

APPLICATION OF RANDOM MATRIX  
THEORY TO WIRELESS BROADCAST  
PROBLEMS AND SEISMIC SIGNAL  
PROCESSING

BY

**BABAR HASAN KHAN**

A Thesis Presented to the  
DEANSHIP OF GRADUATE STUDIES

**KING FAHD UNIVERSITY OF PETROLEUM & MINERALS**

DHAHRAN, SAUDI ARABIA

In Partial Fulfillment of the  
Requirements for the Degree of

**MASTER OF SCIENCE**

In

**ELECTRICAL ENGINEERING**

**December 2010**

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS  
DHAHRAN 31261, SAUDI ARABIA

DEANSHIP OF GRADUATE STUDIES

This thesis, written by **BABAR HASAN KHAN** under the direction of his thesis adviser and approved by his thesis committee, has been presented to and accepted by the Dean of Graduate Studies, in partial fulfillment of the requirements for the degree of **MASTER OF SCIENCE IN ELECTRICAL ENGINEERING**.

Thesis Committee



Dr. Tareq Y. Al Naffouri (Adviser)



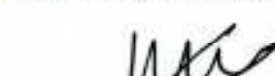
Dr. Samir Al Ghadban (Co-adviser)



Dr. Maan Kousa (Member)



Dr. Wail A. Mousa (Member)



Dr. Melwane Debbah (Member)



Dr. Samir H. Abdul Jauwad  
Department Chairman



Dr. Salam A. Zummo  
Dean of Graduate Studies

15/6/10  
Date

*Dedicated to my loving Mother & Father*

# ACKNOWLEDGMENTS

*In the name of Allah, the Most Beneficent, the Most Merciful*

Praise and gratitude to Allah, the Almighty, with Whose gracious help, I was able to accomplish this work patience and endurance.

Acknowledgement is due to King Fahd University of Petroleum and Minerals, Saudi Arabia and SUPELEC, France for providing additional funding and support to this work.

I am deeply indebted to my thesis advisor Dr. Tareq Al Naffouri for his constant support, guidance and encouragement throughout the course of this research and for many hours, day and night, of attention he devoted to the development of this study. I will always revere his patience, expert guidance and ability to solve intricate problems. He made my pursuit of higher education a truly enjoyable and unforgettable experience.

I also wish to thank Dr. Samir Al Ghadban who served as my thesis co-advisor for his valuable assistance and encouragement in the conduction and successfully completion of this thesis. He also helped me in the write up of the major part of the work I did with him on VBLAST.

I would like to express my profound gratitude and appreciation to Dr. Merouane

Debbah for providing me the opportunity to conduct a part of my research under his supervision in SUPELEC, France. He opened my eyes to a new world of Random Matrix Theory.

Sincere appreciation and grateful thank to my committee member Dr. Wail Mousa for his help and insight in the Seismic Signal Processing part of the thesis and providing me the opportunity to work in a new field of signal processing.

I also would like to thank my committee member Dr. Maan Kousa for his help and support.

Thanks are due to the chairmen of the Electrical Department Dr. Samir H. Abdul Jauwad for his support and assistance.

Special thanks to my dearest mother and father for their emotional support, love, sacrifices, prayers and understanding throughout my academic career. I would not be where I am in life if it weren't for their love.

I also acknowledge my colleagues and friends as I had a pleasant, enjoyable and fruitful company with them especially Salman Yousuf, Ahmed, Zeeshan, Saqib, Asif, Khaja, Zeehasham, Mazhar, Akhlaq, Ibrahim, Saad, Jamal, faraz, Khwaja, Danish and the other members of Mattam Al Miyar.

# TABLE OF CONTENTS

LIST OF TABLES	x
LIST OF FIGURES	xi
ABSTRACT (ENGLISH)	xv
ABSTRACT (ARABIC)	xvi
NOMENCLATURE	xvii
<b>Chapter 1: INTRODUCTION AND MOTIVATION</b>	<b>1</b>
1.1 Introduction . . . . .	1
1.2 Motivation . . . . .	3
1.3 Thesis Objectives . . . . .	3
1.4 Thesis Contributions . . . . .	4
1.5 Thesis Organization . . . . .	5
<b>Chapter 2: SCALING OF THE MINIMUM OF I.I.D. RANDOM VARIABLES</b>	<b>6</b>
2.1 Introduction . . . . .	7
2.2 What does scaling mean? . . . . .	7
2.3 Scaling law for the minimum of a number of random variables . .	10
2.3.1 A More general way to find scaling . . . . .	11
2.3.2 Finding the derivatives of $F(a)$ . . . . .	12
2.3.3 Procedure to find scaling of minimum . . . . .	13

2.3.4	Examples for the Scaling of minimum of a number of random variables . . . . .	14
2.3.5	Scaling of $\ h_i\ ^2$ . . . . .	16
2.3.6	Scaling for Maximum . . . . .	32
2.3.7	Chapter Conclusion . . . . .	39

**Chapter 3: UPLINK SCHEDULING FOR SPATIAL MULTIPLEXING SYSTEMS** **40**

3.1	Introduction . . . . .	41
3.2	System Model . . . . .	45
3.3	Optimal MIMO scheduling . . . . .	45
3.4	V-BLAST Scheduling . . . . .	46
3.5	Bounds on VBLAST Capacity . . . . .	51
3.5.1	MaxMinSV . . . . .	51
3.5.2	MinES . . . . .	54
3.5.3	MaxSNR . . . . .	57
3.5.4	MinInvTrace . . . . .	58
3.6	Simulation Results . . . . .	59
3.6.1	BER Performance and Outage Capacity . . . . .	59
3.6.2	Effect of Suboptimal Detection . . . . .	61
3.6.3	Advantage of V-BLAST over SISO and SIMO systems . . . . .	62
3.6.4	Spatial Multiplexing with Sphere Decoding . . . . .	63
3.7	Chapter Summary . . . . .	67

**Chapter 4: ESTIMATION OF THE DISTRIBUTION OF RANDOMLY DEPLOYED WIRELESS SENSORS** **69**

4.1	Introduction . . . . .	70
4.2	Sensor Deployment . . . . .	71
4.2.1	System Model . . . . .	71
4.2.2	Distribution Estimation . . . . .	73
4.3	Moments Approach . . . . .	73

4.3.1	Step 1: Rectangular additive free deconvolution . . . . .	74
4.3.2	Step 2: Multiplicative free deconvolution . . . . .	76
4.3.3	Step 3: Moments of $\mathbf{V}\mathbf{V}^H$ . . . . .	77
4.3.4	Step 4: Estimation of $P$ . . . . .	81
4.3.5	Step 5: PDF Approximation . . . . .	81
4.4	Simulation Results . . . . .	82
4.4.1	Estimation of the moments of $\mathbf{V}\mathbf{V}^H$ . . . . .	83
4.4.2	Estimation of $P$ . . . . .	85
4.4.3	Estimation of the noise variance $\sigma^2$ . . . . .	86
4.4.4	Special Cases . . . . .	87
4.4.5	PDF Approximation . . . . .	91
4.5	Chapter Conclusion . . . . .	94
 <b>Chapter 5: APPLICATION TO SEISMIC SIGNAL PROCESS-</b>		
<b>ING</b>		<b>96</b>
5.1	Introduction . . . . .	96
5.1.1	Data Acquisition - Common Shot Point . . . . .	97
5.1.2	Multiples . . . . .	99
5.2	Sample Covariance Matrix . . . . .	100
5.3	Asymptotic Behavior of Sample Covariance Matrix . . . . .	101
5.4	Narrow Band Case . . . . .	105
5.4.1	System Model . . . . .	105
5.4.2	High Resolution Spectral Estimators . . . . .	107
5.4.3	Minimum Norm . . . . .	108
5.5	Wide Band Case . . . . .	112
5.5.1	Before Preflattening . . . . .	112
5.5.2	After Preflattening . . . . .	113
5.5.3	Simulation Results . . . . .	114
5.6	Subspace Approach . . . . .	125
5.6.1	System Model . . . . .	125



5.6.2	Estimation of Time Delays . . . . .	127
5.6.3	Estimation of Velocity and Zero offset Time . . . . .	128
5.6.4	Simulation Results . . . . .	129
5.7	Chapter Conclusion . . . . .	130
<b>Chapter 6: CONCLUSIONS AND FUTURE WORK</b>		<b>133</b>
6.1	Conclusion . . . . .	133
6.2	Future Work . . . . .	133
<b>APPENDIX A: Proof of Equation (4.18)</b>		<b>135</b>
<b>APPENDIX B: Proof of Equation (4.19)</b>		<b>138</b>
<b>REFERENCES</b>		<b>140</b>
<b>VITAE</b>		<b>150</b>

# LIST OF TABLES

2.1	Scaling Table for minimum of different bounded distributions . . .	37
5.1	RMSE for MUSIC and GMUSIC at different SNRs . . . . .	129

# LIST OF FIGURES

1.1	The Marcenko Pasture density function for $c=0.1$ , $c=0,5$ and $c=1$ .	2
2.1	Uniform Distribution . . . . .	17
2.2	U-Quadratic Distribution . . . . .	18
2.3	Non-Central Chi Square Distribution . . . . .	19
2.4	Chi Square Distribution . . . . .	20
2.5	Irwin Hall Distribution . . . . .	21
2.6	Gamma Distribution . . . . .	22
2.7	Rayleigh Distribution . . . . .	23
2.8	Pareto Distribution . . . . .	24
2.9	Log Logistic Distribution . . . . .	25
2.10	Half Normal Distribution . . . . .	26
2.11	Folded Normal Distribution . . . . .	27
2.12	Kumaraswamy Distribution . . . . .	28
2.13	Beta Distribution . . . . .	29
2.14	$\ h_i\ ^2$ . . . . .	30
2.15	$\ h_i\ ^2$ , non-central and with different variances . . . . .	31
2.16	Uniform Distribution . . . . .	33
2.17	U-Quadratic Distribution . . . . .	34
2.18	Irwin Hall Distribution . . . . .	35
2.19	Beta Distribution . . . . .	36
3.1	Uplink MIMO Scheduling . . . . .	44
3.2	PDF of $\lambda_{min}$ . . . . .	54

3.3	PDF of $\kappa$ . . . . .	56
3.4	Aggregate BER of 4x4 QPSK V-BLAST users with uplink scheduling	60
3.5	Capacity CCDF of 4x4 QPSK V-BLAST with uplink scheduling .	61
3.6	Capacity versus number of users at 4x4 MIMO channels and at 10% Outage probability . . . . .	62
3.7	Capacity versus number of users at 10% Outage probability for suboptimal detectors . . . . .	63
3.8	Spectral advantage of V-BLAST over receive diversity and SISO systems with uplink scheduling . . . . .	64
3.9	BER Comparison of V-BLAST and MRC with uplink scheduling .	65
3.10	Sphere Decoder scheduling for 4x4 spatial multiplexing uplink users	65
3.11	Capacity versus number of users at 4x4 MIMO channels and at 10% outage probability for SM-SD . . . . .	66
4.1	Relative distance of the moments of $\mathbf{V}\mathbf{V}^H$ and the estimated mo- ments by applying moments approach for varying $\sigma$ . . . . .	83
4.2	Relative distance of the moments of $\mathbf{V}\mathbf{V}^H$ and the estimated mo- ments by applying moments approach for $\sigma^2 = 1$ . . . . .	84
4.3	Estimation of $P$ for Uniform distribution with $K = L = 1000$ for varying $\sigma$ . . . . .	85
4.4	Estimation of $P$ for Uniform distribution with $K = L = 1000$ with $\sigma = 1$ . . . . .	86
4.5	Estimation of $\sigma$ for Uniformly distributed $\omega$ with $K = L = 1000$ .	86
4.6	Estimation of $\sigma$ for Uniformly distributed $\omega$ with $K, L$ and $P$ increasing asymptotically . . . . .	87
4.7	Estimation of $\kappa$ for Von Mises distribution with $K = L = P = 100$	88
4.8	Estimation of $\kappa$ for Von Mises distribution with varying $K, L$ and $P$	89
4.9	Estimation of $\kappa$ for Von Mises distribution with $K = L = P = 100$ and varying $\sigma$ . . . . .	90

4.10	Estimation of $R$ for wigner semi-circle with $K = L = P = 100$ and $\sigma^2 = 1$ . . . . .	90
4.11	Estimation of $R$ for wigner semi-circle with $\sigma^2 = 1$ and varying $K$ , $L$ and $P$ . . . . .	91
4.12	Estimation of $R$ for wigner semi-circle with varying $\sigma^2$ and $K = L = P = 1000$ . . . . .	92
4.13	Estimation of the PDF of $\omega$ , where $\omega$ has a beta distribution with $\alpha = 2$ and $\beta = 5$ . . . . .	93
4.14	Estimation of the PDF of $\omega$ , where $\omega$ has a beta distribution with $\alpha = 1$ and $\beta = 3$ . . . . .	94
4.15	Estimation of the PDF of $\omega$ , where $\omega$ has a beta distribution with $\alpha = 2$ and $\beta = 2$ . . . . .	95
5.1	Common Source Point. . . . .	97
5.2	Seismic Data(wiggle display) courtesy of Yilmaz [50]. . . . .	98
5.3	Multiple Layers. . . . .	99
5.4	Ricker Wavelet . . . . .	111
5.5	Wigger Plot of the synthetic noise free received data with $M = 20$ , $N = 40$ , $f_c = 5\text{Hz}$ , $\theta_1 = 35^\circ$ and $\theta_2 = 37^\circ$ . . . . .	115
5.6	Wigger Plot of the synthetic received data with $M = 20$ , $N = 40$ , $\theta_1 = 35^\circ$ , $f_c = 5\text{Hz}$ , $\theta_2 = 37^\circ$ and $SNR = 10\text{dB}$ . . . . .	116
5.7	Wigger Plot of the synthetic received data with $M = 20$ , $N = 40$ , $\theta_1 = 35^\circ$ , $f_c = 5\text{Hz}$ , $\theta_2 = 37^\circ$ and $SNR = 5\text{dB}$ . . . . .	117
5.8	Estimation of Direction of Arrival (DOA) with $\theta_1 = 35^\circ$ , $\theta_2 = 37^\circ$ , $f_c = 5\text{Hz}$ , $M = 20$ , $N = 20$ , $d/\lambda_c = 0.5$ , Empirical $SNR = 10\text{dB}$ . . .	118
5.9	Estimation of Direction of Arrival (DOA) with $\theta_1 = 35^\circ$ , $\theta_2 = 37^\circ$ , $f_c = 5\text{Hz}$ , $M = 20$ , $N = 100$ , $d/\lambda_c = 0.5$ , Empirical $SNR = 10\text{dB}$ . . .	118
5.10	Estimation of Direction of Arrival (DOA) with varying $N$ , $\theta_1 = 35^\circ$ , $\theta_2 = 37^\circ$ , $M = 20$ and empirical $SNR = 10\text{dB}$ . . . . .	119

5.11	Minimum Square Error for the direction of arrival at different $SNRs$ with $\theta_1 = 35^\circ$ , $\theta_2 = 37^\circ$ , $f_c = 5\text{Hz}$ , $M = 20$ , $N = 20$ , $d/\lambda_c = 0.5$ . . . .	120
5.12	Wigger Plot of data before preflattening with $f_c = 30\text{Hz}$ , $V_1 = 3600\text{m/s}$ , $V_2 = 7000\text{m/s}$ , $M = 20$ , $N = 300$ and $d = 60\text{m}$ . . . . .	121
5.13	Estimate of velocities $V_1 = 3600\text{m/s}$ and $V_2 = 7000\text{m/s}$ before Preflatten- tening . . . . .	122
5.14	Wigger Plot of data after preflattening with $V_1 = 2700\text{m/s}$ , $V_2 = 3600\text{m/s}$ , $V_a = 3000\text{m/s}$ , $f_c = 30\text{Hz}$ , $M = 20$ , $N = 200$ . . . . .	123
5.15	Estimate of velocities $V_1 = 2700\text{m/s}$ and $V_2 = 3600\text{m/s}$ after Preflat- tening with $V_a = 3000\text{m/s}$ , $f_c = 30\text{Hz}$ , $M = 20$ , $N = 200$ . . . . .	124
5.16	Wigger Plot of the seismic data with $M = 20$ , $K = 150$ , $v = 600\text{m/s}$ , $T_0 = 0.1\text{s}$ , $P = 20\text{Hz}$ , $Q = 50\text{Hz}$ , $f_c = 40\text{Hz}$ and the empirical $SNR = 10\text{dB}$ . . . . .	131
5.17	MUSIC Estimate of velocity and zero-offset time . . . . .	132
5.18	GMUSIC Estimate of velocity and zero-offset time . . . . .	132

# THESIS ABSTRACT

**NAME:** Babar Hasan Khan  
**TITLE OF STUDY:** Application of Random Matrix Theory to Wireless Broadcast Problems and Seismic Signal Processing  
**MAJOR FIELD:** Electrical Engineering  
**DATE OF DEGREE:** December 2009

*This thesis explores some of the possible application of Random Matrix Theory (RMT) to the field of wireless broadcast problems and seismic signal processing. In this study, a new scaling law for the bounded minimum random variable was developed directly from the characteristic function of the random variable. An optimal scheduler was devised to maximize the VBLAST (Vertical Bell Labs Layered Space Time Architecture) uplink capacity while reducing the complexity of calculations. By using the moments approach we estimated the distribution of the randomly deployed sensors without any communication between themselves. If the distribution is already known then new parametric estimators are derived. Using the results of Random Matrix Theory (RMT) we were able to derived new estimators for direction of arrivals, velocity and zero offset time for the seismic signals.*

# خلاصة الرسالة

الاسم الكامل : بابر حسن خان

عنوان الرسالة : تطبيقات نظرية المصفوفات العشوائية في مجالات الاتصالات اللاسلكية ومعالجة الاشارات الزلزالية.

التخصص : هندسة الكهربائية

تاريخ الشهادة : ديسمبر 2009

تتضمن هذه الرسالة بحثا في التطبيقات الممكنة لنظرية المصفوفات العشوائية في مجالات الاتصالات اللاسلكية و معالجة الاشارات الزلزالية. هذه الدراسة تشتمل على تطوير قانون التناسب المتعلق بالمتغير العشوائي ذو الحد الأدنى وذلك بالاعتماد على الدالة المميزة للمتغير العشوائي. اعتمادا على هذا التطوير تم اقتراح الجدول المثالي والذي يعمل على زيادة سعة الخط التحميلي لل (VBLAST) (شيفرات مختبرات بل المكانية الزمانية المتعامدة) مع التقليل من كمية الحسابات المعقدة.

اضافة الى ما سبق، واعتمادا على طريقة حساب القيم المتوقعه فلقد قمنا باعطاء توزيع للمجسات العشوائية والتي لا يوجد اي نوع من انواع التصال بينها على الاطلاق. وبالاعتماد على هذا التوزيع الذي اصبح معروفا لدينا فاننا قمنا بتقديم محددات متغيرة جديدة ، وبالاعتماد على نتائج التي حصلنا عليها من تطبيق نظرية المصفوفات العشوائية قمنا باقتراح محددات جديدة للاتجاهات التي تصل من خلالها الاشارة الزلزالية و محددات للسرعة وصول الاشارة الزلزالية بالاضافة الى محددات للوقت الخاص بالاشارة الزلزالية



# Nomenclature

## Abbreviations

MIMO	:	Multiple Input Multiple Output
SISO	:	Single Input Single Output
LOS	:	Line of Sight
RMT	:	Random Matrix Theory
VBLAST	:	Vertical-Bell-Labs LAYered Space Time architecture
SNR	:	Signal to Noise Ratio
AWGN	:	Additive white Gaussian noise
SM	:	Spatial Multiplexing
SM-SD	:	Spatial Multiplexing with Sphere Decoding
DOAs	:	Direction Of Arrivals
PDF	:	Probability Density Function
CDF	:	Cumulative distribution function
QoS	:	Quality of Service
CSP	:	Common Shot Point
MVDR	:	Minimum Variance Distortionless Response
MUSIC	:	MUltiple Signal Classification algorithm
DFT	:	Discrete Fourier Transform

## Notations

- $(.)^H$  : Hermitian transpose
- $\|\cdot\|$  : Euclidean Norm
- $Tr(\cdot)$  : Trace of a matrix
- $Tr_a(\cdot)$  : Normalized trace of a matrix
- $\mathcal{N}(m, R)$  : Gaussian random variable with a mean  $m$  and  
a auto-correlation matrix  $R$
- $\boxplus$  : Rectangular additive free convolution
- $\boxminus$  : Rectangular additive free deconvolution
- $\boxtimes$  : Multiplicative free convolution
- $\boxdiv$  : Multiplicative free deconvolution

# CHAPTER 1

# INTRODUCTION AND MOTIVATION

## 1.1 Introduction

Random Matrix Theory (RMT) initially motivated by practical problems [1] [2] [3] has now evolved into a highly influential field. Random matrix theory now is involved in fields as diverse as communications, seismic signal processing, nuclear physics, neural network, information theory and statistical physics. After the initiation of “Free probability theory” by Dan Voiculescu in the late 1980s, numerous literature exploiting random matrix theory has emerged [4] [5].

Under certain assumptions random matrix theory offers insight into the asymptotic behaviors of the distribution of the eigenvalues and singular values of random matrices. Random matrices of fixed and small order do not offer much information but as the matrices grow at a fixed ratio of its number of columns and rows, the

empirical distributions converge to a non-random deterministic functions.

For example, the empirical distribution of the eigenvalues of  $\mathbf{H}^H\mathbf{H}$  (where  $\mathbf{H}$  is a matrix with zero mean i.i.d Gaussian entries) almost surely converge to Marcenko Pasture distribution as the number of rows and columns of  $\mathbf{H}$  increases with a fixed ratio. Figure 1.1 shows the plot of Marcenko Pasture density for different

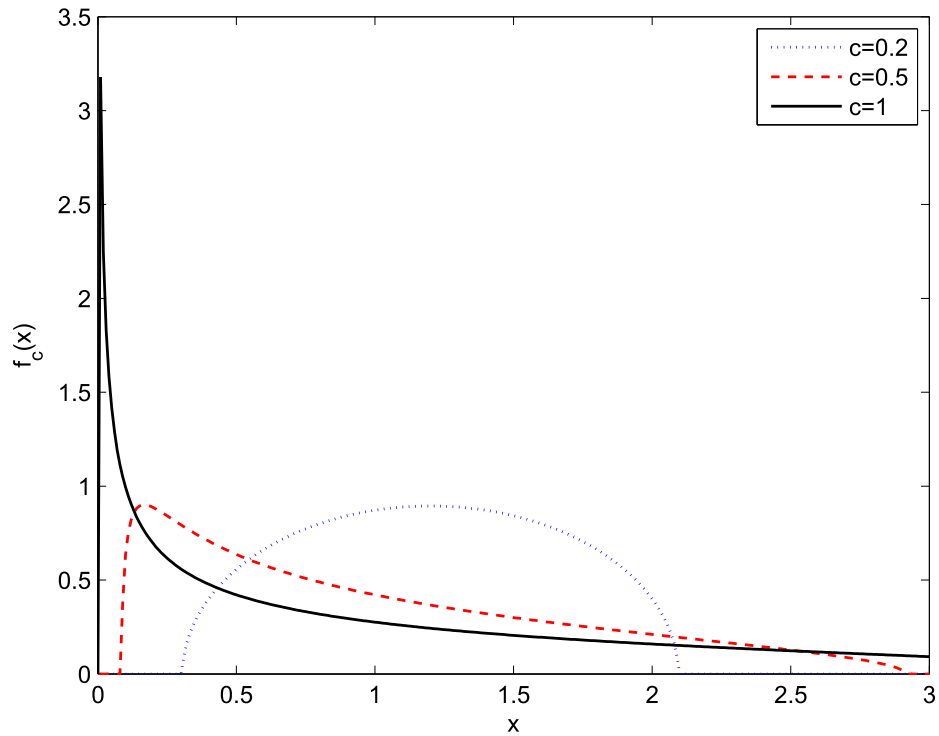


Figure 1.1: The Marcenko Pasture density function for  $c=0.1$ ,  $c=0,5$  and  $c=1$ .

values of  $c$  where  $c$  is the ratio of the number of rows and the number of columns of the matrix  $\mathbf{H}$ .

## 1.2 Motivation

New tools are emerging from the literatures of random matrix theory. These tools can be exploited for solving the problems of wireless communication and seismic signal processing. The randomness found in both fields will prove to be useful when carrying out parametric estimations.

In communications, processing of matrices of large dimensions can prove to be computationally expensive. Any certainty or simplicity of the wireless channel can be used to solve a variety of problems e.g. providing bounds on capacity, making new decoding algorithms and developing new multi-scheduling techniques.

Similarly, the present era of high fuel and energy demand has made mapping a seismic profile of the underground terrain for future explorations more important than ever. Seismic data suffers from computational complexity because of the voluminous size of the data to be processed. Random matrix theory is ideal for applications where the dimensions of matrices are large.

## 1.3 Thesis Objectives

The main goal of this work is to explore different ways we can use or develop tools from random matrix theory (RMT) for the application of wireless communications and seismic signal processing.

## 1.4 Thesis Contributions

The work presents results of investigation related to the objective mentioned in the previous section. The main contributions of this work are summarized as follows,

- This thesis develops scaling laws for the minimum of bounded i.i.d random variables.
- This thesis develops several suboptimal uplink schedulers for a VBLAST spatial multiplexing system. The thesis provides bounds on the rate achieved using these techniques.
- This thesis develops an estimator of the distribution of randomly deployed wireless sensors. As a consequence some parametric estimators are also presented such as number of harmonics and noise variance.
- This thesis develops a new estimator for the direction of arrivals (DOAs) useful for seismic applications. We also provide with estimators for the velocities of the incoming waves. The estimators are applied to narrow band and wideband seismic models and shows significant performance. Subspace approach estimators is also developed for simultaneously estimating velocity and zero offset time delay.

## 1.5 Thesis Organization

The rest of the thesis is organized as follows.

Chapter 2 was an extension of Dr. Naffouri's work. It provides with a scaling law for positive i.i.d random variables and highlights its use in communication problems.

Chapter 3 was written under the supervision of Dr. Samir Al Ghadban. This chapter studies different scheduling schemes for VBLAST algorithm and provides with the best suboptimal scheduling scheme. The chapter also provides with bounds on capacity under different schemes.

Chapter 4 was written under the supervision of Dr. Merouane Debbah. In collaboration between KFUPM and SUPELEC institute I worked with Dr. Merouane during the research internship in France paid by both institutes. Chapter 4 provides an estimator for the distribution of randomly deployed wireless sensors using moments approach from RMT. The chapter also presents other parametric estimators for the wireless sensors.

Chapter 5 was carried out under the supervision and guidance of Dr. Wail Mousa. This chapter presents a new spectral estimator for the estimation of direction of arrivals (DoAs). RMT spectral estimators are also modified to be used in the wideband case for the estimation of velocities. These estimators are also used for the subspace model.

This thesis ends with conclusion and future direction in Chapter 6. An appendix is attached in the end followed by references.

## CHAPTER 2

# SCALING OF THE MINIMUM OF I.I.D. RANDOM VARIABLES

Studying the performance of multiuser MIMO communication systems can be quite challenging. Research in this area has thus focused on the behavior of such systems in the large number of users and antennas regimes which relies on understanding the scaling of a large number of iid random variables. In this chapter, we study the scaling behavior of the max-min norm,  $\max_{B \geq 0} \text{Tr}(B) \leq P \min_{1 \leq i \leq n} \|h_i\|_B^2$ , where  $h_i$  are iid vectors. Such a study is for example important for evaluating the asymptotic performance of multicast systems in which a base station with  $M$  antennas broadcasts common information to  $n$  users. We study this scaling for various regimes including large  $M$  and large  $n$  as well as regimes where both  $n$  and  $M$  grow simultaneously.



## 2.1 Introduction

Multiuser communication is currently an active area of research where one is interested in scheduling users, maximizing system throughput, fairness issues,... etc [24], [23], [28], [21]. Solving multiuser problems for an arbitrary number of users might be too complex. As such, in many instances, researchers resort to studying the asymptotic regimes of large number of users. This usually entails studying the scaling of iid random variables.

A prerequisite to this entails finding closed form expressions for the CDF and pdf of the variables involved which might still be too prohibitive. Moreover, sometimes it might be easier to characterize the behavior of the characteristic function.

In this chapter, we show how the scaling law of the minimum of iid random variables can be obtained by studying the behavior of the CDF and its derivatives at one point. This can also be obtained by studying the behavior of the characteristic function at infinity.

## 2.2 What does scaling mean?

We say that a variable scales for large  $n$  scales if there are sequences  $a_n$  and  $b_n$  such that

$$F^n(a_n X + b_n) \rightarrow G(x) \quad \text{as } n \rightarrow \infty$$

at all continuity points of  $G(x)$ . It has been shown that when such a  $G(x)$  exists, it falls into one of three categories:

$$G_1(x; \alpha) = \begin{cases} 0; & x \leq 0, \alpha < 0 \\ \exp(-x^{-\alpha}); & x > 0 \end{cases}$$

$$G_2(x; \alpha) = \begin{cases} \exp(-(-x)^\alpha); & x \leq 0, \alpha < 0 \\ 0; & x > 0 \end{cases}$$

$$G_3(x) = \exp(-e^{-x}); \quad -\infty < x < \infty$$

Similar asymptotic distributions exist for the minimum of iid random variables. Specifically, there are three limiting distributions  $G_i^*(x)$  defined in terms of the maximum pdf counterpart

$$G_i^*(x) = 1 - G_i(-x)$$

**Example: Beamforming for Multicast** In beamforming [23], the transmitter sends  $M$  beams  $\phi_1, \phi_2, \dots, \phi_M$ , and asks each user to feedback the SINR associated with each beam. For example the SINR associated with beam  $\phi_1$  for user  $i$  is given

by

$$\text{SINR}_i = \frac{|h_i^* \phi_1|^2}{\frac{1}{\rho} + \sum_{m=2}^M |h_i^* \phi_m|^2}$$

where  $\rho = \frac{M}{P}$  is the signal to noise ratio. Since the base station has to appeal to all users in the group, we are constrained by the worst user. For a given beam, the  $\text{SINR}_i$ 's are iid with CDF

$$F_{\text{SINR}}(x) = 1 - \frac{e^{-\frac{x}{\rho}}}{(1+x)^{M-1}} \quad x \geq 0$$

So the CDF of the minimum of  $n$  such SINR's is given by

$$F_{\min \text{ SINR}}(x) = 1 - (1 - F_{\text{SINR}}(x))^n \quad (2.1)$$

$$= 1 - \frac{e^{-\frac{nx}{\rho}}}{(1+x)^{n(M-1)}} \quad x \geq 0 \quad (2.2)$$

Now note that

$$\lim_{n \rightarrow \infty} F_{\min \text{ SINR}} \left( \frac{x}{\left(\frac{1}{\rho} + M - 1\right)n} \right) = \lim_{n \rightarrow \infty} 1 - \frac{e^{(-\frac{x}{\rho} + M - 1)}}{\left(1 + \frac{x}{n(\frac{1}{\rho} + M - 1)}\right)^{n(M-1)}} \quad (2.3)$$

$$= 1 - e^{-x} \quad (2.4)$$

$$= 1 - G_1(x; 1) \quad (2.5)$$

This shows that for large  $n$ ,  $\min_i \text{SINR}_i$  scales as  $\frac{1}{(\frac{1}{\rho} + M - 1)n}$ .

The method of Example 1 might not apply all the time as it is difficult to find the CDF in closed form sometimes.

**Example 2: Scaling of spatially correlated channel norms** Consider the issue of finding the scaling of the  $\min_i \|h_i\|^2$  when  $h_i \sim \mathcal{N}(m, R)$ . When  $R = I$  and  $m = 0$ ,  $\|h\|^2$  is chi-square distributed with  $M$  degrees of freedom. On the other hand, when we deviate from this ideal case, say when  $m \neq 0$ , the CDF has different forms depending on whether some of the eigenvalues of the  $R$  are the same or not. In the case that  $m \neq 0$ , we don't even have a closed expression for the CDF (or pdf).

In the following subsection, we provide a general method for finding the scaling of the minimum.

## 2.3 Scaling law for the minimum of a number of random variables

Let  $X_1, X_2, \dots, X_n$  be iid random variables with pdf  $f(x)$ , CDF  $F(x)$ , and characteristic function  $\phi(x)$ . Let  $a$  also be the infimum of the support of  $X_i$ .<sup>1</sup> We would like to find the scaling law of the minimum,  $x_{\min}(n) = \{x_1, x_2, \dots, x_n\}$ . To this end, note that the CDF of the minimum is given by

$$F_{\min}(x) = 1 - (1 - F(x))^n$$

### 2.3.1 A More general way to find scaling

Let's expand  $F(x)$  in a Taylor series

$$F(x) = \sum_{i=0}^{\infty} F^{(i)}(a) \frac{(x-a)^i}{i!} \quad (2.6)$$

Note that  $F(a) = 0$  and let  $F_{\min}^{(i_0)}(a)$  be the first nonzero derivative of  $F(x)$  around  $a$ . Then

$$F_{\min}(x) = 1 - \left( 1 - \frac{F^{(i_0)}(a)}{i_0!} (x-a)^{i_0} - \sum_{i=i_0+1}^{\infty} F^{(i)}(a) \frac{(x-a)^i}{i!} \right)^n$$

Now, we claim that  $F^{(i_0)}(a) > 0$ . For if it were negative, then  $F^{(i_0-1)}(a) > 0$  would be decreasing in an interval  $(a, a + \epsilon)$ . Or as  $F^{(i_0)}(a) = 0$ , we see that  $F^{(i_0-1)}(x)$  is negative in this interval. Continuing this way, we can show that  $F^{(i_0-1)}(x), \dots, F^{(0)}(x)$  are negative in  $(a, a + \epsilon)$  which contradicts the nonnegative nature of  $F(x)$ .

We can thus replace  $x$  by  $\frac{i_0!^{\frac{1}{i_0}}}{F^{(i_0)}(a)^{\frac{1}{i_0}}} \frac{x}{n^{\frac{1}{i_0}}}$ . Then

$$F_{\min} \left( \frac{i_0!^{\frac{1}{i_0}}}{F^{(i_0)}(a)^{\frac{1}{i_0}}} \frac{x}{n^{\frac{1}{i_0}}} + a \right) = 1 - \left( 1 - \frac{x^{i_0}}{n} + \mathcal{O}\left(\frac{1}{n^{\frac{i_0+1}{i_0}}}\right) \right)^n$$

which for large  $n$  reads

$$\lim_{n \rightarrow \infty} F_{\min} \left( \frac{i_0!^{\frac{1}{i_0}}}{F^{(i_0)}(a)^{\frac{1}{i_0}}} \frac{x}{n^{\frac{1}{i_0}}} + a \right) = 1 - \exp(-x^{i_0}) \quad (2.7)$$

This is of the form

$$\lim_{n \rightarrow \infty} F_{\min}(a_n x + a) = 1 - \exp(-x^{i_0}) \quad (2.8)$$

$$= 1 - G_2(-x; i_0) \quad (2.9)$$

where

$$a_n = \frac{i_0!^{\frac{1}{i_0}}}{F^{(i_0)}(a)^{\frac{1}{i_0}}} \frac{1}{n^{\frac{1}{i_0}}}$$

### 2.3.2 Finding the derivatives of $F(a)$

It remains to find the least  $i_0$  for which  $F^{i_0}(a) \neq 0$ . Fortunately, we can do so without having to explicitly find the CDF and its derivatives by using the characteristic function and relying instead on the initial value theorem. To do so, define

$$D(x) = F(x + a)$$

then

$$\lim_{x \rightarrow 0} D^{(j)}(x) = \lim_{x \rightarrow a} F^{(j)}(x)$$

Now recall that the pdf  $f(x)$  and the characteristic function  $\phi(s)$  form a Laplace transform pair. Then by the time-shift and differentiation properties

$$D^{(j)}(x) \rightarrow s^{(j-1)} e^{-as} \phi(s)$$

Now applying the initial value theorem to the Laplace transform pair above yields

$$\lim_{x \rightarrow 0} D^{(j)}(x) = \lim_{s \rightarrow \infty} s^j e^{-as} \phi(s)$$

i.e.

$$\boxed{\lim_{x \rightarrow a} F^{(j)}(x) = \lim_{s \rightarrow \infty} s^{(j-1)} e^{-as} \phi(s)}$$

We can summarize the results of this section in the following theorem

**Theorem 2.1** *Let  $X_1, X_2, \dots, X_n$  be iid nonnegative random variables with CDF  $F(x)$ , and characteristic function  $\phi(x)$ . Assume that  $X_i$  is bounded from below and let  $a$  be the infimum of the support of  $X_i$ . Let  $X_{\min}(n)$  denote the minimum of these random variables  $\min \{X_1, X_2, \dots, X_n\}$ . Then  $a_n x_{\min}(n) + a$  converges in distribution to random variable  $y$  with CDF  $F_y(y) = 1 - \exp(-y^{i_0})$  where  $i_0$  is the first non-zero derivative of  $F(x)$  at zero, i.e.,  $F^{(i_0)}(a) \neq 0$  and  $F^{(j)}(a) = 0$  for all  $j < i_0$  and where*

$$a_n = \frac{i_0!^{\frac{1}{i_0}}}{F^{(i_0)}(a)^{\frac{1}{i_0}}} \frac{1}{n^{\frac{1}{i_0}}}$$

Furthermore, we can find  $F^{(i_0)}(a)$  using the initial value theorem

$$\boxed{\lim_{x \rightarrow a} F^{(j)}(x) = \lim_{s \rightarrow \infty} s^{(j-1)} e^{-as} \phi(s)}$$

### 2.3.3 Procedure to find scaling of minimum

1. Find the infimum (i.e. least value) of the random variable and denote it  $a$ .

Thus for  $\|h_i\|^2$ ,  $a = 0$  and for  $\det(I + H_i H_i^*)$ ,  $a = \det(I) = 1$ .

2. Find the first derivative such that

$$\lim_{x \rightarrow a} F^{(j)}(x) \neq 0$$

and use this to find the sequences  $a_n$  and  $b_n$  defined above. This gives the scaling behavior.

3. If there is difficulty finding the derivative of  $F(x)$ , use the characteristic function as outlined above.

### 2.3.4 Examples for the Scaling of minimum of a number of random variables

This section includes simulation results for different random distributions when applied to (2.7). The following are a few of the distributions tested

#### Uniform Distribution

Let  $x$  be a uniform random variable,  $x \in (a, b)$ . In this simulation we take  $a = 3$  and  $b = 6$ . The PDF is defined as  $f(x) = \frac{1}{b-a}$ . In this case we find  $i_0 = 1$ , with  $F^{i_0}(a) = \frac{1}{b-a}$ .

#### U-Quadratic Distribution

Let  $x$  be a U-Quadratic random variable,  $x \in [a, b]$ . The CDF is defined as  $F(x) = \frac{\alpha}{3}((x - \beta)^3 + (\beta - a)^3)$ , where  $\alpha = \frac{12}{(b-a)^3}$  and  $\beta = \frac{b+a}{2}$ . In this case we find  $i_0 = 1$ , with  $F^{i_0}(a) = \frac{3}{b-a}$ .



### **Non-Central Chi-Square Distribution**

Let  $x$  be the non central random variable with  $n = 4$ , i.e. sum of 4 non central gaussian variables with mean  $m_x = 1$ . In this case  $i_0 = 1$  and  $F^{i_0}(0) = 1/4 e^{-2 \frac{m_x^2}{\sigma^2}} \sigma^{-4}$ .

### **Chi- Square Distribution**

This distribution is taken same as that of the noncentral distribution but with mean  $m_x = 0$ . Therefore,  $F^{i_0}(0) = 1/4\sigma^{-4}$ .

### **Irwin Hall Distribution**

Irwin Hall random variable is created by summing uniform random variables. In this simulation we are taking the sum of  $M = 15$  uniform random variable. Here  $i_0 = M$  and  $F^{i_0}(a) = 1$

### **Gamma Distribution**

Here  $a = 0$ ,  $i_0 = k$  and  $F^{i_0}(0) = \theta^k$ , where  $\theta$  is the scale parameter and  $k$  is the shape parameter.

### **Rayleigh Distribution**

Let  $x$  follow rayleigh distribution. In this case  $i_0 = 2$  and  $F^{i_0}(0) = \frac{1}{\sigma^2}$

### **Pareto Distribution**

Here  $i_0 = 1$  and  $F^{i_0}(x_m) = \frac{k}{x_m}$

### Log Logistic Distribution

Here  $i_0 = 1$  with  $F^{i_0}(0) = 1$

### Half Normal Distribution

With  $i_0 = 1$  and  $F^{i_0}(0) = \frac{1}{\sigma} \sqrt{\frac{2}{\pi}}$

### Folded Normal Distribution

With  $i_0 = 1$  and  $F^{i_0}(0) = \frac{1}{\sigma} \sqrt{\frac{2}{\pi}} e^{-\frac{\mu^2}{2\sigma^2}}$

### Kumaraswamy Distribution

Let  $x$  be Kumaraswamy random variable  $x \in [0, 1]$ , With  $i_0 = a$  and  $F^{i_0}(0) = ab(a-1)!$

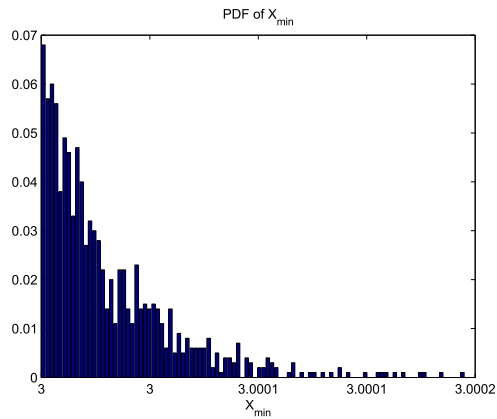
### Beta Distribution

Let  $x$  be Beta random variable  $x \in [0, 1]$ , With  $i_0 = \alpha$  and  $F^{i_0}(0) = \frac{(\alpha-1)!}{B(\alpha, \beta)}$

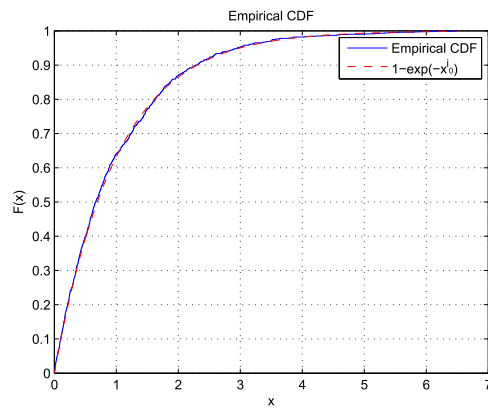
### 2.3.5 Scaling of $\|h_i\|^2$

Considering  $h_i$  to be a complex gaussian zero mean random variable with  $M$  sums each having different variances, in this case we have  $i_0 = M$  and  $F^{i_0}(0) = \frac{1}{2^M \prod_{i=1}^M \sigma_i^2}$ . Figure 2.14 shows the simulation results for  $M = 1$  with  $n = 1000$

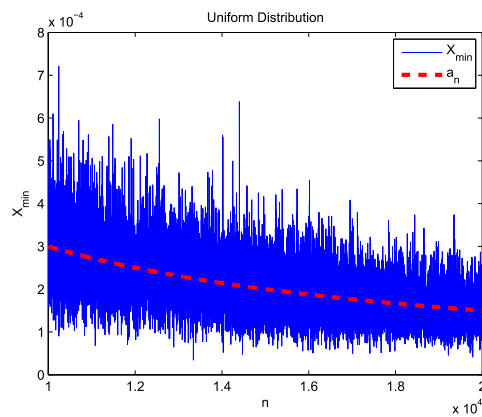
Lets again consider the case with  $M = 5$  each with different variances. As  $i_0$  increases so does the error, so we have to increase  $n$  accordingly. Figure 2.15 shows the simulation result for  $M = 5$  and  $n = 10000000$



(a) PDF of  $X_{min}$  for Uniform Distribution

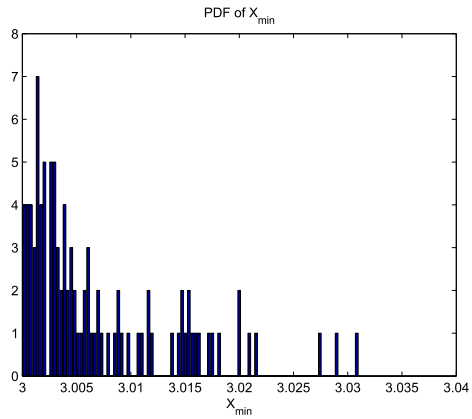


(b) Comparison of Empirical and Theoretical CDF

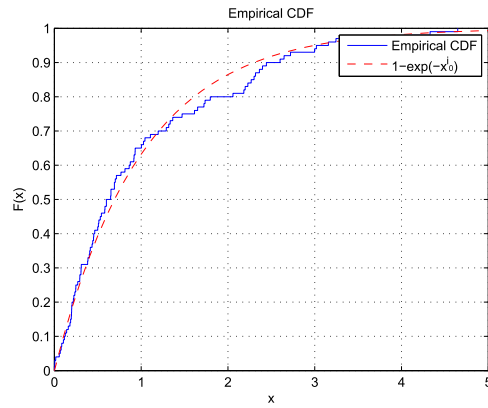


(c)  $X_{min}$  vs  $n$

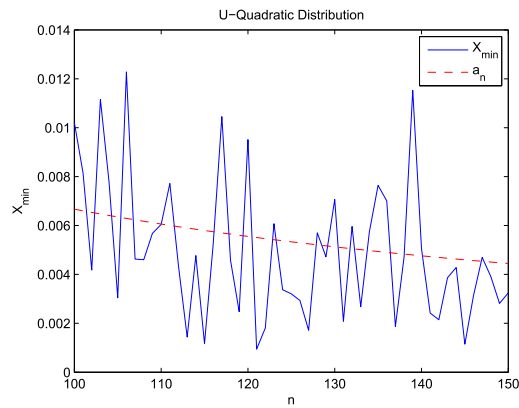
Figure 2.1: Uniform Distribution



(a) PDF of  $X_{min}$  for U- Quadratic Distribution

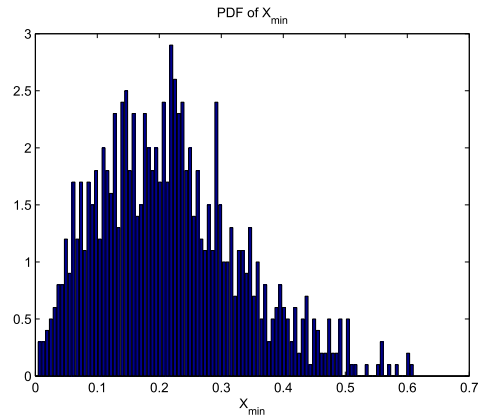


(b) Comparison of Empirical and Theoretical CDF

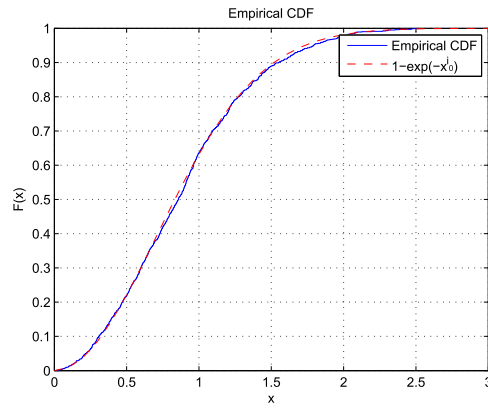


(c)  $X_{min}$  vs  $n$

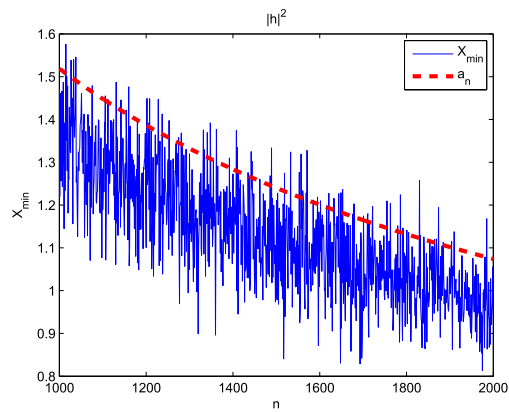
Figure 2.2: U-Quadratic Distribution



(a) PDF of  $X_{min}$  for Non-Central Chi Square Distribution

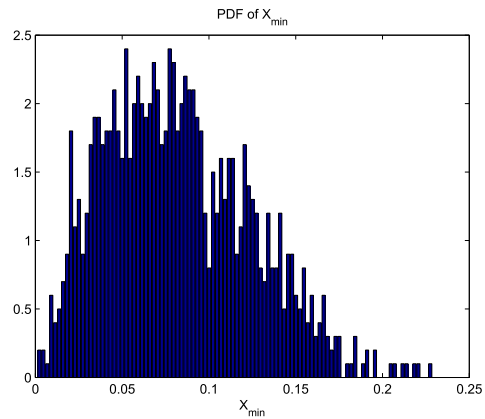


(b) Comparison of Empirical and Theoretical CDF

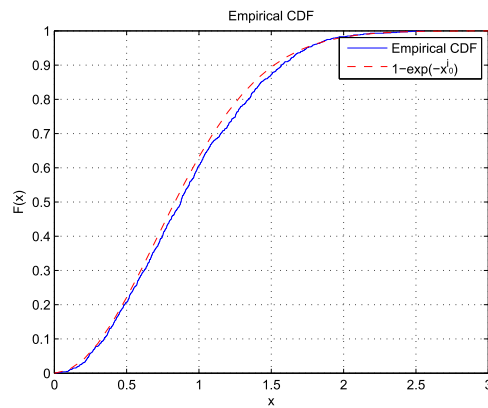


(c)  $X_{min}$  vs  $n$

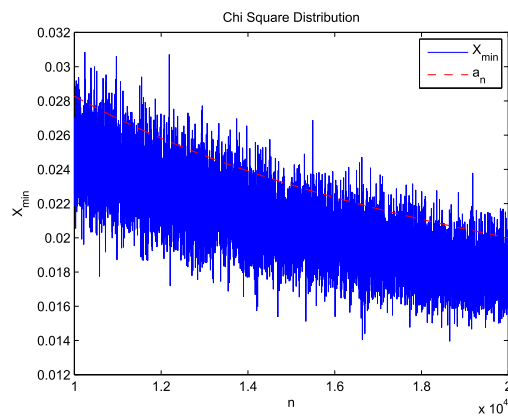
Figure 2.3: Non-Central Chi Square Distribution



(a) PDF of  $X_{min}$  for Chi Square Distribution

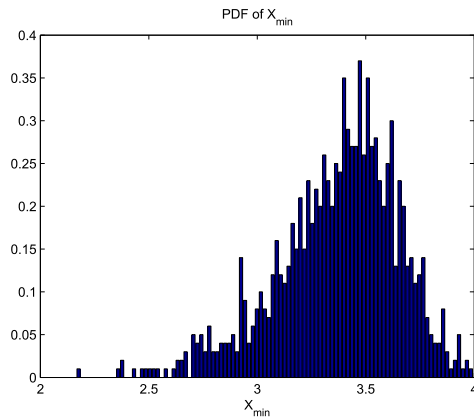


(b) Comparison of Empirical and Theoretical CDF

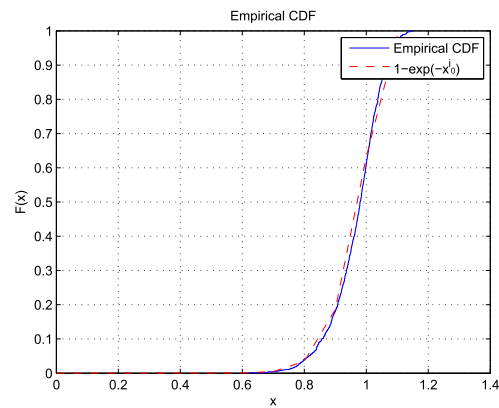


(c)  $X_{min}$  vs  $n$

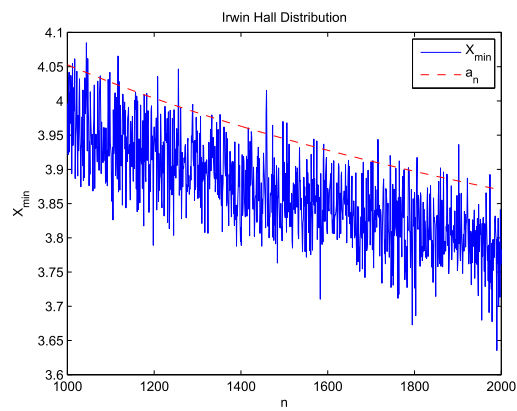
Figure 2.4: Chi Square Distribution



(a) PDF of  $X_{min}$  for Irwin Hall Distribution

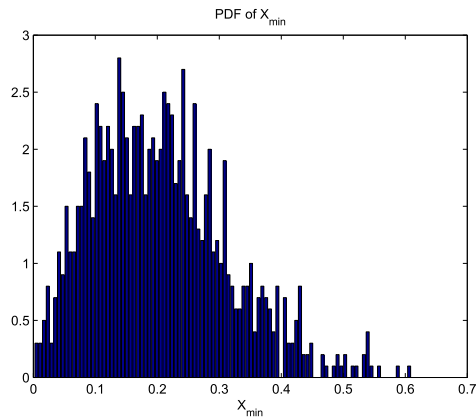


(b) Comparison of Empirical and Theoretical CDF

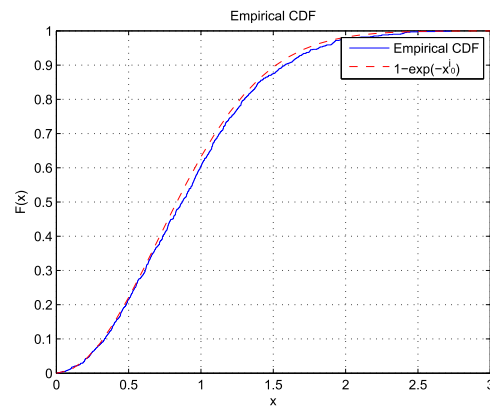


(c)  $X_{min}$  vs  $n$

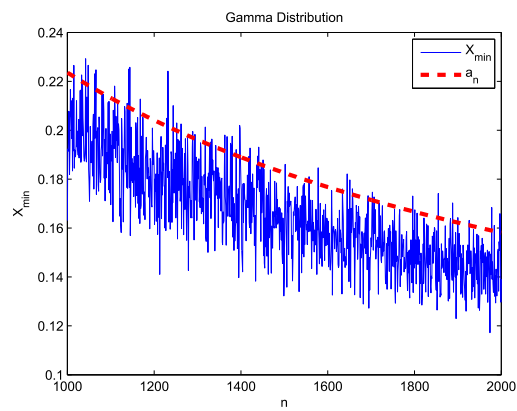
Figure 2.5: Irwin Hall Distribution



(a) PDF of  $X_{min}$  for Gamma Distribution



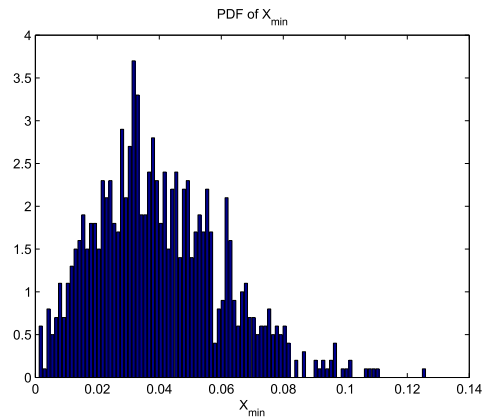
(b) Comparison of Empirical and Theoretical CDF



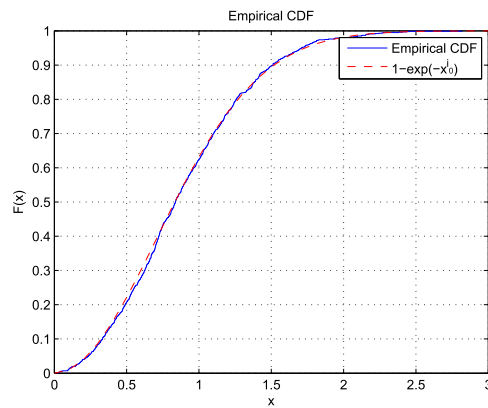
(c)  $X_{min}$  vs  $n$

Figure 2.6: Gamma Distribution

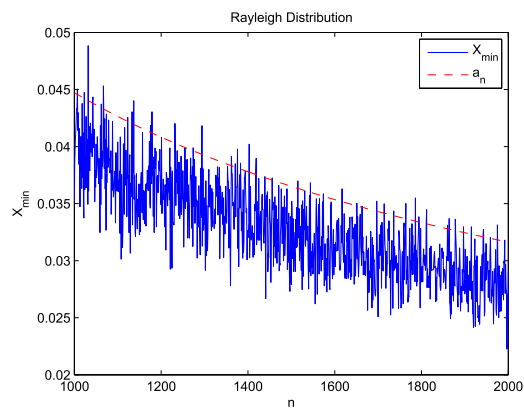




(a) PDF of  $X_{min}$  for Rayleigh Distribution

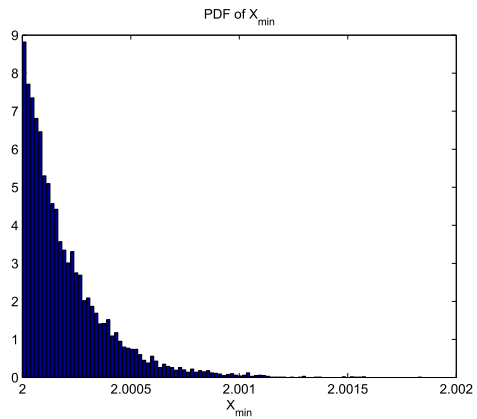


(b) Comparison of Empirical and Theoretical CDF

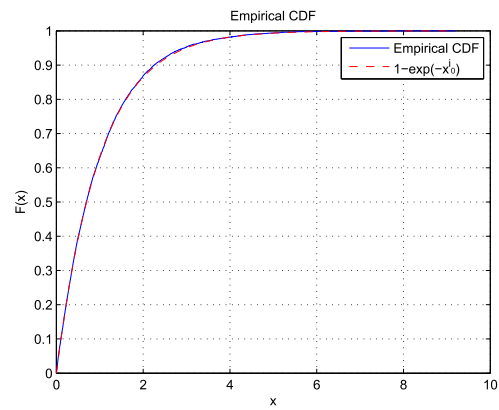


(c)  $X_{min}$  vs  $n$

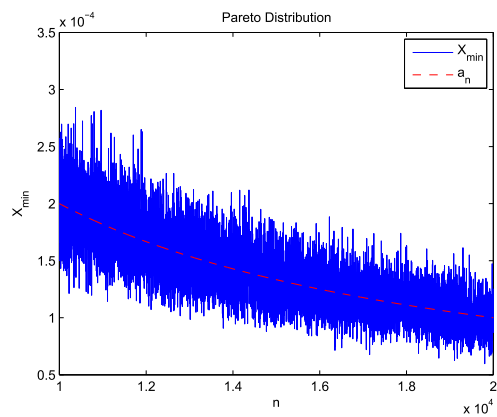
Figure 2.7: Rayleigh Distribution



(a) PDF of  $X_{min}$  for Pareto Distribution

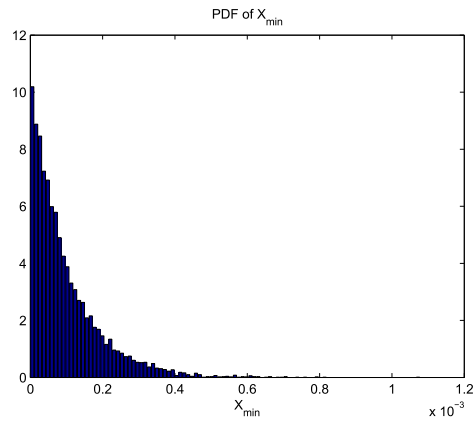


(b) Comparison of Empirical and Theoretical CDF

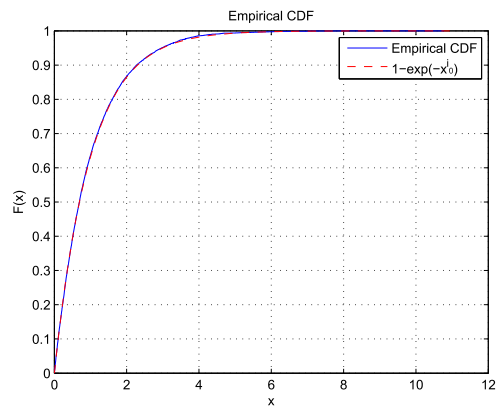


(c)  $X_{min}$  vs  $n$

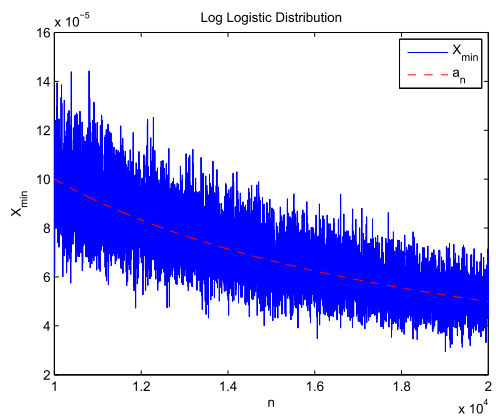
Figure 2.8: Pareto Distribution



(a) PDF of  $X_{min}$  for Log Logistic Distribution

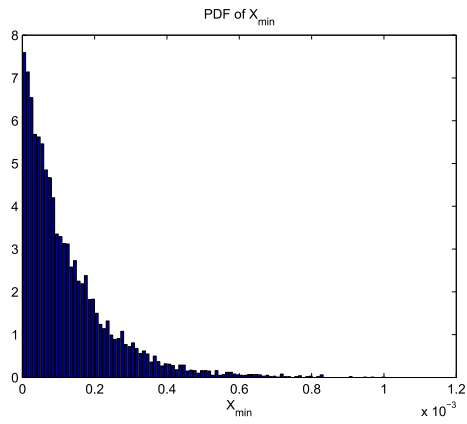


(b) Comparison of Empirical and Theoretical CDF

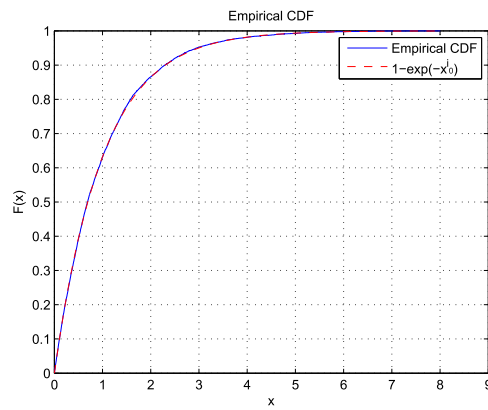


(c)  $X_{min}$  vs  $n$

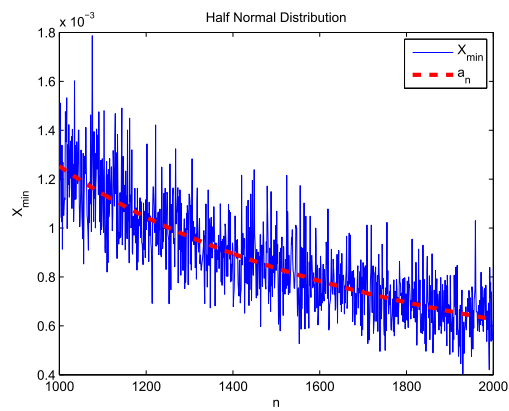
Figure 2.9: Log Logistic Distribution



(a) PDF of  $X_{min}$  for Half Normal Distribution

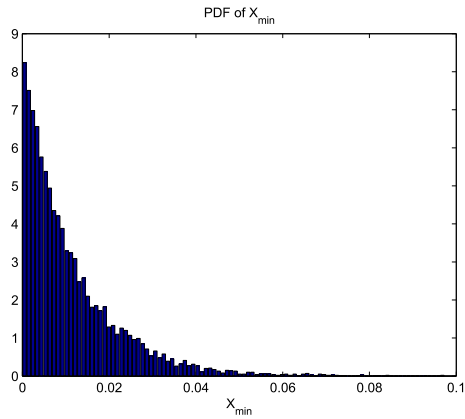


(b) Comparison of Empirical and Theoretical CDF

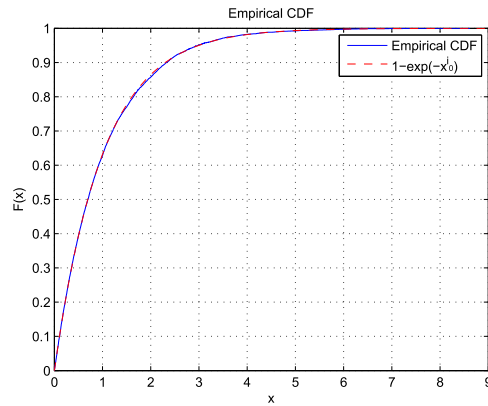


(c)  $X_{min}$  vs  $n$

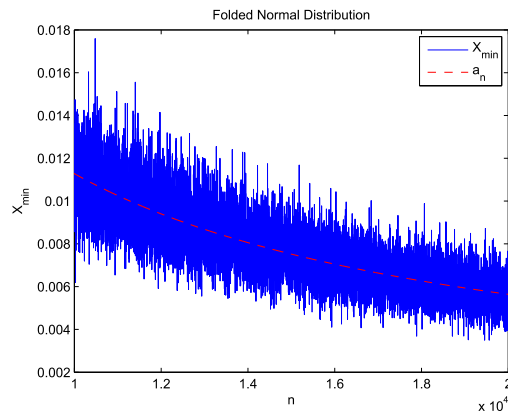
Figure 2.10: Half Normal Distribution



(a) PDF of  $X_{min}$  for Folded Normal Distribution

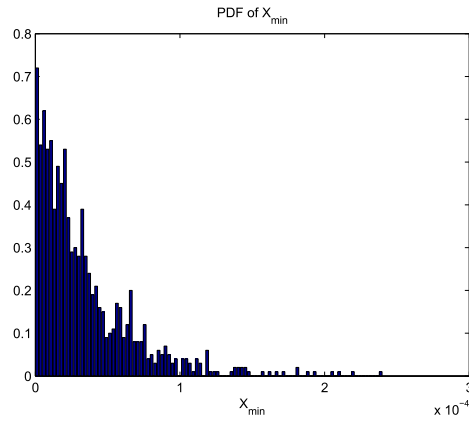


(b) Comparison of Empirical and Theoretical CDF

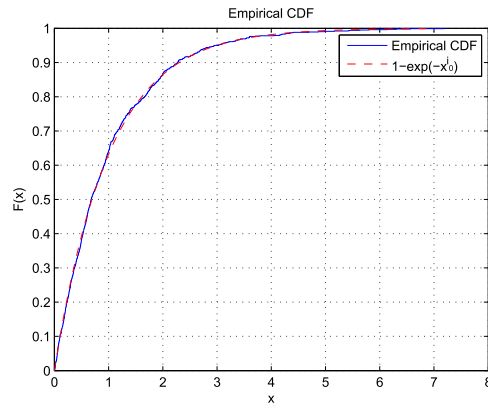


(c)  $X_{min}$  vs  $n$

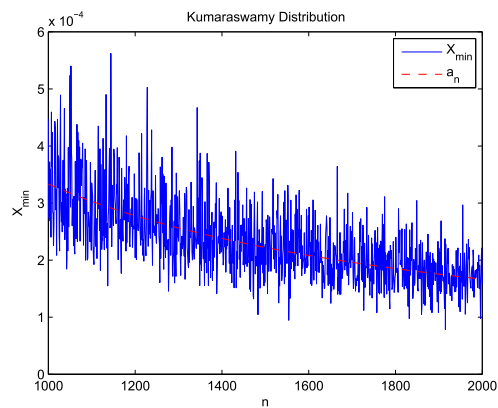
Figure 2.11: Folded Normal Distribution



(a) PDF of  $X_{min}$  for Kumaraswamy Distribution

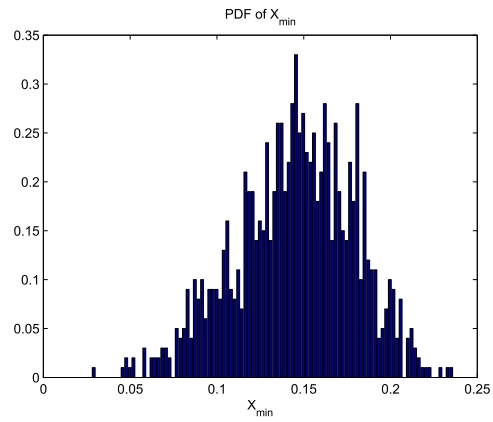


(b) Comparison of Empirical and Theoretical CDF

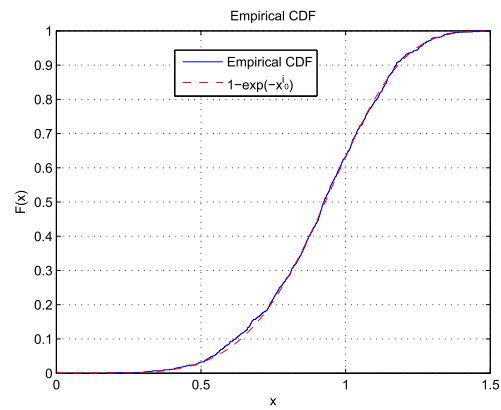


(c)  $X_{min}$  vs  $n$

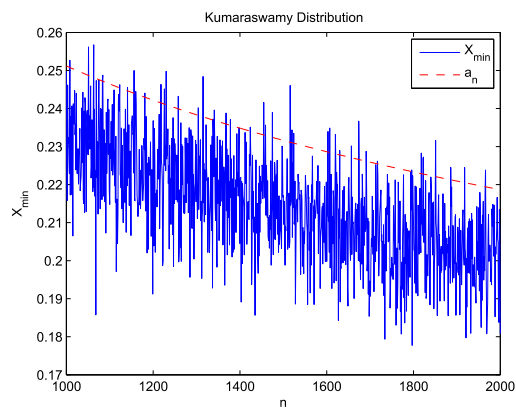
Figure 2.12: Kumaraswamy Distribution



(a) PDF of  $X_{min}$  for Beta Distribution

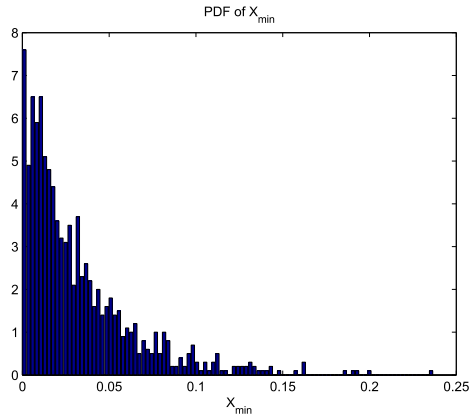


(b) Comparison of Empirical and Theoretical CDF

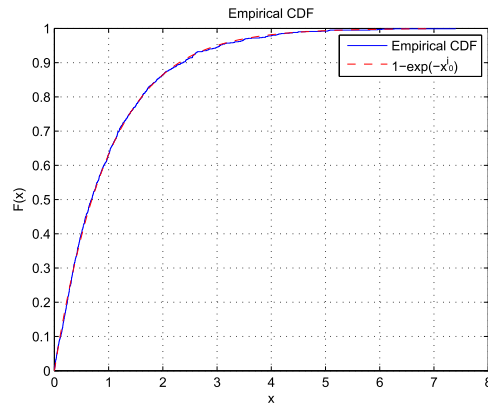


(c)  $X_{min}$  vs  $n$

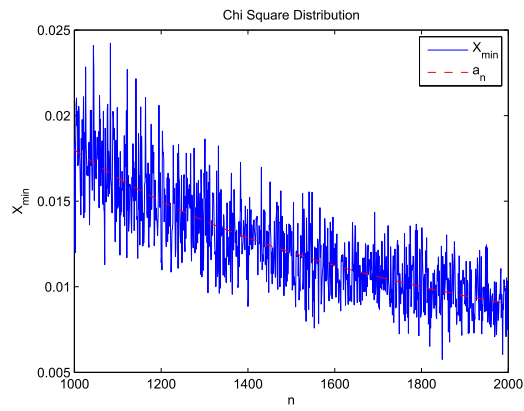
Figure 2.13: Beta Distribution



(a) PDF of  $X_{min}$  for  $\|h_i\|^2$



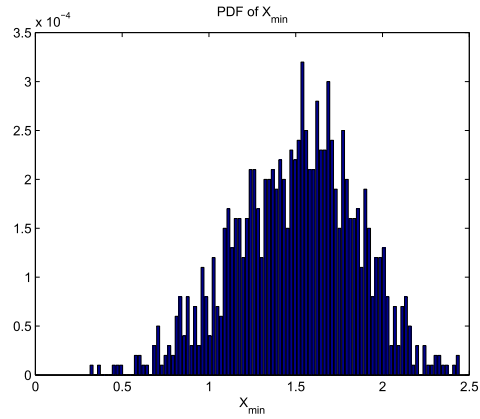
(b) Comparison of Empirical and Theoretical CDF



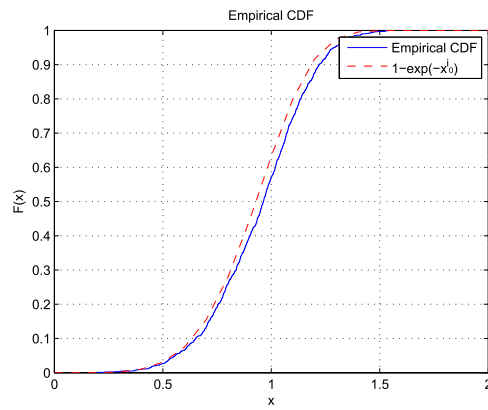
(c)  $X_{min}$  vs  $n$

Figure 2.14:  $\|h_i\|^2$

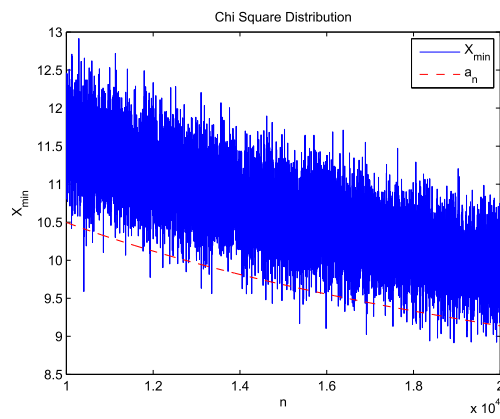




(a) PDF of  $X_{min}$  for  $\|h_i\|^2$



(b) Comparison of Empirical and Theoretical CDF



(c)  $X_{min}$  vs  $n$

Figure 2.15:  $\|h_i\|^2$ , non-central and with different variances

### 2.3.6 Scaling for Maximum

We can easily find the scaling for the maximum of a random variable by assuming  $U = -x$  and then simply finding the minimum of  $U$ , here we take  $b$  to be the supremum of  $x$ . It is important to note that  $f_U(u) = f_X(-x)$  and  $\phi_u(s) = \phi_x(-s)$ , here we have the following

$$\lim_{n \rightarrow \infty} F_{\min}(a_n u + b) = 1 - \exp(-u^{i_0}) \quad (2.10)$$

where

$$a_n = \frac{i_0!^{\frac{1}{i_0}}}{F^{(i_0)}(a)^{\frac{1}{i_0}}} \frac{1}{n^{\frac{1}{i_0}}}$$

The result here is supported by the following simulations,

#### Uniform Distribution

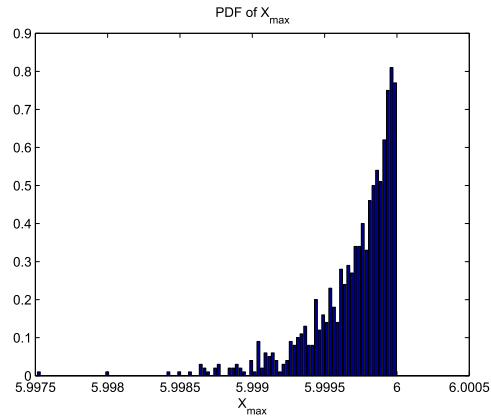
Here  $a = 3$ ,  $b = 6$ ,  $i_0 = \alpha$  and  $F^{i_0}(0) = \frac{(\alpha-1)!}{B(\alpha, \beta)}$

#### U-Quadratic Distribution

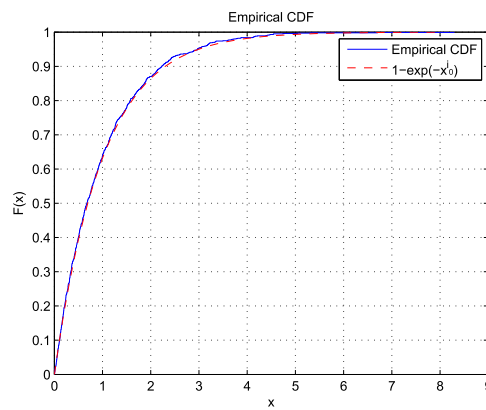
Here  $b = 5$ ,  $i_0 = 1$ , with  $F^{i_0}(a) = \frac{3}{b-a}$ .

#### Irwin Hall Distribution

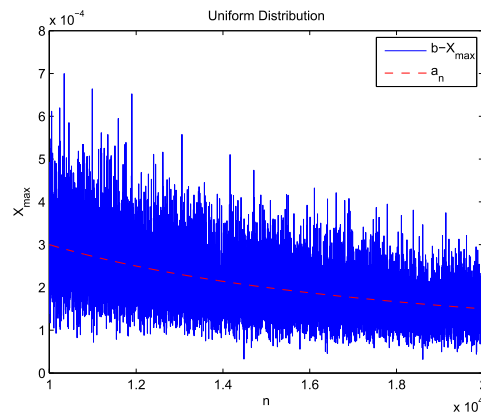
Here  $b = M$ ,  $i_0 = \alpha$  and  $F^{i_0}(0) = \frac{(\alpha-1)!}{B(\alpha, \beta)}$



(a) PDF of  $X_{max}$  for Uniform Distribution



(b) Comparison of Empirical and Theoretical CDF



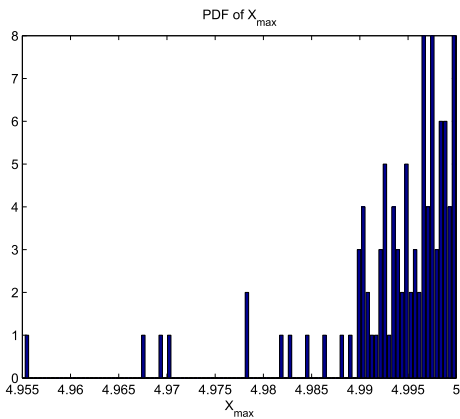
(c)  $b - X_{max}$  vs  $n$

Figure 2.16: Uniform Distribution

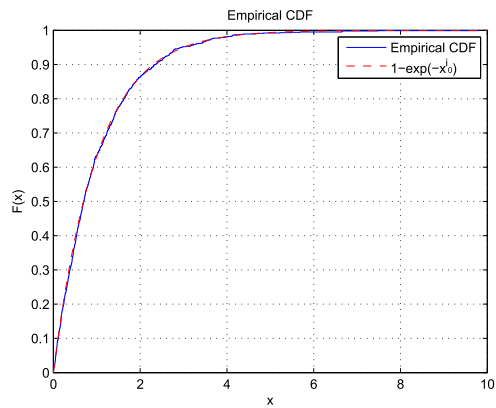
## Beta Distribution

Here we are taking the sum of  $M = 15$  uniform random variable with  $i_0 = M$  and

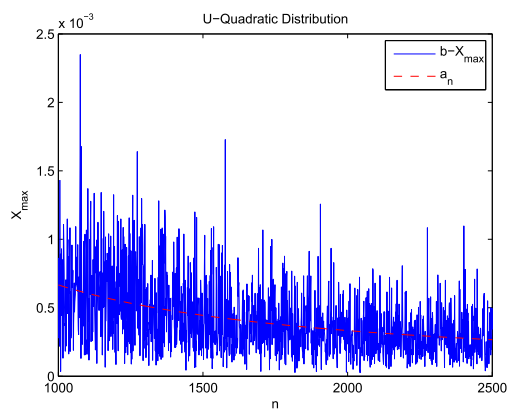
$$F^{i_0}(a) = 1$$



(a) PDF of  $X_{max}$  for U-Quadratic Distribution

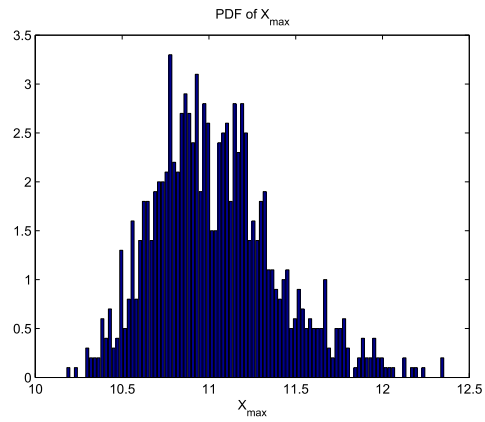


(b) Comparison of Empirical and Theoretical CDF

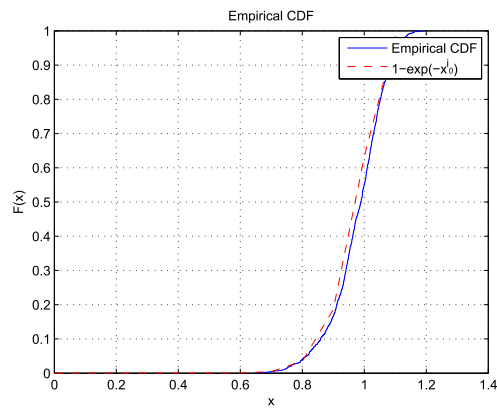


(c)  $b - X_{max}$  vs  $n$

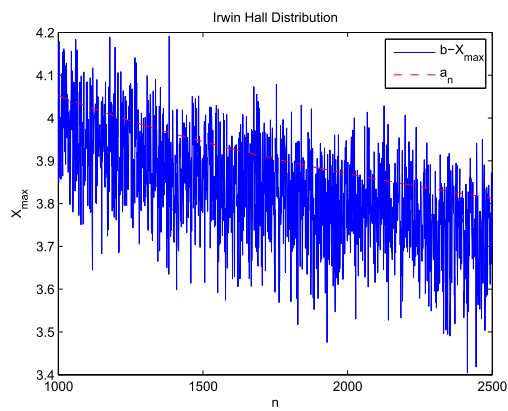
Figure 2.17: U-Quadratic Distribution



(a) PDF of  $X_{max}$  for Irwin Hall Distribution

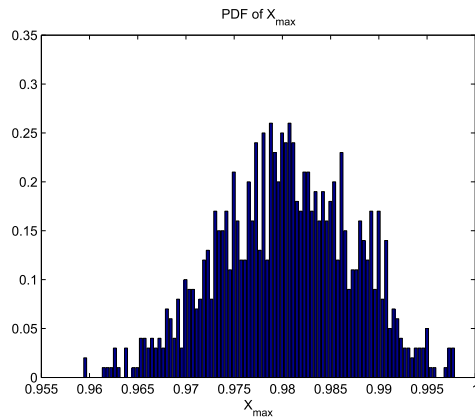


(b) Comparison of Empirical and Theoretical CDF

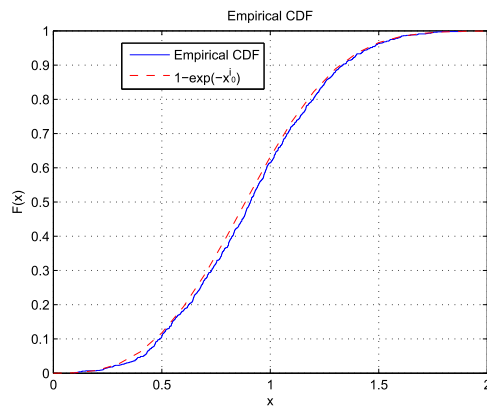


(c)  $b - X_{max}$  vs  $n$

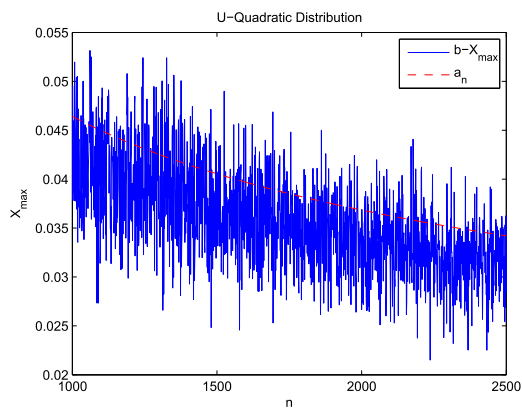
Figure 2.18: Irwin Hall Distribution



(a) PDF of  $X_{max}$  for Beta Distribution



(b) Comparison of Empirical and Theoretical CDF



(c)  $b - X_{max}$  vs  $n$

Figure 2.19: Beta Distribution

The following table list the PDF, CDF, characteristic function, value of  $i_0$  and

$F^{i_0}(0)$

Table 2.1: Scaling Table for minimum of different bounded distributions

	$f(x)$	$F(x)$	$\phi(s)$	$i_0$	$F^{i_0}(0)$
Uniform Dist.	$\frac{1}{b-a}$			1	$\frac{1}{b-a}$
U-Quadratic Dist.	$\alpha(x - \beta)^2$	$\frac{\alpha}{3}((x - \beta)^3 + (\beta - \alpha)^3)$		1	$\frac{3}{b-a}$
Non-Central Chi Square Dist.			$\frac{1}{(1-2\sigma^2s)^{n/2}} e^{\frac{\sum_{i=1}^n m_i^2}{1-2\sigma^2s}}$	$n/2$	$\frac{1}{2^{n/2}\sigma^n} e^{-\frac{\sum_{i=1}^n m_i^2}{2\sigma^2}}$
Chi-Square Dist.			$\frac{1}{(1-2\sigma^2s)^{n/2}}$	$n/2$	$\frac{1}{2^{n/2}\sigma^2}$
Irwin Hall Dist.				M	1
Gamma Dist.	$x^{k-1} \frac{e^{-x/\theta}}{\Gamma(k)\theta^k}$	$\frac{\gamma(k, x/\theta)}{\Gamma(k)}$	$(1 - \theta s)^{-k}$	$k$	$\theta^{-k}$
Rayleigh Dist.	$\frac{x e^{-x^2/2\sigma^2}}{\sigma^2}$	$1 - e^{-x^2/2\sigma^2}$		2	$\frac{1}{\sigma^2}$
Pareto Dist.	$\frac{kx_m^k}{x^{k+1}}$	$1 - \left(\frac{x_m}{x}\right)^k$		1	$\frac{k}{x_m}$
Log-Logistic Dist.	$\frac{(\beta/\alpha)(x/\alpha)^{\beta-1}}{(1+(x/\alpha)^\beta)^2}$	$\frac{1}{1+(x/\alpha)^{-\beta}}$		1	1
Half Normal Dist.		$\int_0^x \frac{1}{\sigma} \sqrt{\frac{2}{\pi}} e^{-y^2/2\sigma^2} dy$		1	$\frac{1}{\sigma} \sqrt{\frac{2}{\pi}}$
Folded Normal Dist.		$\int_0^x \frac{1}{\sigma\sqrt{2\pi}} e^{-(-y-\mu)^2/2\sigma^2} dy + \int_0^x \frac{1}{\sigma\sqrt{2\pi}} e^{(-y-\mu)^2/2\sigma^2} dy$		1	$\frac{1}{\sigma} \sqrt{\frac{2}{\pi}} e^{-\mu^2/2\sigma^2}$
Kumaraswamy Dist.	$abx^{a-1}(1-x^a)^{b-1}$			a	$ab(a-1)!$
Beta Dist.	$\frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)}$			$\alpha$	$\frac{(\alpha-1)!}{B(\alpha, \beta)}$

**Example 1 Revisited** Let's apply the technique just derived to the Example 1

above. Here

$$F_{\text{SINR}}(x) = 1 - \frac{e^{-\frac{x}{\rho}}}{(1+x)^M}$$

$$F'_{\text{SINR}}(x) = f(x) = \frac{e^{-\frac{x}{\rho}}}{(1+x)^M} \left( \frac{1}{\rho}(1+x) + M - 1 \right)$$

Thus,

$$F'_{\text{SINR}}(0) = f_{\text{SINR}}(0) = \left( \frac{1}{\rho} + M - 1 \right)$$

So, the scaling of the minimums SINR is given by

$$\begin{aligned} \min_i \text{SINR}_i &= \frac{C_1}{F^{(1)}(0)} \frac{1}{n} \\ &= \frac{1}{\frac{1}{\rho} + M - 1} \frac{1}{n} \end{aligned} \quad (2.11)$$

which coincides with the result obtained the direct way.

Applying this theorem to find the scaling of  $\min_i \|h_i\|^2$  is equally difficult because we need to find the CDF of  $\|h_i\|^2$  before we actually evaluate its derivatives.

**Example 2: Scaling of spatially correlated channel norms** Let's find the scaling law for  $\min_{h_i} \|h_i\|^2$  when  $h_i$  are iid  $CN(0, R)$ . The pdf and CDF of  $\|h_i\|^2$  will both have different forms depending on whether some of the eigenvalues  $\lambda_l$  of  $R$  are the same or different, and so the direct method for scaling can be quite challenging. On the other hand, the characteristic function takes one form and is given by

$$\phi(s) = \prod_{l=1}^M \frac{1}{1 + \lambda_l s}$$

From this, it is easy to see that

$$\lim_{s \rightarrow \infty} s^i \phi(s) = F^{(i)}(0) = 0 \quad \text{for } i < M$$



and that

$$\lim s^M \phi(s) = F^{(M)}(0) = \frac{1}{\prod_{l=1}^M \lambda_l} = \frac{1}{\det(R)}$$

We thus conclude that

$$\boxed{\min_i \|h_i\|^2 \text{ scales as } C_M \det(R)^{\frac{1}{M}} \frac{1}{n^{\frac{1}{M}}}}$$

### 2.3.7 Chapter Conclusion

This chapter develops a new method for characterizing the scaling of the minimum of the bounded i.i.d. random variables. This method presents a tool for Extreme Value Theory (EVT) to scale the minimum directly from the characteristic function. And we used this method to scale several random variables. The technique developed in this chapter will also be evoked in the following chapter.

## CHAPTER 3

# UPLINK SCHEDULING FOR SPATIAL MULTIPLEXING SYSTEMS

We study in this chapter uplink scheduling for MIMO users. The scheduler selects one user at a time based on a certain criterion that depends on the detection algorithm. Each user spatially multiplexes his data over the transmit antennas. This spatial multiplexing (SM) scheme provides high data rates while a multi-user diversity obtained from scheduling improves the performance of the uplink system. Vertical-Bell-Labs LAYERed Space Time architecture (V-BLAST) is a practical SM scheme that uses a serial nulling and cancellation algorithm to detect each layer. Each set of data transmitted from each antenna is called a layer. The main results of this study show that the scheduler that maximizes the optimal MIMO capacity doesn't work well for a V-BLAST system. Instead, we find a scheduler

that maximizes the V-BLAST capacity which is derived specifically from the V-BLAST detection algorithm. Furthermore, we investigate suboptimal schedulers and their performances. We also define some bounds on capacity for the V-BLAST system. In addition, we look into scheduling for SM with sphere decoding and we find that in this case, using MIMO capacity as the scheduling criterion is the best option.

### **3.1 Introduction**

Until recently, most of the studies on MIMO techniques were focused on optimizing the physical layer. However, in a multiuser environment, optimizing the physical layer for each user doesn't necessary optimize the system performance nor it takes advantage of the statistical independence of the fading channels among the users. Furthermore, different users have different needs in terms of data rates, power limits and Quality of Service (QoS). These requirements make scheduling an important technique for optimizing the performance of a communication system and utilizing the system resources efficiently. Since the physical link of each user is a fading channel, and it is usually independent from one user to another, scheduling transmission to the best user at a time leads to a form of selection diversity known as multiuser diversity. In general, schedulers are designed to maximize system throughput and/or capacity and minimize error rates, but they should also provide fairness to users and minimize packet delays. In addition, it is used to minimize interference, adapt to traffic loads, and satisfy

a quality-of-service (QoS) requirement.

Scheduling is sometimes classified under radio resource management, cross-layer optimization or multiuser diversity. In single-input single-output (SISO) systems, where the base station and each mobile have one antenna only, it has been shown that selecting the user who has the maximum signal to noise ratio (MaxSNR) maximizes the total information capacity of the uplink system [34]. Similar results are also found for the downlink from the base station to the mobile unit [39]. This scheduler is known as MaxSNR scheduling. Over MIMO channels, most of the studies are based on theoretical information capacity [33], [30], [37] and on the downlink, which is the broadcast channel from the base station to the mobile unit. For example, it has been shown in [32] that space time block coding (STBC) and scheduling aren't a good match. In fact, scheduling to a user with single antenna outperforms scheduling using STBC. The reason is that STBC averages the fades while the scheduler tends to benefit from high peaks in the fading channel. In addition, multiuser diversity obtained from scheduling is much higher than the spatial diversity of STBC, so STBC diversity doesn't add much benefit. On the other hand, spatial multiplexing (SM) schemes match perfectly with scheduling. This is because they provide high data rates while the scheduler compensates for the lack of diversity by providing multiuser selection diversity.

In a MIMO system, scheduling could be done to a single user or multiple

users. Scheduling to multiple users, i.e allowing more than one user to transmit or receive at the same time, is shown to be optimal in terms of maximizing system capacity and throughput. In [33], downlink scheduling to multiple users improved the average throughput compared to a single user scheduling. Furthermore, the optimal uplink MIMO scheduling based on an information theoretical approach was considered in [35]. They showed that it should allocate all the power to at most  $M_R$  users, where  $M_R$  is the number of receive antennas at the base station. Also, they found that the optimal power resource allocation is water-filling in space and time. In [30], the authors found that multiuser scheduling reduces the average delay experienced by the users compared to single-user scheduling. In [38], the scheduler selects  $K$  users at a time and it cycles through groups of users in a round robin (RR) fashion. Thus, it provides diversity through multiple antennas while it insures fairness through RR scheduling.

In this chapter, we investigate scheduling for uplink V-BLAST [40] users. V-BLAST is a practical spatial multiplexing (SM) MIMO system. We focus on single-user scheduling. Although it is not optimal, it is more practical and easily implemented. The scheduler selects one user at a time based on a certain criterion. Each user spatially multiplexes their data over the transmit antennas to provide high data rates while the multiuser diversity obtained from scheduling improves the performance of the uplink system. Our main contribution in this work is finding the capacity maximization scheduling criteria for V-BLAST

uplink users. Based on the V-BLAST system we also define some bounds on capacity. Also, we show that the optimal MIMO capacity criterion doesn't work well for V-BLAST. The V-BLAST maximum capacity scheduler is derived specifically from its detection algorithm. As such, we investigate the performance of suboptimal scheduling criteria that are based on the MIMO channel matrix directly. In addition, we look into scheduling for spatial multiplexing with sphere decoding and in this case shows that scheduling based on maximum MIMO capacity provides the best performance.

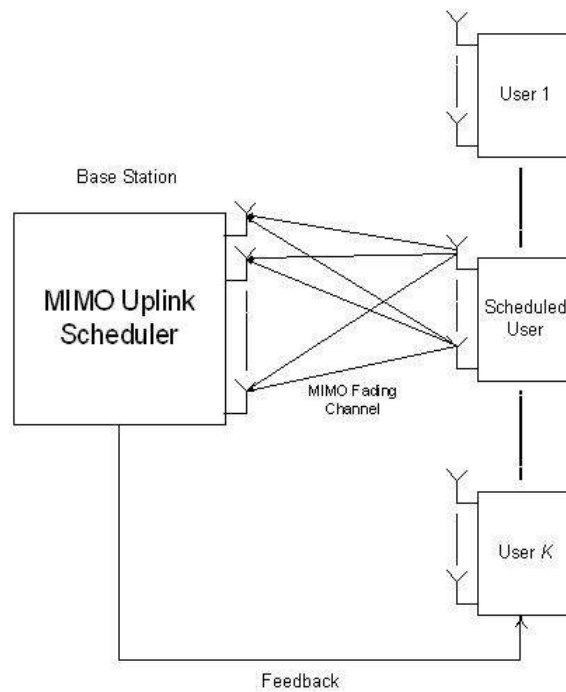


Figure 3.1: Uplink MIMO Scheduling

## 3.2 System Model

We consider scheduling a single user at a time. The average SNR is assumed to be the same for all users; they are either at similar distance or they use power control. There are  $K$  users and each user transmits through  $M_T$  transmit antennas and the receiver has  $M_R$  receive antennas ( $M_R \geq M_T$ ). The MIMO channel is assumed to be an independent Rayleigh flat fading MIMO channel where each coefficient is an i.i.d complex Gaussian random variable with zero mean and unit variance. The received signal from user  $k$  is:

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \eta_k \quad (3.1)$$

where  $\mathbf{y}_k$  is an  $M_R \times 1$  received vector,  $\mathbf{H}_k$  is an  $M_R \times M_T$  MIMO channel matrix for the  $k^{th}$  user,  $\mathbf{x}_k$  is an  $M_T \times 1$  transmitted symbols from user  $k$ , and  $\eta_k$  is an  $M_R \times 1$  i.i.d complex AWGN vector of zero mean and variance  $N_0/2$  per dimension.

## 3.3 Optimal MIMO scheduling

Assuming that an optimal MIMO encoder/ decoder is available, the optimal MIMO scheduler selects the user which has a channel matrix that maximizes

the MIMO capacity:

$$C_{max} = \max_{k=1,2,\dots,K} C_k; \text{ where} \quad (3.2)$$

$$C_k = \log_2(\det(\mathbf{I}_{M_R} + \frac{SNR}{M_T} \mathbf{H}_k \mathbf{H}_k^H)) \quad (3.3)$$

where  $\mathbf{I}_{M_R}$  is the identity matrix and  $\mathbf{A}^H$  is the conjugate-transpose (Hermitian) of  $\mathbf{A}$ . Assuming that  $M_T \leq M_R$ , we can write:

$$C_k = \sum_{n=1}^{M_T} \log_2 \left( 1 + \frac{SNR}{M_T} \lambda_n(\mathbf{H}_k) \right) \quad (3.4)$$

where  $\lambda_n(\mathbf{H}_k)$  is the  $n^{th}$  eigenvalue of  $\mathbf{H}_k \mathbf{H}_k^H$ . The maximum capacity is achieved when the channel is orthogonal [31] and in this case,  $\lambda_n(\mathbf{H}_k) = \|\mathbf{H}_k\|_F^2$ , which correspondences to the squared Frobenius norm. The resulting MIMO capacity will then be:

$$C_k = M_T \log_2 \left( 1 + \frac{SNR}{M_T} \|\mathbf{H}_k\|_F^2 \right) \quad (3.5)$$

### 3.4 V-BLAST Scheduling

V-BLAST [40] is a practical MIMO architecture that spatially multiplexes transmitted data over multiple transmit antennas. Data transmitted from each antenna is called a layer of information. At the receiver, a serial interference nulling and cancellation algorithm is used to detect each layer. Although V-BLAST is a full spatial multiplexing scheme, it has poor energy performance



because of the lack of spatial diversity. Therefore, it makes a good match with scheduling since multiuser diversity obtained improves the performance significantly.

This section proposes and evaluates scheduling algorithms for uplink V-BLAST users. It has been shown that in single antenna systems, selecting the user that has the maximum SNR (MaxSNR) is optimal [34] and it maximizes the system information capacity. However, the MaxSNR scheduler is not optimal for V-BLAST users as will be seen later from the simulation results. The reason is that, unlike SISO systems, the received SNR ( $trace(\mathbf{H}_k \mathbf{H}_k^H)$ ) doesn't reflect directly the capacity of MIMO systems. Moreover, scheduling based on maximization of MIMO channel capacity as in (3.3) is also not optimal for V-BLAST since its detection algorithm is suboptimal. Since V-BLAST is an open loop system and all layers have the same rate, an outage in capacity will occur if an outage happens in at least one layer. Therefore, the V-BLAST capacity of the  $k^{th}$  user is dominated by the weakest layer and it is given by [36]:

$$C_k^{ZF-VBLAST} = M_T \min_{n=1,2,\dots,M_T} \left\{ \log_2 \left( 1 + \frac{SNR}{M_T \|W_{ZF,n}^k\|^2} \right) \right\} \quad (3.6)$$

where  $W_{ZF,n}^k$  is the ZF projection row for the  $n^{th}$  layer of the  $k^{th}$  user and  $M_T$  is the number of layers (transmit antennas). Recall that the V-BLAST detector performs a series of interference nulling and cancellation operations. At the  $n^{th}$  stage, let  $\mathbf{H}_{k,n}$  be the MIMO channel matrix after canceling the  $n - 1$  detected

layers. The order of  $\mathbf{H}_{k,n}$  is  $M_R \times M_T - n + 1$ . Consequently, the ZF matrix at this stage is

$$W_{ZF} = (\mathbf{H}_{k,n}^H \mathbf{H}_{k,n})^{-1} \mathbf{H}_{k,n}^H \quad (3.7)$$

Based on the V-BLAST ordering criteria, the strongest layer for detection at this stage is the one with:

$$\|W_{ZF,n}^k\|^2 = \min \left( \text{diag} \left( [\mathbf{H}_{k,n}^H \mathbf{H}_{k,n}] \right)^{-1} \right) \quad (3.8)$$

and the corresponding post-processing SNR is given by,

$$SNR_{k,n}^{ZF} = \frac{SNR}{M_T \|W_{ZF,n}^k\|^2} \quad (3.9)$$

Therefore, the layer that determines the capacity of V-BLAST is the one with largest norm of the ZF projection row. Let  $w_k = \max_{n=1,2,\dots,M_T} \|W_{ZF,n}^k\|^2$  denote the largest projection value for user  $k$ , then the scheduler that maximizes the V-BLAST capacity will select the user with minimum  $w_k$ . In other words, the capacity maximization scheduling for V-BLAST selects the user with largest post-processing SNR of his weakest layer. However, this scheduler needs to perform ZF nulling and ordering to the channel matrices of all users before selecting the best user, which requires a lot of computations. Therefore, we examine other suboptimal schedulers that are based on the received MIMO channels before V-BLAST processing. These schedulers don't take into account the V-BLAST structure. The first one chooses the user with the largest MIMO channel power ( $\text{trace}(\mathbf{H}_k \mathbf{H}_k^H)$ )

and we will refer to it as MaxSNR scheduler, which mimics the optimal scheduler for single antenna systems. The next scheduler measures the eigenspread of  $\mathbf{H}_k \mathbf{H}_k^H$  and selects the user with the minimum eigenspread (MinES). The eigenspread is defined as:

$$s = \frac{\lambda_{max}}{\lambda_{min}} \quad (3.10)$$

where  $\lambda_{max}$  and  $\lambda_{min}$  are the largest and smallest eigenvalues of  $\mathbf{H}_k \mathbf{H}_k^H$ , respectively. The eigenspread gives insight into the orthogonality of the channel. The smaller the value of  $s$ , the closer the matrix is to be orthogonal. The minimum value of  $s$  is one, and it occurs when the channel matrix is orthogonal. The eigenspread is related to the condition number of  $\mathbf{H}_k$  as:

$$s = \text{cond}(\mathbf{H}_k)^2$$

$$\text{where, } \text{cond}(\mathbf{H}_k) = \frac{\rho_{max}}{\rho_{min}} \quad (3.11)$$

where  $\rho_{max}$  and  $\rho_{min}$  are the largest and smallest singular values of  $\mathbf{H}_k$ . From this relation, we derive a scheduler that takes into account both the channel power and the eigenspread of  $\mathbf{H}_k$ . It selects the user that has the largest minimum singular value of  $\mathbf{H}_k$ . From (3.11), we have:

$$\frac{\rho_{max}}{\rho_{min}} = \sqrt{s}$$

$$\rho_{min} = \frac{\rho_{max}}{\sqrt{s}} \quad (3.12)$$

Thus, selecting the largest  $\rho_{min}$  means selecting a large  $\rho_{max}$ , which measures the norm of  $\mathbf{H}_k$  and hence the power, and a small eigenspread(s). We will refer to this scheduler as MaxMinSV.

Finally, we define the scheduling scheme MinInvTrace. In this scheme we chose the user with the minimum value of  $trace((\mathbf{H}_k^H \mathbf{H}_k)^{-1})$ . This scheme is somewhat similar to that of choosing MaxSNR ( $trace(\mathbf{H}_k \mathbf{H}_k^H)$ ) as can be seen from the following inequality.

$$trace((\mathbf{H}_k^H \mathbf{H}_k)^{-1})(trace(\mathbf{H}_k \mathbf{H}_k^H)) \geq M_T^2$$

To summarize, the scheduling algorithms considered for V-BLAST are:

- MaxVBLAST: select the user with  $\min_{k=1,2,\dots,K}\{w_k\}$ , where  $w_k = \max_{n=1,2,\dots,M_T} \{\|W_{ZF,n}^k\|^2\}$  and  $W_{ZF,n}$  is defined in (3.8).
- MinES: select the user with minimum eigenspread of  $\mathbf{H}_k \mathbf{H}_k^H$  as defined in defined in (3.10).
- MaxMinSV: select the user with maximum minimum singularvalue of  $\mathbf{H}_k$ .
- MaxSNR: select the user with maximum Frobenius norm of  $\mathbf{H}_k$  ( $trace(\mathbf{H}_k \mathbf{H}_k^H)$ ).
- MinInvTrace: select the user with minimum value of  $trace((\mathbf{H}_k^H \mathbf{H}_k)^{-1})$ .
- RR: round-robin scheduling allows each user to transmit in a time-division

fashion.

## 3.5 Bounds on VBLAST Capacity

In this section we will relate each of the scheduling techniques with the bounds on the VBLAST Capacity. The best scheduling technique is the one that is able to provide tighter bounds on

$$\min_k \max_n \min_i \|W_{ZF,n,i}^k\|^2$$

where  $\min_i \|W_{ZF,n,i}^k\|^2$  is the minimum diagonal value of  $(\mathbf{H}_{k,n}^H \mathbf{H}_{k,n})^{-1}$  of the  $n^{th}$  layer of the  $k^{th}$  user. In this section we will find bounds that are independent of  $n$  and  $i$  for each scheduling technique.

### 3.5.1 MaxMinSV

Choosing a user with maximum minimum singular value of  $\mathbf{H}_k$  is the same as choosing a user with maximum minimum eigenvalue of  $\mathbf{H}_k \mathbf{H}_k^H$ . We have

$$\min_i \|W_{ZF,n,i}^k\|^2 = [(\mathbf{H}_{k,n}^H \mathbf{H}_{k,n})^{-1}]_{i,i}$$

this means that we are looking for the minimum  $i^{th}$  diagonal value of  $(\mathbf{H}_{k,n}^H \mathbf{H}_{k,n})^{-1}$ .

We can write it in the form of following inequality,

$$\lambda_{min}^{k,n}((\mathbf{H}_{k,n}^H \mathbf{H}_{k,n})^{-1}) \leq \min_i \|W_{ZF,n,i}^k\|^2 \leq \lambda_{max}^{k,n}((\mathbf{H}_{k,n}^H \mathbf{H}_{k,n})^{-1}) \quad (3.13)$$

Since  $\mathbf{H}_n^H \mathbf{H}_n$  is squared Hermetian matrix, it can be written as:

$$\mathbf{H}_{k,n}^H \mathbf{H}_{k,n} = \mathbf{U} \Lambda_{k,n} \mathbf{U}^H \quad (3.14)$$

where  $\mathbf{U}$  is a unitary matrix with orthonormal eigenvectors and  $\Lambda_{k,n}$  is the diagonal matrix of eigenvalues  $[\lambda_1, \lambda_2, \dots, \lambda_{M_T}]$ . Another approach is to use QR decomposition as done in [41],[42],[43].

Using simple matrix manipulations, the inverse is written as:

$$[\mathbf{H}_{k,n}^H \mathbf{H}_{k,n}]^{-1} = \mathbf{U} \Lambda_{k,n}^{-1} \mathbf{U}^H \quad (3.15)$$

The diagonal elements of  $\Lambda_{k,n}^{-1}$  are  $[\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, \dots, \frac{1}{\lambda_{M_T}}]$ . Therefore,

$$\lambda_{min}^{k,n}((\mathbf{H}_{k,n}^H \mathbf{H}_{k,n})^{-1}) = \frac{1}{\lambda_{max}^{k,n}(\mathbf{H}_{k,n}^H \mathbf{H}_{k,n})} \quad (3.16)$$

$$\lambda_{max}^{k,n}((\mathbf{H}_{k,n}^H \mathbf{H}_{k,n})^{-1}) = \frac{1}{\lambda_{min}^{k,n}(\mathbf{H}_{k,n}^H \mathbf{H}_{k,n})} \quad (3.17)$$

For the sake of simplicity we will henceforth refer to  $\lambda_{max}^{k,n}(\mathbf{H}_{k,n}^H \mathbf{H}_{k,n})$  and  $\lambda_{min}^{k,n}(\mathbf{H}_{k,n}^H \mathbf{H}_{k,n})$  as  $\lambda_{max}^{k,n}$  and  $\lambda_{min}^{k,n}$  respectively. We now have,

$$\frac{1}{\lambda_{max}^{k,n}} \leq \min_i \|W_{ZF,n,i}^k\|^2 \leq \frac{1}{\lambda_{min}^{k,n}} \quad (3.18)$$

$$\frac{1}{\min_n \lambda_{max}^{k,n}} \leq \max_n \min_i \|W_{ZF,n,i}^k\|^2 \leq \frac{1}{\min_n \lambda_{min}^{k,n}} \quad (3.19)$$

According to the inclusion principle for matrices the minimum value of  $\lambda_{min}$  occurs at  $n = 1$  and minimum value of  $\lambda_{max}$  occurs at  $n = M_T$ . Now we have an upper bound given by,

$$\max_n \min_i \|W_{ZF,n,i}^k\|^2 \leq \frac{1}{\lambda_{min}^{(k)}} \quad (3.20)$$

or by taking minimum over  $K$  users,

$$\min_k \max_n \min_i \|W_{ZF,n,i}^k\|^2 \leq \frac{1}{\max_k \lambda_{min}^{(k)}} \quad (3.21)$$

We will use this inequality to establish lower bounds on MaxMinSV.

For the sake of simplicity, we will focus on the case when  $M_T = M_R$ . The pdf of the smallest eigenvalue for the case of  $M_T = M_R$  is given by [44],

$$f_{\lambda_{min}}(x) = M_T e^{-x M_T} u(x) \quad (3.22)$$

In this case,  $\mu = \frac{1}{M_T}$  and  $\sigma = \frac{1}{M_T}$ , where  $\mu$  and  $\sigma^2$  represents the mean and variance. This is demonstrated in Figure 3.2 which plots the empirical and theoretical PDF for  $\lambda_{min}$  for  $M_T = M_R = 10$ . From Chapter 10 of [47] we see that the maximum of  $\lambda_{min}$  will scale as  $\frac{1}{M} \log K$ . Therefore we have a new bound on the capacity of VBLAST as,

$$C_{VBLAST}^{ZF} \geq M_T \log_2 \left( 1 + \frac{SNR \log_2 K}{M_T^2} \right) \quad (3.23)$$

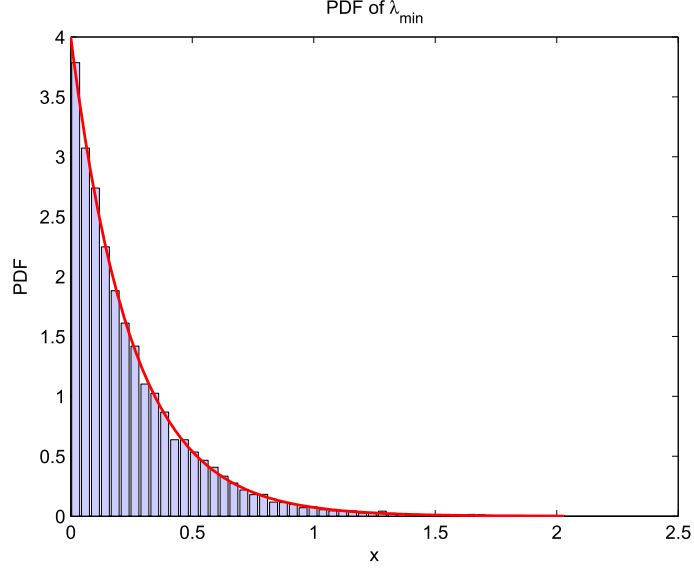


Figure 3.2: PDF of  $\lambda_{min}$

### 3.5.2 MinES

In this case we start with the inequality

$$\min_i \|W_{ZF,n,i}^k\|^2 \leq \frac{1}{\lambda_{min}^{k,n}} \leq \kappa^{(k,n)} \quad (3.24)$$

here  $\kappa^{(k,n)}$  is the condition number of the matrix  $\mathbf{H}_{k,n}\mathbf{H}_{k,n}^H$  defined as  $\lambda_{max}/\lambda_{min}$ .

Now

$$\max_n \min_i \|W_{ZF,n,i}^k\|^2 \leq \max_n \kappa^{(k,n)} \quad (3.25)$$



The condition number  $\kappa$  is maximum when  $n = 1$  a lemma which follows from the inclusion principle of matrices, Therefore we have,

$$\max_n \min_i \|W_{ZF,n,i}^k\|^2 \leq \kappa^{(k)} \quad (3.26)$$

$$\min_k \max_n \min_i \|W_{ZF,n,i}^k\|^2 \leq \min_k \kappa^{(k)} \quad (3.27)$$

The pdf of  $s$  the condition number of  $\mathbf{H}_k$  for the case of  $M_T = M_R$  is given by [44],

$$f_{s/M_T}(x) = \frac{8}{x^3} e^{-\frac{4}{x^2}} u(x) \quad (3.28)$$

As  $\kappa = s^2$ , we can show using a variable transformation argument that,

$$f_{\kappa}(x) = \frac{4M_T^2}{x^2} e^{-\frac{4M_T^2}{x}} u(x) \quad (3.29)$$

The simulated and theoretical PDF of  $\kappa$  is show in Figure 3.3,

Now we find the upper bound on  $\min_k \kappa^{(k)}$ . [46] provides with bounds on inequalities for minimum of random variables. A random variable  $\xi$  is bounded as,

$$\frac{1}{2\alpha K} \leq E \min_k \xi^{(k)} \leq \frac{1}{\beta K} \quad (3.30)$$

where  $\alpha > 0$  and  $\beta > 0$ . These inequalities exist if the random variable  $\xi$  satisfies

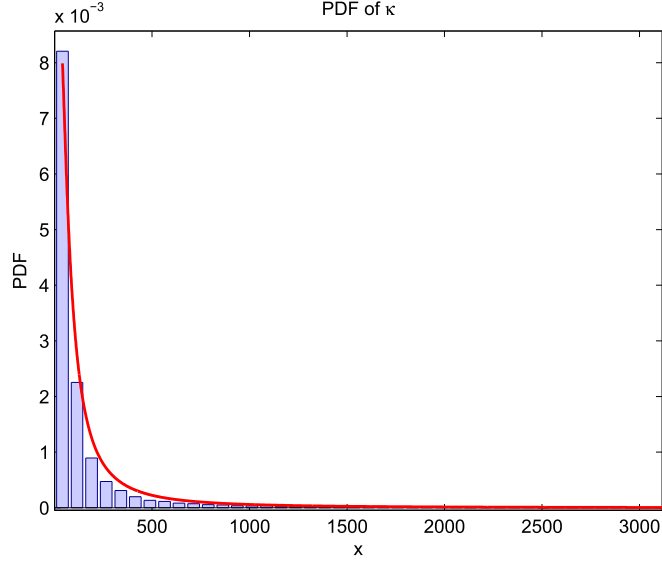


Figure 3.3: PDF of  $\kappa$

$(\alpha, \beta)$  condition defined as,

$$P(|\xi| \leq t) \leq \alpha t \quad (3.31)$$

$$P(|\xi| \geq t) \leq e^{-\beta t} \quad (3.32)$$

Using the PDF of  $\kappa$  we find that  $\alpha = \beta = \frac{1}{4M_T^2}$ . Now the upper bound on the minimum is given by,

$$E \min_k \xi^{(k)} \leq \frac{4M_T^2}{K} \quad (3.33)$$

Now we have,

$$\min_k \max_n \min_i \|W_{ZF,n,i}^k\|^2 \leq \frac{4M_T^2}{K} \quad (3.34)$$

### 3.5.3 MaxSNR

We start with the established inequality,

$$\frac{1}{\max_k \min_n \text{trace}(\mathbf{H}_{k,n} \mathbf{H}_{k,n}^H)} \leq \frac{1}{\max_k \min_n \lambda_{max}^{k,n}} \leq \min_k \max_n \min_i \|W_{ZF,n,i}^k\|^2 \quad (3.35)$$

Again from the inclusion principle the minimum of the trace occurs at  $n = M_T$ . Since we do now know which will be the last layer. The PDF of  $\text{trace}(\mathbf{H}_{k,n} \mathbf{H}_{k,n}^H)$  for last  $n$  becomes conditional. If we solve for the unordered VBLAST algorithm we will get,

$$\frac{1}{\max_k \text{trace}(\mathbf{h}_k^H \mathbf{h}_k)} \leq \min_k \max_n \min_i \|W_{ZF,n,i}^k\|^2 \quad (3.36)$$

where  $\mathbf{h}_k$  is a row vector containing  $M_R$  i.i.d complex gaussian entries. In this case the trace will be a chi-squared random variable with  $2M_R$  degrees of freedom. The PDF in this case is given as,

$$f_{tr}(x) = \frac{x^{M_R-1}}{(M_R-1)!} e^{-x} u(x) \quad (3.37)$$

Now we use an upper bound on the maximum of the trace. [45] provides us with a tight upper bound,

$$E [\max_k \text{trace}(\mathbf{h}_k^H \mathbf{h}_k)] \leq \mu + \sigma \sqrt{K-1} \quad (3.38)$$

where  $\mu$  and  $\sigma^2$  are mean and variances respectively. In this case,  $\mu = 2M_R$  and  $\sigma^2 = 4M_R$ . Now we have,

$$E[\max_k \text{trace}(\mathbf{h}_k^H \mathbf{h}_k)] \leq 2\sqrt{M_R}(\sqrt{M_R} + \sqrt{K-1}) \quad (3.39)$$

Here we define an upper bound on capacity,

$$C_{VBLAST}^{ZF-unorder} \leq M_T \log_2 \left( 1 + \frac{2SNR(\sqrt{M_R} + \sqrt{K-1})}{\sqrt{M_T}} \right) \quad (3.40)$$

### 3.5.4 MinInvTrace

We have,

$$\min_k \max_n \min_i \|W_{ZF,n,i}^k\|^2 \leq \min_k \text{trace}((\mathbf{H}_k^H \mathbf{H}_k)^{-1}) \quad (3.41)$$

(3.41) results from the fact that the minimum diagonal entry of a matrix is less than the trace of the matrix, the inclusion principle maximizes the equation when  $n = 1$ . Our simulation shows that this is the tightest bound that we have but we were not able to prove it theoretically. This is the reason behind its good performance.

## 3.6 Simulation Results

### 3.6.1 BER Performance and Outage Capacity

The BER performance of the wireless uplink system over different SNR with scheduling is shown in Figure 3.4. The MaxSNR scheduler doesn't capture any multiuser diversity but it gains a around 1 dB compared to the RR algorithm due to selecting the maximum SNR. On the other hand, the best scheduler is the one that maximizes V-BLAST capacity (MaxVBLASTCapc) by selecting the user who has the strongest weakest layer as described earlier. The MinInvTrace and the MaxMinSV schedulers capture the multiuser diversity but MinInvTrace provides more gain since it proves to be the tightest bound irrespective of the cancelation layer of VBLAST. These scheduling techniques perform very close to MaxVBLASTCapc which has a higher diversity at high SNRs (sharper slope). The result in this figure also shows that using maximum MIMO capacity as the scheduling criterion, as in (3.3), doesn't perform very well for V-BLAST. The reason for that is the suboptimality of the V-BLAST detection algorithm.

The complementary cumulative distribution function (CCDF) of the capacity of V-BLAST uplink scheduling is shown in Figure (3.5). The result shows that the MaxMinSV and MinInvTrace schedulers perform very close to the MaxVBLAST-Capc scheduler.

Figure (3.6) shows the capacity gains of uplink V-BLAST scheduling at 10% outage versus number of users. The optimal MIMO capacity scheduling, as defined in (3.3), is estimated by assuming the availability of optimal MIMO modems.

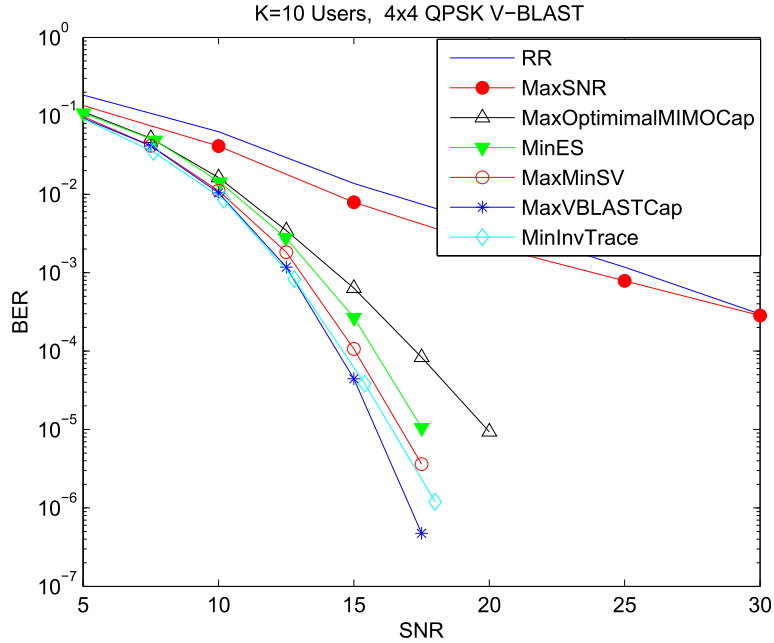


Figure 3.4: Aggregate BER of 4x4 QPSK V-BLAST users with uplink scheduling

Therefore, it provides an upper bound for the V-BLAST scheduling algorithms. However, when using RR scheduling with optimal MIMO modems, the rates provided by V-BLAST with scheduling are higher when the number of users is greater than five. The reason is that V-BLAST with scheduling captures  $K$ -fold diversity, where  $K$  is the number of users, in addition to being a full spatial multiplexing scheme while the optimal MIMO system with RR scheduling has only spatial diversity. Furthermore, the MaxVBLASTCap scheduler approaches optimal MIMO scheduling for large number of users. Thus, scheduling greatly improves the capacity of the uplink system even with suboptimal detectors.

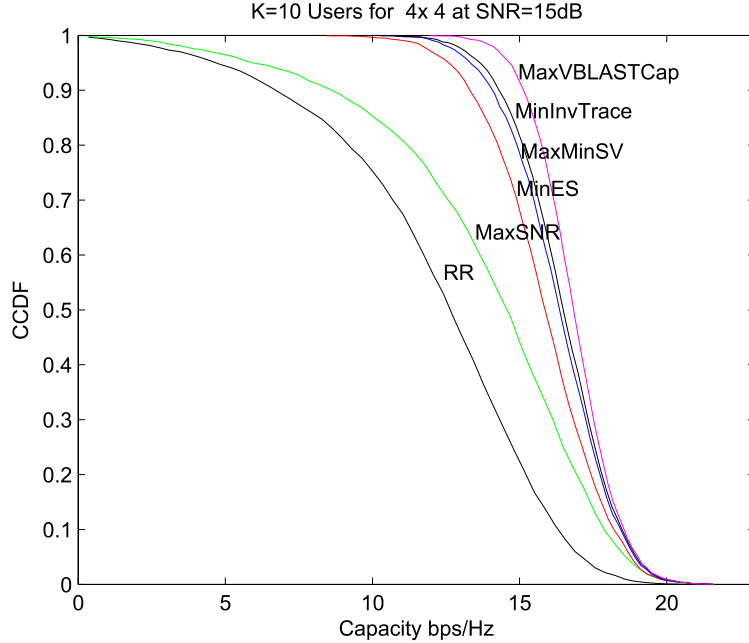


Figure 3.5: Capacity CCDF of 4x4 QPSK V-BLAST with uplink scheduling

### 3.6.2 Effect of Suboptimal Detection

To further examine the effect of using less complex receivers for spatially multiplexed (SM) users with uplink scheduling, we consider two suboptimal receivers. The first one removes the ordering process from the V-BLAST algorithm and the second one is a simple ZF receiver. The rationale is that multiuser diversity may overcome the weakness of the suboptimal algorithms.

Similar to (3.6), the capacity of a spatial multiplexing system with ZF receiver for user  $k$  is:

$$C_{SM-ZF}^k = M_T \left\{ \log_2 \left( 1 + \frac{SNR}{M_T \|W_{SM-ZF}^k\|^2} \right) \right\} \quad (3.42)$$

where,  $\|W_{SM-ZF}^k\|^2 = [(\mathbf{H}_k^H \mathbf{H}_k)^{-1}]_{ii}$

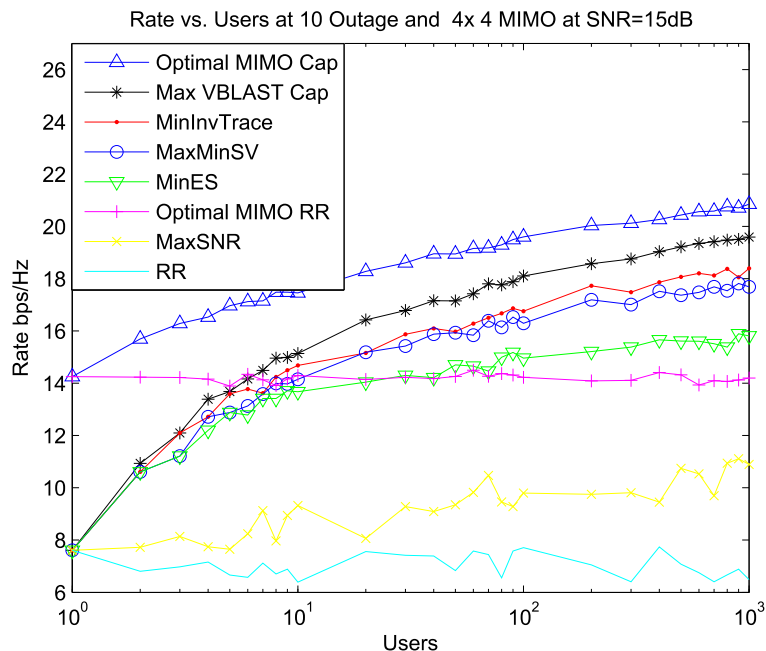


Figure 3.6: Capacity versus number of users at 4x4 MIMO channels and at 10% Outage probability

Figure 3.7 compares the capacity of SM schemes with uplink scheduling versus the number of users for different detectors. The result shows the loss in rate due to using simpler SM detection. Since the multi-user diversity order is high, we expect that the loss in rate will be small. The result indicates a small loss in capacity for large number of users. When the number of transmit and receive antennas is low, such as  $2 \times 2$ , the loss is even smaller.

### 3.6.3 Advantage of V-BLAST over SISO and SIMO systems

The simulation results in Figure 3.8 show the advantage of V-BLAST over receive diversity and single-input single-output (SISO) systems. The optimal scheduler



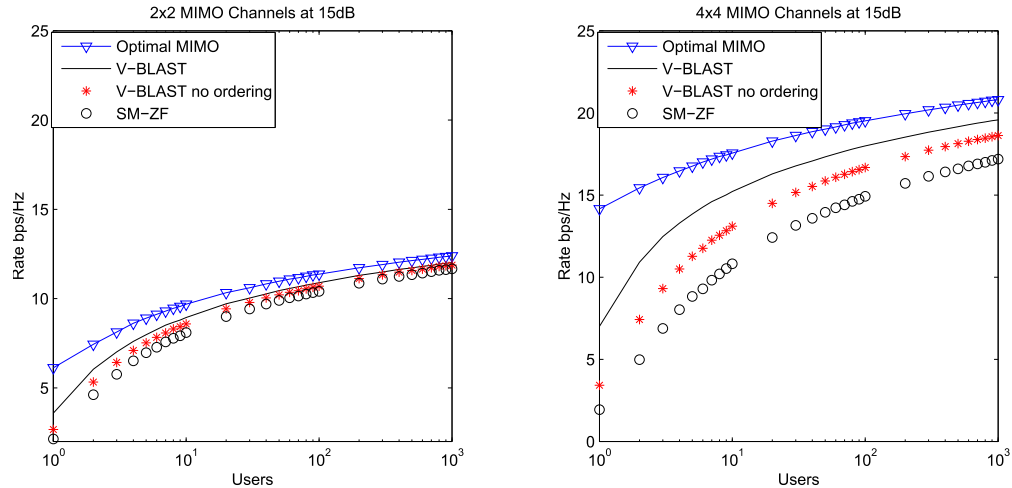


Figure 3.7: Capacity versus number of users at 10% Outage probability for sub-optimal detectors

that maximizes the capacity for SISO and single-input multiple-output (SIMO) systems is to select the user that has maximum SNR  $\max_{k=1,2,\dots,K} \{h_k^H h_k\}$ . Although V-BLAST is a suboptimal detector and it suffers from noise enhancement and error propagation, it boosts capacity by over 100% compared to SISO and MISO systems. Furthermore, its capacity increases logarithmically with number of users, similar to the behavior of MIMO systems when the number of transmit antennas exceeds number of receive antennas. Moreover, the BER performance of V-BLAST outperforms MRC by 8 dB at low BERs as shown in Figure 3.9.

### 3.6.4 Spatial Multiplexing with Sphere Decoding

A sphere decoder (SD) is a maximum likelihood (ML) detector that provides full receive diversity for spatial multiplexing systems with cubic average complexity at high SNRs. In this section, we will examine few scheduling algorithms for spatial multiplexing with sphere decoding (SM-SD). The scheduler that maximizes

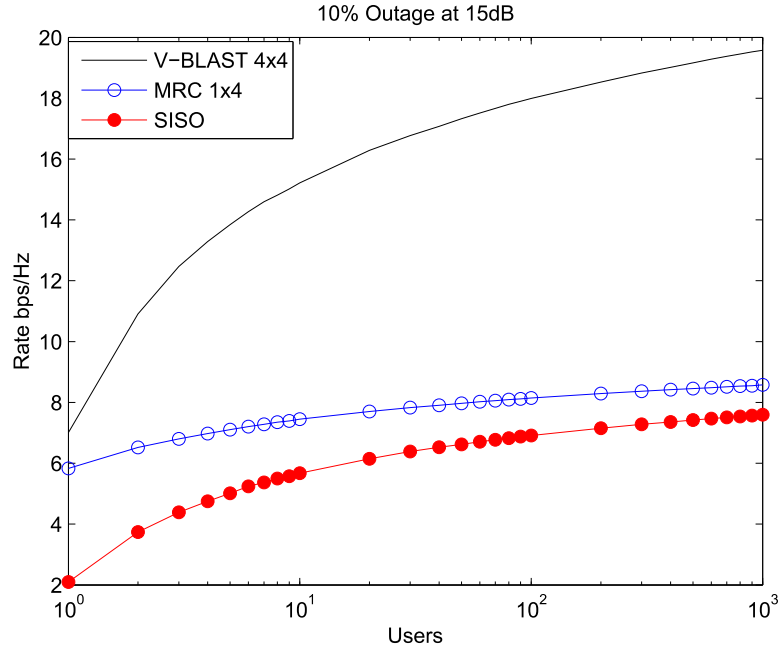


Figure 3.8: Spectral advantage of V-BLAST over receive diversity and SISO systems with uplink scheduling

MIMO capacity, as described in (3.3), is the optimal for SD. Other schedulers are described earlier. The BER performance is shown in Figure 3.10. The key difference between SD and V-BLAST is the receive diversity advantage due to ML detection. This may reduce the importance of scheduling. However, compared to RR scheduling, multiuser diversity still gives significant gains in BER performance. Figure 3.11 shows the capacity gains of SM-SD with different scheduling criteria at 10% outage as a function of number of users. Notice that the performance of the MaxSNR scheduler with SD is much better than with V-BLAST. This is a result of the full receive diversity advantage of SD. However, MinInvTrace gives better results compared to other scheduling techniques.

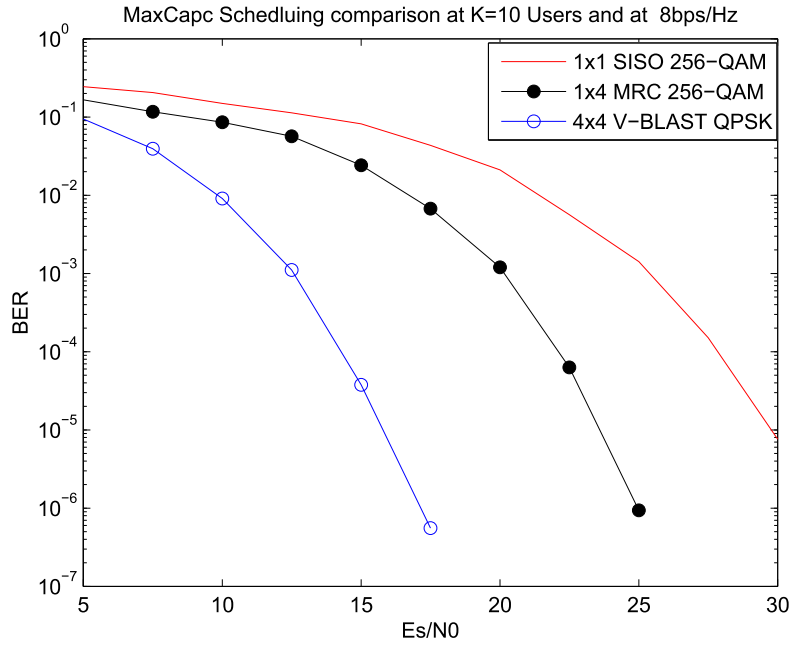


Figure 3.9: BER Comparison of V-BLAST and MRC with uplink scheduling

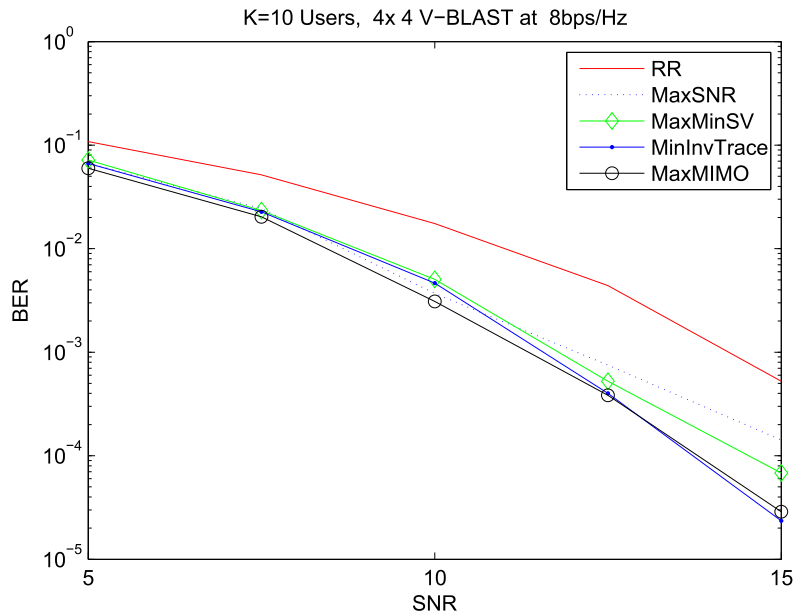


Figure 3.10: Sphere Decoder scheduling for 4x4 spatial multiplexing uplink users

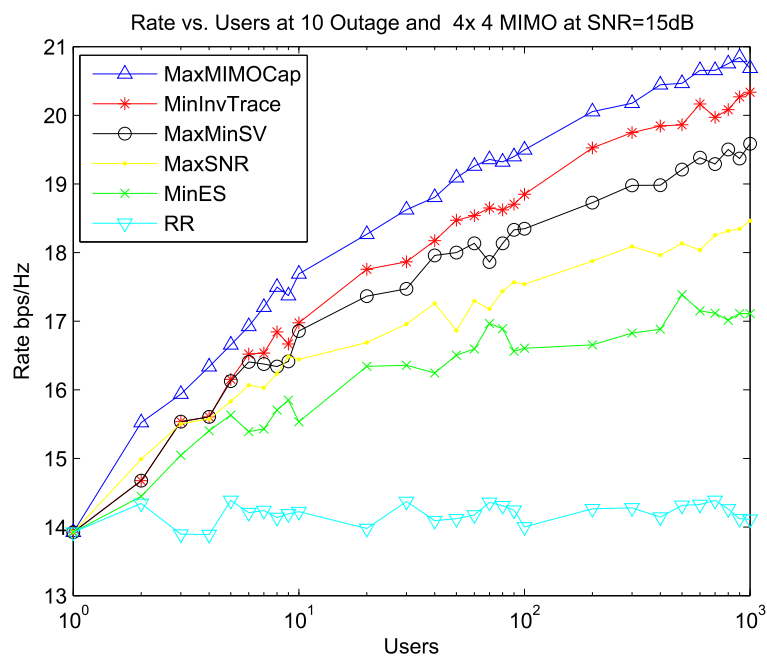


Figure 3.11: Capacity versus number of users at 4x4 MIMO channels and at 10% outage probability for SM-SD

## 3.7 Chapter Summary

This chapter compared and evaluated the performance of several scheduling algorithms for the uplink scheduling of V-BLAST users. The scheduler selects one user at a time and each user spatially multiplexes his data over the transmit antennas. The basic scheduling technique is known as round-robin in which each user is allowed to transmit in an ordered fashion without taking any advantage of the channel variations and statistics.

The main contribution of this chapter is finding the optimal scheduler that maximizes V-BLAST capacity for uplink single user scheduling. It selects the user that has the maximum post-processing SNR of the weakest layer. It needs to perform nulling and ordering for each MIMO channel matrix in order to evaluate this criterion. A suboptimal scheduler that performs very close to the optimal is to schedule to the user that has the maximum minimum singular value of the MIMO channel. This criterion takes into account the eigenspread of the MIMO channel and its power at the same time. The eigenspread of the MIMO channel measures how close it is to orthogonality because orthogonal channels maximize the MIMO capacity. This chapter also includes the different scaling results of the scheduling techniques used.

Furthermore, we showed that scheduling based on maximum MIMO capacity is not optimal for V-BLAST. That is due to the suboptimality of the V-BLAST detection algorithm while the MIMO capacity criterion assumes optimal encoding and decoding. Furthermore, scheduling for V-BLAST users outperforms optimal

MIMO users with round-robin scheduling. This greatly emphasizes the capacity gains that scheduling offers to wireless uplink systems. Moreover, comparing V-BLAST to less complex MIMO receivers, such as a zero forcing filter, shows that there is a small loss in capacity at a large number of users. When the number of transmit-receive antennas is low, such as  $2 \times 2$ , the loss is negligible. In addition, comparing the capacity of scheduling for V-BLAST with SISO and SIMO systems, we find huge gains in capacity. Finally, we studied scheduling for a spatial multiplexing system when using a sphere decoder (SD) at the receiver. Since SD is a maximum likelihood receiver, scheduling to a user who maximizes the MIMO channel capacity performs the best. However, the gains of scheduling in this case are less significant since SD already provides full receive diversity to the MIMO system.

## CHAPTER 4

# ESTIMATION OF THE DISTRIBUTION OF RANDOMLY DEPLOYED WIRELESS SENSORS

The distribution of randomly deployed wireless sensors plays a vital role in the quality of and methods used for data acquisition and signal reconstruction. In this chapter we use moments approach to recover the moments of the transmitted signal and we then relate these moments to the distribution of the randomly deployed sensors.

## 4.1 Introduction

Sensor networks can be used for a variety of important tasks, such as measuring or monitoring temperature, sunlight or seismic activity in an area [6]. The goal of such networks is to extract as much information as possible from an array of randomly deployed wireless sensors.

Consider a case where wireless sensors are released from an airplane at fixed intervals so that when the sensors land on earth their positions follow a random distribution. In such a case, estimating the variance of the distribution of the sensors might help determine the speed of wind in the area while estimating the whole distribution can improve the signal reconstruction [7]-[8].

In environment monitoring different algorithms can be used to estimate the sensors location [9]-[10] but these algorithms require a substantial amount of communication between the sensors in large networks. In this chapter, we represent a way to estimate any information regarding the distribution of the sensors using a noisy signal received from the sensors without any communication between the sensors. This chapter is organized as follows. Section 4.2.1 describes the system model used. Section 4.3 provides an estimate for the PDF of the distribution of the sensors. Simulation results are summarized in Section 4.4.5 followed by the conclusion. Throughout the chapter  $\text{Tr}(\mathbf{A})$  represents a trace of a matrix  $\mathbf{A}$  and  $\text{tr}_L(\mathbf{A})$  represents a normalized trace ( $\frac{1}{L}\text{Tr}(\mathbf{A})$ ) of a matrix  $\mathbf{A}$ .



## 4.2 Sensor Deployment

### 4.2.1 System Model

Consider a one dimensional physical field with  $L$  sensors deployed in the interval  $[0, 1]$ . Let  $d_i \in [0, 1]$  represent the position of the  $i^{th}$  sensor. In the case of Line of Sight (LOS), the continuous-time band-limited signal  $y(\omega_i)$  can be represented as the weighted sum of  $P$  harmonics,

$$y(\omega_i) = \frac{1}{\sqrt{P}} \sum_{k=0}^{P-1} x_k e^{-jk\omega_i} + \sigma n_i \quad (4.1)$$

where  $\omega_i = 2\pi d_i$ ,  $i = 1, 2, \dots, L$ ,  $x_k$  is the corresponding Fourier coefficients up to the  $P^{th}$  harmonic and  $n_i$  is additive white Gaussian noise with unit variance and  $\sigma^2$  is the variance of noise for the  $i^{th}$  sensor. We can write this model in vector form as

$$\mathbf{y} = \mathbf{V}^H \mathbf{x} + \sigma \mathbf{n} \quad (4.2)$$

where  $\mathbf{y}$  is the received signal vector of length  $L$  whose  $i^{th}$  element is  $y(\omega_i)$ ,  $\mathbf{x}$  is the transmitted signal of length  $P$  whose  $k^{th}$  element is  $x_k$ ,  $\mathbf{n}$  is the noise vector of length  $L$  and  $\mathbf{V}$  is a  $P \times L$  Vandermonde matrix given by

$$\mathbf{V} = \frac{1}{\sqrt{P}} \begin{pmatrix} 1 & \dots & 1 \\ e^{-j\omega_1} & \dots & e^{-j\omega_L} \\ \vdots & \ddots & \vdots \\ e^{-j(P-1)\omega_1} & \dots & e^{-j(P-1)\omega_L} \end{pmatrix}. \quad (4.3)$$

Here  $\omega_1, \omega_2, \dots, \omega_L$  are i.i.d random variables representing the position of  $d_i^{th}$  sensor with a certain distribution and are bounded within the interval  $[0, 2\pi)$ . Now suppose that we have  $K$  such observations for the received signal, then the new model can be represented as

$$\mathbf{Y} = \mathbf{V}^H \mathbf{X} + \sigma \mathbf{N} \quad (4.4)$$

Where  $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_K]$ ,  $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_K]$  and  $\mathbf{N} = [\mathbf{n}_1, \mathbf{n}_2 \dots \mathbf{n}_K]$ . The sample covariance matrix is defined as  $\mathbf{Y}\mathbf{Y}^H$ .

In this chapter we assume that the training sequence matrix  $\mathbf{X}$  and the noise matrix  $\mathbf{N}$  are zero mean Gaussian matrix with i.i.d. entries with a variance of  $\frac{1}{K}$ . However the methods used in this chapter can also be used when the training sequence  $\mathbf{X}$  is deterministic. The matrices involved are of large dimension with  $c_1 = \lim_{P \rightarrow \infty} \frac{P}{K}$ ,  $c_2 = \lim_{P \rightarrow \infty} \frac{L}{P}$  and  $c_3 = \lim_{K \rightarrow \infty} \frac{L}{K}$ . Through out this chapter we will consider  $c_1 \in [0, 1]$  and  $c_3 \in [0, 1]$  i.e.  $L, P \leq K$

### 4.2.2 Distribution Estimation

The estimation of the distribution of  $\omega$  in eq (4.3) provides us with the information of the location of the sensors in an area. This information can be either in the form of a certain parameter regarding the distribution, the range of the distribution or it could be the distribution of  $\omega$  itself. This is a difficult task when there is no communication between the sensors as is the case here. However, as we will show, the moments of  $\mathbf{V}\mathbf{V}^H$  can be estimated and related to the distribution of the deployed sensors by using the moments approach. Specifically, we relate the moments of  $\mathbf{V}\mathbf{V}^H$  up to a certain order with a polynomial approximation of the distribution of  $\omega$ .

## 4.3 Moments Approach

The moments approach [13] provides us with a good estimator for the moments of the Vandermonde matrix. The moment ( $m_n$ ) of a  $P \times P$  matrix  $\mathbf{H}$  is defined as

$$m_n = \text{tr}_P(\mathbf{H}^n) \quad (4.5)$$

where  $m_n$  is the normalized trace of the matrix. Free deconvolution [11] relates the eigenvalue distribution of the covariance matrix ( $\mu_{\mathbf{Y}\mathbf{Y}^H}$ ) with the eigenvalue

distribution the Vandermonde matrix  $(\mu_{\mathbf{V}\mathbf{V}^H})$  as

$$\mu_{\mathbf{V}\mathbf{V}^H} = (\mu_{\mathbf{Y}\mathbf{Y}^H} \boxminus_{c_3} \mu_{\mathbf{N}\mathbf{N}^H}) \boxtimes \mu_{\mathbf{X}\mathbf{X}^H} \quad (4.6)$$

Where  $\boxminus_{c_3}$  is Rectangular additive free deconvolution (section 4.3.1) and  $\boxtimes$  is multiplicative free deconvolution (section 4.3.2).

The moments estimation simply follow the following steps

### 4.3.1 Step 1: Rectangular additive free deconvolution

Consider the covariance matrix

$$\mathbf{Y}\mathbf{Y}^H = (\mathbf{A} + \mathbf{N})(\mathbf{A} + \mathbf{N})^H \quad (4.7)$$

where  $\mathbf{A} = \mathbf{V}^H\mathbf{X}$ . Rectangular additive free deconvolution ( $\boxminus_{c_3}$ ) provides us with the moments of  $\mathbf{Y}\mathbf{Y}^H$  in terms of moments of  $\mathbf{A}\mathbf{A}^H$  and moments of  $\mathbf{N}\mathbf{N}^H$ . The moments  $(m_n)$  are related to another sequence of numbers called the rectangular free cumulants  $(t_n)$  via the probability measure  $\epsilon$ . Refer to which [13] gives the following set of equations for the relation between the moments and the rectangular free cumulants.

$$T_\epsilon(z(c_3 M_{\epsilon^2}(z) + 1)(M_{\epsilon^2}(z) + 1)) = M_{\epsilon^2}(z)$$

where

$$T_\epsilon(z) = \sum_{n \geq 1} t_n(\epsilon) z^n \text{ and } M_{\epsilon^2}(z) = \sum_{n \geq 1} m_n(\epsilon) z^n.$$

This equation can be written in a recursive form as

$$\begin{aligned} m_0(\epsilon) &= 1 \\ m_n(\epsilon) &= t_n(\epsilon) \\ &+ \sum_{k=1}^{n-1} c_3^k t_k(\epsilon) \sum_{\substack{l_1, \dots, l_{2k} \geq 0 \\ l_1 + \dots + l_{2k} = n-k}} m_{l_1}(\epsilon) \dots m_{l_{2k}}(\epsilon) \end{aligned}$$

Let  $\gamma$ ,  $\eta$  and  $\tau$  be the probability measure of  $\mathbf{Y}\mathbf{Y}^H$ ,  $\mathbf{A}\mathbf{A}^H$  and  $\mathbf{N}\mathbf{N}^H$  respectively.

In this case free deconvolution the rectangular free cumulant of these probability measure are related to each other as

$$t_n(\eta) = t_n(\gamma) - t_n(\tau)$$

If  $\mathbf{N}$  is a random matrix with independent Gaussian entries with variance  $\frac{1}{K}$  then the eigenvalue distribution of  $\mathbf{N}\mathbf{N}^H$  follows a Marchenko-Pastur distribution with parameter  $\frac{1}{c_3}$ . In this case the rectangular free cumulants of  $\mathbf{N}\mathbf{N}^H$  are given as  $t_n(\tau) = \delta_{n,1}, \forall n \geq 1$  (here  $\delta_{n,1}$  is the dirac delta function) [14]. It is important to note that the rectangular free cumulants of  $\sigma^2 \mathbf{N}\mathbf{N}^H$  are given by  $t_n(\sigma^2 \tau) = (\sigma^2)^n t_n(\tau) = (\sigma^2)^n \delta_{n,1}$

### 4.3.2 Step 2: Multiplicative free deconvolution

Now that rectangular additive free deconvolution has provided us with the moments of  $\mathbf{A}^H \mathbf{A}$ , or equivalently that of  $\mathbf{V}^H \mathbf{X} \mathbf{X}^H \mathbf{V}$  we can extract the moments of  $\mathbf{V} \mathbf{V}^H$ . We have

$$m_n(\mathbf{X} \mathbf{X}^H \mathbf{V} \mathbf{V}^H) = c_2 m_n(\mathbf{V}^H \mathbf{X} \mathbf{X}^H \mathbf{V}) \quad (4.8)$$

Multiplicative free deconvolution ( $\boxtimes$ ) provides us with the moments of  $\mathbf{V} \mathbf{V}^H$  in terms of the moments of  $\mathbf{X} \mathbf{X}^H \mathbf{V} \mathbf{V}^H$  and moments of  $\mathbf{X} \mathbf{X}^H$ . The moments  $m_n$  are related to another sequence of numbers called the multiplicative free cumulants ( $s_n$ ) through the probability measure  $\epsilon$ . Specifically, the relation between these two quantities is given by [12]

$$M_\epsilon(z) S_\epsilon(M_\epsilon(z)) = z(1 + M_\epsilon(z))$$

where,

$$S_\epsilon(z) = \sum_{n \geq 1} s_n(\epsilon) z^{n-1} \text{ and } M_\epsilon(z) = \sum_{n \geq 1} m_n(\epsilon) z^n.$$

These set of equations can be represented in a recursive form as

$$\begin{aligned} m_1(\epsilon) s_1(\epsilon) &= 1, \\ m_n(\epsilon) &= \sum_{k=1}^{n+1} s_k(\epsilon) \sum_{\substack{l_1, \dots, l_k \geq 1 \\ l_1 + \dots + l_k = n+1}} m_{l_1}(\epsilon) \dots m_{l_k}(\epsilon) \end{aligned}$$

Let  $\vartheta$ ,  $\varsigma$  and  $\psi$  be the probability measures of  $\mathbf{X}\mathbf{X}^H\mathbf{V}\mathbf{V}^H$ ,  $\mathbf{X}\mathbf{X}^H$  and  $\mathbf{V}\mathbf{V}^H$  respectively. Then these probability measures are related to each other through the multiplicative free cumulants as

$$\begin{aligned} s_1(\varsigma)s_1(\psi) &= s_1(\vartheta) \\ s_1(\varsigma)s_n(\psi) &= s_n(\vartheta) - s_n(\varsigma)s_1(\psi) \\ &\quad - \sum_{k=2}^{n-1} s_k(\varsigma)s_{n+1-k}(\psi) \end{aligned}$$

For example, if  $\mathbf{X}$  is  $P \times K$  random matrix with independent Gaussian entries with variance  $\frac{1}{K}$  then the eigenvalue distribution of  $\mathbf{X}\mathbf{X}^H$  follows a Marchenko-Pastur distribution with parameter  $\frac{1}{c_1}$ . In this case the multiplicative free cumulants of  $\mathbf{X}\mathbf{X}^H$  are given by  $s_n(\varsigma) = (-c_1)^{n-1}$ ,  $\forall n \geq 1$ .

### 4.3.3 Step 3: Moments of $\mathbf{V}\mathbf{V}^H$

Assume  $\omega_i = 2\pi(i-1) + \omega'_i$  where  $i = 1, 2, \dots, L$ . Here  $\omega'_i$  is a random variable with continuous (not necessarily uniform) distribution and is bounded by  $[0, 2\pi)$ . Then the asymptotic  $n^{\text{th}}$  moment of the Vandermonde matrix is defined as

$$m_n = \lim_{P \rightarrow \infty} E[\text{tr}_P(\mathbf{V}\mathbf{V}^H)^n] \quad (4.9)$$

In this case, the moments of the Vandermonde matrix can be calculated as [12]

$$\sum_{\rho \in \mathcal{P}(n)} K_{\rho, \omega} c_2^{|\rho|} \quad (4.10)$$

where  $\mathcal{P}(n)$  is the set of all partitions of  $\{1, 2, \dots, n\}$  and  $\rho$  is the notation for a particular partition in  $\mathcal{P}(n)$ , it could also be written as  $\rho = \{\rho_1, \dots, \rho_k\}$ , where  $\rho_j$  are the blocks of  $\rho$  and  $|\rho|$  is the number of blocks in  $\rho$ . The coefficient  $K_{\rho, \omega}$  is called the Vandermonde mixed moment expansion coefficient and is defined as [12]

$$K_{\rho, \omega} = \lim_{N \rightarrow \infty} \frac{1}{P^{n+1-|\rho|}} \int_{(0, 2\pi)^{|\rho|}} \prod_{k=1}^n \frac{1 - e^{jP(\omega_{b(k-1)} - \omega_{b(k)})}}{1 - e^{jP(\omega_{b(k-1)} - \omega_{b(k)})}} d\omega_1 \dots d\omega_{|\rho|} \quad (4.11)$$

We can write the moments of the Vandermonde matrix in terms of the PDF of  $\omega$  as

$$m_n = \sum_{\rho \in \mathcal{P}(n)} K_{\rho, u} c_2^{|\rho|} I_{|\rho|} \quad (4.12)$$

where  $I_{|\rho|} = (2\pi)^{|\rho|-1} \left( \int_0^{2\pi} p_\omega(x)^{|\rho|} dx \right)$  and  $u \sim \mathcal{U}(0, 1)$  (here  $\mathcal{U}$  is the uniform distribution). It is extremely difficult an explicit expression of  $K_{\rho, u}$  for any moments (in [12] only the first seven moments were computed). In this chapter we numerically calculate them using the following algorithm.

**Algorithm:**  $K_{\rho, u}$  can alternatively be expressed as the volume of the solution set



of

$$\sum_{k \in \rho_j} l_{k-1} = \sum_{k \in \rho_j} l_k \quad (4.13)$$

here  $0 \leq l_k \leq 1$ . This volume is calculated after expressing  $|\rho| - 1$  variables in terms of  $n + 1 - |\rho|$  free variables and is bounded within  $[0, 1]$ . We have that  $K_{\rho,u} = 1$  when the partitions of  $\rho$  are non-crossing [15] otherwise it is smaller the 1.

As  $I_{|\rho|}$  depends on the block cardinalities  $|\rho_j|$ , we can group  $K_{\rho,u}$  for  $\rho$  with equal block cardinalities together. We can group the cardinalities in descending order  $r_1 \geq r_2 \geq \dots \geq r_k$  and define

$$K_{r_1, r_2, \dots, r_k} = \sum_{\hat{\rho} \in \mathcal{P}(n)} K_{\rho, u} \quad (4.14)$$

where  $|\hat{\rho}_i| = r_i, \forall i$ . Then the moments can be represented as

$$m_n = \sum_{k=1, 2, \dots, n} \sum_{\substack{r_1, r_2, \dots, r_k \\ r_1 + r_2 + \dots + r_k = n}} K_{r_1, r_2, \dots, r_k} c_2^k I_k \quad (4.15)$$

Using this approach the first few moments expand to

$$\begin{aligned}
m_1 &= K_1 c_2 I_1 \\
m_2 &= K_2 c_2 I_1 + K_{1,1} c_2^2 I_2 \\
m_3 &= K_3 c_2 I_1 + K_{2,1} c_2^2 I_2 + K_{1,1,1} c_2^3 I_3 \\
m_4 &= K_4 c_2 I_1 + (K_{3,1} + K_{2,2}) c_2^2 I_2 + K_{2,1,1} c_2^3 I_3 \\
&\quad + K_{1,1,1,1} c_2^4 I_4 \\
m_5 &= K_5 c_2 I_1 + (K_{4,1} + K_{3,2}) c_2^2 I_2 + (K_{3,1,1} + K_{2,2,1}) c_2^3 I_3 \\
&\quad + K_{2,1,1,1} c_2^4 I_4 + K_{1,1,1,1,1} c_2^5 I_5 \\
&\quad \vdots \quad \vdots
\end{aligned} \tag{4.16}$$

**Corollary 4.3.1** *Consider a case when  $\omega_i = a'(i - 1) + \omega'_i$ , where  $a' \neq 2\pi$ . The moments in this case can be represented as*

$$\begin{aligned}
m_n &= \sum_{\rho' \in \mathcal{P}(n), \rho' \geq \rho} \sum_{j_1, \dots, j_n} a^{j_n + \frac{n(1-n)}{2} - 2 + \sum_{k=1}^n (2k-1)j_k} \\
&\quad K_{\rho, u} c_2^{|\rho|} L^{-|\rho|} I_{|\rho|}
\end{aligned} \tag{4.17}$$

where  $a = e^{-ja'}$  and the sequence  $j_1, j_2, \dots, j_n$  gives rise to the partition  $\rho$ .

The proof of (4.17) follows in parallel with the one provided in [12].

#### 4.3.4 Step 4: Estimation of P

Consider the case where  $P \leq L$ . Conventionally the estimation of the number of harmonics  $P$  is done by calculating the eigenvalues of the covariance matrix  $\mathbf{Y}\mathbf{Y}^H$ .

The covariance matrix will have  $L - P$  eigenvalues equal to  $\frac{1}{\sqrt{L}}\sigma$  and  $P$  eigenvalues greater than  $\frac{1}{\sqrt{L}}\sigma$ . The moments approach provides us with an alternative method to estimate  $P$  by only using the  $2^{nd}$  moment of the covariance matrix.

**Theorem 4.3.2** *The number of unknown harmonics  $P$  can be estimated as*

$$P = \frac{Km_2(\gamma) - K - L - K\sigma^4 - 2K\sigma^2 \pm \sqrt{\frac{K^2m_2^2(\gamma) - 2K(K\sigma^4 + 2K\sigma^2 + L + K)m_2(\gamma) + 2KL\sigma^4 + 4KL\sigma^2 + k^2 + 6K^2\sigma^4 + 4K^2\sigma^2 + K^2\sigma^8 + 4K^2\sigma^6 + L^2 + 2LK - 4LK\sigma^2 I_2(\sigma^2 + 2)}{4\sigma^2(\sigma^2 + 2)}}}{4\sigma^2(\sigma^2 + 2)} \quad (4.18)$$

Here  $\gamma$  is the probability measure of the covariance matrix  $\mathbf{Y}\mathbf{Y}^H$ .

The proof of (4.18) can be found in Appendix A.

#### 4.3.5 Step 5: PDF Approximation

Now in order to estimate the PDF of  $\omega$ , where  $\omega$  is bounded within  $[0, 1)$ , we use the Weierstrass approximation and multinomial expansion to derive an alternative

form of  $I_n$ .  $I_n$  is given by

$$\begin{aligned}
I_n = & \lim_{t \rightarrow \infty} (2\pi)^{n-1} \sum_{k_1 + \dots + k_t = n} \binom{n}{k_1, \dots, k_t} \\
& \cdot \left( \frac{\Gamma(1 + tn - \sum_{v=1}^t vk_v) \Gamma(\sum_{v=1}^t vk_v - n + 1)}{\Gamma((t-1)n + 2)} \right) \\
& \cdot \left( \prod_{v=1}^t \left[ p_\omega \left( \frac{v-1}{t-1} \right) \binom{t-1}{v-1} \right]^{k_v} \right) \tag{4.19}
\end{aligned}$$

where  $p_\omega(x)$  represents the unknown PDF of  $\omega$  and  $p_\omega(0), p_\omega(\frac{1}{t-1}), p_\omega(\frac{2}{t-1}), \dots, p_\omega(1)$  are the unknowns of the weierstrass approximation. Now equating (4.19) with (4.16) will give us a set of non-linear equations to solve. As the number of unknowns should be equal to the number of equations, we take  $n = t$ . One can solve for the numerical values of the unknowns by using any optimization algorithm. The proof of (4.19) can be found in appendix B.

Weierstrass approximation considers  $\omega$  to be bounded within  $[0, 1)$  however we can loosen the bound on the range of  $\omega$  from  $[0, 1)$  to  $[0, a)$  where  $a < 2\pi$  by replacing  $t$  by  $t/a$  in (4.19).

## 4.4 Simulation Results

In this section we will show the estimates for the moments of  $\mathbf{V}\mathbf{V}^H$ ,  $P$ , noise variance  $\sigma^2$  and parametric estimates for some special cases. Finally, we will estimate the PDF of the distribution of the wireless sensors.

#### 4.4.1 Estimation of the moments of $\mathbf{V}\mathbf{V}^H$

By applying rectangular additive free deconvolution and multiplicative free deconvolution on the covariance matrix  $\mathbf{Y}\mathbf{Y}^H$  we get the moments of  $\mathbf{V}\mathbf{V}^H$ . To check the validity of the moments approach we plot the relative distance between the estimated moments of  $\mathbf{V}\mathbf{V}^H$  and the actual moments of  $\mathbf{V}\mathbf{V}^H$ . We define the relative distance of the  $n^{\text{th}}$  moment for all  $n$  as

$$\left| \frac{m_n[(\mu_{\mathbf{Y}\mathbf{Y}^H} \boxminus_{c_3} \mu_{\mathbf{N}\mathbf{N}^H}) \boxminus \mu_{\mathbf{X}\mathbf{X}^H}] - 1}{m_n[\mathbf{V}\mathbf{V}^H]} \right|$$

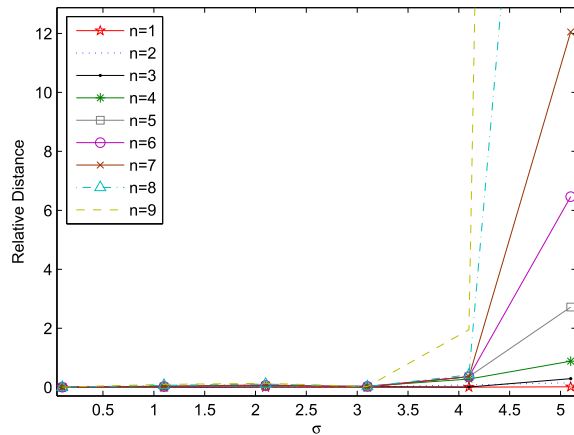


Figure 4.1: Relative distance of the moments of  $\mathbf{V}\mathbf{V}^H$  and the estimated moments by applying moments approach for varying  $\sigma$

Figure 4.1 is a plot of the relative distance for  $n = 1, 2, \dots, 9$  as the noise variance  $\sigma^2$  increases for  $K = L = N = 1000$ . The simulation result is averaged over 10 observations when  $\omega$  has uniform distribution. It is clear from the plot that as the variance of noise increases the estimation of the moments gets worse. This result is easily understood if we define the  $n^{\text{th}}$  rectangular free cumulant

error for  $\mathbf{NN}^H$  (see Section 4.3.1) as  $\varepsilon_n$ . It can be represented as

$$\varepsilon_n = t_n^{actual} - \delta_{n,1}$$

It follows that the error for the  $n^{th}$  free cumulant of  $\sigma\mathbf{NN}^H$  is  $(\sigma^2)^n\varepsilon_n$ . For  $\sigma^2 > 1$  the error exponentially increases with  $n$ .

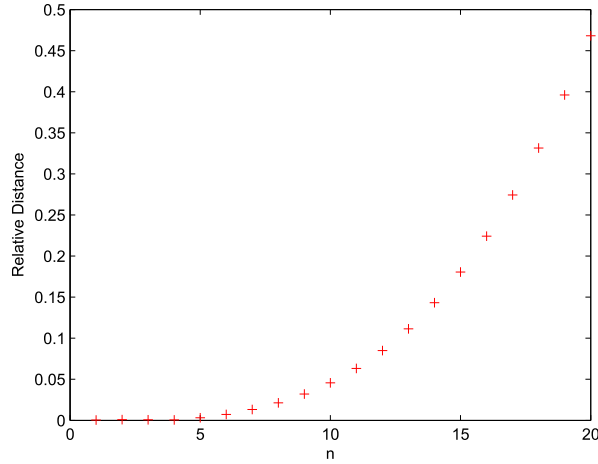


Figure 4.2: Relative distance of the moments of  $\mathbf{VV}^H$  and the estimated moments by applying moments approach for  $\sigma^2 = 1$

Figure 4.2 plots the relative distance when  $\omega$  has a uniform distribution for  $n = 1, 2, \dots, 20$  with  $K = L = P = 1000$  and  $\sigma^2 = 1$ . The small relative distance between the actual moments and the estimated moments of  $\mathbf{VV}^H$  indicate that the moments approach is a good estimate for the moments of the Vandermonde matrix.

#### 4.4.2 Estimation of $P$

Consider here the case that  $\omega$  has a uniform distribution. We use (4.18) to estimate the value of the unknown harmonics. In this case  $I_2 = 2\pi$  and  $P$  is estimated as,

$$P = \frac{Km_2(\gamma) - K - L - K\sigma^4 - 2K\sigma^2 \pm \sqrt{\frac{K^2m_2^2(\gamma) - 2K(K\sigma^4 + 2K\sigma^2 + L + K)m_2(\gamma) + 2KL\sigma^4 + 4KL\sigma^2 + k^2 + 6K^2\sigma^4 + 4K^2\sigma^2 + K^2\sigma^8 + 4K^2\sigma^6 + L^2 + 2LK - 8LK\sigma^2\pi(\sigma^2 + 2)}{2\sigma^2(\sigma^2 + 2)}}}{2\sigma^2(\sigma^2 + 2)} \quad (4.20)$$

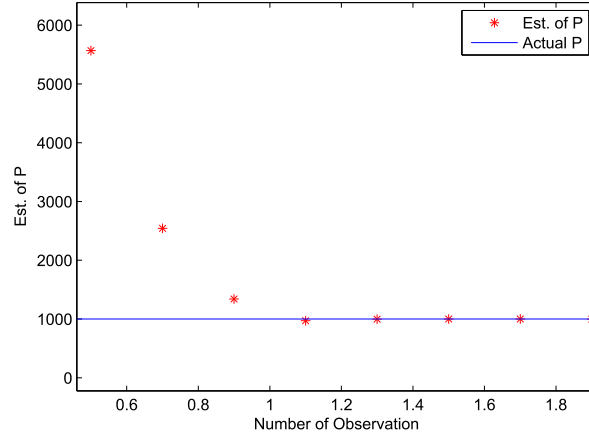


Figure 4.3: Estimation of  $P$  for Uniform distribution with  $K = L = 1000$  for varying  $\sigma$

Figure 4.3 estimates  $P$  as  $\sigma^2$  increases with  $\omega$  having a uniform distribution and  $K = L = 1000$ . The estimation of  $P$  gets worse for small values of  $\sigma$  which follows from (4.18) as  $\lim_{\sigma \rightarrow 0^+} P = \infty$ .

Figure 4.4 provides an estimate of  $P$  with  $K = L = 1000$  and  $\sigma^2 = 1$  with  $\omega$  having uniform distribution. In this figure the estimate of  $P$  is averaged over a number of observations. The equation for estimating  $P$  is derived by using only the first two moment estimates and because the the estimate of the first moments

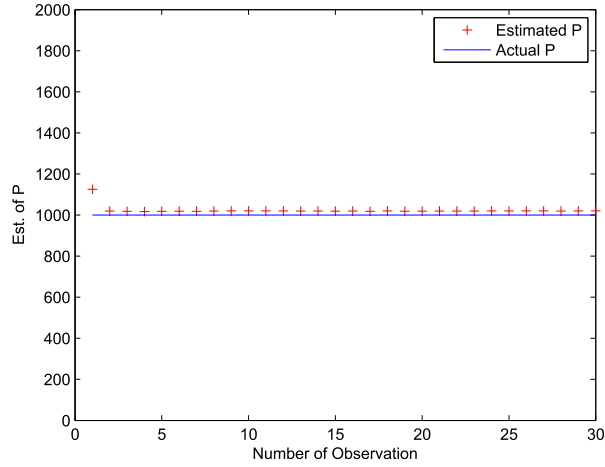


Figure 4.4: Estimation of  $P$  for Uniform distribution with  $K = L = 1000$  with  $\sigma = 1$

are better for higher  $\sigma$  (as seen from Figure 4.1), we end up with a good estimate for  $P$  at higher values of  $\sigma$ .

### 4.4.3 Estimation of the noise variance $\sigma^2$

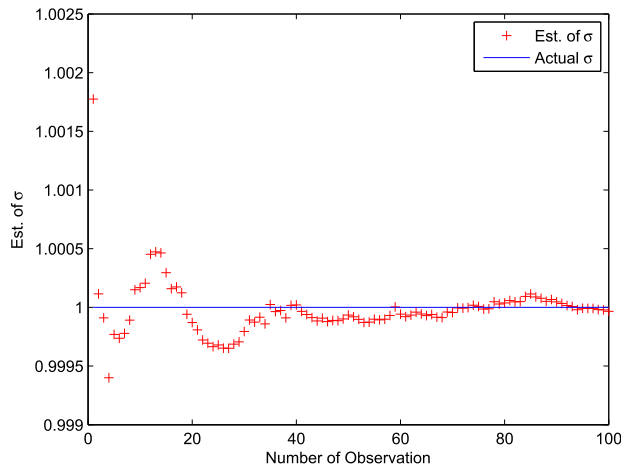


Figure 4.5: Estimation of  $\sigma$  for Uniformly distributed  $\omega$  with  $K = L = 1000$

We can estimate the variance of noise  $\sigma^2$  in a similar way to the way we estimated the value of  $P$ . Appendix A derives an expression for estimation of  $\sigma$ .



Specifically, we show that (see Appendix equation (A.9))

$$\sigma = \sqrt{m_1(\gamma) - 1} \quad (4.21)$$

where  $\gamma$  is the probability measure of  $\mathbf{Y}\mathbf{Y}^H$ . Figure 4.5 estimates the value of  $\sigma$  averaged over different number of observations when  $\omega$  has a uniform distribution with  $K = L = N = 1000$ .

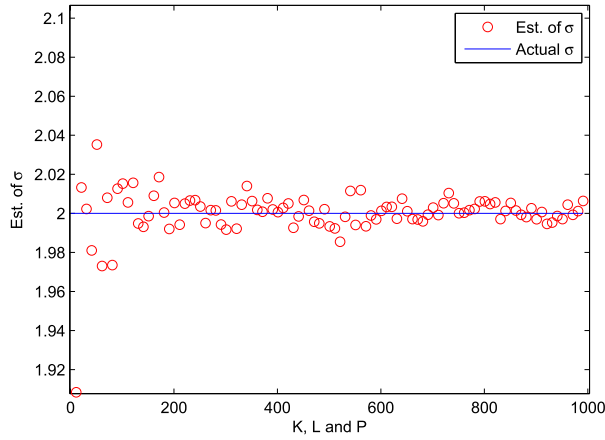


Figure 4.6: Estimation of  $\sigma$  for Uniformly distributed  $\omega$  with  $K$ ,  $L$  and  $P$  increasing asymptotically

Figure 4.6 shows the estimate of  $\sigma$  as  $K$ ,  $L$  and  $N$  increases asymptotically.

We see that the higher the dimension the more accurate are the results.

#### 4.4.4 Special Cases

In this subsection, we consider different distribution for  $\omega$  and try to estimate the underlying parameters of the corresponding distribution.

## Von Mises Distribution

Let us consider the case where  $\omega$  has Von Mises distribution. Von Mises distribution has an inverted bell shape and is bounded within  $(0, 2\pi]$ . Its parameter  $1/\kappa$  is analogous to  $\sigma$  in a Gaussian distribution. The distribution  $p_\omega$  of Von Mises distribution for the angle  $x$  is given by

$$p_\omega(x|\mu, \kappa) = \frac{e^{\kappa \cos(x-\mu)}}{2\pi J_0(\kappa)} \quad (4.22)$$

Here  $J_0$  is the Bessel function of the zero order. The value of  $I_2$  for a von Mises distribution is given by

$$I_2 = \frac{J_0(2\kappa)}{J_0(\kappa)^2} \quad (4.23)$$

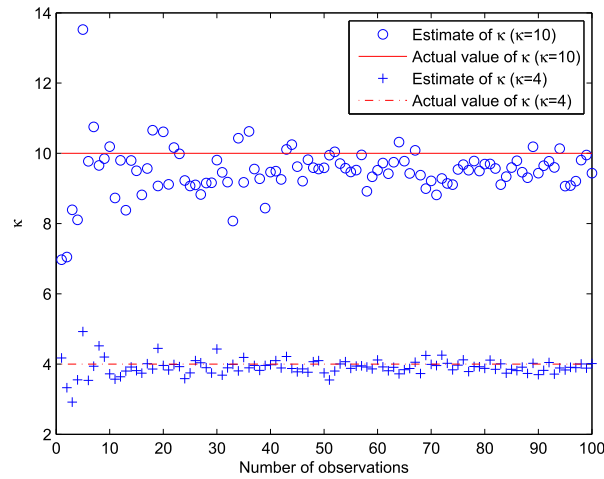


Figure 4.7: Estimation of  $\kappa$  for Von Mises distribution with  $K = L = P = 100$

The value of  $\kappa$  is estimated by solving the  $2^{nd}$  equation of (4.16) and (4.23).

Figure 4.7 shows the estimation result of  $\kappa$  when the trace values are  $\kappa = 4$  and

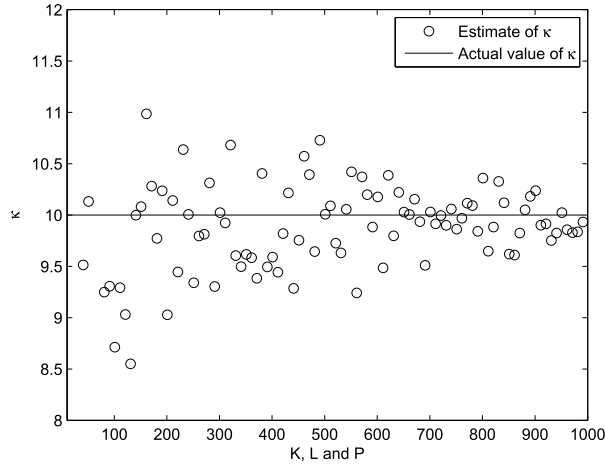


Figure 4.8: Estimation of  $\kappa$  for Von Mises distribution with varying  $K$ ,  $L$  and  $P$

$\kappa = 10$ . In this simulation, we set  $K = L = P = 100$ .

Figure 4.8 simulates the result for the estimation of  $\kappa$  as  $K$ ,  $L$  and  $P$  increase linearly. It can be seen that by asymptotically increasing the value of  $K$ ,  $L$  and  $P$  the estimation gets better. Figure 4.9 estimates the value of  $\kappa$  for  $K = L = P = 100$  for varying  $\sigma$ . From the figure, we conclude that as we increase the variance of the additive noise, the estimate quality gets worst.

### Bounded Wigner Semi-Circle Distribution

When  $\omega$  has a wigner semi-circle distribution bounded within  $[-R, R]$ , where  $R \leq \pi$ , the distribution  $p_\omega$  is given by,

$$p_\omega(x|R) = \frac{2}{\pi R^2} \sqrt{R^2 - x^2} \tag{4.24}$$

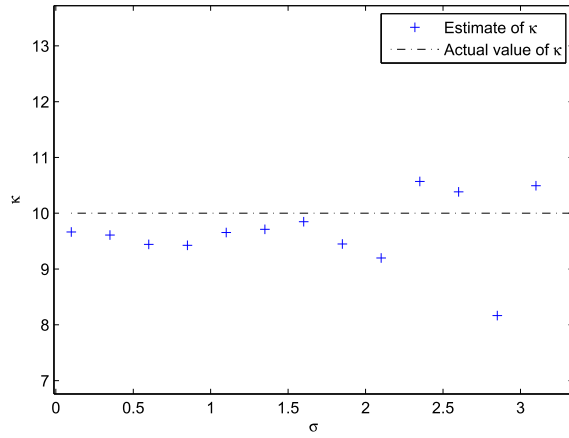


Figure 4.9: Estimation of  $\kappa$  for Von Mises distribution with  $K = L = P = 100$  and varying  $\sigma$

$I_2$  in this case is given by

$$I_2 = \frac{32}{3\pi R} \quad (4.25)$$

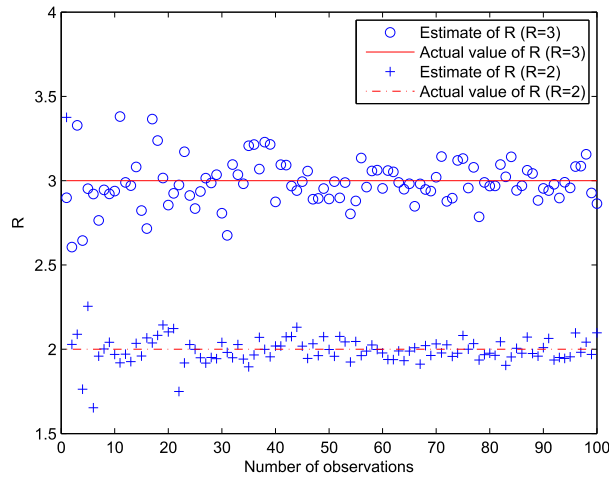


Figure 4.10: Estimation of  $R$  for wigner semi-circle with  $K = L = P = 100$  and  $\sigma^2 = 1$

Again by solving the  $2^{nd}$  equation of (4.16) and (4.25) we can estimate  $R$ .

In this case equation (4.16) is simplified to the equation (A.10) provided in the

Appendix and  $R$  is estimated as,

$$R = \frac{32c_2}{3\pi(m_2(\gamma) - c_1c_2 - 4c_1\sigma^2 - 2c_1\sigma^4 - 1)} \quad (4.26)$$

Figure 4.10 shows the result for  $R = 2$  and  $R = 3$  with  $K = L = P = 100$  and  $\sigma^2 = 1$  averaged over different number of observation.

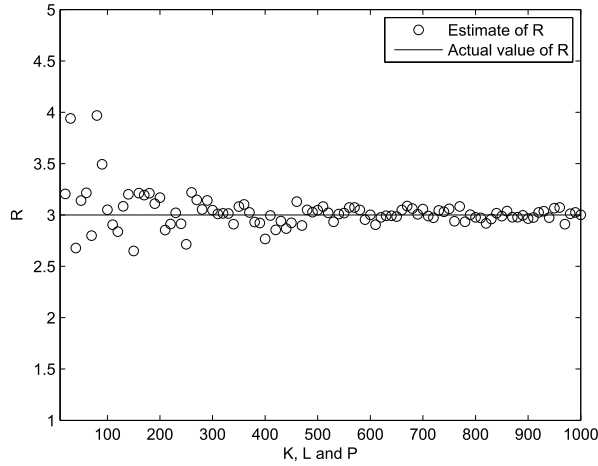


Figure 4.11: Estimation of  $R$  for wigner semi-circle with  $\sigma^2 = 1$  and varying  $K$ ,  $L$  and  $P$

Figure 4.11 presents the estimation of  $R$  as the value of  $K$ ,  $L$  and  $P$  increases asymptotically with  $\sigma^2 = 1$ . Figure 4.12 estimats the value of  $R$  for with varying  $\sigma$  and  $K = L = P = 1000$ .

#### 4.4.5 PDF Approximation

If we have the moments of the Vandermonde matrices we can extract information about the distribution of  $\omega$  in equation (4.3) by using the equation (4.15). We are calculating  $K_{\rho,u}$  statistically from (4.13).

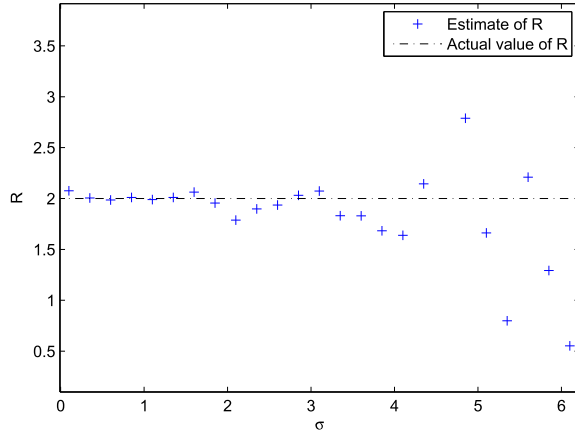


Figure 4.12: Estimation of  $R$  for wigner semi-circle with varying  $\sigma^2$  and  $K = L = P = 1000$

Moments approach give us the moments of the matrix  $\mathbf{V}\mathbf{V}^H$ . In this way we get the estimates for  $I_n$  which are related to the distribution of  $\omega$ .

### Beta Distribution

Now, consider a case where we have no information about the distribution of  $\omega$  apart from the fact that it is bounded between  $[0, 1]$ . Consider  $\omega$  to have a beta distribution with two degree of freedom. In this case the pdf is given by,

$$p_{\omega}(x) = \frac{x^{\alpha}(1-x)^{\beta}}{B(\alpha, \beta)} \quad (4.27)$$

where  $B$  is the beta function and the pdf has two degrees of freedom  $\alpha$  and  $\beta$ . Beta distribution was chosen for this simulation especially because for different values of parameter  $\alpha$  and  $\beta$  we obtain totally different shapes of the pdf. For  $n = t$  equating (4.19) with (4.16) gives us a set of non-linear equations of order  $n$ . Here we use Least Square Mean (LSM) based algorithm to calculate the unknowns.

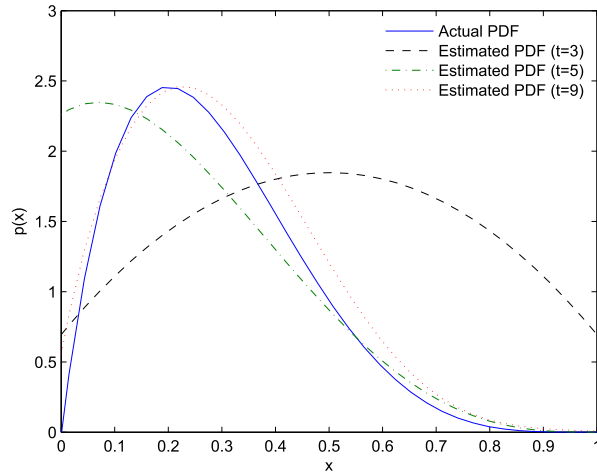


Figure 4.13: Estimation of the PDF of  $\omega$ , where  $\omega$  has a beta distribution with  $\alpha = 2$  and  $\beta = 5$

Figure 4.13 shows the simulation result for the estimation of the PDF of  $\omega$  with  $t = 3, 5, 9$  when  $\omega$  has a Beta distribution with  $\alpha = 2$  and  $\beta = 5$ . It is clear from the simulation result that as we increase the value of  $t$  we get accurate results. However note also that with the increase in  $t$  we are considering more and more moments  $n$  and as higher moments have some small error the estimation of the pdf will degrade after a certain value of  $t$ . In this case while designing and optimization algorithm to solve the unknowns more weightage must be given to the lower moment equations. Figure 4.14 represents the estimation of the PDF when  $\omega$  has beta distribution with  $\alpha = 1$ ,  $\beta = 3$  and  $t = 3, 5, 9$ . Similarly, Figure 4.15 shows the estimate of the beta distribution with  $\alpha = 1$ ,  $\beta = 3$  and  $t = 3, 5$ .

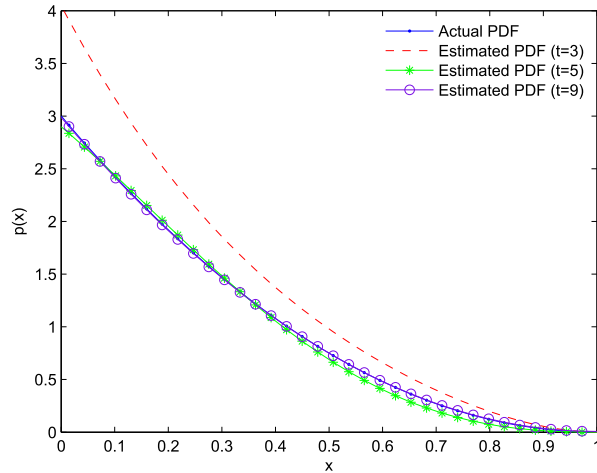


Figure 4.14: Estimation of the PDF of  $\omega$ , where  $\omega$  has a beta distribution with  $\alpha = 1$  and  $\beta = 3$

## 4.5 Chapter Conclusion

In this chapter we showed that by using the moments approach we can find an estimate for the distribution of the sensors without any communication between the sensors themselves. This is helpful when you have a large network of randomly deployed sensors and one needs to know their distribution. In the case where the distribution of the sensors is already known we can estimate the values of the parameters of the distribution or the range of the distribution.



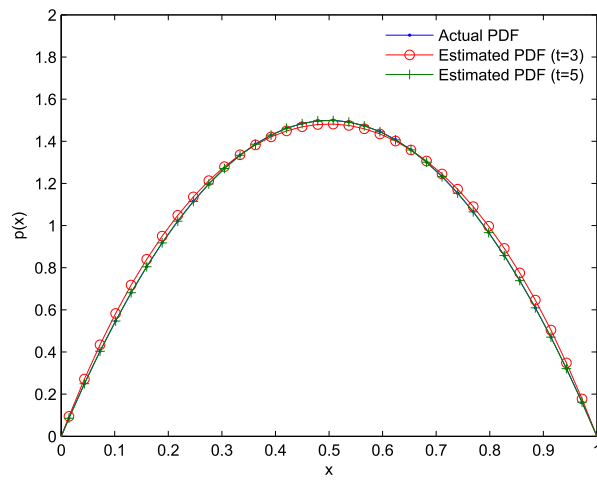


Figure 4.15: Estimation of the PDF of  $\omega$ , where  $\omega$  has a beta distribution with  $\alpha = 2$  and  $\beta = 2$

## CHAPTER 5

# APPLICATION TO SEISMIC SIGNAL PROCESSING

### 5.1 Introduction

Seismology plays a prominent role in the search of hydrocarbons. The seismic data has to be processed in order to reveal under ground images that leads to the identification of potential hydrocarbon reservoirs [49] [50]. Any information obtained about the different layers present below the surface is valuable [53] [52] [57].

Before going into the application of RMT for seismology first let us study the acquisition of seismic data. Then we will revise some of the tools used in RMT followed by the seismic modeling.

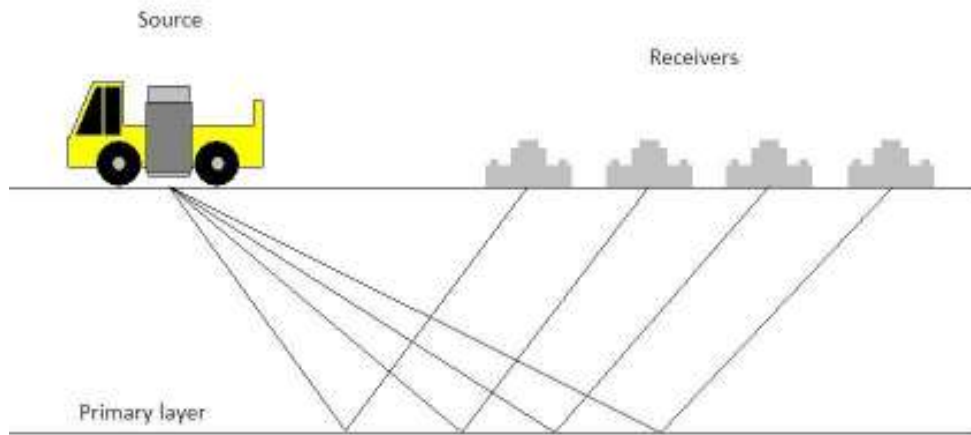


Figure 5.1: Common Source Point.

### 5.1.1 Data Acquisition - Common Shot Point

The seismic data model is based on the measurement of the travel time of seismic waves reflected by the subsurface layers at different velocities. Seismic energy is provided by a source located on the surface also known as a 'Shot'. Now there can be many kinds of shots depending on the topology of the terrain. For example in the shallow applications we use a hammer and a plate i.e. a weight is dropped or an explosive charge is ignited. Nowadays we also use large vibrators as an environmentally friendly source but an explosive source is still preferred. In the case of marine seismic data the source is usually an airgun. This source produces energy. The energy radiates out from the shot point in all directions. This energy is detected on surface using a linear array (or spread) of geophones spaced at regular intervals and the collection of such data obtained from a single source

point is called *Common Shot Point*. Observation of the travel-times of the direct and reflected signals provides information on the depth profile of the subsurfaces. During data acquisition individual shot records are displayed as variable area wiggle traces displaying travel time against distance. Figure 5.2 shows a typical wiggle plot of a seismic data. The data from a single geophone is stacked vertically. The horizontal axis represents the source receiver offset array and the vertical axis represents time.

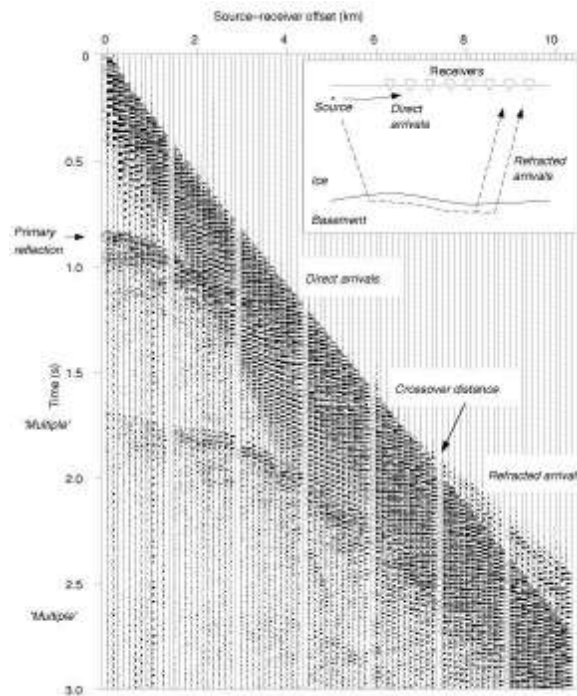


Figure 5.2: Seismic Data(wiggle display) courtesy of Yilmaz [50].

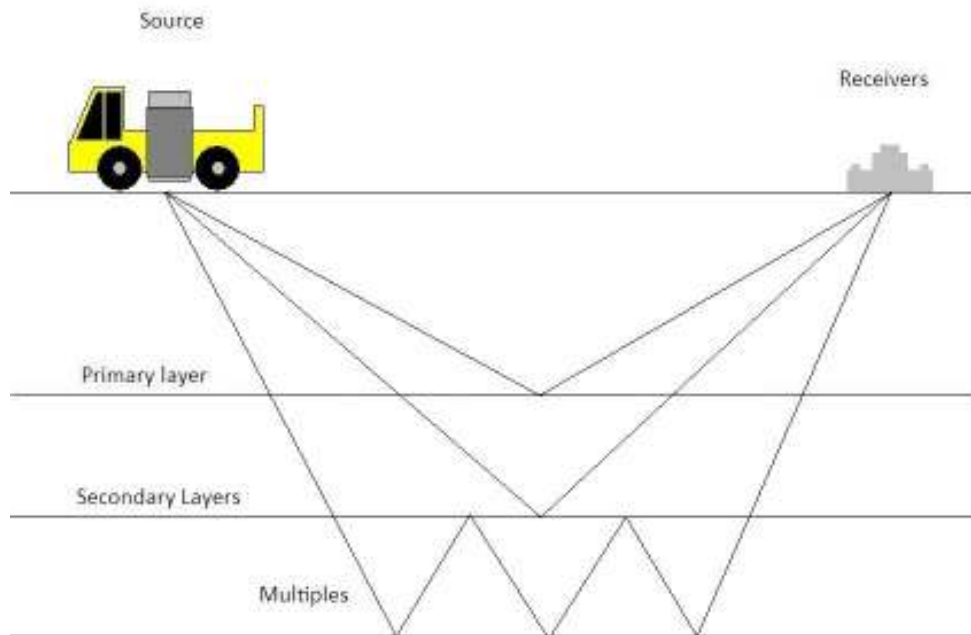


Figure 5.3: Multiple Layers.

### 5.1.2 Multiples

A downgoing event arriving later in time than the primary is defined as a multi-path arrival or simply multiple. Depending on their time delay from the primary events with which they are associated, multiples are characterized as short-path or peg-leg, implying that they interfere with the primary reflection, or long-path, where they appear as separate events. As seen in the Fig (5.3) due to multiple reflections between different layers, we are getting a signal that does not give us the information of the true layer.

## 5.2 Sample Covariance Matrix

Consider Figure 5.1 and assume that we collect  $N$  independent samples from an array of  $M$  equidistant sensors. Then the received signal  $\mathbf{y}(n)$  where  $n = 1, 2, \dots, N$  can be considered time slice of Figure 5.2 along the vertical axis. The sample covariance matrix is defined as the average sum of vector outer products  $\mathbf{y}(n)\mathbf{y}^H(n)$  where  $n = 1, 2, \dots, N$ .

$$\hat{\mathbf{R}}_M = \frac{1}{N} \sum_{n=1}^N \mathbf{y}(n)\mathbf{y}^H(n) \quad (5.1)$$

In seismic signal processing the covariance matrix is a many-to-one mapping [54]. It compresses voluminous data while retaining sufficient information for the estimation of the unknown parameters some of them will be discussed in this chapter. The authors of [48] provides an excellent overview of covariance analysis for seismic signal processing.

The matrix that we have access to  $\hat{\mathbf{R}}_M$  is the sample covariance matrix and does not represents the true covariance matrix  $\mathbf{R}_M$  because of the added noise in the received signal. Lets just say that we want to study the asymptotic behavior of a certain scalar function of  $\hat{\mathbf{R}}_M$ , lets call it  $\phi(\hat{\mathbf{R}}_M)$ . Traditionally, when  $N \rightarrow \infty$  and  $M$  is fixed,  $\hat{\mathbf{R}}_M$  is a good estimate of  $\mathbf{R}_M$ , i.e.,  $|\phi(\hat{\mathbf{R}}_M) - \phi(\mathbf{R}_M)| \rightarrow 0$ . However when  $M$  and  $N$  have the same order of magnitude and when considering the limits  $N, M \rightarrow \infty$ ,  $M/N \rightarrow c$ ,  $0 < c < \infty$  then  $|\phi(\hat{\mathbf{R}}_M) - \phi(\mathbf{R}_M)| \not\rightarrow 0$  [59].

Suppose that we needed to estimate the eigenvalues or the eigenvector of the covariance matrix. Traditionally, we would use the sample covariance matrix to estimate the eigenvalues or eigenvectors but the RMT provides us with other tools to estimate such parameters.

### 5.3 Asymptotic Behavior of Sample Covariance Matrix

Let the eigenvalues and eigenvectors of the sample covariance matrix  $\hat{\mathbf{R}}_M$  be  $\hat{\lambda}_1 \leq \hat{\lambda}_2 \leq \dots \leq \hat{\lambda}_M$  and  $\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \dots, \hat{\mathbf{e}}_M$  respectively. RMT states that if certain assumptions are met, the empirical eigenvalues of the sample covariance matrix converge (as  $N, M \rightarrow \infty$ ) to a non-random distribution. One can then characterize the asymptotic behavior of the eigenvalues of the true covariance matrix  $\mathbf{R}_M$  in terms of the sample covariance matrix  $\hat{\mathbf{R}}_M$ . Consider Stieltjes transforms [55] of the sample covariance matrix which are related to their eigenvalue densities,

$$\hat{m}_M(z) = \mathbf{s}^H (\hat{\mathbf{R}}_M - z\mathbf{I}_M)^{-1} \mathbf{s}, \quad \hat{b}_M(z) = \frac{1}{M} \text{tr} [(\hat{\mathbf{R}}_M - z\mathbf{I}_M)^{-1}] \quad (5.2)$$

And the corresponding transforms of the true covariance matrix will be,

$$m_M(z) = \mathbf{s}^H (\mathbf{R}_M - z\mathbf{I}_M)^{-1} \mathbf{s}, \quad b_M(z) = \frac{1}{M} \text{tr} [(\mathbf{R}_M - z\mathbf{I}_M)^{-1}] \quad (5.3)$$

where  $z \in \mathbb{C}^+ = \{z \in \mathbb{C} : \text{Im}(z) > 0\}$ ,  $\mathbf{s}$  is a deterministic  $M \times 1$  column vector. Consider the following assumptions

**(Asp1)** The covariance matrix  $\mathbf{R}$  has uniformly bounded spectral radius for all  $M$  and  $\sup_M \{\|s\|\} < \infty$

**(Asp2)** The received vector  $\mathbf{y}(n)$  can be expressed as  $\mathbf{y}(n) = \mathbf{R}^{1/2}\mathbf{u}(n)$ , where  $\mathbf{u}(n), n = 1, 2, \dots, N$  are i.i.d. complex random vectors and  $\mathbf{R}^{1/2}$  is a positive definite Hermitian deterministic matrix.

Under the above two assumptions,  $\hat{m}_M(z)$  and  $\hat{b}_M(z)$  are close to their respective deterministic counterparts  $\bar{m}_M(z)$  and  $\bar{b}_M(z)$ , therefore we have,

$$|\hat{m}_M(z) - \bar{m}_M(z)| \rightarrow 0, \quad |\hat{b}_M(z) - \bar{b}_M(z)| \rightarrow 0 \quad (5.4)$$

where  $\hat{b}_M(z) = \hat{b}$  is the unique solution to the following set of equations  $\{b \in \mathbb{C} : -(1-c)/z + c\hat{b} \in \mathbb{C}^+\}$

$$\hat{b} = \frac{1}{M} \sum_{r=1}^M \frac{1}{\lambda(1-c-cz\hat{b}) - z} \quad (5.5)$$

and

$$\hat{m}_M(z) = \sum_{r=1}^M \frac{\mathbf{s}_1^H \mathbf{e}_r \mathbf{e}_r^H \mathbf{s}_2}{\lambda_r(1-c-cz\hat{b}_M(z)) - z} \quad (5.6)$$

Now we will consider another assumption

**(Asp3)** The number of samples per sensor should be greater than a pre-defined value given by,

$$\frac{N}{M} > \frac{1}{M} \sum_{m=1}^M \left( \frac{\hat{\lambda}_m}{\hat{\lambda}_m - \xi} \right)^2 \quad (5.7)$$



where  $\xi$  is the smallest real valued solution to the equation.

$$\frac{1}{M} \sum_{m=1}^M \frac{\hat{\lambda}_M^2}{(\hat{\lambda}_m - \xi)^3} = 0 \quad (5.8)$$

Now lets consider a traditional estimator that is related to the sample eigenvector and is defined as,

$$\hat{\eta}_m = \mathbf{s}_1^H \hat{\mathbf{e}}_k \hat{\mathbf{e}}_k^H \mathbf{s}_2 \quad (5.9)$$

again where  $\mathbf{s}_1$  and  $\mathbf{s}_2$  are two deterministic vectors.

**(Asp4)**  $\mathbf{s}_1$  and  $\mathbf{s}_2$  have uniformly bounded norms for all  $M$ .

Under assumptions Asp1-Asp4 [60] [61] we have,

$$\left| \hat{\eta}_m - \sum_{k=1}^M \left( \frac{1}{2\pi j} \oint_{\partial \mathbb{R}_y^-(m)} d_k(z) dz \right) \mathbf{s}_1^H \mathbf{e}_k \mathbf{e}_k^H \mathbf{s}_2 \right| \rightarrow 0 \quad (5.10)$$

where  $\partial \mathbb{R}_y^-(m)$  is the negatively oriented contour around a rectangle defined as,

$$\mathbb{R}_y(m) = \{z \in \mathbb{C} : \sigma_1 \leq \text{Re}[z] \leq \sigma_2, |\text{Im}[z]| \leq y\} \quad (5.11)$$

here  $\sigma_1$  and  $\sigma_2$  are such that  $(\sigma_1, \sigma_2)$  encloses the eigenvalue cluster relating to  $\lambda_m$

and  $d_k(z)$  is,

$$d_k(z) = \frac{1}{\hat{\lambda}_k(1 - c - cz\bar{b}_M(z)) - z} \quad (5.12)$$

[61] simplifies the expression for  $\hat{\eta}_m$  as,

$$\frac{1}{2\pi j} \oint_{\partial\mathbb{R}_y^-(m)} d_k(z)dz = \frac{1}{2\pi j} \oint_{C^-(m)} \chi_k(\omega)d\omega \quad (5.13)$$

here,

$$\chi_k(\omega) = \frac{1}{\hat{\lambda}_k - \omega} \frac{1 - c \frac{1}{M} \sum_{r=1}^M \left(\frac{\hat{\lambda}_r}{\hat{\lambda}_r - \omega}\right)^2}{1 - c \frac{1}{M} \sum_{r=1}^M \left(\frac{\hat{\lambda}_r}{\hat{\lambda}_r - \omega}\right)} \quad (5.14)$$

we solve the above integral by using the residue theorem [62] and we get,

$$\bar{\eta}_m = \sum_{k=1}^M w_m(k) \mathbf{s}_1^H \mathbf{e}_k \mathbf{e}_k^H \mathbf{s}_2 \quad (5.15)$$

here,

$$w_m(k) = \begin{cases} 1 - \sum_{r=1, r \neq m}^M \left( \frac{\hat{\lambda}_m}{\hat{\lambda}_r - \lambda_m} - \frac{\mu_m}{\hat{\lambda}_r - \mu_m} \right), & k = m \\ \frac{\hat{\lambda}_m}{\hat{\lambda}_r - \lambda_m} - \frac{\mu_m}{\hat{\lambda}_r - \mu_m}, & k \neq m \end{cases} \quad (5.16)$$

here  $\mu_m$  is the  $m^{\text{th}}$  solution to the equation,

$$\frac{1}{M} \sum_{r=1}^M \frac{\hat{\lambda}_r}{\hat{\lambda}_r - \mu} = \frac{1}{c} \quad (5.17)$$

Suppose now that instead of a single eigenvector  $\mathbf{e}_k$  we have two or more eigenvectors stacked together  $\mathbf{E}_k$ , in this case  $\mathbf{e}_k$  in (5.15) will be replaced by  $\mathbf{E}_k$  and we would get a different weighting factor  $w_k(m)$  by solving (5.14) over a different value of  $\sigma_1$  and  $\sigma_2$  spanning the respective eigenvalues of eigenvectors in  $\mathbf{E}_k$ .

Now we will use results from RMT summarized above for seismic signal processing,

such as the estimation of direction of arrival, velocity and zero offset time. We divide the received seismic data into a narrow band signal and a wide band signal and then model them accordingly.

## 5.4 Narrow Band Case

### 5.4.1 System Model

If there is only one reflection then the signal received at the array of  $M$  sensors is just a replica of the original source signal arriving at the receivers at different time delays. We will see later in this chapter that these time delays can be translated into direction of arrival. Lets assume that there are total of  $K$  reflections, i.e., the received signal will be a sum of  $K$  reflected signals each arriving a sensor at a time delay of  $\tau_i$ , where  $i = 1, 2, \dots, K$ . The narrow band models take advantage of the fact that time delays can be translated into phase shifts in frequency domains [63]. These shifts in time and frequency domain of a single trace are defined as,

$$s(t - \tau) \leftrightarrow S(\omega)e^{-j\omega\tau} \quad (5.18)$$

Here,  $s(t - \tau)$  is the shifted version of the reflected signal. In the narrow band case, the bandwidth of the reflected signal is small enough relative to the center frequency  $\omega_c$  that the phase shift is approximately constant over the bandwidth. This approximation can be shown as,

$$S(\omega)e^{-j\omega\tau} \approx S(\omega)e^{-j\omega_c\tau} \leftrightarrow s(t)e^{-j\omega_c t} \quad (5.19)$$

Now, we write the received signal from the linear array of sensors as the weighted sum of steering vectors with added noise.

$$\mathbf{y}(t) = [\mathbf{v}(\theta_1)\mathbf{v}(\theta_2) \dots \mathbf{v}(\theta_K)]s(t) + \sigma\mathbf{n}_s(t) \quad (5.20)$$

Here,  $\mathbf{n}_s(t)$  is a vector of independent gaussian noise at all  $M$  sensors at time  $t$ ,  $\sigma$  is the variance of the noise and  $\mathbf{v}(\theta_i)$  is a  $M \times 1$  steering vector given by,

$$\mathbf{v}(\theta_i) = [1 \ e^{-j\omega_i d/\lambda_c \sin \theta_i} \dots e^{-j\omega_i(M-1)d/c \sin \theta_i}]^T \quad i = 1, 2, \dots, K \quad (5.21)$$

here  $d$  is the distance between the sensors,  $\lambda_c$  is the wavelength,  $\theta_i$  is the direction of arrival (DoA). The discrete model will look something like,

$$\mathbf{y}(n) = \mathbf{V}\mathbf{s}(n) + \sigma\mathbf{n}_s(n), \quad n = 1, 2, \dots, N \quad (5.22)$$

Here,  $\mathbf{V}$  is the Vandermonde matrix. If we stack all the time snaps of the received signal, the new model will then be represented as,

$$\mathbf{Y} = \mathbf{V}\mathbf{S} + \sigma\mathbf{N}_s \quad (5.23)$$

where  $\mathbf{Y}$  is a  $M \times K$  matrix  $[\mathbf{y}(1), \dots, \mathbf{y}(N)]$ .  $\mathbf{S}$  is a  $K \times N$  matrix  $[\mathbf{s}(1), \dots, \mathbf{s}(N)]$

and  $\mathbf{N}_s$  is a  $M \times N$  i.i.d. gaussian matrix  $[\mathbf{n}_s(1), \dots, \mathbf{n}_s(N)]$ . The matrices involved are of large dimension with  $c = \lim_{N \rightarrow \infty} \frac{M}{N}$ ,  $0 < c < \infty$ .

The data covariance matrix is defined as,

$$\mathbf{R} = \mathbf{V}\mathbf{P}_x\mathbf{V}^H + \mathbf{R}_N \quad (5.24)$$

where  $\mathbf{P}_x = E\{\mathbf{ss}^H\}$  is the source covariance matrix, it is a diagonal matrix containing energy of each arriving wavelet. The seismic wavelet is a small wave or ripple produced by the source, here it refers to the attenuated replica of the original source wave as it is reflected by different layers.  $\mathbf{R}_N = E\{\mathbf{n}_s\mathbf{n}_s^H\}$  is the noise covariance matrix. When the noise is white  $\mathbf{R}_N = \sigma^2\mathbf{I}$  we have,

$$\mathbf{R} = \mathbf{V}\mathbf{P}_x\mathbf{V}^H + \sigma^2\mathbf{I} \quad (5.25)$$

but in practice one has only access to the sample covariance matrix given by,

$$\hat{\mathbf{R}} = \frac{\mathbf{Y}\mathbf{Y}^H}{N} \quad (5.26)$$

### 5.4.2 High Resolution Spectral Estimators

When dealing with seismic data, we often consider extracting information from the frequency domain representation of the signal. We apply FFT to either temporal or spacial signal samples. However dealing with FFTs of voluminous seismic data has some drawbacks like we end up with more data than is required, FFT is

not ideal for small transient signals with few samples and the spectral resolution is limited by the sampling time of the data [51]. An other useful tool that we use are spectral estimators. They either estimate the Direction of arrival or the frequency of the incoming signal. There are many such estimators like minimum variance distortionless response (MVDR), the MUltiple SIgnal Classification (MUSIC) algorithm [64], enhanced minimum variance, maximum entropy [66] and minimum norm estimators.

These estimators traditionally use sample covariance matrix to estimate the unknown parameter. In [67], a modified form of MUSIC algorithm called GMUSIC to estimate the direction of arrivals. The results show that when the order of the matrix is small ( $M$  and  $N$  are comparable) GMUSIC performs better but as  $N \rightarrow \infty$  as  $M$  is fixed the results converge for both estimators.

In the next section we are going to apply the same technique on a different spectral estimator and evaluate its performance.

### 5.4.3 Minimum Norm

The minimum norm spectral estimator was initially introduced by Kumaresan and Tufts [56]. It adds a polynomial form constraint by minimizing the norm of a linear combination of noise subspace and scales the first element of the noise space eigenvector to unity. The traditional minimum norm estimator  $P_{MN}$  is given by,

$$P_{MN}(\theta) = \frac{\mathbf{a}(\theta)^H \mathbf{Q}_n \mathbf{Q}_n^H \mathbf{1}_1 \mathbf{1}_1^H \mathbf{Q}_n \mathbf{Q}_n^H \mathbf{a}(\theta)}{(\mathbf{q}^H \mathbf{q})^2} \quad (5.27)$$

where  $\mathbf{a}(\theta) = [1 \ e^{-j\omega d/\lambda_c \sin \theta} \ \dots \ e^{-j\omega(M-1)d/\lambda_c \sin \theta}]^T$ ,  $\mathbf{Q}_n$  is the noise space eigenvectors matrix,  $\mathbf{q} = [q_{(1,1)}q_{(1,2)}, \dots, q_{(1,K)}]^T$  is the vector of the first elements of noise eigenvectors and  $\mathbf{1}_1 = [100 \dots 0]^T$  of length  $M$ .

Now we apply RMT approach to obtain a different estimator. We solve (5.13) where the interval  $(\sigma_1, \sigma_2)$  encloses the eigenvalue cluster only relating to noise eigenvalues. With a little simplification we come up with,

$$P_{GMN}(\theta) = \frac{\mathbf{a}(\theta)^H \mathbf{Q} \mathbf{W} \mathbf{Q}^H \mathbf{1}_1 \mathbf{1}_1^H \mathbf{Q} \mathbf{W} \mathbf{Q}^H \mathbf{a}(\theta)}{\mathbf{1}_1^H \mathbf{Q} \mathbf{W} \mathbf{Q}^H \mathbf{1}_1} \quad (5.28)$$

where  $\mathbf{W}$  is a diagonal matrix with diagonal entries given by,

$$w(m, m) = \begin{cases} 1 + \sum_{k=M-K+1}^M \left( \frac{\hat{\lambda}_m}{\hat{\lambda}_r - \lambda_m} - \frac{\mu_m}{\hat{\lambda}_r - \mu_m} \right), & m \leq M - K \\ - \sum_{k=1}^{M-K} \frac{\hat{\lambda}_m}{\hat{\lambda}_r - \lambda_m} - \frac{\mu_m}{\hat{\lambda}_r - \mu_m}, & m > M - K \end{cases} \quad (5.29)$$

here  $\mu_m$  is the  $m^{\text{th}}$  solution to the equation,

$$\frac{1}{M} \sum_{r=1}^M \frac{\hat{\lambda}_r}{\hat{\lambda}_r - \mu} = \frac{1}{c} \quad (5.30)$$

By searching for the local maxima of  $P_{MN}^{-1}$  and  $P_{GMN}^{-1}$  we estimate the Direction of arrival (DoAs).

## Simulation Results

The reflected signal  $\mathbf{s}(n)$  is generated as a Ricker Wavelet. As seen from Figure 5.4(a) the wavelet starts at a particular time  $t_0$  and after some times dies again.

It is important for a narrow band signal that the reflected wave has a very narrow bandwidth and the whole bandwidth can be approximated by its center frequency. Figure 5.4(b) shows that the center frequency of the wavelet generated is 5Hz.

Seismic data is generated according to the model (5.23) where  $\mathbf{V}$  is the vandermonde matrix and the rows of  $\mathbf{S}$  are the reflected ricker wavelets. Now, consider that there are only two reflections  $K = 2$  coming at an angle of  $35^\circ$  and  $37^\circ$  with a center frequency of 5Hz, in fact they superimpose each other, i.e., their time of arrival is almost the same. Since it is a narrow band signal, the time delay between adjacent receiver is very small. Figure 5.5 shows a noise free synthetic data gathered by 20 sensors with a sampling rate of  $10ms$  and gathers a total of 40 samples. As seen from the figure, it is hard to differentiate or even distinguish the two waves with a naked eye. It looks like a single reflected wave.

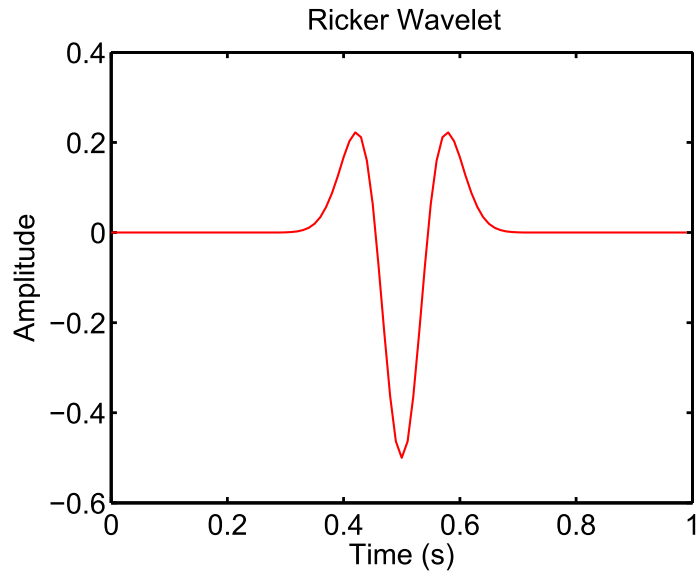
For the sake of comparison Figure 5.6 and Figure 5.7 are noise added models of Figure 5.5 with  $SNR$  of 10dB and 5dB respectively.

Figure 5.8 shows the comparison between the Min Norm and G-Min Norm estimators when  $M = 20$ ,  $N = 20$   $d/\lambda_c = 0.5$  and  $SNR = 10dB$ . As it can be seen from Figure 5.8 when  $M$  and  $N$  are of comparable magnitude, the G-Min norm estimator detects two direction of arrivals while the traditional estimator detects a single direction of arrival.

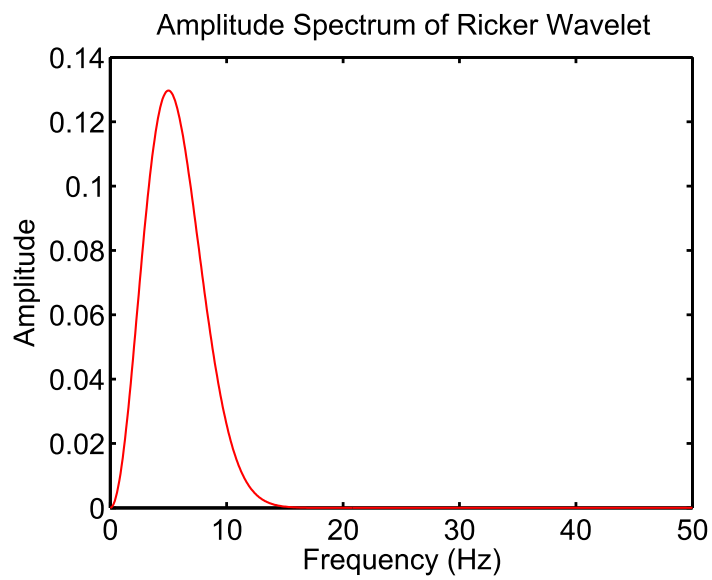
We now increase the sample size from  $N = 20$  to  $N = 100$ . Figure 5.9 shows that the Min Norm and G-Min Norm estimators give the same result.

The convergence of the traditional estimator and G-estimator as  $N$  increases





(a) Temporal representation of Ricker Wavelet



(b) Frequency Spectrum

Figure 5.4: Ricker Wavelet

can be seen from Figure 5.10.

Figure 5.11 plots the MSE (sum of the variance and the squared bias of the estimator) with respect to  $SNR$ . It can be clearly seen that the MSE of G-Min Norm estimator is less than that of Min Norm estimator.

## 5.5 Wide Band Case

We will be comparing the results of MUSIC and GMUSIC algorithms for the wideband seismic case by estimating the velocities of the coming wavefronts. The spectral estimators to be maximized are  $P_{MUSIC}^{-1}$  and  $P_{GMUSIC}^{-1}$  and are given by [67],

$$P_{MUSIC}(V) = \mathbf{a}^H(V) \mathbf{Q}_n \mathbf{Q}_n^H \mathbf{a}(V) \quad (5.31)$$

$$P_{GMUSIC}(V) = \mathbf{a}^H(V) \mathbf{Q} \mathbf{W} \mathbf{Q}^H \mathbf{a}(V) \quad (5.32)$$

where  $\mathbf{W}$  is given by (5.29).

### 5.5.1 Before Preflattening

The seismic wavefronts arrive to the equispaced horizontal geophone array with the model provided by (5.22). However the delay at the  $m^{th}$  sensor is not an integral multiple of any intersensor delay, in fact the delay at the  $m^{th}$  sensor follows a hyperbolic path. The time delay  $\tau_{mk}$  of the  $k^{th}$  wavefront at the  $m^{th}$

sensor is approximated by [48],

$$\tau_{m,k} = \frac{((m-1)d)^2}{2T_0V_k^2} \quad (5.33)$$

where  $T_0$  is the vertical two way travel time between the sensor and the reflected surface and  $V_k$  is the velocity of the  $k^{\text{th}}$  wavefront. The delay column vector in (5.22) becomes  $\mathbf{v}_k = [1 \ e^{-j\omega_k\tau_{2,k}} \ \dots \ e^{-j\omega_k\tau_{M,k}}]^T$ .

If the time delay  $\tau_{m,k}$  from one end of the array to the other is not small enough compared to the temporal correlation time, then the assumption of narrow band fails to hold [58]. In this case the wavefronts lose energy into dimension of space other than the ideal rank- $K$  dimension of the plane wave. This loss in energy appears as colored noise in the near diagonal elements of the covariance matrix.

### 5.5.2 After Preflattening

As stated earlier that the narrowband assumption does not work well when the time delay from one sensor to the other is large. We can reduce this delay by preflattening the data. In preflattening we shift the data set in a hyperbolic manner. Consider that we have two wavefronts with velocities  $V_1$  and  $V_2$  then we preflatten the data to an approximate velocity  $V_a$  which lies in between  $V_1$  and  $V_2$  by a shift given by [48],

$$\tau_m(V) = \sqrt{T_0^2 + (m-1)d/V^2} - \sqrt{T_0^2 + (m-1)d/V_a^2} \quad (5.34)$$

### 5.5.3 Simulation Results

The synthetic data generated for the wideband case was generated by adding the wavelets with a time delay given by equation (5.33). Figure 5.12 shows two ricker wavelets (both with center frequency  $f_c = 30Hz$ ) traveling at velocities  $V_1 = 3600m/s$  and  $V_2 = 7000m/s$  respectively. Here we have  $M = 20$ ,  $N = 300$  and  $d = 60m$ . It is clear from the figure that it does not fall under the category of narrow band signal.

Figure 5.13 shows the two estimators. MUSIC algorithm is performing as expected and shows multiple peaks as if there were many narrowband wavefronts. However it is interesting to note the performance of GMUSIC, the first two peak do give us a good estimate of the velocities.

Now consider that the two incoming wavelets have velocities  $V_1 = 2700m/s$  and  $V_2 = 3600m/s$ , both with ricker venter frequency of  $f_c = 30Hz$ . Here  $M = 20$ ,  $N = 200$  and the data is preflattened with an approximate velocity of  $V_a = 3000m/s$ . Figure 5.14 is the wigger plot after preflattening

Now the performance of MUSIC and GMUSIC algorithm is as expected from Figure 5.15

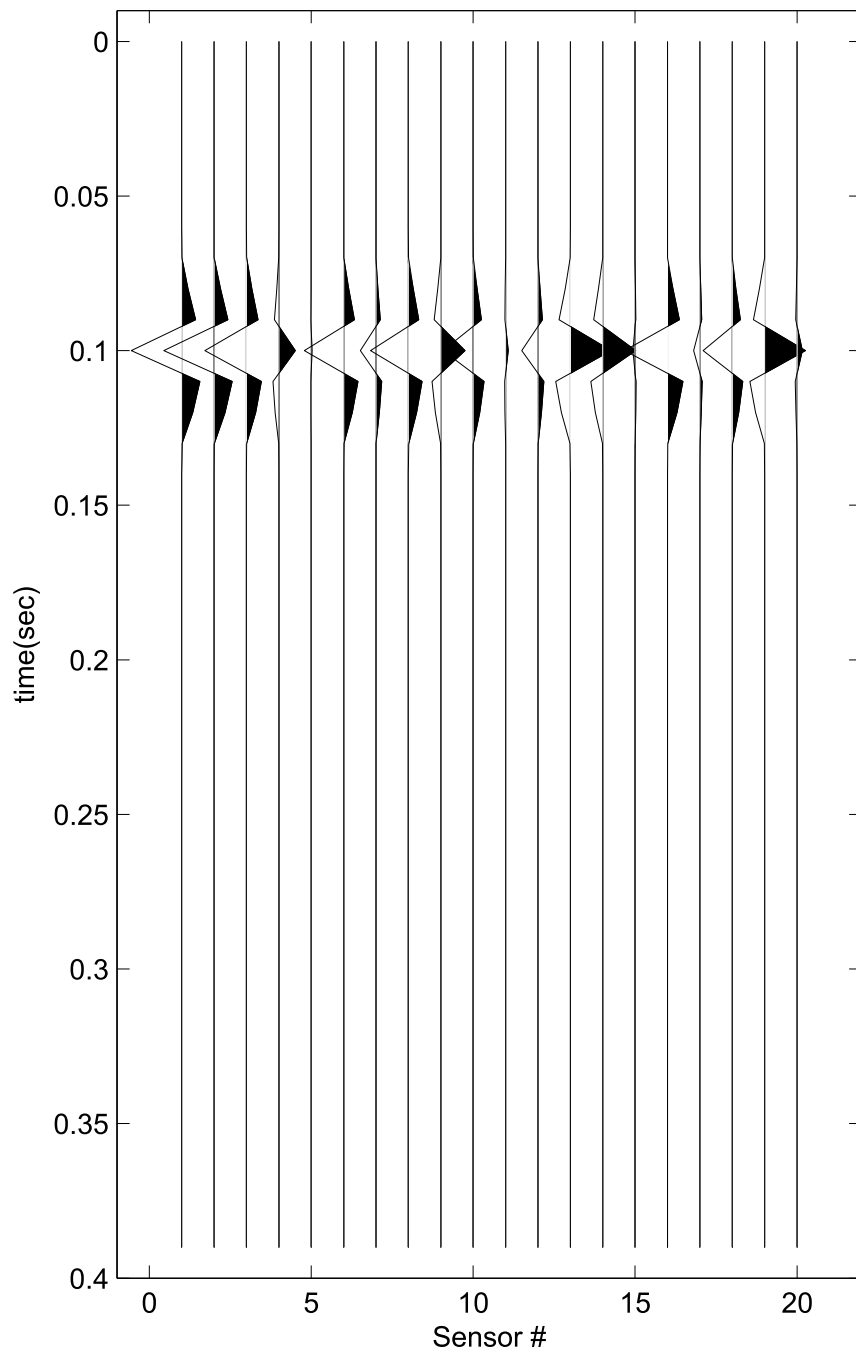


Figure 5.5: Wigner Plot of the synthetic noise free received data with  $M = 20$ ,  $N = 40$ ,  $f_c = 5\text{Hz}$ ,  $\theta_1 = 35^\circ$  and  $\theta_2 = 37^\circ$

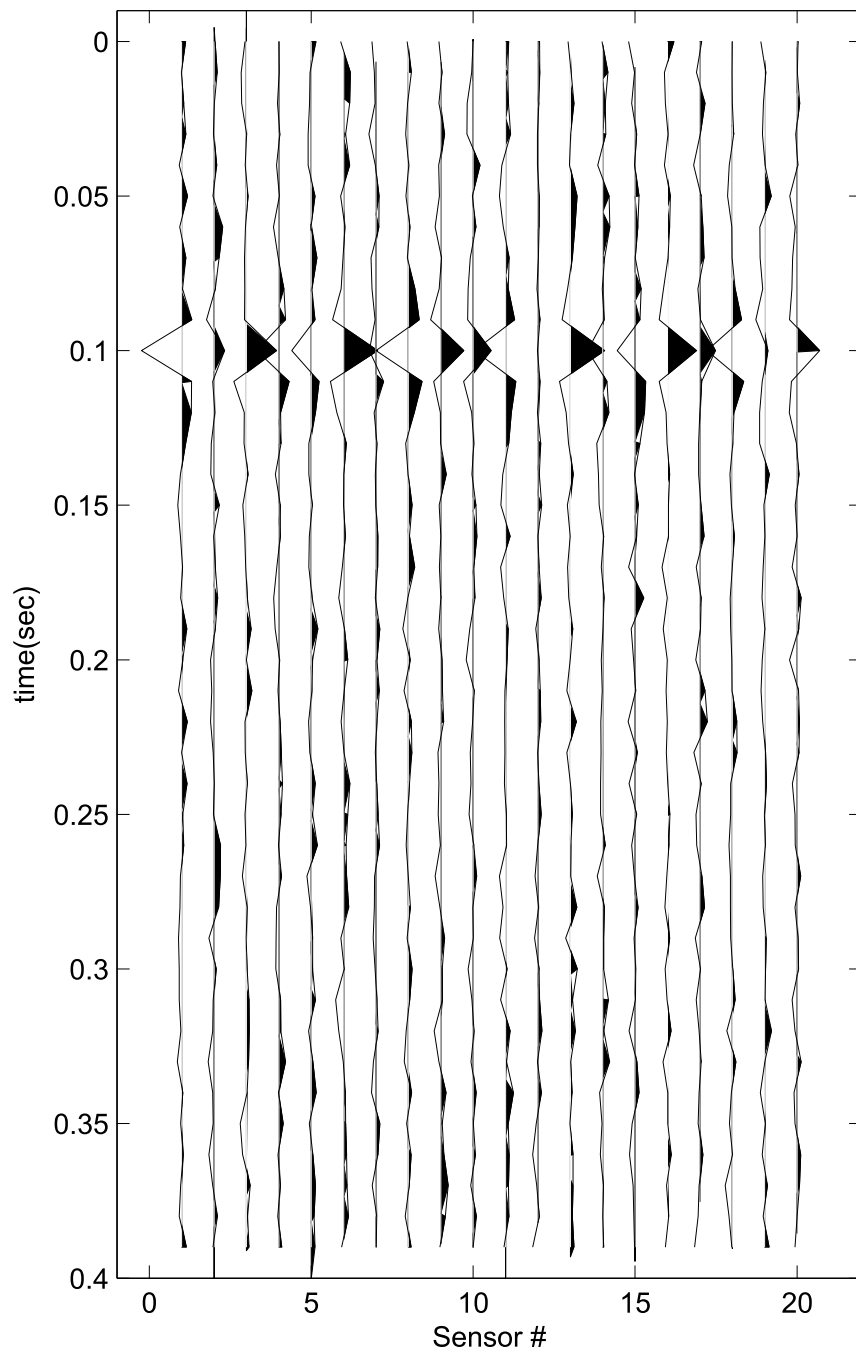


Figure 5.6: Wigner Plot of the synthetic received data with  $M = 20$ ,  $N = 40$ ,  $\theta_1 = 35^\circ$ ,  $f_c = 5\text{Hz}$ ,  $\theta_2 = 37^\circ$  and  $SNR = 10\text{dB}$

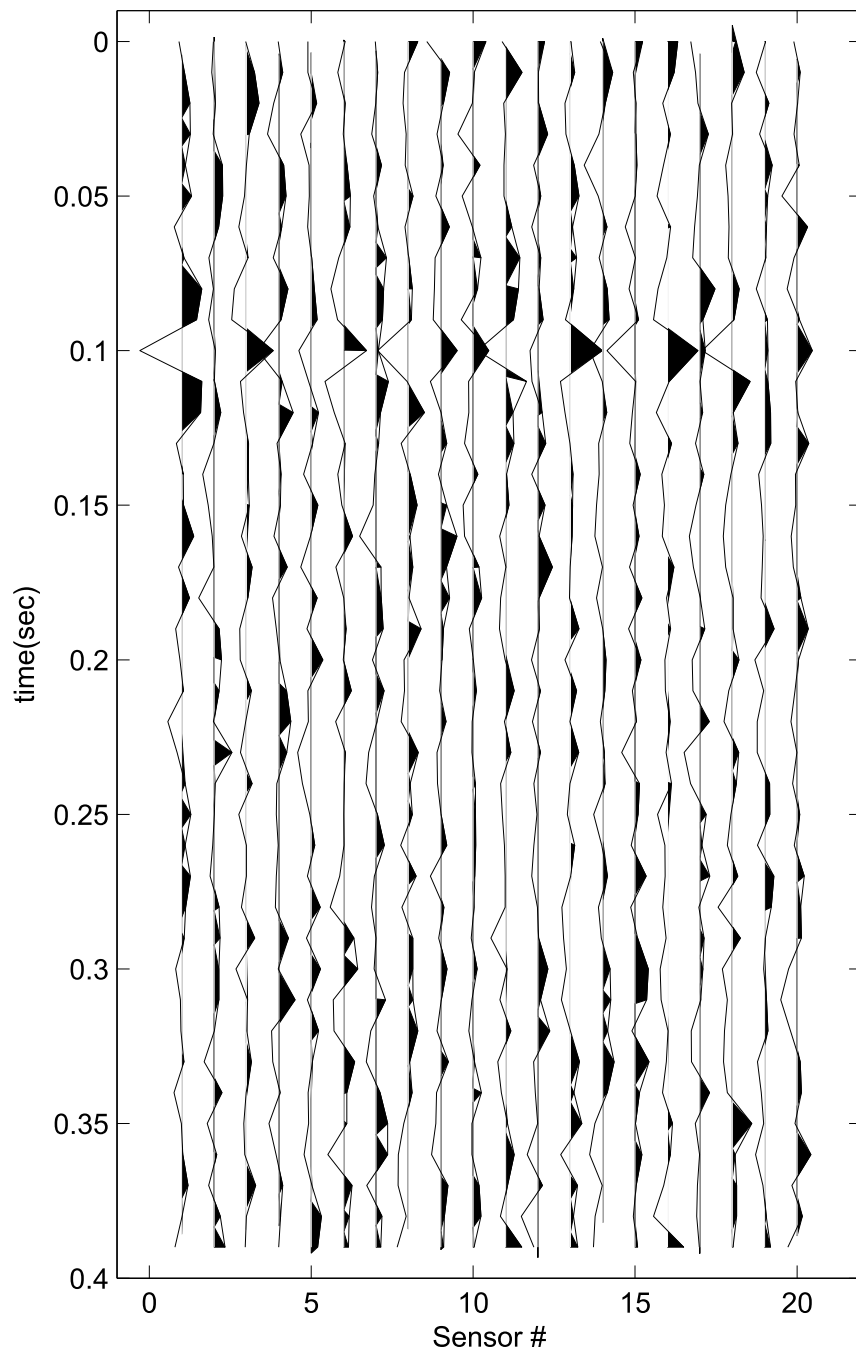


Figure 5.7: Wigner Plot of the synthetic received data with  $M = 20$ ,  $N = 40$ ,  $\theta_1 = 35^\circ$ ,  $f_c = 5\text{Hz}$ ,  $\theta_2 = 37^\circ$  and  $SNR = 5\text{dB}$

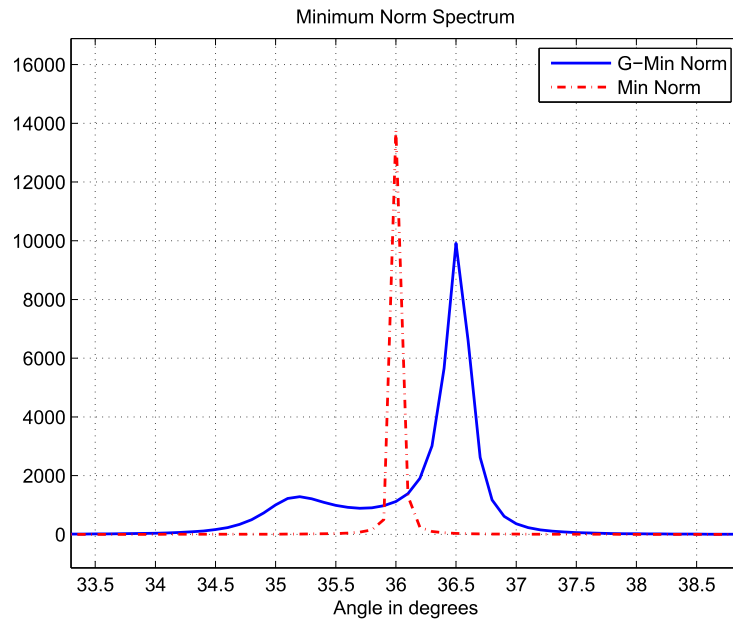


Figure 5.8: Estimation of Direction of Arrival (DOA) with  $\theta_1 = 35^\circ$ ,  $\theta_2 = 37^\circ$ ,  $f_c = 5\text{Hz}$ ,  $M = 20$ ,  $N = 20$ ,  $d/\lambda_c = 0.5$ , Empirical  $SNR = 10\text{dB}$

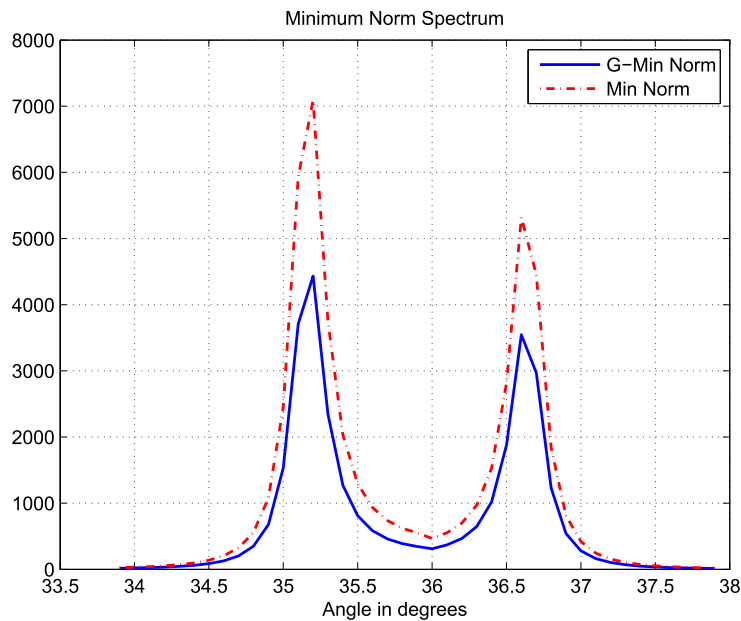


Figure 5.9: Estimation of Direction of Arrival (DOA) with  $\theta_1 = 35^\circ$ ,  $\theta_2 = 37^\circ$ ,  $f_c = 5\text{Hz}$ ,  $M = 20$ ,  $N = 100$ ,  $d/\lambda_c = 0.5$ , Empirical  $SNR = 10\text{dB}$



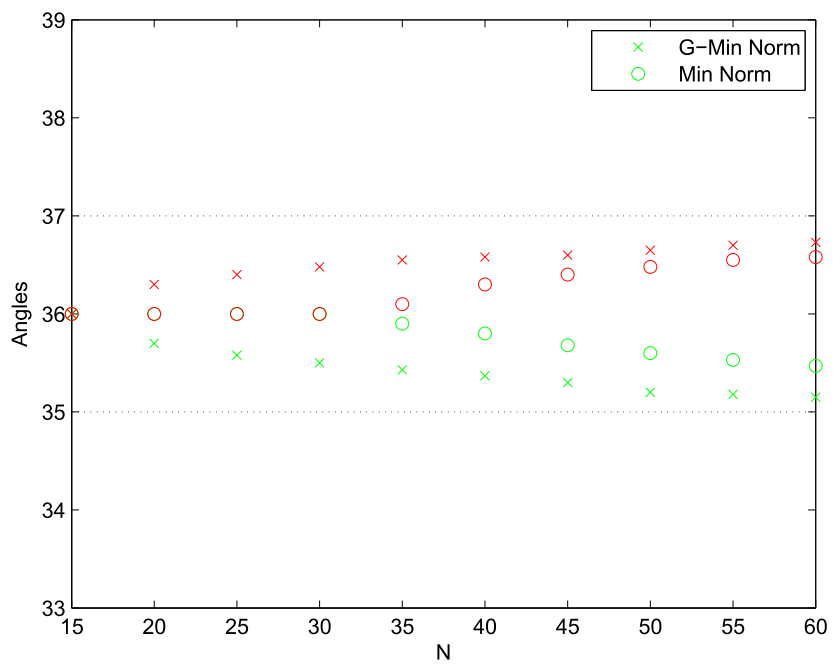


Figure 5.10: Estimation of Direction of Arrival (DOA) with varying  $N$ ,  $\theta_1 = 35^\circ$ ,  $\theta_2 = 37^\circ$ ,  $M = 20$  and empirical  $SNR = 10dB$

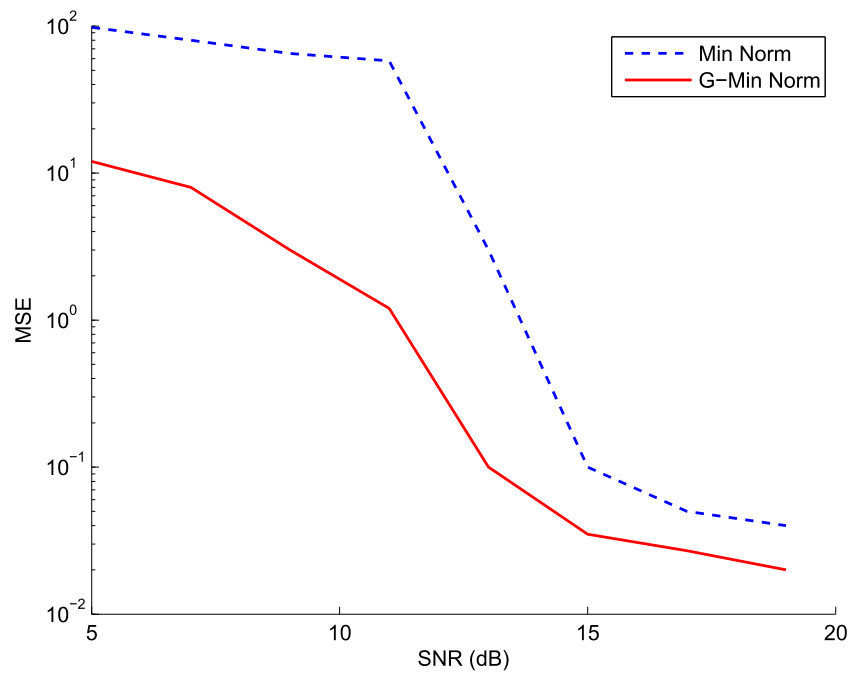


Figure 5.11: Minimum Square Error for the direction of arrival at different  $SNRs$  with  $\theta_1 = 35^\circ$ ,  $\theta_2 = 37^\circ$ ,  $f_c = 5\text{Hz}$ ,  $M = 20$ ,  $N = 20$ ,  $d/\lambda_c = 0.5$

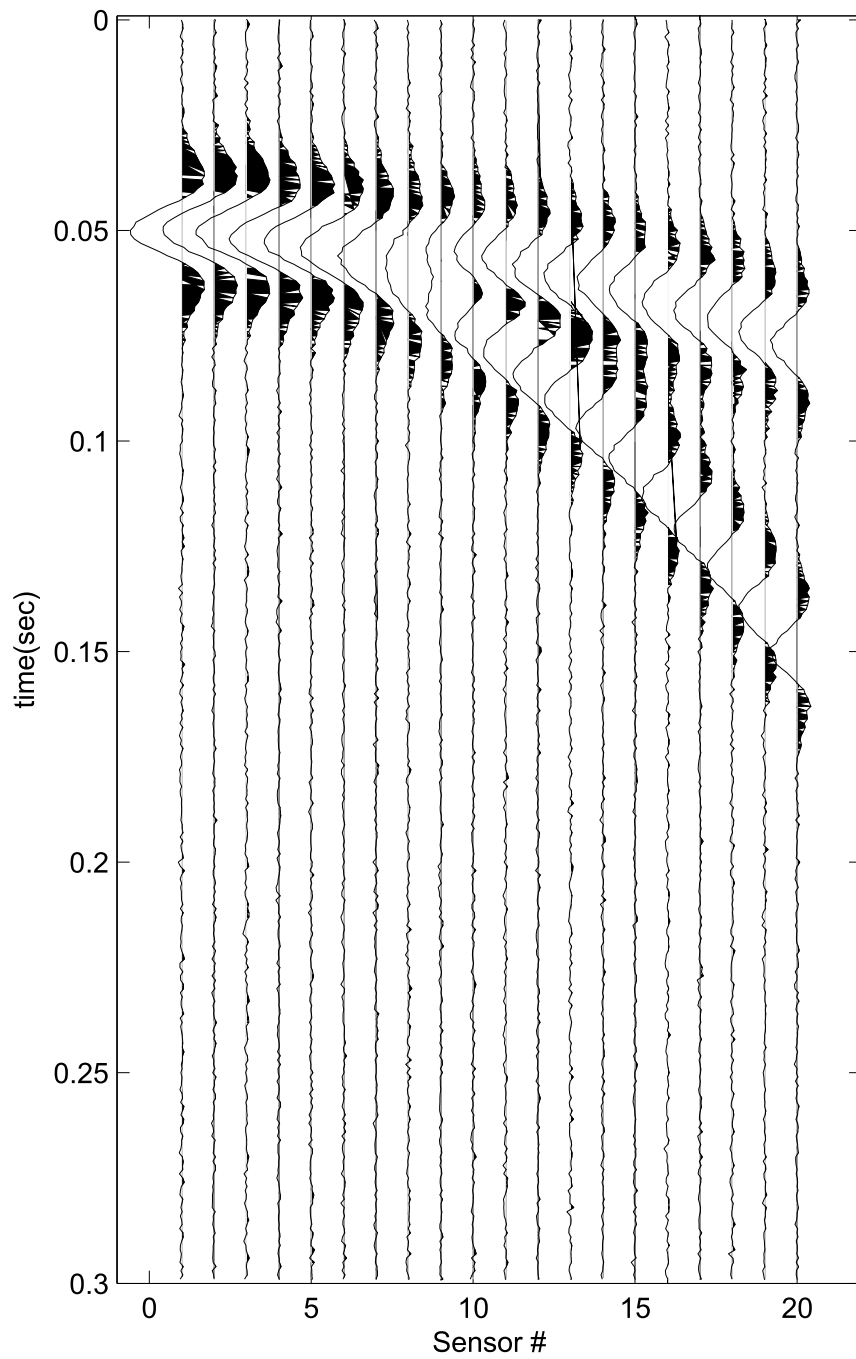


Figure 5.12: Wigner Plot of data before preflattening with  $f_c = 30\text{Hz}$ ,  $V_1 = 3600\text{m/s}$ ,  $V_2 = 7000\text{m/s}$ ,  $M = 20$ ,  $N = 300$  and  $d = 60\text{m}$

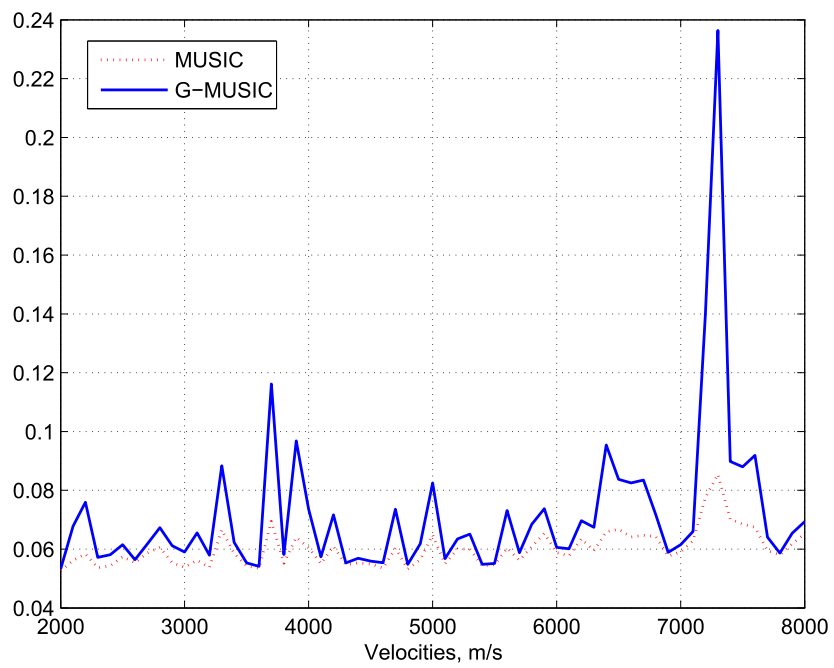


Figure 5.13: Estimate of velocities  $V_1 = 3600m/s$  and  $V_2 = 7000m/s$  before Preflating

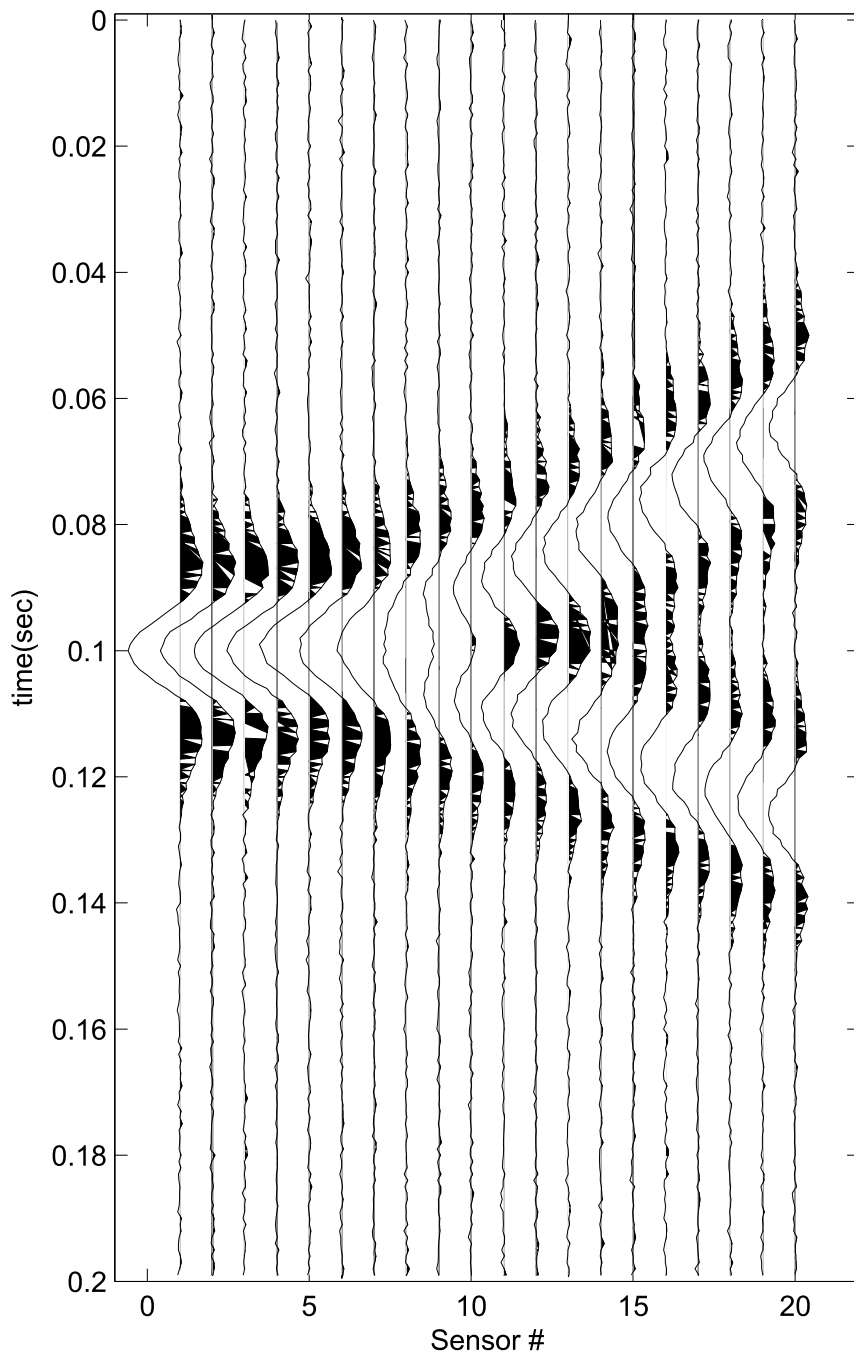


Figure 5.14: Wigner Plot of data after preflattening with  $V_1 = 2700m/s$ ,  $V_2 = 3600m/s$ ,  $V_a = 3000m/s$ ,  $f_c = 30Hz$ ,  $M = 20$ ,  $N = 200$

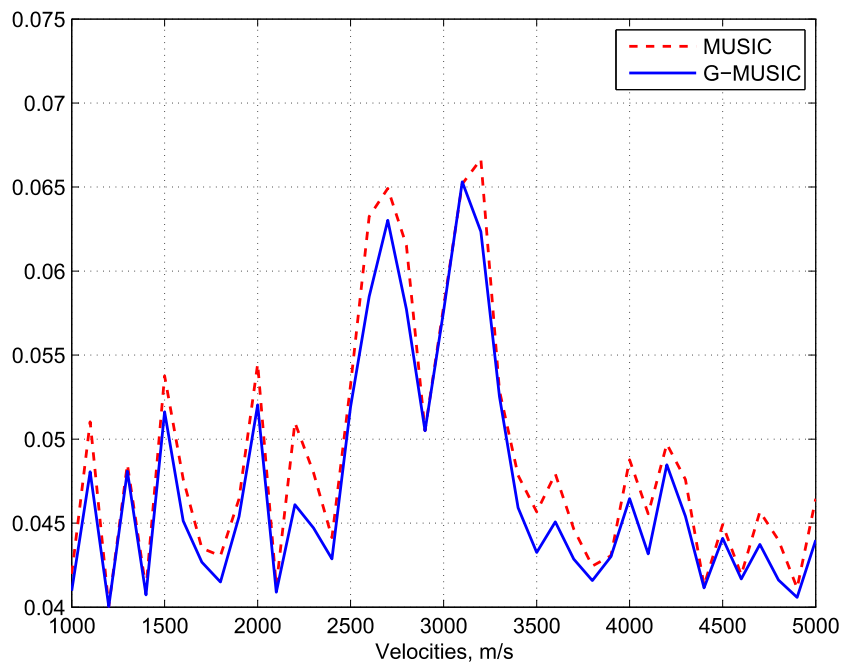


Figure 5.15: Estimate of velocities  $V_1 = 2700m/s$  and  $V_2 = 3600m/s$  after Preflattening with  $V_a = 3000m/s$ ,  $f_c = 30Hz$ ,  $M = 20$ ,  $N = 200$

## 5.6 Subspace Approach

In this section of the chapter we will try to estimate the stacking velocity and zero-offset time simultaneously. The zero offset time  $T_0$  is the time required for a wave to travel vertically from the source to the reflecting surface and back. The stacking velocity will be the average velocity of the wavelet. The Wave velocity often is deemed useful because it gives us a hint of the media through it traveled.

### 5.6.1 System Model

Consider that the seismic data has a single wavefront or multiple wavefronts with the same time delay arriving at  $M + 1$  geophones (sensors). We are taking  $K$  samples of the received data, then the received matrix  $\mathbf{Y}(t)$  can be represented as [48],

$$\mathbf{Y}(t) = \begin{bmatrix} y(t_0 - \tau_0) & y(t_0 - \tau_1) & y(t_0 - \tau_2) & \cdots & y(t_0 - \tau_M) \\ y(t_1 - \tau_0) & y(t_1 - \tau_1) & y(t_1 - \tau_2) & \cdots & y(t_1 - \tau_M) \\ y(t_2 - \tau_0) & y(t_2 - \tau_1) & y(t_2 - \tau_2) & \cdots & y(t_2 - \tau_M) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y(t_K - \tau_0) & y(t_K - \tau_1) & y(t_K - \tau_2) & \cdots & y(t_K - \tau_M) \end{bmatrix}$$

If  $y(t - \tau)$  is a wideband signal then it can not be expressed as  $y(t)e^{-j\omega\tau}$  as we did in the narrow band case, so we apply the subspace techniques in the frequency domain.

The discrete fourier transform (DFT) of  $\mathbf{Y}(t)$  can be written as,

$$\mathbf{Y}(\omega) = \begin{bmatrix} y(\omega_0)e^{-j\omega_0\tau_0} & y(\omega_0)e^{-j\omega_0\tau_1} & y(\omega_0)e^{-j\omega_0\tau_2} & \dots & y(\omega_0)e^{-j\omega_0\tau_M} \\ y(\omega_1)e^{-j\omega_1\tau_0} & y(\omega_1)e^{-j\omega_1\tau_1} & y(\omega_1)e^{-j\omega_1\tau_2} & \dots & y(\omega_1)e^{-j\omega_1\tau_M} \\ y(\omega_2)e^{-j\omega_2\tau_0} & y(\omega_2)e^{-j\omega_2\tau_1} & y(\omega_2)e^{-j\omega_2\tau_2} & \dots & y(\omega_2)e^{-j\omega_2\tau_M} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y(\omega_K)e^{-j\omega_K\tau_0} & y(\omega_K)e^{-j\omega_K\tau_1} & y(\omega_K)e^{-j\omega_K\tau_2} & \dots & y(\omega_K)e^{-j\omega_K\tau_M} \end{bmatrix}$$

here  $y(\omega_k) = |y(\omega_k)|e^{-j\theta}$ ,  $\omega_k = k\frac{2\pi}{K+1}$ ,  $k = 0, 1, 2, \dots, K$ . If we choose the the first sensor to be at the location of the source, then  $\tau_0 = 0$  and we would get,

$$\mathbf{Y}(\omega) = \begin{bmatrix} y(\omega_0) & y(\omega_0)e^{-j\omega_0\tau_1} & y(\omega_0)e^{-j\omega_0\tau_2} & \dots & y(\omega_0)e^{-j\omega_0\tau_M} \\ y(\omega_1) & y(\omega_1)e^{-j\omega_1\tau_1} & y(\omega_1)e^{-j\omega_1\tau_2} & \dots & y(\omega_1)e^{-j\omega_1\tau_M} \\ y(\omega_2) & y(\omega_2)e^{-j\omega_2\tau_1} & y(\omega_2)e^{-j\omega_2\tau_2} & \dots & y(\omega_2)e^{-j\omega_2\tau_M} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y(\omega_K) & y(\omega_K)e^{-j\omega_K\tau_1} & y(\omega_K)e^{-j\omega_K\tau_2} & \dots & y(\omega_K)e^{-j\omega_K\tau_M} \end{bmatrix}$$

Now we define a transformation matrix  $\mathbf{T}(\omega)$  as,

$$\mathbf{T}(\omega) = \begin{bmatrix} \frac{e^{j\theta_0}}{|y(\omega_0)|} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \frac{e^{j\theta_K}}{|y(\omega_K)|} \end{bmatrix}$$

Multiplying  $\mathbf{T}(\omega)$  with  $\mathbf{Y}(\omega)$  we get,

$$\mathbf{T}(\omega)\mathbf{Y}(\omega) = \begin{bmatrix} 1 & e^{-j\omega_0\tau_1} & \dots & e^{-j\omega_0\tau_M} \\ \vdots & \ddots & \vdots & \\ 1 & e^{-j\omega_K\tau_1} & \dots & e^{-j\omega_K\tau_M} \end{bmatrix}$$

Dropping the first column of the matrix as it does not contain any information.

We define a new matrix  $\bar{\mathbf{A}}$



$$\bar{\mathbf{A}}(\omega) = \begin{bmatrix} e^{-j\omega_0\tau_1} & \dots & e^{-j\omega_0\tau_M} \\ \vdots & \ddots & \vdots \\ e^{-j\omega_K\tau_1} & \dots & e^{-j\omega_K\tau_M} \end{bmatrix}$$

Lets assume that the range of the bandwidth of signal is  $(P, Q)$  where  $0 \leq P < Q \leq K$ , we define  $\omega_d = \frac{2\pi}{K+1}$ . Different version of  $\bar{\mathbf{A}}$  is defined as,

$$\mathbf{A}(\omega) = \begin{bmatrix} e^{-jP\omega_d\tau_1} & \dots & e^{-jP\omega_d\tau_M} \\ e^{-j(P+1)\omega_d\tau_1} & \dots & e^{-j(P+1)\omega_d\tau_M} \\ \vdots & \ddots & \vdots \\ e^{-j(Q-1)\omega_d\tau_1} & \dots & e^{-j(Q-1)\omega_d\tau_M} \\ e^{-jQ\omega_d\tau_1} & \dots & e^{-jQ\omega_d\tau_M} \end{bmatrix}$$

This matrix  $\mathbf{A}$  has Vandermonde characteristics, now we can apply the spectral estimation techniques that we discussed in the pervious section.

### 5.6.2 Estimation of Time Delays

We are going to find the time delays  $\tau_i$ ,  $i = 1, 2, \dots, M$  by using MUSIC and GMUSIC algorithm. The matrix  $\mathbf{A}(\omega)$  with a order of  $(Q - P + 1) \times M$  has a rank of  $M$  (when  $Q - P + 1 \geq M$ ) because of this reason we can find the subspace alternatively as [48],

$$\mathbf{\Omega}_s = \mathbf{A}(\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \quad \mathbf{\Omega}_0 = \mathbf{I} - \mathbf{\Omega}_s \quad (5.35)$$

In this case the MUSIC delay spectrum function as [48],

$$P_{MUSIC}(\tau) = \mathbf{a}^H(\tau)\mathbf{\Omega}_0\mathbf{\Omega}_0\mathbf{a}(\tau) \quad (5.36)$$

where  $\mathbf{a}(\tau) = [e^{jP\omega_d\tau} \dots e^{jQ\omega_d\tau}]^H$ . And the GMUSIC delay spectral function is defined as,

$$P_{GMUSIC}(\tau) = \mathbf{a}^H(\tau)\mathbf{Q}\mathbf{W}\mathbf{Q}^H\mathbf{a}(\tau) \quad (5.37)$$

where  $\mathbf{W}$  is given by (5.29). By choosing  $\tau$  corresponding to the  $M$  local maxima of  $P_{MUSIC}^{-1}$  and  $P_{GMUSIC}^{-1}$  we find the time delays.

### 5.6.3 Estimation of Velocity and Zero offset Time

If  $x_i$ ,  $i = 1, 2, \dots, M$  is the distance of the sensors from the source. Then we can estimate the zero offset time ( $T_0$ ) and stacking velocity ( $v$ ) simultaneously by solving the following equation [48],

$$\begin{pmatrix} \frac{1}{v^2} \\ \\ T_0 \end{pmatrix} = \begin{pmatrix} x_1^2 & -2\tau_1 \\ \vdots & \vdots \\ x_M^2 & -2\tau_M \end{pmatrix}^\dagger \begin{pmatrix} \tau_1^2 \\ \vdots \\ \tau_M^2 \end{pmatrix} \quad (5.38)$$

where  $\dagger$  is the left pseudo-inverse of the matrix.

### 5.6.4 Simulation Results

We generate a wide band synthetic data with a single wavefront reflection as shown in Figure 5.16 where  $M = 20$ ,  $K = 150$ ,  $v = 600m/s$ ,  $T_0 = 0.1s$ ,  $P = 20Hz$ ,  $Q = 50Hz$ ,  $f_c = 40Hz$  and the empirical  $SNR = 10dB$ . When you look at the figure it seems that the wavefront starts from  $T_0 = 0.02s$  but actually this is the truncated data and the original value of  $T_0$  is still the same.

Figure 5.17 and Figure 5.18 show the MUSIC and GMUSIC estimate of the velocity and zero-offset time respectively.

It can be seen that GMUSIC gives a better result then MUSIC algorithm in terms of RMSE which is summarized in table 5.1.

Table 5.1: RMSE for MUSIC and GMUSIC at different SNRs

RMSE	MUSIC (SNR=10dB)	GMUSIC (SNR=10dB)	MUSIC (SNR=20dB)	GMUSIC (SNR=20dB)
Velocity (m/s)	132	83	15	15
Zero offset (s)	0.0225	0.0071	0.0057	0.0057

## 5.7 Chapter Conclusion

This chapter uses the results of RMT to demonstrate that for seismic data models of comparable number of rows and columns we have a better estimate of the function of the sample covariance matrix. However, the traditional and RMT estimators converge when the number of rows of data matrix grows linearly while the number of columns remains fixed. This chapter tests few of the tools present in RMT for seismic signal processing, however we can exploit the field of RMT further.

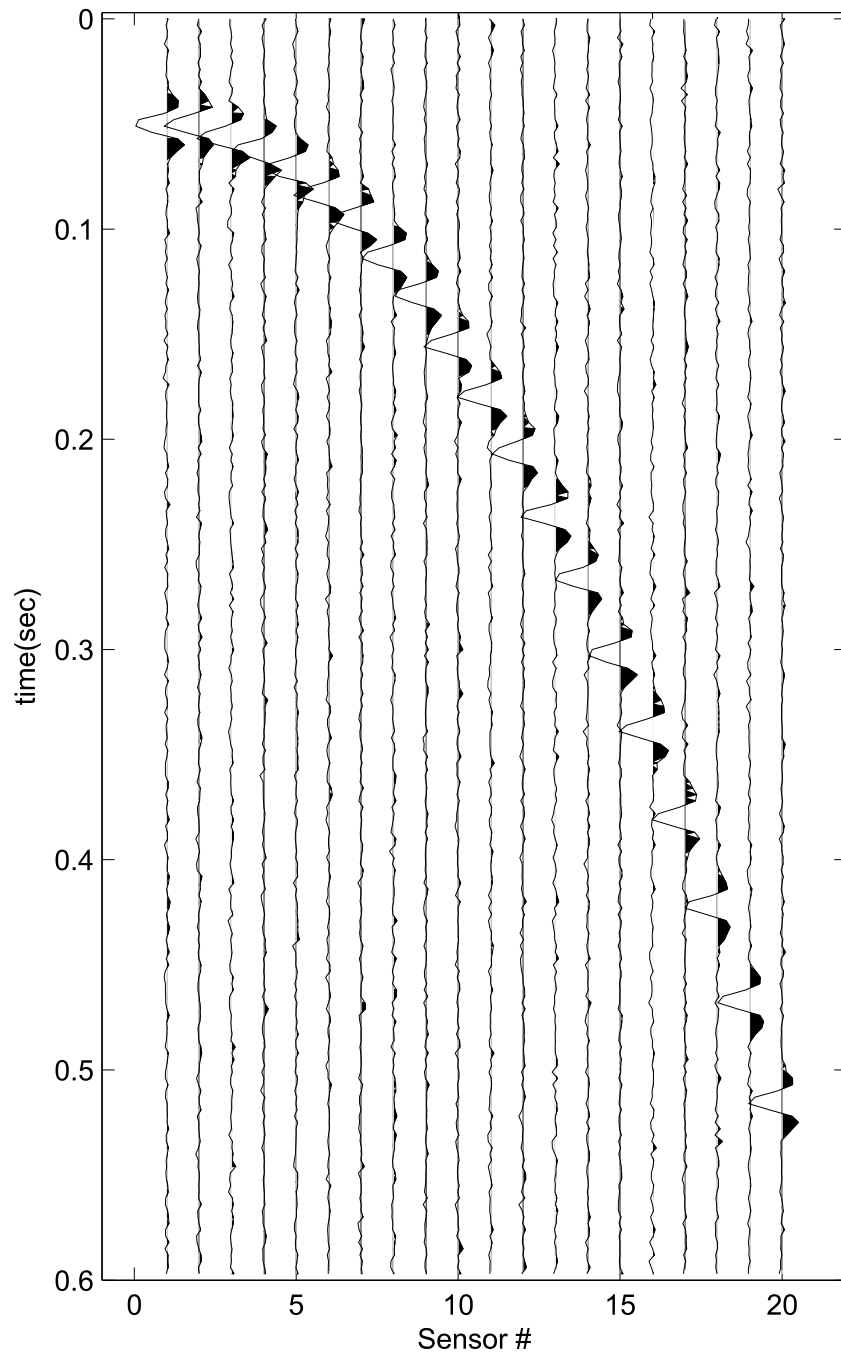


Figure 5.16: Wigner Plot of the seismic data with  $M = 20$ ,  $K = 150$ ,  $v = 600m/s$ ,  $T_0 = 0.1s$ ,  $P = 20Hz$ ,  $Q = 50Hz$ ,  $f_c = 40Hz$  and the empirical  $SNR = 10dB$

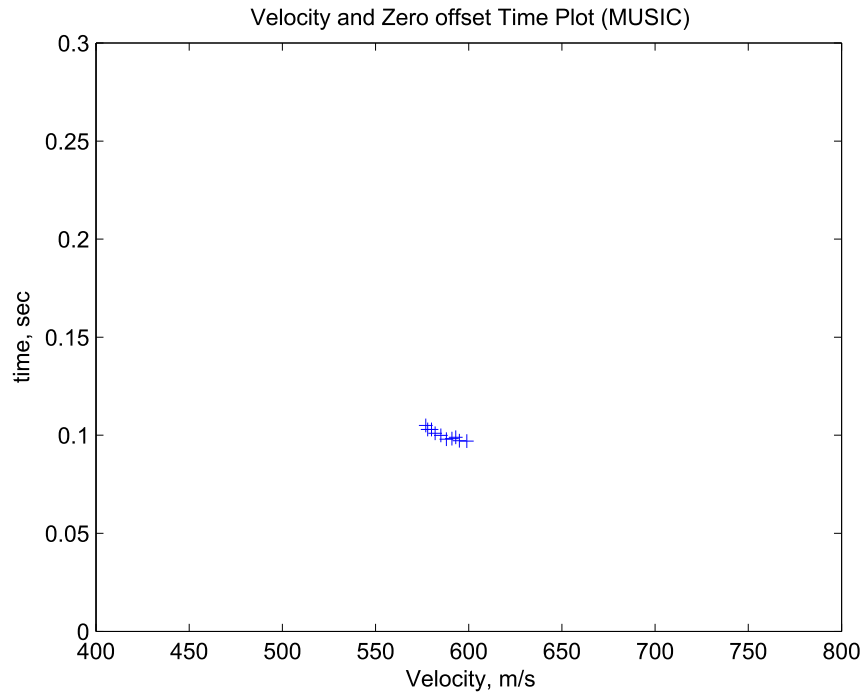


Figure 5.17: MUSIC Estimate of velocity and zero-offset time

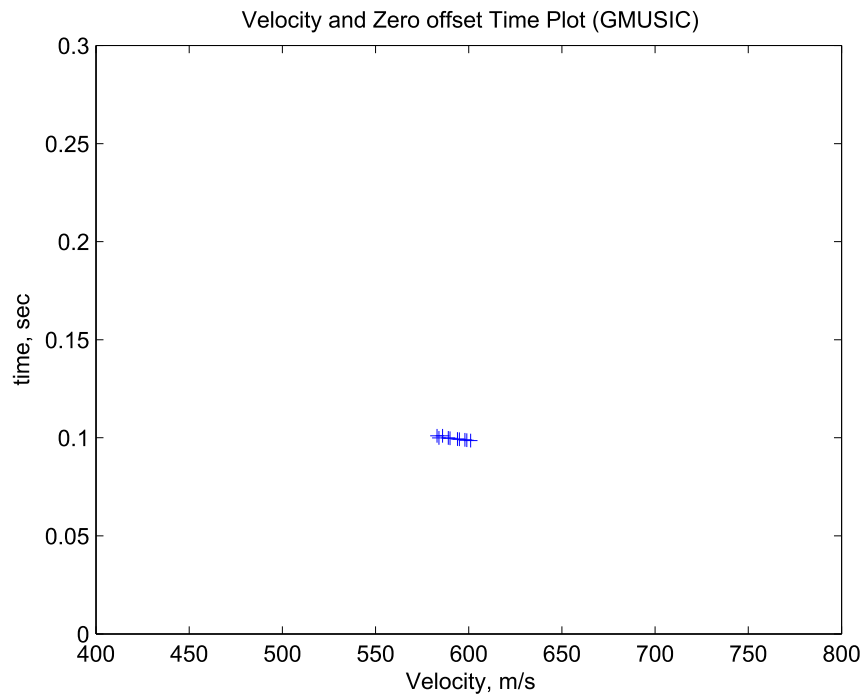


Figure 5.18: GMUSIC Estimate of velocity and zero-offset time

## CHAPTER 6

# CONCLUSIONS AND FUTURE WORK

### 6.1 Conclusion

This thesis highlights some of the uses of Random Matrix Theory (RMT) for wireless communication and seismic signal processing. This thesis started off with a very simple scaling law for bounded i.i.d random variables. We investigated different suboptimal scheduling schemes of VBLAST system and provided a mathematical justification for their behaviors. We invoked the power of RMT for the estimation of the distribution of wireless sensor networks. An finally we explored a few applications of RMT for seismic signal processing.

### 6.2 Future Work

This work can be extended to do future research in the following ways:

- A scaling law could be devised for unbounded i.i.d. random variables. This would result in a more powerful tool for Extreme Value Theory (EVT).
- The results obtained on the Vandermonde matrices could be further enhanced for the case defined as,

$$\mathbf{C} = \mathbf{V}_1 \mathbf{V}_2^H + \mathbf{V}_2 \mathbf{V}_2^H \quad (6.1)$$

where  $\mathbf{V}_1$  and  $\mathbf{V}_2$  are Vandermonde matrices of different distribution.

- For seismology, G-estimation method could be implemented for other high resolution spectral estimators such as ESPRIT.
- There are many other aspects of seismology that are left unexplored. For example we could use the G-estimation technique for eigenimage processing of a seismic data.
- The RMT estimators stated in the seismic chapter could be used as a better estimate of eigenvalues and eigenvectors when implementing facial detection and recognition system.



## APPENDIX A

# PROOF OF EQUATION (4.18)

As defined earlier, let  $\gamma$ ,  $\eta$ ,  $\tau$ ,  $\vartheta$ ,  $\varsigma$  and  $\psi$  be the probability measure of  $\mathbf{Y}\mathbf{Y}^H$ ,  $\mathbf{V}^H\mathbf{X}\mathbf{X}^H\mathbf{V}$ ,  $\mathbf{N}\mathbf{N}^H$ ,  $\mathbf{X}\mathbf{X}^H\mathbf{V}\mathbf{V}^H$ ,  $\mathbf{X}\mathbf{X}^H$  and  $\mathbf{V}\mathbf{V}^H$  respectively. The rectangular additive free deconvolution provides us with the following moments of  $\mathbf{V}^H\mathbf{X}\mathbf{X}^H\mathbf{V}$

$$m_1(\eta) = m_1(\gamma) - \sigma^2 \tag{A.1}$$

$$m_2(\eta) = m_2(\gamma) - 2\sigma^2(1 + c_1)m_1(\gamma) + \sigma^4(1 + c_1) \tag{A.2}$$

The moments of  $\mathbf{V}^H\mathbf{X}\mathbf{X}^H\mathbf{V}$  are related to the moments of  $\mathbf{X}\mathbf{X}^H\mathbf{V}\mathbf{V}^H$  as

$$m_1(\vartheta) = c_2 m_1(\eta) = c_2(m_1(\gamma) - \sigma^2) \tag{A.3}$$

$$\begin{aligned} m_2(\vartheta) = c_2 m_2(\eta) &= c_2 m_2(\gamma) - 2c_2 \sigma^2(1 + c_1)m_1(\gamma) \\ &\quad + c_2 \sigma^4(1 + c_1) \end{aligned} \tag{A.4}$$

Multiplicative free deconvolution can now be applied to get the moments of  $\mathbf{V}\mathbf{V}^H$  and are estimated as

$$m_1(\psi) = c_2(m_1(\gamma) - \sigma^2) \quad (\text{A.5})$$

$$\begin{aligned} m_2(\psi) &= c_2^2(m_1(\gamma) - \sigma^2)^2 - c_2(c_2(1 + c_1)m_1^2(\gamma) \\ &\quad + 2\sigma^2(1 + c_1 - c_2 - c_1c_2)m_1(\gamma) \\ &\quad + \sigma^4(c_2 - c_1) - m_2(\gamma)) \end{aligned} \quad (\text{A.6})$$

The moments of  $\mathbf{V}\mathbf{V}^H$  are also given by (4.16)

$$m_1(\psi) = c_2 \quad (\text{A.7})$$

$$m_2(\psi) = c_2 + c_2^2 I_2 \quad (\text{A.8})$$

This simplification occurs due to the fact  $K_1 = K_2 = K_{1,1} = 1$  and  $I_1 = 1$  irrespective of the distribution of  $\omega$ . By equation the first moments of  $\mathbf{V}\mathbf{V}^H$ , we conclude

$$m_1(\gamma) = \sigma^2 + 1 \quad (\text{A.9})$$

Equation the second moment of  $\mathbf{V}\mathbf{V}^H$  gives us

$$1 + c_2 I_2 = m_2(\gamma) + (-c_1 c_2) m_1^2(\gamma) + 2\sigma^2(c_1 c_2 - c_1 - 1) m_1(\gamma) + \sigma^4(1 + c_1 - c_1 c_2) \quad (\text{A.10})$$

$$1 + \frac{L}{P} I_2 = m_2(\gamma) + \left(-\frac{L}{K}\right) m_1^2(\gamma) + 2\sigma^2 \left(\frac{L}{K} - \frac{P}{K} - 1\right) m_1(\gamma) + \sigma^4 \left(1 + \frac{P}{K} - \frac{L}{K}\right) \quad (\text{A.11})$$

$$1 + \frac{L}{P} I_2 = m_2(\gamma) - \sigma^4 \left(1 + \frac{P}{K}\right) - 2\sigma^2 \left(1 + \frac{P}{K}\right) - \left(\frac{L}{K}\right) \quad (\text{A.12})$$

Solving for  $P$

$$P = \frac{\pm \sqrt{\frac{K m_2(\gamma) - K - L - K \sigma^4 - 2K \sigma^2}{K^2 m_2^2(\gamma) - 2K(K \sigma^4 + 2K \sigma^2 + L + K) m_2(\gamma) + 2KL \sigma^4 + 4KL \sigma^2 + k^2 + 6K^2 \sigma^4 + 4K^2 \sigma^2} + K^2 \sigma^8 + 4K^2 \sigma^6 + L^2 + 2LK - 4LK \sigma^2 I_2(\sigma^2 + 2)}}{2\sigma^2(\sigma^2 + 2)} \quad (\text{A.13})$$

In the case where the estimation of  $P$  gives a complex result then the estimation becomes a maximization problem for the following concave function

$$f_{\max\{P\}} = -1 - \frac{L}{P} I_2 + m_2(\gamma) - \sigma^4 \left(1 + \frac{P}{K}\right) - 2\sigma^2 \left(1 + \frac{P}{K}\right) - \left(\frac{L}{K}\right) \quad (\text{A.14})$$

Similarly, by using higher moments we can derive alternate estimation formulas for  $P$  but as seen from figure 4.2 the estimation of the lower moments is more accurate than higher moments.

## APPENDIX B

# PROOF OF EQUATION (4.19)

Consider the case where  $\omega$  has a distribution of  $p(x)$  bounded within  $[0, 1)$ .  $I_n$  is given by

$$I_n = (2\pi)^{n-1} \int_0^1 p_\omega(x)^n dx \quad (\text{B.1})$$

the probability distribution  $p_\omega(x)$  can be represented alternatively as a Bernstein polynomial

$$p_\omega(x) = \lim_{t \rightarrow \infty} \sum_{v=1}^t \binom{t-1}{v-1} p_\omega \left( \frac{v-1}{t-1} \right) x^{v-1} (1-x)^{t-v} \quad (\text{B.2})$$

we get

$$I_n = \lim_{t \rightarrow \infty} (2\pi)^{n-1} \int_0^1 \left( \sum_{v=1}^t \binom{t-1}{v-1} p_\omega \left( \frac{v-1}{t-1} \right) x^{v-1} (1-x)^{t-v} \right)^n dx \quad (\text{B.3})$$

we can expand the integration by using multinomial expansion defined as

$$(x_1, x_2, \dots, x_t)^n = \sum_{k_1+k_2+\dots+k_t=n} \binom{n}{k_1, k_2, \dots, k_t} x_1 x_2 \dots x_t \quad (\text{B.4})$$

therefore we get

$$\begin{aligned} I_n &= \lim_{t \rightarrow \infty} (2\pi)^{n-1} \sum_{k_1+k_2+\dots+k_t=n} \binom{n}{k_1, k_2, \dots, k_t} \\ &\quad \cdot \left( \prod_{v=1}^t \left[ p_\omega \left( \frac{v-1}{t-1} \right) \binom{t-1}{v-1} \right]^{k_v} \right) \\ &\quad \cdot \int_0^1 x^{\sum_{v=1}^t (v-1)k_v} (1-x)^{\sum_{v=1}^t (t-v)k_v} dx \end{aligned} \quad (\text{B.5})$$

$$\begin{aligned} &= \lim_{t \rightarrow \infty} (2\pi)^{n-1} \sum_{k_1+k_2+\dots+k_t=n} \binom{n}{k_1, k_2, \dots, k_t} \\ &\quad \cdot \left( \frac{\Gamma(1+tn - \sum_{v=1}^t vk_v) \Gamma(\sum_{v=1}^t vk_v - n + 1)}{\Gamma((t-1)n + 2)} \right) \\ &\quad \cdot \left( \prod_{v=1}^t \left[ p_\omega \left( \frac{v-1}{t-1} \right) \binom{t-1}{v-1} \right]^{k_v} \right) \end{aligned} \quad (\text{B.6})$$

# REFERENCES

- [1] J. Wishart, “The generalized product moment distribution in samples from a normal multivariate population,” *Biometrika*, vol. 20 A, pp. 32-52, 1928.
- [2] E. Wigner, “Characteristic vectors of bordered matrices with infinite dimensions,” *The Annals of Mathematics*, vol. 62, pp. 546-564, 1955.
- [3] V. A. Marčenko and L. A. Pastur, “Distributions of eigenvalues for some sets of random matrices,” *Math USSR-Sbornik*, vol. 1, pp. 457-483, 1967.
- [4] Fumio Hiai and Denis Petz, “The Semicircle Law, Free Random Variables, and Entropy”, ISBN 0-8218-2081-8
- [5] Voiculescu, D. V., Dykema, K. J., Nica, A. “Free random variables. A non-commutative probability approach to free products with applications to random matrices, operator algebras and harmonic analysis on free groups.” *CRM Monograph Series*, 1. American Mathematical Society, Providence, RI, 1992. ISBN 0-8218-6999-X

- [6] W. S. Conner, J. Heidemann, L. Krishnamurthy, X. Wang, and M. Yarvis, “Workplace Applications of Sensor Networks,” *Technical Report ISI-TR-2004-591*, USC/Information Sciences Institute, July, 2004.
- [7] A. Nordio, C.F. Chiasserini and E. Viterbo, “Bandlimited Field Reconstruction for Wireless Sensor Networks,” [http://arxiv.org/PS\\_cache/arxiv/pdf/0707/0707.1954v1.pdf](http://arxiv.org/PS_cache/arxiv/pdf/0707/0707.1954v1.pdf)
- [8] P. Marziliano, M. Vetterli, “Reconstruction of Irregularly Sampled Discrete-Time Bandlimited Signals with Unknown Sampling Locations,” *IEEE Transactions on Signal Processing*, vol. 48, no.12, pp. 3462-3471, Dec. 2000.
- [9] L. Hu, D. Evans, “Localization for Mobile Sensor Networks,” *Tenth Annual International Conference on Mobile Computing and Networking (ACM MobiCom 2004)*, Philadelphia, PA, September-October 2004.
- [10] O. Kwon, H. Song, “Localization through Map Stitching in Wireless Sensor Networks,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 19, no. 1, pp. 93-105, Jan. 2008.
- [11] Ø. Ryan and M. Debbah “Free Deconvolution for Signal Processing Applications,” *IEEE International Symposium on Information Theory, ISIT 2007*, June 2007.
- [12] Ø. Ryan and M. Debbah “Asymptotic Behaviour of Random Vandermonde Matrices with Entries on the Unit Circle,” submitted to *IEEE Transactions on Information Theory*, vol. 1, no.1, Jan. 2008.

- [13] F. Benaych-Georges, M. Debbah “Free Deconvolution: from Theory to Practice,” submitted
- [14] F. Benaych-Georges “Infinitely divisible distribution for rectangular free convolution: classification and matricial interpretation,” *Probab. Theory Related Fields*, vol. 139, no. 1-2, pp. 143-189, 2007.
- [15] R. Speicher “Free Probability Theory and Non-Crossing partitions,” *Séminaire Lotharingien de Combinatoire*, B39c (1997), 38 pages, 1997.
- [16] A. Nordio, C.F. Chiasserini, E. Viterbo “On the d-dimensional Quasi-Equally Spaced Sampling,” *submitted to IEEE Transactions on Signal Processing*, Jun. 2008.
- [17] Oyan Ryan and Mérouane Dabbah, “Channel capacity estimation using free probability theory,” *IEEE Trans. on Signal processing*, vol 56, no.11, Nov. 2008.
- [18] Pascal Vallet, “Matrices Aléatoires et Applications au Traitement Statistique du Signal,” *Rapport de Stage*,, Effectué de Mars à Septembre 2008.
- [19] Oyvind Ryan and Merouane Debbah “Random Vandermonde Matrices - Part I: Fundamental results,” to be submitted *IEEE Trans. on Information Theory*, January. 2008.
- [20] X. Mestre “Improved estimation of eigenvalues of covariance matrices. and their associated subspaces using their samples estimates,” submitted *IEEE Trans. on Information Theory*, January. 2008.



- [21] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE Jour. Selec. Areas. Commu.*, vol. 21, no. 5, pp. 684–702, June 2003.
- [22] Q. H. Spencer, C. B. Peel, A. L. Swindlehurst, and M. Haardt, "An introduction to the multi-user MIMO downlink," *IEEE Comm. Magazine*, vol. 42, no. 10, pp. 60–67, Oct. 2004.
- [23] M. Sharif and B. Hassibi "On the capacity of MIMO broadcast channel with partial side information" *IEEE Trans. Info. Theory*, vol. 51, no. 2, Feb. 2005.
- [24] P. Viswanath, D. N. Tse, and R. Laroia, "Opportunistic beamforming using dump antennas," *IEEE Trans. Inform.*, vol. 48, no. 6, pp. 1277–1294, June 2002.
- [25] T. Yoo and A. Goldsmith, "On the Optimality of Multi-Antenna Broadcast Scheduling Using Zero-Forcing Beamforming" *To appear: IEEE JSAC Special Issue on 4G Wireless Systems*.
- [26] T. Yoo and A. Goldsmith, "On the Optimality of Multi-Antenna Broadcast Scheduling Using Zero- Forcing Beamforming" *To appear: IEEE JSAC Special Issue on 4G Wireless Systems*.
- [27] T. Y. Al-Naffouri, M. Sharif, and B. Hassibi " On the Capacity of Correlated MIMO Broadcast Channel with Partial Side Information," *International Symposium on Information Theory (ISIT)*, Seattle, OR, Jul. 2006.

- [28] T. Y. Al-Naffouri, M. Sharif, and B. Hassibi "How Much Does Transmit Correlation Affect the Sum-Rate of MIMO Downlink Channels?" *to be submitted to IEEE Transactions. on Communications.*
- [29] A. Khitsi, Coding Techniques for Multicasting, Master's thesis, MIT Cambridge, MA, 2004.
- [30] M. Airy, S. Shakkattai and R. W. Jr. Heath, "Spatially greedy scheduling in multi-user MIMO wireless systems," *Signals, Systems and Computers, 2003. Conference Record of the Thirty-Seventh Asilomar Conference on*, vol. 1, pp. 982 - 986, 9-12 Nov. 2003.
- [31] G. J. Foschini and M. J. Gan, "On Limits of Wireless Communications in a Fading Environment when using Multiple Antennas," *Wireless Personal Communications*, vol 6, pp. 311-335, March 1998.
- [32] Ran Gozali, R M. Buehrer and B. D. Woerner, "The Impact of Multiuser Diversity on Space-Time Block Coding," *IEEE Comm. Letters*, vol. 7, no 5, pp. 213:215, May 2003.
- [33] R. W. Jr. Heath, M. Airy and A.J. Paulraj, "Multiuser diversity for MIMO wireless systems with linear receivers," *Signal, Systems and Computers 2001, Conference Record of the Thirty-Fifth Asilomar Conference on*, vol. 2, pp. 1194-1199, 2001.

- [34] R. Knopp and P. Humblet, "Information capacity and power control in single cell multiuser communications," *Proc. IEEE Int. Computer Conf. (ICC'95)*, Seattle, WA, June 1995.
- [35] V. K.N. Lau, Y. Liu and T. A. Chen, "Optimal Multi-user Space Time Scheduling for Wireless Communications," , " *IEEE 56th VTC 2002-Fall. 2002*, vol. 4, 24-28 Sept. 2002, pp. 1939-1942.
- [36] C. B. Papadias and G. J. Foschini, "On the Capacity of Certain Space-Time Coding Schemes," *EURASIP Journal on Applied Signal Processing 2002*, vol. 5, pp. 447-458, May 2002.
- [37] M. Realp and A.I Perez-Neira, "Cross-layer MAC scheduling for multiple antenna systems," *Global Telecommunications Conference, 2004. GLOBECOM '04. IEEE*, vol. 5, pp. 3352 - 3356, 29 Nov.-3 Dec. 2004.
- [38] Oh-Soon Shin and Kwang Bok Lee, "Antenna-Assisted Round Robin Scheduling for MIMO Cellular Systems," *IEEE Comm. Letters*, vol. 7, no 3, pp. 109:111, March 2003.
- [39] D. N. C. Tse, "Optimal power allocation over parallel Gaussian channels," *Proc. Int. Symp. Infrmation Theory*, Ulm, Germany, June 1997.
- [40] P. W. Wolniansky, G. J. Foschini, G. D. Golden and R. A. Valenzuela, "V-BLAST: An Architecture for Realizing Very High Data Rates Over the Rich-Scattering Wireless Channel," *Proc. ISSSE-98*, Pisa, Italy, pp. 295-300, Sept. 29, 1998.

- [41] S. Loyka, F. Gagnon, “On BER Analysis of the BLAST without optimal ordering over Rayleigh Fading Channel,” *Vehicular Technology Conference, VTC2004-Fall, IEEE 60th*, Vol.2, pp. 1473-1477, 2004.
- [42] Yi Jiang, Xiayu Zheng, Jian Li, “Asymptotic performance analysis of V-BLAST,” *Global Telecommunications Conference, 2005. GLOBECOM '05. IEEE*, Vol.6, pp. -3886, 2005.
- [43] Heunchul Lee, Inkyu Lee, “Channel Capacity of BLAST based on the zero forcing criterion,” *Vehicular Technology Conference, VTC 2006-Spring. IEEE 63rd*, Vol.4, pp. 1615-1619, 2006.
- [44] Alan Edelman, “Eigenvalues and condition numbers of random matrices,” *SIAM Journal on Matrix Analysis and Applications*, Vol.9, Issue 4, pp. 543-560, Oct 1988.
- [45] D. Bertsimas, K. Natarajan, Chung-Piaw Teo “Tight Bounds On Expected Order Statistics,” *Probability in the Engineering and Informational Sciences*, Vol.20, Issue 4, pp. 667-686, 2006.
- [46] Y. Gordon, A. E. Litvak, C. Schutt, E. Werner “On the minimum of several random variables,” *Proceedings of American Mathematical Society*, Vol.134, No. 12, pp. 3669-3675, Dec. 2006.
- [47] H. A. David *Order Statistics*,, 1st ed. New York: Wiley, 1970.

- [48] Edited by R. Lynn Kirilin, William J. Done “Covariance Analysis for Seismic Signal Processing,” *Geophysical development No.8, Society of Exploration Geophysicists.*
- [49] Scharf, Louis L. “Statistical signal processing : detection, estimation, and time series analysis,” *Addison-Wesley Pub. Co*, 1991.
- [50] Ozdogan Yilmaz, “Seismic Data Processing” *SEG Investigations in Geophysics, Society of Exploration Geophysicists, 1987, VOL 3*, pp. 62–151.
- [51] Ian Naismith Sneddon, “Fourier transforms” *Dover Publications*, 1995.
- [52] Arya V.K, Aggarwal J.K, “Deconvolution of Seismic Data,” *Benchmark Papers in Electrical Engineering and Compute Science*,
- [53] Robinson, Enders A, “Geophysical Signal Processing” TN 269 .R553, 1958
- [54] Mosteller, F., and Tukey, J. W., “Data analysis and regression” Addison Wesley Publ. Co., 1977
- [55] Tulino, A. M., Verdu, S., “Random Matrix Theory and Wireless communications”
- [56] Kumaresan, R., and D. W., Trufts, “Estimating the parameters of exponentially damped sinusoids and pole zeromodeling in the noise” *IEEE Trans. Acoust. Speech Signal Process.* vol. 30, no. 6, pp. 833-840,1982.
- [57] Anton Z, “Deconvolution”, Delft University of Technology, ISBN 0-934634-62-9

- [58] Yeo-Sun Yoon, "Direction-of-arrival estimation of wideband sources using sensor arrays", PhD Thesis, Georgia Institute of Technology, July 2004
- [59] V. Girko, "An Introduction to Statistical Analysis of Random Arrays", The Netherlands: VSP, 1998.
- [60] X. Mestre, "Improved estimation of eigenvalues of covariance matrices and their associated subspaces using their sample estimates", 2006. Preprint. Available at <http://www.cttc.es/drafts/cttc-rc-2006-004.pdf>.
- [61] X. Mestre, "On the Asymptotic Behavior of the Sample Estimates of Eigenvalues and Eigenvectors of Covariance Matrices" accepted at *Signal Processing, IEEE Transaction on*.
- [62] J. Marsden and M. Hoffman, "Basic Complex Analysis", 3rd ed. New York: Freeman, 1987.
- [63] Stoica, P., and R. Moses, "Introduction to Spectral Analysis", Upper Saddle River, NJ: Prentice Hall, 1997.
- [64] Kaveh, M., and Barabell, A. J., "The statistical performance of the MUSIC and minimum-norm algorithms in resolving plane waves in noise" *Acous. Speech and Sig. Proc., IEEE Transaction on*. vol. 34, pp. 331-341, 1986
- [65] Key, S. C., Kirlin, R.L., and Smithson, S.B., "Seismic velocity analysis using maximum likelihood weighted eigenvalue ratios" *57th Annual Internat Mtg., Soc. Expl. Geophys.* Expanded Abstract, 461-464, 1987

- [66] Kirlin, R. L., “New maximum entropy spectrum using uncertain eigenstructure constraints” *Aerosp. Elect. Systems, IEEE Transaction on.* vol. 28, pp. 2-24, 1992.
- [67] X.Mestre, “An improved weighted MUSIC algorithm for small sample size scenarios,” *in Proc. 14th Workshop on Adaptive Sensor Array Processing*, Lexington, MA, 2006

# Vitae

- Babar Hasan Khan
- Born in Karachi, Pakistan on March 28, 1985
- Received Bachelor of Science (B.Sc) in Electronic Engineering from GIK Institute of Engineering Sciences & Technology, Pakistan in June 2006.
- Joined King Fahd University of Petroleum & Minerals, Dhahran, Saudi Arabia as a Research Assistant in February 2007.
- Completed Master of Science (M.S.) in Electrical Engineering in December 2009.
- Email: *babar.hasan.1@gmail.com*
- Present Address: Room 241, Bldg. 903, KFUPM, Dhahran 31261, Saudi Arabia.
- Permanent Address: Azizia, Jeddah, Saudi Arabia.