2007

# A Joint data rate - error rate analysis in correlated space-time-wireless channels

Hui Tong
*Michigan Technological University*

# A Joint Data Rate - Error Rate Analysis in Correlated Space-Time Wireless Channels

by

Hui Tong

A DISSERTATION

Presented to the Faculty of

The Graduate College at the Michigan Technological University

In Partial Fulfillment of Requirements

For the Degree of Doctor of Philosophy

Major: Electrical Engineering

Under the Supervision of Prof. Seyed Alireza Zekavat

Houghton, Michigan

2007

A Joint Data Rate - Error Rate Analysis in Correlated Space-Time Wireless Channels

Hui Tong, Ph.D. Candidate

Michigan Technological University, 2007

Advisor: Prof. Seyed Alireza Zekavat

Future generations of wireless communications demand a) high data rate and b) low error rate. However, information theory has revealed that in general the two performance metrics can not be optimized simultaneously. In other words, there is a tradeoff between data rate and error rate. Recently, this tradeoff is studied in multi-input-multi-output (MIMO) fading channels, where multi-antenna is employed at both transmitter side and receiver side.

Original works on the tradeoff assumes rich scattering environments, i.e., fadings across different antenna pairs are independent. But, in real applications, scattering sources can be poor. For example, rural, on highway, or on the top of a skyscraper. Therefore, practically fadings across different antenna pairs are not independent: they are correlated. It is motivated to study the tradeoff in poor scattering environments. However, solving the interested problem is not as trivial as it looks like. In fact, it has to be decomposed into several simpler problems.

First, the tradeoff is investigated in single-input-single-output (SISO) channels, where temporal correlations are incorporated. A novel concept, spike sharpening effect, is introduced to interpret the tradeoff in SISO channels.

Second, the single-input-multi-output (SIMO) channels are studied, where multiple antennas are employed at the receiver side. Error performance in SIMO channels is analyzed and simulated under perturbations and non-stationarity considerations.

Thirdly, correlated MIMO channel modeling is studied, i.e., spatially correlated fadings are considered. A unified framework is proposed to accommodate several stochastic models that are presented in the literature. Moreover, under this framework, it is depicted that different models are feasible in different physical environments.

Finally, an innovative channel capacity analysis reveals that, in high signal-to-noise ratio regimes, a correlated MIMO channel can be approximately decomposed into a linear summation of multiple parallel SISO and SIMO channels. Then, combining the previous results, the tradeoff in poor scattering MIMO channels is established.

The above results are not only meaningful for abstract analysis, but also useful for practical algorithm development in wireless communications. They help wireless engineer to incorporate practical considerations (Doppler spread, perturbation, stationarity, poor scattering, etc) into their design, and aim to build a system that can approach its ultimate performance in real environments.

# Acknowledgements

I was lucky to meet all the people during the amazing journey of my Ph.D. This thesis would not have been possible without the support of them.

Foremost, I wish to thank my advisor, Prof. Seyed A. Zekavat, for his guidance over the past four years. Not only his broad knowledge in wireless communications, but also his thorough working attitude have helped me to go through the process of carving out a Ph.D. He has certainly invested the most time overseeing the research, followed by thoroughly reviewing this report. I have definitely benefited from his extraordinary motivation, great intuition, and technical insight. I would feel honored if my thinking and working capability has been improved to just a small fraction of his outstanding qualities.

I am grateful to Prof. Michael C. Roggemann, Prof. Jindong Tan, and Prof. Will H. Cantrell for taking the time to serve as my advisory committee members. Their insightful and invaluable comments improve the quality of the dissertation greatly. No doubt this thesis would not contain the quality nor quantity if there were no help from them.

Special thanks go to two professors who I encountered during my Ph.D. studies: Prof. Zhi (Gerry) Tian for introducing me the fascinating world of MIMO, and Prof. Venu V. Veeravalli for pointing out the relationship between diversity-multiplexing tradeoff and my original ideas.

I am indebted to my friends, Xiukui Li and Zhonghai Wang, for their help both technical wise and in personal life. I also want to thank my colleagues, Jafar Pour-

rostam, Wenjie Xu, and Babak A. Bastami, for collaborating with them in writing several papers. My thanks extend to other members of the wireless communications research group at Michigan Tech., Mohsen Pourkhaataoun, Senjuti Basu, Taha A. Khalaf and Xiaofeng Yang, for they maintain a comfortable atmosphere in our lab.

Last, but not the least, I give sincere thanks to my parents, my sister and my wife, for their continuous and unconditional support. It is your love that encouraged me to reach the end of this long journey.

*Hui Tong*

*June 2007*

# Contents

# List of Figures

# List of Tables

# List of Symbols

| | |
|---|---|
| $A$ | Square of amplitude of channel fading |
| $\mathbf{A}$ | Fourier Transform Matrix |
| $B$ | Band Width |
| $c$ | The Speed of Light |
| $\mathcal{C}$ | Channel Capacity |
| $d$ | Diversity Gain (or Antenna Separation) |
| $d_c$ | Duty Cycle |
| $E_b$ | Bit Energy |
| $E_s$ | Symbol Energy |
| $\mathbf{G}$ | Total Correlation Matrix |
| $\mathbf{G}_T$ | Transmit Correlation Matrix |
| $\mathbf{G}_R$ | Receive Correlation Matrix |
| $h$ | Channel Fading |
| $\mathbf{H}$ | Channel Fading Matrix |
| $\mathbf{H}_v$ | Virtual Fading Matrix |
| $\mathbf{I}$ | Identity Matrix |
| $\mathcal{I}$ | Mutual Information |

| | |
|---|---|
| $j$ | Imaginary Unit |
| $N_0$ | White Noise Variance |
| $r$ | Multiplexing Gain |
| $\mathbf{U}$ | Eigen Vector Matrix |
| $\vec{V}$ | Array Response Vector |
| $\vec{w}$ | White Noise Vector |
| $\vec{W}$ | Beamforming Weighting Vector |
| $\vec{x}$ | Channel Input |
| $\vec{y}$ | Channel Output |
| | |
| $\eta$ | Signal-to-Noise Ratio |
| $\tau$ | Bit (or Chip) Duration |
| $\sigma$ | Standard Deviation |
| $\sigma^2$ | Variance |
| $\rho$ | Correlation Coefficient |
| $\zeta$ | Euler-Mascheroni Constant |
| $\mathbf{\Lambda}$ | Eigen Value Matrix |
| $\mathbf{\Omega}$ | Virtual Variance Matrix |
| | |
| $\otimes$ | Kronecker Product |
| Æ | Expectation |
| $\vec{(\cdot)}$ | Vectorization |

| | |
|---|---|
| $(\cdot)^T$ | Transpose |
| $(\cdot)^\dagger$ | Conjugate Transpose |
| $(\cdot)^*$ | Element-wise Conjuate |

# List of Acronyms and Abbreviations

| | |
|---|---|
| AWGN | Additive White Gaussian Noise |
| ARMA | Auto-Regressive Moving-Average |
| BF | Beamforming |
| BPSK | Binary Phase Shift Keying |
| CDF | Cumulative Distribution Function |
| CDI | Channel Distribution Information |
| CDMA | Code Division Multiple Access |
| CSI | Channel State Information |
| DBS | Dynamic Base Station |
| DFT | Discrete Fourier Transform |
| DOS | Direction-of-Arrival |
| EVD | Eigen Value Decomposition |
| FDD | Frequency Division Duplex |

| | |
|---|---|
| FIR | Finite Impulse Response |
| GPS | Global Positioning System |
| HPBW | Half Power Beam Width |
| IBI | Inter Base Station Interference |
| IDFT | Inverse Discrete Fourier Transform |
| IDR | ID Repetition Signal |
| IIR | Infinite Impulse Response |
| IRT | ID Repetition Time |
| ISI | Inter-Symbol-Interference |
| IXI | Inter Transponder Interference |
| KPF | Kronecker Product Form |
| LCMV | Linear Constrained Minimum Variance |
| LHS | Left Hand Side |
| LOS | Line of Sight |
| LPS | Local Positioning System |
| MA | Multiple Access |
| MAI | Multiple Access Interference |
| MANET | Mobile Ad-hoc Networks |
| MIMO | Multi-Input-Multi-Output |

| | |
|---|---|
| MISO | Multi-Input-Single-Output |
| MMSE | Minimum Mean-Square-Error |
| MRC | Maximal Ratio Combining |
| MSE | Mean-Square-Error |
| NLOS | Non Line of Sight |
| OSTBC | Orthogonal Space-Time Coding |
| PDF | Probability Density Function |
| POS | Probability of Detection |
| PSD | Power Spectral Density |
| RCVR | Receiver |
| RHS | Right Hand Side |
| SDMA | Spatial Division Multiple Access |
| SIMO | Single-Input-Multi-Output |
| SISO | Single-Input-Single-Output |
| SNR | Signal-to-Noise Ratio |
| STC | Space-Time Coding |
| STBC | Space-Time Block Code |
| SVD | Singular Value Decomposition |

| | |
|---|---|
| TDD | Time Division Duplex |
| TOA | Time-of-Arrival |
| TRX | Transponder |
| UIU | Unitary-Independent-Unitary |
| ULA | Uniform Linear Array |
| VCR | Virtual Channel Representation |
| WLPS | Wireless Local Positioning System |

# Chapter 1

# Introduction

This dissertation studies wireless communications in *correlated* fading channels. Many realistic wireless channels, such as single-input-single-output (SISO) channels with moderate Doppler spread, and multi-input-multi-output (MIMO) channels in poor scattering environments, are categorized as correlated channels. In these environments, a joint data rate - error rate analysis is conducted. It is demonstrated that there is an optimal tradeoff between data rate and error rate. This tradeoff can be extended to other environments in a similar manner. This study is useful for investigating the capacity limits of mobile adhoc networks, e.g., wireless local positioning system under development at Michigan Tech [Tong and Zekavat(2007)].

## 1.1 Motivation

In general, a communication system consists of: 1) a transmitter, 2) a receiver and 3) a channel that maintains transmitter-receiver connection [Proakis(2000)]. In wireless communications, the channel gains are random, because the transmitted electromagnetic wave propagates through media in a random manner. Historically, the random channel gains are called fadings [Rappaport(2002)].

The simplest fading channel, slow frequency flat SISO channel, has been studied for over thirty years [Jakes(1974)]. In such channels, the transmit antenna sends

information symbols to the receive antenna in a sequential way. Assuming fadings vary slowly, a (relatively) constant fading might be experienced over one or multiple symbol durations. Accordingly, one way to model temporal fadings is the block fading model [Marzetta and Hochwald(1999)], i.e., fadings remain unchanged during a certain number of symbol durations, namely a block, and then they become completely independent in the next block. Longer block corresponds to smaller temporal correlation. Although simple, this model omits certain essential features of fading processes, such as wide sense stationarity. In our work, we view the SISO fading process as a stochastic process, and use its power spectral density (PSD) to describe the correlated fadings.

In addition to SISO (single antenna) channels, if multiple antennas are employed at the receiver side, the system would experience single-input-multi-output (SIMO) channels. Similarly, multi-input-single-output (MISO) and MIMO channels can be defined. In such channels, information symbols can be send not only in sequential, but also in parallel way. In the past decade, multi-antenna systems developed in an astounding speed.

At the beginning stage, multiple antennas are mainly exploited to achieve transmit (or receive) diversity [Alamouti(1999)] [Tarokh et al.(1998)], i.e., error rate performance can be improved. With the development of MIMO capacity analysis , it is gradually realized that multiple antennas may also improve data rate performance, namely multiplexing gain [Telatar(1999)] [Foschini and Gans(1998)]. Both of the two advantages of multi-antenna systems have been commercialized [Golden et al.(1999)] [Behbahani et al.(2000)].

However, recently a benchmark work found that diversity gain and multiplexing gain can not be optimized simultaneously [Zheng and Tse(2003)]. In other words, there is a tradeoff between diversity gain and multiplexing gain. The original works

propose the tradeoff assuming spatially independent fadings, i.e., fadings across different antenna pairs are independent.

Although it is neat and relatively easy to analyze independent fadings, experimental results reveal that spatial fadings are practically correlated [Chizhik et al.(2002)]. A number of models have been proposed for correlated fadings. Examples are [Kermoal et al.(2001)] [Sayeed(2002)] [Weichselberger et al.(2006)]. In our work, we introduce a unified framework that accommodates all of the above channel models. Moreover, it is found that different models are suitable only in certain physical environments. This well explains why each model is supported by different experiments.

With a solid understanding of correlated channel models, the diversity-multiplexing tradeoff are studied for correlated fading channels. Time varying SISO (temporally correlated) and time-invariant MIMO (spatially correlated) channels are examined. Finally time-varying MIMO (both temporally and spatially correlated) channels are preliminarily studied.

## 1.2 Outline

In general terms, this dissertation examines the performance of wireless communication systems under certain practical considerations. The outline of each of the chapter is as follows.

**Chapter 1**, the present chapter, gives the motivation, outline, and contributions of this dissertation.

**Chapter 2** studies the joint data rate - error rate performance in SISO channels in high SNR regimes. The concepts of diversity gain and multiplexing gain are first introduced and analyzed. The diversity-multiplexing tradeoff is then interpreted through a novel spike sharpening effect. Finally, temporal fading correlations are considered using the view of spike sharpening effect.

**Chapter 3** starts incorporating multi-antennas into our system. In this chapter, multi-antennas are employed only at the receiver side (SIMO). With fully correlated fadings, conventional Fourier beamforming techniques are dealt under array perturbations. The perturbations are caused by the imperfect knowledge of array manifolds in real applications. Theoretical analysis matches simulation results very well.

**Chapter 4** extends conventional beamforming to statistically optimal beamforming techniques (in SIMO channels). Particularly the cyclostationarity property of periodical signals is exploited to counter the non-stationarity problems in certain practical environments. This work helps the covariance matrix estimation for optimal beamformers. It should be noted that although the works of Chapter 3 and 4 are conducted in the background of wireless local positioning systems, their concepts and contributions are general: they can be applied into general communication systems as well.

With a fine understanding of SIMO systems, **Chapter 5** finally considers the case where multiple antennas are employed at both transmitter and receiver side (MIMO). Although correlated SIMO channel modeling is trivial, correlated MIMO channel modeling turns out to be much harder, due to its multi-dimension nature. In this chapter, a unified framework that accommodates previous correlated MIMO channel models is introduced. Then, through an eigen analysis of the unified framework, it is shown that each of the previous models are applicable only in certain physical environments. Finally, a preliminary error rate analysis for orthogonal space-time block codes is conducted.

**Chapter 6** explores the diversity-multiplexing tradeoff in correlated MIMO channels. Defining outage probability as the probability that the random channel capacity is smaller than the transmission rate, the tradeoff can be demonstrated through computing the outage probability. In this work, it is found that the outage

probability in MIMO systems with asymptotically high signal-to-noise ratio can be decomposed into the production of multiple conditional outage probabilities. Each conditional outage probability corresponds to a parallel SISO system. Hence, the tradeoff in MIMO system can be decomposed to a summation of the tradeoffs in multiple parallel SISO systems. This outage decomposition approach is applicable in both correlated and independent MIMO channels. The tradeoff with a small number of antennas are explicitly derived.

**Chapter 7** concludes the dissertation and discusses several broader visions.

The organization of this dissertation is summarized in Fig. 1.1.

### 1.3 Research Contributions

The main contribution of this dissertation is the joint error rate - data rate analysis under practical considerations, for example, poor scattering environments. Details of the research contributions in each chapter are as follows.

**Chapter 2** analyzes diversity-multiplexing tradeoff in SISO channels through spike sharpening effect. The mathematical derivations are introduced in a submitted journal paper, while the intuitions are depicted in a submitted conference paper.

- H. Tong and S. A. Zekavat, "On the Fundamental Tradeoff between Diversity Gain and Multiplexing Gain in SISO Channels with Memory," *IEEE Transactions on Information Theory*, in review.

- H. Tong and S. A. Zekavat, "A Novel Perspective on the Diversity-Multiplexing Tradeoff through Spike Sharpening Effect: Multiplicative Noise in Time Varying SISO Fading Channels with High SNR," *IEEE WCNC 2008*, submitted to.

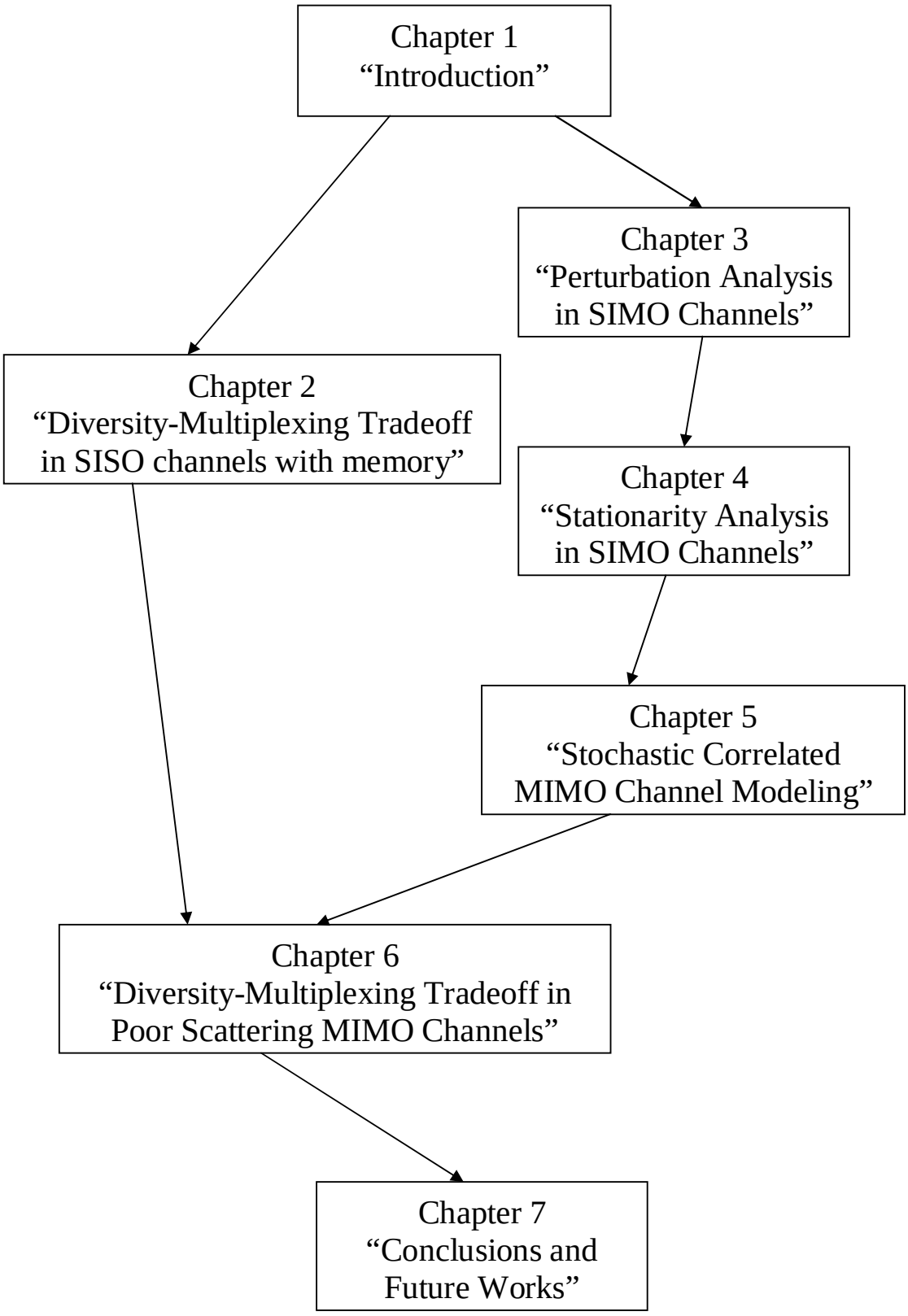


Figure 1.1: Organization of the Dissertation.

**Chapter 3** studies fully correlated SIMO channels. Under array perturbations, the error performance is theoretically derived for conventional Fourier beamformers. The discussion in this chapter was presented in one published journal paper and one published conference paper.

- H. Tong and S. A. Zekavat, "A Novel Wireless Local Positioning System via a Merger of DS-CDMA and Beamforming: Probability-of-Detection Performance Analysis under Array Perturbations," *IEEE Transactions on Vehicular Technology*, no. 3, vol. 56, May 2007, pp. 1307-1320.

- H. Tong and S. A. Zekavat, "Wireless Local Positioning System via DS-CDMA and Beamforming: A Perturbation Analysis," *IEEE WCNC 2005*, New Orleans, LA, Apr. 2005.

**Chapter 4** exploits cyclostationarity to counter non-stationarity problems in covariance matrix estimation for optimal beamformer in SIMO channels. The results are illustrated in one accepted journal paper and one published conference paper.

- H. Tong, J. Pourrostam, and S. A. Zekavat, "Optimum Beamforming for a Novel Wireless Local Positioning System: A Stationarity Analysis and Solution," *EURASIP Journal on Advances in Signal Processing*, to appear in.

- H. Tong and S. A. Zekavat, "Wireless Local Positioning System via DS-CDMA and Beamforming: A Stationarity Analysis," *SPIE 2005*, Orlando, FL, Mar. 2005.

**Chapter 5** reviews existing correlated channel models, organizes their relationships, and maps their eigen structure to specific physical environments. A preliminary error performance analysis is conducted. The results in this chapter are scattered in two journal papers (one published, one in review) and three conference papers (two published/accepted, one in review).

- H. Tong and S. A. Zekavat, "Spatially Correlated MIMO Channel: Generation via Virtual Channel Representation," no. 5, vol. 10, *IEEE Communications Letters*, May. 2006.

- H. Tong and S. A. Zekavat, "Stochastic MIMO Channel Modeling: What Environments Can Be Represented by the Kronecker Product Form?," *IEEE Transactions on Wireless Communications*, in review.

- H. Tong and S. A. Zekavat, "Capacity and Error Performance of Correlated MIMO Systems Via Virtual Channel Representation," *IEEE PIMRC 2007*, Athens, Greece, Sep. 2007, to appear in.

- H. Tong and S. A. Zekavat, "A Simple Beamforming-SIMO Merger in Spatially Correlated Channel via Virtual Channel Representation," *IEEE GLOBECOM 2005*, St. Louis, MO, Nov, 2005.

- H. Tong and S. A. Zekavat, "On the Suitable Environments of the Kronecker Product Form in MIMO Channel Modeling," *IEEE WCNC 2007*, submitted to.

**Chapter 6** demonstrates an approach for computing the diversity-multiplexing tradeoff in correlated MIMO channels. The proposed outage decomposition approach converts MIMO channels to a combination of multiple parallel SISO channels; hence, facilitates the tradeoff computation. The following journal paper is currently in progress.

- H. Tong and S. A. Zekavat, "The Diversity-Multiplexing Tradeoff in Poor Scattering 2x2 MIMO Channels through an Outage Decomposition Approach," in progress.

I am also grateful for the collaborations in the following conference papers.

- S. A. Zekavat, H. Tong and J. Tan, "A Novel Wireless Local Positioning System for Airport (Indoor) Security," *SPIE 2004*, San Diego, CA, Feb. 2004.

- J. Pourrostam, S. A. Zekavat, and H. Tong, "A Novel DOA Estimation in Wireless Local Positioning Systems," *IEEE RADAR 2007*, Waltham, MA, Apr. 2007.

- W. Xu, S. A. Zekavat and H. Tong, "A Novel Approach for Spatially Correlated Multi-User MIMO Channel Modeling: Impact of Surface Roughness and Directional Scattering", *Allerton 2007*, to appear in.

# Chapter 2

# The Diversity-Multiplexing Tradeoff in SISO Channels with Memory

This chapter investigates the tradeoff between multiplexing gain and diversity gain in a Single-Input-Single-Output (SISO) system with *temporally correlated* fadings. More explicitly, we consider an average power constrained, single-user, SISO system in time varying, frequency flat, Rayleigh fading channels, where channel state information is perfectly known at the receiver, but it is not known at the transmitter. Here, the codeword length may span over a time much greater than the channel coherence time.

The new view of Power Spectral Density (PSD), which is a sufficient description of the statistics of fading process, is adopted to describe correlated fadings. Categorization of fading processes based on their PSDs, and the corresponding fading generation methods are discussed in detail. This categorization can be considered as the generalization of Telatar's concepts about ergodic channels.

Multiplexing gain and diversity gain serve as the critical measures of transmission rate and error rate in high signal-to-noise ratio (SNR) regimes, respectively. With the assumption of flat PSD, i.e., independent fadings, prior research has shown

that multiplexing gain and diversity gain can not be optimized simultaneously: the tradeoff between them resembles a simple linear relationship.

In practical systems, the PSD of a fading process is not necessarily flat, but may exhibit any shape with a positive Lebesgue measure. It is demonstrated that the tradeoff between multiplexing gain and diversity gain is scaled by the "spreadness" of the PSD, which is essentially the second order statistics of the instantaneous mutual information in the corresponding fading channel. Using the "spreadness", it is finally shown that although both slow and fast varying channels yield the same channel capacity; in fast varying channels, it is easier to approach such a capacity, i.e., fast varying channels yield better diversity gain assuming the same multiplexing gain.

## 2.1 Introduction

The study of channel capacity has been the interest of communication researchers for over fifty years [Verdu(1998)]. The original Shannon's capacity only deals with Additive White Gaussian Noise (AWGN) channels [Shannon(1948)]. In recent decades, channel capacity in *fading* channels has drawn significant attention due to the rapid development of wireless communication systems, in which we need to deal with fading channels but not AWGN channels. The fundamental difference between AWGN channels and fading channels is the nature of channel gain. In AWGN channels, channel gain is deterministic and time invariant. In fading channels, channel gain (fading) is *random* and (in most cases) time-varying.

In this chapter, we consider an average-power constrained, single-input-single-output (SISO), single user system in time varying, frequency flat, Rayleigh fading channels, where channel state information (CSI) is available at receiver, but not known at transmitter. This is a rather standard wireless communication system configuration. Channel capacity of these systems has been well understood

[Goldsmith and Varaiya(1997)] [Ozarow et al.(1994)]. However, just knowing channel capacity is often not sufficient in practical systems: It is important to know the rate that we can approach to such a capacity.

What is channel capacity? As it is well known, error probability vanishes when code word length approaches to infinity, while a definitely positive transmission rate still can be maintained. The supremum of the transmission rate is channel capacity. In a real system, we do not have an infinitely long code word. Hence, it is also important to study how error probability decays with code word length: If error probability decays quickly with codeword length, it is easy to approach channel capacity.

In general, it is hard to obtain an exact characterization of the decay rate of error probability [Feinstein(1955)]. A famous approach is Gallager's error exponent, which serves as an upper bound of the error probability [Gallager(1968)]. More recently, Zheng and Tse [Zheng and Tse(2003)] point out that the outage probability is the lower bound of the error probability, through Fano's inequality [Cover and Thomas(1991)]. Moreover, they show that, in terms of diversity gain, outage probability is a tight bound for error probability in high SNR regimes.

In this work, we are particularly interested in the error probability analysis in high signal-to-noise ratio (SNR) regimes. In other words, we are interested in the tradeoff between multiplexing gain and diversity gain, which serves as the critical measure of transmission rate and error rate, respectively. First, multiplexing gain is defined as [Zheng and Tse(2003)]:

$$r = \lim_{\eta \to \infty} \frac{R(\eta)}{\log \eta}. \tag{2.1}$$

A transmission strategy would possess a multiplexing gain $r$, if the ratio of its transmission rate $R(\eta)$ and the logarithm of SNR $\eta$ is kept constant in high SNR regimes.

On the other hand, diversity gain is defined as [Zheng and Tse(2003)]:

$$d(r) = -\lim_{N\to\infty}\lim_{\eta\to\infty}\frac{P_e(N,\eta,r)}{N\cdot\log\eta} \tag{2.2}$$

This definition already incorporates the fact that the error probability $P_e$ decays exponentially with both the codeword length $N$ and the logarithm of SNR $\eta$. Moreover, it also shows that diversity gain is a function of multiplexing gain, i.e., error performance is a function of transmission rate.

Intuitively, one would expect that higher transmission rate leads to worse error performance: actually it is true. For *independent* channel fadings, in high SNR regimes with long codeword, the tradeoff between multiplexing gain and diversity gain corresponds to [Zheng and Tse(2003)]:

$$d(r) = 1 - r, \tag{2.3}$$

i.e., roughly speaking, increasing transmission rate would lead to higher error probability.

Please note that in general, Doppler frequency shift is much smaller than the signal bandwidth. Therefore, in most real systems, fadings are not independent but correlated: slowly varying channels creates correlated fadings. One previous approach to model the speed of channel variation is the block fading model: Fadings are assumed to be constant for a certain block of symbols, and then become completely independent for the next block of symbols. Of course this is a very rough approach.

Assuming fading processes are circularly symmetric wide sense stationary ergodic complex Gaussian processes, the new view of power spectral density (PSD) is incorporated in this work to describe the fading processes. Independent fadings exhibit a flat PSD, while correlated fadings exhibit other shapes of PSD. This work shows

that, as long as the PSD possesses a definitely positive Lebesgue measure, the tradeoff between multiplexing gain and diversity gain with *correlated* fadings corresponds to:

$$d(r) = \mathsf{s} \cdot (1 - \mathsf{r}) \tag{2.4}$$

where $\mathsf{s}$ is called the "spreadness" of the PSD, and corresponds to:

$$\mathsf{s} = \left( \frac{6}{\pi^2} \cdot \sum_{\mathsf{k}=1}^{\infty} \int_{-0.5}^{0.5} \frac{\mathsf{S}_\mathsf{k}^2(\mathsf{f})}{\mathsf{k}^2} \, \mathsf{df} \right)^{-1}. \tag{2.5}$$

In (2.5), $S(f)$ denotes the PSD of fading process, and $S_k(f) = S(f) * S_{k-1}(f)$, where $*$ denotes cyclic convolution. For flat PSD, $\mathsf{s} = 1$. For other shapes of PSD, $\mathsf{s} < 1$.

Extensions of the tradeoff between multiplexing gain and diversity gain have been investigated recently, such as [Yang and Belfiore(2006)] and [Gamal et al.(2006)]. However, few of them is dealing with correlated fadings. Particularly, none of them adopts the view of spectral analysis. Other relevant discussions, such as [Lapidoth and Moser(2003)] [Lapidoth(2005)] and [Etkin and Tse(2006)], that adopt the view of linear prediction (which is very close to the view of spectral analysis), only deal with channel capacity but not error performance. Hence, this chapter contributes to the joint data rate and error rate analysis of wireless communications, in the view of spectral analysis.

The result proposed in this work can be easily extended into multi-input-multi-output (MIMO) cases on the concept level. However, it is still hard to obtain an explicit mathematical expression in MIMO systems with correlated fadings: Kronecker product in place of matrix multiplication, and, eigen decomposition in place of diagonalization lead to very high mathematical difficulties. It should be emphasized that, once those mathematical difficulties in MIMO systems are solved, viewing fading process through its PSD inherently leads to the immediate application of many well-established spectral-based signal processing algorithms, e.g., cepstrum, 2D Capon,

discrete prolate spheroidal sequence, in the newly proposed 4G and above wireless communication systems that are most likely based on advanced MIMO techniques. Hence, this work may have a profound impact in the future wireless system design.

Up to now, the main results have been introduced. The rest of this chapter is organized as follow: In Section 2.2, we discuss fading channel models; In Section 2.3, we show the derivation of the tradeoff; Section 2.4 introduces the intuitions behind the tradeoff, and provides numerical examples; Finally Section 2.5 concludes this chapter. Throughout the whole chapter, we try to keep just a minimum amount of math in the text, and refer all the details of derivations to appendices.

## 2.2 Channel Model

The study of temporal channel modeling started in the age of Einstein [Cohen(2005)], and surprisingly in the twenty first century, new results are still being generated [Baddour and Beaulieu(2005)]. In this work, we particularly consider the modeling of time-varying, frequency flat, Rayleigh fading SISO channels that are possibly with memory, i.e., fadings are possibly correlated. Exploiting the discrete-time, base band channel expression [Tse and Viswanath(2005)], this fading process can always be considered as a wide sense stationary ergodic stochastic complex Gaussian process [Goldsmith(2005a)]. The most convenient way to describe such a process is its Power Spectral Density (PSD) [Papoulis and Pillai(2002)]. Hence, our channel model analysis is completely based on the analysis of PSD. Assuming channel estimation process generates a fading coefficient for every symbol, the sampling frequency of channel fadings would be equal to the signal bandwidth. In this chapter, all PSDs are scaled by this sampling frequency.

In this section, we first categorize fading processes: regular processes, bandlimited processes, and periodical processes. Then, we compare different fading generation

techniques and their corresponding suitable fading models. Finally, we show that block fading model, different from its original purpose, is not an adequate approximation of bandlimited processes: It turns out to be a rough approach compared to the PSD-based fading generation techniques.

### 2.2.1 Categorizing Temporal Channel Models

Telatar's seminar work [Telatar(1999)] introduces two extreme cases of temporal channels: 1) *Extremely fast time varying channels*: In this case channel fadings would be independent over each symbol. Assuming wide sense stationarity, the fadings would be independent and identically distributed (i.i.d.). In this work, the associated channels are called i.i.d. channels, while in Telatar's work they are called ergodic channels; 2) *Extremely slow time varying channels*: In this case, channel fadings are generated randomly, but kept invariant during the whole period of observation. In this work, the associated channels are called time-invariant channels, while in Telatar's work they are called non-ergodic channels.

Real channels comprise not only extremely fast and extremely slow cases, but also the cases between these two extreme cases. Here, we propose a *general* categorization of fading processes based on their PSDs, as shown in Fig. 2.1. In the order of speed of channel variation, five types of fading processes are introduced: i.i.d. fading, regular process, bandlimited process, periodical process, and time-invariant process.

Here, two channel categorization approaches are discussed: 1) spectral analysis, and 2) optimum linear prediction. It is shown that those approaches are interrelated: the approach of optimum linear prediction is based on the approach of spectral analysis. Explaining the two approaches and their relationship, we shall have an intuition why this categorization leads to an intrinsic sense of the speed of channel variation.

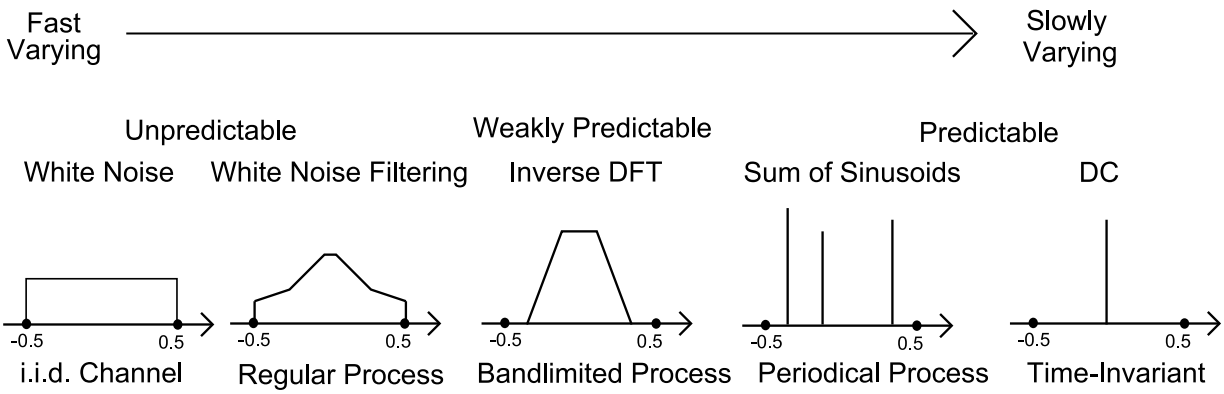


Figure 2.1: Fading Processes Categorization

**Spectral Analysis**

i.i.d. channels are special cases of regular processes, as fadings in regular processes are filtered white noise. The filters associated to regular processes are combinations of finite impulse response (FIR) and/or infinite impulse response (IIR) filters, with a finite number of taps, expressed via an auto regressive moving average (ARMA) model. The PSD and the corresponding filter coefficients are related by spectral factorization [Papoulis and Pillai(2002)].

PSDs of regular processes span over the whole frequency axis, from the $-0.5$ sampling frequency to the 0.5 sampling frequency. Here, the sampling frequency of fading processes is in fact the signal bandwidth. In regular processes, the number of zeros of the PSD is finite, as they represent a finite order filter. Hence, one property of regular processes corresponds to:

$$u(\{f : S(f) = 0\}) = 0. \tag{2.6}$$

where $u(\cdot)$ denotes the Lebesgue measure on the interval $[-0.5, 0.5]$, and $S(f)$ denotes the PSD of a fading process.

On the other hand, time-invariant channels are special cases of periodical processes, as shown in Fig. 2.1. When there is only one sinusoid with $f = 0$, the

periodical process becomes time-invariant process. Note that periodical processes are not ergodic: their fading statistics can not be obtained via the observed fading time series. Hence, it is actually not very adequate to use PSD to describe those processes.

However, we can extend the spectral representation of deterministic signals to represent periodical processes. They can be considered as the summation of multiple sinusoids: the amplitude of each sinusoid is chosen randomly at the beginning, and then held constant for the whole transmission period (possibly infinitely long). A famous example of the periodical processes is the Jake's model [Jakes(1974)], where evenly spaced frequencies are chosen. Assuming the number of sinusoids is finite, the property of periodical property corresponds to:

$$u(\{f : S(f) = 0\}) = 1. \tag{2.7}$$

In addition to regular and periodical processes, the PSD of bandlimited fading processes correspond to:

$$u(\{f : S(f) = 0\}) = u, \quad u \in (0, 1). \tag{2.8}$$

Noting that the PSD of a fading process is essentially its Doppler spectrum and in most real systems the maximum Doppler frequency shift is much smaller than signal bandwidth, fading processes would be practically bandlimited.

It can be seen that regular processes, bandlimited processes and periodical processes are sorted in the order of their Lebesgue measure, while i.i.d channels and time-invariant channels are special cases of regular processes and periodical processes, respectively. Therefore, a generalization of Telatar's ergodic and non-ergodic channel models has been made here.

Intuitively, more spread the PSD has, faster the channel varies. However, a consistent measurement of the "spreadness" of PSD has not been presented in the current literature. In Section 2.3, the explicit mathematical expression of the "spreadness" of PSD is introduced.

**Optimum Linear Prediction**

Another approach for categorizing fading processes is based on optimum linear prediction error [Etkin and Tse(2006)].

*Question*: Can we accurately predict the next future value of a stochastic process, if an infinite number of past value is known? This question has already been answered by mathematicians, and was summarized by Papoulis in [Papoulis(1985)]. Note that the proof in [Papoulis(1985)] is based on spectral analysis, which is consistent with the discussion in Section 2.2.1. Roughly speaking, the proof exploits the fact that FIR filter for bandlimited processes must possess infinite order, while FIR filters for regular processes possesses a finite order.

The optimum linear prediction error corresponds to:

$$\epsilon_{pred} = \arg \min_{a_1,a_2,\cdots} Æ \left[ \left( h_0 - \sum_{k=1}^{T} a_k h_{-k} \right)^2 \right] \tag{2.9}$$

where Æ denotes expectation, $h_{-T}, \cdots, h_{-2}, h_{-1}$ denotes the known fading values, $h_0$ denotes the currently unknown fading value, and $a_1, a_2, \cdots, a_T$ represents the linear prediction filter.

For regular processes, a rough answer[1] is [Papoulis(1985)]:

$$\epsilon_{pred} > 0, \quad \text{for } T \to \infty. \tag{2.10}$$

This corresponds to the fact that, even given an infinitely long past, the next fading can not be predicted accurately.

For bandlimited processes [Papoulis(1985)]:

$$\epsilon_{pred} \to 0, \quad \text{for } T \to \infty. \tag{2.11}$$

[1]We do not discuss the fine answer, i.e., the Paley-Wiener condition, since the goal of this chapter is not to study the exact optimal linear prediction error.

So, in this case, the prediction error decreases as the length of known past increases, but it is never exactly zero with finite $N$. This is the reason why bandlimited processes are also called "weakly predictable processes" [Papoulis(1985)].

For periodical processes, we have:

$$\epsilon_{pred} = 0, \quad \text{for finite}\, T \tag{2.12}$$

So given just a finite number of known values, the next value of a periodical process can be estimated accurately. Actually ,this is easy to understand, since the fading values are periodical in this case.

What does prediction error mean? Larger prediction error corresponds to a harder capability of predicting the future value through its past. This corresponds to a low correlation between future and past values. Accordingly, comparing (2.10) to (2.12), we recognize that regular, bandlimited, and periodical processes are sorted in the order of speed of fading correlations.

### 2.2.2 Fading Generation Algorithms

In addition to categorizing fading processes, the generation techniques corresponding to different categories are different as well.

#### White Noise Filtering and Regular Processes

Regular processes are mainly modeled by spectral factorization: its spectral can always be decomposed to certain polynomials [Saeks(1976)], i.e., poles and zeros structure. Then, a finite order FIR/IIR combined filter can be implemented via pole-zero structure. A regular process is the output of such a minimum phase filter with white noise as its input.

**Sum of Sinusoids and Periodical Processes**

As we have shown, the spectral of periodical processes contains a finite number of lines, i.e., line spectral. It is easily understood that such a process can be generated by sum of sinusoids. A famous example of such a model is the Jake's model.

**Inverse DFT and Bandlimited Processes**

Bandlimited processes can be generated either by a filter with infinite number of zeros, or by an infinite number of sinusoids. But, apparently both of the two methods are not adequate enough for simulation purposes. Bandlimited fading generation has been well studied in a series works conducted by Beaulieu [Beaulieu and Tan(1997)] [Young and Beaulieu(2000)], which corresponds to:

$$\vec{h} \overset{\text{DFT}}{\longleftrightarrow} (s^{0.5} \odot \vec{h}_{iid}). \tag{2.13}$$

Here, $s^{0.5}$ denotes the square root of the PSD, $h_{iid}$ is a vector consisting of i.i.d. complex random variables, and, $\odot$ represents element-wise multiplication. Applying DFT to the right hand side (RHS) of (2.13), the temporal fading process is achieved.

The IDFT generation is a natural expression of periodogram, as described in [Stoica and Moses(1997)]. It works not only with bandlimited fadings, but also with regular and periodical processes. It is particularly applicable to the generation of a large number of channel fadings, in which the aliasing effect is kept minimum. In our work, we consider the cases that the number of fadings approaches infinity. Hence, the inverse DFT method would be our preferred fading generation technique.

### 2.2.3 Block Fading Assumption is Not Adequate

Currently, block fading model is the most widely used temporal model for time varying channels due to its simplicity in simulations. The block fading model assumes that fadings are constant in a block, e.g., $L$ transmitted symbols, while fadings in

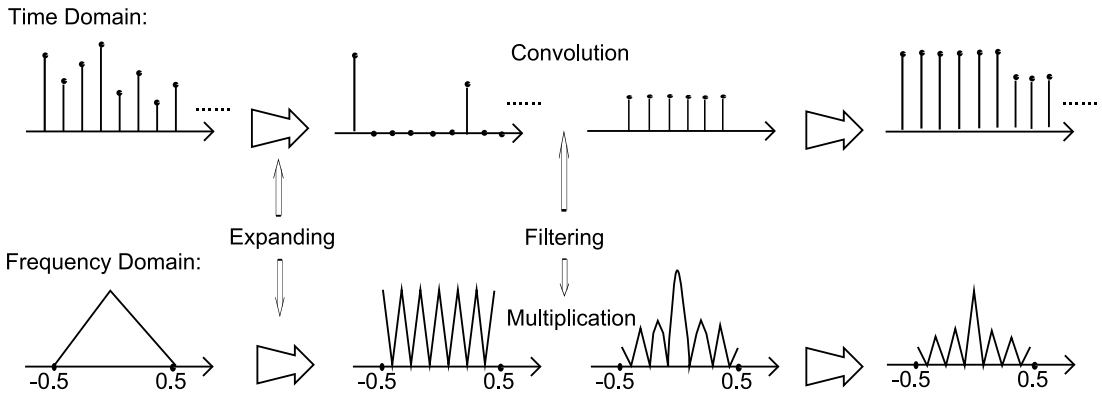


Figure 2.2: Understanding the block fading model in the view of spectral analysis

different blocks are completely independent. Approximately, $L$ is the ratio of channel coherence time and symbol duration, or the ratio of signal bandwidth and maximum doppler frequency.

If $L = 1$, fading process would be a i.i.d. process, i.e., extremely fast varying channels. If $L \to \infty$, fading process shall be a time-invariant process, i.e., extremely slowly varying channels. $L = 2, 3, \cdots$ indicates cases between extremely fast and extremely slowly varying channels. Therefore, block fading models can be considered as another generalization of Telatar's ergodic channel concept. Here, we address that PSD-based generalization is finer than block fading model.

As shown in Fig. 2.2, block fading generation essentially is time domain interpolation. First an i.i.d. fading process is generated. Then, this i.i.d. process is expanded, i.e., $L - 1$ zeros are inserted between each i.i.d fadings. Finally, the expanded i.i.d process convolves with a filter whose $L$ coefficients are the same, and the output would be block fadings.

The frequency domain expression of block fading generation is shown in Fig. 2.2 as well. In frequency domain, the original spectrum is scaled after expanding. Because the frequency response of the all one filter with $L$ taps is a sinc function, the spectrum of block fadings would be the multiplication of the scaled original spectrum

and the sinc function formed by the all one filter. Hence, the spectrum of fadings generated by block fading method includes both main lobe and side lobes.

Ideally, the spectrum of generated fadings should include only the main lobe, while in this case it includes sidelobes. For block fading model, side lobes, which are artificially introduced by the sinc function with length $L$, would not decay when the number of generated fading increases. Levels of side lobe are solely determined by $L$, and they are independent of the number of generated fadings.

In contrast, applying PSD and IDFT method, the generated spectrum would be the convolution of the desired spectrum with the frequency response of an all one filter with length $T$, where $T$ denotes the number of total generated fading [Stoica and Moses(1997)]. When $T \rightarrow \infty$, the spectrum of the all one filter with length $T$ would approach a Dirac delta function; consequently, the generated spectrum would approach to the desired spectrum. However, for block fading model, when $T \rightarrow \infty$, the generated spectrum is still severely distorted. In this work, we are interested in generating a large number of fadings; hence, we adopt the IDFT method to generate correlated fadings.

## 2.3 The Tradeoff between Multiplexing Gain and Diversity Gain

The tradeoff is mainly shown through computing the first and the second order statistics of mutual information. We first define the mutual information, then we discuss the relationship between various concepts and statistics of mutual information: 1) Channel capacity is mainly related to the first order statistics of mutual information, and 2) the tradeoff between diversity gain and multiplexing gain is determined by the second order statistics.

The results shows that the first order statistics does not vary with the PSD of fadings, while the second order statistics varies with the PSD. Hence, although both fast and slowly varying channels yield the same channel capacity, the diversity gain in fast varying channels is higher, assuming the same multiplexing gain.

### 2.3.1 Mutual Information in SISO Fading Channels

Before defining the mutual information and capacity, we first introduce the signal model of frequency flat SISO channels. The frequency flat SISO channel input-output relationship corresponds to:

$$y[n] = h[n] \cdot x[n] + w[n], \quad n \in \{\cdots, -1, 0, 1, 2, \cdots\} \tag{2.14}$$

where $x[n]$ and $y[n]$ denote the transmitted and received symbols, respectively, and $h[n]$ and $w[n]$ denote channel fadings and receiver noise, respectively. It is assumed that $w[n]$ is an i.i.d. complex Gaussian process with variance $\sigma_w^2$, and $h[n]$ is a circularly symmetric complex Gaussian process, which can be regular, bandlimited, or periodical. Representing (2.14) in the vector form:

$$\vec{y} = \mathbf{H} \cdot \vec{x} + \vec{w} \tag{2.15}$$

where $\vec{y}$, $\vec{x}$ and $\vec{w}$ are column vectors, $\mathbf{H}$ denotes a diagonal matrix, whose diagonal elements are channel fadings, i.e., $\mathbf{H} = \text{diag}(\vec{h})$.

Without loss of optimality, we set the input sequence, $\vec{x}$, circularly symmetric complex Gaussian random variables (see details about this assumption in [Telatar(1999)]). Then, the output $\vec{y}$ would comprise circularly symmetric complex Gaussian random variables, and its covariance matrix corresponds to:

$$\mathbf{Q}_{yy} = \text{Æ}\left[\vec{y}\vec{y}^{\dagger}\right] = \sigma_w^2 \cdot \left(\mathbf{I} + \eta \cdot \mathbf{H}\mathbf{Q}_{xx}\mathbf{H}\right), \tag{2.16}$$

where $\mathbf{Q}_{xx}$ represents the input covariance matrix, i.e., $\mathbf{Q}_{xx} = \text{Æ}\left[\vec{x}\vec{x}^{\dagger}\right]$. Note that the input must satisfy an average power constraint, i.e.,

$$\frac{\text{Tr}(\mathbf{Q}_{xx})}{N} \leq P_t, \tag{2.17}$$

where $N$ denotes the length of codeword. Moreover, the signal-to-noise ratio (SNR), $\eta$, is defined as:

$$\eta = \frac{P_t}{\sigma_w^2}. \tag{2.18}$$

Now, we can start deriving the mutual information in SISO fading channels. Assuming perfect CSI at the receiver, this channel has one input sequence $\vec{x}$, and two output sequences $\vec{y}$ and $\vec{h}$. Hence, in this channel, the mutual information between the input and the output corresponds to:

$$\begin{aligned} \mathcal{I}(\vec{x}; \vec{y}, \vec{h}) &= \mathcal{I}(\vec{x}; \vec{h}) + \mathcal{I}(\vec{x}; \vec{y}|\vec{h}) \\ &= \mathcal{I}(\vec{x}; \vec{y}|\vec{h}) \end{aligned} \tag{2.19}$$

The first equality follows the chain rule, and the second equality is due to the fact that CSI is not available at the transmitter, i.e., $\mathcal{I}(\vec{x}; \vec{h}) = 0$. Then, the mutual information in (2.19) corresponds to:

$$\begin{aligned} \mathcal{I}(\vec{x}; \vec{y}|\vec{h}) &= \mathcal{H}(\vec{y}|\vec{h}) - \mathcal{H}(\vec{y}|\vec{x}, \vec{h}) \\ &= \mathcal{H}(\vec{y}|\vec{h}) - \mathcal{H}(\vec{w}) \\ &= \log\det(\pi e \mathbf{Q}_{yy}|\vec{h}) - N \cdot \log(\pi e \sigma^2) \\ &= \log\det\left(\mathbf{I} + \eta \cdot \mathbf{H}\mathbf{Q}_{xx}\mathbf{H}\right) \end{aligned} \tag{2.20}$$

The first equality follows the definition of mutual information; the second equality follows input-output relationship in (2.15); the third exploits the differential entropy rate of Gaussian processes [Cover and Thomas(1991)], where $e$ denotes the Euler's constant; and the last equality is apparent as shown in (2.16).

Mathematically, (2.20) is a simplified version of Telatar's capacity equation: In our work $\mathbf{H}$ is restricted to be diagonal, while this restriction is not applied to Telatar's equation. Here, we discuss a different perspective of (2.20) from that of

Telatar. We say that the mutual information in fading channels, i.e., $\mathcal{I}(\vec{x};\vec{y},\vec{h})$ is *a random variable*, since it is a function of channel fadings. The concept of random mutual information is the key point to understand channel capacity in fading channels.

The next natural question is: What determines the statistical properties of the random mutual information $\mathcal{I}(\vec{x};\vec{y},\vec{h})$? We easily note that they are determined by 1) SNR, $\eta$, 2) fading statistics, $S(f)$, 3) code word length, $N$, and 4) input covariance matrix, $\mathbf{Q}_{xx}$. In this work, we are interested in investigating the effects of the fading statistics $S(f)$ on the first and second order statistics of mutual information.

To study the effects of $S(f)$, we assume an asymptotically high SNR environment with sufficiently large code word length. The assumption of high SNR denotes high transmission power, i.e., sufficient resources at the receiver side. The assumption of large code word length denotes sufficient decoding computation power, i.e., sufficient resources at the receiver side. Hence, the assumptions essentially remove the resource limitation at both the transmitter side and the receiver side.

Now, assuming high SNR and long codeword, the statistics of mutual information is only function of fading PSD $S(f)$, and input covariance matrix $\mathbf{Q}_{xx}$. In Section 2.3.2, it shown that in high SNR regimes, $\mathbf{Q}_{xx} = \mathbf{I}$ is optimal in terms of achieving channel capacity. This result significantly alleviates the mathematical difficulties in the second order statistics analysis.

Through the above three conditions: high SNR, long codeword, and identity input covariance matrix, the final relationship between fading PSD and the second order statistics of mutual information would be established. The relationship turns out to be the final result: the tradeoff between diversity gain and multiplexing gain.

### 2.3.2 The First Order Statistics of Mutual Information

The main goal of this subsection is to show that identity input covariance matrix is optimal in high SNR regimes, which facilitates the second order statistics analysis.

**Theorem 2.3.1.** *Given a fixed code word length $N$, and certain fading statistics $S(f)$, the expectation of the mutual information $\mathcal{I}(\vec{x};\vec{y},\vec{h})$ increases linearly with the logarithm of the SNR $\eta$. Moreover, the input covariance matrix that maximizes the expectation of the mutual information is an identity matrix, i.e., identity input covariance matrix achieves channel capacity in high SNR regimes.*

*In other words, the solution to the following optimization:*

$$\arg\max_{\mathbf{Q}_{xx}} \lim_{\eta\to\infty} \frac{\text{Æ}\left[\log\det\left(\mathbf{I}+\eta\cdot\mathbf{H}\mathbf{Q}_{xx}\mathbf{H}\right)\right]}{\log\eta} \quad s.t. \quad \frac{Tr(\mathbf{Q}_{xx})}{N}\leq P_t \tag{2.21}$$

*is* $\mathbf{Q}_{xx}=\mathbf{I}$.

*Proof.* See appendix 2.A. □

Based on this theorem, We set $\mathbf{Q}_{xx}=\mathbf{I}$ in the rest of this chapter. Since $\mathbf{H}$ is diagonal in SISO cases, the associated mutual information would be in a very simple form by setting $\mathbf{Q}_{xx}=\mathbf{I}$, i.e.,:

$$\log\det\left(\mathbf{I}+\eta\cdot\mathbf{H}\mathbf{H}\right)=\sum_{n=1}^{N}\log(1+\eta A_n), \tag{2.22}$$

where $A_n = h_n h_n^*$. Considering a Rayleigh fading channel, $A_n$ would follow an exponential distribution.

### 2.3.3 Concepts Related to the First Order Statistics: Capacity and Ergodic Capacity

Here, we mainly discuss the concepts that are relevant to the first order statistics of mutual information. The classical definition of channel capacity is the supremum of mutual information per unit time, with sufficiently large code word length, i.e.,

$$\mathcal{C}=\lim_{N\to\infty}\sup\frac{\mathcal{I}(\vec{x};\vec{y},\vec{h})}{N}. \tag{2.23}$$

Particularly, using (2.20), the channel capacity corresponds to:

$$\mathcal{C} = \lim_{N\to\infty} \sup_{\mathbf{Q_{xx}}} \frac{\log\det\left(\mathbf{I} + \eta \cdot \mathbf{H}\mathbf{Q}_{xx}\mathbf{H}\right)}{N}. \tag{2.24}$$

In the asymptotically high SNR regime, we have shown that $\mathbf{Q_{xx}} = \mathbf{I}$ is optimal. Accordingly, the channel capacity corresponds to:

$$\begin{aligned}
\mathcal{C} &= \lim_{N\to\infty} \frac{\sum_{n=1}^{N} \log(1+\eta\cdot A_n)}{N} \\
&= \text{Æ}\left[\log(1+\eta\cdot A_n)\right] &(2.25)\\
&= \int_0^\infty \log(1+\eta\cdot A) p_A(A)\, dA. &(2.26)
\end{aligned}$$

Here, $p_A(A)$ is the PDF of square of fading amplitudes, e.g., exponential distribution $p_A(A) = e^{-A}$ for Rayleigh fading channels. We note that the result in (2.26) is exactly the same as that of [Goldsmith and Varaiya(1997)].

An important point is that $\mathcal{C}$ would be independent of $S(f)$, if we assume $\mathbf{Q_{xx}} = \mathbf{I}$. This means that *slowly and fast varying channels yield the same channel capacity*, in high SNR regimes. This is a somewhat important conclusion which was conveyed implicity in [Biglieri et al.(1998)].

Moreover, the second equality of (2.25) is consistent with the definition of ergodic capacity, i.e., the expectation of mutual information with *finite* $N$ coincides with the classical definition of channel capacity, which is defined via an *infinite* $N$. This observation validates the properness of the concept of ergodic capacity.

Similar as Telatar's discussion, (2.25) only holds for regular and bandlimited processes. For periodical process, we can not replace expectation with limitation, because the effective number of samples of fadings is finite. Thus, for periodical processes, the conventional definition of channel capacity does not exist: (2.24) would not approach a certain limit even if $N$ is infinite.

Up to now, we have discussed the first order statistics of mutual information and the relevant concepts: the channel capacity and the ergodic capacity. The two parameters represents the ultimate performance of communication systems. However, in practical systems, just knowing the ultimate performance is not enough. We also need to know how hard it is to obtain such a performance. This is mainly determined by the second order statistics of mutual information and the relevant concepts: outage capacity, multiplexing gain and diversity gain.

### 2.3.4 Second Order Statistics of Mutual Information

We showed that $\mathbf{Q}_{xx} = \mathbf{I}$ optimizes mean of mutual information in high SNR regimes. Hence, throughout the analysis of the second order statistics, $\mathbf{Q}_{xx}$ is set to $\mathbf{I}$. As discussed in Section 2.3.1, the second order statistics is a function of fading PSD, $S(f)$, SNR, $\eta$, and codeword length, $N$.

In this section, we show that, assuming high SNR and long codeword, the effects of $S(f)$, $\eta$ and $N$ can be easily separated. Considering high SNR and long code word: 1) the variance of mutual information would be independent of SNR and decays linearly with $N$, and 2) the spreadness of $S(f)$ determines how fast the variance decays with $N$, that is, the slope of the linear relationship.

The first theorem introduces the relationship between SNR and the variance of mutual information. This variance approaches a limit which is independent of SNR, when SNR approaches infinity.

**Theorem 2.3.2.** *When SNR approaches infinity, the variance of mutual information approaches a limit that is independent of SNR, i.e.,*

$$\lim_{\eta \to \infty} var\left(\log\det\left(\mathbf{I} + \eta \cdot \mathbf{HH}\right)\right) = var\left(\log \mathbf{HH}\right) \tag{2.27}$$

.

*Proof.* The equation seems easy, but the derivation is quite cumbersome for correlated fadings. See the details in Appendix 2.B. □

Note that LHS of (2.27) is a function of $\eta$ while its RHS is not; hence, the variance of mutual information does not increase with SNR in high SNR regimes. Here, it is noted that the mean of mutual information increases linearly with the logarithm of SNR. Therefore, comparing (2.26) and (2.27), it is evident that the first and second order statistics of mutual information exhibits completely different behaviors as SNR increases.

The next step is to study the effect of codeword length on the variance of mutual information. Here, we introduce two lemmas before introducing the Theorem on the relationship between codeword length and the variance of mutual information.

**Lemma 2.3.3.** *In high SNR regimes, the variance of mutual information is expressed by the dilogarithm of correlation between square of fading amplitudes, i.e.,*

$$var\left(\log \mathbf{HH}\right) = N \cdot \pounds(\rho_0) + 2 \cdot \sum_{l=1}^{N-1} (N-l)\pounds(\rho_l), \tag{2.28}$$

*where $\pounds(\cdot)$ denotes the dilogarithm function which is defined as:*

$$\pounds(\rho_l) = -\int_0^{\rho} \frac{\log(1-t)}{t}\, dt, \tag{2.29}$$

*or equivalently,*

$$\pounds(\rho_l) = \sum_{k=1}^{\infty} \frac{(\rho_l)^k}{k^2} \tag{2.30}$$

*Here, $\rho_l$ represents the correlation coefficient between two exponential random variables, i.e.,*

$$\rho_l = \frac{Æ[A_n A_{n-l}] - Æ[A_n]Æ[A_{n-l}]}{\sqrt{var(A_n)var(A_{n-l})}} \tag{2.31}$$

*Proof.* See the details in Appendix 2.C. □

Through lemma 2.3.3, we can compute LHS of (2.28), if the correlation coefficient between $A_n$ and $A_{n-l}$ is known. Adopting the fading generation process in Section 2.2.2, we have another lemma that maintains the relationship of fading PSD and correlation coefficient between the square of corresponding fading amplitudes:

**Lemma 2.3.4.** *The correlation coefficients of square of fading amplitudes corresponds to the square of amplitudes of correlation between fadings, i.e.,*

$$\rho_l = R_l R_l^*, \tag{2.32}$$

*where*

$$R_l = \frac{1}{N}\sum_{n=1}^{N} S_n e^{j\cdot 2\pi \frac{n}{N} l} \tag{2.33}$$

*In* (2.33)*, $S_n$ denotes the discrete samples of the Power Spectral Density of the fadings, and satisfies:*

$$\sum_{n=1}^{N} S_n = N \tag{2.34}$$

*Proof.* See the details in Appendix 2.D. □

Combining lemmas 2.3.4 and 2.3.3, we can compute the variance of mutual information for any $N$ in high SNR regimes. However, still a closed form expression of the variance of mutual information is hard to obtain. Only when $N$ is sufficiently large, we can express the variance of mutual information in terms of fading PSD, in which we mainly exploit the fact that eigen values of auto-covariance matrix sequence follows the PSD when $N \to \infty$.

**Theorem 2.3.5.** *In high SNR regimes, and sufficiently large codeword length $N$, the variance of normalized mutual information decays linearly with $N$, where the scalar is determined by the doppler spectrum (PSD) of fading, i.e.,*

$$\lim_{N\to\infty}\lim_{\eta\to\infty} var\left(\frac{1}{N}\log\det\left(\mathbf{I}+\eta\cdot\mathbf{HH}\right)\right) = \frac{\sum_{k=1}^{\infty}\left(\int_{-0.5}^{0.5}\frac{S_k^2(f)}{k^2}\,df\right)}{N}\cdot\frac{6}{\pi^2}, \tag{2.35}$$

*where $S_1(f) = S(f)$, $S_k(f) = S_{k-1}(f) * S(f)$, note $*$ denotes cyclic convolution.*

*Proof.* See the details in Appendix 2.E. □

So finally, it is shown that the variance of mutual information decays linearly with $N$. Moreover, we define the spreadness of a PSD:

$$\mathsf{s} = \left( \frac{6}{\pi^2} \sum_{\mathsf{k}=1}^{\infty} \int_{-0.5}^{0.5} \frac{\mathsf{S}_\mathsf{k}^2(\mathsf{f})}{\mathsf{k}^2} \, \mathsf{df} \right)^{-1} \tag{2.36}$$

The spreadness of PSD serves at the scalor that how fast the variance decays with $N$. More spread the PSD has, faster the variance decays with $N$.

For i.i.d. channels, $S_k(f) = 1$ for any $k$. Then, the corresponding spreadness is: $\frac{6}{\pi^2} \sum_{k=1}^{\infty} \frac{1}{k^2} = 1$. For other channels, $\int_{-0.5}^{-0.5} S_k^2(f)\, df > 1$. Hence, we see that the spreadness of PSD is maximum in i.i.d. channels, while slowly varying channels yield smaller spreadness. Specifically, we note that if a PSD contains a dirac delta function, the spreadness would be zero. In those scenarios, reliable communications is not possible, because the variance of mutual information does not vanish even with large codeword length.

### 2.3.5 Concepts Related to the Second Order Statistics of Mutual Information: Multiplexing Gain and Diversity Gain

As discussed in Section 6.1, in high SNR regimes, multiplexing gain and diversity gain are essentially transmission rate and error rate, respectively. It has been depicted that the mean of mutual information is $\log \eta$ [Zheng and Tse(2003)]. Now, suppose we have a transmission scheme with transmission rate $r \cdot \log \eta$, $r < 1$, which corresponds to a transmission rate that is less than channel capacity. Here, $r$ is the multiplexing gain. Now, what is the corresponding outage probability for this multiplexing gain? First, we have the following result for the outage probability:

$$\begin{aligned} \lim_{\eta\to\infty} \lim_{N\to\infty} P_{out}(r \cdot \log \eta) &= \lim_{\eta\to\infty} \lim_{N\to\infty} P\left( \frac{1}{N} \sum_{n=1}^{N} \log(1 + \eta \cdot A_n) < r \cdot \log \eta \right) \\ &= \lim_{\eta\to\infty} \lim_{N\to\infty} P\left( \sum_{n=1}^{N} \log A_n < -(1-r) N \log \eta \right). \end{aligned} \tag{2.37}$$

Now, considering an i.i.d. channel, the above result can be simplified to:

$$\begin{aligned}
&\lim_{\eta\to\infty}\lim_{N\to\infty} P_{out}(r\cdot\log\eta)\\
&=\lim_{\eta\to\infty}\lim_{N\to\infty} P(\log A_1 < -(1-r)\log\eta, \log A_2 < -(1-r)\log\eta,\\
&\qquad\qquad \cdots, \log A_N < -(1-r)\log\eta)\\
&=\lim_{\eta\to\infty}\lim_{N\to\infty} P\left(\log A_1 < -(1-r)\log\eta\right)\cdot P\left(\log A_2 < -(1-r)\log\eta\right)\\
&\qquad\qquad \cdots P\left(\log A_N < -(1-r)\log\eta\right)\\
&=\lim_{\eta\to\infty}\lim_{N\to\infty} \left(P\left(\log A_1 < -(1-r)\log\eta\right)\right)^N . \qquad (2.38)
\end{aligned}$$

Because we have shown that $A_1$ follows exponential distribution, it is easy to show:

$$\begin{aligned}
\lim_{\eta\to\infty} P\left(\log A_1 < -(1-r)\log\eta\right) &= \lim_{\eta\to\infty} P\left(A_1 < \exp\left(-(1-r)\log\eta\right)\right)\\
&= \lim_{\eta\to\infty}\left(1-\exp(\eta^{-(1-r)})\right)\\
&= \lim_{\eta\to\infty}\exp\left(-(1-r)\log\eta\right), \qquad (2.39)
\end{aligned}$$

Accordingly, the outage probability for an i.i.d. channel corresponds to:

$$\lim_{\eta\to\infty}\lim_{N\to\infty} P_{out}(r\cdot\log\eta) = \exp\left(-(1-r)\cdot N\cdot\log\eta\right). \qquad (2.40)$$

To compute the diversity gain defined in (2.2), and we can replace error probability with outage probability when computing diversity gain [Zheng and Tse(2003)], i.e. ,

$$d(r) = -\lim_{\eta\to\infty}\frac{P_e(r\cdot\log\eta)}{\log\eta} = -\lim_{\eta\to\infty}\frac{P_{out}(r\cdot\log\eta)}{\log\eta}. \qquad (2.41)$$

Hence, in i.i.d. channels, diversity gain can be expressed as a function of multiplexing gain:

$$d(r) = -\lim_{N\to\infty}\lim_{\eta\to\infty}\frac{P_{out}(r\cdot\log\eta)}{N\cdot\log\eta} = 1-r. \qquad (2.42)$$

Next, we consider the correlated channels. What is the difference between correlated channels and i.i.d. channels? Note that for bandlimited and regular fading

processes, the capacity is a summation of many correlated random variables. Hence, when $N \to \infty$, the capacity distribution for correlated channels would be the same as i.i.d. channels. The only difference is that the convergence rate in correlated channels is slower, i.e., given the same number of fadings, the variance of correlated fadings is larger than the variance of independent fadings. In other words, variance of $N'$ correlated fadings would be equal to the variance of $N$ independent variables, where:

$$N' = \frac{N}{\mathsf{s}}, \quad \mathsf{s} \leq 1. \tag{2.43}$$

Note that $\mathsf{s} \leq 1$ for all cases, where the equality only holds for i.i.d. channels. Intuitively, this corresponds to the fact that, compared to i.i.d. fadings, in correlated fadings, longer codeword is needed to achieve the same variance (or error probability).

Replacing $N$ in (2.40) with $N' \cdot \mathsf{s}$ in (2.43), the outage probability for correlated channels would correspond to:

$$\lim_{\eta\to\infty} \lim_{N\to\infty} P_{out}(r \cdot \log \eta) = \exp\left(-(1-r) \cdot \log \eta \cdot \mathsf{s} \cdot \mathsf{N}\right) \tag{2.44}$$

Then, based on the definition of diversity gain, the diversity gain is:

$$d(r) = -\lim_{N\to\infty} \lim_{\eta\to\infty} \frac{P_{out}(r \cdot \log \eta)}{N \cdot \log \eta} = \mathsf{s} \cdot (1 - \mathsf{r}) \tag{2.45}$$

which exactly corresponds to the result presented in Section 6.1.

In this section, the derivation process for the tradeoff between diversity gain and multiplexing gain has been demonstrated via a mathematical approach. Numerical results and intuitive explanations on the tradeoff is presented in the next section.

## 2.4 Intuitions and Numerical Examples

First, this section intuitively demonstrates the tradeoff through discussing the effects of 1) SNR, $\eta$, 2) codeword length, $N$, and 3) fading PSD, $S(f)$, on the mean and variance of the normalized mutual information:

$$\mathcal{I}_{norm} = \frac{\mathcal{I}(\vec{x}; \vec{y}, \vec{h})}{N}, \tag{2.46}$$

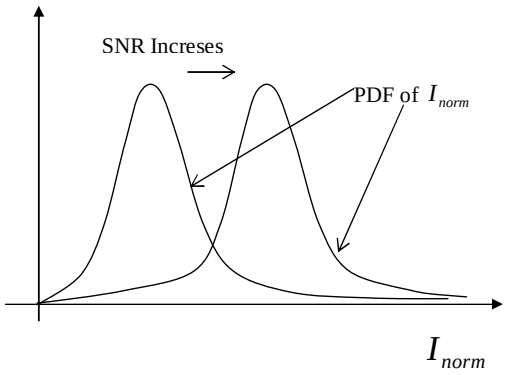


Figure 2.3: The effects of SNR on mean and variance.

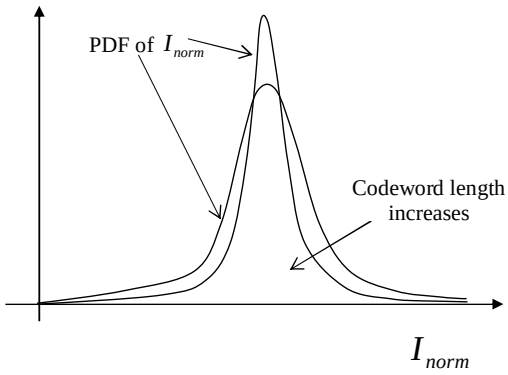


Figure 2.4: The effects of codeword length on mean and variance.

in which the optimal input covariance matrix has been assumed to be an identity matrix. Because the mean of the normalized mutual information is channel capacity, we choose to analyze the normalized mutual information, but not the mutual information.

The effects of $\eta$ are shown in Fig. 2.3. Assuming fixed $N$ and $S(f)$, Theorem 2.3.1 reveals that the mean increases linearly with logarithm of SNR; while Theorem 2.3.2 states that the variance does not increase with SNR. Hence, *the mean increases with SNR while the variance is independent of SNR.*

The effects of $N$ are shown in Fig. 2.4. As stated by Theorem 2.3.2, the variance decays linearly with $N$. Meanwhile, it is easy to compute that the mean does not chage with $N$. Hence, *the variance decreases with $N$ while the mean is independent of $N$.*

Then, how to understand the tradeoff via the effects of SNR and $N$ on the mean and the variance? It is shown in Fig. 2.5. First, it is noted that in high SNR regimes, the mean is $\log \eta$. Next, let a transmission strategy possess a multiplexing gain $r$, i.e., the transmission rate is $r \cdot \log \eta$. Note that the outage probability, $\mathsf{P}(\mathcal{I}_{\mathsf{norm}} < \mathsf{r} \cdot \log \eta)$, corresponds to the shadowed area in Fig. 2.5. In general, it is hard to maintain an

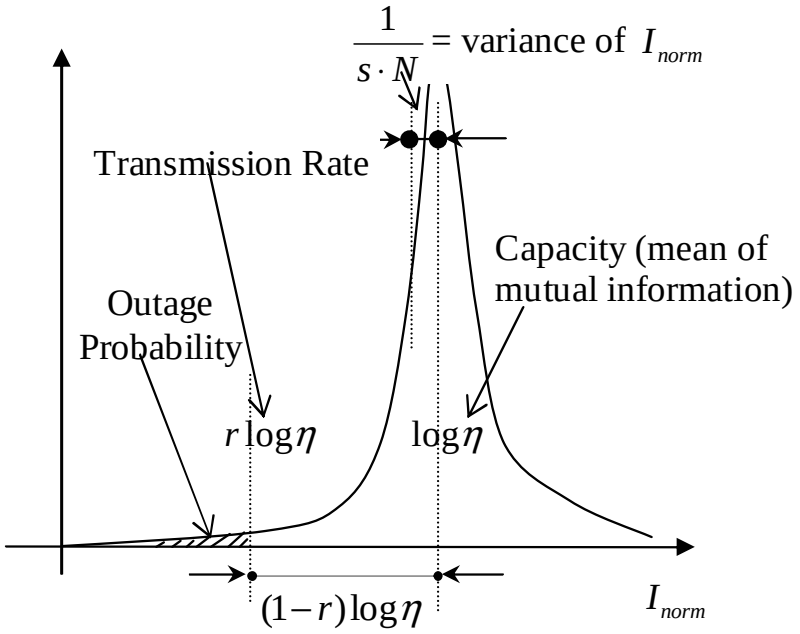


Figure 2.5: Understanding the Tradeoff between Diversity Gain and Multiplexing Gain through the PDF of $\mathcal{I}_{norm}$.

exact closed form expression for outage probability, because it is hard to obtain the exact PDF of $\mathcal{I}_{norm}$, the normalized mutual information.

However, when $\eta \to \infty$ and $N \to \infty$, as long as $r < 1$, the "distance" between the transmission rate and channel capacity (shown in Fig. 2.5) is much larger than the variance [see (2.35)], i.e., $(1-r)\log\eta \gg \frac{1}{\mathsf{s}}N$. In this case, the outage probability is the "tail" of the PDF of the normalized mutual information. Our contribution is to show that the "tail" of the PDF of $\mathcal{I}_{norm}$ resembles an exponential function for Rayleigh Fading channels [see (2.39)]. The parameter of this exponential function is: 1) proportional to the difference between capacity and transmission rate, i.e., $(1-r)\log\eta$; and 2) inversely proportional to the variance of normalized mutual information, $\frac{1}{\mathsf{s}\cdot\mathsf{N}}$. Then, it is shown that $P_{out} = exp(-(1-r)\mathsf{sN}\log\eta)$ [see (2.44)], which directly corresponds to the tradeoff between diversity gain and multiplexing gain.

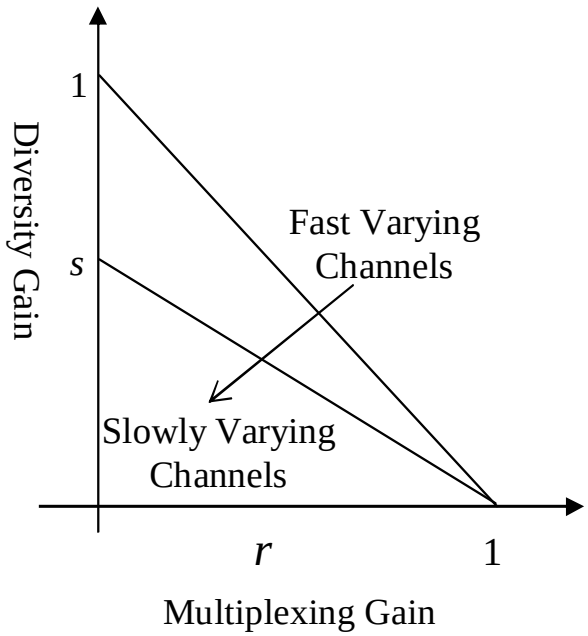


Figure 2.6: Effects of the spreadness of Fading PSD on the Tradeoff between Diversity Gain and Multiplexing Gain

The ultimate goal of this work is to analyze the effects of spreadness of PSD on the tradeoff between diversity gain and multiplexing gain, which is shown in Fig. 2.6. The tradeoff itself has a linear relationship, i.e., increasing multiplexing gain decreases diversity gain linearly. Here, the spreadness of PSD, $\mathsf{s}$, serves as the slope of the linear relationship. Assuming the same multiplexing gain, fast varying channels, which have bigger spreadness, achieve better diversity gain compared to slowly varying channels.

Finally, to understand why $\mathsf{s}$ is called "spreadness", we compute the spreadness for a family of PSDs. The structures of PSDs in this family follows similar shape, as shown in Fig. 2.7. The only parameter of this family, $\kappa$, corresponds to:

$$\kappa = \frac{a}{b}, \tag{2.47}$$

where $a$ and $b$ has been defined in Fig. 2.7. The advantage of this family is that the Lebesgue measure of the members in this family ranges from 0 to 1. Based on

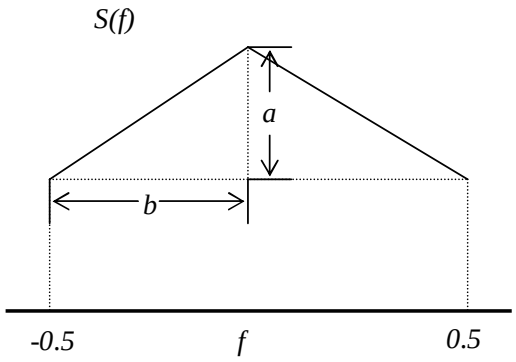


Figure 2.7: A PSD family

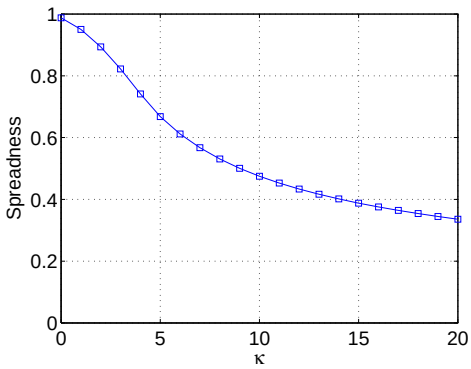


Figure 2.8: Spreadness for the PSD family in Fig. 2.7

the geometry of Fig. 2.7, it is easy to see that: $\kappa = 0$ corresponds to i.i.d channel; $0 < \kappa \leq 4$ corresponds to regular processes; $4 < \kappa < \infty$ corresponds to bandlimited processes. When $\kappa \to \infty$, the PSD corresponds to time-invariant channels.

The spreadness with respect to $\kappa$ is plotted in Fig. 2.8. As expected, larger $\kappa$ leads to smaller spreadness. Hence, as PSD becomes more flat ($\kappa$ decreases), the spreadness increases.

## 2.5 Conclusions

In this work, we discuss the fundamental tradeoff between multiplexing gain and diversity gain in SISO channels with correlated fadings. Different from previous studies, the view of power spectral density (PSD) is adopted to describe correlated fadings. Categorization of fading processes and corresponding fading generation methods are discussed in detail. Then, the tradeoff between multiplexing gain and diversity gain are shown to resemble a simple linear relationship, where the "spreadeness" of the fading PSD serves as the slope of the linear relationship. More spread PSD, i.e., faster varying channels, yield better diversity gain than slowly varying channels, assuming the same multiplexing gain.

## 2.A Proof of Theorem 2.3.1

First, we note that the optimum $\mathbf{Q}_{xx}$ satisfies:

$$Æ\left[\log\det(\mathbf{I}+\eta\cdot\mathbf{HH})\right]\leq Æ\left[\log\det(\mathbf{I}+\eta\cdot\mathbf{HQ}_{xx}\mathbf{H})\right], \tag{2.48}$$

where in the LHS, the optimum $\mathbf{Q}_{xx}$ is replaced by $\mathbf{I}$.

Second, we have:

$$Æ\left[\log\det(\mathbf{I}+\eta\cdot\mathbf{HQ}_{xx}\mathbf{H})\right]\leq Æ\left[\log\det(\mathbf{I}+N\eta\cdot\mathbf{HH})\right]. \tag{2.49}$$

Since all eigen values of $N\mathbf{I}$ is greater than the biggest eigen value of $\mathbf{Q}_{xx}$, due to the average power constraint, then, due to the convexity of the function $\log\det(\cdot)$, the $\log\det$ of the two matrices have the above inequality.

Combining (2.49) and (2.48), the upper bound and lower bound of supremum of mutual information would correspond to:

$$Æ\left[\log\det(\mathbf{I}+\eta\cdot\mathbf{HH})\right]\leq Æ\left[\log\det(\mathbf{I}+\eta\cdot\mathbf{HQ}_{xx}\mathbf{H})\right]\leq Æ\left[\log\det(\mathbf{I}+N\eta\cdot\mathbf{HH})\right] \tag{2.50}$$

The goal of this proof is to show that the upper bound and lower bound approach the same value, when $\eta\to\infty$, i.e.,

$$\lim_{\eta\to\infty}\frac{Æ\left[\log\det(\mathbf{I}+\eta\cdot\mathbf{HH})\right]}{\log\eta}=\lim_{\eta\to\infty}\frac{Æ\left[\log\det(\mathbf{I}+N\eta\cdot\mathbf{HH})\right]}{\log\eta}. \tag{2.51}$$

Then, we cay say the optimal $\mathbf{Q}_{xx}$ achieves the same performance as $\mathbf{I}$ in high SNR regimes, i.e.,

$$\lim_{\eta\to\infty}\frac{Æ\left[\log\det(\mathbf{I}+\eta\cdot\mathbf{HH})\right]}{\log\eta}=\lim_{\eta\to\infty}\frac{Æ\left[\log\det(\mathbf{I}+\eta\cdot\mathbf{HQ}_{xx}\mathbf{H})\right]}{\log\eta}. \tag{2.52}$$

In other words, $\mathbf{I}$ would be optimal in terms of achieving channel capacity in high SNR regimes.

First, in high SNR regimes, the lower bound corresponds to:

$$\lim_{\eta\to\infty} \frac{\text{Æ}\left[\log\det(\mathbf{I}+\eta\cdot\mathbf{HH})\right]}{\log\eta} \tag{2.53}$$

$$= \lim_{\eta\to\infty} \frac{\text{Æ}\left[\sum_{n=1}^{N}\log(1+\eta\cdot A_n)\right]}{\log\eta}$$

$$= \lim_{\eta\to\infty} \frac{\sum_{n=1}^{N}\text{Æ}\left[\log(1+\eta\cdot A_n)\right]}{\log\eta}$$

$$= \lim_{\eta\to\infty} \frac{N\cdot\int_0^\infty \log(1+\eta\cdot A)\cdot P_A(A)\,\mathrm{dA}}{\log\eta}, \tag{2.54}$$

where $A_n = h_n h_n^*$, i.e., $A_n$ is the square of the amplitudes of the channel fadings.

By dominant convergence theorem, we can exchange the integral and the limitation, i.e.,

$$\lim_{\eta\to\infty} \frac{\int_0^\infty \log(1+\eta\cdot A)\cdot p_A(A)\,\mathrm{dA}}{\log\eta} = \int_0^\infty \lim_{\eta\to\infty} \frac{\log(1+\eta\cdot A)}{\log\eta}\cdot p_A(A)\,\mathrm{dA}, \tag{2.55}$$

where $p_A(A)$ denotes the PDF of $A$, the square of fading amplitude.

Note that when $\eta\to\infty$,

$$\frac{\log(1+\eta\cdot A)}{\log\eta} = \frac{\log\eta+\log A+\log(1+\frac{1}{\eta A})}{\log\eta} \to 1. \tag{2.56}$$

Inserting (2.56) in (2.55):

$$\int_0^\infty \lim_{\eta\to\infty} \frac{\log(1+\eta\cdot A)}{\log\eta}\cdot p_A(A)\,\mathrm{dA} = \int_{0^+}^\infty \mathrm{p_A(A)}\,\mathrm{dA}. \tag{2.57}$$

Note that at the RHS of (2.57), the lower limit of the integration is $0^+$ but not 0. Hence, as long as the distribution of the square of fading amplitude has a zero probability at the amplitude of zero, i.e., the PDF does not contain a delta function at the $A=0$, (2.57) would correspond to one in SISO channels. A similar phenomenon has been observed in [Koch and Lapidoth(2006)]. Noting that all current fading models (Rayleigh, Ricean, and Nakagami) satisfy this condition, hence (2.57) equals to one holds for all frequency flat SISO fading channels.

Substituting (2.57) to the expression of the lower bound, (2.53), the lower bound is shown to be:

$$N \int_{0^+}^{\infty} p_A(A)\,\mathrm{dA}, \tag{2.58}$$

which in most cases is simply $N$.

The discussion about the upper bound would be similar. First, the upper bound corresponds to:

$$\begin{aligned} &\lim_{\eta\to\infty} \frac{\text{Æ}\left[\log\det(\mathbf{I}+\eta N\mathbf{HH})\right]}{\log\eta} \\ =&\lim_{\eta\to\infty} \frac{N\cdot\int_0^{\infty}\log(1+\eta NA)p_A(A)\,\mathrm{dA}}{\log\eta} \\ =&N\cdot\int_0^{\infty}\lim_{\eta\to\infty}\frac{\log(1+\eta NA)}{\log\eta}p_A(A)\,\mathrm{dA}. \end{aligned} \tag{2.59}$$

Note that when $\eta\to\infty$,

$$\frac{\log(1+\eta NA)}{\log\eta} = \frac{\log\eta+\log NA+\log(1+\frac{1}{\eta A})}{\log\eta} \to 1. \tag{2.60}$$

Substituting (2.60) into (2.59), the upper bound would correspond to:

$$N \int_{0^+}^{\infty} p_A(A)\,\mathrm{dA}, \tag{2.61}$$

which is the same as the result for lower bound.

Hence, $\mathbf{I}$ would be the optimal input covariance matrix in terms of achieving capacity in high SNR regimes.

## 2.B Proof of Theorem 2.3.2

The theorem states:

$$\lim_{\eta\to\infty}\mathrm{var}\left(\log\det\left(\mathbf{I}+\eta\cdot\mathbf{HH}\right)\right) = \mathrm{var}\left(\log\mathbf{HH}\right) \tag{2.62}$$

Considering $\mathbf{H}$ is diagonal, the above equation corresponds to:

$$\lim_{\eta\to\infty}\mathrm{var}\left(\sum_{n=1}^{N}\log(1+\eta A_n)\right) = \mathrm{var}\left(\sum_{n=1}^{N}\log A_n\right). \tag{2.63}$$

Expanding the LHS of (2.63):

$$N \cdot f_{LHS}(\eta, 0) + 2 \cdot \sum_{l=1}^{N-1} (N-l) \cdot f_{LHS}(\eta, l), \tag{2.64}$$

where $f_{LHS}(\eta, l)$ corresponds to:

$$f_{LHS}(\eta, l) = Æ[\log(1+\eta A_n)\log(1+\eta A_{n-l})] - Æ[\log(1+\eta A_n)]Æ[\log(1+\eta A_{n-l})]. \tag{2.65}$$

Expanding the RHS of (2.63):

$$N \cdot f_{RHS}(0) + 2 \cdot \sum_{l=1}^{N-1} (N-l) \cdot f_{RHS}(l), \tag{2.66}$$

where $f_{RHS}(\eta, l)$ corresponds to:

$$f_{RHS}(l) = Æ[\log(A_n)\log(A_{n-l})] - Æ[\log(A_n)]Æ[\log(A_{n-l})]. \tag{2.67}$$

Comparing (2.64) and (2.66), to prove this theorem, we just need to prove that

$$\lim_{\eta \to \infty} f_{LHS}(\eta, l) = f_{RHS}(l) \tag{2.68}$$

To check if (2.68) holds, we check the detailed structure of $f_{LHS}(\eta, l)$.

It is known that the joint PDF of two exponential variables, $A$ and $A'$, corresponds to [Simon and Alouini(1998)]:

$$p(A, A') = \frac{1}{1-\rho} \exp\left(-\frac{A+A'}{1-\rho}\right) I_0\left(-\frac{2\sqrt{\rho A A'}}{1-\rho}\right), \tag{2.69}$$

where $I_0$ denotes the modified Bessel function of the first kind with zero order, and the parameter $\rho$ corresponds to the correlation coefficient between $A$ and $A'$:

$$\rho = \frac{Æ[AA'] - Æ[A]Æ[A']}{\sqrt{\mathrm{var}(A)\mathrm{var}(A')}}, \tag{2.70}$$

Interestingly, it has been observed that in general the second order statistics, i.e., correlation coefficient, can not determine the joint PDF of multiple exponential

variables. The only exception is the bivariate exponential variables, in which the correlation coefficient determines the joint PDF [Mallik(2003)].

Applying (2.69), the first part of (2.65) corresponds to:

$$\begin{aligned}&\text{Æ}[\log(1+\eta A_n)\log(1+\eta A_{n-l})]\\&=\int_0^\infty\int_0^\infty \log(1+\eta A)\log(1+\eta A')\frac{1}{1-\rho_l}\exp\left(-\frac{A+A'}{1-\rho_l}\right)I_0\left(-\frac{2\sqrt{\rho_l AA'}}{1-\rho_l}\right)\,\mathrm{dA}\,\mathrm{dA}',\end{aligned} \tag{2.71}$$

where $\rho_l$ denotes the correlation coefficient between $A_n$ and $A_{n-l}$. The second part of (2.65) corresponds to:

$$\text{Æ}[\log(1+\eta A_n)]\text{Æ}[\log(1+\eta A_{n-l})]=\left(\int_0^\infty \log(1+\eta A)\exp(-A)\,\mathrm{dA}\right)^2 \tag{2.72}$$

Here, we adopt a similar approach as in [Tan and Beaulieu(1997)] for analyzing (2.71), i.e., we use the infinite series representation of the modified Bessel function [(8.441.1) in [Gradshteyn and Ryzhik(1965)]],

$$I_0\left(-\frac{2\sqrt{\rho_l AA'}}{1-\rho_l}\right)=\sum_{k=0}^{\infty}\frac{(\rho_l AA')^k}{(1-\rho_l)^{2k}(k!)^2}. \tag{2.73}$$

Substituting (2.73) into (2.71):

$$\begin{aligned}&\text{Æ}[\log(1+\eta A_n)\log(1+\eta A_{n-l})]\\&=\int_0^\infty\int_0^\infty \log(1+\eta A)\log(1+\eta A')\frac{1}{1-\rho_l}\exp\left(-\frac{A+A'}{1-\rho_l}\right)\sum_{k=0}^{\infty}\frac{(\rho_l AA')^k}{(1-\rho_l)^{2k}(k!)^2}\,\mathrm{dA}\,\mathrm{dA}'\end{aligned} \tag{2.74}$$

$$=\sum_{k=0}^{\infty}\frac{(\rho_l)^k}{(1-\rho_l)^{2k+1}(k!)^2}\left(\int_0^\infty A^k\log(1+\eta A)\exp(-\frac{A}{1-\rho_l})\,\mathrm{dA}\right)^2 \tag{2.75}$$

To see the effect of high SNR $\eta$, we write:

$$\begin{aligned}\log(1+\eta A)&=\log\eta+\log A+\log(1+\frac{1}{\eta A})\\&=\log\eta+\log A+\mathcal{O}(1/\eta),\end{aligned} \tag{2.76}$$

where the second equality holds only in high SNR regimes.

Substituting (2.76) in (2.75), (2.71) in high SNR regimes would correspond to:

$$\begin{aligned}
&\lim_{\eta\to\infty} Æ[\log(1+\eta A_n)\log(1+\eta A_{n-l})] \\
&= \sum_{k=0}^{\infty} \frac{(\rho_l)^k}{(1-\rho_l)^{2k+1}\cdot(k!)^2}\left(\int_0^{\infty}(\log\eta+\log A)\exp(-\frac{A}{1-\rho_l})A^k\,\mathrm{dA}\right)^2 \qquad (2.77)\\
&= (\log\eta)^2\cdot\underbrace{\sum_{k=0}^{\infty}\frac{(\rho_l)^k}{(1-\rho_l)^{2k+1}(k!)^2}\left(\int_0^{\infty}A^k\exp(-\frac{A}{1-\rho_l})\,\mathrm{dA}\right)^2}_{a_1}+ \\
&\underbrace{2\log\eta\cdot\sum_{k=0}^{\infty}\frac{(\rho_l)^k}{(1-\rho_l)^{2k+1}(k!)^2}\int_0^{\infty}A^k\log A\exp(-\frac{A}{1-\rho_l})\,\mathrm{dA}\int_0^{\infty}\mathrm{A^k}\exp(-\frac{\mathrm{A}}{1-\rho_l})\,\mathrm{dA}}_{b_1}+ \\
&\underbrace{\sum_{k=0}^{\infty}\frac{(\rho_l)^k}{(1-\rho_l)^{2k+1}(k!)^2}\left(\int_0^{\infty}A^k\exp(-\frac{A}{1-\rho_l})\,\mathrm{dA}\right)^2}_{c_1} \qquad (2.78)
\end{aligned}$$

For notation wise simplicity, here we define the coefficients of powers of $\log\eta$ as $a_1, b_1, c_1$, as shown in (2.77).

Similarly, using (2.76), the second part of (2.65) in high SNR regimes corresponds to:

$$\begin{aligned}
&\lim_{\eta\to\infty}\left(Æ[\log(1+\eta A)]\right)^2 \\
&= \left(\int_0^{\infty}(\log\eta+\log A)\exp(-A)\,\mathrm{dA}\right)^2 \\
&= (\log\eta)^2\cdot\left(\int_0^{\infty}\exp(-A)\,\mathrm{dA}\right)^2+2\log\eta\cdot\int_0^{\infty}\log A\exp(-A)\,\mathrm{dA}+\left(\int_0^{\infty}\log\mathrm{A}\exp(-\mathrm{A})\,\mathrm{dA}\right)^2 \\
&= (\log\eta)^2+2\log\eta\cdot\zeta+\zeta^2. \qquad (2.79)
\end{aligned}$$

The last equality is based on the following definite integral [(4.331.1) in [Gradshteyn and Ryzhik(1

$$\int_0^{\infty}\exp(-\mu x)\log x\,\mathrm{dx}=-\frac{1}{\mu}(\zeta+\log\mu), \qquad (2.80)$$

where $\zeta$ is the Euler-Mascheroni Constant[2], defined as:

$$\begin{aligned}\zeta &= \lim_{n\to\infty}\left[\sum_{m=1}^{n}\frac{1}{m} - \log n\right] \\ &= 0.5772156....\end{aligned} \tag{2.81}$$

Moreover, we see the RHS of (2.68) corresponds to:

$$\begin{aligned}&f_{RHS}(l) = \text{Æ}[\log(A_n)\log(A_{n-l})] - \text{Æ}[\log(A_n)]\text{Æ}[\log(A_{n-l})] \\ &= \int_0^\infty\int_0^\infty \frac{\log A \log A'}{1-\rho_l}\exp\left(-\frac{A+A'}{1-\rho_l}\right) I_0\left(-\frac{2\sqrt{\rho_l A A'}}{1-\rho_l}\right)\,\text{dA}\,\text{dA}' - \left(\int_0^\infty \log \text{A}\exp(-\text{A})\,\text{dA}\right)^2 \\ &= \sum_{k=0}^{\infty}\frac{(\rho_l)^k}{(1-\rho_l)^{2k+1}(k!)^2}\left(\int_0^\infty A^k\log A\exp(-\frac{A}{1-\rho_l})\,\text{dA}\right)^2 - \left(\int_0^\infty \log A\exp(-A)\,\text{dA}\right)^2 \\ &= c_1 - \zeta^2.\end{aligned} \tag{2.82}$$

The above results have already incorporated the infinite series representation of modified Bessel functions, (2.73), and the definite integral in (2.80).

Through the results in (2.82), (2.79) and (2.77), to prove (2.68), we just need to show that $a_1 = 1$ and $b_1 = \zeta$.

We first check if $a_1 = 1$.

$$\begin{aligned}a_1 &= \sum_{k=0}^{\infty}\frac{(\rho_l)^k}{(1-\rho_l)^{2k+1}(k!)^2}\left(\int_0^\infty A^k\exp(-\frac{A}{1-\rho_l})\,\text{dA}\right)^2 \\ &= \sum_{k=0}^{\infty}\frac{(\rho_l)^k(1-\rho_l)}{(k!)^2}\left(\int_0^\infty(\frac{A}{1-\rho_l})^k\exp(-\frac{A}{1-\rho_l})\,\text{d}\frac{\text{A}}{1-\rho_l}\right)^2.\end{aligned} \tag{2.83}$$

Let $x = \frac{A}{1-\rho_l}$, the above results correspond to:

$$a_1 = \sum_{k=0}^{\infty}\frac{(\rho_l)^k(1-\rho_l)}{(k!)^2}\left(\int_0^\infty x^k\exp(-x)\,\text{dx}\right)^2. \tag{2.84}$$

We also know that [(3.381.4) in [Gradshteyn and Ryzhik(1965)]]

$$\int_0^\infty x^k\exp(-\mu x)\,\text{dx} = \frac{1}{\mu^{\text{k+1}}}\Gamma(\text{k}+1), \tag{2.85}$$

[2]In most literatures, the notation of Euler-Mascheroni Constant is $\gamma$. However, we have used $\gamma$ to represent the degrees of freedom in this work, so we use another notation, $\zeta$, for this constant.

where $\Gamma(\cdot)$ denotes the Gamma function. It has been well known that

$$\Gamma(k) = (k-1)! \quad \text{for } k = 1, 2, 3... \tag{2.86}$$

Then, $a_1$ finally corresponds to:

$$a_1 = \sum_{k=0}^{\infty} \frac{\rho^k(1-\rho)}{(k!)^2}(k!)^2 = (1-\rho)\sum_{k=0}^{\infty}\rho^k = 1 \tag{2.87}$$

Now, we check if $b_1 = \zeta$.

$$\begin{aligned} b_1 &= \sum_{k=0}^{\infty} \frac{(\rho_l)^k}{(1-\rho_l)^{2k+1}(k!)^2} \int_0^{\infty} A^k \log A \exp(-\frac{A}{1-\rho})\,\mathrm{dA} \int_0^{\infty} \mathrm{A^k} \exp(-\frac{\mathrm{A}}{1-\rho})\,\mathrm{dA} \\ &= \sum_{k=0}^{\infty} \frac{(\rho_l)^k(1-\rho_l)}{k!} \int_0^{\infty} (\frac{A}{1-\rho_l})^k \left(\log\frac{A}{1-\rho_l} + \log(1-\rho_l)\right) \exp(-\frac{A}{1-\rho_l})\,\mathrm{dA} \end{aligned} \tag{2.88}$$

Here, we already exploited the result in (2.85). Again, let $x = \frac{A}{1-\rho_l}$,

$$b_1 = \sum_{k=0}^{\infty} \frac{(\rho_l)^k(1-\rho_l)}{k!} \left(\int_0^{\infty} x^k \exp(-x)\log x\,\mathrm{dx} + \log(1-\rho_l)\cdot\int_0^{\infty}\mathrm{x^k}\exp(-\mathrm{x})\,\mathrm{dx}\right) \tag{2.89}$$

The above can be solved using [(4.352.4) in [Gradshteyn and Ryzhik(1965)]],

$$\int_0^{\infty} x^k \exp(-x)\log x\,\mathrm{dx} = \Gamma'(\mathrm{k}+1), \tag{2.90}$$

where $\Gamma'(\cdot)$ denotes the derivative of Gamma function, and corresponds to [Weisstein(Web)]:

$$\Gamma'(k+1) = k!\left(\zeta - \sum_{m=1}^{k}\frac{1}{m}\right), \quad \text{for } k = 1, 2, 3... \tag{2.91}$$

and $\zeta$ has been defined in (2.81). Then, $b_1$ would correspond to:

$$\begin{aligned} b_1 &= \sum_{k=0}^{\infty} \frac{(\rho_l)^k(1-\rho_l)}{k!} \left(k!\left(\zeta - \sum_{m=1}^{k}\frac{1}{m}\right) + \log(1-\rho_l)\cdot k!\right) \\ &= (1-\rho_l)\left(\zeta + \log(1-\rho_l)\right)\sum_{k=0}^{\infty}(\rho_l)^k + (1-\rho_l)\sum_{k=0}^{\infty}\sum_{m=1}^{k}\frac{(\rho_l)^k}{m} \end{aligned} \tag{2.92}$$

$$
\begin{aligned}
&= \zeta + \log(1-\rho_l) - (1-\rho_l)\sum_{m=1}^{\infty}\sum_{k=m}^{\infty}\frac{(\rho_l)^k}{m} \\
&= \zeta + \log(1-\rho_l) - \sum_{m=1}^{\infty}\frac{(\rho_l)^m}{m} \\
&= \zeta
\end{aligned} \tag{2.93}
$$

Hence, we have shown that $a_1 = 1$ and $b_1 = \zeta$. Then, (2.68) holds. Via (2.68), (2.63) can be proved, i.e., Lemma 2.3.2 is proved.

## 2.C Proof of Lemma 2.3.3

In appendix 2.B, we have shown that

$$\text{var}\,(\log \mathbf{HH}) = N \cdot f_{RHS}(0) + 2 \cdot \sum_{l=1}^{N-1}(N-l) \cdot f_{RHS}(l), \tag{2.94}$$

where

$$f_{RHS}(l) = Æ[\log(A_n)\log(A_{n-l})] - Æ[\log(A_n)]Æ[\log(A_{n-l})]. \tag{2.95}$$

Hence, to prove Lemma 2.3.3, we just need to show that:

$$Æ[\log(A_n)\log(A_{n-l})] - Æ[\log(A_n)]Æ[\log(A_{n-l})] = \pounds(\rho_l), \tag{2.96}$$

where $\rho_l$ is the correlation coefficient between $A_n$ and $A_{n-l}$, and $\pounds(\cdot)$ denotes the dilogarithm function which is defined in (2.29).

Using (2.80), we see that $Æ[\log(A_n)] = \zeta$. Therefore, to prove the Lemma, we just need to show:

$$Æ[\log(A_n)\log(A_{n-l})] = \zeta^2 + \pounds(\rho_l), \tag{2.97}$$

where $\zeta$ has been defined in (2.81).

Applying the joint PDF of two exponential variables (2.69), we write the LHS of (2.97) as:

$$Æ[\log(A_n)\log(A_{n-l})] = \int_0^{\infty}\int_0^{\infty}\frac{\log A \log A'}{1-\rho_l}\exp\left(-\frac{A+A'}{1-\rho_l}\right) I_0\left(-\frac{2\sqrt{\rho_l AA'}}{1-\rho_l}\right)\,\text{dA}\,\text{dA}', \tag{2.98}$$

then using the infinite series representation of modified Bessel functions (2.73),

$$Æ[\log(A_n)\log(A_{n-l})] = \sum_{k=0}^{\infty} \frac{(\rho_l)^k}{(k!)^2(1-\rho_l)} \left( \int_0^{\infty} \left( \frac{A}{1-\rho_l} \right)^k \log A \exp(-\frac{A}{1-\rho_l})\, \mathrm{dA} \right)^2 . \tag{2.99}$$

Letting $x = \frac{A}{1-\rho_l}$,

$$\begin{aligned} &Æ[\log(A_n)\log(A_{n-l})] \\ &= \sum_{k=0}^{\infty} \frac{(\rho_l)^k(1-\rho_l)}{(k!)^2} \left( \int_0^{\infty} x^k \log x \exp(-x)\, \mathrm{dx} + \log(1-\rho_l) \int_0^{\infty} \mathrm{x^k} \exp(-\mathrm{x})\, \mathrm{dx} \right)^2 . \end{aligned} \tag{2.100}$$

Using the definite integral results in (2.80), (2.85) and (2.90),

$$Æ[\log(A_n)\log(A_{n-l})] = \sum_{k=0}^{\infty} (\rho_l)^k (1-\rho_l) \left( \log(1-\rho_l) - \zeta + z_k \right)^2 , \tag{2.101}$$

where

$$z_k = \begin{cases} \sum_{m=1}^{k} \frac{1}{m} & \text{for } k = 1, 2, 3... \\ 0 & \text{for } k = 0 \end{cases} \tag{2.102}$$

Expanding the square of polynomial in (2.101),

$$\begin{aligned} &Æ[\log(A_n)\log(A_{n-l})] \\ &= (1-\rho_l) \cdot (\log(1-\rho_l) - \zeta)^2 \cdot \underbrace{\sum_{k=0}^{\infty} (\rho_l)^k}_{a2} + \\ &(1-\rho_l) \cdot 2\left(\log(1-\rho_l) - \zeta\right) \cdot \underbrace{\sum_{k=1}^{\infty} (\rho_l)^k \sum_{m=1}^{\infty} \frac{1}{m}}_{b2} + (1-\rho_l) \cdot \underbrace{\sum_{k=1}^{\infty} (\rho_l)^k \left( \sum_{m=1}^{k} \frac{1}{m} \right)^2}_{c2} \end{aligned} \tag{2.103}$$

The essential part of this proof is to analyze $a_2$, $b_2$ and $c_2$. First, calculations of $a_2$ and $b_2$ are straightforward:

$$a_2 = \frac{1}{1-\rho_l}, \tag{2.104}$$

and,

$$b_2 = \sum_{k=1}^{\infty}\sum_{m=1}^{k}\frac{(\rho_l)^k}{m} = \sum_{m=1}^{\infty}\sum_{k=m}^{\infty}\frac{(\rho_l)^k}{m} = \frac{1}{1-\rho_l}\sum_{m=1}^{\infty}\frac{(\rho_l)^m}{m} = -\frac{\log(1-\rho_l)}{1-\rho_l}. \tag{2.105}$$

The analysis of $c_2$ is much more difficult:

$$\begin{aligned}
c_2 &= \sum_{k=1}^{\infty}(\rho_l)^k\left(\sum_{m=1}^{k}\frac{1}{m}\right)^2 \\
&= \sum_{k=1}^{\infty}\sum_{m=1}^{k}\sum_{n=1}^{k}\frac{(\rho_l)^k}{mn} \\
&= \sum_{m=1}^{\infty}\sum_{k=m}^{\infty}\sum_{n=1}^{k}\frac{(\rho_l)^k}{mn} \\
&= \sum_{m=1}^{\infty}\sum_{n=1}^{m}\sum_{k=m}^{\infty}\frac{(\rho_l)^k}{mn} + \sum_{m=1}^{\infty}\sum_{n=m+1}^{\infty}\sum_{k=n}^{\infty}\frac{(\rho_l)^k}{mn} \\
&= \frac{1}{1-\rho_l}\left(\sum_{m=1}^{\infty}\sum_{n=1}^{m}\frac{(\rho_l)^m}{mn} + \sum_{m=1}^{\infty}\sum_{n=m+1}^{\infty}\frac{(\rho_l)^n}{mn}\right) \\
&= \frac{1}{1-\rho_l}\left(\sum_{m=1}^{\infty}\sum_{n=1}^{m}\frac{(\rho_l)^m}{mn} + \sum_{n=2}^{\infty}\sum_{m=1}^{n-1}\frac{(\rho_l)^n}{mn}\right) \\
&= \frac{1}{1-\rho_l}\left(\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\frac{(\rho_l)^m}{mn} + \sum_{m=2}^{\infty}\sum_{n=1}^{m-1}\frac{(\rho_l)^m}{mn}\right) \\
&= \frac{1}{1-\rho_l}\left(2\cdot\sum_{m=1}^{\infty}\sum_{n=1}^{m}\frac{(\rho_l)^m}{mn} - \sum_{m=1}^{\infty}\frac{(\rho_l)^m}{m^2}\right)
\end{aligned} \tag{2.106}$$

For convenience, we let:

$$f_{c_2}(\rho_l) = 2\cdot\sum_{m=1}^{\infty}\sum_{n=1}^{m}\frac{(\rho_l)^m}{mn} - \sum_{m=1}^{\infty}\frac{(\rho_l)^m}{m^2} \tag{2.107}$$

By Taylor series expansion of logarithm function, we have the closed form expression of derivative of $f_{c_2}(\rho_l)$:

$$f'_{c_2}(\rho_l) = -2\cdot\frac{\log(1-\rho_l)}{\rho_l} + \frac{\log(1-\rho_l)}{\rho_l(1-\rho_l)} \tag{2.108}$$

Through the definition of dilogarithm function in (2.29), we can show that:

$$f_{c_2}(\rho_l) = \pounds(\rho_l) + (\log(1-\rho_l))^2 \tag{2.109}$$

Finally, we have $c_2$,

$$c_2 = \frac{\pounds(\rho_l) + (\log(1-\rho_l))^2}{1-\rho_l} \tag{2.110}$$

Substituting the values of $a_2$, $b_2$ and $c_2$ to (2.103), we see that (2.97) holds. Hence, lemma 2.3.3 is proved.

## 2.D Proof of Lemma 2.3.4

The correlation coefficient $\rho_l$ is defined as:

$$\rho_l = \frac{Æ[A_n A_{n-l}] - Æ[A_n]Æ[A_{n-l}]}{\sqrt{\mathrm{var}(A_n)\,\mathrm{var}(A_{n-l})}}. \tag{2.111}$$

Easily we know $Æ[A_n] = 1$ and $Æ[A_{n-l}] = 1$, so the only unknown term is $Æ[A_n A_{n-l}]$. Noting that $A_n = h_n h_n^*$, we have:

$$Æ[A_n A_{n-l}] = Æ[h_n h_n^* h_{n-l} h_{n-l}^*] \tag{2.112}$$

By the Inverse DFT fading generation process discussed in Section 2.2, the fadings $h_n$ correspond to:

$$h_n = \frac{1}{\sqrt{N}} \sum_{a=1}^{N} h_a S_a^{\frac{1}{2}} \exp(j2\pi \frac{a}{N} \cdot n), \tag{2.113}$$

where $h_a$ denotes multiple i.i.d. complex Gaussian variables, and $S_a^{\frac{1}{2}}$ is the square root of the discrete sample of the PSD $S(f)$, and for large $N$ satisfies:

$$\sum_{a=1}^{N} S_a = N \tag{2.114}$$

Substituting (2.113) into (2.112), we have the covariance between two joint exponential variables:

$$\begin{aligned} &Æ[A_n A_{n-l}] \\ &= \frac{1}{N^2} \sum_{a=1}^{N} \sum_{b=1}^{N} \sum_{c=1}^{N} \sum_{d=1}^{N} Æ[h_a h_b^* h_c h_d^*] S_a^{\frac{1}{2}} S_b^{\frac{1}{2}} S_c^{\frac{1}{2}} S_d^{\frac{1}{2}} \exp\left(j2\pi\left((a-b)\cdot\frac{n}{N} + (c-d)\cdot\frac{n-l}{N}\right)\right). \end{aligned} \tag{2.115}$$

Note:

$$Æ[h_a h_b^* h_c h_d^*] = \begin{cases} 1 & \text{if } a = b \text{ and } c = d \text{ and } a \neq c \\ 1 & \text{if } a = d \text{ and } b = c \text{ and } a \neq c \\ 2 & \text{if } a = b = c = d \\ 0 & \text{otherwise} \end{cases} \tag{2.116}$$

, Hence, (2.112) corresponds to:

$$\begin{aligned} Æ[A_n A_{n-l}] &= \frac{1}{N^2}\left(\sum_{a=1}^{N}\sum_{\substack{c=1 \\ a\neq c}}^{N} S_a S_c + \sum_{a=1}^{N}\sum_{\substack{c=1 \\ a\neq c}}^{N} S_a S_c \exp\left(j2\pi\frac{a-c}{N}\cdot l\right) + 2\sum_{a=1}^{N}(S_a)^2\right) \\ &= \frac{1}{N^2}\left(\sum_{a=1}^{N}\sum_{c=1}^{N} S_a S_c + \sum_{a=1}^{N}\sum_{c=1}^{N} S_a S_c \exp\left(j2\pi\frac{a-c}{N}\cdot l\right)\right) \end{aligned} \tag{2.117}$$

Substituting (2.117) into (2.111), the final correlation coefficient corresponds to:

$$\rho_l = R_l R_l^*, \tag{2.118}$$

where

$$R_l = \frac{1}{N}\sum_{a=1}^{N} S_a \exp\left(j2\pi\frac{a}{N}\cdot l\right). \tag{2.119}$$

Hence, the lemma 2.3.4 is proved.

A remarkable property of $\rho_l$ is that it is always positive. This is in fact expected: Considering we have a large fading amplitude at one time, at other time we would expect a large fading but not a small fading amplitude; hence, the correlation coefficient is always positive.

## 2.E Proof of Theorem 2.3.5

First of all, we define the norm operation of a matrix:

$$|\mathbf{A}| = \sum_{a=1}^{N}\sum_{b=1}^{N}(\mathbf{A})_{a,b} \tag{2.120}$$

where $(\mathbf{A})_{a,b}$ denote the element of matrix $\mathbf{A}$ at $a^{th}$ row and $b^{th}$ column. Note this operation is not the Frobenius norm operation, here we define the norm simply as the summation of all elements in the matrix.

Now, we can write the variance of the normalized mutual information as:

$$\mathrm{var}\left(\frac{1}{N}\sum_{n=1}^{N}\log A_n\right)=\frac{1}{N^2}\left(N\cdot\pounds(\rho_0)+2\sum_{l=1}^{N-1}(N-l)\cdot\pounds(\rho_l)\right)=\frac{|\mathbf{L}|}{N^2}, \quad (2.121)$$

where $\mathbf{L}$ is defined as:

$$\mathbf{L}=\begin{bmatrix}\pounds(\rho_0) & \pounds(\rho_1) & \cdots & \pounds(\rho_{N-1})\\ \pounds(\rho_1) & \pounds(\rho_0) & \cdots & \pounds(\rho_{N-2})\\ \vdots & \vdots & \ddots & \vdots\\ \pounds(\rho_{N-1}) & \pounds(\rho_{N-2}) & \cdots & \pounds(\rho_0)\end{bmatrix}. \quad (2.122)$$

Using the series representation of the dilogarithm function,

$$\pounds(\rho_l)=\sum_{k=1}^{\infty}\frac{(\rho_l)^k}{k^2}, \quad (2.123)$$

we can write the norm of $\mathbf{L}$ in another form:

$$|\mathbf{L}|=\sum_{k=1}^{\infty}\frac{|\mathbf{L}_k|}{k^2}, \quad (2.124)$$

in which $\mathbf{L_k}$ corresponds to:

$$\mathbf{L}_k=\begin{bmatrix}(\rho_0)^k & (\rho_1)^k & \cdots & (\rho_{N-1})^k\\ (\rho_1)^k & (\rho_0)^k & \cdots & (\rho_{N-2})^k\\ \vdots & \vdots & \ddots & \vdots\\ (\rho_{N-1})^k & (\rho_{N-2})^k & \cdots & (\rho_0)^k\end{bmatrix}. \quad (2.125)$$

In appendix 2.D, we have shown that $\rho_l = R_l R_l^*$ , therefore, we can write:

$$|\mathbf{L}_k|=\mathrm{tr}(\mathbf{R}_k\mathbf{R}_k^\dagger)=\sum_{n=1}^{N}\lambda_n(\mathbf{R}_k\mathbf{R}_k^\dagger) \quad (2.126)$$

where $\lambda_n(\cdot)$ denotes the $n^{th}$ eigen value of the corresponding matrix, and $\mathbf{R}_k$ is defined as:

$$\mathbf{R}_k=\begin{bmatrix}(R_0)^k & (R_1)^k & \cdots & (R_{N-1})^k\\ (R_1)^k & (R_0)^k & \cdots & (R_{N-2})^k\\ \vdots & \vdots & \ddots & \vdots\\ (R_{N-1})^k & (R_{N-2})^k & \cdots & (R_0)^k\end{bmatrix}. \quad (2.127)$$

In general, the eigen values of $\mathbf{R}_k$ is not easy to be computed via the PSD $S(f)$. However, for large values of $N$, some nice mathematical results are available already.

The asymptotic behavior of the product of Toeplitz matrices are well studied in [Grenander and Szegio(1958)] and recently summarized in an engineering manner in [Gray(1977)]. Here, we refer to Theorem 5.3 in [Gray(1977)].

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \lambda_n(\mathbf{R}_k \mathbf{R}_k^*) = \int_{-0.5}^{0.5} \left(S_k(f)\right)^2 \, \mathrm{df} \tag{2.128}$$

where $S_k(f)$ satisfies:

$$(R_l)^k = \int_{-0.5}^{0.5} S_k(f) \exp\left(j2\pi \frac{l}{N} \cdot f\right) \, \mathrm{df}. \tag{2.129}$$

In other words, $S_k(f)$ is the spectrum corresponding to the auto-covariance matrix $\mathbf{R}_k$. We note that the elements of $\mathbf{R}_k$ is the $k^{th}$ power of elements of $\mathbf{R}$, and we know that time-domain multiplication is equivalent to the frequency-domain convolution. In other words, it is noticed that $bR_k$ is multiplication of $bR_{k-1}$ and $\mathbf{R}_1$, then its spectral, $S_k(f)$ would corresponds to convolution of $S_{k-1}(f)$ and $S_1(f)$, i.e.,:

$$S_k(f) = S_{k-1}(f) * S(f), \tag{2.130}$$

where $*$ denotes the cyclic convolution, and $S_1(f) = S(f)$.

Combining the results in (2.130), (2.128) and (2.126), Theorem 2.3.5 is proved.

# Chapter 3

# Perturbation Analysis in SIMO Channels

This chapter[1] investigates the Probability-of-Detection (POD) performance of a novel Wireless Local Positioning System (WLPS[2]) realized via Direct Sequence Code Division Multiple Access (DS-CDMA) and beamforming techniques. The proposed WLPS has unique signaling schemes that discriminate it from the traditional wireless systems and allows the WLPS to have many civilian and military applications. The WLPS consists of two main parts: (1) the detecting unit, a base station carried by a mobile unit defined as Dynamic Base Station (DBS), and (2) the being detected unit, a transponder (TRX) that is mounted on the targets, each assigned a unique identification code (ID code). Each DBS should be capable of detecting and locating all available TRXs in its coverage area. As a result, the main complexity of this system is focused at the DBS receiver.

In this work, we introduce WLPS structure, and both theoretically and numerically compute, and compare the Probability-of-Detection performance of the DBS receiver realized by a merger of DS-CDMA and antenna arrays. We also analyze and simulate the performance of this system under perturbations in the DBS an-

[1]Portions of this chapter were previously reported in [Zekavat(2003)] [Zekavat et al.(2004)] [Tong and Zekavat(2005)].

[2]The WLPS US Patent is pending by Michigan Tech. University.

tenna array vector. The simulations are particularly performed for vehicular collision avoidance applications.

## 3.1 Introduction

Historically, positioning was developed for navigational purposes with a wide variety of civilian and military applications, and they fall into two main categories, Global Positioning Systems (GPS) [Getting(1993)] [Grant(1988)] and Local Positioning Systems (LPS) [Werb and Lanzl(1998)] . GPS is a precise, all-weather, 24 hour satellite based positioning system mainly developed for direction finding and navigation [Bernard(1992)]. However, GPS has the following problems: (1) its signal does not penetrate into the buildings; hence, it does not perform at indoor areas, (2) it looses its precision in rich scattering environments such as urban areas, (3) it is mainly suitable for navigation and for tracking or command purposes, it should be merged with a communication system for transmission of position information from the GPS to a center (e.g., command center in defense applications), and (4) it is yet expensive.

Local positioning systems (LPS) fall into two main categories [Vossiek et al.(2003)]: (1) Self Positioning: A mobile device finds its own instantaneous location with respect to a fixed point, e.g., the starting point or a beacon node, and (2) Remote Positioning: A mobile device finds the instantaneous positions of other objects (mobiles) with respect to its own position. The Wireless Local Positioning System (WLPS) introduced in this chapter is a remote positioning system with active targets capable of detecting and locating targets within several hundred meters in a dynamic environment [Zekavat(2003)] [Zekavat et al.(2004)] [Tong and Zekavat(2005)].

In the quest for wireless positioning, different systems have been developed or are under development. For example, radar systems are used to find the position of targets in the surrounding areas via transmission of a short burst of energy and processing its reflection from the targets [Skolnik(1981)] [Sekine et al.(1992)]. The ability of

radars to detect the desired targets is hindered by clutters or reflections from undesirable objects and interfering radars, which are inevitable in typical indoor and urban areas, rendering radar systems impractical [Hermann et al.(2001)] [Zekavat(2003)]. Another example is a vision system that uses video signals collected from a camera to recognize targets and estimate positions [Saneyoshi(1996)]. Such systems, including Chrysler's mobile positioning system, face major limitations at night and in severe weather conditions such as intense rain, snow and fog.

The Wireless Local Positioning System (WLPS) introduced in this chapter consists of: (1) a base station in each monitoring mobile, which serves as a non-static or Dynamic Base Station (DBS), and (2) a transponder (TRX) in target mobiles, which acts as active targets. Unique IDentification (ID) codes are assigned to each target. DBS transmits a short pulse containing an ID Request (IDR) signal to all targets located in its vicinity, and it does not transmit within two consecutive IDRs called ID request Repetition Time ($IRT$). The targets respond to that signal by transmitting their ID codes back to the DBS. DBS recognizes each target by its ID code, and then positions, tracks and monitors those targets. Positioning is realized via calculating the Time-of-Arrival (TOA) and the Direction-of-Arrival (DOA). TOA is defined as the time difference between transmission of the ID request signal and reception of the corresponding TRX ID.

The performance of WLPS system depends on two main variables: a) Probability-of-Detection (POD) of the ID of each TRX, and b) accuracy of positioning, which is a function of the estimation of TOA and DOA. The positioning performance of WLPS primarily depends on the development of TOA/DOA estimation techniques. In general, TOA can be estimated via RAKE receivers [Ramos et al.(1997)]; DOA can be estimated via super resolution spatial spectral estimation techniques [Al-Ardi et al.(2004)]; or DOA/TOA can be estimated jointly [Wang et al.(2001)].

Novel DOA estimation techniques have been designed for WLPS and be presented in [Pourrostam et al.(2007)].

In this chapter, we focus on the ID detection process of the WLPS which involves a study on the Probability-of-Detection performance. WLPS uses TRX ID IDs to distinguish and track the TRXs in its coverage area. Moreover, some WLPS applications may only need ID detection but not TOA/DOA estimations: For example, as an application of WLPS in stop-free toll stations, when a driver passes by a toll station, the monitoring system (DBS) would just record the vehicle's ID (TRX) for billing purposes.

Many local positioning systems via active targets have been introduced in the literature. Examples include: (1) Airborne traffic alert and collision avoidance (TCAS) systems [Williamson and Spencer(1989)] [Kahne and Frolow(1996)] developed for future air navigation systems (FANS) [Trim(1990)] use transponders at airplanes (active targets) for positioning purposes. These systems use the radar principles for range resolution with a range of 10-40 miles [Ractliffe(1990)] which leads to a limited capacity suitable for those applications. The technique used in TCAS is not feasible for wireless channels, which experience multi-path fading and interference effects, and cover a range of 0.1-1 mile in many applications; (2) Cell phone positioning systems may exploit a triangulation technique [Cedervall(1998)] to estimate subscriber's position. In order to achieve a reasonable positioning accuracy (around $100m$), the subscriber's signal can be received by at least three base stations. But, this number of base stations might not be available at all times. In addition, experimental results show that the capability of this system is limited by multipath environments [Hepsaydir(1999)]. A merger of area power intensity map and directional antennas may also be utilized for cell phone positioning purposes [Kikuchi et al.(2003)]. This approach requires detailed information that

constrains to static (as opposed to dynamic) positioning; (3) In tagged local positioning systems (TLPS), mobiles (active targets) transmit periodic signals at all times, and static base stations receive those signals and locate the targets via a triangulation technique [Werb and Lanzl(1998)]. In order to increase its robustness to multipath effects and increase its accuracy, this system utilizes additional readers and reference tags. These systems are limited to a relatively fixed environment [Ni et al.(2003)]; and (4) In WLAN positioning systems, a triangulation technique is used by mobile nodes for positioning purposes. In order to perform triangulation, WLAN positioning requires a number of nodes to be involved in the positioning process [Singh et al.(2004)]. Hence, positioning process can not be performed independently by each node, which limits the applications of these positioning systems.

In the WLPS introduced in this chapter, each DBS is capable of positioning all TRXs in its coverage area. Moreover, the DBS does not need additional prior environmental information. Hence, WLPS is not limited to the fixed base station. In addition, via the application of wireless signaling schemes such as Code Division Multiple Access (CDMA), diversity combining and beamforming, the interference and multipath fading effects can be reduced and the probability of detection increases. The WLPS can be defined as a node in Mobile Adhoc Network (MANET) with a variety of applications in security systems (e.g., in indoor security via implementation of DBS and TRX on security guards and just TRX on people entering the building), vehicle collision avoidance system and multi robot control (e.g., by implementing the DBS and TRX on all vehicles or robots [Zekavat(2003)]), and defense (e.g., for command control and tracking, by implementation of DBS on the central command and control, DBS and TRX on all commanders and TRX on soldiers). The coverage area of WLPS can be increased via multiple-hop localization techniques [Zekavat et al.(2004)] [Li et al.(2000)] [Gharavi and Ban(2003)].

In this work, we investigate the realization of this system via standard transmitters and receivers (i.e., simple modulators and demodulators) as well as Direct Sequence CDMA (DS-CDMA) receivers with and without antenna arrays and beamforming (BF) techniques, and we compare their Probability-Of-Detection, $P_d$, performance. In DS-CDMA the transmitted symbol is multiplied in a spreading code in the time domain. Orthogonal spreading codes maintain the orthogonality between the signals transmitted from different TRXes. In addition, multipath fading effects are reduced by exploiting path diversity at the DS-CDMA receiver. This technique alone enhances the Probability-of-Detection (POD) performance of the DBS receiver. A merger of DS-CDMA with Spatial Division Multiple Access (SDMA) highly enhances the performance of the DBS. SDMA is accomplished by employing directionality via antenna arrays at the DBS receiver. These antenna arrays are required for DOA estimation as well. Our simulation depicts that by a proper selection of $IRT$, both DS-CDMA and standard Rake receivers with conventional BF lead to a high POD performance. Here, we particularly perform the simulations for vehicular collision avoidance (road safety) applications of WLPS.

In real applications, perturbation of the array weight vector due to effects such as sensor position errors, gain errors, phase errors, mutual coupling between sensors, imperfect channel phase estimation, and different cable length for each sensor degrades the receiver performance [Yang and Swindlehurst(1995)] [Godara(1997)]. Synthesizing all these effects leads to an additive random Gaussian error on the array weight vector [Wax and Anu(1997)]. In this work, we also study the performance degradation (POD and capacity) caused by these perturbation effects.

The chapter is organized as follows: Section II introduces the WLPS structure. Section III discusses the theoretical analysis of the system performance for different system configurations as well as the perturbation effect. Section IV represents the simulation and analytical results. Section V concludes the chapter. A table of abbreviations used throughout this chapter is provided in Table 3.1.

Table 3.1: Table of Abbreviations

| | |
|---|---|
| BF | Beamforming |
| DBS | Dynamic Base Station |
| DOA | Direction of Arrival |
| DS-CDMA | Direct Sequence Code Division Multiple Access |
| GPS | Global Positioning System |
| IDR | ID Request Signal |
| IRT | ID Request Repetition Time |
| LPS | Local Positioning System |
| POD | Probability-of-Detection |
| RCVR | Receiver |
| SDMA | Spatial Division Multiple Access |
| TOA | Time Of Arrival |
| TRX | Transponder |
| WLPS | Wireless Local Positioning System |

## 3.2 A Novel Wireless Local Positioning System

The two WLPS main parts include: A dynamic base station (DBS) and a transponder (TRX). The DBS transmitter generates an ID code request (IDR) signal every $IRT$ (ID request repetition time) to all TRXs in the coverage area; then, it waits to receive a response back from the TRX within $IRT$ (see Fig. 3.1). TRX transmits a unique ID code as soon as it detects the IDR signal transmitted by the DBS. The ID code is selected from simple pseudo random codes which consist of +1 and -1. Hence, the number of bits in the code depicts the maximum capacity of the WLPS. Depending on the application, the ID code can be assigned permanently or can be assigned by the DBS transmitter. The structure of DBS is shown in Fig. 4.1.

In a WLPS structure, each DBS communicates with a number of TRX in its coverage area simultaneously. This is the same as usual cellular communication systems. However, in contrast to cellular systems, in WLSP each TRX communicates with a number of DBS simultaneously as well. In addition, the time of transmission and reception would be different at the DBS, as shown in Fig. 3.1. Moreover, the whole DBS and TRX use different transmission frequencies. Thus, the overall system

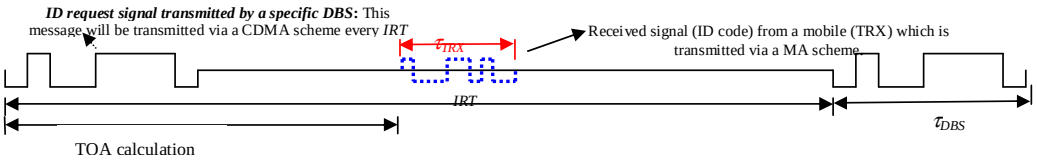


Figure 3.1: Transmission of IDR and reception from TRX in DBS. Assuming a Pseudo Random ID codes, the number of the bits in the code represents the maximum capacity of the WLPS.

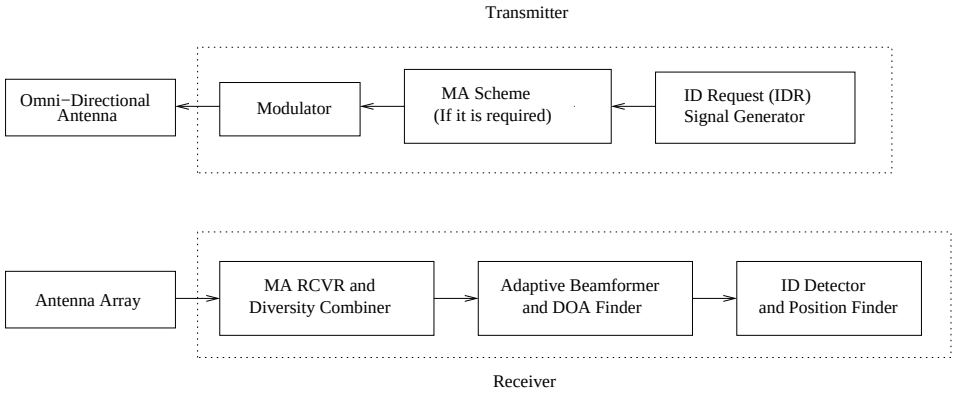


Figure 3.2: DBS structure

is considered as a time division duplex (TDD)-frequency division duplex (FDD), that is, hybrid TDD/FDD communication system (differ from cellular systems that are either TDD or FDD). This allows WLPS to reduce the interference effects via a proper selection of $IRT$.

The minimum allowable value for $IRT$ ($IRT_{min}$) is calculated to avoid range ambiguity: If the response to each IDR signal is received in DBS within $IRT$, mobile range is calculated correctly; however, if it is received after the next IDR transmission, the range is not correctly calculated. Here, $IRT$ is a function of the maximum coverage or the maximum range $R_{max}$. The minimum allowable $IRT$ corresponds to:

$$IRT_{min} = 2T_{max} + T_d + T_G, \tag{3.1}$$

where $T_{max}$ denotes the maximum possible time delay between the TRX transmission and the DBS reception, $T_d$ is the TRX time delay in responding to the IDR signal which determines the minimum time before receiving the first signal back from a TRX, and $T_G$ is the guard band time to avoid range ambiguity, corresponds to [Zekavat(2004)]:

$$T_G = 5T_m + \tau_{DBS} + \tau_{TRX}. \tag{3.2}$$

Here, $T_m$ is the wireless channel delay spread, $\tau_{DBS}$ and $\tau_{TRX}$ are the durations of DBS and TRX transmitted signals, respectively. Considering the maximum uplink antenna array half power beam widths (HPBW) $\beta$ to be less than 90°, using a simple geometry in a scattering environment, $T_{max}$ is determined by $R_{max}$ and $\beta$ via [Zekavat(2004)]:

$$T_{max} = \left(\frac{R_{max}}{2c}\right) \cdot \frac{1+\cos\beta}{\cos\beta}, \tag{3.3}$$

where $c$ denotes the speed of light. Eq. (3.1) refers to a lower limit for $IRT$ (i.e., $IRT_{min}$). The upper limit for $IRT$ ($IRT_{max}$) is a function of the speed of moving TRX and DBS, and accordingly, the required processing speed.

The TRX receiver is subject to inter-DBS interference (IBI), since more than one DBS may transmit IDR signals in the coverage area of a TRX. Large selection of $IRT$ reduces the probability-of-overlap, $p_{ovl}$, or collision of the DBSs transmitted signals at the TRX receiver. In addition, a number of TRXs in the DBS coverage area respond to the IDR signal of one DBS simultaneously, causing inter-TRX-interference (IXI) at the DBS receiver. Both IXI and IBI are functions of the $p_{ovl}$ for the transmitted signals from TRXs and DBSs, respectively. $p_{ovl}$ has a profound effect on the performance of the receiver, is a function of the number of mobiles or transmitters (DBS or TRX), $K$, in their coverage area, and corresponds to:

$$p_{ovl} = 1 - (1 - d_c)^{K-1}, \tag{3.4}$$

where the duty cycle, $d_c$, defined as $d_c = \tau/T$, $\tau = \tau_{DBS}(\tau_{TRX})$ is the duration of DBS (TRX) transmitted signal, and $T = T_{DBS}(T_{TRX})$, where $T_{DBS} = IRT_{min}$, $T_{TRX} = IRT$.

In general, large selection of $IRT$ reduces IBI effects at the TRX receiver and highly enhances its POD performance. However, the large selection of $IRT$ does not affect the IXI, since all of the signals are received by DBS receiver within the $T_{max}$ time frame, which is mainly a function of the maximum coverage range. It is also worth mentioning, each TRX located in the coverage area of more than one DBS may generate ID codes in response to more than one DBS within each $IRT$. This leads to both IXI as well as range ambiguity. Range ambiguity can be resolved via changing code assignments (MA codes, ID codes or both) for different DBS. In addition, in Eq. (4.1), the parameter $d_c$ is a function of $\tau_{TRX}$ and $\tau_{DBS}$, so as the probability-of-overlap and the POD performance of WLPS.

In general, selection of $\tau_{DBS}$ and $\tau_{TRX}$ depends on the POD, desired system capacity (in terms of the number of TRX/DBS accommodated), bandwidth, positioning accuracy and maximum coverage range, and may vary with WLPS application. The duration of the transmitted signal by the DBS ($\tau_{DBS}$) and TRX ($\tau_{TRX}$) should be much smaller than the $IRT$ to reduce $p_{ovl}$ among signals received by receivers of TRX and DBS, respectively. A smaller $p_{ovl}$ decreases both the IBI effects (at TRX) and IXI effects (at DBS), which in turn enhances the POD performance and the capacity of the WLPS system. On the other hand, the system maximum capacity expressed by the maximum number of TRX (DBS) determines the number of bits within each ID code, which is to be transmitted over a period of $\tau_{TRX}$ ($\tau_{DBS}$) . The required bandwidth is inversely proportional to $\tau_{TRX}$ and $\tau_{DBS}$ for a given capacity. A large selection of $IRT$ allows $\tau_{DBS}$ to be selected much larger than $\tau_{TRX}$ without sacrificing $p_{ovl}$ at the TRX receiver. Hence, WLPS bandwidth is mainly determined by the value of $\tau_{TRX}$ (see Section 4.5, the simulation results).

## 3.3 WLPS Implementation And Theoretical Analysis

As discussed in Section 4.2, in general, the interference effect IXI (IBI) at the DBS (TRX) receiver can be mitigated via selecting $d_c$ small enough. Large selection of $IRT$ reduces the $d_c$ and consequently the IBI at the TRX receiver; however, large selection of $IRT$ does not have any effect on the IXI at the DBS receiver. The $d_c$ can also be reduced via small selection of $\tau$ ($\tau_{DBS}$ or $\tau_{TRX}$), that in turn enhances the required bandwidth. Hence, while a standard receiver ensures TRX high performance for a multi-user environment, the DBS performance is improved just via MA schemes. We start with the theoretical investigation of performance for standard receivers and then we continue the discussion for DS-CDMA schemes. The theoretical results discussed here can be equivalently applied to both TRX and DBS receivers. The details of derivations of equations presented in this section are provided in Appendices 3.A-3.C.

### 3.3.1 Standard Receiver System

Assuming a standard receiver at the DBS (TRX), the transmitted signal from the TRX (DBS) corresponds to (see Fig. 3.1):

$$s^k(t) = \sum_{n=0}^{N-1} b^k[n] g_{T_b}(t - nT_b) \cdot \cos(2\pi f_c t), \tag{3.5}$$

where $N$ denotes the number of bit per ID code (that represents the maximum capacity of the WLPS), $b^k[n]$ denotes the $n^{th}$ bit of user $k$'s ID, $T_b = \tau/N$ represents the DBS(TRX) bit duration where $\tau = \tau_{TRX}(\tau_{DBS})$; $g_{T_b}(t)$ are rectangular pulses with the duration $\tau$ and $T_b$.

Assuming a frequency selective channel, the received signal $r(t)$ at the DBS(TRX) receiver is a mixture of signals from different TRXs(DBSs) and different paths, which is given by:

$$r(t) = \sum_{k=1}^{K} \sum_{l=0}^{L^k-1} \sum_{n=0}^{N-1} \alpha_l^k b^k[n] \cdot g_{T_b}(t - \tau_l^k - nT_b) \cdot \cos(2\pi f_c t + \phi_l^k) + n(t), \tag{3.6}$$

where $K$ denotes the total number of TRXs(DBSs), $L^k$ is the number of paths for TRX(DBS) $k$, and $\alpha_l^k$, $\tau_l^k$, $\phi_l^k$ denote the fading factor, time delay and random phase for $k^{th}$ user's $l^{th}$ path, respectively.

After the demodulation, the $n^{th}$ bit output for $j^{th}$ DBS's (TRX's) $q^{th}$ path corresponds to:

$$y_q^j[n] = \int_{\tau_q^j + nT_b}^{\tau_q^j + (n+1)T_b} r(t)\cos(2\pi f_c t + \phi_q^j)\, dt, \tag{3.7}$$

Assuming all TRXs(DBSs) have the same number of paths, i.e., $L^k = L, \forall k$, and fading energy is uniformly distributed in paths, i.e., $E[(\alpha_q^j)^2] = 1/L$, the instantaneous SINR in $j^{th}$ user's $q^{th}$ path, i.e., for any path of any TRX(DBS) is (see Appendix 3.B):

$$r_i = \frac{A_a}{D_a \cdot (K-1) + D_a \cdot (1 - \frac{1}{L}) + \frac{1}{\bar{r}_0}} \cdot \frac{2}{L}, \tag{3.8}$$

where $\bar{r}_0$ is the average SNR, which is defined as the ratio between average bit energy and white noise. In this case,$A_a = 1$, $D_a = d_c$ , and $d_c$ denotes the duty cycle [see Eq. (4.1)].

### 3.3.2 Standard Receiver merger with Conventional Beamforming

BF techniques reduce the signal from other users and other paths as long as they are in different direction from the desired user and path. In this case, with the same transmitted signal as Eq. (3.5), the received signal at the DBS(TRX) with antenna arrays corresponds to:

$$\vec{r}(t) = \sum_{k=1}^{K} \sum_{l=0}^{L^k-1} \sum_{n=0}^{N-1} \alpha_l^k \cdot \vec{V}(\theta_l^k) \cdot b^k[n] \cdot g_{T_b}(t - \tau_l^k - nT_b) \cdot \cos(2\pi f_c t + \phi_l^k) + \vec{n}(t), \tag{3.9}$$

where $\vec{V}(\theta_l^k)$ denotes the array response vector and corresponds to:

$$\vec{V}(\theta_l^k) = \left[1 \quad exp(j \cdot \frac{-2\pi d\cos(\theta_l^k)}{\lambda}) \quad \cdots \quad exp(j \cdot \frac{-2(M-1)\pi d\cos(\theta_l^k)}{\lambda})\right]^T. \tag{3.10}$$

Here, $d$ is the spacing between antenna elements, $M$ is the total number of antennas, $\lambda$ denotes the carrier wavelength and $\theta_l^k$ is the direction of $k^{th}$ user's $l^{th}$ path.

After BF and demodulation, the $n^{th}$ bit for $j^{th}$ user's $q^{th}$ path corresponds to:

$$y_q^j[n] = \vec{W}^H(\theta_q^j) \int_{\tau_q^j+nT_b}^{\tau_q^j+(n+1)T_b} \vec{r}(t)\cos(2\pi f_c t + \phi_q^j)\,dt, \tag{3.11}$$

where $\vec{W}(\theta_q^j) = \vec{v}(\theta_q^j)$ if no perturbation presents, and $H$ denotes Hermition transpose. $\vec{W}(\theta_q^j)$ with perturbation is explained in section 3.3.5.

In this case, the SINR for any path of any user is (see Appendix 3.B):

$$r_i = \frac{A_b}{B_b \cdot D_b \cdot (K-1) + B_b \cdot D_b \cdot (1-\frac{1}{L}) + \frac{M}{\bar{r}_0}} \cdot \frac{2}{L}, \tag{3.12}$$

where $A_b = M^2$, $D_b = d_c$, and

$$B_b = \sum_{m=0}^{M-1}(m+1)J_0(\frac{2\pi dm}{\lambda})J_0(\frac{-2\pi dm}{\lambda}) + \sum_{m=M}^{2M-2}(2M-m-1)J_0(\frac{2\pi dm}{\lambda})J_0(\frac{-2\pi dm}{\lambda}). \tag{3.13}$$

Here, $J_0$ represents the zero$^{th}$ order Bessel function of the first kind.

### 3.3.3 The DS-CDMA System

The transmitted DS-CDMA signal by the $k^{th}$ TRX (DBS) corresponds to:

$$s^k(t) = \sum_{n=0}^{N-1} b^k[n] \cdot g_{T_b}(t-nT_b) \cdot a^k(t-nT_b) \cdot \cos(2\pi f_c t), \tag{3.14}$$

where $a^k(t) = \sum_{i=0}^{G-1} C_i^k g_{T_c}(t-iT_c)$, $C_i^k \in \{-1, 1\}$, denotes the spreading code, and $G$ is the processing gain (code length), $T_c = \tau/(N \cdot G)$ , $\tau = \tau_{TRX}(\tau_{DBS})$, represents the chip duration and, $g_{T_c}(t)$ is a rectangular pulse with the duration of $T_c$.

The received signal corresponds to:

$$r(t) = \sum_{k=1}^{K}\sum_{l=0}^{L^k-1}\sum_{n=0}^{N-1} \alpha_l^k b^k[n] \cdot g_{T_b}(t-\tau_l^k-nT_b) \cdot a^k(t-\tau_l^k-nT_b) \cdot \cos(2\pi f_c t + \phi_l^k) + n(t). \tag{3.15}$$

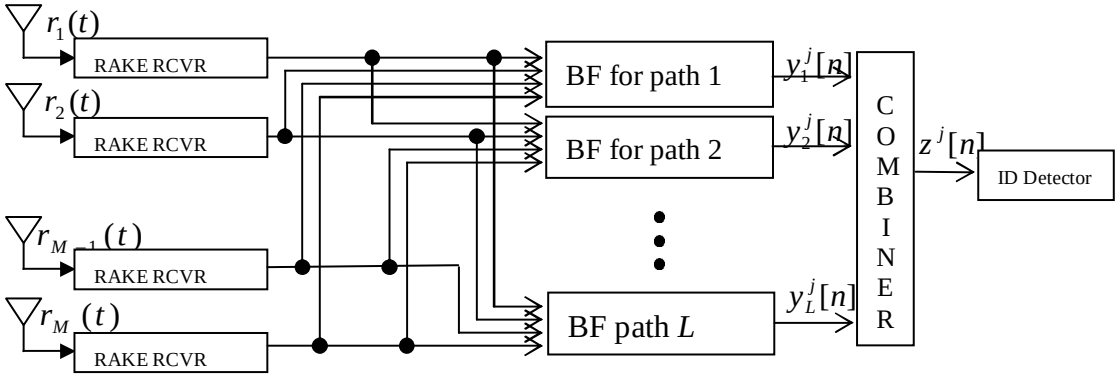


Figure 3.3: DBS(TRX) receiver structure assuming a frequency selective channel and smart antennas with DS-CDMA

After despreading, the $n^{th}$ bit output of $j^{th}$ DBS's(TRX's) $q^{th}$ path corresponds to:

$$y_q^j[n] = \int_{\tau_q^j+nT_b}^{\tau_q^j+(n+1)T_b} r(t) \cdot \cos(2\pi f_c t + \phi_q^j) \cdot a^k(t - \tau_q^j - nT_b)\, dt. \tag{3.16}$$

Finally the SINR for any path of any DBS(TRX) is (see Appendix 3.B):

$$r_i = \frac{A_c}{D_c \cdot (K-1) + D_c \cdot (1 - \frac{1}{L}) + \frac{G}{\bar{r}_0}} \cdot \frac{2}{L}, \tag{3.17}$$

where $A_c = G$, and

$$D_c = d_c - \frac{N^2 - 3N + 3}{2N^2} \cdot d_c^2. \tag{3.18}$$

### 3.3.4 DS-CDMA Merger with Conventional Beamforming

With the same transmitted signal as in Eq. (4.4), the received signal at the DBS (TRX) for the antenna array (see Fig. 4.3), which is a mixture of signals from different TRXs(DBSs) and different paths, is given as:

$$\vec{r}(t) = \sum_{k=1}^{K} \sum_{l=0}^{L^k-1} \sum_{n=0}^{N-1} \alpha_l^k \cdot \vec{V}(\theta_l^k) \cdot b^k[n] \cdot g_{T_b}(t - \tau_l^k - nT_b) \cdot a^k(t - \tau_l^k - nT_b) \cdot \cos(2\pi f_c t + \phi_l^k) + \vec{n}(t). \tag{3.19}$$

The $i^{th}$ bit output of BF for $j^{th}$ user's $q^{th}$ path is given as:

$$y_q^j[n] = \vec{W}^H(\theta_q^j) \int_{\tau_q^j+nT_b}^{\tau_q^j+(n+1)T_b} \vec{r}(t) \cdot \cos(2\pi f_c t + \phi_q^j) \cdot a^k(t - \tau_q^j - nT_b)\, dt. \tag{3.20}$$

Finally, the SINR for any path of any user corresponds to (see Appendix 3.B:

$$r_i = \frac{A_d}{B_d \cdot D_d \cdot (K-1) + B_d \cdot D_d \cdot (1 - \frac{1}{L}) + \frac{MG}{\bar{r}_0}} \cdot \frac{2}{L}, \tag{3.21}$$

where $A_d = M^2 G$, $B_d = B_b$, and $D_d = D_c$.

### 3.3.5 Perturbation effects

Here, we investigate the effects of perturbation in sensor position, channel phase estimation, and the effects of mutual coupling between sensors, receiver fluctuation due to temperature and humidity, quantization effects, different cable length for each sensor, imperfect channel phase estimation, etc. According to Central Limit Theorem, the summation of these effects leads to additive zero mean Gaussian random variables along the elements in the array, i.e., the weighting vector $\vec{W}(\theta_l^k)$ in Eq. (3.16) and (4.9) does not match array response vector $\vec{V}(\theta_l^k)$ exactly, but corresponds to $\vec{W}(\theta_l^k) = \vec{V}(\theta_l^k) + \vec{\varepsilon}_l^k$ , where $\vec{E}$, called error vector, denotes a column vector that contains $M$ independent Gaussian random variables with zero mean and variance $\sigma_\epsilon^2$ [Rao and Jones(2001)].

With this error vector, the peak of the directional beam does not point to the desired user perfectly, but steered away from the desired user. Therefore, the power of the desired user is reduced. In addition, sidelobe's power of the directional beam increases; hence, power of IXI(IBI) increases. As a result, SINR is reduced by perturbation, which is a function of $\sigma_\epsilon^2$.

With perturbations, for standard receiver with antenna arrays, as introduced in Eq. (3.12), the SINR for any path of DBS(TRX) corresponds to (see Appendix 3.B):

$$r_i = \frac{A_e}{B_e \cdot D_e \cdot (K-1) + B_e \cdot D_e \cdot (1 - \frac{1}{L}) + \frac{M+\sigma_\epsilon^2}{\bar{r}_0}} \cdot \frac{2}{L},$$

where $D_e = d_c$, and, (see Appendix 3.C)

$$A_e = M^2 + \sigma_\epsilon^2 \sum_{m=0}^{M-1} J_0(2 \cdot \frac{2\pi dm}{2}), \tag{3.22}$$

$$B_e = \sum_{m=0}^{M-1}(m+1)J_0(\frac{2\pi dm}{\lambda})J_0(\frac{-2\pi dm}{\lambda}) + \sum_{m=M}^{2M-2}(2M-m-1)J_0(\frac{2\pi dm}{\lambda})J_0(\frac{-2\pi dm}{\lambda}) + \sigma_\epsilon^2 \sum_{m=0}^{M-1} J_0(2\cdot\frac{2\pi dm}{2}). \quad (3.23)$$

Similarly, for DS-CDMA system with antenna arrays, as introduced in Eq. (3.21), the SINR for any path of any user corresponds to (see Appendix 3.B):

$$r_i = \frac{A_f}{B_f \cdot D_f \cdot (K-1) + B_f \cdot D_f \cdot (1-\frac{1}{L}) + \frac{(M+\sigma_\epsilon^2)\cdot G}{\bar{r}_0}} \cdot \frac{2}{L},$$

where $D_f = D_e$, $B_f = B_e$ as introduced in Eq. (3.23), and (see Appendix 3.C)

$$A_f = G \cdot (M^2 + \sigma_\epsilon^2 \sum_{m=0}^{M-1} J_0(2\cdot\frac{2\pi dm}{2})). \quad (3.24)$$

### 3.3.6 Path Diversity Combining

Finally, for all of the receivers discussed in parts A-E, we apply Maximal Ratio Combining (MRC) across the path diversity components [Shah and Haimovich(2000)]:

$$z^j[n] = \sum_{l=1}^{L} \alpha_l^j y_l^j[n]. \quad (3.25)$$

Therefore, the final instantaneous SINR expression can be written as:

$$r_0 = r_e \cdot (\alpha_1^2 + \alpha_2^2 + \cdots + \alpha_L^2). \quad (3.26)$$

In Eq. (3.26), the parameter $r_e$ has been defined in Eq. (3.8), (3.12), (3.17), (3.21), corresponding to the receivers introduced in previous subsections. The Bit-Error-Rate for all of the discussed receivers corresponds to (see Appendix 3.A):

$$P_e = \int_0^\infty Q(\sqrt{2r_0}) f(r_0|\bar{r}_0)\, dr_0 = 0.5\left(1 - \sqrt{\frac{\bar{r}}{1+\bar{r}}} \cdot \sum_{l=0}^{L-1} \binom{2l}{l} \cdot \frac{1}{2^{2l}\cdot(1+\bar{r})^l}\right). \quad (3.27)$$

Here, $\bar{r} = L \cdot r_i$, which corresponds to a uniform power distribution over $L$ paths, $r_i$ has been introduced in (8), (12), (17) and (21).

For frequency-selective channel, $L$ is greater than one. For flat-fading channel, $L$ equals to one. If all bits are detected correctly, the ID of the desired user is detected correctly. Therefore, the probability-of-detection is given as:

$$P_d = (1 - P_e)^N, \tag{3.28}$$

and the probability-of-miss-detection corresponds to:

$$P_{md} = 1 - P_d. \tag{3.29}$$

### 3.4 Numerical Results And Analysis

In this section, we evaluate the POD performance and capacity (in terms of number-of-TRXs(DBSs)) of WLPS system under multi-TRX, multi-path environment, via simulations and we compare the results with the theoretical result of Eq. (3.28). This setup is typically useful for vehicle collision avoidance applications, where each vehicle is required to cover the front area. For simulation purposes, we assume:

1. The ID code has 6 bits ($N = 6$);
2. The DS-CDMA code has 64 chips ($G = 64$);
3. Channel delay spread, $T_m$ for a typical street area is $27\,nsec$ [Arowojolu et al.(1994)];
4. Carrier frequency $= 3\,GHz$, $\tau_{TRX} = 1.2\,\mu s$, and $\tau_{DBS} = 24\,\mu s$;
5. The antenna array is linear with 4 elements, and element spacing $d = \frac{\lambda}{2} = 0.05\,m$ ($HPBW = 27^\circ$);
6. Four multipaths lead to $L = 4$ fold path diversity for DS-CDMA system;
7. The TRX distance and angle are uniformly distributed in $[\,0 \quad 1\,]\,km$ and $[\,0 \quad \pi\,]$, respectively;

8. Uniform multi-path intensity profile, i.e., bit energy is distributed in each path identically;

9. Binary Phase Shift Keying (BPSK) modulation;

10. Perfect power control and DOA, TOA estimation; and,

11. The average Signal-to-Noise Ratio (SNR) introduced in Eq. (3.8), (3.12), (3.17) and (3.21) is $\bar{r}_0 = 20\,dB$.

12. The error variance, if perturbation presents, equals to 0.5, i.e., $\sigma_\epsilon^2 = 0.5$. This error variance means the standard deviation of perturbations is around 70% of elements of array response vector without perturbations.

Based on the assumed setup, then we deduce that TRX signal TOA is uniformly distributed in $[\,T_d \quad T_{max}\,]$, at the DBS receiver. Assuming $T_d \ll T_{max}$, the TRX signal TOA is approximately uniformly distributed in $[\,0 \quad T_{max}\,]$. In addition, the minimum $IRT$, $IRT_{min}$, is $32\,\mu sec$ [c.f., Eq. (3.1)]. We select a larger value $IRT = 24\,msec$ in order to reduce the IBI effects. It is worth mentioning that in the assumptions, $\tau_{DBS}$ is selected larger than $\tau_{TRX}$, since the IBI effect at the TRX receiver can be removed via large selection of $IRT$. Hence, a smaller bandwidth is required for the DBS transmitter. With the assumed $\tau_{DBS}$ and $\tau_{TRX}$, the required bandwidth of a DS-CDMA (standard) transmitter is $320MHz$ ($5MHz$) for TRX, and $16MHz$ ($250KHz$) for DBS, respectively. Hence, DBS required bandwidth is much smaller than the TRX and the WLPS bandwidth is mainly determined by the TRX transmission bandwidth, as expected. In addition, using these parameters, the duty cycle for DBS and TRX receivers correspond to $d_{c,DBS} \simeq 0.1$ and $d_{c,TRX} \simeq 0.001$.

As we mentioned earlier, the IBI at the TRX receiver can be considerably reduced by selecting the $IRT$ as large as possible; however, this selection will not affect

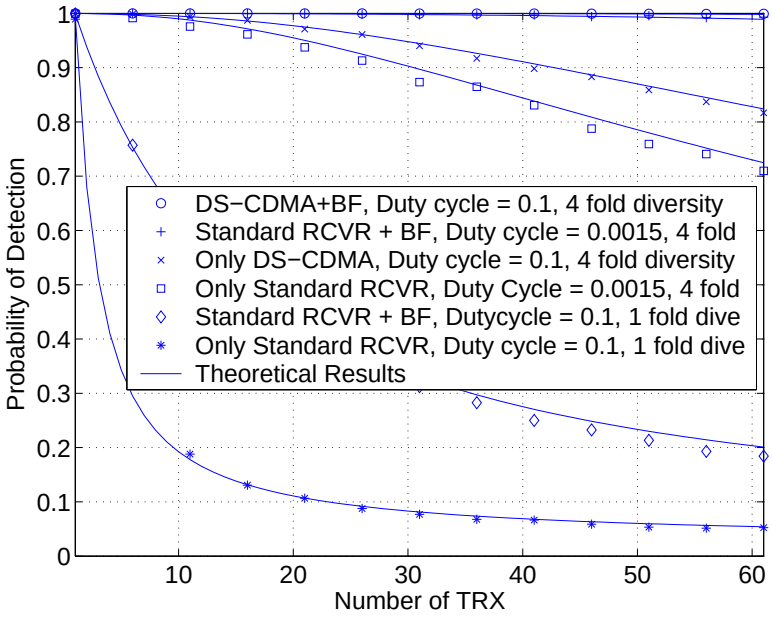


Figure 3.4: Simulation results for DBS receivers, without perturbation.

IXI at the DBS receiver. Hence, a TRX receiver can just be implemented by a simple transceiver (or DS-CDMA) system without employment of BF, while a DBS receiver needs a combination with BF. It should be mentioned that antenna arrays are required at the DBS receiver for DOA estimation as well. A small $d_{c,TRX} \simeq 0.001$ at the TRX receiver leads to a small $p_{ovl}$, which leads to small IBI and high POD. In contrast, a large $d_{c,DBS} \simeq 0.1$ at the DBS receiver leads to a high $p_{ovl}$ that results in high IXI. Both BF and CDMA techniques help to reduce the IXI effects at DBS.

The POD ($P_d$) of the DBS receiver is depicted in Fig. 3.4. This figure compares $P_d$ vs. the number of TRX for a standard transceiver and a DS-CDMA transceiver, with or without antenna arrays. It shows that in general the $P_d$ decreases as the number of TRX increases, which is a direct result of IXI. While beamforming enhances the POD for a standard receiver, yet it does not lead to a high POD performance (see the lower two curves). However, BF considerably enhances the capacity of the DS-CDMA system (see the $3^{rd}$ and $6^{th}$ curves from the bottom). Merging DS-CDMA with BF is thus highly promising for enhancing the $P_d$ performance of WLPS systems.

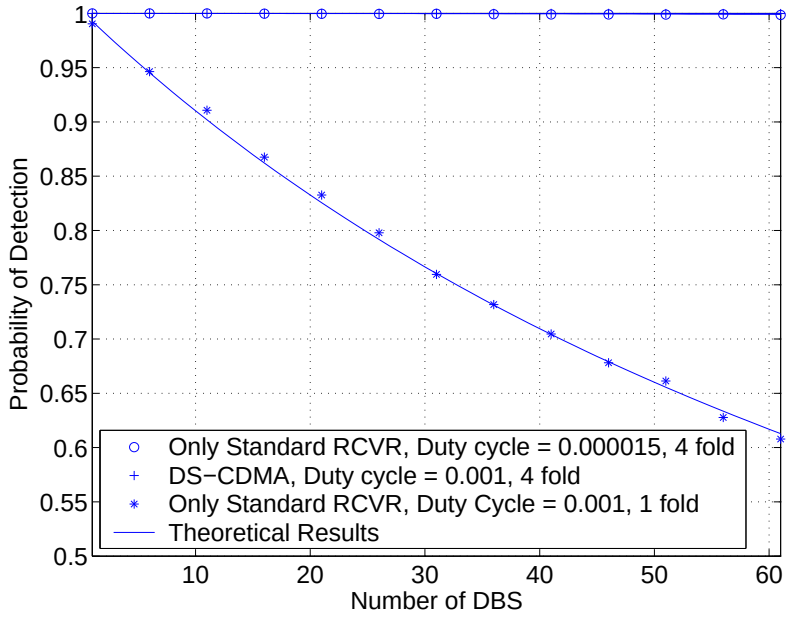


Figure 3.5: Simulation results for TRX receivers, without perturbation.

In Fig. 3.4 the solid lines represent the theoretical results which have a good match with the simulations.

The $P_d$ results for TRX receiver using standard receiver is shown in Fig. 3.5. Although simple, a standard TRX receiver typically achieves good $P_d$ performance. Further improvement is possible by selecting a larger $IRT$ value, or a smaller $\tau_{DBS}$ value. Occupying the same bandwidth as DS-CDMA, a standard receiver should choose $\tau_{TRX}$ ( $\tau_{DBS}$) to be $1/64^{th}$ of that of a DS-CDMA system. In this case, the same number of path diversity as the DS-CDMA receiver (i.e., four fold diversity) is achievable. This corresponds to $d_{c,TRX} \simeq 0.000015$ ($d_{c,DBS} \simeq 0.0015$), which leads to a very small $p_{ovl}$ at the DBS (TRX) receiver and very high $P_d$. This fact has been shown in Fig. 3.4 (see $5^{th}$ curve from the bottom) and Fig. 3.5 (see the top curves). The top curves in Fig. 3.4 and Fig. 3.5 have been redrawn in Fig. 3.6 and 3.8. In addition, the probability-of-miss-detection [defined in Eq. (3.29)] corresponding to the curves in Fig. 3.6 and 3.8 have been sketched in Fig. 3.7 and 3.9, respectively.

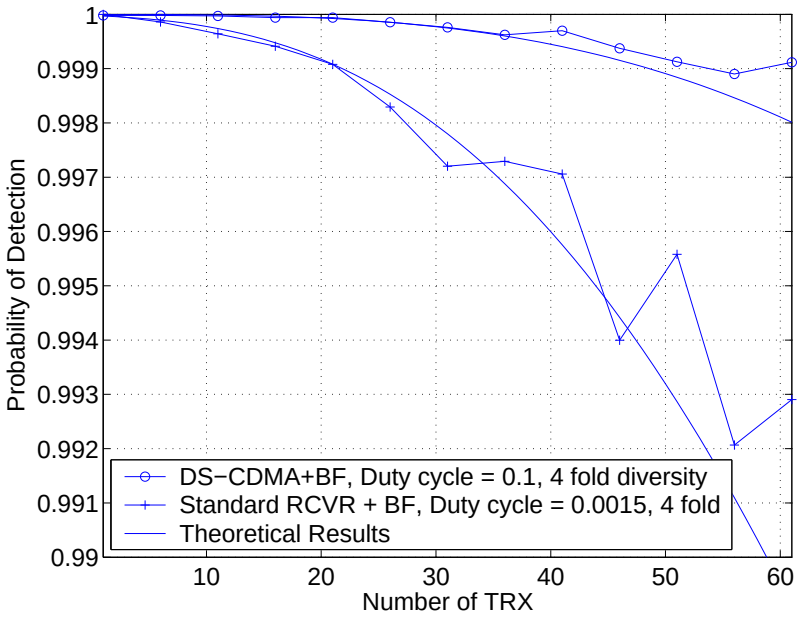


Figure 3.6: The top two curves of Fig. 3.4: Probability of Detection.

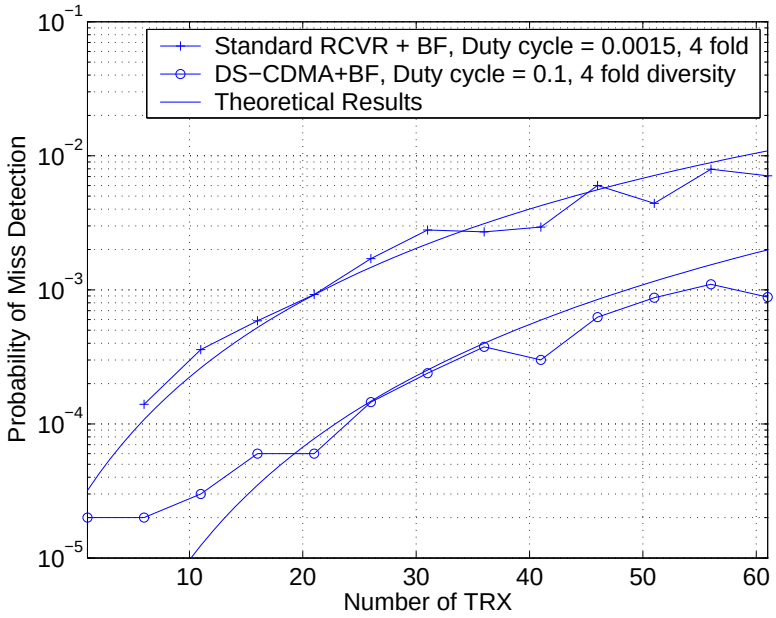


Figure 3.7: Probability of miss Detection corresponds to Fig. 3.6.

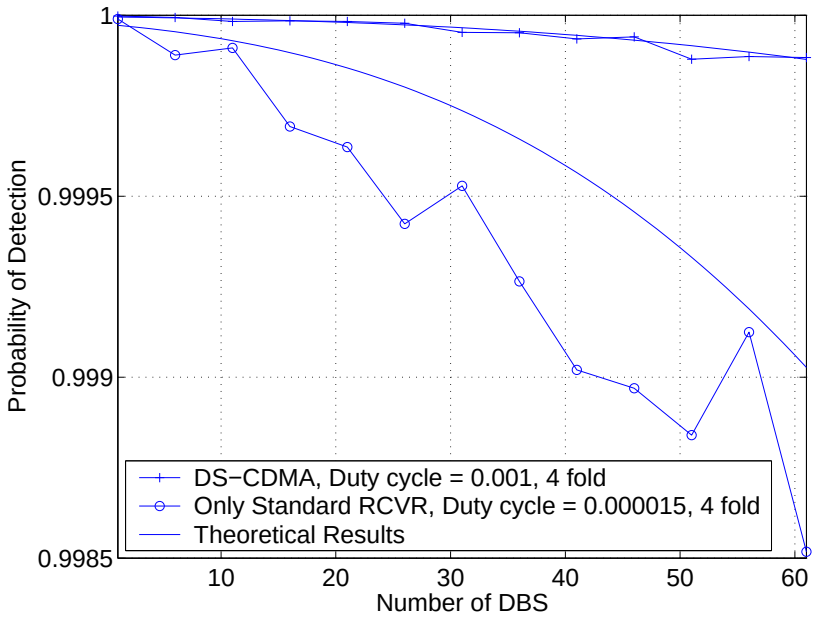


Figure 3.8: The top two curves of Fig. 3.5: Probability of Detection.

Fig. 3.6 and 3.7 show that, at the DBS receiver with similar bandwidths, DS-CDMA receivers outperforms standard receivers. When high number of TRXs are available, (e.g., Number of TRX=60), $P_d$ of standard system is approximately 99%, while $P_d$ of DS-CDMA system is approximately 99.8%. Hence, $P_d$ of DS-CDMA system satisfies most tracking purposes, which is due to a combination of path diversity and orthogonality achieved via these systems.

Fig. 3.9 shows that, at the TRX receiver with similar bandwidths, a DS-CDMA system outperforms standard system too. Because of the large selection of $IRT$ helps reduce IBI, TRX receiver with antenna array performance is even better than DBS receiver with antenna array in similar environments. When high number of TRXs are available,(e.g., Number of TRX=60), $P_d$ of standard system is approximately 99.9%, while $P_d$ of DS-CDMA system is approximately 99.99%. Note that the TRX receiver performance can also be improved via larger selection of $IRT$ (such as $0.24\,sec$). Due to simplicity, a standard receiver system is recommended for TRX receiver.

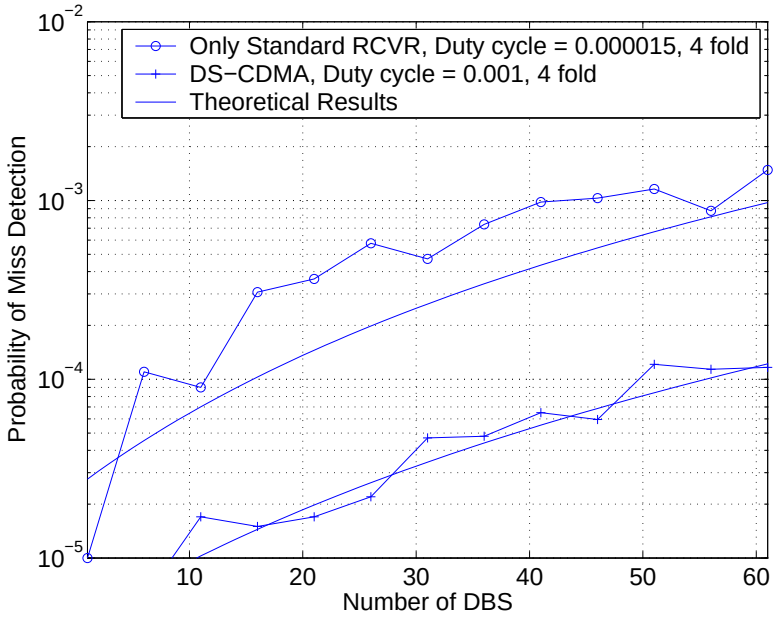


Figure 3.9: Probability of miss Detection corresponds to Fig. 3.8.

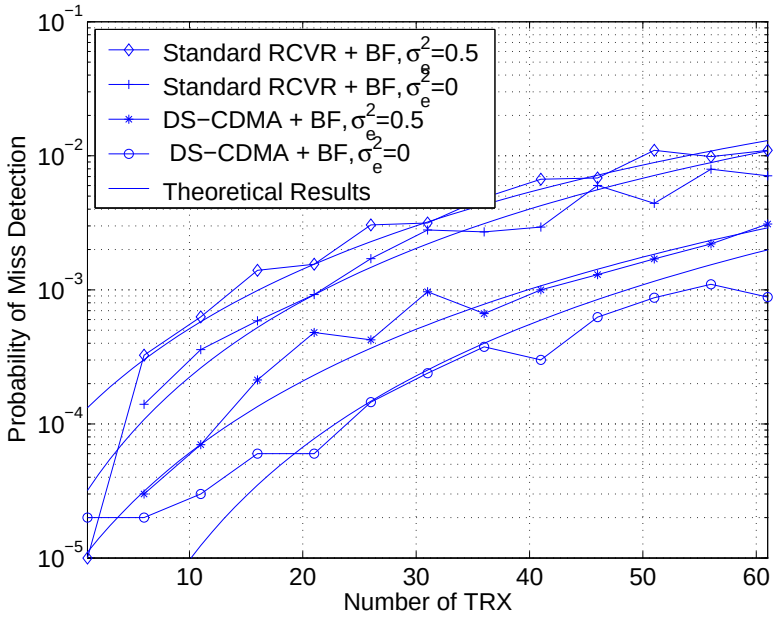


Figure 3.10: DBS performance with perturbations vs. Number of Users.

Fig. 3.10 shows the capacity of the system, i.e., the probability-of-miss-detection as a function of the number of TRX, under perturbations. This figure shows the system capacity decreases as the error variance equals to 0.5, and setting the probability-of-miss-detection threshold at 0.1%, the system capacity is reduced by 20% for DS-CDMA system, and 30% for standard systems.

## 3.5 Conclusions

This chapter presents a novel WLPS system. With a DBS/TRX structure and a novel signaling technique, WLPS would have various applications in road safety (vehicle collision avoidance), multirobot control, defense, law enforcement, and security. We studied and compared the probability-of-detection (POD) performance and the capacity of DBS and TRX with both DS-CDMA and standard systems with and without beamforming (BF) for a special setup suitable for road safety applications. The study shows that with similar bandwidth both DS-CDMA and standard systems with BF ensures high POD performance and capacity while DS-CDMA technique leads to a better performance comparing to standard receivers. We also investigated the perturbation effects in DBS antenna array weights. Perturbation reduces the asymptotic system POD performance and capacity. Future studies will focus on the development of unique DOA estimation for WLPS as well as adaptive BF techniques at the DBS receiver for high scattering environments.

## 3.A Probability of Detection Derivation

The POD is defined as the probability that DBS detects all bits within an ID correctly, and corresponds to:

$$P_d = (1 - P_e)^N, \tag{3.A.30}$$

where $P_e$ denotes the probability that one bit is not detected correctly, and $N$ denotes the number of bits in one ID.

Assuming $L$ path diversities and applying maximal ration combining, the received signal is in the form:

$$s = s_0 \cdot \sum_{l=1}^{L} |h_l|^2 + \sum_{l=1}^{L} h_l^* n_l, \tag{3.A.31}$$

where $s_0$ denotes the transmitted signal, $h_l$ denotes the fading over $l^{th}$ path, which is a circularly symmetric complex Gaussian random variable, and $n_l$ is the corresponding noise.

The instantaneous received signal power corresponds to:

$$P_s = E_s \cdot \left( \sum_{l=1}^{L} |h_l|^2 \right)^2, \tag{3.A.32}$$

where $E_s$ denotes the transmitted signal power.

The instantaneous noise power corresponds to:

$$P_n = \frac{N_0}{2} \cdot \sum_{l=1}^{L} |h_l|^2, \tag{3.A.33}$$

where $N_0$ denotes the thermal noise power.

Combining (3.A.32) and (3.A.33), the instantaneous SNR corresponds to:

$$r = \bar{r} \sum_{l=1}^{L} |h_l|^2, \tag{3.A.34}$$

where $\bar{r}$ denotes the average SNR.

Assuming a Rayleigh fading channel, it is well known that $|h_l|$ follows Rayleigh distribution, and $|h_l|^2$ follows exponential distribution. The summation of i.i.d. exponential random variables follows Erlang distribution, i.e., :

$$p = \sum_{l=1}^{L} |h_l|^2 \sim \frac{p^{(L-1)}}{(L-1)!} \cdot e^{-p}. \tag{3.A.35}$$

The instantaneous SNR follows a scaled Erlang distribution:

$$r = p\bar{r} \sim \frac{r^{L-1}}{(L-1)! \cdot \bar{r}^L} \cdot e^{-\frac{r}{\bar{r}}}. \tag{3.A.36}$$

The BER corresponds to:

$$P_e = \int_0^\infty Q(\sqrt{2r}) P(r|\bar{r})\, dr \tag{3.A.37}$$

$$= \int_0^\infty \int_{\sqrt{2r}}^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \frac{e^{-\frac{r}{\bar{r}}} \cdot r^{L-1}}{(L-1)! \cdot \bar{r}^L}\, dx\, dr. \tag{3.A.38}$$

Let $v = \frac{r}{\bar{r}}$, after certain manipulations, the BER corresponds to:

$$P_e = \frac{1}{(L-1)!} \int_0^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \int_0^{\frac{x^2}{2\bar{r}}} e^{-v} v^{L-1}\, dv\, dx. \tag{3.A.39}$$

Note that

$$\int_0^{\frac{x^2}{2\bar{r}}} e^{-v} v^{L-1}\, dv = -(L-1)! \sum_{l=0}^{L-1} \frac{v^l e^{-v}}{l!} \Big|_{v=0}^{v=\frac{x^2}{2\bar{r}}}$$

$$= (L-1)! \left(1 - \sum_{l=0}^{L-1} \frac{(\frac{x^2}{2\bar{r}})^k e^{-\frac{x^2}{2\bar{r}}}}{l!}\right) \tag{3.A.40}$$

Combining (3.A.40) and (3.A.39), the BER corresponds to:

$$P_e = \int_0^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \left(1 - \sum_{l=0}^{L-1} \frac{(\frac{x^2}{2\bar{r}})^k e^{-\frac{x^2}{2\bar{r}}}}{l!}\right) dx \tag{3.A.41}$$

In 3.A.41:

$$\int_0^\infty \frac{1}{\sqrt{2\pi} \cdot l!} e^{-\frac{x^2}{2} - \frac{x^2}{2\bar{r}}} (\frac{x^2}{2\bar{r}})^l\, dx$$

$$= \frac{1}{l!} \cdot \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{(2\bar{r})^l} \int_0^\infty x^{2l} e^{-\frac{1+\bar{r}}{2\bar{r}} x^2}\, dx$$

$$= \frac{1}{l!} \cdot \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{(2\bar{r})^l} \cdot \frac{(2l-1)!!}{2^{l+1} \left(\frac{1+\bar{r}}{2\bar{r}}\right)^l} \sqrt{\frac{2\pi\bar{r}}{1+\bar{r}}}$$

$$= \binom{2l}{l} \frac{1}{2^{2l+1}(1+\bar{r})^l} \sqrt{\frac{\bar{r}}{1+\bar{r}}} \tag{3.A.42}$$

Hence, the final BER expression is:

$$P_e = \int_0^\infty Q(\sqrt{2r_0}) f(r_0|\bar{r}_0)\, dr_0 = 0.5 \left(1 - \sqrt{\frac{\bar{r}}{1+\bar{r}}} \cdot \sum_{l=0}^{L-1} \binom{2l}{l} \cdot \frac{1}{2^{2l} \cdot (1+\bar{r})^l}\right). \tag{3.A.43}$$

Here, $\bar{r} = L \cdot r_i$, which corresponds to a uniform power distribution over $L$ paths. Please note that the result in (3.A.37) is not only applicable for single user case, but also applicable for multi-user environments, as long as the interference and thermal noise follows Gaussian distribution, which is valid in most cases based on the central limit theorem. Thus, the result in (3.A.43) is applicable for CDMA multi-user case as well, where $r_e$ is in fact Signal to Interference and Noise Ratio (SINR), which is also called "effective SNR". The derivation of effective SNR in WLPS is presented in Appendix 3.B.

Although we derived the above results independently, it is recently found that similar results have been obtained in [Eng and Milstein(1995)].

## 3.B Effective SNR Derivation

As we discussed in Section 3.3.4, the received signal with BF and DS-CDMA corresponds to:

$$\vec{r}(t) = \sum_{k=1}^{K} \sum_{l=0}^{L^k-1} \sum_{n=0}^{N-1} \alpha_l^k \cdot \vec{V}(\theta_l^k) \cdot b^k[n] \cdot g_{T_b}(t-\tau_l^k-nT_b) \cdot a^k(t-\tau_l^k-nT_b) \cdot \cos(2\pi f_c t+\phi_l^k)+\vec{n}(t). \tag{3.B.44}$$

The $i^{th}$ bit output of BF for $j^{th}$ user's $q^{th}$ path is given as:

$$\begin{aligned} y_q^j[n] &= \vec{W}^H(\theta_q^j) \int_{\tau_q^j+nT_b}^{\tau_q^j+(n+1)T_b} \vec{r}(t) \cdot \cos(2\pi f_c t + \phi_q^j) \cdot a^k(t - \tau_q^j - nT_b)\, dt \\ &= \sum_{k=1}^{K} \sum_{l=0}^{L^k-1} \sum_{n=0}^{N-1} \alpha_l^k \cdot C_{ql}^{jk} \cdot b^k[n] \cdot \cos(\phi_l^k - \phi_q^j) \cdot \gamma^{kj}(\tau_h^l - \tau_q^j - nT_b) \\ &+ \int_{\tau_q^j+nT_b}^{\tau_q^j+(n+1)T_b} \vec{W}^H(\theta_q^j) \cdot \vec{n}(t) \cdot a^k(t - \tau_q^j - nT_b) \cdot \cos(2\pi f_c t + \phi_q^j)\, dt, \end{aligned} \tag{3.B.45}$$

where

$$C_{ql}^{jk} = \vec{W}^H(\theta_q^j) \cdot \vec{V}(\theta_l^k), \tag{3.B.46}$$

and

$$\gamma^{kj}(\tau) = \int_0^{T_b+\tau} a^k(t) \cdot a^j(t-\tau)\, dt. \tag{3.B.47}$$

The first term in 3.B.45 can be decomposed to three parts: Desired signal, Inter-Symbol-Interference (ISI), and Multi-Access-Interference (MAI).

The desired signal corresponds to the case that $k = j$ and $q = l$:

$$S_q^j[n] = \alpha_q^j \cdot C_{qq}^{jj} \cdot r^{jj}(0) \cdot b^j[n]. \tag{3.B.48}$$

The ISI corresponds to the case that $k = j$ and $q \neq l$:

$$ISI_q^j[n] = \sum_{l=0, l\neq q}^{L^k-1} \sum_{i=0}^{N-1} \alpha_l^j \cdot C_{ql}^{jj} \cdot b^j[n] \cdot \cos(\phi_l^j - \phi_q^j) \cdot r^{jj}(\tau_l^j - \tau_q^j - nT_b). \tag{3.B.49}$$

The MAI corresponds to the case that $k \neq j$:

$$MAI_q^j[n] = \sum_{k=1, k\neq j}^{K} \sum_{l=0}^{L^k-1} \sum_{i=0}^{N-1} \alpha_l k \cdot C_{ql}^{jk} \cdot b^k[n] \cdot \cos(\phi_l^k - \phi_q^j) \cdot \gamma^{kj}(\tau_l^k - \tau_q^j - nT_b). \tag{3.B.50}$$

Finally, the noise component corresponds to:

$$NI_q^j[n] = \vec{W}^H(\theta_q^j) \cdot \int_{\tau_q^j+nT_b}^{\tau_q^j+(n+1)T_b} \vec{n}(t) \cdot a^k(t - \tau_q^j - nT_b) \cdot \cos(2\pi f_c t + \phi_q^j)\, dt. \tag{3.B.51}$$

The SINR corresponds to:

$$SINR_q^j[n] = \frac{E\left[(S_q^j[n])^2\right]}{E\left[(ISI_q^j[n])^2\right] + E\left[(MAI_q^j[n])^2\right] + E\left[(NI_q^j[n])^2\right]}. \tag{3.B.52}$$

The desired signal power corresponds to:

$$E\left[(S_q^j[n])^2\right] = E[(\alpha_q^j)^2] \cdot E[(C_{qq}^{jj})^2] \cdot E[(r^{jj}(0))^2]/L. \tag{3.B.53}$$

Assuming a Rayleigh fading channel,

$$E\left[(S_q^j[n])^2\right] = E[(C_{qq}^{jj})^2] \cdot E[(r^{jj}(0))^2]/L. \tag{3.B.54}$$

Note that the Rayleigh channel assumption will be constantly used in the following derivations.

The ISI power corresponds to:

$$E\left[(ISI_q^j[n])^2\right] = \sum_{l=0,l\neq q}^{L^k-1} E[(C_{ql}^{jj})^2] \cdot E[\cos(\phi_l^j - \phi_q^k)^2] \cdot \sum_{i=0}^{N-1} E[(r^{jj}(\tau_l^j - \tau_q^j - nT_b))^2] \tag{3.B.55}$$

Because $\phi_l^j \sim U[0, 2\pi]$, $\phi_q^j \sim U[0, 2\pi]$ and they are independent of each other, it is easy to find $E[\cos(\phi_l^j - \phi_q^k)^2] = 0.5$. Assuming all TRX experience the same number of paths ($L^k = L, \forall k$), within a specific wireless environments, the ISI power finally corresponds to:

$$E\left[(ISI_q^j[n])^2\right] = (1 - \frac{1}{L}) \cdot \frac{1}{2} \cdot E[(C_{ql}^{jj})^2] \cdot \sum_{i=0}^{N-1} E[(\gamma^{jj}(\tau_l^j - \tau_q^j - nT_b))^2] \tag{3.B.56}$$

Similarly, the MAI power corresponds to:

$$E[(MAI_q^j[n])^2] = \frac{K-1}{2} \cdot E[(C_{ql}^{jk})^2] \cdot \sum_{i=0}^{N-1} E[(\gamma^{kj}(\tau_l^k - \tau_q^j - nT_b))^2] \tag{3.B.57}$$

Finally, assuming unit transmit energy, the noise power corresponds to:

$$NI_q^j[n] = \frac{E[\vec{W}^H(\theta_q^j)\vec{W(\theta_q^j)}] \cdot G}{M \cdot 2\bar{r}_o} \tag{3.B.58}$$

where $\bar{r}_o$ denotes the average SNR.

Now, knowing $E[(C_{ql}^{jk})^2] = E[(C_{ql}^{jj})^2]$ and $E[(\gamma^{jj}(\tau_l^j - \tau_q^j - nT_b))^2] = E[(\gamma^{kj}(\tau_l^k - \tau_q^j - nT_b))^2]$ from appendix III (see (III.72) and (III.78)), substituting (II.54), (II.56), (II.57) and (II.58) into (II.52), the effective SNR corresponds to:

$$r_i = \frac{E[(C_{qq}^{jj})^2] \cdot G/L}{(K - \frac{1}{L}) \cdot E[(C_{ql}^{jj})^2] \cdot \sum_{i=0}^{N-1} E[(r^{jj}(\tau_l^j - \tau_q^j - nT_b))^2] + \frac{E[\vec{W}^H(\theta_q^j)\vec{W(\theta_q^j)}] \cdot G}{M \cdot 2\bar{r}_o}} \tag{3.B.59}$$

Let $A = E[(C_{qq}^{jj})^2] \cdot G$, $B = E[(C_{ql}^{jj})^2] = E[(C_{ql}^{jk})^2]$, and $D = \sum_{i=0}^{N-1} E[(r^{jj}(\tau_l^j - \tau_q^j - nT_b))^2] = \sum_{i=0}^{N-1} E[(r^{jk}(\tau_l^j - \tau_q^j - nT_b))^2]$, the effective SNR corresponds to:

$$r_i = \frac{A}{(K - \frac{1}{L}) \cdot B \cdot D + \frac{E[\vec{W}^H(\theta_q^j)\vec{W(\theta_q^j)}] \cdot G}{M \cdot \bar{r}_o}} \cdot \frac{2}{L} \tag{3.B.60}$$

This is the general framework that is used in effective SNR computing. The parameters $A\,, B\,, D$ and $E[\vec{W}^H(\theta_q^j)\vec{W(\theta_q^j)}]$ varies with system configuration, i.e., with and without DS-CDMA, and, with and without beamforming.

Before we go to the computation of $A\,, B\,, D$ and $E[\vec{W}^H(\theta_q^j)\vec{W(\theta_q^j)}]$, let us give an intuitive explanation of the framework.

First, we explain the term at the end of (3.B.60), $\frac{2}{L}$. It is divided by $L$ because the signal energy is distributed in $L$ paths in our system. It is multiplied by 2 since all interference and noise are attenuated by 2 because of random channel phase (see explanation of (II.55), i.e., $E[\cos(\phi_l^j - \phi_q^k)^2] = 0.5$), while desired signal is not attenuated due to perfect phase tracking.

Second, it is easy to understand that $A$ represents how well the desired signal is magnified via beamforming and CDMA techniques.

Third, $\frac{E[\vec{W}^H(\theta_q^j)\vec{W(\theta_q^j)}]\cdot G}{M}$ denotes how the noise is amplified due to beamforming and CDMA. We note the amplification of noise is always less than or equal to the amplification of desired signal.

Finally, we explain the interference term $(K - \frac{1}{L}) \cdot B \cdot D$. Assuming $K$ users, each user transmit through $L$ paths, there are totally $KL$ signals. Only one of those signals is desired; All others are interference. So, the number of interferers is $KL - 1$. Note each interferer possesses $\frac{1}{L}$ energy, the raw interference power is $K - \frac{1}{L}$. Here, the term $B$ and $D$ represent how well the multi-access techniques (SDMA and CDMA) suppress those interferers.

Now we see the framework in appendix II is solid. Exact equations for $A$, $B$, $D$, and $E[\vec{W}^H(\theta_q^j)\vec{W(\theta_q^j)}]$ are given in appendix III. Here we give an example, the derivation of (21), to show how to determine $A\,, B\,, D$ and $E[\vec{W}^H(\theta_q^j)\vec{W(\theta_q^j)}]$ via results in appendix III.

First, according to (III.72), with $\sigma_\epsilon^2 = 0$:

$$E[(C_{qq}^{jj})^2] = M^2 \tag{3.B.61}$$

Hence,

$$A = M^2 \cdot G \tag{3.B.62}$$

Second, according to (III.72) as well, with $\sigma_\epsilon^2 = 0$:

$$B = E[(C_{ql}^{jj})^2] = E[(C_{ql}^{jk})^2] \tag{3.B.63}$$
$$= \sum_{m=0}^{M-1}(m+1)J_0(\frac{2\pi dm}{\lambda})J_0(-\frac{2\pi dm}{\lambda}) + \sum_{m=M}^{2M-2}(2M-m-1)J_0(\frac{2\pi dm}{\lambda})J_0(-\frac{2\pi dm}{\lambda})$$

Third, according to (III.78):

$$D = d_c - \frac{N^2-3N+3}{2N^2}d_c^2 \tag{3.B.64}$$

Finally, according to (III.73), with $\sigma_\epsilon^2 = 0$:

$$E[\vec{W}^H(\theta_q^j)\vec{W(\theta_q^j)}] = M^2 \tag{3.B.65}$$

Now, we clearly see how to obtain the result in (21). All other equations are derived in a similar manner. All necessary information regarding computation of $A$, $B$, $D$, and $E[\vec{W}^H(\theta_q^j)\vec{W(\theta_q^j)}]$ are included in appendix III.

## 3.C Spatial and Temporal Correlation Computation

As we discussed in section 3.3.5, $\vec{W}(\theta_l^k) = \vec{V}(\theta_l^k) + \vec{\varepsilon}_l^k$. Hence, using (3.B.46):

$$C_{ql}^{jk} = \vec{W}_q^{jH} \cdot \vec{V}_l^k = \vec{V}_q^{jH} \cdot \vec{V}_l^k + \vec{\varepsilon}_q^{jH} \cdot \vec{V}_l^k, \tag{3.C.66}$$

the spatial correlation corresponds to:

$$E\left[(C_{ql}^{jk})^2\right] = E\left[(\vec{V}_q^{jH} \cdot \vec{V}_l^k)^2\right] + E\left[(\vec{\varepsilon}_q^{jH} \cdot \vec{V}_l^k)^2\right]. \tag{3.C.67}$$

The cross terms are removed because $\vec{\varepsilon}_q^j$ is zero-mean and independent from $\vec{V}_l^k$ and $\vec{V}_q^j$.

In the first term of (3.C.67):

$$\begin{aligned}(\vec{V}_q^{j^H}\cdot\vec{V}_l^k)^2 &= \left[\sum_{m=0}^{M-1} e^{j2\pi dm(\cos_q^j-\cos_l^k)/\lambda}\right]^2\\ &= \sum_{m=0}^{M-1}(m+1)e^{j2\pi dm(\cos_q^j-\cos_l^k)/\lambda}\\ &\quad+\sum_{m=M}^{2M-2}(2M-m-1)e^{j2\pi dm(\cos_q^j-\cos_l^k)/\lambda}.\end{aligned}\tag{3.C.68}$$

Assuming $\theta_q^j$ and $\theta_l^k$ are independent from each other and uniformly distributed in $[0,\pi]$, the first term in (3.C.67) corresponds to:

$$\begin{aligned}E\left[(\vec{V}_q^{j^H}\cdot\vec{V}_l^k)^2\right] &= \int_0^\pi\int_0^\pi \frac{1}{\pi^2}(\vec{V}_q^{j^H}\cdot\vec{V}_l^k)^2\,d\theta_q^j\,d\theta_l^k\\ &= \sum_{m=0}^{M-1}(m+1)J_0(\frac{2\pi dm}{\lambda})J_0(-\frac{2\pi dm}{\lambda})\\ &\quad+\sum_{m=M}^{2M-2}(2M-m-1)J_0(\frac{2\pi dm}{\lambda})J_0(-\frac{2\pi dm}{\lambda})\end{aligned}\tag{3.C.69}$$

In the first term of (3.C.67):

$$(\vec{\varepsilon}_q^{j^H}\cdot\vec{V}_l^k)^2 = \left(\sum_{m=0}^{M-1}\varepsilon_{q,m}^{j^H}\cdot e^{-j2\pi dm\cos\theta_l^k/\lambda}\right)^2,\tag{3.C.70}$$

where $\varepsilon_{q,m}^j$ denotes the $m^{th}$ element of $\vec{\varepsilon}_q^j$. Here, the second term in (3.C.67) corresponds to:

$$\begin{aligned}E\left[(\vec{\varepsilon}_q^{j^H}\cdot\vec{V}_l^k)^2\right] &= E\left[\sum_{m=0}^{M-1}(\varepsilon_{q,m}^{j^H})^2\cdot e^{-j2\cdot2\pi dm\cos\theta_l^k/\lambda}\right]\\ &= \sigma_\epsilon^2\cdot\sum_{m=0}^{M-1}\int_0^\pi\frac{1}{\pi}e^{-j2\cdot2\pi dm\cos\theta_l^k/\lambda}\,d\theta_l^k\\ &= \sigma_\epsilon^2\sum_{m=0}^{M-1}J_0(-\frac{2\cdot2\pi dm}{\lambda})\end{aligned}\tag{3.C.71}$$

Combining (3.C.69) and (3.C.71), the final expression of spatial correlation cor-

responds to:

$$E\left[(C_{ql}^{jk})^2\right] = \begin{cases} M^2 + \sigma_\epsilon^2 \sum_{m=0}^{M-1} J_0(-\frac{2\cdot 2\pi dm}{\lambda}) & \text{if } l = q \text{ and } j = k \\ \sum_{m=0}^{M-1}(m+1)J_0(\frac{2\pi dm}{\lambda})J_0(-\frac{2\pi dm}{\lambda}) & \\ +\sum_{m=M}^{2M-2}(2M-m-1)J_0(\frac{2\pi dm}{\lambda})J_0(-\frac{2\pi dm}{\lambda}) & \\ +\sigma_\epsilon^2 \sum_{m=0}^{M-1} J_0(-\frac{2\cdot 2\pi dm}{\lambda}) & \text{if } l \neq q \text{ or } j \neq k \end{cases} \tag{3.C.72}$$

Following a similar manner, it is easy to find that:

$$E\left[\vec{W}^H(\theta_q^j)\vec{W}(\theta_q^j)\right] = M^2 + \sigma_\epsilon^2, \tag{3.C.73}$$

which will be used in noise power computation.

Assuming a pseudo-noise coded DS-CDMA system, given $k \neq j$, the temporal correlation corresponds to:

$$E\left[\gamma^{kj}||\tau|\right] = \begin{cases} 0 & \text{if } T_b < |\tau| \\ 1 & \text{if } 0 \leq |\tau| \leq T_b \end{cases}, \tag{3.C.74}$$

and,

$$E\left[(\gamma^{kj})^2||\tau|\right] = \begin{cases} 0 & \text{if } T_b < |\tau| \\ 1 & \text{if } 0 \leq |\tau| \leq T_b \end{cases}. \tag{3.C.75}$$

Let $\tau' = \tau_l^k - \tau_q^j$, assuming $\tau_l^k$ and $\tau_q^j$ uniformly distributed in $[0, \tau_0]$, the probability density function of $|\tau'|$ corresponds to:

$$p(|\tau'|) = \frac{2}{\tau_0}\left(1 - \frac{|\tau'|}{\tau_0}\right) \qquad \text{if } |\tau'| < \tau_0, \tag{3.C.76}$$

where $\tau_0 = IRT_{min}$ defined in (3.1) .

Note that

$$E[(\gamma^{kj}(\tau')^2] = \int_0^{\tau_0} E\left[(\gamma^{kj})^2||\tau'|\right] p(|\tau'|)\, d|\tau'| = \int_0^{T_b} \frac{2}{\tau_0}\left(1 - \frac{|\tau'|}{\tau_0}\right) d|\tau'| = 2\cdot\frac{T_b}{\tau_0} - (\frac{T_b}{\tau_0})^2, \tag{3.C.77}$$

.

Finally

$$\sum_{n=0}^{N-1} E[(\gamma^{kj}(\tau_l^k - \tau_q^j - nT_b))^2] = d_c - \frac{N^2 - 3N + 3}{2N^2} d_c^2, \tag{3.C.78}$$

where $d_c$ denotes the duty cycle defined as $N \cdot \frac{T_b}{IRT_{min}}$.

# Chapter 4

# Stationarity Analysis in SIMO Channels

This chapter investigates the implementation of a novel Wireless Local Positioning System (WLPS). WLPS main components are: (a) a dynamic base station (DBS), and (b) a transponder, both mounted on mobiles. The DBS periodically transmits ID request signals. As soon the transponder detects the ID request signal, it sends its ID (a signal with a limited duration) back to the DBS. Hence, the DBS receives *non-continuous* signals *periodically* transmitted by the transponder. The non-continuous nature of the WLPS leads to non-stationary received signals at the DBS receiver, while the periodic signal structure leads to the fact that the DBS received signal is also cyclostationary.

This work discusses the implementation of Linear Constrained Minimum Variance (LCMV) beamforming at the DBS receiver. We demonstrate that the non-stationarity of the received signal causes the sample covariance to be an inaccurate estimate of the true signal covariance. The errors in this covariance estimate limit the applicability of LCMV beamforming. A modified covariance matrix estimator, which exploits the cyclostationarity property of WLPS system is introduced to solve the non-stationarity problem. The cyclostationarity property is discussed in detail theoretically and via simulations. It is shown that the modified covariance matrix estimator significantly improves the DBS performance. The proposed technique can be

applied to periodic-sense signaling structures, such as the WLPS, RFID and reactive sensor networks.

## 4.1 Introduction

This chapter[1][2] investigates how to implement optimal beamforming for a novel wireless local positioning system (WLPS[3]). We focus on how to estimate covariance matrix for optimal beamforming, because the specific signaling scheme in this WLPS, i.e., cyclostationarity, enables a novel covariance matrix estimator.

The WLPS consists of two main components [Tong and Zekavat(2005)]: a Dynamic Base Station (DBS) and a transponder (or possibly a number of transponders), all mounted on mobiles. The DBS periodically transmits ID request signals (a short burst of energy). Each time a transponder detects the ID request signal, it sends its unique ID (a signal with a limited duration) back to the DBS. In the WLPS, the DBS detects and tracks the positions *and* IDs of the transponders in its coverage area. The position of a transponder is determined by the combination of time-of-arrival (TOA) and direction-of-arrival (DOA). TOA is estimated via the time difference between the transmission of ID request signal and the reception of the corresponding ID. DOA estimation would be possible if an antenna array is installed at the DBS receiver [Stoica and Nehorai(1989)].

In WLPS, a *single* unit (the DBS) is capable of positioning transponders located in its coverage area. In systems such as cell phone positioning [Hellebrandt et al.(1997)] and radio frequency ID [Juels(2006)], *multiple* units should cooperate in the process of positioning. Accordingly, the WLPS has many civilian and military applications. For example, in vehicle collision avoidance applications, each vehicle (car) may carry

[1]This work was partially reported in [Tong and Zekavat(2005)].
[2]This work is supported by the US NSF grant ECS-0427430.
[3]WLPS US Patent is Pending at Michigan Tech. University.

a DBS and each pedestrian may carry a transponder. Then, each vehicle is able to position (and identify) pedestrians. Another possible application of the WLPS is airport security, where security guards may carry DBSs and passengers may carry transponders.

The WLPS can be considered as a merger of positioning and communication systems: The TOA/DOA estimation is the primary procedure for positioning, while the ID detection process is supported by communications. This chapter investigates the ID detection performance, i.e., the communication aspect of the WLPS, while the TOA/DOA estimation process is discussed in [Pourrostam et al.(2007)] and [Wang and Zekavat(2006)].

As depicted in [Tong and Zekavat(2007)], the main source of error in the ID detection process is the interference from other transponders. To reduce this interference, direct sequence code division multiple access (DS-CDMA) and beamforming techniques are adopted in the WLPS. The conventional beamforming methods (delay and sum) in the WLPS has been discussed in [Tong and Zekavat(2007)]. In general, linear constrained minimum variance (LCMV) beamforming outperforms conventional beamforming in terms of interference suppression [Frost(1972)]. Therefore, it is natural to extend our study from conventional beamforming to LCMV beamforming.

An important step to perform LCMV beamforming is the estimation of the covariance matrix of the received signal. Considering stationary signals, sample covariance accurately estimates the true signal covariance [Carlson(1988)]. However, in the WLPS, the received signal at the DBS receiver is not stationary, because the DBS transmits ID request signals non-continuously. The non-stationarity of the received signal causes the sample covariance to be an inaccurate estimate of the true signal covariance. The errors in this covariance estimate limit the applicability of LCMV beamforming in the WLPS.

In this work, a modified covariance matrix estimator is proposed. The transponders transmit signals non-continuously and repetitively. Accordingly, the DBS received signal is non-stationary and cyclostationary. The proposed modified covariance matrix estimator exploits the cyclostationarity to counter the non-stationarity problem. A detailed theoretical analysis shows that, in most practical situations, the cyclostationarity duration is sufficiently long to ensure an accurate estimate. Finally, the WLPS ID detection performance is numerically simulated. The numerical results confirm that the modified covariance matrix estimator improves the WLPS performance significantly. It should be further noted that the proposed estimator is not restricted to this particular WLPS system: it is possible to apply this estimator to any system that exhibits repetitive structures. Hence, the proposed covariance matrix estimator has a wide range of applications.

Beamforming [Capon(1969)] and cyclostationarity [Gardner et al.(2006)] have been studied separately for more than fifty years. In recent decades, a joint consideration of beamforming and cyclostationarity (i.e., beamforming for cyclostationary signals) attracted certain attention [Wu and Wong(1996)] [Lee and Lee(1999)]. In those studies, the signals are both stationary and cyclostationary. In other words, continuous signals with repetitive structures are considered. In our work, we study non-continuous signals with repetitive structures. Therefore, this chapter exploits cyclostationarity to counter the non-stationarity problem in optimal beamforming.

The rest of the chapter is organized as follows: Section 4.2 introduces the fundamentals of the WLPS structure; Section 4.3 discusses the implementation of WLPS system and the non-stationarity problem; Section 4.4 demonstrates how to exploit cyclostationarity to counter the non-stationarity problem; Section 4.5 presents numerical results and Section 4.6 concludes the chapter.

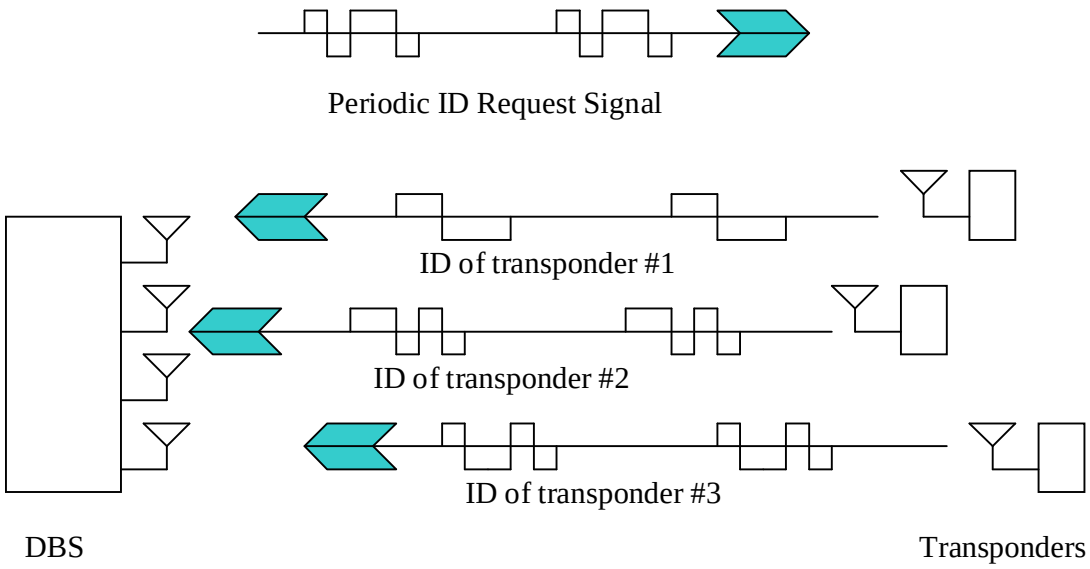


Figure 4.1: WLPS Basic structure.

## 4.2 Non-Continuous and Repetitive Nature of the WLPS

The WLPS comprises of a set of DBS and transponders. In the scope of this chapter, we consider the communication between one DBS and multiple transponders. The DBS transmits ID request signals periodically to all transponders in its coverage area. Once a transponder detects the ID request signal, it sends its unique ID (a signal with limited duration) back to the DBS, as shown in Fig. 4.1. The DBS is equipped with multiple antennas to support DOA estimation and beamforming.

In the WLPS, a DBS communicates with multiple transponders simultaneously. This is the same as standard cellular communication systems. However, different from cellular systems, the DBS received signal in the WLPS is not stationary.

As shown in Fig. 4.1, the signal transmitted by a transponders do not span over the whole time domain. This feature leads to a new performance measure metric: probability-of-overlapping, $p_{ovl}$, which is defined as the probability that the desired ID is overlapped with the ID signals from other transponders. In standard wireless systems, $p_{ovl}$ is always unity for multiple transponders. In the DBS receiver, the

probability-of-overlapping is less than unity and corresponds to:

$$p_{ovl} = 1 - (1 - d_c)^{K-1}, \tag{4.1}$$

where $K$ denotes the number of transponders, and, $d_c$ represents duty cycle, which is defined as:

$$d_c = \frac{\tau}{IRT_{min}}. \tag{4.2}$$

Here, $\tau$ is the duration of the ID of a transponder, and $IRT_{min}$ is the time difference between the first responding transponder and the last responding transponder. A comprehensive results for $IRT_{min}$ has been introduced in [Tong and Zekavat(2005)]; here, roughly,

$$IRT_{min} = \frac{R_{max}}{2c}, \tag{4.3}$$

where $R_{max}$ is the maximum coverage distance of the DBS, and $c$ denotes the speed of light. For vehicle collision avoidance applications, typically $R_{max}$ should not not exceed $1\,km$. The exact value of $R_{max}$ may vary with different environments, e.g., urban or highways. .

In general, through this preliminary study, the non-continuous nature of the WLPS seems alleviate the interference problem: the undesired signals from other transponders may or may not interfere with the desired signal. In contrast, in standard communication systems, the undesired signals always overlap with the desired signal.

However, it is noted that the non-continuous nature of the WLPS is not sufficient in terms of rejecting interference. As shown in Fig. 4.2, the probability-of-overlapping is very high when $d_c = 0.1$ with a moderate number of transponders ($N = 10$). In many applications, e.g., vehicle collision avoidance, the duty cycle might be even larger than 0.1. Therefore, one can not expect to suppress interference reliably through the non-continuous nature of the WLPS.

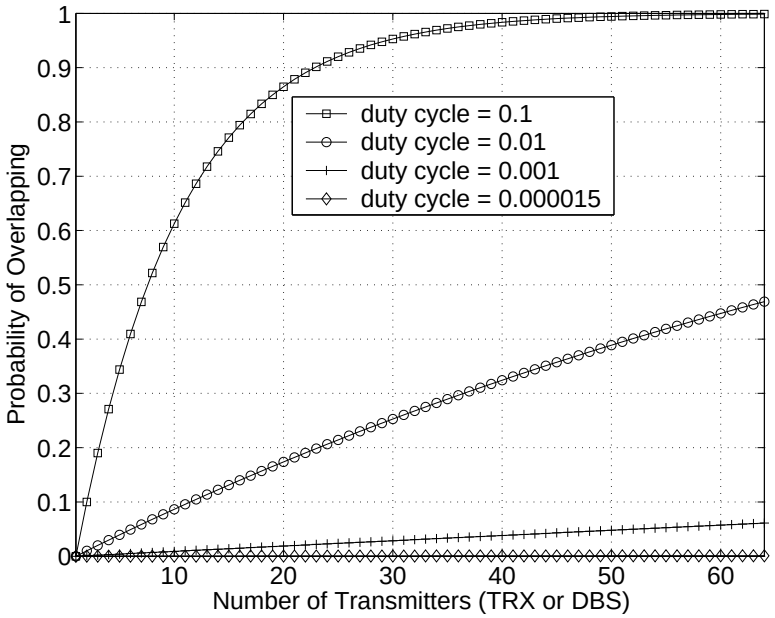


Figure 4.2: The probability of overlapping.

To reduce interference power, DS-CDMA and beamforming techniques are necessary in the WLPS. A detailed analysis for conventional beamforming and DS-CDMA techniques has been presented in [Tong and Zekavat(2007)]. In general, optimal beamformers performs better than the conventional beamformer. Hence, it is natural to extend our study from conventional beamformer to optimal beamformers.

Optimal beamformers generate a statistically optimum estimation of the desired signal through applying a weight vector to the observed data. This weight vector is computed via optimizing a certain cost function. Examples of these cost functions include total power, SINR, entropy, Mean Square Error, or Non-Gaussianity [Godara(1997)][Hyvarinen et al.(2001)]. Here, LCMV beamformer is selected because: 1) it is particularly good at rejecting interference, and 2) it only requires the observations of the received signals and the direction of the desired signal. The former is easy to obtain and the latter has been available via the DOA estimation process, which is prior to the beamforming process.

The basic structure of the WLPS has been introduced in this section. In the next section, we introduce the signal model of the WLPS, and describe the beamforming implementation in a mathematical form. It is emphasized that directly applying LCMV beamforming in the WLPS is not appropriate due to its non-stationary nature. In Section 4.4, cyclostationarity would be exploited to solve the non-stationarity problem.

## 4.3 System Implementation and Non-Stationarity Analysis

Once a transponder detects the ID request signal, it would transmit its unique ID back to the DBS. To suppress interference from other transponders, the bits in the ID are spread by DS-CDMA techniques. Hence, the transponders would periodically transmit DS-CDMA signals that are with a limited duration. In a multi-path (urban) environments, the received signal at the DBS receiver would be the summation of DS-CDMA signals from multiple transponders through multiple paths. Finally, in the DBS receiver, it is possible to apply DS-CDMA despreading and beamforming techniques to extract the ID of the desired transponder, as explained in Section 4.3.1

In this work, the DOA estimation for the paths of the desired transponder is assumed to be perfect: Although the non-stationarity nature does have effect on DOA estimation, the effect turns out to be minimal, and the DOA estimation is accurate enough for most practical applications[Pourrostam et al.(2007)]. Since the only required information for LCMV beamforming is the directions of the paths of the desired transponder and the estimation of covariance matrix, a good estimation of the covariance matrix would ensure a good ID detection performance, as depicted in Section 4.3.2.

In Section 4.3.3, it is shown that the standard sample covariance matrix estimator does not lead to a good quality of covariance matrix estimation. The reason is

that, due to the non-stationary nature of the WLPS, different bits of the ID experience difference interference. Hence, averaging covariance matrix over each bit does not lead to a consistent estimator, i.e., increasing the number of averaged data does not reduce the mean square error (MSE) of the estimation. The consistent covariance matrix estimator, which exploits the cyclostationarity of the WLPS, would be introduced in Section 4.4.

### 4.3.1 Signal Model

The transmitted DS-CDMA signal by the $k^{th}$ transponder corresponds to:

$$s^k(t) = g_\tau(t) \cdot \sum_{n=0}^{N-1} b^k[n] \cdot g_{T_b}(t - nT_b) \cdot a^k(t - nT_b) \cdot \cos(2\pi f_c t) \tag{4.4}$$

where $N$ denotes the number of bits per ID code (that represents the maximum capacity of the WLPS, which is in the order of $2^N$), $b^k[n]$ denotes the $n^{th}$ bit of transponder $k$'s ID, $T_b = \frac{\tau}{N}$ represents the transponder bit duration, $g_\tau(t)$ and $g_{T_b}(t)$ are rectangular pulses with the duration of $\tau$ and $T_b$, respectively. Here, $a^k(t)$ denotes the spreading code for the $k^{th}$ transponder, i.e.,

$$a^k(t) = \sum_{g=0}^{G-1} C_g^k g_{T_c}(t - gT_b), \quad C_g^k \in \{-1, 1\}, \tag{4.5}$$

where $G$ ($G \leq 2^N$) [4] is the processing gain (code length), $T_c = \frac{T_b}{G} = \frac{\tau}{N \cdot G}$ represents the chip duration, and $g_{T_c}(t)$ is a rectangular pulse with the duration of $T_c$.

With an antenna array mounted on the DBS receiver, the received signal at the DBS (see Fig. 4.3), which is the summation of signals from multiple transponders through multiple paths, corresponds to:

$$\vec{r}(t) = \sum_{k=1}^{K} \sum_{l=0}^{L^k-1} \sum_{n=0}^{N-1} \alpha_l^k \vec{V}(\theta_l^k) b^k[n] g_{T_b}(t - \tau_l^k - nT_b) g_\tau(t - \tau_l^k) \cdot \\ \cdot a^k(t - \tau_l^k - nT_b) \cos(2\pi f_c t + \phi_l^k) + \vec{n}(t), \tag{4.6}$$

[4] Note that $2^N$ is the maximum number of transponders that the system can accommodate.

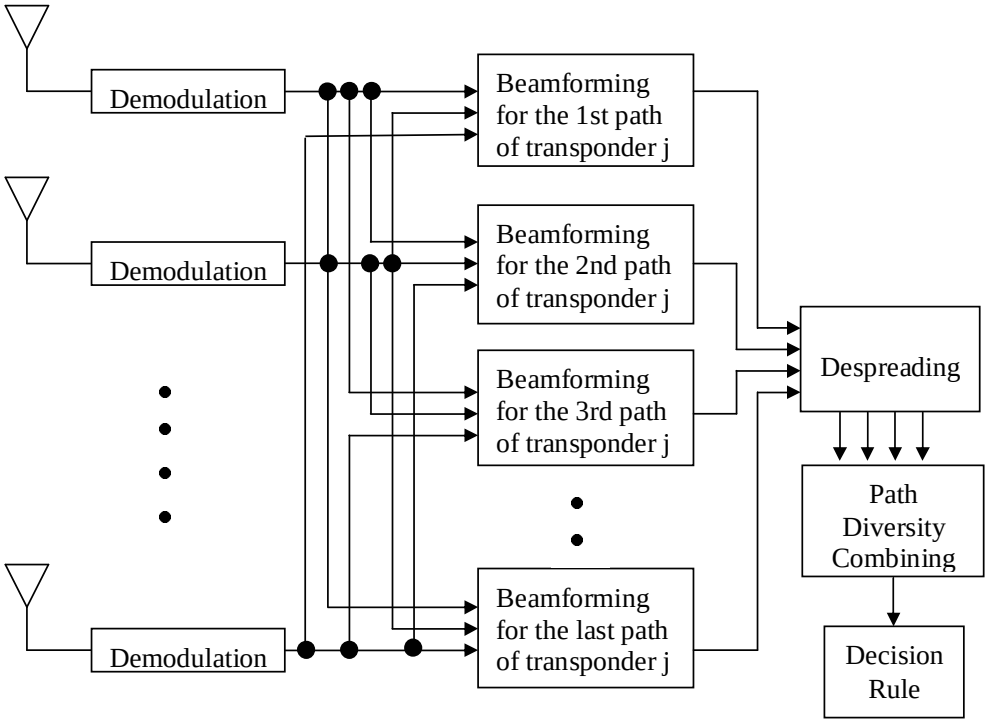


Figure 4.3: DBS receiver implementation via antenna arrays and DS-CDMA systems.

where $K$ denotes the total number of transponders, $L^k$ is the number of paths for the transponder $k$, and $\alpha_l^k$, $\tau_l^k$, $\phi_l^k$ denote the fading factor, time delay and random phase shift for $k^{th}$ transponder's $l^{th}$ path, respectively. Here, for simplicity of presentation, we assume $L^k = L$, $\forall\, k$. $\vec{V}(\theta_l^k)$ denotes the array response vector that corresponds to:

$$\vec{V}(\theta_l^k) = \begin{bmatrix} 1 & \exp(-i \cdot 2\pi d \cos(\theta_l^k)/\lambda) & \cdots & \exp(-i \cdot 2(M-1)\pi d \cos(\theta_l^k)/\lambda) \end{bmatrix}^T . \tag{4.7}$$

Here, $i$ denotes the imaginary unit, $d$ is the spacing between antenna elements, $M$ is the total number of antennas, $(\cdot)^T$ denotes transpose, $\lambda$ denotes the carrier wavelength and $\theta_l^k$ is the direction of $k^{th}$ transponder's $l^{th}$ path. Basically, in (4.7), we assume half wavelength spacing between antennas, and the precise knowledge of array manifold at the DBS receiver.

After demodulation, the $g^{th}$ chip of the $n^{th}$ bit output for the $j^{th}$ transponder's the $q^{th}$ path would correspond to:

$$\vec{y}_q^j[n,g] = \int_{\tau_q^j+nT_b+gT_c}^{\tau_q^j+(n+1)T_b+(g+1)T_c} \vec{r}(t)\cos(2\pi f_c t + \phi_q^j)g(t-\tau_q^j-nT_b-gT_c)\,dt. \quad (4.8)$$

The $g^{th}$ chip of the $n^{th}$ bit output of the beamformer for $j^{th}$ transponder's $q^{th}$ path is given as:

$$z_q^j[n,g] = \vec{W}^H(\theta_q^j)\cdot\vec{y}_q^j[n,g], \quad (4.9)$$

where the weight vector $\vec{W}(\theta_q^j)$ and $\vec{y}_q^j[n,q]$ are both $1\times M$ column vectors, and $H$ denotes Hermitian transpose. Except Each RAKE corresponds to one path. Each path is received from a specific direction. Hence, beamforming on each RAKE is applied to capture the energy from the associated direction.

The receiver in Fig. 4.3 and (4.9) resembles a spatial RAKE-like structure. Here, each RAKE corresponds to one path. Each path is received from a specific direction. Hence, beamforming on each RAKE is applied to capture the energy from the associated direction.

After beamforming, the signals from different paths are combined via Maximal Ratio Combining:

$$z^j[n,g] = \sum_{l=1}^{L}\alpha_l^j z_l^j[n,g] \quad (4.10)$$

Finally, the CDMA despreading is applied and the detected bit is given as:

$$z^j[n] = \sum_{g=1}^{G} z^j[n,g]C_g^j. \quad (4.11)$$

The above description has included all necessary steps of WLPS ID detection process, except the calculation of the weight vector $\vec{W}(\theta_q^j)$ in (4.9), which is the kernel part of this work. Here, we discuss how to determine $\vec{W}(\theta_q^j)$ in Section 4.3.2.

### 4.3.2 Weight Vector Calculation

The conventional beamforming weight vector simply corresponds to:

$$\vec{W}_f(\theta_q^j) = \vec{V}(\theta_q^j). \tag{4.12}$$

Noting that $\vec{V}(\theta_q^j)$ is a predefined linear phase filter, which coincides with the definition of discrete Fourier transform, it is said that the conventional beamforming is equivalent to discrete Fourier transform [Stoica and Moses(1997)].

The LCMV beamforming, which minimizes the total output power, while keeping the desired signal power constant, corresponds to the solution of [Frost(1972)]:

$$\min_{\vec{W}_c(\theta_q^j)} \vec{W}_c^H(\theta_q^j)\mathbf{R}_q^j\vec{W}_c(\theta_q^j) \quad s.t.\ \vec{W}_c^H(\theta_q^j)\vec{V}(\theta_q^j) = 1 \tag{4.13}$$

Using Lagrange Multiplier, the solution of the above equation, i.e., LCMV BF, is given by [Stoica et al.(2003)]:

$$\vec{W}_c(\theta_q^j) = \frac{\mathbf{R}_q^{j^{-1}}\vec{W}_f(\theta_q^j)}{\vec{W}_f^H(\theta_q^j)\mathbf{R}_q^{j^{-1}}\vec{W}_f(\theta_q^j)} \tag{4.14}$$

where $\mathbf{R}_q^j$ is the covariance matrix of $j^{th}$ transponder's $q^{th}$ path's observed signal, i.e., $\mathbf{R}_q^j = E\left[\vec{y}_q^j \cdot \vec{y}_q^{j^H}\right]$.

In this work, precise knowledge of the DOA $\theta_q^j$ and array manifold are assumed, i.e., $\vec{W}_f(\theta_q^j)$ is perfectly known. Then, the only left important implementation issue of the LCMV beamforming is the estimation of $\mathbf{R}_q^j$. In general, the sample covariance matrix estimator corresponds to:

$$\hat{\mathbf{R}}_q^j = \frac{1}{\Gamma}\sum_{n=0}^{\Gamma-1} \vec{y}_q^j[n]\vec{y}_q^{j^H}[n] \tag{4.15}$$

where $\Gamma$, ($\Gamma \in \{1, 2, 3 \cdots N\}$ ), denotes the selected data length for $\mathbf{R}_q^j$ estimation. If $\vec{y}_q^j[n]$ is a stationary and ergodic process, the sample average equals time average, and the sample covariance matrix estimator leads to an accurate estimate of $\mathbf{R}_q^j$. In another word, the sample covariance matrix estimator would be consistent, and increasing the number of data samples reduces the error variance of the sample covariance matrix estimator.

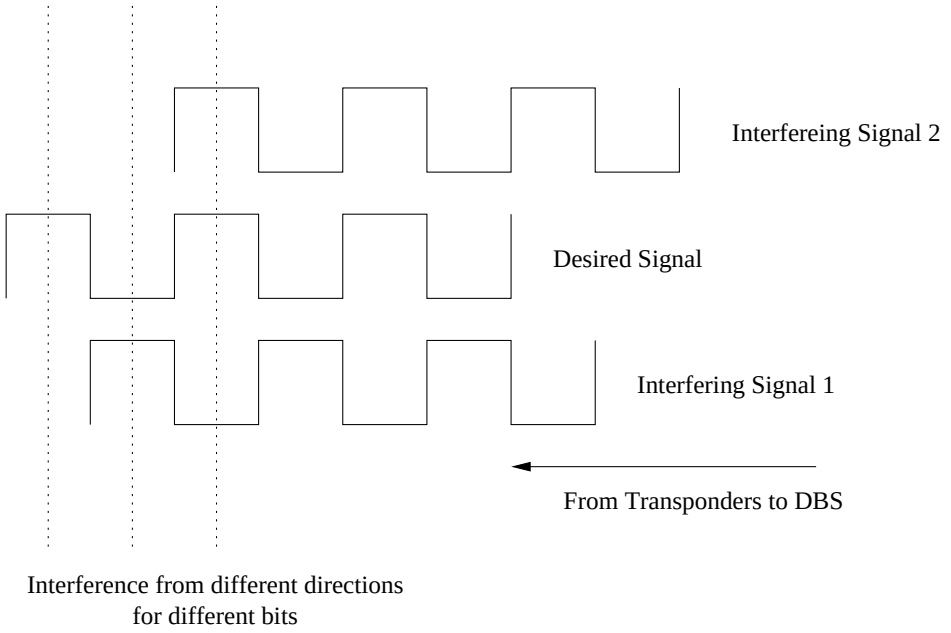


Figure 4.4: Different chips experience different interference.

### 4.3.3 NonStationarity Analysis

Standard wireless communication systems are stationary because of transmission of very long sequences from a large number of users. In other words, in these systems, different chips of the desired signal would experience the same interference. However, because the WLPS transponder transmitted signal is a short burst signal, the interfering signal may only interfere with some, but not all chips of the desired signal (see Fig. 4.4). Hence, the interference changes within each bit of the desired signal. This is especially the case for medium probability-of-overlapping, $p_{ovl}$, values. Therefore, in WLPS, $\mathbf{R_q^j}$ varies for different chips and large selection of $\Gamma$ does not necessarily lead to a high quality of the covariance matrix estimation. To have a better understanding when the received signal is not stationary, we have the following discussion:

- Small values of $d_c$ in (4.1) leads to low $p_{ovl}$ (see Fig. 4.2). In an extreme situation, $p_{ovl} \rightarrow 0$. In this case, since there is no interference at all, $E[\vec{y}_q^j[n]\vec{y}_q^{jH}[n]] =$

$E[\vec{y}_q^j[n+1]\vec{y}_q^{jH}[n+1]]$, and the sample covariance matrix estimator leads to an accurate estimation. However, the main advantage of LCMV beamforming is interference suppression, and in this situation LCMV will not provide better performance than conventional beamforming even with accurate estimation of $\mathbf{R_q^j}$.

- Large values of $d_c$ in dense transponder environment leads to $p_{ovl} \rightarrow 1$. In this case, the sum of interferences becomes would approximately be white noise, and the received signal statistically tends to be stationary, i.e., $E[\vec{y}_q^j[n]\vec{y}_q^{jH}[n]] \simeq E[\vec{y}_q^j[n+1]\vec{y}_q^{jH}[n+1]]$. In this case, the covariance matrix would be an identity matrix and LCMV beamforming becomes equivalent to conventional beamforming.

- Medium $d_c$ values and moderate transponder density lead to a spatial structure for the interference, i.e., several interfering signals are received in different directions. In this case, the received data samples would be non-stationary, large selection of $\Gamma$ does not improve the quality of covariance estimation, and the sample covariance matrix estimator is not consistent.

Fig. 4.5 represents the mean square error (MSE) between the true value and the estimated values of covariance matrix as a measure of non-stationarity, assuming a flat fading channel. The MSE corresponds to:

$$MSE = \sum_{m=1}^{M}\sum_{u=1}^{M}(\mathbf{R}_q^j(m,u) - \hat{\mathbf{R}}_q^j(m,u))^2 \tag{4.16}$$

where $M$ is the number of antenna array elements, $\mathbf{R_q^j}$ and $\hat{\mathbf{R}}_q^j$ denote the true and estimated covariance matrixes via sample covariance matrix, respectively. The direction and distance of the transponders are assumed to be uniformly distributed in

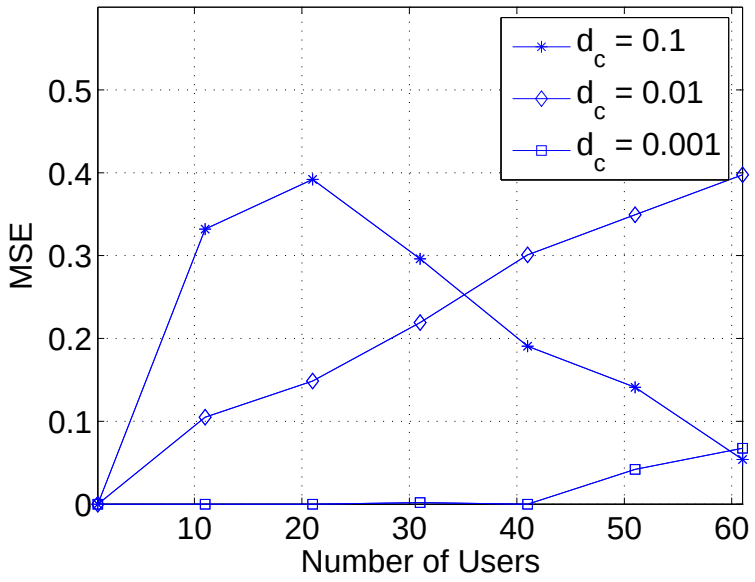


Figure 4.5: Simulation Results: The Mean Square Error of estimated covariance matrix by standard estimation method.

$[0, \pi]$ and $[0, R_{max}]$. The estimated covariance matrix is normalized before comparing it with true covariance matrix.

It is seen that when $d_c$ is small ($= 0.001$) and the number of transponder is small ($< 30$), the MSE is kept minimal, which is consistent with the first case discussed above.

When $d_c$ is large ($= 0.1$) and the number of transponder is large ($> 60$), the MSE is small as well. This corresponds to the second case discussed: A large number of interferences leads to a spatially white structure. In other words, every chip is interfered by signals in many directions. Hence, the interference over different chips would be similar, which leads to a stationary process.

When $d_c$ is moderate ($= 0.01$), the MSE is large, i.e., the non-stationarity problem is severe.

The high MSE shown in Fig. 4.5 leads to low probability-of-detection. As a result, directly applying LCMV beamforming does not improve the system performance

compared to a conventional beamforming. This point is verified by ID detection simulations in Fig. 4.10 (see Section 4.5).

## 4.4 Estimator Based on the Cyclostationarity

Section 4.3.3 introduced the the non-stationarity problem in the WLPS. This section proposes a modified covariance matrix estimator to solve the non-stationarity problem, which exploits the cyclostationarity property of the WLPS.

### 4.4.1 New Estimator Via Cyclostationarity

The non-stationarity is mainly generated by the non-continuous transmission of transponders. However, it should be noted that, in addition to the non-continuousness, the transmission is also periodical. In every period, a transponder retransmits the same ID bits with the same spreading code. Now, assuming all transponders' directions and distances remain the same for a number of periods, same chips of transponder ID in different period experience the same interference (See Fig. 4.6). Here, the period of transponder transmission is called ID request time (IRT).

The repetition property of transponder transmission is also known as cyclostationarity: although different chips in the same period does not experience same interferences, same chips in different periods experience same interferences. Hence, it is possible to apply beamforming to each chip, if the covariance matrix for each chip can be estimated. As shown in Fig. 4.6, the covariance matrix estimation via cyclostationarity for the $g^{th}$ chip of the $n^{th}$ bit corresponds to:

$$\hat{\mathbf{R}}_q^j[n,g] = \frac{1}{\Omega}\sum_{\omega=1}^{\Omega} \vec{y}_q^j[n,g,\omega]\vec{y}_q^{jH}[n,g,\omega] \tag{4.17}$$

where $\Omega$ denotes the number of period within which the cyclostationarity holds. Using (4.17), consequently (4.8) and (4.9) would correspond to:

$$\vec{y}_q^j[n,g,\omega] = \int_{\tau_q^j+nT_b+gT_c+(\omega-1)IRT}^{\tau_q^j+(n+1)T_b+(g+1)T_c+(\omega-1)IRT} \vec{r}(t)\cos(2\pi f_c t + \phi_q^j)\cdot$$

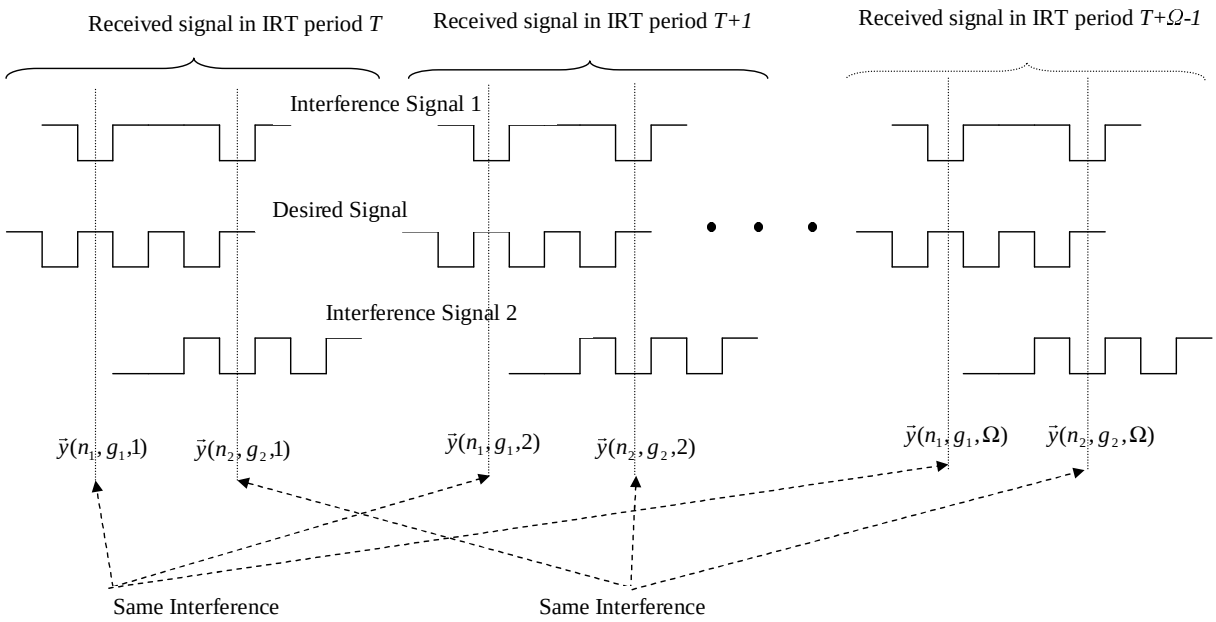


Figure 4.6: Same Chips in different IRT periods have the same interference

$$g(t - \tau_q^j - nT_b - gT_c - \omega IRT)\, dt, \quad \omega \in \{1, 2, \cdots, \Omega\}, \tag{4.18}$$

and

$$z_q^j[n, g] = \frac{1}{\Omega} \sum_{\omega=1}^{\Omega} \vec{W}^H(\theta_q^j) \cdot \vec{y}_q^j[n, g, \omega], \tag{4.19}$$

respectively. (4.19) reflects both beamforming and equal gain time diversity combining processes. Because each frame experiences independent fading, we also achieve time diversity benefits via combining the chips from different IRT. The receiver structure via cyclostationarity is shown in Fig. 4.7. Here, a separate block is considered for the covariance matrix estimator via cyclostationarity, since the new estimator requires a temporary storage of the received signals.

It should be noted the proposed consistent covariance matrix estimator may not be restricted to LCMV beamforming, various optimal [Xia and Giannakis(2006)] or robust beamforming [Lorenz and Boyd(2005)][Vorobyov et al.(2004)] methods may also use this estimator. In this chapter, the application of LCMV beamforming in the WLPS is introduced. The proposed concept may be easily extended to any signal

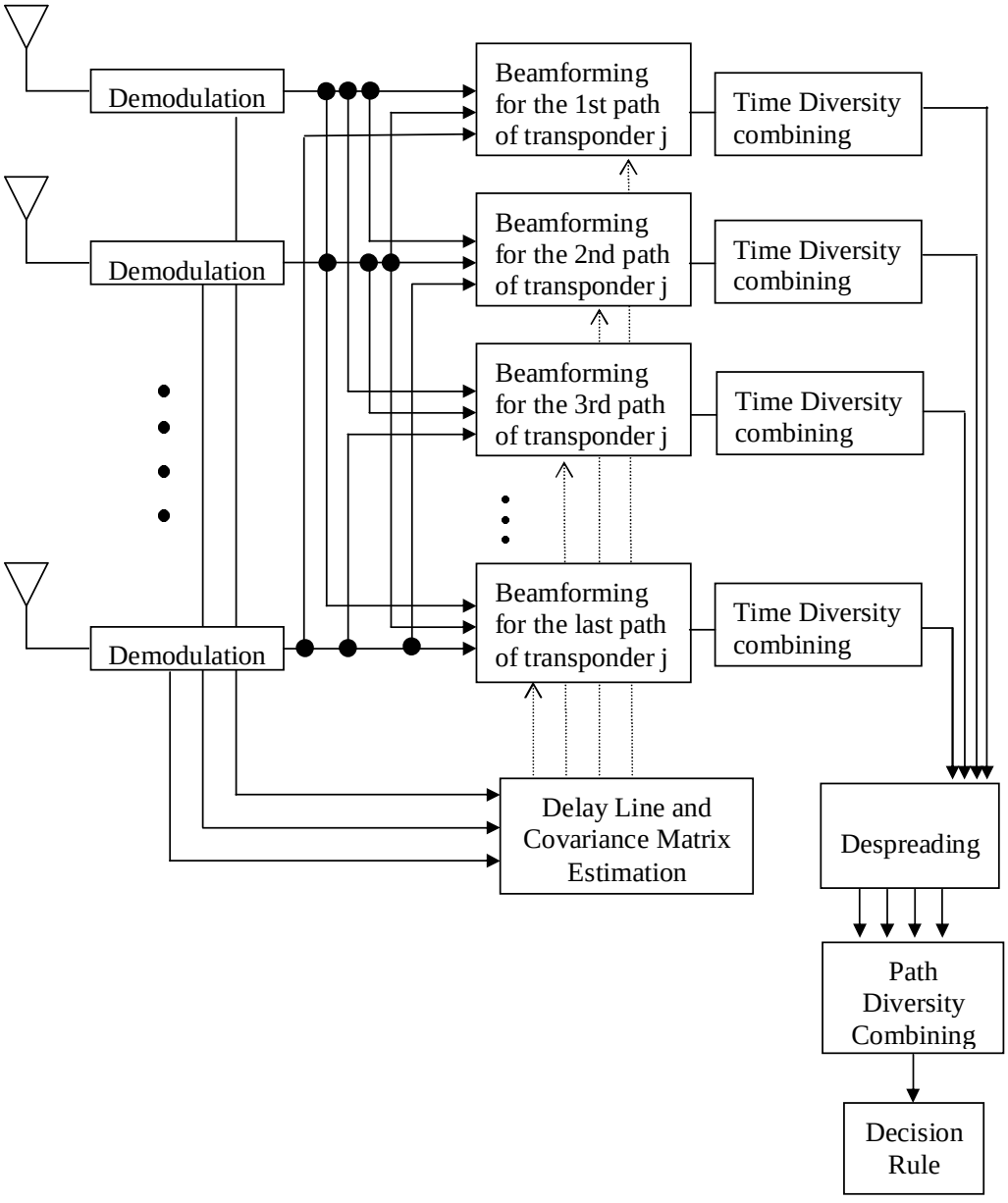


Figure 4.7: Receiver Structure with using cyclostationarity

processing algorithm that requires an estimation of covariance matrix, as long as the system exhibits a repetitive nature.

### 4.4.2 Cyclostationarity Duration

An important issue of the new estimator is the maximum possible value of $\Omega$, i.e., the number of periods that the cyclostationarity holds. A larger value of $\Omega$ leads to better estimation, while a small value of $\Omega$ (e.g., 1 or 2) will render the estimator via cyclostationarity improper.

#### Cyclostationarity Duration for A Single Transponder

Basically, $\Omega$ is determined by IRT and the duration within which the cyclostationarity remains available, and corresponds to:

$$\Omega \leq \frac{T_{cy}}{IRT}, \tag{4.20}$$

where $T_{cy}$ is the time within which cyclostationarity condition holds, and $IRT$ denotes the repetition time of the ID request signal. Two parameters impact the cyclostationarity: The direction and the distance of transponder. Hence, the $T_{cy}$ is the time within which: (a) the direction of the transponder approximately remains constant, *and* (b) the distance of the transponder approximately remains unchanged (see Fig. 4.8).

Therefore, we consider the impact of the movement of the transponder in two directions. The first is in the direction that is parallel to the line connecting transponder and antenna array. In this direction, the variation of the TOA within the duration of $T_{cy}$ should be much smaller than the chip duration $T_{ch}$, i.e., TOA is relatively fixed during $T_{cy}$, which corresponds to:

$$T_{cy} \ll \frac{c}{B \cdot v_{\|}}, \tag{4.21}$$

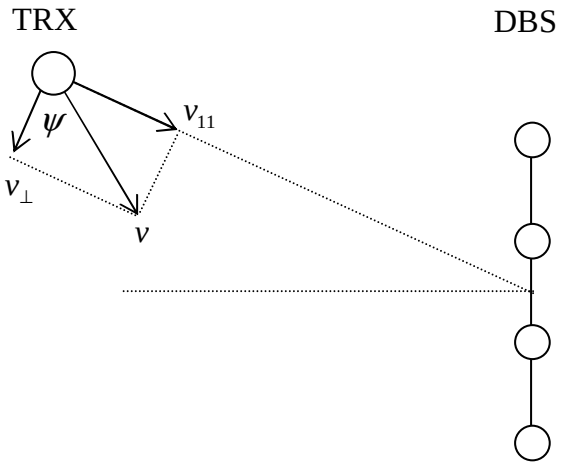


Figure 4.8: Relationship between $v$, $v_\perp$ and $v_\parallel$

where $c$ is the speed of light, $B = 1/T_{ch}$ denotes the transponder signal bandwidth, and $v_\parallel$ represents the doppler velocity of the transponder;

The second direction is the direction that is perpendicular to the line connecting transponder and antenna array. In this direction, the variation of DOA should be much smaller than the antenna array half power beamwidth, i.e., DOA is relatively fixed during $T_{cy}$, which corresponds to:

$$T_{cy} \ll \frac{\frac{\theta_B}{2} \cdot d}{v_\perp}, \tag{4.22}$$

where $\theta_B$ is the half power beam width, $d$ denotes the distance between transponder and DBS, and $v_\perp$ is depicted in Fig. 4.8.

Combining the above two conditions, the final condition corresponds to:

$$T_{cy} \ll min(\frac{c}{B \cdot v_\parallel}, \frac{\theta_B \cdot d}{2v_\perp}) \tag{4.23}$$

Note that the first condition (TOA constraint) is independent of distance, while the second condition (DOA constraint) depends on both velocity and distance.

Equivalent to (4.23), we have the conditions for cyclostationarity doppler frequency, which corresponds to:

$$f_{cy} = \frac{1}{T_{cy}} \gg max(\frac{B \cdot v_\parallel}{c}, \frac{2v_\perp}{\theta_B \cdot d}) \tag{4.24}$$

This means that the changing rate of cyclostationarity should be much larger than DOA/TOA changing rate.

Knowing $v_{\parallel} = v \cdot \cos(\psi)$ and $v_{\perp} = v \cdot \sin(\psi)$ (see Fig. 4.8), and considering $\psi$ a uniform random variable within 0 and $2\pi$, the *cyclostationarity doppler spread* ($B_{cy,d}$), which is the Root-Mean-Square (RMS) value of cyclostationarity doppler frequency, corresponds to:

$$B_{cy,d} = max(\sqrt{Æ\left[\frac{B \cdot v_{\parallel}}{c}\right]^2}, \sqrt{Æ\left[\frac{2v_{\perp}}{B_d \cdot d}\right]^2}), \tag{4.25}$$

where $Æ(\cdot)$ denotes expectation operation.

Applying simple mathematical manipulations, (4.25) would correspond to:

$$B_{cy,d} = max(\frac{B \cdot v}{\sqrt{2}c}, \frac{\sqrt{2}v}{\theta_B \cdot d}). \tag{4.26}$$

Then, using (4.26) and similar to the definition of channel coherence time, we define the cyclostationarity coherence time as [Rappaport(2002)]:

$$T_{cy,c} \cong \frac{1}{B_{cy,d}} \tag{4.27}$$

In order to guarantee cyclostationarity during $T_{cy}$, $T_{cy}$ should be selected smaller than $T_{cy,c}$, or:

$$T_{cy} < T_{cy,c} \tag{4.28}$$

To demonstrate the effects of the two conditions on cyclostationarity, the cyclostationarity coherence time with various velocity and distance values has been computed in Fig. 4.9. Here, we assume $300\,MHz$ bandwidth and 27° half power beamwidth (consistent with four antenna elements). The first area of interest in Fig. 4.9 is low-velocity and short-range area, which is mainly suitable for applications such as indoor and airport security. Note that the cyclostationarity doppler spread varies with distance in this area. Hence, we can conclude that for short range applications,

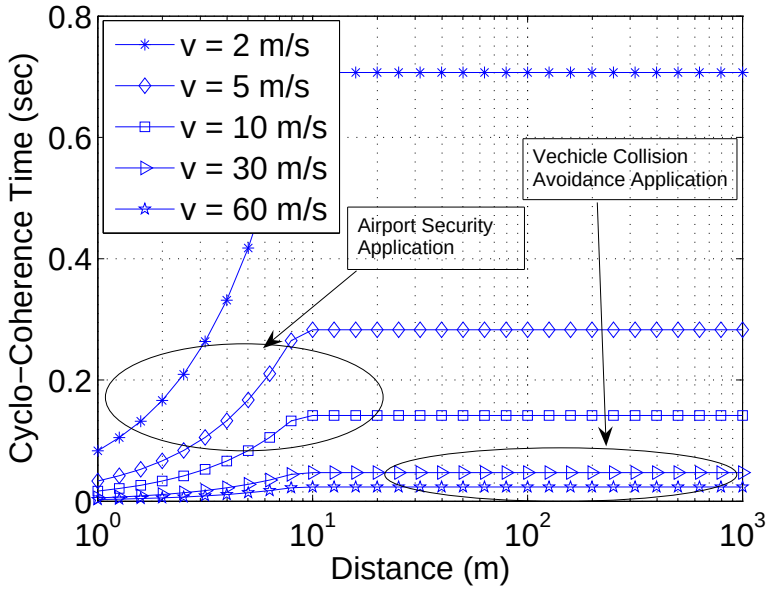


Figure 4.9: Cyclostationarity Coherence Time for Different Applications, Single transponder

DOA would be the dominant condition for cyclostationarity. The second area of interest is high-velocity, long-range area, which is mainly suitable for vehicle collision avoidance system. Note that the cyclostationarity doppler spread is independent of distance in this area. We can conclude that for long range applications, the main constraint is the rate of change of TOA.

## 4.5 Numerical Results

In this section, we use Monte-Carlo simulations to evaluate the ID detection performance of the WLPS system implemented via LCMV beamforming, with and without the newly proposed covariance matrix estimator via the cyclostationarity property. Here, we consider a multi-transponder, multi-path environment. For simulation purposes, we assume:

1. The ID code has 6 bits ($N = 6$);

2. The DS-CDMA code has 64 chips ($G = 64$);

3. Channel delay spread for a typical street area is $27\,nsec$ [Arowojolu et al.(1994)];

4. Carrier frequency $= 3\,GHz$, $\tau_{TRX} = 1.2\,\mu s$, and $\tau_{DBS} = 24\,\mu s$;

5. The antenna array is linear with 4 elements, and element spacing $d = \frac{\lambda}{2} = 0.05\,m$ (half power beamwidth $= 27^{\circ}$);

6. Four multipaths lead to $L = 4$ fold path diversity;

7. The transponder distance and angle are uniformly distributed in $[\,0 \quad 1\,]\,km$ and $[\,0 \quad \pi\,]$, respectively;

8. Uniform multi-path intensity profile, i.e., bit energy is distributed in each path identically;

9. Binary Phase Shift Keying (BPSK) modulation; and,

10. Perfect power control and DOA/TOA estimation.

The above assumptions are particularly suitable for vehicle safety applications. Based on the assumed setup, transponder signal TOA is uniformly distributed in $[\,T_d \quad T_{max}\,]$, at the DBS receiver. Assuming $T_d << T_{max}$, approximately TOA of transponder signal is uniformly distributed in $[\,0 \quad T_{max}\,]$, and the required bandwidth of a DS-CDMA transponder transmitter is $320MHz$. Using these parameters, $IRT_{min} = 12\,\mu s$, then the duty cycle for DBS receivers would correspond to $d_{c,DBS} \simeq 0.1$, which leads to a high probability-of-overlapping (see Fig. 4.2).

Assuming the vehicle speed is $30\,m/s$, the cyclostationarity coherence time (based on an average distance of $500\,m$) would be $47.1\,ms$, as shown in Fig. 4.9. As we mentioned in Section 4.2, usually $IRT$ is selected much larger than $IRT_{min}$ in order to reduce interference power at transponder receiver. Here, we select $IRT = 1.2\,ms$. Using (4.28), $T_{cy} \cong T_{cy,c}/5 = 9.42\,ms$, and using (4.20), finally $\Omega \cong 8$. In other words,

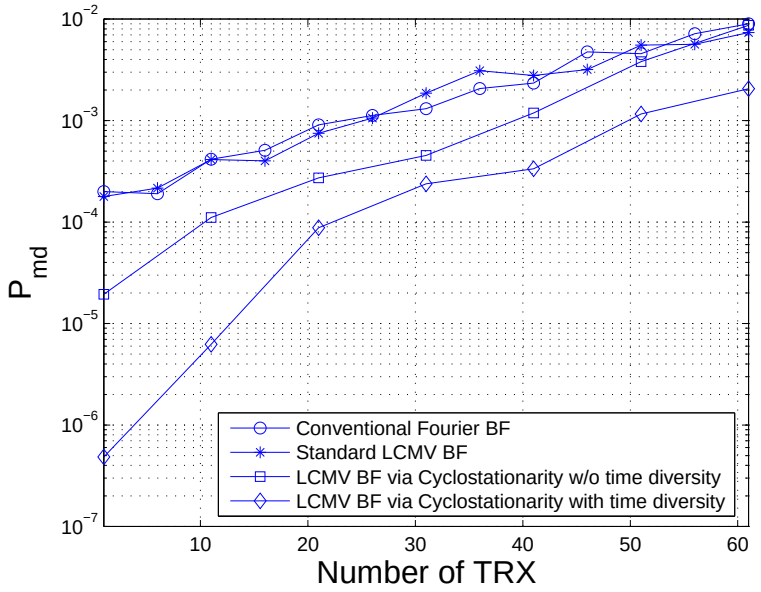


Figure 4.10: LCMV BF result with using cyclostationarity

within 8 IRT frames, the conditions for cyclostationarity would well exist. It should be mentioned that the conditions simulated in this chapter leads to a conservative selection of $\Omega$, and in many applications, higher value than $\Omega = 8$ is expected.

The simulation results are shown in Fig. 4.10. The measurement of ID detection performance is Probability-of-Miss-detection ($P_{md}$), i.e., the probability that the ID of the desired transponder is not detected correctly. Here, $P_{md} = 1-(1-P_d)^N$, where $N$ the number of bits per ID, and $P_d$ denotes the probability that one bit of the ID is detected correctly. As discussed in Section 4.3.3, due to nonstationarity nature of the WLPS, traditional sample covariance matrix estimator computation leads to a high probability of miss detection. It can also be seen that the performance of LCMV BF with the covariance matrix estimator via cyclostationary property leads to a significantly improved performance compared to the standard covariance matrix estimator. It is observed that the proposed technique doubles the capacity of this system at the $P_{md} = 10^{-3}$ (i.e., from 25 to 50).

The result not only benefits from solving non-stationarity problem, but also the time diversity attained over the 8 IRT periods, since the fading is assumed to be independent over chips in different frames (IRT). This diversity improves the performance in conjunction with cyclostationarity. In order to demonstrate the different effects of time diversity combining and optimum beamforming, we also perform the optimum beamforming without using time diversity combining. It can be seen that both of the two techniques contributes to DBS receiver performance. It is observed that as the number of transponders increases, the time diversity has a dominant impact on the performance improvement.

## 4.6 Conclusion

This chapter proposes a novel covariance matrix estimator, which is the critical step for optimal beamforming implementation, in a wireless local positioning system with a periodic signaling structure. Different from standard wireless systems, the standard sample covariance matrix estimator is not consistent in the WLPS system due to its non-stationarity nature. A new consistent estimator, which exploits the cyclostationarity property of the WLPS, is proposed. It is demonstrated that in most application, the cyclostationarity duration is sufficiently large for covariance matrix estimation. Numerical simulations verify that the new estimator improves system performance significantly.

# Chapter 5

# Stochastic Correlated MIMO Channel Modeling

Recent Multi-Input-Multi-Output (MIMO) research demonstrates that real wireless environments exhibit spatially correlated fadings: The independent fading assumption does not hold in all cases. Accordingly, spatially correlated MIMO channel models have attracted considerable attention. The Kronecker product form (KPF) is the currently most widely used MIMO channel model describing spatial correlations. However, its validity in *various* environments have not been well examined.

In this work, first the correlation structure of the KPF is discussed. It is shown that the KPF exhibits a separable correlation structure. Second, we find the physical environments that correspond to separable correlation structure. It is claimed that the separable correlation structure corresponds to separable scatterers in physical environment.

Note that the separable scatterers present in macrocell environments, but they do not in microcell environment. Hence, the KPF is applicable in macrocell environments, but it is not in microcell environments. Numerical simulations, which is based on the ray tracing approach with geometry based channel model, confirm the above conclusion.

A number of stochastic MIMO channel models have been proposed, and each of them is supported by different experimental results. This chapter clarifies the

relationship between various models, and provides a clear, unified framework for understanding spatially correlated MIMO channel modeling.

The most comprehensive description of MIMO fading correlations is called the "total correlation matrix". It is found that, for different channel models, the eigen structure of the total correlation matrix would be different. W-model has separable eigen basis, virtual channel representation predefines fixed and separable eigen basis, and Kronecker product form requires separable eigen basis and separable eigen values.

Moreover, it is noted that different eigen structures correspond to different physical environments, i.e., different stochastic channel models are suitable in different physical environments. A detailed investigation reveals that W-model assumes separable transmitter and receiver; virtual channel representation is suitable for large dimensional arrays; and Kronecker product form works in macrocell environments.

## 5.1 Introduction

In the past decade, Multiple-Input-Multiple-Output (MIMO) systems have attracted significant attention in wireless communications due to its great potential to increase channel capacity [Telatar(1999)]. Originally, MIMO research assumes fadings between different antenna elements are completely independent [Alamouti(1999)]. However, on going studies confirm that the assumption of channel independence is not always satisfied in real applications [Shiu et al.(2000)]. Hence, spatially correlated MIMO channel modeling becomes the interest of many researchers recently [Yu and Ottersten(2002)].

Spatially correlated MIMO channel modeling mainly deals with two dimensional (2-D) correlated random variables generation. The most famous one of such models is the Kronecker product form (KPF) [Kermoal et al.(2001)]. The KPF is a natural extension of the traditional one dimensional (1-D) correlated random variables generation. Consequently, the KPF has been widely accepted and extensively adopted

in most MIMO studies [Shin et al.(r in)]. However, it should be noted that the KPF has its own limitation: It is suitable only in certain environments.

In this work, first the correlation structure of the KPF is discussed. It is shown that the KPF possesses separable correlations [Chuah et al.(2002)]. Next, through a simple ray tracing approach, it is demonstrated that real environments does not necessarily exhibit separable correlations. Then, in which specific environment, correlations would be separable?

Here, it is claimed that separable correlation corresponds to separable scatterers. Noting that in macrocell and microcell environments, the scatterers form two rings and eclipse shapes, respectively [Ertel et al.(1998)], it is said that macrocell environments has separable scatterers while microcell environments does not. Hence, the KPF is suitable for macrocell, but it is not applicable for microcell. Numerical simulations confirm this conclusions. Finally, a generalization of the KPF for Line-of-Sight considerations is discussed.

An example of macrocell environment is the cellular mobile system, while the example of microcell environment would be the Wi-Fi (IEEE 802.11) systems. Hence, the result in this chapter has a considerable impact on MIMO system design in various applications, such as cellular or Wi-Fi systems.

In the past decade, Multiple-Input-Multiple-Output (MIMO) systems have attracted significant attention in wireless communications research due to their great potential to increase channel capacity [Telatar(1999)] [Foschini and Gans(1998)]. Originally, MIMO research assumes fadings across different antenna pairs are completely independent [Alamouti(1999)] [Tarokh et al.(1998)]. However, recent studies confirm that the assumption of channel independence is not always satisfied in real applications [Shiu et al.(2000)]. Hence, after year 2000, researchers initiated the investigation of spatially correlated MIMO channels [Kang and Alouini(2006)] [Shin et al.(r in)].

The most widely used spatially correlated MIMO channel model is the Kronecker product form (KPF) [Kermoal et al.(2001)]. However, the KPF has been questioned by many researchers [Oestges et al.(2005)], both experimentally [Ozcelik et al.(2003)] and theoretically [Abdi and Kaveh(2002)]. Therefore, after the introduction of the KPF, a number of other spatially correlated MIMO channel models have been proposed. Examples of these techniques include the virtual channel representation (VCR) [Sayeed(2002)], and W-model [Weichselberger et al.(2006)]. This work focuses on the above three models. Other MIMO channel models, e.g., Gesbert model [Chizhik et al.(2002)] [Gesbert et al.(2002)] and Muller's finite scatterer model [Debbah and Muller(2005)], are beyond the scope of this dissertation.

Surprisingly, all of the three models are supported by different experiments [Almers et al.(2007)] [Yu and Ottersten(2002)] Therefore, all of them seem equally convincing. Then, for MIMO system designers, it is hard to choose a proper model, since none of them seems superior to another. In this work, it is shown that each channel model is suitable for describing only a certain physical environment. Hence, the selection of channel model should depend on applications. Because space-time signal processing algorithm varies with channel modeling, this work may lead to the first explicit and general guidance on how to design MIMO systems based on physical environments.

We first introduce a unified framework that accommodates all MIMO channel models: the total correlation matrix, which is always a semi-positive definite Hermitian Toeplitz matrix. Then, it is demonstrated that for previously proposed channel models, the total correlation matrix possesses some special eigen structures. W-model has separable eigen basis; the VCR assumes fixed separable eigen basis; and the KPF requires separable eigen basis and separable eigen values.

Second, we point out that different eigen structure corresponds to different physical environment: W-model assumes separable transmitter and receiver; the VCR is

suitable for large dimensional arrays; and the KPF works with macrocell environments.

Finally, although the total correlation matrix framework is good in many environments, it is found that it can not describe all physical environments. For example, the keyhole problem, which is depicted by Gesbert model, is beyond the correlation considerations. Hence, Gesbert model can not be included in the total correlation matrix framework.

The mathematical notations in this chapter are: Bold capital letter denotes matrix, small letter with arrow on its head denotes vector. Moreover, $(\cdot)^\top$ and $(\cdot)^\dagger$ denotes transpose and conjugate transpose, respectively; $(\cdot)^*$ denotes element-wise conjugate operation; Æ$[\cdot]$ denotes element-wise expectation operation; $\otimes$ denotes Kronecker product; vec$(\cdot)$ denotes vectorization operation, i.e., if $\mathbf{H} = [\vec{h_1}\, \vec{h_2} \cdots \vec{h_N}]$, then vec$(\mathbf{H}) = [\vec{h_1}^\top\, \vec{h_2}^\top \cdots \vec{h_N}^\top]^\top$; $\odot$ denotes element-wise matrix multiplication; $\mathcal{D}(\vec{h})$ denotes the diagonal matrix whose diagonal elements corresponds to the vector $\vec{h}$.

The rest of the chapter is organized as follow: Section 5.2 discusses Kronecker Product Form, which is the most widely used model currently. Its correlation structure and feasible physical environments are both analyzed and simulated. Section 5.3 discusses a relatively new channel model: Virtual Channel Representation. The correlation coefficients between channel fadings in VCR model is analytically derived, then it is shown that VCR possess a non-separable correlation structure. Section 5.4 introduces the unified framework for all the MIMO channel models, KPF, VCR and W-model. Their correlation structure and relevant physical environments are compared and summarized. Section 5.5 conducts a preliminary error performance analysis with VCR model; Section 5.6 depicts the keyhole problem and Gesbert model, which is beyond the considerations of correlation structure; Section 6.5 concludes this chapter.

## 5.2 The Kronecker Product Form

The input-output relationship in MIMO systems corresponds to:

$$\vec{y} = \mathbf{H}\vec{x} + \vec{n}, \tag{5.1}$$

where $\vec{y}$, $\vec{n}$ and $\vec{x}$ are column vectors with dimension $N_R$, $N_R$, and $N_T$, respectively. The goal of MIMO channel modeling is to find the structure of $\mathbf{H}$, which is a $N_R \times N_T$ complex matrix. $N_T$ denotes the number of transmit antennas, and $N_R$ denotes the number of receive antennas.

Assuming Rayleigh fading, the elements of $\mathbf{H}$ are zero-mean circularly symmetric [Telatar(1999)] complex Gaussian random variables (r.v.). For correlated fadings, the elements of $\mathbf{H}$ would be correlated. The most famous model that tries to describe the correlations is the Kronecker product form (KPF).

The KPF proposes that the fading matrix, $\mathbf{H}$, is [Kermoal et al.(2001)]:

$$\mathbf{H} = \mathbf{G}_R^{\frac{1}{2}} \mathbf{H}_{iid} \mathbf{G}_T^{\frac{1}{2}\top}. \tag{5.2}$$

Here, $\mathbf{H}_{iid}$ represents a $N_R \times N_T$ matrix consisting of independent and identically distributed (i.i.d.) complex Gaussian random variables (r.v.). The matrices $\mathbf{G}_T^{\frac{1}{2}}$ and $\mathbf{G}_R^{\frac{1}{2}}$ satisfy:

$$\mathbf{G}_T^{\frac{1}{2}} \mathbf{G}_T^{\frac{1}{2}\dagger} = \mathbf{G}_T \quad \text{and} \quad \mathbf{G}_R^{\frac{1}{2}} \mathbf{G}_R^{\frac{1}{2}\dagger} = \mathbf{G}_R, \tag{5.3}$$

where $\mathbf{G}_T$ and $\mathbf{G}_R$ are transmit and receive correlation matrices, with dimensions $N_T \times N_T$ and $N_R \times N_R$, respectively.

### 5.2.1 Correlation Structure

The name "Kronecker product form" is due to the fact that the total correlation matrix, $\mathbf{G}$, is the Kronecker product of $\mathbf{G}_R$ and $\mathbf{G}_T$. The total correlation matrix is defined as:

$$\mathbf{G} = \text{Æ}\left[\text{vec}(\mathbf{H}) \cdot \text{vec}(\mathbf{H})^\dagger\right], \tag{5.4}$$

where vec$(\cdot)$ represent vectorization operation, i.e., if

$$\mathbf{H} = \left[\vec{h}_1\, \vec{h}_2, \cdots, \vec{h}_{N_T}\right], \tag{5.5}$$

then

$$\mathrm{vec}(\mathbf{H}) = \left[\vec{h}_1^{\top}\, \vec{h}_2^{\top}, \cdots, \vec{h}_{N_T}^{\top}\right]^{\top}. \tag{5.6}$$

Roughly speaking, the vectorization operation "stacks" all columns of $\mathbf{H}$ into a single column. Effectively the the vectorization operation converts two dimensional matrix to an one dimensional vector.

$\mathbf{G}$ is said to be a sufficient description of the fading matrix $\mathbf{H}$, since the joint probability density function of elements in $\mathbf{H}$ can be found once the matrix $\mathbf{G}$ is known [Papoulis and Pillai(2002)]. For the KPF, the total correlation matrix corresponds to:

$$\mathbf{G} = \mathbf{G}_T \otimes \mathbf{G}_R. \tag{5.7}$$

Here, $\otimes$ denotes Kronecker product. For example, the Kronecker product of two matrices $\mathbf{A}$ and $\mathbf{B}$ would be:

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \cdots \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}, \tag{5.8}$$

where $a_{11}, a_{12}, a_{21}, a_{22}, \cdots$ are the elements of the matrix $\mathbf{A}$.

Note that the elements of $\mathbf{G}$ are correlations between fadings. If the elements of $\mathbf{G}$ can be arbitrarily selected, then the correlation between fadings would be a function of four parameters, i.e.,

$$Æ\left[h_{pq} h_{st}^{*}\right] = r(p, q, s, t), \tag{5.9}$$

where $h_{pq}$ denotes the element of $\mathbf{H}$ at $p^{th}$ row and $q^{th}$ column, $h_{st}$ is similarly defined.

However, according to (5.7), the elements of $\mathbf{G}$ can not be selected arbitrarily: they are structured. Consequently, $Æ[h_{pq}h_{st}^*]$ would not be a function of four parameter, it would possess some special property. The first property is stationary:

$$Æ[h_{pq}h_{st}^*] = r(p-s, q-t), \tag{5.10}$$

i.e., the correlation between fadings are determined by their difference, but not their absolute position.

In addition to the stationarity property, the KPF has a finer property: separable stationarity property. After some manipulation of (5.7), the correlation between fadings would be:

$$\boxed{Æ[h_{pq}h_{st}^*] = r_R(p-s) \cdot r_T(q-t)} \tag{5.11}$$

Here, $r_R(\cdot)$ and $r_T(\cdot)$ are two arbitrary auto correlation sequences with maximum lags $N_R - 1$ and $N_T - 1$, respectively. (5.11) presents that the correlation can be separated to two parts: transmitter part and receiver part, and both of the two parts are stationary.

It is noted that if the KPF holds, then (5.11) holds. On the other hand, if (5.11) holds, then the KPF holds. Therefore, (5.11) *will be the tool to check the validity of KPF in various environments.* If in one specific environment, (5.11) holds, then the KPF is suitable in this environment. Otherwise, the KPF is not capable of describing this environment.

In section 5.2.2, we check that if (5.11) is satisfied in a general environment. No any restriction on the positions of scatterers is superimposed in section 5.2.2. It will be seen that in general that the KPF is not applicable. In section 5.2.3 and 5.2.4, the scatterers are placed on eclipse and two rings, respectively. Overall, the results in section 5.2.3 and 5.2.4 show that the KPF works only with separable scatterers, i.e., if scatterers are locally clustered around transmitters and receivers, respectively, then the KPF is applicable, otherwise it is not.

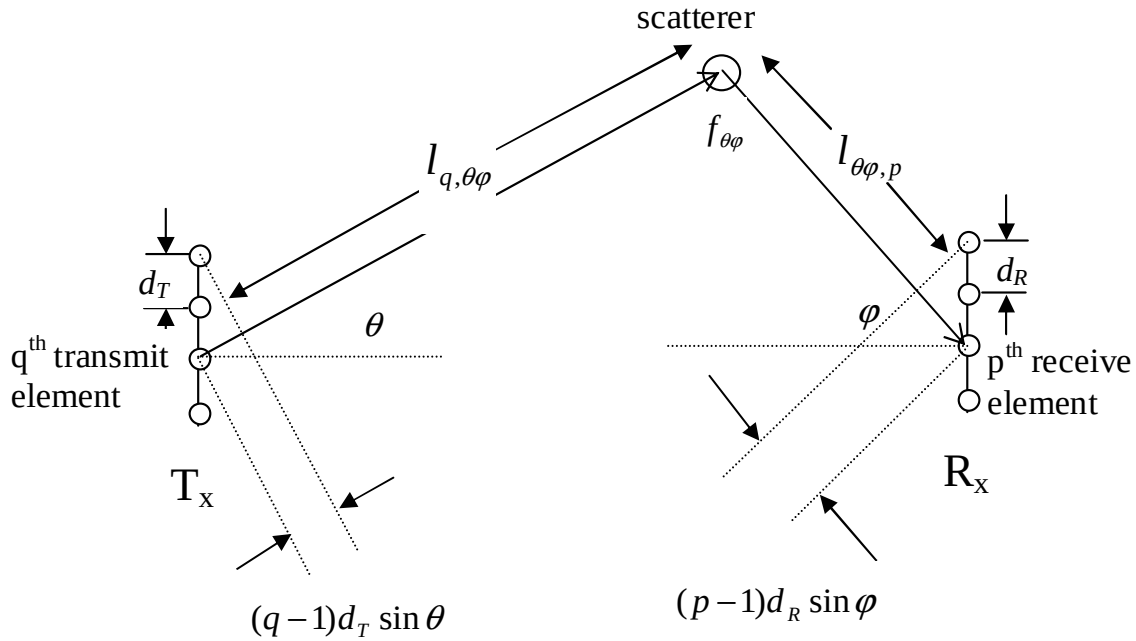


Figure 5.1: Geometry for fading computation

### 5.2.2 General Environments

Here, we assume: 1) the scatterers are pervasively distributed in the space, 2) a single bounce channel model, 3) the distance between scatterer and antenna array is significantly larger than the antenna array size, 4) all scatterers reflects the incident energy omni-directionally, i.e., isotropic fading, and 5) receiver collects all transmitted energy.

Using a technique similar to [Byers and Takawira(2004)], the fading between $q^{th}$ transmitter element and $p^{th}$ receiver element corresponds to:

$$h_{pq} = \lim_{\Delta\theta\to 0}\lim_{\Delta\phi\to 0}\frac{\sqrt{\Delta\theta\,\Delta\phi}}{2\pi}\sum_{\theta=0}^{2\pi}\sum_{\phi=0}^{2\pi} f_{\theta\phi}\cdot\exp(j\cdot(l_{q,\theta\phi} + (q-1)d_T\sin\theta + (p-1)d_R\sin\psi + l_{\theta\phi,p})\cdot\frac{2\pi}{\lambda}). \quad (5.12)$$

The corresponding geometry of this equation is shown in Fig. 5.1, where $f_{\theta\phi}$ denotes the complex reflection gain of the scatter in the direction $(\theta,\phi)$, $l_{p,\theta\phi}$ and $l_{\theta\phi,q}$ represent the distance from the scatter $f_{\theta\phi}$ to the transmitter and receiver, respectively,

and, $d_T$ and $d_R$ represents the antenna element spacing at transmitter and receiver, respectively.

Using (5.12), the correlation between fadings corresponds to:

$$\text{Æ}\left[h_{pq}h_{st}^{*}\right] = \lim_{\Delta\theta\to 0}\lim_{\Delta\psi\to 0}\frac{\Delta\theta\,\Delta\psi}{(2\pi)^2}\sum_{\theta=0}^{2\pi}\sum_{\psi=0}^{2\pi}E\left[\left|f_{\theta\phi}^{2}\right|\right]\cdot$$
$$\cdot\exp(j((q-t)d_T\sin\theta+(p-s)d_R\sin\psi)\frac{2\pi}{\lambda}). \tag{5.13}$$

Please note that in (5.13), the absolute distances $l_{q,\theta\phi}$ and $l_{\theta\phi,p}$ in (5.12) are canceled out, which means the channel statistics are determined by the scatter's direction and antenna spacing, but independent of the absolute distance $l_{\theta\phi}$ and $l_{\theta\phi,p}$.

The integral form of (5.13) corresponds to:

$$\text{Æ}\left[h_{pq}h_{st}^{*}\right] = \int_0^{2\pi}\int_0^{2\pi}p(\theta,\phi)\exp(j((q-t)d_T\sin\theta$$
$$+(p-s)d_R\sin\phi)\frac{2\pi}{\lambda})\,d\theta\,d\phi, \tag{5.14}$$

where $p(\theta,\psi)$ is the scatterer density function, defined as:

$$p(\theta,\psi) = \frac{\text{Æ}\left[f_{\theta\phi}^{2}\right]}{(2\pi)^2}. \tag{5.15}$$

Here, the scatterer density function is analogous to a 2-D power density function, which describes how the power is distributed along $\theta$ and $\psi$ directions.

Observing (5.14), the KPF can not describe the general channel, because in general the fading correlation can not be written in the form of $r_R(p-s)r_T(q-t)$. A similar conclusion can be found in [Abdi and Kaveh(2002)].

For some special case of $p(\theta,\psi)$, the KPF is still capable of describing the channel. For example, if the scatterer density function is in the form:

$$p(\theta,\psi) = p(\theta)p(\psi), \tag{5.16}$$

then the double integral in (5.14) could be decomposed into two separate integrals, and the correlation structure would be separable. Then, the KPF would be capable of describing this particular channel.

### 5.2.3 Microcell Environments

In this section, we consider microcell environments, where scatterer distribution follows an elliptic structure [Ertel et al.(1998)], as shown in Fig. 5.2. The transmitter and receiver antenna arrays are located in the focal of the ellipse. An example of microcell environments is indoor environments, e.g., the environments for Wi-Fi (IEEE 802.11) systems. In such environments, transmitters and receivers would share the same scatterers.

Here, we assume the scatterer-transmitter/receiver distance and symbol duration are significantly larger than the antenna size and the delay created by antenna element spacing, respectively. Now, due to the geometry of ellipse, all scatterers create similar channel delays. This delay difference would be in the order of antenna size, which is much smaller than the symbol duration. This leads to a frequency flat fading channel.

Using the same assumptions in Section 5.2.2 and similar technique as (5.14), the fading correlation between $p^{th}$ transmitter element and $q^{th}$ receiver element corresponds to:

$$\text{Æ}\left[h_{pq}h_{st}^{*}\right] = \int_{0}^{2\pi} p(\psi)\cdot exp(\pi\cdot j((p-s)\frac{b\sin\psi}{c+b\cos\psi} + (q-t)\frac{b\sin\psi}{c-b\cos\psi}))\,d\psi \tag{5.17}$$

where $\psi$ denotes the angle of arrival of scatterers with respect to the center of the ellipse (see Fig. 5.2).

Although the closed form solution of (5.17) does not exist, but for a particular $p(\psi)$, the integral can be numerically evaluated. Obviously the first chosen $p(\psi)$ is:

$$p(\psi) = \frac{1}{2\pi}. \tag{5.18}$$

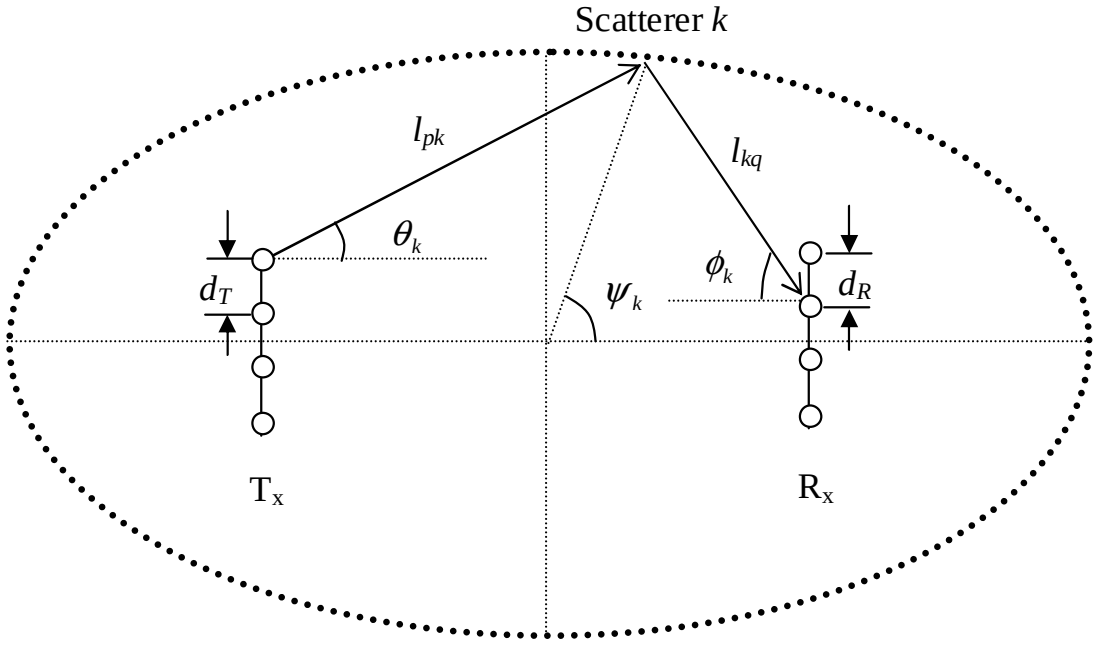


Figure 5.2: Geometry of Scatterers Distribution for Microcell Environments

This scatterer density function corresponds to a uniformly distribution of scatterers along $\psi$, which leads to a rich scattering environment.

The correlation results for (5.18) is shown in Fig. 5.3. In this figure, $\rho(0,d)$ $[\rho(d,0)]$ refers to the correlation between two receiver [transmitter] elements with spacing $d$ and the same transmitter [receiver]. Due to symmetric geometry, $\rho(0,d) = \rho(d,0)$. Moreover, $\rho(d,d)$ refers to the correlation between two receiver elements with spacing $d$ and two transmitter elements with spacing $d$. If the KPF is realistic, according to separable spatial stationarity condition, $\rho(d,d) = \rho(0,d)^2$, which is the "$\rho(d,d)$ prediction" curve in the figure. Analytical results are numerical evaluation of (5.17), and simulation results are generated via (5.12), assuming 1000 scatterers in (5.13) and 256 fading realizations are generated to compute the fading statistics.

First, we see that the simulation results have a good match with analytical results. Second, if the KPF is realistic, the predicted $\rho(d,d)$ should be the same as analytical and simulated $\rho(d,d)$. However, there is a significant difference between

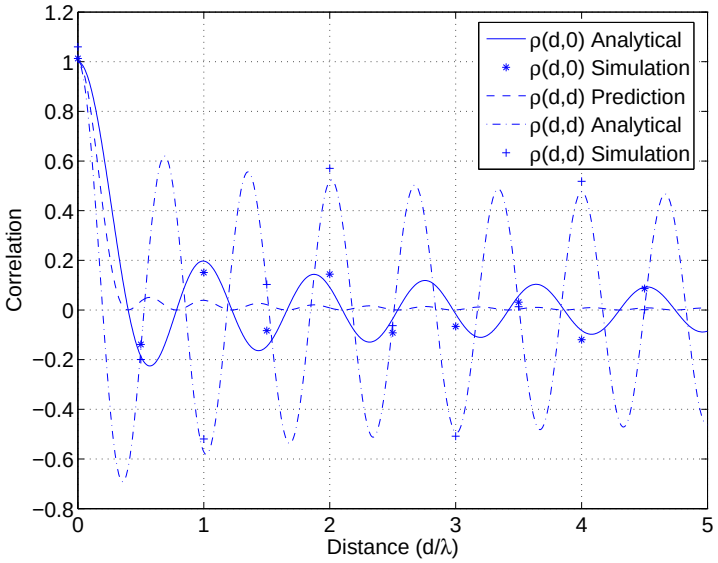


Figure 5.3: Correlation Results for Elliptical Channel with $p(\psi) = \frac{1}{2\pi}$.

predicted $\rho(d, d)$ and analytical/simulated $\rho(d, d)$. This confirms that the the Kronecker product form method does not lead/generate realistic correlation between antennas in microcell environments.

Third, it is shown that the correlation at larger distances is not necessarily smaller, while the envelope of the correlation decreases monotonically. This is because the correlation is a discrete sample of the sinc-like function. Overall, in contrast to common sense, the correlation may not decrease when distance increases, although the envelope of the correlation monotonically decreases with distance.

The second chosen $p(\psi)$ is:

$$p(\psi) = \frac{2}{\pi} |\sin(\psi)|. \tag{5.19}$$

Comparing with the former distribution (5.18), the latter one leads to a non-uniform distribution, because the reflected power is more concentrated in some particular direction, i.e., $\psi = \frac{\pi}{2}$ and $\psi = -\frac{\pi}{2}$.

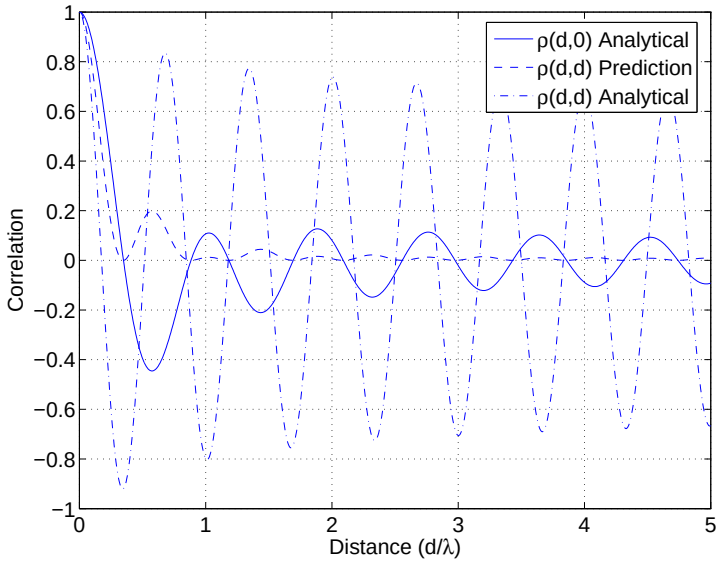


Figure 5.4: Correlation Results for Elliptical Channel with $p(\psi) = \frac{2}{\pi}|sin(\psi)|$.

The analytical results for (5.19) is shown in Fig. 5.4. First, similar to $p(\psi) = \frac{1}{2\pi}$, the KPF prediction is significantly different from real value. Second, comparing Figs. 5.3 and 5.4, the correlation for $p(\psi) = \frac{2}{\pi}|\sin(\psi)|$ is much larger than $p(\psi) = \frac{1}{2\pi}$. This is reasonable since the second scatterer density function leads to a more non-uniform distribution, thus it leads to a more correlated channel.

In this section, we evaluate a particular channel family (elliptical distributed channel), and proved that, in general the KPF does not generate a realistic description in microcell environments. A discussion on other channel models that might be suitable for microcell environments can be found in another paper of the authors [Tong and Zekavat(d toa)].

### 5.2.4 Macrocell Environments

In macrocell environments, the scatterer distribution would follow the two ring model [Ertel et al.(1998)], as shown in Fig. 5.5. An example of macrocell environments is the environments for cellular systems, where both base station and mobile are

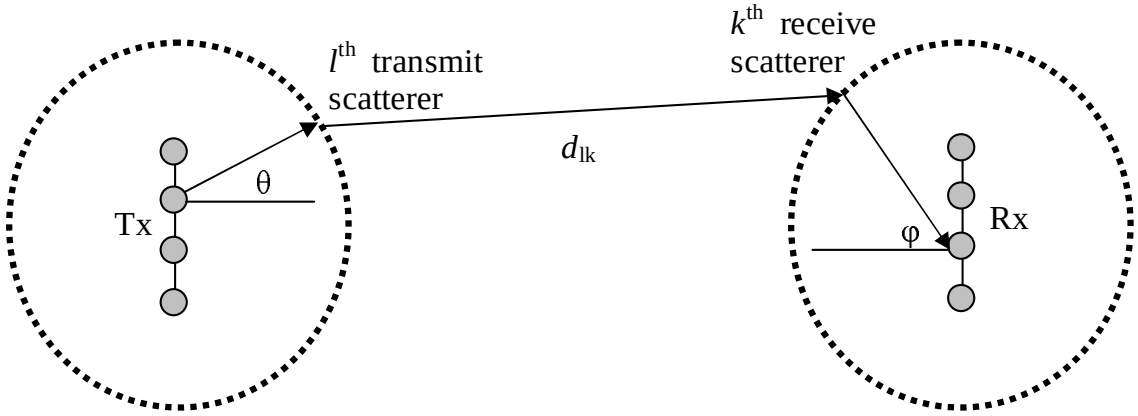


Figure 5.5: Geometry of Scatterers Distribution in Macrocell Environments.

surrounded by multiple reflective objects, and there is not line of sight between the mobile and the base station.

Using similar assumption as (5.12), the fading between $q^{th}$ transmit antenna element and $p^{th}$ receive antenna element corresponds to:

$$h_{pq} = \lim_{K\to\infty}\lim_{L\to\infty}\frac{1}{\sqrt{KL}}\cdot\sum_{l=1}^{L}\sum_{k=1}^{K} g_l\cdot g_k\cdot\exp((R_T+ (q-1)d_T\sin\theta_l + d_{lk} + (p-1)d_R\sin\phi_k + R_R)\frac{2\pi}{\lambda}) \tag{5.20}$$

where $g_l$ and $g_k$ denote $l^{th}$ transmit scatterer and $k^{th}$ receive scatterer, respectively, and, $R_T$ and $R_R$ denote radius of transmit scatterers and receiver scatterers, respectively.

The correlation between two fadings corresponds to:

$$\begin{aligned}
&Æ\left[h_{pq}h_{st}^{*}\right] = \lim_{K\to\infty}\lim_{L\to\infty}\frac{1}{KL}\sum_{l=1}^{L}\sum_{k=1}^{K} Æ\left[|g_l|^2\right] Æ\left[|g_k|^2\right]\cdot \\
&\cdot\exp\left(((q-t)d_T\sin\theta_l + (p-s)d_R\sin\phi_k)\,\frac{2\pi}{\lambda}\right) \\
&= \lim_{L\to\infty}\frac{1}{L}\sum_{l=1}^{L} Æ\left[|g_l|^2\right]\exp\left(((q-t)d_T\sin\theta_l)\,\frac{2\pi}{\lambda}\right)\cdot \\
&\lim_{K\to\infty}\frac{1}{K}\sum_{k=1}^{K} Æ\left[|g_k|^2\right] exp\left(((p-s)d_R\sin\phi_k)\,\frac{2\pi}{\lambda}\right)
\end{aligned} \tag{5.21}$$

The above summation approaches to the integral:

$$\text{Æ}\left[h_{pq}h_{st}^{*}\right] = \int_{0}^{2\pi} p_{\theta}(\theta) exp\left(((q-t)d_T \sin\theta)\,\frac{2\pi}{\lambda}\right) d\theta$$
$$\cdot \int_{0}^{2\pi} p_{\phi}(\phi) \exp\left(((p-s)d_R \sin\phi_k)\,\frac{2\pi}{\lambda}\right) d\phi \tag{5.22}$$

Simply define the first term and second term in (5.22) as $r_T(q-t)$ and $r_R(p-s)$, respectively, it is easy to see that the correlation in macrocell environment satisfies separable spatial stationarity condition, i.e., $\text{Æ}\left[h_{pq}h_{st}^{*}\right] = r_T(q-t)r_R(p-s)$. Hence, in macrocell environments, the KPF is still capable of generating realistic channel statistics.

The simulation results for macrocell environments are shown in Figs. 5.6 and 5.7. The first chosen scatterer density function would be uniform distribution at both sides, i.e, $p(\theta) = p(\phi) = \frac{1}{2\pi}$. The two side correlation analytical results are predicted by the KPF; we see that the simulation results matches with the KPF prediction. Hence, the KPF can describe this macrocell environment.

The second chosen scatterer density function is von Mises distribution at both sides, i.e.,

$$p(\theta) = \frac{e^{\kappa\cos\theta}}{2\pi I_0(\kappa)} \quad \theta \in [-\pi, \pi) \tag{5.23}$$

and

$$p(\phi) = \frac{e^{\kappa\cos\phi}}{2\pi I_0(\kappa)} \quad \phi \in [-\pi, \pi) \tag{5.24}$$

where $I_0(\cdot)$ denotes the modified Bessel function of the first kind of order zero. Here, $\kappa$ is the main parameter. If $\kappa = 0$, the von Mises Distribution corresponds to rich scattering environment, i.e., scatterers are uniformed distributed in $[-\pi, \pi]$. Larger $\kappa$ leads to smaller angle dispersion, i.e., a more correlated channel. $\kappa \to \infty$ leads to a coherent channel. It has been observed that von Mises Distribution matches with many real azimuth power spectrum [Abdi et al.(2002)]. In our simulation, we choose

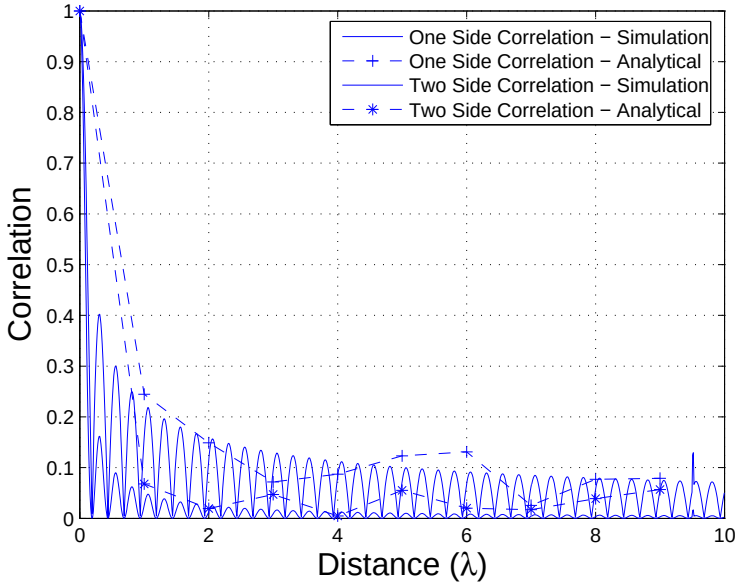


Figure 5.6: Correlation Results for Two Ring Channel with uniformly distributed scatterers.

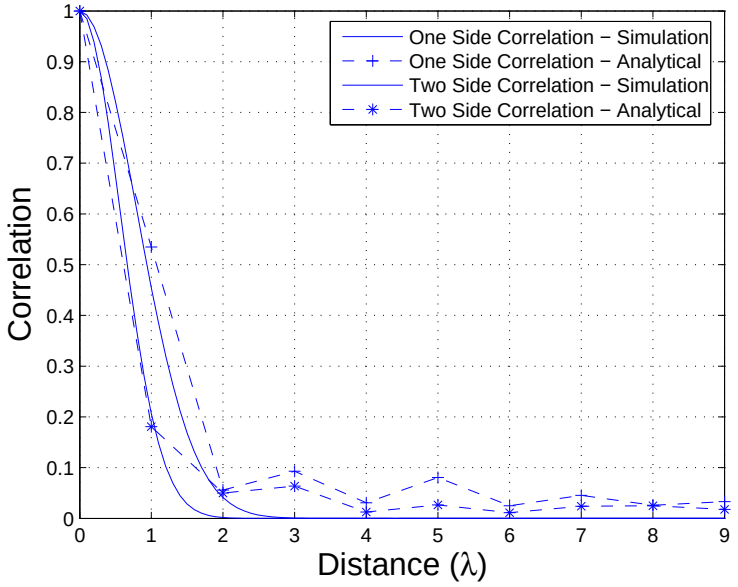


Figure 5.7: Correlation Results for Two Ring Channel with von Mises Distribution at both sides ($\kappa = 10$).

$\kappa = 10$. Here, the simulation results matches with the KPF analytical results for von Mises Distribution. Therefore, via both theoretical and numerical results, we show that the KPF is capable of describing macrocell environments.

### 5.2.5 Line-of-Sight Considerations

In Section 5.2.3 and 5.2.4, the microcell and macrocell environments without Line-of-Sight (LoS) are considered. We show that the KPF is suitable for macrocell environments without LoS. However, in real environments, LoS may present. Hence, it is also necessary to consider the effects of LoS components.

In [Abdi and Kaveh(2002)], it has been clearly stated that macrocell environments with LoS do not possess separable correlation structure. Therefore, directly applying the KPF in macrocell environments with LoS is not feasible. However, in this section, we show that a variation of the KPF can successfully describe macrocell environments with LoS. This variation is motivated by the relevant works in [Tulino et al.(2005)].

First, the eigen value decomposition (EVD) of the transmit covariance matrix corresponds to:

$$\mathbf{G}_T = \mathbf{U}_T \mathbf{\Lambda}_T \mathbf{U}_T^{\dagger}. \tag{5.25}$$

Here, $\mathbf{G}_T$ is assumed to be semi-positive definite. It is well known that each column of $\mathbf{U}_T$ corresponds to a certain direction, and elements of the diagonal matrix $\mathbf{\Lambda}_T$ denotes the reflected signal power in the corresponding directions. Let $\mathbf{\Lambda}_T^{\frac{1}{2}}$ denotes the $N_T \times N_T$ diagonal matrix whose diagonal elements are square roots of diagonal elements of $\mathbf{\Lambda}_T$, the matrix $\mathbf{G}_{\frac{1}{2}}$ in (5.77) would correspond to:

$$\mathbf{G}_T^{\frac{1}{2}} = \mathbf{U}_T \mathbf{\Lambda}_T^{\frac{1}{2}}. \tag{5.26}$$

In the KPF, the matrix $\mathbf{G}_T^{\frac{1}{2}}$ is the critical parameter that determines spatial correlations at the transmitter side. In (5.26), $\mathbf{G}_T^{\frac{1}{2}}$ is decomposed into two parts: directions, and the power in those directions.

Similarly, for receiver covariance matrix, the EVD corresponds to:

$$\mathbf{G}_R = \mathbf{U}_R \mathbf{\Lambda}_R \mathbf{U}_R^{\dagger}. \tag{5.27}$$

Moreover, similarly $\mathbf{G}_R^{\frac{1}{2}}$ can be decomposed into the multiplication of directions and the power in those directions:

$$\mathbf{G}_R^{\frac{1}{2}} = \mathbf{U}_R \mathbf{\Lambda}_R^{\frac{1}{2}}. \tag{5.28}$$

Substituting (5.26) and (5.28) into the KPF generation equation (5.77), (5.77) would turn out to be:

$$\mathbf{H}_{NLoS} = \mathbf{U}_R \mathbf{\Lambda}_R^{\frac{1}{2}} \mathbf{H}_{iid} \mathbf{\Lambda}_T^{\frac{1}{2}^{\top}} \mathbf{U}_T^{\top}. \tag{5.29}$$

Here, the KPF is decomposed into the multiplication of the receive directions, the power in the receive directions, randomness of the scatterers ($\mathbf{H}_{iid}$, the power in the transmit directions, and the transmit directions.

(5.29) describes the non-LoS components in macrocell environments. Now, assuming that there is a LoS component in the $m^{th}$ receive direction and $n^{th}$ transmit direction, the LoS component would be:

$$\mathbf{H}_{LoS} = \mathbf{U}_R \mathbf{\Omega}_{LoS} \mathbf{U}_T^{\top}, \tag{5.30}$$

where $\mathbf{\Omega}_{LoS}$ corresponds to:

$$\mathbf{\Omega}_{LoS} = \begin{bmatrix} 0 & \cdots & \cdots & 0 \\ 0 & \cdots & \cdots & 0 \\ \vdots & \ddots & 1 & \vdots \\ 0 & \cdots & \cdots & 0 \end{bmatrix}. \tag{5.31}$$

The only non-zero element of $\mathbf{\Omega}_{LoS}$ is on its $m^{th}$ row and $n^{th}$ column.

Two remarks are necessary for understanding the fading for LoS component. First, (5.30) does not possess a separable structure. This is why the KPF is not directly feasible for macrocell environments with LoS. Second, there is no random

Table 5.1: Summary of the Suitability of the Kronecker Product Form

| | Without Line-of-Sight | With Line-of-Sight |
|---|---|---|
| Microcell Environments | No | No |
| Macrocell Environments | Yes | Needs Generalization |

variables in (5.30). This is because that the LoS components are deterministic, while non-LoS components are random.

Combining both LoS and Non-LoS components, the fadings in macrocell environments with LoS would correspond to:

$$\mathbf{H} = \mathbf{U}_R \left( \sqrt{\frac{1}{K+1}} \mathbf{\Lambda}_R^{\frac{1}{2}} \mathbf{H}_{iid} \mathbf{\Lambda}_T^{\frac{1}{2}^\top} + \sqrt{\frac{K}{K+1}} \mathbf{\Omega}_{LoS} \right) \mathbf{U}_T^\top, \tag{5.32}$$

where the Rician factor $K$ denotes the ratio of signal power in LoS component over the non-LoS scattered power.

The model in (5.32) would be capable of describing fadings in Macrocell environments with LoS. (5.32) is essentially a generalization of the KPF model. When $K = 0$, i.e., the power of LoS component would be zero, and (5.32) boils down to the original KPF. When $K \to \infty$, (5.32) would correspond to a free space propagation: the scattering phenomenon does not exist. For other values of $K$, both LoS and non-LoS components exist.

Finally, the suitability of the KPF in various environments is summarized in Table 5.1. First, the KPF is not suitable for microcell environments, no matter if the LoS exists or not. Second, the KPF is applicable to macrocell environments without LoS. Thirdly, a generalization of the KPF is feasible in macrocell environments with LoS.

## 5.3 Virtual Channel Representation

This chapter discusses the recently proposed Virtual Channel Representation (VCR) and compares it with traditional 2-D product form method. In [Ertel and Reed(1998)],

the authors presented a production-form correlated fading channel generation. By multiplying a row of i.i.d. (identical independent distribute) zero-mean complex Gaussian variables and a specific matrix $\mathbf{G}^{\frac{1}{2}}$, the amplitudes of the resulted complex Gaussian variables would be correlated, i.e., equal-power correlated Rayleigh variables are generated. The correlation between those Rayleigh variables are determined by the matrix $\mathbf{G}^{\frac{1}{2}}$, which is Cholesky (eigen value) decomposition of the correlation matrix $\mathbf{G}$. The two dimensional (2-D) version of production approach has been widely used in many references as well [Chuah et al.(2002)]. However, the channel capacity predicted by production-form method is significantly lower than experiment results [Ozcelik et al.(2003)].

In [Young and Beaulieu(2000)], an approach other than the production-form is proposed, where, the amplitudes of Discrete Fourier Transform (DFT) of a row (one dimensional DFT) of independent zero-mean complex Gaussian variables are correlated Rayleigh variables. The correlation is determined by the variances of those independent zero-mean complex Gaussian variables. The 2-D version of DFT method, which is called VCR, has been reported in [Sayeed(2002)]. VCR transforms the fading between different antenna elements to the fading between different beams. A recent experiment result shows that VCR provides a better channel capacity prediction than production-form method [Zhou and Sayeed(d to)].

The chapter contributions include: (a) Extending the theoretical fading amplitude and phase correlation result from 1-D DFT to 2-D DFT, i.e., extending the theoretical correlation results from SIMO to MIMO channels; (b) Investigating the relationship between 2-D DFT and 2-D production-form methods, and proving that 2-D DFT is a more general method than the production-form.

Considering: 1) The recent claim in the literature that 2-D production-form method is an oversimplified model [Oestges et al.(2005)] and it contradicts experiment results [Ozcelik et al.(2003)], and, 2) We prove that VCR is a more general model than 2-D production-form method, this work leads to a good hint that

VCR is a more realistic channel model compared with the traditional production-form method. Thus, MIMO spatially correlated channels (such as indoor or urban areas) should be generated with VCR model in order to match with experiments [Zhou and Sayeed(d to)].

### 5.3.1 VCR Fading Amplitude and Phase Correlation

Throughout this chapter, we use the following notation: Superscripts $H$ and $T$ denote Hermitian transpose and transpose, respectively; Subscripts $R$ and $T$ corresponds to receiver and transmitter, respectively; Subscripts $v$ and $P$ means the matrix is used by VCR or production-form methods, respectively.

Assuming $M$ transmit antenna elements and $N$ receive antenna elements, the channel fading matrix is generated by 2-D DFT of VCR channel fading matrix, $\mathbf{H}_v$, via:

$$\mathbf{H} = \mathbf{A}_R \mathbf{H}_v \mathbf{A}_T^H, \tag{5.33}$$

where

$$\mathbf{A}_R = [\vec{a}_R(\theta_{1R})\ \vec{a}_R(\theta_{2R})\ \cdots\ \vec{a}_R(\theta_{MR})] \quad (\mathbf{A}_R \in \mathcal{C}^{M\times M}) \tag{5.34}$$

and

$$\mathbf{A}_T = [\vec{a}_T(\theta_{1T})\ \vec{a}_T(\theta_{2T})\ \cdots\ \vec{a}_T(\theta_{NT})] \quad (\mathbf{A}_T \in \mathcal{C}^{N\times N}) \tag{5.35}$$

.

Here,

$$\vec{a}_R(\theta_{mR}) = \frac{1}{\sqrt{M}} \begin{bmatrix} 1 & e^{-j2\pi\theta_{mR}} & \cdots & e^{-j(M-1)2\pi\theta_{mR}} \end{bmatrix}^T \tag{5.36}$$

and

$$\vec{a}_T(\theta_{nT}) = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 & e^{-j2\pi\theta_{nT}} & \cdots & e^{-j(N-1)2\pi\theta_{nT}} \end{bmatrix}^T \tag{5.37}$$

are the array vector of the receiver and transmitter, respectively; and $m \in \{1, 2, \cdots, M\}$, $n \in \{1, 2, \cdots, N\}$. $\theta_{mR}$ and $\theta_{nT}$ are evenly selected within the range of $[-0.5 \quad 0.5]$,

i.e.,

$$\theta_{mR} = \frac{2m + 1 - M}{2M} \tag{5.38}$$

and

$$\theta_{nT} = \frac{2n + 1 - N}{2N}. \tag{5.39}$$

In this case, $\mathbf{A}_R$ and $\mathbf{A}_T$ are unitary matrixes, i.e., $\mathbf{A}_R\mathbf{A}_R^H = \mathbf{I}_M \in \mathcal{R}^{\mathcal{M}\times\mathcal{M}}$ and $\mathbf{A}_T\mathbf{A}_T^H = \mathbf{I}_N \in \mathcal{R}^{\mathcal{N}\times\mathcal{N}}$, where $\mathbf{I}_M$ and $\mathbf{I}_N$ are identity matrixes. Moreover, $\mathbf{A}_R$ and $\mathbf{A}_T$ are in fact Fourier Transform matrixes, and, $\mathbf{H}$ and $\mathbf{H}_v$ are 2-D Fourier pairs.

The elements of $\mathbf{H}_v$ are zero-mean complex Gaussian variables with variance matrix $\mathbf{\Psi} = \text{var}(\mathbf{H}_v)$. This variance matrix completely characterizes the statistics of $\mathbf{H}_v$. Here, for convenience and without loss of generality, we assume unity transmitted bit energy, which corresponds to: $\sum_{n=1}^{N}\sum_{m=1}^{M}\psi_{m,n} = 1$, where $\psi_{m,n}$ is the $m^{th}$ row $n^{th}$ column element of $\mathbf{\Psi}$. In this section, we show the theoretical results of correlation between any two elements of $\mathbf{H}$ for any value of $\mathbf{\Psi}$.

The fading between $p^{th}$ receiver and $q^{th}$ transmitter elements, or the $p^{th}$ row, $q^{th}$ column element of $\mathbf{H}$, corresponds to: $h_{p,q} = Re(h_{p,q}) + j \cdot Im(h_{p,q})$. Both $Re(h_{p,q})$ and $Im(h_{p,q})$ are zero-mean Gaussian variables since they are linear combinations of independent zero-mean Gaussian variables. It is easy to show that $Re(h_{p,q})$ and $Im(h_{p,q})$ are uncorrelated and their variances correspond to:

$$u = E\left[(Im(h_{p,q}))^2\right] = E\left[(Re(h_{p,q}))^2\right] = \frac{1}{2MN}. \tag{5.40}$$

Now, we study the relationship between $h_{p,q}$ and another fading $h_{s,t}$. The correlations between real and imaginary parts of $h_{p,q}$ and $h_{s,t}$ correspond to:

$$\begin{aligned} u_1 = E\left[Re(h_{p,q})Re(h_{s,t})\right] &= E\left[Im(h_{p,q})Im(h_{s,t})\right] \\ &= \frac{1}{MN}\sum_{n=1}^{N}\sum_{m=1}^{M}\frac{\psi_{m,n}}{2}\cos\theta_{pqstmn}, \end{aligned} \tag{5.41}$$

and,

$$u_2 = E\left[Re(h_{p,q})Im(h_{s,t})\right] = -E\left[Re(h_{s,t})Im(h_{p,q})\right] \\ = \frac{1}{MN}\sum_{n=1}^{N}\sum_{m=1}^{M}\frac{\psi_{m,n}}{2}\sin\theta_{pqstmn}, \tag{5.42}$$

where $\theta_{pqstmn} = 2\pi\left[(n-1)\frac{t-q}{N} - (p-s)\frac{2m-1-M}{2M}\right]$.

The cross-correlation coefficient between the amplitudes of $h_{p,q}$ and $h_{s,t}$ corresponds to [eq. (1.5-26), 8]:

$$\rho_r = E\left[|h_{p,q}||h_{s,t}|\right] = \frac{(1+\lambda)E_i\left(\frac{2\sqrt{\lambda}}{1+\lambda}\right) - \frac{\pi}{2}}{2-\frac{\pi}{2}}, \tag{5.43}$$

where $E_i()$ denotes the complete Elliptic integral of the second kind. Moreover, the cross-correlation coefficient between the phase of $h_{p,q}$ and $h_{s,t}$ corresponds to [eq. (1.5-33), 8]:

$$\rho_\theta = E\left[\angle h_{p,q}\angle h_{s,t}\right] = 3\xi(1+2\xi) - \zeta. \tag{5.44}$$

The parameters $\lambda$, $\xi$ and $\zeta$ in (5.43) and (5.44) correspond to:

$$\lambda = \sqrt{\frac{u_1^2+u_2^2}{u^2}}, \quad \xi = \frac{1}{2\pi}\sin^{-1}\left(\frac{u_1}{u}\right), \quad \zeta = \frac{3}{4\pi^2}\sum_{k=1}^{\infty}\frac{\lambda^{2k}}{k^2}$$

To verify the theoretical results in (5.43) and (5.44), a small simulation is conducted. According to (5.73), the channel fading generation process in this simulation includes three steps: (1) Selecting a virtual variance matrix $\mathbf{\Psi}$, which can be chosen arbitrarily or mapped from a physical channel model; (2) Generating the virtual channel matrix $\mathbf{H}_v$ using $\mathbf{\Psi} = \text{var}(\mathbf{H}_v)$; and, (3) Generating channel fading matrix $\mathbf{H}$ using (5.73).

The comparison of theoretical and simulation results of the correlation coefficient of the amplitudes and phases are shown in Fig. 1. The x-axis in Fig. 1 represents the "dominance factor" $z$ defined as follow: With $\psi_{m,n}$ the element of $\mathbf{\Psi}$, and $p$, $q$

arbitrarily selected dominating beams, we define a simple model for $\mathbf{\Psi}$ as: $\psi_{m,n} = 1$ if $m = p, n = q$, and, $\psi_{m,n} = z$ otherwise. The y-axis of Fig. 1 is the fade amplitude (phase) correlation coefficients over adjacent antenna elements, assuming four transmitters and six receivers.

If $z$ approaches 0, there is only one dominating virtual path between transmitter and receiver, and the antenna element fades become coherent, i.e., the same fade amplitudes and fixed phase shifts over the antenna elements. If $z$ approaches 1, there is no dominating virtual path between transmitter and receiver, and the antenna element fades tend to become independent since the signal received from different directions has equal power due to a rich scattering environment. Hence, as $z$ changes from 0 to 1, the correlation coefficient of the channel changes from 1 to 0, or the channel changes from coherent to independent (see the x-axis of Fig. 1).

In this section, we (1) explained the correlated channel generation process using VCR; and, (2) extended the previous theoretical correlation coefficient results in [Jakes(1974)] from 1-D DFT to 2-D DFT case by recomputing the parameters $u$, $u1$ and $u2$. In the next section, we will use the above results ($u$, $u1$, $u2$) to prove that the generation of $\mathbf{H}$ via VCR in (5.73) is a more general technique than the traditional production-form.

### 5.3.2 VCR versus 2D Production-Form

The traditional production-form process corresponds to:

$$\mathbf{H} = \mathbf{G}_R^{\frac{1}{2}} \mathbf{H}_P \mathbf{G}_T^{\frac{1}{2}H}, \tag{5.45}$$

where $\mathbf{H}_P$ is an $M \times N$ matrix consisting of i.i.d zero-mean complex Gaussian elements. The matrixes $\mathbf{G}_T^{\frac{1}{2}}$ and $\mathbf{G}_R^{\frac{1}{2}}$ in (5.77) satisfy: $\mathbf{G}_T^{\frac{1}{2}} \mathbf{G}_T^{\frac{1}{2}H} = \mathbf{G}_T$ and $\mathbf{G}_R^{\frac{1}{2}} \mathbf{G}_R^{\frac{1}{2}H} = \mathbf{G}_R$, where $\mathbf{G}_T$ and $\mathbf{G}_R$ are transmit and receive correlation matrixes with elements

that respectively satisfy: $E\left[h_{p,q}h_{p,t}^{H}\right] = g_{T_{q,t}}$ and $E\left[h_{p,q}h_{s,q}^{H}\right] = g_{R_{p,s}}$. Here, $g_{T_{q,t}}$ is the $q^{th}$ row $t^{th}$ column element of $\mathbf{G}_T$, and $g_{R_{p,s}}$ is the $p^{th}$ row $s^{th}$ column element of $\mathbf{G}_R$. Note that the correlation between the two channels $h_{p,q}$,$h_{p,t}$ is independent of index of receiver, $p$. This corresponds to the fact that $g_{T_{q,t}}$ only represents local scatterers around transmitter elements. Similar conclusion applies to $h_{p,q}$, $h_{s,q}$ and $g_{R_{p,s}}$. In addition, the most important property of production method is [Chuah et al.(2002)]:

$$E\left[h_{p,q}h_{s,t}^{H}\right] = g_{R_{p,s}}g_{T_{q,t}}. \tag{5.46}$$

Since $g_{T_{q,t}}$ and $g_{R_{p,s}}$ only represent local scatterers around transmitters/receivers, this property confirms that for production form, the correlation between two spatial channels only considers local scatterers. In other words, this method assumes free space in the medium between transmitters and receivers, which is not true in all wireless environments such as indoor and urban areas.

Another common assumption in wireless communications is spatially stationarity of the correlation matrix elements, $g_{R_{p,s}}$, $g_{T_{q,t}}$, which corresponds to:

$$g_{R_{p,s}} = r_R(p-s) \quad \text{and} \quad g_{T_{q,t}} = r_T(q-t), \tag{5.47}$$

where $r_R(p-s)$ is the local correlation coefficient determined by the distance between $r^{th}$ and $s^{th}$ correlation receiver element, and $r_T(q-t)$ is defined similarly. This means that the local correlations $g_{R_{p,s}}$ and $g_{T_{q,t}}$ are solely determined by the distance between two transmitter or receiver elements, and they are independent of their absolute positions in the environment. We will use this assumption in the following derivations as well; and, define local correlation vectors $\vec{r}_R = [r_R(0)\ r_R(1)\ \cdots\ r_R(N-1)]^T$, and $\vec{r}_T = [r_T(0)\ r_T(1)\ \cdots\ r_T(M-1)]^T$ for the receiver, and transmitter, respectively.

Here, we introduce a theorem that clarifies the relationship between 2-D VCR and 2-D production-form method. This theorem shows the fading channels that 2-D production-form generates are a subset of the fading channels that 2-D VCR model would generate.

**Theorem 5.3.1.** *The 2-D production-form is equivalent to 2-D VCR method if and only if the variance matrix of* $\mathbf{H}^v$, $\mathbf{\Psi}$, *can be decomposed into two vectors* $\vec{\psi}_T \in \mathcal{R}^{\mathcal{N}\times\infty}$ *and* $\vec{\psi}_R \in \mathcal{R}^{\mathcal{M}\times\infty}$, *that is:*

$$\mathbf{\Psi} = \vec{\psi}_R \vec{\psi}_T^H. \tag{5.48}$$

*Proof.* Substituting (5.47) into (5.46), the correlation between the two channels with different transmitter and receiver elements in 2-D production-form corresponds to:

$$E\left[h_{p,q}h_{s,t}^H\right] = r_R(p-s)r_T(q-t). \tag{5.49}$$

Please note that $E\left[h_{p,q}h_{s,t}^H\right]$ also fully describes the fading statistics. Using (5.41) and (5.42), the same correlation for VCR corresponds to:

$$E\left[h_{p,q}h_{s,t}^H\right] = \frac{1}{MN}\sum_{n=1}^{N}\sum_{m=1}^{M}\psi_{m,n}e^{j\theta_{pqstmn}}. \tag{5.50}$$

Here, first, we prove the sufficient condition, i.e., if (5.48) is satisfied, 2-D VCR model and 2-D production-form can be translated into each other. If $\mathbf{\Psi} = \vec{\psi}_R\vec{\psi}_T^H$, then $\psi_{m,n} = \psi_R(m)\psi_T(n)$, and (5.50) corresponds to :

$$\begin{aligned} E\left[h_{p,q}h_{s,t}^H\right] = \frac{1}{N}\sum_{n=1}^{N}\psi_T(n)e^{-j2\pi(q-t)\frac{n-1}{N}} \\ \times\frac{1}{M}\sum_{m=1}^{M}\psi_R(m)e^{-j2\pi(p-s)\frac{2m-1-M}{2M}}. \end{aligned} \tag{5.51}$$

Comparing Eqs. (5.51) and (5.49), we find that defining

$$r_R(p-s) = \frac{1}{M}\sum_{m=1}^{M}\psi_R(m)e^{-j2\pi(p-s)\frac{2m-1-M}{2M}}, \tag{5.52}$$

and,

$$r_T(q-t) = \frac{1}{N}\sum_{n=1}^{N}\psi_T(n)e^{-j2\pi(q-t)\frac{n-1}{N}}, \tag{5.53}$$

the 2-D production-form and 2-D VCR method would be equivalent. Equations (5.52) and (5.53) represent two separate discrete Fourier transforms.

Second, we prove the necessary condition, i.e., if 2-D VCR model and 2-D production-form can be translated into each other, then $\mathbf{\Psi} = \vec{\psi}_R \vec{\psi}_T^H$. That is, if:

$$r_T(q-t)r_R(p-s) = \frac{1}{MN}\sum_{n=1}^{N}\sum_{m=1}^{M}\psi_{m,n}e^{j\theta_{pqstmn}} \tag{5.54}$$

then $\mathbf{\Psi}$ could be decomposed into multiplication of two vectors. Note that $\vec{r}_R\vec{r}_T^H$ and $\mathbf{\Psi}$ are 2D DFT pairs if (5.54) holds. According to separability property of 2D DFT, if 2-D Fourier transform of $\mathbf{\Psi}$ is multiplication of two vectors, then $\mathbf{\Psi}$ can be written as multiplication of two vectors as well. Thus, the necessary condition is proved. □

**Corollary 5.3.2.** *If a 2-D VCR model is expressed by 2-D production-form, then ($\vec{r}_R$, $\vec{\psi}_R$) and ($\vec{r}_T$,$\vec{\psi}_T$) would be two discrete Fourier transform pairs, as shown in* (5.52) *and* (5.53).

According to theorem 5.3.1, in the first step of VCR correlated channel generation process, if the chosen $\mathbf{\Psi}$ satisfies (5.48), i.e., the matrix $\mathbf{\Psi}$ can be expressed by multiplication of two vectors, this VCR model can be represented by production-form as well. Otherwise, this VCR channel can not be represented by production-form method. As we discussed, since production-form method describes only local scatters, we reach to the conclusion that VCR model describes both local and non-local scatterers and VCR is a more general channel model than production-form method.

Now, we 1) explained the generation process of 2-D VCR model; 2) extended the theoretical results of the correlation coefficient of 1-D DFT method to 2-D DFT method, i.e., from SIMO to MIMO channels; 3) simulated a simple 2-D VCR model to verify the theoretical results; and, 4) proved that the traditional 2-D production fading channel generation process is a special case of 2-D VCR model. Moreover, combining recent literature and this work, this work leads to a good hint that VCR

is a more realistic channel model compared with the traditional production-form. According to this conclusion, all simulations and theoretical results in spatially correlated channel should be generated with this VCR model in order to match with experiment results.

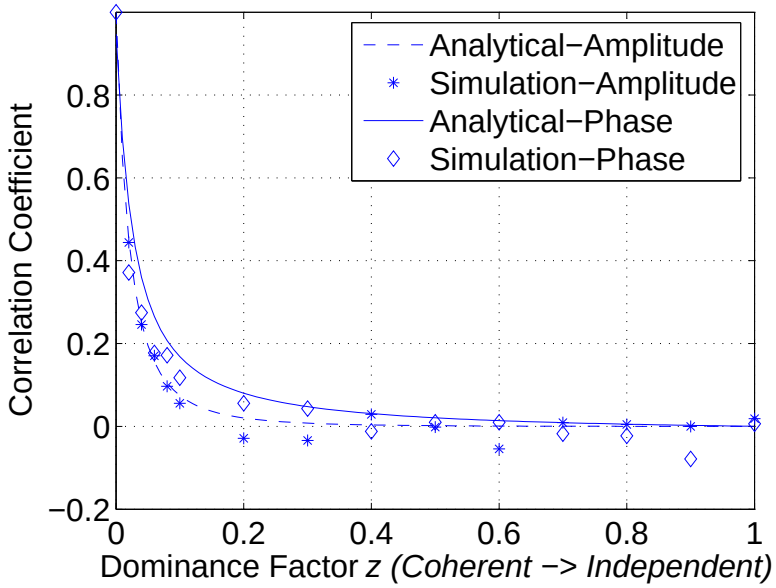


Figure 5.8: Theoretical and simulated amplitude (phase) correlation coefficients comparison.

## 5.4 The Total Correlation Matrix: Comparing KPF, VCR and W-model

The input-output relationship in MIMO systems corresponds to:

$$\vec{y} = \mathbf{H}\vec{x} + \vec{n}, \tag{5.55}$$

where $\vec{y}$, $\vec{n}$ and $\vec{x}$ are column vectors with dimension $N_R$, $N_R$, and $N_T$, respectively. The goal of MIMO channel modeling is to find the structure of $\mathbf{H}$, which is an $N_R \times N_T$ complex matrix. $N_T$ denotes the number of transmit antennas, and $N_R$ denotes the number of receive antennas.

Assuming Rayleigh fading, the elements of **H** would be zero-mean circularly symmetric [Telatar(1999)] complex Gaussian random variables. Non-Rayleigh faded MIMO channels can be generated via an extension of Rayleigh MIMO channels [Tulino and Verdu(2006)].

In spatially correlated fading environments, the elements of **H** would be correlated. The next step is to find how to describe the correlation between those elements. Here, the most comprehensive description of their correlations is the "total correlation matrix", which is defined via vectorization operation.

The column-wise form of the fading matrix, **H**, corresponds to:

$$\mathbf{H} = \left[\vec{h}_1\, \vec{h}_2, \cdots, \vec{h}_{N_T}\right]. \tag{5.56}$$

Here, $\vec{h}_i$ represents the $i^{th}$ column of **H**. Then, the vectorized fading matrix, vec(**H**), would be:

$$\mathrm{vec}(\mathbf{H}) = \left[\vec{h}_1^{\top}\, \vec{h}_2^{\top}, \cdots, \vec{h}_{N_T}^{\top}\right]^{\top}. \tag{5.57}$$

Roughly speaking, the vectorization operation "stacks" all columns of **H** into a single column. Effectively the vectorization operation converts two dimensional matrix to an one dimensional vector.

Here, the total correlation matrix is defined as the covariance matrix of the vectorized fading matrix:

$$\mathbf{G} = Æ\left[\mathrm{vec}(\mathbf{H}) \cdot \mathrm{vec}(\mathbf{H})^{\dagger}\right]. \tag{5.58}$$

**G** is said to be a sufficient description of the fading matrix **H**, since the joint probability density function (PDF) of elements in **H** can be found once the matrix **G** is known[Papoulis and Pillai(2002)].

Note that each element of **G** represents the correlation between two elements in **H**. Because there are $N_R N_T$ elements in **H**, **G** is with dimensions $N_R N_T \times N_R N_T$. It is also easy to verify that **G** is conjugate symmetric, i.e., Hermitian.

In addition to Hermitian form, there are more requirements on the form of $\mathbf{G}$, because of the physical world constraints. Here, we introduce the first main result: $\mathbf{G}$ would be in Toeplitz form if elements $\mathbf{H}$ are samples of a two dimensional stationary ergodic circularly symmetric complex Gaussian random process.

**Theorem 5.4.1.** *If the elements of* $\mathbf{H}$ *are samples of a two dimensional stationary ergodic circularly symmetric complex Gaussian random process, then*

$$g_{m,n} = \begin{cases} g_{m-n} & m \geq n \\ g^*_{n-m} & m < n \end{cases}, \tag{5.59}$$

*where* $g_{m,n}$ *denotes the element of* $\mathbf{G}$ *in the* $m^{th}$ *row and* $n^{th}$ *column,* $(\cdot)^*$ *represents conjugate operation.*

*Proof.* If a two dimensional random process satisfies the conditions claimed in the above theorem, it would always possess a two dimensional (2-D) power spectral density (PSD), $S(\omega, u)$. The relationship between correlation and the 2-D PSD corresponds to:

$$\begin{aligned} &Æ[h_{p,q}h^*_{s,t}] = \\ &\int_{-0.5}^{-0.5}\int_{-0.5}^{-0.5} S(\omega,u)e^{-j2\pi(\omega(p-s)+u(q-t))}\,d\,\omega\,d\,u. \end{aligned} \tag{5.60}$$

Note that $Æ[h_{p,q}h^*_{s,t}]$ is the element of $\mathbf{G}$ at row $(q-1)N_R + p$ and column $(t-1)N_R + s$. Then the difference between the indexes of the row and column would be $(q-t)N_R + (p-s)$.

Now, assume another element of $\mathbf{G}$, $Æ[h_{p',q'}h^*_{s',t'}]$, is on the same diagonal of $Æ[h_{p,q}h^*_{s,t}]$, i.e., $(q-t)N_R + (p-s) = (q'-t')N_R + (p'-s')$. Then, considering $p, s, p', s' \in \{1, 2, \cdots, N_R\}$ and $q, t, q', t' \in \{1, 2, \cdots, N_T\}$, it can be found that $q - t = q' - t'$ and $p - s = p' - s'$. Via (5.60), it is found that $Æ[h_{p',q'}h^*_{s',t'}] = Æ[h_{p,q}h^*_{s,t}]$. In other words, the elements located on the same diagonal would be identical. Therefore, $\mathbf{G}$ would be in Toeplitz form. □

In general, MIMO fading processes satisfy the conditions stated in Theorem 5.4.1 [Goldsmith(2005b)]; Hence, $\mathbf{G}$ is considered in in Hermitian Toeplitz form for the rest of this chapter.

According to this theorem, it seems easy to generate 2-D fading processes via vectorization operation: Generating an $N_R N_T$ dimensional vector with any arbitrary correlations, and then converting this vector to an $N_R \times N_T$ matrix, the fading matrix would be produced. But, we shall immediately see that the physical world applies more constraints on the structure of $\mathbf{G}$, in addition to the Hermitian Toeplitz form.

In the next section, we start analyzing the eigen structure of $\mathbf{G}$. It is well known that Hermitian Toeplitz matrix is semi-positive definite, i.e., $\mathbf{G}$ has the singular value decomposition:

$$\mathbf{G} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{\dagger}, \tag{5.61}$$

where $\mathbf{U}$ and $\mathbf{\Lambda}$ are eigen vector matrix and eigen value matrix, respectively; $\mathbf{\Lambda}$ is diagonal, and the elements on the diagonal are non-negative.

### 5.4.1 The W-model: Separable Eigen Basis

The W-model[Weichselberger et al.(2006)] proposes that the structure of fading matrix, $\mathbf{H}$, corresponds to:

$$\mathbf{H} = \mathbf{U}_R(\mathbf{\Omega} \odot \mathbf{H}_{iid})\mathbf{U}_T^{\dagger}. \tag{5.62}$$

Here, $\mathbf{U}_R$ and $\mathbf{U}_T$ denote two complex unitary matrices with $N_R \times N_R$ and $N_T \times N_T$ dimensions, respectively. $\mathbf{\Omega}$ is an $N_R \times N_T$ matrix with non-negative real elements, and $\mathbf{H}_{iid}$ represents a matrix that consists of $N_R \times N_T$ independent identically distributed (i.i.d.) complex Gaussian random variables, $\odot$ denotes element-wise matrix multiplication.

For W-model, the total correlation matrix, $\mathbf{G}$, has separable eigen basis, in addition to Hermitian Toeplitz form.

**Theorem 5.4.2.** *If a fading matrix* $\mathbf{H}$ *has the structure depicted in* (5.62)*, then the eigen basis of the corresponding total correlation matrix,* $\mathbf{G}$*, would be the Kronecker product of two other unitary matrices,* $\mathbf{U}_R$ *and* $\mathbf{U}_T$*, i.e.,:*

$$\mathbf{G} = (\mathbf{U}_R \otimes \mathbf{U}_T)\mathcal{D}(vec(\mathbf{\Omega}))(\mathbf{U}_{\mathcal{R}} \otimes \mathbf{U}_{\mathcal{T}})^{\dagger}. \tag{5.63}$$

*Here,* $\mathcal{D}(\cdot)$ *corresponds to a diagonal matrix, whose diagonal elements are determined by the elements of* $\mathbf{\Omega}$*.* $\otimes$ *denotes Kronecker product. The Kronecker product of two matrices* $\mathbf{A}$ *and* $\mathbf{B}$ *is defined as:*

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \cdots \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}. \tag{5.64}$$

*Equivalently, this theorem states that for W-model, the eigen basis matrix of* $\mathbf{G}$ *is the Kronecker product of* $\mathbf{U}_R$ *and* $\mathbf{U}_T$*, and the eigen values of* $\mathbf{G}$ *is the elements of* $\mathbf{\Omega}$*.*

*Proof.* Because $\mathbf{U}_R$ and $\mathbf{U}_T$ are both unitary, (5.62) becomes:

$$\mathbf{H}\mathbf{U}_T = \mathbf{U}_R(\mathbf{\Omega} \odot \mathbf{H}_{iid}). \tag{5.65}$$

Applying vectorization on both sides,

$$\mathrm{vec}\left(\mathbf{H}\mathbf{U}_T\right) = \mathrm{vec}\left(\mathbf{U}_R(\mathbf{\Omega} \odot \mathbf{H}_{iid})\right), \tag{5.66}$$

then

$$(\mathbf{U}_T^{\dagger} \otimes \mathbf{I}_{N_R})\mathrm{vec}(\mathbf{H}) = (\mathbf{I}_{N_T} \otimes \mathbf{U}_R)\mathrm{vec}\left((\mathbf{\Omega} \odot \mathbf{H}_{iid})\right). \tag{5.67}$$

Multiplying both sides by their own conjugate, and taking expectation:

$$\begin{aligned} &(\mathbf{U}_T^{\dagger} \otimes \mathbf{I}_{N_R})\text{Æ}\left[\mathrm{vec}(\mathbf{H})\mathrm{vec}(\mathbf{H})^{\dagger}\right](\mathbf{U}_T \otimes \mathbf{I}_{N_R}) \\ &= (\mathbf{I}_{N_T} \otimes \mathbf{U}_R)\text{Æ}[\mathrm{vec}(\mathbf{\Omega} \odot \mathbf{H}_{iid})\mathrm{vec}(\mathbf{\Omega} \odot \mathbf{H}_{iid})^{\dagger}](\mathbf{I}_{N_T} \otimes \mathbf{U}_R^{\dagger}) \end{aligned} \tag{5.68}$$

Noting that $\mathbf{U}_T \otimes \mathbf{I}_{N_R}$ is a unitary matrix, and

$$Æ[\mathrm{vec}(\mathbf{\Omega} \odot \mathbf{H}_{iid})\mathrm{vec}(\mathbf{\Omega} \odot \mathbf{H}_{iid})^{\dagger}] = \mathcal{D}\left(\mathrm{vec}(\mathbf{\Omega})\right) \tag{5.69}$$

the final expression would correspond to:

$$\mathbf{G} = (\mathbf{U}_R \otimes \mathbf{U}_T)\mathcal{D}(\mathrm{vec}(\mathbf{\Omega}))(\mathbf{U}_{\mathcal{R}} \otimes \mathbf{U}_{\mathcal{T}})^{\dagger}. \tag{5.70}$$

□

The above theorem states that W-model does not have any constraint on the eigen values of $\mathbf{G}$: they could be any non-negative numbers. However, it assumes the eigen basis of $\mathbf{G}$ is the Kronecker product of two other unitary matrices, i.e.,

$$\mathbf{U} = \mathbf{U}_R \otimes \mathbf{U}_T, \tag{5.71}$$

where $\mathbf{U}$ is the eigen basis of $\mathbf{G}$, $\mathbf{U}_R$ and $\mathbf{U}_T$ are introduced in (5.62). Through (5.71), W-model is said to have separable eigen basis, i.e., the eigen basis matrix $\mathbf{U}$ can be separated to transmitter side component $\mathbf{U}_T$ and receiver side component $\mathbf{U}_R$.

The next question is, in what physical environment, the total correlation matrix has separable eigen basis? At the first glance, it seems not easy. Fortunately, this question has been answered in [Barton and Fuhrmann(1993)] in a different background. The answer is: When there are $N_R$ antennas clustered together, while another $N_T$ antennas clustered together, and the distance between the two clusters are greatly larger than the cluster size, $\mathbf{G}$ would possess separable eigen basis. In other words, *separable eigen basis corresponds to separable transmitter and receiver.* Since most MIMO systems have separable transmitter and receiver, W-model would be capable of describing most MIMO systems.

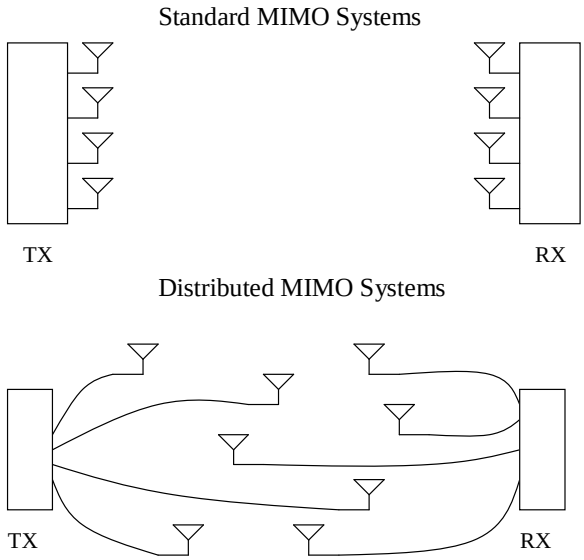


Figure 5.9: Comparison of standard and distributed MIMO systems.

An exception might be the distributed MIMO systems, as shown in Fig. 5.9. In those systems, the transmitting antennas and receiving antennas are not necessarily clustered together. Hence, W-model would not be able to describe such channels. Another possible example is multi-hop mobile ad hoc networks (MANET). Although MANET is not a MIMO system, its channel modeling is closely relevant to MIMO channel modeling. Imagining that there are $N_T$ nodes in the talking distance of node $A$, and $N_R$ nodes in the talking distance of node $B$, then the channel between $A$ and $B$ can be described by the general correlation structure, but can not be described by W-model.

### 5.4.2 The VCR: Fixed Separable Eigen Basis

The discussion of the VCR is slightly different from the discussion of the W-model, because the VCR varies with array manifolds. Originally, the VCR is proposed with uniform linear array (ULA); hence, we start with ULA cases. It is shown that the VCR predefines fixed *and* separable eigen basis. While separable eigen basis corresponds to separable transmitter and receiver, being fixed corresponds to the fact

that the VCR is suitable for only large dimensional arrays. Finally, we demonstrate how to extend the original VCR to non-ULA arrays.

**The ULA Case**

Assuming ULA receiving antenna, the array response vector corresponds to:

$$a_R(\vartheta) = \frac{1}{N_R}\left[1 \;\; \exp(-j\frac{2\pi\lambda\vartheta}{d}) \;\; \cdots \;\; \exp(-j\frac{2\pi\lambda\vartheta \cdot N_R}{d})\right]^\top. \tag{5.72}$$

Here, $\lambda$ is the carrier wavelength, and $d$ denotes the spacing between antenna elements. Taking $a_R(\vartheta)$ as the columns of $\mathbf{A}_R$ and uniformly selecting values of $\vartheta$ in $[-1, 1]$, $\mathbf{A}_R$ would be a Fourier transform matrix with dimensions $N_R \times N_R$: If $\vec{y} = \mathbf{A}_R\vec{x}$, then $\vec{y}$ would be the fast Fourier transform of $\vec{x}$[Stoica and Moses(1997)]. It is noted that $\mathbf{A}_R$ is unitary. $\mathbf{A}_T$ can be defined in a similar manner.

The VCR proposes that the fading matrix corresponds to:

$$\mathbf{H} = \mathbf{A}_R(\mathbf{\Omega} \odot \mathbf{H}_{iid})\mathbf{A}_T^\dagger. \tag{5.73}$$

Comparing (5.73) and (5.62), the VCR is exactly the same as the W-model, except that the adaptive eigen basis $\mathbf{U}_R$ and $\mathbf{U}_T$ are replaced by fixed Fourier transform matrices $\mathbf{A}_R$, $\mathbf{A}_T$. Hence, the VCR is a special case of the W-model.

According to the above conclusion, the eigen basis of $\mathbf{G}$ would be:

$$\mathbf{U} = \mathbf{A}_R \otimes \mathbf{A}_T. \tag{5.74}$$

It can be verified that the Kronecker product of two Fourier transform matrices is still a Fourier transform matrix [Regalia and Mitra(1989)]. Therefore, $\mathbf{U}$ would be a Fourier transform matrix with dimensions $N_RN_T \times N_RN_T$. Hence, in the eigen structure of the total correlation matrix, there is no any constraint on eigen values, but the eigen basis must be a Fourier transform matrix.

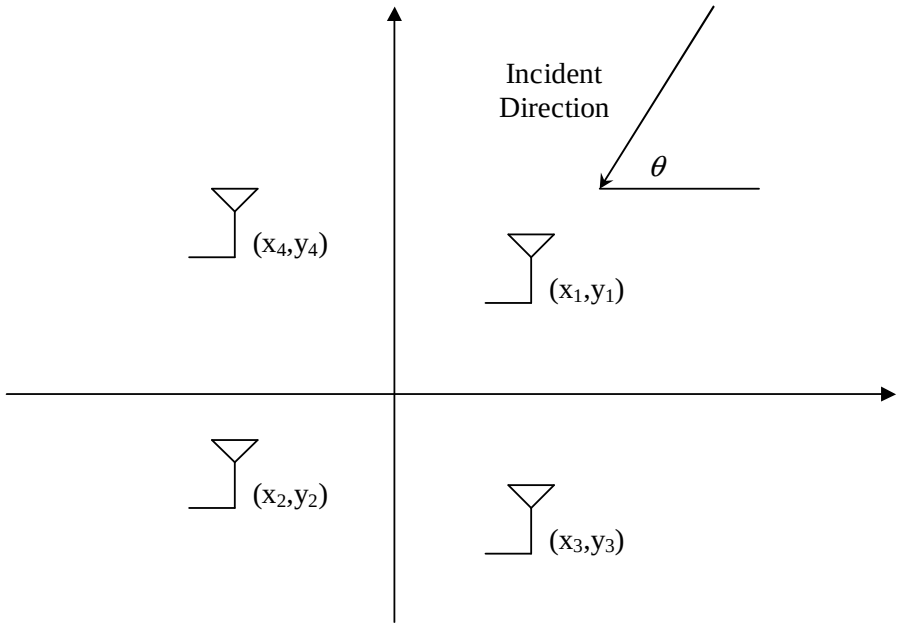


Figure 5.10: An Example of Non-ULA Array

Now the question is, under what condition, the eigen basis of a complex Hermitian Toeplitz matrix would be a Fourier transform matrix? [Gray(1977)] answers this question: The eigen basis matrix of $\mathbf{G}$ would approach a Fourier transform matrix, if and only if the dimension of $\mathbf{G}$ approach infinity, i.e., $N_R N_T \to \infty$. Hence, the VCR, although makes a good intuition, is rather an approximation *when the number of antennas is large*. This conclusion is analogous the conclusion by Slepian that Fourier transform is optimal only if the data length is infinity [Stoica and Moses(1997)].

**Extending the VCR to Non-ULA Cases**

In [Hong et al.(2003)], it is claimed that the VCR is suitable only for ULA, while the KPF is applicable for both ULA and non-ULA; hence, the VCR is a special case of the KPF. However, it should be noted that although the original VCR is proposed with ULA, the essential concept of the VCR does not necessarily assume ULA. To understand how to extend VCR to non-ULA cases, we have the following example.

An $N$-elements 2D antenna array is shown in Fig. 5.10. This array may exhibit any arbitrary shape. The matrix of the positions of the array elements, $\mathbf{D}$, is defined as:

$$\mathbf{D} = \begin{bmatrix} y_1 & y_2 & \cdots & y_N \\ x_1 & x_2 & \cdots & x_N \end{bmatrix}^T \tag{5.75}$$

Now, with a plane wave in an incident direction $\theta$, what is the array response vector corresponding to this direction? With simple geometry, it can be found that:

$$\vec{a}(\theta) = \exp\left(-j2\pi \left(\mathbf{D} \cdot [\cos\theta \ \sin\theta]\right)/\lambda\right). \tag{5.76}$$

When all elements of the array are placed on the $X$-axis with the same spacing, (5.76) is simplified to ULA case, i.e., (5.72).

Here, $\vec{a}(\theta)$ still represents a Fourier transform: it performs a proper sum and delay operation, and can filters out signals in the direction of $\theta$ for even an irregular array. It is determined by only array shape, but independent from scatterers' positions. Setting $\vec{a}(\theta)$ a column vector of $\mathbf{A}_T$, $\mathbf{A}_T$ can be found. Advanced radar systems have been built based on the above analysis [Palmer et al.(1998)]. Overall the key assumption of the VCR is the assumption of a large number of antennas, but it does not necessarily assume ULA.

### 5.4.3 The KPF: Separable Eigen Values and Separable Eigen Basis

The Kronecker product form (KPF) is the most widely used MIMO channel model, mainly because it is a natural extension of one dimensional correlated channel models. However, it may not be a proper selection in certain environments.

For the KPF, the fading matrix corresponds to [Kermoal et al.(2001)]:

$$\mathbf{H} = \mathbf{G}_R^{\frac{1}{2}} \mathbf{H}_{iid} \mathbf{G}_T^{\frac{1}{2}\top}. \tag{5.77}$$

Here, the matrixes $\mathbf{G}_T^{\frac{1}{2}}$ and $\mathbf{G}_R^{\frac{1}{2}}$ satisfy:

$$\mathbf{G}_T^{\frac{1}{2}} \mathbf{G}_T^{\frac{1}{2}\dagger} = \mathbf{G}_T \quad \text{and} \quad \mathbf{G}_R^{\frac{1}{2}} \mathbf{G}_R^{\frac{1}{2}\dagger} = \mathbf{G}_R, \tag{5.78}$$

where $\mathbf{G}_T$ and $\mathbf{G}_R$ are transmit and receive correlation matrices, with dimensions $N_T \times N_T$ and $N_R \times N_R$, respectively.

For KPF, the total correlation matrix has been well known [Kermoal et al.(2001)]:

$$\mathbf{G} = \mathbf{G}_R \otimes \mathbf{G}_T. \tag{5.79}$$

Respectively, the singular value decomposition of $\mathbf{G}_R$ and $\mathbf{G}_T$ corresponds to:

$$\mathbf{G}_R = \mathbf{U}_R \mathbf{\Lambda}_R \mathbf{U}_R^\dagger \quad \text{and} \quad \mathbf{G}_T = \mathbf{U}_T \mathbf{\Lambda}_T \mathbf{U}_T^\dagger. \tag{5.80}$$

Combining (5.80) and (5.79), the total correlation matrix have the following eigen structure:

$$\mathbf{G} = (\mathbf{U}_R \otimes \mathbf{U}_T)(\mathbf{\Lambda}_R \otimes \mathbf{\Lambda}_T)(\mathbf{U}_R \otimes \mathbf{U}_T)^\dagger. \tag{5.81}$$

Note that both its eigen value matrix and eigen basis matrix are Kronecker product of two other matrices. Therefore, it is clear that for KPF the total correlation matrix has *separable eigen basis and separable eigen values.*

It has been known that separable eigen basis corresponds to separable antenna arrays, then: what does separable eigen values correspond to? Here it is claimed that *separable eigen values correspond to separable scatterers*: If all scatterers are distributed close to either transmitter or receiver, then KPF would be capable of describing this environment. Otherwise, KPF would not be able to describe the channel.

In what physical environments, scatterers would be separable? The answer is in macrocell environments, e.g., cellular environments, scatterers would be separable. According to geometrically based channel model [Ertel et al.(1998)], scatterers in macrocell environments form a two ring shape. The two scatterer rings are around base station and mobile, respectively. In this case, KPF is suitable since scatterers are separable.

On contrast, in microcell environments, e.g., Wi-Fi, scatterers would form an elliptic shape [Ertel et al.(1998)], while the transmitter and receiver are on the two focals of the ellipse. In this situation, KPF would not be applicable since scatterers are not separable. The above conclusion, which is about the suitabliity of KPF in macrocell and microcell environments, has been checked via a ray tracing approach [Tong and Zekavat(d tob)].

## 5.5 A Preliminary Error Analysis with VCR

In this section, we numerically evaluate the error performance of OSTBC in spatially correlated channels, with both KPF and VCR. As we discussed, VCR is a realistic model while KPF is not. Now we are interested in a question, assume we have a channel generated via VCR, if this channel is misinterpreted as a KPF channel, what will be the error performance of this particular KPF channel? Answer to this question has good practical values: we will see why KPF was so popular although it is not a correct model

### 5.5.1 Error Probability Evaluation

First, we consider a particular VCR model, in which the virtual variance matrix, $\mathbf{\Psi}$, corresponds to:

$$\mathbf{\Psi} = \frac{N_R}{1 + z(N_T N_R - 1)} \begin{bmatrix} 1 & z & \cdots & z \\ z & z & \cdots & z \\ \vdots & \vdots & \ddots & \vdots \\ z & z & \cdots & z \end{bmatrix}, \tag{5.82}$$

where the dominance factor, $z$, determines the channel statistics. If $z = 1$, the channel is completely independent; If $z = 0$, the channel is completely coherent; if $0 \leq z \leq 1$, the channel is correlated, and, the correlation between antenna elements decreases with $z$.

Second, if the above VCR model is misinterpreted as a KPF channel, the local correlation matrix $\mathbf{G}_R$ would correspond to:

$$\mathbf{G}_R = Æ[\mathbf{H}^\dagger \mathbf{H}] = \mathbf{A}_R \mathbf{\Lambda}_R \mathbf{A}_R^\dagger, \tag{5.83}$$

where

$$\mathbf{\Lambda}_R = \frac{N_R \begin{bmatrix} 1+z(N_T-1) & 0 & \cdots & 0 \\ 0 & zN_T & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & zN_T \end{bmatrix}}{1+z(N_T N_R - 1)}. \tag{5.84}$$

$\mathbf{G}_T$ can be defined similarly.

Here, we also see the reason why we say KPF incorporate two marginal distributions to describe a joint distribution: Considering $\mathbf{\Psi}$ as a joint distribution, the diagonals of $\mathbf{Lambda}_T$ and $\mathbf{\Lambda}_R$ would be the marginal distributions.

Now, we have all the necessary components, $\mathbf{\Lambda}_T$, $\mathbf{\Lambda}_R$, and $\mathbf{\Psi}$, for computing error performance of OSTBC. The error performance results are depicted in Fig. 5.11 and Fig. 5.12.

Observing Fig. 5.11, we notice that: 1)KPF and VCR have the same performance in either independent or coherent channels; this is understandable since both VCR and KPF can describe the two extreme cases; 2) KPF always predicts a lower error probability than the real value (VCR results); and 3) the error probability difference between VCR and KPF results are pretty small; in almost all $z$ regions, the difference is almost negligible. We believe this is the reason why KPF was reported as a valid channel model[Kermoal et al.(2001)]. We notice that the experiment was conducted in an indoor environment, in which low values of $z$ may not be applicable.

From Fig. 5.11, we know that when $0.01 \leq z \leq 0.2$, the difference between VCR and KPF results is significant. Hence, we choose some $z$ value in this range, and plot the error performance, as shown in Fig. 5.12. As we expect, the error probability decreases as $z$ increases. More importantly, we see that KPF greatly overestimate error performance in those cases.

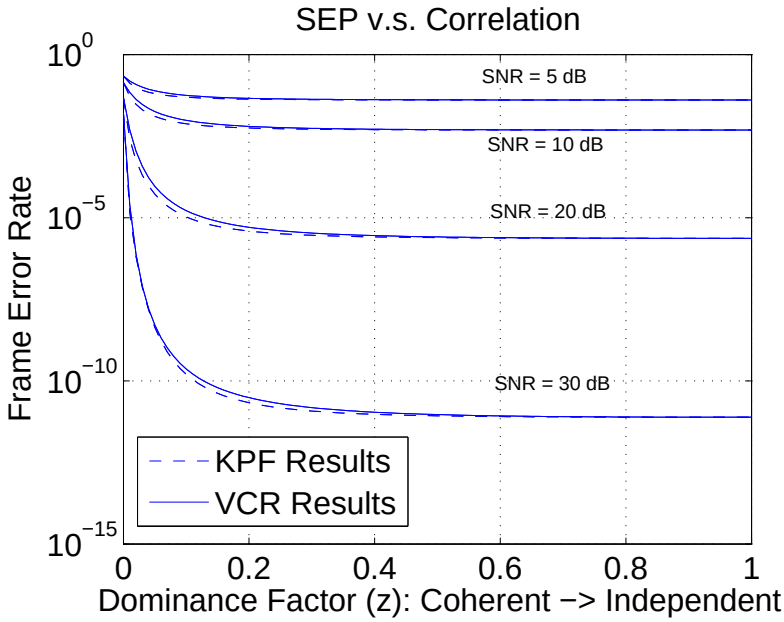


Figure 5.11: Error Probability of STC in Correlated Channels: Error Probability with respect to Correlation ($N_T = N_R = 4$)

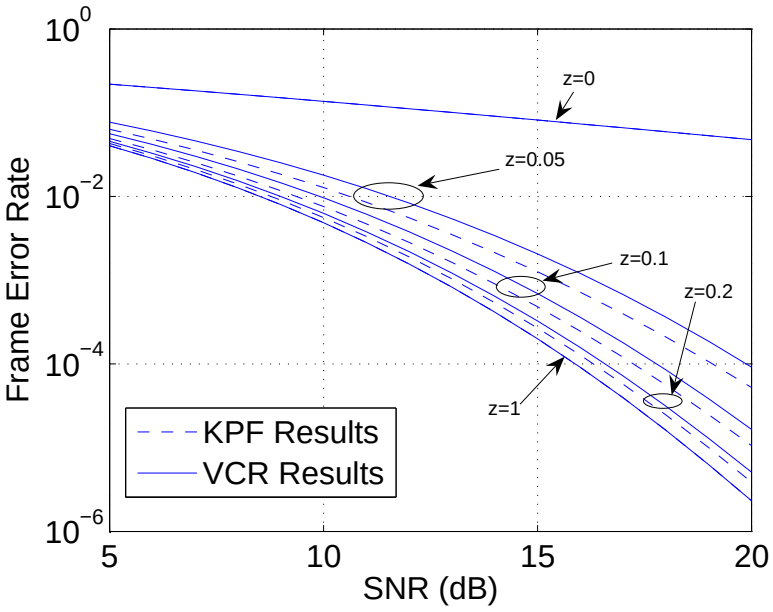


Figure 5.12: Error Probability of STC in Correlated Channels: Error Probability with respect to SNR ($N_T = N_R = 4$)

### 5.5.2 Relating Capacity and Error Performance

In addition to the error performance in both VCR and KPF, the capacity performance of MIMO systems in VCR and KPF has been studied. We first verify the experiment results[Ozcelik et al.(2003)] via a simulation. Assuming channel state information at receiver and no channel information at transmitter, (probably) the most famous instantaneous channel capacity expression corresponds to:

$$\mathcal{C} = \log\det\left(\mathbf{I}_{N_R} + \frac{SNR}{N_T}\mathbf{H}\mathbf{H}^{\dagger}\right) \tag{5.85}$$

This equation is feasible for any spatial correlations.

Now, incorporating the particular VCR model in (5.82), and the corresponding KPF model in (5.84) (if VCR is misinterpreted as a KPF model), we simulate channel fadings using (5.77) and (5.73). Then, the instantaneous capacity is computed via (5.85). The ergodic capacity (mean of instantaneous capacity) with respect to SNR is shown in Fig. 5.13. Analyzing those simulation results, we notice that there is a violent agreement between our simulation and the experiment results[Ozcelik et al.(2003)]. When $z = 0.01$ and $SNR = 25\,dB$, we see the capacity for VCR is about $10\,bps/Hz$ and capacity for KPF is about $9\,bps/Hz$, which matches with 10% capacity reduction conclusion reported in[Ozcelik et al.(2003)].

Now, we see in the same situation, KPF overestimates error performance and underestimates capacity *simultaneously*; this leads to a new question: is there any inherent relationship between error performance and capacity? To answer this question, we study the distribution of $\mathcal{C}$, as shown in Fig. 5.14.

Note that lower C values correspond to worse channel states (deeper fades). Based on the PDF results sketched in Fig. 4, the probability that KPF achieves a low capacity values is lower than VCR. Hence, KPF leads to better error performance rates than VCR as confirmed in Fig. 2. While as seen in Fig. 4 the ergodic (average)

capacity of KPF is lower than VCR. Hence, smaller error probability is achieved in price of lower capacity.

Note that, when $\mathcal{C}$ is low, channel state is bad (deep fading). Analogous to SISO fading systems, the deep fading situation is the major error component. This is verified by results in Fig. 5.14: error probability is consistent with the capacity distribution around $\mathcal{C} = 0$.

We note that a similar conclusion has been made in[Wang and Giannakis(2003)]: it is shown that diversity gain is consistent with SNR distribution at the vicinity of the origin. The main difference between SNR distribution and capacity distribution, is capacity distribution assumes maximum likelihood detection, while SNR distribution may vary with different detection process. Hence, we can safely assume error performance is by determined the capacity distribution at the vicinity at the origin.

Based on these observation, we conjecture that there is an optimal tradeoff between capacity and optimal error performance: In one channel, we can achieve higher throughput capacity than another channel. Then, the optimal error performance we can achieve in the former channel must be worse than the latter one. The mentioned tradeoff matches with our intuitions: high risk (error) leads to high benefit (capacity), and vice versa.

The conjectured tradeoff, if it is true, would bave a profound impact on MIMO system design: In contrast to current MIMO designs, which directly optimizes error performance or capacity, the optimal MIMO system design must consider both capacity and error performance. The optimal design should optimize capacity under the constraint of error performance, or vice versa. A more thorough discussion of the tradeoff is undergoing right away.

## 5.6 Beyond the Correlation Problem: Keyholes

Up to now, it has been shown that the total correlation matrix framework is a comprehensive model for correlation structure of Rayleigh fading matrix. However,

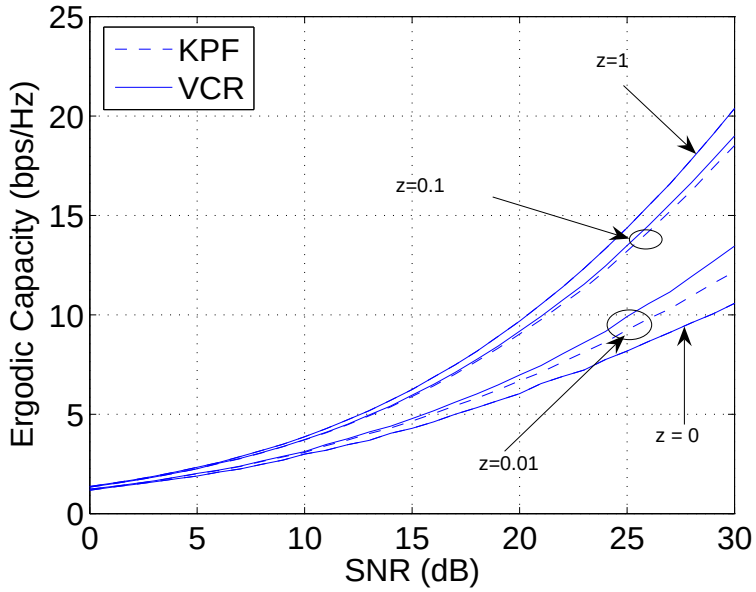


Figure 5.13: Capacity Performance in Correlated Channels ($N_T = N_R = 4$)

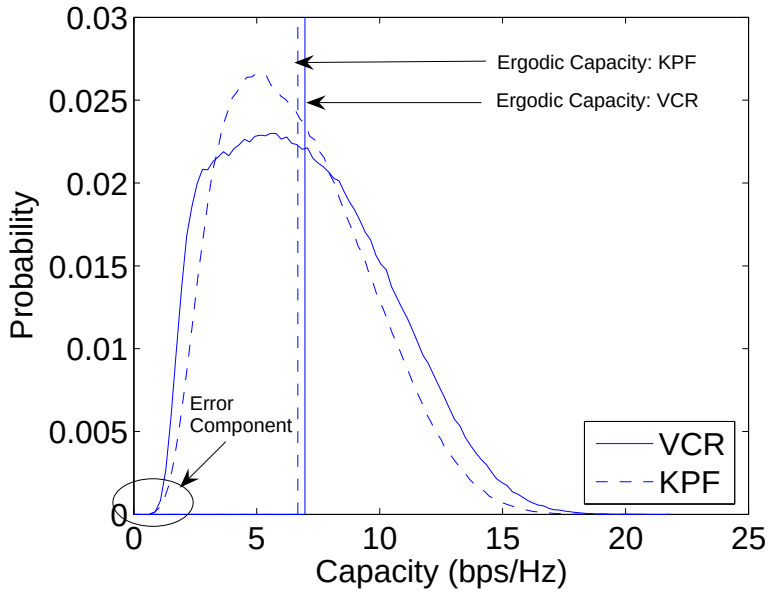


Figure 5.14: Distribution of Instantaneous Capacity ($N_T = N_R = 4$, $SNR = 20\,dB$ and $z = 0.01$)

correlation consideration is not the only issue in physical environment modeling. For example, the keyhole problem is not in the scope of correlation considerations. Consequently, Gesbert model [Gesbert et al.(2002)], which considers the keyhole problem of MIMO channel, can not be included in the total correlation matrix framework.

It has been well known that the throughput capacity gain of MIMO channel is mainly determined by the rank of fading matrix $\mathbf{H}$ [Telatar(1999)]: high rank of $\mathbf{H}$ leads to high throughput capacity gain. Note that the rank of $\mathbf{H}$ is a random variable. It was first believed that channel correlation structure fully determines the statistics of the rank of $\mathbf{H}$: Low correlation leads to high rank and high correlation leads to low rank. But soon it was discovered that correlation structure is not sufficient to determine the statistics of rank of $\mathbf{H}$ [Chizhik et al.(2002)]: Keyhole effect is not included in correlation structures.

The simplest keyhole model corresponds to:

$$\mathbf{H} = \vec{\alpha} \cdot \vec{\beta}^{\top} \tag{5.86}$$

where $\vec{\alpha} = [\alpha_1, \alpha_2, \cdots, \alpha_{N_T}]^{\top}$ $\left(\vec{\beta} = [\beta_1, \beta_2, \cdots, \beta_{N_R}]^{\top}\right)$ is a column vector that represents the channel between the Tx $\left(\text{Rx}\right)$ antenna elements and the hole, as illustrated in Fig. 5.15. Both $\vec{\alpha}$ and $\vec{\beta}$ contain independent zero-mean complex Gaussian random variables. It is easy to verify that the elements of $\mathbf{H}$ are statistically independent, however, the rank of $\mathbf{H}$ is still unity. Here, we observe that the correlation structure does not fully determine the statistics of rank of $\mathbf{H}$. This effect is called keyhole effect. As shown in Table 5.2, only when the channel has both low correlation and no keyhole, the channel fading matrix would possess a high rank, i.e., high throughput capacity. The general keyhole effects can be described by Gesbert model.

The Gesbert channel fading generation corresponds to:

$$\mathbf{H} = \mathbf{G}_R^{\frac{1}{2}} \mathbf{H}_R \boldsymbol{\Upsilon} \mathbf{H}_T^{\dagger} \mathbf{G}_T^{\frac{1}{2}\top} \tag{5.87}$$

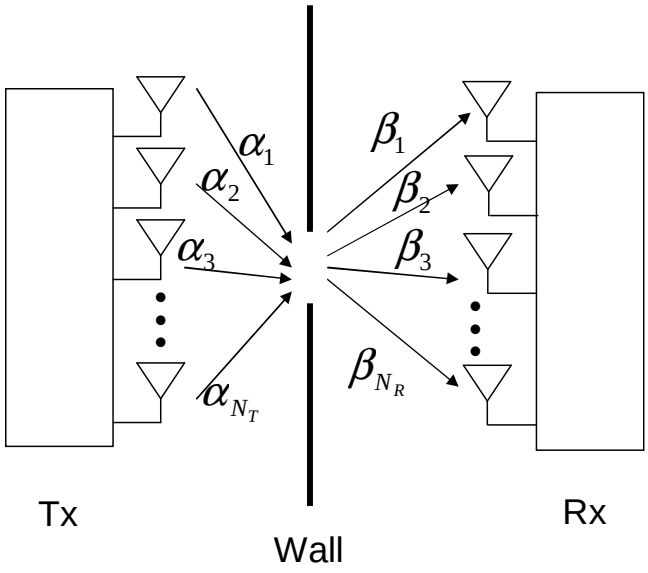


Figure 5.15: Simplest Keyhole Channel

where $\mathbf{G}_R^{\frac{1}{2}}$, and $\mathbf{G}_T^{\frac{1}{2}\dagger}$ are the same as those introduced for the KPF. Moreover, $\mathbf{H_R}$ and $\mathbf{H_T}$, denotes two matrices consists of i.i.d. complex zero-mean Gaussian random variables. Here, $\boldsymbol{\Upsilon}$ denote a rank deficient matrix (detailed structure in [Gesbert et al.(2002)]). When $\boldsymbol{\Upsilon}$ is full rank, it would be an identity matrix; in this case, Gesbert model reduces to the KPF. When $\boldsymbol{\Upsilon}$ has a low rank, Gesbert model would describe keyhole effects.

Observing (5.87), Gesbert model has a double decoupled structure. First, it decouples keyhole and correlation phenomenon into two separate structures: $\boldsymbol{\Upsilon}$ (for keyhole) and $\mathbf{G}_R^{\frac{1}{2}}$, $\mathbf{G}_T^{\frac{1}{2}}$ (for keyhole), which greatly facilitate the modeling process. Second, it decouples scatterers into transmit scatterers and receive scatterers: ($\mathbf{G}_R^{\frac{1}{2}}$, $\mathbf{H_R}$) and ($\mathbf{G}_T^{\frac{1}{2}}$, $\mathbf{H_T}$). As we depict in [Tong and Zekavat(d tob)], the second decoupling, transmit and receive scatterer separation, is not a realistic approach in all cases.

From the above discussion, we can see an interesting duality between W-model and Gesbert model. First, both of them assume separable transmitter and receiver,

Table 5.2: Effects of Correlation and Keyholes on Rank of Fading Matrix

| | With Keyhole | Without Keyhole |
|---|---|---|
| Low Correlated | Low Rank | High Rank |
| Highly Correlated | Low Rank | Low Rank |

Table 5.3: Channel Model Comparison

| | Correlation Only | Correlation and Keyhole |
|---|---|---|
| Local Scatterers Only | Kronecker product form | Gesbert Model |
| Both Local and Non-Local Scatterers | virtual channel representation or W-model | The Most Comprehensive Model is absent yet |

which is generally satisfied in MIMO systems. Second, the W-model considers correlation structure for both local and non-local scatterers; Gesbert model considers correlation structure and keyhole problem only for local scatterers. As shown in Table 5.3, there is no model that is capable describing non-separable correlation and keyhole problem simultaneously yet. Now, an interesting problem would be: is there any application that experiences both non-separable correlation and keyhole effect? If so, what model would describe the environments for this application? To the best of the authors' knowledge, this problem is still open in literature.

## 5.7 Conclusions

This chapter discusses the suitability of the Kronecker product form (KPF) in various environments. First, the Kronecker product form is shown to exhibit separable correlation structures. Next, various physical environments are checked if they possesses separable correlation structure. It is found that macrocell environments without line-of-sight has separable correlation because scatterers in macrocell environments are separable. Therefore, the KPF is capable of describing macrocell environments without line-of-sight. On the other hand, scatterers in microcell environments are not separable, hence microcell environments does not experience separable correlations, i.e., the KPF is not valid in microcell environments. Finally, a generalization of the KPF is proposed to described macrocell environments with line-of-sight.

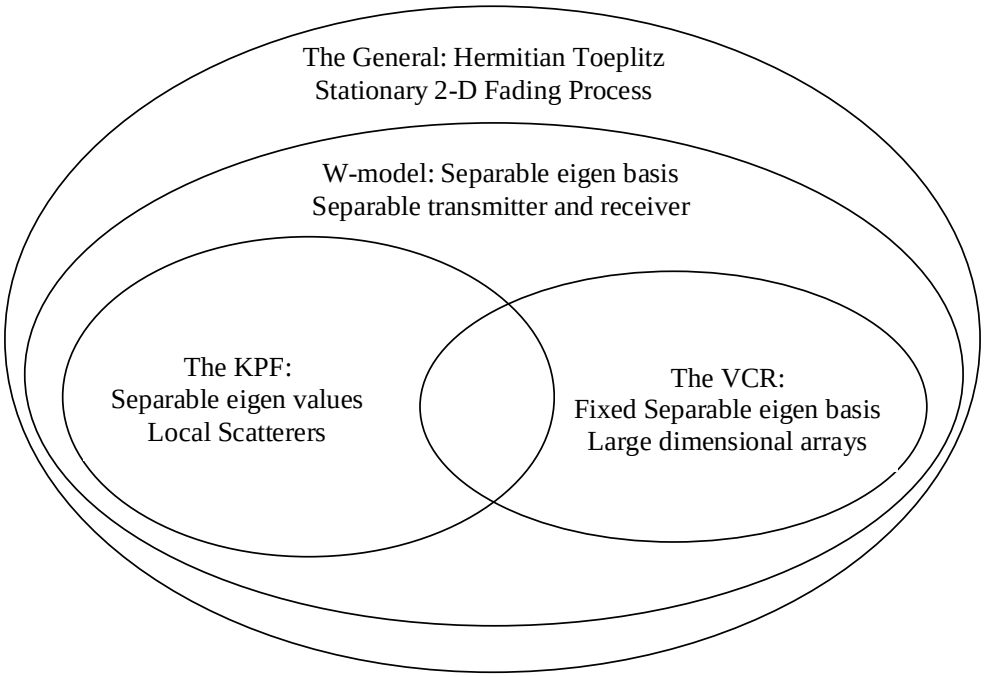


Figure 5.16: Main conclusions of this chapter.

In this chapter, we first introduce the most comprehensive description of MIMO fading matrix description: the total correlation matrix. Then, the eigen structure of the total correlation matrix is investigated for various correlated MIMO channels: W-model, the virtual channel representation, and the Kronecker product form. It is further noted that different eigen structure corresponds to different physical environments. The eigen structure and the corresponding physical environments for the aforementioned models are summarized in Fig. 5.16.

# Chapter 6

# The Diversity-Multiplexing Tradeoff in Correlated MIMO Channels

The diversity-multiplexing tradeoff has been well studied for *rich scattering* multi-input-multi-output (MIMO) channels. Defining outage probability as the probability that the random channel capacity is smaller than transmission rate, the tradeoff can be demonstrated through computing the outage probability. Traditionally outage probability is computed by the joint distribution of eigen values of the fading matrix.

In this work, it is depicted that the outage probability in MIMO systems with asymptotically high signal-to-noise ratio can be computed by decomposing it into the production of multiple conditional outage probabilities. Each conditional outage probability corresponds to a parallel single-input-single-output (SISO) system. Hence, the tradeoff in MIMO system can be decomposed to a summation of the tradeoffs in multiple parallel SISO systems. It is noted that this approach does not necessarily need the knowledge of the joint distribution of the eigen values of the fading matrix.

In the next step, the above approach is extended to *poor scattering* MIMO channels. The Unitary-Independent-Unitary models are adopted to describe the fading matrix in such environments. Because the closed form expression of eigen values

of such fading matrices is mathematically difficult to compute, the tradeoff in poor scattering channels was considered hard to obtain. Now, through the approach that decomposes outage probability, the tradeoff in poor scattering environments can be found. Closed form results for several special types of poor scattering environments are presented in detail as well.

## 6.1 Introduction

Future generations of wireless communications demand a) high data rate and b) low error rate [Paulraj et al.(2004)]. However, as information theory has revealed, in general these two performance metrics can not be optimized simultaneously. Therefore, there is a tradeoff between data rate and error rate.

The tradeoff between data rate and error rate varies with different applications. Therefore, it possesses different names in history. In additive white Gaussian noise (AWGN) channels, it is called "error exponent"[Gallager(1968)]. More recently, it is studied in multi-input-multi-output (MIMO) fading channels with asymptotically high signal-to-noise ratio (SNR) [Zheng and Tse(2003)]. In such environments, it is called "the diversity-multiplexing tradeoff". Here, diversity and multiplexing correspond to error rate and data rate in high SNR regimes, respectively.

Original works on the tradeoff [1] assumes rich scattering environments, i.e., fadings across different antenna pairs are independent. However, in real applications, scattering sources can be poor [Sayeed(2002)]. For example, rural, on highway, or on the top of a skyscraper. Therefore, practically fadings across different antenna pairs are not independent: they are correlated. Hence, it is necessary to study the tradeoff in poor scattering environments.

[1] Hereafter "the tradeoff" refers to "the diversity-multiplexing tradeoff".

There have been a number of models that try to describe poor scattering environments. Gesbert model depicts poor scattering due to lack of remote scatterers, i.e., keyhole problem [Gesbert et al.(2002)]. On contrast, unitary-independent-unitary (UIU) model corresponds to poor scattering due to lack of local scatterers [Sayeed(2002)]. The tradeoff using Gesbert model has been studied in [Yang and Belfiore(2006)]. The goal of this work is to study the tradeoff based on UIU model. It should be noted that the most widely used Kronecker product form correlation structure is a special case of the UIU model [Weichselberger et al.(2006)].

Finding the tradeoff is equivalent to computing the outage probability, which is defined as the probability that the random channel capacity is smaller than the transmission rate. For rich scattering or Gesbert model or Kronecker product form, the outage probability can be calculated through the knowledge of the joint distribution of the eigen values of the fading matrix. However, this approach is not viable for the general UIU model, because the joint distribution of the eigen values in such models is a mathematically open problem [Veeravalli et al.(2005)].

In this work, we first depict that the outage probability can be decomposed into a production of multiple conditional outage probabilities. Furthermore, it is found that each conditional outage probability is equivalent to the outage probability of a parallel SISO channel (in terms of measuring diversity gain). Therefore, the tradeoff in MIMO systems is the summation of the tradeoff in multiple parallel SISO channels. This approach does not need the computation of exact joint distribution of eigen values. The tradeoff based on UIU models is computed using this approach ( outage probability decopos.

The contributions of this work are: 1) showing that the outage probability can be decomposed, which leads to a very novel, interesting and intuitive interpretation of MIMO advantages, as shown in Section 6.2, and 2) establishing the tradeoff via the UIU model.

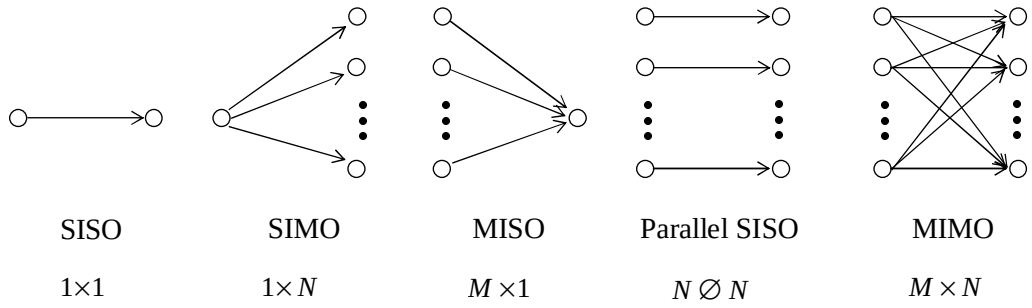


Figure 6.1: Channel Types

The rest of this chapter is organized as follows: Section 6.2 introduces definitions and a preliminary example; Section 6.3 derives the tradeoff in rich scattering environments, using the approach that decomposes outage probability; Section 6.4 establish the tradeoff based on the UIU model; Section 6.5 concludes this chapter.

## 6.2 Preliminaries and An Example

As shown in Fig. 6.1, single-input-single-output (SISO) system is a $1 \times 1$ system. Similarly, single-input-multi-output (SIMO), multi-input-single-output (MISO), and MIMO systems are called $1\times N$, $M\times 1$ and $M\times N$ systems, respectively. Here, $M$ and $N$ represent the number of transmit and receive antennas, respectively. Moreover, to distinguish a parallel SISO systems from an $N \times N$ MIMO system, it is denoted as $N \oslash N$.

In this section, we first introduce the definitions of multiplexing gain and diversity gain. We explain how to obtain the tradeoff through computing outage probability. Then, a simple outage computation shows that the tradeoff in $1 \times 1$, $1 \times N$, $M \times 1$, and $N \oslash N$ systems exhibits a simple linear relationship. Finally, we preliminarily demonstrate how to find the tradeoff in rich scattering MIMO channels through our outage decomposition approach. $1 \times 1$, $1 \times 2$, $2 \times 2$, $2 \times 3$ and $3 \times 3$ systems are studied successively. The tradeoff for the general $M\times N$ MIMO system with independent fadings is computed in Section 6.3, using the consistent outage decomposition approach.

### 6.2.1 Definitions

Multiplexing gain is defined as the ratio between transmission rate and the logarithm of SNR, in asymptotically high SNR regimes, i.e.,

$$r = \lim_{\eta \to \infty} \frac{R}{\log \eta}. \tag{6.1}$$

Here, $r$, $R$, and $\eta$ denote the multiplexing gain, transmission rate and SNR, respectively. Multiplexing gain is essentially the transmission rate in high SNR regimes. Because transmission rate may approach infinity when SNR is high, it has to be normalized by the logarithm of SNR to maintain a meaningful discussion, which yields the definition of multiplexing gain.

Diversity gain is defined as the ratio between the logarithm of error probability and the logarithm of SNR, in asymptotically high SNR regimes, i.e.,

$$d = -\lim_{\eta \to \infty} \frac{\log P_e}{\log \eta}. \tag{6.2}$$

Here, $d$ and $P_e$ denote diversity gain and error probability, respectively. By this definition, diversity gain corresponds to the slope of error rate curve in high SNR regimes. Hence, the diversity gain is the most critical measurement of error probability in our interested area (high SNR regimes).

It should be noted that error probability varies with different transmission/reception schemes, such as modulation constellation, space-time coding, zero-forcing decoding, etc. Then, diversity gain is different for different signal processing schemes. In general, it is very difficult to find the tradeoff through calculating the error probability, due to the large number of available signal processing algorithms. Diversity gain for a channel is defined as the upper bound of the diversity gain for all possible signal processing schemes, given certain multiplexing gain and channel fading statistics. Hereafter "diversity gain" refers to "diversity gain for a channel", but not "diversity gain for a signal processing scheme", unless otherwise stated.

It has been shown that error probability can be replaced by outage probability, in terms of measuring diversity gain. In other words [Zheng and Tse(2003)],

$$d = -\lim_{\eta\to\infty} \frac{\log P_{out}}{\log \eta}. \tag{6.3}$$

The outage probability, $P_{out}$ corresponds to the probability that the random channel capacity in fading channels is smaller than the transmission rate,

$$P_{out} = P(\mathcal{C} < r \log \eta). \tag{6.4}$$

For a single user, average power constrained MIMO system in non-ergodic channels, assuming no channel state information at the transmitter, the random channel capacity corresponds to:

$$\mathcal{C} = \log \det(\mathbf{I} + \frac{\eta}{M}\mathbf{H}\mathbf{H}^{\dagger}), \tag{6.5}$$

where $\mathbf{H}$ denotes MIMO channel fading matrix. Channel capacity in SISO, SIMO, MISO and parallel SISO can be considered as special cases of (6.5).

Combining (6.3), (6.4), and (6.5), we can explicitly express diversity gain as a function of multiplexing gain and fading statistics,

$$d = -\lim_{\eta\to\infty} \frac{P(\det(\mathbf{I} + \frac{\eta}{M}\mathbf{H}\mathbf{H}^{\dagger}) < \eta^r)}{\log \eta}. \tag{6.6}$$

It is clear that diversity gain is a function of multiplexing gain $r$ and the statistics of $\mathbf{H}$.

Assuming Rayleigh fading, entries of $\mathbf{H}$ would be circularly symmetric Gaussian random variables (RV). Then, the correlation structure of $\mathbf{H}$ would fully describe fading statistics. The tradeoff assuming *independent* entries of $\mathbf{H}$ has been well studied in [Zheng and Tse(2003)] through the knowledge of joint distributions of eigen values of $\mathbf{H}$. The goal of this chapter is to study the tradeoff with *correlated* entries of $\mathbf{H}$. Because the joint distribution of eigen values for correlated $\mathbf{H}$ is a

mathematically open problem, it is necessary to find an alternative approach to compute the outage probability in (6.6).

The approach in our work is to decompose MIMO outage probability to the production of multiple simpler outage probabilities, i.e., the outage probability in SISO, SIMO and parallel SISO. Section 6.2.2 studies those simple outage probabilities and the corresponding tradeoff. Section 6.2.3 demonstrates how to decompose MIMO outage probability to those simple outage probabilities.

### 6.2.2 The Tradeoff in SISO, SIMO, MISO and Parallel SISO

For convenience, we use the notation $\doteq$ to represent that two functions are "equivalent in terms of measuring diversity gain". Mathematically, we can write

$$f_1(\eta) \doteq f_2(\eta) \Longleftrightarrow -\lim_{\eta\to\infty} \frac{\log f_1(\eta)}{\log \eta} = -\lim_{\eta\to\infty} \frac{\log f_2(\eta)}{\log \eta} \tag{6.7}$$

It should be noted that the values of $f_1(\eta)$ and $f_2(\eta)$ themselves may not be identical: their difference in fact can be large. For example, consider $f_1(\eta) = 10^6 \cdot \eta^{-1} + \eta^{-2}$, and $f_2(\eta) = 0.01\eta^{-1}$: When the value of $\eta$ is moderate, $f_1(\eta)$ is much large than $f_2(\eta)$. However, both of them yield the same diversity gain. Hence, they are equivalent in terms of measuring diversity gain. Another more important example is that $P_e$ and $P_{out}$, as functions of $\eta$, are equivalent in terms of measuring diversity gain.

Now, we compute the tradeoff in SISO channels. Assuming non-ergodic fadings, the channel capacity in SISO corresponds to:

$$\mathcal{C}_{SISO} = \log(1 + \eta A). \tag{6.8}$$

Here, $A$ denotes the square of amplitude of channel fadings, i.e., $A = hh^*$, where $h$ is the complex channel fading. Subsequently, the outage probability in this channel corresponds to:

$$P(\mathcal{C}_{SISO} < r \log \eta) = P(\log(1 + \eta A) < r \log \eta)$$

$$\doteq P(A < \eta^{-(1-r)}) \tag{6.9}$$

The second equality is due to the fact that the constant 1 is negligible when $\eta \to \infty$, in terms of measuring diversity.

Certainly, the outage probability in (6.9) can be easily obtained because $|h|^2$ follows a simple exponential distribution. However, to keep the SISO derivation consistent with our future derivations (for SIMO and MIMO), we use another approach: we use Taylor series expansion of the cumulative distribution function (CDF) of $|h|^2$.

Because $A$ follows exponential distribution, its CDF corresponds to:

$$P(A < A_0) = 1 - e^{-A_0}.$$

Using Taylor series expansion of exponential functions,

$$e^{-A_0} = 1 - A_0 + \frac{A_0^2}{2} - ...$$

the CDF of $A$ corresponds to

$$P(A < A_0) = A_0 - \frac{A_0^2}{2} + ...$$

For a very small value of $\epsilon$,

$$P(A < \epsilon) \simeq \epsilon \tag{6.10}$$

Using (6.10), because $\eta^{-(1-r)}$ is small when SNR is high and $0 < r < 1$, the outage probability in (6.9) corresponds to:

$$P(A < \eta^{-(1-r)}) \doteq \begin{cases} \eta^{-(1-r)} & \text{if } 0 < r < 1 \\ 1 & \text{if } r > 1 \end{cases} \tag{6.11}$$

Substituting (6.9) into the definition of diversity gain in (6.2), the tradeoff in non-ergodic SISO channels turns out to be:

$$d(r) = (1 - r)^+, \tag{6.12}$$

where $(\cdot)^+$ denotes $\max(\cdot, 0)$.

Extending the SISO tradeoff to non-Rayleigh fadings is possible via the above approach. First, we let $h$ be a (possibly) non-Rayleigh fading, and assume values of $A$ span over $(0, \infty)$. Because $P(A < A_0)$ is a function of $A_0$, we can express its Taylor series expansion as

$$P(A < A_0) = c_0 + c_1 A_0 + c_2 A_0^2 + c_3 A_0^3 + ... \tag{6.13}$$

Because $A$ has zero probability to be exactly zero, $c_0$ is always zero. If $c_1 \neq 0$, then the channels with $h$ would possess the tradeoff in (6.12): Rayleigh fading is an example of this case. If $c_1 = 0$ and $c_2 \neq 0$ , then the tradeoff would be

$$d(r) = 2 \cdot (1 - r)^+, \tag{6.14}$$

Roughly speaking, this channel would have two folds of diversity. In summary, the diversity gain in SISO corresponds to the lowest power of the Taylor series in (6.13). This result matches our intuition, because higher power of $P(A < A_0)$ corresponds to lower deep fading probability in this channel. For convenience, we still use Rayleigh fading in our discussion, but it should be noted that our approach can be extend to arbitrary types of fadings.

Next, we study the tradeoff in SIMO systems. In a $1 \times N$ system, the channel capacity corresponds to:

$$\mathcal{C}_{SIMO} = \log(1 + \eta \cdot \sum_{n=1}^{N} A_n). \tag{6.15}$$

where $A_n$ is the fading power over $n^{th}$ path. Then, assuming independent fadings, the outage probability in SIMO corresponds to:

$$P(\mathcal{C}_{SIMO} < r \log \eta) \doteq P\left(\sum_{n=1}^{N} A_n < \eta^{-(1-r)}\right)$$

$$\begin{aligned}
&\doteq P(A_1 < \eta^{-(1-r)} \cap A_2 < \eta^{-(1-r)} \cap \cdots \cap A_N < \eta^{-(1-r)}) \\
&= \prod_{n=1}^{N} P(A_n < \eta^{-(1-r)}) \\
&\doteq \begin{cases} \eta^{-N\cdot(1-r)} & \text{if } 0 < r < 1 \\ 1 & \text{if } r > 1 \end{cases}
\end{aligned} \tag{6.16}$$

The first equality follows the same manner as the derivation for SISO systems.

The second equality needs more discussion. First, if for any $n \in \{1, 2, \cdots N\}$, $A_n > \eta^{-(1-r)}$, then $\sum_{n=1}^{N} A_n > \eta^{-(1-r)}$. In other words, if $\sum_{n=1}^{N} A_n < \eta^{-(1-r)}$, then $A_n < \eta^{-(1-r)}$ for any $n$. Therefore

$$P\left(\sum_{n=1}^{N} A_n < \eta^{-(1-r)}\right) \dot{\leq} P(A_1 < \eta^{-(1-r)} \cap A_2 < \eta^{-(1-r)} \cap \cdots \cap A_N < \eta^{-(1-r)})$$

Second, if for all $n \in \{1, 2, \cdots N\}$, $A_n < \eta^{-(1-r)}$, then $\sum_{n=1}^{N} A_n < N \cdot \eta^{-(1-r)} \doteq \eta^0 \cdot \eta^{-(1-r)}$, where $\eta^0$ corresponds to a deterministic value (i.e., a value that is not a function of $\eta$). Hence,

$$P\left(\sum_{n=1}^{N} A_n < \eta^{-(1-r)}\right) \dot{\geq} P(A_1 < \eta^{-(1-r)} \cap A_2 < \eta^{-(1-r)} \cap \cdots \cap A_N < \eta^{-(1-r)})$$

Combining the above two results, the second equality is obtained. This result will appear multiple times in our future discussion.

The third equality exploits the independent fading assumption. For correlated fading, the third equality may not hold. We will show how to find the tradeoff for correlated SIMO in Section 6.4.

The fourth equality has been explained in SISO derivation.

Substituting the result in (6.16) into the definition of diversity gain in (6.2), the tradeoff in SIMO corresponds to:

$$d(r) = N(1-r)^{+}. \tag{6.17}$$

Comparing the tradeoff for SIMO and SISO tradeoff, there are three notes that are worth to mention, as shown in Fig. 6.2. First, given a diversity gain $r$, the

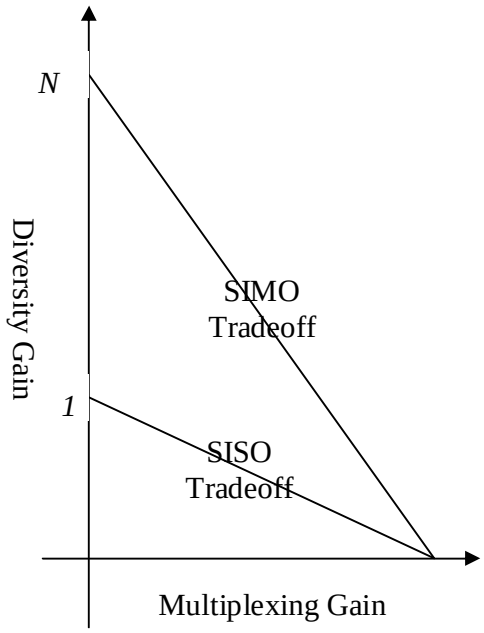


Figure 6.2: Comparison of SIMO and SISO Tradeoffs

SIMO system diversity gain is $N$ times of the diversity gain in SISO systems: This is the reason why SIMO system is more reliable than SISO. Second, it is noted that although the diversity gain is increased, the maximum possible multiplexing gain in SIMO is still the same as SISO. This means that SIMO does increase the maximum attainable transmission rate. Finally, if a given diversity gain is required (and it is smaller than one), SIMO system can use a higher transmission rate than SISO.

MISO system has the same capacity equation except that there are $M$ times SNR reduction due to power splitting between multiple transmitters. Therefore, MISO outage probability is:

$$\begin{aligned} P(\mathcal{C}_{MISO} < r \log \eta) &\doteq \prod_{n=1}^{N} P(A_n < M \cdot \eta^{-(1-r)}) \\ &\doteq \begin{cases} \eta^{-N\cdot(1-r)} & \text{if } 0 < r < 1 \\ 1 & \text{if } r > 1 \end{cases} \end{aligned} \tag{6.18}$$

Hence, SIMO and MISO systems have the same tradeoff. The reason is that the tradeoff is studied in *asymptotically* high SNR regimes: a finite reduction in SNR

does not alter its aysmptocity. The same concept will appear multiple times in our future derivations.

The final tradeoff in Section 6.2.2 is for parallel SISO. This system is not discussed as often as SISO or SIMO in the literature, but it is a key point in our work. An $N \oslash N$ system can be considered as a special case of MIMO system in which the fading matrix is restricted to be diagonal. Then, the random channel capacity in parallel SISO corresponds to:

$$\mathcal{C}_{ParallelSISO} = \log\left(\prod_{n=1}^{N}(1+\eta A_n)\right). \tag{6.19}$$

Thus, the outage probability in parallel SISO is:

$$P(\prod_{n=1}^{N}(1+\eta A_n) < \eta^r). \tag{6.20}$$

To quantitatively study the above outage probability, it is necessary to write $\prod_{n=1}^{N}(1+\eta A_n)$ in an explicit form:

$$\prod_{n=1}^{N}(1+\eta A_n) = 1 + \sum_{n=1}^{N} c_n \eta^n, \tag{6.21}$$

i.e., it is a polynomial of $\eta$. The maximum power of $\eta$ is $N$. Using a simple algebraic manipulation, it can be shown that the coefficient for the power of $n$, $c_n$, corresponds to:

$$c_n = \sum_{\mathcal{S}} \prod_{i=s_1}^{i=s_n} A_i \tag{6.22}$$

Here, $\mathcal{S}$ is a subset of $\{1, 2, \cdots, N\}$ with $n$ elements. $\sum_{\mathcal{S}}$ denotes the summation over all possible $\mathcal{S}$. Therefore, for an arbitrary $n$, there are $C_N^n = \frac{N!}{n!(N-n)!}$ elements in the summation. Please note that in (6.22), $c_n$ is a random variable for any $n$.

Using (6.21), the outage probability in parallel SISO corresponds to:

$$P(\prod_{n=1}^{N}(1+\eta|h_n|^2) < \eta^r)$$

$$
\begin{aligned}
&= P\left(\sum_{n=1}^{N} c_n \eta^n < \eta^r\right) \\
&\doteq P\left(c_N < \eta^{-(N-r)} \cap c_{N-1} < \eta^{-(N-1-r)} \cap \cdots \cap c_1 < \eta^{-(1-r)}\right) \\
&= P\left(c_N < \eta^{-(N-r)}\right) \cdot P\left(c_{N-1} < \eta^{-(N-1-r)} | c_N < \eta^{-(N-r)}\right) \cdots \\
&\quad \cdot P\left(c_1 < \eta^{-(1-r)} | c_N < \eta^{-(N-r)} \cap c_{N-1} < \eta^{-(N-1-r)} \cap \cdots \cap c_2 < \eta^{-(2-r)}\right)
\end{aligned} \tag{6.23}
$$

The first equality is a direct result from (6.21). The second equality follows the similar discussion as SIMO systems. The third equality decomposes one outage probability into the production of multiple $(N-1)$ conditional probabilities. This is the first time that we see how to decompose outage probability explicitly in this chapter. We extend this approach to MIMO in Section 6.2.3. It is noted that this decomposition is applicable in asymptotically high SNR regimes only, because the second equality holds only in high SNR regimes.

Now, we compute the probabilities in (6.23). We show that in parallel SISO systems, the first probability determines diversity, all other conditional probabilities are equivalent to $\eta^0$ in terms of measuring diversity. For MIMO system, those conditional probabilities impact the final diversity expression as well.

Using (6.22), we find that

$$
c_N = A_1 A_2 \cdots A_N. \tag{6.24}
$$

Because the multiplication of two complex Gaussian random variables still follows complex Gaussian distribution, $c_N$ follows exponential distribution. Thus, the first term in (6.23) turns out to be

$$
P\left(c_N < \eta^{-(N-r)}\right) \doteq \eta^{-(N-r)}. \tag{6.25}
$$

Next, we compute the second term in (6.23).

$$
P\left(c_{N-1} < \eta^{-(N-1-r)} | c_N < \eta^{-(N-r)}\right)
$$

$$= P\left(A_1A_2A_{N-1} + A_1A_2\cdots A_{N-2}A_N + \cdots + A_2A_3\cdots A_N < \eta^{-(N-1-r)} | c_N < \eta^{-(N-r)}\right)$$

$$\doteq P(A_1A_2\cdots A_N - 1 < \eta^{-(N-1-r)} \cap A_1A_2\cdots A_{N-2}A_N < \eta^{-(N-1-r)} \cap \cdots$$

$$\cap A_2A_3\cdots A_N < \eta^{-(N-1-r)} | c_N < \eta^{-(N-r)}) \tag{6.26}$$

Therefore, calculating the second probability is equivalent to answering the following question: given $\| h_1h_2\cdots h_N \| < \eta^{-(N-r)}$, what is the probability that $\| h_1h_2\cdots h_{N-1} \|_{\mathcal{F}}^2 < \eta^{-(N-1-r)}$? To answer this question, we calculate

$$P(A_1A_2\cdots A_{N-1} < \eta^{-(N-1-r)} | c_N < \eta^{-(N-r)})$$

$$\doteq P(A_1A_2\cdots A_{N-1} < \eta^{-(N-1-r)} | c_N < \eta^{-(N-r)} \cap A_N < \eta^{-1}) P(A_N < \eta^{-1}) +$$

$$P(A_1A_2\cdots A_{N-1} < \eta^{-(N-1-r)} | c_N < \eta^{-(N-r)} \cap A_N > \eta^{-1}) P(A_N > \eta^{-1})$$

$$= 1 \cdot (1 - \eta^{-1}) + \eta^{-1} \cdot P(A_1A_2\cdots A_{N-1} < \eta^{-(N-1-r)} | c_N < \eta^{-(N-r)} \cap A_N < \eta^{-1})$$

$$\doteq \eta^0, \tag{6.27}$$

where $\eta^0$ has been defined in the explanations of (6.16). The second equality in (6.27) can be explained by the following three expressions:

$$P(A_N > \eta^{-1}) \doteq 1 - \eta^{-1} \tag{6.28}$$

$$P(A_N < \eta^{-1}) \doteq \eta^{-1}, \tag{6.29}$$

and noting that $c_N = A_1A_2\cdots A_N$,

$$P(A_1A_2\cdots A_{N-1} < \eta^{-(N-1-r)} | c_N < \eta^{-(N-r)} \cap A_N > \eta^{-1}) = 1 \tag{6.30}$$

Therefore, if $c_N < \eta^{-(N-r)}$ then all elements in $c_{N-1}$ would be smaller than $\eta^{-(N-1-r)}$ almost surely (i.e., the probability that $c_{N-1} < \eta^{-(N-1-r)}$ approaches to $\eta^0$ when $\eta \to \infty$). Thus, the second term in (6.23) contributes zero to diversity gain calculation. Similarly, it can be shown that all of the the remaining terms contributes zero to diversity gain. The outage probability in parallel SISO corresponds to:

$$P(\prod_{n=1}^{N}(1 + \eta A_n) < \eta^r) \doteq \begin{cases} \eta^{-(N-r)} & \text{if } 0 < r < N \\ 1 & \text{if } r > N \end{cases} \tag{6.31}$$

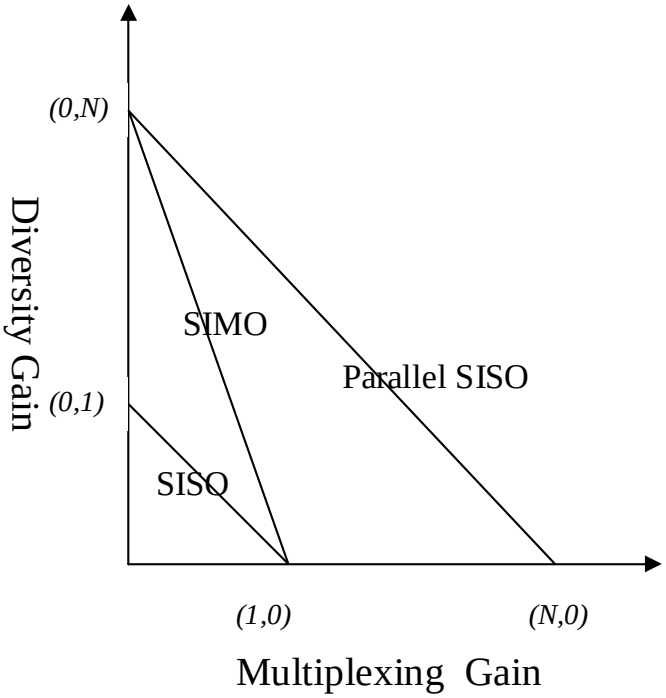


Figure 6.3: Simple Tradeoffs: SISO, SIMO and Parallel SISO

and the corresponding tradeoff in parallel SISO would be:

$$d(r) = (N - r)^{+}. \tag{6.32}$$

As expected, parallel SISO increases both maximum diversity gain and maximum multiplexing gain, while SIMO increases only maximum diversity gain.

The summary of the tradeoff in SISO, SIMO and parallel SISO is presented in Fig. 6.3. Given a certain multiplexing gain, SIMO creates more diversity than SISO, while parallel SISO enables better diversity than SIMO. Therefore, we say parallel SISO is better than SIMO, and SIMO is better than SISO. In Section 6.2.3, we see that MIMO is better than parallel SISO.

### 6.2.3 Preliminary Tradeoff Results in MIMO

Up to now, we have investigated the tradeoff in SISO, SIMO, and parallel SISO systems. Then, we can start studying the tradeoff in MIMO systems. In this section,

we present some specific cases of independent MIMO channels. The first case is the $2 \times 2$ system, which is the simplest MIMO system. We show that the tradeoff in $2 \times 2$ is the summation of the tradeoffs in one $2 \oslash 2$ and two $1 \times 1$. Following cases are $2 \times 3$ and $3 \times 3$ systems.

First, we introduce a lemma to simplify mathematical expression.

**Lemma 6.2.1.** *A finite reduction in SNR does not alter the diversity-multiplexing tradeoff. For example,*

$$P(\det(\mathbf{I} + \eta \mathbf{H}\mathbf{H}^\dagger) < \eta^r) \doteq P(\det(\mathbf{I} + \frac{\eta}{M}\mathbf{H}\mathbf{H}^\dagger) < \eta^r), \tag{6.33}$$

*where $M$ can be the number of transmit antennas.*

*Proof.* See Appendix 6.A. □

Based on this lemma, we can safely ignore the parameter $M$, which reduces SNR by the factor of $M$, due to the power splitting between transmit antennas. This helps us to express derivations concisely.

Now we present the tradeoff in some simple MIMO systems.

*Case1: $2 \times 2$ MIMO Systems.* First of all, the fading matrix in $2 \times 2$ system corresponds to:

$$\mathbf{H} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix}. \tag{6.34}$$

Then, it is easy to find that:

$$\det(\mathbf{I} + \eta \mathbf{H}\mathbf{H}^\dagger) = \eta^2 \det(\mathbf{H}\mathbf{H}^\dagger) + \eta(A_{11} + A_{12} + A_{21} + A_{22}) + 1, \tag{6.35}$$

where $A_{11}, A_{12}, A_{21}, A_{22}$ are square of amplitudes of $h_{11}, h_{12}, h_{21}, h_{22}$, respectively. Then, the outage probability in high SNR regimes can be decomposed:

$$P(\det(\mathbf{I} + \eta \mathbf{H}\mathbf{H}^\dagger) < \eta^r)$$

$$= P(\eta^2 \det(\mathbf{H}\mathbf{H}^\dagger) + \eta(A_{11} + A_{12} + A_{21} + A_{22}) + 1 < \eta^r)$$

$$\doteq P(\eta^2 \det(\mathbf{H}\mathbf{H}^\dagger) < \eta^r \cap \eta(A_{11} + A_{12} + A_{21} + A_{22}) < \eta^r)$$

$$\doteq P(\det(\mathbf{H}\mathbf{H}^\dagger) < \eta^{-(2-r)}) P((A_{11} + A_{12} + A_{21} + A_{22}) < \eta^{-(1-r)} | \det(\mathbf{H}\mathbf{H}^\dagger) < \eta^{-(2-r)}) \tag{6.36}$$

Note that the second equality follows the similar logic as (6.27). Therefore, the MIMO outage probability can be decomposed into the production of one outage probability and another conditional outage probability.

Now, we compute the first probability. Because $\mathbf{H}$ is a square matrix, we can write

$$\det(\mathbf{H}\mathbf{H}^\dagger) = \det(\mathbf{H}) \cdot \det(\mathbf{H}^\dagger) = \det(\mathbf{H}) \cdot (\det(\mathbf{H}))^*. \tag{6.37}$$

Clearly,

$$\det(\mathbf{H}) = h_{11}h_{22} - h_{21}h_{12}, \tag{6.38}$$

It is has been found that $\det(\mathbf{H}\mathbf{H}^*)$ has the same Taylor series expansion as exponential distribution when it is small [Springer and Thompson(1970)], and

$$P(\det(\mathbf{H}\mathbf{H}^\dagger) < \eta^{-(2-r)}) \doteq \begin{cases} \eta^{-(2-r)} & \text{if } 0 < r < 2 \\ 1 & \text{if } r > 2 \end{cases} \tag{6.39}$$

The tradeoff corresponding to the first outage probability is:

$$d_1(r) = (2 - r)^+. \tag{6.40}$$

It should be noted that the tradeoff is the same as a $2 \oslash 2$ system.

The second conditional probability is computed as follows. If $\det(\mathbf{H}\mathbf{H}^\dagger) < \eta^{-(2-r)}$ and $\eta \to \infty$, then the matrix $\mathbf{H}$ is singular, i.e., $[h_{11}h_{12}]$ would be a linearly scaled version of $[h_{21}\, h_{22}]$ [Brookes(2005)]. Then, the probability that $(A_{11} + A_{12} + A_{21} + A_{22}) < \eta^{-(1-r)}$ would be identical to the probability that $(A_{12} + A_{21}) < \eta^{-(1-r)}$, i.e.,

$$P((A_{11} + A_{12} + A_{21} + A_{22}) < \eta^{-(1-r)} | \det(\mathbf{H}\mathbf{H}^\dagger) < \eta^{-(2-r)})$$

$$\begin{aligned}
&\doteq P((A_{12}+A_{21})<\eta^{-(1-r)}) \\
&\doteq P(A_{12}<\eta^{-(1-r)})\cdot P(A_{21}<\eta^{-(1-r)}) \\
&\doteq \begin{cases} \eta^{-2(1-r)} & \text{if } 0<r<2 \\ 1 & \text{if } r>2 \end{cases}
\end{aligned} \tag{6.41}$$

Thus, the tradeoff corresponding to the second conditional outage probability is:

$$d_2(r) = 2\cdot(1-r)^+ = (1-r)^+ + (1-r)^+. \tag{6.42}$$

This tradeoff corresponds to two $1\times 1$ systems.

Because diversity gain corresponds to the logarithm of outage probability, the production of outage probability would turn out to be the summation of diversity. Therefore, the final tradeoff in $2\times 2$ systems is:

$$d(r) = (2-r)^+ + (1-r)^+ + (1-r)^+. \tag{6.43}$$

As we have stated, this tradeoff is the summation of the tradeoff in one $2\oslash 2$ system and two $1\times 1$ systems.

*Case 2:* $2\times 3$ *MIMO Systems.* In this case, the channel fading matrix corresponds to:

$$\mathbf{H} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{22} \end{bmatrix}. \tag{6.44}$$

Again it is easy to show

$$\det(\mathbf{I}+\eta\mathbf{H}\mathbf{H}^\dagger) = \eta^2\cdot\det(\mathbf{H}\mathbf{H}^\dagger) + \eta\cdot\sum_{n=1}^{2}\sum_{m=1}^{3} A_{nm} + 1. \tag{6.45}$$

Similar as $2\times 2$ systems, the outage probability in $2\times 3$ systems is decomposed to:

$$\begin{aligned}
&P(\det(\mathbf{I}+\eta\mathbf{H}\mathbf{H}^\dagger)<\eta^r) \\
&= P\left(\eta^2\cdot\det(\mathbf{H}\mathbf{H}^\dagger) + \eta\cdot\sum_{n=1}^{2}\sum_{m=1}^{3} A_{nm} + 1 < \eta^r\right) \\
&\doteq P\left(\det(\mathbf{H}\mathbf{H}^\dagger)<\eta^{-(2-r)}\right) P\left(\sum_{n=1}^{2}\sum_{m=1}^{3} A_{nm} < \eta^{-(1-r)} \,|\, \det(\mathbf{H}\mathbf{H}^\dagger)<\eta^{-(2-r)}\right)
\end{aligned} \tag{6.46}$$

Again, the first outage probability corresponds to:

$$P(\det(\mathbf{H}\mathbf{H}^{\dagger}) < \eta^{-(2-r)}). \tag{6.47}$$

In $2 \times 3$ systems, $\mathbf{H}$ is not square; hence, $\det(\mathbf{H})$ does not exist. Then, the method in $2 \times 2$ systems is not applicable anymore. We have to use Cauchy-Binet formula to express $\det(\mathbf{H}\mathbf{H}^{\dagger})$ as

$$\det(\mathbf{H}\mathbf{H}^{\dagger}) = \det(\mathbf{H}_1\mathbf{H}_1^{\dagger}) + \det(\mathbf{H}_2\mathbf{H}_2^{\dagger}) + \det(\mathbf{H}_3\mathbf{H}_3^{\dagger}), \tag{6.48}$$

where

$$\mathbf{H}_1 = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix}, \mathbf{H}_2 = \begin{bmatrix} h_{11} & h_{13} \\ h_{21} & h_{23} \end{bmatrix}, \text{ and } \mathbf{H}_3 = \begin{bmatrix} h_{12} & h_{13} \\ h_{22} & h_{23} \end{bmatrix}. \tag{6.49}$$

Here, the two column vectors of $\mathbf{H}_1$ correspond to the first and second columns of $\mathbf{H}$, the two column vectors of $\mathbf{H}_2$ correspond to the second and third columns of $\mathbf{H}$, and finally the two column vectors of $\mathbf{H}_1$ are the first and third columns of $\mathbf{H}$.

Substituting (6.48) into (6.47), the first outage probability turns out to be

$$\begin{aligned} & P(\det(\mathbf{H}\mathbf{H}^{\dagger}) < \eta^{-(2-r)}) \\ &= P(\det(\mathbf{H}_1\mathbf{H}_1^{\dagger}) + \det(\mathbf{H}_2\mathbf{H}_2^{\dagger}) + \det(\mathbf{H}_3\mathbf{H}_3^{\dagger}) < \eta^{-(2-r)}) \\ &\doteq P(\det(\mathbf{H}_1\mathbf{H}_1^{\dagger}) < \eta^{-(2-r)} \cap \det(\mathbf{H}_2\mathbf{H}_2^{\dagger}) < \eta^{-(2-r)} \cap \det(\mathbf{H}_3\mathbf{H}_3^{\dagger}) < \eta^{-(2-r)}). \end{aligned} \tag{6.50}$$

Equation (6.50) states that the event that $\mathbf{H}$ in outage would occur if and only $\mathbf{H}_1$, $\mathbf{H}_2$ and $\mathbf{H}_3$ are all in outage, i.e., one outage event is decomposed into three outage events. It should be noted that the three outage events are not independent: They are mutually independent but not jointly independent.

For example, the first event would happen if the second column of $\mathbf{H}$ is a linearly scaled version of its first column. Similarly, the second event would happen if the third column is a linearly scaled version of the second column. Therefore, if both the

first and second outage events happen, the third column would be a linearly scaled version of the second column. This means that the third event would almost surely happen. Therefore,

$$\begin{aligned} & P(\det(\mathbf{H}_1\mathbf{H}_1^*) < \eta^{-(2-r)} \cap \det(\mathbf{H}_2\mathbf{H}_2^\dagger) < \eta^{-(2-r)} \cap \det(\mathbf{H}_3\mathbf{H}_3^\dagger) < \eta^{-(2-r)}) \\ \doteq\ & P(\det(\mathbf{H}_1\mathbf{H}_1^\dagger) < \eta^{-(2-r)} \cap \det(\mathbf{H}_2\mathbf{H}_2^\dagger) < \eta^{-(2-r)}) \end{aligned} \tag{6.51}$$

Note that the event that $\det(\mathbf{H}_1\mathbf{H}_1^\dagger) < \eta^{-(2-r)}$ is mutually independent of the event $\det(\mathbf{H}_2\mathbf{H}_2^\dagger) < \eta^{-(2-r)}$, then

$$\begin{aligned} & P(\det(\mathbf{H}_1\mathbf{H}_1^\dagger) < \eta^{-(2-r)} \cap \det(\mathbf{H}_2\mathbf{H}_2^\dagger) < \eta^{-(2-r)}) \\ =\ & P(\det(\mathbf{H}_1\mathbf{H}_1^\dagger) < \eta^{-(2-r)}) \cdot P(\det(\mathbf{H}_2\mathbf{H}_2^\dagger) < \eta^{-(2-r)}) \\ \doteq\ & \begin{cases} \eta^{-2(2-r)} & \text{if } 0 < r < 2 \\ 1 & \text{if } r > 2 \end{cases} \end{aligned} \tag{6.52}$$

, accordingly the tradeoff corresponding to the first outage probability is

$$d_1(r) = (2-r)^+ + (2-r)^+. \tag{6.53}$$

Therefore, the first outage probability corresponds to two $2 \oslash 2$ systems.

The second outage probability is relatively easy to obtain. Given $\mathbf{H}$ in outage, each column of $\mathbf{H}$ is a linearly scaled version of another. Therefore,

$$\begin{aligned} & P\left(\sum_{n=1}^{2}\sum_{m=1}^{3} A_{nm} < \eta^{-(1-r)} | \det(\mathbf{H}\mathbf{H}^\dagger) < \eta^{-(2-r)}\right) \\ \doteq\ & P(A_{21} + A_{13} < \eta^{-(1-r)}) \\ \doteq\ & \begin{cases} \eta^{-2(1-r)} & \text{if } 0 < r < 1 \\ 1 & \text{if } r > 1 \end{cases}, \end{aligned} \tag{6.54}$$

and the tradeoff corresponding to the second outage probability would be:

$$d_2(r) = (1-r)^+ + (1-r)^+. \tag{6.55}$$

The final tradeoff is:

$$d(r) = (2-r)^+ + (2-r)^+ + (1-r)^+ + (1-r)^+, \tag{6.56}$$

which means that a $2 \times 3$ system can be considered as the summation of two $2 \oslash 2$ systems and two $1 \times 1$ systems.

*Case 3:* $3 \times 3$ *MIMO Systems.* First, the fading matrix $\mathbf{H}$ corresponds to:

$$\mathbf{H} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}. \tag{6.57}$$

Moreover, let $\vec{h}_1$, $\vec{h}_2$, $\vec{h}_3$ denote the first, the second and the third rows of $\mathbf{H}$. Then, we define

$$\mathbf{H}_1 = \begin{bmatrix} \vec{h}_1 \\ \vec{h}_2 \end{bmatrix}, \quad \mathbf{H}_2 = \begin{bmatrix} \vec{h}_2 \\ \vec{h}_3 \end{bmatrix}, \quad \text{and } \mathbf{H}_3 = \begin{bmatrix} \vec{h}_1 \\ \vec{h}_3 \end{bmatrix}. \tag{6.58}$$

Defining $\mathbf{H}_1, \mathbf{H}_2, \mathbf{H}_3$ helps us to write the following result in an elegant form:

$$\det(\mathbf{I} + \eta \mathbf{H}\mathbf{H}^\dagger) = \eta^3 \cdot \det(\mathbf{H}\mathbf{H}^\dagger) + \eta^2 \sum_{n=1}^{3} \det(\mathbf{H}_n \mathbf{H}_n^\dagger) + \eta \sum_{n=1}^{3}\sum_{m=1}^{3} A_{nm} + 1 \tag{6.59}$$

Similar to $2\times3$ systems, the outage probability in $3\times3$ systems can be decomposed into the production of three (conditional) outage probabilities:

$$\begin{aligned} & P(\det(\mathbf{I} + \eta \mathbf{H}\mathbf{H}^\dagger) < \eta^r) \\ \doteq\ & P(\det(\mathbf{H}\mathbf{H}^\dagger) < \eta^{-(3-r)}) \cdot P(\sum_{n=1}^{3} \det(\mathbf{H}_n \mathbf{H}_n^\dagger) < \eta^{-(2-r)} | \det(\mathbf{H}\mathbf{H}^\dagger) < \eta^{-(3-r)}) \cdot \\ & \cdot P(\sum_{n=1}^{3}\sum_{m=1}^{3} A_{nm} < \eta^{-(1-r)} | \det(\mathbf{H}\mathbf{H}^\dagger) < \eta^{-(3-r)} \cap \sum_{n=1}^{3} \det(\mathbf{H}_n \mathbf{H}_n^\dagger) < \eta^{-(2-r)}) \end{aligned} \tag{6.60}$$

First, outage probability is easy to obtain. Because $\mathbf{H}$ is a square matrix, $\det(\mathbf{H}\mathbf{H}^\dagger) = \det(\mathbf{H}) \cdot (\det(\mathbf{H}))^*$ holds. Because $\det(\mathbf{H}\mathbf{H}^\dagger)$ follows exponential distribution approximately when $\det(\mathbf{H}\mathbf{H}^\dagger)$ is small, we have

$$P(\det(\mathbf{H}\mathbf{H}^\dagger) < \eta^{-(3-r)}) \doteq \begin{cases} \eta^{-(3-r)} & \text{if } 0 < r < 3 \\ 1 & \text{if } r > 3 \end{cases} \tag{6.61}$$

This outage probability is equivalent to a $3 \oslash 3$ system.

Using the similar logic as (6.27), the second outage probability can be equivalently written as

$$\begin{aligned} & P(\sum_{n=1}^{3} \det(\mathbf{H}_n\mathbf{H}_n^\dagger) < \eta^{-(2-r)} | \det(\mathbf{H}\mathbf{H}^\dagger) < \eta^{-(3-r)}) \\ \doteq\ & P(\det(\mathbf{H}_1\mathbf{H}_1^\dagger) < \eta^{-(2-r)} \cap \det(\mathbf{H}_2\mathbf{H}_2^\dagger) < \eta^{-(2-r)} \\ & \cap \det(\mathbf{H}_3\mathbf{H}_3^\dagger) < \eta^{-(2-r)} | \det(\mathbf{H}\mathbf{H}^\dagger) < \eta^{-(3-r)}) \end{aligned} \tag{6.62}$$

Now, we analyze this conditional event. When the event $\det(\mathbf{H}_1\mathbf{H}_1^\dagger) < \eta^{-(2-r)}$ occurs, $\vec{h}_2$ would be a linearly scaled version of $\vec{h}_1$. Similarly, when $\det(\mathbf{H}_2\mathbf{H}_2^\dagger) < \eta^{-(2-r)}$ occurs, $\vec{h}_3$ would be a linearly scaled version of $\vec{h}_1$. Therefore, if both $\det(\mathbf{H}_1\mathbf{H}_1^\dagger) < \eta^{-(2-r)}$ and $\det(\mathbf{H}_2\mathbf{H}_2^\dagger) < \eta^{-(2-r)}$ occur, $\vec{h}_2$ would be a linearly scaled version of $\vec{h}_3$, which means that $\det(\mathbf{H}_3\mathbf{H}_3^\dagger) < \eta^{-(2-r)}$ would occur for sure. Therefore, (6.62) can be simplified to:

$$P(\det(\mathbf{H}_1\mathbf{H}_1^\dagger) < \eta^{-(2-r)} \cap \det(\mathbf{H}_2\mathbf{H}_2^\dagger) < \eta^{-(2-r)} | \det(\mathbf{H}\mathbf{H}^\dagger) < \eta^{-(3-r)}) \tag{6.63}$$

Now, we try to remove the conditioning part. First, $\det(\mathbf{H}\mathbf{H}^\dagger) < \eta^{-(3-r)}$ occurs when $\vec{h}_3$ is a linear combination of $\vec{h}_1$ and $\vec{h}_2$. Under this condition, if $\mathbf{H}_1\mathbf{H}_1^\dagger < \eta^{-(2-r)}$ occurs, i.e., $\vec{h}_1$ is a linearly scaled version of $\vec{h}_2$, then $\vec{h}_3$ would be a linearly scaled version of $\vec{h}_1$, which means that $\mathbf{H}_2\mathbf{H}_2^\dagger < \eta^{-(2-r)}$ would occur for sure. Therefore, the second outage probability can be simplified to be:

$$P(\det(\mathbf{H}_1\mathbf{H}_1^\dagger) < \eta^{-(2-r)}). \tag{6.64}$$

Similar to the computation for $2 \times 3$ system, the second outage probability corresponds to:

$$P(\sum_{n=1}^{3} \det(\mathbf{H}_n\mathbf{H}_n^\dagger) < \eta^{-(2-r)} | \det(\mathbf{H}\mathbf{H}^\dagger) < \eta^{-(3-r)}) \doteq \begin{cases} \eta^{-2(2-r)} & \text{if } 0 < r < 2 \\ 1 & \text{if } r > 2 \end{cases} \tag{6.65}$$

This probability corresponds to two $2 \oslash 2$ systems.

Finally, we compute the third conditional outage probability. Condition $\det(\mathbf{H}\mathbf{H}^\dagger) < \eta^{-(3-r)}$ means that $\vec{h}_3$ is a linear combination of $\vec{h}_1$ and $\vec{h}_2$, therefore:

$$\begin{aligned} & P(\sum_{n=1}^{3}\sum_{m=1}^{3} A_{nm} < \eta^{-(1-r)} | \det(\mathbf{H}\mathbf{H}^\dagger) < \eta^{-(3-r)} \cap \sum_{n=1}^{3} \det(\mathbf{H}_n\mathbf{H}_n^\dagger) < \eta^{-(2-r)}) \\ \doteq\ & P(\sum_{n=1}^{2}\sum_{m=1}^{3} A_{nm} < \eta^{-(1-r)} | \sum_{n=1}^{3} \det(\mathbf{H}_n\mathbf{H}_n^\dagger) < \eta^{-(2-r)}) \\ \doteq\ & P(\sum_{n=1}^{2}\sum_{m=1}^{3} A_{nm} < \eta^{-(1-r)} | \det(\mathbf{H}_1\mathbf{H}_1^\dagger) < \eta^{-(2-r)}). \end{aligned} \tag{6.66}$$

The condition that $\det(\mathbf{H}_1\mathbf{H}_1^\dagger) < \eta^{-(2-r)}$ means that the columns of $\mathbf{H}_1$ are linearly scaled version of each other. Therefore,

$$\begin{aligned} & P(\sum_{n=1}^{2}\sum_{m=1}^{3} A_{nm} < \eta^{-(1-r)} | \det(\mathbf{H}_1\mathbf{H}_1^\dagger) < \eta^{-(2-r)}) \\ \doteq\ & P(A_{11} + A_{21} < \eta^{-(1-r)}) \\ \doteq\ & \begin{cases} \eta^{-2(1-r)} & \text{if } 0 < r < 1 \\ 1 & \text{if } r > 1 \end{cases} . \end{aligned} \tag{6.67}$$

This outage probability is equivalent to two $1 \times 1$ systems.

Combining the first, second and third outage probabilities, the tradeoff in $3 \times 3$ systems would be:

$$d(r) = (3-r)^+ + (2-r)^+ + (2-r)^+ + (1-r)^+ + (1-r)^+. \tag{6.68}$$

In other words, a $3 \times 3$ system is considered the combination of one $3 \oslash 3$ system, two $2 \oslash 2$ systems and two $1 \times 1$ systems.

## 6.3 The Tradeoff with Rich Scattering

In this section, we derive the tradeoff for independent $N \times M$ systems, using the outage decomposition approach. Correlated $N \times M$ systems will be studied in Section and 6.4, using the same outage decomposition approach.

First, let $\Phi(\eta) = \det(\mathbf{I} + \eta\mathbf{H}\mathbf{H}^{\dagger})$, which is a polynomial of $\eta$. The key step of outage decomposition is to find the coefficients of $\Phi(\eta)$. To write the coefficients in a simple form, we first introduce two definitions.

**Definition 6.3.1.** *Row-Wise Sub-Fading-Matrix with Order-***n***: Let the channel fading matrix* $\mathbf{H}$ *be a* $N \times M$ *(*$M \geq N$*) matrix. Denote the row vectors of* $\mathbf{H}$ *as* $\vec{h}_1$, $\vec{h}_2$, $\cdots$, $\vec{h}_N$, *i.e.,*

$$\mathbf{H} = \begin{bmatrix} \vec{h}_1 \\ \vec{h}_2 \\ \vdots \\ \vec{h}_N \end{bmatrix}, \tag{6.69}$$

*then a row-wise sub-fading-matrix with order-*$n$ *of* $\mathbf{H}$ *(*$n \leq N$*) is defined as*

$$\mathbf{H}_{n,row} = \begin{bmatrix} \vec{h}_{s_1} \\ \vec{h}_{s_2} \\ \vdots \\ \vec{h}_{s_n} \end{bmatrix}, \tag{6.70}$$

*where* $\{s_1, s_2, \cdots, s_n\}$ *is a subset of* $\{1, 2, \cdots, N\}$.

By definition, a row-wise sub-fading-matrix is an $n \times M$ matrix, whose $n$ row vectors are selected from the $N$ row vectors. Obviously one fading matrix has $\mathrm{C_N^n} = \frac{\mathrm{N!}}{\mathrm{(N-n)!n!}}$ row-wise sub-fading-matrices.

**Definition 6.3.2.** *Sub-Fading-Matrix with Order-***n***: Let the channel fading matrix* $\mathbf{H}$ *be an* $N \times M$ *(*$M \geq N$*) matrix, and*

$$\mathbf{H} = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1M} \\ h_{21} & h_{22} & \cdots & h_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ h_{N_1} & h_{N_2} & \cdots & h_{NM} \end{bmatrix}, \tag{6.71}$$

*then a sub-fading-matrix with order* $n$ *of* $\mathbf{H}$ *is defined as:*

$$\mathbf{H}_n = \begin{bmatrix} h_{s_{r1}s_{c1}} & h_{s_{r1}s_{c2}} & \cdots & h_{s_{r1}s_{cn}} \\ h_{s_{r2}s_{c1}} & h_{s_{r2}s_{c2}} & \cdots & h_{s_{r2}s_{cn}} \\ \vdots & \vdots & \ddots & \vdots \\ h_{s_{rn}s_{c1}} & h_{s_{rn}s_{c2}} & \cdots & h_{s_{rn}s_{cn}} \end{bmatrix}, \tag{6.72}$$

*where* $\{s_{r1}, s_{r2}, \cdots, s_{rn}\}$ *is a subset of* $\{1, 2, \cdots, N\}$*, and* $\{s_{c1}, s_{c2}, \cdots, s_{cn}\}$ *is a subset of* $\{1, 2, \cdots, M\}$*.*

By definition, a sub-fading-matrix with order-n is an $n \times n$ matrix. For an $N \times M$ fading matrix, there are totally $\mathrm{C}_{\mathrm{N}}^{\mathrm{n}}\mathrm{C}_{\mathrm{M}}^{\mathrm{n}}$ sub-fading-matrices.

Using the above two definitions, the polynomial $\Phi(\eta)$ can be written in an elegant form.

**Theorem 6.3.3.** *Let fading matrix* $\mathbf{H}$ *be an* $N \times M$ *(*$N \leq M$*) matrix. Define* $\Phi(\eta) = \det(\mathbf{I} + \eta\mathbf{H}\mathbf{H}^{\dagger})$*, then* $\Phi(\eta)$ *can always be written as a polynomial of* $\eta$*, where the maximal power is* $N$*, i.e.,*

$$\Phi(\eta) = c_N \cdot \eta^N + c_{N-1} \cdot \eta^{N-1} + \cdots + c_1 \cdot \eta + 1 \tag{6.73}$$

*where the coefficients* $c_n$ $n \in 1, 2, \cdots N$ *corresponds to*

$$c_n = \sum_S \det(\mathbf{H}_{n,row}\mathbf{H}_{n,row}^{\dagger}). \tag{6.74}$$

*Here, the summation over* $S$ *denotes the summation over all possible* $\mathbf{H}_{n,row}$*, i.e., there are totally* $\mathrm{C}_{\mathrm{N}}^{\mathrm{n}}$ *elements in the summation.*

*Proof.* See Appendix 6.B. □

Combining the above theorem and the Cauchy-Binet formula, we have an important corollary.

**Corollary 6.3.4.** *The coefficients of* $\eta^n$ *corresponds to the summation of all possible determinants of sub-fading-matrix with order* $n$*, i.e.,*

$$c_n = \sum_S \det(\mathbf{H}_n\mathbf{H}_n^{\dagger}), \quad n = 1, 2, \cdots N \tag{6.75}$$

*where the summation over* $S$ *denotes the summation over all possible* $\mathbf{H}_n$*. Therefore, there are totally* $\mathrm{C}_{\mathrm{N}}^{\mathrm{n}}\mathrm{C}_{\mathrm{M}}^{\mathrm{n}}$ *elements in the summation.*

*Proof.* The Cauchy-Binet formula considers a non-square matrix $\mathbf{H}_{n,row}$, which is a $n \times M$ matrix ($n < M$). For such a matrix, we have

$$\det(\mathbf{H}_{n,row}\mathbf{H}_{n,row}^{\dagger}) = \sum_{S} \det(\mathbf{H}_{n}\mathbf{H}_{n}^{\dagger}). \tag{6.76}$$

Here,

$$\mathbf{H}_{n} = \left[\vec{h}_{s_{c1}}\vec{h}_{s_{c2}}\cdots\vec{h}_{s_{cn}}\right], \tag{6.77}$$

where $\vec{h}_{s_{c1}}$ denotes the $s_{c1}^{th}$ column of $\mathbf{H}_{n,row}$, and $s_{c1}, s_{c2}, \cdots s_{cn}$ is a subset of $1, 2, \cdots, M$. $\sum_S$ denotes the summation over all possible subsets of $1, 2, \cdots, M$.

Combining the Cauchy-Binet formula and Theorem 6.3.3, corollary 6.3.4 is proved.

□

For future notation simplicity, we also define $c_0 = 1$.

Writing the coefficients in an explicit form is a critical step in our outage decomposition. As we have shown in (6.16), in high SNR regimes, the outage probability corresponds to:

$$\begin{aligned}
&P(\log\Phi(\eta) < r\log\eta) \\
&= P(\Phi(\eta) < \eta^{r}) \\
&= P(\sum_{n=0}^{N} c_n\eta^{n} < \eta^{r}) \\
&= P(c_N\cdot\eta^{N} < \eta^{r} \cap c_{N-1}\cdot\eta^{N-1} < \eta^{r} \cap \cdots c_1\cdot\eta < \eta^{r}) \\
&= P(c_N < \eta^{-(N-r)})P(c_{N-1} < \eta^{-(N-1-r)} | c_N < \eta^{-(N-r)})\cdots \\
&\quad \cdot P(c_1 < \eta^{-(1-r)} | c_N < \eta^{-(N-r)} \cap c_{N-1} < \eta^{-(N-1-r)}\cdots c_2 < \eta^{-(2-r)})
\end{aligned} \tag{6.78}$$

Therefore, for an $N \times M$ system, its outage probability can be decomposed into the production of $N$ (conditional) outage probabilities.

Next, we show that the first outage probability corresponds to $M-N+1$ $N \oslash N$ systems, the second (conditional) outage probability corresponds to two $(N-1) \oslash$

$(N-1)$ systems, the third (conditional) outage probability corresponds to two $(N-2)\oslash N-2$ systems, and the last outage probability corresponds to two $1\oslash 1$ systems.

The first outage probability can be analyzed as following. Using Corollary 6.3.4, it is clear that

$$c_N = \det(\mathbf{H}\mathbf{H}^\dagger) = \sum_S \det(\mathbf{H}_N \mathbf{H}_N^\dagger) \tag{6.79}$$

where $\mathbf{H}_N$ denotes a sub-fading-matrix with order $N$. There are totally $\mathrm{C}_{\mathrm{M}}^{\mathrm{N}}$ elements in the summation. Because all of the elements in the summation are positive, the first outage probability can be written as:

$$P\left(\sum_S \det(\mathbf{H}_N \mathbf{H}_N^\dagger) < \eta^{-(N-r)}\right) \doteq P\left(\bigcap_S \left(\det(\mathbf{H}_N \mathbf{H}_N^\dagger) < \eta^{-(N-r)}\right)\right) \tag{6.80}$$

where $\bigcap(\cdot)$ denotes the union of all events in the bracket. Equation (6.80) states that the total fading matrix in outage is equivalent to all its sub-fading-matrices (with order $N$) are in outage.

We notice that in general the events that sub-fading-matrices are in outage are not jointly independent, as we have shown for the $2\times 3$ case. However, a finer observation shows that selecting $M-N+1$ sub-fading-matrices to be in outage would force all $\mathrm{C}_{\mathrm{M}}^{\mathrm{N}}$ sub-fading-matrices to be in outage. For example, if $\mathbf{H} = [\vec{h}_1, \vec{h}_2, \cdots, \vec{h}_M]$, we can define:

$$\begin{aligned}
&\mathbf{H}_1 = [\vec{h}_1, \vec{h}_2, \cdots, \vec{h}_N] \\
&\mathbf{H}_2 = [\vec{h}_2, \vec{h}_3, \cdots, \vec{h}_{N+1}] \\
&\quad\vdots \qquad\qquad (6.81) \\
&\mathbf{H}_{M-N+1} = [\vec{h}_{M-N+1}, \vec{h}_{M-N+2}, \cdots, \vec{h}_M] \qquad (6.82)
\end{aligned}$$

If $\mathbf{H}_1$ is in outage, then $\vec{h}_1$ is a linear combination of $\vec{h}_2, \cdots \vec{h}_N$. Similarly, if $\mathbf{H}_2$ is in outage, then $\vec{h}_{N+1}$ would be a linear combination of $\vec{h}_2, \cdots \vec{h}_N$. If both $\mathbf{H}_1$ and $\mathbf{H}_2$ are in outage, then $\vec{h}_2, \cdots \vec{h}_N$ would form the basis of the matrix $[\vec{h}_1\, \vec{h}_2, \cdots, \vec{h}_N, \vec{h}_{N+1}]$,

which means this matrix is in outage. Next, in addition to $\mathbf{H}_1$ and $\mathbf{H}_2$, if $\mathbf{H}_3$ is in outage, then $\vec{h}_{N+2}$ would be a linear combination of $\vec{h}_3 \cdots \vec{h}_{N+1}$. Then, $\vec{h}_3 \cdots \vec{h}_{N+1}$ would span the matrix $[\vec{h}_1\, \vec{h}_2, \cdots, \vec{h}_{N+1}, \vec{h}_{N+2}]$, which means this matrix is in outage as well. This process can be repeated $M-N+1$ times to cover the whole fading matrix $[\vec{h}_1, \vec{h}_2, \cdots, \vec{h}_M]$. Therefore, the probability that all sub-fading-matrices are in outage equals to the probability that $\mathbf{H}_1, \mathbf{H}_2, \cdots, \mathbf{H}_{M-N+1}$ are in outage, i.e.,

$$P\left(\bigcap_{S}\left(\det(\mathbf{H}_N\mathbf{H}_N^{\dagger}) < \eta^{-(N-r)}\right)\right) = P\left(\bigcap_{n=1,2,\cdots M-N+1}\left(\det(\mathbf{H}_n\mathbf{H}_n^{\dagger}) < \eta^{-(N-r)}\right)\right). \tag{6.83}$$

Moreover, it is noted that the above $(M-N+1)$ events are independent. For example, $\mathbf{H}_1$ in outage denotes if $\vec{h}_1$ is a linear combination of $\vec{h}_2, \cdots, \vec{h}_N$, while $\mathbf{H}_2$ in outage denotes if $\vec{h}_{N+1}$ is a linear combination of $\vec{h}_2, \cdots, \vec{h}_N$. Because $\vec{h}_1$ is independent of $\vec{h}_{N+1}$, the two outage events are independent. Similarly, we can show that the third outage events is independent of both the first and second outage events, and this process can be repeated multiple times. Therefore, in (6.83):

$$P\left(\bigcap_{n=1,2,\cdots M-N+1}\left(\det(\mathbf{H}_n\mathbf{H}_n^{\dagger}) < \eta^{-(N-r)}\right)\right) = \prod_{n=1}^{M-N+1} P\left(\det(\mathbf{H}_n\mathbf{H}_n^{\dagger}) < \eta^{-(N-r)}\right) \tag{6.84}$$

Because for a square matrix $\mathbf{H}$, $P\left(\det(\mathbf{H}\mathbf{H}^{\dagger}) < \eta^{-(N-r)}\right) \doteq \eta^{-(N-r)}$, we finally have:

$$P(c_N < \eta^{-(N-r)}) \doteq \left(\eta^{-(N-r)}\right)^{(M-N+1)}, \tag{6.85}$$

and the correspond tradeoff is:

$$d_1(r) = \underbrace{(N-r)^+ + (N-r)^+ + \cdots + (N-r)^+}_{\text{totally } (M-N+1)}, \tag{6.86}$$

which means that the first outage probability corresponds to $(M-N+1)$ $N \oslash N$ systems.

In the next step, conditional outage probabilities are computed. Now, the question is: given all sub-fading-matrices with order greater than $n$ in outage, how many sub-fading-matrices with order $n$ in outage would force all sub-fading-matrices with order $n$ in outage? The answer is two. Subsequently, those conditional outage probabilities corresponds to two $(N-1) \oslash (N-1)$, two $(N-2) \oslash (N-2)$ systems, etc. Detailed derivations follow.

Here, we study the case of sub-fading-matrices with order $N-1$. Let $\mathbf{H}_{N-1,a}$ and $\mathbf{H}_{N-1,b}$ are two distinct sub-fading-matrices of $\mathbf{H}$ with order $N-1$, and suppose them to be in outage. Let the rows of $\mathbf{H}_{N-1,a}$ are the $s_1, s_2, \cdots, s_{N-1}$ row of $\mathbf{H}$. Given $\mathbf{H}$ in outage and using similar techniques for sub-fading-matrices with order $N$, it is shown that all sub-fading-matrices with the same rows as $\mathbf{H}_{N-1,a}$ are in outage. Similarly, the sub-fading-matrices with the same columns as $\mathbf{H}_{N-1,b}$ would be in outage. Now, select an arbitrary sub-fading-matrix with order $(N-1)$, $\mathbf{H}_{N-1}$, we can always find a sub-fading-matrix that has the same rows of $\mathbf{H}_{n-1}$ in outage, and another sub-fading-matrix that has the columns of $\mathbf{H}_{n-1}$. This means that $\mathbf{H}_{N-1}$ would be in outage for sure. Hence, given sub-fading-matrices with order greater than $n$ in outage, just two sub-fading-matrices with order $N-1$ would force all sub-fading-matrices with order $N-1$ in outage. Therefore,

$$
\begin{aligned}
&P(c_{N-1} < \eta^{-(N-1-r)} | c_N < \eta^{-(N-r)}) \doteq \begin{cases} 2 \cdot \eta^{-(N-1-r)} & \text{if } 0 < r < N-1 \\ 1 & \text{if } r > N-1 \end{cases} \\
&\qquad \vdots \\
&P(c_1 < \eta^{-(1-r)} | c_N < \eta^{-(N-r)} \cap c_{N-1} < \eta^{-(N-1-r)} \cdots c_2 < \eta^{-(2-r)}) \\
&\doteq \begin{cases} 2 \cdot \eta^{-(1-r)} & \text{if } 0 < r < 1 \\ 1 & \text{if } r > 1 \end{cases}
\end{aligned} \tag{6.87}
$$

Combining the above outage probabilities, the final tradeoff in rich scattering environments is:

$$
d(r) = \sum_{n=1}^{M-N+1} (N-r)^+ + 2 \cdot \sum_{n=1}^{N-1} (n-r)^+ \tag{6.88}
$$

which matches the original expression in Zheng's paper [Zheng and Tse(2003)].

## 6.4 The Tradeoff with Poor Scattering

Because UIU model is the most comprehensive model among the three proposed models, we prefer to study the tradeoff for UIU model. Because channel capacity is invariant with respect to orthogonal transformation, it is found that:

$$\begin{aligned}\mathcal{C} &= \log\det\left(\mathbf{I}+\frac{\eta}{M}\mathbf{H}\mathbf{H}^{\dagger}\right)\\ &= \log\det(\mathbf{I}+\frac{\eta}{M}\mathbf{H}_v\mathbf{H}_v^{\dagger}),\end{aligned} \tag{6.89}$$

where $\mathbf{H}_v$ is defined as (see Chapter 5)

$$\mathbf{H}_v = \mathbf{\Omega}\odot\mathbf{H}_{iid}. \tag{6.90}$$

Therefore, the tradeoff is fully determined by the virtual variance matrix $\mathbf{\Omega}$, i.e., the tradeoff is a function of $\mathbf{\Omega}$.

Next, we present a theorem that states the magnitudes of elements of $\mathbf{\Omega}$ does not alter the tradeoff.

**Theorem 6.4.1.** *Denoting the $(i,j)^{th}$ element of $\mathbf{\Omega}$ as $\omega_{i,j}$, and defining*

$$\omega'_{i,j} = 1_{>0}\left(\omega_{i,j}\right), \tag{6.91}$$

*another matrix $\mathbf{\Omega}'$ can be form with $\omega'_{i,j}$ as its elements. Here, $1_{>0}(\cdot)$ denotes an indicator function, i.e., its value would be one if its argument is greater than zero, otherwise its value would be zero. It is found that the tradeoff for $\mathbf{\Omega}'$ and $\mathbf{\Omega}$ is the same.*

The theorem states that, for computing the tradeoff, the elements of $\mathbf{\Omega}$ can be safely replaced by one, unless the value of the element is zero. For example, a $3\times 3$ system with

$$\mathbf{\Omega} = \begin{bmatrix} 0.1 & 3 & 0.1\\ 0 & 8 & 2\\ 5 & 0.01 & 0\end{bmatrix} \tag{6.92}$$

has the same tradeoff as a system with

$$\mathbf{\Omega}' = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \tag{6.93}$$

while the latter is much easier to analyze.

*Proof.* First, we define

$$\omega_{i,j}^{max} = \max(\mathbf{\Omega}) \cdot 1_{>0}\left(\omega_{i,j}\right), \quad \text{and} \quad \omega_{i,j}^{min} = \min(\mathbf{\Omega}) \cdot 1_{>0}\left(\omega_{i,j}\right), \tag{6.94}$$

where $\mathbf{\Omega}^{\max}$ corresponds to the largest value of $\mathbf{\Omega}$, while $\mathbf{\Omega}^{\min}$ is the smallest non-zero value of $\mathbf{\Omega}$. Matrices $\mathbf{\Omega}^{max}$ and $\mathbf{\Omega}^{min}$ are formed by $\omega_{i,j}^{max}$ and $\omega_{i,j}^{min}$, respectively.

For example, for the variance matrix in (6.92), correspondingly

$$\mathbf{\Omega}^{max} = \begin{bmatrix} 8 & 8 & 8 \\ 0 & 8 & 8 \\ 8 & 8 & 0 \end{bmatrix}, \quad \text{and} \quad \mathbf{\Omega}^{min} = \begin{bmatrix} 0.01 & 0.01 & 0.01 \\ 0 & 0.01 & 0.01 \\ 0.01 & 0.01 & 0 \end{bmatrix}. \tag{6.95}$$

Clearly the capacity for the $\mathbf{\Omega}$, $\mathbf{\Omega}^{max}$ and $\mathbf{\Omega}^{min}$ has the following relationship:

$$\begin{aligned} \mathcal{C}^{min} &= \log\det\left(\mathbf{I} + \frac{\eta}{M}(\mathbf{\Omega}^{min} \odot \mathbf{H}_{iid})(\mathbf{\Omega}^{min} \odot \mathbf{H}_{iid})^{\dagger}\right) \\ &< \mathcal{C} = \log\det\left(\mathbf{I} + \frac{\eta}{M}(\mathbf{\Omega} \odot \mathbf{H}_{iid})(\mathbf{\Omega} \odot \mathbf{H}_{iid})^{\dagger}\right) \\ &< \mathcal{C}^{max} = \log\det\left(\mathbf{I} + \frac{\eta}{M}(\mathbf{\Omega}^{max} \odot \mathbf{H}_{iid})(\mathbf{\Omega}^{max} \odot \mathbf{H}_{iid})^{\dagger}\right) \end{aligned} \tag{6.96}$$

However, note that the capacity $\mathcal{C}^{min}$ is the same as $\mathcal{C}^{max}$ in terms of computing the tradeoff (i.e., when $\eta \to \infty$), because the difference corresponds to a finite reduction in SNR(the reduction ratio is $\left(\frac{\mathbf{\Omega}^{\min}}{\mathbf{\Omega}^{\max}}\right)^2$). Therefore, the matrix $\mathbf{\Omega}'$ defined in (6.91) is the same as $\mathbf{\Omega}$ for computing the diversity-multiplexing tradeoff. □

Through the above discussion, it is clear that correlated MIMO fading matrix is equivalent to an independent matrix with some arbitrary elements set to be zero. For KPF, because the fading matrix is separable (see Chapter 5), the tradeoff is the same as independent MIMO: the only difference is that the number of antenna

elements $(M, N)$ has to be replaced by the rank of the receive and transmit matrix $\Gamma_M, \Gamma_N$, respectively.

The only left non-trivial tradeoff problem is for UIU models. Unfortunately, although the outage decomposition approach is still applicable, the linearity argument used in rich scattering is not applicable in poor scattering environments. For example, given a virtual variance matrix

$$\Omega = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \tag{6.97}$$

the determinants of two sub-fading-matrices with order 2 is zero (in outage), however, the determinant of the third sub-fading-matrix is not necessarily zero.

The reason is that, if all elements are complex Gaussian RVs, the vectors in a sub-fading-matrix span the whole $\mathbb{R}^{2\times 2}$ space (almost surely). However, for poor scattering environments, the vectors in a sub-fading-matrix does not necessarily span the whole $\mathbb{R}^{2\times 2}$ space. Therefore, the third column vector may not be expressed as a linear combination of the other two column vectors. And all following discussions can not be continued.

Fortunately, the outage decomposition is still applicable. Then, it is still possible to compute the tradeoff for some special cases. The first non-trivial (i.e., the case that can not be simplified to SISO, SIMO, and parallel SISP) $2 \times 2$ system in poor scattering environments possesses the virtual variance matrix

$$\mathbf{\Omega} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}. \tag{6.98}$$

The corresponding polynomial is (similar to (5))

$$\Phi(\eta) = \eta^2 \cdot A_{11}A_{22} + \eta \cdot (A_{11} + A_{22} + A_{12}) + 1. \tag{6.99}$$

Recall that $A_{i,j}$ corresponds to the square of the amplitude of corresponding fadings $h_{i,j}$, $i \in 1, 2, \cdots, N$ and $j \in 1, 2, \cdots, M$, then the outage probability is

$$P(\Phi(\eta) < \eta^r) \doteq P(A_{11}A_{22} < \eta^{-(2-r)}) \cdot P((A_{11} + A_{22} + A_{12}) < \eta^{-(1-r)} | A_{11}A_{22} < \eta^{-(2-r)}) \tag{6.100}$$

The first outage probability corresponds to $d_1(r) = (2-r)^+$. Note that if $A_{11}A_{22} < \eta^{-(2-r)}$, then $A_{11}$ and $A_{22}$ are always smaller than $\eta^{-(1-r)}$. Therefore, the second outage probability corresponds to $d_2(r) = (1-r)^+$. The final tradeoff for this $2 \times 2$ system is:

$$d(r) = (2-r)^+ + (1-r)^+. \tag{6.101}$$

which means that it corresponds to a combination of one $2 \oslash 2$ system and one $1 \oslash 1$ system.

For $2 \times 3$ systems, because switch rows and columns of virtual fading matrix does not alter the tradeoff (based on Corollary 6.3.4), there are totally four non-trivial poor-scattering cases:

$$\mathbf{\Omega}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}, \mathbf{\Omega}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}, \mathbf{\Omega}_3 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}, \text{ and } \mathbf{\Omega}_4 = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \tag{6.102}$$

The following computations are similar to (6.36).

For $\mathbf{\Omega}_1$, the outage probability can be decomposed to:

$$\begin{aligned} P(\Phi(\eta) < \eta^r) \doteq & P(A_{11}A_{22} + A_{11}A_{23} < \eta^{-(2-r)}) \\ & \cdot P(A_{11} + A_{22} + A_{23} < \eta^{-(1-r)} | A_{11}A_{22} + A_{11}A_{23} < \eta^{-(2-r)}) \end{aligned} \tag{6.103}$$

Note that if $A_{11}A_{22} < \eta^{-(2-r)}$, then $A_{11} < \eta^{-(2-r)}$, and then $A_{11}A_{23} < \eta^{-(2-r)}$ almost surely. Therefore, the first outage probability corresponds to $d_1(r) = (2-r)^+$. Moreover, if $A_{11}A_{22} < \eta^{-(2-r)}$, then $A_{11} < \eta^{-(1-r)}$ and $A_{22} < \eta^{-(1-r)}$ for sure, therefore, the second outage probability corresponds to:

$$d(r) = (2-r)^+ + (1-r)^+ \tag{6.104}$$

For $\mathbf{\Omega}_2$, the outage probability can be decomposed to:

$$P(\Phi(\eta) < \eta^r) \doteq P(A_{11}A_{22} + A_{11}A_{23} < \eta^{-(2-r)})$$

$$\cdot P(A_{11}+A_{21}+A_{22}+A_{23}<\eta^{-(1-r)}|A_{11}A_{22}+A_{11}A_{23}<\eta^{-(2-r)}) \tag{6.105}$$

The same as $\mathbf{\Omega}_1$, the first outage probability corresponds to $d_1(r)=(2-r)^+$. Because there is one more fading in the second outage probability, the second outage probability would correspond to $(1-r)^+ + (1-r)^+$. The tradeoff for $\mathbf{\Omega}_2$ is

$$d(r)=(2-r)^+ + (1-r)^+ + (1-r)^+ \tag{6.106}$$

For $\mathbf{\Omega}_3$, the outage probability can be decomposed to:

$$\begin{aligned}P(\Phi(\eta)<\eta^r) \doteq & P(A_{11}A_{22}+A_{12}A_{23}+A_{11}A_{23}<\eta^{-(2-r)}) \\ & \cdot P(A_{11}+A_{12}+A_{22}+A_{23}<\eta^{-(1-r)}|A_{11}A_{22}+A_{12}A_{23}+A_{11}A_{23}<\eta^{-(2-r)})\end{aligned} \tag{6.107}$$

First, if $A_{11}A_{22}<\eta^{-(2-r)}$ and $A_{12}A_{23}<\eta^{-(2-r)}$, then $A_{11}<\eta^{-(1-r)}$ and $A_{23}<\eta^{-(1-r)}$, finally $A_{11}A_{23}<\eta^{-(2-r)}$. Therefore, $d_1 r=(2-r)^+ + (2-r)^+$. Moreover, if $A_{11}A_{22}<\eta^{-(2-r)}$ and $A_{12}A_{23}<\eta^{-(2-r)}$, then $A_{11}$, $A_{12}$, $A_{22}$, and $A_{23}$ are all smaller than $\eta^{-(1-r)}$. Therefore, the second outage probability is equivalent to 1 in terms of measuring diversity gain. The tradeoff for $\mathbf{\Omega}_3$ is

$$d(r)=(2-r)^+ + (2-r)^+. \tag{6.108}$$

For $\mathbf{\Omega}_4$, the outage probability can be decomposed to:

$$\begin{aligned}P(\Phi(\eta)<\eta^r) \doteq & P(A_{11}A_{22}+A_{11}A_{23}+\det(\mathbf{H}_3\mathbf{H}_3^\dagger)<\eta^{-(2-r)}) \\ & \cdot P(A_{11}+A_{12}+A_{13}+A_{22}+A_{23}<\eta^{-(1-r)}|A_{11}A_{22}+A_{11}A_{23}+\det(\mathbf{H}_3\mathbf{H}_3^\dagger)<\eta^{-(2-r)}),\end{aligned} \tag{6.109}$$

where

$$\mathbf{H}_3=\begin{bmatrix} h_{12} & h_{13} \\ h_{22} & h_{23} \end{bmatrix} \tag{6.110}$$

Note, if $A_{11}A_{22}<\eta^{-(2-r)}$ and $A_{11}A_{23}<\eta^{-(2-r)}$, then $A_{22}<\eta^{-(2-r)}$ and $A_{22}<\eta^{-(2-r)}$ (i.e., the probability that $c_{N-1}<\eta^{-(N-1-r)}$ approaches to $\eta^0$ when $\eta\to\infty$), then

$\det(\mathbf{H}_3\mathbf{H}_3^\dagger) < \eta^{-(2-r)}$ for sure. Therefore, the first outage probability corresponds to the tradeoff $(2-r)^+ + (2-r)^+$. Moreover, if $A_{11}A_{22} < \eta^{-(2-r)}$ and $A_{11}A_{23} < \eta^{-(2-r)}$, the only fading that is not necessarily in outage is $h_{21}$. In this case, the second outage probability corresponds to the tradeoff $(1-r)^+$. The final tradeoff for $\mathbf{\Omega}_4$ is

$$d(r) = (2-r)^+(2-r)^+ + (1-r)^+. \tag{6.111}$$

There is a number of ways to decompose a poor scattering MIMO channel into multiple parallel SISO channels. Based on the above derivations, we found that the tradeoff for a small number of antennas are the best tradeoff for the decompositions. It is conjectured that it is also true for a large number of antennas, although a manual computation is too complicated to be realized.

## 6.5 Conclusions

In this chapter, an alternative approach for computing the diversity-multiplexing tradeoff in time-invariant MIMO channels is proposed. The outage decomposition approach in this chapter does not necessarily require the precise knowledge of the joint distribution of eigen values of channel fading matrix: Hence, it can be extended to the correlated MIMO channel case, which is described by the UIU models. For a small number of antennas, the tradeoff in such environments is explicitly computed. The tradeoff with large dimensional antenna arrays are conjectured.

## 6.A Proof of Lemma 6.2.1

Let

$$\det(\mathbf{I} + \frac{\eta}{M}\mathbf{H}\mathbf{H}^\dagger) = \sum_{n=0}^{N} b_n \left(\frac{\eta}{M}\right)^n, \tag{6.112}$$

where $b_n, n \in \{1, 2, \cdots, N\}$ are functions of $\mathbf{H}$. Then, the outage probability corresponds to

$$P(\det(\mathbf{I} + \frac{\eta}{M}\mathbf{H}\mathbf{H}^\dagger) < \eta^r)$$

$$= P(b_N < M^N \cdot \eta^{-(N-r)} \cap M^{N-1} b_{N-1} < \eta^{-(N-1-r)} \cap \cdots \cap \cdot b_1 < M\eta^{-(1-r)}) \tag{6.113}$$

However, because

$$-\lim_{\eta\to\infty} \frac{\log\left(M^N\right)}{\log \eta} = 0, \tag{6.114}$$

we can safely write

$$M^N \doteq \eta^0, \tag{6.115}$$

where $\eta \to \infty$. Similarly, we can find that $M^{N-1} \doteq \eta^0$, $M^{N-2} \doteq \eta^0$, $M \doteq \eta^0$.

Substituting (6.115) to (6.113), it is shown that

$$\begin{aligned}
&P(\det(\mathbf{I} + \frac{\eta}{M}\mathbf{H}\mathbf{H}^\dagger) < \eta^r) \\
=&P(b_N < M^N \cdot \eta^{-(N-r)} \cap M^{N-1} b_{N-1} < \eta^{-(N-1-r)} \cap \cdots \cap M \cdot b_1 < \eta^{-(1-r)}) \\
\doteq&P(b_N < \eta^0 \cdot \eta^{-(N-r)} \cap b_{N-1} < \eta^0\eta^{-(N-1-r)} \cap \cdots \cap b_1 < \eta^0\eta^{-(1-r)}) \\
\doteq&P(b_N < \eta^{-(N-r)} \cap b_{N-1} < \eta^{-(N-1-r)} \cap \cdots \cap b_1 < \eta^{-(1-r)}) \\
\doteq&P(\det(\mathbf{I} + \eta\mathbf{H}\mathbf{H}^\dagger) < \eta^r)
\end{aligned} \tag{6.116}$$

i.e., Lemma 6.2.1 is proved.

## 6.B Proof of Theorem 6.3.3

The proof is mainly based on the mathematical studies conducted on some previous studies on characteristic polynomial [Pennisi(1987)]. Let $\mathbf{K}$ be an $N \times N$ matrix, and $\mathbf{I}$ be an $N \times N$ identity matrix, considering the following polynomial [2]:

$$g(t) = \det(\mathbf{K} + t\mathbf{I}) = b_N t^N + b_{N-1} t^{N-1} + b_{N-2} t^{N-2} + \cdots + b_1 t + b_0, \tag{6.117}$$

it is found that the coefficient of $t^n$, $b_n$, corresponds to the summation of "sub-determinants of $\mathbf{K}$ resulting from crossing out the rows and columns numbered $\{i_1, i_2, i_3, \cdots, i_n\}$", where $\{i_1, i_2, i_3, \cdots, i_n\}$ is a subset of $\{1, 2, 3, \cdots, N\}$.

[2] This is a variation of characteristic polynomial. For convenience, the minus sign in characteristic polynomial was intentionally changed to plus sign, in the original work.

Our interested problem is to find the coefficients of the following polynomial:

$$\begin{aligned} g(\eta) &= \det(\mathbf{I} + \eta \mathbf{H}\mathbf{H}^{\dagger}) \\ &= \eta^{N} \left( \det(\mathbf{H}\mathbf{H}^{\dagger} + \eta^{-1}\mathbf{I}) \right) \\ &= \sum_{n=0}^{N} b_n \eta^{N-n} \end{aligned} \tag{6.118}$$

where $b_n$ denotes the summation of "subdeterminants of $\mathbf{K}$ resulting from crossing out the rows and columns numbered $\{i_1, i_2, i_3, \cdots, i_n\}$", where $\{i_1, i_2, i_3, \cdots, i_n\}$ is a subset of $\{1, 2, 3, \cdots, n\}$. The first equality exploits the basic properties of determinant, and the second equality is based on the results in (6.117).

Complimentarily, for the coefficient of $\eta^n$, its coefficient is the summation of "subdeterminants of $\mathbf{H}\mathbf{H}^{\dagger}$ resulting from choosing the rows and columns numbered $\{i_1, i_2, i_3, \cdots, i_n\}$", where $\{i_1, i_2, i_3, \cdots, i_n\}$ is a subset of $\{1, 2, 3, \cdots, n\}$. For convenience, let $\mathbf{H}\mathbf{H}^{\dagger} = \mathbf{\Xi}$, then the coefficient of $\eta^n$, is the summation of the determinants of:

$$\begin{bmatrix} \xi_{i_1,i_1} & \xi_{i_1,i_2} & \cdots & \xi_{i_1,i_n} \\ \xi_{i_2,i_1} & \xi_{i_2,i_2} & \cdots & \xi_{i_2,i_n} \\ \vdots & \vdots & \ddots & \vdots \\ \xi_{i_n,i_1} & \xi_{i_n,i_2} & \cdots & \xi_{i_n,i_n} \end{bmatrix}, \tag{6.119}$$

where $\xi_{i,j}$ denotes the $\{i,j\}$th element of $\mathbf{\Xi}$.

Let the row vectors of $\mathbf{H}$ be $\vec{h}_1\, \vec{h}_2 \cdots \vec{h}_n$, it can be found that

$$\xi_{i,j} = \vec{h}_i \vec{h}_j^{\dagger}. \tag{6.120}$$

Combining (6.119) and (6.120), it is finally shown that the coefficient of $\eta^n$ corresponds to the summation of determinants of all possible row-wise-sub-fading-matrix, which is defined in (6.70), i.e., Theorem 6.3.3 is proved.

# Chapter 7

# Conclusions and Future Works

This dissertation studies SISO, SIMO, and MIMO channels in a successive manner. Both information theoretic and practical algorithm development aspects have been investigated. For the former, a joint error rate and data rate analysis, namely diversity-multiplexing tradeoff, is studied under certain practical considerations, such as temporal correlations and spatial correlations. For the latter, system error performances under array perturbation or non-stationarity effects have been analyzed and simulated.

## 7.1 Conclusions

After giving motivations in Chapter 1, Chapters 2-6 apply multi-antenna techniques in wireless comminations systems performance analysis with multi-antenna techniques cumulatively. Chapter 2 considers SISO, i.e., single antenna system, and investigates the diversity-multiplexing tradeoff with temporal correlations. Chapter 3 and 4 apply multi-antenna at the receiver side (SIMO), under practical considerations such as perturbations and non-stationarity. Chapter 5 and 6 deploy multi-antenna at both transmitter and receiver side, that is, MIMO channels. A two dimensional correlation structure analysis and the relevant error rate - data rate analysis have been presented in the two chapters. Impacts of each chapter are listed as follows.

**Chapter 2** has inspired a novel interpretation of the diversity-multiplexing tradeoff: spike sharpening effect. Because lower correlations correspond to sharper spikes (less randomness in the mutual information between channel input and output), given the same multiplexing gain, better diversity is achievable under smaller correlations. Both qualitative and quantitative results have been presented. This result supports the process of examining the optimality of a designed channel coding scheme: if a channel coding scheme can achieve this tradeoff, it is optimal. Otherwise, it is not optimal on the information theoretical sense.

**Chapter 3** opens the discussions on multi-antennas. Specifically, this chapter considers mutli-antenna at receiver side, and spatial fadings are fully correlated. In this environment, the so called "beamforming" techniques are adopted. The conventional Fourier beamforming performance under array perturbations is theoretically analyzed and numerically simulated. The perturbations are generated by the imperfect knowledge of array manifolds, which happens frequently in real applications. Therefore, the result in this chapter helps wireless engineers to predict the performance loss in real systems, due to certain imperfect prior knowledge of the system parameters.

**Chapter 4** extends the discussions in Chapter 3. Conventional Fourier beamforming is replaced by a statistically optimal beamforming, that is, linear constrained minimum variance (LCMV) beamforming. Although generally LCMV beamforming performs better than conventional beamforming in terms of interference suppression, in certain environment this is not the case. One example is the system that transmits signals non-continuously, for example, wireless local positioning system. Due to its non-stationary nature, standard sample covariance estimator does not work properly in this system. Thus, directly applying LCMV beamforming in such environments is not feasible. An alternative covariance estimator, which exploits

the cyclostationarity property, is proposed to counter the non-stationarity problem. Summarizing, this chapter introduces a novel solution for wireless system designers, when non-stationarity presents in their systems.

**Chapter 5** starts incorporating multi-antenna at both transmitter and receiver side. This chapter considers MIMO channel modeling. Many channel models have been proposed to describe correlated MIMO channels. We first propose a general framework that can accommodate previously proposed channel models. Then, it is shown that each of previously introduced channel model possesses different eigen structures. Finally, via a ray tracing approach, it is verified that different eigen structures correspond to different physical environments. Thus, the previously proposed channel models are feasible in specific environments, such as microcell or macrocell environments. The results in this chapter help designers to select a proper channel model based on their applications, and optimize their designs correspondingly.

**Chapter 6** performs the joint data rate - error rate analysis in correlated MIMO channels. An innovative outage decomposition approach is introduced in this chapter. It is shown that, in high SNR regimes, the outage probability in MIMO channels can be decomposed into the production of multiple conditional outage probabilities, where each conditional outage probability corresponds to a parallel SISO channel. Because diversity corresponds to the logarithm of outage probability, diversity in MIMO channels would be the summation of the diversity in multiple parallel SISO channels. Based on the results in Chapter 5, this approach is applicable for even correlated MIMO channels. Then, incorporating the tradeoff results in Chapter 2, the final tradeoff in correlated MIMO channels can be established, especially for a small number of antennas.

## 7.2 Future Works

Based on the presented works in this dissertation, several meaningful problems can launched in the near future. Moreover, multiple long term extensions are proposed

at the end of this chapter.

The first is the elaborations of MIMO channel modeling results.

- *Two Dimensional Auto-Regressive Moving Average (ARMA) channel modeling:* Two dimensional spectral analysis has been preliminarily studied in this dissertation. Moreover, it would be interesting to know if the two dimensional spectral can be described by a two dimensional FIR/IIR filter. This would help people to transfer results in filter design and spectral analysis MIMO channel modeling areas.

- *Non-Separable Time-Space Variation Considerations:* Temporal correlations and spatial correlations are considered as sealable in current works. In fact, they are not separable practically. Only certain preliminary works tried to model joint spatially-temporally correlated channels, in which frequently transmitter and receiver correlation are assumed separable. A three dimensional model that can depict joint temporal, transmitter and receiver correlations should considered. Such a model can help the design of future Wi-Fi/WiMAX systems.

- *Multi-user MIMO channel modeling"* In this dissertation, only intra-user correlations are considered. With the development of current MIMO systems, multi-user MIMO systems are emerging. Subsequently, inter-user correlations have to dealt with. It would be interesting to know how two users are correlated when they share the same local scatterers.

The second is the variations of the joint error rate - data rate analysis in correlated MIMO channels.

- *Optimal transmit covariance matrix design:* In the present work, neither channel state information nor channel statistical information is available at the transmitter. If just channel statistical information is available at the transmitter, it is shown that a "waterfilling-like" scheme can optimize the data rate. However, minimal data rate - error rate analysis has been conducted for this case. The solution to this problem would certainly help the system optimization when only limited feedback bandwidth is available.

- *Correlated space-time coding:.* Current coding approach in space-time correlated channels is a two-step approach: first designing the code for independent channels, then applying a precoding scheme that takes correlation into consideration. It would be interesting to know if a joint design can outperform the two-step approach, i.e., the code is designed for correlated fading directly, but not through the two-steps approach.

Finally, the works in this dissertation can be extended to multi-user MIMO and mobile ad-hoc networks (MANET) MIMO. The proposed ideas here should be considered as long term goals, due to their complexities and mathematical difficulties.

- Single user MIMO and multi-user SIMO are considered in this dissertation. It is necessary to extend the results to multi-user MIMO channels. Only preliminary results (such as two users) are available in this area, because this is a (at least) three dimensional problem. Henceforth, an extensive usage of Kronecker product is inevitable in this study, which is not very familiar for many engineers. Only partial results might be obtained in correlated multi-user MIMO channels. For example, for a two users and one base scenario, different clusters of scatters are located around each user and the bast station. In this case, the capacity region might be obtainable.

- MANET MIMO cases are even more complicated than multi-user cases, especially when multi-hop is allowed. Nonetheless, it is still important to study this case, because MANET studies is turning to be popular in the foreseeable future. Currently, MANET routing schemes are designed based on topology, while physical scatterers are omitted. It would be interesting to know if it is possible to design a routing scheme based on the knowledge of physical environments. Based on previous studies on simple MIMO channels, it is believed that even a small amount of channel information would improve network throughput significantly.

# Bibliography

[Abdi et al.(2002)] Abdi, A., J. A. Barger, and M. Kaveh, 2002: A parametric model for the distribution of the angle of arrival and the associated correlation function and power spectrum at the mobile station. *IEEE Transaction on Vehicular Technologies*, **51**, 425–434.

[Abdi and Kaveh(2002)] Abdi, A., and M. Kaveh, 2002: A space-time correlation model for multielement antenna systems in mobile fading channels. *IEEE J. Sel. Areas Commun.*, **20**(3), 550 – 560.

[Al-Ardi et al.(2004)] Al-Ardi, E., R. Shubair, and M. Al-Mualla, 2004: Computationally efficient high-resolution doa estimation in multipath environment. *IEEE Electronics Letters*, **40**(14), 908–910.

[Alamouti(1999)] Alamouti, S., 1999: A simple transmit diversity technique for wireless communications. *IEEE Journal on Selected Areas in Communications*, **16**(8).

[Almers et al.(2007)] Almers, P., E. Bonek, A. Burr, N. Czink, M. Debbah, V. Degli-Esposti, H. Hofstetter, P. Kyosti, D. Laurenson, G. Matz, A. F. Molisch, C. Oestges, and H. Ozcelik, 2007: Survey of channel and radio propagation models for wireless MIMO systems. *EURASIP Journal on Wireless Communications and Networking*, **2007**, 1–19.

[Arowojolu et al.(1994)] Arowojolu, A. A., A. M. D. Turkman, and J. D. Parsons, 1994: Time dispersion in urban microcellular enviroments. in *Proceedings IEEE 44th Vehicular Technical Conference*, Vol. 1, pp. 150–154.

[Baddour and Beaulieu(2005)] Baddour, K., and N. Beaulieu, 2005: Autoregressive modeling for fading channel simulation. *IEEE Transactions on Wireless Communications*, **4**(4), 1650– 1662.

[Barton and Fuhrmann(1993)] Barton, T., and D. Fuhrmann, 1993: Covariance structure for multidimensional data. *Multidimensional Systems and Signal Processing*, **4**, 111–123.

[Beaulieu and Tan(1997)] Beaulieu, N., and C. Tan, 1997: An fft method for generating bandlimited gaussian noise variates. in *Global Telecommunications Conference, 1997. (IEEE GLOBECOM '97).*, Vol. 2, pp. 684 – 688.

[Behbahani et al.(2000)] Behbahani, F., J. Leete, Y. Kishigami, A. Roithmeier, K. Hoshino, and A. Abidi, 2000: A 2.4-ghz low-if receiver for wideband wlan in 6-m cmos-architecture and front-end. *IEEE Journal of Solid-State Circuits*, **35**(12), 1908 – 1916.

[Bernard(1992)] Bernard, M. E., 1992: The Global Positioning System. *IEE Review*, **38**, 99–102.

[Biglieri et al.(1998)] Biglieri, E., J. Proakis, and S. Shamai, 1998: Fading channels: Information-theoretic and communications aspects. *IEEE Transactions on Information Theory*, **44**, 2619–2692.

[Brookes(2005)] Brookes, M., 2005: The matrix reference manual.

[Byers and Takawira(2004)] Byers, G., and F. Takawira, 2004: Spatially and temporally correlated MIMO channels: modeling and capacity analysis. *IEEE Transactions on Vehicular Technology*, **53**(3), 634–643.

[Capon(1969)] Capon, J., 1969: High resolution frequency-wavenumber spectrum analysis. *Proceedings of the IEEE*, **57**(8), 1408– 1418.

[Carlson(1988)] Carlson, B. D., 1988: Covariance matrix estimation errors and diagonal loading in adaptive arrays. *IEEE Transactions on Aerospace and Electronic Systems*, **24**(4), 397–401.

[Cedervall(1998)] Cedervall, M., 1998: Mobile positioning for third generation WCDMA systems. in *Proceedings IEEE 1998 International Conference on Universal Personal Communications, (ICUPC '98)*.

[Chizhik et al.(2002)] Chizhik, D., G. Foschini, M. Gans, and R. Valenzuela, 2002: Keyholes, correlations, and capacities of multielement transmit and receive antennas. *IEEE Transactions on Wireless Communications*, **1**, 361 – 368.

[Chuah et al.(2002)] Chuah, C.-N., D. Tse, J. Kahn, and R. Valenzuela, 2002: Capacity scaling in MIMO wireless systems under correlated fading. *IEEE Transactions on Information Theory*, **48**(3), 637 – 650.

[Cohen(2005)] Cohen, L., 2005: The history of noise [on the 100th anniversary of its birth]. *IEEE Signal Processing Magazine*, **2005**(6), 20– 45.

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

[Debbah and Muller(2005)] Debbah, M., and R. Muller, 2005: Mimo channel modelling and the principle of maximum entropy. *IEEE Transactions on Information Theory*, **51**, 1667 – 1690.

[Eng and Milstein(1995)] Eng, T., and L. Milstein, 1995: Coherent ds-cdma performance in nakagami multipath fading. *IEEE Transactions on Communications*, **43**(243), 1134 – 1143.

[Ertel et al.(1998)] Ertel, R., P. Cardieri, K. Sowerby, T. Rappaport, and J. Reed, 1998: Overview of spatial channel models for antenna array communication systems. *IEEE Personal Communications*, **5**(1), 10–22.

[Ertel and Reed(1998)] Ertel, R., and J. Reed, 1998: Generation of two equal power correlated Rayleigh fading envelopes. *IEEE Communications Letters*, **2**(10), 276–278.

[Etkin and Tse(2006)] Etkin, R., and D. Tse, 2006: Degrees of freedom in some underspread mimo fading channels. *IEEE Transactions on Information Theory*, **52**(4).

[Feinstein(1955)] Feinstein, A., 1955: Error bounds in noisy channels without memory. *IRE Transactions on Information Theory*, **1**, 13–14.

[Foschini and Gans(1998)] Foschini, G., and M. Gans, 1998: On limits of wireless communications in a fading environment when using multiple antennas. *Wireless Personal Communications*, **6**(3), 311–335.

[Frost(1972)] Frost, O. L., 1972: An Adaptive Algorithm for Linearly Constrained Adaptive Array Processing. *IEEE Proceedings*, **60**, 926–935.

[Gallager(1968)] Gallager, R. G., 1968: *Information Theory and Reliable Communication*. Wiley.

[Gamal et al.(2006)] Gamal, A. E., J. Mammen, B. Prabhakar, and D. Shah, 2006: Throughput-delay trade-off in wireless networks – part i: The fluid model. *IEEE Transactions on Information Theory*, **52**(6), 2568 – 2592.

[Gardner et al.(2006)] Gardner, W. A., A. Napolitano, and L. Paura, 2006: Cyclostationarity: Half a century of research. *Signal Processing*, **86**, 639–697.

[Gesbert et al.(2002)] Gesbert, D., H. Bolcskei, D. Gore, and A. Paulraj, 2002: Outdoor MIMO wireless channels: models and performance prediction. *IEEE Transactions on Communications*, **50**(12), 1926 – 1934.

[Getting(1993)] Getting, I. A., 1993: Perspective/Navigation-The Global Positioning System. *IEEE Spectrum*, **30**(12), 36–38.

[Gharavi and Ban(2003)] Gharavi, H., and K. Ban, 2003: Multihop sensor network design for wide-band communications. *Proceedings of the IEEE*, **91**(8), 1221–1234.

[Godara(1997)] Godara, L. C., 1997: Application of antenna arrays to mobile communications. II. Beam-forming and direction-of-arrival considerations. *Proceedings of the IEEE*, **85**(8), 1195–1245.

[Golden et al.(1999)] Golden, G., C. Foschini, R. Valenzuela, and P. Wolniansky, 1999: Detection algorithm and initial laboratory results using v-blast space-time communication architecture. *Electronics Letters*, **35**(1), 14 – 16.

[Goldsmith(2005a)] Goldsmith, A., 2005a: *Wireless Communications*. Chap. 3, pp. 71–77. Cambridge.

[Goldsmith(2005b)] Goldsmith, A., 2005b: *Wireless Communications*. Cambridge University Press.

[Goldsmith and Varaiya(1997)] Goldsmith, A., and P. Varaiya, 1997: Capacity of fading channels with channel side information. *IEEE Transactions on Information Theory*, **43**, 1986–1992.

[Gradshteyn and Ryzhik(1965)] Gradshteyn, I., and I. Ryzhik, 1965: *Table of Integrals, Series and Products*. Academic Press, 4 edition.

[Grant(1988)] Grant, L. V., 1988: An overview of the Federal Radionavigation Plan. in *Proceedings IEEE 21st Annual conference on Electronics and Aerospace Conference, (EASCON'88)*, pp. 173–177.

[Gray(1977)] Gray, R., 1977: Toeplitz and circulant matrices: A review.

[Grenander and Szegio(1958)] Grenander, U., and G. Szegio, 1958: *Toeplitz Forms and Their Applications*. University of Calif. Press, Berkeley and Los Angeles.

[Hellebrandt et al.(1997)] Hellebrandt, M., R. Mathar, and M. Scheibenbogen, 1997: Estimating position and velocity of mobiles in a cellular radio network. *IEEE Transactions on Vehicular Technology*, **46**(1), 65 – 71.

[Hepsaydir(1999)] Hepsaydir, E., 1999: Analysis of mobile positioning measurements in CDMA cellular networks. in *Proceedings 1999 IEEE Radio and Wireless Conference, (RAWCON '99)*.

[Hermann et al.(2001)] Hermann, R., M. Marc-Michel, K. Michael, and L. Urs, 2001: Research activities in automotive Radar. in *MSMW'01 Symp.*

[Hong et al.(2003)] Hong, Z., K. Liu, R. W. Heathand, and A. M. Sayeed, 2003: Spatial multiplexing in correlated fading via the virtual channel representation. *IEEE Journal on Selected Areas in Communications*, **21**(5), 856–866.

[Hyvarinen et al.(2001)] Hyvarinen, A., J. Karhunen, and E. Oja, 2001: *Independent Component Analysis*. John Wiley and Sons.

[Jakes(1974)] Jakes, W. C., 1974: *Microwave Mobile Communications*. New York: Wiley.

[Juels(2006)] Juels, A., 2006: Rfid security and privacy: a research survey. *IEEE Journal on Selected Areas in Communications*, **24**(3), 381 – 394.

[Kahne and Frolow(1996)] Kahne, S., and I. Frolow, 1996: Air traffic management: evolution with technology. *IEEE Proceedings*, **16**(4), 12–21.

[Kang and Alouini(2006)] Kang, M., and M.-S. Alouini, 2006: Capacity of correlated MIMO rayleigh channels. *IEEE Transactions on Wireless Communications*, **5**(1), 143 – 155.

[Kermoal et al.(2001)] Kermoal, J. P., L. Schumacher, K. I. Pedersen, P. E. Mogensen, and F. Frederiksen, 2001: A stochastic MIMO radio channel model with experimental validation. *IEEE Journal on Selected Areas in Communications*, **20**(6), 1211 – 1226.

[Kikuchi et al.(2003)] Kikuchi, S., A. Sano, H. Tsuji, and R. Miura, 2003: A novel approach to mobile-terminal positioning using single array antenna in urban environments. in *Proceedings 2003 IEEE 58th Vehicular Technology Conference, VTC 2003-Fall*, Vol. 2.

[Koch and Lapidoth(2006)] Koch, T., and A. Lapidoth, 2006: Gaussian fading is the worst fading. in *Proceedings 2006 IEEE International Symposium on Information Theory (ISIT)*, Seatle, USA.

[Lapidoth(2005)] Lapidoth, A., 2005: On the asymptotic capacity of stationary gaussian fading channels. *IEEE Transactions on Information Theory*, **51**(2), 437 – 446.

[Lapidoth and Moser(2003)] Lapidoth, A., and S. M. Moser, 2003: Capacity bounds via duality with applications to multi-antenna systems on flat fading channels. *IEEE Transactios on Information Theory*, **49**(10), 2426– 2467.

[Lee and Lee(1999)] Lee, J.-H., and Y.-T. Lee, 1999: Robust adaptive array beamforming for cyclostationary signals undercycle frequency error. *IEEE Transactions on Antennas and Propagation*, **47**(2), 233–241.

[Li et al.(2000)] Li, J., J. Jannotti, S. J. Douglas, D. Couto, D. R. Karger, and R. Morris, 2000: A Scalable Location Service for Geographic Ad Hoc Routing. in *Proceedings ACM Mobicom 2000*, pp. 120–130, Boston, MA.

[Lorenz and Boyd(2005)] Lorenz, R. G., and S. P. Boyd, 2005: Robust Minimum Variance Beamforming. *IEEE Transactions on Signal Processing*, **53**(5), 1684 – 1696.

[Mallik(2003)] Mallik, R. K., 2003: On multivariate rayleigh and exponential distributions. *IEEE Transactions on Information Theory*, **49**, 1499–1515.

[Marzetta and Hochwald(1999)] Marzetta, T., and B. Hochwald, 1999: Capacity of a mobile multipleantenna communication link in rayleigh flat fading. *IEEE Transactions on Information Theory*, **45**, 139–157.

[Ni et al.(2003)] Ni, L. M., Y. Liu, Y. C. Lau, and A. P. Patil, 2003: LANDMARC: indoor location sensing using active RFID. in *Proceedings of the First IEEE International Conference on Pervasive Computing and Communications, 2003, (PerCom 2003)*, pp. 407 – 415.

[Oestges et al.(2005)] Oestges, C., B. Clerckx, D. Vanhoenacker-Janvier, and A. Paulraj, 2005: Impact of Fading Correlations on MIMO Communication Systems in Geometry-Based Statistical Channel Models. *IEEE Transactions on Wireless Communications*, **4**(3), 1112 – 1120.

[Ozarow et al.(1994)] Ozarow, L., S. Shamai, and A. Wyner, 1994: Information theoretic considerations for cellular mobile radio. *IEEE Transactions on Vehicular Technology*, **43**(2), 359 – 378.

[Ozcelik et al.(2003)] Ozcelik, H., M. Herdin, W. Weichselberger, J. Wallace, and E. Bonek, 2003: Deficiencies of 'Kronecker' MIMO radio channel model. *IEEE Journal on Seleted Areas in Communications*, **39**(16), 1201 – 1210.

[Palmer et al.(1998)] Palmer, R. D., S. Gopalam, T. Yu, and S. Fukao, 1998: Coherent radar imaging using capon's method. *Radio Science*, **33**(6), 1585–1598.

[Papoulis(1985)] Papoulis, A., 1985: Predictable processes and wold's decomposition: A review. *IEEE Transactions on Acoustics, Speech, and Signal Processing [see also IEEE Transactions on Signal Processing]*, **33**(4), 933– 938.

[Papoulis and Pillai(2002)] Papoulis, A., and S. U. Pillai, 2002: *Probability, Random Variables and Stochastic Processes*. Chap. 11, pp. 499–522. MaGraw-Hill.

[Paulraj et al.(2004)] Paulraj, A., D. Gore, R. Nabar, and H. Bolcskei, 2004: An overview of MIMO communications - a key to gigabit wireless. *Proceedings of the IEEE*, **92**(2), 198–218.

[Pennisi(1987)] Pennisi, L. L., 1987: Coefficients of the characteristic polynomial. *Mathematics Magazine*, **60**(1), 31–33.

[Pourrostam et al.(2007)] Pourrostam, J., S. A. Zekavat, and H. Tong, 2007: Novel direction-of-arrival estimation techniques for periodic-sense local positioning systems. in *Proceeding IEEE RADAR 2007*, Waltham, MA.

[Proakis(2000)] Proakis, J., 2000: *Digital Communications*. McGraw-Hill Science/Engineering/Math, 4th edition.

[Ractliffe(1990)] Ractliffe, S., 1990: Air traffic control and mid-air collisions. *Journal of Electronics and Communication Engineering*, **2**(5), 202 – 208.

[Ramos et al.(1997)] Ramos, J., M. Zoltowski, and H. Liu, 1997: A low-complexity space-time rake receiver for ds-cdma communications. *IEEE Signal Processing Letters*, **4**(9), 262 – 265.

[Rao and Jones(2001)] Rao, A. M., and D. L. Jones, 2001: Efficient quadratic detection in perturbed arrays via Fourier transform techniques. *IEEE Transactions on Signal Processing*, **49**(7), 1269 – 1281.

[Rappaport(2002)] Rappaport, T. S., 2002: *Wireless Communications: Principles and Practice*. Prentice Hall, second edition.

[Regalia and Mitra(1989)] Regalia, P. A., and S. K. Mitra, 1989: Kronecker products, unitary matrices and signal processing applications. *SIAM Review*, **31**, 586–613.

[Saeks(1976)] Saeks, R., 1976: The factorization problem: survey. *Proceedings of the IEEE*, **64**, 90– 95.

[Saneyoshi(1996)] Saneyoshi, K., 1996: Drive assist system using stereo image recognition. in *IEEE Intelligent Vehicle Symp.*, pp. 230–235.

[Sayeed(2002)] Sayeed, A., 2002: Deconstructing multiantenna fading channels. *IEEE Transactions on Signal Processing*, **50**(10), 2563 – 2579.

[Sekine et al.(1992)] Sekine, M., T. Senoo, I. Morita, and H. Endo, 1992: Design method for an automotive laser radar system and future prospects for laser radar. in *Intelligent Vehicles Symposium '92.*, pp. 120–125.

[Shah and Haimovich(2000)] Shah, A., and A. Haimovich, 2000: Performance analysis of maximal ratio combining and comparison with optimum combining for mobile radio communications with cochannel interference. *IEEE Transactions on Vehicular Technology*, **49**(4), 1454–1463.

[Shannon(1948)] Shannon, C., 1948: A mathematical theory of communication. *Bell System Technical Journal*, **27**, 379–423, 623–656.

[Shin et al.(r in)] Shin, H., M. Z. Win, J. H. Lee, and M. Chiani, to appear in: On the capacity of doubly correlated MIMO channels. *IEEE Transactions on Wireless Communications*.

[Shiu et al.(2000)] Shiu, D.-S., G. Foschini, M. Gans, and J. Kahn, 2000: Fading correlation and its effect on the capacity of multielement antenna systems. *IEEE Transactions on Communications*, **48**(3), 502 – 513.

[Simon and Alouini(1998)] Simon, M., and M.-S. Alouini, 1998: A simple single integral representation of the bivariate Rayleigh istribution. *IEEE Communications Letters*, **2**, 128–130.

[Singh et al.(2004)] Singh, R., M. Guainazzo, and C. S. Regazzoni, 2004: Location determination using WLAN in conjuction with GPS network (global positioning system). in *Proceedings 2004 IEEE 59th Vehicular Technology Conference, VTC 2004-Spring*.

[Skolnik(1981)] Skolnik, M. I., 1981: *Introduction to Radar Systems*. Mc-Graw Hill Co.

[Springer and Thompson(1970)] Springer, M. D., and W. E. Thompson, 1970: The distribution of products of beta, gamma and gaussian random variables. *SIAM Journal on Applied Mathematics*, **18**(4), 721–737.

[Stoica and Moses(1997)] Stoica, P., and R. Moses, 1997: *Introduction to Spectral Analysis*. Prentice Hall.

[Stoica and Nehorai(1989)] Stoica, P., and A. Nehorai, 1989: Music, maximum likelihood, and cramer-rao bound. *IEEE Transactions on Signal Processing*, **37**(5), 720–741.

[Stoica et al.(2003)] Stoica, P., Z. Wang, and J. Li, 2003: Robust Capon Beamforming. *Signal Processing Letters*, **10**(6), 172–175.

[Tan and Beaulieu(1997)] Tan, C., and N. Beaulieu, 1997: Infinite series representations of the bivariate Rayleigh and Nakagami-m distributions. *IEEE Transactions on Communications*, **45**, 1159–1161.

[Tarokh et al.(1998)] Tarokh, V., N. Seshadri, and A. Calderbank, 1998: Space-time codes for high data rate wireless communication: performance criterion and code construction. *IEEE Transactions on Information Theory*, **44**(2).

[Telatar(1999)] Telatar, I. E., 1999: Capacity of multi-antenna Gaussian channels. *European Transactions on Telecommunications*, **10**(6), 585–595.

[Tong and Zekavat(2005)] Tong, H., and S. A. Zekavat, 2005: LCMV Beamforming for a Novel Wireless Local Positioning System: A Stationarity Analysis. in *Proceedings The International Society for Optical Engineering (SPIE 2005)*.

[Tong and Zekavat(2007)] Tong, H., and S. A. Zekavat, 2007: A novel wireless local positioning system via a merger of ds-cdma and beamforming: Probability-of-detection performance analysis under array perturbations. *IEEE Transactions on Vehicular Technologies*, **56**(3), 1307–1320.

[Tong and Zekavat(d toa)] Tong, H., and S. A. Zekavat, submitted toa: An eigen analysis of mimo channel models. in *IEEE GlobeCom 2007*.

[Tong and Zekavat(d tob)] Tong, H., and S. A. Zekavat, submitted tob: On the suitable environments of the kronecker product form in mimo channel modeling. in *IEEE GlobeCom 2007*.

[Trim(1990)] Trim, R. M., 1990: Mode S: An introduction and overview [secondary surveillance radar]. *Journal of Electronics and Communication Engineering*, **2**(2), 53–59.

[Tse and Viswanath(2005)] Tse, D., and P. Viswanath, 2005: *Fundamentals of Wireless Communications*. Chap. 2, pp. 25–29. Cambridge.

[Tulino et al.(2005)] Tulino, A., A. Lozano, and S. Verdu, 2005: Impact of antenna correlation on the capacity of multiantenna channels. *IEEE Transactions on Information Theory*, **51**(7), 2491 – 2509.

[Tulino and Verdu(2006)] Tulino, A., and S. Verdu, 2006: Capacity-achieving input covariance for single-user multi-antenna channels. *IEEE Transactions on Wireless Communications*, **5**(3), 662 – 671.

[Veeravalli et al.(2005)] Veeravalli, V., Y. Liang, and A. Sayeed, 2005: Correlated MIMO wireless channels: capacity, optimal signaling, and asymptotics. *IEEE Transactions on Information Theory*, **51**(6), 2058 – 2072.

[Verdu(1998)] Verdu, S., 1998: Fifty years of shannon theory. *IEEE Transactions on Information Theory*, **44**(6), 2057–2078.

[Vorobyov et al.(2004)] Vorobyov, S., A. Gershman, Z.-Q. Luo, and N. Ma, 2004: Adaptive beamforming with joint robustness against mismatched signal steering

vector and interference nonstationarity. *IEEE Signal Processing Letters*, pp. 108 – 111.

[Vossiek et al.(2003)] Vossiek, M., L. Wiebking, P. Gulden, J. Wieghardt, and C. Hoffmann, 2003: Wireless local positioning - Concepts, solutions, applications. in *Proceedings Radio and Wireless Conference, RAWCON '03. Proceedings*, pp. 219–224.

[Wang et al.(2001)] Wang, Y.-Y., J.-T. Chen, and W.-H. Fang, 2001: Joint estimation of doa and delay using tst-music in a wireless channel. *IEEE Signal Processing Letters*, **8**(2), 58 – 60.

[Wang and Giannakis(2003)] Wang, Z., and G. Giannakis, 2003: A simple and general parameterization quantifying performance in fading channels. *IEEE Transactions on Communications*, **51**(8), 1389 – 1398.

[Wang and Zekavat(2006)] Wang, Z., and S. A. Zekavat, 2006: Manet localization via multi-node toa-doa optimal fusion. in *proceedings IEEE MILCOM'06*, Washington DC.

[Wax and Anu(1997)] Wax, M., and Y. Anu, 1997: Performance analysis of the minimum variance beamformer in the presence of steering vector errors. *IEEE Transactions on Signal Processing*, **44**(4), 938–947.

[Weichselberger et al.(2006)] Weichselberger, W., M. Herdin, H. Ozcelik, and E. Bonek, 2006: A stochastic MIMO channel model with joint correlation of both link ends. *IEEE Transactions on Wireless Communications*, **5**(1), 90 – 100.

[Weisstein(Web)] Weisstein, E. W., Web: Gamma function. From MathWorld - A Wolfram Web Resource.

[Werb and Lanzl(1998)] Werb, J., and C. Lanzl, 1998: Designing a positioning system for finding things and people indoors. *IEEE Spectrum*, **35**(9), 71–78.

[Williamson and Spencer(1989)] Williamson, T., and N. A. Spencer, 1989: Development and operation of the Traffic Alert and Collision Avoidance System (TCAS). *IEEE Control Systems Magazine*, **77**(11), 1735–1744.

[Wu and Wong(1996)] Wu, Q., and K. M. Wong, 1996: Blind adaptive beamforming for cyclostationary signals. *IEEE Transactions on Signal Processing*, **44**(11), 2757–2767.

[Xia and Giannakis(2006)] Xia, P., and G. Giannakis, 2006: Design and analysis of transmit-beamforming based on limited-rate feedback. *IEEE Transactions on Signal Processing*, **54**(5), 1853 – 1863.

[Yang and Swindlehurst(1995)] Yang, J., and A. L. Swindlehurst, 1995: The effects of array calibration errors on DF-Based signal copy performance. *IEEE Transactions on Signal Processing*, **43**(11), 2724–2732.

[Yang and Belfiore(2006)] Yang, S., and J.-C. Belfiore, 2006: Diversity-multiplexing tradeoff of double scattering mimo channels. *IEEE Transactions on Information Theory, submitted to*.

[Young and Beaulieu(2000)] Young, D., and N. Beaulieu, 2000: The generation of correlated Rayleigh random variates by inverse discrete Fourier transform. *IEEE Transactions on Communications*, **48**(7), 1114 – 1127.

[Yu and Ottersten(2002)] Yu, K., and B. Ottersten, 2002: Models for MIMO propagation channels: a review. *Special Issue on Adaptive Antennas and MIMO Systems, Wiley Journal on Wireless Communications and Mobile Computing*, **2**(7), 553–666.

[Zekavat(2003)] Zekavat, S. A., 2003: Mobile Base Station (MBS) wireless with applications in vehicle early warning systems. in *Proceedings the University of Texas at Austin 2003 Wireless Networking Symposium.*

[Zekavat(2004)] Zekavat, S. A., 2004: A novel application for wireless communications in vehicle early warning. in *Proceedings IEEE Consumer Communications and Networking Conference*, pp. 715–717.

[Zekavat et al.(2004)] Zekavat, S. A., H. Tong, and J. Tan, 2004: A Novel Wireless Local Positioning System for Airport (Indoor) Security. in *Proceedings The International Society for Optical Engineering (SPIE 2004).*

[Zheng and Tse(2003)] Zheng, L., and D. Tse, 2003: Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels. *IEEE Transactions on Information Theory*, **49**(5), 1073 – 1096.

[Zhou and Sayeed(d to)] Zhou, Y., and A. Sayeed, submitted to: Experimental study of MIMO channel statistics and capacity via the virtual channel representation. *IEEE Transactions on Wireless Communications.*