same approach taken for unknown channels, one can seek to derive a useful lower bound on the capacity by evaluating the information rate achievable with Gaussian signaling.

The capacity of a system with pilot-aided channel estimation is equivalent to that of a system with perfect CSI but with a loss in bandwidth and power reflecting the use of the pilot signals and the CSI estimation errors. The mutual information of the Rayleigh fading channel with Gaussian signaling serves as a lower bound to the capacity with unconstrained signaling. The information rate achieved by Gaussian signaling can be expressed as follows in terms of the capacity with known CSI:

$$R_{\text{P}} = (1-\alpha)\,C_{\text{Rayleigh}}\,(\eta\rho) \quad \text{nats/symbol}, \tag{3.28}$$

where the subscript denotes "pilot", and C_{Rayleigh} is the ergodic capacity of a Rayleigh fading channel with perfect CSI given by (3.17). The information rate (3.28) is also expressed in terms of α and η, which represent the loss in bandwidth and the power

efficiency, respectively, due to the use of pilot symbols. Clearly, $R_{\mathrm{P}} \leq C_{\mathrm{Rayleigh}}$, with equality if and only if $\alpha = 0$, $\eta = 1$ (i.e., perfect CSI and no pilot). It can be seen, from (3.27) and (3.28), that the information rate depends on the resources allocated to the pilot symbols, through the parameters α and β. In the sequel, the resource allocations α and β that optimize the rate (3.28) will be determined, for fixed values of f_D and ρ.

Optimal Pilot Symbol Insertion Frequency It will be shown that for fixed values of f_D, ρ, and β, the achievable information rate (3.28) is a decreasing function of α. The approach taken below is suggested by [24], where the problem is solved for the block fading channel. Take the derivative of R_{P} with respect to α:

$$\begin{aligned}
\frac{\partial R_{\mathrm{P}}}{\partial \alpha} &= -C_{\mathrm{Rayleigh}}(\eta\rho) + (1-\alpha)\frac{\partial}{\partial \alpha} C_{\mathrm{Rayleigh}}(\eta\rho) \\
&= \mathbb{E}\left[-\ln(1+\eta\rho\lambda) + (1-\alpha)\frac{\rho\lambda}{1+\eta\rho\lambda}\frac{\partial \eta}{\partial \alpha}\right] \\
&= \mathbb{E}\left[-\ln(1+\eta\rho\lambda) + \right. \\
&\qquad \left. \frac{\eta\rho\lambda}{1+\eta\rho\lambda}\frac{(1-\alpha)(2f_D+\beta\rho)}{(1-\alpha)(2f_D+\beta\rho)+2f_D(1-\beta)\rho}\right] \\
&< \mathbb{E}\left[-\ln(1+\eta\rho\lambda) + \frac{\eta\rho\lambda}{1+\eta\rho\lambda}\right] \\
&\leq 0,
\end{aligned}$$

where the first inequality holds because

$$\frac{(1-\alpha)(2f_D+\beta\rho)}{(1-\alpha)(2f_D+\beta\rho)+2f_D(1-\beta)\rho} < 1$$

for $f_D > 0$, $\rho > 0$, and $\beta < 1$; the second inequality follows from the fact that

$$\ln(1+x) \geq \frac{x}{1+x}$$

for all nonnegative values of x.

Therefore, the information rate is maximized when α attains its minimum possible value, which is just the minimum frequency satisfying the Nyquist criterion, i.e.,

$$\alpha_{\text{opt}} = 2f_D. \tag{3.29}$$

Clearly, the optimal pilot insertion frequency does not depend on the SNR ρ. With this choice of pilot symbol insertion frequency, the coefficient η can be written as

$$\eta = \frac{(1-\beta)\beta\rho}{(1-2f_D)(2f_D+\beta\rho)+2f_D(1-\beta)\rho}. \tag{3.30}$$

Optimal Pilot Power Allocation Having established that the optimal fraction of bandwidth used for pilot symbols is given by (3.29), one can subsequently find the optimal power allocation for the pilot symbols. That will be the value of β that maximizes η in (3.30).

Expression (3.30) can be rewritten

$$\eta = \frac{1}{1-4f_D}\frac{(1-\beta)\beta}{\beta+\nu}, \tag{3.31}$$

with

$$\nu = \frac{2f_D(1-2f_D+\rho)}{\rho(1-4f_D)}. \tag{3.32}$$

The value of β that maximizes (3.31) is dependent on f_D. Assume that $f_D < 1/4$. This assumption is motivated by the following argument. If the normalized bandwidth of the transmission is 1, and it is required that less than 50% of the bandwidth be allocated to the pilot, then from Nyquist sampling considerations $f_D < 1/4$. Conversely, if more than 50% of the bandwidth needs to be allocated to the pilot, then noncoherent techniques that do not require channel estimation are probably more desirable. Under the assumption $f_D < 1/4$, ν is positive, and the optimal fraction β of power allocated to the pilot symbols, in the sense that the coefficient η given in (3.31) is maximized, can be shown to be

$$\beta_{\text{opt}} = g(\nu), \tag{3.33}$$

where the function $g(x)$ is defined as

$$g(x) = -x + \sqrt{x(x+1)}. \tag{3.34}$$

Setting β equal to this optimal value and substituting in (3.31) leads to

$$\eta_{\text{opt}} = \frac{1}{1-4f_D}\left(\sqrt{\nu+1} - \sqrt{\nu}\right)^2, \tag{3.35}$$

with ν given by (3.32).

It is not hard to verify the following properties of ν, $g(x)$, and η_{opt}:

Lemma 1 *1.* $\nu > 0$, *if* $\rho > 0$ *and* $0 < f_D < 1/4$;

2. ν *is an increasing function of* f_D;

3. ν *is a decreasing function of* ρ, *if* $0 < f_D < 1/4$;

4. $g(x)$ *is a nonnegative, increasing function of* x *for* $x > 0$, *and its value tends to* $1/2$ *as* x *tends to infinity;*

5. η_{opt} *is a decreasing function of* ν *for* $\nu > 0$.

It follows from Lemma 1, that the optimal power allocation to the pilot β_{opt} is a decreasing function of ρ and an increasing function of f_D, while η_{opt} is an increasing function of ρ and a decreasing function of f_D. In other words, if the fading becomes slower or the SNR becomes higher, a smaller fraction of the power should be allocated to pilot symbols, and using pilot symbols becomes more efficient. It can also be shown that $\beta_{\text{opt}} > 2f_D$ for all values of ρ and all values of $f_D < 1/4$, implying that the power allocated to pilot symbols should always be greater than the average power.

Since for a fixed fading rate, β_{opt} and η_{opt} are both functions of the SNR, it is of interest to examine the limiting behavior at high SNR and low SNR. These results will be presented in Section 3.4.

3.3.3 Rates with Constrained Resource Allocation

It have been shown that if it is free to choose the allocations α and β, then the optimal solution is given in (3.29) and (3.33), respectively, and the optimal values of α and β are never equal. More specifically, $\beta_{\text{opt}} > \alpha_{\text{opt}} = 2f_D$ for any SNR ρ and fading rate $f_D < 1/4$, implying that the optimal power of pilot symbols always exceeds the average power of the data symbols.

It is of interest to investigate the special case considered in [26], where the pilot symbol power and average data symbol power are constrained to be equal. To distinguish between the unconstrained and constrained cases, quantities associated with the latter are marked with the superscript '*'. With this constraint in effect, the fractions of bandwidth and power allocated the pilot symbols are equal, i.e., $\alpha^* = \beta^*$. Moreover, the constraint implies that the powers per symbol for pilot and data are equal. It is possible to find the optimal resource allocation to the pilot symbols by maximizing the information rate

$$R_{\text{P}}^{*}(\beta^*) = (1-\beta^*)\, C_{\text{Rayleigh}}(\eta^*\rho) \quad \text{nats/symbol}, \tag{3.36}$$

where the maximization is over β^* in the interval $2f_D \leq \beta^* < 1$. Letting $\alpha^* = \beta^*$ in (3.27), the coefficient η^* is given by

$$\begin{aligned} \eta^* &= \frac{\beta^*\rho}{2f_D(1+\rho)+\beta^*\rho} \\ &= \left[1+2f_D(\beta^*)^{-1}\left(1+\rho^{-1}\right)\right]^{-1}. \end{aligned} \tag{3.37}$$

No simple analytic expression for the optimal value of β^* appears to be available. It is observed, through numerical computations, that the optimal value of β^* is always greater than $2f_D$ [26]. With that, one can write $\beta^* = \alpha^* > 2f_D$ (compare to $\alpha = 2f_D$ for the unconstrained case). The intuition here is that the larger bandwidth required by the pilot, makes up for the pilot power now constrained to be the same as that of the data symbols (i.e., the pilot symbols need to be sent more often).

Although the detailed dependence of the optimal value of β^* on ρ must be determined numerically, simple asymptotic approximations do exist for the high or low SNR regimes. These results are presented in the next section.

3.4 Asymptotic Behavior

This section presents asymptotic approximations for information rates developed in previous sections, as well as the optimal resource allocation for pilot symbol aided system, in the high and low SNR regimes.

3.4.1 High SNR

In the high SNR regime, one can use the following approximation for $C_{\text{Rayleigh}}(\rho)$ [70, 71]

$$C_{\text{Rayleigh}}(\rho) \simeq \ln \rho - \gamma \quad \text{nats/symbol}, \quad \rho \gg 1, \tag{3.38}$$

where $\gamma = 0.577\cdots$ is Eüler's constant.

Mutual Information Lower Bound It can be easily seen from (3.10) that P_Δ is monotonically increasing with ρ, and $P_\Delta \to \infty$ as $\rho \to \infty$. It is of interest to study the asymptotic behavior of the $\widetilde{P}_\Delta / C_{\text{Rayleigh}}$ ratio for high SNR.

For the lowpass spectrum, the information rate penalty is given by

$$\begin{aligned} \lim_{\rho\to\infty} \frac{\widetilde{P}_{\Delta,\text{U}}}{C_{\text{Rayleigh}}} &= \lim_{\rho\to\infty} \frac{2f_D \ln(1+\rho\frac{1}{2f_D})}{\ln\rho - \gamma} \\ &= 2f_D. \end{aligned} \tag{3.39}$$

It is shown in Appendix A.2 that the same relation also holds for Clarke's spectrum. It can readily be seen that in the high SNR regime, the ratios of the capacity lower bound (3.19) to C_{Rayleigh} equal $(1 - 2f_D)$ for both types of spectra. Therefore, for large blocks and high SNR, the fractional loss in information rate due to unknown CSI equals twice the channel normalized fading rate.

The capacity lower bound (3.19) can be approximated

$$R_{\text{LB}} \simeq (1 - 2f_D)\, C_{\text{Rayleigh}}\left(\eta_{\text{LB}}\rho\right), \tag{3.40}$$

where the SNR efficiency η_{LB} can be found by equating the relation above with (3.19) and using (3.38), and is

$$\eta_{\text{LB}} = \exp\left[\frac{2f_D}{1 - 2f_D}\left(\ln\left(2f_D\right) - \gamma\right)\right]. \tag{3.41}$$

Pilot-Aided System with Optimal Allocation From (3.32) and (3.33), one can obtain the limiting value for the power allocation to pilot symbols:

$$\lim_{\rho\to\infty} \beta_{\text{opt}} = \frac{-2f_D + \sqrt{2f_D\left(1 - 2f_D\right)}}{1 - 4f_D}. \tag{3.42}$$

It follows from (3.35) that the SNR efficiency

$$\lim_{\rho\to\infty} \eta_{\text{opt}} = \frac{1}{1 + 2\sqrt{2f_D\left(1 - 2f_D\right)}}. \tag{3.43}$$

Since $1/2 < \lim_{\rho\to\infty} \eta_{\text{opt}} < 1$ provided that $0 < f_D < 1/4$, the loss in effective SNR due to the use of channel estimation is less than 3 dB when the SNR is high. This result confirms the intuition that for high SNR the use of a pilot is preferred to techniques that do not use channel estimates such as noncoherent or differential detection.

Pilot-Aided System with Constrained Allocation From (3.38) and (3.36), for any value of β^* in the interval $2f_D \le \beta^* < 1$,

$$\begin{aligned}\lim_{\rho\to\infty} \frac{R_{\text{P}}^*\left(\beta^*\right)}{R_{\text{P}}^*\left(2f_D\right)} &= \frac{1 - \beta^*}{1 - 2f_D} \lim_{\rho\to\infty} \frac{\ln\rho - \ln\left[1 + 2f_D\left(\beta^*\right)^{-1}\left(1 + \rho^{-1}\right)\right] - \gamma}{\ln\rho - \ln\left[1 + \left(1 + \rho^{-1}\right)\right] - \gamma} \\ &= \frac{1 - \beta^*}{1 - 2f_D},\end{aligned}$$

which is decreasing in β^*. It follows that

$$\lim_{\rho\to\infty} \beta_{opt}^* = 2f_D, \tag{3.44}$$

and from (3.37) that

$$\begin{aligned} \lim_{\rho\to\infty} \eta^*_{opt} &= \lim_{\rho\to\infty} \frac{1}{1+(1+\rho^{-1})} \\ &= \frac{1}{2}. \end{aligned} \tag{3.45}$$

Therefore, for constrained bandwidth and power allocations (or equivalently, equal power pilot and data symbols), the loss in effective SNR is 3 dB in the high SNR regime. From comparison of (3.43) and (3.45), this loss is always greater than that without the constraint on the power of pilot and data symbols.

These results are summarized in the following proposition:

Proposition 2 *As the SNR goes to infinity, the asymptotic SNR efficiencies for the information rates (3.19), (3.28), and (3.36) are given by, respectively,*

$$\eta_\infty = \begin{cases} \exp\left[\frac{2f_D}{1-2f_D}\left(\ln\left(2f_D\right)-\gamma\right)\right], & \textit{for } C_{LB} \\ \frac{1}{1+2\sqrt{2f_D(1-2f_D)}}, & \textit{for } R_P \\ \frac{1}{2}, & \textit{for } R_P^* \end{cases} . \tag{3.46}$$

Correspondingly, for sufficiently high SNR, the information rates (3.19), (3.28), and (3.36) can be approximated by

$$\left(1-2f_D\right) C_{Rayleigh}\left(\eta_\infty \rho\right) \quad \textit{nats/symbol.} \tag{3.47}$$

Since in the high SNR regime with perfect CSI, the capacity has a slope of 1 bit/symbol/3 dB, it is clear from (3.47) that all the information rates (3.19), (3.28), and (3.36) have the same slope of $(1-2f_D)$ bits/symbol/3 dB when the SNR is high.

Fig. 3.3 shows the achievable information rates (3.19), (3.28), and (3.36), as well as their asymptotic approximations (3.47) for high SNR. The curves are shown as functions of SNR for normalized fading rate $f_D = 0.01$. The capacity with perfect CSI, $C_{\text{Rayleigh}}(\rho)$ (given by (3.17)), which serves as a trivial upper bound on the capacity without CSI, is also

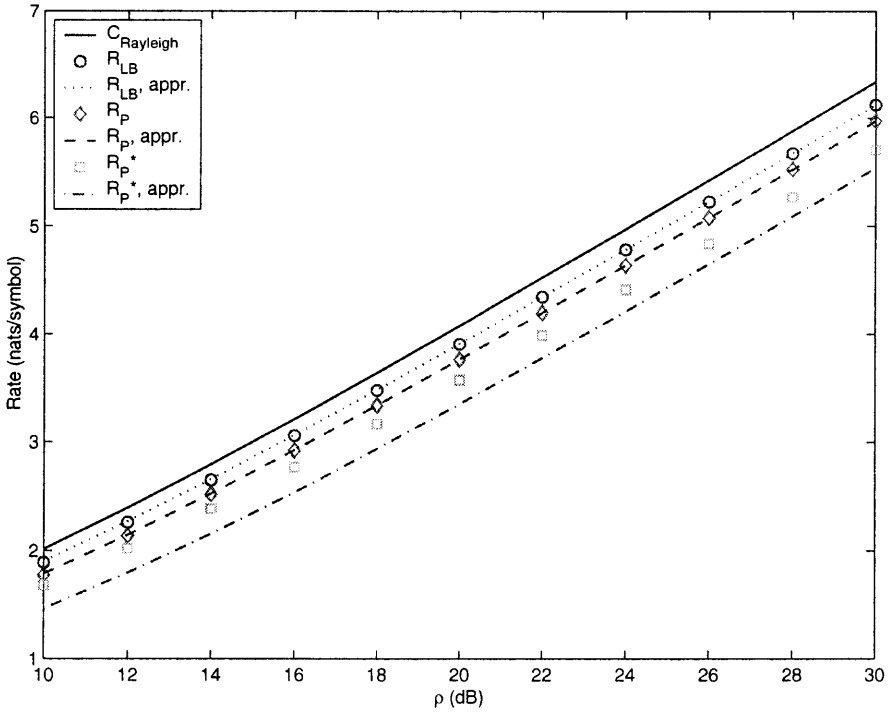

Figure 3.3 Achievable rates versus SNR in the high SNR regime ($f_D = 0.01$).

included. It can be observed that for modestly high SNR, the asymptotic expressions (3.47) provide good approximations of the achievable information rates.

3.4.2 Low SNR

In the following, the following asymptotic formula for $C_{\text{Rayleigh}}(\rho)$ for low SNR [70] will be used

$$C_{\text{Rayleigh}}(\rho) = \rho - \rho^2 + o(\rho^2) \quad \text{nats/symbol}, \ \rho \ll 1. \tag{3.48}$$

Mutual Information Lower Bound Using (3.48) and the Taylor series of the information rate loss $\widetilde{P}_{\Delta,U}$ (3.14) around 0 ($\rho/f_D \ll 1$), and neglecting higher order terms, one can obtain the following approximation for the capacity lower bound (3.19) in the low

SNR regime:

$$\begin{aligned} R_{\text{LB}} &\simeq \rho - \rho^2 - 2f_D \left[\left(\frac{\rho}{2f_D} \right) - \frac{1}{2} \left(\frac{\rho}{2f_D} \right)^2 \right] \\ &= \frac{1 - 4f_D}{4f_D} \rho^2. \end{aligned} \tag{3.49}$$

Pilot-Aided System with Optimal Allocation From (3.32), as the SNR ρ tends to zero, ν tends to infinity. It then follows from Lemma 1 that $g(\nu) \to 1/2$ and from (3.33) that

$$\lim_{\rho \to 0} \beta_{\text{opt}} = \frac{1}{2}. \tag{3.50}$$

This result indicates that in the low SNR regime with Gaussian signaling, half the power should be devoted to pilot symbols.

It can be noted that from (3.35),

$$\begin{aligned} \eta_{\text{opt}} &= \frac{1}{1 - 4f_D} \left(\sqrt{\nu + 1} - \sqrt{\nu} \right)^2 \\ &= \frac{1}{1 - 4f_D} \left(\frac{1}{\sqrt{\nu + 1} + \sqrt{\nu}} \right)^2, \end{aligned} \tag{3.51}$$

which for large ν can be approximated

$$\begin{aligned} \eta_{\text{opt}} &\simeq \frac{1}{(1 - 4f_D)\,\nu} \\ &= \frac{\rho}{8f_D\,(1 - 2f_D)}, \end{aligned} \tag{3.52}$$

where the last expression has been obtained from (3.32). It follows from (3.28) and (3.48) that the achievable information rate with pilot and optimal bandwidth/power allocation is

$$\begin{aligned} R_{\text{P}} &\simeq (1 - 2f_D)\,\eta_{\text{opt}}\rho \\ &= \frac{1}{8f_D} \rho^2 \quad \text{nats/symbol.} \end{aligned} \tag{3.53}$$

Pilot-Aided System with Constrained Allocation Using (3.37) and (3.48) in (3.36) and maximizing with respect to β^*, it can be shown that for low SNR

$$\beta^*_{opt} \simeq g(\delta), \tag{3.54}$$

where the function g is defined in (3.34) and

$$\delta = 2f_D\left(1+\frac{1}{\rho}\right). \tag{3.55}$$

It can be seen from (3.32) and (3.55) that $\delta < \nu$ if $\rho < 1/2$. It follows from the monotonicity of the function g that $\beta^*_{opt} < \beta_{\text{opt}}$. As SNR $\rho \to 0$, it follows from (3.55) that $\delta \to \infty$, and from Lemma 1 that

$$\lim_{\rho\to 0}\beta^*_{\text{opt}} = \frac{1}{2}. \tag{3.56}$$

Hence half of the bandwidth and power should be devoted to pilot symbols when the SNR is low.

For small values of ρ, the information rate when pilot symbols are used can be obtained from (3.37) and (3.48):

$$\begin{aligned} R^*_{\text{P}}\left(\beta^*_{opt}\right) &\simeq \left(1-\beta^*_{opt}\right)\left[1+2f_D\left(\beta^*_{opt}\right)^{-1}\left(1+\rho^{-1}\right)\right]^{-1}\rho \\ &= \left(\sqrt{1+\delta}-\sqrt{\delta}\right)^2\rho, \end{aligned} \tag{3.57}$$

where the last relation was obtained with the help of (3.55). It follows that

$$\begin{aligned} R^*_{\text{P}}\left(\beta^*_{opt}\right) &\simeq \frac{1}{4\delta}\rho \\ &\simeq \frac{1}{8f_D}\rho^2 \quad \text{nats/symbol}, \end{aligned} \tag{3.58}$$

which is the same result as in (3.53). It is concluded that in the low SNR regime, negligible additional performance degradation is caused by constraining the pilot to have equal power and bandwidth allocation.

These results are summarized in the following proposition:

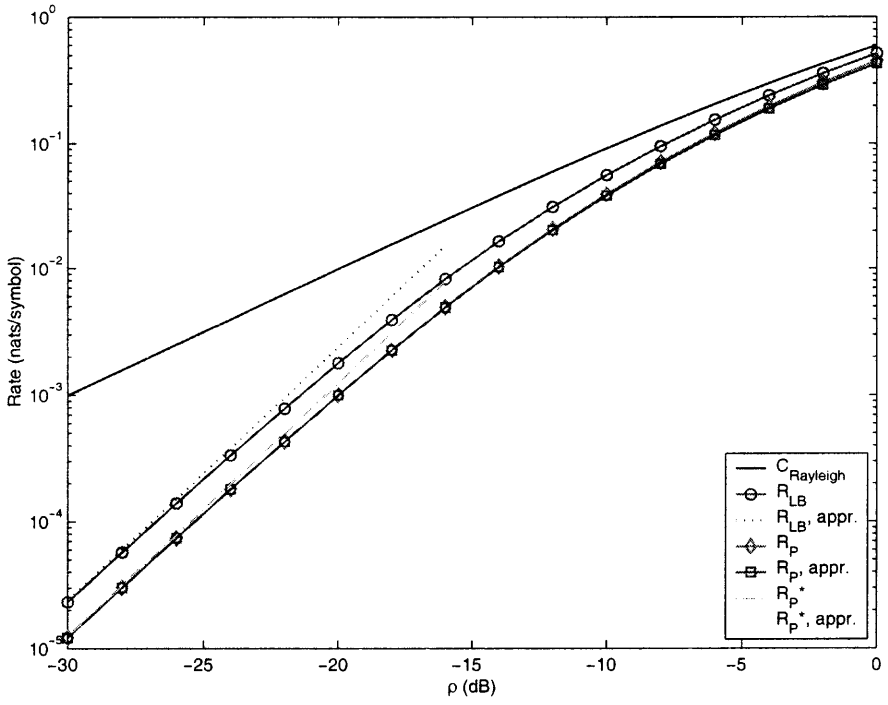

Figure 3.4 Achievable rates versus SNR in the low SNR regime ($f_D = 0.01$).

Proposition 3 *For sufficiently low SNR, the information rates (3.19), (3.28), and (3.36) can be approximated by*

$$\kappa\rho^2 \quad \textit{nats/symbol}, \tag{3.59}$$

with κ given by

$$\kappa = \begin{cases} \frac{1-4f_D}{4f_D}, & \textit{for } C_{LB} \\ \frac{1}{8f_D}, & \textit{for } R_P \textit{ or } R_P^*. \end{cases}$$

This result indicates that it is ineffective to use Gaussian signaling when the SNR is low, since the information rates decrease at a quadratic rate with SNR. This conclusion is not expected to change materially if discrete signaling is used in lieu of Gaussian signaling since in the low SNR regime differences between signaling formats tend to be de-emphasized.

Fig. 3.4 shows the achievable information rates (3.19), (3.28), and (3.36), as well as their asymptotic approximations (3.59) in the low SNR regime. The curves are shown on

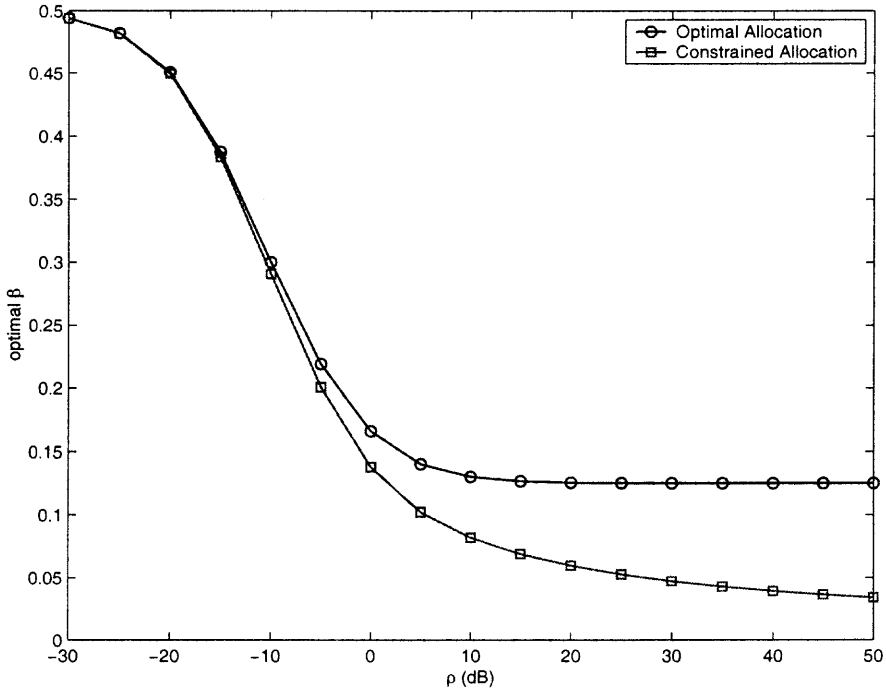

Figure 3.5 Optimal fractional power allocation to pilot versus SNR ($f_D = 0.01$).

log-log axes as functions of SNR for normalized fading rate $f_D = 0.01$. It can be observed that for the pilot-aided system, the loss due to the equal bandwidth-power allocation is negligible for sufficiently low SNR.

The optimal power allocations for pilot symbols are plotted in Fig. 3.5 as a function of SNR for normalized fading rate $f_D = 0.01$. The optimal allocations β_{opt} are calculated from (3.33). For the constrained case, the pilot resource allocations are determined numerically, by optimizing the capacity lower bound (3.36). It can be seen that when the SNR goes to zero, both β_{opt} and β^*_{opt} tend to $1/2$, as per (3.50) and (3.56), respectively. Conversely, when SNR goes to infinity, β_{opt} and β^*_{opt} converge to 0.125 and 0.02, as predicated by (3.42) and (3.44), respectively.

3.5 Low-Duty-Cycle Signaling in the Low SNR Regime

It is well known that the for vanishing SNR, the capacity can be achieved by on-off signaling schemes [13,6]. However the levels of the input symbols and their probabilities are complicated functions of the SNR that can only be determined by intensive numerical computations [13]. Here a simple, but suboptimal approach is investigated that enables closed-form expressions for the threshold level of signaling and the duty cycle.

It have been seen that using i.i.d. Gaussian signaling is efficient in the high SNR regime and inefficient in the low SNR regime. Specifically, on one hand, at high SNR, the ratio of achievable rate to capacity with perfect CSI approaches $(1 - 2f_D)$ as per (3.39); on the other hand, this ratio approaches zero as SNR goes to zero, since in the low SNR regime, the achievable rates with Gaussian signaling decay quadratically with SNR, whereas the capacity with known CSI decays linearly. This motivates us to ask the following questions: Is there a way to use the Gaussian signaling more efficiently? How high need "high SNR" be for Gaussian signaling to be efficient? To find answers to these questions, a signaling scheme with duty cycle $\sigma \leq 1$ is considered in this section. That is, the system transmits blocks of i.i.d. Gaussian signals with average power $\mathcal{E}_s/\sigma$, only a fraction σ of time, so that the average transmit power remains $\mathcal{E}_s$.

3.5.1 Optimal Duty Cycle

It follows from (3.19) that the following information rate is achievable by this signaling scheme:

$$R^{(\mathrm{LDC})} = \sigma \left[C_{\mathrm{Rayleigh}}\left(\frac{\rho}{\sigma}\right) - 2f_D \ln\left(1 + \frac{\rho}{2f_D\sigma}\right)\right] \quad \text{nats/symbol.} \tag{3.60}$$

Now turn to optimize the achievable information rate (3.60) over σ in the interval $0 < \sigma \leq 1$. Note that when the CSI is known, the channel capacity is a decreasing function of the duty cycle σ at *any* SNR due to the loss of signal dimensions. For the

unknown CSI case, the existence of an optimal value for the duty cycle can be motivated by the following intuitive argument: At one extreme, at high SNR, the information rate with Gaussian signaling is dominated by the term $\sigma \ln(\rho/\sigma)$ (see (3.47) and (3.38)). Here, for a decreasing $\sigma < 1$, the increase in SNR through ρ/σ is compressed by the log function and does not make up for the linear loss with σ. Hence, the best solution is $\sigma = 1$. At the other extreme, for low SNR, the information rate is linear with $\sigma(\rho/\sigma)^2$ (see (3.49)). Here it can be seen that if the duty cycle decreases at the same rate as SNR, the fall-off in information rate becomes linear rather quadratic.

It can be shown that $R^{(\mathrm{LDC})}$ is a concave function of σ. The first derivative of $R^{(\mathrm{LDC})}$ with respect to σ is found to be

$$\frac{\partial}{\partial\sigma}R^{(\mathrm{LDC})} = \frac{\rho}{\sigma}\frac{1}{1+\rho/(2f_D\sigma)} - 2f_D\ln\left(1+\frac{\rho}{2f_D\sigma}\right) - \left(1+\frac{\sigma}{\rho}\right)\exp\left(\frac{\sigma}{\rho}\right)\mathrm{Ei}\left(-\frac{\sigma}{\rho}\right) - 1, \tag{3.61}$$

which depends on ρ and σ only through ρ/σ. Therefore the optimal duty cycle, which maximizes the achievable rate (3.60), can be expressed as

$$\sigma_{\mathrm{opt}} = \begin{cases} \frac{\rho}{\rho_{\mathrm{crit}}}, & \rho < \rho_{\mathrm{crit}} \\ 1, & \text{otherwise,} \end{cases} \tag{3.62}$$

where ρ_{crit}, called the *critical SNR*, is the unique solution to the equation

$$\frac{x}{1+x/(2f_D)} - 2f_D\ln\left(1+\frac{x}{2f_D}\right) - \left(1+\frac{1}{x}\right)\exp\left(\frac{1}{x}\right)\mathrm{Ei}\left(-\frac{1}{x}\right) - 1 = 0. \tag{3.63}$$

Interestingly, ρ_{crit} is a function of only the fading rate f_D, and is plotted in Fig. 3.6. From the figure, it can be observed that for a specified SNR ρ, a higher channel Doppler requires a higher ρ_{crit}, i.e., a more peaky signal. A similar relation between signal peakiness and channel coherence time was recently observed for the block-fading channel [73]. This link between peakiness and fading rate adds to the body of literature (reviewed in the Introduction) that focuses on the relation between peakiness and SNR or peakiness and

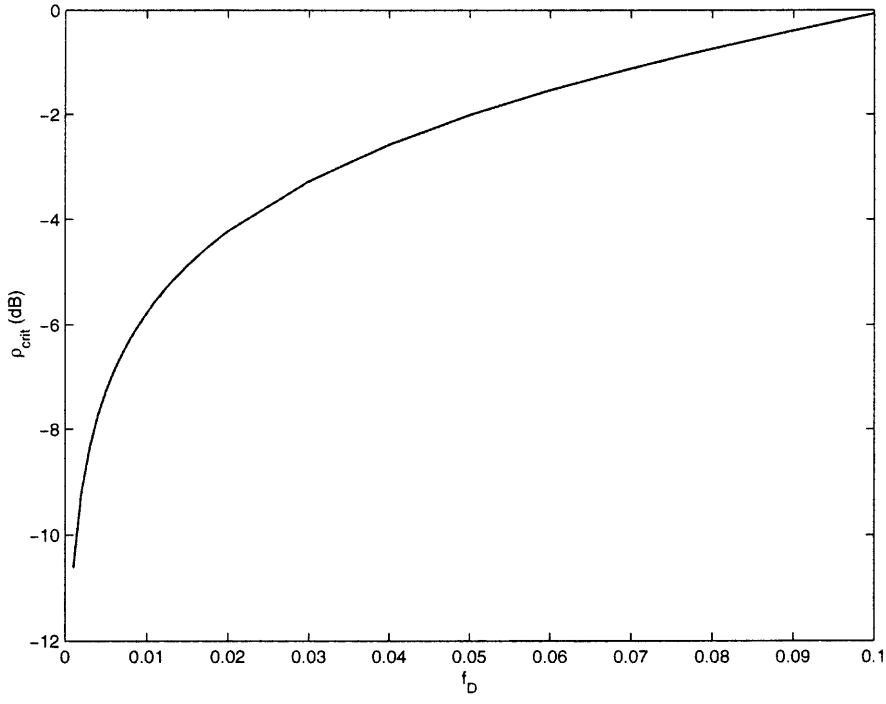

Figure 3.6 Critical SNR versus normalized fading rate.

bandwidth. Here, the relation to SNR can be observed from (3.62), where for a fixed ρ_{crit} (i.e., fixed Doppler spread), the optimal duty cycle σ_{opt} decreases linearly with the SNR ρ.

The critical SNR is the threshold below which a higher information rate than (3.19) can be achieved by signaling with duty cycle less than unity. In this sense, the system can be considered to be operating in the low SNR regime if $\rho < \rho_{\text{crit}}$.

In the low SNR regime, the optimized information rate (3.60) can be written as

$$R_{\text{opt}}^{(\text{LDC})} = \frac{\rho}{\rho_{\text{crit}}} \left[C_{\text{Rayleigh}}(\rho_{\text{crit}}) - 2f_D \ln\left(1 + \frac{\rho_{\text{crit}}}{2f_D}\right) \right], \quad \rho < \rho_{\text{crit}}, \tag{3.64}$$

which is a linear function of SNR, as opposed to the quadratic function of SNR for the information rates given in (3.59). Using the approximation for low SNR (3.48), one can obtain the following proposition.

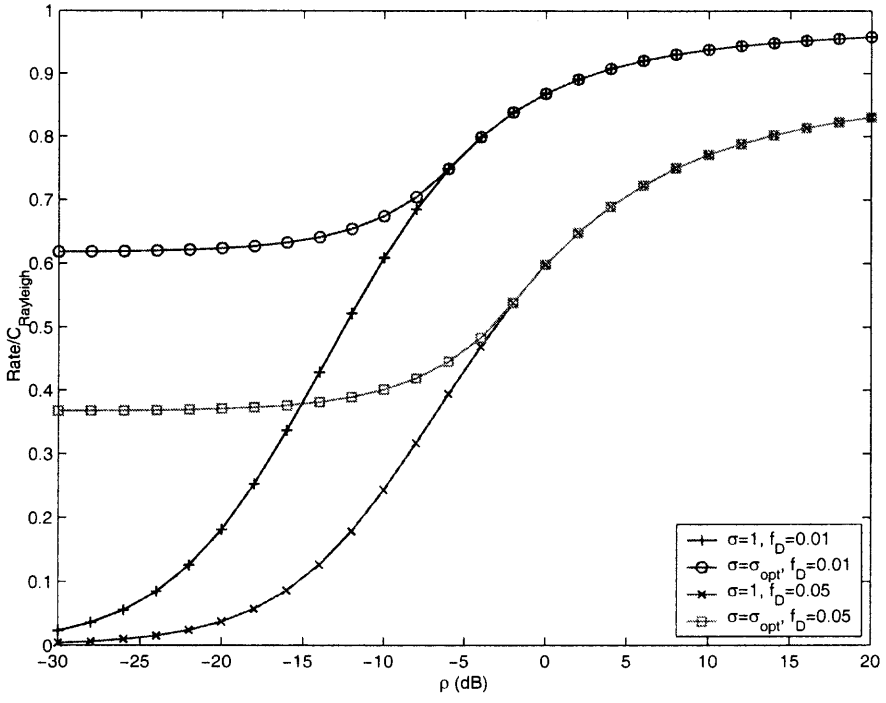

Figure 3.7 Ratio of achievable rate to the capacity with perfect CSI versus SNR; σ_{opt} is given by (3.62).

Proposition 4 *In the low SNR regime,*

$$\lim_{\rho \to 0} \frac{R_{opt}^{(LDC)}}{C_{Rayleigh}} = \frac{1}{\rho_{crit}} \left[C_{Rayleigh}\left(\rho_{crit}\right) - 2f_D \ln\left(1 + \frac{\rho_{crit}}{2f_D}\right) \right]. \tag{3.65}$$

Fig. 3.7 shows the ratios of the information rate (3.60) with optimized duty cycle and the capacity lower bound (3.19) to the capacity with perfect CSI. The curves are shown as functions of SNR for normalized fading rates $f_D = 0.01$ and 0.05. It can be seen that this ratio approaches $(1 - 2f_D)$ as SNR goes to infinity, and approaches the value given by (3.65) as SNR goes to zero.

3.5.2 An Example for BPSK/QPSK Signals

Finally it is noted that the above system parameters designed for Gaussian signals can also be used to other signals. As an example, the resource allocations (3.29) and (3.33),

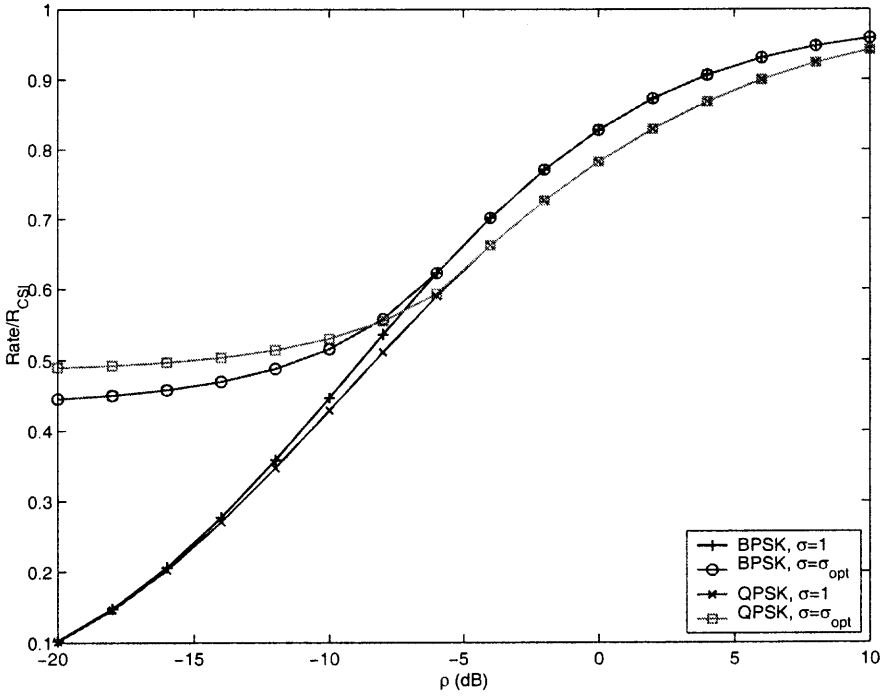

Figure 3.8 Ratio of achievable rate of a pilot-aided system to the mutual information with perfect CSI versus SNR for BPSK and QPSK signals ($f_D = 0.01$); resource allocations of pilot symbols are given by (3.29) and (3.33); σ_{opt} is given by (3.62).

and the duty cycle (3.62) are applied to a pilot-aided system using binary PSK (BPSK) and quaternary PSK (QPSK) signals. For BPSK and QPSK signals, the mutual information with perfect CSI can be calculated numerically by using the techniques in [74]. The ratios of achievable rates to those with perfect CSI are plotted in Fig. 3.8. Although these parameters are generally not optimal for BPSK or QPSK signals, the curves shown in Fig. 3.8 have the same characteristic shapes as those for Gaussian signals. It is clear from the figure that signaling at duty cycle (3.33) also improves the achievable rates for low SNR.

3.6 Chapter Summary

This chapter investigates the achievable information rates for a wireless communication system when neither the transmitter nor the receiver has *a priori* knowledge of the channel state information (CSI). The dynamics of the continuously fading, flat Rayleigh channel

are characterized by a Doppler spectrum with fading rate known to the receiver. The information rate penalty due to unknown CSI, is expressed in terms of the fading rate, the transmission block length and the signal-to-noise ratio (SNR). The asymptotic (large block length) information rate penalty is evaluated for both Clarke's Doppler spectrum and the ideal lowpass spectrum. The performance of a system using pilot-aided channel estimation is studied. The allocation of bandwidth and power to the pilot symbols is determined by optimizing the achievable information rate. Asymptotic approximate expressions of the achievable rates suggest that using independent and identically distributed (i.i.d.) Gaussian signals is efficient when the SNR is high and inefficient when the SNR is low. It is demonstrated that the it is useful to use a simple, low-duty-cycle signaling scheme in the low SNR regime. The optimal duty cycle is determined as a function of the fading rate. It is shown that the peakiness of the signaling depends on the fading rate.

CHAPTER 4

DECISION-DIRECTED ITERATIVE CHANNEL ESTIMATION FOR MIMO SYSTEMS

Due to the high computation complexity of matrix inversion and the fact that the matrix inversion has to be calculated in each iteration, optimal decision directed channel estimation for MIMO channels is difficult to implement in practice. In this chapter, an iterative method is proposed for decision directed channel estimation in MIMO systems. The resulting estimator can be interpreted as applying a data-independent filter to the weighted sum of the minimum norm least square (MNLS) estimate and the estimate obtained from the previous iteration. When the weight parameter is chosen appropriately, for single transmit antenna systems the proposed iterative method reduces to the optimal estimator.

This chapter is organized as follows. The system and channel models are described in Section 4.1. Section 4.2 contains the development of the optimal MAP estimate and an iterative algorithm for its evaluation. Simulation results are provided in Section 4.3 to illustrate the performance of the proposed method. Conclusions are given in Section 4.4.

4.1 System Model

The MIMO system of interest is equipped with N transmit antennas and M receive antennas. The channel between each transmit and receive antenna pair is modeled as flat fading process. The discrete time received signal at the mth receive antenna at time t, $r_{m,t}$, is given by

$$r_{m,t} = \mathbf{s}_t^T \mathbf{h}_{m,t} + n_{m,t} \tag{4.1}$$

where $\mathbf{h}_{m,t} = [h_{1m,t}\ h_{2m,t}\ \cdots\ h_{Nm,t}]^T$ is the $N \times 1$ vector whose nth component represents the complex fading gain between the nth transmit antenna and the mth receive antenna at time t, $\mathbf{s}_t = [s_{1,t}\ s_{2,t}\ \cdots\ s_{N,t}]^T$ is the $N \times 1$ vector of symbols simultaneously transmitted

by the N transmit antennas at time t, $n_{m,t}$ is the complex noise sample with variance $N_0/2$ per dimension.

Consider the transmission of a block (a *frame*) of L symbol vectors, $\mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_L$. The received sequence at the mth receive antenna, denoted by an $L \times 1$ vector $\mathbf{r}_m = [r_{m,1}\ r_{m,2}\ \cdots\ r_{m,L}]^T$, is of the form

$$\mathbf{r}_m = \mathbf{S}\mathbf{h}_m + \mathbf{n}_m \tag{4.2}$$

where $\mathbf{n}_m$ is an $L \times 1$ (possibly colored) complex Gaussian noise vector with zero mean and covariance matrix $\mathbf{R_n} = E[\mathbf{n}_m \mathbf{n}_m^H]$. For white noise, the noise covariance matrix is $\mathbf{R_n} = N_0\mathbf{I}$. The $NL \times 1$ vector $\mathbf{h}_m = [\mathbf{h}_{m,1}^T\ \mathbf{h}_{m,2}^T\ \cdots\ \mathbf{h}_{m,L}^T]^T$ is the vector of channel parameters. The $L \times NL$ data matrix $\mathbf{S}$ is defined as

$$\mathbf{S} = \begin{bmatrix} \mathbf{s}_1^T & \mathbf{0}^T & \cdots & \mathbf{0}^T \\ \mathbf{0}^T & \mathbf{s}_2^T & \cdots & \mathbf{0}^T \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}^T & \mathbf{0}^T & \cdots & \mathbf{s}_L^T \end{bmatrix}. \tag{4.3}$$

For Rayleigh fading the channel coefficients $h_{nm,t}$ can be modeled as zero mean complex Gaussian random variables. If the antennas are spaced sufficiently far apart, the fading is assumed to be uncorrelated across antennas. For two-dimensional isotropic scattering (Clarke's model), the correlation function can be expressed as

$$\mathbb{E}[h_{n_1 m_1, t_1} h^*_{n_2 m_2, t_2}] = \delta_{n_1 n_2} \delta_{m_1 m_2} J_0(2\pi f_D (t_1 - t_2)) \tag{4.4}$$

where $\delta_{n_1 n_2} = 1$ if $n_1 = n_2$ and $\delta_{n_1 n_2} = 0$ otherwise, J_0 is the zeroth order modified Bessel function of the first kind, f_D is the normalized fading rate.

Since the fading is assumed to be uncorrelated across antennas, the observation $\mathbf{r}_m$ provides no information for the channel parameter $\mathbf{h}_{m'}$ if $m \neq m'$. In the sequel the receive

antenna index m is suppressed so that (4.2) becomes

$$\mathbf{r} = \mathbf{S}\mathbf{h} + \mathbf{n}. \tag{4.5}$$

Given the decision on S (from detector or decoder), it is of interest to estimate the channel parameter $\mathbf{h}$ based on the observation $\mathbf{r}$.

Conditioned on $\mathbf{S}$ and $\mathbf{h}$, the received signal vector is distributed as $\mathbf{r} \sim \mathcal{CN}(\mathbf{S}\mathbf{h}, \mathbf{R}_\mathbf{n})$. Therefore the conditional probability density function (pdf) for $\mathbf{r}$ is

$$p(\mathbf{r}|\mathbf{h}, \mathbf{S}) = \pi^{-L}|\mathbf{R}_\mathbf{n}|^{-1} \exp(-(\mathbf{r} - \mathbf{S}\mathbf{h})^H \mathbf{R}_\mathbf{n}^{-1}(\mathbf{r} - \mathbf{S}\mathbf{h})). \tag{4.6}$$

For Rayleigh fading channel, the *a priori* pdf for $\mathbf{h}$ is

$$p(\mathbf{h}) = \pi^{-NL}|\mathbf{R}_\mathbf{h}|^{-1} \exp(-\mathbf{h}^H \mathbf{R}_\mathbf{h}^{-1}\mathbf{h}) \tag{4.7}$$

where the covariance matrix $\mathbf{R}_\mathbf{h}$ is determined by the normalized fading rate via (4.4).

4.2 Channel Estimation

4.2.1 Optimal Estimate

The optimal minimum mean square error (MMSE) estimate of $\mathbf{h}$ based on the received signal vector $\mathbf{r}$ and transmitted data matrix $\mathbf{S}$, is the conditional mean, $E[\mathbf{h}|\mathbf{r}, \mathbf{S}]$. Since $\mathbf{r}$ and $\mathbf{h}$ are jointly Gaussian, the conditional mean is equivalent to the maximum *a posteriori* (MAP) estimate of $\mathbf{h}$, $\hat{\mathbf{h}}_{MAP}(\mathbf{S})$, which maximizes the *a posteriori* pdf $p(\mathbf{h}|\mathbf{r}, \mathbf{S})$ with respect to $\mathbf{h}$. A necessary condition for this maximization is

$$\frac{\partial}{\partial \mathbf{h}} \ln p(\mathbf{h}|\mathbf{r}, \mathbf{S})|_{\mathbf{h}=\hat{\mathbf{h}}_{MAP}(\mathbf{S})} = \mathbf{0}. \tag{4.8}$$

The *a posteriori* pdf for $\mathbf{h}$ is

$$p(\mathbf{h}|\mathbf{r}, \mathbf{S}) = \frac{p(\mathbf{r}|\mathbf{h}, \mathbf{S})p(\mathbf{h}|\mathbf{S})}{p(\mathbf{r}|\mathbf{S})} \tag{4.9}$$

where $p(\mathbf{r}|\mathbf{h},\mathbf{S})$ and $p(\mathbf{h}|\mathbf{S}) = p(\mathbf{h})$ are given in (4.6) and (4.7), respectively. The pdf $p(\mathbf{r}|\mathbf{S})$ does not affect the maximization over $\mathbf{h}$. From (4.6), (4.7) and (4.9),

$$\frac{\partial}{\partial \mathbf{h}} \ln p(\mathbf{h}|\mathbf{r},\mathbf{S}) = \underbrace{(\mathbf{S}^H\mathbf{R}_\mathbf{n}^{-1}\mathbf{S} + \mathbf{R}_\mathbf{h}^{-1})}_{\mathbf{A}(\mathbf{S})}\mathbf{h} - \underbrace{\mathbf{S}^H\mathbf{R}_\mathbf{n}^{-1}\mathbf{r}}_{\mathbf{b}}. \tag{4.10}$$

The MAP estimate of $\mathbf{h}$ is found by setting the derivative in (4.10) equal to zero. Thus,

$$\hat{\mathbf{h}}_{MAP}(\mathbf{S}) = (\mathbf{S}^H\mathbf{R}_\mathbf{n}^{-1}\mathbf{S} + \mathbf{R}_\mathbf{h}^{-1})^{-1}\mathbf{S}^H\mathbf{R}_\mathbf{n}^{-1}\mathbf{r}. \tag{4.11}$$

The computation of (4.11) involves the inversion of an $NL \times NL$ matrix $\mathbf{A}(\mathbf{S})$. By application of the matrix inversion lemma, it can be shown that the cannel estimate can be expressed as

$$\hat{\mathbf{h}}_{MAP}(\mathbf{S}) = \mathbf{R}_\mathbf{h}\mathbf{S}^H(\mathbf{S}\mathbf{R}_\mathbf{h}\mathbf{S}^H + \mathbf{R}_\mathbf{n})^{-1}\mathbf{r} \tag{4.12}$$

where the dimension of the matrix to be inverted is $L \times L$. Expressions (4.11) and (4.12) have been employed as the optimal channel estimator in iterative channel estimation and decoding of space time coded systems [30, 32, 34]. Since computation of (4.12) requires the inversion of an $L \times L$ matrix of complexity $O(L^3)$, the computational cost is very high for practical frame sizes L. Furthermore, the matrix to be inverted is in general dependent on transmitted data $\mathbf{S}$, which is highly undesirable for practical implementation.

For systems with single transmit antenna ($N = 1$), in the case of mutiple phase shift keying (MPSK) modulation ($s_{n,t}^* s_{n,t} = 1$, $n = 1, 2, \cdots, N$, $t = 1, 2, \cdots, L$) and white noise ($\mathbf{R}_\mathbf{n} = N_0\mathbf{I}$), $\mathbf{A}(\mathbf{S})$ in (4.11) is given by $\mathbf{A}(\mathbf{S}) = \frac{1}{N_0}\mathbf{I} + \mathbf{R}_\mathbf{h}^{-1}$, which is independent

on S. The estimate can be simplified to

$$\begin{aligned} \hat{\mathbf{h}}_{MAP}(\mathbf{S}) &= (\tfrac{1}{N_0}\mathbf{I} + \mathbf{R}_{\mathbf{h}}^{-1})^{-1}\mathbf{S}^H \tfrac{1}{N_0}\mathbf{r} \\ &= \underbrace{\mathbf{R}_{\mathbf{h}}(\mathbf{R}_{\mathbf{h}} + N_0\mathbf{I})^{-1}}_{\mathbf{F_W}} \underbrace{\mathbf{S}^H\mathbf{r}}_{\hat{\mathbf{h}}_{ML}(\mathbf{S})} \end{aligned} \quad . \tag{4.13}$$

Equation (4.13) can be interpreted as a Wiener filter $\mathbf{F_W}$ applied to the maximum likelihood (ML) estimate $\hat{\mathbf{h}}_{ML}(\mathbf{S})$ [4]. It has been employed as the optimal channel estimator in iterative channel estimation and decoding for systems with single transmit antenna [35, 36].

4.2.2 Iterative Solution

Instead of carrying out the direct matrix inversion in (4.11), one can use iterative methods to find an approximate solution of the MAP equation

$$\mathbf{A}(\mathbf{S})\mathbf{h} = \mathbf{b}. \tag{4.14}$$

Let $\mathbf{A} = \mathbf{Q} - \mathbf{P}$ be any decomposition of $\mathbf{A}$ such that $\mathbf{Q}$ is nonsingular. Let $\hat{\mathbf{h}}^{(0)}$ be an arbitrary initial vector, then the vector sequence $\hat{\mathbf{h}}^{(0)}, \hat{\mathbf{h}}^{(1)}, \hat{\mathbf{h}}^{(2)}, \cdots$ generated by the following iteration

$$\mathbf{Q}\hat{\mathbf{h}}^{(k+1)} = \mathbf{P}\hat{\mathbf{h}}^{(k)} + \mathbf{b} \tag{4.15}$$

converges to the true solution if and only if the spectral radius of the iteration matrix $\mathbf{B} = \mathbf{Q}^{-1}\mathbf{P}$ satisfies $\rho(\mathbf{B}) < 1$, where the spectral radius $\rho(\mathbf{B})$ is defined as the modulus of the largest eigenvalue of $\mathbf{B}$.

Different decompositions of $\mathbf{A}$ result in iterative methods given by different names, such as Jacobi, Gauss-Seidel, and successive overrelaxation (SOR) iteration. Details of convergence and computation aspects of these methods can be found in mathematical

literature [75] [76]. In order to avoid calculating the inverse of a data dependent matrix in decision directed channel estimation, one can decompose $\mathbf{A}(\mathbf{S})$ into data independent part $\mathbf{Q}$ and data dependent part $-\mathbf{P}(\mathbf{S})$, i.e., $\mathbf{A}(\mathbf{S}) = \mathbf{Q} - \mathbf{P}(\mathbf{S})$. One straightforward way to do so is

$$\mathbf{Q} = C\mathbf{I} + \mathbf{R}_{\mathbf{h}}^{-1}, \tag{4.16}$$

$$\mathbf{P}(\mathbf{S}) = C\mathbf{I} - \mathbf{S}^H\mathbf{R}_{\mathbf{n}}^{-1}\mathbf{S}, \tag{4.17}$$

where C is some constant.

With $\mathbf{Q}$ and $\mathbf{P}$ difined in (4.16) and (4.17), the iteration becomes

$$\begin{aligned}\hat{\mathbf{h}}^{(k+1)} &= \underbrace{(C\mathbf{I} + \mathbf{R}_{\mathbf{h}}^{-1})^{-1}}_{\mathbf{Q}^{-1}} \\ &\quad (\underbrace{(C\mathbf{I} - \mathbf{S}^H\mathbf{R}_{\mathbf{n}}^{-1}\mathbf{S})}_{\mathbf{P}(\mathbf{S})}\hat{\mathbf{h}}^{(k)} + \mathbf{S}^H\mathbf{R}_{\mathbf{n}}^{-1}\mathbf{r}).\end{aligned} \tag{4.18}$$

A sufficient condition for the convergence of (4.18) is given in the following proposition:

Theorem 1 *If the noise is white, the iteration (4.18) converge to the optimal estimate if*

$$C \geq \frac{N}{2N_0}. \tag{4.19}$$

Proof: Since $\mathbf{R}_{\mathbf{h}}$ is positive definite, all the eigenvalues of $\mathbf{Q}$ are greater than C, hence $\rho(\mathbf{Q}^{-1}) < \frac{1}{C}$. Let λ be any eigenvalue of $\mathbf{P}$, then $C - \frac{N}{N_0} \leq \lambda \leq C$. If $C \geq \frac{N}{2N_0}$,

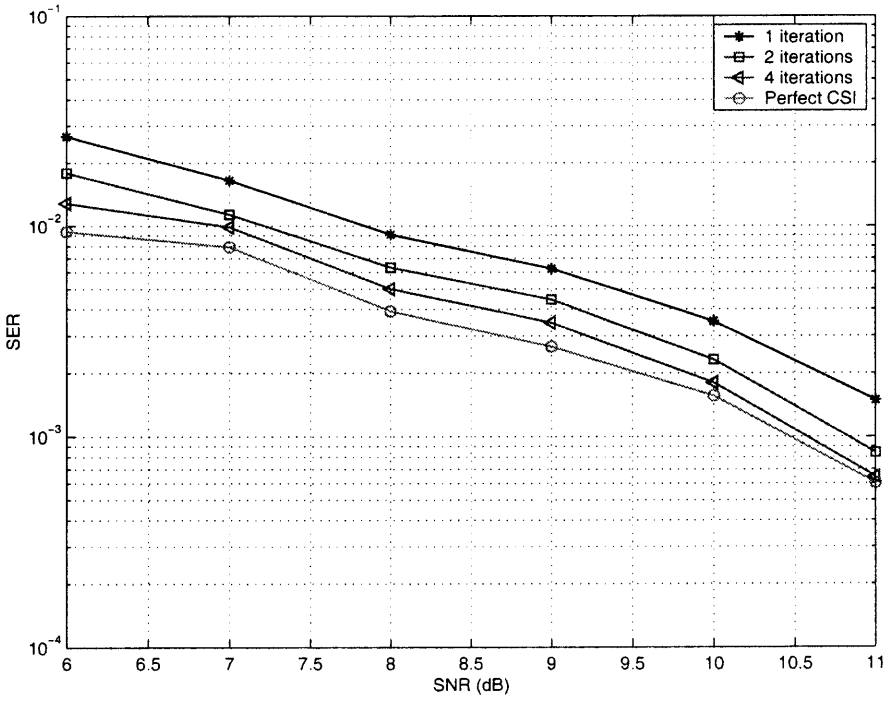

Figure 4.1 Symbol error probability versus SNR ($f_D = 0.001$)

then $-C \leq C - \frac{N}{N_0} \leq \lambda \leq C$ so that $\rho(\mathbf{P}) \leq C$. Therefore

$$\begin{aligned} \rho(\mathbf{Q}^{-1}\mathbf{P}) &\leq \|\mathbf{Q}^{-1}\mathbf{P}\|_2 && (4.20)\\ &\leq \|\mathbf{Q}^{-1}\|_2\|\mathbf{P}\|_2 && (4.21)\\ &= \rho(\mathbf{Q}^{-1})\rho(\mathbf{P}) && (4.22)\\ &< 1, && (4.23) \end{aligned}$$

where $\|\cdot\|_2$ denotes spectral norm, inequalities (4.20), (4.21) and (4.23) are immediately clear, equality (4.22) holds because $\mathbf{Q}^{-1}$ is symmetric and $\mathbf{P}$ is Hermitian. ■

From the signal model (4.5), one can formulate an underdetermined least-squares problem for the channel estimate, which affords the following minimum norm solution

$$\hat{\mathbf{h}}_{LS}(\mathbf{S}) = \mathbf{S}^H(\mathbf{S}\mathbf{S}^H)^{-1}\mathbf{r}. \quad (4.24)$$

Similar to (4.13), (4.18) can be written as

$$\begin{aligned}
\hat{\mathbf{h}}^{(k+1)} &= \mathbf{R_h}(\mathbf{R_h} + \tfrac{1}{C}\mathbf{I})^{-1} \\
&\quad ((\mathbf{I} - \tfrac{1}{C}\mathbf{S}^H\mathbf{R_n}^{-1}\mathbf{S})\hat{\mathbf{h}}^{(k)} + \tfrac{1}{C}\mathbf{S}^H\mathbf{R_n}^{-1}\mathbf{r}) \\
&= \underbrace{\mathbf{R_h}(\mathbf{R_h} + \frac{1}{C}\mathbf{I})^{-1}}_{\mathbf{F}}((\mathbf{I} - \tfrac{1}{C}\mathbf{S}^H\mathbf{R_n}^{-1}\mathbf{S})\hat{\mathbf{h}}^{(k)} \\
&\quad + \underbrace{\frac{1}{C}\mathbf{S}^H\mathbf{R_n}^{-1}\mathbf{S}}_{\mathbf{W}_{LS}(\mathbf{S})}\underbrace{\mathbf{S}^H(\mathbf{S}\mathbf{S}^H)^{-1}\mathbf{r}}_{\hat{\mathbf{h}}_{LS}(\mathbf{S})}) \\
&= \mathbf{F}((\mathbf{I} - \mathbf{W}_{LS}(\mathbf{S}))\hat{\mathbf{h}}^{(k)} + \mathbf{W}_{LS}(\mathbf{S})\hat{\mathbf{h}}_{LS}(\mathbf{S})).
\end{aligned} \tag{4.25}$$

It can be seen from (4.25) that the next estimate, $\hat{\mathbf{h}}^{(k+1)}$, is obtained by applying a filter, $\mathbf{F}$, to the weighted sum of the current estimate, $\hat{\mathbf{h}}^{(k)}$, and the minimum norm least square estimate, $\hat{\mathbf{h}}_{LS}(\mathbf{S})$.

For systems with single transmit antenna, MPSK modulation and white noise, $\mathbf{W}_{LS}(\mathbf{S}) = \frac{1}{C}\frac{1}{N_0}\mathbf{I}$. In particular, if $C = \frac{1}{N_0}$, then $\mathbf{W}_{LS}(\mathbf{S}) = \mathbf{I}$ and $\mathbf{I} - \mathbf{W}_{LS}(\mathbf{S}) = \mathbf{O}$, so that (4.25) reduces to (4.13).

4.3 Numerical Results

The performance of the proposed iterative decision channel estimation is demonstrated by a space time code system with joint channel estimation and data decoding. The system structure is described in [30]. Numerical results were obtained for the 4PSK 8 state space time code presented in [77] with 2 transmit and 2 receive antennas. Each frame consists of 14 pilot symbols and 116 data symbols. The pilot symbols are used to obtain the initial estimate of channel. Soft decision feedback is used to improve the channel after each iteration. The soft decision of each symbol is obtained by averaging with respect to its *a posteriori* probability from the BCJR decoding algorithm [78].

The symbol error probability (SEP) performance versus SNR per symbol (E_S/N_0) is evaluated for normalized fading rates $f_D = 0.001$ and $f_D = 0.01$ in Fig. 4.1 and Fig. 4.2, respectively. Performance of the same code with perfectly known channel state information

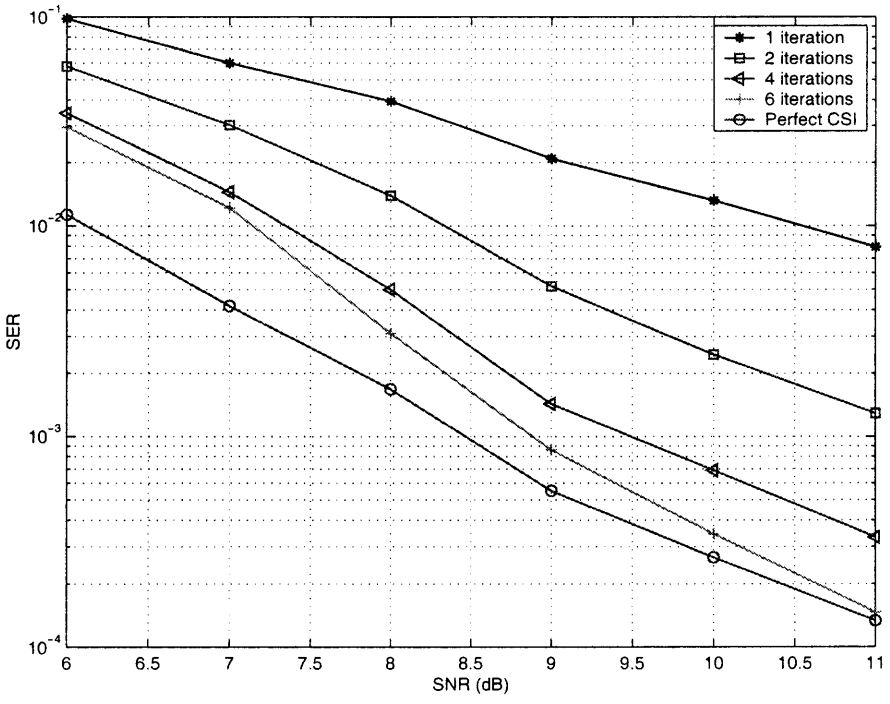

Figure 4.2 Symbol error probability versus SNR ($f_D = 0.01$).

at the receiver is also included for comparison. Significant performance gain from the decision feedback from the decoding can be observed. In both cases, the performance with the decision directed channel estimate at high SNR is within 0.5 dB of that with known CSI.

Fig. 4.3 shows the normalized mean squared error (MSE) versus the number of iterations. It can be seen that the performance improves with the increase in the number of iterations, but the improvement is negligible after 4 and 6 iterations for $f_D = 0.001$ and $f_D = 0.01$, respectively.

4.4 Chapter Summary

A simple iterative method for decision directed channel estimation for MIMO systems is introduced in this chapter. Its application to the joint channel estimation and data decoding for space time coded system is illustrated. Simulation results suggest that near optimal

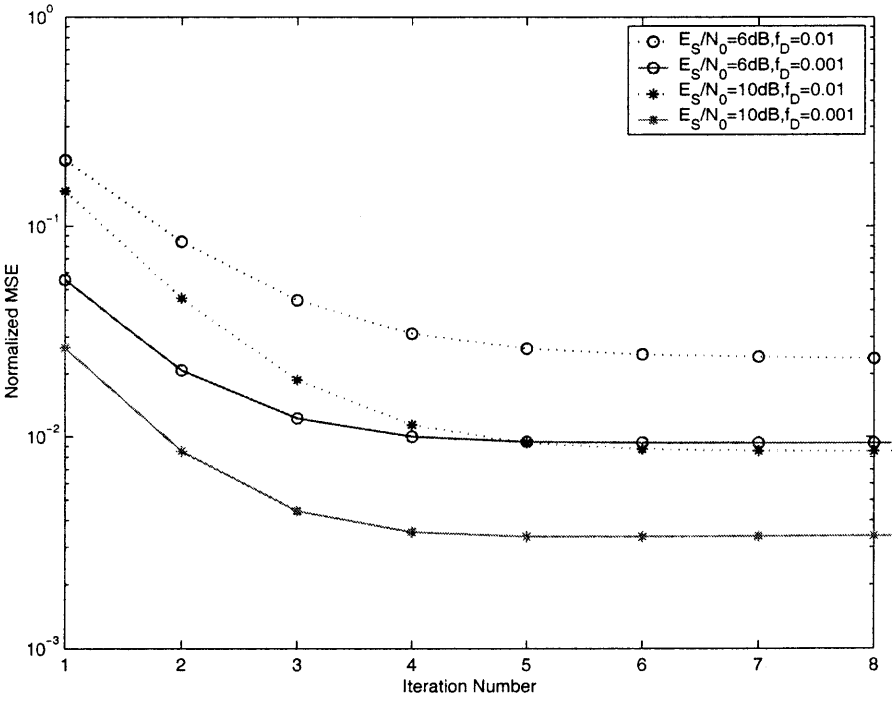

Figure 4.3 Normalized mean squared error versus iteration number.

performance can be achieved with reasonable number of iterations.

CHAPTER 5

COOPERATIVE RELAYING OVER FADING CHANNELS

In this chapter, the average and worst-case performance of cooperative relaying in fading channels will be investigated in terms of the average signal-to-noise ratio (SNR) and outage probability. Using the *long-term* CSI of the channels, the power allocation between the source and relay will be determined. The SNR optimization can be used to minimize transmitted power when the system experiences the ergodic fluctuations of the channels over a sufficiently long time. The outage optimization is useful when the delay and complexity constraints of the system prevent the inherent time diversity of the channels from being exploited. When only *local* knowledge of CSI is available, suboptimal combining techniques can be used at the destination [54].

This chapter is organized as follows. Section 5.1 describes the system and channel model. Sections 5.2 presents the performance for the optimal receivers when the destination has global knowledge of CSI. Power allocation is developed in Section 5.3, under the assumption that the long-term CSI of the channels is available to the source and relay. Section 5.4 considers suboptimal receivers with only local knowledge of CSI. Conclusions are given in Section 5.5.

5.1 System Model

Consider a three-node wireless network illustrated in Fig. 5.1. Information is to be transmitted from the source to the destination with the assistance of the relay. The transmissions are subject to flat fading and additive noise.

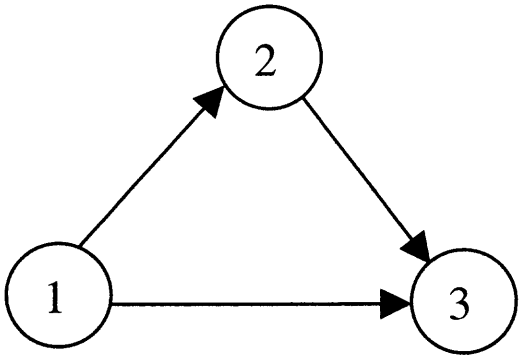

Figure 5.1 Illustration of a three-node network.

5.1.1 Channel Model

Assume that the channels are block fading, i.e., the channel coefficients remain constant for a block of channel uses. In the sequel, g_{ij} is the complex gain of the channel from node i to node j, and $n_{j,k}$ is the additive noise at node j in the k th block, where $i \in \{1,2\}$, $j \in \{2,3\}$, and $k \in \{1,2\}$. The mean strength of the channel from node i to node j is G_{ij}. The noise power at the receiver of each node is N_0. For Rayleigh fading and additive white Gaussian noise (AWGN), the channel gains and noise samples are distributed as $g_{ij} \sim \mathcal{CN}(0, G_{ij})$ and $n_{j,k} \sim \mathcal{CN}(0, N_0)$. Assume that g_{ij} and $n_{j,k}$ are statistically independent for different values of i, j, and k, and that the channel coefficients and the noise samples are independent.

5.1.2 Signal Model

In the first block, the source transmits to the relay and the destination. Let x denote the baseband transmitted signal at the source node, then the received signals at the relay node

and the destination node are respectively given by

$$\begin{aligned} y_{r,1} &= g_{12}x + n_{r,1}, \\ y_{d,1} &= g_{13}x + n_{d,1}. \end{aligned} \tag{5.1}$$

In the second block, for amplify-and-forward relaying, the relay simply amplifies the received signal from the first block by a factor A, and forwards it to the destination, so that the destination nodes receives

$$\begin{aligned} y_{d,2} &= g_{23}Ay_{r,1} + n_{d,2} \\ &= g_{23}Ag_{12}x + g_{23}An_{r,1} + n_{d,2}. \end{aligned} \tag{5.2}$$

The signals from the direct path (5.1) and the relay path (5.2) are combined at the destination, yielding

$$y_c = w_1 y_{d,1} + w_2 y_{d,2} \tag{5.3}$$

where w_1 and w_2 are appropriate weighting coefficients for the direct and relay paths, respectively. The combined signal (5.3) is then used for detection or decoding.

5.1.3 Relaying Strategies

For channel inversion and constant amplification, the amplification factors A are chosen, respectively, subject to the *short-term* (per block) and *long-term* (average) power constraint imposed at the relay node. Correspondingly, knowledge of the *instantaneous* or *average* channel gains of the source-relay channel is required at the respective relay node.

Assume that the average power of the signal transmitted at the source is $\mathbb{E}\left[|x|^2\right] = E_s$, and that the average transmitted power at the relay node is E_r. [1] If $\alpha = E_r/E_s$ and

[1]Channel-inversion relaying has been studied extensively in [44]. Some variants of this problem has been treated in the literature without the power constraint at the relay. For example, $|A| = 1$ is considered in [61]. In [55,56,57,58,59], the noise at the relay is neglected to facilitate performance analysis, and correspondingly $|A|$ is simply set to $E_r/(E_s|g_{12}|)$ (or $|g_{12}|^{-1}$). Another relaying

$\rho = E_s/N_0$, the latter being the average transmit SNR of the source, then the magnitude of the amplification factor can be expressed as

$$|A| = \begin{cases} \sqrt{\frac{\alpha\rho}{\rho|g_{12}|^2+1}}, & \text{Channel Inversion} \\ \sqrt{\frac{\alpha\rho}{\rho G_{12}+1}}, & \text{Constant Amplification.} \end{cases} \tag{5.4}$$

For performance analysis purposes, it can be assumed $E_r = E_s$ ($\alpha = 1$) without loss of generality, and the discrepancies in power can be lumped into the path strengths, as in [44]. Here, α is viewed as a power allocation parameter that can be used to optimize performance under the constraint that the total transmitted power is fixed.

For given values of G_{ij} and α, the instantaneous SNR of the combined signal observed at the receiver is a random variable parameterized by the average SNR ρ. In the sequel, the system performance will be measured by average SNR and outage probability.

5.2 Optimal Receiver with Global CSI

When knowledge of *all* the channel coefficients is available at the destination, the optimal (in the sense that the instantaneous SNR is maximized) combining method is the maximal-ratio combining (MRC), where the weighing coefficients are given by

$$\begin{aligned} w_1 &= C g_{13}^* \\ w_2 &= C \left(|A|^2 |g_{23}|^2 + 1\right)^{-1} A^* g_{12}^* g_{23}^* \end{aligned}$$

with C being some arbitrary constant. The SNR of the combined signal equals the sum of the SNRs in individual paths, i.e.,

$$S_c = S_1 + S_2$$

strategy, constant power (CP), is considered Chapter 6. CP relaying does not require any channel information, but is difficult to analyze.

where S_1 and S_2 are the SNR of the direct and relay path, respectively. From (5.1) and (5.2),

$$S_1 = |g_{13}|^2 \rho \tag{5.5}$$

$$S_2 = \frac{|A|^2 |g_{12}|^2 |g_{23}|^2}{|A|^2 |g_{23}|^2 + 1} \rho. \tag{5.6}$$

5.2.1 Average SNR

It is apparent that the average SNR for the optimal MRC receiver is

$$\mathbb{E}[S_c] = \mathbb{E}[S_1] + \mathbb{E}[S_2]. \tag{5.7}$$

It can be shown that constant amplification relaying performs better than channel inversion relaying, when average SNR is the performance measure.

Proposition 5 *Let S_2^{CI} and S_2^{CA} be the SNR of the relay path for CI and CA relaying, respectively, then*

$$\mathbb{E}\left[S_2^{CI}\right] \leq \mathbb{E}\left[S_2^{CA}\right] \leq \frac{\alpha G_{12} G_{23} \rho^2}{1 + G_{12}\rho + \alpha G_{23}\rho},$$

where the first equality holds if and only if the source-relay link is not fading, and the second equality holds if and only if the relay-destination is not fading.

Proof: Let $U = |g_{12}|^2$, $V = |g_{23}|^2$, then from (5.4) and (5.6), the SNR of the relay path, $S_2(\rho)$, can be rewritten as

$$S_2^{\mathrm{CA}} = \frac{UV}{V + G_{12}/\alpha + 1/(\alpha\rho)} \rho \tag{5.8}$$

and

$$S_2^{\mathrm{CI}} = \frac{UV}{V + U/\alpha + 1/(\alpha\rho)} \rho, \tag{5.9}$$

for constant-amplification and channel-inversion relaying, respectively.

Since the function $f(x) = x/(x+c)$ is a concave function of x for positive values of c, it follows from Jensen's inequality that

$$\begin{aligned}\mathbb{E}_U\left[S_2^{\mathrm{CI}}\right] &\leq \frac{\mathbb{E}[U]V}{V+\mathbb{E}[U]/\alpha+1/(\alpha\rho)}\rho \\ &= \frac{G_{12}V}{V+G_{12}/\alpha+1/(\alpha\rho)}\rho \\ &= \mathbb{E}_U\left[S_2^{\mathrm{CA}}\right],\end{aligned}$$

where the equality holds if U is deterministic, i.e., the source-relay link is not fading.

Hence

$$\begin{aligned}\mathbb{E}\left[S_2^{\mathrm{CI}}\right] &= \mathbb{E}_V\left[\mathbb{E}_U\left[S_2^{\mathrm{CI}}\right]\right] \\ &\leq \mathbb{E}_V\left[\mathbb{E}_U\left[S_2^{\mathrm{CA}}\right]\right] \\ &= \mathbb{E}\left[S_2^{\mathrm{CA}}\right]. \end{aligned} \tag{5.10}$$

The upper bound of $\mathbb{E}\left[\mathsf{S}_2^{\mathrm{CA}}(\rho)\right]$ can be obtained by applying Jensen's inequality again,

$$\begin{aligned}\mathbb{E}\left[S_2^{\mathrm{CA}}\right] &= \mathbb{E}_V\left[\frac{G_{12}V}{V+G_{12}/\alpha+1/(\alpha\rho)}\rho\right] \\ &\leq \frac{G_{12}G_{23}}{G_{23}+G_{12}/\alpha+1/(\alpha\rho)}\rho, \end{aligned} \tag{5.11}$$

where the equality holds if V is deterministic, i.e., the relay-destination link is not fading.

Combining (5.10) and (5.11) completes the proof.

■

For Rayleigh fading and constant-amplification relaying, the average SNR is evaluated in Appendix B.1 and is given by the expression

$$\mathbb{E}\left[S_2^{\mathrm{CA}}\right] = G_{12}\rho\left[1+\mu\exp(\mu)E_1(\mu)\right], \tag{5.12}$$

with

$$\mu = \frac{1}{\alpha G_{23}}\left(G_{12}+\frac{1}{\rho}\right) \tag{5.13}$$

and E_1 denoting the first order exponential integral function (see (B.4) in Appendix B.1).

For a simpler expression, the average SNR of the relay path can be further upper-bounded by

$$\mathbb{E}[S_2] \leq \frac{\alpha G_{12} G_{23} \rho^2}{1 + G_{12}\rho + \alpha G_{23}\rho} < \frac{\alpha G_{12} G_{23}}{G_{12} + \alpha G_{23}} \rho \triangleq \bar{S}_2^{\mathrm{UB}}. \tag{5.14}$$

It is noted that the effect of dropping the term 1 in the denominator is negligible when the SNR is high.

Since

$$\mathbb{E}[S_1] = \bar{S}_1 = G_{13}\rho \tag{5.15}$$

it follows from (5.7) and (5.14) that the average SNR is upper-bounded by

$$\mathbb{E}[S_c] < \bar{S}_c^{\mathrm{UB}} \triangleq \left(G_{13} + \frac{\alpha G_{12} G_{23}}{G_{12} + \alpha G_{23}} \right) \rho. \tag{5.16}$$

Fig. 5.2 shows the SNR gain versus the total SNR, for $\alpha = 1$, $G_{13} = 1$, $G_{12} = 8$, and $G_{23} = 8$ (this corresponds to the case that the relay is located halfway between the source the relay for a path loss exponent of 3). The SNR gain is defined as the ratio of average output SNR $\mathbb{E}[\mathrm{S_c}(\rho)]$ to the total transmit SNR $(1 + \alpha)\rho$. Only simulation results are given for CI relaying, since the analytic expression of the average SNR is intractable. It can be observed that the SNR gain is not sensitive to the transmit SNR when SNR is modestly large. From this figure it is observed that the SNR gain for constant amplification exceeds (for the conditions shown) that for channel inversion by approximately 0.6 dB.

5.2.2 Outage Probability

For CA relaying, the outage probability is evaluated in Appendix B.2. At high SNR, the outage probability is given in (B.7), in which the first term is on the order of $\rho^{-2} \ln \rho$, and the second term is on the order of ρ^{-2}. Hence the first term dominates the outage probability

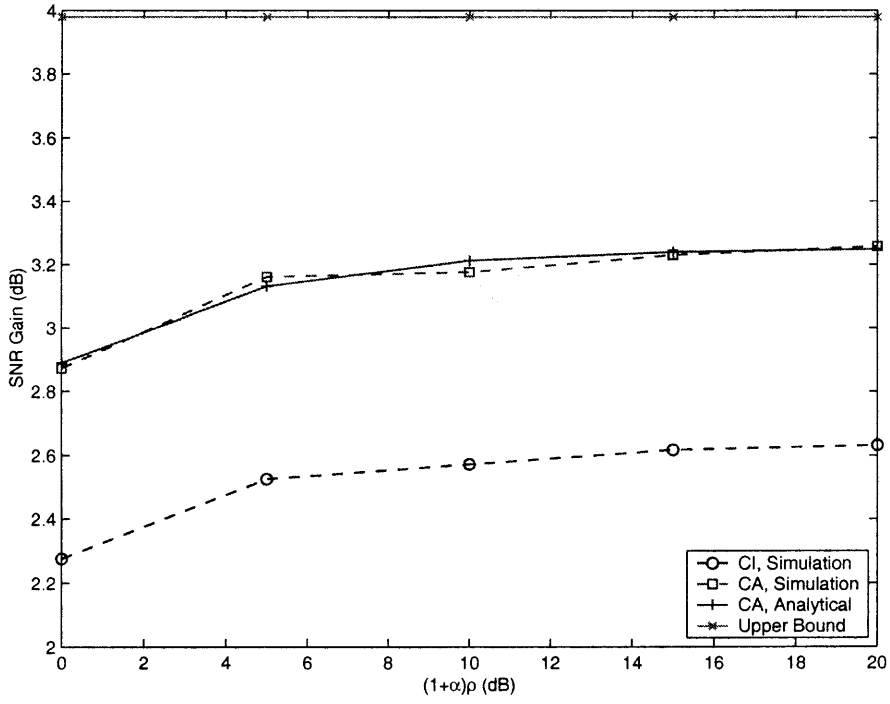

Figure 5.2 Average SNR gain versus total SNR for optimal receivers ($\alpha = 1$, $G_{13} = 1$, $G_{12} = 8$, and $G_{23} = 8$). Dashed curves are obtained via Monte-Carlo simulations. The upper bound is computed from the expression (5.16). The analytic result for constant amplification and Rayleigh fading is computed from (5.12).

when the SNR is sufficiently high. Specifically, it can be shown that

$$P_{\text{out}} \sim \frac{t^2}{2G_{13}G_{23}\alpha}\frac{\ln \rho}{\rho^2}, \tag{5.17}$$

where the sign $\sim$ indicates that the ratio of the two sides tends to unity as $\rho \to \infty$.

For channel-inversion relaying, it is shown in [44] that

$$\begin{aligned} P_{\text{out}} &\sim \frac{t^2}{2G_{13}}\left(\frac{1}{G_{12}} + \frac{1}{\alpha G_{23}}\right)\frac{1}{\rho^2} \\ &= \frac{t^2}{2\bar{S}_1\bar{S}_2^{\text{UB}}}, \end{aligned} \tag{5.18}$$

where the last relation is obtained from (5.14) and (5.15).

The outage probability can be lower-bounded by assuming that the relay knows the message of the source *a priori*, in which case the outage probability can be approximated

at high SNR by

$$P_{\text{out}}^{\text{LB}}(\rho) \sim \frac{t^2}{2G_{13}G_{23}\alpha}\frac{1}{\rho^2}. \tag{5.19}$$

A comparison of (5.17) and (5.18) with (5.19) reveals that the outage probability exceeds the lower bound by a factor of $\left(1+\frac{\alpha G_{23}}{G_{12}}\right)$ for CI relaying, and by a factor of $\ln\rho$ for CA relaying. These asymptotic expressions also indicate that second order diversity is achieved for CI relaying, and is not achieved for CA relaying.

It can be seen from (5.14), (5.15), and (5.18), that for high SNR and channel-inversion relaying, the outage probability is inversely proportional to the *product* of the average SNR of the direct link, $\bar{S}_1$, and the upper bound to the average SNR of the relay link, $\bar{S}_2^{\text{UB}}$. In the next section, the optimal power allocation by maximizing this product will be determined.

Fig. 5.3 shows the outage probabilities as functions of total SNR, $(1+\alpha)\rho$, for $\alpha = 1$ and statistically symmetric network ($G_{ij} = 1$). The threshold SNR is set to $t = 0$ dB, the minimum SNR to maintain a rate above $R = 1$ bit/transmission. From (5.17) and (5.18), for each 10 dB increase in SNR, the outage probability decreases by a factor 100 for channel inversion and by a factor of $100/(\ln 10)$ for constant amplification, and this factor is 10 for direct transmission. However, this comparison only applies to large SNR. It can be seen from Fig. 5.3 that the high SNR approximation is accurate only when the SNR is much higher (say, 20 dB) than the threshold SNR. For an outage probability of 10^{-3}, the difference in required SNR between channel inversion and constant amplification is less than 2 dB.

5.2.3 The Amount of Fading

It have been seen that although the constant-amplification relaying has a higher average output SNR than channel-inversion relaying, it performs worse in terms of outage performance when the SNR is high. This suggests that the SNR for the CA relaying has a higher variance. To see this, one can analyze the *amount of fading* (AF), which is defined as

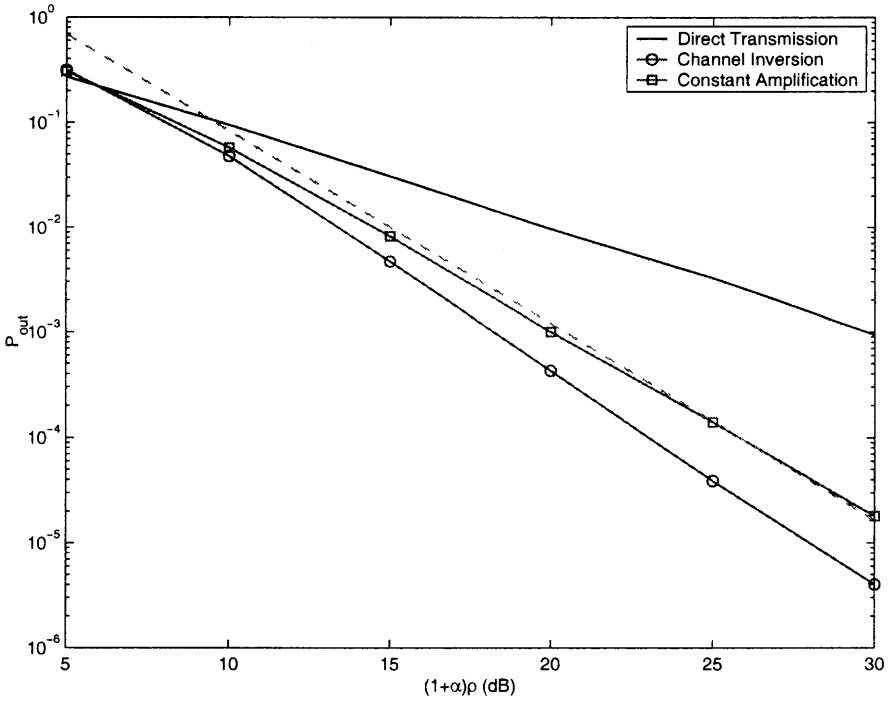

Figure 5.3 Outage probabilities versus total SNR for optimal receivers ($\alpha = 1$, $G_{ij} = 1$, $t = 0$ dB). Solid curves are obtained via Monte-Carlo simulation. The dashed curve is computed from expression (B.7).

the ratio of the variance to the square of the first moment of SNR, and serves as a simple measure of severity of the fading [79] [80]. The AF associated with the relay link is

$$AF\left(S_2\right) = \frac{\mathrm{var}\left[S_2\right]}{\left(\mathbb{E}\left[S_2\right]\right)^2} = \frac{\mathbb{E}\left[\left(S_2\right)^2\right]}{\left(\mathbb{E}\left[S_2\right]\right)^2} - 1. \tag{5.20}$$

For CA relaying and Rayleigh fading, the amount of fading is evaluated in Appendix B.1 and is given by the expression

$$AF\left(S_2^{CA}\right) = 2f\left(\mu\right) - 1 \tag{5.21}$$

with μ given by (5.13) and

$$f\left(x\right) = \frac{1 - x\exp\left(x\right)\left[2E_1\left(x\right) - E_2\left(x\right)\right]}{\left[1 - x\exp\left(x\right)E_1\left(x\right)\right]^2},$$

where E_n is the exponential integral function (Appendix B.1, (B.4)).

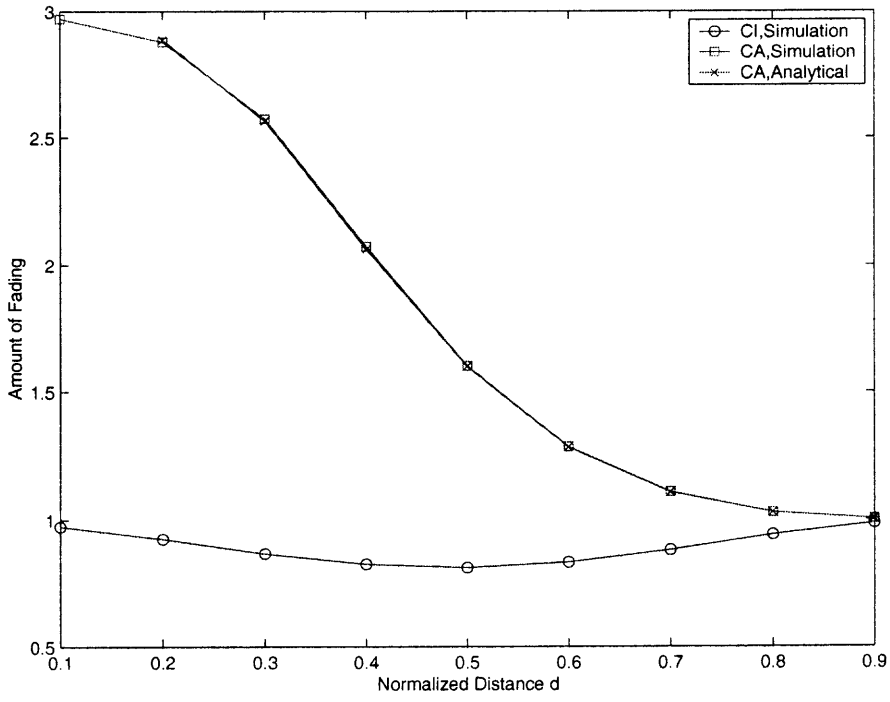

Figure 5.4 The amount of fading for the relay link ($\alpha = 1$, $\rho = 10$ dB).

Fig. 5.4 shows the amount of fading for the relay link, for $\alpha = 1$ and $\rho = 10$ dB. Again, only simulation results are provided for CI relaying because the analytic expression is not available. It can be seen from this figure that CA relaying has a higher AF value than CI relaying. For CA relaying, the AF decreases as the relay moves towards the destination, but is always greater than 1. For CI relaying, the AF appears to be less than 1, and the minimum is attained when the relay is located halfway between the source and relay, i.e., the source-relay and relay-destination links are balanced. Since both the source-relay and relay-destination links are Rayleigh fading and have AF values of 1, the end-to-end relay link is more fading than each individual link for CA relaying, but the fading is mitigated by using CI relaying.

5.3 Power Optimization

In this section the power allocation (the value of α) will be determined that is optimal in some sense. To do so, one can fix the total transmitted power $(E_s + E_r)$ and exploit the knowledge of the mean strengths of the channels, which are assumed to be available to the source and relay. Let $\rho_T = (E_s + E_r)/N_0$, then

$$\rho = \frac{\rho_T}{1+\alpha}. \tag{5.22}$$

In the following, first the upper bound to average SNR of combined signal (5.16) is maximized, which is just the *sum* of $\bar{S}_1$ and $\bar{S}_2^{\text{UB}}$. Then the *product* of $\bar{S}_1$ and $\bar{S}_2^{\text{UB}}$ is maximized. As observed before, this also minimizes the outage probability for high SNR in the case of channel-inversion relaying.

5.3.1 Power Optimization of SNR Gain

Upon substitution of (5.22), the upper bound on the average SNR, (5.16), can be rewritten as

$$\bar{S}_{\text{c}}^{\text{UB}} = \left(G_{13} + \frac{\alpha G_{12} G_{23}}{G_{12} + \alpha G_{23}}\right) \frac{\rho_T}{1+\alpha}. \tag{5.23}$$

The optimal value of α that maximizes (5.23) is found to be

$$\alpha_1 = \begin{cases} 0, & G_{23} < G_{13} \\ \frac{-G_{13}G_{12} + G_{12}\sqrt{\delta}}{G_{23}(G_{13}+G_{12})}, & \text{otherwise,} \end{cases} \tag{5.24}$$

with

$$\delta = G_{13}G_{23} + G_{12}G_{23} - G_{13}G_{12}.$$

With this power allocation, the average output SNR upper bound (5.23) is given by

$$\left(\bar{S}_{\text{c}}^{\text{UB}}\right)_{\text{opt}} = \left(\frac{G_{12}+G_{13}}{G_{12}+\sqrt{\delta}}\right)^2 G_{23}\rho_T. \tag{5.25}$$

Since $\sqrt{\delta} > G_{13}$ for $G_{23} > G_{13}$, it follows from (5.25) that

$$\left(\bar{S}_{\mathrm{c}}^{\mathrm{UB}}\right)_{\mathrm{opt}} < G_{23}\rho_T.$$

This result agrees with the intuition that the SNR can be enhanced, by using cooperative relaying, only if the relay-destination link is stronger than the source-destination link (or $G_{23} > G_{13}$), and that the resulting power gain can never exceed G_{23}/G_{13}.

For CA relaying, the optimal power allocation can also be obtained by maximizing the exact average SNR (per (5.12) and (5.15)) rather than the upper bound (5.23). However, the resulting allocation factor α depends on the SNR, and must be determined numerically since no simple analytical solution seems to be available.

Fig. 5.5 shows the SNR gain versus the normalized distance d between the source and relay, for $\rho = 10$ dB. The relay is assumed to be located on the line that passes through the source and destination, and the mean strengths of the channel are determined by $G_{13} = 1$, $G_{12} = d^{-3}$, $G_{23} = (1-d)^{-3}$. It can be seen that power optimization has a significant impact on the SNR gain mainly when the relay gets closer to the destination. It is noticed that there are crossovers between the curves for $\alpha = 1$ and $\alpha = \alpha_1$. This is because for $\alpha = \alpha_1$ it is not the exact average SNR that is maximized, but its upper bound. However, the difference is very small, as can be observed for CA relaying, for which the exact average SNR can be optimized using (5.12) and (5.15).

5.3.2 Power Optimization of Outage

Now consider the optimization of the product $\bar{S}_1\bar{S}_2^{\mathrm{UB}}$, which according to (5.18), minimizes the outage probability at high SNR. From (5.15), (5.14) and (5.22),

$$\bar{S}_1\bar{S}_2^{\mathrm{UB}} = \frac{\alpha G_{12}G_{23}G_{13}\rho_T^2}{(G_{12}+\alpha G_{23})(1+\alpha)^2} \tag{5.26}$$

The optimal value of α that maximizes (5.26) is given by

$$\alpha_2 = \frac{1}{4}\left(-\lambda + \sqrt{\lambda^2+8\lambda}\right) \tag{5.27}$$

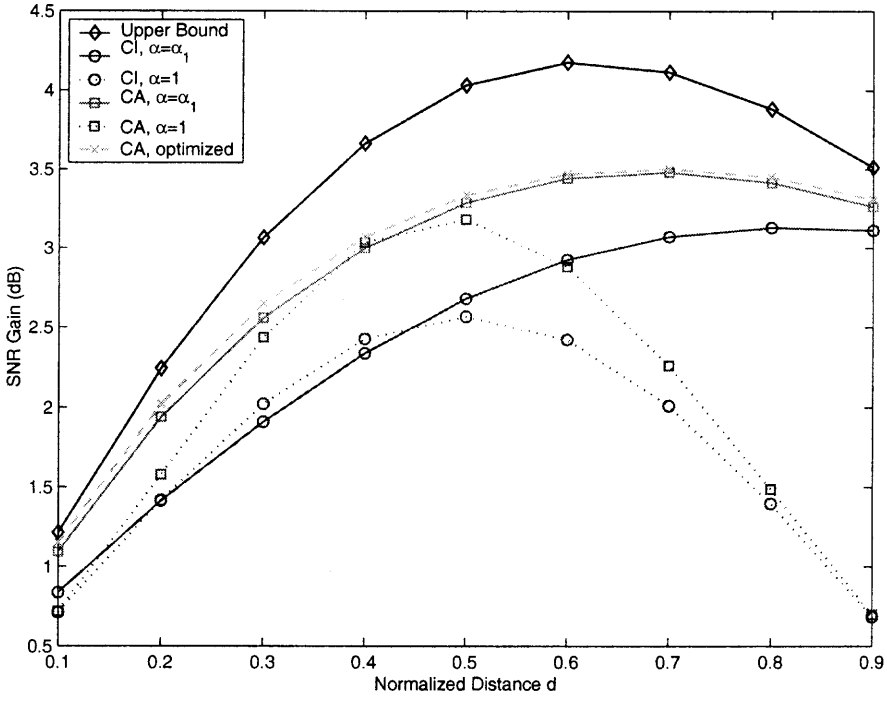

Figure 5.5 SNR gain versus normalized distance for optimal receivers ($\rho = 10$ dB); α_1 is given by (5.24).

where

$$\lambda = \frac{G_{12}}{G_{23}}.$$

Interestingly, this power allocation does not depend on the strength of the direct source-destination link. It can be shown that α_2 is an increasing function of λ, $\lim_{\lambda\to\infty} \alpha_2 = 1$, and $\alpha_2 < 1$ for $\lambda < \infty$, implying that the source is always allocated more power than the relay, and that less fractional power needs to be allocated to the relay when the relay-destination channel becomes relatively weaker.

Now compare the total transmitted power required for a system with and without power optimization. Let $\rho_T|_{\alpha=1}$ and $\rho_T|_{\alpha=\alpha_2}$, respectively, denote the required transmitted SNR for $\alpha = 1$ and $\alpha = \alpha_2$ to maintain the same level of outage performance, and assume that the outage probability depends on the SNR only through $\bar{S}_1\bar{S}_2^{\mathrm{UB}}$, then from (5.26) and

after some algebraic manipulations, it can be shown that

$$\rho_T|_{\alpha=\alpha_2} = g(\lambda)\,\rho_T|_{\alpha=1},$$

where

$$g(\lambda) = \sqrt{\frac{(1+\alpha_2)^2(\lambda+\alpha_2)}{4\alpha_2(\lambda+1)}}. \tag{5.28}$$

It can be further shown that $g(\lambda)$ is an increasing function of λ, with $g(0) = 1/2$, and $\lim_{\lambda\to\infty} g(\lambda) = 1$. Therefore the power saved by using optimal allocation α_2 is at most 3 dB. This occurs when $\lambda \to 0$, or the relay-destination link is too weak so that half of the power allocated to the relay is wasted.

Fig. 5.6 shows the outage probabilities versus $(1+\alpha)\rho$, for $G_{13} = 1$, $G_{12} = 64/27$, $G_{23} = 64$ (or normalized distance $d = 0.75$). The threshold SNR is $t = 0$ dB. It can be observed that for high SNR, the savings due to power optimization are, respectively, about 2 dB and 1 dB, for CI and CA relaying. In this case $\lambda = 1/27$, $\alpha_2 = 0.1271$, and the saving for CI relaying predicted by (5.28) is 2.01 dB .

5.4 Suboptimal Receiver with Local CSI Only

When the destination node knows only the channel coefficients of *local* links, i.e., knowledge of g_{12} is not available, maximal-ratio combining can not to employed. It is noted that coherent combining is still possible if the phase uncertainty of the source-relay is taken care of at the relay: the relay can just set the phase of the amplification factor A opposite to that of g_{12}. At the destination, the phases of the weighting coefficients w_1 and w_2 are set opposite to those of g_{13} and g_{23}, respectively. Without loss of generality, let $|w_1| = 1$, and determine $|w_2|$ without requiring the knowledge of g_{12}. The SNR of the combined signal is then given by

$$S_c = \frac{\left(|g_{13}| + |w_2|\,|A|\,|g_{12}|\,|g_{23}|\right)^2}{1 + |w_2|^2\left(1 + |A|^2\,|g_{23}|^2\right)}\rho. \tag{5.29}$$

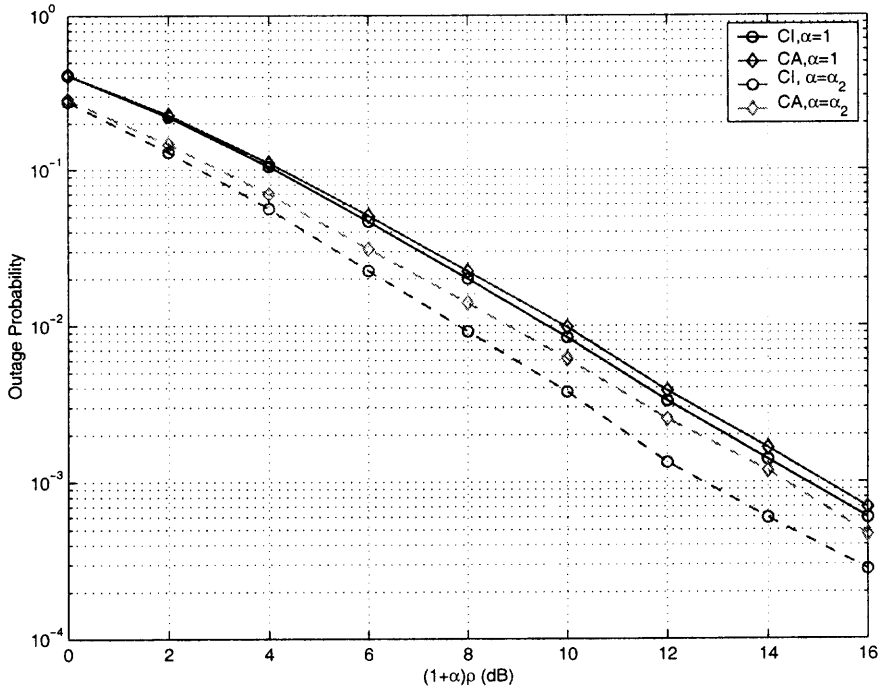

Figure 5.6 Outage probabilities versus total SNR for optimal receivers ($G_{13} = 1$, $G_{12} = 64/27$, $G_{23} = 64$, $t = 0$ dB); α_2 is given by (5.27).

5.4.1 Equal-Gain Combiner (EGC)

Probably the simplest suboptimal combining method is equal-gain combining, for which the magnitude of the weighing coefficients w_1 and w_2 are set equal, i.e., $|w_2| = |w_1| = 1$. In this case, the SNR of the combined signal is given by

$$S_c = \frac{\left(|g_{13}| + |A|\,|g_{12}|\,|g_{23}|\right)^2}{|A|^2\,|g_{23}|^2 + 2}\rho.$$

EGC requires only knowledge of the phase of local CSI and can be applied with both CI and CA relaying.

5.4.2 Maximal-Average-SNR Combiner (MASC)

Without global knowledge of CSI at the receiver, although the instantaneous SNR can not be maximized it is possible to design a combiner that maximizes the average SNR for CA

relaying, conditioned on g_{13} and g_{23}, assuming that the first two moments of $|g_{12}|$ is known at the destination. The SNR of the combined signal is a function of the random variable g_{12}. For CA relaying, averaging (5.29) over g_{12} yields the following conditional average SNR:

$$\mathbb{E}_{g_{12}}[S_c] = \frac{|g_{13}|^2 + 2|A||g_{13}||g_{23}|\mathbb{E}[|g_{12}|]|w_2| + |A|^2\mathbb{E}[|g_{12}|^2]|g_{23}|^2|w_2|^2}{1+|w_2|^2(1+|A|^2|g_{23}|^2)}\rho$$

The optimal magnitude of w_2 that maximizes the average SNR is given by the expression

$$|w_2|_{opt} = \frac{1}{2}\left[\sqrt{\gamma^2 + \frac{4}{1+|A|^2|g_{23}|^2}} + \gamma\right]$$

with

$$\gamma = \frac{|A|^2\mathbb{E}[|g_{12}|^2]|g_{23}|^2 - |g_{13}|^2(1+|A|^2|g_{23}|^2)}{|A||g_{13}||g_{23}|\mathbb{E}[|g_{12}|](1+|A|^2|g_{23}|^2)}.$$

5.4.3 Selective EGC (SEGC)

Unlike MRC, the inclusion of the relay link in EGC does not always strengthen the signals due to lack of knowledge of the source-relay channel. Now consider a simple combiner that only uses the signals in the relay link only when the direct link is weak, specifically

$$|w_2| = \begin{cases} 0, & |g_{13}|\rho \geq t \\ 1, & \text{otherwise.} \end{cases}$$

For either CA or CI relaying, the distribution functions of the suboptimal combiners discussed above are intractable. Instead, performance can be examined by resorting to numerical simulations.

In Fig. 5.7 the SNR gains for receivers with global and local CSI are plotted as functions of the normalized distance for $\alpha = 1$ and $\rho = 10$ dB. It can be seen that the gap between the curves for MRC (global CSI) and EGC (local CSI) receivers becomes wider when the relay moves closer to the destination. For CA relaying, the SNR gain for the MRC receiver is only slightly greater (within 0.5 dB) than for MASC (local CSI).

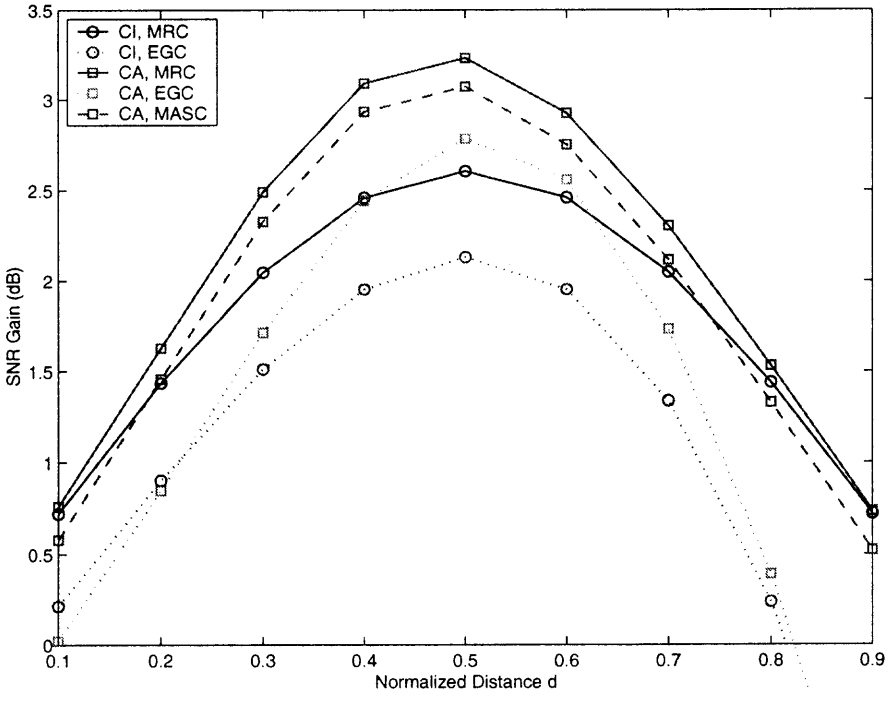

Figure 5.7 Average SNR versus normalized distance for optimal and suboptimal receivers ($\alpha = 1$, $\rho = 10$ dB).

Fig. 5.8 compares the outage probabilities for the optimal and suboptimal receivers for a statistically symmetrical network. The threshold SNR is $t = 0$ dB. It is observed that, due to lack of global knowledge of CSI and the use of EGC, the second-order diversity is completely lost for channel-inversion relaying, while the loss is less significant for constant-amplification relaying. As expected, the outage performance of SEGC is better than EGC.

5.5 Chapter Summary

This chapter investigates the performance of amplify-and-forward cooperative relaying schemes in Rayleigh fading channels, in terms of average signal-to-noise ratio (SNR) and outage probability and for various cases of available channel state information (CSI). For optimal receivers with global knowledge of CSI, it is shown that constant-amplification

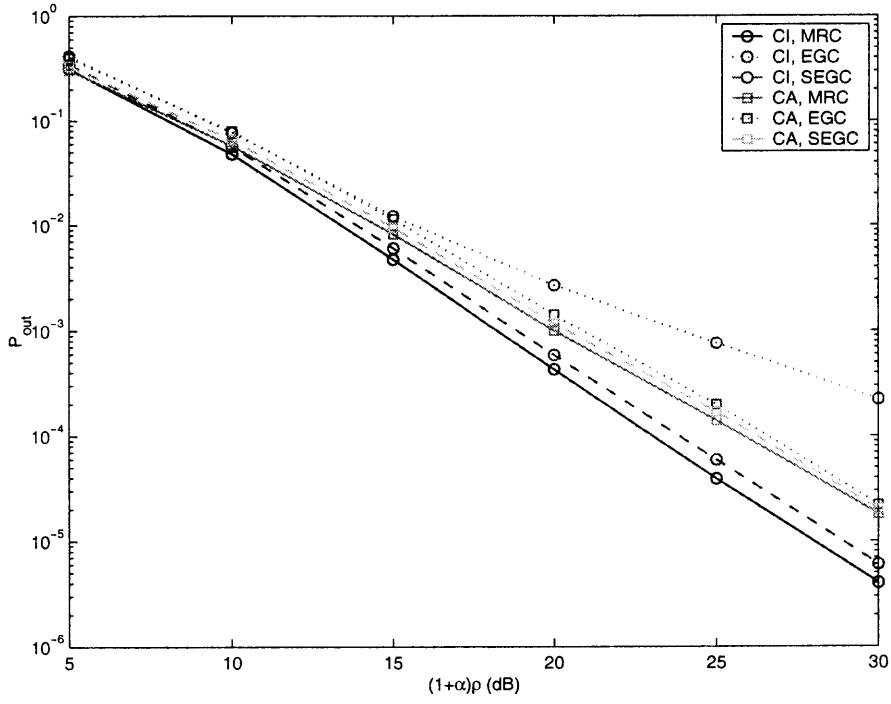

Figure 5.8 Outage probabilities versus total SNR for optimal and suboptimal receivers ($\alpha = 1, G_{ij} = 1, t = 0$ dB).

relaying has a higher average SNR gain and a lower diversity order than channel-inversion relaying. Optimal power allocation strategies are developed that optimize the average SNR and outage performance, respectively. When only local knowledge of CSI is available, it is demonstrated through simulations that suboptimal coherent receivers can still achieve substantial power and diversity gains.

CHAPTER 6

MULTICHANNEL NONCOHERENT DETECTION WITH APPLICATIONS TO COOPERATIVE DIVERSITY

It appears that the Bayesian-based noncoherent detector can not be easily applied to systems with cooperative diversity. First, Bayesian-based noncoherent detector requires the statistical information of the channels, which may not be available since each relayed channel depends on the source-relay and relay-destination channels and the operation of the relay nodes. Second, even if the prior distribution of the channels is available, the Bayesian-based detection may be too complicated to implement since multiple integration is involved. For example, even the statistics of both the source-relay and relay-destination channels are of very simple forms, say Rayleigh, the distribution of end-to-end source-destination can only be expressed in integral functions. Third, because of noise amplification at the relay nodes, the noise powers in a cooperative diversity depend on the gains of the relay-destination channels and the amplification factors of the relay nodes. In [81], a suboptimal noncoherent detection, inspired by the ML nondetection of decode-and-forward cooperative diversity, is suggested for amplify-and-forward cooperative diversity.

In this chapter two noncoherent detection algorithms are derived for multichannel receivers using the GLRT approach, which do not require knowledge of the channel statistics at the receiver. When the noise powers are known, the GLRT-based detector chooses the signal that minimizes the weighed sum of squared errors; otherwise the detector chooses the signal that maximizes the product of squared errors. These detectors are used for the noncoherent detection of systems with AF cooperative diversity.

The remainder of this chapter is organized as follows. Section 6.1 briefly describes the multichannel system model and related results on noncoherent detection. Section 6.2 contains the development of two noncoherent detection algorithms based on the GLRT

rules. In Section 6.3 the developed algorithms are used for noncoherent detection for amplify-and-forward cooperative diversity. Conclusions are given in Section 6.4.

6.1 Multichannel System Model

This section presents the signal model of multichannel reception and briefly summarizes related detections algorithms.

In a wireless communication system operating over L independent flat-fading branches, the discrete-time received signal of the lth branch at time k is given by

$$r_{l,k} = h_l s_k + n_{l,k}$$

where s_k is the transmitted symbol, h_l is the channel gain of the lth branch, and $n_{l,k}$ is sample of the additive Gaussian noise. It is assumed that the channel coefficient remains constant over a duration of K symbol periods (block fading). Then the vector notation for the signal model is

$$\mathbf{r}_l = h_l \mathbf{s} + \mathbf{n}_l \tag{6.1}$$

where $\mathbf{r}_l = (r_{l,1}, r_{l,2}, \cdots, r_{l,K})^T$, $\mathbf{s} = (s_1, s_2, \cdots, s_K)^T$, and $\mathbf{n}_l = (n_{l,1}, n_{l,2}, \cdots, n_{l,K})^T$. Assume the noise samples are independent with respect to l and k. Specifically,

$$E\left[n_{l,k} n^*_{l',k'}\right] = \delta(l - l')\,\delta(k - k')\,\sigma_l^2.$$

Here the noise powers σ_l^2 are not necessarily equal, although the assumption of equal noise powers are reasonable in most diversity systems. For example, in a system with amplify-and-forward cooperative diversity, the noise powers are in general not equal because of the noise amplification at the relay nodes. The case of unequal noise powers may also happen in systems with interference cancellation [82].

The conditional probability density function (pdf) of the received signal can be expressed as

$$\begin{aligned} p(\{\mathbf{r}_l\}_{l=1}^{L}|\mathbf{s},\{h_l\}_{l=1}^{L},\{\sigma_l^2\}_{l=1}^{L}) &= \prod_{l=1}^{L} p\left(\mathbf{r}_l|\mathbf{s},h_l,\sigma_l^2\right) \\ &= \prod_{l=1}^{L} \frac{1}{(\pi\sigma_l^2)^K} \exp\left(-\frac{\|\mathbf{r}_l-\mathbf{s}h_l\|^2}{\sigma_l^2}\right), \end{aligned} \tag{6.2}$$

where $p\left(\mathbf{r}_l|\mathbf{s},h_l,\sigma_l^2\right)$ is the conditional pdf of the received signal in the lth branch.

The optimal noncoherent ML detector chooses the transmitted symbol sequence that maximizes the noncoherent likehood function, which can be obtained by integrating the conditional probability (6.2) over the unknown channel parameters $\{h_l\}_{l=1}^{L}$:

$$p(\{\mathbf{r}_l\}_{l=1}^{L}|\mathbf{s},\{\sigma_l^2\}_{l=1}^{L}) = \int\cdots\int p(\{\mathbf{r}_l\}_{l=1}^{L}|\mathbf{s},\{h_l\}_{l=1}^{L},\{\sigma_l^2\}_{l=1}^{L}) p\left(\{h_l\}_{l=1}^{L}\right) dh_1\cdots dh_L. \tag{6.3}$$

Evaluation of the likehood functions (6.3) requires knowledge of the channel statistics ($p\left(\{h_l\}_{l=1}^{L}\right)$) and knowledge of the noise powers ($\{\sigma_l^2\}_{l=1}^{L}$). For independent Rayleigh fading, it is shown that the maximizing (6.3) is equivalent to maximizing the metric [42,82]

$$\eta'(\mathbf{s}_i) = \sum_{l=1}^{L} \frac{|\mathbf{s}_i^H\mathbf{r}_l|^2}{\sigma_l^2\left(KE_S+\sigma_l^2/\Omega_l\right)} \tag{6.4}$$

where Ω_l is the mean strength of the lth diversity branch.

For more general channels, noncoherent ML detection is complicated [42]. When the mean strengths Ω_l of the channel branches are know, the detector using Rayleigh (6.4) serves as a suboptimal detector for general channels. This detector is referred to as the suboptimal likelihood-ratio test (SLRT) detector in the sequel.

Another suboptimal detector is the noncoherent Equal-Gain Combining (NEGC) (also called postdetection EGC) detector, whose decision metric is given by

$$\eta'(\mathbf{s}_i) = \sum_{l=1}^{L} |\mathbf{s}_i^H\mathbf{r}_l|^2. \tag{6.5}$$

6.2 Noncoherent Detection without Knowledge of Channel Statistics

In this section, detection algorithms are developed using the GLRT rules, in which no knowledge of the channel statistics is required.

6.2.1 Known Noise Powers

When the noise powers $\{\sigma_l^2\}_{l=1}^L$ are known, the GLRT detector is given by

$$\begin{aligned}\hat{i} &= \arg\max_i \max_{\{h_l\}_{l=1}^L} p(\{\mathbf{r}_l\}_{l=1}^L | \mathbf{s}_i, \{h_l\}_{l=1}^L, \{\sigma_l^2\}_{l=1}^L) \\ &= \arg\max_i \prod_{l=1}^L \frac{1}{(\pi\sigma_l^2)^K} \exp\left(-\frac{\left\|\mathbf{r}_l - \mathbf{s}_i \hat{h}_l^{(i)}\right\|^2}{\sigma_l^2}\right)\end{aligned} \tag{6.6}$$

where $\hat{h}_l^{(i)}$ is the maximum-likelihood estimate of $\hat{h}_l$ under the assumption that $\mathbf{s}_i$ is transmitted, i.e.,

$$\begin{aligned}\hat{h}_l^{(i)} &= \arg\max_{h_l} p\left(\mathbf{r}_l | \mathbf{s}_i, h_l, \sigma_l^2\right) \\ &= \frac{\mathbf{s}_i^H \mathbf{r}_l}{|\mathbf{s}_i|^2}.\end{aligned} \tag{6.7}$$

Substitution of (6.7) into (6.6) leads to

$$\hat{i} = \arg\max_i \frac{1}{\prod_{l=1}^L (\pi\sigma_l^2)^K} \exp\left(-\sum_{l=1}^L \frac{1}{\sigma_l^2}\left[|\mathbf{r}_l|^2 - \frac{|\mathbf{s}_i^H \mathbf{r}_l|^2}{|\mathbf{s}_i|^2}\right]\right)$$

Therefore, the GLRT detector chooses the signal $\mathbf{s}_i$ that minimizes

$$\eta(\mathbf{s}_i) = \sum_{l=1}^L \frac{1}{\sigma_l^2} e_l(\mathbf{s}_i) \tag{6.8}$$

where

$$e_l(\mathbf{s}_i) = |\mathbf{r}_l|^2 - \frac{|\mathbf{s}_i^H \mathbf{r}_l|^2}{|\mathbf{s}_i|^2} \tag{6.9}$$

is just the squared error of fitting the received signal vector $\mathbf{r}_l$ by another signal vector $\mathbf{s}h_l$. Therefore this detector is called the minimum weighted sum of squared errors (MWSE) detector.

Minimization of (6.8) is equivalent to the maximization of the following weighted sum of the correlations

$$\sum_{l=1}^{L} \frac{|\mathbf{s}_i^H \mathbf{r}_l|^2}{\sigma_l^2 |\mathbf{s}_i|^2}$$

which simplifies to the following decision matric for signals with constant power ($|\mathbf{s}_i|^2 = KE_s$):

$$\eta'(\mathbf{s}_i) = \sum_{l=1}^{L} \frac{|\mathbf{s}_i^H \mathbf{r}_l|^2}{\sigma_l^2}. \tag{6.10}$$

Equations (6.4) and (6.10) indicate that the knowledge of channel statistics becomes less useful as the observation duration K increases. In the limit when K approaches infinite, the two detectors coincide. It can also be seen that the two detectors are equivalent for a uniform power profile and equal noise powers per branch.

6.2.2 Unknown Noise Powers

When the noise powers $\{\sigma_l^2\}_{l=1}^L$ are known, the GLRT detector is given by

$$\begin{aligned} \hat{i} &= \arg\max_i \max_{\{\sigma_l^2\}_{l=1}^L} \max_{\{h_l\}_{l=1}^L} p(\{\mathbf{r}_l\}_{l=1}^L | \mathbf{s}_i, \{h_l\}_{l=1}^L, \{\sigma_l^2\}_{l=1}^L) \\ &= \arg\max_i \prod_{l=1}^{L} \frac{1}{\left(\pi \hat{\sigma}_l^{2(i)}\right)^K} \exp\left(- \frac{\left\| \mathbf{r}_l - \mathbf{s}_i \hat{h}_l^{(i)} \right\|^2}{\hat{\sigma}_l^{2(i)}} \right) \end{aligned} \tag{6.11}$$

where $\hat{h}_l^{(i)}$ and $\hat{\sigma}_l^{2(i)}$ are the maximum-likelihood estimate of $\hat{h}_l$ and $\hat{\sigma}_l^2$ under the assumption that $\mathbf{s}_i$ is transmitted, which are respectively given by (6.7) and

$$\hat{\sigma}_l^{2(i)} = \frac{1}{K} \left\| \mathbf{r}_l - \mathbf{s}_i \hat{h}_l^{(i)} \right\|^2 \tag{6.12}$$

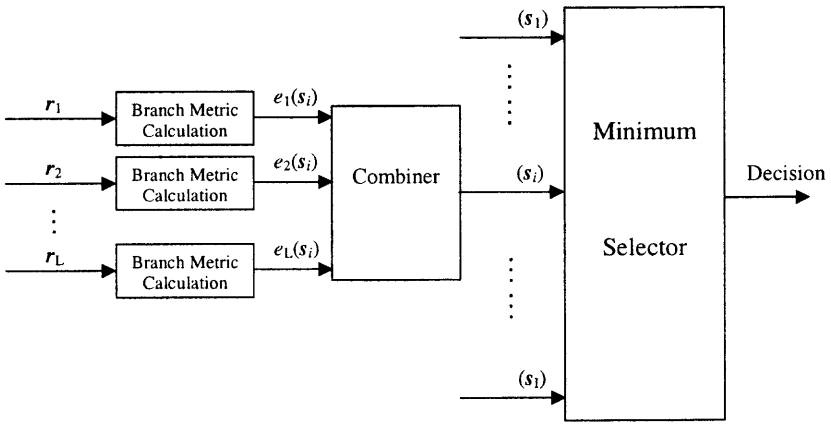

Figure 6.1 Structure of noncoherent receiver for an L-branch channel.

Table 6.1 Noncoherent Detectors with Various Available Information

Detector	Channel Statistics	Noise Powers	Combiner Function
SLRT	Y	Y	$\sum_{l=1}^{L} \frac{1}{\sigma_l^2\left(KE_S+\sigma_l^2/\Omega_l\right)} e_l$
MWSE	N	Y	$\sum_{l=1}^{L} \frac{1}{\sigma_l^2} e_l$
NEGC	N	N	$\sum_{l=1}^{L} e_l$
MPSE	N	N	$\prod_{l=1}^{L} e_l$

It follows from (6.7), (6.11), and (6.12) that the GLRT detector chooses the signal $\mathbf{s}_i$ that minimizes

$$\eta\left(\mathbf{s}_i\right) = \prod_{l-1}^{L} \left[\left|\mathbf{r}_i\right|^2 - \frac{\left|\mathbf{s}_i^H \mathbf{r}_l\right|^2}{\left|\mathbf{s}_i\right|^2} \right] \tag{6.13}$$

which, according to (6.9), is just the *product* of squared errors.

If follows from (6.4), (6.5), (6.8), (6.9) and (6.13) that all the above detectors discussed above can be implemented using the same structure shown in Fig. 6.1. The difference lies in the required knowledge of the channel statistics and noise powers and the combining functions, which is summarized in Table 6.1. It is noted that all these detectors coincide for a single-channel system ($L = 1$).

6.3 Application to Noncoherent Detection for Amplify-and-Forward Cooperative Diversity

This section considers the noncoherent detection for systems with amplify-and-forward cooperative diversity.

6.3.1 Signal Model for AF Cooperative Diversity

Consider an N-node wireless network consisting of a source, a destination, and $N-2$ relay nodes ($N \geq 3$), as illustrated in Fig. 6.2. Information is transmitted from the source to the destination with the assistance of the relay nodes. Without loss of generality, assume that node 1 is the source, node N is the destination, and nodes 2, 3, $\cdots$, $N-1$ are relay nodes. In the sequel, g_{ij} is the complex gain of the channel from node i to node j, and $\mathbf{n}_{ij}$ is the additive noise vector at node j corresponding to the channel g_{ij}, where $i \in \{1, 2, \cdots, N-1\}$, $j \in \{2, 3, \cdots, N\}$. The mean strength of the channel from node i to node j is G_{ij}. Assume that g_{ij} and n_{ij} are statistically independent for different values of i and j, and that the channel coefficients and the noise samples are independent. The noise power at the receiver of each node is N_0.

Cooperative diversity schemes consist of two-stage transmissions. In the first stage, the source transmits and all the other nodes including the destination and relay listen. Let $\mathbf{s}$ denote the $K \times 1$ baseband transmitted signal vector at the source node,then the received signal vector at node i is given by

$$\mathbf{y}_{1i} = g_{1i}\mathbf{s} + \mathbf{n}_{1i}, i = 2, 3, \cdots, N.$$

In the second stage, the relay nodes process and send the received signals to the destination node using mutually orthogonal channels. For amplify-and-forward relaying, relay node i simply amplifies the received signal $\mathbf{y}_{1i}$ by a factor A_i, and forwards it to the

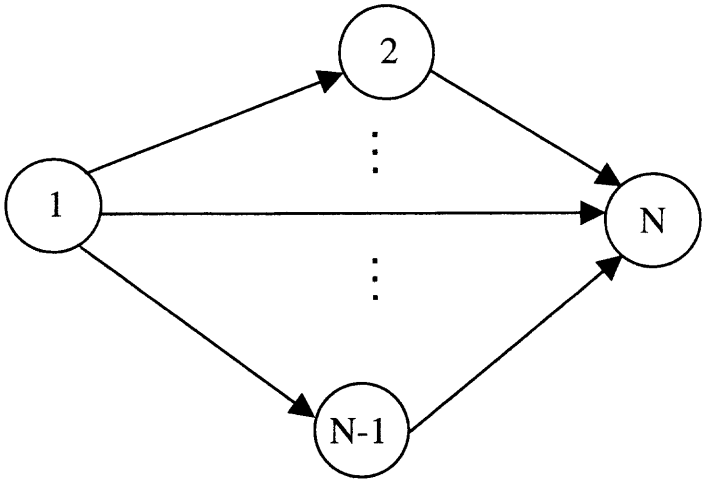

Figure 6.2 Illustration of an N-node wireless network.

destination, so that the destination nodes receives

$$\begin{aligned} \mathbf{y}_{iN} &= g_{iN}A_i\mathbf{y}_{1i} + \mathbf{n}_{iN} \\ &= g_{iN}A_ig_{1i}\mathbf{s} + g_{iN}A_i\mathbf{n}_{1i} + \mathbf{n}_{iN},\ i = 2, 3, \cdots, N-1. \end{aligned} \tag{6.14}$$

The signals received at the destination node can be formulated as the $(N-1)$-fold signal model (6.1), with $\mathbf{r}_l = \mathbf{y}_{lN}$, $l = 1, 2, \cdots, N-1$, and the equivalent channel gains and noise variances respectively given by

$$h_l = \begin{cases} g_{1N}, & l = 1 \\ g_{1i}A_ig_{iN}, & l = 2, 3, \cdots, N-1 \end{cases}$$

and

$$\sigma_l^2 = \begin{cases} N_0, & l = 1 \\ \left(A_i^2\left|g_{iN}\right|^2 + 1\right)N_0, & l = 2, 3, \cdots, N-1. \end{cases} \tag{6.15}$$

It becomes clear from (6.15) that, unlike conventional diversity systems where the noise powers are usually fixed, the noise powers in the relay links in a amplify-and-forward

cooperative diversity system depend on the instantaneous relay-destination channel gain.

The amplification factors A_i are chosen subject to the average power constraint imposed at the relay node. Let the average transmitted power at node i be E_i. There are several relaying strategies with different choices of amplification factors, e.g., the amplification factor at relay node i is can be given by

$$A_i = \sqrt{\frac{E_i}{E_1 |g_{1i}|^2 + N_0}}, \text{ Channel Inversion}$$

$$A_i = \sqrt{\frac{E_i}{E_1 G_{1i} + N_0}}, \text{ Constant Amplification}$$

$$A_i = \sqrt{\frac{KE_i}{|y_{1i}|^2}}, \text{ Constant Power}$$

Note that at the relay nodes, instantaneous channel information is needed for the channel-inversion (CI) relaying, mean channel information is need for the constant-amplification (CA) relaying, and no channel information is required for constant-power (CP) relaying. In this work, CI relaying is not considered since instantaneous channel information is required at the relay nodes.

The remainder of this section investigates the performance of the detectors developed in Section 6.2. The MPSE and NEGC detectors can be used in systems with any of the diversity schemes, since decisions are made based solely on the received signals. The MWSE detectors are suitable for constant-amplification diversity only, and requires information of the noise power N_0, the mean strengths of the source-relay channels G_{1i}, and the instantaneous amplitude of the the relay-destination channels $|g_{iN}|^2$, $i = 2, 3, \cdots, N-1$.

A suboptimal detector for constant-amplification cooperative diversity, inspired by the decision metric of decode-and-forward diversity, is suggested in [81]. It is observed that for equal-energy signaling, this detector has the same form of an SLRT detector (6.4),

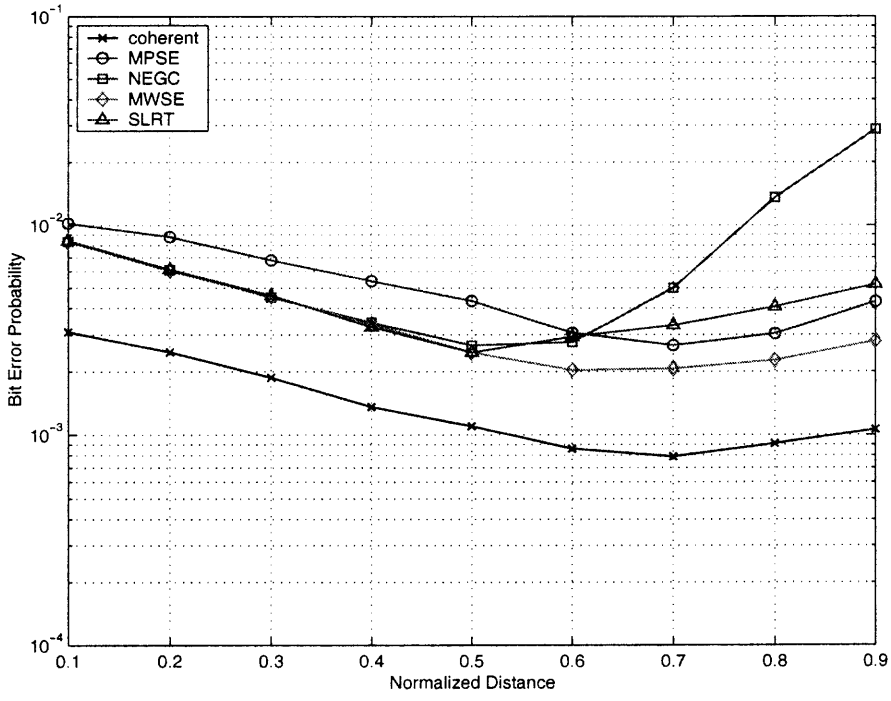

Figure 6.3 Bit error probability for CA relaying (uniform power allocation, binary DPSK, $E_T/N_0 = 13$ dB).

with Ω_l and σ_l^2 replaced by

$$\Omega_l = \begin{cases} G_{1N}, & l = 1 \\ G_{1i}A_i^2 G_{iN}, & l = 2, 3, \cdots, N-1 \end{cases}$$

and

$$\sigma_l^2 = \begin{cases} N_0, & l = 1 \\ \left(A_i^2 G_{iN} + 1\right) N_0, & l = 2, 3, \cdots, N-1. \end{cases}$$

The SLRT detector requires knowledge of the the noise power N_0 and the mean strengths of the source-relay channels G_{1i} and the relay-destination channels G_{iN}, $i = 2, 3, \cdots, N-1$.

In the network considered in this section, all the relay nodes are assumed to be located on the straight line that passes through the source and destination. Without loss of

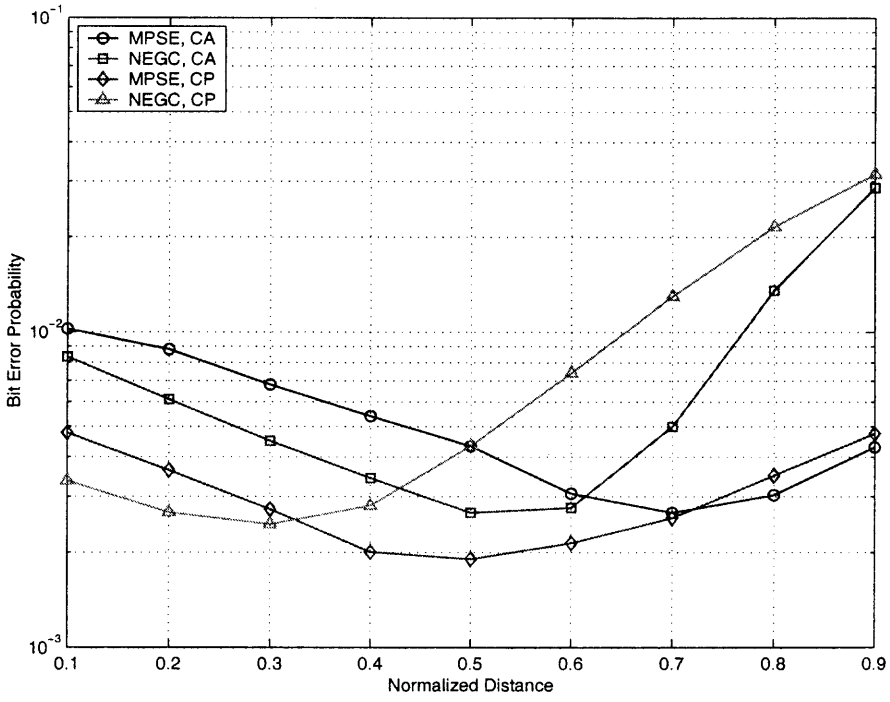

Figure 6.4 Bit error probability versus normalized distance (uniform power allocation, binary DPSK, $E_T/N_0 = 13$ dB).

generality the distance between the source and relay is normalized to be 1. The location of relay node i is determined by the normalized distance d_i between the source and the relay. Assume a path-loss exponent of 3, then the mean strengths of the channels are determined by $G_{1N} = 1$, $G_{1i} = d^{-3}$, and $G_{iN} = (1-d)^{-3}$, for $i = 2, 3, \cdots, N-1$.

6.3.2 Uniform Power Allocation

First consider the cases that all the nodes use the same average transmit power E_s. Fig. 6.3 shows the the bit error probability for a three-node network. Blocks of binary DPSK symbols of length $K = 4$ are transmitted. The total SNR is $E_T/N_0 = 13$ dB. The curves are shown as a function of the normalized distance d_2 between the relay and source, for the case of CA relaying, in which all the algorithms discussed in previous sections can be used. It can be observed that all the noncoherent detection algorithms performs reasonably well

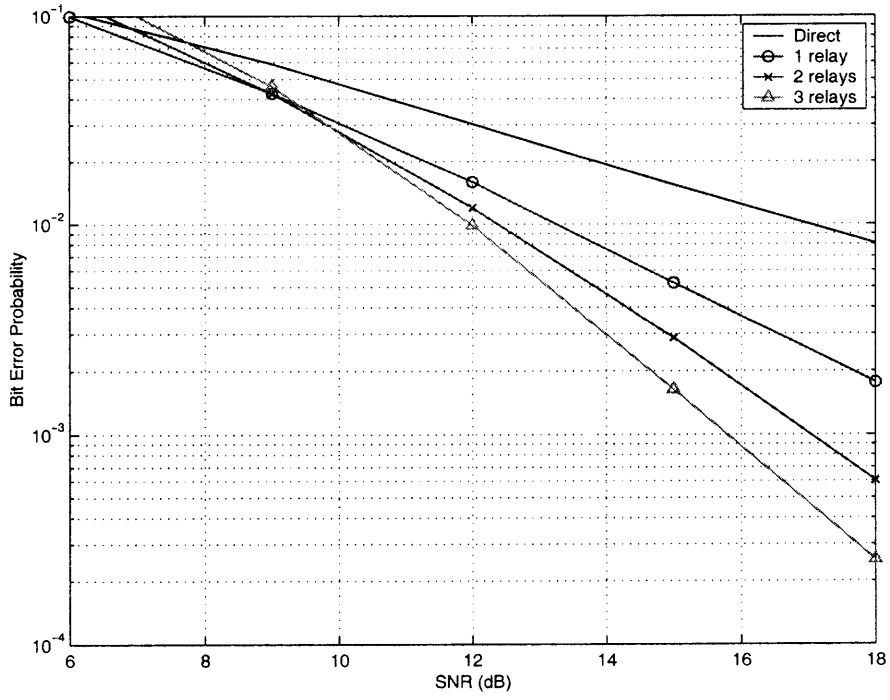

Figure 6.5 Bit error probability versus total SNR for noncoherent MPSE detection of binary DPSK (uniform power allocation).

when the relay is close to the source, but the NEGC and SLRT detectors are more sensitive to the location of the relay. In particular, the MPSE and MWSE detectors outperform the NEGC and SLRT detectors when the relay is close to the destination.

It is of interest to compare the performance of the MPSE and NEGC detectors, since both detectors can be used in any diversity systems since no knowledge of the channel statistics or noise powers is required. Fig. 6.4 compares the performance of these two detectors for CA and CP relaying schemes. These curves indicate that good performance is achieved by the MPSE detector in conjunction with CP relaying over a wide range of relay locations. For this reason the MPSE detection for CP diversity is focused in the following.

Now examine the performance of the MPSE detector in networks with multiple relays. Fig. 6.5 shows the bit error probability versus total SNR in a random network. The normalized distance d_i of relay location nodes are independently, uniformly distributed on

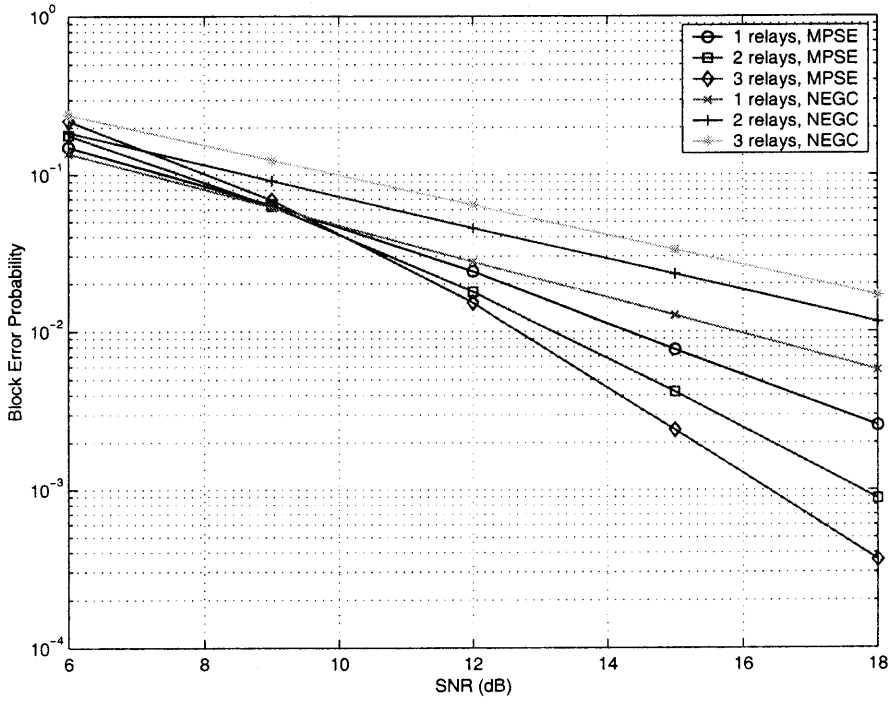

Figure 6.6 Block error probability versus total SNR for noncoherent detection of orthogonal signals (uniform power allocation, Walsh-Hadamard sequences).

the interval (0, 1). It is clear from these curves that the diversity gain for the noncoherent detection manifests only when the SNR is high enough. When the SNR is low, the diversity gain is overwhelmed by the noncoherent combining penalty. For the example considered here, the SNR beyond which using more number of relays is about 9 dB.

The same results are observed in Fig. 6.6 for noncoherent detection of orthogonal signals in a random network. The transmitted signals are Walsh-Hadamad sequences of length 4. For purpose of comparison, the performance of NEGC detector is also shown in this figure. It can be observed that the performance for NEGC detector becomes even worse with increasing number of relays.

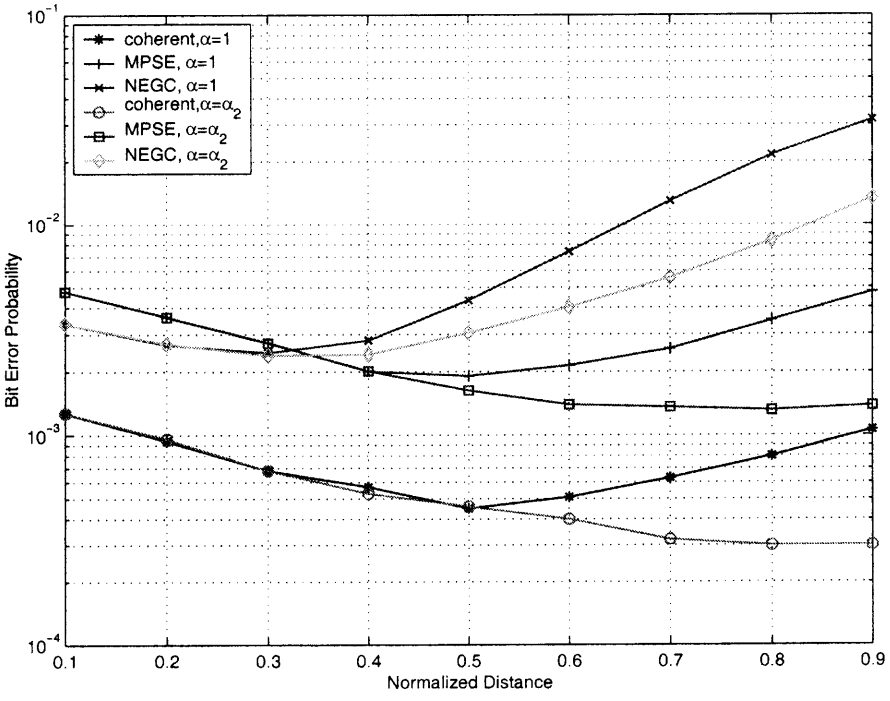

Figure 6.7 Bit error probability versus normalized distance (binary DPSK, $E_T/N_0 = 13$ dB); α_2 is given by (5.27).

6.3.3 Nonuniform Power Allocation

So far it have been supposed that all the nodes use the same transmit power. The extension of our results to account for nonuniform power allocation among nodes is straightforward. However, it is noted that determination of the optimal power allocation for the noncoherent case is very challenging. The power allocation for AF diversity is very difficulty even for coherent cases, and most analytical solutions are approximate or suboptimal, since the exact solutions generally depends on the metrics of interest, the value of SNR, and the relaying strategies. Here no attempt is made to find the optimal power allocation strategies, rather, it is intended to show whether the power strategies developed for the coherent cases could bring in some performance gain when used in systems with noncoherent detection.

Similar to [53], assume that all the relay nodes use the same average transmit power, i.e., $E_2 = E_3 = \cdots E_{N-1}$. Let us denote the ratio of power of each of the relay node to

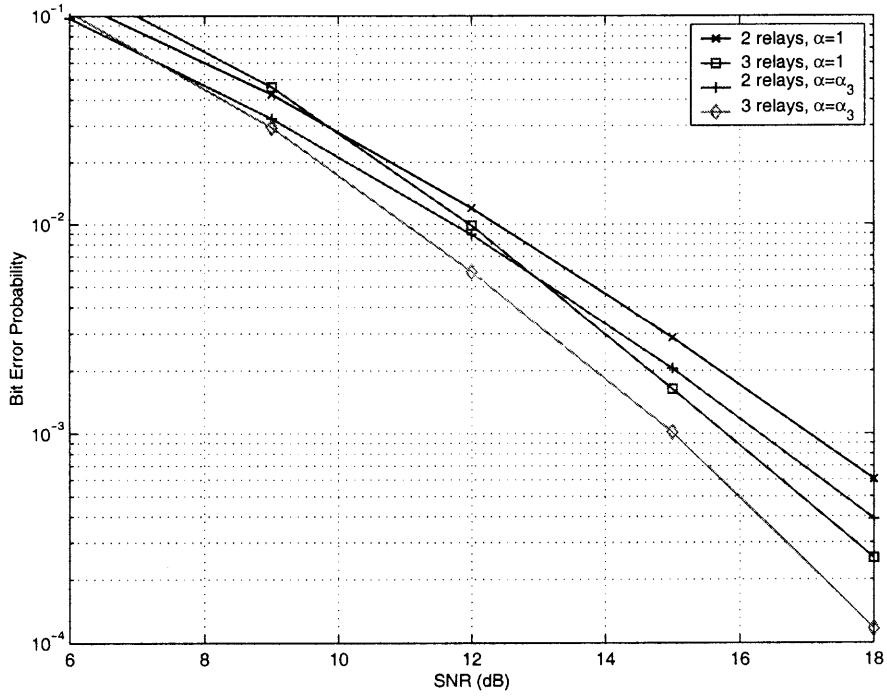

Figure 6.8 Bit error probability versus total SNR for noncoherent MPSE detection of binary DPSK; α_3 is given by (6.16).

that of the source node by $\alpha = E_2/E_1$.

For a three-node network, the value of α that minimizes the outage probability at high SNR is given by (5.27).

For the case of multiple relays ($N > 3$), a suboptimal power allocation is suggested in [53], in which all the relay nodes take half of the total transmit power, i.e.,

$$\alpha_3 = \frac{1}{N-2}. \tag{6.16}$$

As shown in Fig. 6.7, the performance for using power allocation (5.27) is better than that for uniform allocation. The same results are obtained in Fig. 6.8 in systems with multiple relays. In this figure a performance gain of about 1 dB is observed for power allocation (6.16).

6.4 Chapter Summary

Based on the generalized likelihood ratio rules, two algorithms (MWSE and MPSE) were developed for noncoherent detection without requiring knowledge the channel statistics. The detector chooses the signal that minimizes the weighted sum or product of error squares, depending on the availability of the noise powers. The detection algorithms were used for the detection of DPSK in systems with amplify-and-forward cooperative diversity. It was shown that the proposed algorithm is robust to the location of the relay node. When the total transmit power is distributed uniformly among nodes, it was also illustrated that using multiple relay nodes with noncoherent detection is only beneficial when the SNR is high. It is demonstrated that some performance improvement can be obtained by directly applying simple power allocation developed for the coherent cases to the noncoherent cases.

CHAPTER 7

CONCLUSIONS

This chapter summarizes the contribution of this work and points to some topics for future research.

7.1 Contributions

The principal results obtained in the dissertation are summarized below:

In Chapter 3 the achievable rates were investigated for i.i.d. Gaussian signaling over time-varying fading channels with CSI unknown to either the transmitter or the receiver. A lower bound on mutual information obtained by upper-bounding the penalty due to unknown CSI, and the achievable rates for pilot-aided systems with optimized resource allocations, are presented. Asymptotic expressions for the achievable rates provide quantitative measures on well-known properties of Gaussian signaling, namely, that it is efficient for slowly fading channels when the SNR is high and inefficient when the SNR is low. Specifically, when the SNR is high, the loss in bandwidth efficiency is twice the normalized fading rate, typically a small number for channels of practical interest; the loss in power efficiency depends only on the fading rate, and is less than 3 dB for pilot-aided systems. Conversely, when the SNR goes to zero, the achievable rates decrease quadratically with SNR, rather than linearly with SNR as capacity does. This conclusion is not expected to change materially if PSK signals are used in lieu of Gaussian signals since in the low SNR regime differences between signaling formats tend to be vanish.

A simple, low-duty-cycle signaling was suggested for improving information rates in the low SNR regime. The critical SNR, identified as the threshold between the low and high SNR regimes, is a function of only the fading rate. At SNR's below the critical SNR, low-duty-cycle signaling restores the linear relation between capacity and SNR. The duty

cycle is formulated as a function of the SNR and, through the critical SNR, as a function of the fading rate. At the optimal duty cycle, as the SNR goes to zero, the ratio of achievable rate to capacity with perfect CSI approaches a constant that depends only on the normalized fading rate.

Although the system parameters (resource allocations and duty cycles) were optimized for Gaussian signaling, it was shown that they can also be applied to systems using other signaling formats.

In Chapter 4 an iterative algorithm was proposed to avoid the inversion of a data-dependent matrix in decision-directed channel estimation for systems with multiple transmit antennas. Convergence conditions were found as a function of the number of transmit antennas and the noise power.

In Chapter 5 it was shown that constant-amplification relaying outperforms channel-inversion relaying in terms of average output SNR, when global knowledge of CSI is available and maximal-ratio combining is employed at the destination. However, if outage probability is concerned, channel-inversion relaying achieves full second-order diversity in Rayleigh fading channels, while constant-amplification relaying does not (the diversity order is between 1 and 2). Specifically, when the average SNR ρ is large, the outage probability decays as ρ^{-2} for channel-inversion relaying and as $\rho^{-2} \ln \rho$ for constant-amplification relaying. To show that the SNR for CA relaying is more fluctuating, the amount of fading of the relay link was examined. It was found that if both the source-relay and relay-destination links are Rayleigh fading, then the end-to-end relay link is more fading than Rayleigh for CA relaying, and is less fading than Rayleigh for CI relaying. Although the outage probabilities of CI and CA relaying behave significantly different at high SNR, it was observed that for modestly large SNR, CI relaying only slightly outperforms CA relaying.

Power allocation strategies were developed under the constraint imposed on the total transmitted power of the source and relay. Using the knowledge of mean strengths of the channels, the sum and product of the average SNR of the direct link and the upper bound

to the average SNR of the relay link were optimized. It was shown the improvements in average and outage performance resulting from these allocation strategies. It was demonstrated through simulation that when only knowledge of local CSI is available, substantial power and diversity gains are still achievable by employing suboptimal coherent combiners.

Multichannel noncoherent detection was investigated in Chapter 6. Two detection algorithms were formulated based on the generalized likelihood ratio test rules. Depending on the knowledge of the noise powers of the diversity branches, the proposed detector chooses the signal that produces the minimum weighed sum or product of error squares. Numerical results are presented for noncoherent detection of differential phase-shift keying and orthogonal signals in systems with amplify-and-forward cooperative diversity. It was observed that good performance was achieved over a wide range of relay locations by the constant-power relaying in conjunction with the detector based on the minimum product of squared error. Although the optimal power allocation for the systems using noncoherent detection is difficult to determine, numerical results demonstrated that performance gains can be achieved by using power allocation strategies developed for the coherent cases.

7.2 Future Work

Some possible avenues for further research could be:

- Power allocation for cooperative diversity with noncoherent detection. It has been seen that the system benefits from the diversity gains of using multiple relays at high enough SNR, but suffers from the noncoherent penalties at low SNR. It would be useful to determine the optimal number of relays for a particular SNR. It was also observed that the simple power allocation designed for the coherent cases can also be used for noncoherent cases. Determination of the optimal power allocation among nodes is an interesting problem that needs further investigation.

- General Channel Models. Rayleigh fading channels was focused through this dissertation. It would be interesting to extend this work to more general channels models, for example, Nakagami fading or Rician fading.

APPENDIX A

ASYMPTOTIC INFORMATION RATE PENALTY FOR CLARKE'S DOPPLER SPECTRUM

A.1 Evaluation of $\widetilde{P}_{\Delta}$

Here the asymptotic (large block length) $\widetilde{P}_{\Delta}$ for Clarke's Doppler spectrum (2.5) is shown to be given by (3.12) and (3.13).

Substitution of (2.5) into (3.11) gives

$$\begin{aligned}
\widetilde{P}_{\Delta,\mathrm{C}} &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \ln\left(1+\rho H_{\mathrm{C}}\left(\omega\right)\right) d\omega \\
&= \frac{1}{2\pi}\int_{-2\pi f_D}^{2\pi f_D} \ln\left(1+\rho\frac{1}{\pi f_D\sqrt{1-(\omega/2\pi f_D)^2}}\right) d\omega \\
&= \int_{-f_D}^{f_D} \ln\left(1+\frac{\rho}{\pi f_D}\frac{1}{\sqrt{1-(f/f_D)^2}}\right) df
\end{aligned} \tag{A.1}$$

Change of variables $\sin\theta = f/f_D$ yields

$$\begin{aligned}
\widetilde{P}_{\Delta,\mathrm{C}} &= f_D\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \ln\left(1+\frac{\rho}{\pi f_D}\frac{1}{\cos\theta}\right)\cos\theta d\theta \\
&= 2f_D\phi\left(\frac{\rho}{\pi f_D}\right),
\end{aligned} \tag{A.2}$$

where

$$\phi\left(x\right) = \int_{0}^{\frac{\pi}{2}} \ln\left(1+x\frac{1}{\cos\theta}\right)\cos\theta d\theta,\ x\geq 0. \tag{A.3}$$

Let $K(\theta,x)$ denote the integrand defined for all pairs $(\theta,x)\in[0,\frac{\pi}{2})\times[0,\infty)$, then, clearly,

$$\frac{\partial}{\partial x}K\left(\theta,x\right) = 1-\frac{x}{x+\cos\theta} \tag{A.4}$$

is continuous over the region $[0, \frac{\pi}{2}) \times [0, \infty)$. Therefore,

$$\begin{aligned}
\frac{d}{dx}\phi(x) &= \int_0^{\frac{\pi}{2}} \frac{\partial}{\partial x} K(\theta, x) d\theta \\
&= \frac{\pi}{2} - x \int_0^{\frac{\pi}{2}} \frac{1}{x + \cos\theta} d\theta \\
&= \begin{cases} \frac{\pi}{2} - \frac{x}{\sqrt{1-\rho^2}} \ln\left(\frac{1+\sqrt{1-x^2}}{x}\right), & x < 1 \\ \frac{\pi}{2} - 1, & x = 1 \\ \frac{\pi}{2} - \frac{2x}{\sqrt{x^2-1}} \tan^{-1}\left(\sqrt{\frac{x-1}{x+1}}\right), & x > 1, \end{cases}
\end{aligned} \tag{A.5}$$

where the integral ([83, 2.553.3]) has been used.

$$\begin{aligned}
&\int \frac{1}{a + b\cos x} dx \\
&= \begin{cases} \frac{2}{\sqrt{a^2-b^2}} \tan^{-1}\left(\frac{\sqrt{a^2-b^2}\tan(x/2)}{a+b}\right), & a^2 > b^2 \\ \frac{1}{\sqrt{b^2-a^2}} \ln\left(\frac{\sqrt{b^2-a^2}\tan(x/2)+a+b}{\sqrt{b^2-a^2}\tan(x/2)-a-b}\right), & a^2 < b^2 \end{cases}
\end{aligned} \tag{A.6}$$

Since $\phi(x) = \int \frac{d}{dx}\phi(x)dx + C$, where C can be determined directly from (A.3), $\phi(1) = \frac{\pi}{2} - \ln 2$, and the indefinite integrals

$$\int \frac{2x}{\sqrt{x^2-1}} \tan^{-1}\left(\sqrt{\frac{x-1}{x+1}}\right) dx = -\ln x + 2\sqrt{x^2-1}\tan^{-1}\left(\sqrt{\frac{x-1}{x+1}}\right), \; x > 1 \tag{A.7}$$

$$\int \frac{x}{\sqrt{1-x^2}} \ln\left(\frac{1+\sqrt{1-x^2}}{x}\right) dx = \ln x + \sqrt{1-x^2}\ln\left(\frac{1+\sqrt{1-x^2}}{x}\right), \; x < 1 \tag{A.8}$$

resulting in the expression (3.13).

A.2 Evaluation of Asymptotic Ratio $\widetilde{P}_{\Delta}/C_{\text{Rayleigh}}$

In this section the ratio $\widetilde{P}_{\Delta}/C_{\text{Rayleigh}}$ is evaluated for Clarke's Doppler spectrum for large SNR.

From (3.12), (3.38) and (A.5),

$$\lim_{\rho\to\infty} \frac{\widetilde{P}_{\Delta,\text{C}}}{C_{\text{Rayleigh}}} \tag{A.9a}$$

$$= \lim_{\rho\to\infty} \frac{2f_D\phi(\rho\frac{1}{\pi f_D})}{\ln(1+\rho)-\gamma} \tag{A.9b}$$

$$= \lim_{\rho\to\infty} \frac{\frac{2}{\pi}\frac{d}{dx}\phi(x)|_{x=\rho/\pi f_D}}{1/(1+\rho)} \tag{A.9c}$$

$$= \lim_{\rho\to\infty} \left[\left(1-\frac{4}{\pi}\sqrt{\frac{(\rho/\pi f_D)^2}{(\rho/\pi f_D)^2-1}}\tan^{-1}\left(\sqrt{\frac{\rho-\pi f_D}{\rho+\pi f_D}}\right)\right)\right. \tag{A.9d}$$

$$\left.+\rho\left(1-\frac{4}{\pi}\sqrt{\frac{(\rho/\pi f_D)^2}{(\rho/\pi f_D)^2-1}}\tan^{-1}\left(\sqrt{\frac{\rho-\pi f_D}{\rho+\pi f_D}}\right)\right)\right] \tag{A.9e}$$

$$= \lim_{\rho\to\infty} \frac{\sqrt{1-\left(\frac{\pi f_D}{\rho}\right)^2}-\frac{4}{\pi}\tan^{-1}\left(\sqrt{\frac{\rho-\pi f_D}{\rho+\pi f_D}}\right)}{\sqrt{1-\left(\frac{\pi f_D}{\rho}\right)^2}/\rho} \tag{A.9f}$$

$$= \lim_{\rho\to\infty} \frac{\left[1-\left(\frac{\pi f_D}{\rho}\right)^2\right]^{-1/2}\left(\frac{\pi f_D}{\rho}\right)^3-\frac{4}{\pi}\left[1-\left(\frac{\pi f_D}{\rho}\right)^2\right]^{-1/2}\frac{\pi f_D}{2\rho^2}}{\left[1-\left(\frac{\pi f_D}{\rho}\right)^2\right]^{-1/2}\left[\frac{1}{\rho}\left(\frac{\pi f_D}{\rho}\right)^3-\frac{1}{\rho^2}\right]} \tag{A.9g}$$

$$= 2f_D, \tag{A.9h}$$

where the l'Hospital's rules have been used in (A.9c) and (A.9g).

APPENDIX B

PERFORMANCE OF CONSTANT-AMPLIFICATION RELAYING OVER RAYLEIGH FADING CHANNELS

B.1 Average SNR and Amount of Fading

For Rayleigh fading, $|g_{ij}|^2$ are exponentially distributed with parameter $1/G_{ij}$. Therefore the probability density functions of U and V are respectively

$$\begin{aligned} f_U(u) &= \frac{1}{G_{12}} \exp\left(-\frac{u}{G_{12}}\right), u > 0 \\ f_V(v) &= \frac{1}{G_{23}} \exp\left(-\frac{v}{G_{23}}\right), v > 0. \end{aligned} \tag{B.1}$$

Let $a = G_{12}/\alpha + 1/(\alpha\rho)$. Then it follows from (5.8) and (B.1) that

$$\begin{aligned} \mathbb{E}\left[S_2^{\mathrm{CA}}\right] &= \mathbb{E}[U]\,\mathbb{E}\left[\frac{V}{V+a}\right]\rho \\ &= G_{12}\rho\left[1 - \int_0^\infty \frac{a}{v+a}\frac{1}{G_{23}} \exp\left(-\frac{v}{G_{23}}\right) dv\right] \\ &= G_{12}\rho\left[1 - \frac{a}{G_{23}} \exp\left(\frac{a}{G_{23}}\right) E_1\left(\frac{a}{G_{23}}\right)\right] \end{aligned} \tag{B.2}$$

and

$$\begin{aligned} \mathbb{E}\left[\left(S_2^{\mathrm{CA}}\right)^2\right] &= \mathbb{E}\left[U^2\right]\mathbb{E}\left[1 - \frac{2a}{V+a} + \frac{a^2}{(V+a)^2}\right] \\ &= 2\left(G_{12}\right)^2\rho^2\left[1 - \frac{a}{G_{23}} \exp\left(\frac{a}{G_{23}}\right)\left[2E_1\left(\frac{a}{G_{23}}\right) - E_2\left(\frac{a}{G_{23}}\right)\right]\right] \end{aligned} \tag{B.3}$$

where E_n is the exponential integral function

$$E_n(x) = \int_1^\infty t^{-n} \exp(-xt)\, dt. \tag{B.4}$$

Substitution of $a = G_{12}/\alpha + 1/(\alpha\rho)$ into (B.2) leads to the expression of average SNR (5.12). The amount of fading (5.21) follows directly from (5.20), (B.2), and (B.3).

B.2 Outage Probability

The distribution function for S_c is

$$\begin{aligned}\Pr[S_c < t] &= \Pr[S_2 < t - S_1] \\ &= \int_0^t \Pr[S_2 < x] \frac{1}{G_{13}\rho} \exp\left(-\frac{t-x}{G_{13}\rho}\right) dx, \end{aligned} \tag{B.5}$$

where for constant-amplification relaying $\Pr[S_2 < x]$ can be evaluated by using (5.8) and (B.1):

$$\begin{aligned}\Pr\left[S_2^{\mathrm{CA}} < x\right] &= \Pr\left[U < x\left(\frac{1}{\rho} + \frac{\rho G_{12}+1}{\alpha\rho^2}\frac{1}{V}\right)\right] \\ &= \int_0^\infty \left[1 - \exp\left(-\frac{x}{G_{12}}\left(\frac{1}{\rho} + \frac{\rho G_{12}+1}{\alpha\rho^2}\frac{1}{v}\right)\right)\right] \frac{1}{G_{23}} \exp\left(-\frac{v}{G_{23}}\right) dv \\ &= 1 - \exp\left(-\frac{x}{G_{12}\rho}\right) \sqrt{\frac{4(\rho G_{12}+1)x}{\alpha\rho^2 G_{12}G_{23}}} K_1\left(\sqrt{\frac{4(\rho G_{12}+1)x}{\alpha\rho^2 G_{12}G_{23}}}\right), \end{aligned} \tag{B.6}$$

where the last equality follows from the integral in [83, 3.324.1], and where the function K_1 denotes the first order modified Bessel function of the second kind.

To find an approximation expression that provides more insight on the behavior at high SNR, the following series representations can be used [83, 8.446, 1.211.1]

$$\begin{aligned} K_1(x) &= \frac{1}{x} + \frac{x}{2}\ln x + \frac{x}{2}\left(\gamma - \frac{1}{2} - \ln 2\right) + O\left(x^3 \ln x\right) \\ \exp(x) &= 1 + x + O\left(x^2\right) \end{aligned}$$

where $\gamma = 0.577\cdots$ denotes Eüler's constant. It follows from (B.5) and (B.6) (for large ρ)

$$\begin{aligned} &\Pr[S_c < t] \\ = \; & \frac{1}{G_{13}\rho}\int_0^t \left(1 - \frac{t}{G_{13}\rho} + \frac{v}{G_{13}\rho}\right) \\ & \left[-\frac{(G_{12}\rho+1)}{\alpha G_{23}G_{12}\rho^2} v\ln v + \left[\frac{1}{G_{12}\rho} - \frac{(G_{12}\rho+1)}{\alpha G_{23}G_{12}\rho^2}\left(\ln\frac{(G_{12}\rho+1)}{\alpha G_{23}G_{12}\rho^2} + 2\gamma - 1\right)\right] v\right] dv + O\left(\rho^{-3}\ln\rho\right) \\ = \; & \frac{t^2}{2G_{13}\rho}\frac{(G_{12}\rho+1)}{\alpha G_{23}G_{12}\rho^2}\left(\ln\frac{\alpha G_{23}G_{12}\rho^2}{G_{12}\rho+1}\right) + \\ & \frac{t^2}{2G_{13}\rho}\left[\frac{1}{G_{12}\rho} - \frac{(G_{12}\rho+1)}{\alpha G_{23}G_{12}\rho^2}\left(2\gamma - \frac{3}{2} + \ln t\right)\right] + O\left(\rho^{-3}\ln\rho\right). \end{aligned} \tag{B.7}$$

REFERENCES

[1] B. M. Hochwald and T. L. Marzetta, "Unitary space-time modulation for multiple-antenna communications in Rayleigh flat fading," *IEEE Trans. Inform. Theory*, vol. 46, no. 2, pp. 543–564, Mar. 2000.

[2] R. H. Clarke, "A statistical theory of mobile-radio reception," *Bell. Syst. Tech. J.*, vol. 47, no. 6, pp. 957–1000, July/Aug. 1968.

[3] H. J. Landau and H. O. Pollak, "Prolate spheroidal wave functions, Fourier analysis and uncertainty-III: The dimension of the space of essentially time- and band-limited signals," *Bell Syst. Tech. J.*, vol. 41, no. 4, pp. 1295–1336, July 1962.

[4] H. Meyr, M. Moeneclaey, and S. Fechtel, *Digital Communication Receivers: Synchronization, Channel Estimation and Signal Processing*. New York: Wiley, 1998.

[5] E. Biglieri, J. Proakis, and S. Shamai (Shitz), "Fading channels: Information-theoretic and communications aspects," *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2619–2692, Oct. 1998.

[6] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1319–1343, June 2002.

[7] I. E. Telatar and D. N. C. Tse, "Capacity and mutual information of wideband multipath fading channels," *IEEE Trans. Inform. Theory*, vol. 46, no. 4, pp. 1384–1400, July 2000.

[8] M. Médard and R. G. Gallager, "Bandwidth scaling for fading multipath channels," *IEEE Trans. Inform. Theory*, vol. 48, no. 4, pp. 840–852, Apr. 2002.

[9] V. G. Subramanian and B. Hajek, "Broad-band fading channels: Signal burstiness and capacity," *IEEE Trans. Inform. Theory*, vol. 48, no. 4, pp. 809–827, Apr. 2002.

[10] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Trans. Inform. Theory*, vol. 45, no. 1, pp. 139–157, Jan. 1999.

[11] L. Zheng and D. N. C. Tse, "Communication on the Grassmann manifold: A geometric approach to the noncoherent multiple-antenna channel," *IEEE Trans. Inform. Theory*, vol. 48, no. 2, pp. 359–383, Feb. 2002.

[12] A. J. Viterbi, "Performance of an M-ary orthogonal communication system using stationary stochastic signals," *IEEE Trans. Inform. Theory*, vol. 13, no. 3, pp. 414–422, July 1967.

[13] I. C. Abou-Faycal, M. D. Trott, and S. Shamai (Shitz), "The capacity of discrete-time memoryless Rayleigh-fading channels," *IEEE Trans. Inform. Theory*, vol. 47, no. 4, pp. 1290–1301, May 2001.

[14] G. Taricco and M. Elia, "Capacity of fading channels with no side information," *Electron. Lett.*, vol. 33, no. 16, pp. 1368–1370, July 1997.

[15] S. Shamai (Shitz) and T. L. Marzetta, "Multiuser capacity in block fading with no channel state information," *IEEE Trans. Inform. Theory*, vol. 48, no. 4, pp. 938–942, Apr. 2002.

[16] M. Godavarti, T. L. Marzetta, and S. Shamai (Shitz), "Capacity of a mobile multiple-antenna wireless link with isotropically random Rician fading," *IEEE Trans. Inform. Theory*, vol. 49, no. 12, pp. 3330–3334, Dec. 2003.

[17] R.-R. Chen, R. Koetter, U. Madhow, and D. Agrawal, "Joint noncoherent demodulation and decoding for the block fading channel: A practical framework for approaching Shannon capacity," *IEEE Trans. Commun.*, vol. 51, no. 10, pp. 1676–1689, Oct. 2003.

[18] M. Peleg and S. Shamai (Shitz), "On the capacity of blockwise incoherent MPSK channel," *IEEE Trans. Commun.*, vol. 46, no. 5, pp. 603–609, May 1998.

[19] G. Colavolpe and R. Raheli, "The capacity of noncoherent channels," in *Proc. IEEE Int. Conf. Communications (ICC)*, June 1999.

[20] A. J. Goldsmith and P. P. Varaiya, "Capacity, mutual information, and coding for finite-state Markov channels," *IEEE Trans. Inform. Theory*, vol. 42, no. 3, pp. 898–886, May 1996.

[21] C. Komninakis and R. D. Wesel, "Joint iterative channel estimation and decoding in flat correlated Rayleigh fading," *IEEE J. Select. Areas Commun.*, vol. 19, no. 9, pp. 1706–1717, Mar. 2001.

[22] M. Médard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel," *IEEE Trans. Inform. Theory*, vol. 46, no. 3, pp. 933–946, May 2000.

[23] A. Lapidoth and S. Shamai (Shitz), "Fading channels: How perfect need "perfect side information" be?" *IEEE Trans. Inform. Theory*, vol. 48, no. 5, pp. 1118–1134, May 2002.

[24] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Trans. Inform. Theory*, vol. 49, no. 4, pp. 951–963, Apr. 2003.

[25] X. Deng and A. M. Haimovich, "On pilot symbol aided channel estimation for time varying Rayleigh fading channels," in *Proc. 38th Annual Conf. Information Sciences and Systems (CISS)*, Princeton, New Jersey, USA, Mar. 2004.

[26] J. Baltersee, G. Fock, and H. Meyr, "An information theoretic foundation of synchronized detection," *IEEE Trans. Commun.*, vol. 49, no. 12, pp. 2115–2123, Dec. 2001.

[27] A. Bdeir, I. Abou-Faycal, and M. Médard, "Power allocation schemes for pilot symbol assisted modulation over Rayleigh fading channels with no feedback," in *Proc. IEEE Int. Conf. Commun. (ICC)*, France, Paris, June 2004.

[28] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, no. 3, pp. 311–335, Mar. 1998.

[29] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. Telecomm.*, vol. 10, no. 6, pp. 585–595, Nov./Dec. 1999.

[30] Z. B. Baranski, A. M. Haimovich, and J. Garcia-Frias, "Em-based iterative receiver for space-time coded modulation with noise covariance estimation," in *Proc. IEEE Global Telecomm. Conf. (GLOBECOM)*, Nov. 2002.

[31] E. Chiavaccini and G. M. Vitetta, "MAP symbol estimation on frequency-flat Rayleigh faidng channels via a Bayesian EM algorithm," *IEEE Trans. Commun.*, vol. 49, no. 11, pp. 1869–1872, Nov. 2001.

[32] C. Cozzo and B. L. Hughes, "An iterative receiver for space-time communications," in *Proc. Conference on Information Sciences and Systems (CISS)*, Mar. 2000.

[33] A. Grant, "Joint decoding and channel estimation for space-time codes," in *Proc. IEEE Veh. Tech. Conf. (VTC)*, 2000.

[34] Y. Li, C. N. Georghiades, and G. Huang, "Iterative maximum-likelihood sequence estimation for space-time coded systems," *IEEE Trans. Commun.*, vol. 49, no. 6, pp. 948–951, June 2001.

[35] H.-J. Su and E. Geraniotis, "Low-complexity joint iterative channel estimation and decoding for pilot symbol-assisted modulation and multiple differential detection systems with correlated Rayleigh fading," *IEEE Trans. Commun.*, vol. 50, no. 2, pp. 249–261, Feb. 2002.

[36] M. C. Valenti and B. D. Woerner, "Iterative channel estimation and decoding of pilot symbol assisted turbo codes over flat-fading channels," *IEEE J. Select. Areas Commun.*, vol. 19, no. 9, pp. 1697–1705, Sept. 2001.

[37] X. Deng, A. M. Haimovich, and J. Garcia-Frias, "Decision directed iterative channel estimation for MIMO systems," in *Proc. IEEE Int. Conf. Commun.(ICC)*, Anchorage, AK, USA, May 2003.

[38] H. Chen, X. Deng, and A. M. Haimovich, "Layered turbo space-time coded MIMO-OFDM systems for time varying channels," in *Proc. IEEE Global Telecomm. Conf. (GLOBECOM)*, San Francisco, CA, USA, Dec. 2003.

[39] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory.* Upper Saddle River, NJ: Prentice Hall, 1993.

[40] D. Divsalar and M. K. Simon, "Multiple-symbol differential detection of MPSK," *IEEE Trans. Commun.*, vol. 38, no. 3, pp. 300–308, Mar. 1990.

[41] B. M. Hochwald, T. L. Marzetta, T. J. Richardson, W. Sweldens, and R. Urbanke, "Systematic design and unitary space-time constellations," *IEEE Trans. Inform. Theory*, vol. 46, no. 6, pp. 1962–1973, Sept. 2000.

[42] M. K. Simon and M.-S. Alouini, "Multiple symbol differential detection with diversity reception," *IEEE Trans. Commun.*, vol. 49, no. 8, pp. 1312–1319, Aug. 2001.

[43] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperative diversity - Part I: System description," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927–1938, Nov. 2003.

[44] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inform. Theory*, no. 12, pp. 3062–3080, Dec. 2004.

[45] M. O. Hasna and M.-S. Alouini, "A performance study of dual-hop transmissions with fixed gain relays," *IEEE Trans. Wireless Commun.*, vol. 3, no. 6, pp. 1963–1968, Nov. 2004.

[46] I. Hammerström, M. Kuhn, and A. Wittneben, "Cooperative diversity by relay phase rotation in block fading enviroments," in *Proc. IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Lisboa, Portugal, July 2004.

[47] J. Luo, R. S. Blum, L. J. Greenstein, L. J. Cimini, and A. M. Haimovich, "New approaches for cooperative use of multiple antennas in ad hoc wireless networks," in *Proc. IEEE Veh. Tech. Conf. (VTC)*, Los Angeles, CA, USA, Sept. 2004.

[48] A. El Gamal and S. Zahedi, "Minimum energy communication over a relay channel," in *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, Yokohama, Japan, June/July 2003.

[49] I. Maric and R. D. Yates, "Cooperative multihop broadcast for wireless networks," *IEEE J. Select. Areas Commun.*, vol. 22, no. 6, pp. 1080–1088, Aug. 2004.

[50] Z. Jingmei, Z. Qi, S. Chunju, W. Ying, Z. Zhang, and P. Zhang, "Adaptive optimal transmit power allocation for two-hop non-regenerative wireless relaying system," in *Proc. IEEE Veh. Tech. Conf. (VTC)*, Milan, Italy, May 2004.

[51] I. Hammerström, M. Kuhn, and A. Wittneben, "Impact of relay gain allocation on the performance of cooperative diversity networks," in *Proc. IEEE Veh. Tech. Conf. (VTC)*, Los Angeles, CA, USA, Sept. 2004.

[52] M. O. Hasna and M.-S. Alouini, "Optimal power allocation for relayed transmissions over Rayleigh-fading channels," *IEEE Trans. Wireless Commun.*, vol. 3, no. 6, pp. 1999–2004, Nov. 2004.

[53] J. Luo, R. S. Blum, L. J. Cimini, L. J. Greenstein, and A. M. Haimovich, "Power allocaiton in a transmit diversity system with mean channel gain information," submitted for publication, 2004.

[54] X. Deng and A. M. Haimovich, "Power allocation for cooperative relaying in wireless networks," submitted for publication, 2005.

[55] P. A. Anghel and M. Kaveh, "Exact symbol error probability of a cooperative network in a Rayleigh-fading enviroment," *IEEE Trans. Wireless Commun.*, vol. 3, no. 5, pp. 1416–1421, Sept. 2004.

[56] J. Boyer, D. Falconer, and H. Yanikomeroglu, "A theoretical characterization of the multihop wireless communications channel with diversity," in *Proc. Global Telecomm. Conf. (GLOBECOM)*, San Antonio, Texas, USA, Nov. 2001.

[57] M. O. Hasna and M.-S. Alouini, "End-to-end performance of transmission systems with relays over Rayleigh-fading channels," *IEEE Trans. Wireless Commun.*, vol. 2, no. 6, pp. 1126–1131, Nov. 2003.

[58] W. Mo and Z. Wang, "Average symbol error probability and outage probability analysis for general cooperative diversity system at high signal to noise ratio," in *Proc. Conf. Information Sciences and Systems (CISS)*, Princeton, New Jersey, USA, Mar. 2004.

[59] A. Ribeiro, X. Cai, and G. B. Giannakis, "Symbol error probabilities for general cooperative links," in *Proc. Int. Conf. Commun. (ICC)*, Paris, France, June 2004.

[60] D. Chen and J. N. Laneman, "Noncoherent demodulation for cooperative diversity in wireless systems," in *Proc. IEEE Global Telecomm. Conf. (GLOBECOM)*, Dallas, TX, USA, Nov. 2004.

[61] E. Zimmermann, P. Herhold, and G. Fettweis, "On the performance of cooperative diversity proctocols in practical wireless systems," in *Proc. IEEE Veh. Tech. Conf. (VTC)*, Orlando, Florida, USA, Oct. 2003.

[62] X. Deng and A. M. Haimovich, "Cooperative relaying in wireless networks with local channel state information," in *Proc. IEEE Veh. Tech. Conf. (VTC)*, Dallas, TX, USA, Sept. 2005.

[63] Y. Liang and V. V. Veeravalli, "Capacity of noncoherent time-selective block Rayleigh flat-fading channel," in *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, Lausanne, Switzerland, June/July 2002.

[64] A. Lapidoth, "On the high SNR capacity of stationary Gaussian fading channels," in *Proc. 41st Annual Allerton Conf. Communication, Control, and Computing*, Monticello, IL, Oct. 2003.

[65] X. Deng and A. M. Haimovich, "Achievable rates over time-varying Rayleigh fading channels," submitted for publication, 2004.

[66] K. S. Miller, *Multidimensional Gaussian Distributions*. New York: John Wiley & Sons, 1964.

[67] T. Cover and J. Thomas, *Elements of Information Theory*. New York: John Wiley & Sons, 1991.

[68] X. Deng and A. M. Haimovich, "Information rates of time varying Rayleigh fading channels," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Paris, France, June 2004.

[69] U. Grenander and G. Szegö, *Toeplitz Forms and their Applications*. Berkeley and Los Angeles: University of California Press, 1958.

[70] W. C. Y. Lee, "Estimation of channel capacity in Rayleigh fading environment," *IEEE Trans. Veh. Technol.*, vol. 39, no. 3, pp. 187–189, Aug. 1990.

[71] L. H. Ozarow, S. Shamai (Shitz), and A. D. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Trans. Veh. Technol.*, vol. 43, no. 2, pp. 359–378, May 1994.

[72] A. Lapidoth and P. Narayan, "Reliable communication under channel uncertainty," *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2148–2177, Oct. 1998.

[73] L. Zheng, D. N. C. Tse, and M. Médard, "Channel coherence in the low SNR regime," in *Proc. 38th Annual Conf. Information Sciences and Systems (CISS)*, Princeton, New Jersey, USA, Mar. 2004.

[74] T. F. Wong, "Numerical calculation of symmetric capacity of Rayleigh fading channel with BPSK/QPSK," *IEEE Commun. Lett.*, vol. 5, no. 8, pp. 328–330, Aug. 2001.

[75] A. Greenbaum, *Iterative Methods for Solving Linear Systems*. SIAM, 1997.

[76] L. A. Hageman and D. M. Young, *Applied Iterative Methods*. Academic Press, 1981.

[77] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, no. 2, pp. 744–765, Mar. 1998.

[78] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inform. Theory*, pp. 284–287, Mar. 1974.

[79] M. K. Simon and M.-S. Alouini, *Digital Communication over Fading Channels: A Unified Approach to Performance Analysis*. New York: John Wiley & Sons, 2000.

[80] U. Charash, "Reception through Nakagami fading multipath channels with random delays," *IEEE Trans. Commun.*, vol. 27, no. 4, pp. 657–670, Apr. 1979.

[81] D. Chen and J. N. Laneman, "Cooperative diversity for wireless fading channels without channel state information," in *Proc. 38th Asilomar Conference on Signals, Systems and Computers*, Monterey, CA, USA, Nov. 2004.

[82] D. Lao and A. M. Haimovich, "Multiple-symbol differential detection with interference suppression," *IEEE Trans. Commun.*, vol. 51, no. 2, pp. 208–217, Feb. 2003.

[83] I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series, and Products*. New York: Academic Press, 1994.