CHANNEL CHARACTERIZATION AND SYSTEM DESIGN IN WIRELESS COMMUNICATIONS

Do-Sik Yoo

COMMUNICATIONS & SIGNAL PROCESSING LABORATORY
Department of Electrical Engineering and Computer Science
The University of Michigan
Ann Arbor, MI 48109-2122

February 2002

Acknowledgement

Preface

We introduce parameters called *normalized mean square covariance (NMSV)*, *normalized mean square correlation (NMSR)*, *re-centered normalized mean square correlation (RC-NMSR)*, and *stochastic degree of freedom (SDF)* to efficiently characterize the correlation properties of L^2 stochastic processes. We show that these parameters are very useful to characterize wireless communication channels. In particular, we show that there are very close relationships between the parameters and the performance of various wireless communication systems. Due to such close relationship, the parameters are very useful to evaluate the quality of wireless communication channels. We show that the parameters are very effective in wireless communication system design. In particular, we can obtain direct relations between the quality of channel and the performance of various wireless communication systems. Due to the fundamental nature of the parameters, we believe that they will be very helpful in other branches of science and engineering in which stochastic modeling is useful.

Contents

Chapter 1

Introduction

The invention of personal computers has led us to a revolutionary state of transition to an information age. Strong demand for fast exchange of information is driving the rapid developments of high-speed communication networks. Currently, most information is transmitted over wired networks. However, because of the portability, the role of wireless communications is growing very rapidly. While increasing demand in wired networks can be satisfied by installing new wires, this is fundamentally impossible in wireless communications where the same physical medium has to be shared by many users. Consequently, it is particularly important to develop various technologies for optimal utilization of the limited resources in wireless communications. Since the performance depends fundamentally on the channels under consideration, a communication system design must be preceded by the study of channel characteristics. In Section 1.1, we discuss various previous and current efforts to characterize radio propagation channels and the motivation for the research in this report. Then we briefly describe the major contributions of this report in Section 1.2. In the final section, we outline the organization of the report.

1.1 Radio Propagation Channel Characterization

In modern wireless communications, digital information signals are divided into small size packets and then transmitted. Each packet is usually encoded (by source and channel encoders), modulated, amplified and then carried by high-frequency electromagnetic waves over a radio link. At the receiver side, the reverse processes are done by a receiver antenna with a radio frequency tuner, a demodulator, a channel decoder and a source decoder. In this report, we are primarily interested in the characterization of radio links between the transmitter and the receiver antennae that will

be modeled by randomly time-variant linear systems. The major purpose of channel characterization is in the design and planning of communication systems. Wireless communication channels are usually described by considering three separable phenomena, namely, path loss, shadowing, and multipath fading. In the following, we briefly overview various efforts to characterize such aspects of wireless communication channels and describe why we are interested in parametric characterizations of multipath fading.

The estimation of signal decay due to propagation loss is very important in the determination of the necessary transmission power and the coverage area. There have been various efforts to effectively estimate the propagation loss in various environments for wide-range of carrier frequency. The propagation loss can be predicted in principle by considering various fundamental propagation mechanisms such as reflection, diffraction and scattering given enough information about the topographical situations. The sufficiency of the information depends primarily on the complexity of the terrain structures, the wavelength of the carrier, and the height of the antenna. The computational requirements in most practical situations are usually prohibitively high for purely analytical methods. Such high computational requirements can generally be reduced by judicious use of empirical results. Some of the most well-known propagation loss prediction models are the Longley-Rice and its modified models [1, 2, 3], Empirical Propagation Model-73 (EPM-73) [4], the Okumura model [5], and the Hata model [6].

Although propagation loss models are sometimes quite accurate, it generally fails to predict signal fluctuations due to the effect of the terrain near the antenna. Such a phenomenon of signal fluctuations is usually called shadowing. In modern wireless communications, the effect of shadowing is usually compensated in the network layer by power control and/or rate adjustment. For the evaluation of such technologies, statistical description of the shadowing loss by a log-normal distribution provides useful insights and effective analytical channel models. In such a model, the severity of shadowing is characterized by the logarithmic standard deviation, which depends on the size of the obstacles relative to the carrier wavelength.

The result of propagation loss or shadowing can be represented by a single parameter, namely by the loss of power. Consequently, it is relatively easy to determine appropriate strategy to employ to counteract the effect of propagation loss or shadowing. For example, we can either increase the transmitted power or shrink the coverage area with increased propagation loss. However, the effect of multipath fading is generally more complex because it not only changes in time but also varies over frequency.

Hence, a time-variant linear system model [7] is generally employed to describe multipath fading.

One of the most salient characteristics of multipath fading is the unpredictability. Due to the frequency and time selectivity, multipath fading directly affects the effectiveness of packet design techniques such as modulation, interleaving, and channel coding. Not only because of the practical unpredictability but also because of the usefulness in performance evaluations, multipath fading channels are usually modeled as randomly time-variant linear systems [8]. For such stochastic models, the first and the second order statistics play fundamental roles in the channel characterizations. However, such statistics are generally complicated two dimensional functions in most situations. Consequently, unlike the characterization of propagation loss or shadowing, characterization of multipath fading requires an infinite degree of freedom in the stochastic description of multipath fading channel. This can be a severe burden to system design. This is one reason why various simple secondary statistics such as root mean square (rms) delay spread are often used to characterize the quality of multipath fading channels.

To be useful, parameters that attempt to characterize a channel should be closely related to the system performance. The rms delay spread is meaningful not only because it concisely summarizes the physical situations but also because it is related to the frequency selectivity of a channel. Due to convexity of the error function, the performance of a communication system generally depends heavily on the amount of signal strength fluctuations. To reduce the signal strength variations, various diversity combining schemes are widely adopted that exploit the frequency and time selectivity of the fading stochastic process. Consequently, the performance of such a system depends fundamentally on the amount of frequency and time selectivity.

Other parameters such as coherence bandwidth or correlation time are also widely used to characterize channel frequency or time selectivity. However, as will be seen, all these existing parameters are poorly related to the performance of various existing communication systems. This is because the parameters are local in nature. For example, correlation time characterizes the correlation property in the neighborhood of the time of interest. However, if the correlation time is relatively small compared to the packet length, it fails to represent the overall time selectivity of the channel over the packet duration. In this report, we propose parameters that effectively characterize the overall frequency and/or time selectivity of fading stochastic process in a frequency and/or time region of interest.

1.2   Report Overview

In this section, we briefly summarize the the main contributions of this report. In Section 1.2.1, we describe the fundamental ideas of normalized mean square covariance (NMSV) and stochastic degree of freedom (SDF). These parameters are used to characterize wide-sense stationary uncorrelated scattering (WSSUS) channels. We describe how we characterize non-WSSUS channels with simple parameters in Section 1.2.2. In Section 1.2.3, we discuss the importance of the proposed parameters in system design. Next, we describe our contribution to multiuser detection researches in Section 1.2.4.

1.2.1   Characterization of WSSUS Channels

When the direct line of sight is blocked by obstacles in wireless communications, the channel can often be modeled effectively as a wide-sense stationary uncorrelated scattering channel [8, 9]. Because of the lack of a reliable line of sight signal path, such channels usually represent the worst case situation. To guarantee desired quality of service under all situations, it is particular important to study such adverse situations. Consequently, we first study WSSUS channels.

As noted in the previous section, we start from the fact that most existing channel parameters are local in nature. For example, the coherence bandwidth measures the extent over which the channel frequency response is essentially constant from a particular reference point and hence tends to lose its validity as a measure of frequency selectivity as the total system bandwidth increases. We define a simple parameter called the normalized frequency mean square covariance (NFMSV) that characterizes the overall frequency selectivity of a fading stochastic process in a frequency region of interest. Similarly, we define the normalized time mean square covariance (NTMSV) and the normalized frequency-time mean square covariance (NFTMSV) for the time selectivity and the combined frequency-time selectivity, respectively. We refer to these parameters collectively as the normalized mean square covariance (NMSV).

To see the meaning of the NMSV, consider a set $\{X_1, X_2, \cdots, X_n\}$ of real random variables with zero mean. Then, the NMSV $\mathcal{V}$ of the set of random variables is defined by

$$\mathcal{V} = \frac{\displaystyle\sum_{k=1}^{n}\sum_{l=1}^{n} E^2[X_k X_l]}{\left(\displaystyle\sum_{k=1}^{n} E[X_k^2]\right)^2}. \tag{1.1}$$

It is shown in Appendix 2.4 that the number $1/\mathcal{V}$ is closely related to the effective

number of distinguishable random variables in the set. For example, if the n random variables are uncorrelated and have the same variance, $\mathcal{V} = 1/n$ while if all n random variables are perfectly correlated, $\mathcal{V} = 1$. Based on the result, we define the stochastic degree of freedom (SDF) $\mathcal{F}$ of the set $\{X_1, X_2, \cdots, X_n\}$ of random variables by

$$\mathcal{F} = \frac{1}{\mathcal{V}}. \tag{1.2}$$

The NMSV of a channel is defined by the NMSV of the time-variant frequency response of the channel. We show that there are very close relationships between the NMSV and the performances of diversity combining schemes. From the analysis of simple diversity combining scheme, we interpret SDF of a channel as the degree of freedom available for various diversity combining schemes. Since diversity combining plays fundamental roles in communication system design over WSSUS channels, we can regard the NMSV and hence the SDF are among the most important parameters of WSSUS channels.

1.2.2 Characterization of Non-WSSUS Channels

Characterization of such non-WSSUS channels is generally more complex. In particular, it is evident that the NMSV does not provide sufficient characterizations of non-WSSUS channels due to the presence of the line of sight signal path. The system functions of non-WSSUS channels are divided into the specular components and the scattering components (or diffuse components). We usually model the scattering components as WSSUS channels. Consequently, we will assume regard a non-WSSUS channel as a sum of a (deterministic) specular component and a WSSUS component. Since the specular components are mostly due to the line of sight signals, they consists of single strong path and a set of small often negligible paths. Consequently, the line of sight components can essentially be characterized by finding the fractions of the power in the specular components.

As in the case of WSSUS channels, the characterization of the scattering components is generally more difficult. For WSSUS channels, we started from the observation that existing parameters are local parameters, defined the NMSV's, and then showed that the new parameters are closely related to the system performances. Then we noted that the parameters are directly related to the idea of diversity combining. However, as will be shown in Chapter 5, the NMSV's do not provide a satisfactory characterization of non-WSSUS channels. For non-WSSUS, we start from the observation that the NMSV's are directly related to diversity combining techniques, define

parameters called the re-centered (RC) normalized mean square correlation (NMSR), and show that the parameters are closely related to the performance of the simple diversity combining schemes.

Consequently, we propose to characterize a non-WSSUS channel by specifying the fraction of power contained in the specular component and the RC-NMSR of the diffuse component. Sometimes the specular component of a non-WSSUS channel consists of many small strength paths rather than a single strong line of sight path. Such a situation happens when there is no line of sight path but the signals arriving at each particular delay come from some specific direction. We define the normalized arithmetic variance (NAV) to characterize the specular components in such cases. Hence, for the most general situations, we propose to use three parameters, namely, the fraction of power in the specular component, the NAV of the specular component, and the RC-NMSR of the diffuse components. We show how successful this specification is by simulation results for the simple diversity combining scheme.

1.2.3   Applications in Communication System Design

Since there exist very close relations between the parameters we defined and the performance of systems, the parameters are very useful in system design. First of all, the parameters are very useful in performance evaluation and comparison. Since the parameters are highly related to system performance, we can characterize the "goodness" or "badness" of the channel by the parameters. Consequently, it is possible to study the relation between the system performance and the degree of adversity of the channel through the parameters. This is particularly important since we can understand the effectiveness of a particular communication system design for multipath fading channels. For example, we study the problem of choosing suitable rate of frequency hopping in a FHSS system. It is well-known that more diversity gain is expected if we increase the rate of hopping. However, increasing hopping rate involves burdens such as synchronization and channel estimation. So we expect that there exists an optimal hopping rate, which depends on the channel characteristics. We demonstrate how we can use the parameters to determine optimal hopping rate.

More importantly the parameters can be used in resource allocation problems. As an example, we consider the problem of bandwidth allocation of a FHSS system under a WSSUS channel. In a frequency hopping system, a given band of frequency is divided into a number of slots. We show that we can exploit more SDF with the same amount of total bandwidth by separating the slots and assigning the bandwidth between slots to other systems. As the slot separation grows, we tend to have more SDF. Since it

is not possible to increase the slot separation indefinitely, it is important to find the optimal slot spacing. We demonstrate that the NMSV is very useful in solving this problem. Similar idea of frequency separation can be used for orthogonal frequency division multiplexing (OFDM) systems. We propose carrier-separated OFDM systems to support multiple access with OFDM technology. We show that we can achieve huge diversity gain with low complexity by judicious choice of carrier separation.

1.2.4   Multiple Access Interference Treatment

Another very important issue in current commercial wireless communications is the treatment of the multiple access interference (MAI). There have been great efforts to mitigate MAI by employing various interference treatment algorithms. The area of multiuser detection is so active and broad that it is not easy to list even the most important works. Sergio Verdu collected and summarized some of the most important multiuser algorithms for DS-CDMA systems in his recent book [10]. There are several ways to categorize multiuser detection. One way is to divide it into linear and non-linear schemes. Another is to group it into decision directed or non-decision directed algorithms.

In this report, we consider various decision directed algorithms in realistic multi-path fading environments. We first derive an Ungerboeck-type maximum likelihood sequence estimation algorithm for asynchronous DS-CDMA systems under frequency-time selective multipath fading. We then examine the matched-filter decision feed-back sequence estimation (MFDFSE) algorithm and propose a bias-compensated (BC) MFDFSE algorithm [11, 12]. It is shown that BCMFDFSE achieves huge a performance gain with relatively negligible increase in the system complexity compared to MFDFSE. However, we also observe that multistage successive interference cancellation can achieve a significant performance gain with minimal complexity increase from the conventional receiver. We study the effectiveness of single and multiple stage successive interference cancellation (SIC) and parallel interference cancellation (PIC) algorithms under multirate DS-CDMA environment. It is shown that multistage SIC performs very well under multirate environment.

1.3   Report Organization

The remainder of this report is organized as follows. In Chapter 2, we define various parameters for general stochastic processes and study their properties. We start the chapter with brief remarks on the mathematical descriptions. Next we summa-

rize terminologies in probability and statistics. In particular, various notations are introduced to be used consistently throughout this report. We then define various parameters called normalized arithmetic variance (NAV), normalized mean square mean (NMSM), normalized mean square covariance (NMSV), normalized mean square correlation (NMSR), and re-centered normalized mean square correlation (RC-NMSR). In the last section, we study the meaning of the NMSV and define the concept of stochastic degree of freedom (SDF).

In Chapter 3, we describe wireless communication channels that we study in this report. In this report, we confine ourselves to the phenomenon of multipath fading. In the first two sections, we discuss why multipath fading happens and why we can treat it separately from the path loss and the shadowing. Then, we introduce various terminologies to describe randomly time-variant linear systems which we use to describe multipath fading channels.

In Chapter 4, we demonstrate how effective the NMSV is for the characterization of WSSUS channels. We first define the normalized frequency mean square covariance (NFMSV), the normalized time mean square covariance (NTMSV), and the normalized frequency-time mean square covariance (NFTMSV) to characterize the frequency, the time and the frequency and time selectivities of WSSUS channels. Next, we show how they are related to the correlation functions of WSSUS channels. Then, we provide various analytical and simulational results to show how closely they are related to the performance of various communication systems.

In Chapter 5, we consider channels that are not WSSUS. In this case, the NMSV does not provide a satisfactory relationship with system performance. From the analysis of the simple diversity combining scheme, we define a variant of the re-centered normalized mean square correlation (RC-NMSR) of the time-variant frequency response. We also define variants of normalized arithmetic variance (NAV) to characterize the shape of the specular component of the frequency response. Then, we show how well these parameters are related to the performance of the simple diversity combining scheme. Since diversity combining is fundamentally related to many communication techniques, we expect these parameters provide measures of "goodness" or "badness" of given channels for many systems. More importantly, the logical procedures introduced here can be adopted analogously in entirely different systems.

In Chapter 6, we provide examples to show how the parameters can be used in system design. We first consider a frequency resource allocation problem in a FHSS system. Next we study the effect of frequency hopping rate on the performance of FHSS systems. Then, using similar ideas, we propose carrier-separated orthogonal frequency

division multiplexing (CS-OFDM) systems for multirate multiple access wireless communications.

In Chapter 7, we study and compare various decision directed multiuser detection algorithms under realistic multipath fading environments. In particular, we consider maximum likelihood sequence estimation (MLSE) and its variant such as decision feedback sequence estimation (DFSE) and bias-compensated decision feedback sequence estimation (BC-DFSE). We also study the performance of single and multiple stage successive and parallel interference cancellation algorithms under various situations.

We conclude this report in Chapter 8.

Chapter 2

Statistics of L^2 Processes

In this chapter, we introduce simple parameters to extract useful information efficiently from the first and the second order statistics of stochastic processes. In Section 2.1, we briefly discuss the arithmetic mean of a given function that is the underlying concept for various parameters introduced in this report. Since the theory of measure and integration used in this chapter can be unfamiliar to many readers, we start the section with brief remarks on possible easy interpretation of general integrals. In Section 2.2, we collect a series of definitions and notational conventions used throughout this report. In Section 2.3, we define a set of first and second order statistics of L^2 processes. We first define the *normalized mean square mean* (NMSM) to characterize the relative significance of the mean of an L^2 process compared to the deviation from it. We then define the *normalized mean square covariance (NMSV)*, the *normalized mean square correlation (NMSR)*, and the *re-centered normalized mean square correlation (RC-NMSR)* to characterize the overall correlation properties of an L^2 process. In particular, we show that the NMSV is related to the effective number of uncorrelated random variables to describe the process, which we will call the *stochastic degree of freedom* of the process in the last section. To define and study the parameters with generality, we need some degree of mathematical rigor. Since the background mathematical theories are not treated in a usual engineering curriculum, we briefly overview the theories in Appendix A.

2.1 Introduction

The concept of mean is often useful to characterize the overall characteristics of a set of observed data. The parameters we introduce in this report are in fact some kind of mean values of various functions with proper normalization. In this section, we

define the concept of arithmetic mean for measurable functions in a general context of measure theory. For readers unfamiliar to the concepts of measure and integration, we first give alternative perspectives of various formulae of this chapter without using concepts of the general theory of measure.

2.1.1   Measure and Integration

We believe that readers who need quick references to the theory of measure and integration will find Appendix A helpful. For those readers who are not mathematically oriented, we recommend to regard an integration as a weighted summation. For example, we recommend to regard the following integration

$$\int_X f \, d\mu \tag{2.1}$$

of a function f over a measurable space X with respect to a measure μ as the following summation

$$\sum_{n=1}^{N} f(x_n)\mu_n \tag{2.2}$$

by assuming $X = \{X_1, \cdots, X_n\}$. Here, $\mu_1, \cdots, \mu_N$ are weighting factors that are related to the measure μ. In many cases, the underlying measurable space X is an Euclidean space and the measure μ is a Lebesque measure. In such cases, the corresponding weighting factors μ_n are assumed to be 1. However, for a more general measure μ, the weighting factors μ_n are not necessarily 1, which is a power of the introduction of a general measure. We believe that a choice of proper weighting factors (or a proper measure) can extend the applicability of the concepts introduced in this report.

2.1.2   Arithmetic Mean

Assume that μ is a non-zero finite measure on a measurable space $(X, \mathfrak{X})$. Then, we define the *arithmetic mean* (AM) $\langle h \rangle_X$ of an integrable function $h : X \to \mathbb{C}$ by

$$\langle h \rangle_X = \frac{1}{\mu(X)} \int_X h \, d\mu. \tag{2.3}$$

Now assume that μ is σ-finite and that $\mu(X) = \infty$. Then, we define the arithmetic mean $\langle h \rangle_X$ to be zero if h is integrable. More generally, if $h : X \to \bar{\mathbb{R}}$ is measurable and if h is integrable over X_n for each n and the sequence

$$\langle h \rangle_{X_n} = \frac{1}{\mu(X_n)} \int_{X_n} h \, d\mu \tag{2.4}$$

converges to a unique number M as n goes to ∞ for any sequence $\{X_n\}_{n=1}^{\infty}$ of measurable sets such that $0 < \mu(X_n) < \infty$ and $X_n \subseteq X_{n+1}$ for each n and that $\bigcup_{n=1}^{\infty} X_n = X$, then we define the arithmetic mean $\langle h \rangle_X$ by the number M. In the following, we will often write the right hand side of (2.3) even when $\mu(X) = \infty$ as the meaning defined in this paragraph.

2.2 Stochastic Process

Probability is a special kind of measure. In particular, it is a finite measure taking values in $[0, 1]$. Hence, the theory of probability can be regarded as a special theory of finite measures. However, for conceptual reasons, we associate various terminologies of general measure theory with other names in probability theory. First of all, measurable mappings defined on probability space are called *random elements*. Random elements taking numbers as the values are called random variables. We call the arithmetic mean of a random variable the expected value or the mean of the random variable. In many practical situations, we are interested in a collection of random elements which may be related one another. Generally, an indexed collection of random elements are called a stochastic process or a process in short. Among all stochastic processes, we are particularly interested in L^2 processes for which we define the first and the second order statistic functions. In this section, we summarize the formal definitions of these concepts. More thorough treatments can be found in [13, 14, 15].

2.2.1 Probability Space

In probability theory, upper case letters such X, Y, and Z are used, in general, to stand for measurable mappings and the values these mappings take are usually denoted by lower case letters x, y, and z. The measurable space in probability theory is often denoted by $(\Omega, \mathfrak{F})$. A measure P on a measurable space $(\Omega, \mathfrak{F})$ is called a *probability measure* if $P(\phi) = 0$ and $P(\Omega) = 1$. If P is a probability measure on $(\Omega, \mathfrak{F})$, the triplet $(\Omega, \mathfrak{F}, P)$ is called a *probability space*. Here, the set Ω is called the *sample space* and elements of the σ-field $\mathfrak{F}$ are called *events* of the probability space. For example, the Lebesque measure λ on $I = (0, 1]$ is a probability measure and $(I, \mathfrak{B}(I), \lambda)$ is a probability space. In the remainder of this section, let $(\Omega, \mathfrak{F}, P)$ be a given probability space.

Let $(S, \mathfrak{S})$ be a measurable space. Then, a mapping $X : \Omega \to S$ is said to be a *($\mathfrak{S}$-$\mathfrak{F}$) random element* if $X^{-1}(A) \in \mathfrak{F}$ for any $A \in \mathfrak{S}$. Consequently, a measurable mapping defined on a probability space is called a random element. Given a random element

$X : \Omega \to S$, we define a set function $\mu_X : \mathfrak{S} \to [0,1]$ called the *(probability) distribution* of X by

$$\mu_X(E) = P\big(X^{-1}(E)\big), \quad E \in \mathfrak{S}. \tag{2.5}$$

It is easy to verify that the distribution μ_X of X is a probability measure on $(S, \mathfrak{S})$. Consequently, a random element maps a probability space into another probability space possibly for ease of interpretation or manipulation.

   If $S = \mathbb{R}$ and $\mathfrak{S} = \mathfrak{B}(\mathbb{R})$, then X is said to be a *real random variable*. If $S = \overline{\mathbb{R}}$ and $\mathfrak{S} = \mathfrak{B}(\overline{\mathbb{R}})$, then X is said to be an *extended (real) random variable*. If $S = \mathbb{C}$ and $\mathfrak{S} = \mathfrak{B}(\mathbb{C})$, then X is said to be a *complex random variable*. If X is a complex random variable, then the functions defined by the real and the imaginary parts of X are real random variables. If $S = \mathbb{R}^n$ and $\mathfrak{S} = \mathfrak{B}(\mathbb{R}^n)$ or if $S = \mathbb{C}^n$ and $\mathfrak{S} = \mathfrak{B}(\mathbb{C}^n)$, then X is said to be an *n-dimensional real or complex random variable*, respectively. If $S = \mathbb{R}^\infty$ or $S = \mathbb{C}^\infty$, the set of all sequences of real or complex numbers, respectively, then X is said to be a *real or complex random sequence*.

   Let I be any non-empty set and let $\{(S_i, \mathfrak{S}_i)\}_{i \in I}$ be a collection of measurable spaces. Then, a collection $\{X_i : \Omega \to S_i\}_{i \in I}$ of random elements is said to be *independent* if

$$P\Big(X_{i_1}^{-1}(E_1) \cap \cdots \cap X_{i_m}^{-1}(E_m)\Big) = P\big(X_{i_1}^{-1}(E_1)\big) \cdots P\big(X_{i_m}^{-1}(E_m)\big) \tag{2.6}$$

for any $E_1 \in \mathfrak{S}_{i_1}, \cdots, E_m \in \mathfrak{S}_{i_m}$ and for any finite members $i_1, \cdots, i_m$ of I.

2.2.2   Distribution Functions and Density Functions

Let X be a real random variable. Then, the probability distribution μ_X of X is defined by

$$\mu(B) = P\big(X^{-1}(B)\big), \quad B \in \mathfrak{B}(\mathbb{R}) \tag{2.7}$$

and is a probability measure on $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$. Consequently, a real random variable transforms a probability space into an another probability space in which the sample space is the set of real numbers and the events are the Borel sets of real numbers. We define the *(probability) distribution function* $F_X : \mathbb{R} \to [0,1]$ of X by

$$F_X(x) = \mu_X\Big((-\infty, x]\Big), \quad x \in \mathbb{R}. \tag{2.8}$$

A non-negative measurable function $f_X : \mathbb{R} \to \mathbb{R}$ is said to be a *(probability) density function* of X with respect to the Lebesgue measure λ if

$$F_X(x) = \int_{(-\infty, x]} f_X \, d\lambda, \tag{2.9}$$

for all $x \in \mathbb{R}$.

Let $X = (X_1, \cdots, X_m)$ be an m-dimensional real random vector. The (probability) distribution $\mu_X : \mathfrak{B}(\mathbb{R}^m) \to \mathfrak{F}$ of X is similarly defined by

$$\mu(B) = P\big(X^{-1}(B)\big), \quad B \in \mathfrak{B}(\mathbb{R}^m). \tag{2.10}$$

Also the (probability) distribution function $F_X : \mathbb{R}^m \to [0,1]$ is defined by

$$F_X(x_1, \cdots, x_m) = \mu_X\Big((-\infty, x_1] \times \cdots \times (-\infty, x_m]\Big). \tag{2.11}$$

A non-negative function $f_X : \mathbb{R}^m \to \mathbb{R}$ is called the *(probability) density function* if

$$F_X(x_1, \cdots, x_m) = \int_{(-\infty, x_1] \times \cdots \times (-\infty, x_m]} f_X \, d(\lambda \times \cdots \times \lambda), \tag{2.12}$$

for any $x_1, \cdots, x_m \in \mathbb{R}$. We often call μ_X the *joint (probability) distribution*, F_X the *joint (probability) distribution function*, and f_X the *joint (probability) density function* of the real random variables $X_1, \cdots, X_m$.

2.2.3   Expected Values

Let X be an (extended) real or complex random variable. If the arithmetic mean $\langle X \rangle_\Omega$ (see Section 2.1) is well-defined, we call it the expected value or the mean of X. In probability theory, we write $E[X]$ instead of $\langle X \rangle_\Omega$. Often we also write $m(X)$ for $E[X]$. If the mean $m(X)$ is finite, we define an (extended) random variable D_X called the *deviation of X* by

$$D_X = X - m(X). \tag{2.13}$$

A deviation of an (extended) random variable always has zero mean.

The expected value $E[|X|^2]$ of the (extended) real random variable $|X|^2$ is called the *power* of X and is denoted by $P(X)$. If the deviation D_X is well-defined, the power of the deviation D_X is called the *variance* of X and the non-negative square root of it is called the *(standard) deviation* of X. The standard deviation of X is denoted by $\sigma(X)$. It is easy to see that

$$\sigma(X)^2 = P(X) - |m(X)|^2. \tag{2.14}$$

Let X and Y be two (extended) real or complex random variables defined on the same probability space. The expected value $E[XY^*]$ is called the correlation between

X and Y and denoted by $R(X,Y)$ if it exists. If $R(X,Y)$ exists, then we have

$$|R(X,Y)|^2 \leq P(X)P(Y). \tag{2.15}$$

which is a Cauchy-Schwarz inequality. If both $m(X)$ and $m(Y)$ exist and are finite, and if $R(D_X,D_Y)$ exists, then the covariance $V(X,Y)$ between X and Y is defined by

$$V(X,Y) = R(D_X,D_Y) = R(X,Y) - m(X)m^*(Y). \tag{2.16}$$

It follows from (2.15) that

$$|V(X,Y)|^2 \leq \sigma^2(X)\sigma^2(Y). \tag{2.17}$$

The random variables X and Y are said to be *uncorrelated* if $R(X,Y) = 0$.

Assume that X and Y are L^2 random variables. Then, since $|X| < |X|^2 + 1$ and $|Y| < |Y|^2 + 1$, X and Y are integrable so that the deviations D_X and D_Y are well-defined. It immediately follows that $E[|D_X|^2] < \infty$ and $E[|D_Y|^2] < \infty$. Consequently, X and Y have finite variances and standard deviations. It also follows that both $R(X,Y)$ and $V(X,Y)$ exist and are finite.

2.2.4 L^2 Processes

Let $(S,\mathfrak{S})$ be a measurable space and T be a nonempty set. Then, an indexed family $\{X(t), t \in T\}$ of $(\mathfrak{S}\text{-}\mathfrak{F})$ random elements $X(t) : \Omega \to S$, $t \in T$ is said to be a *stochastic process with state space* $(S,\mathfrak{S})$. We often write $X(t)$ to denote the process $\{X(t), t \in T\}$ if there is no cause of confusion. Now let T be a set of numbers. Then, a stochastic process $\{X(t), t \in T\}$ is said to be *(strictly) stationary* if the two random elements $(X(t_1), \cdots, X(t_m))$ and $(X(t_1 + \delta), \cdots, X(t_m + \delta))$ have the same probability distribution for any finite number of indices $t_1, \cdots, t_m \in T$ and any number δ such that $t_1 + \delta, \cdots, t_m + \delta \in T$. A stochastic process $\{X(t), t \in T\}$ with state space $(\mathbb{C}, \mathfrak{B}(\mathbb{C}))$ is said to be an L^2 process if $E[|X(t)|^2] < \infty$ for any $t \in T$. As noted in Section 2.2.3, we can define the first and the second order statistics for L^2 processes. In the remainder of this section, we assume that $\{X(t), t \in T\}$ is a given L^2 process.

First we define the *mean function* $m_X : T \to \mathbb{C}$ of the process by

$$m_X(t) = E[X(t)], \quad t \in T. \tag{2.18}$$

Then, we define an another stochastic process $\{D_X(t), t \in T\}$ called the deviation of

the stochastic process $\{X(t), t \in T\}$ by

$$D_X(t) = X(t) - m_X(t), \quad t \in T. \tag{2.19}$$

It is easy to see that the deviation is also an L^2 process and has zero mean function.

Next we define the *power function* $P_X : T \to \mathbb{C}$ of the process $X(t)$ by

$$P_X(t) = E[|X(t)|^2]. \tag{2.20}$$

The power function P_{D_X} of the deviation D_X is called the *variance function* of the process $X(t)$. The positive square root of the variance function is called the *standard deviation function* and is denoted by σ_X. Hence, it turns out that

$$\sigma_X^2(t) = E[|X(t) - m_X(t)|^2] = P_X(t) - |m_X(t)|^2. \tag{2.21}$$

The *correlation function* $R_X : T \times T \to \mathbb{C}$ of the process $X(t)$ is defined by

$$R_X(t, t') = E[X(t)X^*(t')], \quad t, t' \in T. \tag{2.22}$$

The correlation function R_{D_X} of the deviation $D_X(t)$ is usually called the *covariance function* of the process $X(t)$ and is denoted by V_X. Consequently, it follows that

$$
\begin{aligned}
V_X(t, t') &= E[(X(t) - m_X(t))(X^*(t') - m_X^*(t'))] & (2.23) \\
&= R_X(t, t') - m_X(t)m_X^*(t') & (2.24)
\end{aligned}
$$

for $t, t' \in T$. The L^2 process $X(t)$ is said to be *uncorrelated* if $R_X(t, t') = 0$ for any $t, t' \in T$ with $t \neq t'$.

Let T be a set of numbers. Then, the L^2 process $\{X(t), t \in T\}$ is said to be *wide-sense stationary (WSS)* if $m_X(t) = m_X(t')$ and $R_X(t, t') = R_X(t+\delta, t'+\delta)$ for any $t, t' \in T$ and any number δ such that $t + \delta, t' + \delta \in T$. Consequently, $X(t)$ is WSS if and only if there exist a constant c and a function r_X such that

$$m_X(t) = c \tag{2.25}$$

and

$$R_X(t, t - \delta) = r_X(\delta) \tag{2.26}$$

for any $t \in T$ and any δ with $t - \delta \in T$.

2.3   Statistics of L^2 Processes

In this section, we define various parameters to characterize stochastic processes. Throughout this section, we assume that a measure space $(T, \mathfrak{T}, \mu)$ and an L^2 process $\{X(t), t \in T\}$ are given. We assume that μ is σ-finite. We first define the normalized mean square mean (NMSM) to quantify the relative significance of the mean compared to the deviation. Next, we define, in turn, the normalized mean square covariance (NMSV), the normalized mean square correlation (NMSR), and the re-centered normalized mean square correlation (RC-NMSR) of the process. Throughout this section, we assume that the functions m_X, R_X, and V_X are well-behaved so that the integrals of them are well-defined. Note that the parameters we define depend on the choice of the measure μ. However, for notational convenience, we will not mention the dependence explicitly in the definition.

2.3.1   Normalized Arithmetic Variance of Mean Function

In this subsection, we define parameters to characterize effectively the mean function m_X of the process. First we consider the arithmetic mean $\langle m_X \rangle_T$ and arithmetic variance $\langle |m_X - \langle m_X \rangle_T|^2 \rangle_T$ of the mean function m_X. These parameters are among the most simplest parameters and are usually very effective for the characterization of the mean function m_X. In many cases, we are interested in the overall shape of the mean function m_X rather than the scale. For such cases, we define the *normalized arithmetic mean (NAM)* $\nu_m(m_X)$ and the *normalized arithmetic variance (NAV)* $\nu_v(m_X)$ by

$$\nu_m(m_X) = \frac{\langle m_X \rangle_T}{\sqrt{\langle |m_X|^2 \rangle_T}} \tag{2.27}$$

and

$$\nu_v(m_X) = \frac{\langle |m_X - \langle m_X \rangle_T|^2 \rangle_T}{\langle |m_X|^2 \rangle_T} \tag{2.28}$$

respectively, assuming that $\langle |m_x|^2 \rangle_T \neq 0$. It immediately follows that

$$\nu_v(m_X) = 1 - |\nu_m(m_X)|^2. \tag{2.29}$$

Consequently, it follows that $0 \leq \nu_v(m_X) \leq 1$ and hence that $0 \leq |\nu_m(m_X)| \leq 1$. Often multiplying m_X by a unit magnitude complex number does not make any difference in the situation. In such a case, $\nu_v(m_X)$ becomes the primary parameter for the shape of the mean function.

2.3.2   Normalized Mean Square Mean

We are often interested in characterizing the amount of energy contained in the mean relative to that in the deviation. For example, consider a random variable Y. Assume that $0 < E[|Y|^2] < \infty$. Then, we define the *normalized square mean (NSM)* of the random variable Y by

$$\frac{|m(Y)|^2}{P(Y)} \tag{2.30}$$

to quantify the fraction of energy is stored in the mean. If the NSM is close to 1, the deviation is negligible so that we can approximate the random variable Y with a number $m(Y)$ without severe loss of accuracy.

For the L^2 process $X(t)$, we define similar quantity called the *normalized mean square mean (NMSM)* $\mathcal{M}\{X\}$ by

$$\mathcal{M}\{X\} = \frac{\displaystyle\int_T |m_X|^2 \, d\mu}{\displaystyle\int_T P_X \, d\mu}. \tag{2.31}$$

Note that we can write (2.31) similarly to (2.28) to appreciate the meaning of NMSM more clearly. In other words, we can rewrite the NMSM as $\langle |m_X|^2 \rangle_T / \langle P_X \rangle_T$. With this expression, we see that NMSM is an arithmetic mean of the square of the mean function with a normalization. From now on, we will not use the notation of arithmetic mean again for direct mathematical expression. When the numerator is ∞, the right-hand side of (2.31) will be understood to be the common limit of the sequences

$$\frac{\displaystyle\int_{T_n} |m_X|^2 \, d\mu}{\displaystyle\int_{T_n} P_X \, d\mu} \tag{2.32}$$

for any sequences $\{T_n\}_{n=1}^\infty$ of measurable subsets of T such that $T_n \subseteq T_{n+1}$ and $0 < \int_{T_n} P_X \, d\mu < \infty$ for each n. If the sequence (2.32) does not converge or if it converges to different limits depending on the choice of sequence of subsets $\{T_n\}$, the NMSM is not defined. In all following definitions, we will mean by similar fractions the limiting values when the numerators are ∞.

As a corresponding quantity to the standard deviation, we define the *normalized*

mean square deviation (NMSD) $\mathcal{D}\{X\}$ by

$$\mathcal{D}\{X\} = \frac{\int_T \sigma_X^2 \, d\mu}{\int_T P_X \, d\mu}. \tag{2.33}$$

Since $P_X = |m_X|^2 + \sigma_X^2$, it follows that

$$\mathcal{M}\{X\} + \mathcal{D}\{X\} = 1 \tag{2.34}$$

and that $0 \leq \mathcal{M}\{X\}, \mathcal{D}\{X\} \leq 1$.

2.3.3   Normalized Mean Square Covariance

In many problems, two random variables Y and Z can be regarded to be essentially the same if $Z = Y + c$ for some real or complex number c. Consequently, in studying the similarity between two random variables Y and Z, we are often more interested in the correlation between the deviations D_Y and D_Z. Intuitively speaking, we often want to determine if Y and Z have the tendency of coincident increase or decrease in their values, which is usually determined by the covariance $V(Y, Z)$. Here, we note that the scaling of a random variable by a real or complex number affect the value of covariance. However, as in the addition of constant, we often regard two random variables Y and Z the same if $Z = cY$ for some non-zero real or complex number c. In other words, the tendency of simultaneous increase or decrease are often more important than the amount of increase or decrease. Consequently, to characterize the similarity between random variables Y and Z, we usually use the absolute value or the absolute square of the correlation coefficient $\rho(Y, Z)$ defined by

$$\rho(Y, Z) = \frac{V(Y, Z)}{\sigma(X)\sigma(Y)}. \tag{2.35}$$

if the deviations $\sigma(Y)$ and $\sigma(Z)$ are not zero.

For the given L^2 process, we define the *normalized mean square covariance (NMSV)* $\mathcal{V}\{X\}$ by

$$\mathcal{V}\{X\} = \frac{\int_{T \times T} |V_X(t, t')|^2 \, d(\mu \times \mu)(t, t')}{\left[\int_T V_X(t, t) \, d\mu(t) \right]^2} \tag{2.36}$$

to characterize the degree of freedom among the random variables in the process. Since

$|V_X(t,t')|^2 \le V_X(t,t)V_X(t',t')$ for any t, $t' \in T$,

$$0 \le \mathcal{V}\{X\} \le 1. \tag{2.37}$$

If $\mathcal{V}\{X\}$ is very close to 1, the random variables in the process are highly correlated. As the NMSV decreases, there are more degree of freedom among the random variables in the process.

2.3.4   Normalized Mean Square Correlation

In Section 2.3.3, we considered the situation when the addition of constant is not important. However, sometimes the values the random variables take are important. In this case, we use correlation rather than covariance to compare two random variables. As a corresponding quantity to the NMSV of the given process $X(t)$, we define the *normalized mean square correlation (NMSR)* $\mathcal{R}\{X\}$ by

$$\mathcal{R}\{X\} = \frac{\displaystyle\int_{T\times T} |R_X(t,t')|^2 \, d(\mu \times \mu)(t,t')}{\left[\displaystyle\int_T R_X(t,t) \, d\mu(t)\right]^2}. \tag{2.38}$$

Since $|R_X(t,t')|^2 \le R_X(t,t)R_X(t',t')$ for any t, $t' \in T$,

$$0 \le \mathcal{R}\{X\} \le 1. \tag{2.39}$$

We define an another related quantity called the *re-centered normalized mean square correlation (RC-NMSR)* $\mathcal{C}\{X\}$ of the process $X(t)$ by

$$\mathcal{C}\{X\} = \frac{\displaystyle\int_{T\times T} \left\{ |R_X(t,t')|^2 - |m_X(t)|^2|m_X^*(t')|^2 \right\} d(\mu \times \mu)(t,t')}{\left[\displaystyle\int_T R_X(t,t) \, d\mu(t)\right]^2}. \tag{2.40}$$

It immediately follows that

$$\mathcal{C}\{X\} = \mathcal{R}\{X\} - \mathcal{M}^2\{X\}. \tag{2.41}$$

Consequently, it follows that $\mathcal{C}\{X\} \le 1$. However, it is not evident what the lower

bound is. To find it out, note that

$$\mathcal{C}\{X\} = \frac{\int_{T \times T} \left\{ |V_X(t,t')|^2 + m_X^*(t)V_X(t,t')m_X(t') + m_X(t)V_X^*(t,t')m_X^*(t') \right\} d(\mu \times \mu)(t,t')}{\left[\int_T R_X(t,t) \, d\mu(t) \right]^2}.$$

(2.42)

Here, it can be shown that

$$\int_{T \times T} m_X^*(t)V_X(t,t')m_X(t') \, d(\mu \times \mu)(t,t') \geq 0.$$

(2.43)

In fact, V_X is said to be positive semi-definite. A function $K : T \times T \to \mathbb{C}$ is said to be a *positive semi-definite or non-negative definite* if

$$\int_{T \times T} f^*(t)K(t,t')f(t') \, d(\mu \times \mu)(t,t')$$

(2.44)

for any $f \in L^2(T, \mathfrak{T}, \mu)$. Consequently, it follows that

$$0 \leq \mathcal{C}\{X\} \leq 1.$$

(2.45)

2.4   Stochastic Degree of Freedom

In this section, we study the meaning of the NMSV and define the concept of the stochastic degree of freedom of an L^2 process. We need the Mercer's theorem and the Karhuenen-Loéve expansion theorem for such a purpose. Since we need a bit more general versions than usually available ones, we briefly study the theorems in the first subsection. Then, we study the relation between the NMSV and the effect degree of freedom in the process.

2.4.1   Mercer's Theorem and Karhuenen-Loéve Expansion

We first study bilinear expansions of kernel functions. Although there is a more general theorem (refer to Chapter 10 in [16]), we consider the Mercer's thoerem, which is sufficient for our purpose. Although they are not integral parts of the theorem, we include (2.47) and (2.48) since they are needed in the next subsection.

Theorem 2.1 (Mercer's Theorem).
Let $(T, \mathfrak{T})$ be a compact topological space and let μ be a finite measure on $(T, \mathfrak{B}(\mathfrak{T}))$ such that $\mu(E) > 0$ for any $E \in \mathfrak{T}$. Assume that $K : T \times T \to \mathbb{C}$ is continuous, $(\mu \times \mu)$-almost

everywhere nonzero, and positive semi-definite. Then, there are a sequence $\{\lambda_i\}_{i \in I}$ of positive real numbers and a sequence $\{\psi_i : T \to \mathcal{C}\}_{i \in I}$ of continuous orthonormal functions such that

$$K(t,s) = \sum_{i \in I} \lambda_i \psi_i(t) \psi_i^*(s). \tag{2.46}$$

Here, $I = \{1, 2, \cdots, N\}$ for some finite natural number N, or $I = \mathbb{N}$, the set of all natural numbers. If $I = \mathbb{N}$, by the equality in (2.46), we mean that the series on the right-hand side converges absolutely and uniformly to K.

In particular, we have

$$\int_{T \times T} |K(t,s)|^2 \, d(\mu \times \mu)(t,s) = \sum_{i \in I} \lambda_i^2 \tag{2.47}$$

and

$$\int_T K(t,t) \, d\mu(t) = \sum_{i \in I} \lambda_i. \tag{2.48}$$

To understand the background of the above theorem, we introduce the Hilbert-Schmidt operator $A : L^2(T, \mathfrak{B}(\mathfrak{T}), \mu) \to L^2(T, \mathfrak{B}(\mathfrak{T}), \mu)$ associated with the kernel K that is defined by

$$A(x)(t) = \int_T K(t,s) x(s) \, d\mu(s), \quad t \in T. \tag{2.49}$$

Then, λ_i is an eigenvalue and ψ_i is a corresponding eigenfunction of the operator A for each $i \in I$. It is well-known that $\{\psi_i\}_{i \in I}$ spans $L^2(T, \mathfrak{B}(\mathfrak{T}), \mu) - \text{Ker}(A)$.[1] Moreover, for any given positive real number ε, only finite members of $\{\lambda_i\}_{i \in I}$ can exceed ε, regardless of the cardinality of I. The proof of this form of the Mercer theorem can be found in Chapter 10 of [16].

In short, Mercer's theorem states that a well-behaved kernel function has a bilinear expansion in terms of positive eigenvalues and orthonormal eigenfunctions. We can apply this theorem to the covariance function of a stochastic process and deduce the Karhuenen-Loéve expansion. For example, Chapter 1 of [15] describes the Karhuenen-Loéve expansion of an L^2 stochastic process with a compact interval as the index set. In many cases, we want to consider a more general index set T than just a compact interval such as a finite union of disjoint rectangles. So we restate the Karhuenen-Loéve expansion in a little bit more general form that is suitable for our purpose, which can be proved following the same lines of reasoning as in [15] with a few modifications.

Theorem 2.2 (Karhuenen-Loéve Expansion).

---

[1]The kernel $\text{Ker}(A)$ of a linear map $A : X \to Y$ is defined to be the set $\{x \in X | Ax = 0\}$.

Let $\{X(t),\, t \in T\}$ be a complex valued stochastic process defined on a compact rectifiable subset T of a Euclidean space. Assume that μ is the Lebesque measure restricted to the Borel subsets of T. Assume that the covariance function V_X of X is continuous and $(\mu \times \mu)$-almost everywhere nonzero. Then, by Mercer's theorem, there exist a sequence $\{\lambda_i\}_{i \in I}$ of positive real numbers and a sequence $\{\psi_i : T \to \mathbb{C}\}_{i \in I}$ of orthonormal functions such that (2.46) holds with K replaced by V_X. If we define a sequence $\{X_i\}_{i \in I}$ of random variables by [2]

$$X_i = \int_T \psi_i^*(t) X(t)\, dt, \tag{2.50}$$

then

$$E[X_i X_j^*] - E[X_i]E[X_j^*] = \lambda_i \delta_{ij} \quad \forall\, i,\, j \tag{2.51}$$

and

$$X(t) - m_X(t) = \sum_{i \in I} (X_i - E[X_i])\psi_i(t) \tag{2.52}$$

where $m_X(t)$ denotes the mean function of $X(t)$. If $I = \mathbb{N}$, then the equality in (2.52) means that the summation in the right-hand side converges in mean square to $X(t)$, uniformly in t.

This theorem states that a well-behaved stochastic process can be expanded in terms of a sequence of random variables and orthonormal functions. We can regard each random variable as an orthonormal projection of the stochastic process. In some sense, we can regard each projection as a degree of freedom. The eigenvalues (of the Hilbert-Schmidt operator associated with the covariance function) are the variances of the random variables and determine the relative significance of the random variables. We note that random variables with small variances will have small effects on the random process (except perhaps on the mean function). In the next subsection, we study how the NMSV is related to the eigenvalues.

2.4.2   Definition of Stochastic Degree of Freedom

Let $\{X(t),\, t \in T\}$ be a well-behaved stochastic process that satisfies all the conditions of the Karhuenen-Loéve expansion theorem. Let $\{\lambda_i\}_{i \in I}$ and $\{\psi_i\}_{i \in I}$ be the sequences of positive scalars (or eigenvalues) and orthonormal functions in the bilinear expansion

---

[2] *If T is a compact rectifiable subset of a Euclidean space, $\{X(t), t \in T\}$ is an L^2 stochastic process with continuous covariance function, and $g : T \to \mathbb{C}$ is continuous, then the Riemann integral $\int_T g(t)X(t)dt$ is well defined and converges to a complex random variable in the L^2 sense. Note that we have to consider the limit of the Riemann sum instead of considering the lower limit or the upper limit because of the lack of order property in this case [15].*

of the covariance function K_X. Since K_X is continuous on a compact set T, it is bounded so that $\int_T K_X(t,t)\,d\mu(t) < \infty$. Also, $\int_T K_X(t,t)\,d\mu(t) > 0$ since λ_i's are positive. Hence, the NMSV $\mathcal{V}\{X\}$ of X is well defined and is given by

$$\mathcal{V}\{X\} = \frac{\displaystyle\sum_{i\in I}\lambda_i^2}{\left[\displaystyle\sum_{i\in I}\lambda_i\right]^2} \tag{2.53}$$

which follows from (2.47) and (2.48).

To appreciate the meaning, consider two positive real numbers x and y with $x < y$ and choose any positive real numbers x' and y' such that $x < x' < y' < y$ with $x' + y' = x + y$. Then, we have

$$x'^2 + y'^2 < x^2 + y^2. \tag{2.54}$$

This implies that the NMSV becomes smaller as the the $\lambda_i's$ become more evenly distributed. As shown in the previous subsection, not all the λ_i's are equally significant. In particular, for any given positive real number ε, we can choose an integer N such that $\sum_{i>N}\lambda_i < \varepsilon$ (which implies that $\sum_{i>N}\lambda_i^2 < \varepsilon^2$). Since a function y/x of real variables x and y is continuous in both x and y except at $x = 0$, this implies that we can choose N so that $\mathcal{V}^{(N)}\{X\}$ defined by

$$\mathcal{V}^{(N)}\{X\} = \frac{\displaystyle\sum_{i=1}^{N}\lambda_i^2}{\left[\displaystyle\sum_{i=1}^{N}\lambda_i\right]^2} \tag{2.55}$$

becomes arbitrarily close to $\mathcal{V}\{X\}$, by making $\sum_{i>N}\lambda_i$ and hence $\sum_{i>N}\lambda_i^2$ very small. Now assume that $\lambda_1, \cdots, \lambda_N$ are chosen so that $\sum_{i=1}^{N}\lambda_i$, $\sum_{i=1}^{N}\lambda_i^2$, and $\mathcal{V}^{(N)}\{X\}$ are close enough to $\sum_{i\in I}\lambda_i$, $\sum_{i\in I}\lambda_i^2$, and $\mathcal{V}\{X\}$, respectively. Then,

$$\mathcal{V}\{X\} \approx \mathcal{V}^{(N)}\{X\} = \frac{1}{N} + \sum_{i=1}^{N}\left(\frac{\lambda_i}{\Lambda_N} - \frac{1}{N}\right)^2 \tag{2.56}$$

where $\Lambda_N = \sum_{i=1}^{N}\lambda_i$. This shows more clearly that NMSV indicates how evenly the non-negligible eigenvalues are distributed. Note that NMSV takes the smallest value $1/N$ when all the N eigenvalues happen to be the same. So if we have N roughly equally strong random variables with all other eigenvalues negligible in the Karhuenen-Loéve expansion, then NMSV of the given stochastic process is approximately $1/N$.

Regardless of the distribution of the eigenvalues, the NMSV characterizes the overall covariance of a given stochastic process. If $1/\mathcal{V}$ is close to a natural number N, then it is possible to say that the overall covariance of the stochastic process is similar to that of a stochastic process with equal N non-zero eigenvalues, which we regard to have N degree of freedom. Observing this we define the *stochastic degree of freedom $\mathcal{F}\{X\}$* of an L^2 process X by

$$\mathcal{F}\{X\} = \frac{1}{\mathcal{V}\{X\}}. \tag{2.57}$$

Consequently, by saying that a stochastic process has d stochastic degree of freedom, we mean that it has the same overall covariance property as a stochastic process has with d equal non-zero eigenvalues. We note that the stochastic degree of freedom is not necessarily an integer but that it is never less than 1.

Chapter 3

Wireless Communication Channels

In this chapter, we describe randomly time-variant linear channels to be studied throughout this report. In most wireless communications, narrow-band or wide-band information signals are modulated by high frequency carriers and then transmitted over a radio link. A radio link can be modeled accurately as a linear time-variant system. Due to various reasons, it is necessary and useful to regard a given radio link channel as randomly time-variant. Consequently, we will regard a wireless communication channel as a randomly time-variant linear system. Since a linear system can be effectively described by it's system functions, we describe a linear channel by the system functions. In particular, we will regard the system functions as stochastic processes, which will be assumed to be Gaussian random processes in many situations. In the following, we briefly describe the physical and mathematical backgrounds of wireless communication channel characterizations. In Section 3.1, we briefly introduce basic electromagnetic propagation mechanisms to understand various aspects of wireless communications. In Section 3.2, we discuss the usual paradigm of hierarchical channel description. In particular, we discuss why multipath fading effect is usually discussed independently from that of shadowing and multipath fading. Then, we provide definitions of various terminologies and theoretical backgrounds for linear time-variant channel modeling with complex base-band representation in Section 3.3. Finally, in Section 3.4, we describe the stochastic channel models we employ throughout in this report.

3.1 Electromagnetic Wave Propagation

In this section, we discuss various mechanisms of electromagnetic wave propagation. Although electromagnetic waves are described accurately by Maxwell equations [17, 18], it is generally a very demanding job to solve them in practical wireless communication environments. Consequently, the behaviours of electromagnetic waves in practical environments are analyzed and understood as combined effects of various well-known phenomena such as scattering, reflection, refraction, and diffraction, which are regarded as basic propagation mechanisms in wireless communications.

3.1.1 Free Space Propagation

It is well-known that electromagnetic wave and light are the same physical entity with different interpretation. In general, we call it light if the frequency is very high ($\gtrsim 10^{14}$ Hz). Consequently, they exhibit fundamentally the same behaviors with difference in the scale. Since the light is visible, it is often helpful to recall the behavior of the light in the study of electromagnetic wave propagation. Assume that there is a point light source in a free space emitting P joules per second isotropically. Let's consider a sphere of radius r with center at the light source. Due to conservation of energy, the power passing through the sphere is still P. Since the area of the sphere is $4\pi r^2$, the power per area passing through the sphere is $P/4\pi r^2$. Consequently, the power per unit area is decreasing according to the inverse square law. Exactly the same law of attenuation happens with electromagnetic wave if the distance between the source and the observer is much larger than the dimension of the source. If the transmitter and receiver are well separated in free space, the received power P_r is given by

$$P_r = \frac{P_t G_t G_r \lambda^2}{16 L \pi^2 r^2} \tag{3.1}$$

where P_t is the transmitted power, G_t the transmitter antenna gain, G_r the receiver antenna gain, r the distance between the transmitter and the receiver in meters, L the system losses ($L > 1$), and λ the wavelength in meters [19].

3.1.2 Scattering, Reflection, and Diffraction

Most wireless communication environments are usually much more complex than free space. In usual urban wireless communications, there are various trees, cars, and buildings that block, scatter, reflect, and diffract electromagnetic waves. Even in satellite to ground mobile, we often have to consider the atmospheric effects. Generally,

when electromagnetic waves interact with an object with dimension comparable to the wavelength, we observe the phenomenon of scattering. On the other hand, we observe reflection, refraction, and diffraction when the object is significantly larger than the wavelength. Reflection and refraction are observed when the wave impinges on the middle of the surface, while diffraction happens when the wave meets the edge of the object.

Sometimes we can observe electromagnetic waves away from the transmitted direction due to scattering. One of the most popular example of electromagnetic wave scattering is Rayleigh scattering which explains why the sky becomes blue or red [17]. According to the Rayleigh scattering theory, blue color is scattered more than the red color. Consequently, the light away from the incident direction is heavily weighted on the blue color range. Generally, when electromagnetic wave is incident on an object, it is partly passed and partly scattered. Also some of the energy can be absorbed by the object. Water vapors in the atmosphere often absorb significant portion of transmitted electromagnetic waves and thus attenuate the signal strength greatly.

When electromagnetic waves are incident on an object with dimension much larger than the wavelength, we can observe the phenomena of reflection and diffraction. If the surface of the object has irregularity of size comparable to the wavelength, the scattering effect is also observable. In fact, it is possible to describe the effects of reflection and diffraction as the cumulative effect of scatterings [20]. When the surface is smooth, the strengths of the reflected and diffracted wave depends on the electromagnetic property of the materials, the angle of incident, and the frequency and the polarization of the electromagnetic wave according to the Fresnel's formulas [17]. However, if the surface is irregular, the reflected wave is further attenuated [21, 22, 19].

Consequently, even when the direct path between transmitter and receiver is obstructed, it is often possible to receive the electromagnetic wave signals due to various objects that cause scattering and reflection objects. The effect of diffraction also contributes to signal transmission over a geometrically obstructed area. The phenomenon of diffraction can be explained using Huygen's principle, which states that all points on a wavefront can be considered as point sources for the production of secondary wavelets, and that these wavelets are combined to produce a new wavefront in the direction of propagation. Diffraction is caused by the propagation of secondary wavelets into the shadowed region. The amount of diffracted energy depends fundamentally on the point of observation and the frequency. Generally, the diffracted energy becomes smaller as the observer approaches the obstacles in the shadowed region and the frequency becomes larger. Consequently, it is generally better to choose low carrier

frequency to allow signals to propagate to the shadowed areas.

3.2   Channel Modeling in Wireless Communications

In this section, we briefly describe the roles of path loss, shadowing, and multipath fading in the characterization of wireless communication channels. The effect of path loss and shadowing can be described by the amount of signal attenuation. Moreover, it generally lasts much longer than the duration of a packet in a usual wireless communications. Unlike path loss and shadowing, multipath fading generally results in more complex effects. In the following, we describe the relations among multipath fading, shadowing, and path loss and discuss why multipath fading is described separately from path loss and shadowing.

3.2.1   Multipath Fading, Shadowing, and Path Loss

In usual indoor or outdoor environments where there are various scatterers and reflectors such as trees, mountains, cars, and buildings, the channel response fluctuates in an unpredictable manner, while the transmitter or the receiver moves over a distance on the order of a carrier wavelength. This fluctuation happens because of the phase change of various signal components arriving at the receiver due to scatterers and reflectors. Movement of the receiver by a distance equal to a fraction of the wavelength can make changes in the phases of the signal components by a similar fraction. As a result, the signals can interfere with one another constructively or destructively over a distance on the order of a wavelength. This is the reason why such a rapid fluctuation happens. This phenomenon is known as the multipath fading in wireless communications because severe signal attenuation can happen due to the presence of multiple routes of signal transmission.

By averaging over an area within a radius of several carrier wavelengths, we can get the local or small-scale average power of the received signal. Consequently, by a local or small-scale area, we mean a neighborhood within a distance of several carrier wavelengths. When there exists an obstacle near the receiver, this local average power of the received signal changes when the receiver moves a distance of tens or hundreds of the carrier wavelength depending on the size of the obstacle and the distance between the receiver and the obstacle. Such a fluctuation of local average power is called shadowing. By averaging out the shadowing fluctuations, we obtain the average received power, which is determined by the (larger-scale or long-range) path-loss and the transmitted power. Consequently, given the transmitted power, the received

signal strength is determined by the addition of signal losses due to (large-scale) path loss, (medium-scale) shadowing and (small-scale) multipath fading. The determination of any of the three losses in practical situations is a mathematically demanding job. Because of the analytical simplicity and the effectiveness in performance analysis, stochastic propagation models are widely adapted.

3.2.2 Multipath Fading and Hierarchical Channel Description

Most digital wireless communication systems are packet-based and the packet size is usually kept small so that the effect of shadowing remains constant over each packet duration. Moreover, if the bandwidth of the system is much smaller than the carrier frequency of the signal, the frequency selectivity due to path-loss and shadowing can usually be neglected. Consequently, path loss and shadowing effect do not involve any signal distortion in usual wideband communications, although it affects the receiver signal-to-noise ratio. Hence, in packet design, we generally treat multipath fading separately from that of path loss and shadowing. To evaluate a particular system, we first study the performance of the packet design considering only multipath fading over a range of received signal-to-noise ratio. Based on the results, the performance of the network layer strategies such as power control or retransmission techniques is studied for given path loss and transmitted power level. Then, the results are used to determine the link budget.

To illustrate the point more concretely, consider the problem of determining the probability $P_G(X)$ of error as a function of global average X of the signal-to-noise ratio that is determined by the path-loss, the transmitted power and the noise power level. Since the shadowing is constant over a packet duration and does not involve any frequency selectivity, the local mean x of the signal-to-noise ratio for a given packet can be written as $x = sX$ with the parameter s representing the shadowing loss. Because the effect of shadowing on a packet is no more than a scaling of the signal-to-noise ratio, it is possible to define the probability $P_L(x)$ of symbol or packet error as a function of the local average signal to noise ratio x. If the probability density function $f_S(s)$ of the shadowing is given, then the probability $P_G(X)$ of error as a function of the global average X of the signal-to-noise ratio is given by

$$P_G(X) = \int_0^\infty P_L(sX) f_S(s) \, ds \qquad (3.2)$$

Here, we assume that the random variable S is properly normalized so that the expectation value $E[S]$ is unity. As stated previously, S is usually regarded as a log-normal

distributed random variable with a suitable choice of variance.

3.3   Linear Time-Variant Channel Model

In the remainder of this report, we will confine ourselves on multipath fading. Consequently, we will mean multipath fading channels by wireless communication channels. In this report, we will model multipath fading channels as randomly time-variant linear systems. The linear system modeling of a radio link is generally very accurate unless the carrier frequency becomes extremely high to reach the optical range. Before proceeding with the stochastic modeling, we first introduce some backgrounds and terminologies to describe linear time-variant (LTV) channels. First, we introduce various system functions for LTV channels. Then, we study the complex representations of signals and LTV channels.

3.3.1   System Functions

As is well-known, time-variant or time-invariant linear channels can be characterized effectively by various system functions such as impulse responses [7]. Let $h(\tau,t)$ denote the channel response at time t due to a unit impulse input at time $t - \tau$. Then, the output $y(t)$ of the channel due to an input signal $x(t)$ is given by

$$y(t) = \int_{-\infty}^{\infty} h(\tau,t)x(t - \tau)\, d\tau. \tag{3.3}$$

We call $h(\tau,t)$ the *time-variant impulse response* of the channel. If the function $h(\tau,t)$ is independent of t, we call the channel time-invariant, in which case $h(\tau,t)$ represents the usual channel impulse response. However, if the channel is time-variant, $h(\tau,t)$ is not the channel response due to a unit input impulse at time t. Note that the channel response at time $t + \tau$ due to a unit input impulse at time t is $h(\tau,t+\tau)$.

For studying time variation of the gain $h(\tau,t)$ at a particular delay time τ, it is useful to consider the *delay Doppler spread function $k(\tau,\nu)$* defined by

$$k(\tau,\nu) = \int_{-\infty}^{\infty} h(\tau,t)\, e^{-j2\pi\nu t}\, dt. \tag{3.4}$$

Consequently, the Doppler spread function $k(\tau,\nu)$ characterizes the frequency content of $h(\tau,t)$ with respect to t at a given delay τ.

Since a signal can usually be represented as a sum of complex sinusoids, it is very useful to consider the output of a linear channel due to complex sinusoids. Let $y_s(t)$ be

the channel output due to the complex sinusoid $x_s(t) = e^{j2\pi ft}$. Then,

$$y_s(t) = \int_{-\infty}^{\infty} h(\tau,t)e^{j2\pi f(t-\tau)} \, d\tau \tag{3.5}$$

$$= H(f,t)x_s(t) \tag{3.6}$$

where $H(f,t)$ is the *time-variant transfer function* of the channel defined by

$$H(f,t) = \int_{-\infty}^{\infty} h(\tau,t) \, e^{-j2\pi f\tau} \, d\tau. \tag{3.7}$$

From linearity, it follows that the output $y(t)$ due to a general input signal $x(t)$ is

$$y(t) = \int_{-\infty}^{\infty} X(f)H(f,t) \, e^{j2\pi ft} \, df. \tag{3.8}$$

where $X(f)$ is the Fourier transform of $x(t)$ defined by

$$X(f) = \int_{-\infty}^{\infty} x(t) \, e^{j2\pi ft} \, df. \tag{3.9}$$

If the channel is time-invariant, (3.6) implies that the output $y_s(t)$ is also a monocromatic signal of the same frequency with possible change only in the amplitude and the phase which are characterized by $H(f,t)$. However, if the channel is time-variant, then the transfer function $H(f,t)$ involves time variation so that the output $y_s(t)$ is no longer monocromatic. To describe this phenomenon mathematically, we introduce the *Doppler spread function* $K(f,\nu)$ of the channel defined by

$$K(f,\nu) = \int_{-\infty}^{\infty} H(f,t) \, e^{-j2\pi\nu t} \, dt. \tag{3.10}$$

Using the inverse Fourier transform relation

$$H(f,t) = \int_{-\infty}^{\infty} K(f,\nu) \, e^{j2\pi\nu t} \, d\nu, \tag{3.11}$$

we can rewrite (3.6) as

$$y_s(t) = \int_{-\infty}^{\infty} K(f,\nu) \, e^{j2\pi(f+\nu)t} \, d\nu. \tag{3.12}$$

From (3.10), we see that the output signal may contain signal component at frequency $f + \nu$. In particular, the Doppler spread function $K(f,\nu)$ represents the amplitude and the phase of the output signal component at frequency $f + \nu$ due to unit complex sinusoid $e^{j2\pi ft}$ of frequency f.

More generally, when the channel is time invariant, it follows from (3.10) that the Fourier transform $Y(f)$ of $y(t)$ is given by

$$Y(f) = H(f,t)X(f). \tag{3.13}$$

For time variant systems, by substituting (3.11) into (3.8), we obtain

$$
\begin{aligned}
y(t) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K(f,\nu)X(f)\, e^{j2\pi(f+\nu)t}\, d\nu\, df \tag{3.14}\\
&= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} K(f-\nu,\nu)X(f-\nu)\, d\nu \right] e^{j2\pi ft}\, df \tag{3.15}
\end{aligned}
$$

so that

$$Y(f) = \int_{-\infty}^{\infty} K(f-\nu,\nu)X(f-\nu)\, d\nu. \tag{3.16}$$

3.3.2   Canonical Complex Representation

A real-valued narrowband or wideband signal $s(t)$ can be represented by

$$s(t) = \mathrm{Re}\left[x(t)e^{j2\pi f_c t}\right] \tag{3.17}$$

where $x(t)$ is a complex-valued signal with no spectral components outside frequency region $[-B, B]$ for some B and f_c with $0 < B \ll f_c$. Here, f_c is usually chosen to be the carrier frequency of the signal. In this representation, the complex signal $x(t)$ is called the (complex) envelope of the signal $s(t)$ and the complex signal $\tilde{s}(t) = x(t)e^{j2\pi f_c t}$ is generally called the (complex) pre-envelope of $s(t)$ with respect to $e^{jf_c t}$. The pre-envelope or the envelope of the real signal $s(t)$ is not defined uniquely by (3.17). However, for physical and mathematical reasons [23], we define the *canonical complex representation* $\tilde{s}(t)$ of a real signal $s(t)$ by

$$\tilde{s}(t) = s(t) + j\hat{s}(t) \tag{3.18}$$

where $\hat{s}(t)$ is the Hilbert transform of $s(t)$ defined by

$$\hat{s}(t) = \frac{1}{\pi}\, \mathrm{P} \int_{-\infty}^{\infty} \frac{s(\tau)}{t-\tau}\, d\tau. \tag{3.19}$$

Here the letter P indicates that the improper integral is in the sense of Cauchy Principal Value. In other words,

$$P \int_{-\infty}^{\infty} \frac{x(\tau)}{t-\tau} d\tau = \lim_{\delta \downarrow 0} \left[\int_{t+\delta}^{\infty} \frac{x(\tau)}{t-\tau} d\tau + \int_{-\infty}^{t-\delta} \frac{x(\tau)}{t-\tau} d\tau \right]. \qquad (3.20)$$

The complex canonical representation is a special type of pre-envelope that is also called an *analytic signal* [23, 24].

Some of the most important properties of Hilbert transform and analytic signal representation are given in [25], which we summarize here.

1) Let

$$s(t) = \sum_{n} A_n \cos(2\pi f_n t + \theta_n) \qquad (3.21)$$

where A_n and θ_n are constants. Then,

$$\hat{s}(t) = \sum_{n} A_n \sin(2\pi f_n t + \theta_n) \qquad (3.22)$$

and

$$\tilde{s}(t) = \sum_{n} A_n e^{j\theta_n} e^{j2\pi f_n t}. \qquad (3.23)$$

2) If $S(f)$, $\hat{S}(f)$, and $\tilde{S}(f)$ are the Fourier transforms of $s(t)$, $\hat{s}(t)$, $\tilde{s}(t)$, respectively, then

$$\hat{S}(f) = \begin{cases} -jS(f) & f > 0 \\ 0 & f = 0 \\ +jS(f) & f < 0. \end{cases} \qquad (3.24)$$

and hence

$$\tilde{S}(f) = \begin{cases} 2S(f) & f > 0 \\ S(f) & f = 0 \\ 0 & f < 0. \end{cases} \qquad (3.25)$$

3) Let $z(t)$ be any complex-valued function with Fourier transform $Z(f)$ vanishing for all $f < 0$. Then, the signal $z(t)$ is the pre-envelope of its real part.

4) Let

$$r(t) = h(t) * s(t) = \int_{-\infty}^{\infty} h(\tau)s(t-\tau) d\tau \qquad (3.26)$$

where $h(t)$ is either a real- or a complex-valued function. Then,

$$\hat{r}(t) = h(t) * \hat{s}(t) \tag{3.27}$$

and hence

$$\tilde{r}(t) = h(t) * \tilde{s}(t) \tag{3.28}$$

so that

$$\tilde{R}(f) = H(f)\tilde{S}(f) = \begin{cases} 2H(f)S(f) & f > 0 \\ H(f)S(f) & f = 0 \\ 0 & f < 0 \end{cases}, \tag{3.29}$$

if $\tilde{R}(f)$, $H(f)$, $S(f)$, and $\tilde{S}(f)$ are the Fourier transforms of $\tilde{r}(t)$, $h(t)$, $s(t)$, and $\tilde{s}(t)$, respectively.

5) For any two waveforms x and y, we define their cross correlation R_{xy} by

$$R_{xy}(\tau) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} x(t)y^*(t - \tau)\, dt \tag{3.30}$$

and their cross-power spectrum W_{xy} by

$$W_{xy}(f) = \int_{-\infty}^{\infty} R_{xy}(\tau)e^{-j2\pi f \tau}\, d\tau. \tag{3.31}$$

We usually denote by R_x the autocorrelation R_{xx} of x. Then,

$$R_{xy}(\tau) = R_{yx}^*(-\tau) \tag{3.32}$$

$$R_{x\hat{x}}(\tau) = \hat{R}_x(\tau) \tag{3.33}$$

$$\hat{R}_x(-\tau) = -\hat{R}_x(\tau) \tag{3.34}$$

$$R_{\hat{x}}(\tau) = R_x(\tau) \tag{3.35}$$

$$\tag{3.36}$$

3.3.3   Baseband Representations of Linear Time-Variant Systems

In this subsection, we briefly discuss equivalent complex baseband representation of bandpass linear time-variant channels. The following theorem provides the necessary mathematical backgrounds.

Theorem 3.1 (Equivalent Complex Baseband Representation).
Consider a channel with time variant impulse response $h(\tau, t)$. Let y be the output due

to the input signal x. We assume that all Fourier transforms of h and y exist and are well-behaved. Assume that there exists a positive real number Δ such that $X(f) = 0$ if $|f| \leq \Delta$ and that $K(f, \nu) = 0$ for all $f \in \mathbb{R}$ if $|\nu| \geq \Delta$. Then,

$$\widetilde{y}(t) = \int_{-\infty}^{\infty} h(\tau, t)\widetilde{x}(t - \tau)\, d\tau. \tag{3.37}$$

Moreover, for any $f_c \in \mathbb{R}$, if we define x_B, y_B, and h_B by

$$x_B(t) = \widetilde{x}(t)e^{-j2\pi f_c t}, \tag{3.38}$$
$$y_B(t) = \widetilde{y}(t)e^{-j2\pi f_c t}, \tag{3.39}$$

and

$$h_B(\tau, t) = h(\tau, t)e^{-j2\pi f_c \tau}, \tag{3.40}$$

then

$$y_B(t) = \int_{-\infty}^{\infty} h_B(\tau, t)x_B(t - \tau)\, d\tau. \tag{3.41}$$

Proof. Since the second part of the theorem is immediate from the first part, we prove only the first part. To prove (3.37), we start from the relation

$$Y(f) = \int_{-\infty}^{\infty} K(f - \nu, \nu)X(f - \nu)\, d\nu \tag{3.42}$$

$$= \int_{-\Delta}^{\Delta} K(f - \nu, \nu)X(f - \nu)\, d\nu. \tag{3.43}$$

From the hypothesis, it follows that

$$Y(0) = \int_{-\Delta}^{\Delta} K(-\nu, \nu)X(-\nu)\, d\nu = 0. \tag{3.44}$$

and that

$$\widetilde{Y}(f) = \begin{cases} 2Y(f) & \text{if } f > 0, \\ 0 & \text{if } f \leq 0. \end{cases} \tag{3.45}$$

Now we have

$$\widetilde{y}(t) = 2\int_0^{\infty} \left[\int_{-\Delta}^{\Delta} K(f - \nu, \nu)X(f - \nu)\, d\nu \right] e^{j2\pi ft}\, df \tag{3.46}$$

$$= \int_0^{\infty} \left[\int_{-\Delta}^{\Delta} K(f - \nu, \nu)\widetilde{X}(f - \nu)\, d\nu \right] e^{j2\pi ft}\, df \tag{3.47}$$

$$= \int_{-\infty}^{\infty} \left[\int_{-\Delta}^{\Delta} K(f - \nu, \nu)\widetilde{X}(f - \nu)\, d\nu \right] e^{j2\pi ft}\, df, \tag{3.48}$$

since $\widetilde{X}(f - \nu) = 0$ if $f \leq 0$ and $\nu \geq -\Delta$. Consequently, it follows that

$$\widetilde{y}(t) \;=\; \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} K(f - \nu, \nu) \widetilde{X}(f - \nu)\, d\nu \right] e^{j2\pi ft}\, df \tag{3.49}$$

$$\;=\; \int_{-\infty}^{\infty} h(\tau, t) \widetilde{x}(t - \tau)\, d\tau. \tag{3.50}$$

$\square$

The above theorem implies that the channel can be described by x_B, y_B, and h_B. Since the signals x and y and the time-variant impulse response h can be recovered from x_B, y_B, and h_B and vice versa, the two representations of the channel are equivalent. In the above theorem, f_c can be chosen arbitrarily. However, in practical narrowband or wideband communication systems, we usually choose as f_c the carrier frequency. In this case, there exists a positive real number W with $W \ll f_c$ such that $\widetilde{X}(f) = 0$ for $f \geq W$, which implies that $\widetilde{Y}(f) = 0$ for $f \geq W + \Delta$. Because the spectra of the input and the output signals are localized around 0, we call the resultant representation by x_B, y_B, and h_B an *(equivalent) complex baseband representation of the system*.

3.4   Randomly Time-Variant Linear Channels

Throughout this report, we will assume that a wireless communication channel is a randomly time-variant linear system. More concretely, we will assume the wireless communication channels are linear time-variant systems with sample functions of some stochastic processes as the system functions. The channels are thus specified by the stochastic processes for the system functions, which we will just call the system functions from now on. For example, the time-variant frequency response $H(f, t)$ of a channel will be assumed to be a stochastic process. In this section, we introduce various terminologies to describe randomly time-variant linear channels. First, we briefly discuss the backgrounds of randomly time-variant linear channel models. Next we define the specular and the diffuse components of the channel system functions. Then, we provide notations for the correlation functions and the covariance functions. We define various concepts to characterize the time and the frequency dispersions of a channel. Throughout this report, we will assume that the time-variant frequency responses are L^2 processes with well-behaved first and second order moments. In particular, we will assume that the diffuse component of a time-variant frequency response is wide-sense stationary in both frequency and time variables. Also we will assume that all system

functions are complex representations, namely, either pre-envelop or envelop representations of the channel system functions.

3.4.1 Introduction

In usual digital communications, a short duration packet is transmitted and treated independently from other packets. Consequently, the effectiveness of a particular packet design depends solely on the channel characteristics during a packet duration. Since the channel characteristics change unpredictably from packet to packet and since we are mostly interested in the average system performance, we model a multipath fading channel as a randomly time-varying system. In real situations, the average system performance can be defined to be the average of the system performance over a long period of time. However, in most wireless communication researches, we consider the ensemble average of the system performance by regarding the channel response as a stochastic process. In particular, we will assume, in the remainder of this report, that channels are *randomly time-variant linear systems* [8]. Such randomly time-variant linear channels will be described by regarding the system functions as stochastic processes. Consequently, wireless communication channels will be described by specifying the characteristics of (randomly time-variant) system functions. In many cases, we will assume the system functions are complex Gaussian stochastic processes. Gaussian stochastic processes are completely specified by the mean and the correlation functions. The first and the second order statistics are usually regarded to be among the most important characteristics for other stochastic processes, too. Consequently, a randomly time-variant linear channel is usually specified by the first and the second order statistics.

3.4.2 Specular and Diffuse Components

The means of the channel responses are called the *specular components* while the deviations are referred to as the *diffuse or scattering components* of the responses. We will follow the conventions of Section 2.2.4 in denoting various first and second order statistics. In particular, m, R, and V will stand for the *mean*, the *correlation*, and the *covariance* functions, respectively. The letter D will stand for the diffuse component. For example, we will write $m_H(f,t)$ and $D_H(f,t)$ for the specular and diffuse components of the time-variant frequency response $H(f,t)$, respectively, so that

$$m_H(f,t) = E[H(f,t)] \qquad (3.51)$$

and

$$D_H(f,t) = H(f,t) - m_H(f,t). \tag{3.52}$$

Also we will write, for example, $R_H(f,t;f',t')$ and $V_H(f,t;f',t')$ for the correlation and the covariance functions of $H(f,t)$ so that

$$R_H(f,t;f',t') = E[H(f,t)H^*(f',t')] \tag{3.53}$$

and

$$V_H(f,t;f',t') = E[D_H(f,t)D_H^*(f',t')] = R_H(f,t;f',t') - m_H(f,t)m_H^*(f',t'). \tag{3.54}$$

3.4.3   Wide-Sense Stationary Uncorrelated Scattering Hypothesis

Throughout this report, we will assume the following two conditions. First of all, we will assume that

$$E[D_h(\tau,t)D_h(\tau',t')] = 0 \tag{3.55}$$

for all f,t and f',t'. Secondly, we will assume that there exists a function $p(\tau,t)$ called *delay cross-power spectral density* such that

$$E[D_h(\tau,t)D_h^*(\tau',t')] = p(\tau,t-t')\delta(\tau-\tau'). \tag{3.56}$$

Consequently, the stochastic process $D_h(\tau,t)$ is wide-sense stationary in t variable and uncorrelated in τ variable. Such conditions are applicable to a wide range of channels.

If a channel has a zero specular component and satisfies the above two relations, then we call it a *wide-sense stationary uncorrelated scattering (WSSUS) channel*. Consequently, we will assume that a general channel system function consists of the specular component and the WSSUS diffuse component. For more theoretical justifications and investigations, please refer to [8, 9, 26, 27, 28]. Note that $p(\tau,0)$ is the delay power profile of the diffuse component. The delay power profile describes how the average received power is distributed with respect to the time delay and hence characterizes the time dispersion of the diffuse component. We note that the baseband representation $h_B(\tau,t) = h(\tau,t)e^{j2\pi f_c\tau}$ also has the same delay cross-power spectral density, namely, that

$$V_{h_B}(\tau,t;\tau',t') = p(\tau,t-t')\delta(\tau-\tau'). \tag{3.57}$$

It is easy to see that

$$V_H(f,t;f',t') = P(f-f',t-t') \tag{3.58}$$

where $P(f,t)$ is the *time-frequency correlation function* defined by

$$P(f,t) = \int_{\mathbb{R}} p(\tau,t)e^{-j2\pi f\tau}\,d\tau. \tag{3.59}$$

Note that the uncorrelatedness in the delay τ implies the wide-sense stationarity in the frequency f. The converse is also true, since the Fourier transform is invertible. Note that $D_H(f,t)$ is wide-sense stationary in both time and frequency.

Similarly, the Doppler delay spread function $k(\tau,\nu)$ has the covariance function V_k satisfying

$$V_k(\tau,\nu;\tau',\nu') = q(\tau,\nu)\delta(\nu-\nu')\delta(\tau-\tau') \tag{3.60}$$

where $q(\nu,\tau)$ is the *scattering function* defined by

$$q(\tau,\nu) = \int_{\mathbb{R}} p(\tau,t)e^{-j\pi\nu t}\,dt. \tag{3.61}$$

We see the wide-sense stationarity of $D_h(\tau,t)$ in time t implies the uncorrelatedness of $D_k(\tau,\nu)$ in frequency ν. Since $D_k(\tau,\nu)$ is the weight of the mode of time variation of the diffuse component, for a given delay τ, the scattering function $q(\nu,\tau)$ defines the average power distribution over the modes of time variation of the diffuse component.

Finally, we see that the Doppler-spread function $K(f,\nu)$ satisfies

$$V_K(f,\nu;f',\nu') = Q(f-f',\nu)\delta(\nu-\nu') \tag{3.62}$$

where $Q(f,\nu)$ is the *Doppler cross-power spectral density* defined by

$$Q(f,\nu) = \int_{\mathbb{R}} q(\tau,\nu)e^{-j2\pi f\tau}\,d\tau. \tag{3.63}$$

Note that

$$P(f,t) = \iint_{\mathbb{R}^2} q(\tau,\nu)e^{-j2\pi(f\tau-\nu t)}\,d\tau\,d\nu \tag{3.64}$$

$$= \int_{\mathbb{R}} Q(f,\nu)e^{j2\pi\nu t}\,d\nu. \tag{3.65}$$

3.4.4   Time Selectivity and Correlation Time

We often say that a channel is *time non-selective* if it is time-invariant. Similarly, if a channel is time-variant, we say that it is *time selective*. To characterize how fast the channel changes, we often use the concept of correlation time. For simplicity, we first consider the case in which all reflected, diffracted, and scattered signals arrive at the

same time (without distortion). In such a case, we can represent the impulse response $h(\tau, t)$ as

$$h(\tau, t) = \alpha(t)\delta(\tau). \tag{3.66}$$

In this case, the channel is said to be *frequency flat* and acts as a time varying multiplier. Mathematically, we define the *correlation time $T_c(\kappa, t)$ of level $\kappa \in (0, 1)$ at time t* by

$$T_c(\kappa, t) = \sup\left\{\delta > 0 \,\middle|\, \left(\forall x \in [-\delta, \delta]\right)\left(\frac{|E[\alpha(t)\alpha^*(t+x)]|}{\sqrt{E[|\alpha(t)|^2]E[|\alpha^*(t+x)|^2]}} \geq \kappa\right)\right\} \tag{3.67}$$

and the *correlation time $T_c(\kappa)$ of level κ* by

$$T_c(\kappa) = \inf\{T_c(\kappa, t) \mid t\}. \tag{3.68}$$

Intuitively, the correlation time $T_c(\kappa, t)$ of level κ at time t is the largest time duration over which the channel gain remains correlated higher than κ around time t and the correlation time $T_c(\kappa)$ is the largest time duration over which the channel gain remains correlated higher than κ around any time instant. We see that $T_c(\kappa, t) = T_c(\kappa)$ for any t and κ, if the channel is wide-sense stationary. Usually, κ is chosen to be close to 1, in which case the channel can be regarded to be roughly the same over the duration of correlation time.

For time dispersive channels, we can consider the time selectivity of the impulse response $h(\tau, t)$ at each delay τ as in the above description. Certainly, the situation is more complex because we have correlation time at each delay τ. It is not evident that there is a way to efficiently define the overall correlation time of the channel in this case because channel gains at a particular delay may or may not dominate the whole channel.

3.4.5   Frequency Selectivity, Coherence Bandwidth, and Delay Spread

In addition to the time selectivity, multipath fading also exhibits frequency selectivity. A channel is said to be *frequency selective* if it exhibits time dispersion. Again for simplicity, we first consider a frequency selective but time non-selective channel with impulse response $h(\tau)$. Similarly to the case of time selectivity, we define the *coherence bandwidth $B_c(\kappa, f)$ of level $\kappa \in (0, 1)$ at frequency f* by

$$B_c(\kappa, f) = \sup\left\{\delta \,\middle|\, \left(\forall x \in [-\delta, \delta]\right)\left(\frac{|E[H(f)H^*(f+x)]|}{\sqrt{E[|H(f)|^2]E[|H^*(f+x)|^2]}} \geq \kappa\right)\right\} \tag{3.69}$$

and the *coherence bandwidth $B_c(\kappa)$ of level κ* by

$$B_c(\kappa) = \inf\left\{B_c(\kappa, f) \mid f\right\}. \tag{3.70}$$

Again, $T_c(\kappa, f) = T_c(\kappa)$ for any κ and f if $H(f)$ is wide-sense stationary.

Since the amount of frequency selectivity is generally inversely related to that of time dispersion, we often use measures of time dispersion as measures of frequency selectivity. Maximum delay spread and root mean square (*rms*) delay spread are often used as measures of time dispersion. To define these parameters, we first consider the power function $P_h(\tau)$ of $h(\tau)$, which is defined by

$$P_h(\tau) = E[|h(\tau)|^2]. \tag{3.71}$$

The power function P_h is also called the *delay power profile* or *power delay profile* of the channel. We define the *maximum delay spread $\tau_{max}(\eta)$ with threshold $\eta \in (0, 1)$* by

$$\tau_{max}(\eta) = \inf\left\{b - a \,\middle|\, \int_a^b P_h(\tau)\,d\tau \geq \eta \cdot \int_{-\infty}^\infty P_h(\tau)\,d\tau\right\}. \tag{3.72}$$

The threshold η is usually chosen to be very small so that $\tau_{max}(\eta)$ actually means the relative time delay between the first arriving and the last arriving signal components. Although maximum delay spread captures some aspects of the delay profile, it fails to characterize the detailed structure between the first arriving and the last arriving signal components. As a more compelling measure, the *rms delay spread τ_{rms}* is defined by

$$\tau_{rms} = \sqrt{\frac{\int_{-\infty}^\infty \tau^2 P_h(\tau)\,d\tau}{\int_{-\infty}^\infty P_h(\tau)\,d\tau} - \left(\frac{\int_{-\infty}^\infty \tau P_h(\tau)\,d\tau}{\int_{-\infty}^\infty P_h(\tau)\,d\tau}\right)^2}. \tag{3.73}$$

Since the delay profile $P_h(\tau)$ is a non-negative real valued function, we can regard it as a probability density function by suitable normalization. With this identification, we can interpret the τ_{rms} as the variance of the time delay with respect to power spread. Consequently, τ_{max} indicates the width of the time dispersion.

By definition, the coherence bandwidth depends on the choice of the level κ. Consequently, it is often more simple and convenient to use the *rms* delay spread as the measure of frequency selectivity. Moreover, there have been various efforts to relate the coherence bandwidth and the *rms* dely spread. Due to the inverse relation, the

coherence bandwidth B_c of a channel is often defined by

$$B_c = \frac{k}{\tau_{rms}} \tag{3.74}$$

with a suitably chosen constant k. Typical value of k is $1/8$ [29]. However, there are several other choices for k depending on the situations [30].

Chapter 4

Characterization of WSSUS Channels

The performance of wireless communication systems are generally bad for channels without any direct line of sight path. Performance evaluation under such an adverse situation is important not only because it is important to guarantee quality of service but also because such a situation is typical in wireless communications. As discussed in the previous chapter, such a channel will be assumed to be a WSSUS channel. In this chapter, we define a set of parameters, collectively called, normalized mean square covariance (NMSV). We define the normalized frequency mean square covariance (NFMSV) and the normalized time mean square covariance (NTMSV) to characterize the frequency selectivity and the time selectivity of a channel, respectively. Normalized frequency-time mean square covariance (NFTMSV) is defined to characterize the combined frequency and time selectivity of a channel. We show how these parameters can be computed from the scattering function of a channel and we define the effective number of resolvable paths and the effective Doppler spread by reinterpreting the concept of stochastic degree of freedom. We then show how closely these parameters are related to the system performance.

4.1 Definitions

In this section, we define parameters for linear stochastic channels with well-behaved time varying frequency response $H(f,t)$. We first define the normalized frequency mean square covariance (NFMSV) to characterize the frequency selectivity. Then, normalized time mean square covariance (NTMSV) is defined to characterize the time selectivity. For the combined frequency and time selectivity, we define normalized

frequency-time mean square covariance (NFTMSV). In the following, we denote by $V_H(f,t;f',t')$ the covariance function of the frequency response, i.e.,

$$V_H(f,t;f',t') = E[H(f,t)H^*(f',t')] - E[H(f,t)]E[H^*(f',t')]. \qquad (4.1)$$

4.1.1 Normalized Frequency Mean Square Covariance

Let Θ represent a frequency region of interest. Then, we define the *normalized frequency mean square covariance* (NFMSV) $\mathcal{V}_f(\Theta;t)$ of the channel over the frequency region Θ at time t by

$$\mathcal{V}_f(\Theta;t) = \frac{\iint_{\Theta \times \Theta} |V_H(f,t;f',t)|^2 \, df \, df'}{\left[\int_{\Theta} V_H(f,t;f,t) \, df \right]^2}. \qquad (4.2)$$

According to the definition of Appendix 2.3.3, $\mathcal{V}_f(\Theta;t)$ is the NMSV of the stochastic process $\{H(f,t), f \in \Theta\}$ where t is a given constant and Θ is an index set. Consequently, it follows that $\mathcal{V}_f(\Theta;t)$ lies in $[0,1]$ and measures the stochastic degree of freedom inherent in the channel frequency selectivity over the frequency region Θ. Later we define the effective number of resolvable paths by 1/NFMSV for a WSSUS channel. In (4.2), we assume a limiting process discussed in Appendix 2.3.2 if the numerator is ∞ as usual, which we will no longer mention.

4.1.2 Normalized Time Mean Square Covariance

Let T be a time region of interest. Then, the *normalized time mean-square covariance* (NTMSV) $\mathcal{V}_t(f;T)$ of the channel over the time region T at frequency f is defined by

$$\mathcal{V}_t(f;T) = \frac{\iint_{T \times T} |V_H(f,t;f,t')|^2 \, dt \, dt'}{\left[\int_{T} V_H(f,t;f,t) \, dt \right]^2}. \qquad (4.3)$$

Similarly to NFMSV, $\mathcal{V}_t(f;T)$ belongs to $[0,1]$ and measures the stochastic degree of freedom in the channel time selectivity over the time region T. Later we define the effective Doppler spread by 1/NTMSV for a WSSUS channel.

4.1.3   Normalized Frequency-Time Mean Square Covariance

Let Ω be a frequency time region of interest. Then, we define the *normalized frequency-time mean square covariance* (NFTMSV) $\mathcal{V}_{ft}(\Omega)$ of the channel over the region Ω by

$$\mathcal{V}_{ft}(\Omega) = \frac{\iint_\Omega \iint_\Omega |V_H(f,t;f',t')|^2 \, df \, dt \, df' \, dt'}{\left[\iint_\Omega V_H(f,t;f,t) \, df \, dt \right]^2}. \tag{4.4}$$

The NFTMSV $\mathcal{V}_{ft}(\Omega)$ is the NMSV of the process $\{H(f,t),(f,t) \in \Omega\}$ where Ω is the two dimensional index set. Hence, it follows that $\mathcal{V}_{ft}(\Omega)$ is in $[0,1]$ and measures the stochastic degree of freedom in the channel selectivity in the frequency-time region Ω.

Before preceding, consider the case of $\Omega = \Theta \times T$. If $V_H(f,t;f',t')$ does not change appreciably for times in T or frequencies in Θ, then $\mathcal{V}_{ft}(\Omega)$ is approximately the same as $\mathcal{V}_f(\Theta;t)$ $(t \in T)$ or $\mathcal{V}_t(f;T)$ $(f \in \Theta)$, respectively. In particular, we have

$$\lim_{\delta \downarrow 0} \mathcal{V}_{ft}\{\Theta \times (t-\delta, t+\delta)\} = \mathcal{V}_f(\Theta, t) \tag{4.5}$$

and

$$\lim_{\delta \downarrow 0} \mathcal{V}_{ft}\{(f-\delta, f+\delta) \times T\} = \mathcal{V}_t(f;T). \tag{4.6}$$

4.2   NMSV for WSSUS Channels

Although we defined the parameters for any well-behaved stochastic process in Section 4.1, we are only interested in wide-sense stationary uncorrelated scattering (WSSUS) channels in the remainder of this chapter. Consequently, all channels in the remainder of this chapters will be assumed to be WSSUS channels. In this section, we show how to compute the parameters from the scattering function of a given channel. Then, we associate physical meanings with the stochastic degrees of freedom defined from the parameters. For notations and relations of various scattering functions, please refer to Section 3.4.

4.2.1   NMSV and Time-Frequency Correlation Function

In this subsection, we show how we express $\mathcal{V}_f(\Theta;t)$, $\mathcal{V}_t(f;T)$, and $\mathcal{V}_{ft}(\Omega)$ in terms of various channel correlation functions. First of all, since $V_H(f,t;f',t') = P(f-f',t-t')$,

we have

$$\mathcal{V}_f(\Theta; t) = \frac{\iint_{\Theta \times \Theta} \left| P(f - f', 0) \right|^2 df\, df'}{\left[P(0,0) \int_{\Theta} df \right]^2} \tag{4.7}$$

$$\mathcal{V}_t(f; T) = \frac{\iint_{T \times T} \left| P(0, t - t') \right|^2 dt\, dt'}{\left[P(0,0) \int_{T} dt \right]^2} \tag{4.8}$$

and

$$\mathcal{V}_{ft}(\Omega) = \frac{\iint_{\Omega} \iint_{\Omega} \left| P(f - f', t - t') \right|^2 df\, dt\, df'\, dt'}{\left[P(0,0) \iint_{\Omega} df\, dt \right]^2}. \tag{4.9}$$

Since the channel is wide-sense stationary in both time and frequency, NFMSV and NTMSV are independent of the time t and the frequency f, respectively. Consequently, we write $\mathcal{V}_f(\Theta)$ and $\mathcal{V}_t(T)$ instead of $\mathcal{V}_f(\Theta; t)$ and $\mathcal{V}_t(f; T)$, respectively.

4.2.2   NMSV and Scattering Function

It is not difficult to rewrite (4.7) - (4.9) in terms of $p(\tau, t)$, $q(\tau, \nu)$ or $Q(f, \nu)$. Of these variations, the formulae with the scattering function $q(\tau, \nu)$ are the most useful because of the ease of physical interpretation and evident duality. To manifest the symmetry and the role of the scattering function, we define a set of kernel functions $K_f(\Theta; \tau)$, $K_t(\nu; T)$ and $K_{ft}(\Omega; \nu, \tau)$ by

$$K_f(\Theta; \tau) = \frac{\left| \int_{\Theta} e^{j2\pi f \tau} df \right|^2}{\left| \int_{\Theta} df \right|^2} \tag{4.10}$$

$$K_t(\nu; T) = \frac{\left| \int_{T} e^{-j2\pi \nu t} dt \right|^2}{\left| \int_{T} dt \right|^2} \tag{4.11}$$

and

$$K_{ft}(\Omega; \nu, \tau) = \frac{\left| \iint_{\Omega} e^{j2\pi(f\tau - \nu t)} df\, dt \right|^2}{\left| \iint_{\Omega} df\, dt \right|^2}. \tag{4.12}$$

Note that if $\Omega = \Theta \times T$, then $K_{ft}(\Omega; \nu, \tau) = K_f(\Theta; \tau) K_t(\nu; T)$.

As shown in Appendix B, we have

$$\mathcal{V}_f(\Theta) = \frac{\iint_{\mathbb{R}^2} p^*(\tau,0)K_f(\Theta,\tau-\tau')p(\tau',0)\,d\tau\,d\tau'}{\left[\int_{\mathbb{R}} p(\tau,0)\,d\tau\right]^2} \tag{4.13}$$

$$\mathcal{V}_t(T) = \frac{\iint_{\mathbb{R}^2} Q^*(0,\nu)K_t(\nu-\nu';T)Q(0,\nu')\,d\nu\,d\nu'}{\left[\int_{\mathbb{R}} Q(0,\nu)\,d\nu\right]^2} \tag{4.14}$$

and

$$\mathcal{V}_{ft}(\Omega) = \frac{\iiiint_{\mathbb{R}^4} q^*(\tau,\nu)K_{ft}(\Omega;\nu-\nu',\tau-\tau')q(\tau',\nu')\,d\tau\,d\nu\,d\tau'\,d\nu'}{\left[\iint_{\mathbb{R}^2} q(\tau,\nu)\,d\tau\,d\nu\right]^2}. \tag{4.15}$$

Note that $\int_{\mathbb{R}} q(\tau,\nu)\,d\nu = p(\tau,0)$ is the delay power profile of the channel. Consequently, NFMSV can be easily obtained from the delay power profile. Similarly, $\int_{\mathbb{R}} q(\tau,\nu)\,d\tau = Q(0,\nu)$ is the Doppler Power Spectrum of the channel and hence NTMSV can be obtained from the Doppler Power Spectrum of the channel.

From (4.13) - (4.15), we see that the NMSV's are dependent not only on the scattering function but also on the regions of interest through the kernel functions. Although the kernel functions can be calculated in principle for any well-shaped regions, we are mostly interested in rectangular shaped regions. Consider a frequency region Θ_r defined by

$$\Theta_r = \bigcup_{k=0}^{K-1} [A + k(v+w), A + k(v+w) + w]. \tag{4.16}$$

where A, v and w are non-negative real numbers. So Θ_r is a finite union of finite intervals. The meaning of the choice of these regions will be apparent in the later sections. It is not difficult to show (see Appendix C)

$$K_f(\Theta_r;\tau) = \left[\frac{\sin\{\pi K(w+v)\tau\}}{K\sin\{\pi(w+v)\tau\}}\right]^2 \mathrm{sinc}^2(\pi w\tau). \tag{4.17}$$

In particular,

$$K_f([A,B];\tau) = \mathrm{sinc}^2\{\pi(B-A)\tau\}. \tag{4.18}$$

From the definition, we see that NFMSV and NTMSV are properties at a specific time and at a specific frequency, respectively, while NFTMSV is an overall channel

property. Normally there is not a simple relation between these three parameters. However, one simple case is worth mentioning in which we can factorize the scattering function $q(\tau, \nu)$ into the delay power profile $p(\tau)$ and the Doppler power spectrum $Q(\nu)$ as

$$q(\tau, \nu) = p(\tau)Q(\nu). \tag{4.19}$$

Delay power profile is a function of the distribution of the signal travel distance, while Doppler power spectrum is related to the angular distribution of the received signal power. Consequently, such a factorization is possible only if the angular distribution of the scatters is independent of the signal travel distance. In this case, it follows that

$$\mathcal{V}_f(\Theta) = \frac{\iint_{\mathbb{R}^2} K_f(\Theta; \tau - \tau')p(\tau)p^*(\tau')\, d\tau'\, d\tau}{\left[\int_{\mathbb{R}} p(\tau)\, d\tau \right]^2} \tag{4.20}$$

$$\mathcal{V}_t(T) = \frac{\iint_{\mathbb{R}^2} K_t(\nu - \nu'; T)Q(\nu)Q^*(\nu')\, d\nu'\, d\nu}{\left[\int_{\mathbb{R}} Q(\nu)\, d\nu \right]^2} \tag{4.21}$$

and

$$\mathcal{V}_{ft}(\Theta \times T) = \mathcal{V}_f(\Theta)\mathcal{V}_t(T). \tag{4.22}$$

Although this factorization is usually not applicable to practical channels, this shows product relationship between NFMSV, NTMSV and NFTMSV. This implies that a channel has higher frequency-time selectivity when it exhibits both frequency and time selectivities.

4.2.3   NFMSV and Delay Power Profile

As pointed out previously, NFMSV can be obtained from the delay power profile of the channel. In this subsection, we investigate the intuitive meaning of NFMSV in relation to the delay power profile. Consider a frequency region $\Theta = [A, B]$ of a WSSUS channel with delay power profile

$$p(\tau, 0) = \sum_{l=0}^{L-1} p_l \delta(\tau - \tau_l). \tag{4.23}$$

where the delay times τ_l's are chosen to be equally spaced. Assume that the delay power profile is properly normalized so that $\sum p_l = 1$. Then, the NFMSV $\mathcal{V}_f([A, B])$ is

given by

$$\mathcal{V}_f([A, B]) = \sum_{l=0}^{L-1} \sum_{l'=0}^{L-1} \operatorname{sinc}^2 \{\pi(B-A)(\tau_l - \tau_{l'})\} p_l p_{l'}. \tag{4.24}$$

Often we employ a tapped delay-line channel model choosing $\tau_l = l/(B-A)$ for each l in (4.23). In this case, the NFMSV has the form

$$\mathcal{V}_f = \sum_{l=0}^{L-1} p_l^2, \tag{4.25}$$

which can be rewritten as

$$\mathcal{V}_f = \frac{1}{L} + \sum_{l=0}^{L-1} \left(p_l - \frac{1}{L} \right)^2. \tag{4.26}$$

We clearly see from (4.26) that $\mathcal{V}_f$ is directly related to the fluctuation of p_l over l. Since the NFMSV is smaller for more evenly distributed power profile, we can see the channel has smaller NFMSV if it has more non-negligible paths. Hence we regard $1/\mathcal{V}_f(\Theta)$ as the effective number of dominant paths and define the *effective number L_e of resolvable paths* of a WSSUS channel in the frequency region Θ by

$$L_e = \frac{1}{\mathcal{V}_f(\Theta)}. \tag{4.27}$$

Of course, the number L_e is not necessarily an integer. However, this number gives a simple and intuitive view of the frequency selectivity in the context of time dispersion.

4.2.4   NFMSV and *rms* Delay Spread

To contrast NFMSV with the *rms* delay spread and the coherence bandwidth as an index of frequency selectivity, we consider a frequency band $[A, B]$ of a WSSUS channel. Assume that the delay power profile of the channel is given by

$$p(\tau, 0) = p_1 \delta(\tau) + p_2 \delta(\tau - 10\beta) + p_3 \delta(\tau - 20\beta) \tag{4.28}$$

where $\beta = 1/(B-A)$. Now consider two simple cases :

$$\text{Case-1:} \quad p_1 = 0.5, \quad p_2 = 0.5 \quad \text{and } p_3 = 0 \tag{4.29}$$

$$\text{Case-2:} \quad p_1 = 0.5, \quad p_2 = 0 \quad \text{and } p_3 = 0.5 \tag{4.30}$$

We can immediately tell that, with Case-1, the *rms* delay spread is smaller and the coherence bandwidth is larger. To see the situation more closely, let's plot
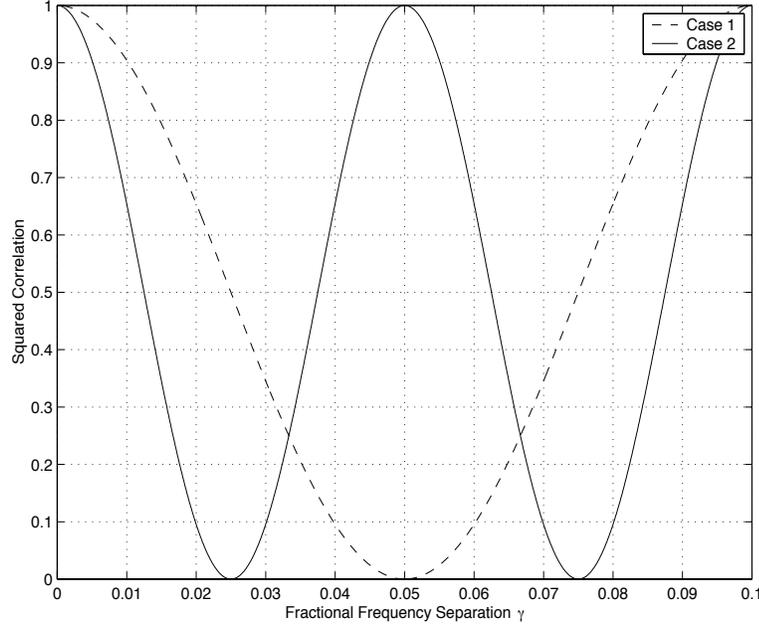
Figure 4.1: Squared Correlation $|E[H(t,f)H^*(t,f')]|^2$ versus the fractional frequency separation $\gamma = (f-f')/(B-A)$. A channel with two equal strength paths is considered. The separation between the two paths in the case 2 is twice as much as that in the case 1. In the case 2, we have larger rms delay spread and smaller coherence bandwidth. However, the NFMSV is the same for the two cases.

$|E[H(t,f)H^*(t,f')]|^2 = |P(f-f',0)|^2$ as a function of the fractional frequency separation $\gamma = (f-f')/(B-A)$. It is easy to show

$$|P(f-f',0)|^2 = \begin{cases} \cos^2(10\pi\gamma), & \text{Case-1,} \\ \cos^2(20\pi\gamma), & \text{Case-2.} \end{cases} \tag{4.31}$$

Note the periodicity of the correlation. We see that Case-1 gives larger coherence bandwidth. However, because the correlation is not monotonic, it is not easy to tell which case exhibits more overall correlation. This is the reason why the coherence bandwidth or the rms delay spread is not a suitable parameter for describing the overall frequency selectivity when the bandwidth is relatively large. In fact, the NFMSV is 0.5 for both cases.

4.2.5   NTMSV and Doppler Spectrum

In the previous subsection, we studied how the NFMSV is related to the delay power profile of a channel. Because of the duality between NFMSV and NTMSV, which is

evident in (4.7) and (4.8) or in (4.13) and (4.14), we can draw a completely similar conclusion about the NTMSV. We define the *effective Doppler spread B_e* of a WSSUS channel in the time region T by

$$B_e = \frac{1}{\mathcal{V}_t(T)}.\tag{4.32}$$

Similarly to NFMSV, the NTMSV is related to the flatness of the Doppler power spectrum. With more flat Doppler power spectrum, we have smaller NTMSV and hence expect more time selectivity. However, there are subtle physical differences between the delay power profile and Doppler power spectrum. The delay power profile is basically determined by the surrounding environments. However, not only the environments but also the movements of the transmitter or the receiver are crucial factors determining the Doppler power spectrum. More specifically, the speed of the movement affects the width of the Doppler power spectrum, while the direction alters the shape. Moreover, the properties of transmitter and receiver antennae play fundamental roles in the determination of the Doppler power spectrum. Consequently, the Doppler power spectrum is not a function only of the surrounding environments.

4.3   Performance of Simple Diversity Combining

In the previous sections, we have shown conceptually how closely the NMSV is related to the selectivity and the degrees of freedom of a WSSUS channel. So it is natural to expect close relationship between the NMSV and the performance of systems designed to exploit the selectivity. Since diversity combining plays fundamental role in most practical systems designed to exploit the channel selectivity, we first consider the relationship between the NFMSV and the performance of a simple diversity combining scheme for frequency selective fading channels. The system we consider can be described as a multicarrier communication system with a repetition code. The receiver detects the symbol with a decision statistic formed by the maximum ratio combining scheme. We show through analysis that there exists close relationship between the NFMSV and the system performance. In particular, we show that the variance of the decision statistic is directly related to the NFMSV. This is important because the fluctuation or the variance of the decision statistic should depend fundamentally on the degrees of freedom available for diversity combining.

   In the remainder of this chapter, we will confine ourselves only to discrete-time circularly symmetric complex Gaussian WSSUS channels with a finite number of resolvable paths. The fading levels at different times for a given path are realizations of a zero mean circularly symmetric complex Gaussian random variable and are cor-

related according to the Doppler power spectrum corresponding to the path. All the systems considered are packet-based and the fading levels corresponding to different packets are assumed to be independent throughout this section.

4.3.1   System Description

We assume that a frequency selective Rayleigh fading channel is given and consider a diversity combining scheme using a repetition code. In this scheme, a binary information symbol is modulated by binary phase shift keying (BPSK) and is transmitted K times over K different narrow-band frequency regions. We assume the fading is constant over each of the narrow-band regions. If the fading levels of the K regions are independent and identically distributed, the diversity order combined by the repetition code is defined to be K. In actual situations, the fading levels are correlated so that the amount of diversity is decreased. We show through system performance analysis that the amount of diversity exploited by the system can be characterized by the NFMSV of the channel. We denote by $H_1, \cdots, H_K$ the fading levels of the K narrow-band regions. The receiver is assumed to have all channel information such as symbol timing and the channel gains $H_1, \cdots, H_K$. If the symbol $+1$ (after BPSK modulation) is transmitted and additive Gaussian noise $\{n_k\}$ is added at the receiver, the decision statistic S can be written as

$$S = \mathrm{Re}\left[\sum_{k=1}^{K} H_k^*(\sqrt{E_b}H_k + n_k) \right] \qquad (4.33)$$

where E_b is the received energy per information bit. We assume, as usual, that the noises are independent and circularly symmetric with mean zero and variance N_0 so that $E[n_i^* n_j] = N_0 \delta_{ij}$ and $E[n_i n_j] = 0$. The receiver decides $+1$ is transmitted if $S > 0$ and -1 otherwise. Note that $H_1, \cdots, H_K$ are correlated complex Gaussian random variables with zero mean. Let $\mathbf{C}$ denote the correlation matrix of the random vector $\mathbf{H} = [H_1, \cdots, H_K]^t$, namely, let $\mathbf{C} = \mathrm{E}[\mathbf{H}\mathbf{H}^\dagger]$. [1] We assume that $\mathrm{E}[\mathbf{H}\mathbf{H}^t] = 0$, which is the case when the channel is an uncorrelated scattering channel. Consequently, we are considering a complex Gaussian wide-sense stationary uncorrelated scattering channel, which is a widely adopted channel model for a mobile communication system evaluation. The fading levels H_k's are normalized so that $\mathrm{tr}\{\mathbf{C}\} = \sum_{k=1}^{K} \mathrm{E}[|H_k|^2] = 1$. The normalized frequency mean square covariance $\mathcal{V}_f(\Theta)$ over the region Θ consisting

---

[1] By $\mathbf{M}^t$ and $\mathbf{M}^\dagger$ we denote the transpose and the Hermitian conjugate of a complex-valued matrix $\mathbf{M}$.

of the K narrow-bands is given by

$$\mathcal{V}_f(\Theta) = \sum_{k=1}^{K} \sum_{l=1}^{K} |C_{kl}|^2. \tag{4.34}$$

4.3.2 Decision Statistic and NFMSV

We first show that the variance of the decision statistic is directly related to the NFMSV over the region. We will assume that $b = +1$ is transmitted.

Lemma 4.1.
The mean $m(S)$ and the variance $V(S)$ of the decision statistics S are given by

$$m(S) = \sqrt{E_b} \tag{4.35}$$

and

$$V(S) = E_b \cdot \mathcal{V}_f(\Theta) + \frac{N_0}{2}. \tag{4.36}$$

Proof. It is easy to see that

$$E[S] = \sqrt{E_b} \cdot \sum_{k=1}^{K} E[|H_k|^2] = \sqrt{E_b} \tag{4.37}$$

and that

$$E[S^2] = E_b \cdot E\left[\left(\sum_{k=1}^{K} |H_k|^2\right)^2\right] + \frac{1}{4} \cdot E\left[\left\{\sum_{k=1}^{K} \left(H_k^* n_k + H_k n_k^*\right)\right\}^2\right]. \tag{4.38}$$

Since $E[n_k n_l] = 0$ for any k and l, the second term of the righthand side of (4.38) reduces to

$$\frac{1}{4} \cdot \sum_{k=1}^{K} \sum_{l=1}^{K} E\left[\left(H_k^* n_k + H_k n_k^*\right)\left(H_l^* n_l + H_l n_l^*\right)\right] \tag{4.39}$$

$$= \frac{1}{4} \sum_{k=1}^{K} \sum_{l=1}^{K} \left\{E[H_k^* H_l]E[n_k n_l^*] + E[H_k H_l^*]E[n_k^* n_l]\right\} \tag{4.40}$$

$$= \frac{N_0}{2}. \tag{4.41}$$

Since $E[H_k H_l] = 0$ for any k and l, we have

$$E\left[\left\{\sum_{k=1}^{K} |H_k|^2\right\}^2\right] = \sum_{k=1}^{K}\sum_{l=1}^{K} E[H_k^* H_k H_l^* H_l] \tag{4.42}$$

$$= \sum_{k=1}^{K}\sum_{l=1}^{K}\left\{E[H_k^* H_k]E[H_l^* H_l] + E[H_k^* H_l^*]E[H_k H_l] + E[H_k^* H_l]E[H_l^* H_k]\right\} \tag{4.43}$$

$$= \left\{\sum_{k=1}^{K} E[|H_k|^2]\right\}^2 + \sum_{k=1}^{K}\sum_{l=1}^{K} |C_{kl}|^2 \tag{4.44}$$

$$= 1 + \mathcal{V}_f(\Theta) \tag{4.45}$$

where we used the fact that

$$E[ABCD] = E[AB]E[CD] + E[AC]E[BD] + E[AD]E[BC] \tag{4.46}$$

for zero mean jointly Gaussian random variables A, B, C, and D [31]. Consequently, it follows

$$V(S) = E[S^2] - E^2[S] = E_b \cdot \mathcal{V}_f(\Theta) + \frac{N_0}{2}. \tag{4.47}$$

$\square$

The performance of the system depends on how often the decision statistic falls below 0, which is generally closely related to the variance of the decision statistic S. Consequently, given the energy E_b and the noise level N_0, the performance of the system will depend heavily on the NMSV. As a special case, consider the asymptotic case when S can be approximated as a Gaussian random variable. Such a situation happens when there is a large degree of freedom in $\mathbf{H}$ which implies small NFMSV. Since S is assumed to be a Gaussian random variable, the probability of error P_b is given by

$$P_b = Q\left(\frac{m(S)}{\sigma(S)}\right) = Q\left(\frac{1}{\sqrt{\mathcal{V}_f(\Theta) + \frac{N_0}{2E_b}}}\right) \tag{4.48}$$

where the function $Q(x)$, which we shall call the *Standard Normal Error Function*, represents the complimentary distribution function of a standard normal random variable that is defined by

$$Q(x) = \int_x^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}\, dt. \tag{4.49}$$

This shows that the probability P_b of bit error is monotonically increasing with $\mathcal{V}_f(\Theta)$. In particular, the performance approaches that under an additive white Gaussian

noise channel as $\mathcal{V}_f(\Theta)$ approaches 0, which is anticipated.

4.3.3   Exact Probability Decision Error

The validity of Gaussian assumption of S depends on the signal-to-noise ratio (SNR) as well as the NFMSV. Smaller NFMSV is required with higher SNR. This means that (4.48) becomes invalid as SNR become larger. More precisely, it is valid only if $\mathcal{V}_f(\Theta) \ll N_0/E_b$. In fact, it is possible to derive the exact bit error rate. As shown in Appendix D, the exact bit error probability P_b is given by

$$P_b = \frac{1}{2\pi} \int_0^\infty \mathrm{Re}\left[\frac{\psi_R(\beta)}{\frac{E_b}{N_0} + j\beta + \sqrt{\frac{E_b}{N_0}\left(\frac{E_b}{N_0} + j\beta\right)}} \right] d\beta \tag{4.50}$$

where $\psi_R(\beta)$ is the characteristic function of the random variable $R = \sum_{k=1}^{K} |H_k|^2$. The characteristics function $\psi_R(\beta)$ is given by [32]

$$\psi_R(\beta) = \prod_{k=1}^{K} \frac{1}{1 - j\mu\lambda_k} \tag{4.51}$$

where $\lambda_1, \cdots, \lambda_K$ are the eigenvalues of the correlation matrix $\mathbf{C}$.

Although (4.50) is exact, it is not evident from this formula how closely P_b is related to $\mathcal{V}_f(\Theta)$. To illustrate the relationship, we evaluate the (4.50) for some practical channels. We consider 71 discrete-time delay power profiles obtained from actual measurements. 41 of them have 240 taps, 3 have 300 taps and remaining 27 have 400 taps. For all 71 profiles, the tap-spacing is 31.25 nsec. We assume that the system uses a contiguous 10.375 MHz bandwidth with 332 equally sliced regions.[2] We assume that the length K of the repetition code used is 332 and each coded symbol is transmitted once over one of the 332 narrow-band regions. Consequently, the whole 10.375 MHz frequency band is utilized for the single information bit transmission. The frequency region Θ of interest is the whole frequency band of 10.375 MHz.

We tend to expect less overall correlation with larger rms delay spread. However, the relation is generally poor for wide-band channels as illustrated in Figure 4.2, in which each dark circle corresponds to one of 71 delay profiles. In particular, rms delay spread is a parameter independent of the bandwidth of the frequency region under consideration. We will show that there exists very close relationship between NFMSV

---

[2]Since we assume that the channel is a WSSUS channel, the situation does not depend on the carrier frequency.
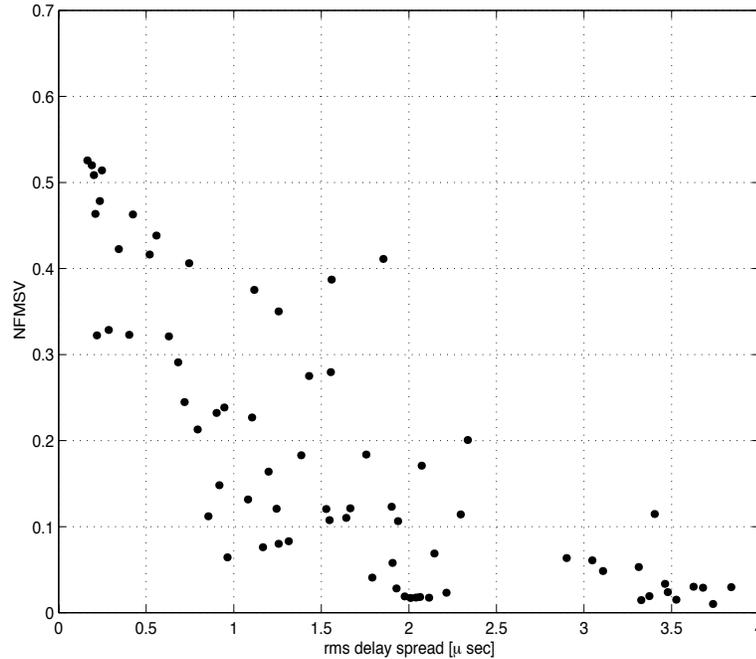
Figure 4.2: Relation between NFMSV and rms delay spread. Each point corresponds to the (rms delay spread, NFMSV) pair for one of 71 frequency selective fading channels. There is not a close relation between rms delay spread and NFMSV.

and the performance of various systems. Consequently, such very poor relationship between rms delay spread and NFMSV implies similarly poor relationship between rms delay spread and system performance. To obtain the probability P_b of decision error, we first compute the correlation matrix $\mathbf{C}$ from the Fourier transform of the delay power profile. Then, we compute the eigenvalues and the characteristic function in turn. Finally, the probability of error is computed by (4.50).

Figure 4.3 shows some of the simulation results for various signal-to-noise ratios (SNR's). Again, each dark circle corresponds to the simulation results for one of 71 channels at a particular SNR. We see that the relation between NFMSV and the probability of error is impressive for relatively low SNR. As the SNR grows, the diversity gain differentiates one channel from another in system performance and the relation becomes less ideal. In particular, the relation is relatively weak for channels with relatively mild diversity. As the NFMSV become larger, the channel becomes more like a flat Rayleigh fading channel, while it becomes more close to Gaussian as the NFMSV becomes smaller. Consequently, we usually expect more ideal NFMSV versus performance relationship for extreme values of NFMSV. However, as we see in the figure, the relation is still good for most cases.
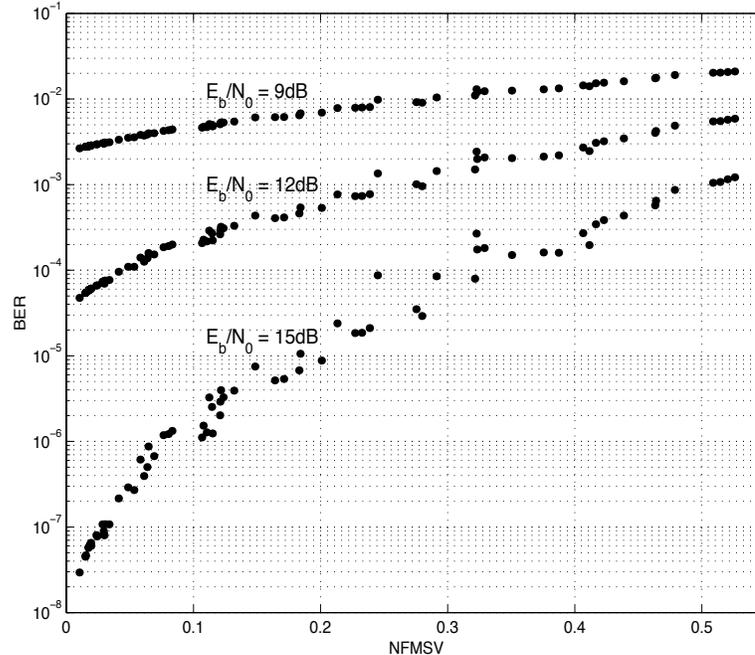
Figure 4.3: BER vs NFMSV for a simple diversity combining scheme. Each point corresponds to the (NFMSV, BER) pair for one of 71 frequency selective fading channel at $E_b/N_0 = 6$, 9 or 12 dB. We observe a roughly monotonic relation between the NFMSV and the BER.

4.4 Peformance of Frequency Hopping Spread Spectrum Systems

In the previous section, we considered the relation between NFMSV and the performance of a simple diversity combining scheme. In this and the next section, we consider more practical systems exploiting the diversity in a fading channel. First, we study the relations between NFMSV and the performance of frequency hopping spread spectrum (FHSS) systems through simulations.

We again consider the 71 WSSUS Rayleigh fading channels considered in the later part of the previous section. We assume that during a packet duration the carrier hops many times among the 332 narrow-band regions comprising the contiguous 10.375 MHz band considered previously. In particular, a pseudo-randomly generated frequency hopping pattern is used. We assume that the system uses BPSK as the modulation. We assume that the frequency selectivity in each of the narrowband regions is negligible. Since frequency hopping alone cannot exploit the frequency diversity, we also consider channel coding schemes. A $(31, 15)$ Reed-Solomon (RS) code and a rate $1/2$ convolutional code are used as examples of block and trellis codes, respectively. In the
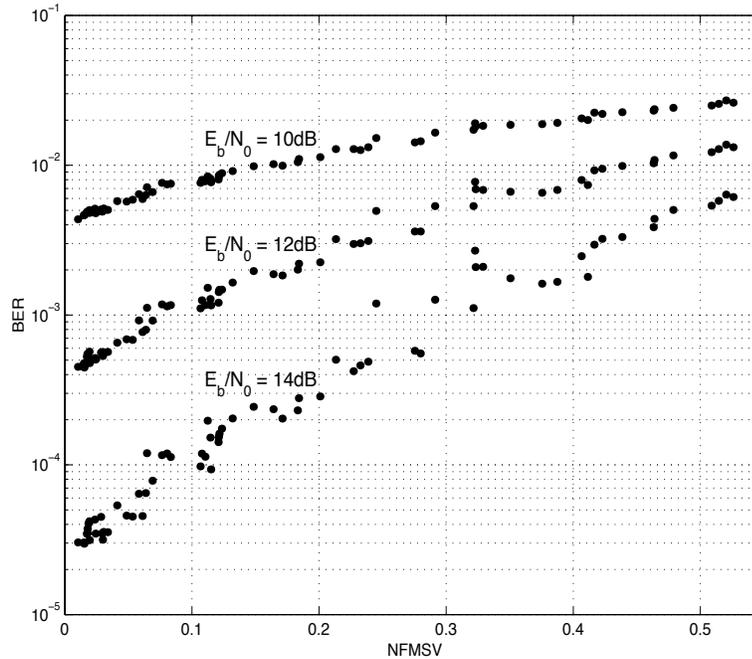
Figure 4.4: BER vs NFMSV for an FHSS system with a (31, 15) RS code. Each point corresponds to the (NFMSV, BER) pair for one of 71 frequency selective fading channel at E_b/N_0 = 10, 12 or 14 dB. We observe a roughly monotonic relation between the NFMSV and the BER.

following, we describe the detailed coding and packet design schemes and provide the simulation results.

4.4.1   Reed-Solomon Code

First, we consider a $(31, 15)$ RS code [33] as an example of a block coding scheme. In this coding scheme, 15 32-ary symbols are encoded to 31 32-ary symbols. Consequently, each codeword consists of 155 coded bits for each 75 information bits. We assume that each packet consists of 20 codewords so that there are 3100 coded bits in a packet, which is RS-symbol interleaved by a block interleaver of size 62 by 10. The coded symbols are written column-wise and are read row-wise. Each row of coded symbols is transmitted over the same frequency slot but different rows are transmitted over randomly and independently chosen frequency slots. This implies the hopping rate is 62 hops/packet and that the hopping pattern is random. Again, we assume that the receiver has full knowledge of the channel. We use, for simplicity, a bounded distance decoding algorithm. In our simulations, we transmit all zero information bits and the receiver makes a hard-decision on each coded bit. We count the number of symbol
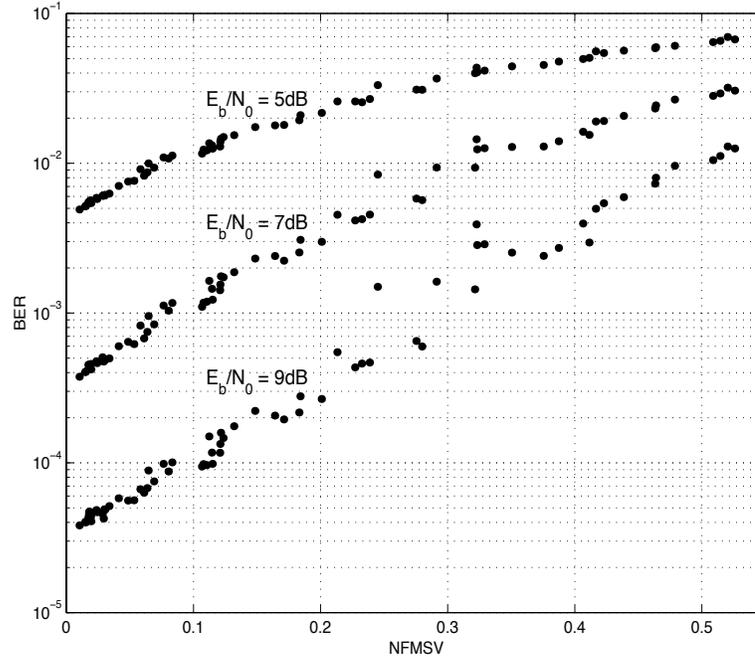
Figure 4.5: BER vs NFMSV for an FHSS system with a (r=1/2, K=5) convolutional code. Each point corresponds to the (BER, NFMSV) pair for one of 71 frequency selective fading channel at E_b/N_0 = 5, 7 or 9 dB.We observe a roughly monotonic relation between the NFMSV and the BER.

errors and assume the errors are not corrected by the decoder if the number of symbol errors exceeds the error correcting capability. We generate at least $20,000$ packets, that is, at least $20,000$ realizations of channel fading levels, and then proceed to generate more until we count $1,000$ bit errors.

Figure 4.4 depicts the relationship between NFMSV and BER of the system. Again, each point corresponds to the (NFMSV, BER) pair for one of the 71 channels at a particular signal-to-noise ratio. We clearly see that there is also very a close relationship between NFMSV and BER performance.

4.4.2   Convolutional Code

In this subsection, as an example of trellis codes, we consider the rate 1/2, constraint length 5 convolutional code with 23 and 35 as the generators in octal from [26]. Each packet is again assumed to have 3100 coded bits which correspond to 1550 information bits, and the output of the convolutional encoder is bit interleaved by a block interleaver of size 62 by 50. Coded bits are written column-wise and are read row-wise. Again each row of coded bits is transmitted over the same frequency slot and differ-

ent rows will be transmitted over randomly and independently chosen frequency slots. Because our goal is not performance optimization, no attempt is made for the trellis termination. Consequently, the receiver chooses the most probable sequence considering all memory states at the end because the trellis does not terminate at a specific memory state. The receiver is assumed to have all information about the channel and uses the soft-decision Viterbi decoding algorithm.

Figure 4.5 shows the simulation results. Each point corresponds to the (NFMSV, BER) pair for one of the 71 channels at a particular signal-to-noise ratio. Again we find there is also very a strong relation between NFMSV and the BER performance.

4.5   Performance of a Direct Sequence Spread Spectrum System

In this section, we investigate, through simulations, the relationship between the NFTMSV and the performance of a direct sequence spread spectrum (DSSS) system with successive interference cancellation (SIC). By spreading, DSSS systems obtain more resolvable paths for rake reception. This is how DSSS benefits from the frequency selectivity of a fading channel. However, frequency selectivity implies intersymbol interference, too. To mitigate the intersymbol interference, we use multi-stage successive interference cancellation. In the simulations, each packet consists of 2,000 binary symbols with spreading gain 25. Each set of 2,000 binary symbols are interleaved by a random interleaver and then spread by random spreading code. The packet of 50,000 chips is then modulated by binary phase shift keying. The signaling rate is assumed to be 5 mega-chips per second. Each packet is then passed to the receiver through a complex Gaussian WSSUS fading channel. At the receiver, additive white Gaussian noise is added to the signal received from the fading channel. The receiver is assumed to have all the channel state information and employs rake reception. After detecting the first symbol using the usual threshold test, its (estimated) contribution is canceled from the received signal (assuming the decision on the first symbol is correct) and then the second symbol is detected. After detecting the second symbol, its contribution is canceled similarly and then detection is made for the third symbol and so on. This is the first stage of the SIC employed in our system. After detecting the last (i.e. 2000th) symbol and subtracting its contribution, we obtain noise estimates which equals the actual noise if all symbols were correctly detected. With this noise estimate, we start the second stage SIC. In the second stage, we first add to the noise estimate the contribution due to the first symbol assuming that the detection in the

first stage is correct. Then we have noise estimates plus the signals due to the first symbol regardless whether the decision for the first symbol in the first stage is correct or not. Now, assuming this signal as the received signal, the rake receiver detects the first symbol and subtracts its contribution again. Then we have new noise estimate to which we add the contribution due to the second symbol and proceed similarly. By repeating this process, we can perform as many stages of SIC as we want. In this work, we stop at the end of the 2nd stage and calculate the bit error rate (BER) as the performance measure of the system under a given channel.

For simulations, we consider 200 discrete-time complex Gaussian WSSUS channels, for which we randomly generate 200 scattering functions. The number of resolvable paths are chosen between 1 and 30 and the delay spreads are chosen between 0 and 20 microseconds. The separations between adjacent paths are chosen not to be uniform but random. The relative amplitude gain and Doppler spectrum are also chosen randomly for each path. To generate the Doppler spectrum, we first consider the Jakes' spectrum [34] of uniformly distributed scatters with Doppler frequency $f_m = 200$ Hz and randomly choose a scale factor s between 0 and 1. Then, we truncate from the Jakes' spectrum the region outside the frequencies between $-sf_m$ and sf_m, which is then properly normalized. The resultant Doppler Spectrum $S(f)$ is

$$S(f) = \begin{cases} \dfrac{2\arcsin s}{\sqrt{f_m^2 - (f - f_c)^2}} & \text{for } |f - f_c| < sf_m \\ 0 & \text{otherwise,} \end{cases} \tag{4.52}$$

where f_c is the carrier frequency of the signal. When the carrier frequency is 2 GHz, the Doppler frequency 200 Hz corresponds to the vehicular speed 30 m/s. Among the 200 channels, 50 are chosen to be time non-selective and 50 are frequency non-selective. And 100 channels are chosen to have both frequency and time selectivities. For these 200 channels, we calculate the NFTMSV using the (4.15) assuming the system occupies exactly 5 MHz bandwidth. To calculate the BER of the system under each of the 200 channels, we perform Monte Carlo simulations. We calculate the number of symbol errors in the first 1000 packets received. If the number of errors exceeds 1000, then we calculate the BER. Otherwise we proceed to generate more packets until we count at least 1000 symbol errors and then calculate the BER.

The results of the simulations are depicted in Figure 4.6. Simulations are performed for three different signal-to-noise ratios (SNR's). In the figure, E_b stands for the energy per bit and N_0 for the one-sided noise power spectral density (PSD) of the additive white Gaussian noise at the receiver. Each point corresponds to the (NFTMSV,
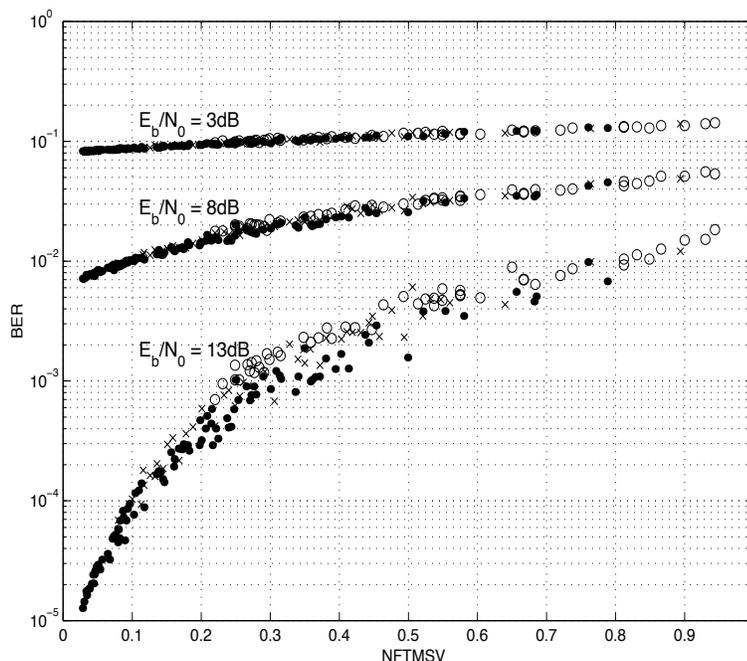
Figure 4.6: BER vs NFTMSV for a DSSS system with 2-stage SIC. Each point corresponds to the (NFTMSV, BER) pair for one of 200 channels at $E_b/N_0 = 3$, 8 or 13 dB. Circles correspond to the channels with time selectivity only, while x-mark to the channels with frequency selectivity only. The solid dots are results for the channels with both frequency and time selectivities. We observe a roughly monotonic relation between the NFTMSV and the BER

BER) pair for each of the 200 channel at $E_b/N_0 = 3, 8$, or 13 dB. We observe that the BER does not depend heavily on the detailed shape of the scattering functions if the NFTMSV's are the same. In particular, the performance does not depend much on whether the channel has time or frequency selectivity if the NFTMSV is the same. We can also see there exists a roughly monotonically increasing relationship between the NFTMSV and the BER. This supports the claim that the NFTMSV represents the diversity of a fading channel and the system exploits the diversity. As a minor point, we note that the BER is slightly higher with the channel with time selectivity only. As mentioned earlier the system exploits the time selectivity by the interleaver and the repetition. Hence, the results imply the random interleaver considered is not effective. In contrast, we can conclude that the 2-stage SIC is quite effective in treating the intersymbol interference.

4.6   Summary and Conclusion

In this chapter, we defined three parameters called normalized frequency mean square covariance (NFMSV), normalized time mean square covariance (NTMSV), and normalized frequency-time mean square covariance (NFTMSV) for multipath fading channels. It is not difficult to see directly from the definitions that these parameters characterize the overall frequency and/or time selectivities of a channel.

As is well known, many communication systems such as spread spectrum systems exploit the selectivity to combat the effect of multipath fading. We showed that there exist very close relationship between the NMSV's and the performance of such systems. First we illustrated the NFMSV is directly related to the variance of the decision statistic of a simple diversity combing scheme under frequency selective fading channels. Since the fluctuations of the decision statistic of the diversity combining scheme should fundamentally depend on the amount of stochastic degree of freedom available, the relation shows how well the NFMSV quantifies the degrees of freedom in frequency selective multipath fading. We also verified that there exists very close relationship between NMSV and the performance of more practical communications systems such as FHSS or DSSS systems.

Chapter 5

Characterization of Non-WSSUS Channels

In the previous chapter, we defined a set of parameters to characterize wide-sense stationary uncorrelated scattering (WSSUS) channels. Although there exist a close relationship between the parameters and the performance of a system under WSSUS channels, they don't exhibit such close relationship under more general channels. In this chapter, by studying the diversity combining schemes, we propose a set of second order statistics collectively called re-centered normalized mean square correlation (RC-NMSR) to summarize the fading characteristics. We illustrate that the system performance is strongly related to the RC-NMSR for channels with identical specular components. Since the specular component consists usually of the direct line of sight path, this implies that RC-NMSR effectively characterizes a multipath fading channel. In the cases when the specular component is more complicated, we use another set of parameters collectively called, normalized arithmetic variance (NAV) in addition to the RC-NMSR.

5.1 Definition

In this section, we define two sets of parameters collectively called re-centered normalized mean square correlation (RC-NMSR) and normalized arithmetic variance (NAV). For channels with the same specular component, RC-NMSR's effectively characterize multipath fading channels. It is natural to expect that the performance of systems is not appreciably different under two channels with similar specular components and fading characteristics. As the RC-NMSR's provide measures of similarities of fading characteristics, the NAV's are defined to quantify the similarity among specular com-

ponents.

Throughout this section, we assume a stochastic channel is given with well behaved time-variant frequency response $H(f,t)$. We denote by $m_H(f,t)$ and $R_H(f,t;f',t')$ the mean function and the correlation function of $H(f,t)$, respectively. Consequently,

$$m_H(f,t) = E[H(f,t)] \tag{5.1}$$

and

$$R_H(f,t;f',t') = E[H(f,t)H^*(f',t')]. \tag{5.2}$$

5.1.1   Re-Centered Normalized Frequency Mean Square Correlation

Let Θ denote a frequency region of interest. We define the *re-centered normalized frequency mean square correlation (RC-NFMSR)* $\mathcal{C}_f(\Theta,t)$ of the channel over the frequency region Θ at time t by

$$\mathcal{C}_f(\Theta,t) = \frac{\int_\Theta \int_\Theta \left\{ |R_H(f,t;f',t)|^2 - |m_H(f,t)|^2 |m_H^*(f',t)|^2 \right\} df\, df'}{\left[\int_\Theta R_H(f,t;f,t)\, df \right]^2} \tag{5.3}$$

Note that the RC-NFMSR coincides with the normalized frequency mean square covariance (NFMSV) if the specular component is zero.

When the specular component is given, the RC-NFMSR characterizes the frequency selectivity of the channel over the frequency region Θ at time t. In many important cases, the specular component represents just the direct line of sight path, in which case we don't need any further specification for the specular component. When the specular component is more complicated, we attempt to characterize its frequency selectivity by normalized frequency arithmetic variance. We define the *normalized frequency arithmetic variance (NFAV)* $\nu_f(\Theta,t)$ of the specular component over Θ at time t by

$$\nu_f(\Theta,t) = \frac{\left\langle \left| m_H(f,t) - \langle m_H(f,t) \rangle_\Theta \right|^2 \right\rangle_\Theta}{\left\langle \left| m_H(f,t) \right|^2 \right\rangle_\Theta} \tag{5.4}$$

$$= 1 - \frac{\left| \int_\Theta m_H(f,t)\, df \right|^2}{\int_\Theta df \int_\Theta \left| m_H(f',t) \right|^2 df'}. \tag{5.5}$$

From Sections 2.3.1 and 2.3.4, it follows that

$$0 \leq \nu_f(\Theta, t), \, \mathcal{C}_f(\Theta, t) \leq 1. \tag{5.6}$$

If $\mathcal{C}_f(\Theta, t)$ or $\nu_f(\Theta, t)$ doesn't depend on time t, then we suppress the time index and write $\mathcal{C}_f(\Theta)$ or $\nu_f(\Theta)$.

5.1.2 Re-Centered Normalized Time Mean Square Correlation

Let T denote a time region of interest. Then, we define the *re-centered time mean square correlation (RC-NTMSR)* $\mathcal{C}_t(f, T)$ of the channel over the time region T at frequency f by

$$\mathcal{C}_t(f, T) = \frac{\int_T \int_T \left\{ |R_H(f, t; f; t')|^2 - |m_H(f, t)|^2 |m_H^*(f, t')|^2 \right\} dt \, dt'}{\left[\int_T R_H(f, t; f, t) \, dt \right]^2}. \tag{5.7}$$

Note that the RC-NTMSR coincides with the normalized time mean square covariance (NTMSV) when the specular component is zero.

If the specular component is given, the RC-NTMSR characterizes the time selectivity of the channel over the time region T at frequency f. To characterize the time selectivity of the specular component, we define the normalized time arithmetic variance (NTAV). We define the *normalized time arithmetic variance (NTAV)* $\nu_t(f, T)$ of the specular component over the time region T at frequency f by

$$\nu_t(f, T) = 1 - \frac{\left| \int_T m_H(f, t) \, dt \right|^2}{\int_T dt \int_T |m_H(f, t')|^2 \, dt'}. \tag{5.8}$$

From Sections 2.3.1 and 2.3.4, it follows that

$$0 \leq \nu_t(f, T), \, \mathcal{C}_t(f, T) \leq 1. \tag{5.9}$$

If $\mathcal{C}_t(f, T)$ or $\nu_t(f, T)$ doesn't depend on frequency f, then we suppress the time index and write $\mathcal{C}_t(T)$ or $\nu_t(T)$.

5.1.3   Re-Centered Normalized Frequency-Time Mean Square Correlation

Let Ω denote a frequency-time region of interest. Then, we define the *re-centered normalized frequency-time mean square correlation (RC-NFTMSR)* $\mathcal{C}_{ft}(\Omega)$ of the channel over the frequency time region Ω by

$$\mathcal{C}_{ft}(\Omega) = \frac{\iint_\Omega \iint_\Omega \left\{ |R_H(f,t;f',t')|^2 - |m_H(f,t)|^2 |m_H^*(f',t')|^2 \right\} df\, dt\, df'\, dt'}{\left[\int_\Omega R_H(f,t;f,t)\, df\, dt \right]^2}. \tag{5.10}$$

Note that the RC-NFTMSR coincides with the normalized frequency-time mean square correlation (NFTMSV) when the specular component is zero.

If the specular component is given, the RC-NFTMSR characterizes the combined frequency and time selectivity of the channel over the frequency-time region Ω. To characterize the frequency and time selectivity of the specular component, we define the *normalized frequency-time arithmetic variance (NFTAV)* $\nu_{ft}(\Omega)$ of the specular component over Ω by

$$\nu_{ft}(\Omega) = 1 - \frac{\left| \iint_\Omega m_H(f,t)\, df\, dt \right|^2}{\iint_\Omega df\, dt \ \iint_\Omega |m_H(f',t')|^2\, df'\, dt'}. \tag{5.11}$$

From Sections 2.3.1 and 2.3.4, it follows that

$$0 \leq \nu_{ft}(\Omega), \mathcal{C}_{ft}(\Omega) \leq 1. \tag{5.12}$$

5.2   Complex Guassian WSSUS Deviation

Although the definitions in the previous section are applicable for any randomly time-variant linear channel, we study the relations between the parameters and the system performance for practically most important channels. In the remainder of this chapter, we shall assume the diffuse component satisfies the circularly symmetric complex Gaussian wide-sense stationary uncorrelated scattering channel model. The justifications of these assumptions can be found in many books. For example, refer to Chapter 2 of [9]. More precisely we assume that the diffuse component $D_h(\tau,t)$ of the time

variant impulse response $h(\tau,t)$ satisfies

$$E[D_h(\tau,t)D_h(\tau',t)] = 0 \tag{5.13}$$

and that there is a function $p(\tau,t)$ called the delay cross-power spectral density of the scattering part such that

$$E[D_h(\tau,t)D_h^*(\tau',t')] = p(\tau,t-t')\delta(\tau-\tau'). \tag{5.14}$$

Finally, we assume that both the real and imaginary parts of the diffuse component are Gaussian random processes.

For channels satisfying these assumptions, it is possible to express RC-NMSR's in terms of the specular component and the delay cross-power spectral density of the diffuse component. In particular, $\mathcal{C}_f(\Theta,t)$ can be obtained from the deterministic part and the delay power profile $p(\tau) = p(\tau,0)$ of the diffuse component. However, $\mathcal{C}_t(f,T)$ can be expressed most conveniently by the Doppler power spectrum $Q(f)$ of the scattering part. To derive such relations, we first express $\mathcal{C}_{ft}(\Omega)$ using the scattering function $q(\tau,\nu)$ of the diffuse component defined by

$$q(\tau,\nu) = \int_{-\infty}^{\infty} p(\tau,t)e^{-j2\pi\nu t}\,dt. \tag{5.15}$$

In Appendix E, we show that [1]

$$
\begin{aligned}
\mathcal{C}_{ft}(\Omega) \;=\;& \frac{\iiiint_{\mathbb{R}^4} q^*(\tau,\nu)K_{ft}(\Omega;\nu-\nu',\tau-\tau')q(\tau',\nu')\,d\tau\,d\nu\,d\tau'\,d\nu'}{\left[\iint_{\mathbb{R}^2} q(\tau,\nu)\,d\tau\,d\nu + \left\langle|m_H|^2\right\rangle_\Omega\right]^2} \\
&- 2\,\frac{\mathrm{Re}\left[\iiiint_{\mathbb{R}^4} v_{ft}^*(\Omega;\nu,\tau,\tau')q(\tau,\nu)v_{ft}(\Omega;\nu,\tau,\tau'')\,d\nu\,d\tau\,d\tau'\,d\tau''\right]}{\left[\iint_{\mathbb{R}^2} q(\tau,\nu)\,d\tau\,d\nu + \left\langle|m_H|^2\right\rangle_\Omega\right]^2}
\end{aligned}
\tag{5.16}
$$

where

$$K_{ft}(\Omega;\nu,\tau) = \frac{\left|\iint_\Omega e^{j2\pi(f\tau-\nu t)}\,df\,dt\right|^2}{\left|\iint_\Omega df\,dt\right|^2}. \tag{5.17}$$

---

[1] $\mathbb{R}$ ($\mathbb{C}$) denotes the set of all real (complex) numbers. For any set S and any positive integer q, S^q means the Cartesian product $S \times \cdots S$ (q times).

and

$$v_{ft}(\Omega; \nu, \tau, \tau') = \frac{\iint_\Omega m_h(\tau', t) e^{-j2\pi(f\tau' + \nu t)} e^{j2\pi f\tau}\, df\, dt}{\iint_\Omega df\, dt} \tag{5.18}$$

To calculate $\mathcal{C}_f(\Theta, t)$, we set $\Omega = \Theta \times (t - \delta, t + \delta)$ in (5.16) and take the limit $\delta \to 0$ noting

$$p(\tau) = p(\tau, 0) = \int_{-\infty}^{\infty} q(\tau, \nu)\, d\nu \tag{5.19}$$

Then, we have

$$\mathcal{C}_f(\Theta, t) = \frac{\iint_{\mathbb{R}^2} p^*(\tau) K_f(\Theta, \tau - \tau') p(\tau')\, d\tau\, d\tau'}{\left[\int_{\mathbb{R}} p(\tau)\, d\tau + \langle |m_H|^2 \rangle_\Theta \right]^2}$$
$$- 2 \frac{\mathrm{Re}\left[\iiint_{\mathbb{R}^3} v_f^*(\Theta; \tau, \tau') p(\tau) v_f(\Theta; \tau, \tau'')\, d\tau''\, d\tau'\, d\tau \right]}{\left[\int_{\mathbb{R}} p(\tau)\, d\tau + \langle |m_H|^2 \rangle_\Theta \right]^2} \tag{5.20}$$

where

$$K_f(\Theta; \tau) = \frac{\left| \int_\Theta e^{j2\pi f\tau}\, df \right|^2}{\left| \int_\Theta df \right|^2} \tag{5.21}$$

and

$$v_f(\Theta; \tau, \tau') = \frac{m_h(\tau', t) \int_\Theta e^{j2\pi(\tau - \tau')f}\, df}{\int_\Theta df}. \tag{5.22}$$

Similarly, we put $\Omega = (f - \delta, f + \delta) \times T$ in (5.16) and take the limit $\delta \to 0$ noting

$$Q(\nu) = \int_{-\infty}^{\infty} q(\tau, \nu)\, d\tau. \tag{5.23}$$

Then, we obtain

$$
\mathcal{C}_t(f,T) = \frac{\iint_{\mathbb{R}^2} Q^*(\nu) K_t(\nu - \nu'; T) Q(\nu') \, d\nu d\nu'}{\left[\int_{\mathbb{R}} Q(\nu) \, d\nu + \langle |m_H|^2 \rangle_T \right]^2}
$$
$$
- 2 \frac{\mathrm{Re}\left[\iiint_{\mathbb{R}^3} v_t^*(\nu, \tau'; T) Q(\nu) v_t(\nu, \tau''; T) \, d\tau' \, d\tau'' \, d\nu \right]}{\left[\int_{\mathbb{R}} Q(\nu) \, d\nu + \langle |m_H|^2 \rangle_T \right]^2}
\tag{5.24}
$$

where

$$
K_t(\nu; T) = \frac{\left| \int_T e^{-j2\pi \nu t} \, dt \right|^2}{\left| \int_T dt \right|^2}
\tag{5.25}
$$

and

$$
v_t(\nu, \tau; T) = \frac{e^{-j2\pi f \tau} \int_T m_h(\tau, t) e^{-j2\pi \nu t} \, dt}{\int_T dt}.
\tag{5.26}
$$

5.3   RC-NMSR and Diversity Combining

Normalized mean square covariances (NMSV's) are useful to characterize WSSUS channels because they have close relations with system performance as shown in Chapter 4. However, it is not evident if they exhibit such relations under channels with non-zero specular component. In fact, they don't exhibit a satisfactory relation with system performance as will be illustrated below. In this section, we study the performance of a simple diversity combining scheme, which motivates the definition of RC-NMSR. Diversity combining is probably the most important scheme to combat multipath fading (ref. Chapter 14 of [35]). Direct sequence (DS) or frequency hopping (FH) spread spectrum systems or antenna diversity schemes are among the most popular examples of diversity combing schemes. In a direct sequence spread spectrum, a rake receiver can be used to combine multiple signal components resolved by the wide bandwidth. In a frequency-hopped-spread-spectrum system, the diversity is achieved by using error correction coding. Basically these schemes combine redundant signals to combat multipath fading. This is the motivation for considering the diversity combining scheme with a simple repetition code.

5.3.1   System Description

We consider basically the same diversity combining scheme as in Section 4.3.1. But, this time the channel is assumed to be Ricean. In the following, we assume that a BPSK-modulated signal is transmitted K times over K different carriers. The signaling rate is assumed to be low enough so that each of the K carriers undergoes flat fading and is high enough to assume that the fading is time non-selective. Then the fading levels of the K carriers are described by K complex numbers $H_1, \cdots, H_K$ which are assumed to be K circularly symmetric complex Gaussian random variables. The receiver is assumed to have enough knowledge about the fading levels to combine the K redundant signals using a maximal ratio combining rule. In this case, the decision statistic S, when the modulated symbol $+1$ is transmitted, is given by

$$S = \mathrm{Re}\left[\sum_{k=1}^{K} H_k^*(\sqrt{E_b}H_k + n_k)\right]. \tag{5.27}$$

where E_b is the average received energy per information bit and n_k is the complex representation of the additive noises at the receiver for the kth carrier. We assume that $\{n_k\}$ are independent identically distributed circularly symmetric complex Gaussian random variables with variance N_0 so that

$$E[n_k n_l^*] = N_0 \delta_{kl}, \quad E[n_k n_l] = 0 \tag{5.28}$$

for any k and l. For proper normalization, we assume that $\sum_{k=1}^{K} E[|H_k|^2] = 1$.

The receiver decides $+1$ is transmitted if $S > 0$, -1 otherwise. The basic idea of this (or any diversity combining) scheme lies in the law of large numbers. In other words, the system attempts to get more reliable (or more deterministic) decision variables by combining multiple signal components with diverse fading levels. Consequently, the performance of such a system tends to depend heavily on the variance of the decision statistic.

Let Θ be the frequency region occupied by the K carriers. Then, the RC-NFMSR $\mathcal{C}_f(\Theta)$ over the frequency region Θ is given by

$$\mathcal{C}_f(\Theta) = \sum_{k=1}^{K}\sum_{l=1}^{K}\left\{\left|E[H_k H_l^*]\right|^2 - \left|E[H_k]\right|^2\left|E[H_l]\right|^2\right\}. \tag{5.29}$$

Because the channel is assumed to be time non-selective, we suppress the time index t in $\mathcal{C}_f(\Theta, t)$.

5.3.2   Decision Statistic and RC-NFMSR

In this subsection, we show that the variance of the decision statistic S is directly related to the RC-NFMSR.

Lemma 5.1.

The mean $m(S)$ and the variance $V(S)$ of the decision statistic S are given by

$$m(S) = \sqrt{E_b} \tag{5.30}$$

and

$$V(S) = E_b \cdot \mathcal{C}_f(\Theta) + \frac{N_0}{2}. \tag{5.31}$$

Proof. First, we rewrite the decision statistic S as

$$S = \sqrt{E_b} \, \|\mathbf{H}\|^2 + \frac{1}{2} \{\mathbf{H}^\dagger \mathbf{n} + \mathbf{n}^\dagger \mathbf{H}\} \tag{5.32}$$

where $\mathbf{H} = (H_1, \cdots, H_K)^t$ and $\mathbf{n} = (n_1, \cdots, n_K)^t$ [2].

Before proceeding, we define $\mathbf{M} = E[\mathbf{H}]$ and $\mathbf{D} = \mathbf{H} - \mathbf{M}$. Since $\mathbf{H}$ and $\mathbf{n}$ are independent with $E[\|\mathbf{H}\|^2] = 1$ and $E[\mathbf{n}] = 0$,

$$E[S] = \sqrt{E_b} \, E[\|\mathbf{H}\|^2] = \sqrt{E_b}. \tag{5.33}$$

and

$$E[S^2] = E_b \, E[\|\mathbf{H}\|^4] + \frac{1}{4} E[\{\mathbf{H}^\dagger \mathbf{n} + \mathbf{n}^\dagger \mathbf{H}\}^2] \tag{5.34}$$

Since $E[n_k n_l] = 0$ for any k and l, the second term of (5.34) is given by

$$\frac{1}{4} \sum_{k=1}^{K} \sum_{l=1}^{K} E[(H_k^* n_k + H_k n_k^*)(H_l^* n_l + H_l n_l^*)] \tag{5.35}$$

$$= \frac{1}{4} \sum_{k=1}^{K} \sum_{l=1}^{K} \{E[H_k^* H_l] E[n_k n_l^*] + E[H_k H_l^*] E[n_k^* n_l]\} \tag{5.36}$$

$$= \frac{N_0}{2}. \tag{5.37}$$

---

[2] For a row vector $\mathbf{v} = (a, \cdots, b)$, $\|\mathbf{v}\|$ and $\|\mathbf{v}^t\|$ shall denote $\sqrt{|a|^2 + \cdots + |b|^2}$

Before manipulating the first term of (5.34), note that

$$E[||\mathbf{D}||^4] \tag{5.38}$$

$$= \sum_{k=1}^{K}\sum_{l=1}^{K} E[D_k^* D_k D_l^* D_l] \tag{5.39}$$

$$= \sum_{k=1}^{K}\sum_{l=1}^{K} \left\{ E[D_k^* D_k]E[D_l^* D_l] + E[D_k^* D_l^*]E[D_k D_l] + E[D_k^* D_l]E[D_l^* D_k] \right\} \tag{5.40}$$

$$= \left\{ E[||\mathbf{D}||^2] \right\}^2 + \sum_{k=1}^{K}\sum_{l=1}^{K} |C_{kl}|^2. \tag{5.41}$$

where we used the fact that

$$E[ABCD] = E[AB]E[CD] + E[AC]E[BD] + E[AD]E[BC] \tag{5.42}$$

for zero mean jointly Gaussian random variables A, B, C, and D [31]. Now we have

$$E[||\mathbf{H}||^4] = E[(||\mathbf{M} + \mathbf{D}||^2)^2] \tag{5.43}$$

$$= E[(\mathbf{M}^\dagger\mathbf{M} + \mathbf{D}^\dagger\mathbf{D} + \mathbf{M}^\dagger\mathbf{D} + \mathbf{D}^\dagger\mathbf{M})^2] \tag{5.44}$$

$$= E[(\mathbf{M}^\dagger\mathbf{M} + \mathbf{D}^\dagger\mathbf{D})^2] + E[(\mathbf{M}^\dagger\mathbf{D} + \mathbf{D}^\dagger\mathbf{M})^2] \tag{5.45}$$

$$= \left\{ (\mathbf{M}^\dagger\mathbf{M})^2 + 2\mathbf{M}^\dagger\mathbf{M}E[\mathbf{D}^\dagger\mathbf{D}] + E[(\mathbf{D}^\dagger\mathbf{D})^2] \right\} + 2\mathbf{M}^\dagger E[\mathbf{D}\mathbf{D}^\dagger]\mathbf{M} \tag{5.46}$$

$$= \left\{ \mathbf{M}^\dagger\mathbf{M} + E[\mathbf{D}^\dagger\mathbf{D}] \right\}^2 + \sum_{k=1}^{K}\sum_{l=1}^{K}\{|C_{kl}|^2 + 2M_k^* C_{kl} M_l\} \tag{5.47}$$

$$= E[||\mathbf{H}||^2] + \mathcal{C}_f(\Theta). \tag{5.48}$$

Consequently, the variance of S is given by

$$V(S) = E_b \cdot \mathcal{C}_f(\Theta) + \frac{N_0}{2}. \tag{5.49}$$

$\square$

5.3.3 RC-NFMSR, NFAV, and Probability of Decision Error

Lemma 5.1 shows that the variance of the decision statistic depends on the RC-NFMSR rather than the NFMSV. Consequently, we expect that the performance of the receiver will depend heavily on the RC-NFMSR $\mathcal{C}_f(\Theta)$. This is particularly true when the $\mathcal{C}_f(\Theta)$ is very close to 0 or $1 - \sum_{k=1}^{K} |E[H_k]|^2$. To see this more clearly, we represent the detec-

tion error probability P_e of the receiver by

$$P_e = \int_0^\infty f_R(r) Q\left(\sqrt{\frac{2E_b r}{N_0}}\right) dr \tag{5.50}$$

where $R = \sum_{k=1}^K |H_k|^2$, $f_R(r)$ is the probability density function of the random variable R, and $Q(\cdot)$ denotes the standard normal error function [3]. This shows that the probability of error is directly related to the random variable R, which has a $2K$ dimensional chi-square distribution. Since small $\mathcal{C}_f(\Theta)$ means large degree of freedom in $H_1, \cdots, H_K$, we can expect R to become closer to a Gaussian random variable. In other words, for two channels with the same small $\mathcal{C}_f(\Theta)$, the distribution functions of R under the two channels are expected to be similar, which is not generally true when $\mathcal{C}_f(\Theta)$ is not small. If $\mathcal{C}_f(\Theta) = 1 - \sum_{k=1}^K \left|E[H_k]\right|^2$, R is a one-dimensional non-central chi-square random variable that is completely determined by $\sum_{k=1}^K \left|E[H_k]\right|^2$. Consequently, as $\mathcal{C}_f(\Theta)$ approaches either extremal values, we can expect stronger performance vs. RC-NFMSR relation.

To analyze further, let $\mathbf{H}$ denote the column vector $(H_1, \cdots, H_K)^t$ [4] and define $\mathbf{M}$ and $\mathbf{D}$ by $\mathbf{M} = E[\mathbf{H}]$ and as before $\mathbf{D} = \mathbf{H} - \mathbf{M}$, respectively. Let $\mathbf{C}$ denote the correlation matrix of $\mathbf{D}$, i.e., let $\mathbf{C} = E[\mathbf{D}\mathbf{D}^\dagger]$. Then, $\mathbf{C}$ is a positive semi-definite matrix so that there exist a unitary matrix $\mathbf{U}$ such that

$$\mathbf{U}\mathbf{C}\mathbf{U}^\dagger = \text{diag}(\lambda_1, \cdots, \lambda_K) \tag{5.52}$$

where $\text{diag}(a, \cdots, b)$ denotes a diagonal matrix with diagonal elements $a, \cdots, b$. Here $\lambda_1, \cdots, \lambda_K$ are eigenvalues of $\mathbf{C}$ and are non-negative real numbers. Now, we define $\mathbf{M}' = \mathbf{U}\mathbf{M}$ and $\mathbf{D}' = \mathbf{U}\mathbf{D}$ so that $E[\mathbf{D}'(\mathbf{D}')^\dagger] = \text{diag}(\lambda_1, \cdots, \lambda_K)$. Moreover, (5.13) implies $E[\mathbf{D}'\mathbf{D}'^t] = 0$. These two correlation properties imply that the $2K$ Gaussian random variables $\text{Re}[D_1'], \cdots, \text{Re}[D_K']$ and $\text{Im}[D_1'], \cdots, \text{Im}[D_K']$ are independent with zero mean and variance

$$E[(\text{Re}[D_k'])^2] = E[(\text{Im}[D_k'])^2] = \frac{\lambda_k}{2}, \tag{5.53}$$

---

[3]The function $Q(x)$, which we shall call the *Standard Normal Error Function*, represents the complimentary distribution function of a standard normal random variable that is defined by

$$Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt. \tag{5.51}$$

[4]For a matrix $\mathbf{M}$, $\mathbf{M}^t$ and $\mathbf{M}^\dagger$ denote the transpose and the adjoint matrices, respectively.

for each k. Now we observe that

$$R = \sum_{k=1}^{K} |M'_k + D'_k|^2. \tag{5.54}$$

Consequently, R is a sum of $2K$ independent chi-square random variables. When the specular component, $\mathbf{M'}$, is zero, the random variable R depends solely on the eigenvalues $\lambda_1, \cdots, \lambda_K$. However, when the specular component is not zero, the specular component also affects the distribution of the random variable R. Consequently, we not only have to observe the distribution of the eigenvalues but also have to take into account the specular component. In fact, as shown in Section 4.3.3, P_e can be expressed as

$$P_e = \frac{1}{2\pi} \int_0^\infty \mathrm{Re}\left[\frac{\psi_R(\beta)}{\frac{E_b}{N_0} + j\beta + \sqrt{\frac{E_b}{N_0}\left(\frac{E_b}{N_0} + j\beta\right)}} \right] d\beta. \tag{5.55}$$

Here $\phi_R(\beta)$ is the characteristic function of R and is given by

$$\psi_R(\beta) = \prod_{k=1}^{K} \left[\frac{1}{1 - j\beta\lambda_k} e^{j\frac{|M'_k|^2\beta}{1 - j\beta\lambda_k}} \right] \tag{5.56}$$

as shown in the Appendix F. Note that $|M'_k|$ (and not $|M_k|$) appear in (5.56).

In many cases, the specular component consists of a single strong line of sight path. In such a case, if the ratio of the powers between the specular and the diffuse components is given, the probability of error is usually very closely related to the RC-NFMSR. In other words, if we test the system under two channels with similar RC-NFMSR's and with similar specular v.s. diffuse parts power ratios, then similar system performances are expected. The situation is more complex when the specular component is more complicated. Generally speaking, if the specular component is given, then the system performance is generally very closely related to the RC-NFMSR. If this is true, it is natural to expect similar performances under two channels with the same or similar RC-NFMSR if they have similar looking specular component. So our next job is to characterize the similarity between two specular components. We attempt to do this by observing the numerical distribution of $M_1, \cdots, M_K$. The first order statistic for these numbers is the arithmetic average, which is dependent upon the overall phase change of $M_1, \cdots, M_K$. In other words, if we multiply $M_1, \cdots, M_K$ by a unit amplitude complex number $e^{i\theta}$ where θ is a real number, then the arithmetic average is multiplied by $e^{i\theta}$. However, it is not difficult to see that this phase change does not affect the system performance. The similarity of distribution is usually characterized

by the variance rather than the mean value. Consequently, we attempt to use the normalized frequency arithmetic variance (NFAV) in Section 5.1.1 to characterize the similiarity of the numerical distribution. Because we are more interested in the shape of distribution, we normalize the variance by average power. Note that the normalized arithmetic variance is directly related to the magnitude of the normalized arithmetic average given the average power.

Consequently, we attempt to describe a multipath fading channel by three parameters, namely, the average power of the specular component [5], the normalized arithmetic variances, and the re-centered normalized mean square correlations. The NFAV $\nu_f(\Theta)$ of $M_1, \cdots, M_K$ is given by

$$\nu_f(\Theta) = \frac{\frac{1}{K}\sum_{k=1}^{K}|M_k - \langle \mathbf{M} \rangle_\Theta|^2}{\frac{1}{K}\sum_{k=1}^{K}|M_k|^2} = 1 - \frac{\left|\sum_{k=1}^{K} M_k\right|^2}{K\sum_{k=1}^{K}|M_k|^2} \tag{5.57}$$

where $\langle \mathbf{M} \rangle_\Theta$ denote the arithmetic mean of $M_1, \cdots, M_K$ defined by

$$\langle \mathbf{M} \rangle_\Theta = \frac{1}{K}\sum_{k=1}^{K} M_k. \tag{5.58}$$

Consequently, NFAV measures the frequency selectivity of the specular component of the channel.

5.3.4   Performance Simulations

It is not yet evident how closely the system performance depends on the second order statistics defined in this paper, although they provide a low order conceptually meaningful characterization for multipath fading channels as described above. We attempt to verify through simulations that the parameters capture essential characteristics of fading channels. For the simulations, we have chosen $K = 64$ and generated 400 delay power profiles from which we have obtained 400 (normalized) correlation matrices denoted by $\mathbf{C}_0$. Also we have obtained 20 arbitrarily generated (normalized) specular components $\mathbf{M}_0's$ with diverse NFAV. From these, we obtain a fading channel by choosing a scale factor $\alpha \in \{0.25, 0.5, 0.75\}$, a (normalized) specular component $\mathbf{M}_0$ and a (normalized) correlation matrix $\mathbf{C}_0$. A realization of the fading level vector $\mathbf{H}$ is then

---

[5]Note that this determines the average power of the diffuse component since we assume that $\sum_{k=1}^{K} E[|H_k|^2] = 1$
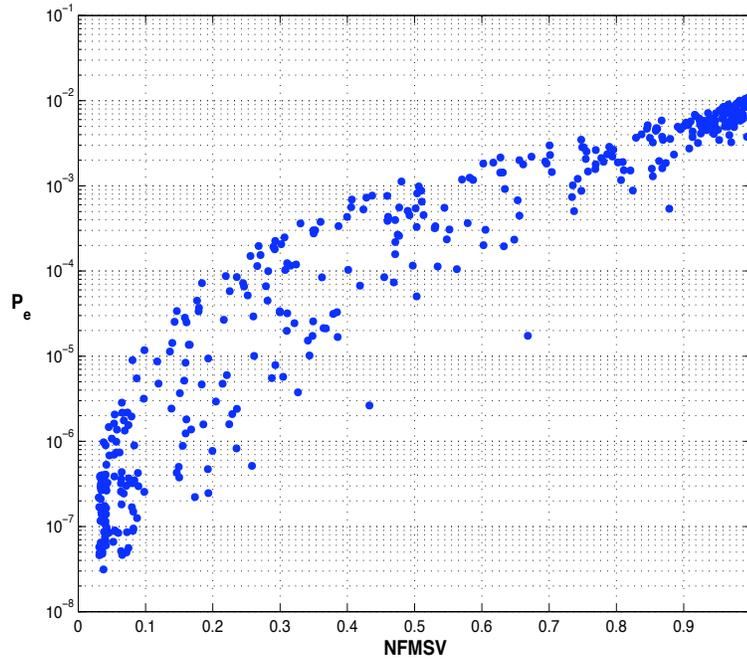
Figure 5.1: The probability P_e of bit error vs. the normalized frequency mean square covariance (NFMSV). Each dot corresponds to the (NFMSV,P_e) pair under one of 400 channels with single line of sight components ($\alpha = 0.5$).
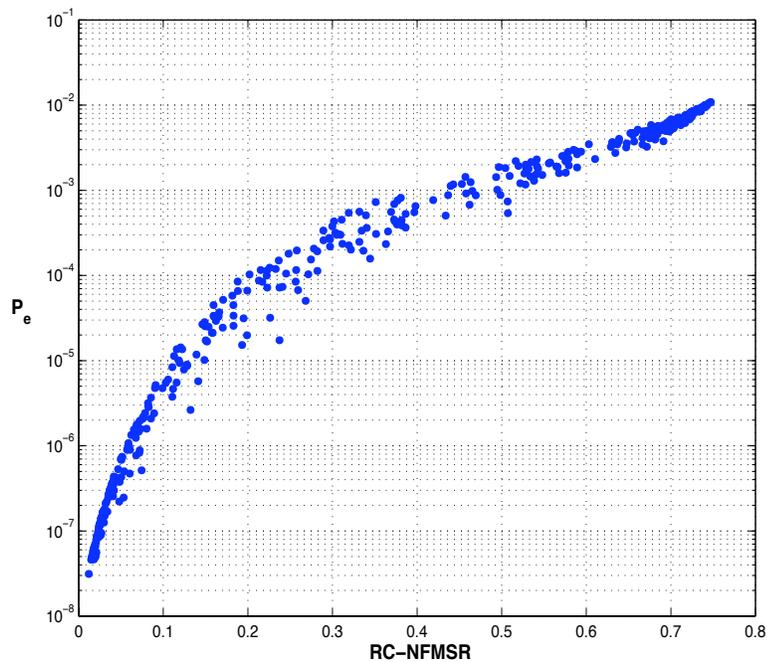
Figure 5.2: The probability P_e of bit error vs. the re-centered normalized frequency mean square correlation (RC-NFMSR). Each dot corresponds to the (NFMSV,P_e) pair under one of 400 channels with single line of sight components ($\alpha = 0.5$).
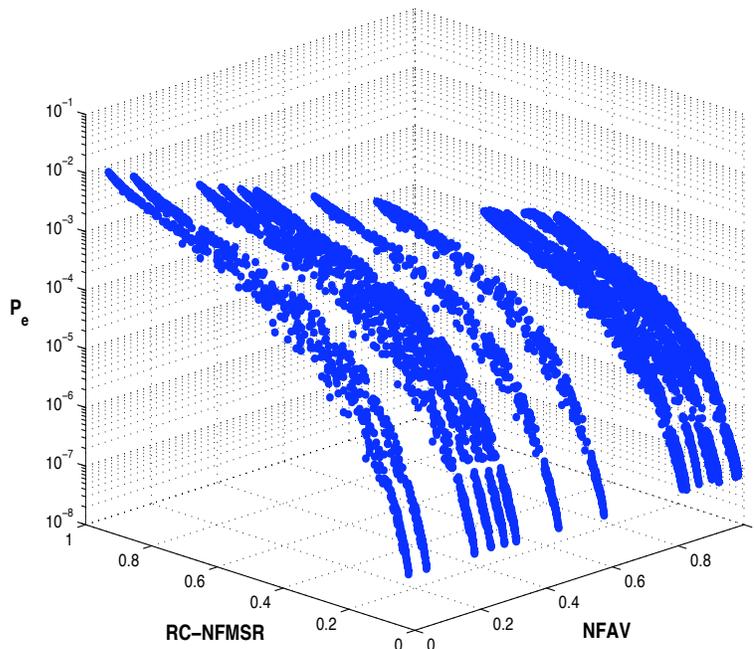
Figure 5.3: The probability P_e of bit error vs. (NFAV,RC-NFMSR). for multipath fading channels $\alpha = 0.25$. Each dot corresponds to the (NFAV,RC-NFMSR,P_e) pair under one of $20 \times 400 = 8,000$ channels.

obtained by

$$\mathbf{H} = \sqrt{\alpha}\,\mathbf{M}_0 + \sqrt{1-\alpha}\,\mathbf{D}_0 \tag{5.59}$$

where $\mathbf{D}_0$ is a realization of a circularly symmetric complex Gaussian random vector with zero mean and covariance matrix $\mathbf{C}_0$. Consequently, we consider total of $3 \times 20 \times 400 = 24,000$ different multipath fading channels.

First, Figure 5.1 depicts the relation between normalized frequency mean square covariance (NFMSV) and the probability error P_e for channels with the same specular component with $\alpha = 0.5$, while Figure 5.2 describes the relation between RC-NFMSR and the system performance. The particular specular component chosen consists of single line of sight path so that $M_1 = \cdots = M_K$ so that $\nu_f(\Theta) = 0$. In these and the following figures, each point corresponds to the probability of error under one of the $24,000$ fading channels. The error probability P_e of each of the channels is obtained by numerical evaluation of (5.55). We see the relation is relatively poor with NFMSV compared to that with RC-NFMSR. For this case, we observe there is a very close relationship between RC-NFMSR and the system performance.

Figures 5.3 - 5.5 shows the performance vs. (NFAV,RC-NFMSR) relations for $\alpha = 0.25, 0.50$ and 0.75, respectively. Although there are some fluctuations, we observe very
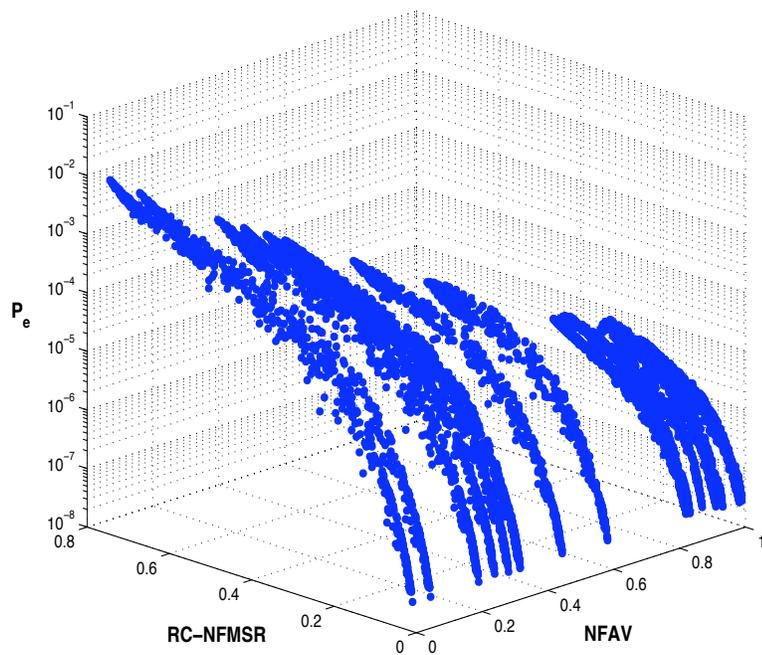
Figure 5.4: The probability P_e of bit error vs. (NFAV,RC-NFMSR). for multipath fading channels $\alpha = 0.50$. Each dot corresponds to the (NFAV,RC-NFMSR,P_e) pair under one of $20 \times 400 = 8,000$ channels.
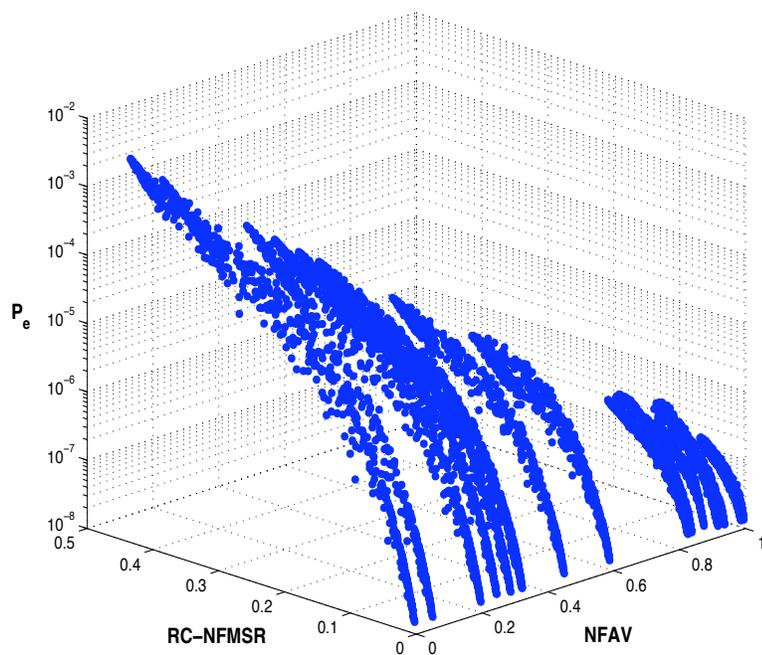
Figure 5.5: The probability P_e of bit error vs. (NFAV,RC-NFMSR). for multipath fading channels $\alpha = 0.75$. Each dot corresponds to the (NFAV,RC-NFMSR,P_e) pair under one of $20x400 = 80,000$ channels.
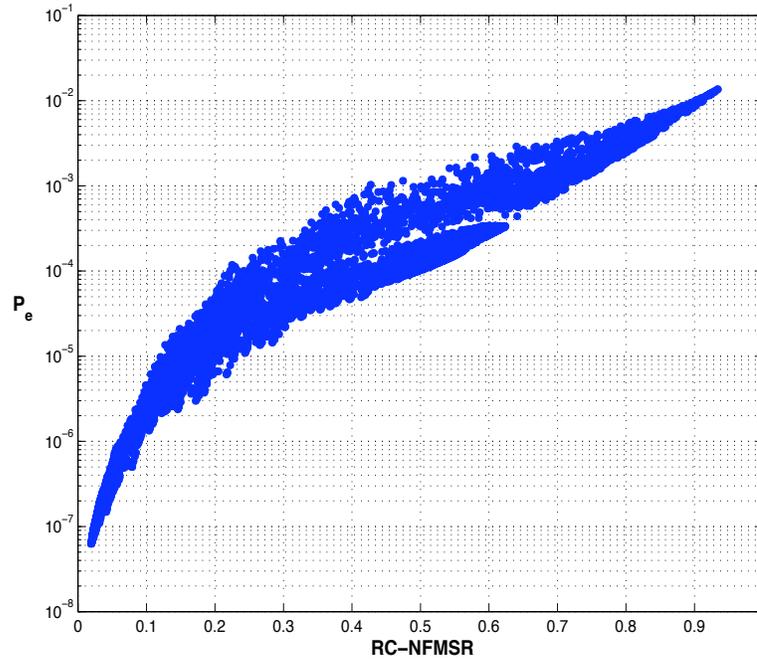
Figure 5.6: The probability P_e of bit error vs. the the re-centered normalized frequency mean square correlation (RC-NFMSR). Each dot corresponds to the (RC-NFMSR,P_e) pair for a channel with $\alpha = 0.25$.
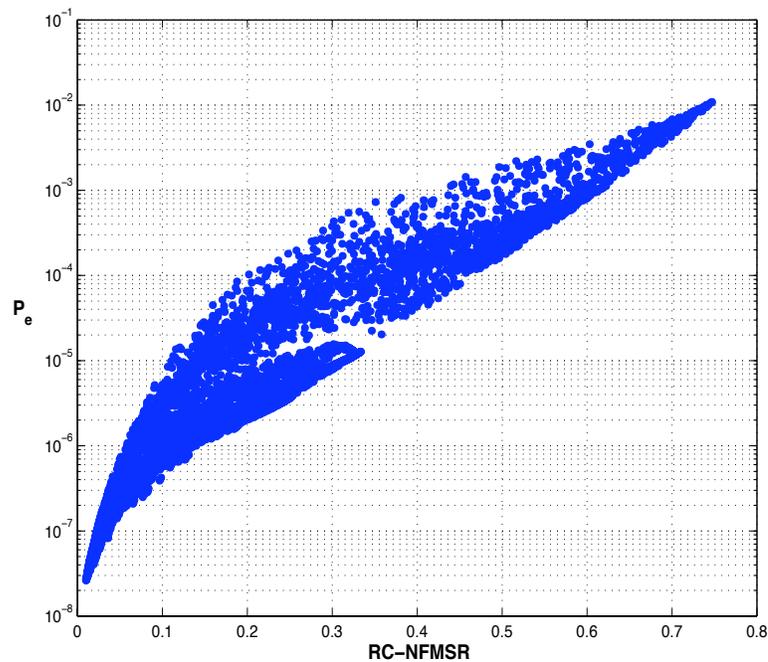
Figure 5.7: The probability P_e of bit error vs. the the re-centered normalized frequency mean square correlation (RC-NFMSR). Each dot corresponds to the (RC-NFMSR,P_e) pair for a channel with $\alpha = 0.50$.
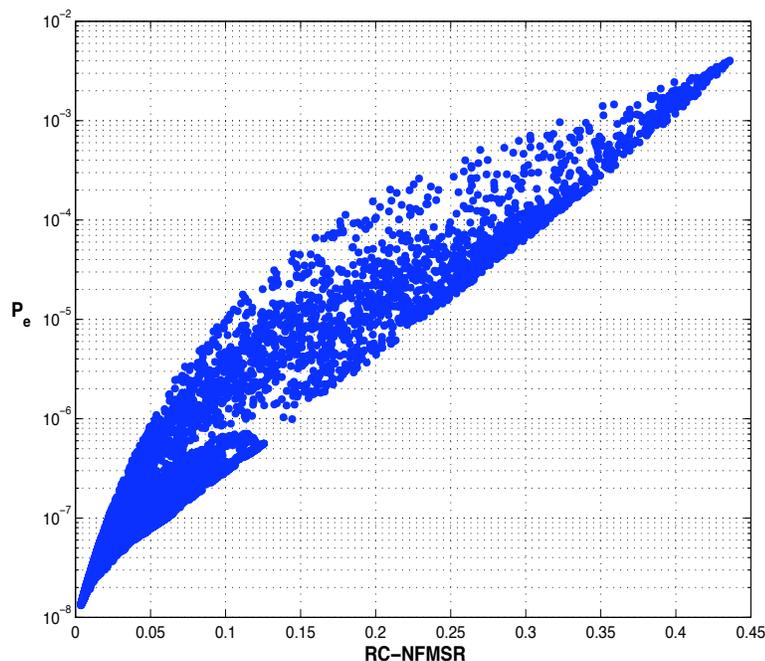
Figure 5.8: The probability P_e of bit error vs. the the re-centered normalized frequency mean square correlation (RC-NFMSR). Each dot corresponds to the (RC-NFMSR,P_e) pair for a channel with $\alpha = 0.75$.

close relation which makes the parameters meaningful. We observe that the effect of NFAV becomes more salient as α increases, which is not surprising. In Figure 5.6 - 5.8, we plot the P_e vs. RC-NFMSR relation for $\alpha = 0.25$, 0.50, and 0.75, respectively. We can still observe there is close relationship between P_e and RC-NFMSR even without specification of NFAV for each choice of α. In particular, we see the relation becomes stronger as RC-NFMSR approaches either of the extremal values.

5.4   Conclusion

In this chapter, we defined statistics that characterize multipath fading channels effectively in relation to the system performance by considering a simple diversity combining scheme. In Section 5.1, we provided formal mathematical definition of the parameters called re-centered (RC) normalized mean square correlation (NMSR) and normalized arithmetic variance (NAV). Re-centered normalized mean square correlation was defined to genearalize the applicability of normalized mean square covariance to non-WSSUS channels. Normalized arithmetic variance was defined to measure similarity or the difference between the shape of the specular part of the frequency response.

We provided theoretical arguments and simulation results to justify those parameters in Section 5.3. We found that there are very close relations between the system performance and the parameters for the simple diversity combining system. Although such results were obtained with a simple and ideal scheme, we can still expect that the parameters will provide meaning characterizations for multipath fading channels because many practical systems designed to combat multipath fading are very closely related to the simple diversity coming scheme in this chapter. We also expect that the line of reasoning itself will provide an exemplary role in studying systems which don't have conceptual similarity with the simple diversity combing scheme considered in this chapter.

Chapter 6

NMSV and System Design

In this chapter, we discuss the use of NMSV in wireless communication system design. In wireless communications, the frequency and the time selectivities of the channel fading are widely exploited in the system design. Spread spectrum technologies are among the most popular examples. We have shown that the performance of such systems are very closely related to the NMSV of the channel in question. Such a close relationship between the NMSV and the system performance is very useful in system design and planning. In Section 6.1, we consider the frequency allocation problem for a FHSS system. We show that we can improve system performance by separating the frequency slots with the same total bandwidth. In Section 6.2, we show how the NMSV can be used in choosing the optimal frequency hopping rate. In particular, we show that it is not necessary to evaluate the performance of a FHSS system under realistic situations. In Section 6.3, we propose carrier-separated (CS) orthogonal frequency division multiplexing (OFDM) systems for multirate multiple access wireless communications. Motivated from the result of Section 6.1, we consider carrier separation to achieve more frequency selectivity. We also show that it is possible to reduce the system complexity by judicious choice of carrier separation.

6.1   Frequency Allocation for a FHSS System

In this section, we consider the problem of frequency band allocation to an FHSS system. In Section 4.4, we assumed that the 332 frequency slots allocated to the system constitute a contiguous interval of frequency band of size 10.375MHz. Now we consider the possibility of allocating 332 non-contiguous frequency slots to the system.
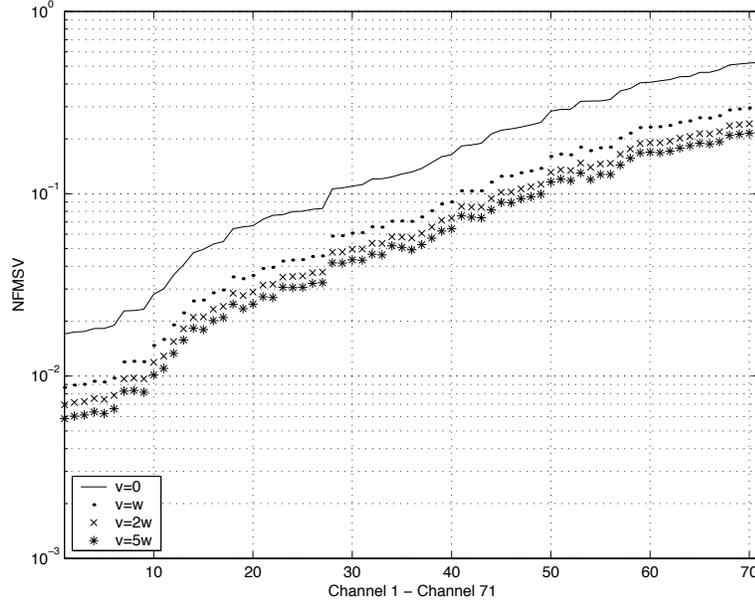
Figure 6.1: The usage of NFMSV in the frequency allocation. We see that we can lower the overall NFMSV of the channel by separating each hop slot by inserting some amount of guard band. However, the NFMSV does not decrease dramatically as the separation bandwidth become greater than the bandwidth of each hop slot. In this figure, the horizontal axis means nothing more than the order of the 71 channels.

Consider the frequency region Θ_r given by

$$\Theta_r = \bigcup_{k=0}^{K-1} [A + k(v+w), A + k(v+w) + w] \tag{6.1}$$

where A, v, and w are non-negative real numbers. Then, Θ_r is a union of K intervals of length w with a space of length v between adjacent intervals. We assume that each interval represents a frequency hop slot. Consequently, as the space v grows larger, the system tends to hop more space with the same frequency pattern and can generally achieve more diversity gain. The NFMSV $\mathcal{V}_f(\Theta_r)$ can be calculated from (4.13) with (4.17). By calculating $\mathcal{V}_f(\Theta_r)$ for the 71 channels introduced in Section 4.4, we obtain Figure 6.1. Here, we assume that $K = 332$ and $w = 31.25\text{kHz}$. From the figure, we can tell that we have smaller NFMSV with larger v, namely, with larger spacing. Since the BER performance is closely related to NFMSV, we can also tell, from the BER v.s. NFMSV curves (in Section 4.4), how much gain is obtained by introducing space between slots. The space between slots can be allocated to other systems. For example, if we choose $v = w$, then we can allocate slots to uplink and downlink systems alternately. We can allow more space by considering neighboring cells together in the
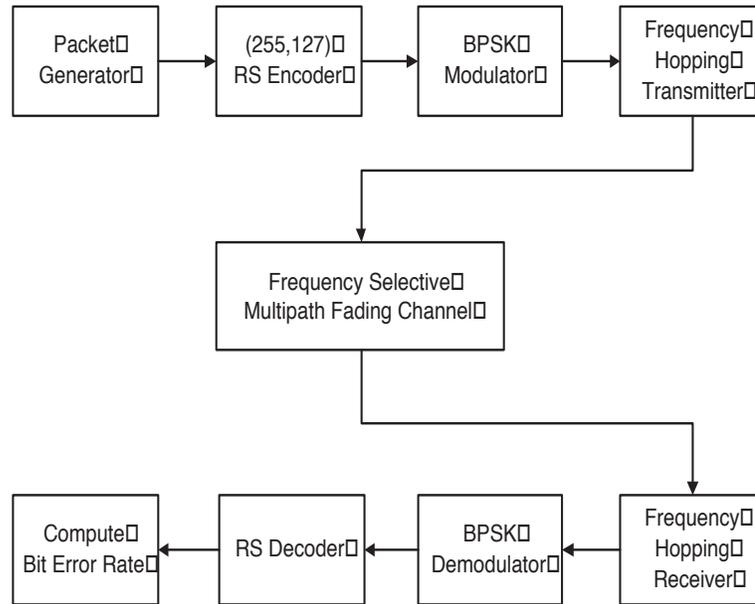
Figure 6.2: Block diagram of an FHSS system with (255,127) Reed-Solomon code.

allocation of frequency slots without wasting frequency resources. However, increasing v larger than $5w$ does not decrease $\mathcal{V}_f(\Theta_r)$ appreciably. Consequently, we expect saturation of performance enhancement after v becomes larger than $5w$.

6.2 Determination of Frequency Hopping Rate

Typically, a frequency hopping system is allowed to hop from one frequency to another within a given region of the frequency band of a physical channel. It is natural to expect an increasing diversity gain with higher hopping rate. However, the actual diversity gain fundamentally depends on the frequency selectivity in the frequency band of the channel allocated to the system. In this section, we investigate the relationship between the NFMSV of the given frequency band and the diversity gain by increasing the frequency hopping rate, which is helpful in determining optimal frequency hopping rate.

6.2.1 NFMSV and the Performance of a FHSS system with a Reed-Solomon Code

To illustrate the relationship between the NFMSV of a channel and possible diversity gain by frequency hopping, we consider a frequency hopping system using a (255, 127) Reed-Solomon code [33] depicted in Figure 6.2. Assume that each packet consists of

one codeword, namely, 2040 coded bits that are modulated by a BPSK modulator before transmitted by a frequency hopping transmitter. The total bandwidth of the channel is assumed to be 10.24MHz which is divided into $N = 255$ slots of bandwidth 40kHz. For definiteness, we assume the system is allocated with the frequency region between $A = 1.0$GHz and $B = 1.01024$GHz. The channel is assumed to be frequency selective but time non-selective. The number of frequency hopping per packet, which we shall denote by K, is chosen between 1 and 255. The hopping is roughly uniformly spaced. For example, if $K = 7$, the 1st, the 37th, the 73th, the 110th, the 146th, the 183th, and the 219th slots are used for transmission. The 255 symbols are divided roughly equally so that the symbols $1 \sim 36$ are transmitted over the 1st slot, the symbols $37 \sim 72$ over the 37th slot and so on.

As a performance measure, we consider the bit error rate (BER) of the system. For simulations, we use a systematic (255,127) Reed Solomon code and assume that the all zero information bits are transmitted. The receiver first makes a hard-decision on each coded symbol. Then, it counts the number of coded-symbol errors and assume the errors are not corrected by the decoder if the number of symbol errors exceeds the error correcting capability, in which case the number of information bit errors in the systematic parts are counted toward the total errors. For the simulations, we consider 300 randomly generated delay profiles and calculated the NFMSV of the frequency band between 1000MHz and 1010.24MHz. Among the randomly generated delay profiles, we select 4 models for illustration. We call them channel 1 - 4.

In Figures 6.3 and 6.5, the delay power profiles of the four channels are depicted. As shown in Figure 6.3, channel 1 and channel 2 have quite different delay profiles. In particular, the rms delay spread of channel 1 is 0.28μsec, while that of channel 2 is 1.26μsec. However, both channel 1 and channel 2 have the same NFMSV, namely, 0.2. Roughly speaking, channels with NFMSV = 0.2 can be thought to have 5 well separated, equal strength paths. We can regard such a channel to have a fair amount of degree of freedom or frequency selectivity. A channel with this amount of degree of freedom is not rare but there are many channels that have smaller amounts of degree of freedom. However, channels with smaller amount of degree of freedom are not very interesting in relation to the frequency hopping rate choice. Consequently, we have chosen NFMSV 0.2 for channels with a small amount of frequency selectivity in our simulations. As an another choice, we have selected NFMSV of 0.06. We see channels 3 and 4 exhibit much more time dispersion than channels 1 and 2. While they have the same NFMSV, channel 4 has larger rms delay spread than channel 3: 4.12μsec vs. 3.16μsec. As the delay profiles indicate these four channels are chosen to show how
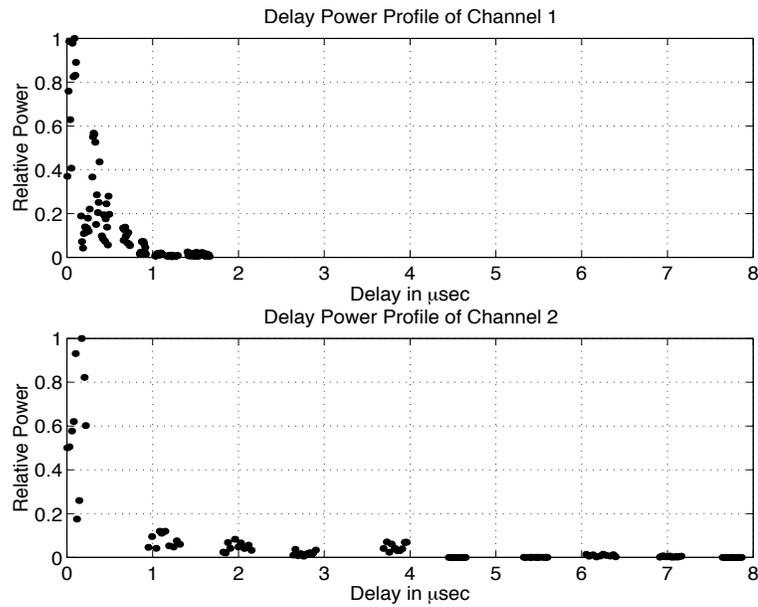
Figure 6.3: Delay power profiles of two different channels with NFMSV=0.2. Channel 1 has *rms* delay spread 0.28 μsec while channel 2 has 1.26 μsec.
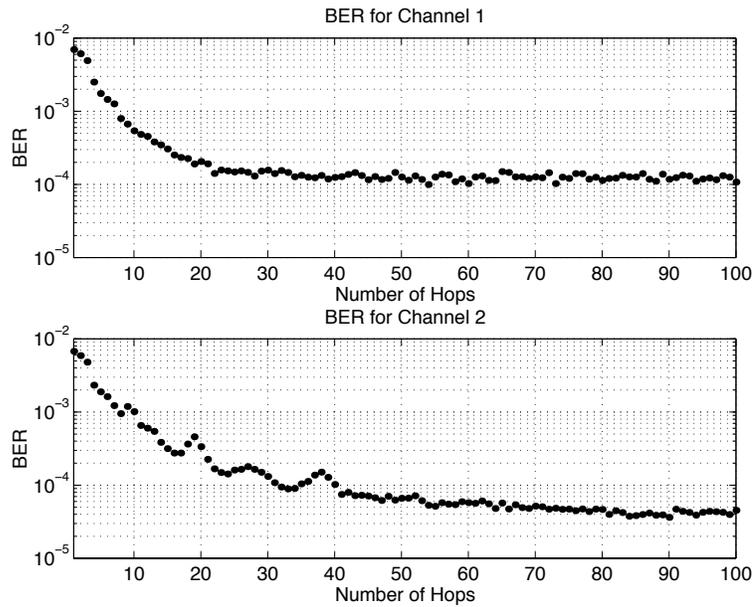
Figure 6.4: Bit error rate (BER) versus number of hops. The rate of change in the BER is slow after 20 hops per packet for both channels.
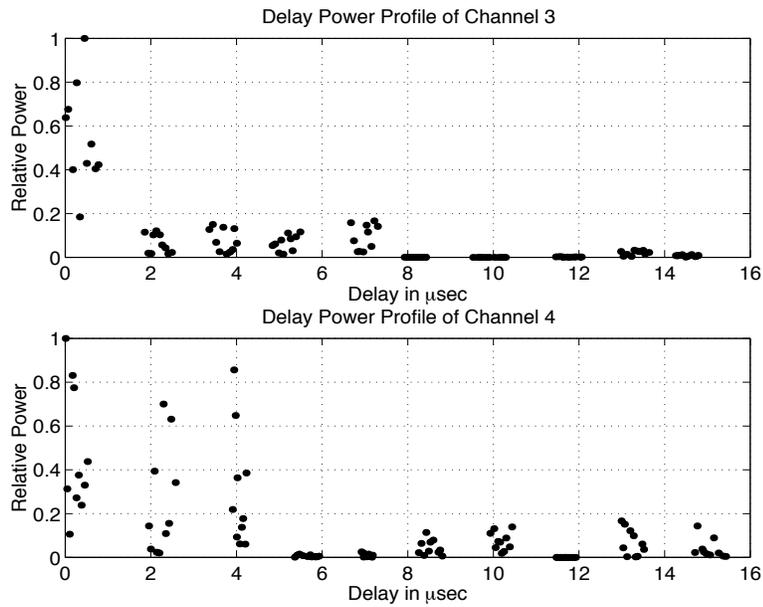
Figure 6.5: Delay power profiles of two different channels with NFMSV=0.06. Channel 3 has *rms* delay spread 3.16 μsec while channel 4 has 4.12 μsec.
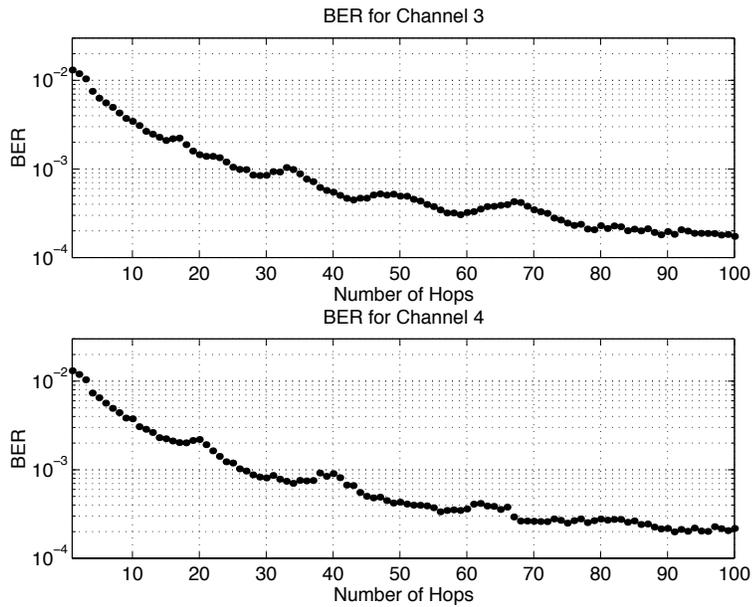
Figure 6.6: Bit error rate (BER) versus number of hops. The rate of change in the BER is slow after 50 hops per packet for both channels.

different two channels can be with the same NFMSV. Channels 3 and 4 are examples of channels with high frequency selectivity.

Figure 6.4 shows the performances versus the hopping rates in channel 1 and channel 2 at signal to noise ratio $E_b/N_0 = 18$ dB where E_b is the energy per information bit and N_0 is the one sided noise power spectral density of the additive white Gaussian noise at the receiver. We see that the performance of the system is slightly better under channel 2. As discussed in Chapter 4, the performance versus NFMSV relation is relatively loose when the channel has relatively moderate degree of freedom. In fact, channels 1 and 2 are chosen to illustrate extremes cases of performance difference with the same NFMSV. Under both channels, we observe noticeable performance enhancements as we increase the hopping rates upto 20 hops per packet. Under channel 2, we have mild increases in the performance with higher hopping rates.

The simulations results for channel 3 and channel 4 are given in Figure 6.6. In this case, we observe that the performances under the two channels are closer. This is in fact a general tendency. In other words, the performance is more closely related with NFMSV for channels with smaller NFMSV. So under two different channels with very small NFMSV, the performance of a frequency hopping system are usually quite close. From the Figure 6.6, it is not easy to determine the point of saturation. However, we observe that the performance enhancements are not great by increasing the hopping rate from 50 to 100 hops per packet.

6.2.2   NFMSV of Frequency Region and System Performance

In the previous subsection, we showed that the diversity gain by frequency hopping depends heavily on the channel characteristics, especially on the NFMSV of the channel by considering a particular coding schemes. Although it is intuitively not unreasonable to expect similar tendency with different coding technology, we give some more justifications by considering the NFMSV of the frequency region of interest. As discussed in Chapter 4, the performance of the system is directly related to the NFMSV of the particular frequency region over which the signals are actually transmitted. In other words, the performance gain should parallel the increase in the stochastic degree of freedom by increasing the frequency hopping rate. Again we consider the same frequency band between 1000MHz and 1010.24MHz with 255 frequency slots and the particular hopping pattern discussed in the previous subsection. Consequently, the hopping pattern is chosen to be as uniform as possible. Let Θ_K be the frequency region experienced by the signals when the hopping rate is chosen to be K hops per packet.

If the system does not hop during a packet duration, the region Θ_1 of interest con-
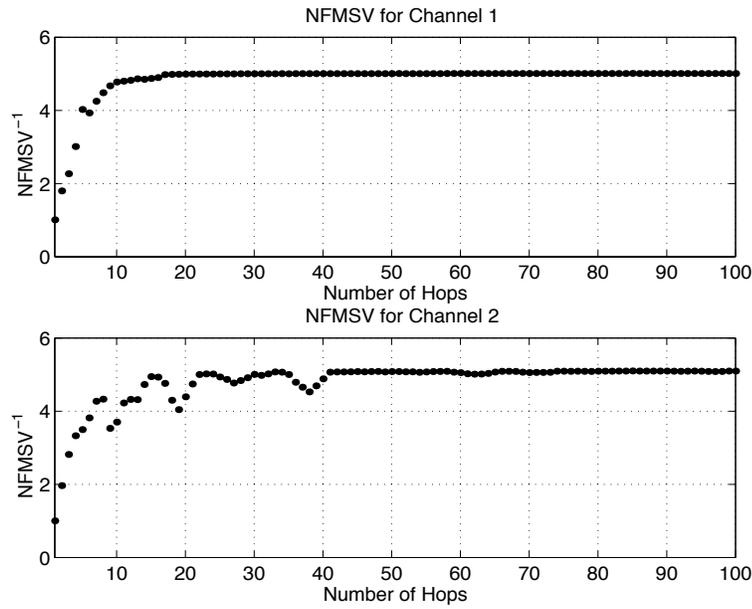
Figure 6.7: NFMSV versus number of hops. NFMSV is saturated after 20 hops for both channels.
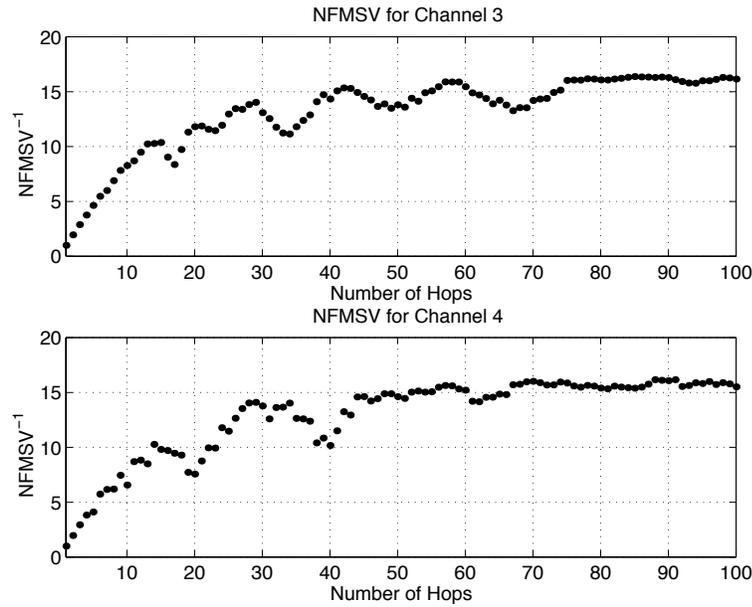
Figure 6.8: NFMSV versus number of hops. NFMSV is saturated after 50 hops for both channels.

sists of just one slot and the NFMSV $\mathcal{V}_f(\Theta_1)$ is 1. If the system hops once during a packet duration, then the system utilizes a region Θ_2 consisting of two slots and the NFMSV will be $\mathcal{V}_f(\Theta_2)$. This implies that the system performance will generally get better if the system hops once. Continuing this way, we consider the case when the system hops $K-1$ times and uses a region Θ_K consisting of K slots where K is between 1 and 255.

In Figure 6.7 and 6.8, we plot $1/\mathcal{V}_f(\Theta_K)$ as a function of K instead of $\mathcal{V}_f(\Theta_K)$ for better visualization. We truncated the plot after $K = 100$ because the NFMSV's saturates at K=100. First consider channel 1. It is not evident from the figure but $\mathcal{V}_f(\Theta)$ decreases until $K = 16$ and then becomes practically saturated. This tendency is reflected in the BER performance in Figure 6.4. In the case of channel 2, $\mathcal{V}_f(\Theta)$ fluctuates until around $K = 40$, which is also reflected in the second part of Figure 6.4. In interpreting the relation between the NFMSV and the system performance, it is important to note the fact that the system performance can be affected by a slight decrease in the NFMSV. For example, doubling NFMSV can imply a change of several orders of magnitude in BER. By comparing Figure 6.8 and 6.6, we can observe similar matches for channel 3 and channel 4. From these observations, we see that the performance gain by a higher hopping rate actually comes from the enhanced diversity which is effectively characterized by the NFMSV of the region of interest. Consequently, we can expect similar diversity gain with different encoding and decoding schemes.

6.3 Carrier Separated OFDM System

6.3.1 Introduction

Most wireless communication systems are based on single carrier spread spectrum technology. For a single carrier system, high rate communication can be achieved either by a high symbol rate or by a complex signal constellation with many bits per symbol [35, 36]. Using a dense signal constellation is undesirable for a wireless system since fading makes it difficult to reliably detect which constellation point was sent. Signaling at a high symbol rate is equally undesirable because the intersymbol interference due to the time dispersion of a multipath fading channel would require a complex high-speed equalizer or a similar device. For these reasons, OFDM (orthogonal frequency division multiplexing) is used to provide acceptable performance in a multipath fading environment promising high peak data rates [37, 38]. OFDM techniques have been used in high-speed wireless LANs, digital audio broadcast systems and wireline high-speed data communications systems. Recently, OFDM is considered
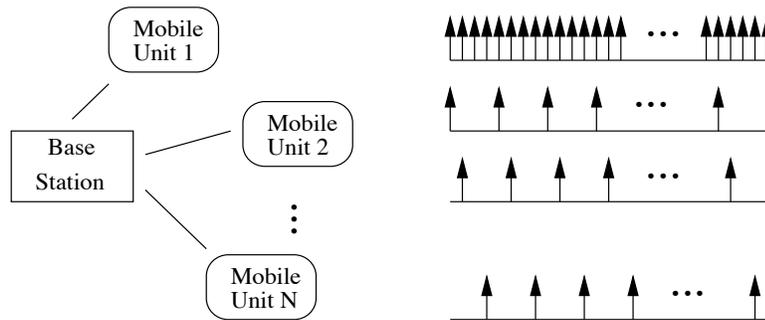
Figure 6.9: Multiaccess communications using M subcarriers. Each user uses K subcarriers. This paper assumes $M = 1024$ and $K = 16$.

for 4th generation wireless communication systems.

As usual in wireless communications, multiple access support is an important issue in the future generation wireless communications. There are various ways to support multiple access with OFDM technology. Frequency division multiple access (FDMA) method is probably the most simple and natural way to support multiple access with OFDM technology. For illustration, consider a base-station with 10.24 MHz bandwidth allocated for multiple users. Assume the 10.24 MHz frequency band is divided into 1024 carriers of 10kHz bandwidth and that each user is allowed to use 16 carriers at a time. Hence, the base-station allocates 160kHz for each user. In the usual FDMA-OFDM, each user is allocated a continuous 160kHz frequency band. At the mobile receiver side, the desired signal looks just like a 16-carrier OFDM signal instead of a 1024-carrier OFDM signal.

However, the usual FDMA-OFDM system is subject to severe performance degradation in a fading channel. An immediate remedy for this drawback is to exploit the diversity of the frequency selectivity in multipath fading by employing frequency hopping. However, there are several drawbacks with frequency hopping. First, to be useful, the frequency hopping should be made fast enough. In such a system, the synchronization and the channel estimation are difficult and the computational overhead is not negligible especially because of the multiple carrier usage. Secondly, for an uplink system, collisions between users give rise to non-negligible performance degradation. An alternative that can achieve similar performance with moderate increase in computational requirements is a carrier-separated (CS) OFDM system we propose. For example, we allocate the 1st, the 65th, $\cdots$, and the 961th carriers to the first user, the 2nd, the 66th, $\cdots$, and the 962th to the second user, and so on. In this case, the 16 carriers allocated to each user are separated by 64 other carriers. Because carriers are uniformly separated, we call this type of system a uniformly carrier-separated
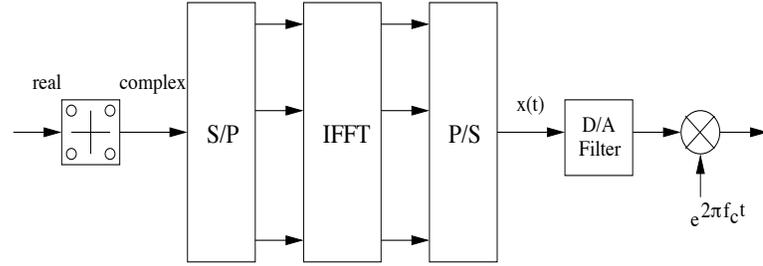
Figure 6.10: Transmitter model of a carrier-separated multiaccess OFDM wireless communication system.

(UCS) OFDM system. In particular, in our example, the maximum possible separation for all users is 64 and hence the system will be called a uniformly-maximum carrier-separated (UMCS) OFDM system.

6.3.2   Fourier Coefficients Computation

Figure 6.10 illustrates the transmitter model of a carrier separated OFDM communication system with total of M subcarriers where K subcarriers are reserved for each user. We have assumed that $M = 1024$ and $K = 16$ throughout this section. The transmitter model consists of a serial-to-parallel (S/P) converter, an inverse FFT (IFFT), and a parallel-to-serial (P/S) converter. For a specific user n out of M/K possible users at any given time k, real-valued binary symbols $d_k \in \{1, -1\}$ is mapped to form a complex sequence $a_k = a_k^I + a_k^Q$. These complex sequences are then mapped to $K \leq M$ subcarriers and combined with the sequences of the other users for the transmission. We denote by $a_{k,n,m}$ the complex symbol of user n at time k and at carrier location m. Each user transmits 32 real-valued binary symbols (16 complex-valued samples) at the same time. Thus, for 1024 carrier system, the total number of real samples is 2048 which equivalent to 1024 complex samples. The aggregated complex samples $\{a_{k,n,m}\}$ for user n at time k are processed by M-point inverse Fourier transform (IFFT) and converted to serial sequences constituting the following OFDM baseband signal

$$x_{k,n}(t) = \sum_{m=0}^{M-1} a_{k,n,m} e^{j2\pi m f_0 t} \quad 0 \leq t \leq T_c \tag{6.2}$$

where M is the size of IFFT corresponding to the number of subcarriers, f_0 is the frequency spacing between adjacent carriers, and $T_c = 1/f_0$ is the time duration of a frame. (i.e., for 1024 subcarriers occupying 10.24MHz bandwidth, $f_0 = 10KHz$ and $T_c = 0.1msec$).
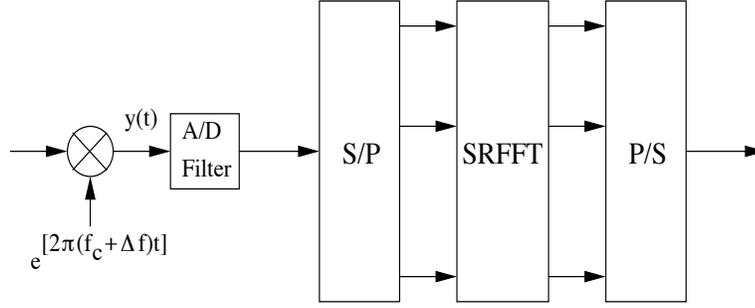
Figure 6.11: The decimation-in-frequency Split radix FFT algorithm based receiver model of a multiaccess OFDM wireless communication system.

Figure 6.11 depicts the receiver model of a uniformly-maximum carrier-separated multiaccess OFDM communication system with total of M subcarriers. Only K subcarriers are demodulated at the receiver for a user. The basic building blocks consists of a serial-to-parallel converter (S/P), a decimation-in-frequency split-radix FFT (SRFFT) module, and a parallel-to-serial converter (P/S). The received signal $y(t)$ after the carrier frequency down-conversion and the low-pass filtering is equal to the transmitted signal $x(t)$ attenuated by multipath fading and then corrupted by some unknown delay τ, local oscillator frequency offset Δf_c, and local oscillator phase offset θ as well as the thermal noise. For simplicity, we assume that the thermal noise is negligible in the following. If we denote by $\{X_m\}$ the transmitted complex symbols attenuated by multipath fading, the received signal $y(t)$ can be written as

$$y(t) = \frac{1}{\sqrt{M}} \sum_{m=0}^{M-1} X_m e^{j2\pi m f_0(t-\tau)} e^{j2\pi \Delta f_c(t-\tau)} e^{j\theta} + n(t) \tag{6.3}$$

in the continuous time domain and

$$y(l) = y(lT_c/M) = \frac{e^{j\theta}}{\sqrt{M}} \sum_{m=0}^{M-1} X_m e^{j2\pi(m+\Delta f_c/f_0)(l/M) - j2\pi(mf_0 + \Delta f_c)\tau} + n(l) \tag{6.4}$$

in the discrete sampled time domain sampled at $t = lT_c/M$. After the Fourier transformation, the FFT coefficients recovered at the receiver are given by

$$Y_m = \frac{1}{\sqrt{M}} \sum_{l=0}^{M-1} y(l) e^{-j2\pi lm/M} \tag{6.5}$$

$$= \frac{e^{j\theta - j2\pi \Delta f_c \tau}}{M} \sum_{p,q=0}^{M-1} X_q e^{j2\pi\{(q-m)p/M + \frac{\Delta f_c}{f_0} p/M - q f_0 \tau\}} \tag{6.6}$$

digitally converted
input samples
0,1,2,...,2047

demodulated samples
corresponding to
carriers 0,63,...,960
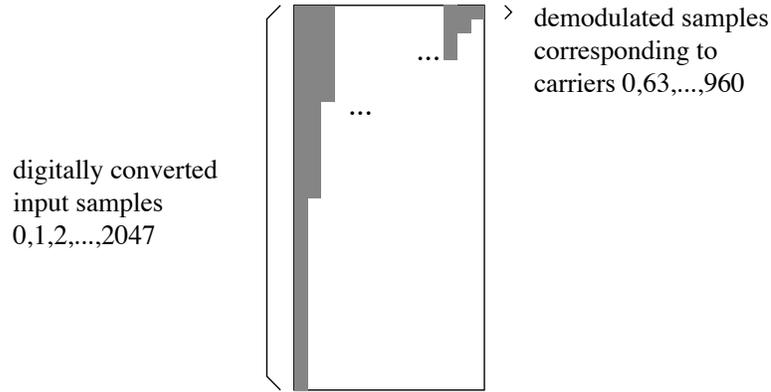
Figure 6.12: Illustration of active computational complexity of the decimation-in-frequency SRFFT algorithm employed in the carrier separated multiaccess OFDM system. Shaded portion represents actual required computation for demodulation. The horizontal axis represents the stages of FFT.

that can be further simplified to

$$Y_m = X_m e^{-j2\pi m f_0 \tau} e^{-j2\pi \Delta f_c \tau} e^{j\theta} \tag{6.7}$$

by assuming that $\Delta f_c / f_0$ is very small. The minimum required sampling rate at the receiver is twice the Nyquist rate ($l = 2$) to eliminate the aliasing during the analog to digital conversion. Thus, the size of FFT at the receiver is equal to 2048 assuming Nyquist sampling rate for the carrier separated multiaccess OFDM system employing 1024 carriers.

The computational complexity of the FFT can be reduced significantly if a decimation-in-frequency algorithm is employed in the demodulation. There are many different ways to implement the FFT. Radix-2, radix-4, and split-radix algorithms are by far the most widely used in practice and hence are good candidates for the implementation of the demodulator. Among these algorithms, the SRFFT is computationally more efficient than the radix-2 or the radix-4 FFT in terms of the total number of necessary multiplications and additions. However, the main drawback of the SRFFT implementation is that its computational flow is highly irregular when compared to the other radix FFT. For DSP processors, radix-2 or radix-4 decimation-in-frequency FFT algorithms are preferable in terms of speed and accuracy. The irregular structure of the SRFFT may render it less suitable for implementation on a digital signal processor. Thus, when SRFFT is adopted for low-power and low-complexity implementation in an ASIC, and its execution timing must be precisely controlled for correct operation.

Since each user uses 16 carriers for data transmission, not all of $Y_1, \cdots, Y_{M-1}$ are

needed to be demodulated at the receiver. When demodulating a signal that consists of only 16 uniformly-maximum separated carriers that were assigned to a user at the transmitter, the active computation needed at the receiver is much less than the complete FFT computation (approximately 18% of total computation complexity or comparable to 128 FFT). The actual computations required by the decimation-in-frequency SRFFT is illustrated as a shaded portion as shown in Figure 6.12 (i.e., processing 2048 samples on the first stage, 1024 samples on the second stage, ..., down to 16 samples on the last stage). In the figure, the 16 carriers correspond to the output of the decimation-in-frequency SRFFT are the 1st, the 65th,..., the 961th, whereas the complete FFT operation generates 2048 outputs (i.e., because of Nyquist sampling rate) where half of those outputs correspond to the carriers located between the carriers which is not needed for proper demodulation. Thus, approximately one half of total computation is not necessary. Moreover, only the first 16 outputs actually corresponds to the transmitted 16 carriers. This simplification cannot be achieved if decimation-in-time FFT algorithms are incorporated for demodulation.

Due to frequency hopping operation or assignment of different set of carriers at the transmitter, it is possible that the actual received carriers for the user may correspond to the carriers located at 2nd, 66th, ...,962th. In this case, the frequency of the local oscillator at the receiver is shifted by f_0 which is equal to 10KHz such that the total amount of computation remains the same regardless of the set of carriers chosen at the transmitter. Even though only 16 outputs are required, it is also possible to compute more carriers with for accurate synchronization purpose.

6.3.3   Performance of a Carrier Separated OFDM System

In the previous subsection, we showed that the carrier separation involves just a mild increase in the computational complexity compared to FDMA OFDM if we emply uniformly-maximum carrier-separation. In this subsection, we study the diversity gain achievable by the uniformly-maximum carrier-separation. It is not hard to see that the performance of the system should be comparable to that of a frequency hopping OFDM system with 16 hops per packet with uniform frequency hopping pattern.

To study the possible performance gain with carrier separation, we need to specify the other parts of the system design such as encoding and decoding methods. However, the performance of a diversity combining scheme should depend fundamentally on the degree of freedom inherent in the multipath fading channel. Or more precisely, it depends on the amount of degree of freedom it exploits out of the total degree of freedom in the channel. For example, in a direct sequence spread spectrum system,

the performance of a rake receiver fundamentally depends on the method of diversity combining. Consequently, selection combining usually underperforms a hybrid selection combining or the maximal ratio combining that exploits more degree of freedom in the multipath fading than the selection combing.

As discussed in Section 4.2, given a channel, the NFMSV $\mathcal{V}_f(\Theta)$ is a function of the frequency region Θ of interest. For the evaluation of the diversity gain by a particular system design, we consider the NFMSV $\mathcal{V}_f(\Theta)$ of the particular frequency region Θ over which signals are actually transmitted. In particular, in a carrier separated OFDM system, it is very important to design the carrier separation to allocate to each user a region with small NFMSV. However, the maximum achievable diversity gain by a judicious choice of the frequency region depends fundamentally on the diversity of the given fading channel. Consequently, it is necessary to investigate the diversity gains of the carrier separation over wide variety of fading channels.

To study the relation between achievable diversity gain by carrier separation and the NFMSV of the whole frequency band of 10.24MHz, we have generated 300 delay power profiles randomly and have selected, among them, 100 practically shaped power profiles with varying amount of time dispersion. While some of the 100 channels have sub-μsec rms (root mean square) delay spread, some channels have $3-4\mu$sec of rms delay spread. The NFMSV $\mathcal{V}_f(\Theta)$ of the region Θ of interest can be calculated from (4.13) and (4.17) in Chapter 4. The regions of our interest are the whole frequency region Θ_w allocated to the 1024 carriers, a region Θ_m consisting of 16 uniformly maximum separated carriers and a region Θ_n consisting of 16 adjacent carriers.

The NFMSV's for the 100 selected channels are plotted in Figure 6.13. For better visualization, we plot $\mathcal{V}_f^{-1}(\Theta)$, the stochastic degree of freedom (SDF), instead of $\mathcal{V}_f(\Theta)$. The NFMSV $\mathcal{V}_f(\Theta_w)$ of the whole frequency band ranges from 0.035 to 0.407. Consequently, the SDF available in the whole frequency band ranges from 2 to 30. The NFMSV $\mathcal{V}_f(\Theta_n)$ of 16 adjacent carriers are from 0.373 to 0.997. Consequently, without carrier separation, the SDF is no greater than 3. In contrast, the NFMSV $\mathcal{V}_f(\Theta_m)$ of the 16 uniformly-maximum separated carriers is between 0.077 and 0.439. Consequently, uniformly-maximum carrier separation achieves SDF from 2 to 13 depending on the SDF of the whole allocated frequency band. Note from the Figure 6.13, that the gain in SDF by 16 uniformly maximum separated carriers generally depends on the NFMSV of the whole band. However, the SDF with any 16 carriers cannot exceed 16. Consequently, we observe there is a large gap between $\mathcal{V}_f(\Theta_w)$ and $\mathcal{V}_f(\Theta_m)$ for channels with small $\mathcal{V}_f(\Theta_w)$. In such a case, increasing the number of carriers to 32 or 64 can achieve higher SDF gain. However, increasing the number of carriers does
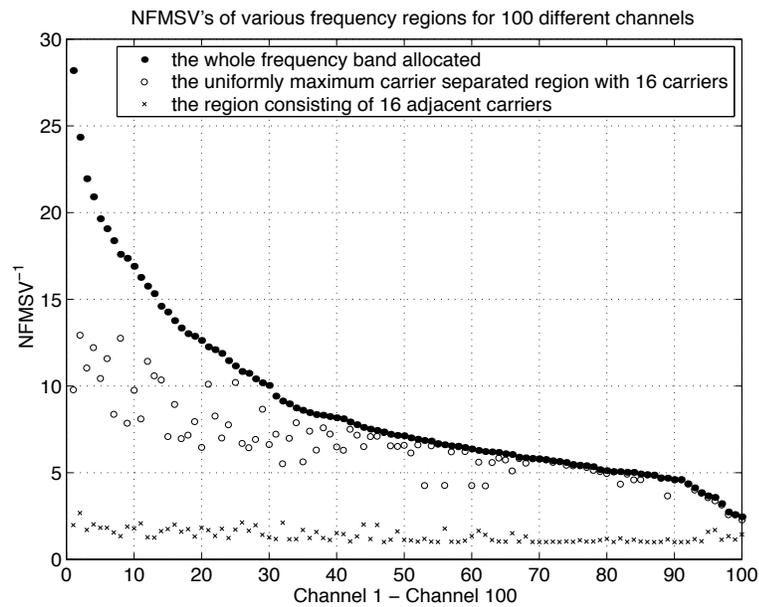
Figure 6.13: NFMSV's of various regions of interest are computed for 100 randomly generated frequency selective fading channels. The diversity gain achieved by the carrier separation is bounded above by the number of carrier and depends on the NFMSV of the whole allocated frequency region. However, it is generally much higher than that by system without carrier separation.

involve problems such as an increased peak to average envelope ratio. Consequently, there are tradeoffs in increasing the number of carriers. The performance enhancements by increased amount of SDF by employing carrier separation depends on the detailed specification of the system. However, doubling the diversity easily makes several orders of magnitude change in the bit error rate at reasonably high signal to noise ratios.

Chapter 7

Multiple Access Interference Cancellation

Multiple access interference (MAI) treatment is one of the most important problem in direct-sequence (DS) code-division multiple access (CDMA) communication systems. In this chapter, we study and compare various multiple access interference treatment algorithms under multipath fading channels in which inter-symbol interference (ISI) exists in addition to MAI. There have been very diverse efforts [10] to reduce MAI and ISI. In Section 7.1, we briefly summarize some of the most important multiuser detection algorithms. Then, in Section 7.2, we consider various multiuser detection algorithms for single-rate asynchronous DS-CDMA systrems under frequency selective fading channels. We first derive maximum likelihood sequence estimation (MLSE) algorithm, which is optimal in the sense of minimum multiuser packet error rate. Based on the MLSE, we consider matched-filter (MF) decision feedback sequence estimation (DFSE) algorithm to reduce the computational complexity. The most significant drawback of MF-DFSE is the performance degradation due to the untreated anti-causal interference due to the reduction of treated memory size. Noting this, an algorithm called bias-compensated (BC) MF-DFSE is considered that compensates the estimate of the untreated interference in the decision stages. These algorithms were compared with other more simple interference cancellation algorithms such as single- and multi-stage successive interference cancellation (SIC) and parallel interference cancellation (PIC) algorithms. Multi-stage SIC is especially notable because of its high performance and low complexity. Moreover, it is very easy to apply SIC algorithms to multi-rate systems. In Section 7.3, we consider SIC and PIC for multi-rate asynchronous DS-CDMA systems under frequency selective fading channels. We compare the performance of multi-stage SIC and PIC for variable-spreading gain and multi-code multi-rate sys-

tems.

7.1   Introduction

Maximum likelihood sequence estimation (MLSE) algorithms are optimal in the sense of the lowest packet error rate. There are two MLSE algorithms, one by Forney [39] and the other by Ungerboeck [40]. Forney's algorithm requires a noise-whitening filter that have to adapt to the channel variations. This can be a severe computational burden for DS-CDMA systems with long spreading codes. Consequently, the systems based on Forney's MLSE are not desirable in mobile communications. In contrast, Ungerboeck's MLSE algorithm does not require whitening process. We generalize the Ungerboeck's algorithm to asynchronous DS-CDMA systems under frequency and time selective multipath fading channels.

The complexity of MLSE can be reduced greatly by employing Viterbi algorithm [39]. However, it is still too huge in practical wireless communication environments. Consequently, various suboptimal algorithms are considered. Decision feedback sequence estimation (DFSE) algorithms [41, 42, 43, 44] are among the most natural variations of the MLSE algorithm. In DFSE, the number of state in the Viterbi algorithm is reduced and the effect of interference not considered is compensated by decision feedback. Although DFSE based on Forney's MLSE algorithm performs very well under causal channels, it is not desirable due to the whitening process. DFSE based on Ungerboeck's formulation, which we call matched-filter (MF) DFSE, is computationally less complex but is poor in performance. Hafeez and Stark [11] observed that such a poor performance stems from the fact that there exist untreated anti-causal interference due to the reduction of state. The modified algorithm with the compensation of the anti-causal interference with rudimentary interference estimation is called bias-compensated (BC) MFDFSE. We compare MFDFSE and BCMFDFSE for asynchronous frequency selective fading channels.

Algorithms that utilize previous decisions in the current or the future decision stages are generally called decision directed algorithms. DFSE is one of the most important decision direct algorithms. Other more ad-hoc decision direct algorithms are successive interference cancellation (SIC) and parallel interference cancellation algorithms. These two algorithms are easy to implement and computationally simple compared to other algorithms. In SIC [45, 46, 47, 48], signal is detected from the most reliable ones and then the estimated contribution is subtracted. If the detection is correct, then the cancellation removes the contribution of the interferer. Otherwise,

the interference is doubled. In PIC, tentative decisions by conventional single user detectors are made first and then the interference due to all other users are canceled before the detection. For a single-stage implementation, PIC generally outperforms SIC. However, for multi-stage realizations, SIC is better than PIC, which is also shown in our performance evaluations.

There are many other multiuser detection algorithms including various variations of aforementioned algorithms. Linear decorrelators [49, 50] and linear minimum mean square error (MMSE) detectors [51, 52, 53] are among the most important multiuser algorithms not considered in this work. Linear decorrelators demands the inversion of correlation matrix, which is not desirable in practical asynchronous DS-CDMA systems under frequency and/or time selective multipath fading environments due to computational complexity. By employing adaptive algorithms [54, 55], it is possible to reduce the complexity of matrix inversion greatly with maintaining excellent performance. Least square criterion is usually employed as the objective criterion in adaptive algorithms [56]. Adaptive detectors minimizing the expected squared estimation error are called adaptive minimum mean square error detectors [52, 57, 58]. Such adaptive algorithms often exhibits nice performance versus complexity tradeoff with and without the knowledge of various information about other users. Consequently, adaptive algorithms are useful in achieving good performance without severe computational burdens.

7.2   Multiuser Detection for Single-Rate Systems

In this section, we consider a single-rate asynchronous DS-CDMA system under frequency selective multipath fading channels. After describing the system model, we derive maximum likelihood sequence estimation (MLSE) based on Ungerboeck's algorithm [40]. Then, we introduce the matched-filter decision feedback sequence estimation (MFDFSE) and bias-compensated (BC) MFDFSE algorithms. After describing these algorithms, we compare the performance of various systems under realistic frequency selective multipath fading channels.

7.2.1   System Model

We consider a BPSK-modulated multiuser asynchronous direct sequence CDMA (DS-CDMA) system. Since the arrivals of symbols of different users are not synchronized, received symbols can be reordered linearly in terms of the beginning of the arrivals at the receiver. By doing this, the system can be described as a single user system in

which another symbol is being transmitted before the end of the transmissions of the previous symbols.

Let $h(t, \tau)$ be the baseband representation of the time varying impulse response of the overall system between symbol generator and receiver. To make the discussion simple, let's assume that there is no delay between transmitter and receiver. Then, $b \cdot h(t, \tau)$ is the received signal at time $t + \tau$ (except for the noise) when symbol b is generated at the transmitter at time t. Then, the baseband representation of the received signal $r(t)$ due to N consecutive data symbols is

$$r(t) = \sum_{i=1}^{N} b_i h(T_i, t - T_i) + n(t) \tag{7.1}$$

where b_i is i^{th} transmitted symbol, T_i is the time of the beginning of ith symbol reception and $n(t)$ is zero mean additive white complex Gaussian noise with two sided spectral density $N_0/2$. Note that T_i may not be replaced by $i\delta$ with some δ because the timing of the transmitted symbols of different users may not be equally separated. This formula gives a system model for the most general intersymbol interference channel including multiple access system.

7.2.2   MLSE

Maximum likelihood decision rule is an optimal rule for the detection of equally probable symbols that minimizes packet error probability. For the system described in the previous section, the maximum likelihood decision rule is to choose the set of symbols $\mathbf{b}^{(N)} = (b_1, \cdots, b_N)$ that maximizes the metric $M(\mathbf{b}^{(N)})$ defined by

$$M(\mathbf{b}^{(N)}) = \|r(t)\|^2 - \left\| r(t) - \sum_{i=1}^{N} b_i h(T_i, t - T_i) \right\|^2 \tag{7.2}$$

where $\|x(t)\|^2$ means

$$\|x(t)\|^2 = \int_{-\infty}^{\infty} |x(t)|^2 \, dt. \tag{7.3}$$

In Appendix G, this is proved using Karhuenen-Loève expansion. Manipulating the integrations, we get the simplification for the metric

$$M(\mathbf{b}^{(N)}) = \sum_{i=1}^{N} (b_i^* y_i + b_i y_i^*) - \sum_{i=1}^{N} \sum_{j=1}^{N} b_i^* S_{i,j} b_j \tag{7.4}$$

where

$$y_i = \int\limits_{-\infty}^{\infty} h^*(T_i, t - T_i) r(t)\, dt \tag{7.5}$$

$$S_{i,j} = \int\limits_{-\infty}^{\infty} h^*(T_i, t - T_i) h(T_j, t - T_j)\, dt \tag{7.6}$$

Here, we observe that $S_{i,j}$ does not depend on the transmitted symbols and depends only on the system characteristics. So the set of matched filter outputs $\{y_i\}$ forms sufficient statistics for the detection. Here, we note that the matched filter output y_i depends on the future symbols in the presence of intersymbol interference.

By observing the fact that $S_{j,i} = S_{i,j}^*$, we can rewrite the metric $M(\mathbf{b}^{(N)})$ as

$$M(\mathbf{b}^{(N)}) = \sum_{i=1}^{N} m_i \tag{7.7}$$

where

$$m_i = Re\left\{ b_i^* \Big(2y_i - S_{i,i} b_i - 2 \sum_{j<i} S_{i,j} b_j \Big) \right\} \tag{7.8}$$

We assume that there exist a number l such that $S_{i,j} = 0$ for all i and j such that $|i - j| > l$. The smallest such number L is called the memory of the system. In the case of multiuser access system under a frequency non-selective channel, the memory is the number of users minus 1.

Now since the metric $M(\mathbf{b}^{(N)})$ to be maximized is represented as a summation of m_i involving only the current statistic y_i, we can use Viterbi algorithm using m_i as our branch metric. Since the information of the previous L symbols is required for the evaluation of m_i, the number of states in the Viterbi algorithm is 2^L.

In m_i, the last term $\sum_{j<i} S_{i,j} b_j$ represents the interference between the ith and the previous symbols. Note that $\sum_{j<i} S_{i,j} b_j$ in $\sum_{j=1}^{i} m_j$ cancels the interference of b_i on the previous symbols only, while y_i introduces an anticausal interference component. In MLSE, this anticausal interference is eventually accounted for when a decision is made. However, in MFDFSE described below, this interference is only partially accounted for in the conditional decision of transmitted symbol.

7.2.3   DFSE

Decision feedback sequence estimation (DFSE) is a sub-optimal reduced-complexity detection scheme derived from MLSE. In DFSE, only the most recent $J(J < L)$ symbols are hypothesized in the Viterbi algorithm yielding M^J states. In other words, when the ith symbol is received, the $(i-J)^{th}$ symbol is decided conditionally on the most recent J symbols. The number J is called the memory order of DFSE. Traditionally DFSE was applied to whitened statistics. Hence, in our discussion, a similar algorithm applied directly to matched filter outputs is called matched filter decision-feedback sequence estimation (MFDFSE).

MFDFSE

In the MFDFSE algorithm, the branch metric u_i is defined by

$$u_i(b_i, \beta_{i-1,J}) = Re\left\{ b_i^*\left(2y_i - S_{i,i} - 2\sum_{j=i-J}^{i-1} S_{i,j}b_j - 2\sum_{j=i-L}^{i-J-1} S_{i,j}\hat{b}_j(\beta_{i-1,J}) \right) \right\} \qquad (7.9)$$

where $\beta_{i-1,J}$ represents the previous state $(b_{i-1}, \cdots, b_{i-J})$ in the Viterbi algorithm and $\hat{b}_{i-j}(\beta_{i,J})$ is the tentative decisions on the symbols conditionally on the state $\beta_{i,J}$. Let $U_{i-1}(\beta_{i-1,J})$ be the accumulated metric for the previous state $\beta_{i-1,J}$. Then, the tentative conditional decision $\hat{b}_{i-J}(\beta_{i,J})$ is given by

$$\hat{b}_{i-J}(\beta_{i,J}) = \mathrm{argmax}_{b_{i-J}}\left\{ U(\beta_{i-1,J}) + u_i(b_i, \beta_{i-1,J}) \right\} \qquad (7.10)$$

and the accumulated metric $U_i(\beta_{i,J})$ is calculated by

$$U_i(\beta_{i,J}) = U_{i-1}(\beta_{i-1,J}) + u_i(\hat{b}_{i-J}(\beta_{i,J}), \beta_{i-1,J}) \qquad (7.11)$$

BCMFDFSE

MFDFSE has two merits. First of all it does not involve computationally expensive whitening process. Secondly the performance of MFDFSE does not depend on the channel phase unlike the case of the original DFSE based on Forney's algorithm. However, MFDFSE generally does not perform as well as the original DFSE. This is because of the fact that the anticausal inference introduced by y_k for $k = i-J, i-J-1, \cdots$ is only partially compensated by $\sum_{k=1}^{i}\sum_{l=i-L}^{i-1} b_k^* S_{k,l}b_l$ upon the conditional decision of

b_{i-J}. The uncompensated component is

$$\sum_{j=i-L+1}^{i-J} Re\left\{2b_j^* \sum_{k=i+1}^{j+L} S_{j,k}b_k\right\} \tag{7.12}$$

where $b_{i+1}, \cdots, b_{i-J+L}$ are yet to appear. In the BCMFDFSE proposed by Hafeez and Stark [11], these future symbols are estimated using conventional detector outputs $\tilde{b}_k = \text{sign}(y_k)$ and the estimation of the uncompensated component using these conventional detector outputs is included in the decision step. So the modified decision rule becomes

$$\hat{b}_{i-J}(\beta_{i,J}) = \text{argmax}_{b_{i-J}}\left\{U(\beta_{i-1,J}) + u_i(b_i, \beta_{i-1,J}) - \text{bias}(\beta_{i-1,J})\right\}$$

where

$$\text{bias}(\beta_{i-1,J}) = \sum_{j=i-L+1}^{i-J} Re\left\{2b_j^* \sum_{k=i+1}^{j+L} S_{j,k}\tilde{b}_k\right\} \tag{7.13}$$

Note that the bias term is not included in the accumulated metric, because the anti-causal interference part is eventually contributed after the decision. The bias may be simplified by maintaining only the term involving b_{i-J}, i.e.,

$$\text{bias}(\beta_{i-1,J}) \approx \text{bias}(b_{i-J}) = 2b_{i-J}^* \sum_{k=i+1}^{i-J+L} S_{i-J,k}\tilde{b}_k \tag{7.14}$$

The simplified bias, $\text{bias}(b_{i-J})$, does not depend on the state because the terms $\tilde{b}_k$ for $k = i - L + 1, \cdots, i - J - 1$ are ignored.

The computational complexity of BCMFDFSE is comparable to that of MFDFSE and is roughly the same as that of MFDFSE when the approximate bias is used. BCMFDFSE can also be used in a multistage approach by feeding back the decisions of the previous stage for the calculation of the bias term.

7.2.4 Performance Evaluation

It is not easy to derive analytical performance results under complicated channel models. So our method of performance evaluation is simulation. First we describe the testing environments and then give simulation results.
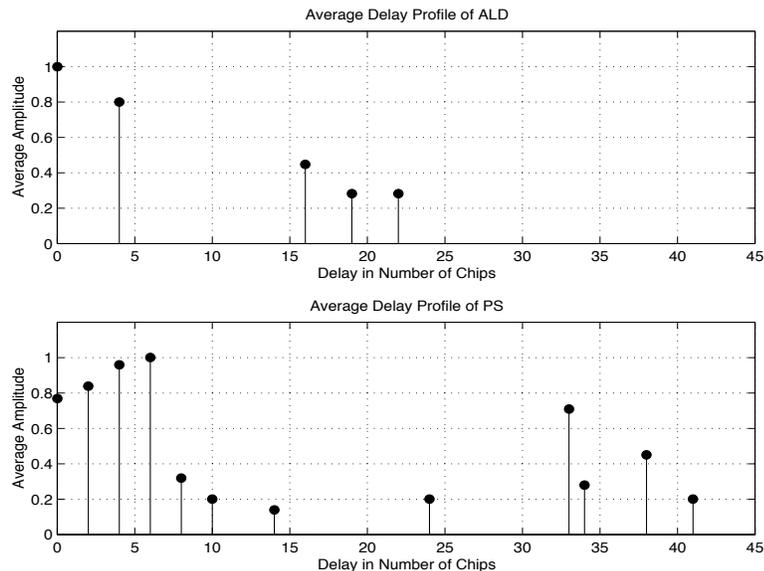
Figure 7.1: Average delay profile of ALD and PS

Testing Environments

Our simulation environments are based on the recently developed ETSI WCDMA standard. We assume the carrier frequency is 2GHz and that the chip rate is 4.096Mcps. For the uplink communication, the frame duration is 10 ms and each frame is divided into 16 power control slots. For each power control slot, the channel fading levels are virtually constant for any terrestrial vehicular speed at the carrier frequency of 2GHz. We tested our system under two different fading channels called American Legion Drive (ALD) and Pine Street (PS). D. C. Cox [59, 60] performed channel measurements at ALD and PS, from which we derived simplified channel models. ALD is a suburban area in New Jersey and PS is a street next to Wall street in New York. We simplified ALD and PS as 5- and 12-path models where the delays are multiples of chip duration and the fading levels of the paths are independent, complex and Gaussian distributed. The average delay profiles of ALD and PS are shown in Fig. 7.1.

In our simulations, we assumed 8 users with spreading gain 32 having the same average strength. We used random spreading codes for chip sequences of the users. Users are assumed to be symbol asynchronous but chip-synchronized and symbols of the $(i+1)^{th}$ $(i = 1, \cdots, 7)$ user are $4i-1$ chips delayed compared to the corresponding symbol of the first user. We assumed that each packet consists of 100 symbols and did not attempt any power control. We assumed that the channel fading levels are constant over a packet duration but independent from packet to packet and user to user. The channel memory of ALD and PS under these settings are 13 and 18. Because the

chip rate is fast enough compared to the change of the baseband signal and everything is assumed to be chip synchronized, we sampled, for discrete time representation, the signal once per each chip duration. And we assume that the receiver has an estimate of the chip-sampled version of $h(t, \tau)$.

We compared BCMFDFSE with conventional matched filter detection, MFDFSE, m-stage parallel interference cancellation (PIC) and m-stage successive interference cancellation (SIC) for $m = 1, 2$, and 4. For an asynchronous DS-CDMA system, there are many different ways to implement PIC and SIC. Some of them are discussed in [61] and [62]. For PIC, we first have to define which symbols are to be considered to correspond to the same symbol interval and we should cancel all interference from the past and future symbols as well as the current symbols. For example, for the single stage PIC, we first make tentative decisions on the symbols using conventional detector and use them to cancel the interference. For subsequent stages, decisions from the previous stage are used for the cancellation.

In the case of SIC, it first orders the users according to the average power of the matched filter outputs over the entire packet. Let's say that user 1 has the highest average power, user 2 has the next and so on. Now the SIC makes a (conventional hard) decision on the 1st symbol of user 1 and cancels the contribution of that symbol from the received signal. It then detects the 2nd symbol of user 1 and cancels the contribution of that symbol from the received signal. If there is interference between the 3rd symbol of user 1 and the 1st symbol of user 2, then the SIC detects the 3rd symbol of user 1 and cancels the contribution as well. After cancelling all interference between the symbols of user 1 and the 1st symbol of user 2, the SIC proceeds to detect the 1st symbol of user 2 and so on. Higher stage SIC starts with the cancelled received signal which is just the noise component if all decisions in the previous stage are correct. The canceled component corresponding to the symbol in decision is added back to the canceled signal to make a decision by matched filter output. The newly estimated contribution is cancelled again before proceeding to the next symbol.

Simulation Results

First we obtained the bit error rates (BER) for ALD and PS with the assumption that the receiver has perfect channel estimation. Figures 7.2 and 7.3 show the results. For various decision feedback sequence estimation algorithms, we denote by (J, D) the memory order J and decision lag D. For ALD, we obtained the BER for MLSE while it was not possible for PS due to complexity. The single stage BCMFDFSE algorithm outperforms all other single stage detection algorithms except for the MLSE algorithm.
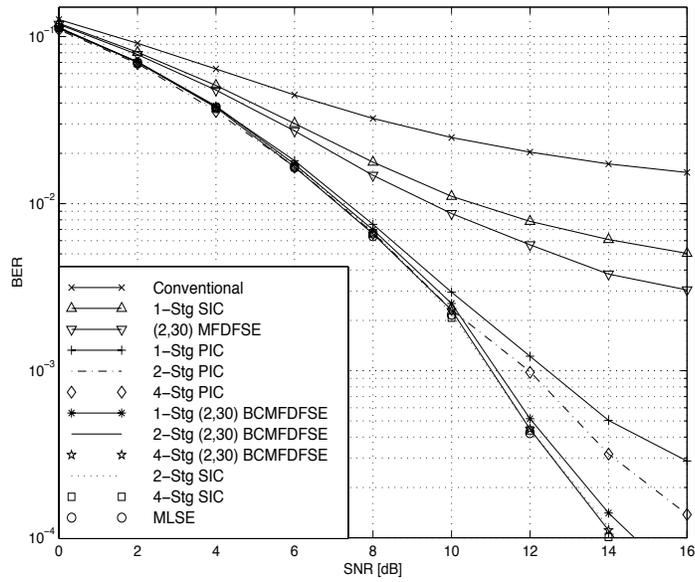
Figure 7.2: BER Performances under ALD channel with Perfect Channel Side Information
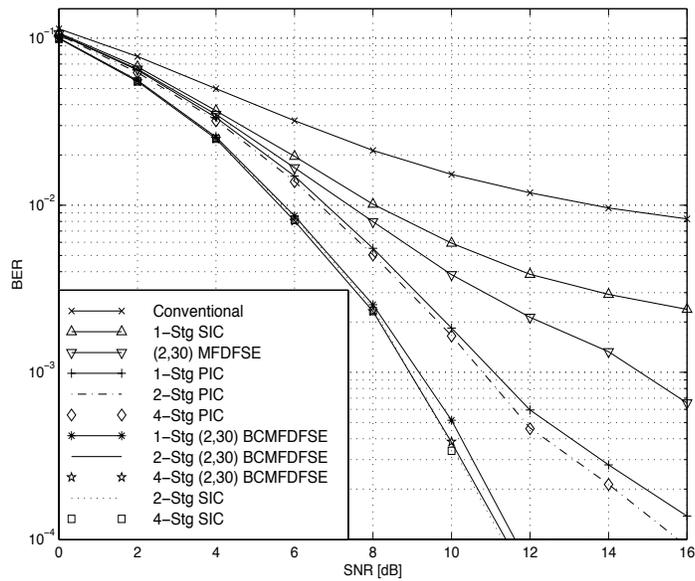
Figure 7.3: BER Performances under PS channel with Perfect Channel Side Information
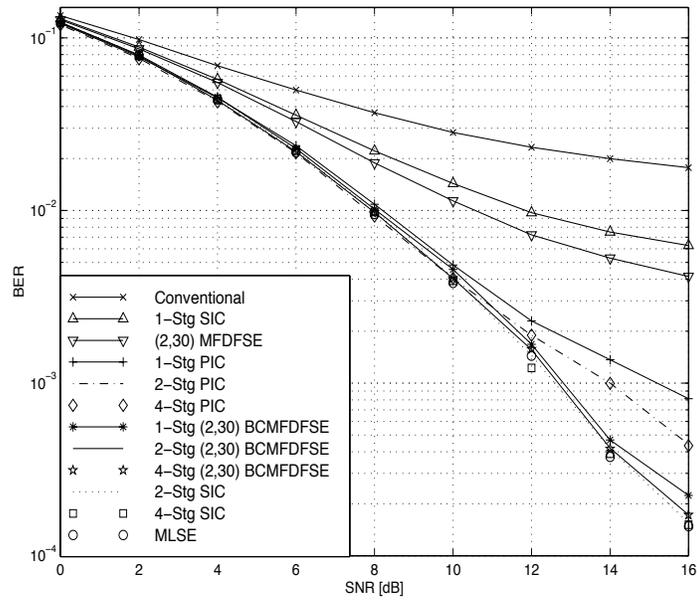
Figure 7.4: BER Performances under ALD channel with Imperfect Channel Side Information ($\sigma = 0.20$)
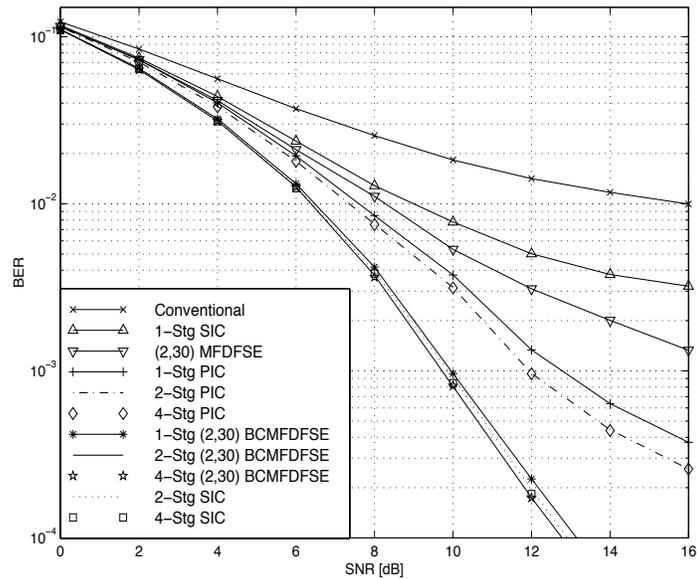
Figure 7.5: BER Performances under PS channel with Imperfect Channel Side Information ($\sigma = 0.20$)

In ALD channel, we see that the performance of BCMFDFSE is very close to that of MLSE up to the BER of 10^{-4}. Except for PIC, all multistage algorithms, in ALD channel, show near optimal performances upto the BER of 10^{-4}. The multistage SIC is especially notable because of the excellent performance and the minimal computational requirements. The excellent performance of a multistage SIC stems from the fact that the false interference cancellation is always compensated in the next-stage of the decision process.

We considered the effect of imperfect channel estimation on the algorithms. To simulate this situation, we multiplied the sampled system response by a Gaussian random variable with mean 1.0 and standard deviation $\sigma = 0.20$ and $\sigma = 0.05$. For $\sigma = 0.05$, the performance is almost the same as with perfect channel estimation, while we have small degradation of performance with $\sigma = 0.20$, which is shown in Figures 7.4 and 7.5. It was observed that the performance degradation due to imperfect channel information is roughly the same for all considered algorithms.

7.3   Interference Cancellation for Multi-Rate Systems

In this section, we examine the effectiveness of SIC and PIC in multirate environments. We first discuss the current research activities for multiuser detection under multirate DS-CDMA environments. Next we describe the system model and interference cancellation algorithms we consider. Then, we provide simulation results for the system.

7.3.1   Introduction

The advent of the third generation wireless communications initiated investigations of multiuser detection for multirate systems [63, 64]. The study of multiuser detection algorithms for multirate systems should be preceded by the specification of the access methods. There are many ways to support multirate transmission in a CDMA system. The two most widely considered schemes are variable spreading gain and multicode systems. With variable spreading gain, high rate is supported by choosing small spreading gain while keeping the chip rate the same. High rate service in a multicode scheme is supported by assigning multiple spreading codes to a single user. Varying the chip rate or assigning multiple carriers can also be considered. In this work, we consider variable spreading gain and multicode schemes for multi-rate support.

Recently, various decorrelators [65, 66] and decision feedback detectors [67] have been studied for multirate systems. In [65, 66], variable spreading gain schemes are

considered, while in [67] a comparison is made between variable spreading gain and multicode schemes with decision feedback detectors. In this work, we consider interference cancellation (IC) algorithms for the two access schemes in asynchronous DS-CDMA systems. There are various reasons for choosing simple IC algorithms for multirate systems. First of all, IC algorithms can easily be integrated with other communication technologies such as error correction coding or channel estimation. Secondly, when applied iteratively, they achieve very high performance in many cases. Thirdly, the computational complexity of IC schemes are relatively low compared to other systems. Fourthly, IC algorithms are easily applicable to multirate systems. Moreover, they are very robust under various non-idealistic situations.

7.3.2   System Model

We consider a dual-rate asynchronous DS-CDMA system under a frequency selective but time non-selective wide-sense stationary uncorrelated scattering Rayleigh fading channel. We assume the system is uncoded and that the signaling rate is fixed to be W chips per second for all users. There are K users accessing a common base-station at R_L bits per second. Each of these K users transmit signals with a single spreading code of processing gain $G_L = W/R_L$. In addition, we assume that there is one more user who accesses the base-station at R_H bits per second where R_H is much higher than R_L. We assume for ease of discussion that R_H/R_L is an integer.

We are interested in the multiuser detection algorithm at the base-station. The received signals from the $K + 1$ users are assumed to be chip-synchronous but not to be symbol synchronous. The $K + 1$ users are assumed to transmit signals over independent multipath channels which have the same statistical properties, namely, the same delay power profile. Typically, high rate service requires high quality of service and hence, we assume that transmitted power is maintained so that the average received power is high enough compared to that of low rate users to guarantee the quality of service.

We consider two schemes of multirate provision. In the first scheme, the high rate user uses a spreading gain $G_H = W/R_H$. In this case, all $K + 1$ users are assumed to be assigned random spreading codes. In the second scheme, the high rate user uses $K' = R_H/R_L$ different spreading codes. Again the spreading codes for the low rate users are assumed to be random spreading codes. As for the K' spreading codes assigned, they are assumed to be orthogonal one another.

We assume that all the low rate users maintain the transmitted signals so that their expected received signal strength are the same. The high rate user is assumed to

maintain higher average received signal strength level. Consequently, the situation is one strong high rate user and K weak low-rate users. We shall also consider various power control schemes, namely, perfect power control, imperfect power control, and no power control.

7.3.3   SIC and PIC

As usual, IC algorithms are preceded by a conventional matched filter detection. We assume that the receiver has perfect or imperfect knowledge about the channel state. Multi-stage SIC starts with power ranking algorithms. We call this stage the 0th stage of the SIC. During the 0th stage, the receiver evaluates the matched filter outputs of a conventional rake filter with maximal ratio combining over the whole packet and then ranks users according to the average of the square of the matched filter output. From now on, we call user 1 the user with highest average squared matched filter output. Similarly user $K + 1$ is the one with lowest average squared matched filter output. The first stage starts with the decision on the first symbol of the first user by the usual rake receiver. After detection, the contribution to the received signal due to that particular symbol is estimated and then canceled. Now, the receiver makes a similar decision on the second bit of the first user and cancels the estimated contribution. The receiver continues until the last symbol of the first user is detected and then canceled. Then, the receiver starts with the first bit of the second user. After completing the last symbol of the second user, the receiver starts with the first symbol of the third user. After the receiver detects and cancels the last symbol of the last user, we are left with a noise estimation and the first stage of SIC is complete.

The second stage starts with adding back to the noise estimation the estimated contribution of the first stage due to the first symbol of the first user. In other words, the canceled component is added back to the noise estimation. Even if the decision at the first stage is wrong, this newly constructed signal contains the signal component due to that particular symbol. Now the receiver makes a decision on the first symbol with this new signal and cancels the newly estimated contribution due to the first symbol of the first user. Then, it proceeds in a similar way to the second symbol of the first user. After finishing with the last symbol of the first user, the receiver proceeds toward the first symbol of the second user. Continuing this way, the receiver completes its second stage when it decides the last symbol of the last user. The third and higher stages are similarly done. Obviously the receiver algorithm described above is overly demanding in terms of delay and hardware requirements such as memory. There are many ways to reduce the complexity without having any difference in the performance

by judiciously choosing a windowing operation of the suggested algorithm.

The zeroth stage of our PIC algorithm is the conventional matched filter detection. After the decisions on all the symbols of the $K + 1$ users by the conventional receiver, the receiver obtains the noise estimation by canceling the estimated contributions due to the detected symbols. The situation is similar to the first stage of the SIC. The difference lies in the fact the receiver in this case proceeds detecting all the symbols of the system without canceling any contributions of other symbols. This difference stems from the parallelism in the application of the matched filter. The second stage is similar. To make a decision on a symbol of a user, the receiver just adds to the noise estimation of the first stage the previously canceled estimated contribution of the particular symbol of the particular user and then makes a decision using the matched filter. Note that the same noise estimation is used for all users and for all symbols. After completing the last symbol of the last user, the receiver recalculates the noise estimation by reconstructing the received signal used at the 0th stage and then by canceling the newly obtained estimated contributions of all users as in the 0th stage. This completes the first stage of the PIC. The second and third stages are similarly done.

7.3.4   Simulations

Rather than considering mathematical analysis under idealistic situations, we study the performance by simulations under more realistic situations. For the simulations, we have chosen the chip rate $W = 2.4$Mcps (mega chips per second). We assume that there are $K = 8$ low rate users with spreading gain $G_L = 24$. The spreading codes of the low rate users are random spreading codes. The packet duration of each users are chosen to be 10msec so that each packet consists of 1000 binary symbols. Both variable spreading gain and multicode schemes are considered. In the variable spreading gain scheme, the high rate user is assigned a spreading gain $G_H = 3$ with random spreading code. Again the packet duration is assumed to be 10ms and so the packet of the high rate user consists of 8000 bits. In the multicode scheme, the high rate user is assigned $K' = 8$ spreading codes. The 8 spreading codes are constructed as follows. First a master random spreading code is constructed and then 8 orthogonal spreading codes are obtained by multiplying by the Walsh Hadamard spreading code of period 8 which are the spreading codes constructed by repeating the rows of 8×8 Hadamard matrix. The high rate user is assumed to maintain the orthogonality between codes by keeping the synchronization, which is perfectly feasible. Because the 8 codes are mixed at the modulation stage of the high rate user, they are transmitted over the same fading

channel.

The modulated signals of the users are transmitted over independent WSSUS Rayleigh fading channels, all of which have the same delay power profile given in the following table. The channel is assumed to have 8 equally spaced taps with roughly exponentially decaying amplitude.

Relative Delay (μsec)	0	1	2	3	4	5	6	7
Relative Amplitude	1.0	0.9	0.8	0.7	0.6	0.5	0.4	0.3

In our simulations, the average signal to noise ratio of the high rate user is assumed to be 5dB above that of the low rate users. In the figures, E_b denotes the average energy (per bit) of low rate users and N_0 the one sided noise power spectral density. We assume that the power control against the shadowing loss is perfect. However, we consider the case in which power control cannot match multipath fading loss effectively. Figure 7.6 to Figure 7.9 are obtained with perfect power control and perfect channel side information. Figure 7.10 to Figure 7.13 are obtained with imperfect power control and imperfect channel side information. In this case, to simulate imperfect power control, we first randomly generate (following Rayleigh fading statistics and the delay power profile) the channel impulse response of each user and calculate the power by summing the squared fading levels at all delays, which we denote by P. In a perfect power control scheme, we re-scale the impulse response by multiplying by $1/\sqrt{P}$. In an imperfect power control situation, the re-scaling factor $(0.8 + 0.2\sqrt{P})/\sqrt{P}$ is chosen instead of $1/\sqrt{P}$. To simulate, the imperfect channel side information, for each delay, we multiply the fading level by a complex Random number $1 + X + jY$ ($j = \sqrt{-1}$) where X and Y are zero mean real Gaussian random variable with variance 0.1. By this process, we generate imprecise channel information supplied to the receiver. Figure 7.14 to Figure 7.17 are obtained without power control but with perfect channel side information. By without power control, we mean no power control for multipath fading. In all cases, the single user performance is obtained assuming single user transmission with 4-stage SIC for ISI cancellation. For all cases, the performance of 4-stage SIC is the best. Moreover, we find that it is roughly the same for both variable spreading gain and multicode schemes. However, the performance of 4-stage PIC is heavily dependent upon what access method is used. We see 4-stage PIC is far better with the multicode scheme, which is expected. In particular, we observe that PIC performs worse under no power control. As a side observation, we verify again that power control does still contribute to the performance enhancements.
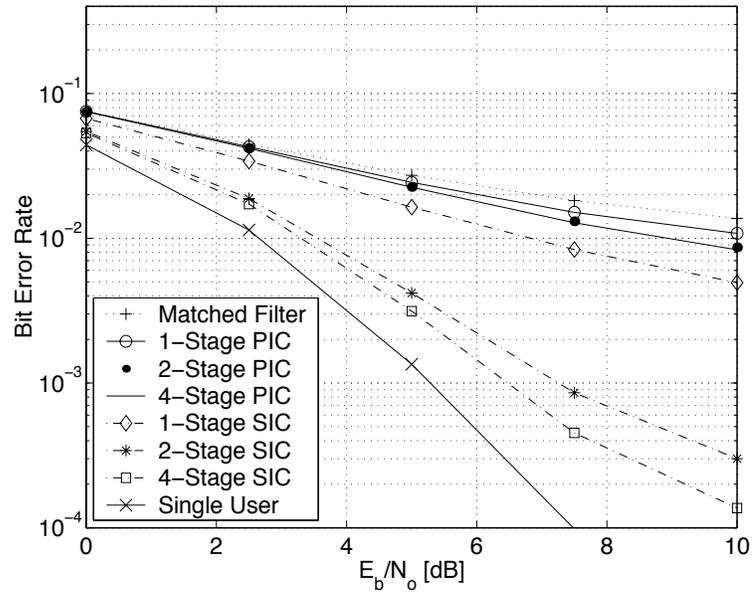
Figure 7.6: High rate user, Perfect power control, Perfect channel information, Variable spreading gain
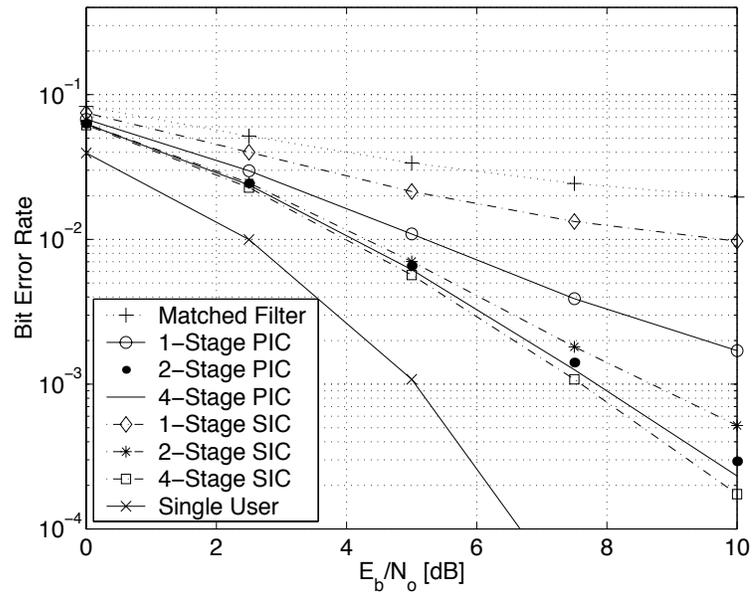
Figure 7.7: High rate user, Perfect power control, Perfect channel information, Multi-code.
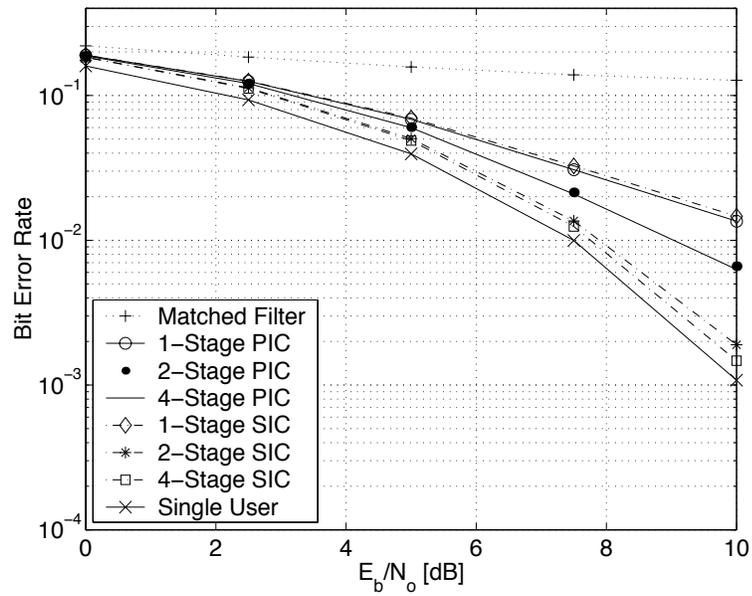
Figure 7.8: Low rate user, Perfect power control, Perfect channel information, Variable spreading gain
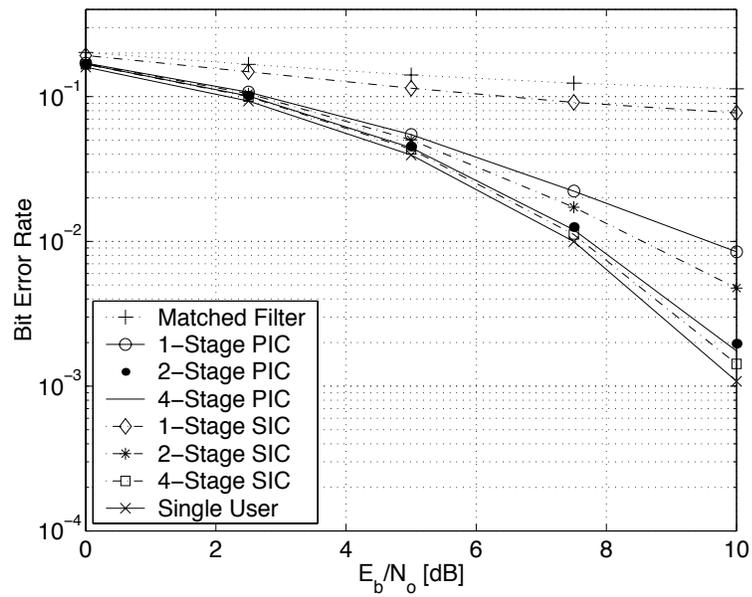
Figure 7.9: Low rate user, Perfect power control, Perfect channel information, Multi-code
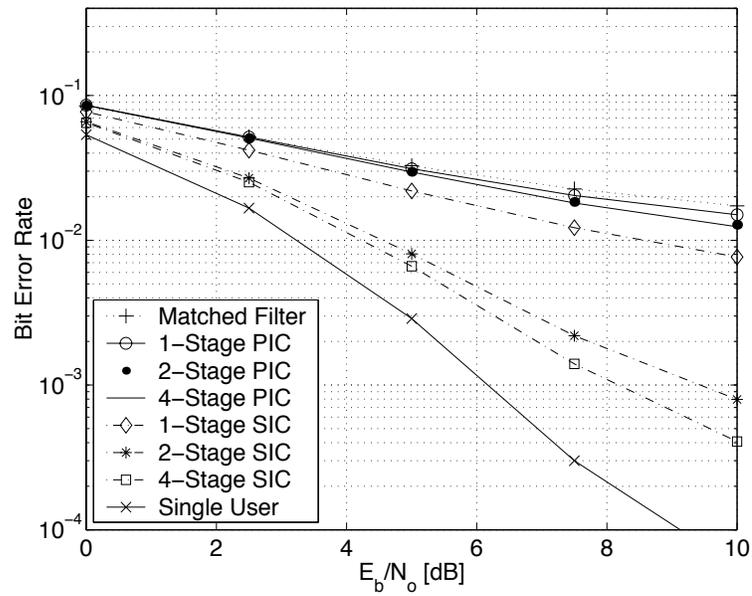
Figure 7.10: High rate user, Imperfect power control, Imperfect channel information, Variable spreading gain
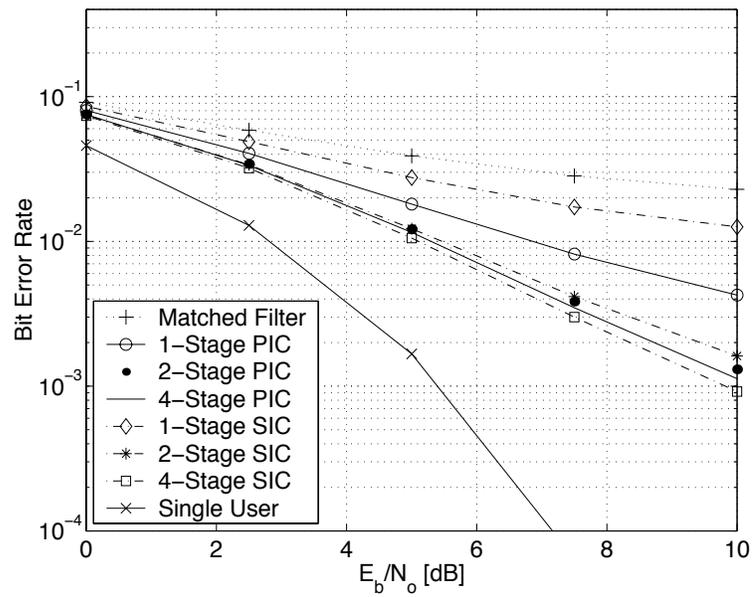
Figure 7.11: High rate user, Imperfect power control, Imperfect channel information, Multicode
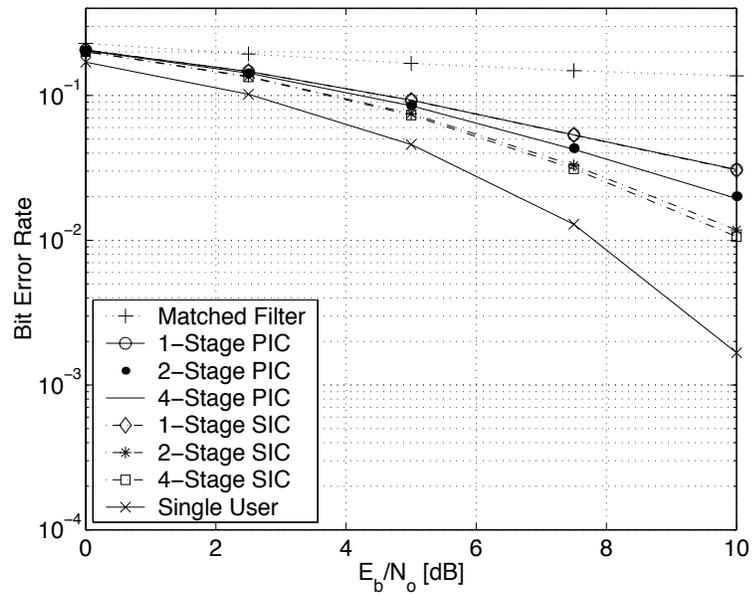
Figure 7.12: Low rate user, Imperfect power control, Imperfect channel information, Variable spreading gain
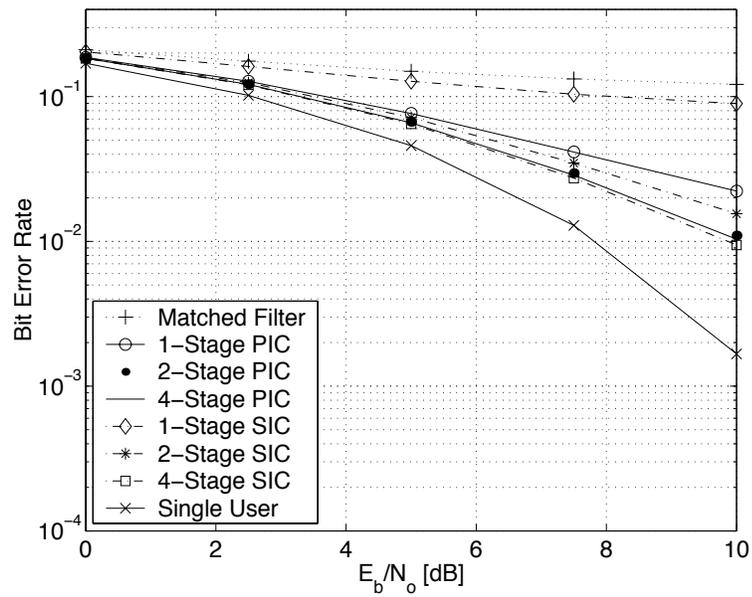
Figure 7.13: Low rate user, Imperfect power control, Imperfect channel information, Multicode
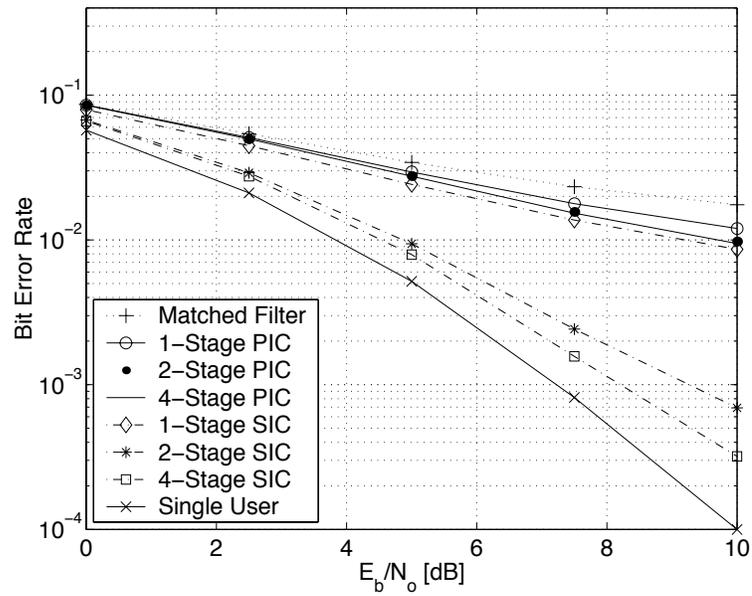
Figure 7.14: High rate user, No power control, Perfect channel information, Variable spreading gain
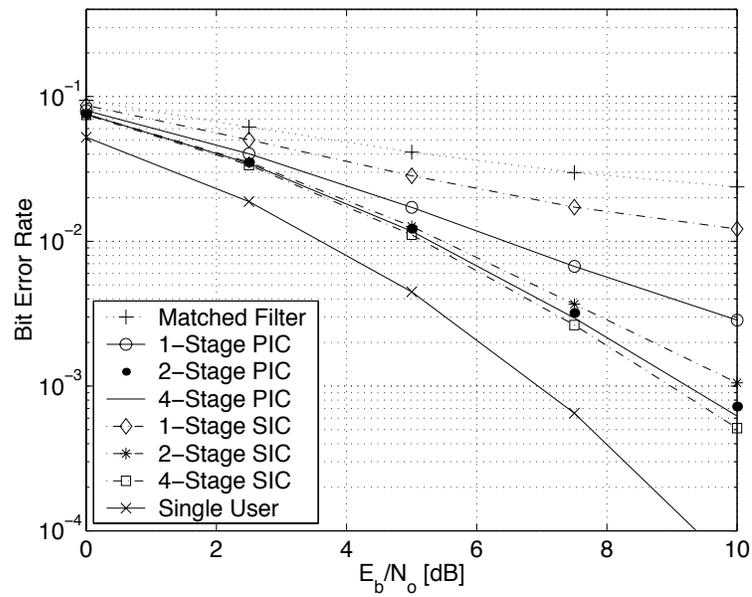
Figure 7.15: High rate user, No power control, Perfect channel information, Multicode
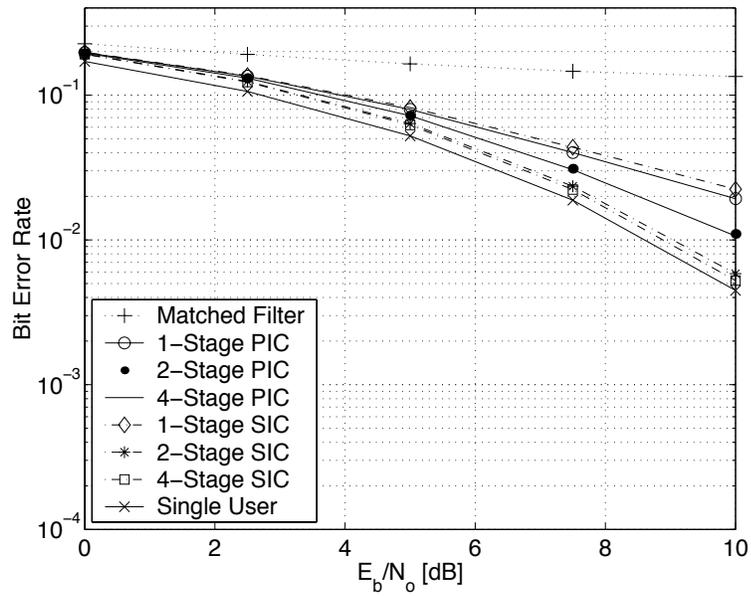
Figure 7.16: Low rate user, No power control, Perfect channel information, Variable spreading gain
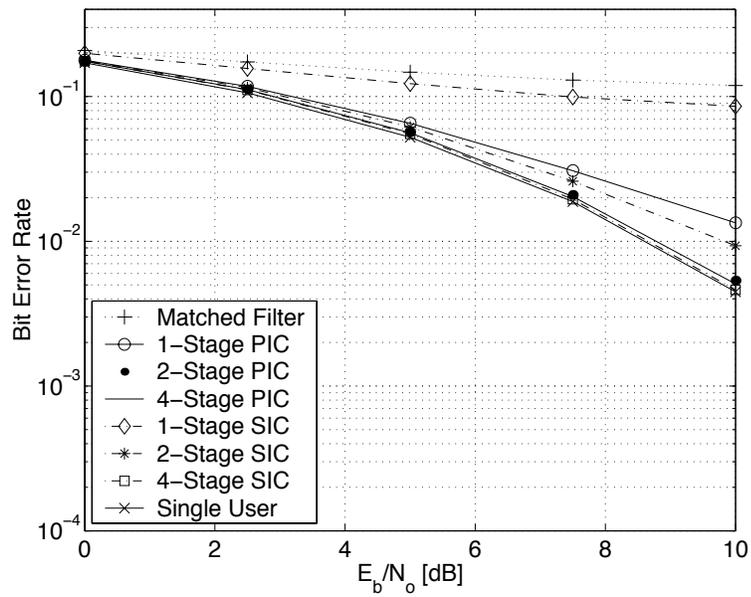
Figure 7.17: Low rate user, No power control, Perfect channel information, Multicode

7.4 Conclusion

In this chapter, we studied and compared various multiuser detection algorithms. We first derived a MLSE algorithm and introduced MFDFSE and BCMFDFSE. We found that BCMFDFSE achieve large performance gain compared with MFDFSE under various situations. We also considered simple IC algorithms. In particular, it was found that multi-stage SIC provides esepecially excellent performance versus complexity tradeoff. Observing this, we applied the multi-stage IC algorithms to multirate environments. We verified the multistage SIC performans very well under various situations. We found that the performance of multi-stage SIC is nearly the same for variable spreading gain and multicode schemes supporting the analysis by Ramakrishna and Holtzman [68]. Also it was found that power control is still desirable for the systems.

Chapter 8

Conclusion and Future Reseach

In this chapter, we conclude this report by summarizing the content of this report and discussing possible future research directions.

8.1 Summary of Contributions

The most important contribution of this report is the introduction of the normalized mean square covariance (NMSV) and the stochastic degree of freedom (SDF). The parameters are defined for general L^2 stochastic processes and studied in the context of wide-sense stationary uncorrelated scattering (WSSUS) channel characterization. By various analytical and experimental investigations, we verified very close relationship between the parameters and the performance of various communication systems. Due to such strong relationship, the parameters are very useful in characterizing the quality of a WSSUS channel.

The existence of such parameters are certainly beneficial to system design. Performance evaluation can be speculated directly in terms of the quality of channels, which makes the system design much more economical. As an example, we studied the diversity gain variation by frequency hopping rate change. We also provided an example to show that the parameters play important roles in optimal channel resource allocation. We showed that resource allocation problem can be studied by obtaining the relation between the parameters and particular resource allocation schemes without actual performance evaluations of underlying systems.

To generalize the validity of the parameters to non-WSSUS channels, we defined the re-centered normalized mean square correlation (RC-NMSR) which is the same as NMSV for WSSUS channels. The RC-NMSR was defined from the analysis of a simple diversity combining scheme. It was illustrated that the RC-NMSR also exhibits very

close relations with system performance for practical non-WSSUS channels. For most practical non-WSSUS channels, the specular components consist of a single strong impulse due to the line of sight path. However, supplementary parameters called normalized arithmetic variance (NAV) were defined to characterize the shape of the specular components of non-WSSUS channels. We showed that NAV is very effective in characterizing the specular components by verifying a strong relation with system performance.

We also studied and compared various multiple access interference (MAI) treatment algorithms. In particular, we derived Ungerboeck-type maximum likelihood sequence estimation algorithm (MLSE) for general asynchronous DS-CDMA systems under frequency and time selective multipath fading. Based on MLSE algorithm, we obtained matched filter decision feedback sequence estimation (MFDFSE) and bias-compensated (BC) MFDFSE. BC-MFDFSE was found to achieve large performance gain over MFDFSE without much increase in the system complexity. Single and multiple stage successive interference cancellation (SIC) and parallel interference cancellation (PIC) algorithms are also studied. Multiple stage SIC was especially impressive in achieving excellent performance with minimal computational complexity.

8.2   Future Research Direction

Due to the fundamental nature of the research presented in this report, it is possible to continue the research in various directions. First of all, the investigations of the information theoretic significance of the NMSV is among the most impending research topics. We will investigate the relations between NMSV and channel capacity and/or cutoff rate under various circumstances. Another important topic is the generalization of NMSV to multiple input multiple output (MIMO) channels, which is very closely related to information theoretic investigations. It will also be important to find out applications of the parameters in other branches of science and engineering.

The continuation of the research on the carrier-separated (CS) orthogonal frequency division multiplexing (OFDM) systems has a special important of practice. We believe they are very strong contenders for the next generation wireless communications due to its high performance, simplicity, and flexibility. Detailed specifications of adaptive coding and modulation schemes and dynamic resource allocation schemes should be followed.

The research on multiuser detection algorithms can be directed toward wide areas of equalization including cancellation of intersymbol interference under severely time-

dispersive channels or applications in spacetime decoder. In realistic applications, estimation and adaptation algorithms are especially important because the validity of the various algorithms depends on the knowledge of the channel status. Introduction of power control to the training sequences might result in an interesting topic.

Appendix A

Measure and Integration

The theory of measure and integration lies on the heart of the mathematical theory of probability and statistics. In this section, we review definitions and necessary mathematical terminologies in measure and integration, most of which can be found in standard analysis and probability textbooks [69, 13, 14]. We start with a brief introduction to the historical background for quick motivation of the theory of measure and integration. At the end of this section, we introduce the concept of the arithmetic mean for general measurable functions. The symbols $\mathbb{R}$, $\mathbb{R}^+$, $\bar{\mathbb{R}}$ and $\mathbb{C}$ will stand for the sets of all real numbers, all positive real numbers, all extended real numbers, and all complex numbers, respectively.

A.1   Introduction

Around the beginning of the twentieth century, it was one of the most important topics in mathematics to generalize the Riemann integral. Among various attempts, Lebesque's approach turned out to be the most successful [69]. To illustrate the idea, consider a function $f : [0, 1] \to [0, \infty)$. Roughly speaking, the Riemann integral $\int_0^1 f(t)\, dt$ is a limiting value of the sum

$$f(a_1)(t_1 - t_0) + f(a_2)(t_2 - t_1) + \cdots + f(a_n)(t_n - t_{n-1}) \tag{A.1}$$

where $0 = t_0 < a_1 < t_1 < a_2 < t_2 < \cdots < t_{n-1} < a_n < t_n = 1$. In other words, to define Riemann integral, we first divide the domain $[0, 1]$ into disjoint sub-intervals and choose a point in each interval. Then consider the sum of the form (A.1). If the sum converges to some unique number as we divide the sub-intervals into sub-intervals again and again, then we define the Riemann integral $\int_0^1 f(t)\, dt$ by the number. For example,

it is well-known that the sums converge if the function f is piecewise continuous in which case the value $f(a_k)$ tends to represent the function over the interval $[t_{k-1}, t_k]$ well enough in the limiting process. However, if the function f is not well-behaved, the Riemann integral is not defined in general.

In Lebesque's theory, more general partitions of the domain are considered rather than partitions into sub-intervals. Intuitively speaking, the idea starts from the division of the image of f into sub-intervals [70]. For example, partition the image of f into n disjoint sub-intervals $I_1, I_2, \cdots, I_n$ and let $E_k = f^{-1}(I_k)$ for $k = 1, 2, \cdots, n$. We then consider the sum

$$b_1 \lambda(E_1) + b_2 \lambda(E_2) + \cdots + b_n \lambda(E_n) \tag{A.2}$$

where $b_k \in I_k$ and $\lambda(E_k)$ is some measure of the size of the set E_k for each k. We note that the value b_k's represent typical values of I_k's as the intervals become narrow, regardless of the continuity of f. The Lebesque integral denoted by $\int_0^1 f(t)\, d\lambda(t)$ is defined by judiciously limiting the above sum with proper definition of the set function λ. Note that the set function λ appears in the notation of the integral since the value of the integral depends fundamentally on the choice of the set function λ. To quantify the sizes of sets properly, the set function must satisfy some conditions. A set function satisfying those conditions will be called a *measure*. However, it is often not possible to define a measure on all subsets of $[0, 1]$ (see for example, Chapter 3 in [71]). Hence, we have to restrict the domain of a measure to special collections of subsets of $[0, 1]$. Generally, a measure is defined on a σ-field of subsets. We note that $\lambda(f^{-1}(I_k))$ should be defined for each interval I_k for the sum (A.2) to be defined. We will define *measurable functions* to satisfy this condition.

Lebesque defined the function λ on the σ-field generated by all open sets in $[0, 1]$ to satisfy that $\lambda([a, b]) = b - a$ for any real numbers a and b with $a \le b$. Consequently, the measure λ is a generalization of the length function l defined on intervals by $l([a, b]) = b - a$. The measure λ is usually called the *Lebesque measure (in [0, 1])*. If the Riemann integral of f exists, then the Lebesque integral also exists and satisfies that

$$\int_{[0,1]} f(t)\, d\lambda(t) = \int_0^1 f(t)\, dt. \tag{A.3}$$

Lebesque measure and integral can be defined on more general subsets of Euclidean spaces [13]. Lebesque measure defined in the q-dimensional Euclidean space will be denoted by λ^q.

A.2   Measurable Spaces and Measurable Functions

Mathematical theories often become the most obvious in abstract settings. The theory of measure and integration is probably one such example. We consider a general non-empty set X. Similarly to the Lebesque measure λ, we want to define a set function μ called a *measure* on the subsets of X. As remarked previously, the measure cannot be defined suitably for all subsets of X in general. Instead we consider a collection $\mathfrak{X}$ of subsets of X such that

1. $E \in \mathfrak{X}$.

2. $X - E \in \mathfrak{X}$, if $E \in \mathfrak{X}$.

3. $\bigcup_{n=1}^{\infty} E_n \in \mathfrak{X}$, if $E_1, E_2, \cdots, \in \mathfrak{X}$.

We call such a collection $\mathfrak{X}$ a *σ-field* or *σ-algebra in X*. If $\mathfrak{X}$ is a σ-field in X, then we call the pair $(X, \mathfrak{X})$ a measurable space. The elements of $\mathfrak{X}$ are called *measurable sets in $(X, \mathfrak{X})$* or *$\mathfrak{X}$-sets*. Often we write X to stand for $(X, \mathfrak{X})$ if it does not cause any confusion. For example, the power set 2^X, namely, the class of all subsets of X is a σ-field in X. Let $\mathfrak{C}$ be any collection of subsets of X. Then, always there exists a unique σ-field in X denoted by $\sigma_X(\mathfrak{C})$ such that $\mathfrak{C} \subseteq \sigma_X(\mathfrak{C})$ and $\sigma_X(\mathfrak{C}) \subseteq \mathfrak{Y}$ for any σ-field $\mathfrak{Y}$ with $\mathfrak{C} \subseteq \mathfrak{Y}$. The σ-field $\sigma_X(\mathfrak{C})$ is called the *σ-field in X generated by $\mathfrak{C}$*.

We call the pair of a set X and a σ-field $\mathfrak{X}$ a measurable space because we can define a reasonably good set function called measure on $\mathfrak{X}$. A set function $\mu : \mathfrak{X} \to \mathbb{R}$ is called a *measure on $(X, \mathfrak{X})$* if

1. $\mu(\emptyset) = 0$.

2. $\mu(E) \geq 0$ for each $E \in \mathfrak{X}$.

3. $\mu(\bigcup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} \mu(E_n)$, if $E_1, E_2, \cdots$ are disjoint sets in $\mathfrak{X}$.

If μ is a measure on $(X, \mathfrak{X})$, then we call the triplet $(X, \mathfrak{X}, \mu)$ a *measure space*. Sometimes, a measure is called a *positive measure* to emphasize that it takes only non-negative values. If $\mu(X)$ is finite, then μ is said to be *finite*. If there exists a countable collection $\{E_n\}_{n=1}^{\infty}$ of sets in $\mathfrak{X}$ such that $\mu(E_n)$ is finite for each n and $\bigcup_{n=1}^{\infty} E_n = X$, then μ is said to be *σ-finite*. For example, the Lebesque measure defined in a compact set $[a, b]$ is a finite measure while the Lebesque measure defined in $\mathbb{R}$ is σ-finite.

Let $(X, \mathfrak{X}, \mu)$ be a measure space. Let $p(x)$ be a propositional function defined for all x in X. If $p(x)$ is true for all x in some $E \ (\in \mathfrak{X})$ with $\mu(X - E) = 0$, then $p(x)$ is said to hold *almost everywhere* in X with respect to the measure μ, for which we often write

'$p(x)$ a.e. x' or 'p a.e.' For example, consider two real valued functions $f, g : X \to \mathbb{R}$ and let E be a $\mathfrak{X}$-set with $\mu(X - E) = 0$. If $f(x) \le g(x)$ for any x in E, then we say '$f(x) \le g(x)$ almost everywhere (in X with respect to μ)' and write '$f \le q$ a.e.'

Let $(X, \mathfrak{X})$ and $(Y, \mathfrak{Y})$ be measurable spaces. Let $\mathfrak{C} = \{A \times B : A \in \mathfrak{X} \text{ and } B \in \mathfrak{Y}\}$. The elements of $\mathfrak{C}$ are called measurable rectangles in $\mathfrak{X} \times \mathfrak{Y}$. We usually write $\mathfrak{X} \otimes \mathfrak{Y}$ for the σ-field in $\mathfrak{X} \times \mathfrak{Y}$ generated by $\mathfrak{C}$. Now let μ and ν be σ-finite measures on $(X, \mathfrak{X})$ and $(Y, \mathfrak{Y})$, respectively. Then, there exists a unique σ-finite measure τ on $\mathfrak{X} \otimes \mathfrak{Y}$ such that

$$\tau(A \times B) = \mu(A)\nu(B), \tag{A.4}$$

for any $A \in \mathfrak{X}$ and $B \in \mathfrak{Y}$. We usually write $\mu \times \nu$ for τ and call it the product measure of μ and ν.

Let X be any set. Then, a collection $\mathfrak{T}$ of subsets of X is called a *topology in X* if

1. $\phi \in \mathfrak{T}$ and $X \in \mathfrak{T}$.

2. $E \cap F \in \mathfrak{T}$ if $E, F \in \mathfrak{T}$.

3. $\bigcup_{i \in I} E_i \in \mathfrak{T}$ for any collection $\{E_i\}_{i \in I}$ of sets in $\mathfrak{T}$.

If $\mathfrak{T}$ is a topology in X, we call the pair $(X, \mathfrak{T})$ a *topological space*. The elements of $\mathfrak{T}$ are called *open sets (in X relative to $\mathfrak{T}$)* or *$\mathfrak{T}$-open sets*. If $X - E \in \mathfrak{T}$, then E is called a *closed set*. A subset C of X is said to be *compact* if we can find a finite sub-collection $E_{i_1}, \cdots, E_{i_n}$ of $\{E_i\}_{i \in I}$ such that $C \subseteq \bigcup_{i=1}^{n} E_i$, for any collection $\{E_i\}_{i \in I}$ of open sets in X such that $C \subseteq \bigcup_{i \in I} E_i$.

If $(X, \mathfrak{T})$ is a topological space, then the σ-field $\sigma_X(\tau)$ in X generated by the topology τ is called the *collection of Borel set in $(X, \mathfrak{T})$* and is also denoted by $\mathfrak{B}(X, \mathfrak{T})$. We often write $\mathfrak{B}(X)$ for $\mathfrak{B}(X, \mathfrak{T})$ if there is no risk of confusion. If $X = \mathbb{R}^n$ or $X = \mathbb{C}^n$, the topology is usually defined by open-balls and is called the *usual topology*. Hence, when we write $\mathbb{R}^n$, for example, the topology is understood to be the usual topology studied in advanced calculus level [72].

Let $(X, \mathfrak{X})$ and $(Y, \mathfrak{Y})$ be measurable spaces. Then, a mapping $f : X \to Y$ is said to be *($\mathfrak{Y}$-$\mathfrak{X}$) measurable* if $f^{-1}(A) \in \mathfrak{X}$ for any $A \in \mathfrak{Y}$ [13, 14]. Let Y be $\mathbb{R}$, $\bar{\mathbb{R}}$, or $\mathbb{C}$. Then, f is said to be a *measurable function* if it is ($\mathfrak{B}(Y)$-$\mathfrak{X}$) measurable. Let $(W, \mathfrak{W})$ and $(Z, \mathfrak{Z})$ be an topological spaces. Then, a mapping $g : W \to Z$ is said to be *continuous* if $g^{-1}(A) \in \mathfrak{W}$ for any $A \in \mathfrak{Z}$.

A.3 Definition of Integrals

As discussed in Section A.1, the theory of measure was motivated by theory of integration. In this subsection, we briefly review how the integrals of measurable functions are defined. Throughout this subsection we assume that $(X, \mathfrak{X}, \mu)$ is a measure space. Let $f : X \to \bar{\mathbb{R}}$ be a measurable function. We first consider the case in which $f \geq 0$. Let $A_1, A_2, \cdots, A_n$ be nonempty disjoint $\mathfrak{X}$-sets such that $\bigcup_{k=1}^n A_k = X$ and consider the sum

$$\sum_{k=1}^n \left\{ \inf_{x \in A_k} f(x) \right\} \mu(A_k). \tag{A.5}$$

Such a class $\{A_k\}_{k=1}^n$ of measurable subsets is called a measurable partition of X. We define the supremum of the sum over all measurable partitions as the integral of f with respect to μ and write $\int_X f \, d\mu$, $\int_X f(x) \, d\mu(x)$, or $\int_X f(x) \, \mu(dx)$. Note that the supremum and hence the integral $\int_X f \, d\mu$ always exists, although it may be ∞, if f is a non-negative measure function. Hence, we define, for a non-negative measurable function $f : X \to \bar{\mathbb{R}}$,

$$\int_X f \, d\mu = \sup_{\{A_k\}} \left[\sum_{k=1}^n \left\{ \inf_{x \in A_k} f(x) \right\} \mu(A_k) \right]. \tag{A.6}$$

Now we consider more general case in which $f \geq 0$ does not hold. We first define the positive part f^+ of f

$$f^+(x) = \begin{cases} f(x) & \text{if } f(x) \geq 0, \\ 0 & \text{if otherwise,} \end{cases} \tag{A.7}$$

and the negative part f^- of f by

$$f^-(x) = \begin{cases} -f(x) & \text{if } f(x) \leq 0, \\ 0 & \text{if otherwise.} \end{cases} \tag{A.8}$$

Then, $f = f^+ - f^-$. If $\int_X f^+ \, d\mu < \infty$ or $\int_X f^- \, d\mu < \infty$, then we define $\int_X f \, d\mu$ by

$$\int_X f \, d\mu = \int_X f^+ \, d\mu - \int_X f^- \, d\mu. \tag{A.9}$$

If $\int_X f \, d\mu$ is finite, then we call f integrable. If f is measurable, then $|f|$ is also measurable and f is integrable if and only if $\int_X |f| \, d\mu < \infty$.

Now consider a complex-valued measurable function $h : X \to \mathbb{C}$. Let f and g be the

real and imaginary parts of h. We call h integrable if $\int_X |h|\,d\mu < \infty$. Both the functions f and g are integrable if and only if h is integrable. If h is integrable, we define

$$\int_X h\,d\mu = \int_X f\,d\mu + j\int_X g\,d\mu. \tag{A.10}$$

Let E be a measurable subset of X and let $1_E : X \to \mathbb{R}$ be the *characteristic function of* E defined by

$$1_E(x) = \begin{cases} 1 & \text{if } x \in E, \\ 0 & \text{if } x \notin E. \end{cases} \tag{A.11}$$

Then, 1_E and $h\cdot 1_E$ are measurable for any measurable function h. If $\int_X h\cdot 1_E\,d\mu$ exists, then we write $\int_E h\,d\mu$ for $\int_X h\cdot 1_E\,d\mu$.

For each integer p, we denote by $L^p(X,\mathfrak{X},\mu)$ the set of all measurable functions $f : X \to \mathbb{C}$ such that

$$\int_X |f|^p\,d\mu < \infty \tag{A.12}$$

and call it the L^p-space in $(X,\mathfrak{X},\mu)$.

A.4   Properties of Integrals

The integrals defined in Section A.3 are more general and flexible than Riemann integrals. In this subsection, we summarize various useful properties of the integrals. Throughout this subsection, we assume that a measure space $(X,\mathfrak{X},\mu)$ is given. We will denote by $\mathfrak{F}$ the collection of all extended real valued measurable functions $f : X \to \bar{\mathbb{R}}$ and by $\mathfrak{F}^+$ the collection of all non-negative functions in $\mathfrak{F}$. Also we will denote by $\mathfrak{G}$ the set of all complex-valued measurable functions on X.

Theorem A.1 (Integrals for Functions in $\mathfrak{F}^+$).
We first remark that the integral is defined in Section A.3 for any non-negative measurable function. In this theorem, we collect various important properties of integrals for non-negative measurable functions.

1. *For $f, g \in \mathfrak{F}^+$, if $f \le g$ a.e., then*

$$\int_X f\,d\mu \le \int_X g\,d\mu. \tag{A.13}$$

2. *(Monotone Convergence Theorem) Let $f, f_1, f_2, \cdots \in \mathfrak{F}^+$. Assume that $f_n \le f_{n+1}$*

a.e. for each n and that $\lim_n f_n = f$ a.e. Then,

$$\lim_{n \to \infty} \int_X f_n \, d\mu = \int_X f \, d\mu. \tag{A.14}$$

3. *Let $f_1, \cdots, f_m \in \mathfrak{F}^+$ and let $\alpha_1 \cdots, \alpha_m \in [0, \infty]$. Then, $\alpha_1 f_1 + \cdots + \alpha_m f_m \in \mathfrak{F}^+$ and*

$$\int_X (\alpha_1 f_1 + \cdots + \alpha_m f_m) \, d\mu = \alpha_1 \int_X f_1 \, d\mu + \cdots + \alpha_m \int_X f_m \, d\mu. \tag{A.15}$$

4. *Let $f_1, f_2, \cdots \in \mathfrak{F}^+$. Then, $\sum_{n=1}^{\infty} f_n \in \mathfrak{F}^+$ and*

$$\int_X \left(\sum_{n=1}^{\infty} f_n \right) d\mu = \sum_{n=1}^{\infty} \left(\int_X f_n \, d\mu \right). \tag{A.16}$$

5. *(Fatou's Lemma) Let $f_1, f_2, \cdots \in \mathfrak{F}^+$. Then,*

$$\int_X \left(\liminf_{n \to \infty} f_n \right) d\mu \leq \liminf_{n \to \infty} \int f_n \, d\mu. \tag{A.17}$$

In particular, if there exists a function $f \in \mathfrak{F}^+$ such that $\lim_n f_n = f$ a.e., then

$$\int_X f \, d\mu \leq \liminf_{n \to \infty} \int f_n \, d\mu. \tag{A.18}$$

Theorem A.2 (Integrals for Functions in $\mathfrak{F}$).
Integrals in Section A.3 are not defined for all measurable functions. In this theorem, we collect various important properties of integrals of general measurable functions.

1. *Let $f, g \in \mathfrak{F}$. If $f \leq g$ a.e., then*

$$\int_X f \, d\mu \leq \int_X g \, d\mu \tag{A.19}$$

   if the integrals are defined.

2. *Assume that $f \in \mathfrak{F}$ and $\alpha \in \mathbb{R}$ or that $f \in \mathfrak{G}$ and $\alpha \in \mathbb{C}$. If $\int_X f \, d\mu$ is defined, then $\int_X (\alpha f) \, d\mu$ is defined and*

$$\int_X (\alpha f) \, d\mu = \alpha \int_X f \, d\mu. \tag{A.20}$$

3. *Assume that $f_1, \cdots, f_m \in \mathfrak{F}$ and $\alpha_1, \cdots, \alpha_m \in \mathbb{R}$ or that $f_1, \cdots, f_m \in \mathfrak{G}$ and $\alpha_1, \cdots, \alpha_m \in \mathbb{C}$. If $f_1, \cdots, f_m$ are integrable, then $\alpha_1 f_1 + \cdots + \alpha_m f_m$ is integrable*

and

$$\int_X \left(\alpha_1 f_1 + \cdots + \alpha_m f_m \right) d\mu = \alpha_1 \int_X f_1 \, d\mu + \cdots + \alpha_m \int_X f_m \, d\mu. \tag{A.21}$$

4. *Let $f \in \mathfrak{F}$ or $f \in \mathfrak{G}$. Then, $|f| \in \mathfrak{F}^+$. If $\int_X f \, d\mu$ is defined, then*

$$\left| \int_X f \, d\mu \right| \leq \int_X |f| \, d\mu. \tag{A.22}$$

5. *(Dominated Convergence Theorem) Let $f, f_1, f_2, \cdots \in \mathfrak{F}$ or $f, f_1, f_2, \cdots \in \mathfrak{G}$. Assume that there exists an integrable function $g \in \mathfrak{F}^+$ such that $|f_n| \leq g$ a.e. for all n and that $f = \lim_n f_n$ a.e. Then, f and f_n are integrable for all n and*

$$\lim_{n \to \infty} \left(\int_X f_n \, d\mu \right) = \int_X f \, d\mu. \tag{A.23}$$

Theorem A.3 (Fubini's Theorem).
Let $(X, \mathfrak{X}, \mu)$ and $(Y, \mathfrak{Y}, \nu)$ be measure spaces. Assume that μ and ν are σ-finite. Let f be an extended real-valued measurable function or a complex valued measurable function on $X \times Y$.

1. *Assume that $f \geq 0$ and define $F : X \to \bar{\mathbb{R}}$ by*

$$F(x) = \int_Y f(x, y) \, d\nu(y). \tag{A.24}$$

   Then, F is a measurable function on X and

$$\int_X F \, d\mu = \int_{X \times Y} f \, d(\mu \times \nu). \tag{A.25}$$

2. *Assume that f is integrable (with respect to $\mu \times \nu$). Then, there exists a set $E \in \mathfrak{X}$ with $\mu(E) = 0$ such that $\int_Y f(x, y) \, d\nu(y)$ exists for all $x \in X - E$. If we define a function F on X by*

$$F(x) = \begin{cases} \int_Y f(x, y) \, d\mu(y), & x \in X\text{-}E, \\ 0, & x \in E, \end{cases} \tag{A.26}$$

   then F is a measurable function on X with well-defined integral $\int_X F \, d\mu$ such that

$$\int_X F \, d\mu = \int_{X \times Y} f \, d(\mu \times \nu). \tag{A.27}$$

Appendix B

NMSV's in terms of Scattering Function

In this section, we derive the formulae in Section 4.2.1 for a type-2 restriction of the WSSUS channel of Section 4.2. From the defintion of $P(f,t)$ in (3.59), we have

$$\iint_{\Omega} \iint_{\Omega} P(f - f', t - t') P^*(f - f', t - t') \, df \, dt \, df' \, dt' \tag{B.1}$$

$$= \iint_{\Omega} \iint_{\Omega} \iiiint_{\mathbb{R}^4} e^{-j2\pi(f-f')(\tau-\tau')+j2\pi(\nu-\nu')(t-t')} \tag{B.2}$$

$$\cdot q(\tau,\nu) q^*(\tau',\nu') \, d\tau \, d\nu \, d\tau' \, d\nu' \, df \, dt \, df' \, dt' \tag{B.3}$$

$$= \iiiint_{\mathbb{R}^4} \left| \iint_{\Omega} e^{-j2\pi f(\tau-\tau')+j2\pi(\nu-\nu')t} \, df \, dt \right|^2 q(\tau,\nu) q^*(\tau',\nu') \, d\tau \, d\nu \, d\tau' \, d\nu' \tag{B.4}$$

and

$$P(0,0) = \iint_{\mathbb{R}^2} q(\tau,\nu) \, d\tau \, d\nu. \tag{B.5}$$

Consequently,

$$\mathcal{V}_{ft}(\Omega) = \frac{\iint_{\Omega} \iint_{\Omega} |P(f - f', t - t')|^2 \, df \, dt \, df' \, dt'}{\left[P(0,0) \iint_{\Omega} df \, dt \right]^2} \tag{B.6}$$

$$= \frac{\iiiint_{\mathbb{R}^4} \left| \int_{\Omega} e^{-j2\pi f(\tau-\tau')+j2\pi(\nu-\nu')t} \, df \, dt \right|^2 q(\tau,\nu) q^*(\tau',\nu') \, d\tau \, d\nu \, d\tau' \, d\nu'}{\left[\iint_{\mathbb{R}^2} q(\tau,\nu) \, d\tau \, d\nu \iint_{\Omega} df \, dt \right]^2}. \tag{B.7}$$

For $\mathcal{V}_f(\Theta;t)$ and $\mathcal{V}_t(f;T)$, choose $\Omega = \Theta \times (t - \delta, t + \delta)$ and $\Omega = (f - \delta, f + \delta) \times T$ and

take the limit $\delta \to 0$ in (B.7).

Appendix C

Calculations of a Kernel

Let $I_k = [A + k(v + w), A + k(v + w) + w]$. Then,

$$\left| \int_\Theta e^{j2\pi f\tau} df \right|^2 = \sum_{k=1}^{K-1} \sum_{k'=1}^{K-1} \int_{I_k} e^{-j2\pi f\tau} df \int_{I_{k'}} e^{-j2\pi f'\tau} df' \qquad (C.1)$$

$$= \sum_{k=1}^{K-1} \sum_{k'=1}^{K-1} w^2 \text{sinc}^2(\pi w\tau) e^{-j2\pi\tau(w+v)(k-k')} \qquad (C.2)$$

$$= w^2 \text{sinc}^2(\pi w\tau) \left| \sum_{k=0}^{K-1} e^{j2\pi\tau(w+v)k} \right|^2 \qquad (C.3)$$

$$= w^2 \text{sinc}^2(\pi w\tau) \left[\frac{\sin\{K\pi(w+v)\tau\}}{\sin\{\pi(w+v)\tau\}} \right]^2. \qquad (C.4)$$

Since $\int_\Theta df = Kw$, we have (4.17).

Appendix D

BER for Simple Diversity Combining

Let $H_k = H_k^R + jH_k^I$, $n_k = n_k^R + jn_k^I$. Then,

$$s = \sqrt{E_c} \sum_{k=1}^{K} \left[(H_k^R)^2 + (H_k^I)^2 \right] + \sum_{k=1}^{K} \left[H_k^R n_k^R + H_k^I n_k^I \right]. \tag{D.1}$$

Here, given all H_k^R's and H_k^I's, $\sum_{k=1}^{K} \left[H_k^R n_k^R + H_k^I n_k^I \right]$ is a Gaussian random variable with mean zero and variance $\frac{N_0}{2} \sum_{i=1}^{K} \left[(H_i^R)^2 + (H_i^I)^2 \right]$ so that

$$P(s < 0 | \mathbf{H} = \mathbf{h}) = Q\left(\sqrt{\frac{2E_c \|\mathbf{h}\|^2}{N_0}} \right) \tag{D.2}$$

where $\mathbf{H}$ denotes the $2K-$dimensional random vector $(H_1^R, H_1^I, \cdots, H_K^R, H_K^I)$ and $\mathbf{h}$ denotes a $2K-$ dimensional vector realization of $\mathbf{H}$. Consequently, we can write

$$P_b = P(s < 0) = \int_{\mathbf{R}^{2K}} f_{\mathbf{H}}(\mathbf{h}) Q\left(\sqrt{\frac{2E_c \|\mathbf{h}\|^2}{N_0}} \right) d\mathbf{h} \tag{D.3}$$

where $f_{\mathbf{H}}$ is the probability density function of $\mathbf{H}$. Here, observing that $Q(\sqrt{\frac{2E_c \|\mathbf{h}\|^2}{N_0}})$ depends only on $\|\mathbf{h}\|^2$, define $R = \|\mathbf{H}\|^2$. Then, we have

$$P_b = \int_0^\infty f_R(r) Q\left(\sqrt{\frac{2E_c r}{N_0}} \right) dr \tag{D.4}$$

where f_R is the probability density function of the random variable R.

Now from the relation between f_R and the characteristic function ψ_R, namely, from

$$f_R(r) = \frac{1}{2\pi} \int_0^\infty \psi_R(\beta) e^{-j\beta r} d\beta, \tag{D.5}$$

we get

$$P_b = \frac{1}{2\pi} \int_{-\infty}^\infty \psi_R(\beta) \int_0^\infty e^{-j\beta r} Q\left(\sqrt{\frac{2E_b r}{N_0}}\right) dr \, d\beta \tag{D.6}$$

in which

$$\int_0^\infty e^{-j\beta r} Q\left(\sqrt{\frac{2E_b r}{N_0}}\right) dr = \int_0^\infty e^{-j\frac{\beta N_0}{2E_b} x^2} \left(\int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt\right) \frac{N_0}{E_b} x \, dx \tag{D.7}$$

$$= \frac{N_0}{\sqrt{2\pi} E_c} \int_0^\infty e^{-\frac{1}{2}t^2} \left(\int_0^t x \, e^{-j\frac{\beta N_0}{2E_c} x^2} dx\right) dt \tag{D.8}$$

$$= \frac{j}{\sqrt{2\pi}\beta} \int_0^\infty \left[e^{-\frac{1}{2}(1+j\frac{\beta N_0}{E_c})t^2} - e^{-\frac{1}{2}t^2}\right] dt \tag{D.9}$$

$$= \frac{1}{\frac{2E_c}{N_0} + 2j\beta + \sqrt{\frac{2E_c}{N_0}\left(\frac{2E_c}{N_0} + 2j\beta\right)}}. \tag{D.10}$$

Here, (D.7) follows from the change of variables ($x = \sqrt{2E_b r/N_0}$) and the definition of $Q(\cdot)$. Then, the Fubini's theorem with two different representations of the same domain of integration provides (D.8). By performing the integration in the parenthesis of (D.8), we obtain (D.9), which yields (D.10) after direct evaluations of the integrations and rearrangement. Finally, after substituting (D.10) into (D.5), we can obtain (4.50) by rearranging the real and imaginary parts.

Appendix E

Various Expressions for RC-NMSR's

In this section, we derive (5.16) - (5.24). To make the descriptions concise, we define the frequency-time correlation function $P(f,t)$ by

$$P(f,t) = \int_{-\infty}^{\infty} p(\tau,t)e^{-j2\pi f\tau}\,d\tau. \tag{E.1}$$

Then, we have

$$V_H(f,t;f',t') = P(f-f',t-t') \tag{E.2}$$

and

$$P(f,t) = \iint_{\mathbb{R}^2} q(\tau,\nu)e^{-j2\pi(f\tau-\nu t)}\,d\tau\,d\nu. \tag{E.3}$$

First note

$$R_H(f,t;f',t') = P(f-f',0) + m_H(f,t)m_H^*(f',t') \tag{E.4}$$

and hence that

$$|R_H(f,t:f',t')|^2 - |m_H(f,t)|^2|m_H^*(f',t')|^2 \tag{E.5}$$

$$= |P(f-f',t-t')|^2 + 2\operatorname{Re}[P(f-f',t-t')m_H^*(f,t)m_H(f',t')] \tag{E.6}$$

where $\operatorname{Re}[\cdot]$ denotes the real part of a complex number $\cdot$.

Consequently,

$$\iint_\Omega R_H(f,t';f,t)\,df\,dt = \iint_\Omega \left[P(0,0)+|D(f,t)|^2\right]df\,dt \tag{E.7}$$

$$= \iint_{\mathbb{R}^2} q(\tau,\nu)\,d\tau\,d\nu \iint_\Omega df\,dt + \iint_\Omega |D(f,t)|^2\,df\,dt \tag{E.8}$$

$$= \iint_\Omega df\,dt \cdot \left(\iint_{\mathbb{R}^2} q(\tau,\nu)\,d\tau\,d\nu + \langle |m_H|^2\rangle_\Omega\right) \tag{E.9}$$

$$\iint_\Omega \iint_\Omega P(f-f',t-t')P^*(f-f',t-t')\,df'\,dt'\,df\,dt \tag{E.10}$$

$$= \iint_\Omega \iint_\Omega \iiiint_{\mathbb{R}^4} e^{-j2\pi(f-f')(\tau-\tau')+j2\pi(\nu-\nu')(t-t')} \tag{E.11}$$

$$\cdot q(\tau,\nu)q^*(\tau',\nu')\,d\tau'\,d\nu'\,d\tau\,d\nu\,df'\,dt'\,df\,dt \tag{E.12}$$

$$= \iiiint_{\mathbb{R}^4} \left|\iint_\Omega e^{-j2\pi f(\tau-\tau')+j2\pi(\nu-\nu')t}\,df\,dt\right|^2 q(\tau,\nu)q^*(\tau',\nu')\,d\tau'\,d\nu'\,d\tau\,d\nu \tag{E.13}$$

and

$$\iint_\Omega \iint_\Omega P(f-f',t-t')D^*(f,t)D(f,t')\,df\,dt\,df'\,dt' \tag{E.14}$$

$$= \iint_\Omega \iint_\Omega \iiiint_{\mathbb{R}^4} q(\tau,\nu)e^{-j2\pi\{(f-f')\tau-(t-t')\nu\}}d^*(\tau',t)e^{j2\pi f\tau'} \tag{E.15}$$

$$\cdot d(\tau'',t')e^{-j2\pi f'\tau''}\,d\tau\,d\nu\,d\tau'\,d\tau''\,df\,dt\,df'\,dt' \tag{E.16}$$

$$= \iiiint_{\mathbb{R}^4} q(\tau,\nu)\left[\iint_\Omega \iint_\Omega d^*(\tau',t)d(\tau'',t')\right. \tag{E.17}$$

$$\left.\cdot e^{j2\pi\{(\tau'-\tau)f-(\tau''-\tau)f'+t\nu-t'\nu\}}\,df\,dt\,df'\,dt'\right]d\nu\,d\tau\,d\tau'\,d\tau''. \tag{E.18}$$

By plugging (E.9-E.18) into (5.10), we obtain (5.16). To obtain (5.20) and (5.24), we set $\Omega = \Theta \times (t-\delta,t+\delta)$ and $\Omega = (f-\delta,f+\delta) \times T$ in (5.16) and take the limit $\delta \to 0$.

Appendix F

The Characteristic Function of R

To obtain the expression (5.56), we start from (5.54) with the definition $R_k = |D'_k + S'_k|^2$. Then we have $\psi_R(\beta) = \prod_{k=1}^K \psi_{R_k}(\beta)$ since $R_1, \cdots, R_K$ are independent and $R = \sum_{k=1}^K R_k$. Here ψ_{R_k} denotes the characteristic function of R_k. By the way, $R_k = R_k^R + R_k^I$ where

$$R_k^R = \left\{ \mathrm{Re}[D'_k] + \mathrm{Re}[S'_k] \right\}^2 \tag{F.1}$$

and

$$R_k^I = \left\{ \mathrm{Im}[D'_k] + \mathrm{Im}[S'_k] \right\}^2. \tag{F.2}$$

Note R_k^R and R_k^I are independent non-central chi-square random variables so that $\psi_{R_k} = \psi_{R_k^R} \psi_{R_k^I}$. Now by direct computation of the characteristic functions $\psi_{R_k^R}$ and $\psi_{R_k^I}$, we can show

$$\psi_{R_k}(\beta) = \frac{1}{1 - j\beta\lambda_k} e^{j\frac{|D'_k|^2\beta}{1 - j\beta\lambda_k}}. \tag{F.3}$$

The characteristic function of a non-central chi-square random variable can also be found from (2.1-116) of [35]. Note that we need to use (5.53).

Appendix G

Maximum Likelihood Sequence Estimation

Assume that a signal $r(t)$ starts at t_i and ends at t_f at the receiver. Let $\{\phi_k\}_{k=1}^{\infty}$ be a complete set of orthonormal functions on the compact interval $[t_i, t_f]$. Then, $r(t)$, $h(T_i, t - T_i)$, and $n(t)$ can be expanded as

$$r(t) \;=\; \sum_{k=1}^{\infty} r_k \phi_k(t) \tag{G.1}$$

$$h(T_i, t - T_i) \;=\; \sum_{k=1}^{\infty} h_{i,k} \phi_k(t) \tag{G.2}$$

$$n(t) \;=\; \sum_{k=1}^{\infty} n_k \phi_k(t) \tag{G.3}$$

Here, $\{n_k\}_{k=1}^{\infty}$ is a sequence of independent identically distributed complex Gaussian random variables with mean zero and variance $N_0/2$.

Let $p(\mathbf{r}^{(n)})$ and $p(\mathbf{r}^{(n)}|\mathbf{b}^{(N)})$ be the probability density functions of the random vector $\mathbf{r}^{(n)} = (r_1, \cdots, r_n)$ when nothing is transmitted and $\mathbf{b}^{(N)}$ are transmitted, respectively. Then the likelihood ratio $\Lambda_n(\mathbf{b}^{(N)})$ defined by

$$\Lambda_n(\mathbf{b}^{(N)}) = \frac{p(\mathbf{r}^{(n)}|\mathbf{b}^{(N)})}{p(\mathbf{r}^{(n)})} \tag{G.4}$$

is given by

$$\Lambda_n(\mathbf{b}^{(N)}) = \frac{\exp\left\{ -\frac{1}{N_0} \sum_{k=1}^{n} \left| r_k - \sum_{i=1}^{N} b_i h_{i,k} \right|^2 \right\}}{\exp\left\{ -\frac{1}{N_0} \sum_{k=1}^{n} |r_k|^2 \right\}} \tag{G.5}$$

so that

$$\Lambda_n(\mathbf{b}^{(N)}) = \exp\left\{ \frac{1}{N_0} \sum_{k=1}^{n} \left(|r_k|^2 - \left| r_k - \sum_{i=1}^{N} b_i h_{i,k} \right|^2 \right) \right\}. \tag{G.6}$$

Now let $\Lambda(\mathbf{b}^{(N)}) = \lim_{n\to\infty} \Lambda_n(\mathbf{b}^{(N)})$. Then the maximum likelihood decision rule is to choose $\mathbf{b}^{(N)}$ that maximizes $\Lambda(\mathbf{b}^{(N)})$. But, since

$$\Lambda(\mathbf{b}^{(N)}) = \exp\left\{ \frac{1}{N_0} \sum_{k=1}^{\infty} \left(|r_k|^2 - \left| r_k - \sum_{i=1}^{N} b_i h_{i,k} \right|^2 \right) \right\}, \tag{G.7}$$

the maximum likelihood decision rule is to choose $\mathbf{b}^{(N)}$ that maximizes the metric $M(\mathbf{b}^{(N)})$ defined by

$$\begin{aligned}
M(\mathbf{b}^{(N)}) &= \sum_{k=1}^{\infty} \left(|r_k|^2 - \left| r_k - \sum_{i=1}^{N} b_i h_{i,k} \right|^2 \right) \tag{G.8} \\
&= \int_{-\infty}^{\infty} |r(t)|^2 \, dt - \int_{-\infty}^{\infty} \left| r(t) - \sum_{i=1}^{N} b_i h(T_i, t - T + i) \right|^2 dt. \tag{G.9}
\end{aligned}$$

Bibliography

[1] A. G. Longley and P. L. Rice, *Prediction of tropospheric transmission loss over irregular terrain, a computer method*, ESSA Technical Report ERL 79-ITS 67, U.S. Department of Commerce, Boulder, CO, July 1968.

[2] A. G. Longley and R. K. Reasoner, *Comparison of propagation measurements with predicted values in the 20 to 10,000 MHz Range*, ESSA Technical Report ERL 148-ITS 97, Department of Commerce, Boulder, CO, Jan. 1970.

[3] A. G. Longley, *Radio propagation in Urban Areas*, U.S. Department. of Commerce OTR 78-144, Apr. 1978.

[4] M. N. Lustgarten and J. A Madison, "An empirical propagation model(EPM-73)," *IEEE Transactions on Electromagnetic Compatibility*, vol. EMC-19, no. 3, Aug. 1977.

[5] Y. Okumura, E. Ohmori, T. Kawano, and K. Fukuda, "Field strength and its variability in VHF and UHF land mobile service," *Review -Electrical Communication Laboratory*, vol. 16, pp. 825–873, September-October 1968.

[6] M. Hata, "Code combining – a maximum-likelihood decoding approach for combining an arbitrary number of noisy packets," *IEEE Transactions on Vehicular Technology*, vol. VT-29(3), pp. 317–325, 1980.

[7] T. Kailath, "Time-variant communication channels," *IEEE Transactions on Information Theory*, pp. 233–237, 1963.

[8] P. A. Bello, "Characterization of randomly time-variant linear channels," *IEEE Transactions on Communication Systems*, vol. CS-11, pp. 360–393, Dec. 1963.

[9] R. S. Kennedy, *Fading Dispersive Communication Channels*, John Wiley & Sons, 1969.

[10] S. Verdu, *Multiuser Detection*, Cambridge University Press, 1998.

[11] A. Hafeez and W. E. Stark, "Decision feedback sequence estimation for un-whitened isi channels with applications to multiuser detection," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 0, pp. 1785 – 1795, Dec. 1998.

[12] D.-S. Yoo, A. Hafeez, and W. E. Stark, "Trellis-based multiuser detection for ds-cdma systems with frequency-selective fading," *IEEE Wireless Communications and Networking Conference*, vol. 2, pp. 829 – 833, 1999.

[13] P. Billingsley, *Probability and Measure*, Wilely-Interscience Publication, 1995.

[14] K. Jacobs, *Measure and Integral*, Academic Press, Inc., 1978.

[15] R. B. Ash, *Topics in Stochastic Processes*, Academic Press, Inc, 1975.

[16] Y. M. Berezansky, Z. G. Sheftel, and G. F. Us, *Functional Analysis, Vol. I*, Birkhäuser Verlag, 1996.

[17] J. D. Jackson, *Classical Electrodyanamics, 2nd ed.*, John Wiley & Sons, 1990.

[18] J. S. Schwinger, L. L. DeRaad, Jr, K. A. Milton, and W.-Y. Tsai, *Classical Electro-dynamics*, Perseus Books, 1998.

[19] T. S. Rappaport, *Wireless Communications*, IEEE Press, 1996.

[20] M. Schwartz, *Principles of Electrodynamics*, McGraw-Hill, 1972.

[21] W. S. Ament, "Toward a theory of reflection by a rough surface," *Proceedings of the IRE*, vol. 41, no. 21, pp. 1343–1349, January 1953.

[22] L. Boithias, *Radio Wave Propagation*, McGraw-Hill, Inc., 1987.

[23] D. Gabor, "Theory of communication," *Journal of IEE*, vol. 93, pp. 429–457, November 1946.

[24] A. D. Poularikas, *The Transforms and Applications Handbook*, CRC Press, Inc., 1996.

[25] J. Dugundji, "Envelopes and pre-envelopes of real waveforms," *IRE Transactions on Information Theory*, pp. 53–57, 1958.

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

[27] K. Pahlavan and A. H. Levesque, *Wireless Information Networks*, John Wiley & Sons, Inc., 1995.

[28] R. L. Peterson, R. E. Ziemer, and D. E. Borth, *Introduction to Spread Spectrum Communications*, Prentice Hall, 1995.

[29] M. J. Gans, "A power spectral theory of propagation in the mobile raido environment," *IEEE Transactions on Vehicular Technology*, vol. 21, pp. 27 – 38, Feb. 1972.

[30] W. C. Y. Lee, *Mobile Communications Engineering, 2nd Edition*, McGraw-Hill Book Company, 1997.

[31] R. J. Brockwell and R. A. Davis, *Time Series - Theory and Methods, 2nd ed.*, Springer Verlag, 1991.

[32] J. Omura and T. Kailath, "Some useful probability distribution," *Technical Report No. 7050-6*, Sept. 1965.

[33] F. J. MacWilliams and N. J. A. Sloane, *The Thoery of Error-Correcting Codes*, North-Holland, 1977.

[34] W. C. Jakes, *Microwave Mobile Communications, IEEE ed.*, IEEE Press, 1994.

[35] J. G. Proakis, *Digital Communications, 4th ed.*, McGraw-Hill, 2000.

[36] A. J. Viterbi, *CDMA: Principles of Spread Spectrum Communication*, Addision-Wesley, 1995.

[37] J. A. C. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come," *IEEE Communications Magazine*, vol. 28, pp. 5–14, May 1990.

[38] B. Hirosaki, "An orthogonal multiplexed qam system using the discrete fourier transform," *IEEE Tranasctions on Communications*, jul 1981.

[39] G. D. Forney, "Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference," *IEEE Transactions on Information Theory*, vol. 18, no. 3, pp. 363 – 378, May 1972.

[40] G Ungerboeck, "Adaptive maximum-likelihood receiver for carrier-modulated data-transmission systems," *IEEE Transactions on Communication Technology*, vol. 22, no. 3, pp. 624 – 636, May 1974.

[41] A. Duel-Hallen and C. Heegard, "Delayed decision-feedback sequence estimation," *IEEE Transactions on Communications*, vol. 37, no. 5, pp. 428 – 436, May 1989.

[42] A. Duel-Hallen, "Decorrelating decision-feedback multiuser detector for synchronous CDMA," *IEEE Transactions on Communications*, vol. 41, pp. 285–290, Feb. 1993.

[43] A. Duel-Hallen, "A family of multiuser decision-feedback detectors for asynchronous code-division multiple-access channels," *IEEE Transactions on Communications*, vol. 43, pp. 421–434, Feb./Mar./Apr. 1995.

[44] M. K. Varanasi, "Decision feedback multiuser detection: A systematic approach," *IEEE Transactions on Information Theory*, vol. 45, no. 1, pp. 219–240, 1999.

[45] R. Kohno, M. Hatori, and H. Imai, "Cancellation techniques of cochannel interference in asynchronous spread spectrum multiple access systems," *Electronics and Communications*, vol. 60-A, pp. 20 − 29, 1983.

[46] R. Kohno, "Pseudo-noise sequences and interference cancellation techniques for spread-spectrum systems-spread spectrum theory and techniques in japan," *IEICE Transactions*, vol. E.74, pp. 1083 − 1092, May 1991.

[47] P. Dent, B. Gudmundson, and M. Ewerbring, "CDMA-IC: a novel code division multiple access scheme based on interference cancellation," *Proceedings of 3rd IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, pp. 98–102, Sept. 1992.

[48] P. Patel and J. Holtzman, "Analysis of a ds/cdma successive interference cancellation scheme using correlations," *IEEE Journal on Selected Areas in Communications*, vol. 12, pp. 796 − 807, June 1994.

[49] D. Shnidman, "A generalized nyquist criterion and an optimum linear receiver for a pulse modulation system," *Bell System TEchnical Journal*, vol. 46, pp. 2163–2177, nov 1967.

[50] R. Lupas and S. Verdu, "Near-far resitatance of multiuser detectors in asynchronous channels," *IEEE Transactions on Communications*, vol. 38, pp. 496–508, Apr. 1990.

[51] Z. Xie, C. Rushforth, R. Short, and T. Moon, "A family of suboptimum detectors for coherence multiuser communications," *IEEE Journal on Selected Areas in Communications*, vol. 8, pp. 683–690, May 1990.

[52] U. Madhow and M. L. Honig, "MMSE interference suppression for direct-sequence spread spectrum CDMA," *IEEE Transactions on Communications*, vol. 42, pp. 3178–3188, Dec. 1994.

[53] H. V. Poor and S. Verdu, "Proability of error in MMSE multiuser detection," *IEEE Transactions on Information Theory*, vol. 43, pp. 858–871, May 1997.

[54] D. Chen and S. Roy, "And adaptive multiuser receiver for cdma systems," *IEEE Journal on Selected Areas in Communications*, vol. 12, no. 5, pp. 808–816, June 1994.

[55] U. Mitra and H. V. Poor, "Adaptive receiver algorithms for near-far resistant cdma," *IEEE Transactions on Communications*, vol. 43, pp. 1713–1724, Feb./Mar./Apr. 1995.

[56] V. Solo and X. Kong, *Adaptive Signal Processing Algorithms*, Prentice-Hall Inc., 1995.

[57] S. Miller, "An adaptive direct-sequence code-division multiple-access receiver for multi-user interference rejection," *IEEE Transactions on Communications*, vol. 43, pp. 1746–1755, Apr. 1995.

[58] R. Rapajic and B. Vucetic, "Adaptive receiver structures for asynchronous CDMA systems," *IEEE Journal on Selectred Areas in Communications*, vol. 12, pp. 685–697, May 1994.

[59] D. C. Cox, "Delay-doppler characteristics of multipath propagation at 910 MHz in a suburban mobile radio environment," *IEEE Transactions on Antennas and Propagation*, vol. 20, pp. 625–635, Sept. 1972.

[60] D. C. Cox, "910 MHz urban mobile radio propagation: Multipath characteristics in new york city," *IEEE Transactions on Communications*, vol. 21, pp. 1188–1194, Nov. 1973.

[61] R. M. Buehrer, N. S. Correal, and B. D. Woerner, "Comparison of multiuser receivers for cellular cdma," *IEEE Global Telecommunications Conference*, vol. 3, pp. 1571 – 1577, 1996.

[62] H. Sugimoto, L. K. Rasmussen, T. J. Lim, and T. Oyama, "Mapping functions for successive interference cancellation in cdma," *IEEE Vehicular Technology Conference*, vol. 3, pp. 2301 – 2305, 1998.

[63] A. Baier, U.-C. Fiebig, W. Granzow, W. Koch, P. Teder, and J. Thielecke, "Design study for a CDMA-based third-generation mobile radio system," *IEEE Journal on Selected Areas in Communications*, vol. 12, no. 4, pp. 733–743, May 1994.

[64] T. Ottosson and A. Svensson, "Multi-rate schemes in DS/CDMA systems," *Proceedings of 1994 IEEE Vehicular Technology Conference*, vol. 2, pp. 1006 – 1010, 1995.

[65] M. Saquib, R. Yates, and N. Mandayam, "A decision feedback decorrelator for a dual rate synchronous DS/CDMA system," *Wireless Networks*, vol. 4, no. 6, pp. 497 – 506, Oct. 1998.

[66] M. Saquib, R. Yates, and A. Ganti, "An asynchronous multirate decorrelator," *IEEE Transactions on Communications*, vol. 48, no. 5, pp. 739–742, may 2000.

[67] U. Mitra, "Comparison of ML-based detection for two multirate access schemes for CDMA signals," *IEEE Transactions on Communications*, vol. 47, no. 1, pp. 64–77, Jan. 1999.

[68] S. Ramakrishna and J. M. Holtzman, "A comparison between single code and multiple code transmission schemes in a CDMA system," *Proceeding of 1998 IEEE Vehicular Technology Conference*, pp. 791–795, 1998.

[69] W. Rudin, *Real and Complex Analysis*, McGraw-Hill Book Company, 3rd edition, 1987.

[70] M. Reed and B. Simon, *Methods of Modern Mathematical Physics I, Functional Analysis*, Academic Press, Inc., 1980.

[71] H. L. Royden, *Real Analysis, 3rd ed.*, Prentice Hall, Inc., 1988.

[72] W. Rudin, *Principles of Mathematical Analysis*, McGraw-Hill Book Company, 3rd edition, 1976.