UNIVERSITY OF CALIFORNIA,
IRVINE

Resource Allocation in Communication, Quantization, and Localization

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Electrical and Computer Engineering

by

Jie Chen

Dissertation Committee:
Professor A. Lee Swindlehurst, Chair
Professor Syed Ali Jafar
Professor Hamid Jafarkhani

2015

DEDICATION

To my family for all the motivation, inspiration, and support

Contents

Page

LIST OF FIGURES **vi**

LIST OF TABLES **viii**

LIST OF ALGORITHMS **ix**

ACKNOWLEDGMENTS **x**

CURRICULUM VITAE **xi**

ABSTRACT OF THE DISSERTATION **xiv**

1 Introduction **1**
1.1 Organization of the Thesis 3

2 Bargaining Based Resource Allocation in MIMO-OFDMA Systems **6**
2.1 Introduction 6
2.2 System Model 10
2.3 Game Theoretic Bargaining Solutions 12
2.3.1 Nash Bargaining Solution 12
2.3.2 Kalai-Smorodinsky Bargaining Solution 13
2.4 Finding the Bargaining Solutions 14
2.4.1 Bisection Search 15
2.4.2 Preference Function Formulation 15
2.4.3 KSBS associated with average rate 17
2.5 Application to the MIMO-OFDMA Downlink with Block Diagonalization . . 19
2.5.1 Block Diagonalization 19
2.5.2 Time Sharing 21
2.5.3 Convex Optimization for NBS 23
2.5.4 Bisection Search for KSBS 24
2.5.5 Preference Function Method for KSBS 25
2.5.6 Algorithm Simplification 26
2.6 Numerical Results 28

3 The Gaussian CEO Problem for Scalar Sources with Arbitrary Memory 33
3.1 Introduction 33
3.2 System Model and Preliminaries 36
3.2.1 Berger-Tung Inner Bound and Sum Rate 37
3.2.2 Extension of the Toeplitz Distribution Theorem 39
3.3 Problem Formulation 40
3.3.1 Mutual Information 40
3.3.2 MSE and Optimal Estimator 44
3.3.3 Full Formulation 45
3.4 Necessary Condition 46
3.5 Sufficient Condition 48
3.5.1 Analysis and Discussion 50
3.6 Special Case - Two Terminals 51
3.6.1 Formulation and Solution 52
3.6.2 Examples 53
3.6.3 Sufficiency Verification 58

4 Application of Massive Antenna Arrays in STAP 59
4.1 Introduction 59
4.2 System Model 62
4.2.1 Simplified Separable Algorithm 64
4.2.2 Asymptotic Orthogonality of Steering Vectors – ULA 65
4.2.3 Asymptotic Orthogonality of Steering Vectors – URA 67
4.3 SINR Analysis 70
4.3.1 SINR Performance Analysis for ULAs 70
4.3.2 SINR Loss Scaling Law 73
4.3.3 SINR Performance Analysis for URAs 76
4.4 Secondary Sample Size Analysis 77
4.5 Simulations 80

5 Positioning in NLOS Environments with a Minimal Set of Measurements 87
5.1 Introduction 87
5.2 NLOS Positioning Problems 90
5.3 Scattering Model 92
5.3.1 One MD-TD Pair 93
5.3.2 Multiple MDs 99
5.3.3 Extensions 103
5.3.4 Improve the Quality of TOA Measurements 104
5.4 Reflection Model 104
5.5 Tracking 113
5.6 Simulations 116

6 Conclusions 120

Bibliography 123

A Proofs **132**
A.1 Proof of Convergence for Theorem 2.1 132
A.2 Convexity of the Achievable Rate Region 134
A.3 Proof of Theorem 3.1 137
A.4 Proof of Lemma 3.1 139
A.5 Proof of Theorem 3.2 141
A.6 Proof of Theorem 4.1 144

List of Figures

Page

1.1 Resource allocation: the subject of study. 2

2.1 Performance comparison of allocation schemes for one channel realization. . . 28
2.2 Average sum rate vs. SNR: equal pathloss case (left: full algorithm implementation, right: simplified algorithms.) 29
2.3 Average minimum rate vs. SNR: equal pathloss case (left: full algorithm implementation, right: simplified algorithms.) 29
2.4 Average sum rate vs. SNR: unequal pathloss case (left: full algorithm implementation, right: simplified algorithms.) 30
2.5 Average minimum rate vs. SNR: unequal pathloss case (left: full algorithm implementation, right: simplified algorithms.) 30
2.6 Average sum rate for various numbers of users: equal pathloss case (left: full algorithm implementation, right: simplified algorithms.) 31
2.7 Average minimum rate for various numbers of users: equal pathloss case (left: full algorithm implementation, right: simplified algorithms.) 31
2.8 Empirical average rates over 150 intervals for 3 users. 32

3.1 Indirect multiterminal source coding (the CEO problem.) 37
3.2 Forward test channels. 42
3.3 Achievable sum rate versus target distortion (first-order Gauss-Markov processes.) 54
3.4 PSDs of the source and the error processes (first-order Gauss-Markov process, symmetric noise.) 55
3.5 PSDs of the source and the error processes (first-order Gauss-Markov process, asymmetric noise.) 55
3.6 Achievable sum rate versus target distortion (band-limited Gaussian processes.) 56
3.7 PSDs of the source and the error processes (band-limited Gaussian processes, symmetric noise.) 57
3.8 PSDs of the source and the error processes (band-limited Gaussian processes, asymmetric noise.) 57
3.9 Eigenvalues of $\mathbf{Z}$ in ω. 58

4.1 Depiction of an airborne radar scenario. 63
4.2 Illustration of a uniform rectangular array. 68

4.3 Target is moderately above the ground clutter patch. 69
4.4 Target sharing the same spatial frequency with clutter patches. 71
4.5 Max. SINR as a function of the number of antenna elements. 82
4.6 SINR as a function of normalized Doppler frequency. 82
4.7 Max. SINR as a function of the total number of antenna elements. 83
4.8 SINR as a function of normalized Doppler frequency. 83
4.9 Max. SINR as a function of the total number of antenna elements. 84
4.10 SINR as a function of normalized Doppler frequency. 84
4.11 Scaling law of SINR loss. 85
4.12 SINR as a function of the number of secondary data samples. 85

5.1 Wireless signal propagation in outdoor and indoor environments. 91
5.2 Geometrical relationship of one MD-TD pair. 93
5.3 A localization scenario with K MDs, each associated with one scatterer relative to the MD. 100
5.4 A localization scenario with one MD, associated with multiple scatterers. . . 103
5.5 The trajectory of the target. 105
5.6 The ambiguity in the trajectory of the target. 106
5.7 RMSE of target position (scattering case.) 116
5.8 RMSE of target speed (scattering case.) 116
5.9 RMSE of target position (reflection case.) 117
5.10 RMSE of target speed (reflection case.) 117
5.11 True and estimated target trajectory. 118

List of Tables

Page

1.1 Resource Allocation Tradeoffs 4

4.1 STAP System Parameters 81

LIST OF ALGORITHMS

Page

1 Bisection Search Algorithm . . . 25

2 Preference Function Based Algorithm . . . 26

3 Reduced Complexity Algorithm . . . 27

ACKNOWLEDGMENTS

In the past few years, I have been very fortunate to have the extraordinary privilege of being able to study and research in a pleasant and intellectual environment. My dissertation advisor, Professor A. Lee Swindlehurst, is an unparalleled scholar and teacher to me, and has been enormously helpful all the time. His teaching style is absolutely fascinating. His suggestions on my research work are keenly insightful. His generous spirit and kindness are amazingly supportive. It has been a true honor to be guided by him on the journey toward knowledge and creation.

For the good times at UC Irvine, I received thoughtful mentoring and advice from a number of academic researchers. My thanks firstly go to the members of my qualifying and thesis committee at the EECS department: Professor Ender Ayanoglu, Professor Syed Ali Jafar, and Professor Hamid Jafarkhani, who have also taught me several courses in class. From them, I have gained a deeper understanding of various topics in communication and information theories.

Furthermore, I would like to thank the former and current fellow Ph.D. students at my group, Dr. Ali Fakoorian, Dr. Jing Huang, Dr. Feng Jiang, Misagh Khayambashi, Dr. Amitav Mukherjee, Sean O'Rourke, Dr. Shun Chi Wu, and Xinjie Yang, who have created wonderful intellectual atmosphere in the research group. The discussions with them play a very important role in my work.

I also thank my friends, Fei Huang, Dr. Yulin Shi, Dr. Jianqi Wang, and Yuchen Yao, who shared the enjoyable moments with me at southern California, and helped me when I needed to cope with stress. In addition, I want to show my appreciation to Shankar Venkatraman and Dr. Sundar Krishnamurthy for the unforgettable glad experience during my internship.

Without the encouragement of my family, I could not have finished my dissertation. The endless support from them is what motivates me to proceed with my research. I can think of no words to adequately express the gratitude to them for their unconditional love.

Last but not least, I am grateful to the financial support from the National Science Foundation under grant CCF-0916073 and the Air Force Office of Scientific Research under grant FA9550-10-1-0310, which have funded my research for the past years.

CURRICULUM VITAE

Jie Chen

EDUCATION

Doctor of Philosophy in Electrical and Computer Engineering **2015**
University of California, Irvine *Irvine, California*

Master of Science in Electronic Engineering **2002**
Shanghai Jiao Tong University *Shanghai, China*

Bachelor of Science in Electronic Engineering **1999**
Shanghai Jiao Tong University *Shanghai, China*

RESEARCH EXPERIENCE

Graduate Research Assistant **2008–2014**
University of California, Irvine *Irvine, California*

TEACHING EXPERIENCE

Teaching Assistant **2014–2015**
University of California, Irvine *Irvine, California*

REFEREED JOURNAL PUBLICATIONS

[1] Jie Chen and A. Lee Swindlehurst, “Applying Bargaining Solutions to Resource Allocation in Multiuser MIMO-OFDMA Broadcast Systems,” *IEEE Journal of Selected Topics in Signal Processing*, vol.6, no.2, pp. 127-139, April 2012.

[2] Jie Chen, Feng Jiang and A. Lee Swindlehurst, “The Gaussian CEO Problem for Scalar Sources with Arbitrary Memory,” submitted to *IEEE Transactions on Information Theory.*

[3] Jie Chen, Feng Jiang and A. Lee Swindlehurst, “On the Performance of a Separable STAP Algorithm for Massive Antenna Arrays,” to be submitted.

[4] Feng Jiang, Jie Chen and A. Lee Swindlehurst, “Estimation in Phase-Shift and Forward Wireless Sensor Network,” *IEEE Transactions on Signal Processing*, vol. 61, no. 15, pp. 3840-3851, Aug. 2013.

[5] Feng Jiang, Jie Chen and A. Lee Swindlehurst, “Optimal Power Allocation for Parameter Tracking in a Distributed Amplify-and-Forward Sensor Network,” *IEEE Transactions on Signal Processing*, vol. 62, no. 9, pp. 2200-2211, May 2014.

[6] Feng Jiang, Jie Chen, A. Lee Swindlehurst and Jose Lopez-Salcedo, “Massive MIMO for Wireless Sensing with a Coherent Multiple Access Channel,” accepted by *IEEE Transactions on Signal Processing.*

REFEREED CONFERENCE PUBLICATIONS

[1] Jie Chen and A. Swindlehurst, “Downlink Resource Allocation for Multi-user MIMO-OFDMA Systems: The Kalai-Smorodinsky Bargaining Approach,” in *Proc. 3rd IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, pp. 380-383, Dec. 13-16, 2009.

[2] Jie Chen and A. Lee Swindlehurst, “On the Achievable Sum Rate of Multiterminal Source Coding for a Correlated Gaussian Vector Source,” in *Proc. IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, March 25-30, 2012.

[3] Feng Jiang, Jie Chen and A. Lee Swindlehurst, “Phase-only Analog Encoding for a Multi-antenna Fusion Center,” in *Proc. IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, March 25-30, 2012.

[4] Jie Chen, Feng Jiang and A. Lee Swindlehurst, “The Gaussian CEO Problem for a Scalar Source with Memory: A Necessary Condition,” in *Proc. 46th Asilomar Conference on Signals, Systems, and Computers*, Nov. 4-7, 2012. (Best Student Paper Award)

[5] Feng Jiang, Jie Chen and A. Lee Swindlehurst, “Parameter Tracking via Optimal Distributed Beamforming in an Analog Sensor Network,” in *Proc. 46th Asilomar Conference on Signals, Systems, and Computers*, Nov. 4-7, 2012.

[6] Jie Chen, Feng Jiang, A. Lee Swindlehurst and Jose Lopez-Salcedo, “Localization of Mobile Equipment in Radio Environments with No Line-of-Sight Path,” in *Proc. IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, May 26-31, 2013.

[7] Feng Jiang, Jie Chen and A. Lee Swindlehurst, “Linearly Reconfigurable Kalman Filtering for a Vector Process,” in *Proc. IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, May 26-31, 2013.

[8] Feng Jiang, Jie Chen and A. Lee Swindlehurst, “Detection in Analog Sensor Networks with a Large Scale Antenna Fusion Center,” in *Proc. 8th IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, June 22-25, 2014.

[9] Jie Chen, Feng Jiang and A. Lee Swindlehurst, “Analysis of a Separable STAP Algorithm for Very Large Arrays,” in *Proc. 48th Asilomar Conference on Signals, Systems, and Computers*, Nov. 2-5, 2014.

[10] Rohan Ramlall, Jie Chen and A. Lee Swindlehurst, “Non-Line-of-Sight Mobile Station Positioning Algorithm using TOA, AOA, and Doppler-Shift,” in *Proc. 3nd International Conference and Exhibition on Ubiquitous Positioning Indoor Navigation and Location Based Service (UPINLBS)*, Nov. 20-21, 2014.

ABSTRACT OF THE DISSERTATION

Resource Allocation in Communication, Quantization, and Localization

By

Jie Chen

Doctor of Philosophy in Electrical and Computer Engineering

University of California, Irvine, 2015

Professor A. Lee Swindlehurst, Chair

With the advancement of signal processing and other enabling technologies, new products and services have appeared in large numbers over the past decade, and are changing people's daily lives quickly and profoundly. They could not have occurred without the rapid development in areas like digital communications, information theory, detection and estimation, where resource allocation plays a crucial role.

In this work, we study resource allocation for two classes of problems. The first class is rate allocation in digital systems. The functionality of modern digital systems can be broadly divided into two parts: communications and source coding. For communications, we systematically study the allocation problem from a game theory perspective for the multiuser downlink broadcast channel, and apply the solutions to the special case where spatial block diagonalization is combined with time-sharing to multiplex a subset of the users. For source coding, we consider the achievable sum-rate/distortion tradeoff for the Gaussian central estimation officer problem with a scalar source having arbitrary memory. We formulate the variational problem of minimizing the sum rate subject to a distortion constraint, and the conventional Lagrange method is extended to solve the problem. A sufficient condition is also found that can be used to verify if the necessary solution results in the minimal sum rate.

The second class of problems is target localization. We analyze the performance of a reduced-dimension separable space-time adaptive processing algorithm for radar systems under the large array assumption. The study shows that in the asymptotic sense the simplified scheme performs as well as the fully adaptive algorithm with a significant saving in computational complexity. For target localization in wireless systems, we propose a two-stage approach in order to handle non-line-of-sight transmission based on mild assumptions regarding the propagation environments. For the first stage of positioning, we analyze the cases of scattering and reflection respectively, and propose methods to estimate the position and velocity of the moving target. Once the estimation is done, the results can be used as the initial values of extended Kalman filters in the second stage to track the subsequent movements of the target.

Chapter 1

Introduction

With the advancement of signal processing and other enabling technologies, new products and services have appeared in large numbers over the past two decades, and are changing people's daily lives quickly and profoundly. For example, high-speed wireless communications allow people to watch videos from websites like YouTube on mobile terminals in real time. Location based services help people find local restaurants and attractions easily, and thus create new business models across a broad range of industries. High-definition television and digital broadcasting completely redefine the viewing experience with revolutionary video and audio compression techniques. The application of automotive radar makes driving safer and even ignites the interest of manufacturing computer-controlled driverless cars. Satellite based navigation technologies are wildly used in automobiles and maritime vessels to replace conventional maps and guidance instruments. The list is endless.

These products and services are what people see and use today, which could not have occurred without the rapid development in areas like signal processing, digital communications, information theory, detection and estimation. In a general sense, resource allocation can be defined as how productive assets are distributed among different uses. As it plays a cru-

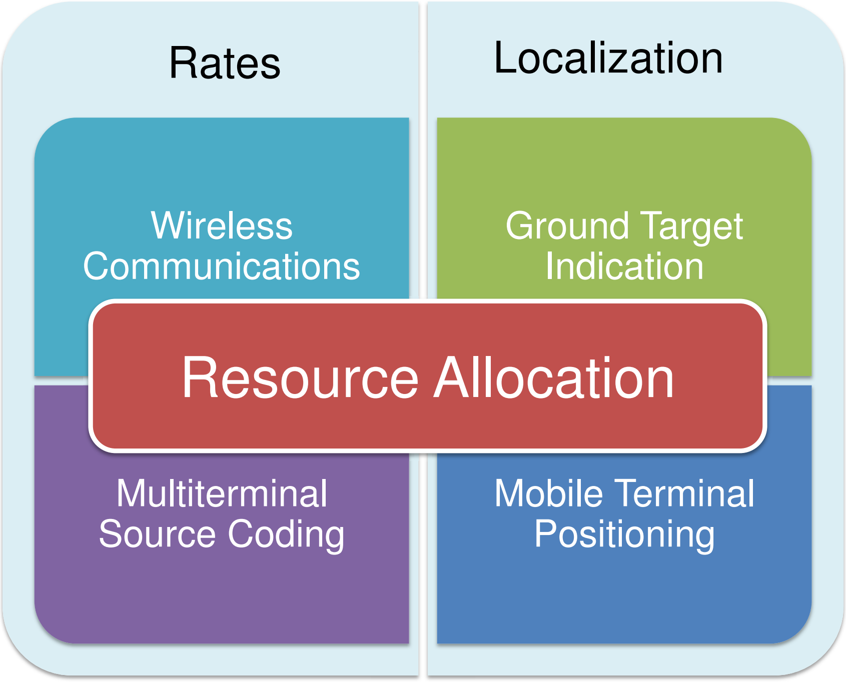


Figure 1.1: Resource allocation: the subject of study.

cial role in economics, health care and education, resource allocation is of great importance in the technological areas, where resources can be power, rate, bandwidth, accuracy, etc. How to allocate resources efficiently is fundamental and pivotal to the performance of all aforementioned systems.

As illustrated in Fig. 1.1, we study resource allocation for two important classes of problems in this work. The first class is rate allocation in digital systems. The functionality of modern digital systems can be broadly divided into two parts: communications and source coding. For both of them, allocated rates compete with other factors in systems, and naturally tradeoffs have to be made among the competing requirements. For example, user multiplexing is regarded as an important technology for increasing the flexibility and efficiency of wireless communication systems, where the sum rate and the fairness for all users in the system can hardly be maximized simultaneously. A well-behaved resource allocation strategy is crucial for the performance of such systems. For source coding, the competing pair is the quantization rate and the distortion in recovered data, where more accurate recovery of the original data requires higher quantization resolution. The second class of problems is target localization. Radar is a target detection technology dealing with line-of-sight (LOS) signals. Motivated by the studies of massive multi-input multi-output (MIMO) in wireless communications which have recently attracted significant research interest, the benefits of very large arrays can also be exploited in other applications such as space-time adaptive processing (STAP) for radar systems in order to save computing resources at the cost of significantly increased number of antenna elements. For target localization in civilian uses, radio positioning has received increasing attention in recent years and found applications in various areas. However, the existence of non-line-of-sight (NLOS) paths introduces considerable positioning errors. To combat NLOS effects, we can either deploy more measuring devices in positioning systems in the hope of obtaining LOS observation of signals, or we can keep the infrastructure unchanged and design smart algorithms to exploit the information contained in all LOS and NLOS paths. Investing on the deployment of more nodes or the enhancement of measuring and estimation capability of existing nodes is well worth consideration. Table 1.1 summarizes all the tradeoffs we consider in this work.

1.1 Organization of the Thesis

In Chapter 2, we systematically study the allocation problem from a game theory perspective for the multiuser downlink broadcast channel. First, we investigate the application of the Nash and Kalai-Smorodinsky bargaining games to a general resource allocation problem and propose algorithms to find the corresponding solutions. Then we apply the general solutions to the special case where spatial block diagonalization is combined with time-sharing to multiplex a subset of the users on every subcarrier. To reduce the computational com-

Table 1.1: Resource Allocation Tradeoffs

Domain	Factor I	Factor II
Wireless Communications	Sum rate in communication	User fairness
Multiterminal Source Coding	Sum rate in quantization	Distortion
Ground Target Indication	Size of antenna array	Requirement on computing resources
Mobile Terminal Positioning	Measuring and positioning capability of a single node	Number of deployed nodes

plexity, a framework for simplifying the resulting algorithms is also given. Numerical results and analysis are provided to compare the performance of the different resource allocation methods.

In Chapter 3, we consider the achievable sum-rate/distortion tradeoff for the Gaussian central estimation officer (CEO) problem with a scalar source having arbitrary memory. We describe how the arbitrary memory problem can be fully characterized by using known results for the vector CEO problem, and then we formulate the variational problem of minimizing the sum rate subject to a distortion constraint. To solve the problem, we extend the conventional Lagrange method and show that if the solution exists, it should consist of a zero part and a non-zero part, where the non-zero part is determined by solving a set of Euler equations. By calculating the second variation of the min-sum-rate problem, a sufficient condition is also found that can be used to determine if the necessary solution results in the minimal sum rate. The special case of two terminals is examined in detail, and it is shown that an analytical solution is possible in this case. Analysis and discussion with examples are provided to illustrate the theoretical results. The general solution obtained in this chapter is shown to be compatible with previous results for cases such as the problem of rate evaluation for sources without memory.

In Chapter 4, we analyze the performance of a reduced-dimension separable STAP algorithm under the large array assumption. We study its performance for clairvoyant interference covariance matrices based on the asymptotic orthogonality of the steering vectors, and we also provide a scaling law for the signal-to-interference plus noise ratio (SINR) loss as the number of antennas grows. The study shows that in the asymptotic sense the simplified scheme performs as well as the fully adaptive STAP method with a significant saving in computational complexity. Appealing to random matrix theory, we finally perform an analysis for the SINR as a function of the number of training samples when the covariance matrix is estimated using a finite collection of secondary data.

In Chapter 5, we propose a two-stage approach in order to handle NLOS scenarios based on mild assumptions regarding the propagation environments. For the first stage of positioning, we analyze the cases of scattering and reflection respectively, and use least squares method to estimate the position and velocity of the moving target. Once the estimation is done, the results can be used as the initial values of extended Kalman filters in the second stage to track the subsequent movements of the target. Compared with previous studies, a minimal set of measurements is required for positioning in our work, *i.e.*, no meaningful estimation could be done with less measurements. Simulation results are provided to show the performance of both stages.

Conclusions are presented in Chapter 6, as well as a description of additional high level questions and technical extensions that are motivated by this work.

Chapter 2

Bargaining Based Resource Allocation in MIMO-OFDMA Systems

2.1 Introduction

The general broadcast or downlink MIMO channel has been studied by many researchers, and the corresponding rate region has been rigorously defined [33, 112]. When the channel state information (CSI) is known at the transmitter, capacity can be achieved by multiuser (MU)-MIMO techniques based on dirty paper coding (DPC) [23]. However, such techniques are computationally prohibitive and not currently suited for application in real systems. Suboptimal but less complex algorithms based on linear processing (*e.g.*, beamforming) have been considered instead for implementation in current wireless standards. A comprehensive discussion of MU-MIMO techniques can be found in [96, 51].

Because of its ability to combat fading in a straightforward way, Orthogonal Frequency Division Modulation (OFDM) has become the basis for most wireless communication standards proposed for the future. Orthogonal Frequency Division Multiple Access (OFDMA) refers to the use of OFDM in allowing multiple users access to a wireless channel, through allocation of the available subcarriers to them. OFDMA provides considerable flexibility in multiuser scenarios, supports various quality of service (QoS) levels, and allows for efficient exploitation of diversity in both the time and frequency domains. Due to the advantages of OFDMA and the potential increase in spectral efficiency from MIMO techniques, many current and almost all newly proposed wireless systems, such as 3GPP LTE [1], LTE-Advanced [3], WiMAX [49] and IEEE 802.16m [50], base their air interfaces on MIMO-OFDMA.

For a multiuser MIMO-OFDMA system, a reasonable allocation of available resources such as power, subcarriers and spatial channels, is crucial to system performance, and there has been considerable research on this topic. Many papers have focused on allocating resources to maximize the sum rate of the system [104, 103, 98], while others have attempted to maintain fairness in terms of QoS among the users, usually according to some heuristic metrics. For instance, [92] tackles the fair allocation problem by assigning different priorities to users and [67] treats requested data rates as weighting factors and then schedules users via a weighted proportional-fair algorithm.

Recently, researchers have begun to interpret wireless communication problems from a game theory perspective, which provides a more formal mechanism for solving resource allocation problems. As a branch of game theory, bargaining games and their corresponding axiomatic solutions have been applied to wireless networks [52, 42] in order to attain a useful tradeoff between overall system efficiency and user fairness. In [69], the scheduling problem for the multiple-input single-output (MISO) interference channel was studied. In [122], the Nash Bargaining Solution (NBS) [68] is applied to a two-user relay setting. The authors of [99] develop a distributed algorithm for spectrum sharing that reasonably approximates the NBS. In [9], the authors show that the NBS can be extended to log-convex utility sets and then study a general wireless scenario where the inter-user interference is the dominating factor for transmission performance. Application of the NBS to 2- and N-player interference channels has been studied in [60, 61, 120, 62]. By generalizing the Kalai-Smorodinsky Bargaining Solution (KSBS) [54] to the multi-player case, the authors of [45] provided load allocation strategies for virtual network sharing.

In this chapter, we focus on the use of bargaining techniques for the MIMO-OFDMA downlink, where a base station (BS) must allocate available resources in order to simultaneously communicate with multiple users. Although the BS sets the transmission parameters in a cellular downlink setting, and the users do not directly cooperate or negotiate with each other, the use of game-theoretic bargaining is still a relevant concept. As stated in [86], a bargaining solution "can be interpreted as an arbitration procedure; *i.e.*, a rule which tells an arbitrator what outcome to select. So long as the arbitration procedure is intended to reflect the relative advantages which the game gives to the players, this interpretation need not be at odds with the interpretation of a solution as a model of the bargaining process." In other words, the basic assumption behind a game-theoretic bargaining solution between several players is that it be identical to what an impartial arbitrator would recommend. In our cellular network application, the BS acts as an arbitrator. It is aware of the payoffs (downlink rates) for each user, and it can enforce the selected outcome. The bargaining solutions serve as the mechanism for making the arbitration decision. In our work, we consider the use of the Nash and Kalai-Smorodinsky bargaining approaches as vehicles for providing a systematic and axiomatic way to address fairness in the multiuser MIMO-OFDMA downlink. An earlier version of this approach was presented in [19].

In [16] and [121], heuristic approximations of bargaining solutions for downlink OFDMA resource allocation are developed, and these papers tackle problems similar to the one we address in this chapter. However, the problem we consider here is more complicated. We attempt to determine the solution in a multi-antenna, multi-carrier setting, which adds significant complexity to the original OFDMA resource allocation problem. Furthermore, we are interested in finding the exact NBS and KSBS (under certain constraints) instead of heuristic approximations. In [48], the empirical performance of the NBS and KSBS is studied for the single-antenna case. The authors of [26] study resource allocation in a similar scenario and derive elegant analytical expressions for both the Nash and KS bargaining solutions. However these closed-form results only hold for very special rate region geometries in which every point on the Pareto boundary corresponds to a max sum-rate solution.

To solve the more general problem, we first establish a mathematical formulation for the bargaining game applied to the downlink MIMO-OFDMA problem. Next, we show that if the resource set is convex and the performance metric with respect to the resource set is concave, then the NBS can be immediately obtained via standard convex optimization techniques. However, the KSBS case is more difficult, and consequently we devise two algorithms with guaranteed convergence that can be used to find the true KSBS. We also present a method for extending KSBS to handle long-term average rate allocation problems, similar to how the proportional-fair algorithm implements a long-term average for the NBS [56, 59]. We demonstrate the use of these algorithms for a special case where the transmitter employs the so-called block diagonalization (BD) algorithm [97] for the transmit precoders on each subcarrier. We show that this special case meets the necessary convexity conditions, and provide details on how to implement the algorithms for this case. Finally we develop a suboptimal but low-complexity algorithmic framework that provides performance close to that obtained with the exact solutions.

The rest of the chapter is organized as follows. In Section 2.2, we describe the system model and formulate the resource allocation problem for the MIMO-OFDMA broadcast channel. In Section 2.3, we provide a brief introduction to the bargaining techniques used in the chapter, and discuss how they can be applied to the resource allocation problem. Section 2.4 proposes methods to solve the problem by using convex optimization techniques, and illustrates the long-term average implementation of the KSBS. Section 2.5 discusses the application of the bargaining solutions for the special case of BD MIMO-OFDMA, followed by a complexity discussion and a simplified algorithmic framework for computing the bargaining solutions in a suboptimal but more practical manner. Numerical simulation results are presented in Section 2.6 and the tradeoffs between efficiency and equity for cases with equal or non-equal pathloss are studied.

2.2 System Model

We consider an MU-MIMO downlink channel with n_T transmit antennas and K users, where user k is equipped with $n_R^{(k)}$ receive antennas. We also assume an N-subcarrier OFDM modulation scheme and that each subcarrier experiences flat fading. This models a typical cellular downlink transmission scenario. With the assumption that linear transmit precoding is performed, the signal received by user k on subcarrier n is

$$\hat{\mathbf{d}}_{k,n} = \mathbf{D}_{k,n} \left(\mathbf{H}_{k,n} \sum_{l=1}^{K} \mathbf{M}_{l,n} \mathbf{d}_{l,n} + \mathbf{n}_{k,n} \right), \tag{2.1}$$

where $\mathbf{d}_{k,n} \in \mathbb{C}^{m_{k,n}}$ carries $m_{k,n}$ data symbols for user k on subcarrier n, $\mathbf{H}_{k,n} \in \mathbb{C}^{n_R^{(k)} \times n_T}$ is the channel matrix, $\mathbf{M}_{k,n} \in \mathbb{C}^{n_T \times m_{k,n}}$ is the transmit precoding matrix, $\mathbf{D}_{k,n} \in \mathbb{C}^{m_{k,n} \times n_R^{(k)}}$ is the receive decoding matrix, and $\mathbf{n}_{k,n} \in \mathbb{C}^{n_R^{(k)}}$ is a complex white Gaussian noise vector. Note that, in general, only a subset of the users will actually use a given subcarrier n for data transmission. Users who are not allocated power for subcarrier n will set $\mathbf{d}_{k,n} = 0$ and no precoder will be computed.

The generic MIMO-OFDMA resource allocation problem consists of determining which subcarriers are assigned to each user (in the MIMO case, several users will in general share each subcarrier), how much power is allocated to each user on each subcarrier, and what transmit precoders will be used. We use the general variable G to denote the possible resource assignments. For example, G could be a tuple $(\mathbf{p}, \boldsymbol{\varphi})$ where $\mathbf{p}$ is the power allocation over all users and subcarriers and $\boldsymbol{\varphi}$ is the subcarrier occupancy indicator. Given an allocation strategy G, we use R_k^G to denote the achievable rate for user k and $\mathbf{R}^G \subset \mathbb{R}^K$ to denote the entire rate region, which in general may or may not be convex.

Based on this system model, the K-user resource allocation problem can be mathematically generalized into a class of optimization problems. Each of these problems has the same constraints, but different objective functions, as follows:

$$\max_{G} \quad U(R_1^G, \ldots, R_K^G) \tag{2.2}$$

$$\text{s.t.} \quad (R_1^G, \ldots, R_K^G) \in \mathbf{R}^G \ , \tag{2.3}$$

where the utility U can have various definitions depending on the specific adopted criterion, such as

$$\text{Max-sum rate: } U = \sum_{k=1}^{K} R_k^G \tag{2.4}$$

$$\text{Max-min rate: } U = \min_{k=1,\ldots,K} R_k^G \ . \tag{2.5}$$

For the max-sum-rate case, a larger share of resources are typically allocated to users with better channel conditions, and weak users often end up with very little throughput. On the other hand, in the max-min case, fairness in the strictest sense becomes the major concern, which typically results in an inefficient resource utilization. In Section 2.3 we will see that game-theoretic bargaining approaches for this problem can be cast in the general framework of (2.2)–(2.3), and used to obtain a meaningful trade-off between overall performance and fairness for individual users.

2.3 Game Theoretic Bargaining Solutions

In this section, we begin by briefly explaining two well-known bargaining solutions [68, 54], and then show how these solutions can be applied to the MIMO-OFDMA resource allocation problem.

2.3.1 Nash Bargaining Solution

A bargaining problem is defined as a pair $(S, \mathbf{d})$, where $S \subset \mathbb{R}^K$ is the set of feasible payoffs and $\mathbf{d} \in S$ is the status-quo or disagreement point. The disagreement point corresponds to the payoffs that the users can obtain in the absence of bargaining, or in our case in the absence of any arbitration by the BS. A single K-dimensional point $\mathbf{u} = (u_1, \ldots, u_K)$ represents a utility vector for K players. For our downlink communications problem, the elements of this vector will correspond to the downlink rates achieved for each user. The axiomatic definition of the NBS is introduced in [68], and for our problem the NBS is equivalent to (2.2)–(2.3) with the following cost function U:

$$\text{NBS: } U = \prod_{k=1}^{K} (R_k^G - R_k^{SQ}) \ , \tag{2.6}$$

where R_k^{SQ} represents the status quo rate of user k. The status quo rate R_k^{SQ} for our application can be chosen to be the minimum rate requirement for user k, which can reasonably be assumed to lie in the rate region through the use of mechanisms such as call admission control.

2.3.2 Kalai-Smorodinsky Bargaining Solution

Others have formulated the game-theoretic bargaining problem with different axiomatic definitions, leading to alternative approaches. One of these is the so-called Kalai-Smorodinsky bargaining solution [54]. In the KSBS formulation, the payoff for the players satisfies

$$\frac{u_1 - u_1^{SQ}}{u_1^{max} - u_1^{SQ}} = \cdots = \frac{u_K - u_K^{SQ}}{u_K^{max} - u_K^{SQ}} \ , \tag{2.7}$$

where u_k^{SQ} denotes the status quo utility for player k, u_k^{max} denotes the maximum possible payoff for player k, and $\{u_k^{max}\}_{k=1}^{K}$ is referred to as the "utopia" point. For our problem, achieving u_k^{max} corresponds to allowing user k to occupy all resources, and thus it is easy to determine. Thus, in the KSBS solution, every user gets the same fraction of his/her maximum possible rate. This interpretation has considerable intuitive appeal, and makes KSBS an attractive approach in situations where one wishes to balance individual fairness with overall system performance. Like the other allocation algorithms considered, the KSBS can also be formulated as the solution to an optimization problem like that in (2.2). The corresponding objective function U is

$$\text{KSBS: } U = r \text{ where } r = \frac{R_k^G - R_k^{SQ}}{R_k^{max} - R_k^{SQ}}, \ \forall k \ . \tag{2.8}$$

Note that we have switched notation from the general utility u_k for user k to the specific value R_k^G denoting user k's rate under allocation G, and $u_k^{max} = R_k^{max}$. For notational simplicity, from now on we will assume that the status quo point for each user is zero in both the NBS and KSBS cases; generalizing our approach for a non-zero status quo is straightforward. When the status quo point is zero, the KSBS is similar to the weighted rate balancing problem, which is discussed in [109] for downlink MIMO beamforming.

2.4 Finding the Bargaining Solutions

Although the task of finding the bargaining solutions can be transformed into optimization problems such as (2.6) and (2.8), the resulting problems will in general be difficult to solve for arbitrary transmission schemes. However, if we can make the following two assumptions:

- (A1) the rate function R_k^G is strictly concave with respect to G, and
- (A2) any constraints on G are convex,

then the optimization problems become easier to solve. By the function composition argument [11], it is straightforward to verify that the negative logarithm of (2.6) is strictly convex with respect to G. Under these assumptions, the NBS can be found using standard numerical techniques, such as the primal-dual interior point method. Note that while assumptions (A1) and (A2) are strong, there are many non-trivial cases where they are satisfied, and hence could be solved using the techniques outlined in this chapter. Some examples include downlink orthogonal CDMA [32], downlink channel inversion MISO-SDMA [79], and downlink MIMO zero-forcing dirty paper coding [12].

Finding the KSBS is more challenging, even with assumptions (A1) and (A2). To see this we rewrite the optimization problem for KSBS here as

$$\begin{aligned} \max_{G} \quad & r \\ \text{s.t.} \quad & r = \frac{R_k^G}{R_k^{max}}, \quad k = 1, 2, \ldots, K\,. \end{aligned}$$

We cannot obtain the KSBS by directly solving this problem with convex optimization techniques due to the fact that the additional equality constraints $r = \frac{R_k^G}{R_k^{max}}, k = 1, \cdots, K$, are not affine with respect to r and G [11]. To circumvent this problem, in the sections below we propose two different approaches that allow us to find the KSBS in an efficient way.

2.4.1 Bisection Search

The first approach is based on the bisection search method, similar to that used in [69] for applying KSBS to MISO inference networks. To see how this method is implemented, recall that the KSBS corresponds to the intersection of the rate region boundary and the line segment from the origin to the utopia point. The goal of the bisection method is to efficiently search along this line segment until the point of intersection is found. The line segment is sequentially bisected, and a feasibility test is conducted to determine if the current bisection point $r' = \frac{R_k^G}{R_k^{max}}$ corresponds to a rate pair that can be achieved for some G. Fortunately, this feasibility test corresponds to a convex optimization problem that is solvable. If the point is feasible, it becomes the new left endpoint of the line segment and the process is repeated. If the point is not feasible, it becomes the new right endpoint of the segment instead. The process continues until the difference between the rate ratios at the endpoints of the line segment is less than some prespecified tolerance, indicating that we are at least that close to the KSBS solution on the boundary of the rate region. The feasibility test at each iteration can be formulated as:

$$\frac{R_k^G}{R_k^{max}} \geq r', \quad k = 1, 2, \ldots, K\,.$$

Due to the assumption that R_k^G is strictly concave with respect to G, we can see that the inequality constraints are strictly convex and therefore this test can be performed by a standard numerical method.

2.4.2 Preference Function Formulation

In [13], the two-user bargaining problem was analyzed, and it was shown that a number of well-known bargaining problems can be unified under one mathematical umbrella through the introduction of the so-called *preference function*. In [45] and [27], the authors show that the preference function concept can be extended to cases involving more than two users. For our problem, we can accordingly define the preference function for user k as

$$Q_k = \frac{R_k^G}{R_k^{max}} + \frac{\beta}{K-1} \sum_{s=1, s\neq k}^{K} \left(1 - \frac{R_s^G}{R_s^{max}}\right), \tag{2.9}$$

and the overall objective function as

$$U = \prod_{k=1}^{K} Q_k = \prod_{k=1}^{K} \left[\frac{R_k^G}{R_k^{max}} + \frac{\beta}{K-1} \sum_{s=1, s\neq k}^{K} \left(1 - \frac{R_s^G}{R_s^{max}}\right)\right]. \tag{2.10}$$

We can utilize the preference function concept to find the KSBS for our resource allocation problem by maximizing (2.10) for $\beta = 1$ and finding the corresponding optimal r and G.

A further investigation reveals that (2.10) is strictly concave with respect to $\{R_k^G\}_{k=1}^K$ for $0 \leq \beta < 1$, but not strictly concave when $\beta = 1$. This results because, in the $\beta = 1$ case, multiple $\{R_k^G\}_{k=1}^K$ tuples can maximize (2.10) only if the ratios $\{\frac{R_k^G}{R_k^{max}}\}_{k=1}^K$ are all equal to r. This means that the different choices of initial point in the numerical search process may lead to different $\{R_k^G\}_{k=1}^K$ tuples that have the same ratio r but are not necessarily Pareto optimal. Among these $\{R_k^G\}_{k=1}^K$ tuples, only the one on the rate region boundary, *i.e.*, the one that is Pareto optimal, is the true KSBS. To get around this difficulty, we propose the following iterative approach to find the KSBS:

1. Select a suitable positive and increasing sequence $\{\beta_i\}_{i=1}^{\infty}$ satisfying $\beta_i < 1$ for all i, but with the sequence converging to 1.

2. At the ith iteration, plug β_i into (2.10) and solve the optimization problem. Since U is strictly concave when $0 \leq \beta < 1$, we can uniquely find the solution $\{R_k^{G*}(\beta_i)\}_{k=1}^K$.

3. Increase i and repeat step 2 until the distance between $\{R_k^{G*}(\beta_{i-1})\}_{k=1}^K$ and $\{R_k^{G*}(\beta_i)\}_{k=1}^K$ is below a predefined tolerance.

The following theorem indicates that this iterative approach converges to the KSBS.

Theorem 2.1. *As β goes to 1, $\{R_k^{G*}(\beta)\}_{k=1}^K$ converges to the KSBS.*

Proof. See Appendix A.1. □

2.4.3 KSBS associated with average rate

The discussion thus far has focused on finding the KSBS for the instantaneous rate allocation problem. In some applications, it is more practical to base the scheduling decisions on long-term rate averages instead. The proportional fair scheduling algorithm is a good example; it is well known that this algorithm results in an average rate allocation that is equivalent to the NBS [56, 59]. In this section, we show that a similar type of solution can be formulated to implement the KSBS for the average user rates.

Assume T_w is the length of the time window over which the rate averaging is to occur, so that the average rate of user k at time n can be expressed as:

$$\overline{R}_k^G(n) = \left(1 - \frac{1}{T_w}\right)\overline{R}_k^G(n-1) + \frac{1}{T_w}R_k^G(n)\,, \tag{2.11}$$

where $R_k^G(n)$ is the rate allocated to user k at time n. We want to find a rule to schedule users for transmission so that in the long run the average rate allocation is the KSBS. Again we use the preference function concept to find the rule. Define

$$f\left(\{\overline{R}_k^G(n)\}_{k=1}^K\right) = \log\prod_{k=1}^K\left[\frac{\overline{R}_k^G(n)}{\overline{R}_k^{max}} + \frac{1}{K-1}\sum_{s=1,s\neq k}^K\left(1 - \frac{\overline{R}_s^G(n)}{\overline{R}_s^{max}}\right)\right]. \tag{2.12}$$

It is easy to verify that f is concave with respect to G. Let $\overline{\mathbf{R}}^{G*}$ denote the long term average KSBS. Thus, $\overline{\mathbf{R}}^{G*}$ should maximize f and fulfill the following optimality condition [11] for arbitrary $\mathbf{R}^G$:

$$\nabla f(\overline{\mathbf{R}}^{G*})(\mathbf{R}^G - \overline{\mathbf{R}}^{G*}) \leq 0, \tag{2.13}$$

where ∇ is the gradient operator, *i.e.*, $\nabla f = \left[\frac{\partial f}{\partial R_1^G}, \ldots, \frac{\partial f}{\partial R_K^G}\right]$.

The first order derivative of f is given by

$$\frac{\partial f}{\partial R_k^G(n)} = \frac{\frac{1/T_w}{\overline{R}_k^{max}}}{\frac{\overline{R}_k^G(n)}{\overline{R}_k^{max}} + \frac{1}{K-1}\sum_{s=1,s\neq k}^K \left(1 - \frac{\overline{R}_s^G(n)}{\overline{R}_s^{max}}\right)} + \sum_{s=1,s\neq k}^{K} \frac{\frac{1}{K-1}\left(-\frac{1/T_w}{\overline{R}_k^{max}}\right)}{\frac{\overline{R}_s^G(n)}{\overline{R}_s^{max}} + \frac{1}{K-1}\sum_{t=1,t\neq s}^K \left(1 - \frac{\overline{R}_t^G(n)}{\overline{R}_t^{max}}\right)} . \tag{2.14}$$

Plugging (2.14) into (2.13) yields the optimization condition

$$\sum_{k=1}^{K} \left\{ \frac{\frac{1/T_w}{\overline{R}_k^{max}}}{\frac{\overline{R}_k^G(n)}{\overline{R}_k^{max}} + \frac{1}{K-1}\sum_{s=1,s\neq k}^K \left(1 - \frac{\overline{R}_s^G(n)}{\overline{R}_s^{max}}\right)} + \sum_{s=1,s\neq k}^{K} \frac{\frac{1}{K-1}\left(-\frac{1/T_w}{\overline{R}_k^{max}}\right)}{\frac{\overline{R}_s^G(n)}{\overline{R}_s^{max}} + \frac{1}{K-1}\sum_{t=1,t\neq s}^K \left(1 - \frac{\overline{R}_t^G(n)}{\overline{R}_t^{max}}\right)} \right\} \times \left(R_k^G(n) - R_k^{G*}(n)\right) \leq 0 \, . \tag{2.15}$$

Note that $\overline{R}_k^G(n)$ is a function of $R_k^G(n)$. The length of the time window T_w is usually set to be very large, so the filtered rate changes very slowly and we can approximately assume that $\overline{R}_k^G(n) = \overline{R}_k^G(n-1)$. Denote this filtered average rate by $\overline{R}_k$. For every scheduling interval n, the optimality condition (2.15) leads to the following rate allocation rule:

$$\mathbf{R}^{G*}(n) = \underset{\mathbf{R}^G(n)}{\arg\max} \sum_{k=1}^{K} \left\{ \frac{\frac{1}{\overline{R}_k^{max}}}{\frac{\overline{R}_k}{\overline{R}_k^{max}} + \frac{1}{K-1}\sum_{s=1,s\neq k}^K \left(1 - \frac{\overline{R}_s}{\overline{R}_s^{max}}\right)} + \sum_{s=1,s\neq k}^{K} \frac{\frac{1}{K-1}\left(-\frac{1}{\overline{R}_k^{max}}\right)}{\frac{\overline{R}_s}{\overline{R}_s^{max}} + \frac{1}{K-1}\sum_{t=1,t\neq s}^K \left(1 - \frac{\overline{R}_t}{\overline{R}_t^{max}}\right)} \right\} \times R_k^G(n) \, . \tag{2.16}$$

If there is more than one rate vector $\mathbf{R}^G(n)$ that maximizes (2.16), we choose the one that is component-wise greater than the others, since the KSBS resides on the boundary of the rate region. In practical wireless systems the candidate rate vector set is discrete and the size of the set is usually small, so an exhaustive search can be used to find the solution. Note that in this case we may not be able to find the exact KSBS, but simulations show that this rule can approximate the KSBS quite well. The form of the KSBS rule is compared with the proportional-fair approach in the equations below for the two-user case:

$$\mathbf{R}^{G*}(n) = \arg\max_{\mathbf{R}^G(n)} \sum_{k=1}^{2} \frac{\frac{1}{\overline{R}_k^{max}} \left(\frac{\overline{R}_{\{1,2\}\backslash\{k\}}}{\overline{R}_{\{1,2\}\backslash\{k\}}^{max}} - \frac{\overline{R}_k}{\overline{R}_k^{max}} \right)}{1 - \left(\frac{\overline{R}_2}{\overline{R}_2^{max}} - \frac{\overline{R}_1}{\overline{R}_1^{max}} \right)^2} \times R_k^G(n) \qquad (KSBS) \tag{2.17}$$

$$\mathbf{R}^{G*}(n) = \arg\max_{\mathbf{R}^G(n)} \sum_{k=1}^{2} \frac{1}{\overline{R}_k} \times R_k^G(n) \qquad (NBS)\,. \tag{2.18}$$

2.5 Application to the MIMO-OFDMA Downlink with Block Diagonalization

The approaches introduced in Section 2.4 are quite general, and can be used to find bargaining solutions for any resource allocation problems provided the rate function R_k^G is strictly concave with respect to G and the constraints on G are convex. In this section we apply the proposed approaches to a special scenario where we use the BD method [97] to calculate the transmit precoders on each subcarrier.

2.5.1 Block Diagonalization

To describe the BD scheme implemented on a given subcarrier n, suppose L_n users have been assigned to this subcarrier, and let $\mathcal{L}_n = \{l_1, \ldots, l_{L_n}\}$ denote the indices corresponding to these users. To compute the precoder for user $l_j \in \mathcal{L}_n$, we form the following matrix:

$$\tilde{\mathbf{H}}_{j,n} = [\mathbf{H}^T_{l_1,n} \;\cdots\; \mathbf{H}^T_{l_{j-1},n} \quad \mathbf{H}^T_{l_{j+1},n} \;\cdots\; \mathbf{H}^T_{l_{L_n},n}]^T \;, \tag{2.19}$$

which is of dimension $\sum_{l \in \mathcal{L}_n - l_j} n_R^{(l)} \times n_T$. For the BD approach [97], $\mathbf{M}_{j,n}$ must lie in the null space of $\tilde{\mathbf{H}}_{j,n}$, which can be found by the singular value decomposition (SVD), provided that n_T is large enough:

$$\tilde{\mathbf{H}}_{j,n} = \tilde{\mathbf{U}}_{j,n}\tilde{\mathbf{\Sigma}}_{j,n}[\tilde{\mathbf{V}}^{(1)}_{j,n} \; \tilde{\mathbf{V}}^{(0)}_{j,n}]^H, \tag{2.20}$$

where $(\cdot)^H$ denotes the complex conjugate transpose, $\tilde{\mathbf{V}}^{(1)}_{j,n} \in \mathbb{C}^{n_T \times \text{rank}(\tilde{\mathbf{H}}_{j,n})}$ holds the first $\text{rank}(\tilde{\mathbf{H}}_{j,n})$ right singular vectors, and $\tilde{\mathbf{V}}^{(0)}_{j,n} \in \mathbb{C}^{n_T \times (n_T - \text{rank}(\tilde{\mathbf{H}}_{j,n}))}$ forms a basis for the null space of $\tilde{\mathbf{H}}_{j,n}$.

With another SVD operation on the matrix $\mathbf{H}_{j,n}\tilde{\mathbf{V}}^{(0)}_{j,n}$, we can find the basis of user j's precoder. The concatenation of all the precoders can be expressed as

$$\mathbf{M}_n = [\mathbf{M}_{l_1,n} \;\; \mathbf{M}_{l_2,n} \;\cdots\; \mathbf{M}_{l_{L_n},n}] = [\tilde{\mathbf{V}}^{(0)}_{l_1,n}\mathbf{V}^{(1)}_{l_1,n} \quad \tilde{\mathbf{V}}^{(0)}_{l_2,n}\mathbf{V}^{(1)}_{l_2,n} \;\cdots\; \tilde{\mathbf{V}}^{(0)}_{l_{L_n},n}\mathbf{V}^{(1)}_{l_{L_n},n}]\mathbf{\Lambda}_n^{\frac{1}{2}} \;,$$

where $\mathbf{\Lambda}_n$ is a diagonal matrix of size $\sum_{j \in \mathcal{L}_n} \text{rank}(\mathbf{H}_{j,n}\tilde{\mathbf{V}}^{(0)}_{j,n})$, whose elements are the power loading factors for each spatial stream, and $\mathbf{V}^{(1)}_{j,n} \in \mathbb{C}^{(n_T - \text{rank}(\tilde{\mathbf{H}}_{j,n})) \times \text{rank}(\mathbf{H}_{j,n}\tilde{\mathbf{V}}^{(0)}_{j,n})}$ are the right singular vectors of $\mathbf{H}_{j,n}\tilde{\mathbf{V}}^{(0)}_{j,n}$. In [97], the authors show that maximizing the sum capacity for the system under the zero-interference constraint requires water-filling on the power loading factors $\mathbf{\Lambda}_n$. On the other hand, for our problem, the elements of $\mathbf{\Lambda}_n$ are adjustable parameters to be allocated by the bargaining solution.

The resulting rate for user k under BD in a MIMO-OFDMA setting will be

$$R_k^G = \sum_{n=1}^{N} \log_2 \left| \mathbf{I} + \frac{1}{N_{k,n}} \mathbf{\Sigma}^2_{k,n}\mathbf{\Lambda}_{k,n} \right| , \tag{2.21}$$

where $\mathbf{I}$ is the identity matrix, $N_{k,n}$ is the noise power, $\mathbf{\Lambda}_{k,n}$ is the submatrix of $\mathbf{\Lambda}_n$ corresponding to user k's power loading factors, and $\mathbf{\Sigma}_{k,n}$ is the diagonal matrix containing the singular values of user k's channel on subcarrier n.

In general, for the required nullspace to exist, the BD scheme assumes that $n_T \geq \sum_{l \in \mathcal{L}_n} n_R^{(l)}$ on every subcarrier n. For situations where n_T is relatively small, a suboptimal approach can be used in order to still implement BD [97]. In this approach, the receiver uses a beamformer to reduce the effective number of spatial channels prior to water-filling. For example, the receiver could choose a subset of the principal left singular vectors of the channel to limit the number of data streams it can receive. In effect, this is like reducing the number of receive antennas, and provides BD with additional degrees of freedom to find a nullspace. In this case the channel $\mathbf{H}_{k,n}$ is regarded as the effective channel formed by the product of the fixed receive beamformers with the actual channel. For our problem, such an approach would have to be implemented suboptimally, with fixed rather than optimized receive beamformers, since the convexity of the problem would be lost if the dimension-reducing beamformers were included as parameters in G.

2.5.2 Time Sharing

In this section, we apply a relaxation to the original model and show that, together with the restriction to BD precoding, a convex programming problem results. Note that, with K users, there are a total of 2^K different user combinations that could be assigned to a given subcarrier n. Some of these combinations will not be feasible for BD, since the sum of the number of receive antennas for users on a given subcarrier cannot exceed n_T. Suppose that after eliminating these infeasible cases, we are left with I possible user combinations on any given subcarrier, and let $\{\varphi_{n,i} : n = 1, \ldots, N \text{ and } i = 1, \ldots, I\}$ denote the set of all possible user combinations over all subcarriers. Furthermore, let $0 \leq \omega_{n,i} \leq 1$ represent the fraction of the time that user combination $\varphi_{n,i}$ is used on subcarrier n. To interpret the physical meaning of $\omega_{n,i}$, consider a block fading transmission scenario in which the channel condition remains unchanged for M consecutive OFDM symbols. During this period, the active users in combination $\varphi_{n,i}$ are allocated $\lfloor M\omega_{n,i}\rfloor$ symbols by the BS. As we will see later, introducing the time sharing factor $\omega_{n,i}$ and allowing a variable power allocation over the time slots [32] make the problem convex and thus more tractable. Similar approaches for modeling subcarrier allocations have been adopted in [14, 83]. Under these assumptions, the rate for user k can now be expressed as

$$\begin{aligned} R_k^G &= \sum_{n=1}^{N}\sum_{i=1}^{I}\omega_{n,i}\log_2\left|\mathbf{I}+\frac{1}{\omega_{n,i}N_{k,n}}\mathbf{\Sigma}_{k,n,i}^2\mathbf{\Lambda}_{k,n,i}\right| \\ &= \sum_{n=1}^{N}\sum_{i=1}^{I}\sum_{t=1}^{n_R^{(k)}}\omega_{n,i}\log_2\left(1+\frac{\sigma_{k,n,i,t}^2\lambda_{k,n,i,t}}{\omega_{n,i}N_{k,n}}\right) \ . \end{aligned} \tag{2.22}$$

Each term in the above sum is strictly concave with respect to $G = [\boldsymbol{\omega}\ \boldsymbol{\lambda}]$, and thus R_k^G is strictly concave in G. Note here that $\boldsymbol{\omega} = \{\omega_{n,i}\}|_{\forall n,i}$ and $\boldsymbol{\lambda} = \{\lambda_{k,n,i,t}\}|_{\forall k,n,i,t}$.

Based on this system model, the K-user resource allocation problem can be generalized into a class of optimization problems with the same constraints, but different objective functions, as follows:

$$\max_{\boldsymbol{\omega},\boldsymbol{\lambda}} \quad U \tag{2.23}$$

$$\text{s.t.} \quad \omega_{n,i} \geq 0, \forall n, i \tag{2.24}$$

$$\lambda_{k,n,i,t} \geq 0, \forall k, n, i, t \tag{2.25}$$

$$\sum_{i=1}^{I}\omega_{n,i} \leq 1, \forall n \tag{2.26}$$

$$\sum_{n=1}^{N}\sum_{i=1}^{I}\sum_{k\in\varphi_{n,i}}\sum_{t=1}^{n_R^{(k)}}\lambda_{k,n,i,t} \leq P\,, \tag{2.27}$$

where U is the objective function of the optimization problem, (2.26) is the time-sharing constraint, and (2.27) is the constraint on the total transmitted power. A similar description can also be found in [42].

We provide a proof in Appendix A.2 that shows the rate region of $\{R_k^G\}_{k=1}^K$ is convex. Now we can apply the approaches introduced in Section 2.4 to the MIMO-OFDMA problem based on BD.

2.5.3 Convex Optimization for NBS

The NBS can be obtained by solving the following optimization problem, which is equivalent to (2.10) implemented with $\beta = 0$:

$$\min_{\boldsymbol{\omega},\boldsymbol{\lambda}} \; -\log\Big(\prod_{k=1}^{K} \frac{R_k^G}{R_k^{max}}\Big) \tag{2.28}$$

$$\text{s.t.} \; -\omega_{n,i} \le 0, \forall n, i \tag{2.29}$$

$$-\lambda_{k,n,i,t} \le 0, \forall k, n, i, t \tag{2.30}$$

$$\sum_{i=1}^{I} \omega_{n,i} - 1 \le 0, \forall n \tag{2.31}$$

$$\sum_{n=1}^{N}\sum_{i=1}^{I}\sum_{k\in\varphi_{n,i}}\sum_{t=1}^{n_R^{(k)}} \lambda_{k,n,i,t} - P \le 0\;. \tag{2.32}$$

Since the logarithm function is monotonic and we know R_k^G is strictly concave, (2.28) is strictly convex and therefore can be iteratively solved using a technique such as the primal-dual interior point method.

2.5.4 Bisection Search for KSBS

The feasibility test for a given r can be formulated as:

$$\frac{R_k^G}{R_k^{max}} \geq r, \quad k = 1, 2, \ldots, K$$
$$\omega_{n,i} \geq 0, \forall n, i$$
$$\lambda_{k,n,i,t} \geq 0, \forall k, n, i, t$$
$$\sum_{i=1}^{I} \omega_{n,i} \leq 1, \forall n$$
$$\sum_{n=1}^{N}\sum_{i=1}^{I}\sum_{k\in\varphi_{n,i}}\sum_{t=1}^{n_R^{(k)}} \lambda_{k,n,i,t} \leq P \ .$$

To put this problem into a more standard form for convex optimization, we introduce an artificial variable s and restate the problem as follows:

$$\min_{\boldsymbol{\omega},\boldsymbol{\lambda}} \quad s \tag{2.33}$$
$$\text{s.t.} \quad r - \frac{R_k^G}{R_k^{max}} - s \leq 0, \forall k \tag{2.34}$$
$$-\,\omega_{n,i} \leq 0, \forall n, i \tag{2.35}$$
$$-\,\lambda_{k,n,i,t} \leq 0, \forall k, n, i, t \tag{2.36}$$
$$\sum_{i=1}^{I} \omega_{n,i} - 1 \leq 0, \forall n \tag{2.37}$$
$$\sum_{n=1}^{N}\sum_{i=1}^{I}\sum_{k\in\varphi_{n,i}}\sum_{t=1}^{n_R^{(k)}} \lambda_{k,n,i,t} - P \leq 0 \ . \tag{2.38}$$

We can see that this new optimization problem is convex by checking the objective function and the constraints. The objective function s and the left-hand side of the constraints except (2.34) are affine, so they are trivially convex. For (2.34), we already know that R_k^G is concave with respect to G. If the resulting solution for s is no greater than 0, the resource allocation $G = [\boldsymbol{\omega}\ \boldsymbol{\lambda}]$ is feasible. Otherwise, the feasibility test fails. Pseudo-code for this

Algorithm 1 Bisection Search Algorithm

```
INPUT: Channel matrices $\mathbf{H}_{n,k}$, power constraint $P$, noise power $N_{k,n}$, and tolerance $\epsilon$.
OUTPUT: Optimal KSBS ratio $r^*$ and corresponding resource allocation result $\boldsymbol{\omega}^*$, $\boldsymbol{\lambda}^*$.
Pick suitable initial feasible vectors $\boldsymbol{\omega}_0$ and $\boldsymbol{\lambda}_0$.
For all $k$, $R_k^{max} \leftarrow \sum_{n=1}^{N}\sum_{i=1}^{I}\omega_{n,i}\log_2|\mathbf{I}+\frac{1}{\omega_{n,i}N_{k,n}}\boldsymbol{\Sigma}_{k,n,i}^2\boldsymbol{\Lambda}_{k,n,i}|$ {To compute the utopia point, all resources are being allocated to user $k$.}
$r^{max} \leftarrow 1$
$r^{min} \leftarrow 0$
while $r^{max} - r^{min} > \epsilon$ do
    $r \leftarrow \frac{1}{2}(r^{max} - r^{min})$
    Optimize $s$ using $\boldsymbol{\omega}_0$ and $\boldsymbol{\lambda}_0$ as an initialization point. The optimum is attained at $\boldsymbol{\omega}$ and $\boldsymbol{\lambda}$.
    if $s > 0$ then
        $r^{max} \leftarrow r$ {infeasible branch}
    else
        $r^{min} \leftarrow r$ {feasible branch}
        $r^* \leftarrow r$, $\boldsymbol{\omega}^* \leftarrow \boldsymbol{\omega}$, $\boldsymbol{\lambda}^* \leftarrow \boldsymbol{\lambda}$
    end if
end while
```

approach is outlined in Algorithm 1.

2.5.5 Preference Function Method for KSBS

Applying the negative logarithm to (2.10), the preference function optimization for our problem can be written in standard form as

$$\min_{\boldsymbol{\omega},\boldsymbol{\lambda}} \; -\log\prod_{k=1}^{K}\left[\frac{R_k^G}{R_k^{max}}+\frac{\beta}{K-1}\sum_{s=1,s\neq k}^{K}(1-\frac{R_s^G}{R_s^{max}})\right] \tag{2.39}$$

$$\text{s.t.} \; -\omega_{n,i}\leq 0, \forall n, i \tag{2.40}$$

$$-\lambda_{k,n,i,t}\leq 0, \forall k, n, i, t \tag{2.41}$$

$$\sum_{i=1}^{I}\omega_{n,i}-1\leq 0, \forall n \tag{2.42}$$

$$\sum_{n=1}^{N}\sum_{i=1}^{I}\sum_{k\in\varphi_{n,i}}\sum_{t=1}^{n_R^{(k)}}\lambda_{k,n,i,t}-P\leq 0\,. \tag{2.43}$$

Algorithm 2 Preference Function Based Algorithm

```
INPUT: Channel matrices $\mathbf{H}_{n,k}$, power constraint $P$, noise power $N_{k,n}$, tolerance $\epsilon$, initial guess $\beta_0$, and scale factor $t$.
OUTPUT: Optimal KSBS ratio $r^*$ and corresponding resource allocation result $\boldsymbol{\omega}^*$, $\boldsymbol{\lambda}^*$.
Pick suitable initial feasible vectors $\boldsymbol{\omega}_0$ and $\boldsymbol{\lambda}_0$.
For all $k$, $R_k^{max} \leftarrow \sum_{n=1}^{N}\sum_{i=1}^{I} \omega_{n,i} \log_2 |\mathbf{I} + \frac{1}{\omega_{n,i}N_{k,n}}\mathbf{\Sigma}_{k,n,i}^2\mathbf{\Lambda}_{k,n,i}|$ {To compute the utopia point, all resources are being allocated to user $k$.}
$\beta \leftarrow \beta_0$
while $\max\{\frac{R_k^G}{R_k^{max}}\} - \min\{\frac{R_k^G}{R_k^{max}}\} > \epsilon$ do
    Optimize the objective function in (2.39) using $\boldsymbol{\omega}_0$ and $\boldsymbol{\lambda}_0$ as an initialization point. The optimum is attained at $\boldsymbol{\omega}$ and $\boldsymbol{\lambda}$.
    $r^* \leftarrow r$, $\boldsymbol{\omega}^* \leftarrow \boldsymbol{\omega}$, $\boldsymbol{\lambda}^* \leftarrow \boldsymbol{\lambda}$, $\beta \leftarrow 1 - t(1-\beta)$
end while
```

We know from the above that this is a convex problem, except when $\beta = 1$. The method introduced in Section 2.4 can be exploited as summarized in Algorithm 2. First we choose an arbitrary $\beta \in [0, 1)$ and check whether it is close enough to the actual KSBS by checking to see if the ratio requirement in (2.7) is satisfied. If not, we increase β and repeat the optimization process. If it is, we can safely claim the solution is good enough and may be used as the KSBS. Comparing this algorithm to Algorithm 1, the first approach is a search along the line segment from the origin to the utopia point, while the second is a search along the boundary of the rate region.

2.5.6 Algorithm Simplification

The solutions presented above have reasonable complexity for situations involving a relatively small number of downlink users. However, the total number I of possible user combinations per subcarrier grows exponentially fast with K, and can make the algorithms computationally intractable when the number of users is large. Although the algorithms can still be used to find performance bounds, simpler approaches may be required for practical implementation. To simplify the algorithms, we can limit the number of possible candidates on each subcarrier. In this section, we discuss how to achieve this goal in two steps. First, for each

Algorithm 3 Reduced Complexity Algorithm

```
INPUT: Channel matrices $\mathbf{H}_{n,k}$, maximum number of users on a single subcarrier $N_u$ and maximum number of combinations $N_c$.
OUTPUT: Optimal KSBS ratio $r^*$ and corresponding resource allocation result $\boldsymbol{\omega}^*$, $\boldsymbol{\lambda}^*$.
for $n = 1$ to $N$ do
    Select the $N_u$ users whose channel matrix has the largest norm on subcarrier $n$.
    Generate all valid combinations of the $N_u$ users, and let $C_n$ denote the number of combinations.
    for $u_n = 1$ to $N_u$ do
        Compute the eigenvalues of the channel matrix of user $u_n$.
    end for
    for $c_n = 1$ to $C_n$ do
        Compute the product of the eigenvalues of every user in combination $c_n$.
    end for
    Select the $N_c$ combinations whose eigenvalue product is the greatest.
end for
Use Algorithm 1 or Algorithm 2 with the selected user combinations on each subcarrier to finish the optimization process and obtain $r^*$, $\boldsymbol{\omega}^*$ and $\boldsymbol{\lambda}^*$.
```

subcarrier, we sort the users based on the strength of their channel on that subcarrier, and we select only the N_u users with the strongest channels for consideration. With sufficient frequency diversity, a different set of users will be chosen for each subcarrier. Second, we only consider a subset of the possible user combinations for that subcarrier by examining the eigenvalues of the channel matrices of all users assigned to a given combination. In particular, we only select the N_c combinations that yield the largest eigenvalue product, considering only those channels that are active in each combination. With these two steps, the optimal algorithms described above can be simplified as outlined in Algorithm 3, which has polynomial complexity in the number of users. Note that if the number of subcarriers N is large enough, the chance for weak users to end up never being one of the N_u users selected on any subcarrier is very small. The algorithm is outlined below for the KSBS case, but it can be used for the NBS as well with obvious modifications.

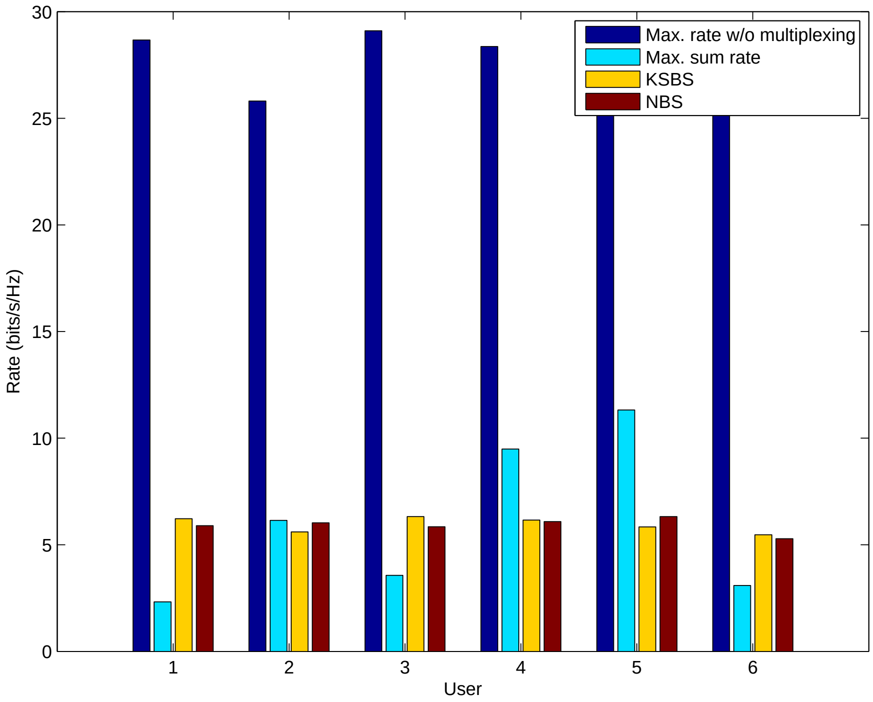


Figure 2.1: Performance comparison of allocation schemes for one channel realization.

2.6 Numerical Results

In this section we present some numerical results to illustrate the performance of the proposed algorithms. To evaluate the effectiveness of the bargaining solutions, we simulated four resource allocation schemes, namely NBS, KSBS, max-sum-rate and the round robin approach, assuming i.i.d. Gaussian MIMO channels. For NBS and KSBS, both the complete and the simplified algorithms are simulated. The max-sum-rate allocation is obtained by solving the optimization problem (2.4). For the round robin allocation, users are sequentially selected to form groups for which the total number of antennas is less than or equal to n_T, then the subcarriers are allocated to each group one after the other. In all of the simulations, the number of antennas at the transmit side is $n_T = 4$, while the number of antennas at the receive side is $n_R = 2$. The total number of subcarriers is $N = 8$ and in the first set of plots there are 6 users in the system. We choose $N_u = 2$ and $N_c = 2$ for the complexity-reduced algorithm. In this simulation, we assume the noise power density for all

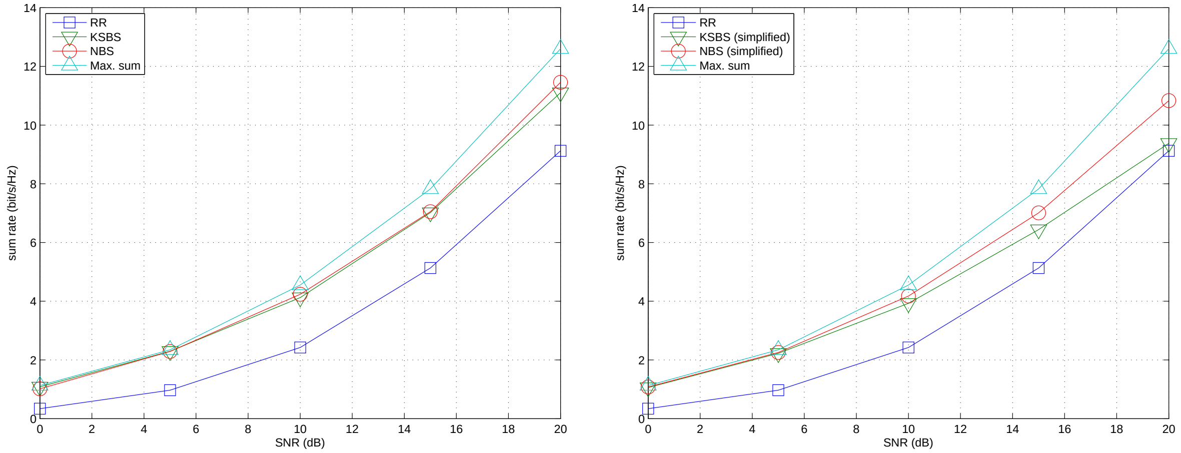


Figure 2.2: Average sum rate vs. SNR: equal pathloss case (left: full algorithm implementation, right: simplified algorithms.)

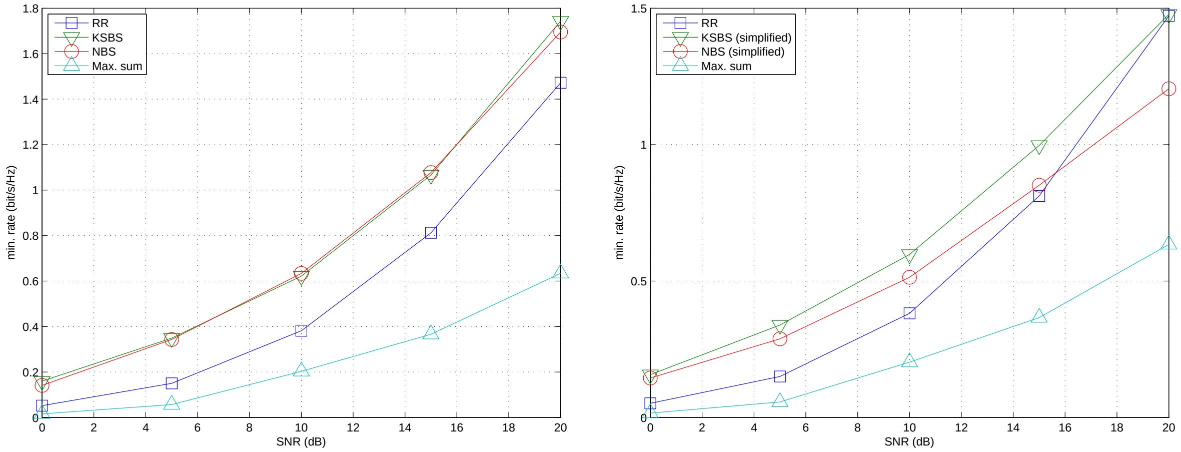


Figure 2.3: Average minimum rate vs. SNR: equal pathloss case (left: full algorithm implementation, right: simplified algorithms.)

users is the same, *i.e.*, $N_{k,n} = N_0$ for all k and n.

Fig. 2.1 shows the results of the respective allocation schemes for one representative channel realization. The theoretical maximum rate R_k^{max} for each user is calculated and is also depicted in the figure. Fig. 2.2 and Fig. 2.3 respectively show the average sum and minimum rate for an SNR range from 0 to 20 dB, where all users experience the same relative pathloss, *i.e.*, where the expected value of the channel norm is the same for all users. As expected, the max-sum-rate algorithm always outperforms the others in terms of sum rate, while the NBS

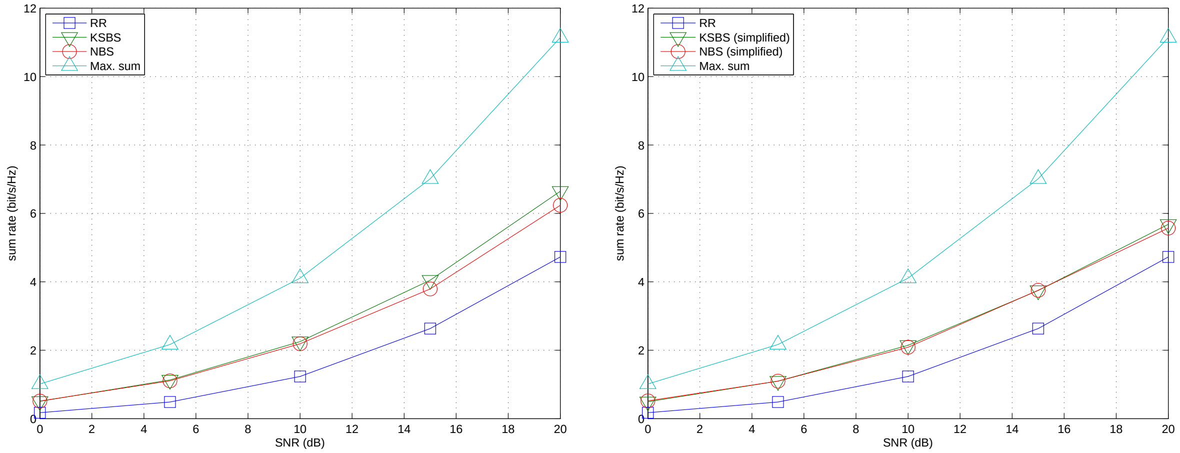


Figure 2.4: Average sum rate vs. SNR: unequal pathloss case (left: full algorithm implementation, right: simplified algorithms.)

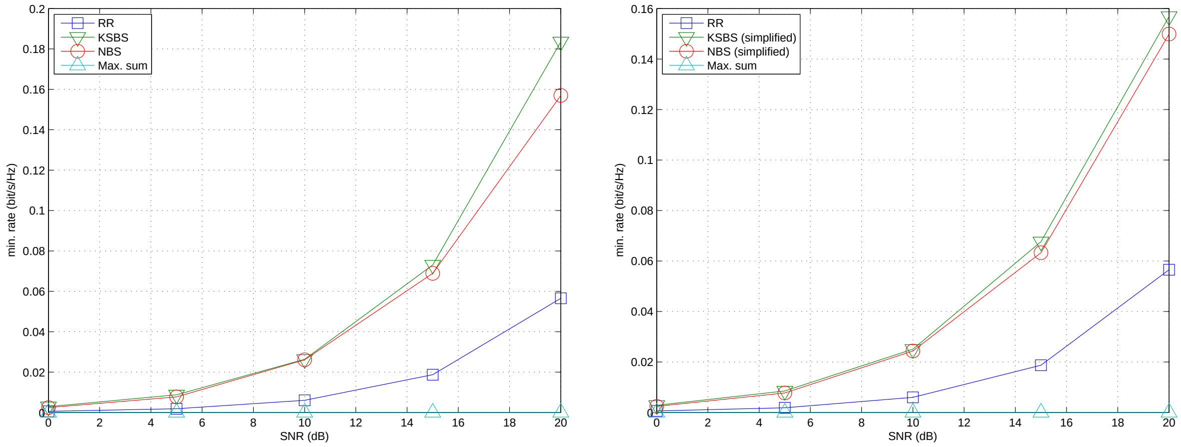


Figure 2.5: Average minimum rate vs. SNR: unequal pathloss case (left: full algorithm implementation, right: simplified algorithms.)

and the KSBS provide a tradeoff between the sum and minimum rate. For this case the rate region is symmetric, which explains why there is little difference in performance between the full implementations of NBS and KSBS. We also notice that the simplified versions of the algorithms do not incur much performance degradation in the low SNR regime, but a more pronounced performance loss at high SNR. This situation can be improved by choosing larger values for N_u and N_c.

Using the same simulation parameters except for the channel gains, Fig. 2.4 and Fig. 2.5

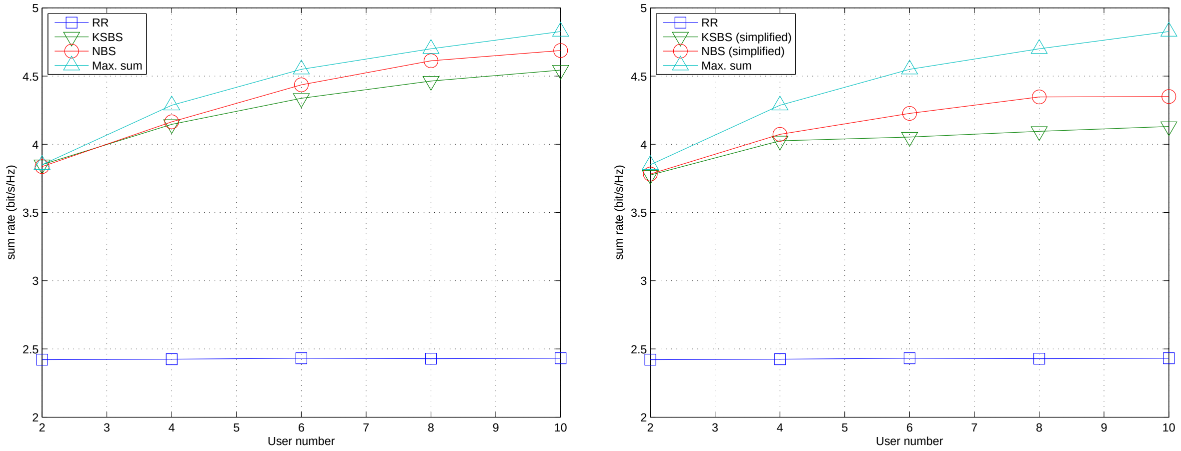


Figure 2.6: Average sum rate for various numbers of users: equal pathloss case (left: full algorithm implementation, right: simplified algorithms.)

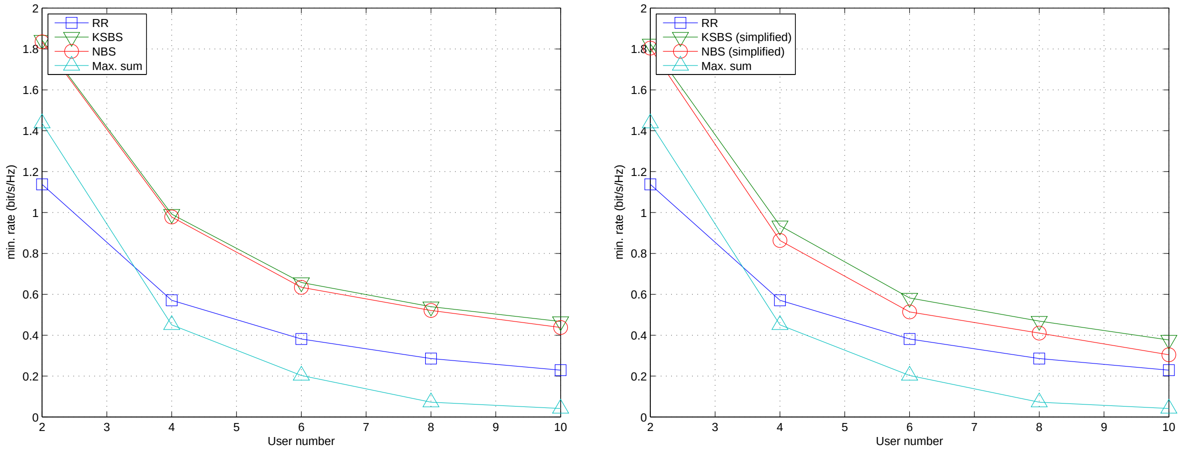


Figure 2.7: Average minimum rate for various numbers of users: equal pathloss case (left: full algorithm implementation, right: simplified algorithms.)

show the performance of the various allocation schemes for the case that half of the users experience an additional 20 dB pathloss. Here, the bargaining solutions provide a more obvious gain in terms of minimum rate.

In Figs. 2.6 and 2.7 we observe the performance of the bargaining solutions from a different perspective. Here we fix the SNR to 10 dB and assume equal pathloss for all users, but we let the number of users vary from 2 to 10. Fig. 2.6 shows the average sum rate and Fig. 2.7 shows the average minimum rate. Clearly, the average sum rate of the round robin

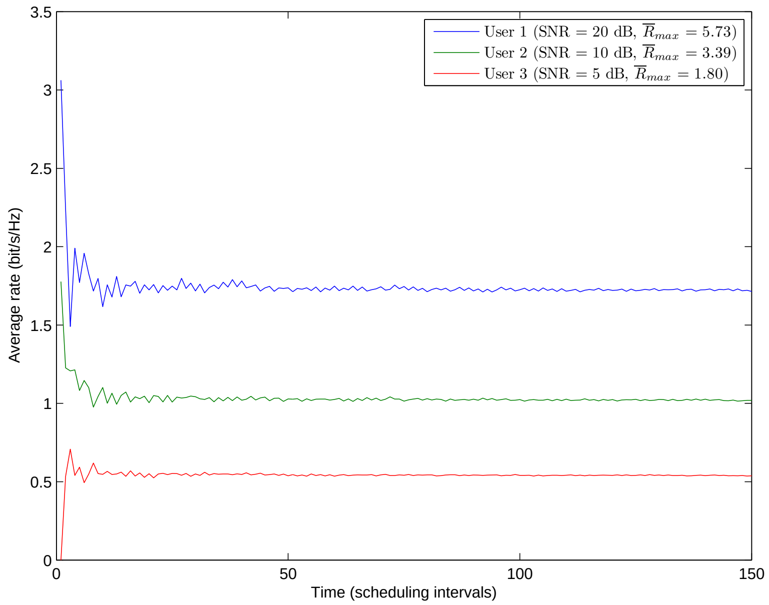


Figure 2.8: Empirical average rates over 150 intervals for 3 users.

algorithm does not change, while the other three algorithms yield a much better sum rate performance. We can also see that the NBS tends slightly more towards the max-sum rate solution, while KSBS tends slightly towards a more equitable solution. Both bargaining solutions significantly outperform the simple round robin algorithm.

In our final simulation example, we implement the scheduling rule described in Section 2.4.3 with recursive average rate updating. The parameter settings are the same as those in the first set of simulation results. Fig. 2.8 shows the evolution of the average rates for 3 users over 150 scheduling intervals. We can see that the average allocated rates become stable quite quickly and the resulting ratios of $\frac{\overline{R}_k}{\overline{R}_k^{max}}$ for the different users are nearly equal, as expected: $(0.299, 0.300, 0.300)$. In this case, all three users get approximately 30% of the maximum rate they could achieve in the single-user scenario.

Chapter 3

The Gaussian CEO Problem for Scalar Sources with Arbitrary Memory

3.1 Introduction

Wireless sensor networks have been the subject of active research for the past decade, and they find many uses in civil, industrial, commercial, and military applications. Such networks are often used for distributed sensing, in which geographically distributed sensors make measurements or local estimates and forward them to a fusion center, which conducts further processing to extract useful information from the data. In practice, the local measurements are typically quantized prior to transmission, and there is clearly a trade-off between the level of quantization (or equivalently the sensors' transmission rates) and the final estimation accuracy. With knowledge of the required accuracy and the statistical characteristics of the source and noise, the fusion center can optimally determine the sensors' individual transmission rates and feed this information back to the sensors in order to efficiently use the available computing and communication resources.

This type of system is equivalent to indirect multiterminal source coding, first studied in [8] and referred to as the CEO problem. In contrast to the direct multiterminal source coding problem [7], where sensors separately measure different but correlated sources and the fusion center attempts to rebuild every source as accurately as possible subject to a sum-rate constraint, each of the sensors in the CEO problem receives a noisy observation of the same source, which is later reconstructed at the fusion center. In order to characterize the rate region for the CEO problem with a Gaussian source, Yamamoto and Itoh first studied the problem in [116] for the two-terminal case, and later it was also discussed independently by Flynn and Gray in [29]. The rate region of the CEO problem was completely characterized for a memoryless Gaussian scalar source with an arbitrary number of observers in [70, 80, 21]. As an extension to the scalar problem, the CEO model with a vector source was also considered in [102, 113, 20].

In recent years, researchers have further attempted to extend the CEO problem in various ways. Pandya et al. generalized the Gaussian CEO problem to the case of a Gaussian vector source where the observation is a noise-corrupted linear transformation of the source signal, and derived an upper bound on the sum rate [74]. Compared to [74], Oohama's model in [71] is more general, allowing the dimension of the source and observation vectors to be different. In addition, [71] provides explicit inner and outer bounds for the rate-distortion region, as well as a sufficient condition under which the lower and upper bounds are equal. This result was later improved on by Yang et al. in [119], who provided a new outer bound using the entropy power inequality for a class of generalized Gaussian CEO problems, together with two sufficient conditions for equality between the outer and inner regions.

The end-to-end rate-distortion performance for different types of source-to-destination communication networks has also been considered. In [114], Xiao et al. discussed the multiterminal source and channel coding problem where quantized observations are sent through an orthogonal multiple access channel to a fusion center, and showed that separate source and channel coding strictly outperforms the use of uncoded transmissions. A variant of the traditional CEO problem was studied in [95] where two terminals observe the same source and deliver noisy measurements to a remote destination through an intermediate relay node. A lower bound on the sum rate was found, and the achievable rates for two special strategies, compute-and-forward and compress-and-forward, were also derived. In [57], Kochman et al. discussed a joint source-channel coding scheme called rematch-and-forward for the scenario where a Gaussian parallel relay network lies between the source and the remote fusion center.

Initiated by the proposal of designs based on generalized coset codes in [81] by Pradhan and Ramchandran, practical coding schemes for the CEO problem also have attracted attention in recent years. While the code design in [81] can be used to arbitrarily allocate rates among source encoders in the achievable rate-distortion region, the gap between its performance and the theoretical limit is still somewhat large. In [117], Yang et al. proposed a high-performance asymmetric coding scheme for the CEO problem which is based on trellis-coded quantization followed by LDPC channel coding, and they showed the resulting code rate is very close to the theoretical bound for a Gaussian source. In [118], they extended the code design to attain the entire rate region through source splitting and channel code partitioning.

Except for a brief discussion in [106], previous studies all assume the sample sequence generated by the source is *memoryless*. In this chapter, we study the achievable sum-rate problem for a Gaussian scalar source with *arbitrary memory*. First, we describe the system model for a source with memory and we characterize the problem using known results for the vector CEO problem. We then formulate the sum-rate calculation as a variational calculus problem with a distortion constraint, and show how to find a necessary condition which the solution to the problem must satisfy. Furthermore, we provide a sufficient condition for determining if the necessary solution achieves the minimal sum rate. A discussion of how to compute the rate-distortion curve is then included, and we note that the solution is compatible with previous findings in rate-distortion theory, which supports the validity of our results. For the special case of a system with two sensor nodes, we derive an analytic expression for the solution to the sum-rate problem and provide further examples to illustrate the theoretical results provided by the necessary and sufficient conditions.

The rest of the chapter is organized as follows. In Section 3.2, we describe the system model and the necessary background and preliminary theorems found in previous work. In Section 3.3, we formulate the sum-rate optimization problem for the L-terminal source coding case. In Section 3.4 and Section 3.5, we present our main results. Section 3.4 provides the necessary condition which can be used to find the solution to the sum-rate problem, while Section 3.5 derives the sufficient condition that can be used to check the optimality of the solution obtained in Section 3.4. Analyses and discussions of our theoretical results can also be found in Section 3.5. Section 3.6 focuses on the two-terminal case, for which we derive an analytic solution and present some representative examples.

3.2 System Model and Preliminaries

The indirect multiterminal source coding problem, or the so-called CEO problem, refers to separate lossy encoding and joint decoding for multiple noise-corrupted observations of a single data source. A block diagram model for the L-terminal CEO problem is illustrated in Fig. 3.1.

The fusion center is interested in recovering the real-valued discrete-time source sequence $x(t), t = 0, \pm 1, \ldots$, where every $x(t)$ is a Gaussian random variable. Without loss of generality, we can take the mean value of $x(t)$ to be zero for all t. The source sequence is assumed to be a stationary stochastic process with arbitrary memory; *i.e.*, in contrast to an

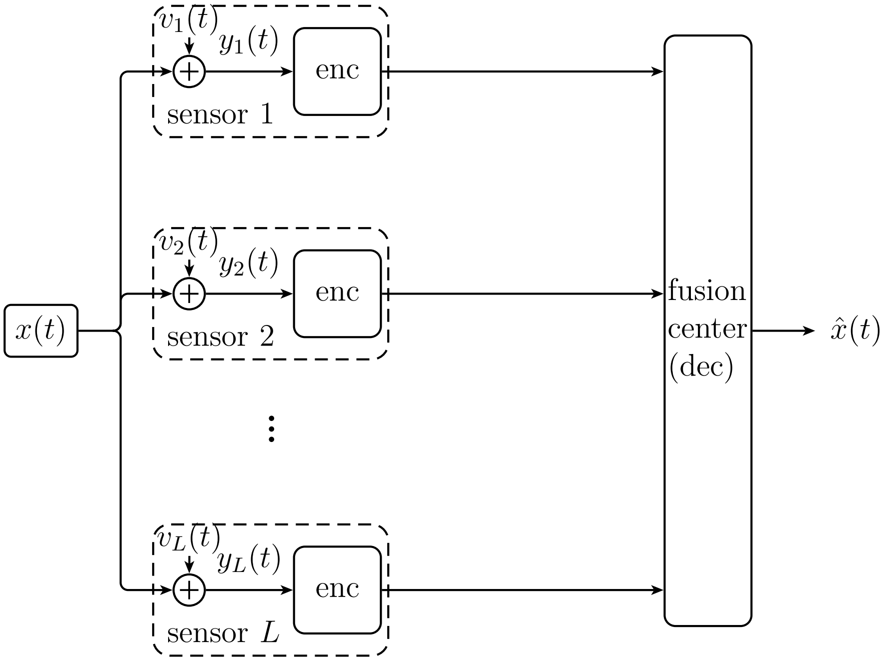


Figure 3.1: Indirect multiterminal source coding (the CEO problem.)

i.i.d. Gaussian source, the power spectral density (PSD) of $x(t)$, denoted by $\Phi_x(\omega)$, is not necessarily constant over frequency. We use $v_i(t)$ to indicate the measurement noise (or local estimation error) at sensor node i, and we assume it is a real-valued zero-mean Gaussian process and i.i.d. in time. The observation at sensor i is given by $y_i(t) = x(t) + v_i(t)$, and $\hat{x}(t)$ is the estimate of $x(t)$ obtained at the fusion center. The error process is defined as $\tilde{x}(t) = x(t) - \hat{x}(t)$. Naturally, the more accuracy required in recovering the source sequence $x(t)$, the higher the source coding rate has to be at the L terminals. We are interested in studying the trade-off between the final fusion error and the sum source coding rate in the Berger-Tung achievable rate sense [21].

3.2.1 Berger-Tung Inner Bound and Sum Rate

Now we describe the calculation of the Berger-Tung achievable sum rate for a scalar Gaussian source with arbitrary memory. We use the method introduced in [6] to evaluate the Berger-Tung achievable sum rate, $R(D)$, for a given target distortion D. Instead of directly calculating the sum rate for an infinitely long Gaussian source sequence with arbitrary memory, we begin with the evaluation of the achievable sum rate, $R_n(D)$, for a Gaussian source word $\mathbf{x} = (x(1), \ldots, x(n))$ and the corresponding reconstructed word $\hat{\mathbf{x}} = (\hat{x}(1), \ldots, \hat{x}(n))$ with the distortion constraint $E\left\{\frac{1}{n}\sum_{t=1}^{n}(x(t) - \hat{x}(t))^2\right\} \leq D$. Letting the size of the source word $\mathbf{x}$ go to infinity, we then have $R(D) = \lim_{n\to\infty} R_n(D)$. Similarly, the distortion constraint becomes $E\{(x(t) - \hat{x}(t))^2\} \leq D$ due to the stationarity of the source process.

For completeness, we briefly introduce the evaluation of the Berger-Tung sum rate for the problem of limited-size vectors. Let $\mathbf{y}_i, i = 1, \ldots, L$ denote the noise-corrupted observation vectors at the sensor nodes. From the discussion above we know that all $\mathbf{y}_i$ are also Gaussian distributed. Based on [21, 102, 113], if there exist auxiliary random vectors $\mathbf{w}_i, i = 1, \ldots, L$ such that $\mathbf{w}_i \to \mathbf{y}_i \ \to \left(\mathbf{x}, \mathbf{y}_{\{i\}^c}, \mathbf{w}_{\{i\}^c}\right)$ forms a Markov chain for all i, and if the distortion between $\mathbf{x}$ and $\hat{\mathbf{x}}$ is no greater than D, the Berger-Tung achievable rate region is the convex hull of

$$\mathbf{R}(\mathbf{w}_1, \ldots, \mathbf{w}_L) = \left\{(R_1, \ldots, R_L)\Big|\sum_{i\in\mathcal{A}} R_i \geq I(\mathbf{y}_{\mathcal{A}}; \mathbf{w}_{\mathcal{A}}|\mathbf{w}_{\mathcal{A}^c}), \forall\mathcal{A} \subseteq \mathcal{I}_L\right\}, \tag{3.1}$$

where $\mathcal{I}_L = \{1, \ldots, L\}$. Thus, the minimal sum rate is

$$R_n(D) = \min_{\mathbf{w}_1,\ldots,\mathbf{w}_L} \frac{1}{n} I(\mathbf{y}_1, \ldots, \mathbf{y}_L; \mathbf{w}_1, \ldots, \mathbf{w}_L). \tag{3.2}$$

The Gaussianity of $\mathbf{y}_i$ and $\mathbf{w}_i$ results in [20]

$$R_n(D) = \min_{\mathbf{w}_1,\ldots,\mathbf{w}_L} \frac{1}{2n} \log \frac{\det(\mathbf{C}_{\mathbf{y}})\det(\mathbf{C}_{\mathbf{w}})}{\det(\mathbf{C}_{\mathbf{yw}})}, \tag{3.3}$$

where $\mathbf{C}_{\mathbf{y}}$, $\mathbf{C}_{\mathbf{w}}$, and $\mathbf{C}_{\mathbf{yw}}$ are the covariance matrices of vectors $(\mathbf{y}_1^T, \ldots, \mathbf{y}_L^T)^T$, $(\mathbf{w}_1^T, \ldots, \mathbf{w}_L^T)^T$ and $(\mathbf{y}_1^T, \ldots, \mathbf{y}_L^T, \mathbf{w}_1^T, \ldots, \mathbf{w}_L^T)^T$, respectively. It is easy to show that $\mathbf{C}_{\mathbf{y}}$, $\mathbf{C}_{\mathbf{w}}$, and $\mathbf{C}_{\mathbf{yw}}$ are all block-symmetric matrices, and that each of their sub-blocks is Toeplitz. Throughout the chapter, all logarithms are assumed to take the natural base, and thus rates are always measured in nats.

3.2.2 Extension of the Toeplitz Distribution Theorem

In [6], evaluation of the rate-distortion function for the point-to-point problem relies on the Toeplitz distribution theorem [35]. Due to the block structure of the mutual information matrices in (3.3), we need to apply an extension [30] of the Toeplitz distribution theorem to our problem, as discussed next.

Let $\mathbf{T}$ denote a $c \times c$ block matrix with $n \times n$ Toeplitz submatrix blocks, *i.e.*,

$$\mathbf{T} = \begin{bmatrix} T_{1,1} & \dots & T_{1,c} \\ \vdots & \ddots & \vdots \\ T_{c,1} & \dots & T_{c,c} \end{bmatrix}$$

and

$$T_{i,j} = \begin{bmatrix} T_{i,j}(1) & \dots & \dots & T_{i,j}(-n) \\ T_{i,j}(2) & T_{i,j}(1) & \dots & T_{i,j}(-n+1) \\ \vdots & \ddots & \ddots & \vdots \\ T_{i,j}(n) & \dots & T_{i,j}(2) & T_{i,j}(1) \end{bmatrix}.$$

If $\mathbf{T}$ is also Hermitian, then from Theorem 3 of [30] we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{cn} F\left(\lambda_k(\mathbf{T})\right) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{u=1}^{c} F(\lambda_u(\mathbf{\Theta}(\omega)))d\omega, \tag{3.4}$$

where F is an arbitrary function, λ_k for $k = 1, \dots, cn$ and λ_u for $u = 1, \dots, c$ are respectively the eigenvalues of $\mathbf{T}$ and $\mathbf{\Theta}(\omega)$, and

$$\mathbf{\Theta}(\omega) = \begin{bmatrix} \sum_{k=-\infty}^{\infty} T_{1,1}(k)e^{-jk\omega} & \dots & \sum_{k=-\infty}^{\infty} T_{1,c}(k)e^{-jk\omega} \\ \vdots & \ddots & \vdots \\ \sum_{k=-\infty}^{\infty} T_{c,1}(k)e^{-jk\omega} & \dots & \sum_{k=-\infty}^{\infty} T_{c,c}(k)e^{-jk\omega} \end{bmatrix}.$$

In the next section, we will see how letting the size of the word $\mathbf{x}$ go to infinity can help us obtain a closed-form expression for the mutual information and the distortion, and hence the problem formulation can be fully derived.

3.3 Problem Formulation

In this section, we provide a full formulation of our problem. In order to calculate the mutual information, we appeal to the extension of the Toeplitz distribution theorem described above. We also describe the optimal estimator structure and the corresponding mean squared error (MSE), which is needed to define the distortion constraint. Finally we will see that the rate evaluation is an infinite-dimensional variational problem.

3.3.1 Mutual Information

The calculation of mutual information is based on (3.3) and (3.4). We first assign the log function to F in (3.4), and then let the size of the matrices in (3.3) go to infinity so that the extension of the Toeplitz distribution theorem can be applied. After applying (3.4) to (3.3), we get the expression of the sum rate for the source-with-memory problem:

$$I(y_1(t), \dots, y_L(t); w_1(t), \dots, w_L(t)) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \log \frac{\det \mathbf{\Phi}_y(\omega) \det \mathbf{\Phi}_w(\omega)}{\det \mathbf{\Phi}_{yw}(\omega)} d\omega \,, \qquad (3.5)$$

where

$$\mathbf{\Phi}_y(\omega) = \begin{bmatrix} \Phi_{y_1}(\omega) & \Phi_{y_1y_2}(\omega) & \cdots & \Phi_{y_1y_L}(\omega) \\ \Phi_{y_2y_1}(\omega) & \Phi_{y_2}(\omega) & \cdots & \Phi_{y_2y_L}(\omega) \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_{y_Ly_1}(\omega) & \Phi_{y_Ly_2}(\omega) & \cdots & \Phi_{y_L}(\omega) \end{bmatrix},$$

$$\mathbf{\Phi}_w(\omega) = \begin{bmatrix} \Phi_{w_1}(\omega) & \Phi_{w_1w_2}(\omega) & \cdots & \Phi_{w_1w_L}(\omega) \\ \Phi_{w_2w_1}(\omega) & \Phi_{w_2}(\omega) & \cdots & \Phi_{w_2w_L}(\omega) \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_{w_Lw_1}(\omega) & \Phi_{w_Lw_2}(\omega) & \cdots & \Phi_{w_L}(\omega) \end{bmatrix},$$

$$\mathbf{\Phi}_{yw}(\omega) = \left[\begin{array}{ccc|ccc} \Phi_{y_1}(\omega) & \cdots & \Phi_{y_1y_L}(\omega) & \Phi_{y_1w_1}(\omega) & \cdots & \Phi_{y_1w_L}(\omega) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \Phi_{y_Ly_1}(\omega) & \cdots & \Phi_{y_L}(\omega) & \Phi_{y_Lw_1}(\omega) & \cdots & \Phi_{y_Lw_L}(\omega) \\ \hline \Phi_{w_1y_1}(\omega) & \cdots & \Phi_{w_1y_L}(\omega) & \Phi_{w_1}(\omega) & \cdots & \Phi_{w_1w_L}(\omega) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \Phi_{w_Ly_1}(\omega) & \cdots & \Phi_{w_Ly_L}(\omega) & \Phi_{w_Lw_1}(\omega) & \cdots & \Phi_{w_L}(\omega) \end{array}\right],$$

and $\Phi_{y_i}(\omega)$, $\Phi_{y_iy_j}(\omega)$, $\Phi_{w_i}(\omega)$, $\Phi_{w_iw_j}(\omega)$, $\Phi_{y_iw_j}(\omega)$, $\Phi_{w_iy_j}(\omega)$ are the auto- and cross-PSDs of the corresponding stochastic processes.

To further evaluate the mutual information expression of (3.5), we need to use the concept of *forward test channels* [25]. Assuming the outputs of the dequantizers at the fusion center, $w_i(t), i = 1, \ldots, L$ are also Gaussian, we can write the whole system in the forward test

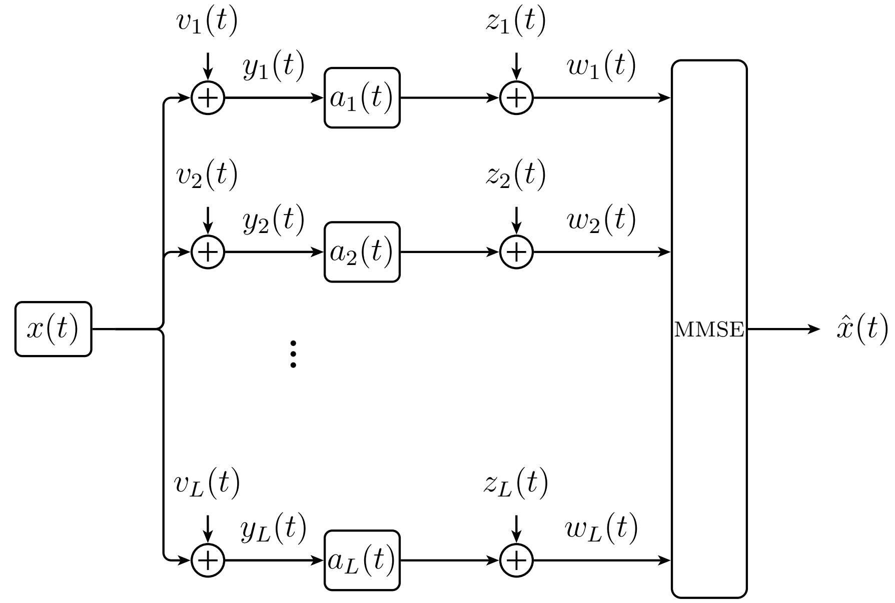


Figure 3.2: Forward test channels.

channel form [6]:

$$w_i(t) = a_i(t) * y_i(t) + z_i(t), \quad i = 1, \ldots, L, \tag{3.6}$$

where $*$ represents convolution, $a_i(t), i = 1, \ldots, L$ are variable real-valued functions to be determined, and $z_i(t), i = 1, \ldots, L$ represent quantization noise processes, which are assumed to be i.i.d. in time and independent from all $y_i(t)$. The forward test channels are depicted in Fig. 3.2. Unlike problems with memoryless sources, convolution is required here instead of simple multiplication.

Given (3.6), the auto- and cross-PSD functions are given by:

$$\begin{cases} \Phi_{y_i}(\omega) = \Phi_x(\omega) + \Phi_{v_i}(\omega) & \\ \Phi_{w_i}(\omega) = |A_i(\omega)|^2 \Phi_{y_i}(\omega) + \Phi_{z_i}(\omega) & \\ \Phi_{w_i y_i}(\omega) = A_i(\omega)\Phi_{y_i}(\omega) & \\ \Phi_{y_i w_i}(\omega) = A_i^*(\omega)\Phi_{y_i}(\omega) & \\ \Phi_{x y_i}(\omega) = \Phi_x(\omega) & \\ \Phi_{x w_i}(\omega) = A_i^*(\omega)\Phi_x(\omega) & \\ \Phi_{w_i y_j}(\omega) = A_i(\omega)\Phi_x(\omega) & i \neq j \\ \Phi_{y_i w_j}(\omega) = A_j^*(\omega)\Phi_x(\omega) & i \neq j \\ \Phi_{y_i y_j}(\omega) = \Phi_x(\omega) & i \neq j \\ \Phi_{w_i w_j}(\omega) = A_i(\omega)A_j^*(\omega)\Phi_x(\omega) & i \neq j \end{cases} , \tag{3.7}$$

where i and j take values in $\mathcal{I}_L$, $A_i(\omega)$ is the discrete time Fourier transform of $a_i(t)$, and $A_i^*(\omega)$ is its conjugate. Plugging (3.7) into (3.5) and after a long series of mathematical manipulations, we can simplify (3.5) to the following expression:

$$I(y_1(t), \ldots, y_L(t); w_1(t), \ldots, w_L(t)) =$$
$$\frac{1}{4\pi} \int_{-\pi}^{\pi} \log \left\{ \frac{\prod_{i=1}^{L} \left[\Phi_{v_i}(\omega) + \Phi_{m_i}^{-1}(\omega)\right]}{\prod_{i=1}^{L} \Phi_{m_i}^{-1}(\omega)} \cdot \left(1 + \sum_{i=1}^{L} \Phi_x(\omega) \left[\Phi_{v_i}(\omega) + \Phi_{m_i}^{-1}(\omega)\right]^{-1} \right) \right\} d\omega \, , \tag{3.8}$$

where $\Phi_{m_i}(\omega) = \left(|A_i(\omega)|^{-2} \Phi_{z_i}(\omega)\right)^{-1}$, and $\Phi_{z_i}(\omega)$ is the PSD of $z_i(t)$. The sum rate is now parameterized in terms of the unknown functions $\Phi_{m_i}(\omega), i = 1, \ldots, L$, from which all $A_i(\omega)$ can be found.

3.3.2 MSE and Optimal Estimator

For memoryless sources, minimizing the MSE between $x(t)$ and $\hat{x}(t)$ can be achieved by means of a simple estimator. Our problem involving sources with arbitrary memory is more complicated, since we need to estimate an entire stochastic process using multiple infinite-length sequences as the input. The estimator will have the form:

$$\hat{x}(t) = \sum_{i=1}^{L} h_i(t) * w_i(t), \tag{3.9}$$

where $h_i(t)$ for $i = 1, \ldots, L$ are optimal infinite-length filters. For known $h_1(t), \ldots, h_L(t)$, the MSE is given by

$$E\left\{\tilde{x}^2(t)\right\} = E\left\{(x(t) - \hat{x}(t))^2\right\} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\Phi_x(\omega) - \sum_{i=1}^{L} \Phi_{xw_i}(\omega) H_i^*(\omega) \right) d\omega\,, \tag{3.10}$$

where $H_i(\omega)$ is the discrete-time Fourier transform of $h_i(t)$, and $\Phi_{xw_i}(\omega)$ is the cross-PSD for $x(t)$ and $w_i(t)$.

At first glance, it appears we may need to find an explicit expression for $H_i(\omega), i = 1, \ldots, L$ in order to connect the MSE solution with the sum-rate criterion in (3.8). The optimal estimator should satisfy the frequency domain Wiener-Hopf equations [53]:

$$\mathbf{H}(\omega)\mathbf{\Phi}_w(\omega) = \mathbf{\Phi}_{xw}(\omega), \tag{3.11}$$

where $\mathbf{H}(\omega) = [H_1(\omega), H_2(\omega), \ldots, H_L(\omega)]$ and $\mathbf{\Phi}_{xw}(\omega) = [\Phi_{xw_1}(\omega), \Phi_{xw_2}(\omega), \ldots, \Phi_{xw_L}(\omega)]$. Fortunately, we do not need to solve (3.11). We simply need an analytic expression for

$\sum_{i=1}^{L}\Phi_{xw_i}(\omega)H_i^*(\omega)$, which can be found as follows:

$$\begin{aligned}
\sum_{i=1}^{L}\Phi_{xw_i}(\omega)H_i^*(\omega) &= \mathbf{\Phi}_{xw}(\omega)\mathbf{H}^H(\omega)\\
&\overset{(a)}{=} \mathbf{\Phi}_{xw}(\omega)\mathbf{\Phi}_w^{-1}(\omega)\mathbf{\Phi}_{xw}^H(\omega)\\
&\overset{(b)}{=} \det\left[1+\mathbf{\Phi}_{xw}(\omega)\mathbf{\Phi}_w^{-1}(\omega)\mathbf{\Phi}_{xw}^H(\omega)\right]-1\\
&\overset{(c)}{=} \frac{\det\left[\mathbf{\Phi}_w(\omega)+\mathbf{\Phi}_{xw}^H(\omega)\mathbf{\Phi}_{xw}(\omega)\right]}{\det\mathbf{\Phi}_w(\omega)}-1\\
&= \frac{\sum_{i=1}^{L}[\Phi_x^2(\omega)]\left[\Phi_{v_i}(\omega)+\Phi_{m_i}^{-1}(\omega)\right]^{-1}}{1+\sum_{i=1}^{L}\Phi_x(\omega)\left[\Phi_{v_i}(\omega)+\Phi_{m_i}^{-1}(\omega)\right]^{-1}},
\end{aligned} \tag{3.12}$$

where (a) follows from (3.11), (b) uses the fact that the determinant of a scalar is the scalar itself, and (c) follows from the matrix determinant lemma [43].

Plugging (3.12) back into (3.10), we end up with the final MSE expression in terms of the unknown functions $\Phi_{m_i}(\omega), i = 1, \ldots, L$:

$$E\left\{\tilde{x}^2(t)\right\} = \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{\Phi_x^{-1}(\omega)+\sum_{i=1}^{L}\left[\Phi_{v_i}(\omega)+\Phi_{m_i}^{-1}(\omega)\right]^{-1}}d\omega. \tag{3.13}$$

3.3.3 Full Formulation

With the above derivations, we can now give the full formulation of the minimum sum-rate problem. Assuming that the rate is minimized when the distortion target is achieved with equality, the rate evaluation problem is given by:

$$\min_{\Phi_{m_i}} \frac{1}{4\pi}\int_{-\pi}^{\pi} \log\left\{ \frac{\prod_{i=1}^{L}\left[\Phi_{v_i}(\omega)+\Phi_{m_i}^{-1}(\omega)\right]}{\prod_{i=1}^{L}\Phi_{m_i}^{-1}(\omega)} \left(1+\sum_{i=1}^{L}\Phi_x(\omega)\left[\Phi_{v_i}(\omega)+\Phi_{m_i}^{-1}(\omega)\right]^{-1}\right)\right\} d\omega \tag{3.14}$$

$$\text{s.t. } \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{1}{\Phi_x^{-1}(\omega)+\sum_{i=1}^{L}\left[\Phi_{v_i}(\omega)+\Phi_{m_i}^{-1}(\omega)\right]^{-1}} d\omega = D \tag{3.15}$$

$$\Phi_{m_i} \geq 0, \quad i=1,\ldots,L\,. \tag{3.16}$$

This is a functional optimization problem and can be tackled via the calculus of variations.

3.4 Necessary Condition

In this section we derive a necessary condition that the solution to the sum-rate problem (3.14)-(3.16) must satisfy. As we have seen in Section 3.3, the optimization in its final form is a constrained variational problem. In fact it can be regarded as an isoperimetric problem with additional inequality constraints. The following theorem shows that in our formulation the solution consists of a term that is zero together with a non-zero term that is determined by solving a set of Euler equations. It is well known in the theory of the calculus of variations that solving the Euler equations results in a necessary condition.

Theorem 3.1 (Isoperimetric problem with additional inequality constraints)**.** *Let u_i and ψ_i, $i=1,\ldots,n$ be functions of ω. For the following functional optimization problem,*

$$\min_{u_1,\ldots,u_n} J(u_1,\ldots,u_n) = \int_a^b f(\omega,u_1,\ldots,u_n,u_1',\ldots,u_n')d\omega \tag{3.17}$$

$$\text{s.t. } \int_a^b g(\omega,u_1,\ldots,u_n,u_1',\ldots,u_n')d\omega = D \tag{3.18}$$

$$u_i \geq \psi_i, i=1,\ldots,n, \tag{3.19}$$

the solution u_i should consist of two parts: ψ_i and a function no less than ψ_i. The latter can be obtained through solving the Euler equations with Lagrange multiplier λ:

$$f_{u_i} - \frac{d}{d\omega} f_{u_i'} + \lambda \left(g_{u_i} - \frac{d}{d\omega} g_{u_i'} \right) = 0, \; i = 1 \ldots, n. \tag{3.20}$$

Proof. See Appendix A.3. □

We can apply Theorem 3.1 to the sum-rate problem (3.14)–(3.16), where the corresponding functions and the upper and lower bounds of integration are

$$\begin{aligned} u_i(\omega) &= \Phi_{m_i}(\omega), \; i = 1, \ldots, L\,, \\ f &= \frac{1}{4\pi} \log \left\{ \frac{\prod_{i=1}^{L} \left[\Phi_{v_i}(\omega) + \Phi_{m_i}^{-1}(\omega) \right]}{\prod_{i=1}^{L} \Phi_{m_i}^{-1}(\omega)} \left(1 + \sum_{i=1}^{L} \Phi_x(\omega) \left[\Phi_{v_i}(\omega) + \Phi_{m_i}^{-1}(\omega) \right]^{-1} \right) \right\}, \\ g &= \frac{1/(2\pi)}{\Phi_x^{-1}(\omega) + \sum_{i=1}^{L} \left[\Phi_{v_i}(\omega) + \Phi_{m_i}^{-1}(\omega) \right]^{-1}}, \\ \psi_i(\omega) &= 0, \; i = 1, \ldots, L\,, \\ a &= -\pi\,, \\ b &= \pi\,. \end{aligned}$$

We note that in our problem the functions f and g do not depend on the derivatives of $u_1, \ldots, u_n$, which means the Euler equations are not differential equations as often encountered in variational problems. This fact eases the solving of the Euler equations, which are

$$\begin{aligned} &\left(\Phi_{v_i}(\omega) \Phi_{m_i}(\omega) + 1 \right) \Phi_{v_i}(\omega) \cdot \left(\frac{1}{\Phi_x(\omega)} + \sum_{i=1}^{L} \frac{1}{\Phi_{v_i}(\omega) + \frac{1}{\Phi_{m_i}(\omega)}} \right)^2 \\ &+ \left(\frac{1}{\Phi_x(\omega)} + \sum_{i=1}^{L} \frac{1}{\Phi_{v_i}(\omega) + \frac{1}{\Phi_{m_i}(\omega)}} \right) = \lambda, \; i = 1, \ldots, L. \end{aligned} \tag{3.21}$$

Problem (3.14)–(3.16) can also be formulated in another way, where the objective is to minimize the distortion (3.15) for a target sum rate (3.14) equal to a given R. Based on the proof for Theorem 3.1, we find that the two formulations are essentially the same with the trivial difference that the Lagrange multiplier in the dual problem becomes $1/\lambda$.

3.5 Sufficient Condition

In this section we find a sufficient condition for problem (3.14)–(3.16) to have a local minimum. To exploit the sufficient condition, we can first use Theorem 3.1 in Section 3.4 to obtain a solution based on the necessary condition, and then use the sufficient condition discussed in this section to check if the solution minimizes the sum rate.

As a first step, we prove the following lemma.

Lemma 3.1 (Degenerate isoperimetric problem)**.** *Let u_i, $i = 1, \ldots, n$ be functions of ω. For the following functional optimization problem to have a local minimum at $(u_1, \ldots, u_n)$,*

$$\min_{u_1,\ldots,u_n} J(u_1, \ldots, u_n) = \int_a^b f(\omega, u_1, \ldots, u_n) d\omega \tag{3.22}$$

$$\text{s.t.} \int_a^b g(\omega, u_1, \ldots, u_n) d\omega = D, \tag{3.23}$$

the sufficient condition is

$$\mathbf{Z}(u_1, \ldots, u_n) \succ 0, \tag{3.24}$$

where

$$\mathbf{Z}(u_1,\ldots,u_n) = \begin{bmatrix} \frac{\partial^2\zeta}{\partial u_1^2} & \cdots & \frac{\partial^2\zeta}{\partial u_1 \partial u_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2\zeta}{\partial u_n \partial u_1} & \cdots & \frac{\partial^2\zeta}{\partial u_n^2} \end{bmatrix},$$

$$\zeta(\omega, u_1,\ldots,u_n) = f(\omega, u_1,\ldots,u_n) + \lambda g(\omega, u_1,\ldots,u_n),$$

and λ is the associated Lagrange multiplier.

Proof. See Appendix A.4. □

Lemma 3.1 is referred to as a degenerate isoperimetric problem because the derivatives of $u_1,\ldots,u_n$ are explicitly excluded in (3.22) and (3.23), which is appropriate for our minimum sum-rate problem. Using Lemma 3.1, we can further obtain the following theorem, which is directly applicable to problem (3.14)–(3.16).

Theorem 3.2 (Degenerate isoperimetric problem with additional inequality constraints)**.** *Let u_i and ψ_i, $i = 1,\ldots,n$ be functions of ω. For the following functional optimization problem,*

$$\min_{u_1,\ldots,u_n} J(u_1,\ldots,u_n) = \int_a^b f(\omega, u_1,\ldots,u_n)d\omega \tag{3.25}$$

$$\text{s.t.} \int_a^b g(\omega, u_1,\ldots,u_n)d\omega = D \tag{3.26}$$

$$u_i \geq \psi_i, i = 1,\ldots,n, \tag{3.27}$$

the tuple $(u_1,\ldots,u_n)$ is a local minimizer for $J(u_1,\ldots,u_n)$ if there exist λ and $u_i, i = 1,\ldots,n$

such that the following matrix is positive-definite:

$$\mathbf{Z}(z_1,\ldots,z_n) =$$
$$\begin{bmatrix} f_{u_1} + 2z_1^2 f_{u_1u_1} + \lambda g_{u_1} + 2\lambda z_1^2 g_{u_1u_1} & \cdots & 2z_1z_n f_{u_1u_n} + 2\lambda z_1 z_n g_{u_1u_n} \\ \vdots & \ddots & \vdots \\ 2z_nz_1 f_{u_nu_1} + 2\lambda z_n z_1 g_{u_nu_1} & \cdots & f_{u_n} + 2z_n^2 f_{u_nu_n} + \lambda g_{u_n} + 2\lambda z_n^2 g_{u_nu_n} \end{bmatrix}, \quad (3.28)$$

and z_i in (3.28) *should fulfill the following conditions:*

1. *for all i, $z_i(f_{u_i} + \lambda g_{u_i}) = 0$ holds;*

2. *for a given i, if the Euler equation $f_{u_i} + \lambda g_{u_i} = 0$ holds, then $z_i = \pm(u_i - \psi_i)^{1/2}$. Otherwise $z_i = 0$.*

Proof. See Appendix A.5. □

The problem in Theorem 3.2 is a degenerate isoperimetric problem with additional inequality constraints, because it is merely the same problem described in Theorem 3.1 but enhanced with further requirements on u_i.

3.5.1 Analysis and Discussion

Global Optimality – Theorem 3.2 only guarantees the existence of a locally optimal solution. Since functional (3.14) is not convex, a given local minimum is not guaranteed to be a global minimum. Thus we cannot safely claim that solving the equations in (3.21) will lead to the true Berger-Tung achievable sum rate, even if the sufficiency of the solution is verified by Theorem 3.2. However, since the number of potential solutions is usually limited, theoretically it is possible to exhaustively calculate all rates associated with the corresponding solutions, and then choose the smallest of these as the global minimum achievable sum rate.

Evaluation of the Achievable Sum-Rate Function – To calculate the achievable sum rate as a function of the distortion D, one would normally substitute the solution to (3.21) into (3.15), and then solve it to get the value of λ. With knowledge of λ, the intermediate optimization variables, $\Phi_{m_1}(\omega), \ldots, \Phi_{m_L}(\omega)$, can be determined and then used in (3.14) to completely determine the achievable rate R. An alternative approach, similar to that explained in [6], is to first pick a value for λ, then use it in (3.21) to get $\Phi_{m_1}(\omega), \ldots, \Phi_{m_L}(\omega)$, which are further substituted into (3.14) and (3.15) to get the corresponding sum rate R and distortion D. The entire achievable rate-distortion curve can be found in this fashion.

Compatibility with Known Results – It is straightforward to show that in the memoryless source case, our general solution reduces to that obtained, for example, in [21]. This can be done via a simple notation change to the functional optimization problem in (3.14)–(3.16). Numerical computations further verify that both approaches generate the same rate-distortion curve.

3.6 Special Case - Two Terminals

In this section we first discuss the problem formulation and corresponding solution for a special case where the number of terminals is $L = 2$. Then some examples are given to illustrate the theoretical results obtained in previous sections. By numerically evaluating the sufficient condition in Section 3.5, we also show that our solution is sufficient for a first-order Gauss-Markov source process.

3.6.1 Formulation and Solution

In [17], a special case where there are only two terminals in the system was considered. The optimal Wiener filters can be explicitly obtained in this case:

$$\begin{aligned}\begin{bmatrix} H_1(\omega) \\ H_2(\omega) \end{bmatrix}^T &= \big[|A_1|^2|A_2|^2(\Phi_x\Phi_{v_1}+\Phi_x\Phi_{v_2}+\Phi_{v_1}\Phi_{v_2}) \\ &\quad +|A_1|^2\Phi_{y_1}\Phi_{z_2}+|A_2|^2\Phi_{y_2}\Phi_{z_1}+\Phi_{z_1}\Phi_{z_2}\big]^{-1}\cdot\begin{bmatrix} A_1^*\Phi_x(|A_2|^2\Phi_{v_2}+\Phi_{z_2}) \\ A_2^*\Phi_x(|A_1|^2\Phi_{v_1}+\Phi_{z_1}) \end{bmatrix}^T,\end{aligned} \tag{3.29}$$

and the minimum sum-rate problem becomes

$$\min_{\Phi_{m_1},\Phi_{m_2}} \frac{1}{4\pi}\int_{-\pi}^{\pi}\log\frac{(\Phi_{y_1}+\Phi_{m_1}^{-1})(\Phi_{y_2}+\Phi_{m_2}^{-1})-\Phi_x^2}{\Phi_{m_1}^{-1}\Phi_{m_2}^{-1}}d\omega \tag{3.30}$$

$$\text{s.t. } \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1}{\frac{1}{\Phi_x}+\frac{1}{\Phi_{v_1}+\frac{1}{\Phi_{m_1}}}+\frac{1}{\Phi_{v_2}+\frac{1}{\Phi_{m_2}}}}d\omega = D \tag{3.31}$$

$$\Phi_{m_1} \geq 0 \tag{3.32}$$

$$\Phi_{m_2} \geq 0. \tag{3.33}$$

Note that A_i and all Φ's are functions of ω, which has simply been dropped to save space.

While it is difficult to obtain a closed-form solution to (3.14)–(3.16) for the general L-terminal case, an analytical solution is possible when $L = 2$. In this case, $n = 2$, $u_i = \Phi_{m_i}$ and $\psi_i = 0$, and we can directly apply Theorem 3.1. Due to the fact that there are no derivatives of $\Phi_i(\omega)$ in the integrand of the objective function, the Euler equations turn out to be merely algebraic rather than differential equations. After eliminating the infeasible solutions to these equations, we end up with the solution:

$$\Phi_{m_1} = \max\left(0, \frac{-\left(\Phi_{v_1}^2(\Phi_x + \Phi_{v_2})^2 + \frac{1}{2}\Phi_x\Phi_{v_1}\Phi_{v_2}(-\lambda\Phi_x\Phi_{v_2} + \Phi_x + \Phi_{v_2}) - \frac{1}{2}\Phi_x^2\Phi_{v_2}^2 - \frac{1}{2}\Phi_x\Phi_{v_2}\Xi\right)}{\Phi_{v_1}(\Phi_x\Phi_{v_1} + \Phi_x\Phi_{v_2} + \Phi_{v_1}\Phi_{v_2})^2}\right)$$

$$\Phi_{m_2} = \max\left(0, \frac{-\left(\Phi_{v_2}^2(\Phi_x + \Phi_{v_1})^2 + \frac{1}{2}\Phi_x\Phi_{v_1}\Phi_{v_2}(-\lambda\Phi_x\Phi_{v_1} + \Phi_x + \Phi_{v_1}) - \frac{1}{2}\Phi_x^2\Phi_{v_1}^2 - \frac{1}{2}\Phi_x\Phi_{v_1}\Xi\right)}{\Phi_{v_2}(\Phi_x\Phi_{v_1} + \Phi_x\Phi_{v_2} + \Phi_{v_1}\Phi_{v_2})^2}\right)$$

$$\begin{aligned}\Xi = \pm(&\lambda^2\Phi_x^2\Phi_{v_1}^2\Phi_{v_2}^2 + 6\lambda\Phi_x^2\Phi_{v_1}^2\Phi_{v_2} + 6\lambda\Phi_x^2\Phi_{v_1}\Phi_{v_2}^2 + 6\lambda\Phi_x\Phi_{v_1}^2\Phi_{v_2}^2 \\ &+\Phi_x^2\Phi_{v_1}^2 + 2\Phi_x^2\Phi_{v_1}\Phi_{v_2} + \Phi_x^2\Phi_{v_2}^2 + 2\Phi_x\Phi_{v_1}^2\Phi_{v_2} + 2\Phi_x\Phi_{v_1}\Phi_{v_2}^2 + \Phi_{v_1}^2\Phi_{v_2}^2)^{\frac{1}{2}}\,,\end{aligned}$$

where the Lagrange multiplier λ should be carefully picked so that (3.31) is fulfilled.

3.6.2 Examples

The form of the solution for some special source processes is discussed below.

First-order Gauss-Markov Process

A scalar discrete-time first-order Gauss-Markov process can be recursively defined as

$$x(t) = \rho x(t-1) + u(t), \tag{3.34}$$

where $u(t) \sim \mathcal{N}(0, \sigma^2)$ is the driving noise. The autocorrelation function of this process is known to be

$$\phi_x(t) = \frac{\sigma^2\rho^{|t|}}{1-\rho^2}, \tag{3.35}$$

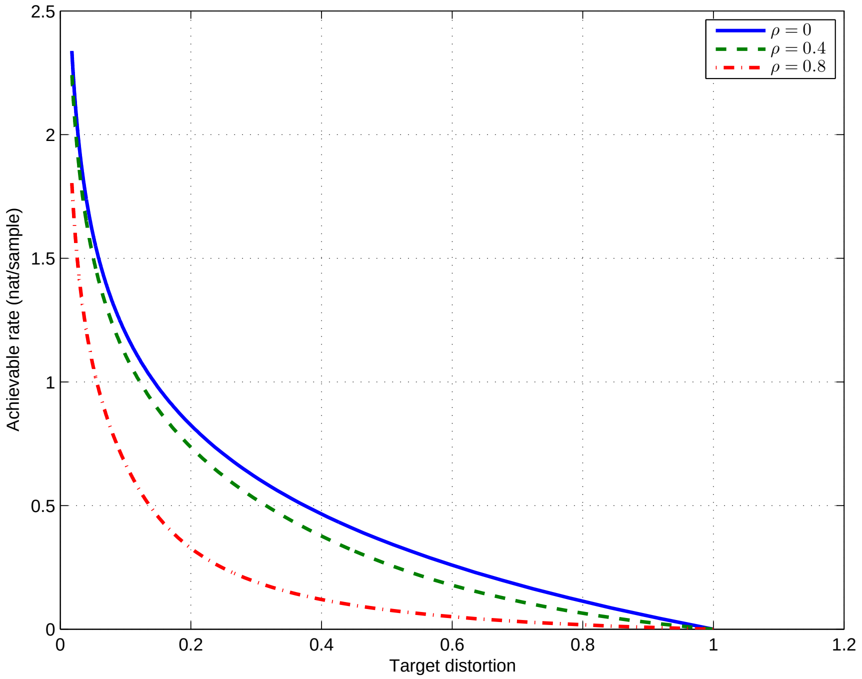


Figure 3.3: Achievable sum rate versus target distortion (first-order Gauss-Markov processes.)

where $|\rho|$ should be less than 1 or otherwise the process is not stationary. The corresponding PSD is

$$\Phi_x(\omega) = \frac{\sigma^2}{1 - 2\rho\cos\omega + \rho^2}. \tag{3.36}$$

Fig. 3.3 shows the rate-distortion curves that result from our solution for the three values $\rho = 0, 0.4,$ and 0.8. The power of the source process is fixed at 1, and the noise power at both sensors is 0.01 (when the noise power is equal at both sensors, we refer to this as the "symmetric" case). As expected, the rate requirement decreases as the correlation increases.

We also show the PSDs of the source process, $\Phi_x(\omega)$, and the error process, $\Phi_{\tilde{x}}(\omega)$, in Fig. 3.4 for a target distortion level $D = 0.7$ when $\rho = 0.8$. In this case $\Phi_{m_1}(\omega)$ is equal to $\Phi_{m_2}(\omega)$.

PSDs for an asymmetric noise case where the variance of the noise at sensor 2 is increased to 0.0135 are also provided. Fig. 3.5 shows the PSDs of $\Phi_x(\omega)$. In this case, the solution yields

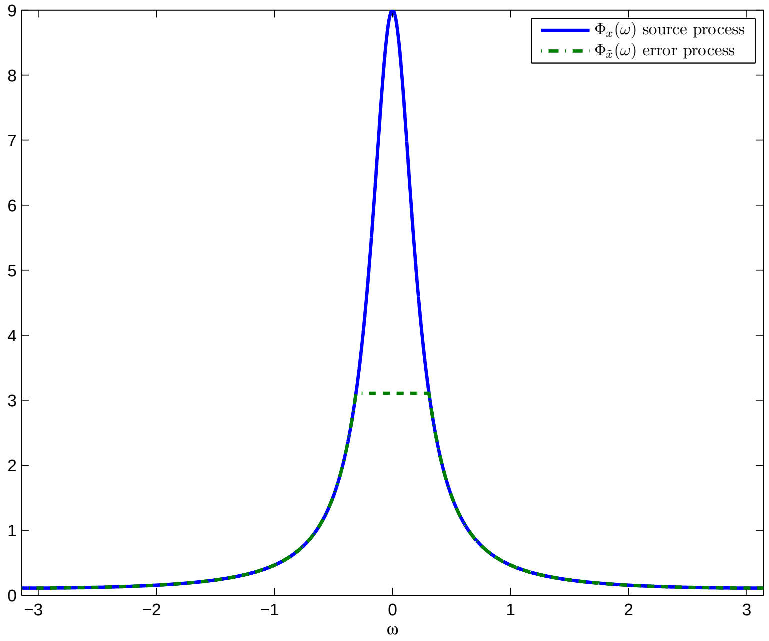


Figure 3.4: PSDs of the source and the error processes (first-order Gauss-Markov process, symmetric noise.)

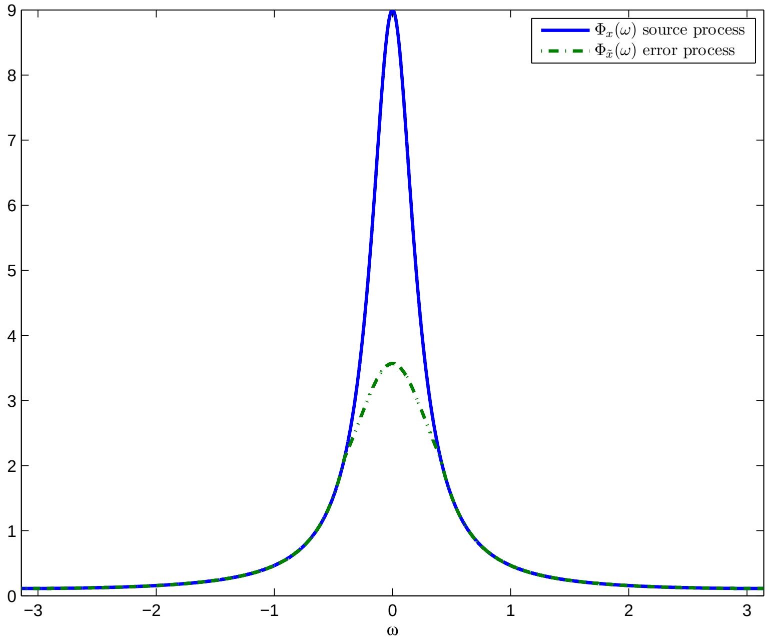


Figure 3.5: PSDs of the source and the error processes (first-order Gauss-Markov process, asymmetric noise.)

$\Phi_{m_2}(\omega) = 0$, and thus sensor node 2 has no contribution to the final estimation performance, and is simply switched off. From Figs. 3.4 and 3.5 we can see that for first-order Gauss-

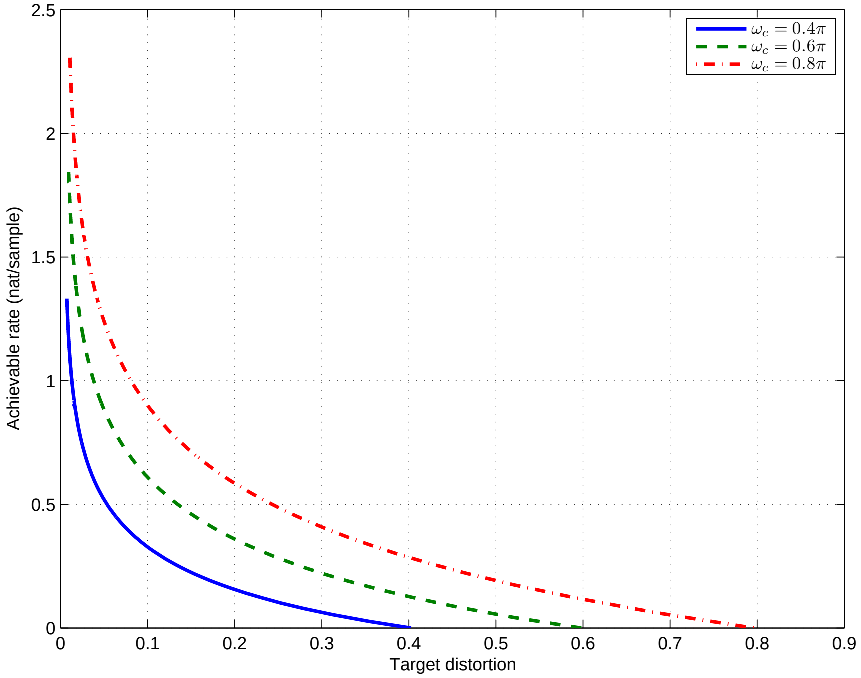


Figure 3.6: Achievable sum rate versus target distortion (band-limited Gaussian processes.)

Markov processes, optimal performance is achieved by using most of the rate to preserve the information in the central area of the source spectrum; there is no need to reconstruct portions of the source process whose information is contained in the off-center areas of the spectrum, which make relatively little contribution to the overall MSE.

Band-limited Gaussian Process

In the second example, the source is a band-limited Gaussian process obtained by passing white Gaussian noise through a fourth-order Butterworth low-pass filter. The power of the unfiltered source process is equal to 1 and the noise power at each sensor is 0.01 for the symmetric noise case. The rate-distortion curves derived from our analysis are shown in Fig. 3.6 for low-pass cut-off frequencies $\omega_c = 0.4\pi, 0.6\pi,$ and 0.8π. Fig. 3.7 shows the PSDs of the source process, $\Phi_x(\omega)$, and the error process, $\Phi_{\tilde{x}}(\omega)$, for the symmetric case, where the target distortion level D is 0.2 and the cut-off frequency is $\omega_c = 0.6\pi$.

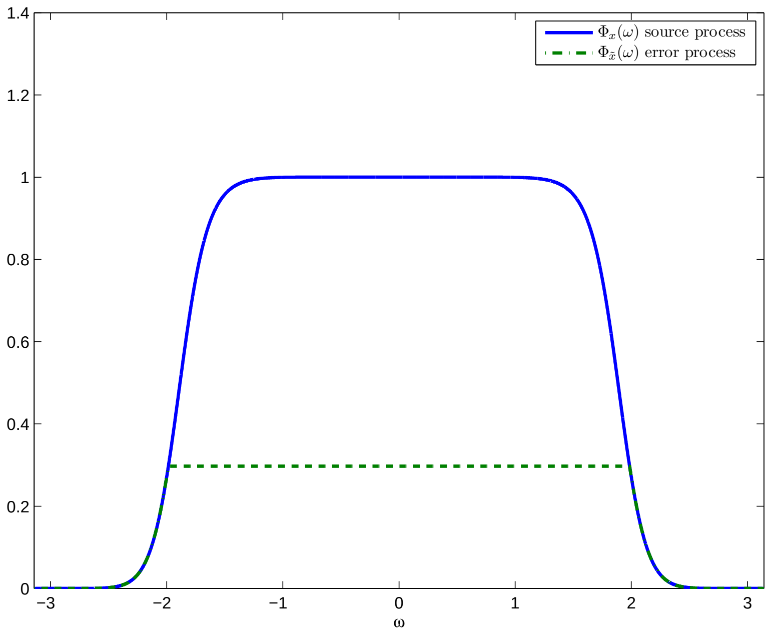


Figure 3.7: PSDs of the source and the error processes (band-limited Gaussian processes, symmetric noise.)

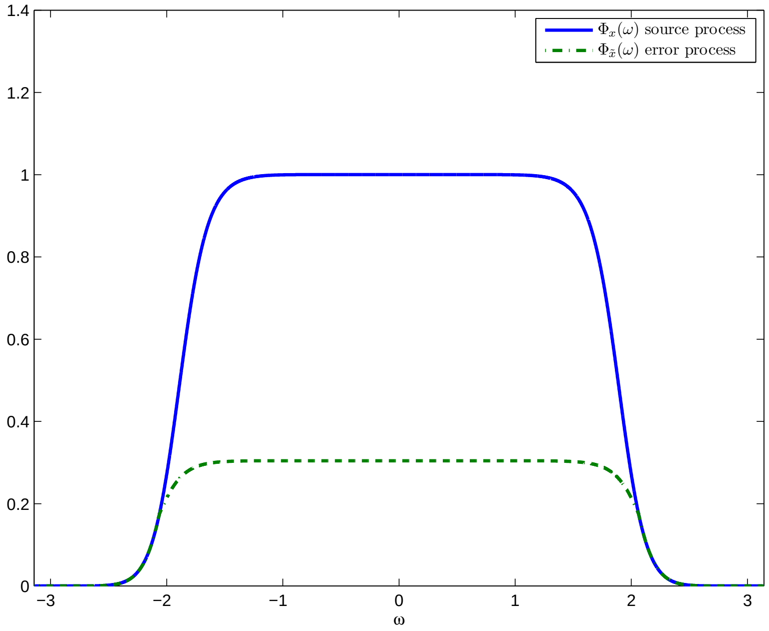


Figure 3.8: PSDs of the source and the error processes (band-limited Gaussian processes, asymmetric noise.)

For the asymmetric noise case, where the noise variance at sensor 2 is increased to 0.0135, Fig. 3.8 shows the PSDs of $\Phi_x(\omega)$ and $\Phi_{\tilde{x}}(\omega)$. As in the previous case, $\Phi_{m_2}(\omega)$ is equal to

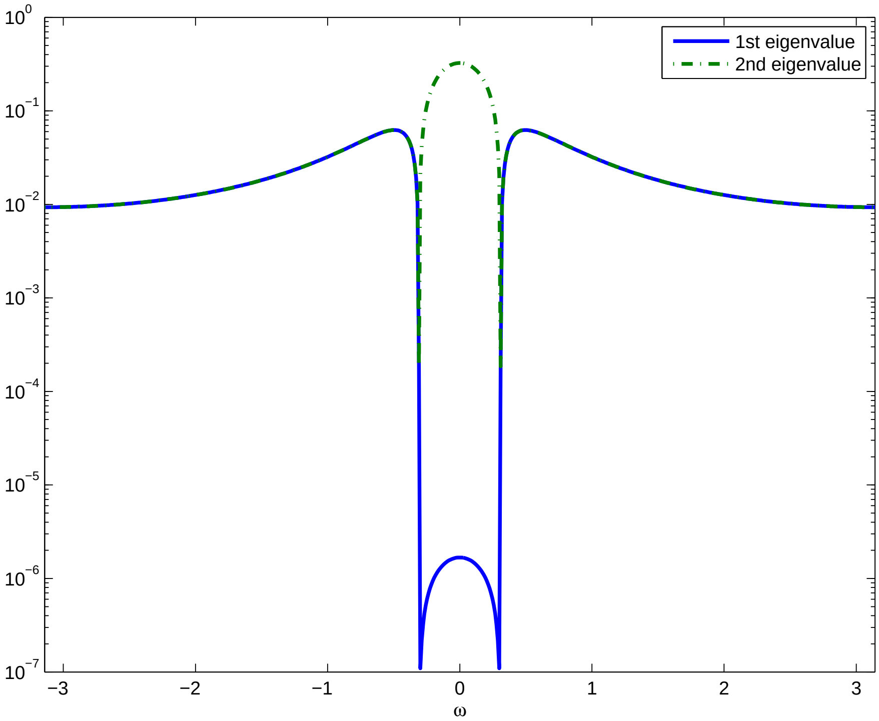


Figure 3.9: Eigenvalues of $\mathbf{Z}$ in ω.

zero and sensor node 2 is switched off.

3.6.3 Sufficiency Verification

Here we use Theorem 3.2 to numerically verify that the solution obtained through Theorem 3.1 is also sufficient for the first-order Gauss-Markov source. Only a numerical verification is feasible due to the complexity of matrix $\mathbf{Z}$ in Theorem 3.2 when (3.36) is substituted in. For the two-terminal case, (3.28) is a 2×2 matrix, and we plot its eigenvalues as a function of ω in Fig. 3.9 for the parameter settings used for Fig. 3.4. We can see that the two eigenvalues are always positive, and thus that the matrix (3.28) is always positive-definite. Based on Theorem 3.2, we know that the solution is not only necessary but also sufficient. In addition, due to the fact that there is only one solution obtained through the use of Theorem 3.1 for the first-order Gauss-Markov case, we know the solution corresponds to the true achievable Berger-Tung sum rate for the given parameters.

Chapter 4

Application of Massive Antenna Arrays in STAP

4.1 Introduction

Studies of “massive” MIMO systems in wireless communications, in which base stations are equipped with a very large number of antennas in order to dramatically increase system capacity, have recently attracted significant attention. The resulting capacity gains can often be achieved with relatively simple signal processing, due to the asymptotic orthogonality of the wireless channels. Hence massive MIMO is regarded as a promising technology for next generation wireless communication systems. Of course, the benefits of very large arrays can also be exploited in other areas as well, such as STAP for radar [110].

In this chapter, we analyze the performance of a particular reduced-dimension separable STAP algorithm, taking the effects of array size and finite secondary data support into account. We study the performance of this simple low-complexity algorithm for clairvoyant interference covariance matrices with orthogonality assumptions on the steering vectors, and show that in the asymptotic sense this scheme performs as well as the "fully adaptive" STAP method, in which the entire interference-plus-noise covariance matrix must be inverted. A scaling law for the SINR loss as a function of the number of antennas is provided. Appealing to random matrix theory, we finally provide an analysis of the SINR as a function of the number of training samples when the covariance matrix is estimated using a finite collection of secondary data. As in the case of MIMO wireless communications, the goal is to demonstrate that the availability of a massive number of antennas allows optimal performance to be obtained with relatively simple signal processing.

Fully adaptive STAP methods [64] are well known for their high computationally complexity, so one must often resort to reduced-dimension or reduced-rank algorithms [64, 36, 78, 37] in order to achieve a reasonable trade-off between performance and complexity. In the past decades, many low complexity algorithms have been proposed, based on choosing different temporal and spatial data combinations, beamspace processing and Doppler filtering [110]. Other approaches for complexity reduction include the eigencanceler [40, 38] which exploits the low-rank structure of the interference-plus-noise covariance matrix, recursive computation of the inverse covariance [72], the multistage Wiener filter [34, 47], and parametric clutter modeling using autoregressive filters [85, 75]. Algorithms that exploit knowledge of the interference-plus-noise covariance matrix usually do so by first estimating it using secondary data samples, and the impact of the size of the available secondary sample support has been studied as far back as [82], in which the probability distribution of the sample matrix inversion (SMI) approach was analyzed. More detailed results were later obtained in [84], and extensions, for example, to cases involving secondary data that contain the signal of interest [10] or to other algorithms such as the eigencanceler [39] have been considered.

Random matrix theory (RMT) [5, 105, 24] provides rich theoretical results that have been useful in the analysis of very large wireless communication systems. Recently, RMT has also been applied to the study of STAP algorithms and other sample covariance matrix problems [66]. For example, the authors of [101] studied STAP in amplitude heterogeneous clutter environments, where the power of interference in the secondary data is different from that in the range bin of detection. They used RMT to derive an asymptotic closed-form expression for the SINR loss. In [100], the asymptotic SINR loss for knowledge-aided STAP algorithms for both accurate and inaccurate *a priori* knowledge cases was analyzed using RMT.

In this chapter, we analyze the performance of a simple partially adaptive STAP algorithm with low computational complexity. In this approach, the received data are first processed by a spatial-only non-adaptive beamformer, followed by an adaptive filter in the temporal domain, and thus can be regarded as a reduced-dimension separable approach. The algorithm itself is not new, and its performance is known to be inferior to others when the number of antenna elements is relatively small. However, we demonstrate that in the massive antenna regime, this algorithm approaches optimal performance, and we provide scaling laws that illustrate its behavior as the number of antennas or the number of secondary data samples grows. More specifically, the main contributions are:

- We analyze the asymptotic performance of the reduced-dimension separable STAP algorithm. The analysis is conducted in detail for one-dimensional antenna arrays, and extensions to two-dimensional cases are presented.
- We find a scaling law for the SINR loss of the reduced-dimension algorithm as an asymptotic function of the number of antennas.
- We use RMT to find a scaling law for the SINR as a function of the number of the secondary data samples when the covariance matrix is not known and must be estimated.

The rest of the chapter is organized as follows. In Section 4.2, we describe the system model and necessary background found in previous work, and we introduce the separable STAP algorithm to be considered. In Section 4.3, we analyze the performance of the algorithm using an approximate but effective method for uniform linear arrays (ULAs). The performance is also compared to that of the fully adaptive STAP algorithm for massive antennas. A scaling law for the SINR loss compared to the fully adaptive STAP algorithm is provided as well. Later in this section, the SINR results are extended to planar antenna array configurations. Appealing to random matrix theory, Section 4.4 provides an analysis of the algorithm when the interference-plus-noise covariance matrix is estimated using secondary data. Simulation results validating our analyses are found in Section 4.5.

4.2 System Model

The problem considered is depicted in Fig. 4.1, in which an airborne platform is moving parallel to the x-axis at speed v_a and a target is located anywhere in space and moving in an arbitrary direction. Angles θ and ϕ respectively represent the elevation and azimuth angles of the target relative to the airborne platform. The platform is assumed to be equipped with an antenna array of N elements, and the transmitter on the platform is assumed to emit M pulses per coherent processing interval (CPI). For a given range gate, the received data at pulse m is denoted by $\mathbf{x}_m$. Defining $\boldsymbol{\chi} = \begin{bmatrix}\mathbf{x}_1^T & \ldots & \mathbf{x}_M^T\end{bmatrix}^T$, the STAP detection problem is defined by the two hypotheses

$$\begin{cases} H_0 : \boldsymbol{\chi} = \boldsymbol{\chi}_c + \boldsymbol{\chi}_n \\ H_1 : \boldsymbol{\chi} = \boldsymbol{\chi}_t + \boldsymbol{\chi}_c + \boldsymbol{\chi}_n \end{cases},$$

where $\boldsymbol{\chi}_t$ is the signal due to the target, $\boldsymbol{\chi}_c$ is due to clutter, and $\boldsymbol{\chi}_n$ represents other noise and interference. The terms $\boldsymbol{\chi}_c$ and $\boldsymbol{\chi}_n$ are assumed to be random vectors with covariance matrices $\mathbf{R}_c$ and $\mathbf{R}_n$, respectively.

The detection statistic for the fully adaptive STAP method is $z = \mathbf{w}^H\boldsymbol{\chi}$, where $\mathbf{w} = \mathbf{R}^{-1}\mathbf{v}_t$ is

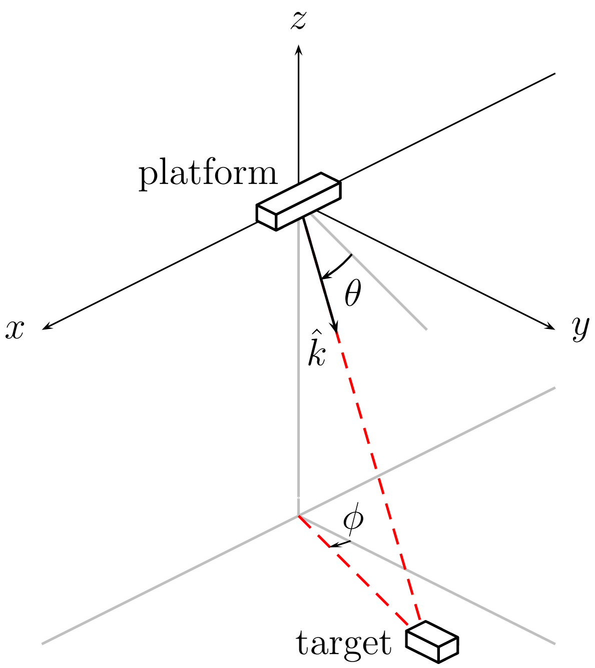


Figure 4.1: Depiction of an airborne radar scenario.

the weighting vector, $\mathbf{R} = \mathbf{R}_c + \mathbf{R}_n$, and $\mathbf{v}_t$ is the target steering vector. The post-processing SINR at the platform determines the overall detection performance. From [111], we know the general expression for the SINR is

$$SINR = \frac{\sigma^2 \xi_t \left| \mathbf{w}^H \mathbf{v}_t \right|^2}{\mathbf{w}^H \mathbf{R} \mathbf{w}}, \tag{4.1}$$

and thus the SINR of the fully adaptive STAP method is

$$SINR = \sigma^2 \xi_t \mathbf{v}_t^H \mathbf{R}^{-1} \mathbf{v}_t, \tag{4.2}$$

where σ^2 is the noise power per element, and ξ_t is the single-pulse SNR for a single antenna element at the receive side.

4.2.1 Simplified Separable Algorithm

The difficulty associated with the full implementation of the STAP algorithm comes from the fact that the dimensionality of $\mathbf{R}$ is $MN \times MN$, which makes the calculation of the matrix inverse difficult even for a system with a moderate number of antennas. In this chapter, we consider the performance of a simple separable STAP algorithm that effectively exploits the benefits of very large antenna arrays. In this approach, the received data are first processed by a spatial-only non-adaptive beamformer $\mathbf{f}$, followed by an adaptive filter in the temporal domain. The spatial-only filter leads to the transformed data vector

$$\tilde{\boldsymbol{\chi}} = (\mathbf{I} \otimes \mathbf{f})^H \boldsymbol{\chi}\,, \tag{4.3}$$

where $\mathbf{I}$ is the identity matrix and $\otimes$ denotes the Kronecker product. Letting $\mathbf{F} = \mathbf{I} \otimes \mathbf{f}$, the transformed covariance matrix is

$$\tilde{\mathbf{R}} = \mathbf{F}^H \mathbf{R} \mathbf{F}\,. \tag{4.4}$$

Similarly, the transformed target steering vector is $\tilde{\mathbf{v}}_t = \mathbf{F}^H \mathbf{v}_t$. Without loss of generality we assume $\mathbf{f}$ is a normalized vector, *i.e.*, $\|\mathbf{f}\| = 1$, and hence we have $\mathbf{F}^H\mathbf{F} = (\mathbf{I} \otimes \mathbf{f})^H (\mathbf{I} \otimes \mathbf{f}) = \mathbf{I}$. For the reduced-dimension algorithm, the detection statistic is

$$\tilde{z} = \tilde{\mathbf{w}}^H \tilde{\boldsymbol{\chi}}\,, \tag{4.5}$$

where $\tilde{\mathbf{w}} = \tilde{\mathbf{R}}^{-1}\tilde{\mathbf{v}}_t = \left(\mathbf{F}^H \mathbf{R} \mathbf{F}\right)^{-1} \mathbf{F}^H \mathbf{v}_t$. The SINR of this algorithm can be calculated from (4.1) and is found to be

$$SINR = \sigma^2 \xi_t \mathbf{v}_t^H \mathbf{F} (\mathbf{F}^H \mathbf{R} \mathbf{F})^{-1} \mathbf{F}^H \mathbf{v}_t\,. \tag{4.6}$$

In practice, matrix $\mathbf{R}$ can only be estimated using secondary data. Letting $\hat{\mathbf{R}}$ denote the estimate, the weighting vector in this case is

$$\mathbf{w} = \mathbf{F}\left(\mathbf{F}^H\hat{\mathbf{R}}\mathbf{F}\right)^{-1}\mathbf{F}^H\mathbf{v}_t\,, \tag{4.7}$$

and from (4.1) we know the corresponding SINR is given by

$$SINR = \frac{\sigma^2\xi_t\left|\mathbf{v}_t^H\mathbf{F}\left(\mathbf{F}^H\hat{\mathbf{R}}\mathbf{F}\right)^{-1}\mathbf{F}^H\mathbf{v}_t\right|^2}{\mathbf{v}_t^H\mathbf{F}\left(\mathbf{F}^H\hat{\mathbf{R}}\mathbf{F}\right)^{-1}\mathbf{F}^H\mathbf{R}\mathbf{F}\left(\mathbf{F}^H\hat{\mathbf{R}}\mathbf{F}\right)^{-1}\mathbf{F}^H\mathbf{v}_t}\,. \tag{4.8}$$

The most frequently used method to estimate $\mathbf{R}$ is SMI, in which the estimated covariance matrix is $\hat{\mathbf{R}} = \frac{1}{K}\mathbf{X}\mathbf{X}^H$, where $\mathbf{X} = [\boldsymbol{\chi}_1, \ldots, \boldsymbol{\chi}_K]$ contains training samples from K different range intervals. We will assume SMI for covariance matrix estimation throughout the chapter.

4.2.2 Asymptotic Orthogonality of Steering Vectors – ULA

Parts of the analysis in subsequent sections rely on assumptions regarding the asymptotic orthogonality of the steering vectors of objects in different locations as the number of antenna elements goes to infinity. The general steering vector is given by $\mathbf{v} = \boldsymbol{\beta}(\varpi) \otimes \boldsymbol{\alpha}(\vartheta)$, where the temporal steering vector is defined by

$$\boldsymbol{\beta}(\varpi) = \begin{bmatrix} 1 \\ e^{j2\pi\varpi} \\ \vdots \\ e^{j(M-1)2\pi\varpi} \end{bmatrix}, \tag{4.9}$$

and where, assuming for example a ULA, the spatial steering vector is

$$\boldsymbol{\alpha}(\vartheta) = \begin{bmatrix} 1 \\ e^{j2\pi\vartheta} \\ \vdots \\ e^{j(N-1)2\pi\vartheta} \end{bmatrix} . \tag{4.10}$$

In the definition of these steering vectors, $\vartheta = \frac{d}{\lambda_0} \cos\theta \sin\phi$ is the spatial frequency, $\varpi = \frac{2v}{\lambda_0 f_r}$ is the normalized Doppler frequency, λ_0 is the signal wavelength, f_r is the pulse repetition frequency, v is the radial speed, and d is the separation between antenna elements.

Let $\hat{\boldsymbol{\beta}} = \frac{1}{\sqrt{M}}\boldsymbol{\beta}$, $\hat{\boldsymbol{\alpha}} = \frac{1}{\sqrt{N}}\boldsymbol{\alpha}$ and $\hat{\mathbf{v}} = \frac{1}{\sqrt{MN}}\mathbf{v}$ be the normalized versions of $\boldsymbol{\beta}$, $\boldsymbol{\alpha}$ and $\mathbf{v}$, respectively. The spatial filtering vector is then $\mathbf{f} = \frac{1}{\sqrt{N}}\boldsymbol{\alpha}$. The inner product of an arbitrary pair of steering vectors $\hat{\mathbf{v}}_i$ and $\hat{\mathbf{v}}_j$ is given by

$$\hat{\mathbf{v}}_j^H \hat{\mathbf{v}}_i = \left(\hat{\boldsymbol{\beta}}(\varpi_j) \otimes \hat{\boldsymbol{\alpha}}(\vartheta_j)\right)^H \left(\hat{\boldsymbol{\beta}}(\varpi_i) \otimes \hat{\boldsymbol{\alpha}}(\vartheta_i)\right) = \left(\hat{\boldsymbol{\beta}}^H(\varpi_j)\hat{\boldsymbol{\beta}}(\varpi_i)\right) \otimes \left(\hat{\boldsymbol{\alpha}}^H(\vartheta_j)\hat{\boldsymbol{\alpha}}(\vartheta_i)\right) , \tag{4.11}$$

where

$$\hat{\boldsymbol{\alpha}}^H(\vartheta_j)\hat{\boldsymbol{\alpha}}(\vartheta_i) = \frac{1}{N}\sum_{n=1}^{N} e^{j(n-1)2\pi(\vartheta_i - \vartheta_j)} . \tag{4.12}$$

Clearly, the quantity $\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} e^{j(n-1)2\pi(\vartheta_i-\vartheta_j)}$ is non-zero only when $\vartheta_i = \vartheta_j$, or equivalently, when $\cos\theta_i \sin\phi_i = \cos\theta_j \sin\phi_j$. This latter equality represents the well-known cone of ambiguity that exists when a ULA is used in scenarios where the signals can have distinct elevation angles of arrival; any two signals on the cone will share an identical spatial steering vector. For a single target observed with ground clutter, the intersection of this cone with the ground and the sphere corresponding to the range bin of interest means that for a ULA, there will be at most two clutter patches that share the steering vector of the target.

With the above in mind, we have

$$
\begin{aligned}
\hat{\mathbf{v}}_i^H \mathbf{F} &= \frac{1}{\sqrt{MN}} \left(\boldsymbol{\beta}_i^H \otimes \boldsymbol{\alpha}_i^H\right) (\mathbf{I} \otimes \mathbf{f}) \\
&= \frac{1}{\sqrt{MN}} \boldsymbol{\beta}_i^H \otimes \left(\boldsymbol{\alpha}_i^H \mathbf{f}\right) \\
&= \begin{cases} \hat{\boldsymbol{\beta}}_c^H & \text{clutter patch } i = c \text{ shares the same } \vartheta \text{ with the target} \\ 0 & \text{otherwise} \end{cases} && (4.13) \\
\hat{\mathbf{v}}_c^H \hat{\mathbf{v}}_t &= \left(\hat{\boldsymbol{\beta}}_c^H \otimes \hat{\boldsymbol{\alpha}}_c^H\right) \left(\hat{\boldsymbol{\beta}}_t \otimes \hat{\boldsymbol{\alpha}}_t\right) = \hat{\boldsymbol{\beta}}_c^H \hat{\boldsymbol{\beta}}_t && (4.14) \\
\mathbf{F} \hat{\boldsymbol{\beta}}_t &= (\mathbf{I} \otimes \mathbf{f}) \hat{\boldsymbol{\beta}}_t = \hat{\mathbf{v}}_t \,, && (4.15)
\end{aligned}
$$

where the subscript t denotes the index of the target, c denote the index of a clutter patch (if any) that shares the same spatial frequency with the target. Note that the radial speed v for clutter patches is equal to $\left(\frac{2v_a}{df_r}\right)\vartheta$, and thus $\hat{\boldsymbol{\beta}}_{c_i} = \hat{\boldsymbol{\beta}}_{c_j}$ and $\hat{\mathbf{v}}_{c_i} = \hat{\mathbf{v}}_{c_j}$ if clutter patches i and j have the same spatial frequency.

4.2.3 Asymptotic Orthogonality of Steering Vectors – URA

Consider the uniform rectangular array (URA) configuration shown in Fig. 4.2. Let d_x and d_z denote the inter-element distances alongside the x and z axes. The displacement vector toward the element at (p, q) is $\mathbf{d}_{p,q} = pd_x\hat{\mathbf{x}} + qd_z\hat{\mathbf{z}}$. A unit vector $\hat{\mathbf{k}}$ in the (ϕ, θ) direction is given by $\hat{\mathbf{k}} = \cos\theta\sin\phi\hat{\mathbf{x}} + \cos\theta\cos\phi\hat{\mathbf{y}} + \sin\theta\hat{\mathbf{z}}$, so the spatial frequency in the URA case is

$$
\vartheta = \frac{\hat{\mathbf{k}}(\phi,\theta)\cdot\mathbf{d}_{p,q}}{\lambda_0} = \frac{pd_x\cos\theta\sin\phi + qd_z\sin\theta}{\lambda_0}\,.
$$

The two-dimensional spatial frequency can also be written as $\vartheta = p\vartheta_x + q\vartheta_z$, where $\vartheta_x = \frac{d_x\cos\theta\sin\phi}{\lambda_0}$ and $\vartheta_z = \frac{d_z\sin\theta}{\lambda_0}$.

The temporal steering vector is defined as in (4.9), and the azimuth and elevation steering

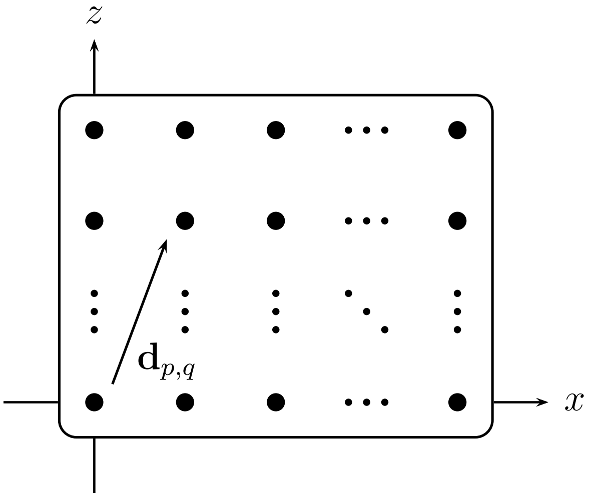


Figure 4.2: Illustration of a uniform rectangular array.

vectors are correspondingly

$$\mathbf{a}(\vartheta_x) = \begin{bmatrix} 1 \\ e^{j2\pi\vartheta_x} \\ \vdots \\ e^{j(P-1)2\pi\vartheta_x} \end{bmatrix} \quad , \quad \mathbf{e}(\vartheta_z) = \begin{bmatrix} 1 \\ e^{j2\pi\vartheta_z} \\ \vdots \\ e^{j(Q-1)2\pi\vartheta_z} \end{bmatrix} .$$

The overall steering vector is $\mathbf{v}(\varpi, \vartheta_x, \vartheta_z) = \boldsymbol{\beta}(\varpi) \otimes \boldsymbol{\alpha}(\vartheta)$, where $\boldsymbol{\alpha} = \mathbf{a}(\vartheta_x) \otimes \mathbf{e}(\vartheta_z)$ is the full spatial steering vector. The target signal is $\boldsymbol{\chi}_t = \alpha_t \boldsymbol{\beta}(\varpi_t) \otimes \mathbf{a}(\vartheta_{xt}) \otimes \mathbf{e}(\vartheta_{zt})$, and the normalized spatial filtering vector now is $\mathbf{f} = \frac{1}{\sqrt{PQ}}\boldsymbol{\alpha}$.

The orthogonality of steering vectors for a URA is analogous to that for a ULA, with certain differences. For a URA, the inner product of the steering vectors is

$$\begin{aligned} \hat{\mathbf{v}}_j^H \hat{\mathbf{v}}_i &= \left(\hat{\boldsymbol{\beta}}(\varpi_j) \otimes \hat{\mathbf{a}}(\vartheta_{xj}) \otimes \hat{\mathbf{e}}(\vartheta_{zj})\right)^H \left(\hat{\boldsymbol{\beta}}(\varpi_i) \otimes \hat{\mathbf{a}}(\vartheta_{xi}) \otimes \hat{\mathbf{e}}(\vartheta_{zi})\right) \\ &= \left(\hat{\boldsymbol{\beta}}^H(\varpi_j)\hat{\boldsymbol{\beta}}(\varpi_i)\right) \otimes \left(\hat{\mathbf{a}}^H(\vartheta_{xj})\hat{\mathbf{a}}(\vartheta_{xi})\right) \otimes \left(\hat{\mathbf{e}}^H(\vartheta_{zj})\hat{\mathbf{e}}(\vartheta_{zi})\right) , \end{aligned}$$

where $\hat{\mathbf{a}} = \frac{1}{\sqrt{P}}\mathbf{a}$ and $\hat{\mathbf{e}} = \frac{1}{\sqrt{Q}}\mathbf{e}$ are the corresponding normalized steering vectors. Then we

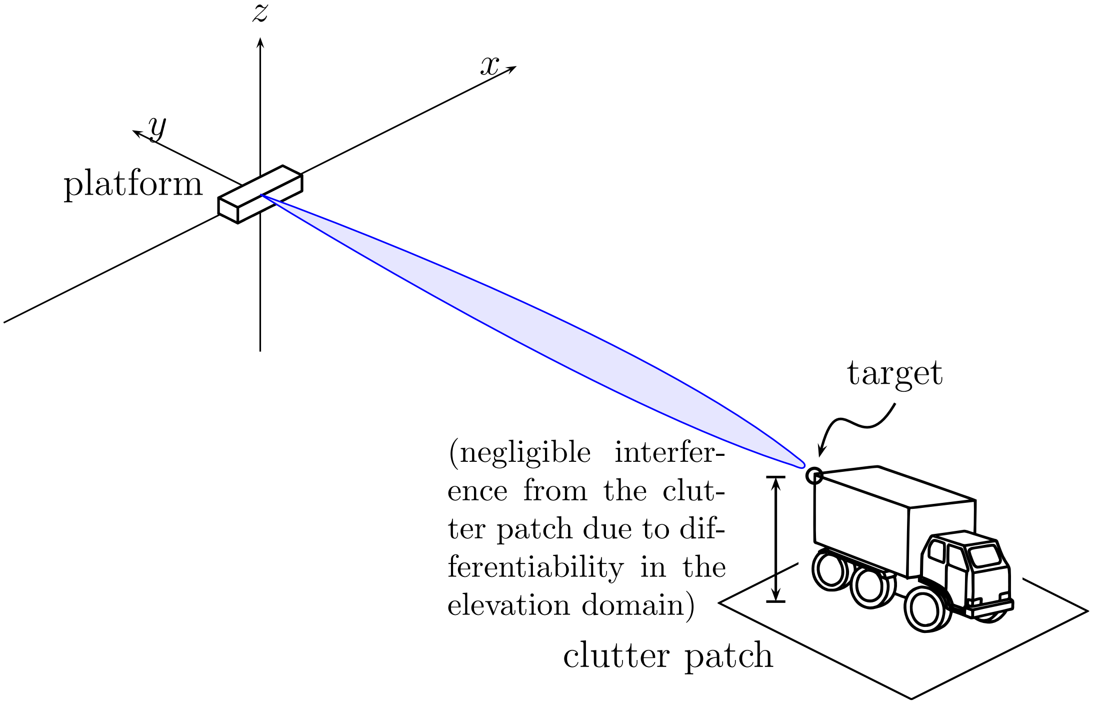

z
x
y
platform
target
(negligible interference from the clutter patch due to differentiability in the elevation domain)
clutter patch

Figure 4.3: Target is moderately above the ground clutter patch.

have

$$\hat{\mathbf{a}}^H(\vartheta_{xj})\hat{\mathbf{a}}(\vartheta_{xi}) = \frac{1}{P}\sum_{n=1}^{P} e^{j(n-1)2\pi(\vartheta_{xi}-\vartheta_{xj})}$$

$$\hat{\mathbf{e}}^H(\vartheta_{zj})\hat{\mathbf{e}}(\vartheta_{zi}) = \frac{1}{Q}\sum_{n=1}^{Q} e^{j(n-1)2\pi(\vartheta_{zi}-\vartheta_{zj})}\,.$$

Thus, we obtain orthogonality for URAs if either P or Q grows to infinity and not the other. Unlike a ULA, planar arrays provide differentiability in the elevation domain. As Q goes to infinity, the main receive beam becomes increasingly narrow in elevation, and in the limit the interference from the clutter patches becomes negligible if the target is just moderately above the ground, as illustrated for example in Fig. 4.3.

4.3 SINR Analysis

In this section, we find a closed-form expression of the asymptotic SINR for the case of a ULA, which can be used to quickly calculate the required number of antenna elements for a given performance target in STAP system designs. To more thoroughly characterize the performance of the separable STAP scheme, a scaling law for the SINR loss compared with the fully adaptive STAP is derived as a function of the number of antennas. Then the asymptotic SINR result is extended to URAs to reflect the differentiability in the elevation domain.

4.3.1 SINR Performance Analysis for ULAs

In general, the signal reflected from the target is contaminated by echoes from clutter. Echoes from clutter patches that share the spatial steering vector of the target are not distinguishable from the target using spatial processing alone, regardless of the number of antennas, as illustrated in Fig. 4.4. In such cases, one must rely on the difference in Doppler between the target and interfering clutter patch. Here we analyze the SINR performance of the reduced-dimension separable STAP algorithm for these cases under the orthogonality assumptions discussed above.

For simplicity, we assume there is no range ambiguity. Thus, for N_c clutter patches, the clutter interference covariance matrix is

$$\mathbf{R}_c = \sigma^2 \sum_{i=1}^{N_c} \xi_i \mathbf{v}_i \mathbf{v}_i^H = \sigma^2 MN \sum_{i=1}^{N_c} \xi_i \hat{\mathbf{v}}_i \hat{\mathbf{v}}_i^H \,, \tag{4.16}$$

where $\mathbf{v}_i$ is the steering vector toward the ith clutter patch, $\hat{\mathbf{v}}_i$ is the corresponding normalized steering vector, and ξ_i the strength of the ith clutter return. We assume a noise covariance matrix of $\mathbf{R}_n = \sigma^2\mathbf{I}$, and we want to find an analytic expression for (4.6) when

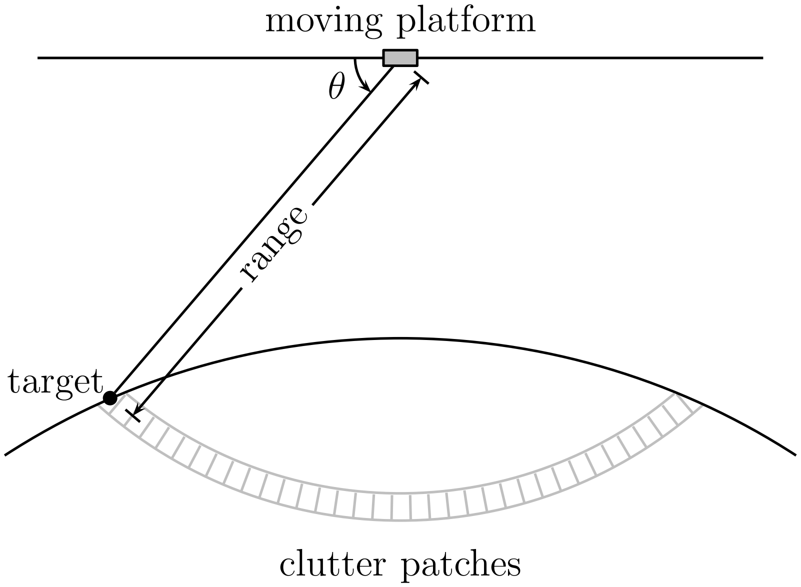


Figure 4.4: Target sharing the same spatial frequency with clutter patches.

$N \to \infty$. Note that $\mathit{span}\{\hat{\mathbf{v}}_1, \ldots, \hat{\mathbf{v}}_{N_c}\}$ forms an N_c-dimensional subspace of $\mathbb{C}^{MN}$, where $N_c \ll MN$. We can find other orthonormal vectors $\hat{\mathbf{u}}_i, i = 1, \ldots, MN - N_c$ in $\mathbb{C}^{MN}$ which together with $\hat{\mathbf{v}}_i, i = 1, \ldots, N_c$ form an orthonormal basis for $\mathbb{C}^{MN}$. With the help of these vectors, the eigendecomposition of the sum covariance matrix $\mathbf{R} = \mathbf{R}_c + \mathbf{R}_n$ can be written as

$$\mathbf{R} = \sum_{i=1}^{N_c} \sigma^2 \left(MN\xi_i + 1\right) \hat{\mathbf{v}}_i \hat{\mathbf{v}}_i^H + \sigma^2 \sum_{i=1}^{MN-N_c} \hat{\mathbf{u}}_i \hat{\mathbf{u}}_i^H \, . \tag{4.17}$$

Hence the square root of $\mathbf{R}$ is

$$\begin{aligned}\mathbf{R}^{\frac{1}{2}} &= \sum_{i=1}^{N_c} \left(\sigma^2 \left(MN\xi_i + 1\right)\right)^{\frac{1}{2}} \hat{\mathbf{v}}_i \hat{\mathbf{v}}_i^H + \left(\sigma^2\right)^{\frac{1}{2}} \sum_{i=1}^{MN-N_c} \hat{\mathbf{u}}_i \hat{\mathbf{u}}_i^H \\ &= \sum_{i=1}^{N_c} \left(\left(\sigma^2 \left(MN\xi_i + 1\right)\right)^{\frac{1}{2}} - \sigma\right) \hat{\mathbf{v}}_i \hat{\mathbf{v}}_i^H + \sigma \mathbf{I} \, , \end{aligned} \tag{4.18}$$

and clearly

$$\mathbf{R}^{-\frac{1}{2}} = \sum_{i=1}^{N_c} \left(\left(\sigma^2 \left(MN\xi_i + 1\right)\right)^{-\frac{1}{2}} - \frac{1}{\sigma} \right) \hat{\mathbf{v}}_i \hat{\mathbf{v}}_i^H + \frac{1}{\sigma}\mathbf{I}\,. \tag{4.19}$$

Using (4.13) and the fact that $\hat{\mathbf{v}}_i, i = 1, \ldots, N_c$ and $\hat{\mathbf{u}}_i, i = 1, \ldots, MN - N_c$ form an orthonormal basis, it is straightforward to see

$$\mathbf{R}^{\frac{1}{2}}\mathbf{F} = \left(\sum_{i=1}^{n} \left(\sigma^2 \left(MN\xi_{c_i} + 1\right)\right)^{\frac{1}{2}} - n\sigma \right) \hat{\mathbf{v}}_c \hat{\boldsymbol{\beta}}_c^H + \sigma\mathbf{F}\,. \tag{4.20}$$

where $c_i, i = 1, \ldots, n$ are the indices of the clutter patches sharing the same spatial frequency, n is the total number of these patches, and $\hat{\mathbf{v}}_c$ and $\hat{\boldsymbol{\beta}}_c$ are the corresponding full and temporal steering vectors. As indicated in Section 4.1, the clutter patches sharing the same spatial frequency have the same full and temporal steering vectors, so $\hat{\mathbf{v}}_{c_i} = \hat{\mathbf{v}}_c$ and $\hat{\boldsymbol{\beta}}_{c_i} = \hat{\boldsymbol{\beta}}_c$ hold for all i and we can use $\hat{\mathbf{v}}_c$ and $\hat{\boldsymbol{\beta}}_c$ to represent the vectors for all patches.

Similarly, applying (4.14) we have

$$\mathbf{R}^{-\frac{1}{2}}\hat{\mathbf{v}}_t = \left(\sum_{i=1}^{n} \left(\sigma^2 \left(MN\xi_{c_i} + 1\right)\right)^{-\frac{1}{2}} - \frac{n}{\sigma} \right) \hat{\mathbf{v}}_c \hat{\boldsymbol{\beta}}_c^H \hat{\boldsymbol{\beta}}_t + \frac{1}{\sigma}\hat{\mathbf{v}}_t\,. \tag{4.21}$$

For the term $(\mathbf{F}^H\mathbf{R}\mathbf{F})^{-1}$, we have

$$\begin{aligned} \mathbf{F}^H\mathbf{R}\mathbf{F} &= \sigma^2 MN \sum_{i=1}^{N_c} \xi_i \mathbf{F}^H \hat{\mathbf{v}}_i \hat{\mathbf{v}}_i^H \mathbf{F} + \sigma^2 \mathbf{F}^H\mathbf{F} \\ &= \sigma^2 MN \hat{\boldsymbol{\beta}}_c \hat{\boldsymbol{\beta}}_c^H \sum_{i=1}^{n} \xi_{c_i} + \sigma^2\mathbf{I}\,, \end{aligned} \tag{4.22}$$

where we have applied (4.13). Using the Sherman-Morrison formula [90], we obtain

$$\left(\mathbf{F}^H\mathbf{R}\mathbf{F}\right)^{-1} = \frac{1}{\sigma^2} \left(\mathbf{I} - \frac{MN\sum_{i=1}^{n} \xi_{c_i}}{1 + MN\sum_{i=1}^{n} \xi_{c_i}} \hat{\boldsymbol{\beta}}_c \hat{\boldsymbol{\beta}}_c^H \right)\,. \tag{4.23}$$

Putting everything together, we finally have an analytic expression for the asymptotic SINR:

$$\begin{aligned} SINR =& \sigma^2 \xi_t \mathbf{v}_t^H \mathbf{F} (\mathbf{F}^H \mathbf{R} \mathbf{F})^{-1} \mathbf{F}^H \mathbf{v}_t \\ =& \sigma^2 \xi_t \mathbf{v}_t^H \mathbf{R}^{-\frac{1}{2}} \mathbf{R}^{\frac{1}{2}} \mathbf{F} (\mathbf{F}^H \mathbf{R} \mathbf{F})^{-1} \mathbf{F}^H \mathbf{R}^{\frac{1}{2}} \mathbf{R}^{-\frac{1}{2}} \mathbf{v}_t \\ =& MN\xi_t \left(1 - \frac{MN \sum_{i=1}^{n} \xi_{c_i}}{1 + MN \sum_{i=1}^{n} \xi_{c_i}} \left| \hat{\boldsymbol{\beta}}_t^H \hat{\boldsymbol{\beta}}_c \right|^2 \right) . \end{aligned} \tag{4.24}$$

Again using (4.21), the SINR of the fully adaptive STAP algorithm can be calculated as

$$\begin{aligned} SINR &= \sigma^2 \xi_t \mathbf{v}_t^H \mathbf{R}^{-1} \mathbf{v}_t \\ &= \sigma^2 \xi_t \mathbf{v}_t^H \mathbf{R}^{-\frac{1}{2}} \mathbf{R}^{-\frac{1}{2}} \mathbf{v}_t \\ &= MN\xi_t \left(1 - \frac{MN \sum_{i=1}^{n} \xi_{c_i}}{1 + MN \sum_{i=1}^{n} \xi_{c_i}} \left| \hat{\boldsymbol{\beta}}_t^H \hat{\boldsymbol{\beta}}_c \right|^2 \right) , \end{aligned} \tag{4.25}$$

and we see that the two expressions in (4.24) and (4.25) are thus asymptotically identical, despite the significant computational savings afforded by the separable algorithm.

4.3.2 SINR Loss Scaling Law

SINR loss is defined as the difference between the SINR of the fully adaptive STAP method in (4.2) and the separable algorithm in (4.6). In this section we derive a scaling law to show how the SINR loss is reduced as the number of antennas N is increased. In particular, we will show that the SINR loss of the separable algorithm approaches zero as $O(N^{-2})$. We begin with further study of (4.12). For an arbitrary N, equation (4.12) can be written as

$$\hat{\boldsymbol{\alpha}}^H(\vartheta_j) \hat{\boldsymbol{\alpha}}(\vartheta_i) = \frac{1}{N} \sum_{n=1}^{N} e^{j(n-1)2\pi(\vartheta_i - \vartheta_j)} = \frac{1}{N} e^{j\pi(N-1)(\vartheta_i - \vartheta_j)} \frac{\sin\left(\pi N \left(\vartheta_i - \vartheta_j\right)\right)}{\sin\left(\pi \left(\vartheta_i - \vartheta_j\right)\right)} . \tag{4.26}$$

Define $\Theta_i = \vartheta_t - \vartheta_i$ and $f(\Theta_i) = \hat{\boldsymbol{\alpha}}^H(\vartheta_i)\hat{\boldsymbol{\alpha}}(\vartheta_t)$. Then

$$f(0) = \lim_{\vartheta_i \to \vartheta_t} \frac{1}{N} e^{j\pi(N-1)(\vartheta_t - \vartheta_i)} \frac{\sin\left(\pi N\left(\vartheta_t - \vartheta_i\right)\right)}{\sin\left(\pi\left(\vartheta_t - \vartheta_i\right)\right)} = 1\,, \tag{4.27}$$

and hence $f(\Theta_{c_i}) = f(\Theta_t) = f(0) = 1$. Expressed in terms of of Θ_i, equation (4.13) is

$$\hat{\mathbf{v}}_i^H \mathbf{F} = \frac{1}{\sqrt{MN}} \left(\boldsymbol{\beta}_i^H \otimes \boldsymbol{\alpha}_i^H\right)\left(\mathbf{I} \otimes \mathbf{f}\right) = f(\Theta_i)\hat{\boldsymbol{\beta}}_i^H\,. \tag{4.28}$$

From (4.28) we see that

$$\mathbf{F}^H\mathbf{R}\mathbf{F} = \sigma^2 MN \sum_{i=1}^{N_c} \xi_i \mathbf{F}^H \hat{\mathbf{v}}_i \hat{\mathbf{v}}_i^H \mathbf{F} + \sigma^2\mathbf{I} = \sigma^2 MN \sum_{i=1}^{N_c} \xi_i |f(\Theta_i)|^2 \hat{\boldsymbol{\beta}}_i \hat{\boldsymbol{\beta}}_i^H + \sigma^2\mathbf{I}\,. \tag{4.29}$$

In order to find the inverse of $\mathbf{F}^H\mathbf{R}\mathbf{F}$, we need to define several auxiliary matrices:

$$\begin{cases} \mathbf{A} & = MN\left(\sum_{i=1}^n \xi_{c_i}|f(\Theta_{c_i})|^2\right)\hat{\boldsymbol{\beta}}_c\hat{\boldsymbol{\beta}}_c^H + \mathbf{I} = MN\hat{\boldsymbol{\beta}}_c\hat{\boldsymbol{\beta}}_c^H \sum_{i=1}^n \xi_{c_i} + \mathbf{I} \\ \mathbf{B} & = \left[\hat{\boldsymbol{\beta}}_1,\ \dots\ ,\ \hat{\boldsymbol{\beta}}_{c_i-1},\ \hat{\boldsymbol{\beta}}_{c_i+1},\ \dots\ ,\ \hat{\boldsymbol{\beta}}_{c_j-1},\ \hat{\boldsymbol{\beta}}_{c_j+1},\ \dots\ ,\ \hat{\boldsymbol{\beta}}_{N_c}\right] \\ \mathbf{C} & = MNdiag(\xi_1|f(\Theta_1)|^2,\ \dots\ , \xi_{c_i-1}|f(\Theta_{c_i-1})|^2, \xi_{c_i+1}|f(\Theta_{c_i+1})|^2,\ \dots\ , \\ & \qquad \xi_{c_j-1}|f(\Theta_{c_j-1})|^2, \xi_{c_j+1}|f(\Theta_{c_j+1})|^2,\ \dots\ , \xi_{N_c}|f(\Theta_{N_c})|^2) \end{cases}.$$

Applying the matrix inversion lemma [90], we obtain

$$\begin{aligned} \left(\mathbf{F}^H\mathbf{R}\mathbf{F}\right)^{-1} &= \left(\sigma^2 MN \sum_{i=1}^{N_c} \xi_i |f(\Theta_i)|^2 \hat{\boldsymbol{\beta}}_i \hat{\boldsymbol{\beta}}_i^H + \sigma^2\mathbf{I}\right)^{-1} \\ &= \frac{1}{\sigma^2}\left(\mathbf{B}\mathbf{C}\mathbf{B}^H + \mathbf{A}\right)^{-1} \\ &= \frac{1}{\sigma^2}\left(\mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}\left(\mathbf{C}^{-1} + \mathbf{B}^H\mathbf{A}^{-1}\mathbf{B}\right)^{-1}\mathbf{B}^H\mathbf{A}^{-1}\right)\,. \end{aligned} \tag{4.30}$$

The SINR of (4.6) can be written as

$$\begin{aligned}&\sigma^2 \xi_t MN \hat{\mathbf{v}}_t^H \mathbf{F}(\mathbf{F}^H \mathbf{R}\mathbf{F})^{-1}\mathbf{F}^H \hat{\mathbf{v}}_t \\ &\quad = \xi_t MN \hat{\boldsymbol{\beta}}_t^H \left(\mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}\left(\mathbf{C}^{-1} + \mathbf{B}^H \mathbf{A}^{-1}\mathbf{B}\right)^{-1}\mathbf{B}^H \mathbf{A}^{-1}\right)\hat{\boldsymbol{\beta}}_t \,. \end{aligned} \tag{4.31}$$

Similarly, the SINR of the fully adaptive STAP algorithm in (4.2) can be written as

$$\sigma^2 \xi_t \mathbf{v}_t^H \mathbf{R}^{-1} \mathbf{v}_t = \xi_t MN \hat{\boldsymbol{\beta}}_t^H \mathbf{A}^{-1} \hat{\boldsymbol{\beta}}_t \,. \tag{4.32}$$

Calculating the difference of the two SINRs provides the following expression for the SINR loss:

$$\xi_t MN \hat{\boldsymbol{\beta}}_t^H \left(\mathbf{A}^{-1}\mathbf{B}\left(\mathbf{C}^{-1} + \mathbf{B}^H \mathbf{A}^{-1}\mathbf{B}\right)^{-1}\mathbf{B}^H \mathbf{A}^{-1}\right)\hat{\boldsymbol{\beta}}_t \,,$$

and we can proceed with the derivation of the scaling law.

Based on the definition in [110], we know ξ_i is $\mathcal{O}\left(N^{-2}\right)$ if the same antenna array is used for both transmission and reception. Furthermore $f(\Theta_i)$ is clearly $\mathcal{O}\left(N^{-1}\right)$, so it can be observed that

$$MN\xi_i |f(\Theta_i)|^2 = \mathcal{O}\left(N^{-3}\right) \,, \tag{4.33}$$

which means that the diagonal elements of $\mathbf{C}$ are all $\mathcal{O}\left(N^{-3}\right)$. Applying the Sherman-Morrison formula again yields

$$\mathbf{A}^{-1} = \mathbf{I} - \frac{MN\sum_{i=1}^{n} \xi_{c_i}}{1 + MN\sum_{i=1}^{n}\xi_{c_i}} \hat{\boldsymbol{\beta}}_c \hat{\boldsymbol{\beta}}_c^H \,. \tag{4.34}$$

We substitute these results into the SINR loss expression and take the limit:

$$\begin{aligned}
&\lim_{N\to\infty} \frac{1}{N^{-2}} \xi_t M N \hat{\boldsymbol{\beta}}_t^H \left(\mathbf{A}^{-1}\mathbf{B} \left(\mathbf{C}^{-1} + \mathbf{B}^H \mathbf{A}^{-1} \mathbf{B} \right)^{-1} \mathbf{B}^H \mathbf{A}^{-1} \right) \hat{\boldsymbol{\beta}}_t \\
&= \lim_{N\to\infty} \xi_t M \hat{\boldsymbol{\beta}}_t^H \left(\mathbf{A}^{-1}\mathbf{B} \left(\frac{1}{N^3}\mathbf{C}^{-1} + \frac{1}{N^3}\mathbf{B}^H \mathbf{A}^{-1} \mathbf{B} \right)^{-1} \mathbf{B}^H \mathbf{A}^{-1} \right) \hat{\boldsymbol{\beta}}_t \\
&\overset{(a)}{=} \lim_{N\to\infty} \xi_t M \hat{\boldsymbol{\beta}}_t^H \left(\mathbf{A}^{-1}\mathbf{B} \left(\frac{1}{N^3}\mathbf{C}^{-1} \right)^{-1} \mathbf{B}^H \mathbf{A}^{-1} \right) \hat{\boldsymbol{\beta}}_t \\
&= \lim_{N\to\infty} \xi_t M \hat{\boldsymbol{\beta}}_t^H \left(\mathbf{A}^{-1}\mathbf{B} \left(N^3 \mathbf{C} \right) \mathbf{B}^H \mathbf{A}^{-1} \right) \hat{\boldsymbol{\beta}}_t \\
&\overset{(b)}{=} \kappa \,,
\end{aligned} \tag{4.35}$$

where (a) follows from (4.34), (b) follows from (4.33), and κ is a constant. This means that the SINR loss follows a scaling law of $\mathcal{O}\left(N^{-2}\right)$ with respect to the number of antennas N, which will be verified by simulations.

4.3.3 SINR Performance Analysis for URAs

The SINR analysis above is applicable to URAs with trivial modifications, and the SINR in (4.24) is also valid for URAs, where $N = PQ$. Furthermore, as mentioned in Section 4.2, URAs can provide differentiability in the elevation domain when Q goes to infinity, and in this case the received data are clutter-free. Calculation of the SINR for (4.6) in the clutter-free case is essentially the same as the calculation for ULAs. The necessary quantities are computed as:

$$\begin{cases}
\mathbf{R}^{\frac{1}{2}}\mathbf{F} = \sigma \mathbf{F} \\
\mathbf{R}^{-\frac{1}{2}}\hat{\mathbf{v}}_t = \left(\sigma^2\right)^{-\frac{1}{2}} \hat{\mathbf{v}}_t \\
\left(\mathbf{F}^H \mathbf{R} \mathbf{F}\right)^{-1} = \left(\sigma^2 \mathbf{F}^H \mathbf{F}\right)^{-1} = \frac{1}{\sigma^2}\mathbf{I}
\end{cases} .$$

Thus the SINR of the separable scheme can be written as

$$\mathbf{v}_t^H \mathbf{R}^{-\frac{1}{2}} \mathbf{R}^{\frac{1}{2}} \mathbf{F} (\mathbf{F}^H \mathbf{R} \mathbf{F})^{-1} \mathbf{F}^H \mathbf{R}^{\frac{1}{2}} \mathbf{R}^{-\frac{1}{2}} \mathbf{v}_t = \xi_t MN \,, \tag{4.36}$$

which is identical to the SINR of fully adaptive STAP:

$$\mathbf{v}_t^H \mathbf{R}^{-1} \mathbf{v}_t = \mathbf{v}_t^H \mathbf{R}^{-\frac{1}{2}} \mathbf{R}^{-\frac{1}{2}} \mathbf{v}_t = \xi_t MN \,. \tag{4.37}$$

4.4 Secondary Sample Size Analysis

In the previous sections we assume matrix $\mathbf{R}$ is known *a priori*, but in practice it must be estimated, for example through the use of secondary data. The usually adopted method is SMI, which produces an unbiased estimate of the true covariance matrix. For this method, the number of secondary data samples used to form the estimate has a significant impact on the performance of the system. In this section we study the scaling law of the SINR for the reduced-dimension separable STAP algorithm as a function of the number of secondary samples assuming it is sufficiently large. With the help of this scaling law, designers can pick a proper value to meet a certain predefined SINR target.

The calculation of the scaling law depends on the following theorem.

Theorem 4.1. *Let $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \quad \dots \quad \mathbf{x}_K]$, where $\mathbf{x}_k \in \mathbb{C}^{L\times 1}$ is a zero-mean complex Gaussian random vector with distribution $\mathcal{N}(\mathbf{0}, \mathbf{\Sigma})$. Random vectors $\mathbf{x}_i$ and $\mathbf{x}_j$ are independent when $i \neq j$. Define a random matrix $\mathbf{A} = \frac{1}{K}\mathbf{U}^H\mathbf{X}\mathbf{X}^H\mathbf{U}$, where $\mathbf{U} \in \mathbb{C}^{L\times M}$, M is a fixed positive integer, and the empirical distribution function of the eigenvalues of $\mathbf{U}\mathbf{U}^H$ converges. Let $\mathbf{P} = \mathbf{U}^H\mathbf{\Sigma}\mathbf{U}$ and $\lambda_i, i = 1, \dots, M$ be the eigenvalues of $\mathbf{P}$. As $L, K \to \infty$, we have* *almost surely*

$$\left|\mathbf{a}^H \left(\mathbf{A} - z\mathbf{I}\right)^{-1} \mathbf{a} - \mathbf{a}^H \left(x(z)\mathbf{P} - z\mathbf{I}\right)^{-1} \mathbf{a}\right| \longrightarrow 0\,, \tag{4.38}$$

where $\mathbf{a} \in \mathbb{C}^{M\times 1}$ *is an arbitrary fixed vector,* $x(z) = 1 - \frac{M}{K} - \frac{M}{K} z m(z)$*, and* $m(z)$ *is the unique solution to*

$$m(z) = \frac{1}{M}\left\{\sum_{n=1}^{M} \frac{1}{\lambda_n \left(1 - \frac{M}{K} - \frac{M}{K} z m(z)\right) - z}\right\}\,. \tag{4.39}$$

Proof: This theorem is similar to one in [65]. The key difference is that in our case the dimensionality of $\mathbf{A}$ is always limited, thus the theorem in [65] is not directly applicable and a proof is required, which is provided in the appendix.

With the help of this theorem, the scaling law can then be calculated. Let K be the number of secondary data samples, $L = MN$, $\mathbf{U} = \mathbf{F}$, and $\mathbf{a} = \mathbf{F}^H \mathbf{v}_t$. Applying Theorem 4.1 to (4.8), we have that when K is sufficiently large, the SINR of the separable scheme is approximately equal to

$$\frac{\sigma^2 \xi_t}{1 - x'(0)} \mathbf{v}_t^H \mathbf{F} \left(\mathbf{F}^H \mathbf{R} \mathbf{F}\right)^{-1} \mathbf{F}^H \mathbf{v}_t\,, \tag{4.40}$$

where $x'(0) = -\frac{1}{K}\left\{\sum_{m=1}^{M} \frac{1}{\lambda_m \left(1 - \frac{M}{K}\right)}\right\}$, and λ_m are the eigenvalues of $\mathbf{F}^H\mathbf{R}\mathbf{F}$. Equation (4.40) is the main result of this section, and its derivation is outlined below.

We need to separately evaluate the numerator and denominator of (4.8). For the numerator, define

$$\begin{cases} g(z) = \mathbf{v}_t^H \mathbf{F} \left(\mathbf{F}^H \hat{\mathbf{R}} \mathbf{F} - z\mathbf{I} \right)^{-1} \mathbf{F}^H \mathbf{v}_t \\ \gamma(z) = \mathbf{v}_t^H \mathbf{F} \left(x(z) \mathbf{F}^H \mathbf{R} \mathbf{F} - z\mathbf{I} \right)^{-1} \mathbf{F}^H \mathbf{v}_t \end{cases} .$$

Then the numerator of (4.8) is equal to $\sigma^2 \xi_t g^2(z)|_{z=0}$. Based on Theorem 4.1 we know $|g(z) - \gamma(z)| \to 0$, so asymptotically the numerator approaches $\sigma^2 \xi_t \gamma^2(z)|_{z=0}$, which is nonrandom and equal to

$$\frac{\sigma^2 \xi_t}{x^2(0)} \left(\mathbf{v}_t^H \mathbf{F} \left(\mathbf{F}^H \mathbf{R} \mathbf{F} \right)^{-1} \mathbf{F}^H \mathbf{v}_t \right)^2 . \tag{4.41}$$

For the denominator, define

$$\begin{cases} h(z) = \mathbf{v}_t^H \mathbf{F} \left(\mathbf{F}^H \mathbf{R} \mathbf{F} \right)^{-\frac{1}{2}} \left(\left(\mathbf{F}^H \mathbf{R} \mathbf{F} \right)^{-\frac{1}{2}} \mathbf{F}^H \hat{\mathbf{R}} \mathbf{F} \left(\mathbf{F}^H \mathbf{R} \mathbf{F} \right)^{-\frac{1}{2}} - z\mathbf{I} \right)^{-1} \left(\mathbf{F}^H \mathbf{R} \mathbf{F} \right)^{-\frac{1}{2}} \mathbf{F}^H \mathbf{v}_t \\ \eta(z) = \mathbf{v}_t^H \mathbf{F} \left(\mathbf{F}^H \mathbf{R} \mathbf{F} \right)^{-\frac{1}{2}} \left(x(z) \left(\mathbf{F}^H \mathbf{R} \mathbf{F} \right)^{-\frac{1}{2}} \mathbf{F}^H \mathbf{R} \mathbf{F} \left(\mathbf{F}^H \mathbf{R} \mathbf{F} \right)^{-\frac{1}{2}} - z\mathbf{I} \right)^{-1} \left(\mathbf{F}^H \mathbf{R} \mathbf{F} \right)^{-\frac{1}{2}} \mathbf{F}^H \mathbf{v}_t \end{cases} .$$

Since

$$\begin{aligned} & \left(\mathbf{F}^H \hat{\mathbf{R}} \mathbf{F} \right)^{-1} \mathbf{F}^H \mathbf{R} \mathbf{F} \left(\mathbf{F}^H \hat{\mathbf{R}} \mathbf{F} \right)^{-1} \\ & = \left(\mathbf{F}^H \mathbf{R} \mathbf{F} \right)^{-\frac{1}{2}} \left(\left(\mathbf{F}^H \mathbf{R} \mathbf{F} \right)^{-\frac{1}{2}} \mathbf{F}^H \hat{\mathbf{R}} \mathbf{F} \left(\mathbf{F}^H \mathbf{R} \mathbf{F} \right)^{-\frac{1}{2}} \right)^{-2} \left(\mathbf{F}^H \mathbf{R} \mathbf{F} \right)^{-\frac{1}{2}} , \end{aligned}$$

the denominator can be rewritten as

$$\begin{aligned} & \mathbf{v}_t^H \mathbf{F} \left(\mathbf{F}^H \hat{\mathbf{R}} \mathbf{F} \right)^{-1} \mathbf{F}^H \mathbf{R} \mathbf{F} \left(\mathbf{F}^H \hat{\mathbf{R}} \mathbf{F} \right)^{-1} \mathbf{F}^H \mathbf{v}_t \\ & = \frac{d}{dz} \mathbf{v}_t^H \mathbf{F} \left(\mathbf{F}^H \mathbf{R} \mathbf{F} \right)^{-\frac{1}{2}} \left(\left(\mathbf{F}^H \mathbf{R} \mathbf{F} \right)^{-\frac{1}{2}} \mathbf{F}^H \hat{\mathbf{R}} \mathbf{F} \left(\mathbf{F}^H \mathbf{R} \mathbf{F} \right)^{-\frac{1}{2}} - z\mathbf{I} \right)^{-1} \left(\mathbf{F}^H \mathbf{R} \mathbf{F} \right)^{-\frac{1}{2}} \mathbf{F}^H \mathbf{v}_t \Bigg|_{z=0} \\ & = \frac{d}{dz} h(z) \Bigg|_{z=0} . \end{aligned}$$

Applying Theorem 4.1 again, we know $|h(z)-\eta(z)| \to 0$, so we can approximate the derivative of $h(z)$ using $\frac{d}{dz}\eta(z)$. After some simplifications, we obtain

$$\frac{d}{dz}\eta(z) = \frac{1-x'(z)}{(x(z)-z)^2}\mathbf{v}_t^H\mathbf{F}\left(\mathbf{F}^H\mathbf{R}\mathbf{F}\right)^{-1}\mathbf{F}^H\mathbf{v}_t\,, \tag{4.42}$$

so the denominator is asymptotically equal to

$$\frac{1-x'(0)}{x^2(0)}\mathbf{v}_t^H\mathbf{F}\left(\mathbf{F}^H\mathbf{R}\mathbf{F}\right)^{-1}\mathbf{F}^H\mathbf{v}_t\,.$$

Now we have the approximate expressions for both the numerator and the denominator, and the ratio is

$$\frac{\sigma^2\xi_t}{1-x'(0)}\mathbf{v}_t^H\mathbf{F}\left(\mathbf{F}^H\mathbf{R}\mathbf{F}\right)^{-1}\mathbf{F}^H\mathbf{v}_t\,, \tag{4.43}$$

where $x'(0) = -\frac{1}{K}\left\{\sum_{m=1}^{M}\frac{1}{\lambda_m\left(1-\frac{M}{K}\right)}\right\}$, and λ_m are the eigenvalues of $\mathbf{F}^H\mathbf{R}\mathbf{F}$.

4.5 Simulations

The simulations in this section verify the theoretical results obtained above. The SINRs for both ULA and rectangular array cases are simulated, and a constant gamma model is assumed for the clutter [110]. The other simulation parameters can be found in Table 4.1.

Simulation results for the ULA case can be found in Fig. 4.5 and Fig. 4.6. Fig. 4.5 shows the change in SINR as a function of the number of antennas, and we observe that the performance of the fully adaptive STAP algorithm is always very close the theoretical upper bound, while the performance of the reduced-dimension separable STAP algorithm is rather poor when the number of antenna elements is small, but it improves and eventually approaches the fully adaptive STAP algorithm as N grows. In this case only $N > 50$ is required for the

Table 4.1: STAP System Parameters

SNR for a single antenna element ξ_t	4 dB
Reflectivity γ	-3 dB
Noise power per element σ^2	1
Range	6000 m
Number of clutter patches	360
Backlobe level	0 dB
Pulse repetition frequency (PRF) f_r	300 Hz
Interelement spacing	$\lambda/2$
Carrier frequency	0.45 GHz
Platform altitude	3000 m
Platform velocity	50 m/s
Number of pulses M	7 for ULA arrays, 4 for rectangular arrays

performance of the two algorithms to converge. Fig. 4.6 shows the change in SINR as a function of the target speed for $N = 146$ antennas, and we see only a slight reduction in minimum detectable velocity for the separable algorithm.

Fig. 4.7 and Fig. 4.8 contain similar simulation results for $P \times 3$ rectangular arrays where P grows from 14 to 278. For Fig. 4.8, a 278×3 antenna array is used. Fig. 4.9 and Fig. 4.10 also show the change in SINR for rectangular arrays, but for the case that the target is above the ground. For Fig. 4.9, the antenna configuration is $3 \times Q$ rectangular arrays where Q grows from 14 to 110. We see that again, for a large enough array, the separable algorithm performs nearly the same as the fully adaptive algorithm. For Fig. 4.10, the configuration of antenna elements is 3×110. As such, there is no significant SINR loss even when the target is relatively stationary. The fully adaptive algorithm exhibits virtually no SINR loss, while the loss of the separable algorithm is less than 2dB.

Fig. 4.11 verifies the scaling law for SINR loss analyzed in Section 4.3. A ULA is considered and the single-trial SINR loss is calculated for a target at the relative radial speed of 50 m/s.

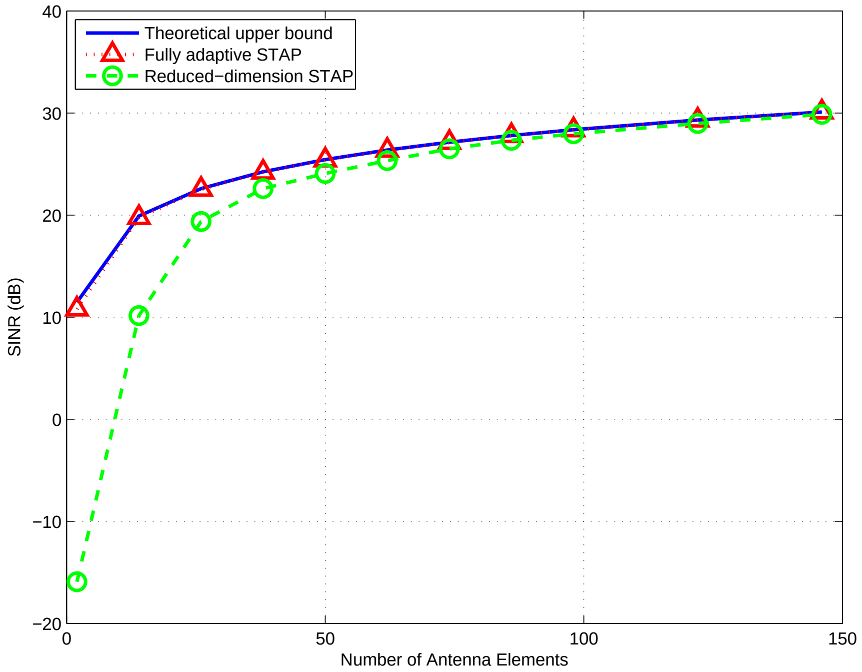


Figure 4.5: Max. SINR as a function of the number of antenna elements.

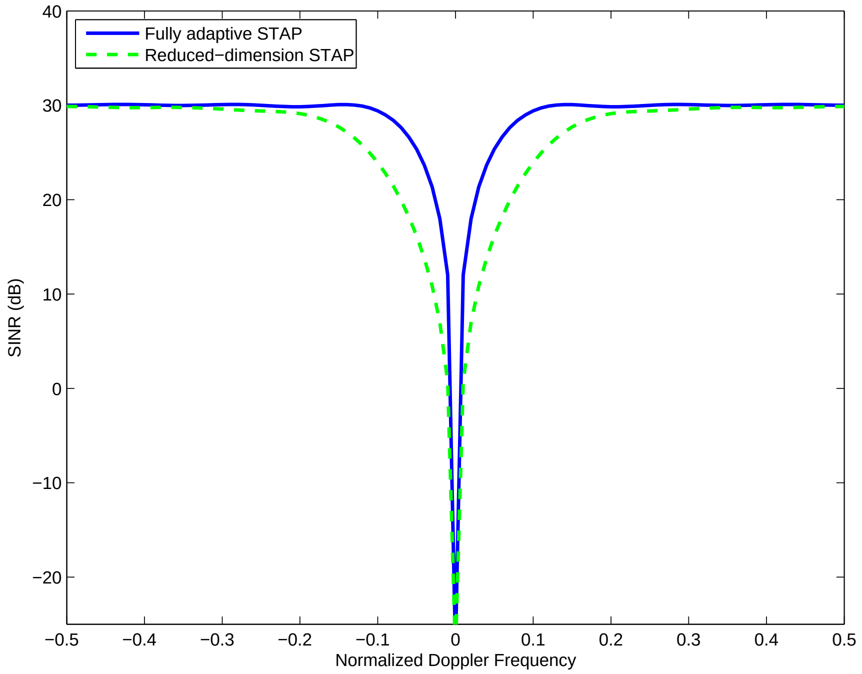


Figure 4.6: SINR as a function of normalized Doppler frequency.

For comparison, a curve showing κ/N^2 is also provided, and we observe that the SINR loss obeys the scaling law revealed in Section 4.3 when N is large.

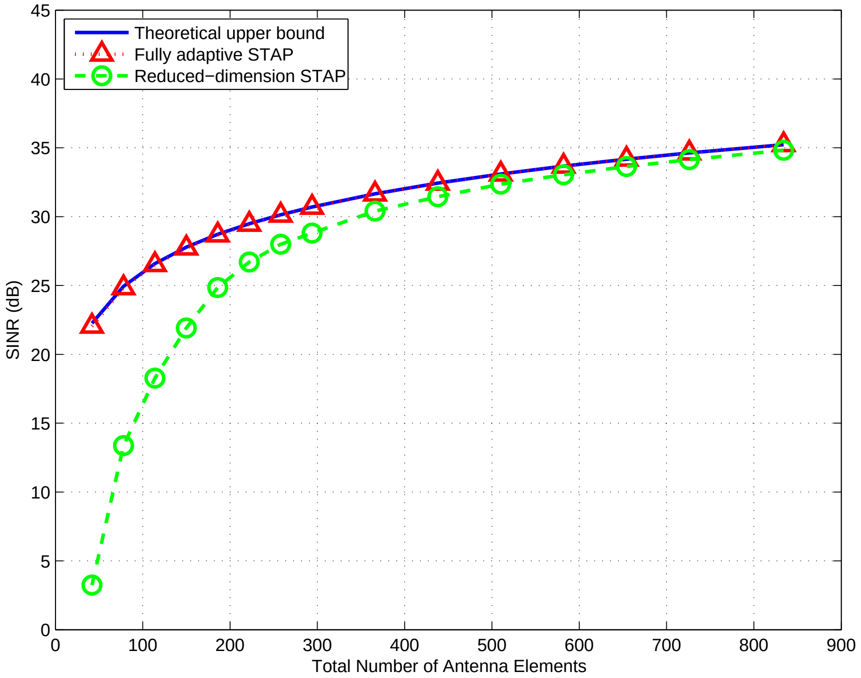


Figure 4.7: Max. SINR as a function of the total number of antenna elements.

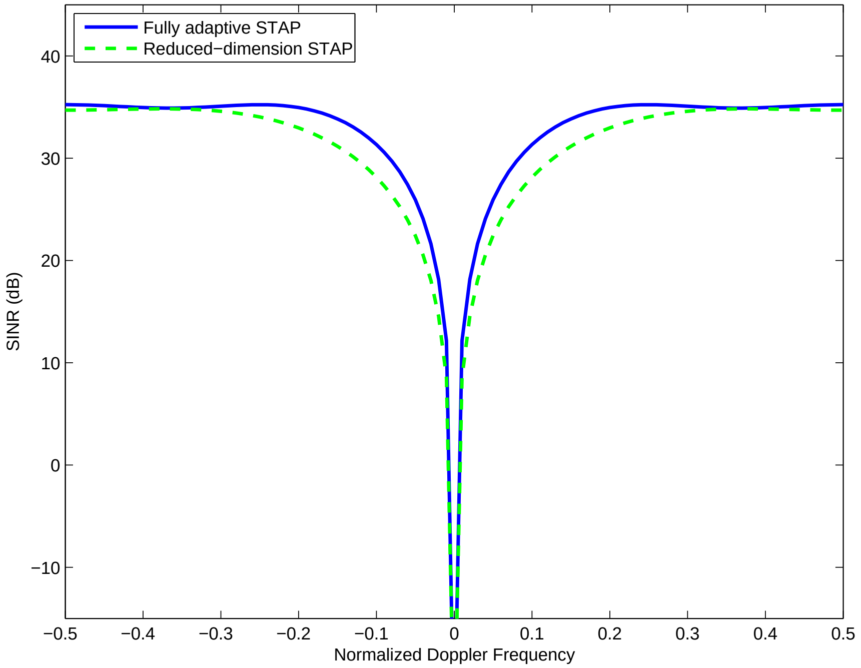


Figure 4.8: SINR as a function of normalized Doppler frequency.

Fig. 4.12 shows the SINR of the separable algorithm for a 2 × 73 URA with $M = 8$ pulses using a single sample estimate of the interference and noise covariance matrix $\mathbf{R}$ (blue curve),

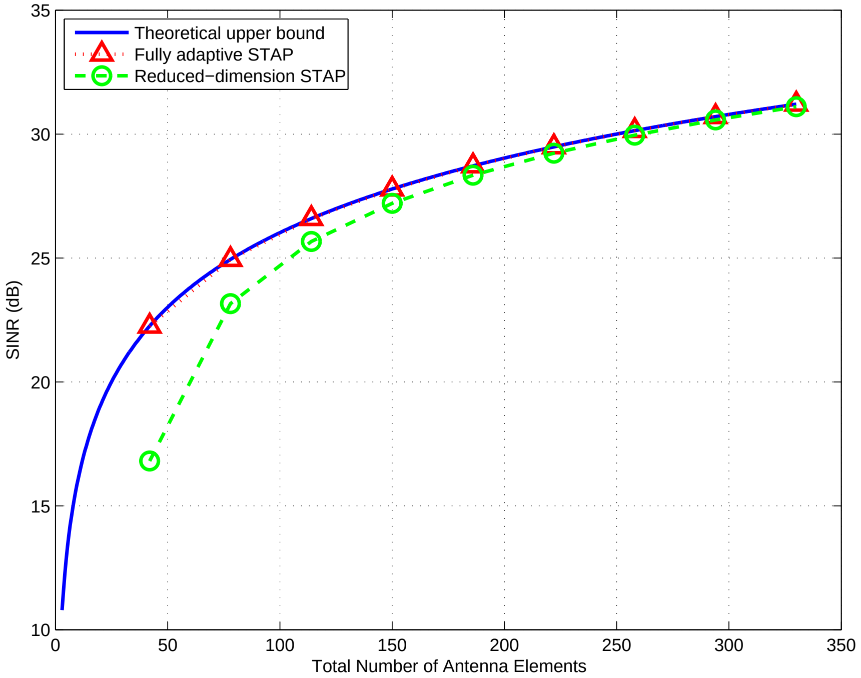


Figure 4.9: Max. SINR as a function of the total number of antenna elements.

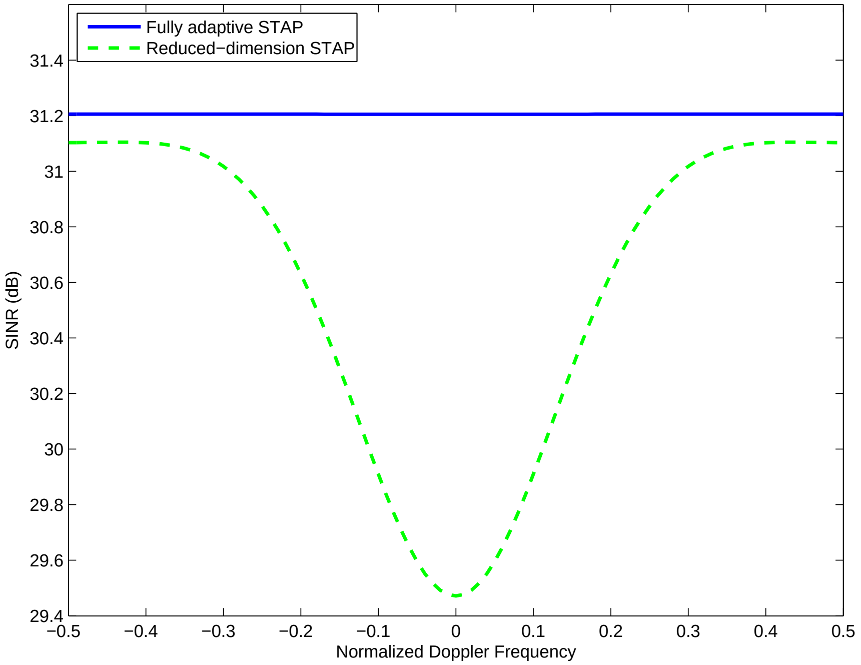


Figure 4.10: SINR as a function of normalized Doppler frequency.

together with the scaling law of (4.40) predicted by Theorem 4.1 (red curve). The green plot shows the average SINR achieved over 200 Monte Carlo simulations with different clutter

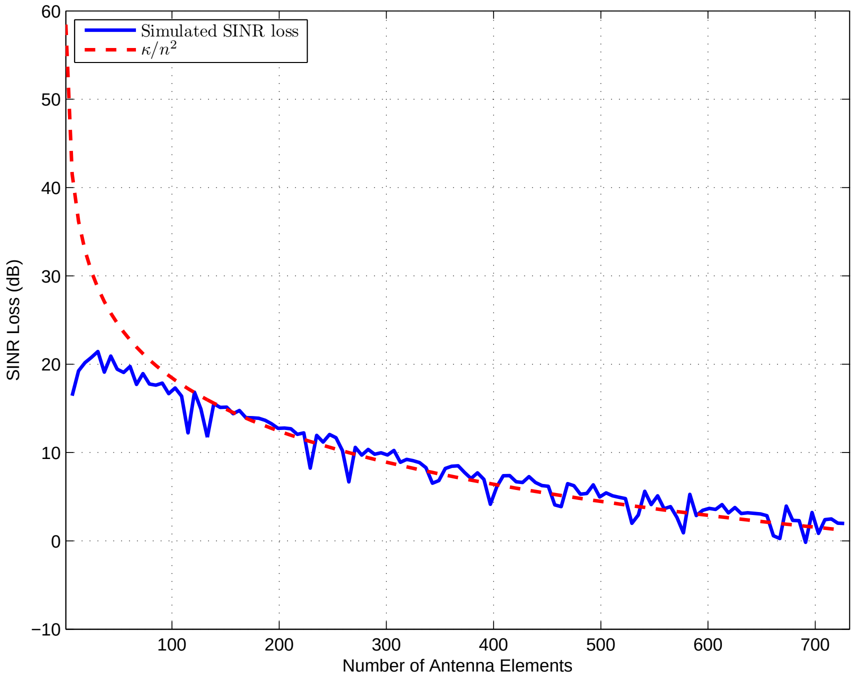


Figure 4.11: Scaling law of SINR loss.

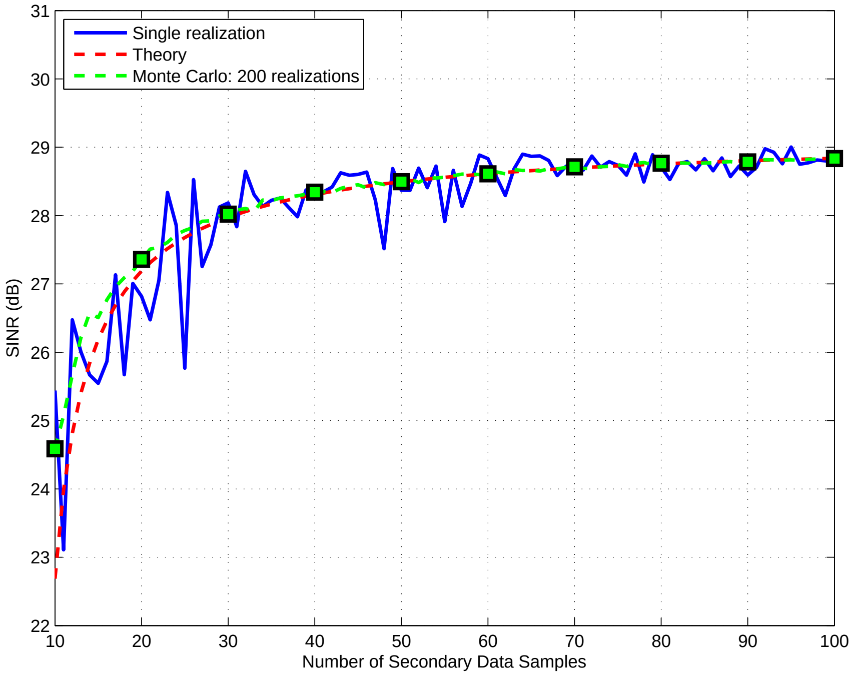


Figure 4.12: SINR as a function of the number of secondary data samples.

realizations. Both the single-trial and Monte Carlo simulation results match the theoretical prediction when K is large enough; in this case only $K > 30$ secondary vectors are enough for the theoretical results to hold with high accuracy.

Chapter 5

Positioning in NLOS Environments with a Minimal Set of Measurements

5.1 Introduction

In recent years, mobile positioning technology has received increasing attention and found many applications in industrial, medical, public safety, and entertainment areas. For example, the enhanced emergency call service in the United States, E-911 [28], requires positioning accuracy of 50 meters for 67% of calls and 150 meters for 90% of calls with handset based localization, and 100 meters for 67% of calls and 300 meters for 90% of calls with network based localization. Besides public-safety applications, another overwhelming driving force for the wide deployment of positioning technology is the strikingly high and still growing penetration of smart phones. Nowadays people use location based services like Yelp, Google Maps and other applications to get information of local businesses, news and weather conditions, which is almost an indispensable part of daily lives.

Various types of technologies can be used for positioning. These technologies are significantly different in physical nature and mathematical theories. The most successful and well-known positioning technology is the global positioning system (GPS) [76], which is a medium-earth orbit satellite based navigation system that provides location and timing information. Positioning in radio networks like cellular communication systems [123] and wireless local area networks (WLANs) [58] is also widely used all over the world. Other interesting but not extensively deployed technologies include radio frequency identification (RFID) [4], infrared (IR) [115], Bluetooth [41], ultrasound identification [44], and optical locating [63]. These methods have their own advantages and disadvantages. For example, RFID and other beacon based methods can provide very high positioning accuracy, whereas they are not scalable for massive deployment. Positioning using cameras, one type of optical localization, requires good illumination and quality images, which sometimes may be unavailable.

For radio wave based technologies, the positioning principles can be divided into three types. The technologies of the first type determine the position of the target based merely on the presence of the target in a particular area, which is within the range of an anchor device transmitting beacon signals. This includes most RFID and Bluetooth based methods, and Cell ID based positioning in UMTS [123] can also be regarded as a variant of this type. The second one is fingerprinting [46], which is most well-known for Wi-Fi networks. In this case, the received signal strength (RSS) is measured at different points within an area covered by WLANs. Thus the radio map of this area is drawn and stored in a database. When positioning is requested, the RSS of the target is reported and compared to various points on the radio map, and the closest one is chosen as the position of the target. The last and most widely used technology is geometric positioning, *e.g.*, GPS and observed time difference of arrival (OTDOA) [2] positioning, where geometrical relationship is exploited to calculate the position of the target based on angle-of-arrival (AOA), time-of-arrival (TOA) [22, 108], time-difference-of-arrival (TDOA) [15], and RSS [77] measurements. Recently this type of methods is even considered for fine-grained RFID localization [107].

However, successful positioning by geometric methods ordinarily requires that there exist unobstructed LOS paths from transmitting to receiving devices. In many cases, radio signals are obstructed by physical obstacles like trees, buildings, and mountains in outdoor environments, and walls and furniture in indoor environments, where the dominant paths are highly likely to be NLOS. This is because in such complicated wireless communication environments, transmitted waves suffer various radiation phenomena such as diffraction, refraction, reflection and scattering before they arrive at destinations. In addition, radio propagation is even more complicated in densely populated urban areas due to the mobility of targets. For a moving car or person, simply turning at a corner into a side street may radically change the wireless environment and the moving object may lose its LOS path to the controlling base station.

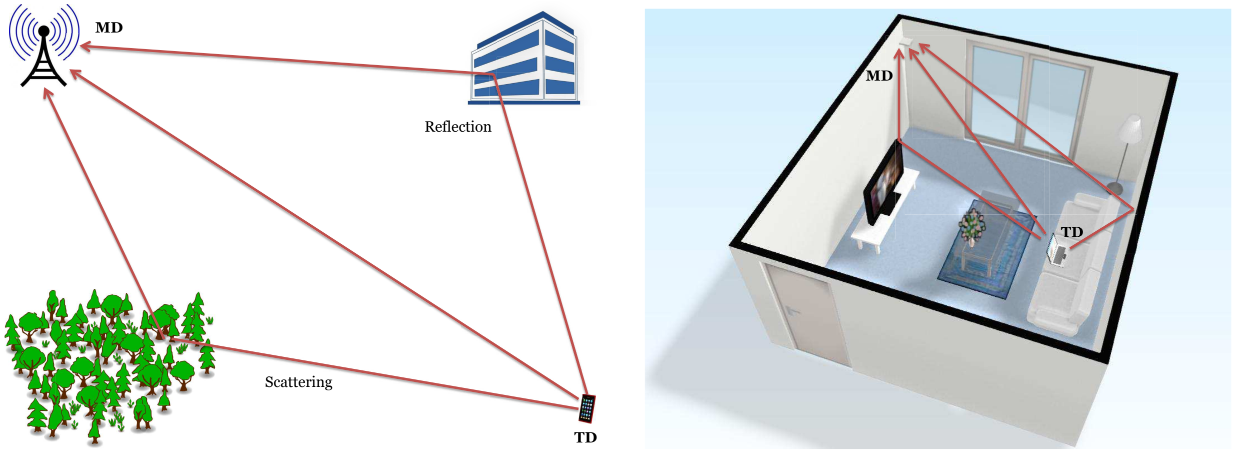
Many localization schemes have been proposed for wireless communication systems in the past years, and most of them assume that three or more LOS propagation paths exist between the transmitters and the receivers. Even in cases where these LOS paths exist, additional NLOS arrivals are almost unavoidable, and the NLOS signals act as the major source of interference that lowers the reliability of geometrical measurements, resulting in considerable positioning errors. Thus far, most research efforts on combating the effects of NLOS signals have focused on error mitigation, *i.e.*, how to detect multipaths that could be mistakenly perceived as LOS paths and then remove their impact [73].

In this chapter, we deal with NLOS signals from a different perspective. Instead of treating them as detrimental factors, we try to extract useful information from them. Furthermore, we want to challenge ourselves with the use of a minimal set of measurements. Obviously, the more types of geometrical measurements we have, the better positioning performance we can expect. However, due to the limitations of hardware in civilian communication systems, sometimes we may only be able to have certain types of measurements but not all. This fact motivates us to see to what extent positioning can be done with minimal input information. A two-stage approach is proposed in the chapter to localize the target. One scenario considered in the first stage is where NLOS paths are due to scattering. We exploit a single-bounce assumption on propagation to deal with localization in pure scattering NLOS environments where no direct LOS path exists, and then extend the method to take LOS paths into account. A similar scenario was discussed in [91], though the method of [91] requires additional measurements such as AOA at mobile stations, which is often difficult to obtain in the absence of a considerable array aperture and a stationary reference orientation for the device. The other NLOS scenario analyzed for the first stage is about reflection, and we find that in this case positioning is only possible with the help of two or more anchor points. In the second stage of the proposed approach, we appeal to the extended Kalman filter (EKF) to track subsequent changes in the target position and velocity, and the output of the first stage are used as the initial values of EKF.

The rest of the chapter is organized as follows. In Section 5.2, we describe the positioning problem in detail and introduce the notation. Section 5.3 illustrates the system models for scattering scenarios and proposes methods to estimate the position and velocity of the target. Section 5.4 analyzes the reflection case. Section 5.5 explains how to formulate the EKF for the tracking problem. Simulation results are presented in Section 5.6.

5.2 NLOS Positioning Problems

In this work, we study two types of NLOS positioning problems, *i.e.*, positioning in scattering and reflection environments, with minimal sets of measurements. Illustrative examples are depicted in Fig. 5.1. The results in this study is applicable to both Wi-Fi and cellular communication cases, so we use the general term “measuring device (MD)” to represent any device that is doing measuring, and “target device (TD)” to represent the target to be localized.



(a) Outdoor environments

(b) Indoor environments

Figure 5.1: Wireless signal propagation in outdoor and indoor environments.

The quantities used for positioning in this chapter are TOA and AOA measurements. Accurate AOA measurements require large aperture antenna arrays, which are feasible at the base station or access point side, but may not be feasible at the moving terminal side. Therefore we restrict ourselves to utilize *uplink* signals only, *i.e.*, just TOA and AOA information at MDs are available for positioning. An MD in our case can be a base station in cellular systems, or an access point in WLANs, and a TD can be a cell phone on a street or a laptop computer inside a room. For positioning in NLOS environments with both uplink and downlink measurements, a good discussion can be found in [91].

We use $(x_{md}^{(i)}, y_{md}^{(i)})$ to denote the coordinate of MD i, which is assumed to be known *a priori* for all MDs because they are stationary and the coordinates can be decided off-line. The measurements we have are $\tau_i(t)$, the TOA measured at MD i with respect to the TD, and $\theta_i(t)$, the AOA of the signal arriving at MD i. These quantities are the only inputs to our algorithms. We use them to estimate the position of the target, $(x_t(t), y_t(t))$, and its velocity, $\mathbf{v}(t) = (v_x(t), v_y(t))$. This problem is difficult because we have no knowledge about all other parameters such as the initial position of the target, the coordinates of the scatterers, the orientation and the location of the reflection wall, and the AOAs of signals arriving at the target. This fact adds significant difficulty to the localization problem. If we had these data, the problem could be quite easy to solve, but now some of them are also intermediate parameters that we need to estimate before they can be used for positioning.

With such a limited set of measurements in mind, we adopt a two-stage process to localize the target. In the first stage, we assume that the target is moving on a straight line at a constant speed for a short period of time, and the propagation environment is not changing. Under this assumption we propose methods to find the position and velocity of the target. We think the assumption is mild because simulation reveals that several seconds of measuring should suffice, which is reasonable for either pedestrians moving in rooms or cars moving on streets. In the second stage, we discard the short-time constant velocity assumption and use EKF to track the movements, using the estimates in the first stage as the initial values of tracking.

5.3 Scattering Model

In this section, we consider scattering scenarios for the first positioning stage, where the signal transmitted by the moving target is scattered by the scatterers near MDs. We will develop the solutions for two different scatterer-only models, for both of which no LOS path is observed at the MDs. Then we extend the methods to include LOS paths.

In the first model, several MDs are simultaneously monitoring the signals from the target, where each MD is associated with one major scatterer. While the signal may arrive at each MD via many multipaths, here we assume that the path with the smallest delay is due to a single bounce from a scatterer at a fixed but unknown location. The TOA and AOA of this path at the MD are measured and forwarded to a centralized controller where the target location estimate is obtained.

In the second model, only one MD is doing measuring and calculation. In contrast to the

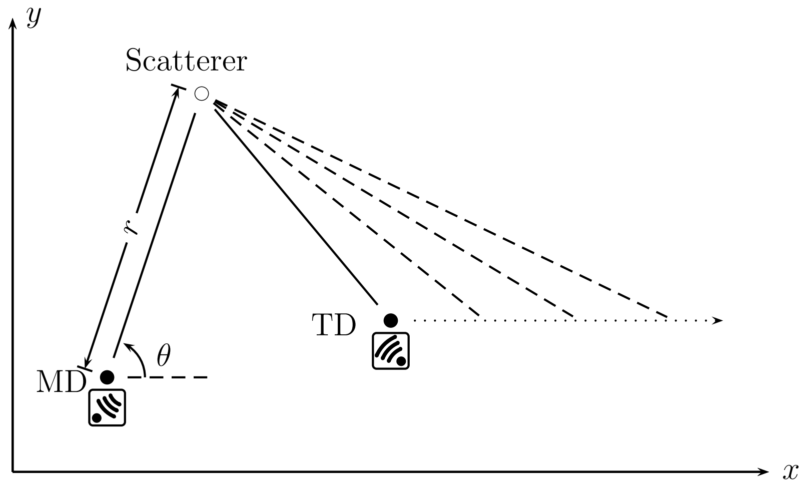


Figure 5.2: Geometrical relationship of one MD-TD pair.

first model where just the data of the path with the smallest delay are used, in this case all multipath components of the wireless channel are utilized. Therefore, the MD is effectively associated with multiple scatterers.

The positioning methods for both models depend on the fact that we can find the exact expression of the distance between the scatterer and the target for one MD-TD pair. So we do an extensive analysis for the one MD-TD pair case first, then explain how the exact expression can be used for positioning in the two models. Furthermore, the proposed methods can handle not only the pure NLOS problems, but are also applicable to LOS and NLOS mixed problems. We include the extension in this section for completeness.

5.3.1 One MD-TD Pair

In this subsection, we consider the case of one MD-TD pair only, because the corresponding analysis acts as the building block for the solution to the positioning problems we are studying. The case is illustrated in Fig. 5.2. Since we only have one MD and one associated scatterer, for the brevity of notation we ignore all indices with respect to MDs and scatterers, and thus (x_s, y_s) is the coordinate of the scatterer, r is the distance between the MD and the scatterer, and $\rho(t)$ is the distance between the scatterer and the target at time t, which are all unknown. Let θ be the AOA measurement of the signal arriving at the MD, from simple geometry we have the following relationships:

$$\begin{cases} x_s = x_{md} + r\cos\theta \\ y_s = y_{md} + r\sin\theta \end{cases} , \tag{5.1}$$

$$\tan\theta = \frac{y_s - y_{md}}{x_s - x_{md}} , \tag{5.2}$$

and

$$\theta = \begin{cases} \arctan\frac{y_s - y_{md}}{x_s - x_{md}} & \text{if } x_s \geq x_{md} \\ \pi + \arctan\frac{y_s - y_{md}}{x_s - x_{md}} & \text{if } x_s < x_{md} \end{cases} . \tag{5.3}$$

The TOA measurement equation is given by

$$\sqrt{(x_s - x_t(t))^2 + (y_s - y_t(t))^2} + \sqrt{(x_s - x_{md})^2 + (y_s - y_{md})^2} = c\tau(t) , \tag{5.4}$$

where $\tau(t)$ is the TOA of the signal from the target to the MD due to the associated scatterer. As we mentioned before, in the first stage we have the short-time constant velocity assumption, so the velocity of the target does not change with time and is denoted by $\mathbf{v} = (v_x, v_y)$. The initial position of the target is $(x_0, y_0) = (x_t(0), y_t(0))$.

At first glance, it might seem that little information can be extracted from the single-bounce model and hence the NLOS positioning problem might not be identifiable, but this is not so. In [18], we proposed a nonlinear least squares (NLS) method to approximately solve the positioning problem. In this chapter we move one step forward and show that an exact solution for r can be obtained, which is done by taking the third-order difference of TOA equations.

Based on Eq. (5.4), we have

$$\rho^2(t) = (x_t(t) - x_s)^2 + (y_t(t) - y_s)^2 \tag{5.5}$$

$$\rho^2(t) = (\tau(t)c - r)^2 \ . \tag{5.6}$$

We assume that the propagation environment does not change and the scatterer is stationary in the first stage, so r is not a function of t. This assumption is relaxed in the second stage.

Equating (5.5) and (5.6), we obtain

$$(x_t(t) - x_s)^2 + (y_t(t) - y_s)^2 = (\tau(t)c - r)^2 \ .$$

If the measurement is done on a regular basis with a time interval of Δt, at the nth measuring instant the above equation becomes

$$(x_t(n\Delta t) - x_s)^2 + (y_t(n\Delta t) - y_s)^2 = (\tau(n\Delta t)c - r)^2 \ .$$

Due to the constant velocity assumption in the first stage, we know $x_t(n\Delta t) = x_0 + nv_x\Delta t$ and $y_t(n\Delta t) = y_0 + nv_y\Delta t$. To save space we use the notation $\tau_n = \tau(n\Delta t)$. For two consecutive time instants n and $n+1$, we have

$$(x_0 + nv_x\Delta t - x_s)^2 + (y_0 + nv_y\Delta t - y_s)^2 = (c\tau_n - r)^2 \ , \tag{5.7}$$

and

$$(x_0 + (n+1)v_x\Delta t - x_s)^2 + (y_0 + (n+1)v_y\Delta t - y_s)^2 = (c\tau_{n+1} - r)^2 \ . \tag{5.8}$$

Taking the difference of both sides of (5.7) and (5.8), we have

$$\begin{aligned} & v_x\Delta t\left(2x_0+(2n+1)v_x\Delta t-2x_s\right)+v_y\Delta t\left(2y_0+(2n+1)v_y\Delta t-2y_s\right) \\ & =\left(c\tau_{n+1}+c\tau_n-2r\right)\left(c\tau_{n+1}-c\tau_n\right)\ , \end{aligned} \tag{5.9}$$

which is the first-order difference of the original TOA measurement equations. At time instant $n+1$, we similarly have

$$\begin{aligned} & v_x\Delta t\left(2x_0+(2n+3)v_x\Delta t-2x_s\right)+v_y\Delta t\left(2y_0+(2n+3)v_y\Delta t-2y_s\right) \\ & =\left(c\tau_{n+2}+c\tau_{n+1}-2r\right)\left(c\tau_{n+2}-c\tau_{n+1}\right)\ . \end{aligned} \tag{5.10}$$

Taking the difference of both sides of (5.9) and (5.10), we obtain the second-order difference for instants n and $n+1$:

$$\begin{aligned} & 2\left(v_x\Delta t\right)^2+2\left(v_y\Delta t\right)^2= \\ & \left(c\tau_{n+2}+c\tau_{n+1}-2r\right)\left(c\tau_{n+2}-c\tau_{n+1}\right)-\left(c\tau_{n+1}+c\tau_n-2r\right)\left(c\tau_{n+1}-c\tau_n\right)\ , \end{aligned} \tag{5.11}$$

and

$$\begin{aligned} & 2\left(v_x\Delta t\right)^2+2\left(v_y\Delta t\right)^2= \\ & \left(c\tau_{n+3}+c\tau_{n+2}-2r\right)\left(c\tau_{n+3}-c\tau_{n+2}\right)-\left(c\tau_{n+2}+c\tau_{n+1}-2r\right)\left(c\tau_{n+2}-c\tau_{n+1}\right)\ . \end{aligned} \tag{5.12}$$

As the final step, we take the third-order difference and get

$$\begin{aligned} & \left(c\tau_{n+2}+c\tau_{n+1}-2r\right)\left(c\tau_{n+2}-c\tau_{n+1}\right)-\left(c\tau_{n+1}+c\tau_n-2r\right)\left(c\tau_{n+1}-c\tau_n\right) \\ & =\left(c\tau_{n+3}+c\tau_{n+2}-2r\right)\left(c\tau_{n+3}-c\tau_{n+2}\right)-\left(c\tau_{n+2}+c\tau_{n+1}-2r\right)\left(c\tau_{n+2}-c\tau_{n+1}\right)\ , \end{aligned}$$

from which we can find the exact solution

$$\begin{aligned} r &= \frac{c}{2} \frac{\left(\tau_{n+3}^2 - 2\tau_{n+2}^2 + \tau_{n+1}^2\right) - \left(\tau_{n+2}^2 - 2\tau_{n+1}^2 + \tau_n^2\right)}{\left(\tau_{n+3} - 2\tau_{n+2} + \tau_{n+1}\right) - \left(\tau_{n+2} - 2\tau_{n+1} + \tau_n\right)} \\ &= \frac{c}{2} \left(\frac{\tau_{n+3}^2 - 3\tau_{n+2}^2 + 3\tau_{n+1}^2 - \tau_n^2}{\tau_{n+3} - 3\tau_{n+2} + 3\tau_{n+1} - \tau_n} \right) . \end{aligned} \tag{5.13}$$

In other words, if there is no measurement error, four consecutive delay measurements will suffice for determining the distance between the scatterer and the target, and the time needed for taking four delay measurements is $3\Delta t$, a really small period in practice. However, the measurements are not perfect in real systems and we need more samples to average out the errors. Therefore the time period of measuring should be longer than $3\Delta t$. Simulation results show that several seconds are necessary for acceptable positioning performance.

Because we know the TOA of the signal from the target to the MD via the scatterer, we indirectly know $\rho(t)$ by Eq. (5.6). Furthermore, with the AOA measurement at the MD and Eq. (5.1) we can determine the position of the scatterer, which replaces the MD to be the new anchor point. With these information, the triangulation method used in LOS positioning [89] can also be adopted to solve our problem if three or more anchor points are present.

Remark 5.1 (Interpretation of the Third-Order Difference)**.** *So far we have seen that by taking the difference of the original TOA measurement equations three times, we can find a closed-form and exact solution for r, and thus $\rho(t)$. Now we give an interpretation to show why this method is effective.*

In the continuous time domain, the delay measurement equation is

$$(x_t(t) - x_s)^2 + (y_t(t) - y_s)^2 = (\tau(t)c - r)^2 \ . \tag{5.14}$$

The first-order derivative of Eq. (5.14) *with respect to* t *is*

$$(x_t(t) - x_s)\,\dot{x}_t(t) + (y_t(t) - y_s)\,\dot{y}_t(t) = c\,(\tau(t)c - r)\,\dot{\tau}(t)\,,$$

and the second-order derivative is

$$\dot{x}_t^2(t) + (x_t(t) - x_s)\,\ddot{x}_t(t) + \dot{y}_t^2(t) + (y_t(t) - y_s)\,\ddot{y}_t(t) = c^2\dot{\tau}^2(t) + c\,(\tau(t)c - r)\,\ddot{\tau}(t)\,.$$

Due to the short-time constant velocity assumption, we know the acceleration of the target is zero, so $\ddot{x}_t(t) = 0$ *and* $\ddot{y}_t(t) = 0$ *and we get*

$$\dot{x}_t^2(t) + \dot{y}_t^2(t) = c^2\dot{\tau}^2(t) + c\,(\tau(t)c - r)\,\ddot{\tau}(t)\,.$$

Finally, taking the third-order derivative gives

$$2\dot{x}_t(t)\ddot{x}_t(t) + 2\dot{y}_t(t)\ddot{y}_t(t) = \frac{d}{dt}\left(c^2\dot{\tau}^2(t) + c\,(\tau(t)c - r)\,\ddot{\tau}(t)\right)\,,$$

which is further simplified as

$$\frac{d}{dt}\left(c\dot{\tau}^2(t) + (\tau(t)c - r)\,\ddot{\tau}(t)\right) = 0\,.$$

We can solve the above equation and find the expression of r *in the continuous time domain:*

$$r = \frac{3c\dot{\tau}(t)\ddot{\tau}(t) + c\tau(t)\dddot{\tau}(t)}{\dddot{\tau}(t)} = \frac{c}{2}\left(\frac{\frac{d^3}{dt^3}\tau^2(t)}{\frac{d^3}{dt^3}\tau(t)}\right)\,. \tag{5.15}$$

It tells us that the distance between the scatterer and the MD is obtainable if the acceleration of the target is zero. The solution in Eq. (5.13) *can be regarded as a discrete time counterpart of Eq.* (5.15)*.*

Remark 5.2 (Speed and Doppler Estimation)**.** *Besides the distance between the scatterer* *and the target, solving Eq.* (5.11) *gives the expression of target speed:*

$$\|\mathbf{v}\| = \frac{\sqrt{2}}{2\Delta t}\sqrt{\left(c\tau_{n+2}+c\tau_{n+1}-2r\right)\left(c\tau_{n+2}-c\tau_{n+1}\right)-\left(c\tau_{n+1}+c\tau_{n}-2r\right)\left(c\tau_{n+1}-c\tau_{n}\right)}\,,$$

though we cannot obtain the heading information of the target.

Furthermore, once the speed is obtained, we can plug it back to Eq. (5.9) *and get*

$$\begin{aligned}&v_x\Delta t\left(2x_0+2nv_x\Delta t-2x_s\right)+v_y\Delta t\left(2y_0+2nv_y\Delta t-2y_s\right)\\&=\left(c\tau_{n+1}+c\tau_{n}-2r\right)\left(c\tau_{n+1}-c\tau_{n}\right)-\left(\|\mathbf{v}\|\Delta t\right)^2\ .\end{aligned}$$

Therefore the Doppler frequency shift of the target at time instant n is

$$\begin{aligned}f_d&=\frac{1}{2\pi}\langle\mathbf{k},\mathbf{v}\rangle=\frac{v_x\left(x_0+nv_x\Delta t-x_s\right)+v_y\left(y_0+nv_y\Delta t-y_s\right)}{\rho(n\Delta t)\lambda}\\&=\frac{\left(c\tau_{n+1}+c\tau_{n}-2r\right)\left(c\tau_{n+1}-c\tau_{n}\right)-\left(\|\mathbf{v}\|\Delta t\right)^2}{2\rho(n\Delta t)\lambda\Delta t}\,,\end{aligned}$$

where λ is the wavelength of the signal, $\mathbf{k}$ is the vector pointing from the scatterer to the target with the norm of $2\pi/\lambda$, and $\langle\cdot\,,\cdot\rangle$ denotes the operation of inner product.

5.3.2 Multiple MDs

With the help of the exact solution of r, we are ready to find the position of the target for the case depicted in Fig. 5.3, where a K-MD system observes uplink signals from a single target, and we want to know where the target is at time instant n. Instead of the MDs, the scatterers now act as new anchor points because of the pure NLOS nature of this scenario.

We want to fuse the data from all MDs and all measuring instants, and the method can be established as an extension to the conventional triangulation [89], where one anchor point

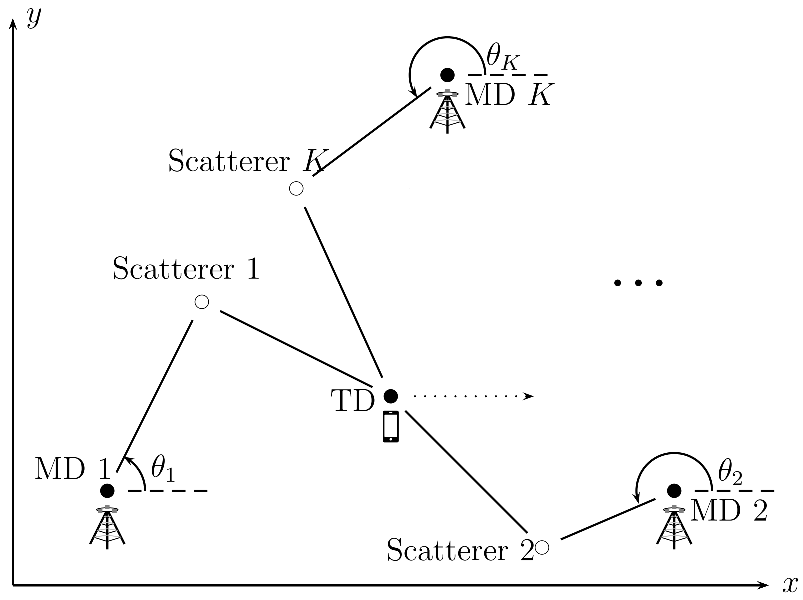


Figure 5.3: A localization scenario with K MDs, each associated with one scatterer relative to the MD.

is treated as the reference point, and others are used to form a set of supplementary linear equations. For our problem, first we look at the kth MD and have

$$\rho_k^2(t) = \left(x_t(t) - x_s^{(k)}\right)^2 + \left(y_t(t) - y_s^{(k)}\right)^2 \ , \tag{5.16}$$

where $\rho_k(t)$ is the distance between scatterer k and the target at time t.

If we choose the scatterer associated with MD 1 as the reference point, we have

$$\rho_1^2(t) = \left(x_t(t) - x_s^{(1)}\right)^2 + \left(y_t(t) - y_s^{(1)}\right)^2 \ . \tag{5.17}$$

Subtracting the left- and right-hand sides of (5.17) from (5.16) for $k \neq 1$, we get

$$\rho_k^2(t) - \rho_1^2(t) = \left(x_s^{(1)} - x_s^{(k)}\right)\left(2x_t(t) - x_s^{(1)} - x_s^{(k)}\right) + \left(y_s^{(1)} - y_s^{(k)}\right)\left(2y_t(t) - y_s^{(1)} - y_s^{(k)}\right) \ . \tag{5.18}$$

Rearranging the terms in Eq. (5.18), we obtain

$$\begin{aligned}&\frac{1}{2}\left(\rho_k^2(t)-\rho_1^2(t)+\left(\left(x_s^{(1)}\right)^2-\left(x_s^{(k)}\right)^2\right)+\left(\left(y_s^{(1)}\right)^2-\left(y_s^{(k)}\right)^2\right)\right)\\&\quad=\left(x_s^{(1)}-x_s^{(k)}\right)x_t(t)+\left(y_s^{(1)}-y_s^{(k)}\right)y_t(t)\,.\end{aligned}\tag{5.19}$$

Let t take values of $n\Delta t$, $n=0,1,2,\ldots,N-1$. The resulted equations can be rewritten as $\mathbf{y}_k=\mathbf{H}_k\mathbf{x}$ for k other than 1, where $\mathbf{x}=[\,x_t(n\Delta t),\ y_t(n\Delta t),\ v_x,\ v_y\,]^T$,

$$\mathbf{H}_k=\begin{pmatrix}x_s^{(1)}-x_s^{(k)} & y_s^{(1)}-y_s^{(k)} & -n\Delta t\left(x_s^{(1)}-x_s^{(k)}\right) & -n\Delta t\left(y_s^{(1)}-y_s^{(k)}\right)\\ & \vdots & & \\ x_s^{(1)}-x_s^{(k)} & y_s^{(1)}-y_s^{(k)} & 0 & 0\\ & \vdots & & \\ x_s^{(1)}-x_s^{(k)} & y_s^{(1)}-y_s^{(k)} & (N-n-1)\Delta t\left(x_s^{(1)}-x_s^{(k)}\right) & (N-n-1)\Delta t\left(y_s^{(1)}-y_s^{(k)}\right)\end{pmatrix},\tag{5.20}$$

and

$$\mathbf{y}_k=\frac{1}{2}\begin{pmatrix}\rho_k^2(0)-\rho_1^2(0)+\left(\left(x_s^{(1)}\right)^2-\left(x_s^{(k)}\right)^2\right)+\left(\left(y_s^{(1)}\right)^2-\left(y_s^{(k)}\right)^2\right)\\ \vdots\\ \rho_k^2(n\Delta t)-\rho_1^2(n\Delta t)+\left(\left(x_s^{(1)}\right)^2-\left(x_s^{(k)}\right)^2\right)+\left(\left(y_s^{(1)}\right)^2-\left(y_s^{(k)}\right)^2\right)\\ \vdots\\ \rho_k^2((N-1)\Delta t)-\rho_1^2((N-1)\Delta t)+\left(\left(x_s^{(1)}\right)^2-\left(x_s^{(k)}\right)^2\right)+\left(\left(y_s^{(1)}\right)^2-\left(y_s^{(k)}\right)^2\right)\end{pmatrix}.\tag{5.21}$$

In order to utilize the information from all MDs, we stack $K-1$ sets of linear equations together:

$$\mathbf{H} = \begin{pmatrix} \mathbf{H}_2 \\ \vdots \\ \mathbf{H}_k \\ \vdots \\ \mathbf{H}_K \end{pmatrix}, \qquad \mathbf{y} = \begin{pmatrix} \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_k \\ \vdots \\ \mathbf{y}_K \end{pmatrix},$$

and

$$\mathbf{y} = \mathbf{H}\mathbf{x}\,. \tag{5.22}$$

If the measurements are free of error, equation (5.22) is an overdetermined system and solving it results in the exact solution for the position and velocity of the target. Practically the measurements are contaminated with noise, so we use the celebrated least squares form to get an approximate $\mathbf{x}$:

$$\hat{\mathbf{x}} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{y}\,.$$

Note that the proposed method requires three or more MDs. Otherwise the problem is not identifiable.

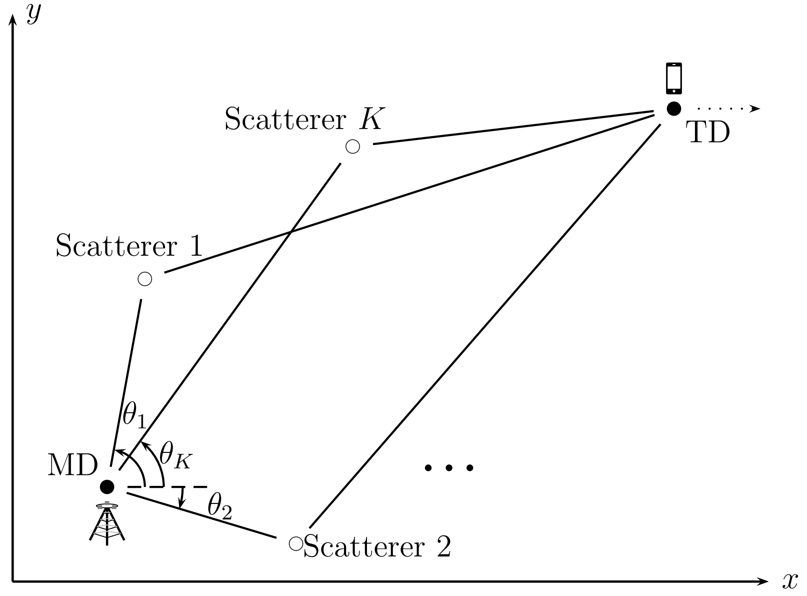


Figure 5.4: A localization scenario with one MD, associated with multiple scatterers.

5.3.3 Extensions

Fig. 5.4 depicts the second model: single MD associated with multiple scatterers. Similar ideas in the previous subsections are also applicable to this model, where all paths of the wireless channel from the target to the MD are utilized instead of one. The distance between scatterer k and the MD can be calculated by using the set of TOA measurements associated with the path passing through scatterer k, so the position of scatterer k is determinable. Then we can keep using the same equation, Eq. (5.22), and the least squares form to estimate the position and velocity of the target. In this model, different paths of the channel usually have different strength, which may have a considerable impact on the reliability of measurements.

The method can be further extended to include cases where both LOS and NLOS paths are observed. To take LOS paths into account, the only change needed is that $\rho_k(t)$ and $(x_s^{(k)}, y_s^{(k)})$ in (5.20) and (5.21) shall be replaced by the TOA of the corresponding LOS path and the position of the corresponding MD, because they are no longer unknown parameters.

Proper application of the proposed method is contingent on that we can classify the paths appropriately, *i.e.*, we need to be able to detect if a given path is LOS or is associated with scatterers. If the signals are from scatterers, we also need to be able to differentiate the single- and multi-scatterer cases. The classification/detection itself is a subject of great interest, but is beyond the scope of this chapter, which focuses on the positioning algorithm per se.

5.3.4 Improve the Quality of TOA Measurements

The quality of TOA measurements certainly has a great effect on the accuracy of final estimates. Motivated by the proposal in [89], we can use a nonlinear optimization method to improve the quality. For a set of TOA measurements obtained at a given MD, we formulate the following optimization problem:

$$\begin{aligned} &\min_{r,\tau_n} \quad \sum_{n=0}^{N-1} (\tau_n - \tilde{\tau}_n)^2 \\ &\text{s.t. } \ r = \frac{c}{2}\left(\frac{\tau_{n+3}^2 - 3\tau_{n+2}^2 + 3\tau_{n+1}^2 - \tau_n^2}{\tau_{n+3} - 3\tau_{n+2} + 3\tau_{n+1} - \tau_n}\right), \qquad n = 0, \dots, N-4 \ , \end{aligned}$$

where $\tilde{\tau}_n$ are the measured TOAs, and τ_n are optimization variables which can be regarded as refined TOAs. The optimization problem is non-convex and can only be solved by numerical methods. Our simulation reveals that the optimization improves the accuracy of measurements by almost one order of magnitude. As a side product of the optimization process, a refined r is also obtained.

5.4 Reflection Model

In this section we study another NLOS case for the first stage of positioning, which is about the important physical phenomenon of reflection. Imagine the signal received by the MD is reflected by a wall. Similar to the scatterer case, we just have the uplink TOA and AOA

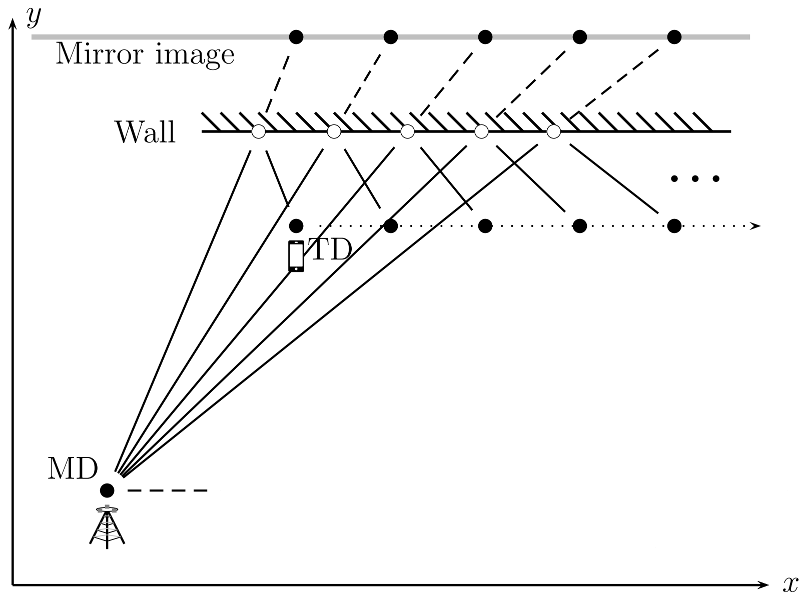


Figure 5.5: The trajectory of the target.

information at the MD side. We do not have maps of the underlying areas, so no knowledge about the wall is ready for use. The analysis shows that we cannot extract any meaningful information from the measurements for the single MD-TD pair due to the ambiguity in the model, which is significantly different from scattering scenarios.

The ambiguity in the single-pair reflection model can be explained by two illustrations. In Fig. 5.5, we qualitatively show the target's trajectory and the signal paths between the MD and the target. To ease the analysis, we also depict the mirror image of the trajectory with respect to the wall. If the TOA and AOA measurements are free of error, the mirror image is completely determinable. In Fig. 5.6 we keep everything the same, but change the orientation of the wall. It is clear that in this case all the TOA and AOA measurements can be chosen to be the same due to symmetry, but the trajectory of the target is totally different from the one in Fig. 5.5. In fact the number of candidate trajectories is infinite for the same set of measurements, only if the mirror image is kept unchanged. This observation tells us that there is no way to determine the orientation and location of the wall in the single-pair model, though the information about the wall is very important for correct positioning. Later we will see only if the measurements are used together with the help of two or more

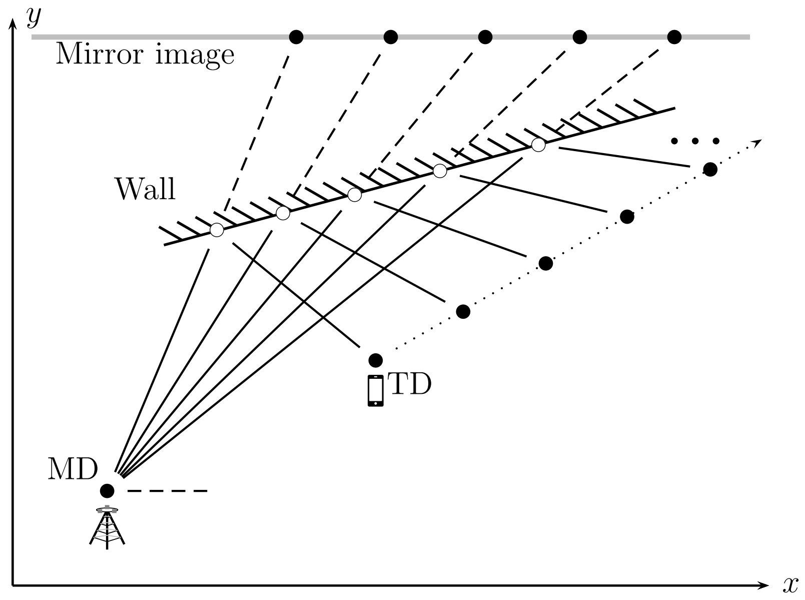


Figure 5.6: The ambiguity in the trajectory of the target.

anchor points, the ambiguity can be removed from the model and the positioning problem is solvable.

For the analysis of the reflection case, we use the same notation as before, and the only new quantity is $(x_m(t), y_m(t))$, which refers to the mirror image of the target's trajectory. Clearly we have

$$\begin{cases} x_m(t) = x_{md} + c\tau(t)\cos\theta(t) \\ y_m(t) = y_{md} + c\tau(t)\sin\theta(t) \end{cases}, \tag{5.23}$$

where $\tau(t)$ is the TOA and $\theta(t)$ is the AOA at the MD side, both changing with time. The mirror image characterized by $(x_m(t), y_m(t))$ can be regarded as known.

The method we propose to solve the positioning problem in the reflection case relies on two facts. Firstly, the lines determined by points on the trajectory of the target and the corresponding mirror points are parallel to each other, *i.e.*, the line passing through $(x_t(t_i), y_t(t_i))$ and $(x_m(t_i), y_m(t_i))$ and the line passing through $(x_t(t_j), y_t(t_j))$ and $(x_m(t_j), y_m(t_j))$ are parallel for all t_i and t_j, $t_i \neq t_j$. Secondly, all these lines are perpendicular to the wall because

$(x_m(t), y_m(t))$ is the mirror of $(x_t(t), y_t(t))$ with respect to the wall.

The equation of the line passing through $(x_t(t), y_t(t))$ and $(x_m(t), y_m(t))$ is

$$\frac{y - y_m(t)}{x - x_m(t)} = \frac{y - y_t(t)}{x - x_t(t)}, \tag{5.24}$$

which can be further converted to

$$y = \frac{y_t(t) - y_m(t)}{x_t(t) - x_m(t)} x + \frac{x_t(t) y_m(t) - x_m(t) y_t(t)}{x_t(t) - x_m(t)},$$

and the slope is

$$\frac{y_t(t) - y_m(t)}{x_t(t) - x_m(t)}.$$

On the other hand, the line equation to characterize the wall is given by

$$\frac{y - \frac{y_m(t)+y_t(t)}{2}}{x - \frac{x_m(t)+x_t(t)}{2}} = \frac{y - \frac{y_m(0)+y_0}{2}}{x - \frac{x_m(0)+x_0}{2}}, \tag{5.25}$$

which can be rewritten as

$$\begin{aligned} y =& \frac{y_m(t) - y_m(0) + y_t(t) - y_0}{x_m(t) - x_m(0) + x_t(t) - x_0} x \\ &+ \frac{(x_m(t) + x_t(t))(y_m(0) + y_0) - (x_m(0) + x_0)(y_m(t) + y_t(t))}{2(x_m(t) - x_m(0) + x_t(t) - x_0)}. \end{aligned}$$

So the slope of the line is

$$\frac{y_m(t) - y_m(0) + y_t(t) - y_0}{x_m(t) - x_m(0) + x_t(t) - x_0}.$$

We know the two lines defined by (5.24) and (5.25) are perpendicular to each other, which requires

$$\frac{y_m(t) - y_m(0) + y_t(t) - y_0}{x_m(t) - x_m(0) + x_t(t) - x_0} \cdot \frac{y_t(t) - y_m(t)}{x_t(t) - x_m(t)} = -1\,.$$

Due to the parallel property, it is equivalent to

$$\frac{y_m(t) - y_m(0) + y_t(t) - y_0}{x_m(t) - x_m(0) + x_t(t) - x_0} \cdot \frac{y_0 - y_m(0)}{x_0 - x_m(0)} = -1\,.$$

The assumption of constant velocity in a short period of time leads to

$$\frac{y_m(n\Delta t) - y_m(0) + nv_y\Delta t}{x_m(n\Delta t) - x_m(0) + nv_x\Delta t} \cdot \frac{y_0 - y_m(0)}{x_0 - x_m(0)} = -1\,.$$

For $n = 0, 1, 2, 3\ldots$, we define $x_{m_n} = x_m(n\Delta t)$ and $y_{m_n} = y_m(n\Delta t)$ for convenience. After simplifications we get

$$(nv_x\Delta t + x_{m_n} - x_{m_0})(x_0 - x_{m_0}) + (nv_y\Delta t + y_{m_n} - y_{m_0})(y_0 - y_{m_0}) = 0\,, \tag{5.26}$$

which summarizes the information we can get from the reflection wall. Furthermore it is easy to see

$$(x_{m_n} - x_{m_0})^2 + (y_{m_n} - y_{m_0})^2 = (nv_x\Delta t)^2 + (nv_y\Delta t)^2\,. \tag{5.27}$$

As we mentioned before, only together with the information provided by anchor points, we are able to find the position and velocity of the target in the reflection model. An anchor point is this section is defined as a point which effectively has LOS observation of the target, and it can be either an MD with known location and TOA, or a scatterer whose location and TOA are estimated as we do in Section 5.3. We do not require AOA knowledge at the anchor points, otherwise the localization process is simple and straightforward. Among these anchor points, we select one located at (x_a, y_a) as the reference and its TOA measurement at time t is denoted by $\tau_a(t)$. The criterion for selection can be based on the reliability of measurement provided by the anchor points. The coordinates of other anchor points are $(x_b^{(k)}, y_b^{(k)})$, $k = 1, 2, \ldots, K-1$. Clearly we have

$$(x_t(t) - x_a)^2 + (y_t(t) - y_a)^2 = (c\tau_a(t))^2 \ .$$

Applying the same technique we have used for the scatterer case, at time instant n we have

$$(x_0 + nv_x\Delta t - x_a)^2 + (y_0 + nv_y\Delta t - y_a)^2 = (c\tau_n)^2 \ ,$$

where $\tau_n = \tau_a(n\Delta t)$. For time instant $n+1$, we have

$$(x_0 + (n+1)v_x\Delta t - x_a)^2 + (y_0 + (n+1)v_y\Delta t - y_a)^2 = (c\tau_{n+1})^2 \ .$$

Taking the difference of both sides of the two equations, we get

$$\begin{aligned} & v_x\Delta t\,(2x_0 + (2n+1)v_x\Delta t - 2x_a) + v_y\Delta t\,(2y_0 + (2n+1)v_y\Delta t - 2y_a) \\ & = (c\tau_{n+1})^2 - (c\tau_n)^2 \ , \end{aligned} \tag{5.28}$$

which is the first-order difference of the delay equations at instant n. At time instant $n+1$, we again have

$$\begin{aligned} & v_x\Delta t\,(2x_0 + (2n+3)v_x\Delta t - 2x_a) + v_y\Delta t\,(2y_0 + (2n+3)v_y\Delta t - 2y_a) \\ & = (c\tau_{n+2})^2 - (c\tau_{n+1})^2 \ . \end{aligned}$$

Taking the second-order difference, we get

$$(v_x\Delta t)^2 + (v_y\Delta t)^2 = \frac{1}{2}\left((c\tau_{n+2})^2 - (c\tau_{n+1})^2 - ((c\tau_{n+1})^2 - (c\tau_n)^2)\right) .$$

Plugging it back to Eq. (5.28), we then get

$$\begin{aligned}
&v_x\Delta t\,(x_0 - x_a) + v_y\Delta t\,(y_0 - y_a) \\
&= -\frac{2n+1}{4}\left((c\tau_{n+2})^2 - (c\tau_{n+1})^2\right) + \frac{2n+3}{4}\left((c\tau_{n+1})^2 - (c\tau_n)^2\right) \\
&\triangleq \zeta_n \,.
\end{aligned} \tag{5.29}$$

It is difficult to completely linearize the positioning problem for the reflection case. Instead, we adopt a one-dimensional line search approach to find the solution. The angle of departure at the selected anchor point at time 0, $\varphi = \varphi(t)|_{t=0}$, is chosen to be the search variable. Equations (5.26) and (5.29) can be rewritten in φ as

$$\begin{aligned}
&(nv_x\Delta t + x_{m_n} - x_{m_0})(c\tau_0\cos\varphi + x_a - x_{m_0}) + \\
&\qquad (nv_y\Delta t + y_{m_n} - y_{m_0})(c\tau_0\sin\varphi + y_a - y_{m_0}) = 0 \,,
\end{aligned} \tag{5.30}$$

and

$$v_x\Delta t\,(c\tau_0\cos\varphi) + v_y\Delta t\,(c\tau_0\sin\varphi) = \zeta_n \,, \tag{5.31}$$

because the simple geometrical relationship holds:

$$\begin{cases} x_t(t) = x_a + c\tau_a(t)\cos\varphi(t) \\ y_t(t) = y_a + c\tau_a(t)\sin\varphi(t) \end{cases} , \tag{5.32}$$

With (5.27), (5.30) and (5.31), we are ready to show the full problem formulation, which incorporates the measurements at instant n, $n = 1, 2, \ldots, N$ from the reflection MD and all anchor points.

To make expressions concise, we define

$$\begin{cases} f_n = (x_{m_0} - x_{m_n})\,(c\tau_0 \cos\varphi + x_a - x_{m_0}) + (y_{m_0} - y_{m_n})\,(c\tau_0 \sin\varphi + y_a - y_{m_0}) \\ \mathbf{g}_n = n\Delta t\; [\, c\tau_0 \cos\varphi + x_a - x_{m_0},\;\; c\tau_0 \sin\varphi + y_a - y_{m_0}\,]^T \\ \mathbf{r}_n = c\tau_0 \Delta t\, [\,\cos\varphi,\;\; \sin\varphi\,]^T \end{cases},$$

and a matrix form relationship can be established as

$$\mathbf{y} = \mathbf{H}\mathbf{x}\,, \tag{5.33}$$

where

$$\mathbf{y} = [\, f_1, \cdots, f_n, \cdots, f_N, \zeta_1, \cdots, \zeta_n, \cdots, \zeta_N\,]^T\,,$$

$$\mathbf{H} = [\,\mathbf{g}_1^T, \cdots, \mathbf{g}_n^T, \cdots, \mathbf{g}_N^T, \mathbf{r}_1^T, \cdots, \mathbf{r}_n^T, \cdots, \mathbf{r}_N^T\,]^T\,,$$

and

$$\mathbf{x} = \begin{pmatrix} v_x \\ v_y \end{pmatrix}.$$

Note this is also an overdetermined system like the one in the scattering model. If there is no error in measurements, $N = 1$ should suffice for the solution of the system. For measurements with noise, we need a larger N to mitigate the effect of errors.

In order to define the objective function of the optimization problem, we further define

$$\begin{cases} \alpha_n & = (x_{m_n} - x_{m_0})^2 + (y_{m_n} - y_{m_0})^2 - \left((nv_x\Delta t)^2 + (nv_y\Delta t)^2\right) \\ \beta_n & = c\tau_n - \left((c\tau_0 \cos\varphi + nv_x\Delta t)^2 + (c\tau_0 \sin\varphi + nv_y\Delta t)^2\right)^{\frac{1}{2}} \\ \gamma_{k,n} & = c\,\varsigma_n^{(k)} - \left(\left(c\tau_0 \cos\varphi + x_a + nv_x\Delta t - x_b^{(k)}\right)^2 + \left(c\tau_0 \sin\varphi + y_a + nv_y\Delta t - y_b^{(k)}\right)^2\right)^{\frac{1}{2}} \end{cases},$$

where $\varsigma_n^{(k)}$ is the TOA associated with the kth non-reference anchor point at time instant n, α_n follows from Eq. (5.27), β_n and $\gamma_{k,n}$ reflect the difference of measured and calculated TOAs.

The optimization problem is

$$\begin{aligned} & \min_{\varphi} \; \sum_{n=1}^{N} \left(\alpha_n^2 + \beta_n^2 + \sum_{k=1}^{K-1} \gamma_{k,n}^2 \right) \\ & \text{subject to} \quad \mathbf{y} = \mathbf{H}\mathbf{x} \; . \end{aligned}$$

The only optimization variable in the objective function is φ. Simulation shows that the objective function has very limited number of local minima, and numerical algorithms can quickly and efficiently find the optimal φ. Once it is obtained, the position and velocity of the target can be calculated following (5.32) and (5.33).

Remark 5.3 (Requirement on the Number of Anchor Points)**.** *As we have seen, the reflection case is even more challenging compared to the scattering case. Without the help from other anchor points, the positioning problem itself is not identifiable. If the number of helping anchor points is one, there is still ambiguity in the model and the trajectory of the target is not uniquely determinable. For the case of two anchor points, at first glance it seems that the mirror image of the MD can act as the third anchor, and thus the conventional triangulation method is adoptable. However, a noticeable difference is that the position of the mirror in* *our model is unknown, because we have no knowledge about the orientation and the location of the wall. So we need to appeal to the proposed one-dimensional search method to localize the target. For the case of more than two anchor points, if the TOA measurements are accurate enough, triangulation using only the anchor points is the straightforward solution. The proposed method is still applicable though. Measuring the signal reflected by the wall is no longer necessary, but can be treated as a supplementary source of geometrical information.*

5.5 Tracking

The previous sections discuss the details about estimation of the position and velocity of TDs with the short-time constant velocity assumption. Once the estimation in the first stage is done, we appeal to the extended Kalman filter to track subsequent movements of the target, under which the constant velocity assumption is relaxed. The results from the first stage are used as the initial values of the EKF, which is able to handle the scattering, reflection and LOS cases simultaneously.

Kalman filters generally consist of two parts: the dynamic model and the observation model. For the dynamic model, we define:

$$\begin{cases} \boldsymbol{\eta}_1(n) = [\ x_t(n),\ y_t(n),\ v_x(n),\ v_y(n)\ ]^T \\ \boldsymbol{\eta}_2(n) = \left[\ x_s^{(1)}(n),\ y_s^{(1)}(n),\ \cdots,\ x_s^{(K_s)}(n),\ y_s^{(K_s)}(n)\ \right]^T \\ \boldsymbol{\eta}_3(n) = [\ x_w(n),\ y_w(n)\ ]^T \end{cases} ,$$

where $(x_t(n), y_t(n))$ is the position of the target, $(v_x(n), v_y(n))$ is the velocity of the target, $(x_s^{(k)}(n), y_s^{(k)}(n))$ is the coordinate of scatterer k, $(x_w(n), y_w(n))$ represents the mirror image of the MD with respect to the wall, K_s is the number of scatterers, and n is the time index.

Further defining $\boldsymbol{\eta}(n) = \left[\ \boldsymbol{\eta}_1^T(n),\ \boldsymbol{\eta}_2^T(n),\ \boldsymbol{\eta}_3^T(n)\ \right]^T$, we have the dynamic model:

$$\boldsymbol{\eta}(n) = \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \boldsymbol{\eta}(n-1) + \mathbf{w}(n), \tag{5.34}$$

where

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

and $\mathbf{w}(n)$ is the noise vector with covariance matrix $\mathbf{M}_w$.

The components of the observation vector are the uplink TOA and AOA measurements at the corresponding MDs. We define

$$\begin{cases} \boldsymbol{\mu}_1(n) = \left[\ \tau_s^{(1)}(n),\ \cdots,\ \tau_s^{(K_s)}(n),\ \theta_s^{(1)}(n),\ \cdots,\ \theta_s^{(K_s)}(n)\ \right]^T \\ \boldsymbol{\mu}_2(n) = [\ \tau_w(n),\ \theta_w(n)\ ]^T \\ \boldsymbol{\mu}_3(n) = \left[\ \tau_d^{(1)}(n),\ \cdots,\ \tau_d^{(K_d)}(n),\ \theta_d^{(1)}(n),\ \cdots,\ \theta_d^{(K_d)}(n)\ \right]^T \end{cases},$$

where $\tau_s^{(i)}(n)$ and $\theta_s^{(i)}(n)$ are the TOA and AOA measurements associated with scatterer i, τ_w and θ_w are with the reflection wall, $\tau_s^{(j)}(n)$ and $\theta_s^{(j)}(n)$ are with LOS path j, and K_d is the number of LOS paths. Similarly we define $\boldsymbol{\mu}(n) = \left[\ \boldsymbol{\mu}_1^T(n),\ \boldsymbol{\mu}_2^T(n),\ \boldsymbol{\mu}_3^T(n)\ \right]^T$, and $\mathbf{z}(n)$ as the noise vector with covariance matrix $\mathbf{M}_z$. Then the observation model is

$$\boldsymbol{\mu}(n) = \mathbf{h}\left(\boldsymbol{\eta}(n)\right) + \mathbf{z}(n),$$

where $\mathbf{h}$ is a stack of TOA and AOA equations for scattering, reflection and LOS cases, respectively. For the case of scattering, the TOA and AOA can be obtained from (5.3) and

(5.4). For the case of reflection, we have

$$\theta = \begin{cases} \arctan \frac{y-y_{md}}{x-x_{md}} & \text{if } x \geq x_{md} \\ \pi + \arctan \frac{y-y_{md}}{x-x_{md}} & \text{if } x < x_{md} \end{cases},$$

and

$$\tau = \frac{1}{c}\sqrt{(x - x_{md})^2 + (y - y_{md})^2},$$

where (x, y) is the coordinate of the reflection point on the wall, and can be calculated as

$$\begin{cases} x = \frac{\frac{1}{2}(x_w^2 - x_{md}^2 + y_w^2 - y_{md}^2)(x_w - x_t) + (y_w x_t - y_t x_w)(y_w - y_{md})}{(y_w - y_t)(y_w - y_{md}) + (x_w - x_{md})(x_w - x_t)} \\ y = \frac{\frac{1}{2}(x_w^2 - x_{md}^2 + y_w^2 - y_{md}^2)(y_w - y_t) - (y_w x_t - y_t x_w)(x_w - x_{md})}{(y_w - y_t)(y_w - y_{md}) + (x_w - x_{md})(x_w - x_t)} \end{cases}.$$

For the case of LOS transmission, the relationship is immediate:

$$\theta = \begin{cases} \arctan \frac{y_t-y_{md}}{x_t-x_{md}} & \text{if } x_t \geq x_{md} \\ \pi + \arctan \frac{y_t-y_{md}}{x_t-x_{md}} & \text{if } x_t < x_{md} \end{cases}. \tag{5.35}$$

$$\tau = \frac{1}{c}\sqrt{(x_t - x_{md})^2 + (y_t - y_{md})^2}, \tag{5.36}$$

Due to the nonlinear nature of $\mathbf{h}$, we use the EKF [55] to do the tracking. The computation of the coefficients in EKF is standard and straightforward, so we skip the detailed calculation steps. Note that the EKF we consider in this section is able to handle scattering, reflection and LOS scenarios simultaneously. If one of them is absent in reality, we just need to remove the corresponding part from the Kalman filter. For example, if there is no LOS path observed then we just need to remove (5.35) and (5.36) from the observation model.

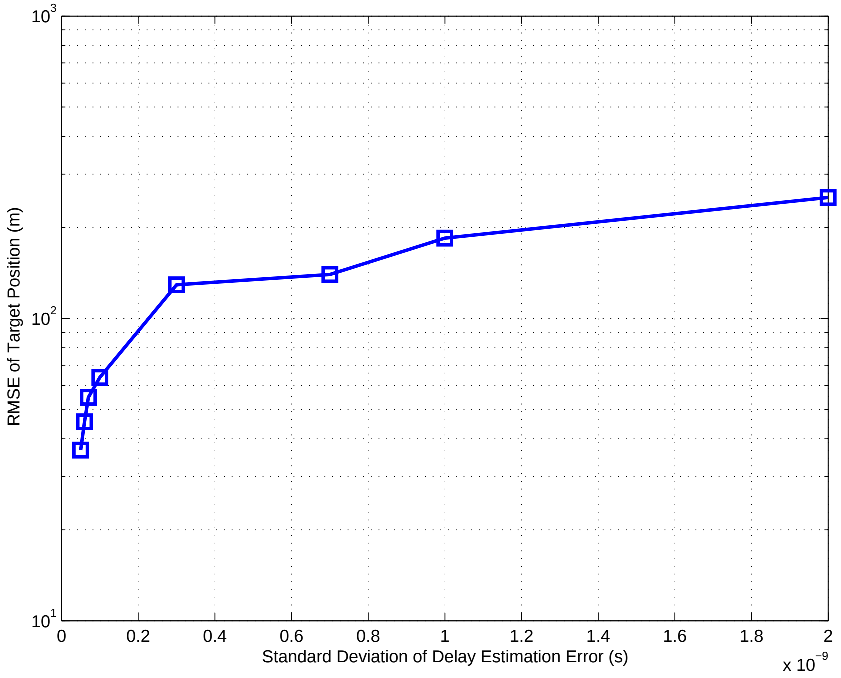


Figure 5.7: RMSE of target position (scattering case.)

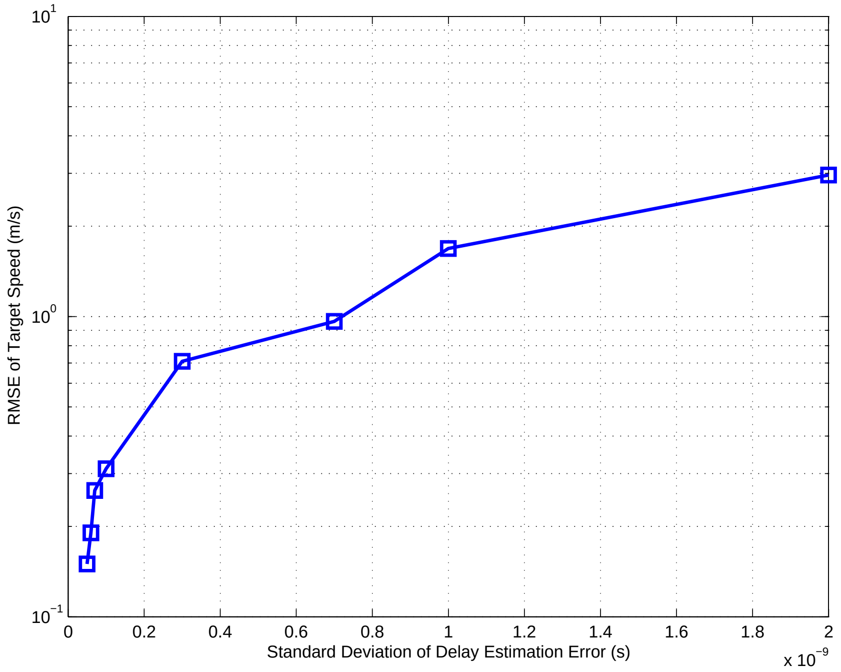


Figure 5.8: RMSE of target speed (scattering case.)

5.6 Simulations

In this section, we show the simulation results of the first and the second stage separately.

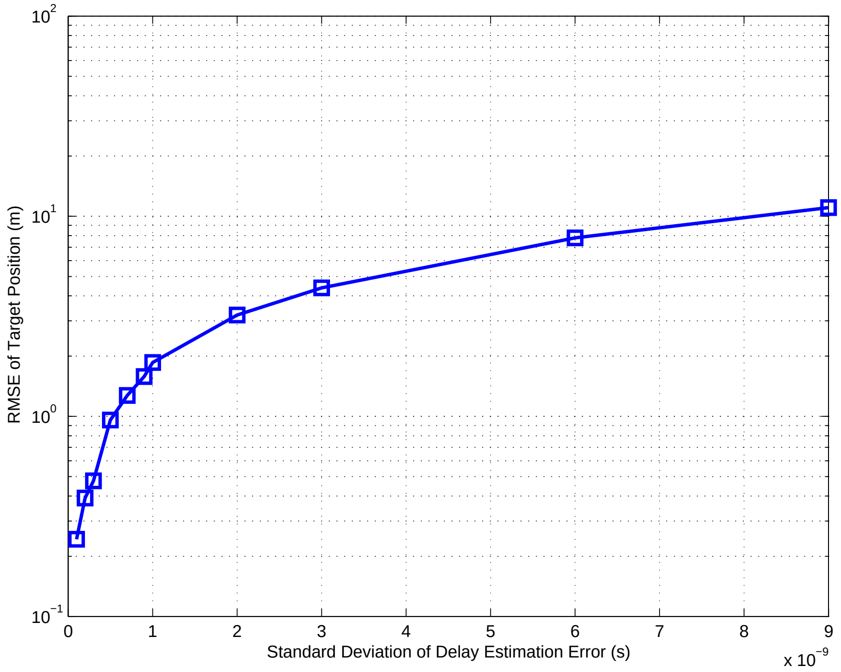


Figure 5.9: RMSE of target position (reflection case.)

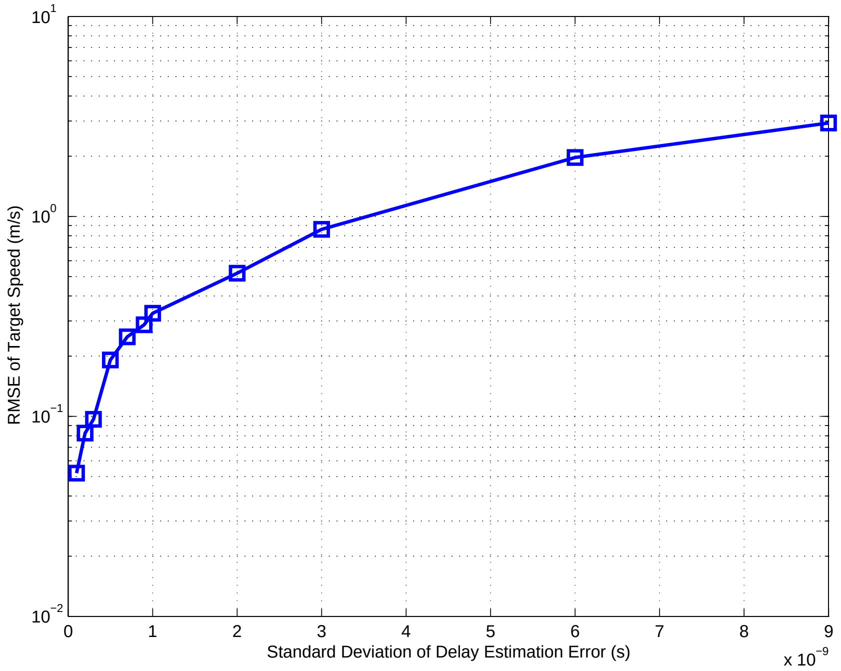


Figure 5.10: RMSE of target speed (reflection case.)

For the scattering model in the first stage, we assume that three MDs are located at $(10, 10)$, $(120, 200)$, and $(200, 20)$ in units of meters, with associated scatterers at $(30, 5)$, $(100, 195)$, and $(215, 35)$, respectively. The initial position of the target is $(116, 80)$, and the target

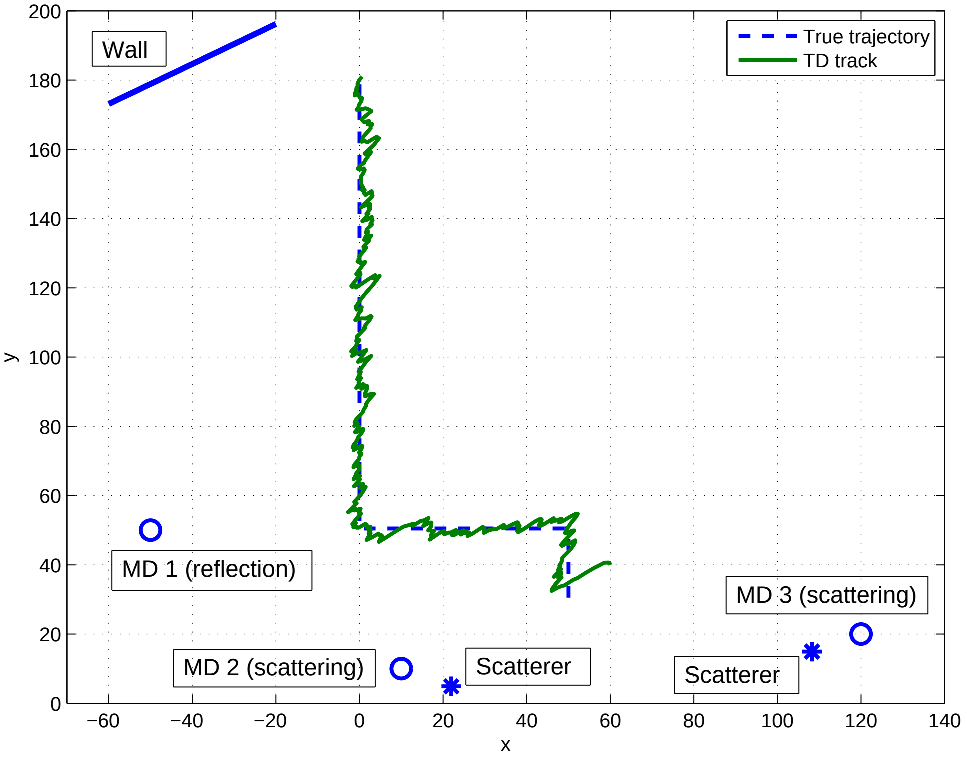


Figure 5.11: True and estimated target trajectory.

moves eastbound at the speed of 3 km/hr for 10 meters. This simulates pedestrians in indoor environments like terminal halls in airports. Except for the coordinates of the MDs, all other parameters are unknown *a priori* and need to be estimated. Figs. 5.7-5.8 show the RMS error of the target position and velocity obtained as a function of the standard deviation of the delay estimation error.

For the reflection model in the first stage, an MD is at $(10, 10)$ and a wall is located at $(100, 100)$ with an orientation angle of $30°$ with respect to the x-axis. The initial position of the target is $(116, 80)$, and its velocity is $\mathbf{v} = (0.5, 0.5)$ m/s. Two anchor points are located at $(200, 20)$, and $(300, 80)$. We assume the anchor points are MDs, so their positions are known in advance. This is again an indoor setting. Like the scattering case, all parameters are unknown and need to be estimated except the coordinates of the MDs. The RMS error of the target position and velocity are shown in Figs. 5.9-5.10.

Based on the analysis in previous sections, we know that if the measurements are error free the proposed methods can generate the exact solution to the positioning problems for both scattering and reflection cases. In the real world, the measurements are certainly contaminated with noise. Looking at the simulation results for both cases, we can see that a reasonably good performance for positioning requires sub-nanosecond accuracy of time measurement. Furthermore, for the same delay deviation the estimation accuracy in the reflection case is better compared to the scattering case. This is because in scattering models the positions of all scatterers need to be estimated, and the effect of estimation error is exacerbated in triangulation.

Fig. 5.11 shows the performance of the EKF tracker once the initial position and velocity of the target are obtained in the first stage. Besides three scatterers located at $(10, 10)$, $(120, 20)$, and $(-50, 50)$, there is also a wall lying at $(-100, 150)$ with an orientation angle of $30°$. The initial position estimate of the target is in error by around 15 m, but we see that the EKF is able to reduce the error below 1 m during periods of straight-line motion.

Chapter 6

Conclusions

This work mainly studies resource allocation problems in communication, quantization, and localization systems, and various types of applications have been considered and analyzed.

In Chapter 2, we studied the problem of downlink MIMO-OFDMA resource allocation from a game theoretic bargaining perspective. For the NBS case, we showed that the solution can be found using conventional convex optimization techniques. For the KSBS case, the problem is not directly solvable with a single convex optimization, so instead we proposed two algorithms that find the solution through a series of convex optimization steps. One of the algorithms was based on a bisection search and the other on the concept of preference functions. We also proposed a scheduling rule to find the KSBS associated with the long-term average rate. To show the effectiveness of the bargaining solutions, We studied a specific example where the users are multiplexed using a block diagonalization scheme, and with time-sharing we show how the allocation problem can be formulated as a convex optimization problem based on both the NBS and KSBS. Using simple heuristics to focus on a subset of the users on each subcarrier, a simplified algorithmic framework is also proposed, which has a polynomial complexity and is more practical for implementation in real systems.

To gain insight into the effectiveness of the application of the bargaining solutions, we simulated different resource allocation schemes for cases with both equal and unequal pathloss. As expected, the simulation results show that the bargaining solutions can systematically achieve a useful tradeoff between overall system efficiency and user fairness.

In Chapter 3, we characterized the sum rate for the Gaussian CEO problem with a scalar source having arbitrary memory. We formulated the sum-rate calculation as a variational calculus problem with a distortion constraint, and extended the conventional Lagrange method to show that if a solution exists, it should satisfy a set of Euler equations. A sufficient condition was also found that can be used to check if the necessary solution actually results in the minimal sum rate. Analysis and discussions with examples were provided to illustrate the theoretical results.

In Chapter 4, we examined STAP systems for cases involving arrays with a massive number of antennas. We studied orthogonality conditions for the spatial steering vectors for both uniform linear and rectangular arrays, and we investigated the performance of a specific low-complexity reduced-dimension separable STAP algorithm. The SINR difference between the algorithm and the fully adaptive method is analyzed, and is found to converge to zero as the number of antennas goes to infinity according to the scaling law $O(N^{-2})$. We also used random matrix theory to study the case where the covariance matrix is estimated using secondary data, and we derived the scaling law for SINR as a function of the number of training samples. Simulation results indicate the validity of the scaling laws, and effectively illustrate how the performance of the simplified separable algorithm approaches the optimal upper bound as the number of antennas grows.

In Chapter 5, we proposed a two-stage method for localizing a mobile terminal in scattering and reflection NLOS environments based on the measurements at several MDs. With the short-time constant velocity assumption in the first stage, an exact solution for the distance between the scatterer and the MD is found for a single MD-TD pair, which can be used to estimate the position and velocity of the target. A nonlinear optimization method is also suggested to improve the quality of TOA measurements. For the reflection case, a one dimensional search method is proposed and we also show that auxiliary information provided by at least two anchor points is necessary to remove ambiguity. The output of the first stage is used as the initial estimation in the second stage, where the short-time constant velocity assumption is relaxed and an EKF is used to track the subsequent movements of the target. Compared to previous studies, only uplink measurements made by the MDs are required to localize the target. Simulation results are provided to show the localization performance of the methods. Areas of future work include using methods other than least squares to overcome the effect of measurement error, and applying the two-stage idea to more complicated environments like multi-hop and multi-reflection models, though more types of measurements are possibly needed.

Besides the potential research directions of each area, some questions can be asked from a more abstract point of view. For systems with layered structures like cellular communications, inter-layer resource allocation can be a problem of interest, because changing the allocation of resources among different layers may have different impact on the overall system performance, provided that the total resources available in the system are fixed. A even more challenging question is how to do resource allocation simultaneously for systems with totally different nature. For a world where technology fusion is the trend and new emerging systems like cloud computing and internet of things are all unprecedented comprehensive and complex in history, even partial answers to the question can be of great importance.

Bibliography

[1] 3GPP 36.201 version 9.0.0. Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer; General description. *3GPP Technical Specification*, Dec. 2009.

[2] 3GPP 36.355 version 12.4.0. LTE Positioning Protocol (LPP). *3GPP Technical Specification*, March 2015.

[3] 3GPP 36.913 version 9.0.0. Requirements for further advancements for Evolved Universal Terrestrial Radio Access (E-UTRA). *3GPP Technical Report*, Dec. 2009.

[4] P. Bahl and V. N. Padmanabhan. RADAR: an in-building RF-based user location and tracking system. In *Proceedings of the 19th Annual Joint Conference of the IEEE Computer and Communications Societies*, volume 2, pages 775–784, 2000.

[5] Z. D. Bai, Y. Chen, and Y.-C. Liang. *Random Matrix Theory and Its Applications: Multivariate Statistics and Wireless Communications.* World Scientific Publishing Company, 2009.

[6] T. Berger. *Rate Distortion Theory: A Mathematical Basis for Data Compression.* Prentice-Hall, 1971.

[7] T. Berger and R. W. Yeung. Multiterminal source encoding with one distortion criterion. *IEEE Trans. Inf. Theory*, 35(2):228–236, 1989.

[8] T. Berger, Z. Zhang, and H. Viswanathan. The CEO problem. *IEEE Trans. Inf. Theory*, 42(3):887–902, 1996.

[9] H. Boche and M. Schubert. Nash bargaining and proportional fairness for wireless systems. *IEEE/ACM Transactions on Networking*, 17(5):1453 –1466, Oct. 2009.

[10] D. M. Boroson. Sample size considerations for adaptive arrays. *IEEE Trans. Aerosp. Electron. Syst.*, AES-16(4):446–451, 1980.

[11] S. Boyd and L. Vandenberghe. *Convex Optimization.* Cambridge University Press, 2004.

[12] G. Caire and S. Shamai. On the achievable throughput of a multiantenna Gaussian broadcast channel. *IEEE Trans. Inf. Theory*, 49(7):1691–1706, 2003.

[13] X. Cao. Preference functions and bargaining solutions. In *Proc. 21st IEEE Conf. Decision and Control*, volume 21, pages 164–171, 1982.

[14] P. Chan and R. Cheng. Capacity maximization for zero-forcing MIMO-OFDMA downlink systems with multiuser diversity. *Wireless Communications, IEEE Transactions on*, 6(5):1880–1889, May 2007.

[15] Y. T. Chan and K. C. Ho. A simple and efficient estimator for hyperbolic location. *Signal Processing, IEEE Transactions on*, 42(8):1905 –1915, Aug. 1994.

[16] T. K. Chee, C.-C. Lim, and J. Choi. A cooperative game theoretic framework for resource allocation in OFDMA systems. In *Proc. 10th IEEE Singapore Int. Conf. Communication systems ICCS 2006*, pages 1–5, 2006.

[17] J. Chen, F. Jiang, and A. L. Swindlehurst. The Gaussian CEO problem for a scalar source with memory: A necessary condition. In *Forty-Sixth Asilomar Conference on Signals, Systems and Computers*, Nov. 2012.

[18] J. Chen, F. Jiang, A. L. Swindlehurst, and J. A. Lopez-Salcedo. Localization of mobile equipment in radio environments with no line-of-sight path. In *Proceedings of 2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 4081–4085, 2013.

[19] J. Chen and A. L. Swindlehurst. Downlink resource allocation for multi-user MIMO-OFDMA systems: The Kalai-Smorodinsky bargaining approach. In *Proc. 3rd IEEE Int Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP) Workshop*, pages 380–383, 2009.

[20] J. Chen and A. L. Swindlehurst. On the achievable sum rate of multiterminal source coding for a correlated Gaussian vector source. In *Proc. IEEE International Conf. on Acoustics Speech and Signal Processing (ICASSP)*, March 2012.

[21] J. Chen, X. Zhang, T. Berger, and S. B. Wicker. An upper bound on the sum-rate distortion function and its corresponding rate allocation schemes for the CEO problem. *IEEE J. Sel. Areas Commun.*, 22(6):977–987, 2004.

[22] K. W. Cheung, H. C. So, W.-K. Ma, and Y. T. Chan. Least squares algorithms for time-of-arrival-based mobile location. *Signal Processing, IEEE Transactions on*, 52(4):1121 – 1130, April 2004.

[23] M. Costa. Writing on dirty paper (corresp.). *Information Theory, IEEE Transactions on*, 29(3):439–441, May 1983.

[24] R. Couillet and M. Debbah. *Random matrix methods for wireless communications.* Cambridge University Press, 2011.

[25] T. Cover and J. Thomas. *Elements of Information Theory.* Wiley Series in Telecommunications and Signal Processing. Wiley-Interscience, 2006.

[26] B. Da and C.-C. Ko. Utility-based dynamic resource allocation in multi-user MIMO-OFDMA cellular systems. In *Communications, 2009. APCC 2009. 15th Asia-Pacific Conference on*, pages 113 –117, Oct. 2009.

[27] Z. Dziong and L. G. Mason. Fair-efficient call admission control policies for broadband networks—a game theoretic framework. *IEEE/ACM Trans. Netw.*, 4(1):123–136, 1996.

[28] Federal Communications Commission. Wireless E911 location accuracy requirements. *FCC Document*, Jan. 2015.

[29] T. Flynn and R. Gray. Encoding of correlated observations. *IEEE Trans. Inf. Theory*, 33(6):773–787, 1987.

[30] H. Gazzah, P. A. Regalia, and J.-P. Delmas. Asymptotic eigenvalue distribution of block Toeplitz matrices and application to blind SIMO channel identification. *IEEE Trans. Inf. Theory*, 47(3):1243–1251, 2001.

[31] I. Gelfand and S. Fomin. *Calculus of Variations*. Dover Books on Mathematics. Dover Publications, 2000.

[32] A. Goldsmith. *Wireless Communications*. Cambridge University Press, New York, NY, USA, 2005.

[33] A. Goldsmith, S. Jafar, N. Jindal, and S. Vishwanath. Capacity limits of MIMO channels. *Selected Areas in Communications, IEEE Journal on*, 21(5):684–702, June 2003.

[34] J. S. Goldstein, I. S. Reed, and L. L. Scharf. A multistage representation of the Wiener filter based on orthogonal projections. *IEEE Trans. Inf. Theory*, 44(7):2943–2959, 1998.

[35] U. Grenander and G. Szegö. *Toeplitz Forms and Their Applications*. Chelsea Publishing Company, 1964.

[36] S. D. Greve, P. Ries, F. D. Lapierre, and J. G. Verly. Framework and taxonomy for radar space-time adaptive processing (STAP) methods. *IEEE Trans. Aerosp. Electron. Syst.*, 43(3):1084–1099, 2007.

[37] J. R. Guerci. *Space-time adaptive processing for radar*. Artech House, 2003.

[38] A. Haimovich. The eigencanceler: adaptive radar by eigenanalysis methods. *IEEE Trans. Aerosp. Electron. Syst.*, 32(2):532–542, 1996.

[39] A. M. Haimovich. Asymptotic distribution of the conditional signal-to-noise ratio in an eigenanalysis-based adaptive array. *IEEE Trans. Aerosp. Electron. Syst.*, 33(3):988–997, 1997.

[40] A. M. Haimovich and Y. Bar-Ness. An eigenanalysis interference canceler. *IEEE Trans. Signal Process.*, 39(1):76–84, 1991.

[41] J. Hallberg, M. Nilsson, and K. Synnes. Positioning with Bluetooth. In *Proceedings of the 10th International Conference on Telecommunications*, volume 2, pages 954–958, 2003.

[42] Z. Han, Z. Ji, and K. Liu. Fair multiuser channel allocation for OFDMA networks using Nash bargaining solutions and coalitions. *Communications, IEEE Transactions on*, 53(8):1366–1376, Aug. 2005.

[43] D. Harville. *Matrix Algebra From a Statistician's Perspective.* Springer, 2008.

[44] M. Hazas and A. Ward. A novel broadband ultrasonic location system. In *UbiComp 2002: Ubiquitous Computing*, pages 264–280. Springer, 2002.

[45] S.-L. Hew and L. White. Cooperative resource allocation games in shared networks: symmetric and asymmetric fair bargaining models. *Wireless Communications, IEEE Transactions on*, 7(11):4166 –4175, Nov. 2008.

[46] J. Hightower and G. Borriello. Location systems for ubiquitous computing. *Computer*, 34(8):57–66, Aug. 2001.

[47] M. L. Honig and J. S. Goldstein. Adaptive reduced-rank interference suppression based on the multistage Wiener filter. *IEEE Trans. Commun.*, 50(6):986–994, 2002.

[48] A. Ibing and H. Boche. Fairness vs. efficiency: Comparison of game theoretic criteria for OFDMA scheduling. In *Proc. Conf. Record of the Forty-First Asilomar Conf. Signals, Systems and Computers ACSSC 2007*, pages 275–279, 2007.

[49] IEEE 802.16 Broadband Wireless Access Working Group. IEEE Standard for Local and Metropolitan Area Networks Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems. *IEEE Std 802.16e-2005*, 2006.

[50] IEEE 802.16 Broadband Wireless Access Working Group. IEEE 802.16m System Requirements. *802 Standards*, Sep. 2009.

[51] M. Jiang and L. Hanzo. Multiuser MIMO-OFDM for next-generation wireless systems. *Proceedings of the IEEE*, 95(7):1430–1469, July 2007.

[52] E. A. Jorswieck and E. G. Larsson. The MISO interference channel from a game-theoretic perspective: A combination of selfishness and altruism achieves Pareto optimality. In *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing ICASSP 2008*, pages 5364–5367, 2008.

[53] T. Kailath, A. Sayed, and B. Hassibi. *Linear Estimation.* Prentice Hall, 2000.

[54] E. Kalai and M. Smorodinsky. Other solutions to Nash's bargaining problem. *Econometrica*, 43(3):513–518, 1975.

[55] S. M. Kay. *Fundamentals Of Statistical Signal Processing.* Prentice Hall, 2001.

[56] F. P. Kelly, A. K. Maulloo, and D. K. H. Tan. Rate control for communication networks: Shadow prices, proportional fairness and stability. *The Journal of the Operational Research Society*, 49(3):237–252, 1998.

[57] Y. Kochman, A. Khina, U. Erez, and R. Zamir. Rematch and forward: Joint source/channel coding for communications. In *Proc. IEEE 25th Convention of Electrical and Electronics Engineers in Israel IEEEI 2008*, pages 779–783, 2008.

[58] A. Kotanen, M. Hannikainen, H. Leppakoski, and T. D. Hamalainen. Positioning with IEEE 802.11b wireless LAN. In *the 14th IEEE Proceedings on Personal, Indoor and Mobile Radio Communications*. Institute of Electrical & Electronics Engineers (IEEE), 2003.

[59] V. K. N. Lau. Proportional fair space–time scheduling for wireless communications. *IEEE Trans. Commun.*, 53(8):1353–1360, 2005.

[60] A. Leshem and E. Zehavi. Bargaining over the interference channel. In *Proc. IEEE Int Information Theory Symp*, pages 2225–2229, 2006.

[61] A. Leshem and E. Zehavi. Cooperative game theory and the Gaussian interference channel. *IEEE J. Sel. Areas Commun.*, 26(7):1078–1088, 2008.

[62] A. Leshem and E. Zehavi. Game theory and the frequency selective interference channel. *IEEE Signal Process. Mag.*, 26(5):28–40, 2009.

[63] O. Maye, J. Schaeffner, and M. Maaser. An optical indoor positioning system for the mass market. In *Proc. of the 3rd Workshop on Positioning, Navigation and Communication*, pages 111–116, 2006.

[64] W. L. Melvin. A STAP overview. *IEEE Aero. El. Sys. Mag.*, 19(1):19–35, 2004.

[65] X. Mestre. On the asymptotic behavior of quadratic forms of the resolvent of certain covariance-type matrices. Tech. Rep. CTTC/RC/2006-01, Centre Tecnològic de Telecomunicacions de Catalunya, 2006.

[66] X. Mestre and M. A. Lagunas. Finite sample size effect on minimum variance beamformers: optimum diagonal loading factor for large arrays. *IEEE Trans. Signal Process.*, 54(1):69–82, 2006.

[67] P. D. Morris and C. R. N. Athaudage. Fairness based resource allocation for multi-user MIMO-OFDM systems. In *Proc. VTC 2006-Spring Vehicular Technology Conf. IEEE 63rd*, volume 1, pages 314–318, 2006.

[68] J. Nash. The bargaining problem. *Econometrica*, 18:155–162, Apr. 1950.

[69] M. Nokleby and A. L. Swindlehurst. Bargaining and the MISO interference channel. *EURASIP Journal on Advances in Signal Processing*, 2009:2, 2009.

[70] Y. Oohama. The rate-distortion function for the quadratic Gaussian CEO problem. *IEEE Trans. Inf. Theory*, 44(3):1057–1070, 1998.

[71] Y. Oohama. Distributed source coding of correlated Gaussian sources. *arXiv preprint arXiv:1007.4418*, 2010.

[72] D. A. Pados and G. N. Karystinos. An iterative algorithm for the computation of the MVDR filter. *IEEE Trans. Signal Process.*, 49(2):290–300, 2001.

[73] K. Pahlavan, F. O. Akgul, M. Heidari, A. Hatami, J. M. Elwell, and R. D. Tingley. Indoor geolocation in the absence of direct path. *Wireless Communications, IEEE*, 13(6):50 –58, Dec. 2006.

[74] A. Pandya, A. Kansal, G. J. Pottie, and M. B. Srivastava. Fidelity and resource sensitive data gathering. In *42nd Allerton Conference*, June 2004.

[75] P. Parker and A. L. Swindlehurst. Space-time autoregressive filtering for matched subspace STAP. *IEEE Trans. Aerosp. Electron. Syst.*, 39(2):510–520, 2003.

[76] B. W. Parkinson and J. J. Spilker. *Progress In Astronautics and Aeronautics: Global Positioning System: Theory and Applications.* AIAA, 1996.

[77] N. Patwari, A. O. Hero III, M. Perkins, N. S. Correal, and R. J. O'Dea. Relative location estimation in wireless sensor networks. *Signal Processing, IEEE Transactions on*, 51(8):2137 – 2148, Aug. 2003.

[78] C. D. Peckham, A. M. Haimovich, T. F. Ayoub, J. S. Goldstein, and I. S. Reid. Reduced-rank STAP performance analysis. *IEEE Trans. Aerosp. Electron. Syst.*, 36(2):664–676, 2000.

[79] C. B. Peel, B. M. Hochwald, and A. L. Swindlehurst. A vector-perturbation technique for near-capacity multiantenna multiuser communication-part I: channel inversion and regularization. *IEEE Trans. Commun.*, 53(1):195–202, 2005.

[80] V. Prabhakaran, D. Tse, and K. Ramachandran. Rate region of the quadratic Gaussian CEO problem. In *Proc. Int. Symp. Information Theory ISIT 2004*, 2004.

[81] S. S. Pradhan and K. Ramchandran. Generalized coset codes for distributed binning. *IEEE Trans. Inf. Theory*, 51(10):3457–3474, 2005.

[82] I. S. Reed, J. D. Mallett, and L. E. Brennan. Rapid convergence rate in adaptive arrays. *IEEE Trans. Aerosp. Electron. Syst.*, AES-10(6):853–863, 1974.

[83] W. Rhee and J. M. Cioffi. Increase in capacity of multiuser OFDM system using dynamic subchannel allocation. In *Proc. IEEE 51st VTC 2000-Spring Tokyo Vehicular Technology*, volume 2, pages 1085–1089, 2000.

[84] C. D. Richmond. Derived PDF of maximum likelihood signal estimator which employs an estimated noise covariance. *IEEE Trans. Signal Process.*, 44(2):305–315, 1996.

[85] J. Roman, M. Rangaswamy, D. Davis, Q. Zhang, B. Himed, and J. Michels. Parametric adaptive matched filter for airborne radar applications. *IEEE Trans. Aerosp. Electron. Syst.*, 36(2):677–692, 2000.

[86] A. Roth. *Axiomatic models of bargaining.* Springer-Verlag, 1979.

[87] F. Rubio. *Generalized consistent estimation in arbitrarily high dimensional signal processing.* PhD thesis, Universitat Politècnica de Catalunya, Barcelona, 2008.

[88] F. Rubio and X. Mestre. Spectral convergence for a general class of random matrices. *Statistics & Probability Letters*, 81(5):592–602, May 2011.

[89] A. H. Sayed, A. Tarighat, and N. Khajehnouri. Network-based wireless location: challenges faced in developing techniques for accurate wireless location information. *IEEE Signal Process. Mag.*, 22(4):24–40, 2005.

[90] G. A. F. Seber. *A Matrix Handbook for Statisticians.* John Wiley & Sons, Inc., Nov. 2007.

[91] C. K. Seow and S. Y. Tan. Non-line-of-sight localization in multipath environments. *IEEE Trans. Mobile Comput.*, 7(5):647–660, 2008.

[92] S. Shi, M. Schubert, and H. Boche. Rate optimization for multiuser MIMO systems with linear processing. *Signal Processing, IEEE Transactions on*, 56(8):4020–4030, Aug. 2008.

[93] J. W. Silverstein. Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices. *Journal of Multivariate Analysis*, 55(2):331–339, 1995.

[94] J. W. Silverstein and Z. D. Bai. On the empirical distribution of eigenvalues of a class of large dimensional random matrices. *Journal of Multivariate analysis*, 54(2):175–192, 1995.

[95] R. Soundararajan and S. Vishwanath. Multi-terminal source coding through a relay. In *Proc. (ISIT) Symp. IEEE Int Information Theory*, pages 1866–1870, 2011.

[96] Q. Spencer, C. Peel, A. Swindlehurst, and M. Haardt. An introduction to the multi-user MIMO downlink. *Communications Magazine, IEEE*, 42(10):60–67, Oct. 2004.

[97] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt. Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels. *Signal Processing, IEEE Transactions on*, 52(2), Feb. 2004.

[98] M. Stojnic, H. Vikalo, and B. Hassibi. Rate maximization in multi-antenna broadcast channels with linear preprocessing. In *Proc. IEEE Global Telecommunications Conf. GLOBECOM '04*, volume 6, pages 3957–3961, 2004.

[99] J. Suris, L. Dasilva, Z. Han, A. Mackenzie, and R. Komali. Asymptotic optimality for distributed spectrum sharing using bargaining solutions. *Wireless Communications, IEEE Transactions on*, 8(10):5225 –5237, Oct. 2009.

[100] B. Tang, J. Tang, and Y. Peng. Performance of knowledge aided space time adaptive processing. *IET Radar, Sonar & Navigation*, 5(3):331, 2011.

[101] B. Tang, J. Tang, and Y. Peng. Clutter nulling performance of SMI in amplitude heterogeneous clutter environments. *IEEE Trans. Aerosp. Electron. Syst.*, 49(2):1366–1373, 2013.

[102] S. Tavildar and P. Viswanath. On the sum-rate of the vector Gaussian CEO problem. In *Proc. Conf Signals, Systems and Computers Record of the Thirty-Ninth Asilomar Conf*, pages 3–7, 2005.

[103] P. Tejera, W. Utschick, G. Bauch, and J. A. Nossek. Subchannel allocation in multiuser multiple-input-multiple-output systems. *Information Theory, IEEE Transactions on*, 52(10), Oct. 2006.

[104] A. J. Tenenbaum and R. S. Adve. Improved sum-rate optimization in the multiuser MIMO downlink. In *Proc. 42nd Annual Conf. Information Sciences and Systems CISS 2008*, pages 984–989, 2008.

[105] A. M. Tulino and S. Verdú. *Random matrix theory and wireless communications*, volume 1. Now Publishers Inc, 2004.

[106] S. Tung. *Multiterminal Rate-Distortion Theory.* PhD thesis, Cornell University, School of Elec. Eng., Ithaca, NY, 1978.

[107] J. Wang and D. Katabi. Dude, where's my card?: RFID positioning that works with multipath and non-line of sight. In *Proceedings of the ACM SIGCOMM 2013 Conference on SIGCOMM*, SIGCOMM '13, pages 51–62, New York, NY, USA, 2013. ACM.

[108] X. Wang, Z. Wang, and B. O'Dea. A TOA-based location algorithm reducing the errors due to non-line-of-sight (NLOS) propagation. *Vehicular Technology, IEEE Transactions on*, 52(1):112 – 116, Jan. 2003.

[109] X. Wang and X.-D. Zhang. Linear transmission for rate optimization in MIMO broadcast channels. *IEEE Trans. Wireless Commun.*, 9(10):3247–3257, 2010.

[110] J. Ward. Space-time adaptive processing for airborne radar. Technical Report 1015, MIT Lincoln Laboratory, Lexington, MA, 1994.

[111] J. Ward. Space-time adaptive processing for airborne radar. In *International Conference on Acoustics, Speech, and Signal Processing*, volume 5, pages 2809–2812, 1995.

[112] H. Weingarten, Y. Steinberg, and S. Shamai. The capacity region of the Gaussian MIMO broadcast channel. In *Proc. Int. Symp. Information Theory ISIT 2004*, 2004.

[113] J.-J. Xiao and Z.-Q. Luo. Optimal rate allocation for the vector Gaussian CEO problem. In *Proc. 1st IEEE Int Computational Advances in Multi-Sensor Adaptive Processing Workshop*, pages 56–59, 2005.

[114] J.-J. Xiao and Z.-Q. Luo. Multiterminal source–channel communication over an orthogonal multiple-access channel. *IEEE Trans. Inf. Theory*, 53(9):3255–3264, 2007.

[115] A. Yamaguchi, M. Ogawa, T. Tamura, and T. Togawa. Monitoring behavior in the home using positioning sensors. In *Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, volume 4, pages 1977–1979, 1998.

[116] H. Yamamoto and K. Itoh. Source coding theory for multiterminal communication systems with a remote source. *Trans. IECE Jpn.*, E63, no. 10:700–706, Oct. 1980.

[117] Y. Yang, V. Stankovic, Z. Xiong, and W. Zhao. Asymmetric code design for remote multiterminal source coding. In *Proc. Data Compression Conf DCC 2004*, 2004.

[118] Y. Yang, V. Stankovic, Z. Xiong, and W. Zhao. On multiterminal source code design. *IEEE Trans. Inf. Theory*, 54(5):2278–2302, 2008.

[119] Y. Yang, Y. Zhang, and Z. Xiong. The generalized quadratic Gaussian CEO problem: New cases with tight rate region and applications. In *Proc. (ISIT) Symp. IEEE Int Information Theory*, pages 21–25, 2010.

[120] E. Zehavi and A. Leshem. Alternative bargaining solutions for the interference channel. In *Proc. 3rd IEEE Int Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP) Workshop*, pages 9–12, 2009.

[121] G. Zhang and H. Zhang. Adapative resource allocation for downlink OFDMA networks using cooperative game theory. In *Proc. 11th IEEE Singapore Int. Conf. Communication Systems ICCS 2008*, pages 98–103, 2008.

[122] Z. Zhang, J. Shi, H.-H. Chen, M. Guizani, and P. Qiu. A cooperation strategy based on Nash bargaining solution in cooperative relay networks. *Vehicular Technology, IEEE Transactions on*, 57(4):2570 –2577, July 2008.

[123] Y. Zhao. Standardization of mobile phone positioning for 3G systems. *IEEE Commun. Mag.*, 40(7):108–116, 2002.

Appendix A

Proofs

A.1 Proof of Convergence for Theorem 2.1

Let $\{R_k^{G*}(\beta)\}_{k=1}^K$ denote the optimal rate allocation for a given β. We need to show that $\{R_k^{G*}(\beta)\}_{k=1}^K$ actually converges to the KSBS as $\beta \to 1$.

Proof. Let $r_k = \frac{R_k^G}{R_k^{max}}$ and substitute it into (2.10) to obtain

$$f(\{r_k\}_{k=1}^K, \beta) = \sum_{k=1}^{K} \log \left[r_k + \frac{\beta}{K-1} \sum_{s=1, s \neq k}^{K} (1 - r_k) \right] = U. \tag{A.1}$$

Let $r_k^*(\beta)$ denote user k's rate when $f(\{r_k\}_{k=1}^K, \beta)$ achieves its optimum. Because the optimum is unique for $\beta \neq 1$, there exists a K-dimensional hyperplane containing the point $\{r_k^*(\beta)\}_{k=1}^K$ that is tangent to the rate region boundary. The hyperplane's equation is

$$\sum_{k=1}^{K} \frac{\partial f}{\partial r_k}\Big|_{r_k = r_k^*(\beta)} (r_k - r_k^*(\beta)) = 0. \tag{A.2}$$

The derivatives can be directly calculated as

$$\begin{aligned}\frac{\partial f}{\partial r_k} =& \frac{1}{r_k + \frac{\beta}{K-1}\sum_{s=1,s\neq k}^{K}(1-r_s)} + \\ & \sum_{i=1,i\neq k}^{K} \frac{-\frac{\beta}{K-1}}{r_i + \frac{\beta}{K-1}\sum_{s=1,s\neq i}^{K}(1-r_s)} .\end{aligned} \tag{A.3}$$

Now we calculate the intersection point of the tangent hyperplane and the line segment from the origin to the utopia point. Because $r_k = \frac{R_k^G}{R_k^{max}}$ is actually a normalized rate ratio, the components of any point on the latter line should all be equal, *i.e.*, $r_i = r_j, \forall i, j$. We use r to represent this value, and thus the intersection point should satisfy

$$\sum_{k=1}^{K} \frac{\partial f}{\partial r_k}\Big|_{r_k = r_k^*(\beta)} (r - r_k^*(\beta)) = 0, \tag{A.4}$$

and from (A.4) we can get a closed-form expression for r:

$$r = \frac{\sum_{k=1}^{K} \frac{\partial f}{\partial r_k}\Big|_{r_k = r_k^*(\beta)} r_k^*(\beta)}{\sum_{k=1}^{K} \frac{\partial f}{\partial r_k}\Big|_{r_k = r_k^*(\beta)}}. \tag{A.5}$$

Since we assume the problem domain is convex and r is on the tangent hyperplane, it resides outside the rate region. If we let r_0 represent the intersection point of the rate region boundary and the line segment from the origin to the utopia point, we can see that $0 \leq r_0 \leq r$. Substituting (A.5) into this inequality, after some mathematical manipulations we have

$$\begin{aligned}(1-\beta) r_0 \leq & \sum_{k=1}^{K} \left[\frac{r_k^*(\beta)}{r_k^*(\beta) + \frac{\beta}{K-1}\sum_{s=1,s\neq k}^{K}(1-r_s^*(\beta))} \right. \\ & \left. + \sum_{i=1,i\neq k}^{K} \frac{-\frac{\beta}{K-1} r_k^*(\beta)}{r_i^*(\beta) + \frac{\beta}{K-1}\sum_{s=1,s\neq i}^{K}(1-r_s^*(\beta))} \right] \\ & \times \frac{1}{\sum_{k=1}^{K} \frac{1}{r_k^*(\beta) + \frac{\beta}{K-1}\sum_{s=1,s\neq k}^{K}(1-r_s^*(\beta))}}.\end{aligned} \tag{A.6}$$

Let $r^* = r^*(\beta)|_{\beta=1}$. As $\beta \rightarrow 1$, the RHS of (A.6) becomes

$$\begin{aligned}
&\sum_{k=1}^{K}\left[\frac{r_k^*}{r_k^* + \frac{1}{K-1}\sum_{s=1,s\neq k}^{K}(1-r_s^*)}\right.\\
&\left.+\sum_{i=1,i\neq k}^{K}\frac{-\frac{1}{K-1}r_k^*}{r_i^* + \frac{1}{K-1}\sum_{s=1,s\neq i}^{K}(1-r_s^*)}\right]\\
&\times\frac{1}{\sum_{k=1}^{K}\frac{1}{r_k^*+\frac{1}{K-1}\sum_{s=1,s\neq k}^{K}(1-r_s^*)}} && \text{(A.7)}\\
&=\frac{K}{\sum_{k=1}^{K}\frac{1}{r_k^*+\frac{1}{K-1}\sum_{s=1,s\neq k}^{K}(1-r_s^*)}}-1 && \text{(A.8)}\\
&\leq\frac{\sum_{k=1}^{K}\left[r_k^*+\frac{1}{K-1}\sum_{s=1,s\neq k}^{K}(1-r_s^*)\right]}{K}-1=0, && \text{(A.9)}
\end{aligned}$$

where from (A.8) to (A.9) we have used the inequality

$$\frac{n}{\sum_{i=1}^{n}\frac{1}{x_i}} \leq \frac{\sum_{i=1}^{n}x_i}{n}\ .$$

In short we have

$$0 \leq (1-\beta)r_0 \leq \frac{K}{\sum_{k=1}^{K}\frac{1}{r_k^*+\frac{1}{K-1}\sum_{s=1,s\neq k}^{K}(1-r_s^*)}}-1 \leq 0. \qquad \text{(A.10)}$$

This means $\sum_{k=1}^{K}\frac{1}{r_k^*+\frac{1}{K-1}\sum_{s=1,s\neq k}^{K}(1-r_s^*)} = K$, where the equality is achievable only when $r_i^* = r_j^*, \forall i, j$. In other words, r_0 and $r^*(\beta)$ coincides when $\beta = 1$. Therefore we have shown that the optimum $r^*(\beta)$ converges to the KSBS as $\beta \rightarrow 1$. □

A.2 Convexity of the Achievable Rate Region

In this appendix, we show that the K-user rate region for the MIMO-OFDMA problem based on BD is convex. As shown in Section 2.5, every subcarrier is time-shared by all users.

For the fixed-power time division scheme, it is well known that the closure of the convex hull of all rate tuples is achievable, in which case the convexity of the rate region is obvious. But in our case we assume a variable-power time division use of the subcarriers, so the convex hull argument does not apply. Here we provide a proof that guarantees the convexity of the rate region.

Theorem A.1. *The achievable rate region of (2.22) is convex.*

Proof. Let $\mathbf{R}^{(1)} = (R_1^{(1)}, R_2^{(1)}, \ldots, R_K^{(1)})$ and $\mathbf{R}^{(2)} = (R_1^{(2)}, R_2^{(2)}, \ldots, R_K^{(2)})$ be two points in the rate region. According to the definition of a convex set, we need to prove $\theta\mathbf{R}^{(1)} + (1-\theta)\mathbf{R}^{(2)}$ is also a point in the rate region for all $\theta \in [0, 1]$. We can write the following:

$$\begin{aligned}
&\theta\mathbf{R}^{(1)} + (1-\theta)\mathbf{R}^{(2)} \\
=&\Bigg(\theta \sum_{n=1}^{N}\sum_{i=1}^{I}\omega_{n,i}^{(1)}\sum_{t=1}^{n_R^{(1)}}\log_2\left(1+\frac{\sigma_{1,n,i,t}^2\lambda_{1,n,i,t}^{(1)}}{\omega_{n,i}^{(1)}N_{1,n}}\right)+ \\
&(1-\theta)\sum_{n=1}^{N}\sum_{i=1}^{I}\omega_{n,i}^{(2)}\sum_{t=1}^{n_R^{(1)}}\log_2\left(1+\frac{\sigma_{1,n,i,t}^2\lambda_{1,n,i,t}^{(2)}}{\omega_{n,i}^{(2)}N_{1,n}}\right), \\
&\qquad\qquad\qquad \ldots, \\
&\theta \sum_{n=1}^{N}\sum_{i=1}^{I}\omega_{n,i}^{(1)}\sum_{t=1}^{n_R^{(K)}}\log_2\left(1+\frac{\sigma_{K,n,i,t}^2\lambda_{K,n,i,t}^{(1)}}{\omega_{n,i}^{(1)}N_{K,n}}\right)+ \\
&(1-\theta)\sum_{n=1}^{N}\sum_{i=1}^{I}\omega_{n,i}^{(2)}\sum_{t=1}^{n_R^{(K)}}\log_2\left(1+\frac{\sigma_{K,n,i,t}^2\lambda_{K,n,i,t}^{(2)}}{\omega_{n,i}^{(2)}N_{K,n}}\right)\Bigg).
\end{aligned} \tag{A.11}$$

Let $\omega_{n,i}^{(\theta)} = \theta\omega_{n,i}^{(1)} + (1-\theta)\omega_{n,i}^{(2)}$ and

$$\lambda_{k,n,i,t}^{(\theta)} = \Bigg[\left(1 + \frac{\sigma_{k,n,i,t}^2 \lambda_{k,n,i,t}^{(1)}}{\omega_{n,i}^{(1)} N_{k,n}}\right)^{\frac{\theta\omega_{n,i}^{(1)}}{\omega_{n,i}^{(\theta)}}} \times$$

$$\left(1 + \frac{\sigma_{k,n,i,t}^2 \lambda_{k,n,i,t}^{(2)}}{\omega_{n,i}^{(2)} N_{k,n}}\right)^{\frac{(1-\theta)\omega_{n,i}^{(2)}}{\omega_{n,i}^{(\theta)}}} \frac{\omega_{n,i}^{(\theta)} N_{k,n}}{\sigma_{k,n,i,t}^2} - \frac{\omega_{n,i}^{(\theta)} N_{k,n}}{\sigma_{k,n,i,t}^2} \Bigg].$$

Here $\omega_{n,i}^{(\theta)}$ and $\lambda_{k,n,i,t}^{(\theta)}$ serve as the new time and power allocation. Then (A.11) can be further written as

$$\begin{aligned} &\theta\mathbf{R}^{(1)} + (1-\theta)\mathbf{R}^{(2)} \\ =&\Bigg(\sum_{n=1}^{N}\sum_{i=1}^{I} \omega_{n,i}^{(\theta)} \sum_{t=1}^{n_R^{(1)}} \log_2\left\{1 + \lambda_{1,n,i,t}^{(\theta)} \frac{\sigma_{1,n,i,t}^2}{\omega_{n,i}^{(\theta)} N_{1,n}}\right\}, \\ &\qquad\qquad \ldots, \\ &\sum_{n=1}^{N}\sum_{i=1}^{I} \omega_{n,i}^{(\theta)} \sum_{t=1}^{n_R^{(K)}} \log_2\left\{1 + \lambda_{K,n,i,t}^{(\theta)} \frac{\sigma_{K,n,i,t}^2}{\omega_{n,i}^{(\theta)} N_{K,n}}\right\}\Bigg). \end{aligned} \tag{A.12}$$

Since $\sum_{n=1}^{N}\sum_{i=1}^{I}\omega_{n,i} \leq 1$, it is easy to show that $\sum_{n=1}^{N}\sum_{i=1}^{I}\omega_{n,i}^{(\theta)} \leq 1$, which means $\omega_{n,i}^{(\theta)}$ is also a valid time allocation. Now we need to prove that the new power allocation also satisfies the power constraint. We start the proof by calculating the sum of the allocated powers:

$$
\begin{aligned}
&\sum_{n=1}^{N}\sum_{i=1}^{I}\sum_{k\in\varphi_{n,i}}\sum_{t=1}^{n_R^{(k)}}\lambda_{k,n,i,t}^{(\theta)}\\
=&\sum_{n=1}^{N}\sum_{i=1}^{I}\sum_{k\in\varphi_{n,i}}\sum_{t=1}^{n_R^{(k)}}\left[\left(1+\frac{\sigma_{k,n,i,t}^2\lambda_{k,n,i,t}^{(1)}}{\omega_{n,i}^{(1)}N_{k,n}}\right)^{\frac{\theta\omega_{n,i}^{(1)}}{\omega_{n,i}^{(\theta)}}}\times\right.\\
&\left.\left(1+\frac{\sigma_{k,n,i,t}^2\lambda_{k,n,i,t}^{(2)}}{\omega_{n,i}^{(2)}N_{k,n}}\right)^{\frac{(1-\theta)\omega_{n,i}^{(2)}}{\omega_{n,i}^{(\theta)}}}\frac{\omega_{n,i}^{(\theta)}N_{k,n}}{\sigma_{k,n,i,t}^2}-\frac{\omega_{n,i}^{(\theta)}N_{k,n}}{\sigma_{k,n,i,t}^2}\right] &\text{(A.13)}\\
\leq&\sum_{n=1}^{N}\sum_{i=1}^{I}\sum_{k\in\varphi_{n,i}}\sum_{t=1}^{n_R^{(k)}}\left\{\left[\frac{\theta\omega_{n,i}^{(1)}}{\omega_{n,i}^{(\theta)}}\left(1+\frac{\sigma_{k,n,i,t}^2\lambda_{k,n,i,t}^{(1)}}{\omega_{n,i}^{(1)}N_{k,n}}\right)+\right.\right.\\
&\left.\left.\frac{(1-\theta)\omega_{n,i}^{(2)}}{\omega_{n,i}^{(\theta)}}\left(1+\frac{\sigma_{k,n,i,t}^2\lambda_{k,n,i,t}^{(2)}}{\omega_{n,i}^{(2)}N_{k,n}}\right)\right]\frac{\omega_{n,i}^{(\theta)}N_{k,n}}{\sigma_{k,n,i,t}^2}-\frac{\omega_{n,i}^{(\theta)}N_{k,n}}{\sigma_{k,n,i,t}^2}\right\} &\text{(A.14)}\\
=&\sum_{n=1}^{N}\sum_{i=1}^{I}\sum_{k\in\varphi_{n,i}}\sum_{t=1}^{n_R^{(k)}}\left[\theta\lambda_{k,n,i,t}^{(1)}+(1-\theta)\lambda_{k,n,i,t}^{(2)}\right] &\text{(A.15)}\\
\leq&\theta P+(1-\theta)P=P, &\text{(A.16)}
\end{aligned}
$$

where (A.14) follows from the fact that $x^{\theta}y^{1-\theta}\leq\theta x+(1-\theta)y$ if $x,y>0$ and $\theta\in[0,1]$.

In conclusion, $\theta\mathbf{R}^{(1)}+(1-\theta)\mathbf{R}^{(2)}$ is also a point in the rate region and hence the region is convex. □

A.3 Proof of Theorem 3.1

Proof. Since $u_i\geq\psi_i, i=1,\ldots,n$, there exist real functions $z_i, i=1,\ldots,n$ such that we can write $u_i=\psi_i+z_i^2$ for all i. Substituting $\psi_i+z_i^2$ for u_i into the functional optimization problem in Theorem 3.1, we end up with

$$\min_{z_1,\ldots,z_n} \int_a^b F(\omega, z_1, \ldots, z_n, z'_1, \ldots, z'_n) d\omega \tag{A.17}$$

$$\text{s.t.} \int_a^b G(\omega, z_1, \ldots, z_n, z'_1, \ldots, z'_n) d\omega = D \ , \tag{A.18}$$

where

$$F(\omega, z_1, \ldots, z_n, z'_1, \ldots, z'_n) = f(\omega, u_1, \ldots, u_n, u'_1, \ldots, u'_n)|_{u_i=\psi_i+z_i^2} \tag{A.19}$$

and

$$G(\omega, z_1, \ldots, z_n, z'_1, \ldots, z'_n) = g(\omega, u_1, \ldots, u_n, u'_1, \ldots, u'_n)|_{u_i=\psi_i+z_i^2}. \tag{A.20}$$

It is clear that solving this problem for z_i is equivalent to finding the optimal u_i.

The recast problem is an isoperimetric problem, which can be solved by using the method of Lagrange multipliers. Letting λ denote the multiplier, we have the Lagrangian as

$$\int_a^b F(\omega, z_1, \ldots, z_n, z'_1, \ldots, z'_n) d\omega + \lambda \int_a^b G(\omega, z_1, \ldots, z_n, z'_1, \ldots, z'_n) d\omega \ . \tag{A.21}$$

Calculating the derivative of the Lagrangian with respect to z_i for all i and setting them to zero yields

$$F_{z_i} - \frac{d}{d\omega} F_{z'_i} + \lambda \left(G_{z_i} - \frac{d}{d\omega} G_{z'_i} \right) = 0. \tag{A.22}$$

Since

$$F_{z_i} = 2 z_i f_{u_i} + 2 z'_i f_{u'_i} \tag{A.23}$$

and

$$F_{z_i'} = 2z_i f_{u_i'}, \tag{A.24}$$

equation (A.22) becomes

$$z_i \left(f_{u_i} - \frac{d}{d\omega} f_{u_i'} + \lambda \left(g_{u_i} - \frac{d}{d\omega} g_{u_i'} \right) \right) = 0, \tag{A.25}$$

which means either $f_{u_i} - \frac{d}{d\omega} f_{u_i'} + \lambda \left(g_{u_i} - \frac{d}{d\omega} g_{u_i'} \right) = 0$ or $z_i = 0$. If $z_i = 0$, u_i should be equal to ψ_i. Otherwise it should be the solution to

$$f_{u_i} - \frac{d}{d\omega} f_{u_i'} + \lambda \left(g_{u_i} - \frac{d}{d\omega} g_{u_i'} \right) = 0, i = 1, \ldots, n, \tag{A.26}$$

which are exactly the Euler equations. This concludes the proof. □

A.4 Proof of Lemma 3.1

We shall prove Lemma 3.1 by calculating the second variation of (3.25). As indicated in [31], if the second variation is strongly positive in a neighborhood of $(u_1, \ldots, u_n)$, then (3.25) has a local minimum at $(u_1, \ldots, u_n)$.

Proof. All admissible $u_i, i = 1, \ldots, n$ must satisfy the constraint

$$\int_a^b g(\omega, u_1, \ldots, u_n) d\omega = D.$$

Utilizing this fact, we can rewrite (3.25) as

$$\begin{aligned}
&J(u_1, \ldots, u_n) \\
&= \int_a^b f(\omega, u_1, \ldots, u_n) d\omega \\
&= \int_a^b \left[f(\omega, u_1, \ldots, u_n) + \lambda g(\omega, u_1, \ldots, u_n) \right] d\omega - \lambda D, \qquad \text{(A.27)}
\end{aligned}$$

where λ is the associated Lagrange multiplier. Let

$$\zeta(\omega, u_1, \ldots, u_n) = f(\omega, u_1, \ldots, u_n) + \lambda g(\omega, u_1, \ldots, u_n), \qquad \text{(A.28)}$$

so that (A.27) can be written as

$$J(u_1, \ldots, u_n) = \int_a^b \zeta(\omega, u_1, \ldots, u_n) d\omega - \lambda D. \qquad \text{(A.29)}$$

Suppose we give $(u_1, \ldots, u_n)$ an admissible increment $(h_1, \ldots, h_n)$, where the selection of $(h_1, \ldots, h_n)$ is arbitrary as long as at least one of the h_i is not identically zero in ω. We want to calculate the second variation of $J(u_1 + h_1, \ldots, u_n + h_n)$, which can be done using Taylor's formula. After a series of mathematical simplifications, the second variation of J can be written as

$$\delta^2 J(u_1, \ldots, u_n) = \frac{1}{2} \int_a^b \mathbf{h}^T \mathbf{Z}(u_1, \ldots, u_n) \mathbf{h} \; d\omega, \qquad \text{(A.30)}$$

where

$$\mathbf{h} = \begin{bmatrix} h_1 \\ \vdots \\ h_n \end{bmatrix} \qquad \text{(A.31)}$$

and

$$\mathbf{Z}(u_1,\ldots,u_n)=\begin{bmatrix}\frac{\partial^2\zeta}{\partial u_1^2} & \cdots & \frac{\partial^2\zeta}{\partial u_1\partial u_n}\\ \vdots & \ddots & \vdots\\ \frac{\partial^2\zeta}{\partial u_n\partial u_1} & \cdots & \frac{\partial^2\zeta}{\partial u_n^2}\end{bmatrix}. \tag{A.32}$$

The integrand of the second variation is a quadratic form in $(h_1,\ldots,h_n)$. Since the selection of $(h_1,\ldots,h_n)$ is arbitrary, the requirement that $\delta^2 J(u_1,\ldots,u_n)$ is strongly positive is equivalent to requiring that $\mathbf{Z}(u_1,\ldots,u_n)$ is positive-definite. So the sufficient condition is

$$\mathbf{Z}(u_1,\ldots,u_n)\succ 0. \tag{A.33}$$

Due to the fact that (3.25) does not contain the derivatives of $u_i, i=1,\ldots,n$, in this sufficiency proof we do not require $h_i, i=1,\ldots,n$ to be continuously differentiable. So the local minimum guaranteed by Lemma 3.1 is not only a weak minimum but also a strong minimum [31].

□

A.5 Proof of Theorem 3.2

Proof. First we appeal to the mathematical technique used in the proof of Theorem 3.1 to recast the problem. Since $u_i\geq\psi_i, i=1,\ldots,n$, there exist real functions $z_i, i=1,\ldots,n$ such that we can write $u_i=\psi_i+z_i^2$ for all i. Substituting $\psi_i+z_i^2$ for u_i into the functional optimization problem in Theorem 3.2, we end up with

$$\min_{z_1,\dots,z_n} \int_a^b F(\omega, z_1, \dots, z_n) d\omega \tag{A.34}$$

$$\text{s.t.} \int_a^b G(\omega, z_1, \dots, z_n) d\omega = D, \tag{A.35}$$

where

$$F(\omega, z_1, \dots, z_n) = f(\omega, u_1, \dots, u_n)|_{u_i = \psi_i + z_i^2}, \tag{A.36}$$

and

$$G(\omega, z_1, \dots, z_n) = g(\omega, u_1, \dots, u_n)|_{u_i = \psi_i + z_i^2}. \tag{A.37}$$

Applying Lemma 3.1 to this recast problem, we know the sufficient condition is

$$\mathbf{Z}(z_1, \dots, z_n) \succ 0, \tag{A.38}$$

where

$$\mathbf{Z}(z_1, \dots, z_n) = \begin{bmatrix} \frac{\partial^2 \zeta}{\partial z_1^2} & \cdots & \frac{\partial^2 \zeta}{\partial z_1 \partial z_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 \zeta}{\partial z_n \partial z_1} & \cdots & \frac{\partial^2 \zeta}{\partial z_n^2} \end{bmatrix} \tag{A.39}$$

$$\zeta(\omega, z_1, \dots, z_n) = F(\omega, z_1, \dots, z_n) + \lambda G(\omega, z_1, \dots, z_n), \tag{A.40}$$

and λ is the associated Lagrange multiplier.

It is easy to show that for $i = 1, \ldots, n$ the first-order derivatives are

$$\frac{\partial \zeta}{\partial z_i} = 2z_i f_{u_i} + 2\lambda z_i g_{u_i}, \tag{A.41}$$

and the second-order derivatives are

$$\frac{\partial^2 \zeta}{\partial z_i^2} = 2f_{u_i} + 4z_i^2 f_{u_i u_i} + 2\lambda g_{u_i} + 4\lambda z_i^2 g_{u_i u_i} \tag{A.42}$$

$$\frac{\partial^2 \zeta}{\partial z_i \partial z_j} = 4z_i z_j f_{u_i u_j} + 4\lambda z_i z_j g_{u_i u_j}. \tag{A.43}$$

Plugging (A.42) and (A.43) back into (A.39), we get

$$\mathbf{Z}(z_1, \ldots, z_n) =$$
$$\begin{bmatrix} 2f_{u_1} + 4z_1^2 f_{u_1 u_1} + 2\lambda g_{u_1} + 4\lambda z_1^2 g_{u_1 u_1} & \cdots & 4z_1 z_n f_{u_1 u_n} + 4\lambda z_1 z_n g_{u_1 u_n} \\ \vdots & \ddots & \vdots \\ 4z_n z_1 f_{u_n u_1} + 4\lambda z_n z_1 g_{u_n u_1} & \cdots & 2f_{u_n} + 4z_n^2 f_{u_n u_n} + 2\lambda g_{u_n} + 4\lambda z_n^2 g_{u_n u_n} \end{bmatrix} \tag{A.44}$$

Based on Lemma 3.1, we know if (A.44) is positive-definite then $(u_1, \ldots, u_n)$ is guaranteed to be a local minimum. Since multiplication of a matrix by a non-negative constant does not change the sign of its eigenvalues, it is equivalent to require (3.28) to be positive-definite.

In addition, from Theorem 3.1 we know

$$z_i \left(f_{u_i} + \lambda g_{u_i} \right) = 0, \tag{A.45}$$

which means either $z_i = 0$ or $f_{u_i} + \lambda g_{u_i} = 0$. In the latter case, we can let z_i be either $(u_i - \psi_i)^{1/2}$ or $-(u_i - \psi_i)^{1/2}$ due to the fact $u_i = \psi_i + z_i^2$. This completes the proof.

□

A.6 Proof of Theorem 4.1

Proof. We follow the proofs in [24, 65, 94, 88, 93], though the details are somewhat different. First, we rewrite the left hand side of (4.38) as

$$\begin{aligned}&\left|\mathbf{a}^H\left((x(z)\mathbf{P}-z\mathbf{I})^{-1}-(\mathbf{A}-z\mathbf{I})^{-1}\right)\mathbf{a}\right|\\ =&\left|\mathbf{a}^H(x(z)\mathbf{P}-z\mathbf{I})^{-1}(\mathbf{A}-x(z)\mathbf{P})(\mathbf{A}-z\mathbf{I})^{-1}\mathbf{a}\right|. \qquad (\text{A.46})\end{aligned}$$

Also define $\mathbf{A}_{(k)}=\mathbf{A}-\frac{1}{K}\mathbf{U}^H\mathbf{x}_k\mathbf{x}_k^H\mathbf{U}$. Using Equation 2.1 in [93], we obtain

$$\begin{aligned}\mathbf{x}_k^H\mathbf{U}(\mathbf{A}-z\mathbf{I})^{-1}&=\mathbf{x}_k^H\mathbf{U}\left(\mathbf{A}_{(k)}-z\mathbf{I}+\frac{1}{K}\mathbf{U}^H\mathbf{x}_k\mathbf{x}_k^H\mathbf{U}\right)^{-1}\\ &=\frac{\mathbf{x}_k^H\mathbf{U}\left(\mathbf{A}_{(k)}-z\mathbf{I}\right)^{-1}}{1+\frac{1}{K}\mathbf{x}_k^H\mathbf{U}\left(\mathbf{A}_{(k)}-z\mathbf{I}\right)^{-1}\mathbf{U}^H\mathbf{x}_k}. \qquad (\text{A.47})\end{aligned}$$

Based on (A.47), the product of $\mathbf{A}$ and $(\mathbf{A}-z\mathbf{I})^{-1}$ can be decomposed as

$$\begin{aligned}\mathbf{A}(\mathbf{A}-z\mathbf{I})^{-1}&=\frac{1}{K}\sum_{k=1}^{K}\mathbf{U}^H\mathbf{x}_k\mathbf{x}_k^H\mathbf{U}(\mathbf{A}-z\mathbf{I})^{-1}\\ &=\frac{1}{K}\sum_{k=1}^{K}\frac{\mathbf{U}^H\mathbf{x}_k\mathbf{x}_k^H\mathbf{U}\left(\mathbf{A}_{(k)}-z\mathbf{I}\right)^{-1}}{1+\frac{1}{K}\mathbf{x}_k^H\mathbf{U}\left(\mathbf{A}_{(k)}-z\mathbf{I}\right)^{-1}\mathbf{U}^H\mathbf{x}_k}.\end{aligned}$$

Using the Sherman-Morrison formula, we know that

$$(\mathbf{A}-z\mathbf{I})^{-1}=\left(\mathbf{A}_{(k)}-z\mathbf{I}\right)^{-1}-\frac{\frac{1}{K}\left(\mathbf{A}_{(k)}-z\mathbf{I}\right)^{-1}\mathbf{U}^H\mathbf{x}_k\mathbf{x}_k^H\mathbf{U}\left(\mathbf{A}_{(k)}-z\mathbf{I}\right)^{-1}}{1+\frac{1}{K}\mathbf{x}_k^H\mathbf{U}\left(\mathbf{A}_{(k)}-z\mathbf{I}\right)^{-1}\mathbf{U}^H\mathbf{x}_k},$$

so the following equation holds

$$\left(\mathbf{A}_{(k)}-z\mathbf{I}\right)^{-1}-(\mathbf{A}-z\mathbf{I})^{-1}=\frac{\frac{1}{K}\left(\mathbf{A}_{(k)}-z\mathbf{I}\right)^{-1}\mathbf{U}^H\mathbf{x}_k\mathbf{x}_k^H\mathbf{U}\left(\mathbf{A}_{(k)}-z\mathbf{I}\right)^{-1}}{1+\frac{1}{K}\mathbf{x}_k^H\mathbf{U}\left(\mathbf{A}_{(k)}-z\mathbf{I}\right)^{-1}\mathbf{U}^H\mathbf{x}_k}. \qquad (\text{A.48})$$

With the help of these auxiliary equations, equation (A.46) can be rewritten as (A.49).

$$\left| \underbrace{\frac{1}{K}\sum_{k=1}^{K}\frac{\mathbf{a}^H\left(x(z)\mathbf{P}-z\mathbf{I}\right)^{-1}\left(\mathbf{U}^H\mathbf{x}_k\mathbf{x}_k^H\mathbf{U}\left(\mathbf{A}_{(k)}-z\mathbf{I}\right)^{-1}-\mathbf{P}\left(\mathbf{A}_{(k)}-z\mathbf{I}\right)^{-1}\right)\mathbf{a}}{1+\frac{1}{K}\mathbf{x}_k^H\mathbf{U}\left(\mathbf{A}_{(k)}-z\mathbf{I}\right)^{-1}\mathbf{U}^H\mathbf{x}_k}}_{\text{(i)}} \right.$$
$$+\underbrace{\frac{1}{K}\sum_{k=1}^{K}\frac{\mathbf{a}^H\left(x(z)\mathbf{P}-z\mathbf{I}\right)^{-1}\left(\mathbf{P}\left(\mathbf{A}_{(k)}-z\mathbf{I}\right)^{-1}-\mathbf{P}\left(\mathbf{A}-z\mathbf{I}\right)^{-1}\right)\mathbf{a}}{1+\frac{1}{K}\mathbf{x}_k^H\mathbf{U}\left(\mathbf{A}_{(k)}-z\mathbf{I}\right)^{-1}\mathbf{U}^H\mathbf{x}_k}}_{\text{(ii)}}$$
$$\left. +\underbrace{\mathbf{a}^H\left(x(z)\mathbf{P}-z\mathbf{I}\right)^{-1}\left(\frac{1}{K}\sum_{k=1}^{K}\frac{1}{1+\frac{1}{K}\mathbf{x}_k^H\mathbf{U}\left(\mathbf{A}_{(k)}-z\mathbf{I}\right)^{-1}\mathbf{U}^H\mathbf{x}_k}-x(z)\right)\mathbf{P}\left(\mathbf{A}-z\mathbf{I}\right)^{-1}\mathbf{a}}_{\text{(iii)}}\right|. \tag{A.49}$$

If the absolute values of (i), (ii) and (iii) converge to 0, then the validity of Theorem 4.1 can be proved using the triangle inequality. We will prove the convergence for the three terms individually below.

Convergence of (i)

The whitened version of the kth column of $\mathbf{X}$ is $\mathbf{y}_k = \mathbf{\Sigma}^{-\frac{1}{2}}\mathbf{x}_k$. Thus matrix $\mathbf{Y} = [\mathbf{y}_1 \ \ldots \ \mathbf{y}_K]$ can be written as $\mathbf{Y} = \mathbf{\Sigma}^{-\frac{1}{2}}\mathbf{X}$. Therefore, the numerator in (i) can be rewritten as a function of $\mathbf{y}_k$, *i.e.*,

$$\begin{aligned}
&\mathbf{a}^H\left(x(z)\mathbf{P}-z\mathbf{I}\right)^{-1}\left(\mathbf{U}^H\mathbf{x}_k\mathbf{x}_k^H\mathbf{U}\left(\mathbf{A}_{(k)}-z\mathbf{I}\right)^{-1}-\mathbf{P}\left(\mathbf{A}_{(k)}-z\mathbf{I}\right)^{-1}\right)\mathbf{a}\\
&=\mathbf{a}^H\left(x(z)\mathbf{P}-z\mathbf{I}\right)^{-1}\left(\mathbf{U}^H\mathbf{\Sigma}^{\frac{1}{2}}\mathbf{y}_k\mathbf{y}_k^H\mathbf{\Sigma}^{\frac{1}{2}}\mathbf{U}\left(\mathbf{A}_{(k)}-z\mathbf{I}\right)^{-1}-\mathbf{P}\left(\mathbf{A}_{(k)}-z\mathbf{I}\right)^{-1}\right)\mathbf{a}\\
&=\mathbf{a}^H\left(x(z)\mathbf{P}-z\mathbf{I}\right)^{-1}\mathbf{U}^H\mathbf{\Sigma}^{\frac{1}{2}}\mathbf{y}_k\mathbf{y}_k^H\mathbf{\Sigma}^{\frac{1}{2}}\mathbf{U}\left(\mathbf{A}_{(k)}-z\mathbf{I}\right)^{-1}\mathbf{a}-\mathbf{a}^H\left(x(z)\mathbf{P}-z\mathbf{I}\right)^{-1}\mathbf{P}\left(\mathbf{A}_{(k)}-z\mathbf{I}\right)^{-1}\mathbf{a}\\
&\overset{(a)}{=}\mathbf{y}_k^H\mathbf{\Sigma}^{\frac{1}{2}}\mathbf{U}\left(\mathbf{A}_{(k)}-z\mathbf{I}\right)^{-1}\mathbf{a}\mathbf{a}^H\left(x(z)\mathbf{P}-z\mathbf{I}\right)^{-1}\mathbf{U}^H\mathbf{\Sigma}^{\frac{1}{2}}\mathbf{y}_k-\mathbf{a}^H\left(x(z)\mathbf{P}-z\mathbf{I}\right)^{-1}\mathbf{P}\left(\mathbf{A}_{(k)}-z\mathbf{I}\right)^{-1}\mathbf{a},
\end{aligned} \tag{A.50}$$

where (a) follows from the fact that $\mathbf{tr}\,(AB) = \mathbf{tr}\,(BA)$.

We then apply Lemma 4 in [88], which says that as $K \to \infty$, almost surely

$$\left| \frac{1}{K} \sum_{k=1}^{K} \left(\mathbf{y}_k^H \mathbf{C}_k \mathbf{y}_k - \mathbf{tr}\,(\mathbf{C}_k) \right) \right| \longrightarrow 0, \tag{A.51}$$

if the entries of $\mathbf{y}_k$ are i.i.d. and have zero mean and unit variance, and $\mathbf{C}_k$ does not depend on $\mathbf{y}_k$. Define

$$\mathbf{C}_k = \frac{\mathbf{\Sigma}^{\frac{1}{2}} \mathbf{U} \left(\mathbf{A}_{(k)} - z\mathbf{I} \right)^{-1} \mathbf{a}\mathbf{a}^H \left(x(z)\mathbf{P} - z\mathbf{I} \right)^{-1} \mathbf{U}^H \mathbf{\Sigma}^{\frac{1}{2}}}{1 + \frac{1}{K} \mathbf{x}_k^H \mathbf{U} \left(\mathbf{A}_{(k)} - z\mathbf{I} \right)^{-1} \mathbf{U}^H \mathbf{x}_k}, \tag{A.52}$$

and note that

$$\mathbf{tr}\,(\mathbf{C}_k) = \frac{\mathbf{a}^H \left(x(z)\mathbf{P} - z\mathbf{I} \right)^{-1} \mathbf{P} \left(\mathbf{A}_{(k)} - z\mathbf{I} \right)^{-1} \mathbf{a}}{1 + \frac{1}{K} \mathbf{x}_k^H \mathbf{U} \left(\mathbf{A}_{(k)} - z\mathbf{I} \right)^{-1} \mathbf{U}^H \mathbf{x}_k}. \tag{A.53}$$

Based on (A.50) it can be observed that the addend in (i) for k is equal to $\mathbf{y}_k^H \mathbf{C}_k \mathbf{y}_k - \mathbf{tr}\,(\mathbf{C}_k)$, so we can apply Lemma 4 and substituting (A.52) and (A.53) into (A.51) results in the convergence of (i).

Convergence of (ii)

Appealing to Lemma A.2 in [87], we just need to show that for any given z the following inequality holds:

$$\max_{1 \le k \le K} E \left\{ \left| \frac{\mathbf{a}^H \left(x(z)\mathbf{P} - z\mathbf{I} \right)^{-1} \mathbf{P} \left(\left(\mathbf{A}_{(k)} - z\mathbf{I} \right)^{-1} - \left(\mathbf{A} - z\mathbf{I} \right)^{-1} \right) \mathbf{a}}{1 + \frac{1}{K} \mathbf{x}_k^H \mathbf{U} \left(\mathbf{A}_{(k)} - z\mathbf{I} \right)^{-1} \mathbf{U}^H \mathbf{x}_k} \right|^p \right\} \le \frac{C}{K^{1+\delta}},$$

where $\delta > 0$, $p \ge 1$ and C are constants. Based on (A.48), it is equivalent to show that the expectation of

$$\left| \frac{\mathbf{a}^H \left(x(z)\mathbf{P} - z\mathbf{I}\right)^{-1} \mathbf{P} \left(\frac{\frac{1}{K}\left(\mathbf{A}_{(k)} - z\mathbf{I}\right)^{-1} \mathbf{U}^H \mathbf{x}_k \mathbf{x}_k^H \mathbf{U} \left(\mathbf{A}_{(k)} - z\mathbf{I}\right)^{-1}}{1 + \frac{1}{K}\mathbf{x}_k^H \mathbf{U}\left(\mathbf{A}_{(k)} - z\mathbf{I}\right)^{-1} \mathbf{U}^H \mathbf{x}_k} \right) \mathbf{a}}{1 + \frac{1}{K}\mathbf{x}_k^H \mathbf{U}\left(\mathbf{A}_{(k)} - z\mathbf{I}\right)^{-1} \mathbf{U}^H \mathbf{x}_k} \right|^p$$

is bounded by $\frac{C}{K^{1+\delta}}$ for all k. From Corollary 3.2 in [24], we know

$$\left| \frac{1}{1 + \frac{1}{K}\mathbf{x}_k^H \mathbf{U}\left(\mathbf{A}_{(k)} - z\mathbf{I}\right)^{-1} \mathbf{U}^H \mathbf{x}_k} \right|$$

is always bounded by $|z|/\mathbf{Im}\left\{z\right\}$. So we just need to prove the expectation of

$$\left| \mathbf{a}^H \left(x(z)\mathbf{P} - z\mathbf{I}\right)^{-1} \mathbf{P} \left(\mathbf{A}_{(k)} - z\mathbf{I}\right)^{-1} \mathbf{U}^H \mathbf{x}_k \mathbf{x}_k^H \mathbf{U} \left(\mathbf{A}_{(k)} - z\mathbf{I}\right)^{-1} \mathbf{a} \right|^p \tag{A.54}$$

is bounded.

For (A.54), we have

$$\begin{aligned}
&\left| \mathbf{a}^H \left(x(z)\mathbf{P} - z\mathbf{I}\right)^{-1} \mathbf{P} \left(\mathbf{A}_{(k)} - z\mathbf{I}\right)^{-1} \mathbf{U}^H \mathbf{x}_k \mathbf{x}_k^H \mathbf{U} \left(\mathbf{A}_{(k)} - z\mathbf{I}\right)^{-1} \mathbf{a} \right| \\
\overset{(a)}{\leq}& \left\|\mathbf{a}\right\|^2 \left\| \left(x(z)\mathbf{P} - z\mathbf{I}\right)^{-1} \mathbf{P} \left(\mathbf{A}_{(k)} - z\mathbf{I}\right)^{-1} \mathbf{U}^H \mathbf{x}_k \mathbf{x}_k^H \mathbf{U} \left(\mathbf{A}_{(k)} - z\mathbf{I}\right)^{-1} \right\|_F \\
\overset{(b)}{\leq}& \left\|\mathbf{a}\right\|^2 \left\| \left(x(z)\mathbf{P} - z\mathbf{I}\right)^{-1} \mathbf{P} \right\|_F \left\| \left(\mathbf{A}_{(k)} - z\mathbf{I}\right)^{-1} \right\|_F^2 \left\| \mathbf{U}^H \mathbf{x}_k \mathbf{x}_k^H \mathbf{U} \right\|_F \\
\overset{(c)}{\leq}& \left\|\mathbf{a}\right\|^2 \left\| \left(x(z)\mathbf{P} - z\mathbf{I}\right)^{-1} \mathbf{P} \right\|_F \left\| \left(\mathbf{A}_{(k)} - z\mathbf{I}\right)^{-1} \right\|_F^2 \left| \mathbf{x}_k^H \mathbf{U} \mathbf{U}^H \mathbf{x}_k \right| \\
\overset{(d)}{\leq}& \frac{1}{\left|\mathbf{Im}\left\{z\right\}\right|^2} \left\|\mathbf{a}\right\|^2 \left\| \left(x(z)\mathbf{P} - z\mathbf{I}\right)^{-1} \mathbf{P} \right\|_F \left| \mathbf{x}_k^H \mathbf{U} \mathbf{U}^H \mathbf{x}_k \right|,
\end{aligned}$$

where (a) follows from Cauchy-Schwarz inequality, (b) follows from the submultiplicative property of Frobenius norm, (c) follows from the definition of Frobenius norm, and (d) follows from the fact that $\left\| \left(\mathbf{A}_{(k)} - z\mathbf{I}\right)^{-1} \right\|_F \leq \frac{1}{\left|\mathbf{Im}\left\{z\right\}\right|}$. So the expectation of (A.54) is bounded by

$$\frac{1}{|\mathbf{Im}\{z\}|^2}\|\mathbf{a}\|^2\left\|(x(z)\mathbf{P}-z\mathbf{I})^{-1}\mathbf{P}\right\|_F E\left\{\left|\mathbf{x}_k^H\mathbf{U}\mathbf{U}^H\mathbf{x}_k\right|\right\},$$

which is a limited value due to the fact that the rank of $\mathbf{U}\mathbf{U}^H$ is fixed to be M and the empirical distribution function of its eigenvalues converges.

Expression for $m(z)$

In this step, we want to find $m(z)$. The Stieltjes transform of random matrix $\mathbf{A}$ is

$$m_{F_{\mathbf{A}}^K}(z)=\frac{1}{M}\mathbf{tr}\left(\left[\frac{1}{K}\mathbf{U}^H\mathbf{X}\mathbf{X}^H\mathbf{U}-z\mathbf{I}\right]^{-1}\right).$$

In order to find $m(z)$, we start with the Stieltjes transform of $\mathbf{B}=\frac{1}{K}\mathbf{X}\mathbf{X}^H\mathbf{U}\mathbf{U}^H$, which is

$$m_{F_{\mathbf{B}}^K}(z)=\frac{1}{L}\mathbf{tr}\left(\left[\frac{1}{K}\mathbf{X}\mathbf{X}^H\mathbf{U}\mathbf{U}^H-z\mathbf{I}\right]^{-1}\right).$$

Let $\lambda_n, n=1,\ldots,L$ be the eigenvalues of $\Sigma^{\frac{1}{2}}\mathbf{U}\mathbf{U}^H\boldsymbol{\Sigma}^{\frac{1}{2}}$ and $c=L/K$. The main theorem of [93] leads to $\left|m_{F_{\mathbf{B}}^K}(z)-m_b(z)\right|\to 0$ when $L,K\to\infty$ and c converges to a constant, and $m_b(z)$ is the unique solution to

$$m_b(z)=\frac{1}{L}\sum_{n=1}^{L}\frac{1}{\lambda_n\left(1-c-czm_b(z)\right)-z}. \tag{A.55}$$

Note that we can rewrite (A.55) as

$$m_b(z)=\frac{1}{L}\left\{\sum_{n=1}^{M}\frac{1}{\lambda_n\left(1-c-czm_b(z)\right)-z}+\frac{L-M}{-z}\right\}, \tag{A.56}$$

due to the fact that matrix $\mathbf{U}\mathbf{U}^H$ has $L-M$ zero eigenvalues.

On the other hand, Lemma 3.1 in [24] shows the following relationship between $m_{F^K_{\mathbf{A}}}(z)$ and $m_{F^K_{\mathbf{B}}}(z)$:

$$\frac{M}{L}m_{F^K_{\mathbf{A}}}(z) = m_{F^K_{\mathbf{B}}}(z) + \frac{L-M}{L}\frac{1}{z}. \tag{A.57}$$

Combining (A.56) and (A.57), we obtain

$$\left| m_{F^K_{\mathbf{A}}}(z) - \frac{1}{M}\left\{\sum_{n=1}^{M}\frac{1}{\lambda_n\left(1-c-czm_b(z)\right)-z}\right\}\right| \longrightarrow 0. \tag{A.58}$$

Define

$$m(z) = \frac{L}{M}m_b(z) + \frac{L-M}{M}\frac{1}{z}, \tag{A.59}$$

and thus we have

$$\begin{aligned}
&\frac{1}{M}\left\{\sum_{n=1}^{M}\frac{1}{\lambda_n\left(1-c-czm_b(z)\right)-z}\right\}\\
=&\frac{1}{M}\left\{\sum_{n=1}^{M}\frac{1}{\lambda_n\left(1-c-cz\left(\frac{M}{L}m(z)-\frac{L-M}{L}\frac{1}{z}\right)\right)-z}\right\}\\
=&\frac{1}{M}\left\{\sum_{n=1}^{M}\frac{1}{\lambda_n\left(1-\frac{M}{K}-\frac{M}{K}zm(z)\right)-z}\right\}.
\end{aligned} \tag{A.60}$$

From (A.56), (A.58), (A.59), and (A.60), we can conclude that

$$\left| m_{F^K_{\mathbf{A}}}(z) - m(z)\right| \longrightarrow 0,$$

and $m(z)$ is the unique solution to

$$m(z) = \frac{1}{M}\left\{\sum_{n=1}^{M}\frac{1}{\lambda_n\left(1-\frac{M}{K}-\frac{M}{K}zm(z)\right)-z}\right\}.$$

Note that the non-zero $\lambda_n, n = 1, \dots, M$ are also the eigenvalues of $\mathbf{U}^H\mathbf{\Sigma}\mathbf{U}$.

Convergence of (iii)

The proof of the convergence of (iii) is similar to those in [24, 65]. For completeness we briefly write the procedure here. For (iii) we want to show

$$\left|\frac{1}{K}\sum_{k=1}^{K}\frac{1}{1+\frac{1}{K}\mathbf{x}_k^H\mathbf{U}\left(\mathbf{A}_{(k)}-z\mathbf{I}\right)^{-1}\mathbf{U}^H\mathbf{x}_k}-x(z)\right| \longrightarrow 0. \tag{A.61}$$

We introduce $\mathbf{Z} = \frac{1}{\sqrt{K}}\mathbf{U}^H\mathbf{X}$, such that $\mathbf{A} = \mathbf{Z}\mathbf{Z}^H$. Matrix $\mathbf{Z}$ can be rewritten as

$$\mathbf{Z} = \begin{bmatrix}\mathbf{z}_1 & \mathbf{Z}_{(1)}\end{bmatrix},$$

where $\mathbf{z}_1$ is the first column of $\mathbf{Z}$, and $\mathbf{Z}_{(1)}$ refers to the remaining part of $\mathbf{Z}$. So we have

$$\frac{1}{K}\sum_{k=1}^{K}\frac{1}{1+\frac{1}{K}\mathbf{x}_k^H\mathbf{U}\left(\mathbf{A}_{(k)}-z\mathbf{I}\right)^{-1}\mathbf{U}^H\mathbf{x}_k} = \frac{1}{K}\sum_{k=1}^{K}\frac{1}{1+\mathbf{z}_k^H\left(\mathbf{Z}_{(k)}\mathbf{Z}_{(k)}^H-z\mathbf{I}\right)^{-1}\mathbf{z}_k}.$$

The proof of the equality in (A.61) can be made by calculating the inverse of $\left(\mathbf{Z}^H\mathbf{Z}-z\mathbf{I}\right)$ twice in different ways and then equating them. On one hand, the inverse can be written as

$$\left(\mathbf{Z}^H\mathbf{Z}-z\mathbf{I}\right)^{-1} = \left(\begin{bmatrix}\mathbf{z}_1^H \\ \mathbf{Z}_{(1)}^H\end{bmatrix}\begin{bmatrix}\mathbf{z}_1 & \mathbf{Z}_{(1)}\end{bmatrix}-z\mathbf{I}\right)^{-1} = \begin{bmatrix}\mathbf{z}_1^H\mathbf{z}_1-z & \mathbf{z}_1^H\mathbf{Z}_{(1)} \\ \mathbf{Z}_{(1)}^H\mathbf{z}_1 & \mathbf{Z}_{(1)}^H\mathbf{Z}_{(1)}-z\mathbf{I}\end{bmatrix}^{-1},$$

so the top left element of $\left(\mathbf{Z}^H\mathbf{Z} - z\mathbf{I}\right)^{-1}$ is

$$\begin{aligned}
&\left(\mathbf{z}_1^H\mathbf{z}_1 - z - \mathbf{z}_1^H\mathbf{Z}_{(1)}\left(\mathbf{Z}_{(1)}^H\mathbf{Z}_{(1)} - z\mathbf{I}\right)^{-1}\mathbf{Z}_{(1)}^H\mathbf{z}_1\right)^{-1} \\
&= \left(-z + \mathbf{z}_1^H\left(\mathbf{I} - \frac{-1}{z}\mathbf{Z}_{(1)}\left(\frac{-1}{z}\mathbf{Z}_{(1)}^H\mathbf{Z}_{(1)} + \mathbf{I}\right)^{-1}\mathbf{Z}_{(1)}^H\right)\mathbf{z}_1\right)^{-1} \\
&\overset{(a)}{=} \left(-z - z\mathbf{z}_1^H\left(\mathbf{Z}_{(1)}\mathbf{Z}_{(1)}^H - z\mathbf{I}\right)^{-1}\mathbf{z}_1\right)^{-1} \\
&= -\frac{1}{z}\frac{1}{1 + \mathbf{z}_1^H\left(\mathbf{Z}_{(1)}\mathbf{Z}_{(1)}^H - z\mathbf{I}\right)^{-1}\mathbf{z}_1},
\end{aligned} \tag{A.62}$$

where (a) follows from the matrix inversion lemma. By rearranging the columns and rows of $\left(\mathbf{Z}^H\mathbf{Z} - z\mathbf{I}\right)^{-1}$, the conclusion in (A.62) can be extended to the kth diagonal element; *i.e.*, the kth diagonal element of $\left(\mathbf{Z}^H\mathbf{Z} - z\mathbf{I}\right)^{-1}$ is equal to

$$-\frac{1}{z}\frac{1}{1 + \mathbf{z}_k^H\left(\mathbf{Z}_{(k)}\mathbf{Z}_{(k)}^H - z\mathbf{I}\right)^{-1}\mathbf{z}_k}. \tag{A.63}$$

On the other hand, the inverse of $\left(\mathbf{Z}^H\mathbf{Z} - z\mathbf{I}\right)$ can be computed in an alternate way. By directly applying the matrix inversion lemma we obtain

$$\left(\mathbf{Z}^H\mathbf{Z} - z\mathbf{I}\right)^{-1} = -\frac{1}{z}\left(\mathbf{I} - \mathbf{Z}^H\left(\mathbf{Z}\mathbf{Z}^H - z\mathbf{I}\right)^{-1}\mathbf{Z}\right). \tag{A.64}$$

Clearly, the kth diagonal element of (A.64) is

$$-\frac{1}{z}\left(1 - \mathbf{z}_k^H\left(\mathbf{Z}\mathbf{Z}^H - z\mathbf{I}\right)^{-1}\mathbf{z}_k\right). \tag{A.65}$$

By equating (A.63) and (A.65), we have

$$\frac{1}{1 + \mathbf{z}_k^H\left(\mathbf{Z}_{(k)}\mathbf{Z}_{(k)}^H - z\mathbf{I}\right)^{-1}\mathbf{z}_k} = 1 - \mathbf{z}_k^H\left(\mathbf{Z}\mathbf{Z}^H - z\mathbf{I}\right)^{-1}\mathbf{z}_k\,. \tag{A.66}$$

Thus, the trace of $-\frac{1}{K}z\left(\mathbf{Z}^H\mathbf{Z}-z\mathbf{I}\right)^{-1}$ is equal to

$$\begin{aligned}\frac{1}{K}\sum_{k=1}^{K}\frac{1}{1+\mathbf{z}_k^H\left(\mathbf{Z}_{(k)}\mathbf{Z}_{(k)}^H-z\mathbf{I}\right)^{-1}\mathbf{z}_k} =& \frac{1}{K}\sum_{k=1}^{K}\left(1-\mathbf{z}_k^H\left(\mathbf{Z}\mathbf{Z}^H-z\mathbf{I}\right)^{-1}\mathbf{z}_k\right)\\ =&1-\frac{1}{K}\mathbf{tr}\left(\left(\mathbf{A}-z\mathbf{I}\right)^{-1}\mathbf{A}\right)\\ =&1-\frac{1}{K}\mathbf{tr}\left(\left(\mathbf{A}-z\mathbf{I}\right)^{-1}\left(\mathbf{A}-z\mathbf{I}+z\mathbf{I}\right)\right)\\ =&1-\frac{M}{K}-\frac{1}{K}\mathbf{tr}\left(\left(\mathbf{A}-z\mathbf{I}\right)^{-1}\left(z\mathbf{I}\right)\right)\\ =&1-\frac{M}{K}-\frac{M}{K}zm_{F_{\mathbf{A}}^K}(z),\end{aligned}$$

which converges to $x(z)$ due to the result $\left|m_{F_{\mathbf{A}}^K}(z)-m(z)\right|\to 0$ obtained previously.

□