Title
Physical Layer Optimization for Wireless Sensing and Network Connectivity

Author
Jiang, Feng

Publication Date
2015

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,
IRVINE

Physical Layer Optimization for Wireless Sensing and Network Connectivity

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Electrical and Computer Engineering

by

Feng Jiang

Dissertation Committee:
Professor A. Lee Swindlehurst, Chair
Professor Ender Ayanoglu
Professor Athina Markopoulou

2015

DEDICATION

to my parents

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

There is an old Chinese saying "Fixed barrack, floating soldiers". In the recent quarters, there are more and more new graduate students beginning to work in the fourth floor of Engineering Hall, and this reminds me that my leaving time is closer. When I first met with Prof. A. Lee Swindlehust in Sep. 2008, I said "Irvine is a quiet and clean place", and Prof. Swindlehurst replied "But someone complains it's too quiet". Time flies fast, and until now I have studied and worked in this quiet campus for almost seven years. With Prof. Swindlehurst's recommendation, I obtained the opportunity to enroll in the Phd program of UCI and I truly believe that without Prof. Swindlehurst's supervision and guidelines, all the results in this dissertation will not be achieved. First of all, I would like to express my deepest gratitude to Prof. Swindlehurst.

I have attended several graduate courses at UCI, and what impressed me most are the Digital Communication taught by Prof. Ender Ayanoglu and Network Coding given by Prof. Athina Markopoulou. I would like to extend my sincere thanks and appreciation to Prof. Ayanoglu and Prof. Markopoulou for the teaching and for serving in my dissertation committee. Also, I would like to thank Prof. Kenneth Mease and Prof. Animashree Anandkumar for serving in my qualify exam committee.

My research work has been funded by the National Science Foundation CCF-0916073 and Air Force Office of Scientific Research grant FA9550-10-1-0310, and I also get support from Broadcom Corporation during my internships in the Broadband, WPAN and WLAN groups. These financial supports are greatly appreciated. Additionally, I have received a lot of help, advices and encouragement from my colleagues and friends, especially in the early stage of my Phd study. I would like to take this opportunity to thank Chiufai Wong, Jie Chen, Shunchi Wu, Xinjie Yang, Sean O'Rourke, Yuchen Yao, Jianqi Wang, Chenwei Wang, Liangbin Li and Chun Meng.

In the last I want to thank my parents for the continuous support and patience over the past years, and this work is dedicated to them.

CURRICULUM VITAE

Feng Jiang

EDUCATION

University of California at Irvine, Irvine, CA
Ph.D. in ELectrical and Computer Engineering, Sep. 2008-May 2015
Advisor: Prof. A. Lee Swindlehurst

Beijing University of Posts and Telecommunications, Beijing, China
M.E in Communication and Information System, Sep. 2005-Mar. 2008

Beijing University of Posts and Telecommunications, Beijing, China
B. E in Communication Engineering, Sep. 2000-Jun. 2004

WORK EXPERIENCE

Broadcom Corporation
Intern at WLAN Group, San Diego, CA Sep. 2014-Mar. 2015
Intern at WPAN Group, Irvine, CA Feb. 2014-Jun. 2014
Intern at Broadband Group, Irvine, CA Oct. 2012-Jun. 2013

University of California at Irvine
Research Assistant at EECS Dept. Sep. 2008-Feb. 2014

PUBLICATIONS

1. Feng Jiang, Jianqi Wang and A. Lee Swindlehurst, Interference-Aware Scheduling for MIMO Networks with Rate-Constrained Connectivity Requirements, *IEEE Trans. on Veh. Tech.*, vol. 61, no. 4, pp. 1762-1778, May 2012.
2. Feng Jiang and A. Lee Swindlehurst, Optimization of UAV Heading for Ground-to-Air Uplink, *IEEE JSAC special issue "Communications Challenges and Dynamics for Unmanned Autonomous Vehicles"*, vol. 30, no. 5, pp. 993-1005, June 2012.
3. Feng Jiang, Jie Chen and A. Lee Swindlehurst, Estimation in Phase-Shift and Forward Wireless Sensor Network, *IEEE Trans. on Signal Process.*, vol. 61, no. 15, pp. 3840-3851, Aug. 2013.
4. Feng Jiang, Jie Chen and A. Lee Swindlehurst, Optimal Power Allocation for Parameter Tracking in a Distributed Amplify-and-Forward Sensor Network, *IEEE Trans. on Signal Process.*, vol. 62, no. 9, pp. 2200-2211, May. 2014.

5. Feng Jiang, Jie Chen, A. Lee Swindlehurst and Jose A. Lopez-Salcedo, Massive MIMO for Wireless Sensing with a Coherent Multiple Access Channel, *accepted by IEEE Trans. on Signal Process.*, Feb. 2015.

6. Feng Jiang, Jie Chen and A. Lee Swindlehurst, Optimal Distributed Detection with Massive MIMO, *IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM) 2014*, pp. 245-248, Jun. 2014.

7. Feng Jiang, Jie Chen and A. Lee Swindlehurst, Linearly Reconfigurable Kalman Filtering for a Vector Process, *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2013*, pp. 5725-5729, May 2013.

8. Feng Jiang, Jie Chen and A. Lee Swindlehurst, Parameter Tracking via Optimal Distribute Beamforming in an Analog Sensor Network, *Asilorma Conference on Signals, Systems and Computers 2012*, pp.1397-1401, Nov. 2012.

9. Feng Jiang, Jie Chen and A. Lee Swindlehurst, Phase-Only Analog Encoding for a Multi-Antenna Fusion Center, *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2012*, pp. 2645-2648, Mar. 2012.

10. Feng Jiang and A. Lee Swindlehurst, Dynamic UAV Relay Positioning for the Ground-to-Air Uplink, *IEEE Globecom Workshop on Wireless Networking and Control for Unmanned Autonomous Vehicles (Wi-UAV) 2010*, pp. 1766-1770, Dec. 2010.

11. Feng Jiang, Jianqi Wang, and A. Lee Swindlehurst, Scheduling for MIMO Networks with Rate- Constrained Connectivity Requirements, *IEEE Vehicular Technology Conference (VTC) 2010 Spring*, pp. 1-5, May 2010.

PROFESSIONAL SERVICES

Paper reviewer for:
IEEE Trans. on Signal Processing
IEEE Trans. on Communications
IEEE Trans. on Wireless Communications
IEEE Trans. on Vehicular Technology
IEEE Trans. on Mobile Computing
IEEE Trans. on Aerospace and Electronic Systems
IEEE Communication Letters

ABSTRACT OF THE DISSERTATION

Physical Layer Optimization for Wireless Sensing and Network Connectivity

By

Feng Jiang

Doctor of Philosophy in Electrical and Computer Engineering

University of California, Irvine, 2015

Professor A. Lee Swindlehurst, Chair

Wireless sensor networks (WSNs) have been widely studied for detection and estimation problems. When a coherent multiple access channel is employed between the sensor nodes and fusion center (FC), each sensor takes a noisy measurement of the signal of interest, amplifies and forwards the measurement to a FC through a wireless fading channel, and the FC makes a decision about the presence of the signal and estimates its parameters based on the coherent sum of the signals from all the sensor nodes. To minimize estimation error or maximize probability of detection, the transmit power at the sensors is optimized under either sum or individual power constraints.

Most of the existing works assume that the FC is configured with a single antenna. It is well-known that multiple antennas can effectively increase the throughput of a wireless link, and in this thesis, we investigate how to exploit the benefit of the multiple antennas in WSN, and we study the detection and estimation performance of a coherent amplify-and-forward WSN, in which the sensor node has single antenna and the FC is configured with a massive number of antennas. When the perfect channel state information (CSI) is available at FC, we derive optimal closed-form sensor transmission gains to optimize the performance of Neyman-Pearson (NP) detector and the

linear minimum mean-squared error estimator (LMMSE), and if CSI is unknown at FC, we find the optimal sensor transmission gains to maximize the deflection coefficient of the energy detector (ED). Regarding the energy efficiency, our analysis show that the performance of NP detector and LMMSE estimator remain asymptotically constant if the sensor transmit power decreases proportionally with the increase in the number of antennas, and for the ED which does not require CSI, we show that a constant deflection can be asymptotically achieved if the sensor transmit power scales as the inverse square root of the number of FC antennas.

Additionally, we consider the problem of optimize the sensor phase to minimize the estimation error at FC, when the FC has a limited number of antennas. Two phase optimization algorithms are proposed and the sensor selection problem is formulated and solved. In addition to the case with multi-antenna FC, we also investigate the optimal power allocation for the WSN with single-antenna FC, when the FC use sensor measurements as input for a Kalman filter to track a dynamic parameter of interest.

When a fixed network infrastructure is not available (e.g. in military or disaster response scenarios), we investigate how to use the multi-antenna unmanned aerial vehicles (UAVs) as a relay to improve the connectivity between the mobile sensor nodes and the FC, which may be separated by a distances greater than their communication range. Several algorithms are proposed to optimize the trajectory of UAV.

Chapter 1

Introduction

1.1 Challenges in Wireless Sensing with a Coherent MAC

Wireless sensor networks (WSNs) have been widely studied for detection and estimation problems. Recently, considerable research has focused on the fusion of *analog* rather than encoded digital data in a distributed sensor network to improve estimation or detection performance. The advantages of analog WSNs have been established in [1–3], where it was shown that when using distortion between the source and recovered signal as the performance metric, digital transmission (separate source and channel coding) achieves an exponentially worse performance than analog signaling. In a distributed analog amplify-and-forward sensor network, the sensor nodes multiply their noisy observations by a complex factor and transmit the result to a fusion center (FC). In a coherent multiple access channel (MAC), the FC uses the coherent sum of the received signals to detect the signal and estimate the signal parameter. The key problem in this setting is designing the multiplication factor for each sensor to meet

some goal in terms of sensing performance or power consumption. Furthermore, for an optimal solution, these multipliers would have to be updated in scenarios where the parameter or wireless channels are time-varying. After the sensor network is deployed, the FC can estimate the wireless channels periodically. Based on the channel coefficients, the FC can first optimize the complex multipliers at sensor nodes under constraint of the sensing performance or power consumptions requirement, then feedback these complex multipliers to sensor nodes for future signal transmission.

Most of the existing results assume that the FC is equipped with only one antenna. Just as multi-antenna receivers can provide significant capacity or diversity gains in communication systems, the estimation or detection performance of a WSN should also benefit from the use of a multi-antenna FC, though prior work on this scenario is limited. Recently researchers have investigated the use of arrays with a massive number of antennas in wireless communication systems in order to improve spectral and energy efficiency [4–7]. Most of the research on so-called "massive MIMO" systems has been focused on cellular networks where the base station (BS) is configured with many antennas while the individual mobile stations have a single antenna. When perfect channel state information (CSI) is available at the BS, it has been shown that the transmit power of the mobile terminals can be reduced proportionally to the increase in the number of antennas without impacting the asymptotic rate of the users in the system [4]. The benefit is somewhat less when the BS uses an imperfect channel estimate; in this case the mobile users' transmit power can be inversely proportional to the square root of the number of antennas in order to achieve a constant rate [5]. For parameter detection or estimation problems in WSNs, an important question is how to design the complex multipliers at sensor nodes to exploit a multi-antenna FC to improve the probability of detection or estimation error.

For the network with fixed infrastructure, the network configuration (e.g., sensor

transmission gain and phase) can be optimized at FC, and then feedback the to the sensor nodes. When a fixed network infrastructure is not available (e.g., in military or disaster response scenarios), the sensor nodes (user) and the FC may be separated by distances greater than their communication range, or the signals may be shadowed due to mountainous terrain or dense surroundings (forests, buildings, etc.). Furthermore, since the users are mobile, the communications environment is constantly changing and thus connectivity between the user nodes and FC is often only sporadic. In such a scenario, the unmanned aerial vehicles (UAVs) can act as airborne relays (essentially "flying basestations") to provide an attractive solution since their altitude allows them to get above the ground-based shadowing and obtain line-of-sight (LOS) or near LOS communication channels over a large area. The UAV can collect the messages from the co-channel users on the ground and then forward them to some remote FC. Also and perhaps most importantly, the inherent mobility of UAVs allows their position to be adjusted in order to best accommodate the evolving network topology. To further improve the communication performance, the UAV can be configured with multiple antennas, and an interesting problem is how to control the motion of the UAV so as to optimize the uplink communications performance.

1.2 General Approach and Contributions

1.2.1 Transmit Power Allocation for Single-antenna FC

Most prior work on estimation in distributed amplify-and-forward sensor networks has focused on the situation where the parameter(s) of interest are time-invariant, and either deterministic or i.i.d. Gaussian. For the single-antenna FC, we consider a scenario in which the parameter of interest is dynamic, and at the sensors, the

observation of this dynamic signal is corrupted by measurement noise. The sensors amplify and forward the measurement to the FC through a coherent MAC, and a Kalman filter is used to track this signal at FC. Utilizing SDP, the optimal power allocations of sensor nodes to the following four problems have been found [8]:

- Minimize MSE under constraint of the sum transmit power of all sensor nodes
- Minimize the sum transmit power under the constraint of MSE
- Minimize MSE under individual power constraint of sensor nodes
- Minimize the maximum individual transmit power of sensor nodes under MSE constraint

Additionally, we generalize the tacking of a dynamic scalar parameter to the vector process. We assume the state-space observations linearly depend on a set of gain parameters, and these gain parameters can be optimized to improve the MSE of the Kalman filter. For the vector-valued dynamic model, the following two problems have been solved [9]:

- Minimize sum MSE under constraint of the sum transmit power
- Minimize the maximum MSE under the sum transmit power

1.2.2 Phase or Power Optimization for Multi-antenna FC

When the FC is configured with multiple antennas, the estimation or detection performance of WSN should be improved due to the diversity gain from additional antennas. For the multi-antenna FC, we consider a distributed WSN with single-antenna sensors that observe an unknown signal corrupted by noise. The sensors apply a phase shift

or gain to their observation and then simultaneously transmit the result to a multi-antenna FC over a coherent MAC. The FC will detect the existence of the signal and estimate its parameters. Based on the availability of the channel state information (CSI), the FC determines the optimal value of the phase or gain for each sensor in order to improve the probability of detection and estimation error, and then feeds this information back to the sensors so that they can apply the appropriate phase shift or gain for future transmission.

If the number of antennas at FC is limited, we assume sensors apply a phase shift (rather than both gain and phase) to their observation and then simultaneously transmit the result to a multi-antenna FC over a coherent MAC. The advantage of a phase-shift-only transmission is that it leads to a simpler analog implementation at the sensor. The detailed contributions regarding the multi-antenna FC are listed below [10]:

- Propose two algorithms to minimize maximum-likelihood (ML) estimation error at FC through optimizing phase-shift of sensor nodes
- Derive conditions to determine whether multiple antennas can provide more benefit than single-antenna FC
- Analyze the estimation performance degradation due to phase error at sensor nodes
- Investigate sensor selection problems under different assumptions of measurement noise

The benefit of the phase-shift and forward scheme is not always achievable. When the FC have a massive number of antennas, we will show that the detection or estimation performance at FC only depends on the gain at the sensor nodes and the phase can

be arbitrary. After receiving the signals from sensor nodes, the FC can optimize the sensor gain to maximize probability of detection or minimize estimation error. For the massive MIMO sensor network, the contributions are summarized as [11]:

- When perfect CSI is available at FC, we derive the optimal closed-form "water-filling" sensor gain allocation solution for Neyman-Pearson (NP) detection or linear minimum mean-squared error estimation(LMMSE)

- When CSI is unknown at FC, we study the performance of energy detector, and to maximize the deflection of the energy detector, optimal sensor gain have been found using quadratic constraint linear programming (QCLP).

- Analyze the energy efficiency at sensor nodes under different CSI assumptions. When perfect CSI is known at FC, the sensor transmit power can decrease proportionally to the antenna number, and meanwhile the NP detector and LMMSE estimator can achieve a constant detection performance or estimation error; when the CSI is unknown, the sensor transmit power can be scaled by the inverse square root of FC antenna number to achieve a constant deflection.

- Derive performance bounds to show when the multi-antenna FC have better detection or estimation performance than the single-antenna FC.

1.2.3 Connectivity Optimization for UAV Assisted Mobile Network

For the UAV assisted mobile sensor network, we assume several ground-based sensors (users) communicate simultaneously with a multi-antenna UAV relay. The users are assumed to transmit signal with a single antenna and the UAV uses beamforming to separate the co-channel data streams. There exist co-channel interference between

different users' data streams and we quantify the *connectivity* of the relay network using the ergodic achievable sum rate of the uplink communication, assuming that the UAV uses a maximum signal-to-interference-plus-noise ratio (SINR) beamformer for interference mitigation. The strength of the mutual interference depends on the correlation between the users' channel vectors, which in a channel with a strong line-of-sight (LOS) component is a function of the signals' angle of arrival (AoA). The AoAs depend in turn on the UAV's heading and the relative positions of the UAV and the ground nodes. Consequently, we study the problem of how to adjust the heading of the UAV to minimize the users' mutual interference and improve the uplink communications performance. The main results of the UAV heading optimization problem are summarized as follows [12]:

- We investigate the case with two static ground nodes and the corresponding heading optimization problem is solved using line search.

- For the general mobile network, we formulate a heading optimization problem and develop a line-search algorithm for dynamically adjusting the UAV heading to maximize the sum rate of the uplink channel, using a prediction filter to track the positions of the mobile ground nodes.

- When a strong LOS channel exists between the ground nodes and UAV, we derive asymptotic closed-form solutions for the heading optimization problem for low and high SNR cases, the performance of which are very close to that of the optimal algorithm.

1.3 Outlines

The rest part of this dissertation is organized as follows. Chapter 2 provides a general overview of the system models used in this thesis. In Chapter 3, we investigate the scalar dynamic parameter tracking problem in amplify-and-forward WSN with a single-antenna FC. Problems of MSE minimization under transmit power constraint or transmit power minimization under MSE constraint have been formulated and optimal solutions have been found using SDP. The power allocation for the tracking of a vector process is studied in Chapter 4, and problems of minimizing sum MSE or maximum MSE under power constraint have been formulated and solved.

Chapter 5 focuses on the deterministic parameter estimation in phase-shift and forward WSN with a multi-antenna FC. Based on the perfect knowledge of CSI, to minimize the estimation error at FC, we propose two algorithms for optimizing sensor phase. Asymptotic performance analyses for cases involving large numbers of sensors or large numbers of FC antennas are derived. The impact of phase errors at the sensor transmitters and the sensor selection problem are also studied.

When the FC is configured with a large antenna array, the detection or estimation performance of a zero-mean Gaussian signal are analyzed in Chapter 6. For the NP detector and LMMSE estimator which requires perfect CSI, we show that detection or estimation performance remains asymptotically constant with the number of FC antennas if the sensor transmit power is scaled by the inverse of the number of antennas. For the energy detector which does not require CSI, we derive optimal gains that maximize the deflection coefficient of the detector, and we show that a constant deflection can be asymptotically achieved if the sensor transmit power is multiplied by the increase of square root of the number of FC antennas.

Chapter 7 considers a collection of single-antenna ground sensor nodes communicating

with a multi-antenna UAV relay over a multiple-access ground-to-air communications link. First, a simple scenario with two static ground nodes is investigated. We then study a more general system setting with multiple mobile ground-based users, and develop an algorithm for dynamically adjusting the UAV heading to maximize the sum rate of the uplink channel. Last, the asymptotic closed-form solutions for UAV heading optimization are derived under the assumptions of high or low SNR. The conclusions of this thesis are summarized in Chapter 8.

The notation used in this dissertation is summarized as follows. Lower-case and upper-case bold letters represent vectors and matrices respectively, and $\mathbb{C}^{M\times 1}$ denotes the space of M-element complex vectors. We use $(\cdot)^T$ and $(\cdot)^H$ for transpose and conjugate transpose respectively. The $M \times M$ identity matrix and zero matrix are denoted as $\mathbf{I}_M$ and $\mathbf{0}_M$ respectively, and $\text{diag}\{d_1 \ \cdots \ d_N\}$ is a $N \times N$ diagonal matrix with d_i as the ith diagonal element. Probabilities and conditional probabilities are denoted by $\Pr(\cdot)$ and $\Pr(\cdot|\cdot)$, and $p(\cdot|\cdot)$ represents a conditional probability density function. The functions $\mathbb{E}\{\cdot\}$ and $\text{Var}\{\cdot\}$ denote the expectation and variance of a random variable, and $\mathcal{CN}(\mathbf{0}, \mathbf{\Sigma})$ denotes the complex Gaussian distribution with zero mean and covariance matrix $\mathbf{\Sigma}$. The ith eigenvalue of a matrix is written as $\lambda_i(\cdot)$, and for two Hermitian matrices $\mathbf{A}$ and $\mathbf{B}$, $\mathbf{A} \succeq \mathbf{B}$ means that $\mathbf{A} - \mathbf{B}$ is positive semidefinite. The operators $\text{tr}(\cdot)$, $\odot$ and $\otimes$ respectively denote the trace of a matrix, the Hadamard (element-wise) product and the Kronecker product of two matrices .

Chapter 2

System Models and Assumptions

In this chapter we introduce the general network and signal models used in this thesis. For the distributed analog amplify-and-forward sensor network, we assume the single-antenna sensor nodes can independently measure the signal of interest and then scale the measurements with a complex multiplier and forward the signal to the FC. Based on the detection or estimation performance, the FC can optimzie the gain or phase at each sensor node, and feedback these information to the sensor nodes for future signal transmission. For the UAV assisted mobile network, the mobile nodes on the ground are assumed to have single antenna, and the UAV is configured with an antenna array. The UAV can track the positions of the ground users through Kalman filtering. The connectivity of the network can be optimized through adjusting the UAV's relative position to the ground nodes, and the trajectory of UAV is controlled by its heading. The details of the system models are described below.

2.1 Parameter Tracking Using a Single-antenna FC

For the parameter tracking problem, we model the evolution of a complex-valued dynamic parameter θ_n using a first-order Gauss-Markov process:

$$\theta_n = \alpha\theta_{n-1} + u_n \ ,$$

where n denotes the time step, α is the correlation parameter and the process noise u_n is zero-mean complex normal variable with variance σ_u^2 (denoted by $\mathcal{CN}(0, \sigma_u^2)$). We assume that θ_0 is zero mean and that the norm $|\alpha| < 1$, so that θ_n is a stationary process. Thus, the variance of θ_n is constant and given by $\sigma_\theta^2 = \sigma_u^2/\left(1 - |\alpha|^2\right)$. A set of N sensors measures θ_n in the presence of noise; the measurement for the ith sensor at time n is described by

$$s_{i,n} = \theta_n + v_{i,n} \ ,$$

where the measurement noise $v_{i,n}$ is distributed as $\mathcal{CN}(0, \sigma_{v,i}^2)$. In an amplify-and-forward sensor network, each sensor multiplies its observation by a complex gain factor and transmits the result over a wireless channel to a FC. The FC receives a coherent sum of the signals from all N sensors in additive noise:

$$\begin{aligned} y_n &= \sum_{i=1}^{N} h_{i,n} a_{i,n} s_{i,n} + w_n \\ &= \sum_{i=1}^{N} (h_{i,n} a_{i,n} \theta_n + h_{i,n} a_{i,n} v_{i,n}) + w_n \ , \end{aligned} \tag{2.1}$$

where $h_{i,n}$ is the gain of the wireless channel between sensor i and the FC, $a_{i,n}$ is the complex transmission gain of sensor i, and w_n is noise distributed as $\mathcal{CN}(0, \sigma_w^2)$.

Based on the received signal y_n, the FC can use a standard Kalman filter to get an estimate $\hat{\theta}_{n|n}$, and the MSE $\mathbb{E}\{|\theta_n - \hat{\theta}_{n|n}|^2\}$ will be a function of the complex gain a_i

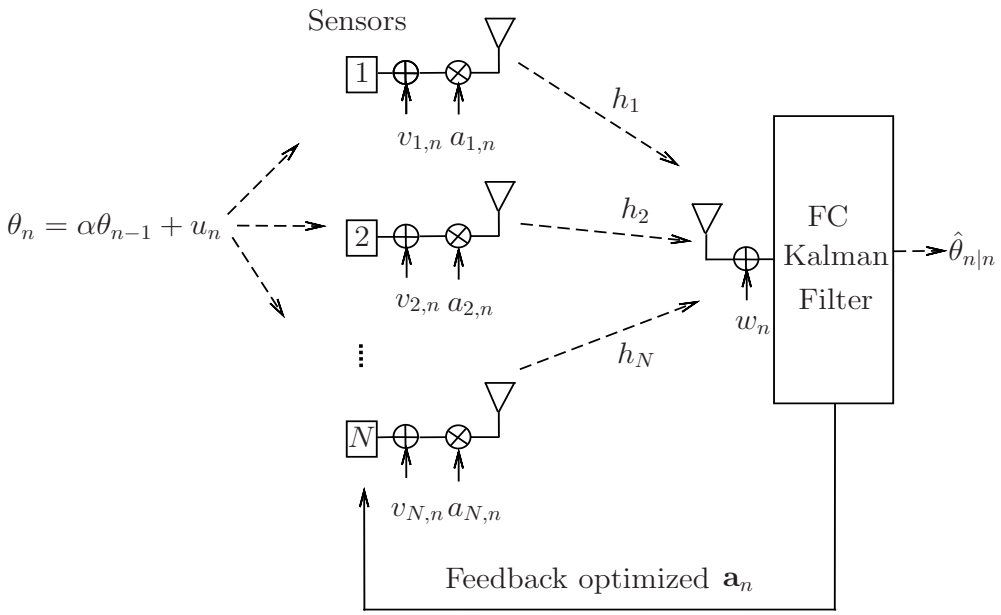

Figure 2.1: Block diagram for parameter tracking with a single-antenna FC

at the sensor nodes, thus the FC can minimize the MSE through optimizing a_i. A block diagram for the WSN with single-antenna FC is provided in figure 2.1.

2.2 Wireless Sensing with a Multi-antenna FC

When the FC is configured with multiple antennas, the received signal from different antennas can be combined coherently to improve the detection or estimation performance of the WSN. For the detection problem, the signal of interest θ is modeled as a zero-mean circular complex Gaussian variable[1] with variance σ_θ^2, a distribution we denote by $\mathcal{CN}(0, \sigma_\theta^2)$. The measurement available at the ith of N sensor nodes is

$$s_i = \theta + v_i \ ,$$

[1] Although we use a Bayesian framework, our approach can be also used for the deterministic case, in which θ is assumed to be a deterministic signal.

where v_i is measurement noise distributed as $\mathcal{CN}(0, \sigma_{v,i}^2)$. The ith sensor multiplies the measurement with a complex gain a_i and coherently forwards the result to the FC through a wireless fading channel. The received signal at the M-antenna FC under the two hypotheses is

$$\begin{aligned} \mathcal{H}_0 &: \mathbf{y} = \mathbf{HDv} + \mathbf{n} \\ \mathcal{H}_1 &: \mathbf{y} = \mathbf{Ha}\theta + \mathbf{HDv} + \mathbf{n} \,, \end{aligned} \tag{2.2}$$

where

$$\begin{aligned} \mathbf{v} &= [v_1 \ \cdots \ v_N]^T \\ \mathbf{a} &= [a_1 \ \cdots \ a_N]^T \\ \mathbf{D} &= \text{diag}\{a_1 \ \cdots \ a_N\} \\ \mathbf{H} &= [\mathbf{h}_1 \ \cdots \ \mathbf{h}_N] \,, \end{aligned}$$

$\mathbf{h}_i \in \mathbb{C}^{M\times 1}$ is the channel gain between the ith sensor and the FC, and the vector $\mathbf{n} \in \mathbb{C}^{M\times 1}$ represents additive Gaussian noise at the FC and has the distribution $\mathcal{CN}(0, \sigma_n^2 \mathbf{I}_M)$.

Assuming that the FC has perfect knowledge of signal variance σ_θ^2, the measurement noise power $\sigma_{v,i}^2$ and the CSI in $\mathbf{H}$, the NP criterion can be used to distinguish between the hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$, and based on $\mathcal{H}_1$, the value of θ can be estimated using LMMSE estimator. When the CSI is unknown at FC, the energy detector can be used to detect the existence of the signal θ. The block diagram for wireless sensing with a M-antenna FC is described in figure 2.2.

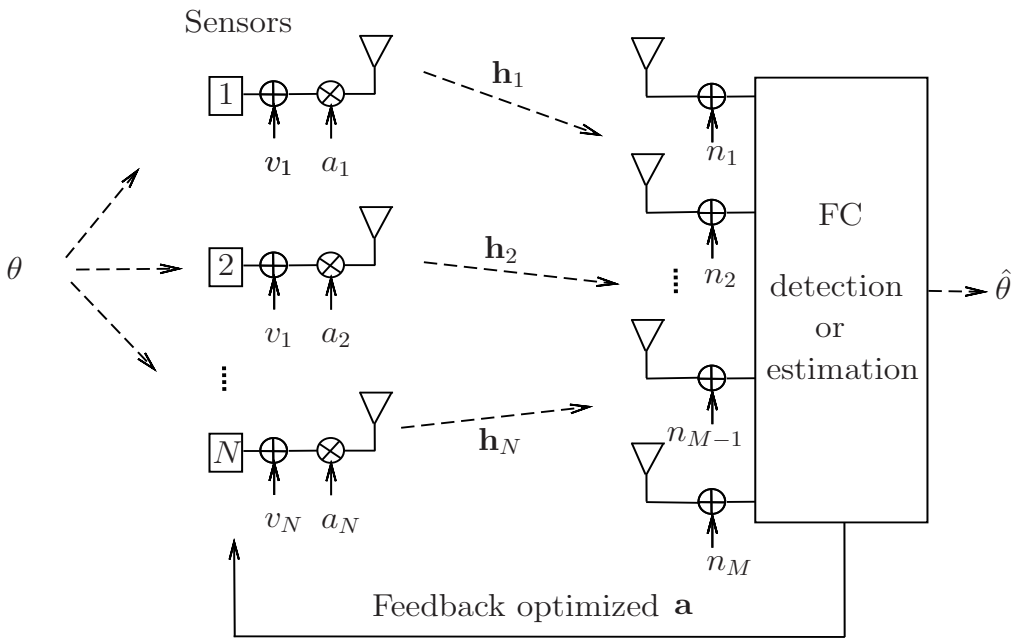

Figure 2.2: Block diagram of wireless sensing with a multi-antenna FC

2.3 Connectivity of UAV Assisted Mobile Sensor Network

When the UAV relay is used to improve the connectivity of the mobile network, the mobile users first transmit the data to the UAV relay and the UAV relay will decode and forward the data to the FC. We assume a UAV configured with an array of M antennas, and a collection of N ground nodes each equipped with a single antenna. We restrict attention to fixed-wing (non-hovering) UAVs that must maintain a certain forward velocity to remain airborne. Fixed-wing UAVs have two advantages for our application: (1) they tend to be somewhat larger than hovering UAVs and allow more flexibility in deploying an antenna array with a larger aperture, and (2) the rotary blade motion on hovering aircraft can lead to high-Doppler reflections of the communications signals that are difficult to compensate for. We assume that, during the period of time in which the UAV is receiving uplink data from the ground

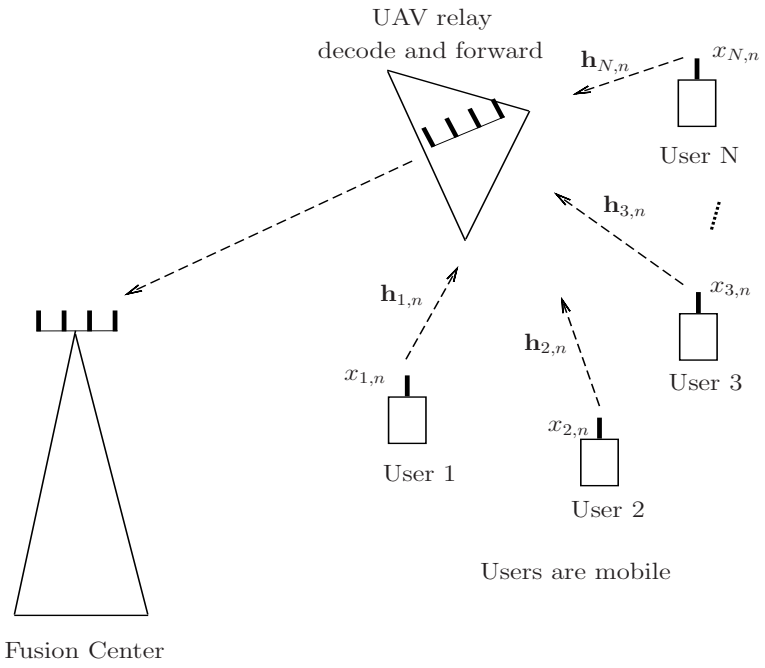

Figure 2.3: Block diagram of UAV relay assisted mobile network

nodes, the UAV maintains a constant altitude h_u and a constant velocity v_u. For simplicity, we assume that each ground node transmits with the same power P_t, but this assumption is easily relaxed. The signal received at the UAV array at time n can thus be written as

$$\mathbf{y}_n = \sum_{i=1}^{N} \sqrt{P_t}\mathbf{h}_{i,n}x_{i,n} + \mathbf{n}_n,$$

where $\mathbf{h}_{i,n} \in \mathbb{C}^{M\times 1}$ is the channel vector between node i and the UAV, the data symbol $x_{i,n}$ is a complex scalar with zero mean and unit magnitude, $\mathbf{n}_n \in \mathbb{C}^{M\times 1}$ is zero-mean additive Gaussian noise with covariance $\mathbb{E}\{\mathbf{n}_n\mathbf{n}_n^H\} = \sigma^2\mathbf{I}_M$. The UAV uses a beamformer $\mathbf{w}_{i,n}$ to isolate the data from the i-th node: $\hat{x}_{i,n} = \mathbf{w}_{i,n}^H\mathbf{y}_n$, and the corresponding SINR is calculated as

$$SINR_{i,n} = \frac{P_t|\mathbf{w}_{i,n}^H\mathbf{h}_{i,n}|^2}{\mathbf{w}_{i,n}\mathbf{Q}_{i,n}\mathbf{w}_{i,n}}\,, \tag{2.3}$$

where $\mathbf{Q}_{i,n} = \sum_{j=1, j\neq i}^{N} P_t \mathbf{h}_{j,n} \mathbf{h}_{j,n}^H + \sigma^2 \mathbf{I}_M$. The connectivity of the network is evaluated using the average sum rate of the uplink transmission which is calculated as

$$C_n = \sum_{i=1}^{N} \mathbb{E}\left\{\log_2(1 + SINR_{i,n})\right\}. \tag{2.4}$$

From (2.3) and (2.4), we can observe that the uplink communication performance depends on the channel vector $\mathbf{h}_{i,n}$ through the $SINR_{i,n}$. In this work, we assume the channel between the ground nodes and the UAV is Rician fading and the LOS component of the channel $\mathbf{h}_{i,n}$ is decided by the AoA which is a function of the positions of the UAV and the ground nodes. After estimating the mobile ground users' positions, the UAV can adjust its heading direction to change its relative position to the ground nodes such that the average uplink sum rate can be maximized.

Chapter 3

Optimal Power Allocation for Tracking of a Scalar Parameter

3.1 Introduction

In a distributed analog amplify-and-forward sensor network, the sensor nodes multiply their noisy observations by a complex factor and transmit the result to an FC. In a coherent MAC, the FC uses the coherent sum of the received signals to estimate the parameter. Most prior work on estimation in distributed amplify-and-forward sensor networks has focused on the situation where the parameter(s) of interest are time-invariant, and either deterministic or i.i.d. Gaussian. The case of an orthogonal MAC, where the FC has access to the individual signals from each sensor, has been studied in [13–19]. For a coherent MAC, relevant work includes [10, 14, 18, 20–22]. In [13–15, 17], two kinds of problems were considered: minimizing the estimation error under sum or individual power constraints, and minimizing the sum transmit power under a constraint on the estimation error. Scaling laws for the estimation

error with respect to the number of sensors were derived in [16, 18] under different access schemes and for different power allocation strategies. In [10, 22], the authors exploited a multi-antenna FC to minimize the estimation error.

More relevant to this chapter is interesting recent work by Leong et al, who model the (scalar) parameter of interest using a dynamic Gauss-Markov process and assume the FC employs a Kalman filter to track the parameter [23, 24]. In [23], both the orthogonal and coherent MAC were considered and two kinds of optimization problems were formulated: MSE minimization under a global sum transmit power constraint, and sum power minimization problem under an MSE constraint. An asymptotic expression for the MSE outage probability was also derived assuming a large number of sensor nodes. The problem of minimizing the MSE outage probability for the orthogonal MAC with a sum power constraint was studied separately in [24].

In this chapter, we consider scenarios similar to those in [23]. In particular, we focus on the coherent MAC case assuming a dynamic parameter that is tracked via a Kalman filter at the FC. As detailed in the list of contributions below, we extend the work of [23] for the case of a global sum power constraint, and we go beyond [23] to study problems where either the power of the individual sensors is constrained, or the goal is to minimize the peak power consumption of individual sensors:

1. We find a closed-form expression for the optimal complex transmission gains that minimize the MSE under a constraint on the sum power of all sensor transmissions. While this problem was also solved in [23] using the KKT conditions derived in [14], our approach results in a simpler and more direct solution. We also examine the asymptotic form of the solution for high total transmit power or high noise power at the FC.

2. We find a closed-form expression for the optimal complex transmission gain

that minimizes the sum power under a constraint on the MSE. In this case, the expression depends on the eigenvector of a particular matrix. Again, while this problem was also addressed in [23], the numerical solution therein is less direct than the one we obtain. In addition, we find an asymptotic expression for the sum transmit power for a large number of sensors.

3. We show how to find the optimal transmission gains that minimize the MSE under individual sensor power constraints by relaxing the problem to a SDP problem, and then proving that the optimal solution can be constructed from the SDP solution.

4. We show how to find the optimal transmission gains that minimize the maximum individual power over all of the sensors under a constraint on the maximum MSE. Again, we solve the problem using SDP, and then prove that the optimal solution can be constructed from the SDP solution.

5. For the special case where the sensor nodes use equal power transmission, we derive an exact expression for the MSE outage probability.

The rest of this chapter is organized as follows. Section 3.2 describes the system model for the parameter tracking problem and provides an expression for the MSE obtained at the FC using a standard Kalman filter. Section 3.3 investigates the MSE minimization problem under the assumption that the sensor nodes have a sum transmit power constraint. The MSE minimization problem with individual sensor power constraints is formulated and solved in Section 3.3.2. The problems of minimizing the sum power or the maximum individual sensor power with MSE constraints are formulated and optimally solved in Section 3.4. In Section 3.5, the MSE outage probability for equal power allocation is derived. Numerical results are presented in Section 3.6 and the conclusions are summarized in Section 3.7.

3.2 Signal Model

As introduced in section 2.1, the signal model in equation (2.1) can be written more compactly in matrix-vector form, as follows:

$$y_n = \mathbf{a}_n^H \mathbf{h}_n \theta_n + \mathbf{a}_n^H \mathbf{H}_n \mathbf{v}_n + w_n \ ,$$

where $\mathbf{h}_n = [h_{1,n} \cdots h_{N,n}]^T$, $\mathbf{a}_n = [a_{1,n} \cdots a_{N,n}]^H$ is a vector containing the conjugate of the sensor transmission gains, $\mathbf{H}_n = \text{diag}\{h_{1,n} \cdots h_{N,n}\}$ is a diagonal matrix, and the measurement noise vector $\mathbf{v}_n = [v_{1,n} \cdots v_{N,n}]^T$ has covariance $\mathbf{V} = \mathbb{E}\{\mathbf{v}_n \mathbf{v}_n^H\} = \text{diag}\left\{\sigma_{v,1}^2 \cdots \sigma_{v,N}^2\right\}$.

The FC is assumed to know the statistics of the various noise processes, the current channel state $\mathbf{h}_n$, and the transmission gains $\mathbf{a}_n$, and it uses a standard Kalman filter to track the parameter θ_n according to the equations below [25]:

- Prediction Step: $\hat{\theta}_{n|n-1} = \alpha \hat{\theta}_{n-1|n-1}$
- Prediction MSE: $P_{n|n-1} = \alpha^2 P_{n-1|n-1} + \sigma_u^2$
- Kalman Gain:

$$k_n = \frac{P_{n|n-1} \mathbf{h}_n^H \mathbf{a}_n}{\mathbf{a}_n^H \mathbf{H}_n \mathbf{V} \mathbf{H}_n^H \mathbf{a}_n + P_{n|n-1} \mathbf{a}_n^H \mathbf{h}_n \mathbf{h}_n^H \mathbf{a}_n + \sigma_w^2}$$

- Measurement Update:

$$\hat{\theta}_{n|n} = \hat{\theta}_{n|n-1} + k_n \left(y_n - \mathbf{a}_n^H \mathbf{h}_n \hat{\theta}_{n|n-1} \right)$$

- Filtered MSE:

$$P_{n|n} = (1 - k_n \mathbf{a}_n^H \mathbf{h}_n) P_{n|n-1} \ . \tag{3.1}$$

The goal is to determine an optimal choice for the gains $\mathbf{a}_n$ that minimizes the filtered MSE under a power constraint, or that minimizes the power consumed in transmitting the data to the FC under an MSE constraint. The optimal gains are then fed back to the individual sensors to use at time n.

3.3 Minimizing MSE under a Power Constraint

3.3.1 Global Sum Power Constraint

In this section, we briefly consider the problem of minimizing the MSE under the assumption that the sensor nodes have a sum power constraint. As mentioned earlier, this problem has already been studied in [23], but the solution we provide here is simpler and more direct. The optimization problem can be written as

$$\begin{aligned} \min_{\mathbf{a_n}} \quad & P_{n|n} \\ s.t. \quad & \mathbf{a}_n^H \mathbf{D} \mathbf{a}_n \leq P_T \ , \end{aligned} \tag{3.2}$$

where $\mathbf{a}_n^H \mathbf{D} \mathbf{a}_n$ and P_T respectively represent the actual and total available transmit power, with $\mathbf{D} = \text{diag}\{\sigma_\theta^2 + \sigma_{v,1}^2 \cdots \sigma_\theta^2 + \sigma_{v,N}^2\}$. From (3.1), minimizing the MSE $P_{n|n}$ is equivalent to maximizing

$$k_n \mathbf{a}_n^H \mathbf{h}_n = \frac{P_{n|n-1} \mathbf{a}_n^H \mathbf{h}_n \mathbf{h}_n^H \mathbf{a}_n}{\mathbf{a}_n^H \mathbf{H}_n \mathbf{V} \mathbf{H}_n^H \mathbf{a}_n + P_{n|n-1} \mathbf{a}_n^H \mathbf{h}_n \mathbf{h}_n^H \mathbf{a}_n + \sigma_w^2} \ ,$$

and after a simple manipulation, the optimization problem in (3.2) is equivalent to

$$\begin{aligned} \max_{\mathbf{a}_n} \quad & \frac{\mathbf{a}_n^H \mathbf{h}_n \mathbf{h}_n^H \mathbf{a}_n}{\mathbf{a}_n^H \mathbf{H}_n \mathbf{V} \mathbf{H}_n^H \mathbf{a}_n + \sigma_w^2} \\ s.t. \quad & \mathbf{a}_n^H \mathbf{D} \mathbf{a}_n \leq P_T \ . \end{aligned} \tag{3.3}$$

Denote the optimal solution to (3.3) as $\mathbf{a}_n^*$. It is easy to verify that the objective function of (3.3) is monotonically increasing in the norm of $\mathbf{a}_n$, which implies that at the optimal solution, the sum transmit power constraint should be met with equality $\mathbf{a}_n^{*H}\mathbf{D}\mathbf{a}_n^* = P_T$. Thus (3.3) becomes a Rayleigh quotient under a quadratic equality constraint. Since the numerator involves a rank-one quadratic term, a simple closed-form solution is possible. If we define $\mathbf{B} = \mathbf{H}_n\mathbf{V}\mathbf{H}_n^H + \frac{\sigma_w^2}{P_T}\mathbf{D}$, the optimal solution is given by

$$\mathbf{a}_n^* = \sqrt{\frac{P_T}{\mathbf{h}_n^H\mathbf{B}^{-1}\mathbf{D}\mathbf{B}^{-1}\mathbf{h}_n}}\mathbf{B}^{-1}\mathbf{h}_n \ . \tag{3.4}$$

Note that the phase of each sensor transmission gain is the conjugate of the channel to the FC (recall that $\mathbf{a}_n$ contains the conjugate of these transmission gains). In [23], this property was assumed from the beginning in order to get an optimization problem with only real-valued variables; however, we see that this phase-matched solution results even without this assumption.

The maximum value of the objective function in (3.3) can be expressed as

$$\frac{\mathbf{a}_n^{*H}\mathbf{h}_n\mathbf{h}_n^H\mathbf{a}_n^*}{\mathbf{a}_n^{*H}(\mathbf{H}_n\mathbf{V}\mathbf{H}_n^H + \frac{\sigma_w^2}{P_T}\mathbf{D})\mathbf{a}_n^*} = \mathbf{h}_n^H\mathbf{B}^{-1}\mathbf{h}_n \ .$$

Given that

$$\begin{aligned} \mathbf{h}_n^H\mathbf{B}^{-1}\mathbf{h}_n &\overset{(a)}{<} \mathbf{h}_n^H(\mathbf{H}\mathbf{V}\mathbf{H}^H)^{-1}\mathbf{h}_n & (3.5)\\ &= \sum_{i=1}^{N}\frac{1}{\sigma_{v,i}^2} \ , & (3.6) \end{aligned}$$

where (a) follows from $\mathbf{B}^{-1} \prec (\mathbf{H}\mathbf{V}\mathbf{H}^H)^{-1}$, a lower bound on the MSE can be obtained

by plugging (3.6) into (3.1):

$$\begin{aligned} P_{n|n} &> \left(1 - \frac{1}{1 + \frac{1}{\left(\sum_{i=1}^{N} \frac{1}{\sigma_{v,i}^2}\right) P_{n|n-1}}}\right) P_{n|n-1} \\ &= \frac{P_{n|n-1}}{1 + \left(\sum_{i=1}^{N} \frac{1}{\sigma_{v,i}^2}\right) P_{n|n-1}} . \end{aligned} \tag{3.7}$$

Equation (3.5) becomes an equality when $\sigma_w^2/P_T \to 0$ or when the signal-to-noise-ratio (SNR) at the FC is very high, and the resulting optimal sensor transmission gains become

$$\mathbf{a}_n^* = \sqrt{\frac{P_T}{\sum_{i=1}^{N} \frac{\sigma_\theta^2 + \sigma_{v,i}^2}{\sigma_{v,i}^4 |h_{n,i}|^2}}} \left[\frac{1}{h_{1,n}\sigma_{v,1}^2} \cdots \frac{1}{h_{N,n}\sigma_{v,N}^2}\right]^H . \tag{3.8}$$

In this case, sensors with small channel gains or low measurement noise are allocated more transmit power. On the other hand, for low SNR at the FC where $\sigma_w^2/P_T \to \infty$, we have $\mathbf{B} \approx \frac{\sigma_w^2}{P_T}\mathbf{D}$, and hence from (3.4) the optimal gain vector is proportional to

$$\mathbf{a}_n^* \propto \left[\frac{h_{1,n}}{\sigma_\theta^2 + \sigma_{v,1}^2} \cdots \frac{h_{N,n}}{\sigma_\theta^2 + \sigma_{v,N}^2}\right]^T . \tag{3.9}$$

Interestingly, unlike the high SNR case, for low SNR the sensors with large channel gains are assigned higher power. This observation will be highlighted later in the simulations of Section 3.6.

3.3.2 Individual Power Constraints

In a distributed sensor network, it is more likely that the power of the individual sensors would be constrained, rather than the total sum power of the network. As seen in the previous section, when the SNR at the FC is high (low), a weak (strong)

channel for a given sensor can lead to a high transmission power that the sensor may not be able to support. Thus, in this section we address the problem of minimizing the MSE under individual sensor power constraints, as follows:

$$\begin{aligned} \min_{\mathbf{a}_n} \quad & P_{n|n} \\ s.t. \quad & |a_{i,n}|^2(\sigma_\theta^2 + \sigma_{v,i}^2) \leq P_{T,i} \ , \quad i = 1, \cdots, N \ , \end{aligned} \tag{3.10}$$

where $P_{T,i}$ is the maximum transmit power available at the ith sensor node. Similar to (3.2), problem (3.10) can be rewritten as

$$\begin{aligned} \max_{\mathbf{a}_n} \quad & \frac{\mathbf{a}_n^H \mathbf{h}_n \mathbf{h}_n^H \mathbf{a}_n}{\mathbf{a}_n^H \mathbf{H}_n \mathbf{V} \mathbf{H}_n^H \mathbf{a}_n + \sigma_w^2} \\ s.t. \quad & |a_{i,n}|^2(\sigma_\theta^2 + \sigma_{v,i}^2) \leq P_{T,i} \ , \quad i = 1, \cdots, N \ . \end{aligned} \tag{3.11}$$

Problem (3.11) is a quadratically constrained ratio of quadratic functions (QCRQ), and as explained below we will use the approach of [26] to transform the QCRQ problem into a relaxed SDP problem. Introduce a real auxiliary variable t and define $\tilde{\mathbf{a}}_n = t\mathbf{a}_n$, so that problem (3.11) is equivalent to

$$\begin{aligned} \max_{\mathbf{a}_n, t} \quad & \frac{\tilde{\mathbf{a}}_n^H \mathbf{h}_n \mathbf{h}_n^H \tilde{\mathbf{a}}_n}{\tilde{\mathbf{a}}_n^H \mathbf{H}_n \mathbf{V} \mathbf{H}_n^H \tilde{\mathbf{a}}_n + \sigma_w^2 t^2} \\ s.t. \quad & \tilde{\mathbf{a}}_n^H \mathbf{D}_i \tilde{\mathbf{a}}_n \leq t^2 P_{T,i} \ , \quad i = 1, \cdots, N \\ & t \neq 0 \ , \end{aligned} \tag{3.12}$$

where $\mathbf{D}_i = \text{diag}\{0 \cdots 0 \ \sigma_\theta^2 + \sigma_{v,i}^2 \ 0 \cdots 0\}$. We can further rewrite problem (3.12) as

$$\begin{aligned} \max_{\mathbf{a}_n, t} \quad & \tilde{\mathbf{a}}_n^H \mathbf{h}_n \mathbf{h}_n^H \tilde{\mathbf{a}}_n \\ s.t. \quad & \tilde{\mathbf{a}}_n^H \mathbf{H}_n \mathbf{V} \mathbf{H}_n^H \tilde{\mathbf{a}}_n + \sigma_w^2 t^2 = 1 \\ & \tilde{\mathbf{a}}_n^H \mathbf{D}_i \tilde{\mathbf{a}}_n \leq t^2 P_{T,i} \ , \quad i = 1, \cdots, N \ . \end{aligned} \tag{3.13}$$

Note that the constraints in problem (3.13) already guarantee that $t \neq 0$, so this constraint is removed.

Define $\bar{\mathbf{a}}_n = [\tilde{\mathbf{a}}_n^H \; t]^H$ and the matrices

$$\bar{\mathbf{H}}_n = \begin{bmatrix} \mathbf{h}_n\mathbf{h}_n^H & \mathbf{0} \\ \mathbf{0}^T & 0 \end{bmatrix}, \qquad \bar{\mathbf{C}}_n = \begin{bmatrix} \mathbf{H}_n\mathbf{V}\mathbf{H}_n^H & \mathbf{0} \\ \mathbf{0}^T & \sigma_w^2 \end{bmatrix}, \qquad \bar{\mathbf{D}}_i = \begin{bmatrix} \mathbf{D}_i & \mathbf{0} \\ \mathbf{0}^T & -P_{T,i} \end{bmatrix},$$

so that problem (3.13) can be written in the compact form

$$\begin{aligned} \max_{\bar{\mathbf{a}}_n} \quad & \bar{\mathbf{a}}_n^H \bar{\mathbf{H}}_n \bar{\mathbf{a}}_n \\ s.t. \quad & \bar{\mathbf{a}}_n^H \bar{\mathbf{C}}_n \bar{\mathbf{a}}_n = 1 \\ & \bar{\mathbf{a}}_n^H \bar{\mathbf{D}}_i \bar{\mathbf{a}}_n \leq 0 \; , \quad i = 1, \cdots, N \; . \end{aligned} \tag{3.14}$$

Defining the $(N+1) \times (N+1)$ matrix $\bar{\mathbf{A}} = \bar{\mathbf{a}}_n \bar{\mathbf{a}}_n^H$, problem (3.14) is equivalent to

$$\begin{aligned} \max_{\bar{\mathbf{A}}} \quad & \mathrm{tr}(\bar{\mathbf{A}}\bar{\mathbf{H}}_n) \\ s.t. \quad & \mathrm{tr}(\bar{\mathbf{A}}\bar{\mathbf{C}}_n) = 1 \\ & \mathrm{tr}(\bar{\mathbf{A}}\bar{\mathbf{D}}_i) \leq 0 \; , \quad i = 1, \cdots, N \\ & \mathrm{rank}(\bar{\mathbf{A}}) = 1 \\ & \bar{\mathbf{A}} \succeq 0 \; . \end{aligned} \tag{3.15}$$

Were it not for the rank constraint, the problem in (3.15) would be a standard SDP problem and could be solved in polynomial time using (for example) the interior point method. Given the difficulty of handling the rank constraint, we choose to relax it

and solve the simpler problem

$$\begin{aligned} \max_{\bar{\mathbf{A}}} \quad & \mathrm{tr}(\bar{\mathbf{A}}\bar{\mathbf{H}}_n) \\ s.t. \quad & \mathrm{tr}(\bar{\mathbf{A}}\bar{\mathbf{C}}_n) = 1 \\ & \mathrm{tr}(\bar{\mathbf{A}}\bar{\mathbf{D}}_i) \leq 0 \ , \quad i = 1, \cdots, N \\ & \bar{\mathbf{A}} \succeq 0 \ , \end{aligned} \tag{3.16}$$

which would provide an upper bound on the optimal value of problem (3.11), and would in general lead to a suboptimal solution for the vector $\mathbf{a}_n$ of transmission gains. However, in the following we show that the optimal solution to the original problem in (3.10) can be constructed from the solution to the relaxed SDP problem in (3.16). The optimality of a rank-relaxed SDP problem similar to the one we consider here has previously been noted in [27], but for a different problem related to physical layer security. To describe how to find the optimal solution from the rank-relaxed problem in (3.16), define $\bar{\mathbf{A}}^*$ to be the solution to (3.16), $\bar{\mathbf{A}}^*_{l,m}$ as the (l,m)th element of $\bar{\mathbf{A}}^*$, and $\bar{\mathbf{A}}^*_N$ as the Nth order leading principal submatrix of $\bar{\mathbf{A}}^*$ formed by deleting the $(N+1)$st row and column of $\bar{\mathbf{A}}^*$. Then the optimal solution can be found via the following theorem.

Theorem 1. *Define the optimal solution to problem (3.16) as $\bar{\mathbf{A}}^*$. Then $\bar{\mathbf{A}}^*_N = \mathbf{a}\mathbf{a}^H$ is rank-one and the optimal solution to problem (3.10) is given by*

$$\mathbf{a}^*_n = \frac{1}{\sqrt{\bar{\mathbf{A}}^*_{N+1,N+1}}}\mathbf{a} \ .$$

Proof: We first utilize the strong duality between problem (3.16) and its dual to find

properties of the optimal solution $\bar{\mathbf{A}}^*$. The dual of problem (3.16) is given by [28]:

$$\begin{aligned} \min_{y_i,z} \quad & z \\ s.t. \quad & \sum_{i=1}^{N} y_i \bar{\mathbf{D}}_i + z\bar{\mathbf{C}}_n - \bar{\mathbf{H}}_n \succeq 0 \\ & y_1, \ldots, y_N, z \geq 0 \ . \end{aligned} \tag{3.17}$$

It is easy to verify that there exist strictly feasible points for problems (3.16) and (3.17). In particular, for (3.16), we can construct

$$\bar{\mathbf{A}}^f = \text{diag}\{ab \cdots ab\ b\} \ ,$$

where $0 < a < \min_i \frac{P_{T,i}}{\sigma_\theta^2 + \sigma_{v,i}^2}$, and $b = \frac{1}{\sum_{i=1}^{N} a|h_{n,i}|^2\sigma_{v,i}^2 + \sigma_w^2}$.

For (3.17), we can randomly select $y_i^f > 0$, and set z^f large enough such that

$$z^f > \max\left\{ \frac{\mathbf{h}_n^H\mathbf{h}_n + \sum_{i=1}^{N} y_i^f P_{T,i}}{\sigma_w^2}, \frac{\mathbf{h}_n^H\mathbf{h}_n - y_i^f(\sigma_\theta^2 + \sigma_{v,i}^2)}{|h_{n,i}|^2\sigma_{v,i}^2} \right\} \ .$$

Then, according to Slater's theorem, strong duality holds between the primal problem (3.16) and the dual problem (3.17) and we have the following complementary condition:

$$\text{tr}(\bar{\mathbf{A}}^*\mathbf{G}^*) = 0 \ , \tag{3.18}$$

where $\mathbf{G}^* = \sum_{i=1}^{N} y_i^* \bar{\mathbf{D}}_i + z^*\bar{\mathbf{C}}_n - \bar{\mathbf{H}}_n$ and y_i^* and z^* denote the optimal solution to problem (3.17). Due to the special structure of $\bar{\mathbf{D}}_i$, $\bar{\mathbf{C}}_n$ and $\bar{\mathbf{H}}_n$, $\mathbf{G}^*$ can be expressed as

$$\mathbf{G}^* = \begin{bmatrix} \mathbf{G}_N^* & \mathbf{0} \\ \mathbf{0}^T & \mathbf{G}_{N+1,N+1}^* \end{bmatrix} \ ,$$

where $\mathbf{G}_N^* = \sum_{i=1}^{N} y_i^*\mathbf{D}_i + z^*\mathbf{H}_n\mathbf{V}\mathbf{H}_n^H - \mathbf{h}_n\mathbf{h}_n^H$ and $\mathbf{G}_{N+1,N+1}^* = z^*\sigma_w^2 - \sum_{i=1}^{N} y_i^* P_{T,i}$.

Since both $\bar{\mathbf{A}}^*$ and $\mathbf{G}^*$ are positive semidefinite, (3.18) is equivalent to

$$\bar{\mathbf{A}}^*\mathbf{G}^* = 0 \; .$$

Additionally, with consideration of the structure of $\mathbf{G}^*$, we have

$$\bar{\mathbf{A}}_N^*\mathbf{G}_N^* = 0 \; .$$

Define $\mathbf{V}_G$ as a set of vectors orthogonal to the column space of $\mathbf{G}_N^*$. Then the row vectors of $\bar{\mathbf{A}}_N^*$ must belong to span($\mathbf{V}_G$) and rank($\bar{\mathbf{A}}_N^*$) $\leq$ rank($\mathbf{V}_G$). For any two matrices $\mathbf{M}$ and $\mathbf{N}$, we have [29] that rank($\mathbf{M}+\mathbf{N}$) $\geq$ |rank($\mathbf{M}$) $-$ rank($\mathbf{N}$)|, so

$$\text{rank}(\mathbf{G}_N^*) \geq \text{rank}\left(\sum_{i=1}^{N} y_i^*\mathbf{D}_i + z^*\mathbf{H}_n\mathbf{V}\mathbf{H}_n^H\right) - \text{rank}(\mathbf{h}_n\mathbf{h}_n^H) = N-1 \; .$$

and

$$\text{rank}(\mathbf{V}_G) = N - \text{rank}(\mathbf{G}_N^*) \leq 1 \; . \tag{3.19}$$

Since $\text{tr}(\bar{\mathbf{A}}^*\bar{\mathbf{H}}) = \mathbf{h}_n^H\bar{\mathbf{A}}_N^*\mathbf{h}_n$ and $\text{tr}(\bar{\mathbf{A}}^*\bar{\mathbf{H}}) > \text{tr}(\bar{\mathbf{A}}^f\bar{\mathbf{H}}) > 0$, we have

$$\bar{\mathbf{A}}_N^* \neq 0 \; , \qquad \text{rank}(\bar{\mathbf{A}}_N^*) \geq 1 \; . \tag{3.20}$$

Combining (3.19) and (3.20) then leads to

$$\text{rank}(\bar{\mathbf{A}}_N^*) = 1 \; .$$

Although at this point we don't know whether the optimal solution $\bar{\mathbf{A}}^*$ is rank-one, we can construct a rank-one optimal solution based on $\bar{\mathbf{A}}^*$. Define the rank-one decomposition of $\bar{\mathbf{A}}_N^*$ as $\bar{\mathbf{A}}_N^* = \mathbf{a}\mathbf{a}^H$, so that the optimal rank-one solution to problem

(3.16) is

$$\bar{\mathbf{A}}' = \bar{\mathbf{a}}^* \bar{\mathbf{a}}^{*H}, \tag{3.21}$$

where $\bar{\mathbf{a}}^* = \left[\mathbf{a}^H \sqrt{\bar{\mathbf{A}}^*_{N+1,N+1}}\right]^H$. It is easy to verify that the rank-one matrix $\bar{\mathbf{A}}'$ can achieve the same result for problem (3.16) as $\bar{\mathbf{A}}^*$.

Since (3.11) is equivalent to problem (3.10) and (3.15), and (3.16) is realized from problem (3.15) by relaxing the rank-one constraint, in general the solution to (3.16) provides an upper bound on the optimal value achieved by (3.11). If the optimal solution to (3.10) is $\mathbf{a}_n^*$, then

$$\frac{\mathbf{a}_n^{*H}\mathbf{h}_n\mathbf{h}_n^H\mathbf{a}_n^*}{\mathbf{a}_n^{*H}\mathbf{H}\mathbf{V}\mathbf{H}^H\mathbf{a}_n^* + \sigma_w^2} \leq \mathrm{tr}(\bar{\mathbf{A}}^*\bar{\mathbf{H}}) \ , \tag{3.22}$$

where $\mathbf{a}_n^*$ and $\bar{\mathbf{A}}^*$ are the optimal solutions to problems (3.10) and (3.16) respectively. Equality can be achieved in (3.22) provided that an optimal rank-one solution exists for (3.16), and (3.21) indicates that such a rank-one solution exists. In the following, we will show how to construct $\mathbf{a}_n^*$ based on $\bar{\mathbf{A}}^*$. According to problem (3.16), since $\mathrm{tr}(\bar{\mathbf{A}}^*\bar{\mathbf{C}}_n) = 1$ and $\bar{\mathbf{A}} \succeq 0$, then we have $\bar{\mathbf{A}}^* \neq 0$ and further $\bar{\mathbf{A}}^*_{N+1,N+1} > 0$. Based on $\bar{\mathbf{a}}^*$, the optimal solution to (3.10) is given by

$$\mathbf{a}_n^* = \frac{\bar{\mathbf{a}}^*}{\sqrt{\bar{\mathbf{A}}^*_{N+1,N+1}}} \ , \tag{3.23}$$

and plugging (3.23) into (3.22) we have

$$\frac{\mathbf{a}_n^{*H}\mathbf{h}_n\mathbf{h}_n^H\mathbf{a}_n^*}{\mathbf{a}_n^{*H}\mathbf{H}\mathbf{V}\mathbf{H}^H\mathbf{a}_n^* + \sigma_w^2} = \mathrm{tr}(\bar{\mathbf{A}}^*\bar{\mathbf{H}}) \ ,$$

which verifies the optimality of $\mathbf{a}_n^*$. ■

3.4 Minimizing Transmit Power under an MSE Constraint

In this section, we consider the converse of the problems investigated in Section 3.3. We first look at the problem addressed in [23], where the goal is to minimize the sum power consumption of all the sensors under the constraint that the MSE is smaller than some threshold. The asymptotic behavior of the solution is then characterized for a large number of sensors, N. Next we study the case where the maximum individual transmit power of any given sensor is minimized under the MSE constraint.

3.4.1 Minimizing Sum Transmit Power

We can express the problem of minimizing the sum transmit power under the constraint that the MSE is smaller than ϵ as follows:

$$\begin{aligned} \min_{\mathbf{a}_n} \quad & \mathbf{a}_n^H \mathbf{D} \mathbf{a}_n \\ s.t. \quad & P_{n|n} \leq \epsilon \ . \end{aligned} \tag{3.24}$$

To make (3.24) feasible, according to (3.1) and (3.7) the value of ϵ should satisfy

$$\frac{P_{n|n-1}}{1 + \left(\sum_{i=1}^{N} \frac{1}{\sigma_{v,i}^2} \right) P_{n|n-1}} \leq \epsilon \leq P_{n|n-1} \ . \tag{3.25}$$

As discussed earlier, the MSE is monotonically decreasing in the norm of $\mathbf{a}_n$, so it is clear that setting $P_{n|n} = \epsilon$ results in the minimum possible transmit power, which we refer to as P_T^*. Conceptually, the problem can be solved by finding the value of P_T^*

for which $P_{n|n} = \epsilon$, and then substituting this value into the solution found in (3.4):

$$\mathbf{a}_n^* = \sqrt{\frac{P_T^*}{\mathbf{h}_n^H \mathbf{B}^{-1}\mathbf{D}\mathbf{B}^{-1}\mathbf{h}_n}}\mathbf{B}^{-1}\mathbf{h}_n \ .$$

Unlike [23], where an unspecified numerical procedure was required to solve this problem, in the following we present a direct "closed-form" solution that finds the result in terms of the eigenvalue and eigenvector of a particular matrix.

Assuming that ϵ satisfies the feasibility constraint of (3.25), we use (3.1) and $P_{n|n} = \epsilon$ to convert (3.24) to the following form:

$$\begin{aligned} \min_{\mathbf{a}_n} \quad & \mathbf{a}_n^H \mathbf{D}\mathbf{a}_n \\ s.t. \quad & \mathbf{a}_n^H \mathbf{E}_n \mathbf{a}_n \geq \left(\frac{P_{n|n-1}}{\epsilon} - 1\right)\sigma_w^2 \ , \end{aligned} \tag{3.26}$$

where $\mathbf{E}_n = P_{n|n-1}\mathbf{h}_n\mathbf{h}_n^H - \left(\frac{P_{n|n-1}}{\epsilon} - 1\right)\mathbf{H}_n\mathbf{V}\mathbf{H}_n^H$. It's obvious that the constraint in problem (3.26) should be active at the optimal solution and we can rewrite problem (3.26) as

$$\begin{aligned} \min_{\mathbf{a}_n} \quad & \frac{\mathbf{a}_n^H \mathbf{D}\mathbf{a}_n}{\mathbf{a}_n^H \mathbf{E}_n \mathbf{a}_n} \\ s.t. \quad & \mathbf{a}_n^H \mathbf{E}_n \mathbf{a}_n = \left(\frac{P_{n|n-1}}{\epsilon} - 1\right)\sigma_w^2 \ . \end{aligned} \tag{3.27}$$

Since both of $\mathbf{a}_n^H \mathbf{D}\mathbf{a}_n$ and $\mathbf{a}_n^H \mathbf{E}_n \mathbf{a}_n$ are positive, problem (3.27) is equivalent to

$$\begin{aligned} \max_{\mathbf{a}_n} \quad & \frac{\mathbf{a}_n^H \mathbf{E}_n \mathbf{a}_n}{\mathbf{a}_n^H \mathbf{D}\mathbf{a}_n} \\ s.t. \quad & \mathbf{a}_n^H \mathbf{E}_n \mathbf{a}_n = \left(\frac{P_{n|n-1}}{\epsilon} - 1\right)\sigma_w^2 \ . \end{aligned} \tag{3.28}$$

Setting $\mathbf{y} = \mathbf{D}^{\frac{1}{2}}\mathbf{a}_n$, problem (3.28) becomes a Rayleigh quotient maximization:

$$\begin{aligned} \max_{\mathbf{y}} \quad & \frac{\mathbf{y}^H \mathbf{D}^{-\frac{1}{2}} \mathbf{E}_n \mathbf{D}^{-\frac{1}{2}} \mathbf{y}}{\mathbf{y}^H \mathbf{y}} \\ s.t. \quad & \mathbf{y}^H \mathbf{D}^{-\frac{1}{2}} \mathbf{E}_n \mathbf{D}^{-\frac{1}{2}} \mathbf{y} = \left(\frac{P_{n|n-1}}{\epsilon} - 1 \right) \sigma_w^2 \ , \end{aligned}$$

whose solution is given by

$$\mathbf{y}^* = \sqrt{\frac{\left(\frac{P_{n|n-1}}{\epsilon} - 1 \right) \sigma_w^2}{\mathbf{v}_1^H \mathbf{D}^{-\frac{1}{2}} \mathbf{E}_n \mathbf{D}^{-\frac{1}{2}} \mathbf{v}_1}} \mathbf{v}_1 \ ,$$

where $\mathbf{v}_1$ denotes the unit-norm eigenvector corresponding to the largest eigenvalue of $\mathbf{D}^{-\frac{1}{2}} \mathbf{E}_n \mathbf{D}^{-\frac{1}{2}}$. The optimal solution to the original problem in (3.24) is thus

$$\mathbf{a}_n^* = \sqrt{\frac{\left(\frac{P_{n|n-1}}{\epsilon} - 1 \right) \sigma_w^2}{\mathbf{v}_1^H \mathbf{D}^{-\frac{1}{2}} \mathbf{E}_n \mathbf{D}^{-\frac{1}{2}} \mathbf{v}_1}} \mathbf{D}^{-\frac{1}{2}} \mathbf{v}_1 \ .$$

The minimum transmit power required to achieve $P_{n|n} = \epsilon$ can be expressed as

$$P_T^* = \mathbf{a}_n^{*H} \mathbf{D} \mathbf{a}_n^* = \frac{(P_{n|n-1} - \epsilon)\sigma_w^2}{\epsilon \lambda_{\max}\{\mathbf{D}^{-\frac{1}{2}} \mathbf{E}_n \mathbf{D}^{-\frac{1}{2}}\}} \ , \tag{3.29}$$

where $\lambda_{\max}(\cdot)$ represents the largest eigenvalue of its matrix argument. A more precise expression for P_T^* can be found when the number of sensors N is large, as shown in Theorem 2 below. The theorem assumes that the channel coefficients are described by the following model:

$$h_{i,n} = \frac{\tilde{h}_{i,n}}{d_i^{\gamma}} \ , \quad \tilde{h}_{i,n} \sim \mathcal{CN}(0,1) \ , \tag{3.30}$$

where d_i denotes the distance between sensor i and the FC, and γ is the propagation path-loss exponent.

Theorem 2. *Assume the channels between the sensors and FC obey the model of (3.30). When the number of sensors is large, the minimum sum transmit power P_T^* that achieves $P_{n|n} = \epsilon$ is bounded by*

$$\frac{\left(P_{n|n-1} - \epsilon\right)\sigma_w^2}{\epsilon(P_{n|n-1}\mathbf{h}_n^H\mathbf{D}^{-1}\mathbf{h}_n - \xi)} \leq P_T^* \leq \frac{\left(P_{n|n-1} - \epsilon\right)\sigma_w^2}{\epsilon P_{n|n-1}\mathbf{h}_n^H\mathbf{D}^{-1}\mathbf{h}_n(1-\zeta)} ,$$

where random variables ζ, ξ are defined as

$$\begin{aligned} \xi &= \left(\frac{P_{n|n-1}}{\epsilon} - 1\right)\min_i \left\{\frac{|h_{i,n}|^2\sigma_{v,i}^2}{\sigma_\theta^2 + \sigma_{v,i}^2}\right\} \\ \zeta &= \frac{\left(\frac{P_{n|n-1}}{\epsilon} - 1\right)\max_i \left\{\frac{|h_{i,n}|^2\sigma_{v,i}^2}{\sigma_\theta^2+\sigma_{v,i}^2}\right\}}{P_{n|n-1}\mathbf{h}_n^H\mathbf{D}^{-1}\mathbf{h}_n} , \end{aligned}$$

and ζ, ξ converge to 0 in probability.

Proof: Since $\mathbf{D}^{-\frac{1}{2}}\mathbf{E}_n\mathbf{D}^{-\frac{1}{2}}$ is the sum of a rank-one and a diagonal matrix, we have the following bounds for $\lambda_{\max}\{\mathbf{D}^{-\frac{1}{2}}\mathbf{E}_n\mathbf{D}^{-\frac{1}{2}}\}$:

$$\begin{aligned} \lambda_{\max}\{\mathbf{D}^{-\frac{1}{2}}\mathbf{E}_n\mathbf{D}^{-\frac{1}{2}}\} &\leq P_{n|n-1}\mathbf{h}_n^H\mathbf{D}^{-1}\mathbf{h}_n - \left(\frac{P_{n|n-1}}{\epsilon} - 1\right)\min_i \left\{\frac{|h_{i,n}|^2\sigma_{v,i}^2}{\sigma_\theta^2 + \sigma_{v,i}^2}\right\} \\ &= P_{n|n-1}\mathbf{h}_n^H\mathbf{D}^{-1}\mathbf{h}_n - \xi , \qquad (3.31) \\ \lambda_{\max}\{\mathbf{D}^{-\frac{1}{2}}\mathbf{E}_n\mathbf{D}^{-\frac{1}{2}}\} &\geq P_{n|n-1}\mathbf{h}_n^H\mathbf{D}^{-1}\mathbf{h}_n - \left(\frac{P_{n|n-1}}{\epsilon} - 1\right)\max_i \left\{\frac{|h_{i,n}|^2\sigma_{v,i}^2}{\sigma_\theta^2 + \sigma_{v,i}^2}\right\} \\ &= P_{n|n-1}\mathbf{h}_n^H\mathbf{D}^{-1}\mathbf{h}_n(1-\zeta) , \qquad (3.32) \end{aligned}$$

where we define $\xi = \left(\frac{P_{n|n-1}}{\epsilon} - 1\right)\min_i \left\{\frac{|h_{i,n}|^2\sigma_{v,i}^2}{\sigma_\theta^2+\sigma_{v,i}^2}\right\}$, and $\zeta = \frac{\left(\frac{P_{n|n-1}}{\epsilon}-1\right)\max_i\left\{\frac{|h_{i,n}|^2\sigma_{v,i}^2}{\sigma_\theta^2+\sigma_{v,i}^2}\right\}}{P_{n|n-1}\mathbf{h}_n^H\mathbf{D}^{-1}\mathbf{h}_n}$.

For any positive constant ν, we have

$$\begin{aligned}
\Pr\{\xi \geq \nu\} &\leq \Pr\left\{\eta \min_i \left\{|\tilde{h}_{i,n}|^2\right\} \geq \tilde{\nu}\right\} \\
&= \Pr\left\{\min_i \left\{|\tilde{h}_{i,n}|^2\right\} \geq \frac{\tilde{\nu}}{\eta}\right\} \\
&= \left(1 - \Pr\left\{|\tilde{h}_{i,n}|^2 \leq \frac{\tilde{\nu}}{\eta}\right\}\right)^N \\
&\overset{(b)}{=} e^{-\frac{N\tilde{\nu}}{\eta}} ,
\end{aligned}$$

where $\tilde{\nu} = \frac{\nu}{\frac{P_{n|n-1}}{\epsilon}-1}$, $\eta = \max_i \left\{\frac{\sigma_{v,i}^2}{(\sigma_{v,i}^2+\sigma_\theta^2)d_i^\gamma}\right\}$ and (b) is due to the fact that $2|\tilde{h}_{i,n}|^2$ is a chi-square random variable with degree 2. When $N \to \infty$, we have

$$\lim_{N\to\infty} \Pr\{\xi \geq \nu\} \leq \lim_{N\to\infty} e^{-\frac{N\tilde{\nu}}{\eta}} = 0 ,$$

and thus ξ converges to 0 in probability.

From the definition of ζ,

$$\begin{aligned}
\zeta &= \left(\frac{P_{n|n-1}-\epsilon}{P_{n|n-1}\epsilon}\right) \max_i \left\{\frac{\frac{|h_{i,n}|^2\sigma_{v,i}^2}{\sigma_\theta^2+\sigma_{v,i}^2}}{\sum_{k=1}^N \frac{|h_{k,n}|^2}{\sigma_\theta^2+\sigma_{v,k}^2}}\right\} \\
&\leq \left(\frac{P_{n|n-1}-\epsilon}{P_{n|n-1}\epsilon}\right) \max_i \left\{\frac{\frac{|h_{i,n}|^2\sigma_{v,i}^2}{\sigma_\theta^2+\sigma_{v,i}^2}}{\sum_{k=1,k\neq i}^N \frac{|h_{k,n}|^2}{\sigma_\theta^2+\sigma_{v,k}^2}}\right\} \\
&\leq \tau \max_i \left\{\frac{|\tilde{h}_{i,n}|^2}{\sum_{k=1,k\neq i}^N |\tilde{h}_{k,n}|^2}\right\},
\end{aligned}$$

where

$$\tau = \left(\frac{P_{n|n-1}-\epsilon}{P_{n|n-1}\epsilon}\right) \frac{\max_i \left\{\frac{\sigma_{v,i}^2}{(\sigma_\theta^2+\sigma_{v,i}^2)d_i^\gamma}\right\}}{\min_i \left\{\frac{1}{(\sigma_\theta^2+\sigma_{v,i}^2)d_i^\gamma}\right\}} .$$

For any positive constant μ, we have

$$\begin{aligned}
\Pr\{\zeta \geq \mu\} &= 1 - \Pr\{\zeta \leq \mu\} \\
&\leq 1 - \Pr\left\{\max_i \left\{\frac{|\tilde{h}_{i,n}|^2}{\sum_{k=1,k\neq i}^{N} |\tilde{h}_{k,n}|^2}\right\} \leq \tilde{\mu}\right\} \\
&= 1 - \left(\Pr\left\{\frac{|\tilde{h}_{i,n}|^2}{\sum_{k=1,k\neq i}^{N} |\tilde{h}_{k,n}|^2} \leq \tilde{\mu}\right\}\right)^N \\
&= 1 - \left(\Pr\left\{\frac{\sum_{k=1,k\neq i}^{N} |\tilde{h}_{k,n}|^2}{|\tilde{h}_{i,n}|^2} \geq \frac{1}{\tilde{\mu}}\right\}\right)^N \\
&= 1 - \left(\Pr\left\{\frac{\sum_{k=1,k\neq i}^{N} |\tilde{h}_{k,n}|^2}{(N-1)|\tilde{h}_{i,n}|^2} \geq \frac{1}{(N-1)\tilde{\mu}}\right\}\right)^N, \qquad (3.33)
\end{aligned}$$

where $\tilde{\mu} = \mu/\tau$.

In (3.33), the random variable $X = \frac{\sum_{k=1,k\neq i}^{N} |\tilde{h}_{k,n}|^2}{(N-1)|\tilde{h}_{i,n}|^2}$ has an F-distribution with parameters $N-1$ and 2. Thus, the cumulative density function of X is given by [28]

$$F(x) = \left(\frac{(N-1)x}{(N-1)x+1}\right)^{N-1},$$

and thus

$$\begin{aligned}
\left(\Pr\left\{X \geq \frac{1}{(N-1)\tilde{\mu}}\right\}\right)^N &= \left(1 - \Pr\left\{X \leq \frac{1}{(N-1)\tilde{\mu}}\right\}\right)^N \\
&= \left(1 - \frac{1}{(1+\tilde{\mu})^{N-1}}\right)^N \\
&= \left(1 - \frac{1}{(1+\tilde{\mu})^{N-1}}\right)^{(1+\tilde{\mu})^{N-1}\frac{N}{(1+\tilde{\mu})^{N-1}}}.
\end{aligned}$$

Since $\tilde{\mu} > 0$ and hence $\lim_{N\to\infty}(1+\tilde{\mu})^{N-1} = \infty$, we have

$$\begin{aligned}
\lim_{N\to\infty}\left(\Pr\left\{X \geq \frac{1}{(N-1)\tilde{u}}\right\}\right)^N &= \lim_{N\to\infty}\left(1 - \frac{1}{(1+\tilde{\mu})^{N-1}}\right)^{(1+\tilde{u})^{N-1}\frac{N}{(1+\tilde{\mu})^{N-1}}} \\
&= \lim_{N\to\infty} e^{\frac{N}{(1+\tilde{u})^{N-1}}}.
\end{aligned}$$

Furthermore,

$$\lim_{N\to\infty} \frac{N}{(1+\tilde{\mu})^{N-1}} = \lim_{N\to\infty} \frac{1}{(1+\tilde{\mu})^{N-1}\ln(1+\tilde{\mu})} = 0 \ ,$$

and thus

$$\lim_{N\to\infty} \left(\Pr\left\{ X \geq \frac{1}{(N-1)\tilde{\mu}} \right\} \right)^N = 1 \ . \tag{3.34}$$

Substituting (3.34) into (3.33) yields

$$\lim_{N\to\infty} \Pr\left\{ \zeta \geq \mu \right\} = 0 \ , \tag{3.35}$$

and we conclude that when $N \to \infty$, ζ converges to 0 in probability. The proof of the theorem is completed by substituting the results of (3.31), (3.32), (3.33) and (3.35) into (3.29). ■

According to the above theorem, when $N \to \infty$, the term $P_{n|n-1}\mathbf{h}_n^H \mathbf{D}^{-1}\mathbf{h}_n$ is the dominant factor in the denominator of the bounds on the sum transmit power, and we have the following asymptotic expression

$$\lim_{N\to\infty} P_T^* \simeq \frac{(P_{n|n-1}-\epsilon)\sigma_w^2}{\epsilon P_{n|n-1}\mathbf{h}_n^H \mathbf{D}^{-1}\mathbf{h}_n} \ . \tag{3.36}$$

This expression illustrates that to achieve the same MSE, increasing the number of sensors reduces the total required transmit power of the network, as well as the required transmit power per sensor. A similar observation was made in [23]. As shown later, our simulation results show that (3.36) provides an accurate approximation to (3.29) as long as ϵ is not too small.

As a final comment on this problem, we note that (3.24) is equivalent to

$$\begin{aligned} \min_{\mathbf{A}} \quad & \mathrm{tr}(\mathbf{AD}) \\ s.t. \quad & \mathrm{tr}(\mathbf{AE}_n) \geq \left(\frac{P_{n|n-1}}{\epsilon} - 1 \right) \sigma_w^2 \\ & \mathrm{rank}(\mathbf{A}) = 1 \\ & \mathbf{A} \succeq 0 \end{aligned} \tag{3.37}$$

for $\mathbf{A} = \mathbf{a}_n \mathbf{a}_n^H$. Relaxing the rank-one constraint on $\mathbf{A}$, problem (3.37) becomes

$$\begin{aligned} \min_{\mathbf{A}} \quad & \mathrm{tr}(\mathbf{AD}) \\ s.t. \quad & \mathrm{tr}(\mathbf{AE}_n) \geq \left(\frac{P_{n|n-1}}{\epsilon} - 1 \right) \sigma_w^2 \\ & \mathbf{A} \succeq 0 \; . \end{aligned} \tag{3.38}$$

Based on the complementary conditions between the dual and primal problems, we can prove that the solution to (3.38) is rank one, and hence that the relaxed SDP yields the optimal $\mathbf{a}_n^*$.

3.4.2 Minimizing Maximum Individual Transmit Power

Here we focus on the problem of minimizing the maximum transmit power of the individual sensors while attempting to meet an MSE objective:

$$\begin{aligned} \min_{\mathbf{a}_n} \max_{i} \quad & |a_{i,n}|^2 (\sigma_\theta^2 + \sigma_{v,i}^2) \\ s.t. \quad & P_{n|n} \leq \epsilon \; . \end{aligned} \tag{3.39}$$

As in Section 3.3.2, we will convert the problem to a rank-relaxed SDP whose solution nonetheless obeys the rank constraint and hence provides the optimal result. To

proceed, introduce an auxiliary variable t and rewrite (3.39) as

$$\begin{aligned} \min_{\mathbf{a}_n, t} \quad & t \\ s.t. \quad & P_{n|n} \leq \epsilon \\ & |a_{i,n}|^2(\sigma_\theta^2 + \sigma_{v,i}^2) \leq t \ , \quad i = 1, \cdots, N \ . \end{aligned} \tag{3.40}$$

Problem (3.40) is equivalent to

$$\begin{aligned} \min_{\mathbf{A}, t} \quad & t \\ s.t. \quad & \mathrm{tr}(\mathbf{A}\mathbf{E}_n) - \left(\frac{P_{n|n-1}}{\epsilon} - 1 \right) \sigma_w^2 \geq 0 \\ & \mathrm{tr}(\mathbf{A}\mathbf{D}_i) - t \leq 0 \ , \ i = 1, \cdots, N \\ & \mathbf{A} \succeq 0 \\ & \mathrm{rank}\,(\mathbf{A}) = 1 \ , \end{aligned} \tag{3.41}$$

where $\mathbf{A} = \mathbf{a}_n \mathbf{a}_n^H$, $\mathbf{E}_n$ is defined as in (3.26), and $\mathbf{D}_i = \mathrm{diag}\{0 \cdots 0 \ \sigma_\theta^2 + \sigma_{v,i}^2 \ 0 \cdots 0\}$, as before.

Relaxing the rank constraint and rewriting the problem to be in standard form, problem (3.41) becomes

$$\begin{aligned} \min_{\tilde{\mathbf{A}}} \quad & \mathrm{tr}(\tilde{\mathbf{A}}\mathbf{T}) \\ s.t. \quad & \mathrm{tr}(\tilde{\mathbf{A}}\tilde{\mathbf{E}}_n) - \left(\frac{P_{n|n-1}}{\epsilon} - 1 \right) \sigma_w^2 \geq 0 \\ & \mathrm{tr}(\tilde{\mathbf{A}}\tilde{\mathbf{D}}_i) \leq 0 \ , \quad i = 1, \cdots, N \\ & \tilde{\mathbf{A}} \succeq 0 \ , \end{aligned} \tag{3.42}$$

where

$$\tilde{\mathbf{A}} = \begin{bmatrix} \mathbf{A} & \mathbf{w} \\ \mathbf{w}^H & t \end{bmatrix}, \qquad \mathbf{T} = \begin{bmatrix} \mathbf{0}_N & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix}, \qquad \tilde{\mathbf{E}}_n = \begin{bmatrix} \mathbf{E}_n & \mathbf{0} \\ \mathbf{0}^T & 0 \end{bmatrix}, \qquad \tilde{\mathbf{D}}_i = \begin{bmatrix} \mathbf{D}_i & \mathbf{0} \\ \mathbf{0}^T & -1 \end{bmatrix},$$

and $\mathbf{w}$ is otherwise arbitrary. Theorem 3 establishes that the optimal solution to (3.39) can be constructed from the solution to the above relaxed SDP.

Theorem 3. *Define the optimal solution to problem (3.42) as* $\tilde{\mathbf{A}}^*$. *Then* $\tilde{\mathbf{A}}_N^* = \tilde{\mathbf{a}}\tilde{\mathbf{a}}^H$ *is rank-one and the optimal solution to problem (3.39) is given by* $\mathbf{a}_n^* = \tilde{\mathbf{a}}$.

Proof: The dual of problem (3.42) is given by

$$\begin{aligned} \max_{y_i, z} \quad & \left(\frac{P_{n|n-1}}{\epsilon} - 1 \right) \sigma_w^2 z \\ s.t. \quad & \mathbf{T} + \sum_{i=1}^{N} y_i \tilde{\mathbf{D}}_i - z\tilde{\mathbf{E}}_n \succeq 0 \\ & y_1, \ldots, y_N, z \geq 0 \; . \end{aligned} \tag{3.43}$$

Using an approach similar to the proof of Theorem 1, one can verify that both (3.42) and (3.43) are strictly feasible, and that strong duality holds between the dual problem (3.43) and the primal problem (3.42). Based on the complementary conditions, it can be shown that $\text{rank}(\tilde{\mathbf{A}}_N^*) = 1$. For brevity the details of the proof are omitted.■

Similar to problems (3.2) and (3.24), duality also exists between (3.10) and (3.39). Define the optimal solution to problem (3.10) as $\mathbf{a}_n^*$ and the corresponding minimum MSE as $P_{n|n}^*$. If we set $\epsilon = P_{n|n}^*$ in (3.39), the optimal solution is also $\mathbf{a}_n^*$.

3.5 MSE Outage Probability for Equal Power Allocation

Here we calculate the MSE outage probability for the suboptimal solution in which each sensor transmits with the same power. The outage probability derived here can serve as an upper bound for the outage performance of the optimal algorithm with individual power constraints. For equal-power transmission, the transmit gain vector is given by

$$\mathbf{a}_e = \sqrt{\frac{P_T}{N}} \left[\frac{1}{\sqrt{\sigma_\theta^2 + \sigma_{v,1}^2}} \cdots \frac{1}{\sqrt{\sigma_\theta^2 + \sigma_{v,N}^2}} \right]^T ,$$

and the corresponding MSE is

$$P_{n|n} = \left(1 - \frac{P_{n|n-1} \mathbf{a}_e^H \mathbf{h}_n \mathbf{h}_n^H \mathbf{a}_e}{\mathbf{a}_e^H \mathbf{H}_n \mathbf{V} \mathbf{H}_n \mathbf{a}_e + P_{n|n-1} \mathbf{a}_e^H \mathbf{h}_n \mathbf{h}_n^H \mathbf{a}_e + \sigma_w^2} \right) P_{n|n-1} \, .$$

As in Theorem 2, we will assume the Gaussian channel model of (3.30). The outage probability $P_{out} = \Pr\left\{ P_{n|n} > \epsilon \right\}$ is evaluated as follows:

$$\begin{aligned} P_{out} &= \Pr\left\{ \frac{\mathbf{a}_e^H \mathbf{h}_n \mathbf{h}_n^H \mathbf{a}_e}{\mathbf{a}_e^H \mathbf{H}_n \mathbf{V} \mathbf{H}_n^H \mathbf{a}_e + \sigma_w^2} < \frac{P_{n|n-1} - \epsilon}{\epsilon P_{n|n-1}} \right\} \\ &= \Pr\left\{ \mathbf{a}_e^H \mathbf{h}_n \mathbf{h}_n^H \mathbf{a}_e - \beta \mathbf{a}_e^H \mathbf{H}_n \mathbf{V} \mathbf{H}_n^H \mathbf{a}_e < \beta \sigma_w^2 \right\} \\ &= \Pr\left\{ \tilde{\mathbf{h}}_n^H \left(\mathbf{M} \tilde{\mathbf{a}}_e \tilde{\mathbf{a}}_e^H \mathbf{M} - \beta \mathbf{Q} \right) \tilde{\mathbf{h}}_n \leq \frac{\beta \sigma_w^2}{P_T} \right\} , \end{aligned}$$

where

$$\begin{aligned} \beta &= \frac{P_{n|n-1} - \epsilon}{\epsilon P_{n|n-1}} \, , \; \tilde{\mathbf{a}}_e = \frac{1}{\sqrt{P_T}} \mathbf{a}_e \, , \; \tilde{\mathbf{h}}_n = [\tilde{h}_{1,n} \cdots \tilde{h}_{N,n}]^T \, , \\ \mathbf{M} &= \operatorname{diag}\left\{ \frac{1}{d_1^\alpha} \cdots \frac{1}{d_N^\alpha} \right\} \, , \\ \mathbf{Q} &= \operatorname{diag}\left\{ \frac{\sigma_{v,1}^2}{N(\sigma_\theta^2 + \sigma_{v,1}^2) d_1^{2\alpha}} \; \cdots \; \frac{\sigma_{v,N}^2}{N(\sigma_\theta^2 + \sigma_{v,N}^2) d_N^{2\alpha}} \right\} \, . \end{aligned}$$

If we define $\mathbf{R} = \mathbf{M}\tilde{\mathbf{a}}_e\tilde{\mathbf{a}}_e^H\mathbf{M} - \beta\mathbf{Q}$, and label the eigenvalues of $\mathbf{R}$ as $\lambda_1, \cdots, \lambda_N$, then the random variable $\tilde{\mathbf{h}}_n^H\mathbf{R}\tilde{\mathbf{h}}_n$ can be viewed as the weighted sum of independent chi-squared random variables $\sum_{i=1}^N \frac{\lambda_i}{2}\chi_i(2)$. From [30], we have

$$P_{out} = 1 - \sum_{i=1}^{N} \frac{\lambda_i^N}{\prod_{l\neq i}(\lambda_i - \lambda_l)} \frac{1}{|\lambda_i|} e^{-\frac{(P_{n|n-1}-\epsilon)\sigma_w^2}{\epsilon P_{n|n-1} P_T \lambda_i}} u(\lambda_i) \ , \tag{3.44}$$

where $u(\cdot)$ is the unit step function. Let $e_1 \geq \cdots \geq e_N$ denote the eigenvalues of $\mathbf{Q}$, so that from Weyl's inequality [31] we have the following bounds for the λ_i:

$$\tilde{\mathbf{a}}_e^H\mathbf{M}^2\tilde{\mathbf{a}}_e - \beta e_1 \leq \lambda_1 \leq \tilde{\mathbf{a}}_e^H\mathbf{M}^2\tilde{\mathbf{a}}_e - \beta e_N \ , \tag{3.45}$$

$$-\beta e_{N-i+1} \leq \lambda_i \leq -\beta e_{N-i+2} \ , \quad 2 \leq i \leq N \ , \tag{3.46}$$

where $\tilde{\mathbf{a}}_e^H\mathbf{M}^2\tilde{\mathbf{a}}_e = \sum_{i=1}^N \frac{1}{N(\sigma_\theta^2+\sigma_{v,i}^2)d_i^{2\alpha}}$. From (3.45), when β is large, λ_1 is negative, and when β is small enough, λ_1 is positive. Meanwhile, since all the eigenvalues of $\mathbf{Q}$ is positive, then according to (3.46) we have that $\lambda_i < 0$ for $2 \leq i \leq N$. Since only λ_1 can be positive, equation (3.44) can be simplified as

$$P_{out} = \begin{cases} 1 - \frac{\lambda_1^{N-1}}{\prod_{l\neq 1}(\lambda_1-\lambda_l)} e^{-\frac{(P_{n|n-1}-\epsilon)\sigma_w^2}{\epsilon P_{n|n-1} P_T \lambda_1}} & \lambda_1 > 0 \\ 1 & \lambda_1 \leq 0 \ . \end{cases} \tag{3.47}$$

From (3.47), when the threshold ϵ is too small, $\beta = \frac{P_{n|n-1}-\epsilon}{\epsilon P_{n|n-1}}$ will be very large and $\lambda_1 \leq 0$, then the outage probability P_{out} equals 1, which means the MSE $P_{n|n}$ is larger than ϵ for every channel realization $\mathbf{h}_n$. For $P_T \to \infty$, the outage probability converges to

$$P_{out} = \begin{cases} 1 - \frac{\lambda_1^{N-1}}{\prod_{l\neq 1}(\lambda_1-\lambda_l)} & \lambda_1 > 0 \\ 1 & \lambda_1 \leq 0 \ . \end{cases}$$

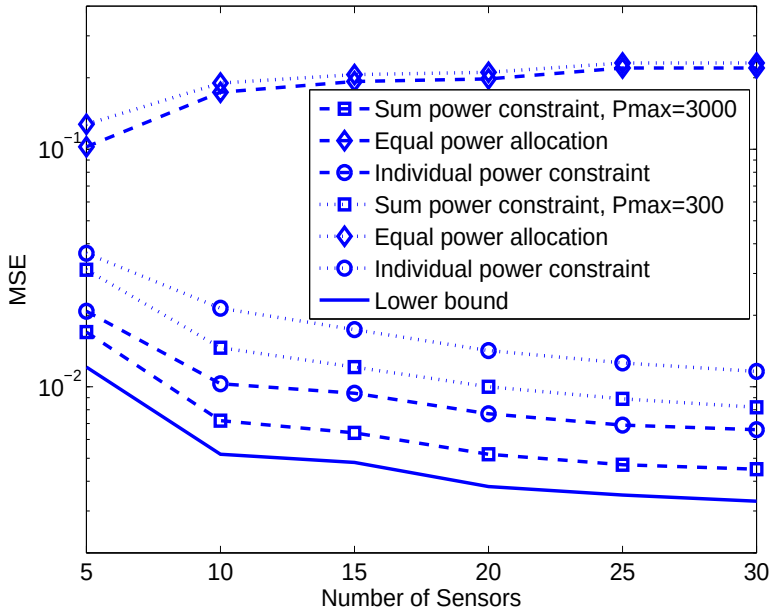

Figure 3.1: MSE vs. number of sensors for $P_{\max} = 300$ or 3000.

3.6 Simulation Results

To investigate the performance of the proposed optimization approaches, the results of several simulation examples are described here. Unless otherwise indicated, the simulations are implemented with the following parameters: distance from the FC to the sensors d_i is uniformly distributed over the interval $[2, 8]$, path loss exponent is set to $\gamma = 1$, the observation noise power $\sigma_{v,i}^2$ at the sensors is uniformly distributed over $[0, 0.5]$, the power of the additive noise at the FC is set to $\sigma_w^2 = 0.5$, the parameter θ is assumed to satisfy $\sigma_\theta^2 = 1$, and the initial MSE is given by $P_{0|-1} = 0.5$. The MSE shown in the plots is obtained by averaging over 300 realizations of $\mathbf{h}_n$. Two different sum power constraints are considered in the simulations: $P_T = 300$ and $P_T = 3000$. To fairly compare the results under sum and individual power constraints, we set $P_{T,i} = \frac{P_T}{N}$, which means that all sensors have the same maximum power when individual constraints are imposed.

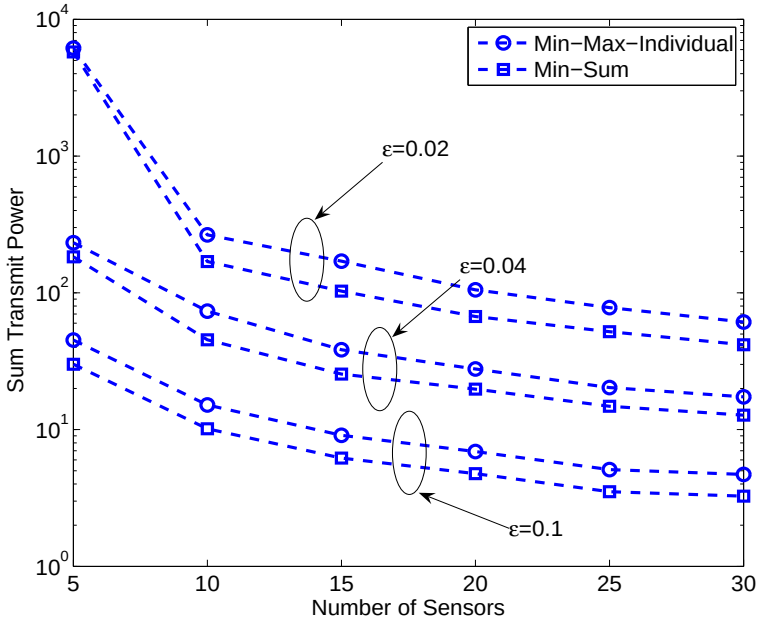

Figure 3.2: Required sum transmit power vs. number of sensors for various MSE constraints.

Fig. 3.1 plots the MSE as a function of the number of sensors in the network for both sum and individual power constraints. The results demonstrate that compared with equal power allocation, the optimized power allocation significantly reduces the MSE; in fact, adding sensors with equal power allocation actually increases the MSE, while the MSE always decreases for the optimal methods. The extra flexibility offered by the global power constraint leads to better performance compared with individual power constraints, but the difference in this case is not large. The lower bound on MSE in (3.7) is also plotted to indicate the performance that that could be achieved with $P_T \to \infty$.

Figs. 3.2 and 3.3 respectively examine sum and peak transmit powers required to achieve MSE values of 0.02, 0.04 and 0.1 for varying numbers of sensors. As expected, individual power constraints lead to higher sum power requirements, while sum power constraints result in higher peak power. Interestingly, the individual power constraints lead to roughly a doubling of the required total sum power to achieve the same

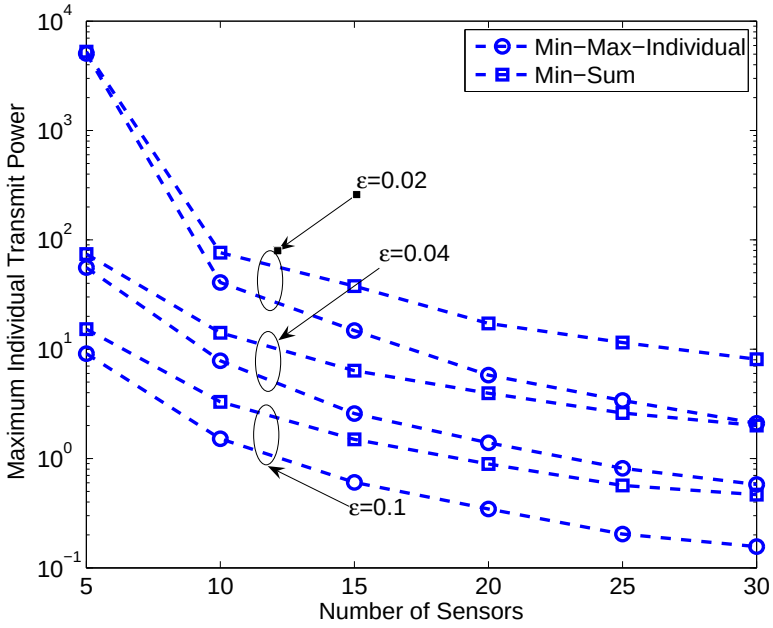

Figure 3.3: Maximum individual transmit power for various MSE constraints.

MSE regardless of the number of sensors, whereas the increase in peak power for the sum constraint relative to individual power constraints grows with N, reaching a factor of 4 to 5 on average when $N = 30$. Fig. 3.4 compares the minimum required sum transmit power to achieve various MSE values in (3.29) with the approximate expression obtained in (3.36). When $\epsilon \geq 0.1$, the approximation is reasonably good even when N is on the order of only 20 to 40. The approximation is less accurate for tighter requirements on ϵ, and requires a larger value of N for the approximation to be valid.

The impact of the SNR at the FC on the sensor power allocation is illustrated in Fig. 3.5 for a given channel realization and $N = 30$ sensors. The x-axis of each plot is ordered according to the channel gain of the sensors, which is shown in the upper left subfigure. The upper right subfigure shows the variance of the measurement noise for each sensor, which for this example was uniformly drawn from the interval $[0.4, 0.5]$ to better illustrate the effect of the channel gain. The optimal power allocation

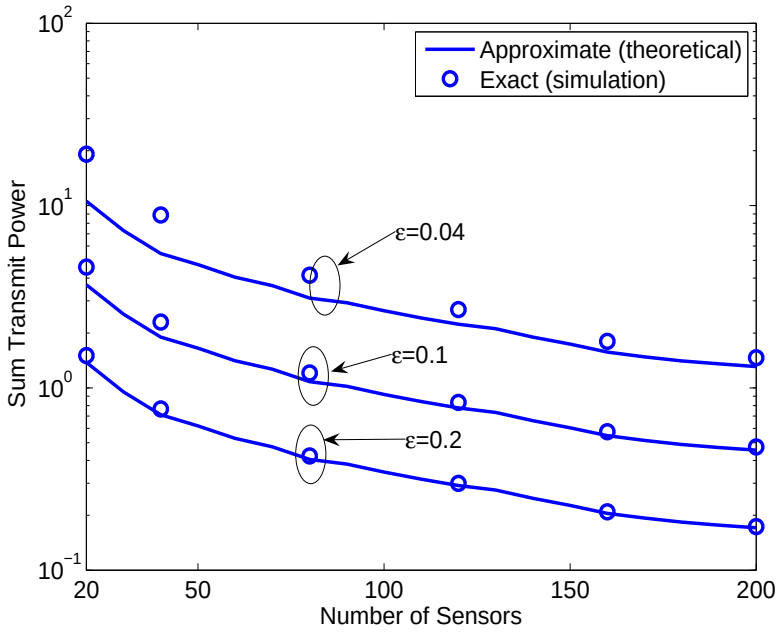

Figure 3.4: Exact and approximate sum transmit power vs. number of sensors.

for this scenario was found assuming both sum and individual power constraints under both low and high SNRs at the FC. The middle subfigures show the power allocation for minimizing MSE assuming a low SNR case with $P_T = 5$, while the bottom subfigures show the allocation for high SNR with $P_T = 1000$. Note that, as predicted by (3.9), the power allocated to the sensors under the sum power constraint for low SNR tends to grow with the channel gain, while as predicted by (3.8), the allocated power is reduced with channel gain under high SNR. The explanation for the different behavior at low and high SNR can be explained as follows: when the SNR is high, the measurement noise will dominate the estimation error at the FC, and the higher the channel gain, the more the measurement noise is amplified, so the sensor nodes with higher channel gains will be allocated less power. When the SNR is low, the additive noise at the FC will dominate the estimation error, the effect of the measurement noise can be neglected, so the nodes with higher channel gains will be allocated more power to increase the power of the desired signal. For individual power constraints, we see that all of the sensors transmit with a maximum power of

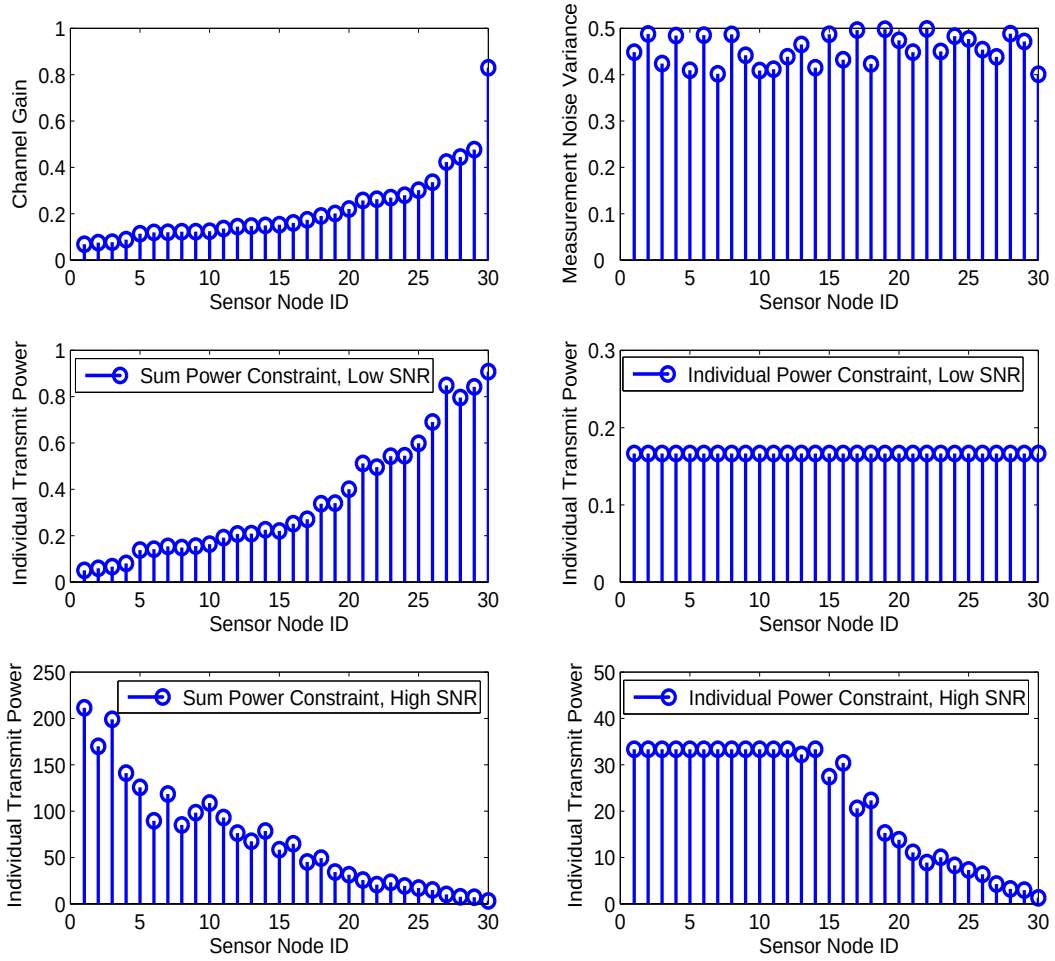

Figure 3.5: Stem plot for the channel gain, measurement noise variance and the individual transmit power allocated to the sensor nodes. The x-axis denotes the sensor node ID and the sensor nodes are indexed according to their channel gain, in ascending order. For the high SNR case the total transmit power is set to $P_T = 1000$ and for the low SNR case the total transmit power is $P_T = 5$.

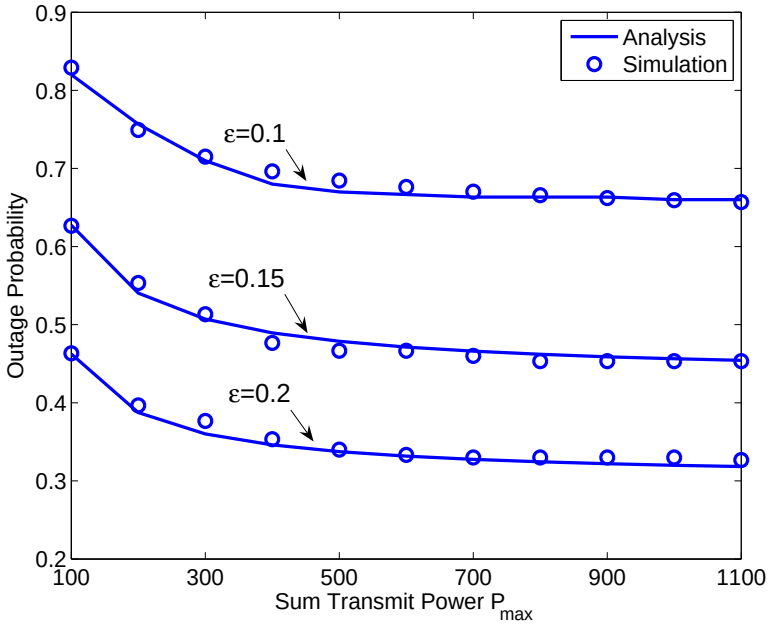

Figure 3.6: MSE outage probability for equal power allocation vs. sum transmit power for $N = 10$ sensors.

$P_T/N = 5/30$ at low SNR, while at high SNR only the sensors with small channel gains use maximum power (in this case $P_T/N = 1000/30$), and the power allocated to sensors with large channel gains decreases, as with the sum power constraint.

Finally, in Fig. 3.6, we show that our analytical expression in (3.47) for the outage probability under equal power allocation closely follows the simulation results for various transmit power levels for a case with $N = 10$ sensors. While these outage probabilities represent upper bounds for the optimal (and generally unequal) transmission gains, we note that these bounds are not particularly tight. The outage probabilities achieved by the optimal algorithms are typically much lower than predicted by (3.47).

3.7 Summary

In this chapter, we considered the problem of optimally allocating power in an analog sensor network attempting to track a dynamic parameter via a coherent multiple access channel. We analyzed problems with either constraints on power or constraints on achieved MSE, and we also examined cases involving global sum and individual sensor power constraints. While prior work had been published for minimizing MSE under a sum power constraint and minimizing sum power under an MSE constraint, we were able to derive closed-form solutions that were simpler and more direct. Going beyond the prior work, we derived new asymptotic expressions for the transmission gains that illustrated their limiting behavior for both low and high SNR at the fusion center, and we found a simple expression for the required sum transmit power when the number of sensors is large. Furthermore, we showed how to minimize MSE under individual power constraints, or minimize peak sensor power under MSE constraints, cases that had not been previously considered. In particular, we demonstrated that solutions to these problems could be found by solving a rank-relaxed SDP using standard convex optimization methods. Finally, we derived an exact expression for the MSE outage probability for the special case where the sensors transmit with equal power, and presented a number of simulation results that confirmed our analysis and the performance of the proposed algorithms.

Chapter 4

Power Allocation for Tracking of a Vector Process

4.1 Introduction

Dynamic state-space models in which the observation equation depends on parameters that can be adaptively tuned to improve performance have recently been proposed by several authors. For example, in [32], dynamic wireless channel parameters such as the delay of arrival, the angle of arrival, the angle of departure, etc, are tracked via a Kalman filter whose performance depends on properties of the antenna array. In [33], the parameters to be estimated are the position and velocity of a target and the observations at a set of mobile sensors are the time delay and Doppler shift of the signal reflected by the target. The positions of the mobile sensors are adjusted in order to minimize the tracking error of a standard extended Kalman filter. In [34], a distributed sensor network problem is considered where the observed signal is a linear function of the transmission gain of each sensor, and optimal values for these gains

are found under different power constraints to minimize the MSE of a scalar variable at the fusion center.

In this chapter, we consider a scenario that generalizes the one assumed in [34] by allowing the estimated parameters and the observations to be vector- rather than scalar-valued. The state-space observation matrix is assumed to depend linearly on a set of parameters, and we consider the problem of optimizing these parameters in order to minimize the MSE obtained by a Kalman filter that tracks the unknown state. Two different optimization problems are considered: one that minimizes the sum MSE (Min-Sum-MSE) over all the parameters, and another that minimizes the maximum MSE (Min-Max-MSE) of all parameters. In the general case, we divide the overall problem into two sub-problems whose optimal solution can be found. The first sub-problem estimates the optimal observation matrix without taking the linear structure into account, and the second finds the set of parameters that are closest to the resulting observation matrix under a quadratic constraint on the parameters themselves. Dividing the problem into these two steps will cause a performance loss, but simulations demonstrate that the loss is minimal, and performance is close to the lower bound given by the solution to the unconstrained problem. We also consider the special case of a scalar observation, and show that in this case the Min-Sum-MSE problem is converted to a Rayleigh quotient maximization problem, for which an optimal closed-form solution is obtained, and we show that for the Min-Max-MSE problem, a relaxed version of the problem leads to a simple SDP feasibility test that can be solved via the bisection algorithm. Simulation results show that in most cases, the solution to the relaxed and unrelaxed problems are the same.

The rest of this chapter is organized as follows. Section 4.2 describes the signal model for the vector observation case and formulates the sum MSE minimization problem and maximum MSE minimization problem. The system model and problem

formulation for the scalar observation case are presented in Section 4.3. Numerical results are provided in Section 4.4 and the conclusions are summarized in Section 4.5.

4.2 Vector Observation Model

We assume the dynamic parameter to be estimated is a complex-valued vector that obeys the following state-space model:

$$\boldsymbol{\theta}_{n+1} \;=\; \mathbf{F}\boldsymbol{\theta}_n + \mathbf{u}_n, \tag{4.1}$$

where $\boldsymbol{\theta}_{n+1} \in \mathbb{C}^{M\times 1}$ is the parameter at time step $n+1$, $\mathbf{F} \in \mathbb{C}^{M\times M}$ is the state transition matrix, and $\mathbf{u}_n \sim \mathcal{CN}(\mathbf{0}, \mathbf{Q})$ is the process noise. The observed signal vector is given by

$$\mathbf{y}_n = \mathbf{C}\boldsymbol{\theta}_n + \mathbf{v}_n, \tag{4.2}$$

where $\mathbf{v}_n \sim \mathcal{CN}(\mathbf{0}, \sigma_v^2\mathbf{I})$ is the observation noise and $\mathbf{C} \in \mathbb{C}^{L\times M}$ is the observation matrix. We assume that $\mathbf{C}$ is a linear function of some parameters $\mathbf{a} \in \mathbb{C}^{N\times 1}$ such that $\text{vec}[\mathbf{C}] = \mathbf{G}\mathbf{a}$ for a given $\mathbf{G} \in \mathbb{C}^{LM\times N}$.

The MSE of the state estimate is found via the standard Kalman filtering [25]:

- Prediction MSE Matrix

$$\mathbf{M}_{n|n-1} = \mathbf{F}\mathbf{M}_{n-1|n-1}\mathbf{F}^H + \mathbf{Q} \tag{4.3}$$

- Kalman Gain Matrix

$$\mathbf{K}_n = \mathbf{M}_{n|n-1}\mathbf{C}^H(\sigma_v^2\mathbf{I}_L + \mathbf{C}\mathbf{M}_{n|n-1}\mathbf{C}^H)^{-1} \tag{4.4}$$

- MSE matrix

$$\begin{aligned}\mathbf{M}_{n|n} &= (\mathbf{I}_M - \mathbf{K}_n\mathbf{C})\mathbf{M}_{n|n-1} \\ &= \left(\mathbf{M}_{n|n-1}^{-1} + \frac{1}{\sigma_v^2}\mathbf{C}^H\mathbf{C}\right)^{-1} . \end{aligned} \tag{4.5}$$

4.2.1 Minimize Sum MSE

In this section we consider the problem of minimizing the sum-MSE under a quadratic constraint of $\mathbf{a}$. The ideal optimization problem is formulated as

$$\begin{aligned} \min_{\mathbf{a}} \quad & \mathrm{tr}(\mathbf{M}_{n|n}) \\ s.t. \quad & \|\mathbf{C}(\mathbf{a})\|_F^2 \leq P . \end{aligned} \tag{4.6}$$

The solution to (4.6) is difficult to obtain directly, so instead we divide the optimization problem into two subproblems. We first find an unconstrained $\mathbf{C}^*$ that minimizes $\mathrm{tr}(\mathbf{M}_{n|n})$, and then based on $\mathbf{C}^*$, we obtain the approximate solution $\mathbf{a}^*$.

The first step is to solve

$$\begin{aligned} \min_{\mathbf{C}} \quad & \mathrm{tr}(\mathbf{M}_{n|n}) \\ s.t. \quad & \|\mathbf{C}\|_F^2 \leq P . \end{aligned} \tag{4.7}$$

Defining $\tilde{\mathbf{C}} = \mathbf{C}^H\mathbf{C}$, we can rewrite (4.7) as

$$\begin{aligned} \min_{\tilde{\mathbf{C}}} \quad & \mathrm{tr}(\mathbf{D}) \\ s.t. \quad & \mathbf{D}^{-1} = \mathbf{M}_{n|n-1}^{-1} + \frac{1}{\sigma_v^2}\tilde{\mathbf{C}} \\ & \mathrm{tr}(\tilde{\mathbf{C}}) \leq P \\ & \tilde{\mathbf{C}} \succeq 0. \end{aligned} \tag{4.8}$$

Replacing the equality in the first constraint of (4.8) with an inequality yields an equivalent optimization problem:

$$\begin{aligned} \min_{\tilde{\mathbf{C}},\mathbf{D}} \quad & \mathrm{tr}(\mathbf{D}) \\ s.t. \quad & \mathbf{D}^{-1} \preceq \mathbf{M}_{n|n-1}^{-1} + \frac{1}{\sigma_v^2}\tilde{\mathbf{C}} \\ & \mathrm{tr}(\tilde{\mathbf{C}}) \leq P \\ & \tilde{\mathbf{C}} \succeq 0. \end{aligned} \tag{4.9}$$

The problem in (4.9) is equivalent to (4.8) in the sense that for the optimal solution of problem (4.9) the equality of the first constraint must hold. According to the Schur complement [35], the first constraint in (4.9) is equivalent to :

$$\begin{bmatrix} \mathbf{M}_{n|n-1}^{-1} + \frac{1}{\sigma_v^2}\tilde{\mathbf{C}} & \mathbf{I}_M \\ \mathbf{I}_M & \mathbf{D} \end{bmatrix} \succeq 0. \tag{4.10}$$

Plugging (4.10) into (4.9), we have

$$\begin{aligned} \min_{\tilde{\mathbf{C}},\mathbf{D}} \quad & \mathrm{tr}(\mathbf{D}) \\ s.t. \quad & \begin{bmatrix} \mathbf{M}_{n|n-1}^{-1} + \frac{1}{\sigma_v^2}\tilde{\mathbf{C}} & \mathbf{I}_M \\ \mathbf{I}_M & \mathbf{D} \end{bmatrix} \succeq 0 \\ & \mathrm{tr}(\tilde{\mathbf{C}}) \leq P \\ & \tilde{\mathbf{C}} \succeq 0. \end{aligned} \tag{4.11}$$

By converting the constraints of problem (4.11) into a large block diagonal linear matrix inequality, we can transform the problem into a standard SDP form, which can be efficiently solved using the interior point method.

Denote the optimal solution to problem (4.11) as $\tilde{\mathbf{C}}^*$, and define the singular value decomposition of $\tilde{\mathbf{C}}^*$ as $\tilde{\mathbf{C}}^* = \mathbf{U}\boldsymbol{\Sigma}\mathbf{U}^H$, so that $\mathbf{C}^* = \mathbf{U}\boldsymbol{\Sigma}^{\frac{1}{2}}$. The performance of $\mathbf{C}^*$ provides a lower bound for problem (4.6). To estimate $\mathbf{a}^*$, we solve[1]

$$\begin{aligned} \min_{\mathbf{a}} \quad & \|\mathrm{vec}(\mathbf{C}^*) - \mathbf{G}\mathbf{a}\|_2^2 \\ s.t. \quad & \mathbf{a}^H\mathbf{G}^H\mathbf{G}\mathbf{a} = P, \end{aligned} \tag{4.12}$$

which directly leads to

$$\mathbf{a}^* = \gamma(\mathbf{G}^H\mathbf{G})^{-1}\mathbf{G}^H\mathrm{vec}(\mathbf{C}^*), \tag{4.13}$$

where γ is defined as $\gamma = \sqrt{\frac{P}{\mathrm{vec}(\mathbf{C}^*)^H\mathbf{G}(\mathbf{G}^H\mathbf{G})^{-1}\mathbf{G}^H\mathrm{vec}(\mathbf{C}^*)}}$.

4.2.2 Minimize the Maximum MSE

When the maximum MSE is to be minimized, the parameter optimization problem can be stated as

$$\min_{\mathbf{a}} \max_{i} \quad [\mathbf{M}_{n|n}]_{i,i} \tag{4.14}$$

$$s.t. \quad \|\mathbf{C}(\mathbf{a})\|_F^2 \leq P\ . \tag{4.15}$$

Similar to (4.6), when treating $\mathbf{C}$ as the variable to be optimized, we can rewrite problem (4.14) as

$$\begin{aligned} \min_{\mathbf{C}} \max_{i} \quad & [\mathbf{M}_{n|n}]_{i,i} \\ s.t. \quad & \|\mathbf{C}\|_F^2 \leq P\ . \end{aligned} \tag{4.16}$$

[1] At the optimal solution of (4.6), the constraint should attain equality.

Introducing an auxiliary variable t, we can rewrite (4.16) as

$$\begin{aligned} \min_{\mathbf{C},t} \quad & t \\ s.t. \quad & t \geq [\mathbf{M}_{n|n}]_{i,i} \ , \\ & \|\mathbf{C}\|_F^2 \leq P \ . \end{aligned} \tag{4.17}$$

Define $\mathbf{e}_i$ as the vector with all zeros except for a 1 in the ith position, so that (4.17) is equivalent to

$$\begin{aligned} \min_{\mathbf{C},t} \quad & t \\ s.t. \quad & t \geq \mathbf{e}_i^T \mathbf{M}_{n|n} \mathbf{e}_i, \\ & \|\mathbf{C}\|_F^2 \leq P \ . \end{aligned} \tag{4.18}$$

Again, we utilize the Schur complement to rewrite the first constraint in (4.18) and we have

$$\begin{aligned} \min_{\tilde{\mathbf{C}},t} \quad & t \\ s.t. \quad & \begin{bmatrix} t & \mathbf{e}_i \\ \mathbf{e}_i^T & \mathbf{M}_{n|n-1}^{-1} + \frac{1}{\sigma_v^2}\tilde{\mathbf{C}} \end{bmatrix} \succeq 0, \ i = 1, \cdots, N \\ & \mathrm{tr}(\tilde{\mathbf{C}}) \leq P \\ & \tilde{\mathbf{C}} \succeq 0 \end{aligned} \tag{4.19}$$

Similar to (4.11), we can write the constraints of problem (4.19) in a large block diagonal linear inequality and convert the problem to a standard SDP form. After obtaining $\tilde{\mathbf{C}}^*$, we can use (15) to find the solution $\mathbf{a}^*$.

4.3 Scalar Observation Model

For a scalar observation, we have

$$y_n = \mathbf{c}^H \boldsymbol{\theta}_n + v_n, \tag{4.20}$$

where $\mathbf{c} = \mathbf{Ga}$ and as before $\mathbf{G} \in \mathbb{C}^{M \times N}$. The MSE of the estimated state is given by

$$\mathbf{M}_{n|n} = \mathbf{M}_{n|n-1} - \frac{\mathbf{M}_{n|n-1}\mathbf{c}\mathbf{c}^H\mathbf{M}_{n|n-1}}{\sigma_n^2 + \mathbf{c}^H\mathbf{M}_{n|n-1}\mathbf{c}}. \tag{4.21}$$

4.3.1 Minimize Sum MSE

For a scalar observation, the sum MSE is computed as

$$\mathrm{tr}(\mathbf{M}_{n|n}) = \mathrm{tr}(\mathbf{M}_{n|n-1}) - \frac{\mathbf{c}^H\mathbf{M}_{n|n-1}^2\mathbf{c}}{\sigma_n^2 + \mathbf{c}^H\mathbf{M}_{n|n-1}\mathbf{c}}, \tag{4.22}$$

and we formulate the following optimization problem:

$$\begin{aligned} \max_{\mathbf{a}} \quad & \frac{\mathbf{a}^H\mathbf{G}^H\mathbf{M}_{n|n-1}^2\mathbf{Ga}}{\sigma_n^2 + \mathbf{a}^H\mathbf{G}^H\mathbf{M}_{n|n-1}\mathbf{Ga}} \\ s.t. \quad & \mathbf{a}^H\mathbf{G}^H\mathbf{Ga} \leq P. \end{aligned} \tag{4.23}$$

Since the objective function in (4.23) is monotonically increasing with the norm of $\mathbf{c}$, the constraint must be active at the optimal solution and we can rewrite the problem as

$$\begin{aligned} \max_{\mathbf{a}} \quad & \frac{\mathbf{a}^H\mathbf{G}^H\mathbf{M}_{n|n-1}^2\mathbf{Ga}}{\mathbf{a}^H(\frac{\sigma_n^2}{P}\mathbf{G^HG} + \mathbf{G}^H\mathbf{M}_{n|n-1}\mathbf{G})\mathbf{a}} \\ s.t. \quad & \mathbf{a}^H\mathbf{G}^H\mathbf{Ga} = P. \end{aligned} \tag{4.24}$$

The solution to the above problem can be found directly as

$$\mathbf{c}^* = \sqrt{\frac{P}{\mathbf{u}^H \mathbf{B} \mathbf{u}}} \mathbf{B}^{-\frac{1}{2}} \mathbf{u}, \tag{4.25}$$

where $\mathbf{B} = \frac{\sigma_n^2}{P} \mathbf{G}^H \mathbf{G} + \mathbf{G}^H \mathbf{M}_{n|n-1} \mathbf{G}$, and $\mathbf{u}$ is the eigenvector corresponding to the largest eigenvalue of $\mathbf{B}^{-\frac{1}{2}} \mathbf{G}^H \mathbf{M}_{n|n}^2 \mathbf{G} \mathbf{B}^{-\frac{1}{2}}$.

4.3.2 Minimize Maximum MSE

When the maximum MSE is to be minimized, the optimization problem becomes

$$\begin{aligned} \min_{\mathbf{a}} \max_{i} \quad & [\mathbf{M}_{n|n}]_{i,i} \\ s.t. \quad & \mathbf{a}^H \mathbf{G}^H \mathbf{G} \mathbf{a} \leq P. \end{aligned} \tag{4.26}$$

Substituting (4.21) into (4.26), we have

$$\begin{aligned} \min_{\mathbf{a}} \max_{i} \quad & [\mathbf{M}_{n|n-1}]_{i,i} - \frac{\mathbf{e}_i^T \mathbf{M}_{n|n-1} \mathbf{G} \mathbf{a} \mathbf{a}^H \mathbf{G}^H \mathbf{M}_{n|n-1} \mathbf{e}_i}{\sigma_n^2 + \mathbf{a}^H \mathbf{G}^H \mathbf{M}_{n|n-1} \mathbf{G} \mathbf{a}} \\ s.t. \quad & \mathbf{a}^H \mathbf{G}^H \mathbf{G} \mathbf{a} \leq P. \end{aligned}$$

Introducing the auxiliary variable t, we have

$$\begin{aligned} \min_{\mathbf{a}} \quad & t \\ s.t. \quad & [\mathbf{M}_{n|n-1}]_{i,i} - \frac{\mathbf{e}_i^T \mathbf{M}_{n|n-1} \mathbf{G} \mathbf{a} \mathbf{a}^H \mathbf{G}^H \mathbf{M}_{n|n-1} \mathbf{e}_i}{\sigma_n^2 + \mathbf{c}^H \mathbf{M}_{n|n-1} \mathbf{c}} \leq t, \ i = 1, \cdots, N, \\ & \mathbf{a}^H \mathbf{G}^H \mathbf{G} \mathbf{a} \leq P \ . \end{aligned} \tag{4.27}$$

After some mathematical manipulation, we can rewrite problem (4.27) into the fol-

lowing form

$$\begin{aligned} \min_{\mathbf{a}} \quad & t \\ s.t. \quad & \big([\mathbf{M}_{n|n-1}]_{i,i} - t\big)\sigma_n^2 \leq \mathbf{a}^H \mathbf{E}_i \mathbf{a},\ i = 1, \cdots, N\ , \\ & \mathbf{a}^H \mathbf{G}^H \mathbf{G} \mathbf{a} \leq P, \end{aligned} \tag{4.28}$$

where $\mathbf{E}_i$ is defined as $\mathbf{E}_i = \mathbf{G}^H(\mathbf{M}_{n|n-1}\mathbf{e}_i\mathbf{e}_i^T\mathbf{M}_{n|n-1} - ([\mathbf{M}_{n|n-1}]_{i,i} - t)\mathbf{M}_{n|n-1})\mathbf{G}$. Problem (4.28) is equivalent to

$$\begin{aligned} \min_{\mathbf{a}} \quad & t \\ s.t. \quad & \big([\mathbf{M}_{n|n-1}]_{i,i} - t\big)\sigma_n^2 \leq \mathrm{tr}\big(\mathbf{A}\mathbf{E}_i\big),\ i = 1, \cdots, N\ , \\ & \mathrm{tr}(\mathbf{A}\mathbf{G}^H\mathbf{G}) \leq P\ , \\ & \mathrm{rank}(\mathbf{A}) = 1\ , \\ & \mathbf{A} \succeq 0\ . \end{aligned} \tag{4.29}$$

At this point, we see that the problem could be efficiently solved were it not for the rank constraint. So we relax this constraint to yield a quasi-convex problem:

$$\begin{aligned} \min_{\mathbf{a}} \quad & t \\ s.t. \quad & \big([\mathbf{M}_{n|n-1}]_{i,i} - t\big)\sigma_n^2 \leq \mathrm{tr}\big(\mathbf{A}\mathbf{E}_i\big),\ i = 1, \cdots, N\ , \\ & \mathrm{tr}(\mathbf{A}\mathbf{G}^H\mathbf{G}) \leq P, \\ & \mathbf{A} \succeq 0\ . \end{aligned} \tag{4.30}$$

In the above problem, given a deterministic t, all the constraints are convex. Denote the optimal value of problem (4.30) as t^*, so that for a $\tilde{t}$ that makes the problem (4.30) feasible, we have $t^* \leq \tilde{t}$, while if the problem is infeasible, we have $t^* > \bar{t}$. To find t^*, we search over t using the bisection method [36]. For a given t, we solve the

SDP feasibility problem:

$$\begin{aligned} \text{find} \quad & \mathbf{A} \\ s.t. \quad & \big([\mathbf{M}_{n|n-1}]_{i,i} - t\big)\sigma_n^2 \leq \text{tr}\big(\mathbf{A}\mathbf{E}_i\big),\ i = 1, \cdots, N\ , \\ & \text{tr}(\mathbf{A}\mathbf{G}^H\mathbf{G}) \leq P\ , \\ & \mathbf{A} \succeq 0 \end{aligned} \tag{4.31}$$

to obtain the optimal t^*, then we plug this t^* into problem (4.30) to find $\mathbf{A}^*$. If $\text{rank}(\mathbf{A}^*) = 1$, then $\mathbf{A}^*$ is the optimal solution to problem (4.26), otherwise, a rank-one solution $\mathbf{a}^*$ can be reconstructed [28]. The optimal value t^* of problem (4.30) can be used as a lower bound for the minimum MSE of problem (4.26). Our simulations indicate that in most cases, the rank of $\mathbf{A}^*$ is one, which indicates the performance of $\mathbf{a}^*$ is very close to the lower bound provided by $\mathbf{A}^*$.

4.4 Simulation Results

In the following simulations, the dimension of $\boldsymbol{\theta}_k$ and $\mathbf{a}$ are $M = 4$ and $N = 3$ respectively and L is set to 4. The observation noise variance is set to $\sigma_v^2 = 0.5$, and the covariance $\mathbf{Q}$ is assumed to be an identity matrix. The matrix $\mathbf{F}$ and the $\mathbf{G}$ are generated as complex Gaussian matrices with independent unit variance elements, and $\mathbf{F}$ is scaled to guarantee convergence of the Kalman filter. Once they are generated, $\mathbf{F}$ and $\mathbf{G}$ are kept constant in the simulation. The MSE performance for different constraints P are calculated after convergence of the Kalman filter.

In Figs. 4.1-4.2, for the vector-observation case the lower bound corresponds to the MSE calculated using $\mathbf{C}^*$. The performance gap between the lower bound and the proposed method represents the performance loss introduced by reducing the number

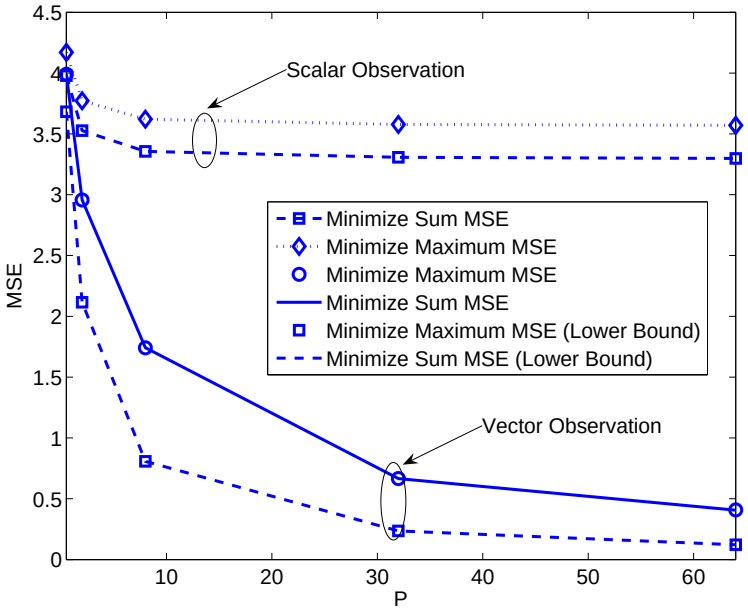

Figure 4.1: Sum MSE vs. the value of constraint P

of control parameters from 16 to 3. When the P is small, the proposed method can achieve a performance close to the lower bound, e.g., when $P = 0.5$, the performance degradation in the sum MSE is less than 10%. For the scalar observation case, we see as expected that the Min-Sum algorithm has the lowest sum-MSE and the Min-Max algorithm has the lowest maximum MSE. Surprisingly, however, in the vector case the performance of the Min-Sum and Min-Max algorithms is essentially identical. The performance of the Min-Max-MSE algorithm appears to be equal to the lower bound which indicates that the solution $\mathbf{A}^*$ to problem (4.30) is very likely rank-one even with the constraint relaxed. With increasing P, a performance floor exists for the scalar-observation case, while in the vector-observation case the MSE performance continues to improve.

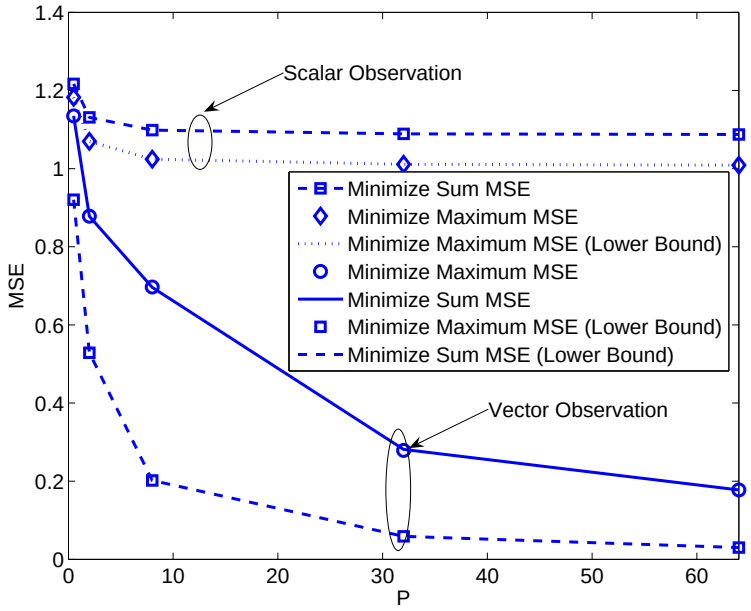

Figure 4.2: Maximum MSE vs. the value of constraint P

4.5 Summary

In this chapter, we investigated the problem of a Kalman filter with a linearly reconfigurable observation matrix. Two kinds of problems were formulated: Min-Sum-MSE or Min-Max-MSE. For the vector observation model, both of the optimization problems are difficult to solve directly, and we divided the problem into two simpler sub-problems that are easier to solve. Simulation results show that when the quadratic constraint is small, the proposed approach provides performance close to the MSE lower bound. For the scalar observation model, the Min-Sum-MSE problem is converted to a Rayleigh quotient maximization problem, for which an optimal closed-form solution is obtained, and for the Min-Max-MSE problem, we relax the rank-one constraint on the observation parameters and transform the optimization problem to an SDP feasibility problem. Based on the solution to the SDP feasibility problem, a rank-one solution can be reconstructed. Simulation results show that with a very high probability the solution of the relaxed problem is indeed rank-one.

Chapter 5

Sensor Phase Optimization for Multi-antenna Fusion Center

5.1 Introduction

The detection and estimation problems in analog WSNs have been widely investigated and a number of studies have focused on algorithm development and analysis for WSNs with a single-antenna FC. In [13], the sensors amplify and forward their observations of a scalar source to the FC via fading channels, and algorithms are developed to either minimize estimation error subject to transmit power constraints or minimize power subject to estimation error constraints. The scalar source model for this problem was generalized to correlated vector sources in [15]. An opportunistic power allocation approach was proposed in [16], and the scaling law with respect to the number of sensors was shown to be the same as the optimal power allocation proposed in [13]. In [22], the asymptotic variance of the best linear unbiased estimator of an analog WSN is derived, together with an analysis of the effect of different

assumptions regarding channel knowledge at the sensors. Scaling laws with respect to the number of sensors have been studied in [18] for a diversity-based method (where only the sensor with the best channel transmits), as well as for the coherent multiple access channel (MAC) and orthogonal channel cases, assuming a Gaussian source. In [20], a power optimization problem was formulated to minimize the outage probability of the MSE for the coherent MAC channel. More complicated settings involving analog WSNs with nonlinear measurement models [17] or relays [37, 38] have also been studied.

The results described above all assume that the FC is equipped with only one antenna. Just as multi-antenna receivers can provide significant capacity or diversity gains in communication systems, the estimation performance of a WSN should also benefit from the use of a multi-antenna FC, though prior work on this scenario is limited. A general scenario is investigated in [14], involving vector observations of a vector-valued random process at the sensors, and linearly precoded vector transmissions from the sensors to a multi-antenna FC. Optimal solutions for the precoders that minimize the mean-squared error (MSE) at the FC are derived for a coherent MAC under power and bandwidth constraints. In [39], single-antenna sensors amplify and forward their observations to a multi-antenna FC, but it is shown that for Rayleigh fading channels, the improvement in estimate variance is upper bounded by only a factor of two compared to the case of a single-antenna FC. The performance of two heuristic algorithms for choosing the gain and phase of the sensor transmissions is also studied. Subsequent results by the same authors in [40, 41], have demonstrated that when the channel undergoes (zero-mean) Rayleigh fading, there is a limit to the improvement in detection performance for a multi-antenna FC as well, but when the channel is Rician, performance improves monotonically with respect to number of antennas.

The term "amplify and forward" is often used to describe analog sensor networks like those discussed above, since each sensor applies a complex gain to the observation before sending it to the FC. For a coherent MAC, one can think of this as a type of distributed transmit beamforming, although it is distinguished from distributed beamforming applications such as those in communications since in a WSN the observed noise is transmitted together with the signal of interest. Some prior research in radar and communications has focused on scenarios where the beamformer weights implement only a phase shift rather than both a gain and a phase. The advantage of using phase shifting only is that it simplifies the implementation and is easily performed with analog hardware. Phase-shift-only beamformers have most often been applied to receivers that null spatial interference [42, 43], but it has also been considered on the transmit side for MISO wireless communications systems [44], which is similar to the problem considered here. For the distributed WSN estimation problem, phase-only sensor transmissions have been proposed in [45], where the phase is a scaled version of the observation itself. Phase-only transmissions were also considered in the context of distributed detection in [40], leading to a problem similar to one of those we consider here.

In addition to the work outlined above, other WSN research has focused on sensor selection problems, particularly in situations where the sensors have limited battery power. In these problems, only a subset of the sensors are chosen to transmit their observations, while the others remain idle to conserve power. The sensor selection problem has been tackled from various perspectives, with the goal of optimizing the estimation accuracy [37, 46, 47] or some heuristic system utility [48, 49]. In [46], the authors investigated maximum likelihood (ML) estimation of a vector parameter by selecting a fixed-size subset of the sensors. An approximate solution was found by relaxing the original Boolean optimization to a convex optimization problem. A dynamic model is used to describe the parameter of interest in [47], and sensors use

the Kalman filter to estimate the parameter. At each time step, a single sensor is selected and the measurement at the selected sensor is shared with all other sensors. A numerical sensor selection algorithm was proposed to minimize an upper bound on the expected estimation error covariance. Instead of the estimation accuracy, a utility function that takes into account the measurement quality or energy cost can also be used as the metric for sensor selection. In [49], each sensor independently optimizes its own operation status based on a utility function which depends on the sensor's own measurement and the predicted operation status of other sensors. A threshold is then found to enable the sensor to switch its status for either energy efficiency or energy consumption, and a power allocation algorithm was proposed to minimize the MSE at FC.

In this chapter we consider a distributed WSN with single-antenna sensors that observe an unknown deterministic parameter corrupted by noise. The low-complexity sensors apply a phase shift (rather than both a gain and phase) to their observation and then simultaneously transmit the result to a multi-antenna FC over a coherent MAC. One advantage of a phase-shift-only transmission is that it leads to a simpler analog implementation at the sensor. The FC determines the optimal value of the phase for each sensor in order to minimize the ML estimation error, and then feeds this information back to the sensors so that they can apply the appropriate phase shift. The estimation performance of the phase-optimized sensor network is shown to be considerably improved compared with the non-optimized case, and close to that achieved by sensors that can adjust both the transmit gain and phase. We analyze the asymptotic behavior of the algorithm for a large number of sensors and a large number of antennas at the FC. In addition, we analyze the impact of phase errors at the sensors due, for example, to errors in the feedback channel, a time-varying main channel or phase-shifter drift. We also consider a sensor selection problem similar to that in [46], and analyze its asymptotic behavior as well. Some additional details

regarding the contributions of this chapter are listed below.

1. We present two algorithms for determining the phase factors used at each sensor. In the first, we use the semidefinite relaxation presented in [40, 50] to convert the original problem to a SDP problem that can be efficiently solved by interior-point methods. For the second algorithm, we apply the analytic constant modulus algorithm (ACMA) [51], which provides a considerably simpler closed-form solution. Despite the reduction in complexity, the performance of ACMA is shown via simulation to be only slightly worse than the SDP solution, and close to the theoretical lower bound on the estimate variance. This is especially encouraging for networks with a large number of sensors N, since the SDP complexity is on the order of $N^{3.5}$, while that for ACMA is only on the order of N^2.

2. We separately derive performance scaling laws with respect to the number of antennas and the number of sensors assuming non-fading channels that take path loss into account. For both cases, we derive conditions that determine whether or not the presence of multiple antennas at the FC provides a significant benefit to the estimation performance. Prior work in [39–41] has focused on either AWGN channels with identical channel gains, or on fading channels where the channel gains are identically distributed, corresponding to the case where the distances from the sensors to the FC are roughly the same. References [39–41] also assume a special case where the noise at each of the sensors has the same variance, although [41] examines how certain upper bounds on performance change when the sensor noise is arbitrarily correlated.

3. Using our model for the non-fading case, we are able to elucidate detailed conditions under which the asymptotic estimation performance will improve with the addition of more antennas M at the FC. While [39, 40] showed that perfor-

mance always improves with increasing M for AWGN channels with identical gains and identically distributed sensor noise, we derive more detailed conditions that take into account the possibility of non-uniform distances between the sensors and FC and non-uniform noise at the sensors.

4. We conduct an analysis of the impact of phase errors at the sensors assuming relatively small phase errors with variance $\sigma_p^2 \ll 1$ (square-radians). In particular, we show that the degradation to the estimate variance is bounded above by a factor of $1+\sigma_p^2$. We note that the effect of errors in the transmit phase at the sensors has previously been considered for the case of $M = 1$ in [22], although using a different phase error model.

5. We consider the sensor selection problem separately for low and high sensor measurement noise. For the low measurement noise scenario, we relax the sensor selection problem to a standard linear programming (LP) problem, and we also propose a reduced complexity version of the algorithm. For the high measurement noise scenario, we show that the estimation error is lower bounded by the inverse of the measurement noise power, which motivates the use of a simple selection method based on choosing the sensors with the lowest measurement noise.

This chapter is organized as follows. Section 5.2 describes the assumed system model. Section 5.3 formulates the phase optimization problem and proposes a numerical solution based on SDP as well as a closed-form solution based on the algebraic constant modulus algorithm. In Section 5.4, the asymptotic performance of the algorithm is analyzed for a large number of sensors and antennas. The effect of phase errors is analyzed in Section 5.5 and the sensor selection problem is investigated in Section 5.6. Simulation results are then presented in Section 5.7 and our conclusions can be found in Section 5.8.

5.2 Signal Model

We assume that N single-antenna sensors in a distributed sensor network independently observe an unknown but deterministic complex-valued parameter θ. Each sensor phase shifts its observation and transmits the signal to the FC. According to the signal model in (2.2), the vector signal received at the FC can be expressed as

$$\mathbf{y} = \mathbf{H}\mathbf{a}\theta + \mathbf{H}\mathbf{D}\mathbf{v} + \mathbf{n} \, , \tag{5.1}$$

where $\mathbf{H} = [\mathbf{h}_1 \cdots \mathbf{h}_N]$ and $\mathbf{h}_i \in \mathbb{C}^{M\times 1}$ is the channel vector between the ith sensor and the FC, $\mathbf{a} = [a_1 \cdots a_N]^T$ contains the adjustable phase parameters and $|a_i| = 1$, $\mathbf{D} = \text{diag}\{a_1 \cdots a_N\}$, $\mathbf{v}$ is the sensor measurement noise vector with covariance $\mathbf{V} = \mathbb{E}\{\mathbf{v}\mathbf{v}^H\} = \text{diag}\left\{\sigma_{v,1}^2, \cdots, \sigma_{v,N}^2\right\}$, and $\mathbf{n}$ is complex Gaussian noise at the FC with covariance $\mathbb{E}\{\mathbf{n}\mathbf{n}^H\} = \sigma_n^2\mathbf{I}_M$. Note that since the sensors can only phase shift their observation prior to transmission, we ignore the issue of power control and assume that the sensors have sufficient power to forward their observation to the FC.

The combined noise term $\mathbf{H}\mathbf{D}\mathbf{v}+\mathbf{n}$ in (5.1) is Gaussian with covariance $\mathbf{H}\mathbf{V}\mathbf{H}^H+\sigma_n^2\mathbf{I}$, since $\mathbf{D}\mathbf{V}\mathbf{D}^H = \mathbf{V}$ due to the phase-only assumption. Assuming the FC is aware of the channel matrix $\mathbf{H}$, the noise covariance $\mathbf{V}$ and σ_n^2, it can calculate the ML estimate of θ using [25]

$$\hat{\theta}_{ML} = \frac{\mathbf{a}^H\mathbf{H}^H(\mathbf{H}\mathbf{V}\mathbf{H}^H + \sigma_n^2\mathbf{I}_M)^{-1}\mathbf{y}}{\mathbf{a}^H\mathbf{H}^H(\mathbf{H}\mathbf{V}\mathbf{H}^H + \sigma_n^2\mathbf{I}_M)^{-1}\mathbf{H}\mathbf{a}} \, .$$

The estimator $\hat{\theta}_{ML}$ is unbiased with variance

$$\text{Var}(\hat{\theta}_{ML}) = \left(\mathbf{a}^H\mathbf{H}^H(\mathbf{H}\mathbf{V}\mathbf{H}^H + \sigma_n^2\mathbf{I}_M)^{-1}\mathbf{H}\mathbf{a}\right)^{-1} \, . \tag{5.2}$$

Furthermore, since the Euclidean norm $\|\mathbf{a}\| = N$ when only phase shifts are used at

the sensors, it is easy to see that the variance is lower bounded by

$$\mathrm{Var}(\hat{\theta}_{ML}) \geq \frac{1}{N\lambda_{\max}\left(\mathbf{H}^H(\mathbf{H}\mathbf{V}\mathbf{H}^H + \sigma_n^2\mathbf{I}_M)^{-1}\mathbf{H}\right)}\ , \tag{5.3}$$

where $\lambda_{\max}(\cdot)$ denotes the largest eigenvalue of its matrix argument. Note that the bound in (5.3) is in general unachievable, since with probability one the given matrix will not have an eigenvector with unit modulus elements.

5.3 Optimizing the Sensor Phase

In this section we consider the problem of choosing $\mathbf{a}$ to minimize $\mathrm{Var}(\hat{\theta}_{ML})$ in (5.2). The unit modulus constraint prevents a trivial solution, but as we note below, a direct solution is not possible even without this constraint since the noise covariance would then depend on $\mathbf{a}$. The general optimization problem is formulated as

$$\begin{aligned} \min_{\mathbf{a}} \quad & \mathrm{Var}(\hat{\theta}_{ML}) \\ s.t. \quad & |a_i| = 1,\ i = 1, \ldots, N\ . \end{aligned} \tag{5.4}$$

Defining $\mathbf{B} = \mathbf{H}^H(\mathbf{H}\mathbf{V}\mathbf{H}^H + \sigma_n^2\mathbf{I}_M)^{-1}\mathbf{H}$, the problem can be rewritten as

$$\begin{aligned} \max_{\mathbf{a}} \quad & \mathbf{a}^H\mathbf{B}\mathbf{a} \\ s.t. \quad & |a_i| = 1,\ i = 1, \ldots, N\ . \end{aligned} \tag{5.5}$$

Note that this optimization can only determine $\mathbf{a}$ to within an arbitrary phase shift $e^{j\phi}$, but this scaling has no impact on the estimate of θ. In other words, the vector $\mathbf{a}$ and the vector $\mathbf{a}e^{j\phi}$ for arbitrary ϕ will both yield the same estimate $\hat{\theta}_{ML}$. Since the FC is aware of the vector $\mathbf{a}$ determined by the optimization in (5.5), any arbitrary

phase factor present in the $\mathbf{H}\mathbf{a}\theta$ term of the model in (5.1) will be canceled when the ML estimate of θ is computed. This is also clear from the variance expression in (5.2), which is insensitive to any phase shift to $\mathbf{a}$.

If there are only two sensors in the network, a simple closed-form solution to (5.5) can be obtained. Defining $\mathbf{B} = \begin{bmatrix} a & be^{j\beta} \\ be^{-j\beta} & c \end{bmatrix}$ with $a, b, c > 0$ and $\mathbf{a} = [e^{j\beta_1}, e^{j\beta_2}]$, then $\mathbf{a}^H\mathbf{B}\mathbf{a}$ is calculated as

$$\begin{aligned} \mathbf{a}^H\mathbf{B}\mathbf{a} &= a + c + 2b\cos(\beta_1 - \beta_2 - \beta) \\ &\leq a + c + 2b \,, \end{aligned} \tag{5.6}$$

and the equality in (5.6) can be achieved for any β_1, β_2 that satisfy $\beta_1 - \beta_2 = \beta$. For the general situation where $N > 2$, a solution to (5.5) appears to be intractable. Instead, in the discussion that follows we present two suboptimal approaches in order to obtain an approximate solution. The first approach is based on an SDP problem obtained by relaxing a rank constraint in a reformulated version of (5.5), similar to the approach proposed in [40, 50]. The second converts the problem to one that can be solved via the ACMA of [51]. It is worth emphasizing here that if the transmission gain of the sensors was also adjustable, then the corresponding problem would be

$$\begin{aligned} \max_{\mathbf{a}} \quad & \mathbf{a}^H\mathbf{H}^H(\mathbf{H}\mathbf{D}\mathbf{V}\mathbf{D}^H\mathbf{H}^H + \sigma_n^2\mathbf{I}_M)^{-1}\mathbf{H}\mathbf{a} \\ s.t. \quad & \mathbf{a}^H\mathbf{a} \leq N \,, \end{aligned} \tag{5.7}$$

which also has no closed-form solution due to the dependence on $\mathbf{a}$ (through the matrix $\mathbf{D}$) inside the matrix inverse. While in general both our SDP solution and (5.7) require numerical optimizations, we will see in Sections 5.4-5.6 that the theoretical analysis of performance and the solution to the sensor selection problem is consider-

ably simpler with the phase-only constraint. The simulations of Section 5.7 will also demonstrate that there is often little performance loss incurred by using phase-shift-only transmissions.

5.3.1 SDP Formulation

To begin, we rewrite (5.5) as follows:

$$\begin{aligned} \max_{\mathbf{a}} \quad & \mathrm{tr}\left(\mathbf{B}\mathbf{a}\mathbf{a}^H\right) \\ s.t. \quad & |a_i| = 1,\ i = 1, \ldots, N\ . \end{aligned} \tag{5.8}$$

Making the association $\mathbf{A} = \mathbf{a}\mathbf{a}^H$, problem (5.8) is equivalent to:

$$\begin{aligned} \max_{\mathbf{A}} \quad & \mathrm{tr}(\mathbf{B}\mathbf{A}) \\ s.t. \quad & \mathbf{A}_{i,i} = 1,\ i = 1, \ldots, N \\ & \mathrm{rank}(\mathbf{A}) = 1 \\ & \mathbf{A} \succeq 0\ , \end{aligned} \tag{5.9}$$

where $\mathbf{A}_{i,i}$ denotes the ith diagonal element of $\mathbf{A}$. Following the approach of [40, 50], we then relax the rank-one constraint, so that the problem becomes a standard SDP:

$$\begin{aligned} \max_{\mathbf{A}} \quad & \mathrm{tr}(\mathbf{B}\mathbf{A}) \\ s.t. \quad & \mathbf{A}_{i,i} = 1,\ i = 1, \ldots, N \\ & \mathbf{A} \succeq 0\ . \end{aligned} \tag{5.10}$$

Defining $\mathbf{B}_r = \text{real}\{\mathbf{B}\}$, $\mathbf{B}_i = \text{imag}\{\mathbf{B}\}$, and similarly for $\mathbf{A}_r$ and $\mathbf{A}_i$, we can convert (5.10) to the equivalent real form

$$\begin{aligned} \max_{\{\mathbf{A}_r, \mathbf{A}_i\}} \quad & \text{tr}(\mathbf{B}_r\mathbf{A}_r - \mathbf{B}_i\mathbf{A}_i) \\ s.t. \quad & \mathbf{A}_{r\ i,i} = 1,\ i = 1, \ldots, N \\ & \begin{bmatrix} \mathbf{A}_r & -\mathbf{A}_i \\ \mathbf{A}_i & \mathbf{A}_r \end{bmatrix} \succeq 0\ . \end{aligned} \tag{5.11}$$

Problem (5.11) can be efficiently solved by a standard interior-point method [52].

In general, the solution to (5.11) will not be rank one, so an additional step is necessary to estimate $\mathbf{a}$. Let $\mathbf{A}_r^*$, $\mathbf{A}_i^*$ denote the solution to problem (5.11), then the solution to problem (5.10) is given by $\mathbf{A}^* = \mathbf{A}_r^* + j\mathbf{A}_i^*$. If $\text{rank}(\mathbf{A}^*) > 1$, we can use a method similar to Algorithm 2 in [28] to extract a rank-one solution, as follows:

1. Decompose[1] $\mathbf{A}^* = \mathbf{C}^H\mathbf{C}$, define $\tilde{\mathbf{B}} = \mathbf{C}\mathbf{B}\mathbf{C}^H$, and find a unitary matrix $\mathbf{U}$ that can diagonalize $\tilde{\mathbf{B}}$.

2. Let $\mathbf{r} \in \mathbb{C}^{N\times 1}$ be a random vector whose ith element is set to $e^{j\omega_i}$, where ω_i is uniformly distributed over $[0, 2\pi)$.

3. Set $\tilde{\mathbf{a}} = \mathbf{C}^H\mathbf{U}\mathbf{r}$, and the solution is given by $\mathbf{a}^* = [a_1^* \ \cdots \ a_N^*]^T$, where $a_i^* = e^{j\angle\tilde{a}_i}$ and $\angle z$ represents the phase of a complex number z.

A detailed discussion of the reasoning behind the above rank-one modification can be found in [28].

[1]Since $\mathbf{A}^*$ is the solution to problem (5.10), $\mathbf{A}^*$ is positive semidefinite.

5.3.2 ACMA Formulation

For this discussion, we will assume that $N > M$, which represents the most common scenario. Thus, the $N \times N$ matrix $\mathbf{B}$ in the quadratic form $\mathbf{a}^H\mathbf{B}\mathbf{a}$ that we are trying to maximize is low rank; in particular, $\text{rank}(\mathbf{B}) \leq M < N$. Clearly, any component of $\mathbf{a}$ orthogonal to the columns or rows of $\mathbf{B}$ will not contribute to our goal of minimizing the estimate variance. In particular, if we define the singular value decomposition (SVD) $\mathbf{B} = \mathbf{U}\mathbf{\Sigma}\mathbf{U}^H$, we ideally seek a vector $\mathbf{a}$ such that

$$\begin{aligned} \mathbf{a} &= \sum_{k=1}^{m} w_k \mathbf{u}_k = \mathbf{U}_m \mathbf{w} \\ |a_i| &= 1 \,, \end{aligned} \tag{5.12}$$

where $\mathbf{U}_m = [\mathbf{u}_1 \ \cdots \ \mathbf{u}_m]$ contains the first $m \leq \text{rank}(\mathbf{B}) \leq M$ singular vectors of $\mathbf{B}$ and $\mathbf{w} = [w_1 \ \cdots \ w_m]^T$. The problem of finding the coefficient vector $\mathbf{w}$ of a linear combination of the columns of a given matrix $\mathbf{U}_m$ that yields a vector with unit modulus elements is precisely the problem solved by the ACMA [51].

Our problem is slightly different from the one considered in [51], since there will in general be no solution to (5.12) even in the absence of noise. However, in our simulation results we will see that the ACMA solution provides performance close to that obtained by the SDP formulation above. Note also that there is a trade-off in the choice of m, the number of vectors in $\text{span}(\mathbf{B})$ to be included in the linear combination of (5.12). A small value of m allows us to focus on forming $\mathbf{a}$ from vectors that will tend to increase the value of $\mathbf{a}^H\mathbf{B}\mathbf{a}$, while a larger value for m provides more degrees of freedom in finding a vector whose elements satisfy $|a_i| = 1$. Another drawback to choosing a larger value for m is that the ACMA solution can only be found if $N > m^2$. As long as M is not too large, one could in principle try all values of $m = 1, \cdots, M$ that satisfy $N > m^2$ and choose the one that yields the smallest estimate variance.

We will see later in the simulations that a small value for m already provides good performance, so the choice of m is not a significant issue.

The general ACMA approach can be formulated to find multiple solutions to (5.12), but in our case we only need a single solution, and thus a simplified version of ACMA can be used, as outlined here for a given m. The ACMA is obtained by defining the rows of $\mathbf{U}_m$ as $\mathbf{U}_m^H = [\tilde{\mathbf{u}}_1 \ \cdots \ \tilde{\mathbf{u}}_N]$, and then rewriting the constraint $|a_i| = |\tilde{\mathbf{u}}_i^H \mathbf{w}| = 1$ as

$$\left(\bar{\tilde{\mathbf{u}}}_i \otimes \tilde{\mathbf{u}}_i\right)^H (\bar{\mathbf{w}} \otimes \mathbf{w}) = 1 \ ,$$

where $\bar{(\cdot)}$ denotes the complex conjugate. Stacking all N such constraints into a single equation results in

$$\mathbf{P}[(\bar{\mathbf{w}} \otimes \mathbf{w})^T \ \ 1]^T = 0 \ , \tag{5.13}$$

where

$$\mathbf{P} = \begin{bmatrix} \left(\bar{\tilde{\mathbf{u}}}_1 \otimes \tilde{\mathbf{u}}_1\right)^H & -1 \\ \vdots & \vdots \\ \left(\bar{\tilde{\mathbf{u}}}_N \otimes \tilde{\mathbf{u}}_N\right)^H & -1 \end{bmatrix} . \tag{5.14}$$

If an exact solution to (5.13) existed, then a vector in the null space of $\mathbf{P}$ would have the form $\left[(\bar{\mathbf{w}} \otimes \mathbf{w})^T \ \ 1\right]^T$, and $\mathbf{w}$ could be found by stripping away the 1 and then unstacking the resulting vector into a rank-one matrix (see [51] for more details). In our problem, an exact solution to (5.13) does not exist, so we use the following approach to obtain an approximation:

1. Let $\mathbf{q}$ represent the right singular vector of $\mathbf{P}$ associated with the smallest singular value, and define the vector $\tilde{\mathbf{q}}$ to contain the first m^2 elements of $\mathbf{q}$.

2. Set $\mathbf{w}$ equal to the singular vector of $\tilde{\mathbf{Q}} + \tilde{\mathbf{Q}}^H$ with largest singular value, where the $m \times m$ matrix

$$\tilde{\mathbf{Q}} = \mathrm{vec}^{-1}(\tilde{\mathbf{q}}) \tag{5.15}$$

is formed by dividing $\tilde{\mathbf{q}}$ into sub-vectors of length m and stacking them together in a matrix.

3. Set $\hat{\mathbf{a}} = \mathbf{U}_m\mathbf{w}$. The vector $\mathbf{a}$ is then found by setting the magnitude of all the elements of $\hat{\mathbf{a}}$ equal to unity. In particular, the i-th element of $\mathbf{a}$ is given by

$$a_i^* = e^{j\angle\hat{a}_i} .$$

5.3.3 Comparison of Computational Complexity

As discussed in [50], the computational load of the SDP problem in (5.10) is of the order $O(N^{3.5})$. The additional steps required to take the SDP result and find a rank-one solution require an $O(N^3)$ eigenvalue decomposition, so the overall complexity is dominated by the SDP. For ACMA, the dominant computational step occurs in finding the m principal eigenvectors of the Hermitian matrix $\mathbf{B}$, which requires only an order $O(mN^2)$ computation [53]. Finding the least dominant singular vector of $\mathbf{P}$ is an $O(N^2)+O(m^4)$ operation, and the remaining steps have relatively trivial complexity. Since $m \ll N$ in typical scenarios, we see that ACMA enjoys a significantly lower computational load compared to the SDP approach. Despite this, we will see that ACMA has performance that is only slightly inferior to using the SDP solution.

5.4 Asymptotic Performance Analysis

In this section, we analyze the asymptotic performance achievable using only phase-shifts for the sensor transmissions. We will separately study cases where the number of sensors is large ($N \to \infty$) or the number of FC antennas is large ($M \to \infty$). Our analysis will be based on a non-fading channel model that takes path loss into

account, similar to models used in [54, 55]. In particular, for the channel between the FC and sensor i, we assume

$$\mathbf{h}_i = \frac{1}{d_i^{\alpha}} \tilde{\mathbf{h}}_i \ ,$$

where d_i denotes the distance between the ith sensor and the FC, α is the path loss exponent and $\tilde{\mathbf{h}}_i$ is given by

$$\tilde{\mathbf{h}}_i = [e^{j\gamma_{i,1}} \ \ e^{j\gamma_{i,2}} \ \cdots \ e^{j\gamma_{i,M}}]^T \ ,$$

where $\gamma_{i,k}$ is uniformly distributed over $[0, 2\pi)$.

5.4.1 Estimation Performance for Large N

From (5.3) we know that the lower bound on $\mathrm{Var}(\hat{\theta}_{ML})$ depends on the largest eigenvalue of $\mathbf{H}^H(\mathbf{H}\mathbf{V}\mathbf{H}^H + \sigma_n^2\mathbf{I}_M)^{-1}\mathbf{H}$. We begin by deriving a lower bound for this eigenvalue. The (m,n)th element of $\mathbf{H}\mathbf{V}\mathbf{H}^H$ can be expressed as

$$\left(\mathbf{H}\mathbf{V}\mathbf{H}^H\right)_{m,n} = \sum_{i=1}^{N} \frac{e^{j(\gamma_{i,m}-\gamma_{i,n})}\sigma_{v,i}^2}{d_i^{2\alpha}} \ .$$

According to the strong law of large numbers, as $N \to \infty$ we have

$$\begin{aligned} \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \frac{e^{j(\gamma_{i,m}-\gamma_{i,n})}\sigma_{v,i}^2}{d_i^{2\alpha}} &\overset{(a)}{=} \mathbb{E}\left\{\frac{\sigma_{v,i}^2}{d_i^{2\alpha}}\right\} \mathbb{E}\left\{e^{j(\gamma_{i,m}-\gamma_{i,n})}\right\} \\ &\overset{(b)}{=} \begin{cases} \mathbb{E}\left\{\frac{\sigma_{v,i}^2}{d_i^{2\alpha}}\right\} & m = n \\ 0 & m \neq n \ , \end{cases} \end{aligned} \tag{5.16}$$

where (a) follows from the assumption that $\gamma_{i,m}$, d_i and $\sigma_{v,i}^2$ are independent and (b) is due to the fact that $\gamma_{i,m}$ and $\gamma_{i,n}$ are independent and uniformly distributed over

$[0, 2\pi)$. Thus, for sufficiently large N we have

$$\lim_{N\to\infty} \mathbf{H}\mathbf{V}\mathbf{H}^H = N\mathbb{E}\left\{\frac{\sigma_{v,i}^2}{d_i^{2\alpha}}\right\}\mathbf{I}_M \ . \tag{5.17}$$

Based on (5.17), we have

$$\begin{aligned}\lim_{N\to\infty} \lambda_{\max}\left(\mathbf{H}^H(\mathbf{H}\mathbf{V}\mathbf{H}^H + \sigma_n^2\mathbf{I}_M)^{-1}\mathbf{H}\right) &= \frac{1}{N\mathbb{E}\left\{\frac{\sigma_{v,i}^2}{d_i^{2\alpha}}\right\} + \sigma_n^2}\left[\lim_{N\to\infty}\lambda_{\max}(\mathbf{H}^H\mathbf{H})\right] \\ &\overset{(c)}{=} \frac{N\mathbb{E}\left\{\frac{1}{d_i^{2\alpha}}\right\}}{N\mathbb{E}\left\{\frac{\sigma_{v,i}^2}{d_i^{2\alpha}}\right\} + \sigma_n^2}\ , \end{aligned} \tag{5.18}$$

where (c) is due to the fact that $\lambda_{\max}(\mathbf{H}^H\mathbf{H}) = \lambda_{\max}(\mathbf{H}\mathbf{H}^H)$. Substituting (5.18) into (5.3), we have the following asymptotic lower bound on the estimate variance:

$$\mathrm{Var}(\hat{\theta}_{ML}) \geq \frac{N\mathbb{E}\left\{\frac{\sigma_{v,i}^2}{d_i^{2\alpha}}\right\} + \sigma_n^2}{N^2\mathbb{E}\left\{\frac{1}{d_i^{2\alpha}}\right\}}\ . \tag{5.19}$$

For large enough N, the lower bound can be approximated using sample averages:

$$\mathrm{Var}(\hat{\theta}_{ML}) \geq \frac{\sum_{i=1}^{N}\frac{\sigma_{v,i}^2}{d_i^{2\alpha}} + \sigma_n^2}{N\sum_{i=1}^{N}\frac{1}{d_i^{2\alpha}}}\ . \tag{5.20}$$

Next, we derive an upper bound on the estimate variance and compare it with the lower bound obtained above. The upper bound is obtained by calculating the variance obtained when only a single antenna is present at the FC. For the given channel model, the optimal choice for the vector of sensor phases is just the conjugate of the channel

phases: $\mathbf{a} = [e^{-j\gamma_{1,1}} \ \cdots \ e^{-j\gamma_{N,1}}]^T$, which when applied to (5.2) leads to

$$\mathrm{Var}(\hat{\theta}_{ML}) \leq \frac{\sum_{i=1}^{N} \frac{\sigma_{v,i}^2}{d_i^{2\alpha}} + \sigma_n^2}{\left(\sum_{i=1}^{N} \frac{1}{d_i^{\alpha}}\right)^2} \ . \tag{5.21}$$

When $N \to \infty$, both the upper and lower bounds converge to 0, but the ratio of the lower bound in (5.20) to the upper bound in (5.21) converges to

$$\lim_{N\to\infty} \frac{\left(\sum_{i=1}^{N} \frac{1}{d_i^{\alpha}}\right)^2}{N \sum_{i=1}^{N} \frac{1}{d_i^{2\alpha}}} = \frac{\left(\mathbb{E}\left\{\frac{1}{d_i^{\alpha}}\right\}\right)^2}{\mathbb{E}\left\{\frac{1}{d_i^{2\alpha}}\right\}} = 1 - \frac{\mathrm{Var}\left\{\frac{1}{d_i^{\alpha}}\right\}}{\mathbb{E}\left\{\frac{1}{d_i^{2\alpha}}\right\}} \ . \tag{5.22}$$

Interestingly, we see that if $\mathrm{Var}\left\{\frac{1}{d_i^{\alpha}}\right\} \ll \mathbb{E}\left\{\frac{1}{d_i^{2\alpha}}\right\}$, the gap between the upper and lower bound is very small, and the availability of multiple antennas at the FC does not provide much benefit compared with the single antenna system when $N \to \infty$. On the other hand, if $\mathrm{Var}\left\{\frac{1}{d_i^{\alpha}}\right\} \to \mathbb{E}\left\{\frac{1}{d_i^{2\alpha}}\right\}$, the potential exists for multiple antennas to significantly lower the estimate variance.

5.4.2 Estimation Performance for Large M

Using the matrix inversion lemma, we have

$$\begin{aligned} \mathbf{H}^H(\mathbf{H}\mathbf{V}\mathbf{H}^H + \sigma_n^2\mathbf{I}_M)^{-1}\mathbf{H} &= \mathbf{H}^H\left(\frac{1}{\sigma_n^2}\mathbf{I}_M - \frac{1}{\sigma_n^4}\mathbf{H}\left(\mathbf{V}^{-1} + \frac{1}{\sigma_n^2}\mathbf{H}^H\mathbf{H}\right)^{-1}\mathbf{H}^H\right)\mathbf{H} \\ &= \frac{1}{\sigma_n^2}\mathbf{H}^H\mathbf{H} - \frac{1}{\sigma_n^4}\mathbf{H}^H\mathbf{H}\left(\mathbf{V}^{-1} + \frac{1}{\sigma_n^2}\mathbf{H}^H\mathbf{H}\right)^{-1}\mathbf{H}^H\mathbf{H} \ . \end{aligned} \tag{5.23}$$

Furthermore, the (m,n)th element of $\mathbf{H}^H\mathbf{H}$ is given by

$$\left(\mathbf{H}^H\mathbf{H}\right)_{m,n} = \frac{1}{d_m^{\alpha} d_n^{\alpha}} \sum_{i=1}^{M} e^{j(\gamma_{n,i} - \gamma_{m,i})} \ . \tag{5.24}$$

Similar to (5.16), as $M \to \infty$ we have

$$\lim_{M\to\infty} \frac{1}{M} \sum_{i=1}^{M} e^{j(\gamma_{n,i}-\gamma_{m,i})} = \begin{cases} 1 & m = n \\ 0 & m \neq n \ , \end{cases} \tag{5.25}$$

and thus

$$\lim_{M\to\infty} \mathbf{H}^H\mathbf{H} = M\text{diag}\left\{ \frac{1}{d_1^{2\alpha}} \ \cdots \ \frac{1}{d_N^{2\alpha}} \right\} \ . \tag{5.26}$$

Substituting (5.26) into (5.23), we have

$$\lim_{M\to\infty} \mathbf{H}^H(\mathbf{H}\mathbf{V}\mathbf{H}^H + \sigma_n^2\mathbf{I}_M)^{-1}\mathbf{H} = \text{diag}\left\{ \frac{M}{d_1^{2\alpha}\sigma_n^2 + M\sigma_{v,i}^2} \cdots \frac{M}{d_N^{2\alpha}\sigma_n^2 + M\sigma_{v,N}^2} \right\} \ ,$$

and thus

$$\lim_{M\to\infty} \text{Var}(\hat{\theta}_{ML}) = \frac{1}{M\sum_{i=1}^{N} \frac{1}{d_i^{2\alpha}\sigma_n^2 + M\sigma_{v,i}^2}} \ . \tag{5.27}$$

Note that this asymptotic expression is independent of the choice of $\mathbf{a}$. Here, for large M, the benefit of having multiple antennas at the FC hinges on the relative magnitude of $M\sigma_{v,i}^2$ versus $d_i^{2\alpha}\sigma_n^2$. If $M\sigma_{v,i}^2 \ll d_i^{2\alpha}\sigma_n^2$, a reduction in variance by a factor of M is possible. In this case, where the SNR at the FC is low but the signals sent from the sensors are high quality, the coherent gain from the combination of the relatively noise-free sensor signals helps increase the SNR at the FC. On the other hand, when $M\sigma_{v,i}^2 \gg d_i^{2\alpha}\sigma_n^2$, performance is asymptotically independent of M. Here, the coherent gain not only applies to θ but also to the sensor noise, which is stronger in this case.

5.5 Impact of Imperfect Phase

The previous sections have assumed that the FC can calculate the vector $\mathbf{a}$ and feed the phase information back to the sensors error free. Whether the feedback channel is digital or analog, there are about to be errors either in the received feedback at the sensors or in how the phase shift is actually implemented. Furthermore, the wireless channel may change during the time required for calculation and feedback of $\mathbf{a}$, so even if the phase shifts are implemented perfectly at the sensors, they may no longer be valid for the current channel. In this section, we evaluate the impact of errors in the sensor phase shifts on the estimation accuracy.

Define the phase shift for the ith sensor as $a_i = e^{j\alpha_i}$, and assume that

$$\alpha_i = \alpha_i^* + \Delta_i \ ,$$

where α_i^* is the optimized phase and Δ_i is a Gaussian perturbation (in radians) with zero mean and variance σ_p^2. Define $\mathbf{E} = \mathbf{H}^H(\mathbf{H}\mathbf{V}\mathbf{H}^H + \sigma_n^2\mathbf{I}_M)^{-\frac{1}{2}}$, so that $Var(\hat{\theta}_{ML})$ can be expressed as

$$Var(\hat{\theta}_{ML}) = \frac{1}{\|\mathbf{a}^H\mathbf{E}\|^2} = \frac{1}{\sum_{i=1}^{M}|\mathbf{a}^H\mathbf{e}_i|^2} \ , \tag{5.28}$$

where $\mathbf{e}_i$ is the ith column of $\mathbf{E}$. Let $e_{i,k}e^{j\beta_k}$ be a polar coordinate representation of the jth element of $\mathbf{e}_i$, so that

$$\begin{aligned} |\mathbf{a}^H\mathbf{e}_i|^2 &= \left|\sum_{k=1}^{N} e_{i,k}e^{j(\alpha_j^*+\Delta_j+\beta_k)}\right|^2 \\ &= \sum_{k=1}^{N} e_{i,k}^2 + \sum_{l=1}^{N}\sum_{\substack{m=1\\m\neq l}}^{N} e_{i,l}e_{i,m}\cos(\alpha_l^* + \Delta_l + \beta_l - \alpha_m^* - \Delta_m - \beta_m). \end{aligned} \tag{5.29}$$

Define $\delta^i_{l,m} = \Delta_l - \Delta_m$ and $\tau^i_{l,m} = \alpha^*_l + \beta_l - \alpha^*_m - \beta_m$. If we assume $\sigma_p^2 \ll 1$, (5.29) may be approximated via a 2nd order Taylor series as follows:

$$
\begin{aligned}
&|\mathbf{a}^H \mathbf{e}_i|^2 \\
&\approx \sum_{k=1}^{N} e_{i,k}^2 + \sum_{l=1}^{N} \sum_{\substack{m=1,\\ m\neq l}}^{N} e_{i,l} e_{i,m} \left(\cos(\tau^i_{l,m}) - \sin(\tau^i_{l,m})\delta^i_{l,m} - \frac{\cos(\tau^i_{l,m})}{2} \left(\delta^i_{l,m}\right)^2 \right) \\
&= \sum_{k=1}^{N} e_{i,k}^2 + \sum_{l=1}^{N} \sum_{\substack{m=1,\\ m\neq l}}^{N} e_{i,l} e_{i,m} \cos(\tau^i_{l,m}) - \sum_{l=1}^{N} \sum_{\substack{m=1,\\ m\neq l}}^{N} e_{i,l} e_{i,m} \left(\sin(\tau^i_{l,m})\delta^i_{l,m} + \frac{\cos(\tau^i_{l,m})}{2} \left(\delta^i_{l,m}\right)^2 \right). \quad (5.30)
\end{aligned}
$$

Substituting (5.30) into (5.28), we have

$$
Var(\hat{\theta}_{ML}) \approx \frac{1}{\sum_{i=1}^{M} (a_i - b_i)}, \quad (5.31)
$$

where

$$
\begin{aligned}
a_i &= \sum_{k=1}^{N} e_{i,k}^2 + \sum_{l=1}^{N} \sum_{\substack{m=1\\ m\neq l}}^{N} e_{i,l} e_{i,m} \cos(\tau^i_{l,m}) \\
b_i &= \sum_{l=1}^{N} \sum_{\substack{m=1\\ m\neq l}}^{N} e_{i,l} e_{i,m} \left(\sin(\tau^i_{l,m})\delta^i_{l,m} + \frac{\cos(\tau^i_{l,m})}{2} \left(\delta^i_{l,m}\right)^2 \right).
\end{aligned}
$$

In equation (5.31), the effect of the phase error is confined to the second double sum inside the outermost parentheses. If we define $\hat{\theta}^P_{ML}$ to be the estimate obtained with no phase errors, then

$$
Var(\hat{\theta}^P_{ML}) = \frac{1}{\sum_{i=1}^{M} \left(\sum_{k=1}^{N} e_{i,k}^2 + \sum_{l=1}^{N} \sum_{\substack{m=1\\ m\neq l}}^{N} e_{i,l} e_{i,m} \cos(\tau^i_{l,m}) \right)}, \quad (5.32)
$$

which is deterministic and does not depend on the random phase errors. We can then

obtain the following approximation

$$
\begin{aligned}
& Var(\hat{\theta}_{ML}) \\
& \overset{(f)}{\approx} Var(\hat{\theta}^P_{ML}) \left(1 + \frac{\sum_{i=1}^{M} \left(\sum_{l=1}^{N} \sum_{\substack{m=1 \\ m \neq l}}^{N} e_{i,l} e_{i,m} \left(\sin(\tau^i_{l,m}) \delta^i_{l,m} + \frac{\cos(\tau^i_{l,m})}{2} \left(\delta^i_{l,m} \right)^2 \right) \right)}{\sum_{i=1}^{M} \left(\sum_{k=1}^{N} e^2_{i,k} + \sum_{l=1}^{N} \sum_{\substack{m=1 \\ m \neq l}}^{N} e_{i,l} e_{i,m} \cos(\tau^i_{l,m}) \right)} \right),
\end{aligned}
$$

where (f) is due to the first order Taylor approximation $(1 - \frac{x}{y})^{-1} \approx 1 + \frac{x}{y}$ for $x \ll y$. We use the ratio of $Var(\hat{\theta}_{ML})$ to $Var(\hat{\theta}^P_{ML})$ to measure the effect of the phase error, which yields

$$
\frac{Var(\hat{\theta}_{ML})}{Var(\hat{\theta}^P_{ML})} \approx \left(1 + \frac{\sum_{i=1}^{M} \left(\sum_{l=1}^{N} \sum_{\substack{m=1 \\ m \neq l}}^{N} e_{i,l} e_{i,m} \left(\sin(\tau^i_{l,m}) \delta^i_{l,m} + \frac{\cos(\tau^i_{l,m})}{2} \left(\delta^i_{l,m} \right)^2 \right) \right)}{\sum_{i=1}^{M} \left(\sum_{k=1}^{N} e^2_{i,k} + \sum_{l=1}^{N} \sum_{\substack{m=1 \\ m \neq l}}^{N} e_{i,l} e_{i,m} \cos(\tau^i_{l,m}) \right)} \right) .
$$

Note that the only term in the above expression that is random is the numerator on the right-hand side. Taking the expectation of the ratio with respect to the phase perturbations Δ_i, we have

$$
\begin{aligned}
& \mathbb{E} \left\{ \frac{Var(\hat{\theta}_{ML})}{Var(\hat{\theta}^P_{ML})} \right\} \\
& = \left(1 + \frac{\sum_{i=1}^{M} \left(\sum_{l=1}^{N} \sum_{\substack{m=1 \\ m \neq l}}^{N} e_{i,l} e_{i,m} \left(\sin(\tau^i_{l,m}) \mathbb{E} \left\{ \delta^i_{l,m} \right\} + \frac{\cos(\tau^i_{l,m})}{2} \mathbb{E} \left\{ \left(\delta^i_{l,m} \right)^2 \right\} \right) \right)}{\sum_{i=1}^{M} \left(\sum_{k=1}^{N} e^2_{i,k} + \sum_{l=1}^{N} \sum_{\substack{m=1 \\ m \neq l}}^{N} e_{i,l} e_{i,m} \cos(\tau^i_{l,m}) \right)} \right) \\
& \overset{(h)}{=} \left(1 + \frac{\sum_{i=1}^{M} \sum_{l=1}^{N} \sum_{\substack{m=1 \\ m \neq l}}^{N} e_{i,l} e_{i,m} \cos(\tau^i_{l,m}) \sigma^2_p}{\sum_{i=1}^{M} \left(\sum_{k=1}^{N} e^2_{i,k} + \sum_{l=1}^{N} \sum_{\substack{m=1 \\ m \neq l}}^{N} e_{i,l} e_{i,m} \cos(\tau^i_{l,m}) \right)} \right) ,
\end{aligned} \tag{5.33}
$$

where in (h) we exploit the fact that $\mathbb{E} \left\{ \delta^i_{l,m} \right\} = 0$ and $\mathbb{E} \left\{ \left(\delta^i_{l,m} \right)^2 \right\} = 2\sigma^2_p$. Since

$$
\sum_{l=1}^{N} \sum_{\substack{m=1 \\ m \neq l}}^{N} e_{i,l} e_{i,m} \cos(\tau^i_{l,m}) \leq (N-1) \sum_{l=1}^{N} e^2_{i,l} ,
$$

the ratio in (5.33) is approximately upper bounded by

$$\mathbb{E}\left\{\frac{Var(\hat{\theta}_{ML})}{Var(\hat{\theta}^{P}_{ML})}\right\} \leq 1 + \left(1 - \frac{1}{N}\right)\sigma_p^2 \,. \tag{5.34}$$

We see from (5.34) that the impact of the phase errors increases with N, but in all cases the degradation in the estimate variance is approximately bounded above by a factor of $1 + \sigma_p^2$.

5.6 Sensor Selection

As mentioned earlier, in situations where it is desired to use only a subset of the sensors to estimate the parameter (e.g., in order to conserve power at the sensors), the FC needs a method to perform the sensor selection. Assuming only $K < N$ of the sensors are to be selected for transmission to the FC, an optimal solution to the problem would require solving the following maximization:

$$\begin{aligned} \max_{\mathbf{a},\mathbf{x}} \quad & \mathbf{x}^T\mathbf{D}^H\mathbf{H}^H\left(\mathbf{HVXH}^H + \sigma_n^2\mathbf{I}_M\right)^{-1}\mathbf{HDx} \\ s.t. \quad & \sum_{i=1}^{N} x_i = K \\ & x_i = \{0,1\} \\ & |a_i| = 1 \,, \end{aligned} \tag{5.35}$$

where $\mathbf{D} = \text{diag}\,\{a_1, \cdots, a_N\}$, $\mathbf{x} = [x_1, \cdots, x_N]^T$ is the selection vector and $\mathbf{X} = \text{diag}\{x_1, \cdots, x_N\}$. Even if one chooses one of the suboptimal approaches described in Section III for estimating $\mathbf{a}$, solving for $\mathbf{x}$ in (5.35) requires an exhaustive search over all possible K-sensor combinations and is in general NP-hard. Instead, in this section we derive conditions under which much simpler selection strategies can be applied.

We consider the following two cases: (1) low sensor noise relative to the noise at the FC, $\sigma_{v,i}^2 \ll \sigma_n^2$, and (2) relatively high sensor noise $\sigma_{v,i}^2 \gg \sigma_n^2$. For (1), we derive a LP solution as well as a simpler greedy algorithm, and for (2) we show that the problem reduces to choosing the sensors with the lowest measurement noise.

5.6.1 Algorithms for High FC Noise

Let $\mathbf{a}$ be the phase vector obtained using one of the algorithms in Section III assuming all N sensors are active. When $\sigma_{v,i}^2 \ll \sigma_n^2$, we ignore the term $\mathbf{HVXH}^H$ in (5.35), and the problem simplifies to

$$\begin{aligned} \max_{\mathbf{x}} \quad & \mathbf{x}^T \mathbf{D}^H \mathbf{H}^H \mathbf{H} \mathbf{D} \mathbf{x} \\ s.t. \quad & \sum_{i=1}^{N} x_i = K \\ & x_i = \{0, 1\} \ . \end{aligned} \tag{5.36}$$

Define $\mathbf{F} = \mathbf{D^H H^H H D}$, so that (5.36) can be rewritten as

$$\begin{aligned} \max_{\mathbf{x}} \quad & \mathbf{x}^T \mathrm{Re}\{\mathbf{F}\} \mathbf{x} \\ s.t. \quad & \sum_{i=1}^{N} x_i = K \\ & x_i = \{0, 1\} \ . \end{aligned} \tag{5.37}$$

Since $x_i^2 = x_i$, (5.37) is equivalent to

$$\begin{aligned} \max_{x_i} \quad & \sum_{i=1}^{N} \mathbf{F}_{i,i} x_i + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \mathrm{Re}\{\mathbf{F}_{i,j}\} x_i x_j \\ s.t. \quad & \sum_{i=1}^{N} x_i = K \\ & x_i = \{0, 1\} \ , \end{aligned} \tag{5.38}$$

where $\mathbf{F}_{i,j}$ denotes the (i,j)th element of matrix $\mathbf{F}$. By linearizing the term $x_i x_j$ [56], (5.38) is equivalent to

$$\max_{x_i, y_{ik}} \sum_{i=1}^{N} \mathbf{F}_{i,i} x_i + 2 \sum_{i=1}^{N-1} \sum_{k=i+1}^{N} \mathrm{Re}\{\mathbf{F}_{i,k}\} y_{ik} \tag{5.39a}$$

$$s.t. \sum_{i=1}^{N} x_i = K \tag{5.39b}$$

$$1 - x_i - x_j + y_{ik} \geq 0 \tag{5.39c}$$

$$x_i - y_{ik} \geq 0 \tag{5.39d}$$

$$x_j - y_{ik} \geq 0 \tag{5.39e}$$

$$y_{ik} \geq 0 \tag{5.39f}$$

$$x_i = \{0, 1\} \ , \tag{5.39g}$$

where the constraints (5.39c)-(5.39g) lead to $y_{ik} = x_i x_k$.

Note that all of the constraints in (5.39) are linear, except for (5.39g). If we relax the constraint in (5.39g), the condition $0 \leq x_i \leq 1$ is implicitly included in (5.40c)-(5.40f), and we are left with a LP problem in standard form [56]:

$$\max_{x_i, y_{ik}} \sum_{i=1}^{N} \mathbf{F}_{i,i} x_i + 2 \sum_{i=1}^{N-1} \sum_{k=i+1}^{N} \mathrm{Re}\{\mathbf{F}_{i,k}\} y_{ik} \tag{5.40a}$$

$$s.t. \sum_{i=1}^{N} x_i = K \tag{5.40b}$$

$$1 - x_i - x_j + y_{ik} \geq 0 \tag{5.40c}$$

$$x_i - y_{ik} \geq 0 \tag{5.40d}$$

$$x_j - y_{ik} \geq 0 \tag{5.40e}$$

$$y_{ik} \geq 0 \ . \tag{5.40f}$$

To find the $x_i = \{0, 1\}$ solution needed for sensor selection, one can take the result

of (5.40) and simply set the K largest elements to one and the rest to zero. If desired, once the K sensors have been selected, the phase vector $\mathbf{a}$ for these K sensors can be recomputed based on a reduced dimension version of the algorithms in Section III.

The above LP problem has $\frac{N(N-1)}{2}+N$ variables and $2N(N-1)+1$ constraints, and thus will require on the order of $\left(\frac{N(N-1)}{2}+N\right)^2(2N(N-1)+1)$ arithmetic operations [52]. A simpler greedy algorithm is presented below that only requires $O(KN)$ operations, and that achieves performance close to the LP approach. The greedy algorithm is based on the following observation [2]:

$$\begin{aligned}\mathbf{x}^T\mathbf{D}^H\mathbf{H}^H\mathbf{H}\mathbf{D}\mathbf{x} &= \sum_{i=1}^{K}\sum_{l=1}^{K}\bar{a}_i a_l \mathbf{h}_i^H\mathbf{h}_l \\ &= \sum_{i=1}^{K-1}\sum_{l=1}^{K-1}\bar{a}_i a_l \mathbf{h}_i^H\mathbf{h}_l + \|\mathbf{h}_K\|^2 + 2\mathrm{Re}\left\{\sum_{l=1}^{K-1}\bar{a}_K a_l \mathbf{h}_K^H\mathbf{h}_l\right\} .\end{aligned}$$

The idea behind the greedy algorithm is to add sensors one at a time based on those for which the last two terms in the above sum are the largest. The steps of the algorithm are detailed below.

Greedy Sensor Selection Algorithm

1. Select the first sensor as the one with the strongest channel: $i = \arg\max_k \|\mathbf{h}_k\|^2$, and initialize the active sensor set as $\mathcal{S} = \{i\}$.

2. While $|\mathcal{S}| \leq K$, perform the following:

   (a) Solve
   $$i = \arg\max_{k\notin\mathcal{S}} \|\mathbf{h}_k\|^2 + 2\mathrm{Re}\left\{\sum_{l\in\mathcal{S}}\bar{a}_k a_l \mathbf{h}_k^H\mathbf{h}_l\right\} .$$

   (b) Update $\mathcal{S} = \mathcal{S}\bigcup i$.

[2]For the simplicity of expression, we assume that the first K sensors are selected out.

As with the LP algorithm, once the K sensors are selected, an updated solution for the associated K elements of $\mathbf{a}$ can be obtained.

5.6.2 Algorithm for High Sensor Noise

When $\sigma_{v,i}^2 \gg \sigma_n^2$ and assuming that $N > M$ (the case of interest when sensor selection is necessary), the original criterion can be simplified to

$$\begin{aligned} \mathbf{a}^H\mathbf{H}^H\left(\mathbf{H}\mathbf{V}\mathbf{H}^H\right)^{-1}\mathbf{H}\mathbf{a} &= \mathbf{a}^H\mathbf{V}^{-\frac{1}{2}}\mathbf{V}^{\frac{1}{2}}\mathbf{H}^H\left(\mathbf{H}\mathbf{V}\mathbf{H}^H\right)^{-1}\mathbf{H}\mathbf{V}^{\frac{1}{2}}\mathbf{V}^{-\frac{1}{2}}\mathbf{a} \\ &= \mathbf{a}^H\mathbf{V}^{-\frac{1}{2}}\mathbf{P}_{VH}\mathbf{V}^{-\frac{1}{2}}\mathbf{a}\,, \end{aligned}$$

where $\mathbf{P}_{VH} = \mathbf{V}^{\frac{1}{2}}\mathbf{H}^H\left(\mathbf{H}\mathbf{V}\mathbf{H}^H\right)^{-1}\mathbf{H}\mathbf{V}^{\frac{1}{2}}$ is a rank M projection matrix. Ideally, to maximize the criterion function, one should attempt to find a vector of the form $\mathbf{V}^{-\frac{1}{2}}\mathbf{a}$ that lies in the subspace defined by $\mathbf{P}_{VH}$. Assuming the vector $\mathbf{a}$ can approximately achieve this goal, the lower bound on variance is approximately achieved and we have

$$\frac{1}{\mathbf{a}^H\mathbf{V}^{-\frac{1}{2}}\mathbf{P}_{VH}\mathbf{V}^{-\frac{1}{2}}\mathbf{a}} \approx \frac{1}{\mathbf{a}^H\mathbf{V}^{-1}\mathbf{a}} = \frac{1}{\sum_{i=1}^{N}\frac{1}{\sigma_{v,i}^2}}\,. \tag{5.41}$$

With respect to the sensor selection problem, this suggests that when $\sigma_{v,i}^2 \gg \sigma_n^2$, the K sensors with the smallest values of $\sigma_{v,i}^2$ should be chosen.

5.7 Simulation Results

Here we present the results of several simulation examples to illustrate the performance of the proposed algorithms. In all cases, the path loss exponent α is set to 1, and each result is obtained by averaging over 300 channel realizations. The sensors are assumed to lie in a plane at random angles with respect to the FC, uniformly

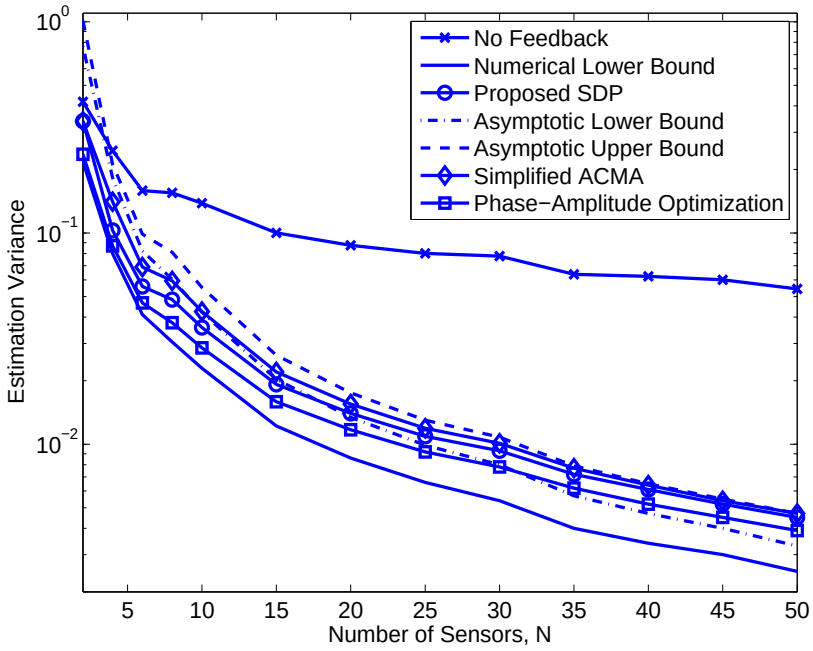

Figure 5.1: Performance of the proposed algorithms with an increasing number of sensors for a low measurement noise scenario ($\sigma_n^2 = 0.1$, $\sigma_{v,i}^2$ uniformly distributed over $[0.01, 0.1]$, d_i uniformly distributed over $[3, 20]$ and $M = 4$). In the plot, the Numerical Lower Bound, Asymptotic Lower Bound and Asymptotic Upper Bound are defined in equations (5.3), (5.20) and (5.21) respectively.

distributed over $[0, 2\pi)$. The distances of the sensors to the FC will be specified separately below. To evaluate the performance without feedback, $\mathbf{a}$ is set to a vector of all ones. In some of the simulations, we will compare the performance of the proposed algorithms with that obtained by (5.7), where both the sensor gain and phase can be adjusted. In these simulations, we use the active-set method to optimize (5.7), and we use several different initializations in order to have a better chance of obtaining the global optimum. When the ACMA algorithm is implemented, the subspace dimension was set at $m = 2$.

In the first two examples, we study the estimation performance for $M = 4$ FC antennas with increasing N for a case where the sensor measurement noise $\sigma_{v,i}^2$ is uniformly distributed over $[0.01, 0.1]$ and the FC noise σ_n^2 is set to 0.1. Fig. 5.1 shows the re-

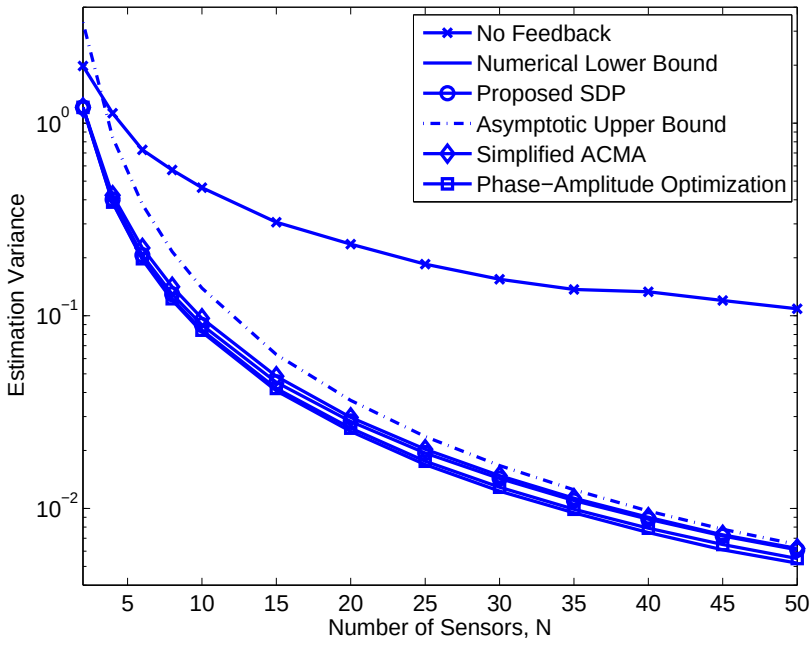

Figure 5.2: Performance of the proposed algorithms with an increasing number of sensors for a low measurement noise scenario ($\sigma_n^2 = 0.1$, $\sigma_{v,i}^2$ uniformly distributed over $[0.01, 0.1]$, $d_i = 11.5$ and $M = 4$).

sults assuming that the sensor distances d_i are uniformly distributed in the interval $[3, 20]$, while in Fig. 5.2 $d_i = 11.5$ for all sensors. In both cases, even though the lower bound of (5.3) is not achievable, we see that the performance of the proposed SDP and ACMA methods is nonetheless reasonably close to the bound, and not significantly worse than the performance obtained by optimizing both the phase and gain. As N gets larger in Fig. 5.1, the estimation error for all of the methods (except the no-feedback case) falls within the asymptotic lower and upper bounds of (5.20) and (5.21). When $N = 50$, the ratio $\mathrm{Var}\left\{\frac{1}{d_i^\alpha}\right\} / \mathbb{E}\left\{\frac{1}{d_i^{2\alpha}}\right\}$ is 0.304 for Fig. 5.1, and the ratio between the lower and upper bound is 0.702, which is in excellent agreement with the value of $1 - 0.304$ predicted by Eq. (5.22). Since the upper bound in (5.21) corresponds to the case of $M = 1$, one may suppose that the gap in Fig. 5.1 between the bounds of (5.20) and (5.21) indicates that the presence of multiple antennas at the FC could provide a benefit for large N. However, the performance of SDP and

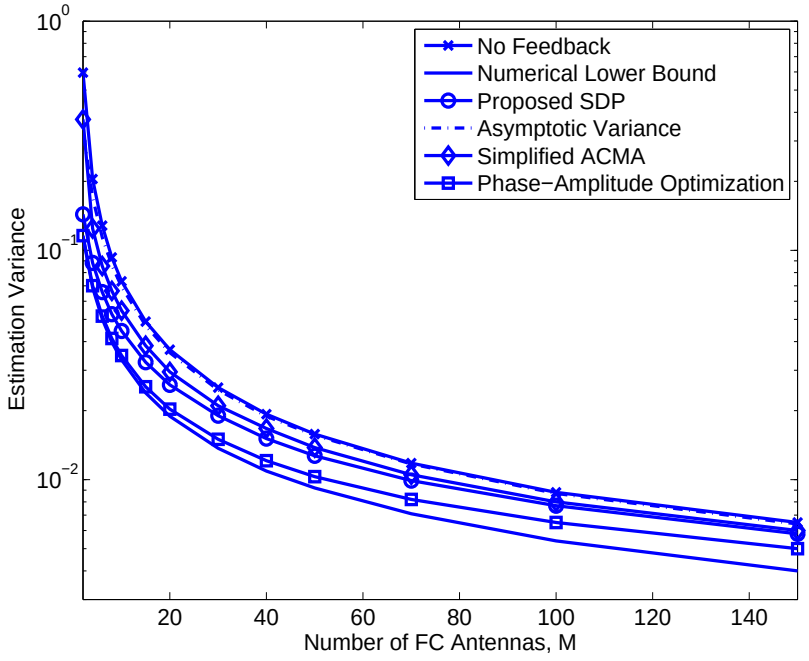

Figure 5.3: Performance of the proposed algorithms with an increasing number of antennas ($\sigma_n^2 = 0.1$, $\sigma_{v,i}^2$ uniformly distributed over $[0.001, 0.01]$, d_i uniformly distributed over $[3, 20]$ and $N = 4$). In the simulation results, the Asymptotic Variance is calculated using equation (5.27).

ACMA is approaching the upper bound more tightly, indicating that there is no benefit from having multiple antennas in this case. In Fig. 5.2 where the d_i are all equal, the asymptotic bounds in (5.20) and (5.21) are identical, and asymptotically we expect no benefit from multiple antennas at the FC. We see again that for large N the performance of the SDP and ACMA methods is essentially at the predicted bound. When the d_i are equal and $\frac{\sigma_{v,i}^2}{d_i^\alpha} \ll \sigma_n^2$, the matrix $\mathbf{H}^H(\mathbf{H}\mathbf{V}\mathbf{H}^H + \sigma_n^2\mathbf{I}_M)^{-1}\mathbf{H}$ asymptotically approaches a scaled identity matrix, so in this case the performance of the proposed phase-shift only algorithms even approaches the lower bound of Eq. (5.3).

Fig. 5.3 illustrates the performance for $N = 4$ with an increasing number of FC antennas M when $\sigma_{v,i}^2$ is uniformly distributed over $[0.001, 0.01]$ and $\sigma_n^2 = 0.1$. In this example, for most of the sensors we have $M\sigma_{v,i}^2 \ll d_i^{2\alpha}\sigma_n^2$, so in this case we see an improvement as the number of FC antennas increases. However, the benefit of

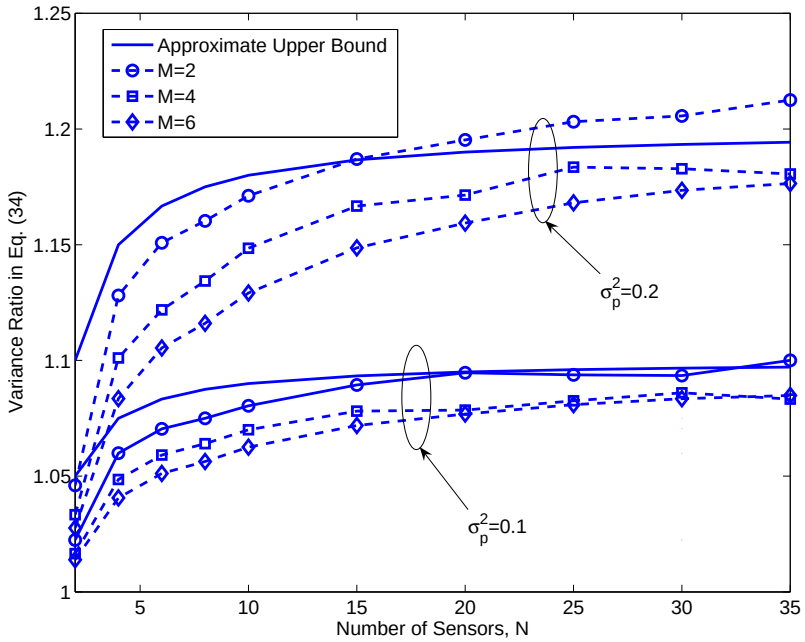

Figure 5.4: Effect of phase errors on algorithm performance ($\sigma_n^2 = 0.1$, $\sigma_{v,i}^2$ uniformly distributed over $[0.01, 0.1]$ and d_i uniformly distributed over $[3, 20]$).

optimizing the transmit phase (and gain for that matter) is reduced as M increases.

In Fig. 5.4, we investigate the effect of phase errors for two cases, $\sigma_p^2 = 0.1$ and $\sigma_p^2 = 0.2$ assuming the same noise parameter settings as in the first two examples. For each channel realization, results for 3000 different phase error realizations were obtained and averaged to obtain the given plot. The ratio of the variance obtained by the SDP algorithm with and without phase errors is plotted for $M = 2, 4, 6$ for both values of σ_p^2, and the approximate bound of (5.34) is also shown. The results show that the performance degradation increases with N, and that (5.34) provides a reasonable indication of performance for large N. Fig. 5.4 also shows that increasing the number of FC antennas improves the robustness of the algorithm to imprecise sensor phase.

In Fig. 5.5, we compare the performance of the three different sensor selection al-

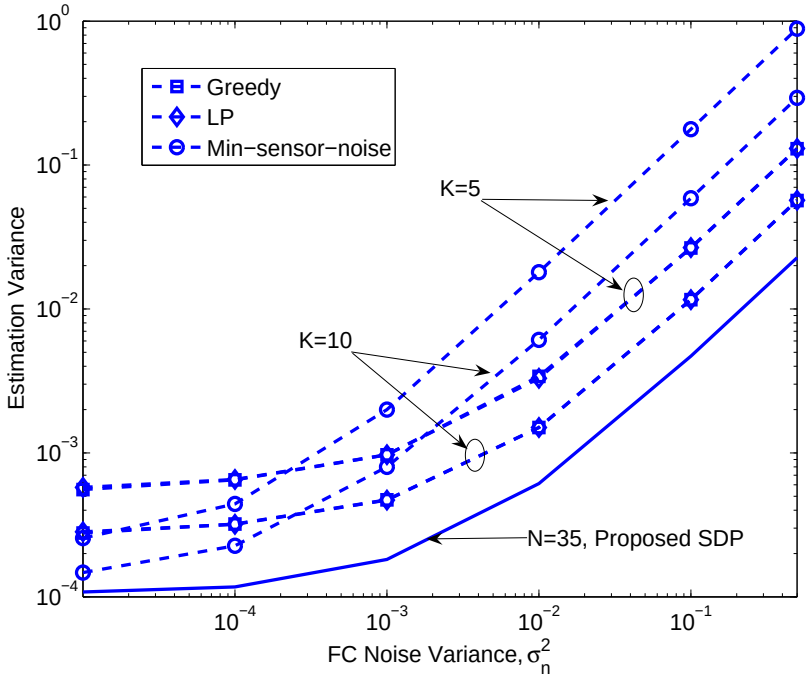

Figure 5.5: Performance comparison between different sensor selection algorithms ($N = 35$, $M = 4$, $\sigma_{v,i}^2$ uniformly distributed over $[0.001, 0.01]$ and d_i uniformly distributed over $[3, 20]$).

gorithms discussed in this chapter (LP, greedy and min-sensor-noise) as a function of σ_n^2 assuming $M = 4$ antennas, $N = 35$ sensors and the sensor noise is uniformly distributed over $[0.001, 0.01]$. The sensor distances d_i are uniformly distributed in the interval $[3, 20]$. Three sets of curves are plotted, one for $K = 5$ selected sensors, one for $K = 10$, and one corresponding to when all the sensor nodes are used (the solid curve, obtained using the SDP algorithm). After the sensor selection, the proposed SDP is used to re-optimize the selected sensor nodes' phase parameters. For small σ_n^2 such that $\sigma_{v,i}^2 \gg \sigma_n^2$, we see as predicted that the best performance is obtained by simply selecting the K sensors with the smaller measurement noise. On the other hand, again in agreement with our analysis, the LP and greedy algorithms achieve the lowest estimation error for larger values of σ_n^2. Interestingly, the greedy algorithm provides performance essentially identical to the LP approach at a significantly reduced computational cost.

5.8 Summary

In this chapter, we investigated a distributed network of single antenna sensors employing a phase-shift and forward strategy for sending their noisy parameter observations to a multi-antenna FC. We presented two algorithms for finding the sensor phase shifts that minimize the variance of the estimated parameter, one based on a relaxed SDP and a closed-form heuristic algorithm based on the ACMA approach. We analyzed the asymptotic performance of the phase-shift and forward scheme for both large numbers of sensors and FC antennas, and we derived conditions under which increasing the number of FC antennas will significantly benefit the estimation performance. We also analyzed the performance degradation that results when sensor phase errors of variance σ_p^2 are present, and we showed that for large N the variance will approximately increase by a factor of $1 + \sigma_p^2$ provided that $\sigma_p^2 \ll 1$ square radian. The sensor selection problem was studied assuming either low or high sensor noise with respect to the noise at the FC. For low sensor noise, two algorithms were proposed, one based on linear programming with a relaxed integer constraint, and a computationally simpler greedy approach. For high sensor noise, we showed that choosing the sensors with the smallest noise variances was approximately optimal. Simulation studies of the proposed algorithms illustrate their advantages and the validity of the asymptotic analyses.

Chapter 6

Optimal Power Allocation for Multi-antenna Fusion Center

6.1 Introduction

The use of WSNs for detection and parameter estimation has been widely studied (e.g., [1, 8, 13, 23, 57–62]). When a coherent MAC is employed between the sensor nodes and fusion center [1, 8, 13, 23, 59–62], each sensor takes a noisy measurement of the signal of interest, amplifies and forwards the measurement to a FC through a wireless fading channel, and the FC makes a decision about the presence of the signal and estimates its parameters based on the coherent sum of the signals from all the sensor nodes. To minimize the detection or estimation errors, the transmit power at the sensors is optimized under either sum or individual power constraints.

The aforementioned works all assume that the FC is configured with a single antenna. It is well-known that multiple antennas can effectively increase the throughput of a wireless link, and recently researchers have investigated the use of arrays with a

massive number of antennas in wireless communication systems in order to improve spectral and energy efficiency [4–7]. When the BS has an array with many antennas and meanwhile the individual mobile stations have a single antenna, the energy efficiency at the mobile stations can be significantly increased. It has been shown that when perfect CSI is available at the BS, the transmit power of the mobile stations can be scaled by inverse of the number of antennas to asymptotic achieve constant user rate in the system [4]; when BS uses an imperfect CSI through LMMSE estimation, a constant user rate is asymptotically achievable, if the mobile users' transmit power is inversely proportional to the square root of the number of antennas[5].

For parameter detection or estimation problems in WSNs, an important question is how to exploit a multi-antenna FC to improve the probability of detection or estimation error. Several recent papers have studied the benefit provided by multiple antennas in the WSN context [10, 39, 40, 63–65]. In [64], the sensors use a fixed transmission gain to forward the measured signal to the multi-antenna FC, and the probabilities of detection and false alarm are derived under different assumptions for the CSI. Power allocation problems for signal detection and estimation are formulated in [39, 40] for a multi-antenna FC under a Rayleigh fading channel, but the performance benefit of a multiple- versus single-antenna FC is shown to be bounded by a constant that is unrelated to the number of antennas. For signal estimation using a phase-shift and forward WSN with a multi-antenna FC, it is shown in [10] that as the number of antennas M grows large, in certain cases the estimation error will decrease by a factor of M. Antenna arrays at the FC are also considered in [63, 65], where each sensor node first makes a local binary decision about the measured signal, and then forwards the decisions to the multi-antenna FC using uniform transmit power. In [63], a number of sub-optimal but low complexity fusion rules at the FC are derived and analyzed, and the results indicate the benefit of using multiple antennas in terms of detection performance. The results in [65] show that when the number of

antennas is very large, the FC can asymptotically achieve the detection performance upper bound even using a linear receiver with imperfect CSI.

While the benefits of massive numbers of antennas have been carefully studied for communication systems, we see above that relatively little work has analyzed their impact for WSNs. In this chapter, we investigate the gains in energy efficiency that can be obtained in a coherent multiple-access WSN when the FC has a large number of antennas, and we show how to determine optimal values for the sensor gains when the CSI is either perfectly known or unknown at the FC. In particular, our motivation is to demonstrate that FC antennas can be traded for sensor power; this is an important observation for WSNs where the sensors must conserve energy (e.g., due to the use of batteries or energy harvesting). The specific contributions of this chapter are detailed in the next section.

In this chapter, we study the detection and estimation performance of a coherent amplify-and-forward WSN with single antenna sensors and a massive number M of antennas at the FC. We assume the parameter of interest is a zero-mean circular complex Gaussian variable and that the wireless channels between the sensor nodes and FC undergo Rayleigh fading. Under these assumptions, we investigate the performance of the Neyman-Pearson (NP) and energy detectors and the linear minimum mean-squared error estimator (LMMSE). Our contributions are summarized below.

(1) For the case where CSI for the sensor nodes is available at the FC and the NP detector can be implemented, we derive the dependence of both probability of detection (PD) and probability of false alarm (PFA) on the sensor transmit power and show that as $M \to \infty$, the sensor power can be reduced by $1/M$ to achieve a constant PD for the same fixed PFA. This is similar in spirit to the results for massive MIMO in wireless cellular communications with perfect CSI [5]. However, unlike [5] which assumes each user transmits with equal power, we derive the optimal

transmission gains for the sensors that maximize PD for a fixed PFA under a sum power constraint. We show that this problem is independent of the sensor phase and convex with respect to the magnitude squared of the sensor gain as $M \to \infty$, and we formulate a simple closed-form "water-filling" solution to calculate the optimal gains. In our simulations, we demonstrate that compared with a uniform power allocation, the optimal gains result in significantly improved PD performance when the sensors transmit with low power, which is the case of interest for energy efficiency.

(2) For the NP detector, we also derive asymptotic performance bounds for cases where the available sum transmit power P satisfies either $P \to \infty$ or $P \to 0$. When $P \to 0$, we show that PD approaches PFA in the single antenna case, but PD is strictly greater than PFA (and potentially significantly greater than PFA) as long as P decreases at a rate of $O(1/M)$ or slower as $M \to \infty$. However, when $P \to \infty$, we show that both the single- and multiple-antenna FC asymptotically achieve the same detection performance, and hence the use of multiple antennas asymptotically provides no benefit for the NP detector at very high signal-to-noise ratios.

(3) For the case where the CSI is unknown or a computationally simpler solution is desired, we study the performance of the energy detector. The *deflection* of the energy detector is used as the performance metric, which generally serves as an accurate indicator of a detector's performance. Our results show that if the sensor transmit power decreases as $1/\sqrt{M}$ when $M \to \infty$, a constant deflection can be achieved. Based on this, we show how to choose the sensor transmission gains to maximize the deflection under a sum power constraint. In particular, we show that when $M \to \infty$, the optimal gains can be found in the general case via a quadratically constrained linear program, and we also show that closed-form solutions are possible for limiting values of the power constraint P. As in the NP detector case, the optimal solution is independent of the sensor phase. Simulation results demonstrate that reducing

transmit power by $1/\sqrt{M}$ to maintain a constant deflection as M grows results in a constant PD. Note that although this result is superficially similar to a result in [5], the case we consider is considerably different since it involves the energy detector which requires no CSI, unlike [5] which assumes a minimum mean-squared error channel estimate obtained using pilot signals. Also, unlike [5], we do not assume a uniform power allocation, but as mentioned above we instead derive optimal sensor transmit gains and illustrate when these optimal gains provide significantly better detection performance.

(4) For the LMMSE estimator, we prove that a constant MSE can be achieved by decreasing the transmit power as $1/M$ as the number of FC antennas M grows. This result is obtained by generalizing the asymptotic results for the NP detector to the LMMSE estimator, and showing that the PD of the NP detector and the LMMSE mean-squared error (MSE) both obey a similar rule as $M \to \infty$. We also derive bounds on the MSE for the limiting cases $P \to 0$ and $P \to \infty$, and show similar behavior for these bounds as in the case of PD for the NP detector.

The remainder of this chapter is organized as follows. In Section 6.2, we introduce the signal model and derive basic results for PD and PFA of NP detector. In Section 6.3, we prove the main results for the NP detector and LMMSE estimator, and we formulate and solve the sensor transmission gain optimization problem to maximize PD for a given PFA under a sum transmit power constraint. The deflection of the energy detector is analyzed in Section 6.4, and the problem of calculating the transmission gains that maximize the deflection is solved. The results of several simulation studies are provided in Section 6.5 to validate the theoretical derivations, and the conclusions of this chapter are summarized in Section 6.6.

6.2 Signal Model and Neyman-Pearson Detector

Based on the signal mode in (2.2), when the NP criterion is used to distinguish between the hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$, the NP detector decides $\mathcal{H}_1$ if [66]

$$L(\mathbf{y}) = \frac{p(\mathbf{y};\mathcal{H}_1)}{p(\mathbf{y};\mathcal{H}_0)} > \gamma \tag{6.1}$$

for a given threshold γ, where $p(\mathbf{y};\mathcal{H}_1)$ and $p(\mathbf{y};\mathcal{H}_0)$ are the conditional probability density functions (PDFs) of $\mathbf{y}$ under $\mathcal{H}_1$ and $\mathcal{H}_0$, respectively. Assume the measurement noise at the sensors is independent, so that the covariance of $\mathbf{v}$ is given by $\mathbf{V} = \text{diag}\{\sigma_{v,1}^2 \ \cdots \ \sigma_{v,N}^2\}$. Since $\mathbf{y}$ is Gaussian under both $\mathcal{H}_1$ and $\mathcal{H}_0$, we have [66]

$$\begin{aligned} p(\mathbf{y};\mathcal{H}_1) &= \frac{1}{\pi^M \text{det}(\mathbf{C}_s+\mathbf{C}_w)} \exp\left(-\mathbf{y}^H(\mathbf{C}_s+\mathbf{C}_w)^{-1}\mathbf{y}\right) \\ p(\mathbf{y};\mathcal{H}_0) &= \frac{1}{\pi^M \text{det}(\mathbf{C}_w)} \exp\left(-\mathbf{y}^H\mathbf{C}_w^{-1}\mathbf{y}\right) \ , \end{aligned} \tag{6.2}$$

where $\mathbf{C}_w = \mathbf{HDVD}^H\mathbf{H}^H + \sigma_n^2\mathbf{I}_M$ is the covariance of $\mathbf{y}$ under $\mathcal{H}_0$, $(\cdot)^H$ is the conjugate transpose, $\mathbf{C}_s = \sigma_\theta^2\mathbf{Haa}^H\mathbf{H}^H$ and $\mathbf{C}_w + \mathbf{C}_s$ is the covariance of $\mathbf{y}$ under $\mathcal{H}_1$.

Lemma 1. *Based on the signal model in (2.2) and the conditional PDFs in (6.2), the NP detector in (6.1) is equivalent to deciding $\mathcal{H}_1$ if*

$$\sigma_\theta^2 |\mathbf{a}^H\mathbf{H}^H\mathbf{C}_w^{-1}\mathbf{y}|^2 > \gamma' \ , \tag{6.3}$$

where

$$\begin{aligned} \gamma' &= (1+\sigma_\theta^2 g(\mathbf{a})) \ln\left[\gamma(1+\sigma_\theta^2 g(\mathbf{a}))\right] \\ g(\mathbf{a}) &= \mathbf{a}^H\mathbf{H}^H\mathbf{C}_w^{-1}\mathbf{Ha} \ . \end{aligned}$$

Proof: Substituting $p(\mathbf{y};\mathcal{H}_1)$ and $p(\mathbf{y};\mathcal{H}_0)$ from (6.2) into (6.1) and calculating the logarithm of (6.1), we have

$$\mathbf{y}^H(\mathbf{C}_w^{-1} - (\mathbf{C}_s+\mathbf{C}_w)^{-1})\mathbf{y} > \ln\left(\gamma(1+\sigma_\theta^2 g(\mathbf{a}))\right) \ , \tag{6.4}$$

where $g(\mathbf{a}) = \mathbf{a}^H\mathbf{H}^H\mathbf{C}_w^{-1}\mathbf{H}\mathbf{a}$, and in the derivation we used the following equality

$$\begin{aligned}
&\ln(\gamma) + \text{lndet}(\mathbf{C}_s + \mathbf{C}_w) - \text{lndet}(\mathbf{C}_w) \\
&= \ln(\gamma) + \text{lndet}(\mathbf{C}_s\mathbf{C}_w^{-1} + \mathbf{I}_M) \\
&\overset{(a)}{=} \ln(\gamma) + \ln\left(1 + \lambda_{\max}(\mathbf{C}_s\mathbf{C}_w^{-1})\right) \\
&= \ln\left(\gamma(1 + \sigma_\theta^2 g(\mathbf{a}))\right) \ ,
\end{aligned}$$

where (a) is due to the fact that $\mathbf{C}_s\mathbf{C}_w^{-1}$ is a rank-one matrix and $\lambda_{\max}(\cdot)$ is the largest eigenvalue of its matrix argument. Using the matrix inversion lemma, the left hand side of (6.4) is calculated as

$$(\mathbf{C}_s + \mathbf{C}_w)^{-1} - \mathbf{C}_w^{-1} = \frac{\sigma_\theta^2}{1 + \sigma_\theta^2 g(\mathbf{a})}\mathbf{C}_w^{-1}\mathbf{H}\mathbf{a}\mathbf{a}^H\mathbf{H}^H\mathbf{C}_w^{-1} \ , \tag{6.5}$$

and substituting (6.5) into (6.4) will produce the desired result. ■

For the NP detector in (6.3), the probability of detection P_D and probability of false alarm P_{FA} are defined as

$$\begin{aligned}
P_D &= \Pr\left(\sigma_\theta^2|\mathbf{a}^H\mathbf{H}^H\mathbf{C}_w^{-1}\mathbf{y}|^2 > \gamma'|\mathcal{H}_1\right) \\
P_{FA} &= \Pr\left(\sigma_\theta^2|\mathbf{a}^H\mathbf{H}^H\mathbf{C}_w^{-1}\mathbf{y}|^2 > \gamma'|\mathcal{H}_0\right) \ .
\end{aligned}$$

To evaluate P_D, we first rewrite it as

$$P_D = \Pr\left(\sigma_\theta^2\tilde{\mathbf{y}}^H\mathbf{W}\tilde{\mathbf{y}} > \gamma'|\mathcal{H}_1\right) \ , \tag{6.6}$$

where $\tilde{\mathbf{y}} = (\mathbf{C}_s + \mathbf{C}_w)^{-\frac{1}{2}}\mathbf{y}$ and

$$\mathbf{W} = (\mathbf{C}_s + \mathbf{C}_w)^{\frac{1}{2}}\mathbf{C}_w^{-1}\mathbf{H}\mathbf{a}\mathbf{a}^H\mathbf{H}^H\mathbf{C}_w^{-1}(\mathbf{C}_s + \mathbf{C}_w)^{\frac{1}{2}} \ .$$

Since $\mathbf{y} \sim \mathcal{CN}(0, \mathbf{C}_s + \mathbf{C}_w)$ under $\mathcal{H}_1$, $\tilde{\mathbf{y}} = (\mathbf{C}_s + \mathbf{C}_w)^{-\frac{1}{2}}\mathbf{y}$ is distributed as $\mathcal{CN}(0, \mathbf{I}_M)$. Defining the eigendecomposition of $\mathbf{W}$ as

$$\mathbf{W} = \mathbf{U}\mathbf{G}\mathbf{U}^H$$

where $\mathbf{G} = \text{diag}\{g(\mathbf{a}) + \sigma_\theta^2 g(\mathbf{a})^2, 0, \cdots, 0\}$, equation (6.6) becomes

$$\begin{aligned} P_D &= \Pr\left(\sigma_\theta^2 \tilde{\mathbf{y}}^H \mathbf{U}\mathbf{G}\mathbf{U}^H \tilde{\mathbf{y}} > \gamma^{'} | \mathcal{H}_1\right) \\ &\overset{(b)}{=} \Pr\left(\sigma_\theta^2 \tilde{\mathbf{y}}^H \mathbf{G} \tilde{\mathbf{y}} > \gamma^{'} | \mathcal{H}_1\right) \\ &\overset{(c)}{=} \exp\left(-\frac{\gamma^{'}}{\sigma_\theta^4 g(\mathbf{a})^2 + \sigma_\theta^2 g(\mathbf{a})}\right) \ , \end{aligned} \tag{6.7}$$

where (b) results since the unitary transformation $\mathbf{U}$ does not change the distribution of $\tilde{\mathbf{y}}$, and (c) holds since $\tilde{\mathbf{y}}^H \mathbf{G} \tilde{\mathbf{y}}$ has a scaled chi-square distribution with two degrees of freedom. In a similar way, P_{FA} can be derived as

$$P_{FA} = \exp\left(-\frac{\gamma^{'}}{\sigma_\theta^2 g(\mathbf{a})}\right) \ . \tag{6.8}$$

6.3 Neyman-Pearson Detector Optimization and Analysis

Both P_D and P_{FA} are functions of the sensor transmission gains $\mathbf{a}$, and thus it is natural to find values for the entries of $\mathbf{a}$ that optimize detection performance. Here

we will show how to find $\mathbf{a}$ such that P_D is maximized for a given P_{FA}. According to (6.8), the threshold required to achieve $P_{FA} = \epsilon$ is

$$\gamma' = -\sigma_\theta^2 g(\mathbf{a}) \ln \epsilon \ .$$

When substituted into (6.7), this threshold yields

$$P_D = \exp\left(\frac{\ln \epsilon}{\sigma_\theta^2 g(\mathbf{a}) + 1}\right) \ . \tag{6.9}$$

Since $\ln \epsilon < 0$, P_D is maximized when the signal-to-noise ratio (SNR) $g(\mathbf{a})$ is maximized. Thus, the problem becomes

$$\begin{aligned} \max_{\mathbf{a}} \quad & g(\mathbf{a}) = \mathbf{a}^H \mathbf{H}^H (\mathbf{H}\mathbf{D}\mathbf{V}\mathbf{D}^H \mathbf{H}^H + \sigma_n^2 \mathbf{I}_M)^{-1} \mathbf{H}\mathbf{a} \\ s.t. \quad & \mathbf{a}^H \mathbf{a} = P \ , \end{aligned} \tag{6.10}$$

where P denotes the constraint on the sum sensor transmit power. This result was derived in [40] by examining the behavior of the error exponent as the number of sensors went to infinity. Here we see the result holds for fixed and finite values of N. The role of $g(\mathbf{a})$ in determining estimation performance for θ has also been noted in [10, 39]. In general, finding a solution to (6.10) is difficult due to its nonlinear and non-convex dependence on $\mathbf{a}$. A simpler solution was found to be possible in [10] if the sensor gains were restricted to all have the same magnitude and only the phase was optimized. In this case, the solution was shown to be found via a relaxed SDP. In this chapter, we show that a closed-form "water-filling" type of solution for (6.10) is possible under the assumption that $M \to \infty$.

6.3.1 Energy Efficiency

For our analysis, we assume the wireless fading channel between the sensor node i and FC is modeled as

$$\mathbf{h}_i = \frac{\tilde{\mathbf{h}}_i}{\sqrt{d_i^{\alpha}}} \,, \tag{6.11}$$

where d_i is the distance between the sensor node and FC, α is the path loss exponent, and $\tilde{\mathbf{h}}_i \in \mathbb{C}^{M\times 1}$ is a complex Gaussian vector with distribution $\mathcal{CN}(0, \mathbf{I}_M)$. Note that the assumption here of independent identically distributed channel coefficients is made primarily to enable the asymptotic analysis of the detection performance at the FC. The following theorem characterizes the energy efficiency of the NP detector for large M.

Theorem 4. *Assuming Rayleigh fading wireless channels, as the number of FC antennas M tends to infinity, the transmit gain $|a_i|^2$ at each sensor can be reduced by $1/M$ to* almost surely *achieve the same optimal P_D for a given fixed P_{FA}.*

Proof: We will show that as $M \to \infty$, the function $g(\mathbf{a})$ in (6.9) and (6.10) remains constant if the product $M|a_i|^2$ is held constant. We first use the matrix inversion lemma to show that

$$\left(\mathbf{HDVD}^H\mathbf{H}^H + \sigma_n^2\mathbf{I}_M\right)^{-1} = \frac{1}{\sigma_n^2}\mathbf{I}_M - \frac{1}{\sigma_n^4}\mathbf{H}\left(\mathbf{E}^{-1} + \frac{1}{\sigma_n^2}\mathbf{H}^H\mathbf{H}\right)^{-1}\mathbf{H}^H \,, \tag{6.12}$$

where $\mathbf{E} = \mathbf{DVD}^H$. Note that we have assumed that $|a_i| > 0$ to guarantee the matrix inverse $\mathbf{E}^{-1}$ exists, but we will see that the final solution allows $|a_i| \to 0$. Substituting (6.12) into $g(\mathbf{a})$ yields

$$g(\mathbf{a}) = \frac{1}{\sigma_n^2}\mathbf{a}^H\mathbf{H}^H\mathbf{H}\mathbf{a} - \frac{1}{\sigma_n^4}\mathbf{a}^H\mathbf{H}^H\mathbf{H}\left(\mathbf{E}^{-1}+\frac{1}{\sigma_n^2}\mathbf{H}^H\mathbf{H}\right)^{-1}\mathbf{H}^H\mathbf{H}\mathbf{a} \,. \tag{6.13}$$

For large M, the product $\mathbf{H}^H\mathbf{H}$ converges almost surely to [5]:

$$\lim_{M\to\infty} \frac{1}{M}\mathbf{H}^H\mathbf{H} = \text{diag}\left\{\frac{1}{d_1^\alpha}\cdots\frac{1}{d_N^\alpha}\right\} , \tag{6.14}$$

and substituting (6.14) into (6.13) yields, after some calculations,

$$\lim_{M\to\infty} g(\mathbf{a}) = \lim_{M\to\infty} \sum_{i=1}^{N} \frac{M|a_i|^2}{\sigma_n^2 d_i^\alpha + \sigma_{v,i}^2 M|a_i|^2} . \tag{6.15}$$

We see that $g(\mathbf{a})$ remains asymptotically unchanged as long as the product $M|a_i|^2$ is held constant, and thus asymptotically equivalent detection performance can be achieved if any decrease in sensor transmit power is balanced by a corresponding increase in the number of FC antennas. ■

6.3.2 Sensor Gain Optimization

Based on (6.15), when $M \to \infty$, the original problem (6.10) can be rewritten as

$$\begin{aligned} \max_{|a_i|^2} \quad & \sum_{i=1}^{N} \frac{M|a_i|^2}{\sigma_n^2 d_i^\alpha + \sigma_{v,i}^2 M|a_i|^2} \\ s.t. \quad & \sum_{i=1}^{N} |a_i|^2 = P . \end{aligned} \tag{6.16}$$

We see from this formulation that as $M \to \infty$, only the magnitude of a_i is important in determining the detection performance, and we see that there is no problem if $|a_i| \to 0$ for some i. As M grows, eventually we reach the point where $\sigma_{v,i}^2 M|a_i|^2 \gg \sigma_n^2 d_i^\alpha$, in which case the choice of the sensor gains no longer matters. However, we will see in the simulations that for moderately large values of M, optimizing (6.16) over $|a_i|$ provides a significant benefit, especially when P is relatively small.

Define a new variable $x_i = |a_i|^2$, so that problem (6.16) is equivalent to

$$\begin{aligned} \min_{x_i} \quad & \sum_{i=1}^{N} \frac{-Mx_i}{\sigma_n^2 d_i^\alpha + \sigma_{v,i}^2 M x_i} \\ s.t. \quad & \sum_{i=1}^{N} x_i = P \\ & 0 \leq x_i \ . \end{aligned} \tag{6.17}$$

In problem (6.17), the objective function is the sum of N convex functions of x_i, and the constraints are linear with respect to the variable x_i, so (6.17) is a convex problem and we can find a "closed-form" solution using the Karush-Kuhn-Tucker (KKT) conditions [52]. The Lagrangian of (6.17) is given by:

$$\mathcal{L}(x_i, \lambda, \mu_i) = \sum_{i=1}^{N} \frac{-Mx_i}{\sigma_n^2 d_i^\alpha + \sigma_{v,i}^2 M x_i} + \lambda \left(\sum_{i=1}^{N} x_i - P \right) - \sum_{i=1}^{N} \mu_i x_i \ , \tag{6.18}$$

and the corresponding KKT conditions are as follows:

$$\begin{aligned} \frac{-\sigma_n^2 d_i^\alpha M}{(\sigma_n^2 d_i^\alpha + \sigma_{v,i}^2 M x_i)^2} + \lambda - \mu_i &= 0 \\ \lambda \left(\sum_{i=1}^{N} x_i - P \right) &= 0 \\ \sum_{i=1}^{N} x_i - P &= 0 \\ x_i \mu_i &= 0 \\ x_i, \ \mu_i, \ \lambda &\geq 0 \ . \end{aligned}$$

After some simple manipulations, we arrive at the following optimal solution to (6.16):

$$|a_i^*| = \sqrt{\frac{\left(\sqrt{\frac{\sigma_n^2 d_i^\alpha M}{\lambda}} - \sigma_n^2 d_i^\alpha \right)^+}{\sigma_{v,i}^2 M}} \ , \tag{6.19}$$

where $\lambda > 0$ is chosen such that $\sum_{i=1}^{N} |a_i^*|^2 = P$. Lower and upper bounds for λ are given by

$$\lambda_u = \frac{M}{\sigma_n^2 \min_i\{d_i^\alpha\}}$$
$$\lambda_l = \min_i \left\{ \frac{\sigma_n^2 d_i^\alpha M}{(\sigma_n^2 d_i^\alpha + \sigma_{v,i}^2 PM)^2} \right\} ,$$

and the unique value of λ can be found via a simple bisection search over $[\lambda_l, \lambda_u]$.

Note that while implementing the NP detector in (6.3) requires instantaneous CSI, the large M assumption allows the optimal gains in (6.19) to be computed using only the channel statistics, determined in this case by the distances of the FC to the sensors. This is of interest since it means the sensors will not require frequent feedback from the FC to update their transmit gains.

6.3.3 Single-Antenna FC

It is of interest to consider the single-antenna FC case separately, both for purposes of comparison and because in this case an exact solution can be obtained. When $M = 1$, the signal model reduces to

$$\mathcal{H}_0 : y = \mathbf{a}^H \mathbf{F} \mathbf{v} + n$$
$$\mathcal{H}_1 : y = \mathbf{a}^H \mathbf{h} \theta + \mathbf{a}^H \mathbf{F} \mathbf{v} + n \, , \tag{6.20}$$

where $\mathbf{a} = [a_1 \cdots a_N]^H$, $\mathbf{h} = [h_1 \cdots h_N]^T$, $\mathbf{F} = \text{diag}\{h_1 \cdots h_N\}$ and h_i denotes the scalar channel gain between the ith sensor and the FC. The conditional PDFs of y

under $\mathcal{H}_1$ and $\mathcal{H}_0$ are given by

$$p(y;\mathcal{H}_1) = \frac{1}{\pi(\sigma_s^2+\sigma_w^2)}\exp\left(-\frac{|y|^2}{\sigma_s^2+\sigma_w^2}\right)$$
$$p(y;\mathcal{H}_0) = \frac{1}{\pi\sigma_w^2}\exp\left(-\frac{|y|^2}{\sigma_w^2}\right) ,$$

where $\sigma_s^2 = \sigma_\theta^2\mathbf{a}^H\mathbf{h}\mathbf{h}^H\mathbf{a}$ and $\sigma_w^2 = \mathbf{a}^H\mathbf{F}\mathbf{V}\mathbf{F}^H\mathbf{a}+\sigma_n^2$.

For a given threshold $\tilde{\gamma}$, the NP detector decides $\mathcal{H}_1$ if

$$L(y) = \frac{p(y;\mathcal{H}_1)}{p(y;\mathcal{H}_0)} > \tilde{\gamma} ,$$

which results in deciding $\mathcal{H}_1$ if

$$|y|^2 > \ln\left(\tilde{\gamma}\left(1+\frac{\sigma_s^2}{\sigma_w^2}\right)\right)\left(1+\frac{\sigma_w^2}{\sigma_s^2}\right)\sigma_w^2 . \tag{6.21}$$

Following an analysis similar to the multi-antenna case, the probability of detection P_D^s and the probability of false alarm P_{FA}^s for the single-antenna FC are given by

$$P_D^s = \exp\left(-\frac{\tilde{\gamma}'}{\sigma_s^2+\sigma_w^2}\right) \tag{6.22}$$
$$P_{FA}^s = \exp\left(-\frac{\tilde{\gamma}'}{\sigma_w^2}\right) , \tag{6.23}$$

where $\tilde{\gamma}' = \ln\left(\tilde{\gamma}\left(1+\frac{\sigma_s^2}{\sigma_w^2}\right)\right)\left(\sigma_w^2+\frac{\sigma_w^4}{\sigma_s^2}\right)$. To fix $P_{FA}^s = \epsilon$, we set $\tilde{\gamma}' = -\sigma_w^2\ln\epsilon$, and maximizing P_D^s for a fixed P_{FA}^s is equivalent to

$$\begin{aligned} \max_{\mathbf{a}} \quad & \frac{\sigma_s^2}{\sigma_w^2} = \frac{\sigma_\theta^2\mathbf{a}^H\mathbf{h}\mathbf{h}^H\mathbf{a}}{\mathbf{a}^H\mathbf{F}\mathbf{V}\mathbf{F}^H\mathbf{a}+\sigma_n^2} \\ s.t. \quad & \mathbf{a}^H\mathbf{a} = P . \end{aligned} \tag{6.24}$$

Problem (6.24) is essentially identical to problem (3) in [8], and using the same

solution method derived in [8] leads to

$$\tilde{\mathbf{a}}^* = \sqrt{\frac{P}{\mathbf{h}^H \mathbf{R}^{-2} \mathbf{h}}} \mathbf{R}^{-1} \mathbf{h} \,, \tag{6.25}$$

where $\mathbf{R} = \mathbf{F}\mathbf{V}\mathbf{F}^H + \frac{\sigma_n^2}{P}\mathbf{I}_N$, and the maximum value of $\frac{\sigma_s^2}{\sigma_w^2}$ is

$$\left.\frac{\sigma_s^2}{\sigma_w^2}\right|_{\tilde{\mathbf{a}}^*} = \sigma_\theta^2 \mathbf{h}^H \mathbf{R}^{-1} \mathbf{h} \,. \tag{6.26}$$

In the following theorem, we compare the detection performance of single- and multi-antenna FCs under low and high transmit power scenarios.

Theorem 5. *Assume $P_{FA} = \epsilon$ and $M \to \infty$. When $P = O(1/M) \to 0$, the NP detector implemented by an FC with M antennas achieves a P_D lower bounded by*

$$P_D > \epsilon^{\frac{1}{1+\frac{\sigma_\theta^2}{3}\sum_{i=1}^{N}\frac{1}{\sigma_{v,i}^2}}} \,, \tag{6.27}$$

while the P_D^s for a single-antenna FC is bounded by

$$\epsilon < P_D^s < \epsilon^{\frac{1}{1+\zeta}}, \tag{6.28}$$

where $\zeta = \frac{1}{2M}\sum_{i=1}^{N}\frac{\sigma_\theta^2 d_i^\alpha}{\sigma_{v,i}^2}\mathbf{h}^H\mathbf{h} \to 0$ in probability. When $P \to \infty$, both P_D and P_D^s converge from below to the same upper bound:

$$\{P_D, P_D^s\} \uparrow \epsilon^{\frac{1}{1+\sigma_\theta^2\sum_{i=1}^{N}\frac{1}{\sigma_{v,i}^2}}} \,. \tag{6.29}$$

Proof: Beginning with the low transmit power case, assume the following suboptimal

choice for the transmission gains: $|\bar{a}_i| = \sqrt{\frac{\sigma_n^2 d_i^\alpha}{2\sigma_{v,i}^2 M}}$, which results in

$$P = \sum_{i=1}^{N} |\bar{a}_i|^2 = \frac{1}{2M} \sum_{i=1}^{N} \frac{\sigma_n^2 d_i^\alpha}{\sigma_{v,i}^2} = O(1/M) \ , \tag{6.30}$$

and hence $P \to 0$ as $M \to \infty$. Substituting $|\bar{a}_i|$ into (6.15), we have

$$g(\bar{\mathbf{a}}) = \frac{1}{3} \sum_{i=1}^{N} \frac{1}{\sigma_{v,i}^2} \ ,$$

where $\bar{\mathbf{a}} = [\bar{a}_1 \cdots \bar{a}_N]^T$. The value for $g(\bar{\mathbf{a}})$ can serve as a lower bound for $g(\mathbf{a})$ when evaluated at the optimal solution $\mathbf{a}^*$ obtained using (6.19) and using P in (6.30) as the power constraint:

$$g(\mathbf{a}^*) \geq \frac{1}{3} \sum_{i=1}^{N} \frac{1}{\sigma_{v,i}^2} \ . \tag{6.31}$$

Substituting (6.31) into (6.9), we have the lower bound for the multi-antenna FC:

$$P_D \geq \epsilon^{\frac{1}{1+\frac{\sigma_\theta^2}{3}\sum_{i=1}^{N}\frac{1}{\sigma_{v,i}^2}}} > \epsilon \ . \tag{6.32}$$

For the single-antenna FC, according to (6.26) we have the following upper bound since $\frac{P}{\sigma_n^2}\mathbf{I}_N \succeq \mathbf{R}^{-1}$:

$$\frac{\sigma_s^2}{\sigma_w^2} \overset{(d)}{\leq} \frac{\sigma_\theta^2 P}{\sigma_n^2} \mathbf{h}^H \mathbf{h} \ , \tag{6.33}$$

where $\frac{P}{\sigma_n^2}\mathbf{I}_N \succeq \mathbf{R}^{-1}$ denotes $\frac{P}{\sigma_n^2}\mathbf{I}_N - \mathbf{R}^{-1}$ is a positive semidefinite matrix. Using (6.33) and (6.30) together with (6.23) and (6.22), it is easy to show that

$$P_D^s \leq \epsilon^{\frac{1}{1+\zeta}}, \tag{6.34}$$

where $\zeta = \frac{1}{2M}\sum_{i=1}^{N} \frac{\sigma_\theta^2 d_i^\alpha}{\sigma_{v,i}^2} \mathbf{h}^H \mathbf{h}$. According to the Rayleigh channel model, $\mathbf{h}^H \mathbf{h}$ is the sum of weighted chi-squared random variables, and for an arbitrary positive number

τ we have

$$\lim_{M\to\infty} \Pr\left(\zeta > \tau\right) \leq \lim_{M\to\infty} \Pr\left(\frac{\sigma_\theta^2 N}{4M} \frac{\max_i \frac{d_i^\alpha}{\sigma_{v,i}^2}}{\min_i d_i^\alpha} \chi^2_{(2N)} > \tau\right) = 0 \ ,$$

where $\chi^2_{(2N)}$ denotes a chi-square variable with $2N$ degrees of freedom. Thus, ζ converges to 0 in probability and hence P_D^s converges to ϵ in probability.

From (6.15), it is clear that for very large M, $g(\mathbf{a})$ is upper bounded by

$$g(\mathbf{a}) \leq \sum_{i=1}^{N} \frac{1}{\sigma_{v,i}^2} \ . \tag{6.35}$$

Note that the lower bound in (6.31) is one third the upper bound in (6.35). When $P \to \infty$ and hence $|a_i|$ is large, the upper bound in (6.35) can be asymptotically achieved even with an equal power allocation $|a_i| = \sqrt{P/N}$. Also, we see that to maximize the upper bound for $g(\mathbf{a})$ in this case, all the sensors should transmit. Plugging (6.35) into (6.9), we have the following upper bound for P_D:

$$P_D \leq \epsilon^{\frac{1}{1+\sigma_\theta^2 \sum_{i=1}^{N} \frac{1}{\sigma_{v,i}^2}}} \ . \tag{6.36}$$

For the single-antenna FC, according to (6.26), we have the following bound as $P \to \infty$ since $(\mathbf{F}\mathbf{V}\mathbf{F}^H)^{-1} \succeq \mathbf{R}^{-1}$:

$$\frac{\sigma_s^2}{\sigma_w^2} \overset{(e)}{\leq} \sigma_\theta^2 \sum_{i=1}^{N} \frac{1}{\sigma_{v,i}^2} \ . \tag{6.37}$$

Using (6.37) together with (6.23) and (6.22) yields

$$P_{PD}^s \leq \epsilon^{\frac{1}{1+\sigma_\theta^2 \sum_{i=1}^{N} \frac{1}{\sigma_{v,i}^2}}} \ . \tag{6.38}$$

Note that for both (6.36) and (6.38), the inequality is asymptotically achieved as $P \to \infty$, and this completes the proof. ■

Theorem 5 shows that when the transmit power P goes to zero, P_D^s for a single-antenna FC converges to P_{FA}^s regardless of the sensor network scenario, while P_D for a multi-antenna FC is strictly greater than P_{FA}, provided that $M \to \infty$ and $P \to 0$ no faster than $O(1/M)$. When σ_θ^2 is large and the $\sigma_{v,i}^2$ are small, P_D can in fact still converge to a value near unity. On the other hand, when P is large, both P_D and P_D^s converge to the same upper bound, and there is no benefit to having multiple antennas at the FC.

6.3.4 LMMSE Estimation

While this chapter is focused on detection, we show here that similar results hold for LMMSE estimation. According to the Gauss-Markov Theorem [25], the LMMSE estimator of θ is

$$\hat{\theta} = \frac{\mathbf{a}^H \mathbf{H}^H (\mathbf{H}\mathbf{D}\mathbf{V}\mathbf{D}^H \mathbf{H}^H + \sigma_n^2 \mathbf{I}_M)^{-1} \mathbf{y}}{\sigma_\theta^{-2} + \mathbf{a}^H \mathbf{H}^H (\mathbf{H}\mathbf{D}\mathbf{V}\mathbf{D}^H \mathbf{H}^H + \sigma_n^2 \mathbf{I}_M)^{-1} \mathbf{H}\mathbf{a}} \,, \tag{6.39}$$

and the mean-squared error is calculated as

$$\mathrm{MSE}(\hat{\theta}) = \mathbb{E}\{|\theta - \hat{\theta}|^2\} = \frac{1}{\sigma_\theta^{-2} + g(\mathbf{a})} \,, \tag{6.40}$$

where $g(\mathbf{a}) = \mathbf{a}^H \mathbf{H}^H (\mathbf{H}\mathbf{D}\mathbf{V}\mathbf{D}^H \mathbf{H}^H + \sigma_n^2 \mathbf{I}_M)^{-1} \mathbf{H}\mathbf{a}$, as defined in (6.10). Thus, the problem of choosing the gains $\mathbf{a}$ to minimize the MSE is identical to the problem of maximizing P_D for a fixed P_{FA} in (6.10), and the same conclusions drawn above regarding energy efficiency and the optimal sensor gains apply here as well. This is also true for the single-antenna FC, as it can be easily shown that minimizing MSE requires maximization of σ_s^2/σ_w^2, as with the NP detector.

The following corollary to Theorem 5 can also be established.

Corollary 1. *When $M \to \infty$ and $P = O(1/M) \to 0$, the MSE of the LMMSE estimator of θ is upper bounded by*

$$\mathrm{MSE}(\hat{\theta}) < \frac{1}{\sigma_\theta^{-2} + \frac{1}{3}\sum_{i=1}^{N} \frac{1}{\sigma_{v,i}^2}} , \tag{6.41}$$

while the MSE achieved by the single-antenna FC is bounded by

$$\frac{\sigma_\theta^2}{1+\zeta} < \mathrm{MSE}(\hat{\theta}_s) < \sigma_\theta^2 , \tag{6.42}$$

where $\zeta = \frac{1}{2M}\sum_{i=1}^{N} \frac{\sigma_\theta^2 d_i^\alpha}{\sigma_{v,i}^2}\mathbf{h}^H\mathbf{h} \to 0$ in probability. When $P \to \infty$, both MSEs converge from above to the same lower bound:

$$\mathrm{MSE}(\hat{\theta}, \hat{\theta}_s) \geq \frac{1}{\sigma_\theta^{-2} + \sum_{i=1}^{N} \frac{1}{\sigma_{v,i}^2}} . \tag{6.43}$$

Proof: The proof essentially follows that for Theorem 5 and is thus omitted. ■

6.4 Energy Detector Analysis and Sensor Gain Optimization

Obtaining the instantaneous CSI required for the NP detector consumes sensor power and could be difficult in fast fading scenarios. Computing the NP test statistic also requires the inverse of the $M \times M$ channel-dependent matrix $\mathbf{C}_w$, which may be challenging when M is large. Consequently, it is of interest to study computationally simpler approaches for detection in sensor networks that can be applied when the CSI for the sensors is unknown. In this section, we examine the performance of the

energy detector (ED), which decides $\mathcal{H}_1$ if

$$T = \frac{1}{M}\mathbf{y}^H\mathbf{y} > \hat{\gamma} \, , \tag{6.44}$$

for some predefined threshold $\hat{\gamma}$.

Under either $\mathcal{H}_0$ or $\mathcal{H}_1$, the test statistic T can be expressed as

$$T = \frac{1}{M}\sum_{i=1}^{M}\frac{\lambda_i}{2}\chi_i^2(2) \, , \tag{6.45}$$

where λ_i is the ith eigenvalue of the covariance matrix $\mathbf{C}_w$ (under $\mathcal{H}_0$) or $\mathbf{C}_s + \mathbf{C}_w$ (under $\mathcal{H}_1$) and the $\chi_i^2(2)$ terms represent independent chi-squared random variables with two degrees of freedom. Thus, while the ED test statistic does not require CSI, computing the ED probability of detection P_D^e and false alarm P_{FA}^e does. When M is large, one could consider approximating T as a normal random variable using the Central Limit Theorem. However, because the largest N eigenvalues of λ_i will increase with M, Lindeberg's condition is not satisfied and the normal distribution can not provide a good approximation for T. Even if the distribution of T could be computed, it would be a complicated function of the transmit gains $\mathbf{a}$ and would be difficult to optimize. Instead, in the following we will use the so-called *deflection* [66–69] of T as the metric of detection performance, which will allow us to obtain an optimal value for $\mathbf{a}$ that does not depend on CSI as $M \to \infty$.

6.4.1 Energy Efficiency

The deflection coefficient for a given test statistic T is defined as [66]

$$D(T) = \frac{\left(\mathbb{E}\{T|\mathcal{H}_1\} - \mathbb{E}\{T|\mathcal{H}_0\}\right)^2}{\mathrm{Var}\{T|\mathcal{H}_0\}} \, , \tag{6.46}$$

where $\mathbb{E}\{\cdot\}$ and $\mathrm{Var}\{\cdot\}$ denote the expectation and variance of a random variable, respectively. The deflection metric in (6.46) can be viewed as the normalized distance between the distributions of T under $\mathcal{H}_0$ or $\mathcal{H}_1$, and is generally regarded as an accurate metric for characterizing detection performance [67]. Note that a *modified deflection* is proposed in [69], which replaces $\mathrm{Var}\{T|\mathcal{H}_0\}$ in (6.46) with $\mathrm{Var}\{T|\mathcal{H}_1\}$. As mentioned below, both deflection statistics yield very similar problem formulations that can be solved via the same approach. As derived in the following theorem, one of the key properties of the energy detector for our WSN application is that the sensor transmit power can be reduced by a factor of $1/\sqrt{M}$ to maintain a constant deflection as $M \to \infty$.

Theorem 6. *Assuming Rayleigh fading channels, the deflection of the test statistic* $T = \frac{1}{M}\mathbf{y}^H\mathbf{y}$ almost surely *remains constant as* $M \to \infty$ *provided that the sensor transmit power satisfies* $|a_i|^2 = \frac{P_i}{\sqrt{M}}$ *for arbitrary constant* P_i *.*

Proof: Using the definition in (6.46),

$$\lim_{M\to\infty} D\left(T\right) = \frac{\left(\mu_{e,1} - \mu_{e,0}\right)^2}{\sigma_{e,0}^2} = \frac{\sigma_\theta^4 \left(\sum_{i=1}^N \frac{|a_i|^2}{d_i^\alpha}\right)^2}{\sum_{i=1}^N \left(\frac{\sigma_{v,i}^2|a_i|^2}{d_i^\alpha} + \frac{\sigma_n^2}{M}\right)^2 + \frac{M-N}{M^2}\sigma_n^4} \,, \tag{6.47}$$

where the parameters $\mu_{e,1}$, $\mu_{e,0}$ and $\sigma_{e,0}^2$ are defined and calculated below. For $\mu_{e,1}$,

$$\begin{aligned}
\mu_{e,1} &= \lim_{M\to\infty} \mathbb{E}\left\{ \frac{1}{M}\mathbf{y}^H\mathbf{y} \middle| \mathcal{H}_1 \right\} \\
&= \lim_{M\to\infty} \frac{1}{M}\mathbb{E}\left\{\tilde{\mathbf{y}}^H(\mathbf{C}_w + \mathbf{C}_s)\tilde{\mathbf{y}}\right\} \\
&= \lim_{M\to\infty} \frac{1}{M}\mathrm{tr}\left(\mathbf{C}_w + \mathbf{C}_s\right) \\
&= \lim_{M\to\infty} \frac{1}{M}\mathrm{tr}\left(\sigma_\theta^2\mathbf{H}^H\mathbf{H}\mathbf{a}\mathbf{a}^H + \mathbf{H}^H\mathbf{H}\mathbf{D}\mathbf{V}\mathbf{D}^H\right) + \sigma_n^2 \\
&\overset{(i)}{=} \sum_{i=1}^N \frac{\left(\sigma_\theta^2 + \sigma_{v,i}^2\right)|a_i|^2}{d_i^\alpha} + \sigma_n^2 \,,
\end{aligned}$$

where $\tilde{\mathbf{y}}$ has distribution $\mathcal{CN}(0, \mathbf{I}_M)$ and in (i) we used (6.14). Similarly, we have

$$\begin{aligned}
\mu_{e,0} &= \lim_{M\to\infty} \mathbb{E}\left\{ \frac{1}{M}\mathbf{y}^H\mathbf{y} \middle| \mathcal{H}_0 \right\} \\
&= \sum_{i=1}^{N} \frac{\sigma_{v,i}^2 |a_i|^2}{d_i^{2\alpha}} + \sigma_n^2 \,, \\
\sigma_{e,0}^2 &= \lim_{M\to\infty} \mathrm{Var}\left\{ \frac{1}{M}\mathbf{y}^H\mathbf{y} \middle| \mathcal{H}_0 \right\} \\
&= \lim_{M\to\infty} \frac{1}{M^2}\mathrm{Var}\{\tilde{\mathbf{y}}^H\mathbf{C}_w\tilde{\mathbf{y}}\} \\
&\overset{(h)}{=} \lim_{M\to\infty} \frac{1}{M^2}\mathrm{tr}(\mathbf{C}_w^2) \\
&= \lim_{M\to\infty} \sum_{i=1}^{N} \left(\frac{\sigma_{v,i}^2|a_i|^2}{d_i^{\alpha}} + \frac{\sigma_n^2}{M} \right)^2 + \frac{(M-N)}{M^2}\sigma_n^4 \,,
\end{aligned}$$

where in (h) we used the results in Lemma 2.

Introducing new variables $x_i = |a_i|^2$, (6.47) is equivalent to

$$\begin{aligned}
\lim_{M\to\infty} D(T) &= \frac{\sigma_\theta^4 \left(\sum_{i=1}^{N} \frac{x_i}{d_i^{\alpha}} \right)^2}{\sum_{i=1}^{N} \left(\frac{\sigma_{v,i}^2 x_i}{d_i^{\alpha}} + \frac{\sigma_n^2}{M} \right)^2 + \frac{M-N}{M^2}\sigma_n^4} \\
&= \frac{\sigma_\theta^4 \mathbf{x}^T\mathbf{d}\mathbf{d}^T\mathbf{x}}{\mathbf{x}^T\mathbf{B}\mathbf{x} + \frac{2\sigma_n^2}{M}\mathbf{b}^T\mathbf{x} + \frac{\sigma_n^4}{M}} \,,
\end{aligned} \tag{6.48}$$

where the variables $\mathbf{x}, \mathbf{d}, \mathbf{B}, \mathbf{b}$ are defined in (6.52)-(6.55). Substituting $x_i = \frac{P_i}{\sqrt{M}}$ into (6.48), we obtain

$$\begin{aligned}
\lim_{M\to\infty} D(T) &= \lim_{M\to\infty} \frac{\sigma_\theta^4 \mathbf{p}^T\mathbf{d}\mathbf{d}^T\mathbf{p}}{\mathbf{p}^T\mathbf{B}\mathbf{p} + \frac{2\sigma_n^2}{\sqrt{M}}\mathbf{b}^T\mathbf{p} + \sigma_n^4} \\
&= \frac{\sigma_\theta^4 \mathbf{p}^T\mathbf{d}\mathbf{d}^T\mathbf{p}}{\mathbf{p}^T\mathbf{B}\mathbf{p} + \sigma_n^4} \,,
\end{aligned} \tag{6.49}$$

where $\mathbf{p} = [P_1 \cdots P_N]$, and we see that $D(T)$ is asymptotically independent of M. We also observe from (6.48) that an asymptotically non-zero deflection requires that $|a_i|^2$ not decrease faster than $\frac{1}{\sqrt{M}}$. ■

Lemma 2. *Given a complex Gaussian random vector* $\mathbf{z} \in \mathbb{C}^{M\times 1}$ *with distribution* $\mathcal{CN}(0, \mathbf{I}_M)$*, and a Hermitian matrix* $\mathbf{A} \in \mathbb{C}^{M\times M}$*, the variable* $\mathbf{z}^H\mathbf{A}\mathbf{z}$ *has a variance* $\mathrm{Var}\{\mathbf{z}^H\mathbf{A}\mathbf{z}\} = \mathrm{tr}(\mathbf{A}^2)$.

Proof: We first rewrite $\mathbf{z}^H\mathbf{A}\mathbf{z}$ as

$$\mathbf{z}^H\mathbf{A}\mathbf{z} = \sum_{i=1}^{N} \frac{\lambda_i(\mathbf{A})}{2}\chi_i^2(2) \ ,$$

where $\lambda_i(\mathbf{A})$ are the non-zero eigenvalues of $\mathbf{A}$ and $\chi_i^2(2)$ are independent chi-squared variables with 2 degrees of freedom, which can be expressed as

$$\chi_i^2(2) = z_{i,1}^2 + z_{i,2}^2 \ ,$$

where the independent variables $z_{i,1}$ and $z_{i,2}$ have normal distribution $\mathcal{N}(0, 1)$. Since $\mathbf{z}^H\mathbf{A}\mathbf{z}$ can be viewed as the sum of N independent variables, the variance of $\mathbf{z}^H\mathbf{A}\mathbf{z}$ is calculated as

$$\begin{aligned}
\mathrm{Var}\{\mathbf{z}^H\mathbf{A}\mathbf{z}\} &= \sum_{i=1}^{N} \frac{\lambda_i^2(\mathbf{A})}{4}\mathrm{Var}\{\chi_i^2(2)\} \\
&= \sum_{i=1}^{N} \frac{\lambda_i^2(\mathbf{A})}{4}\left(\mathrm{Var}\{z_{i,1}^2\} + \mathrm{Var}\{z_{i,2}^2\}\right) \\
&\overset{(u)}{=} \sum_{i=1}^{N} \lambda_i^2(\mathbf{A}) \\
&\overset{(t)}{=} \mathrm{tr}(\mathbf{A}^2) \ ,
\end{aligned}$$

where (u) follows from

$$\mathrm{Var}\{z_{i,k}^2\} = \mathbb{E}\{z_{i,k}^4\} - \left(\mathbb{E}\{z_{i,k}^2\}\right)^2 = 2 \ ,$$

and (t) is due to the fact that $\lambda_i^2(\mathbf{A})$ are the non-zero eigenvalues of the matrix $\mathbf{A^2}$.■

6.4.2 Sensor Gain Optimization

As with the NP detector, the proof of Theorem 6 shows that as $M \to \infty$, only the magnitude $|a_i|$ of the sensor transmission gains influences the deflection. In this section, we address the problem of finding the $|a_i|$ that maximize the deflection under a sum power constraint. The power allocation problem is formulated as

$$\begin{aligned} \max_{|a_i|^2} \quad & D\left(T\right) \\ s.t. \quad & \mathbf{a}^H\mathbf{a} = P \ . \end{aligned} \tag{6.50}$$

According to (6.48), we can rewrite (6.50) as

$$\begin{aligned} \max_{x_i} \quad & \frac{\mathbf{x}^T\mathbf{d}\mathbf{d}^T\mathbf{x}}{\mathbf{x}^T\mathbf{B}\mathbf{x} + \frac{2\sigma_n^2}{M}\mathbf{b}^T\mathbf{x} + \frac{\sigma_n^4}{M}} \\ s.t. \quad & \mathbf{e}^T\mathbf{x} = P \\ & 0 \le x_i \ , \ i = 1, \cdots, N \ , \end{aligned} \tag{6.51}$$

where

$$\mathbf{x} = [|a_1|^2 \ \cdots \ |a_N|^2]^T \tag{6.52}$$

$$\mathbf{d} = \left[\frac{1}{d_1^\alpha} \ \cdots \ \frac{1}{d_N^\alpha}\right]^T \tag{6.53}$$

$$\mathbf{B} = \text{diag}\left\{\frac{\sigma_{v,1}^4}{d_1^{2\alpha}} \ \cdots \ \frac{\sigma_{v,N}^4}{d_N^{2\alpha}}\right\} \tag{6.54}$$

$$\mathbf{b} = \left[\frac{\sigma_{v,1}^2}{d_1^\alpha} \ \cdots \ \frac{\sigma_{v,N}^2}{d_N^\alpha}\right]^T \tag{6.55}$$

$$\mathbf{e} = [1 \ \cdots \ 1]^T \ .$$

We note here that if the modified deflection of [69] is used instead, then the resulting problem is identical to (6.51), except for the definitions of $\mathbf{B}$ and $\mathbf{b}$, which become

$$\mathbf{B}' = \text{diag}\left\{ \frac{\sigma_{v,1}^4 + \sigma_{v,1}^2\sigma_\theta^2}{d_1^{2\alpha}} \ \cdots \ \frac{\sigma_{v,N}^4 + \sigma_{v,N}^2\sigma_\theta^2}{d_N^{2\alpha}} \right\}$$
$$\mathbf{b}' = \left[\frac{\sigma_{v,1}^2 + \sigma_\theta^2}{d_1^{\alpha}} \ \cdots \ \frac{\sigma_{v,N}^2 + \sigma_\theta^2}{d_N^{\alpha}} \right]^T .$$

Thus, the solution to (6.51) described below can be applied directly to the modified deflection as well.

Problem (6.51) is the maximization of the ratio of two quadratic functions under quadratic constraints, which is referred to as a QCRQ problem. In [26], a solution to the QCRQ problem is found by converting it to a SDP via rank relaxation, followed by an eigendecomposition to find a rank-one result. However, in general, the optimality of the rank-one solution to the original problem can not be guaranteed. Consequently, here we take a different approach and find an asymptotically optimal solution by maximizing an upper bound for (6.51) that is tight when $M \to \infty$. In particular, we consider

$$\begin{aligned} \max_{x_i} \quad & \frac{\mathbf{x}^T\mathbf{d}\mathbf{d}^T\mathbf{x}}{\mathbf{x}^T\mathbf{B}\mathbf{x} + \frac{\sigma_n^4}{M}} \\ s.t. \quad & \mathbf{e}^T\mathbf{x} = P \\ & 0 \le x_i \ , \ i = 1, \cdots, N \ . \end{aligned} \tag{6.56}$$

It is easy to verify that (6.56) provides an upper bound for (6.51) and that the bound is asymptotically achieved when $M \to \infty$. Since $M \to \infty$, we could eliminate the second term in the denominator of (6.56) as well, but we will see in the simulations that it is advantageous to keep it, especially in situations where P is small. The simplification that arises when this term is dropped will be discussed later, when

asymptotic solutions for large P are investigated. In the following, we will show that (6.56) can be converted to a quadratically constrained linear program (QCLP) [70] and solved via standard convex optimization methods.

First, we rewrite (6.56) as

$$\max_{x_i} \quad \frac{\mathbf{x}^T\mathbf{d}\mathbf{d}^T\mathbf{x}}{\mathbf{x}^T\tilde{\mathbf{B}}\mathbf{x}} \tag{6.57}$$

$$s.t. \quad \mathbf{e}^T\mathbf{x} = P \tag{6.58}$$

$$0 \leq x_i \ , \ i = 1, \cdots, N \ ,$$

where $\tilde{\mathbf{B}} = \mathbf{B} + \frac{\sigma_n^4}{MP^2}\mathbf{e}\mathbf{e}^T$. Since the objective function in (6.57) is unchanged by a simple scaling of $\mathbf{x}$, we do not need to explicitly consider the constraint in (6.58) in maximizing (6.57), and the optimal solution can be found via the following two steps:

1. Solve

$$\max_{x_i} \quad \frac{\mathbf{x}^T\mathbf{d}\mathbf{d}^T\mathbf{x}}{\mathbf{x}^T\tilde{\mathbf{B}}\mathbf{x}} \tag{6.59}$$

$$s.t. \quad 0 \leq x_i \ , \ i = 1, \cdots, N \ .$$

2. Denote the result of (6.59) as $\tilde{\mathbf{x}}^*$, then the optimal solution to (6.57) is given by

$$\mathbf{x}^* = \frac{1}{\mathbf{e}^T\tilde{\mathbf{x}}^*}\tilde{\mathbf{x}}^* \ . \tag{6.60}$$

To solve problem (6.59), we first rewrite it in the equivalent form

$$\max_{x_i} \quad \mathbf{x}^T\mathbf{d}\mathbf{d}^T\mathbf{x} \tag{6.61}$$

$$s.t. \quad \mathbf{x}^T\tilde{\mathbf{B}}\mathbf{x} = 1 \tag{6.62}$$

$$0 \leq x_i \ , \ i = 1, \cdots, N \ .$$

To convert (6.61) to a QCLP, we make the following two observations: (1) since the elements of $\mathbf{x}$ and $\mathbf{d}$ are non-negative, maximizing $\mathbf{x}^T\mathbf{d}\mathbf{d}^T\mathbf{x}$ is equivalent to maximizing $\mathbf{x}^T\mathbf{d}$, and (2) we can relax the equality constraint in (6.62) to an inequality $\mathbf{x}^T\tilde{\mathbf{B}}\mathbf{x} \leq 1$, since we can always increase the objective function in (6.61) by scaling $\mathbf{x}$ up to meet the constraint with equality. Thus, solving (6.59) is equivalent to solving the QCLP

$$\begin{aligned} \min_{x_i} \quad & -\mathbf{x}^T\mathbf{d} \\ s.t. \quad & \mathbf{x}^T\tilde{\mathbf{B}}\mathbf{x} \leq 1 \\ & 0 \leq x_i \ , \ i = 1, \cdots, N \ , \end{aligned} \tag{6.63}$$

for which straightforward convex optimization methods exist. The final result for the original problem in (6.56) is found by scaling the optimal solution to (6.63) according to (6.60) to satisfy the power constraint.

Our simulation results in Section 6.5 validate the use of the deflection to optimize detection performance. In particular, we will see that performance improves as the deflection is increased and that with the a_i chosen to maximize the deflection, detection performance remains asymptotically constant as $M \to \infty$ if the power constraint P is scaled by $1/\sqrt{M}$.

6.4.3 Single-Antenna FC

For comparison purposes, we derive the deflection for the case of a single-antenna FC. Based on the signal model in equation (6.20), the single-antenna deflection is given

by

$$D(T_s) = \frac{\left(\mathbb{E}\{T_s|\mathcal{H}_1\} - \mathbb{E}\{T_s|\mathcal{H}_0\}\right)^2}{\mathrm{Var}\{T_s|\mathcal{H}_0\}} = \left(\frac{\sigma_\theta^2 \mathbf{a}^H \mathbf{h}\mathbf{h}^H \mathbf{a}}{\mathbf{a}^H \mathbf{F}\mathbf{V}\mathbf{F}^H \mathbf{a} + \sigma_n^2}\right)^2 , \tag{6.64}$$

where $T_s = |y|^2$ and y, $\mathbf{a}$, $\mathbf{h}$ and $\mathbf{F}$ are as defined in equation (6.20). Unlike the deflection in (6.51) when $M \to \infty$, it is easy to verify that $D(T_s)$ in (6.64) decreases monotonically as the norm of the transmission gain $\mathbf{a}$ decreases. If channel state information is available at the FC, then the optimal gains that maximize $D(T_s)$ are given by (6.25). A different approach is required in the single-antenna case without CSI; for example, in the simulations later we assume the sensor nodes transmit with equal power. We will also observe in the simulation results that when the sum transmission power decreases, the probability of detection for the single-antenna FC will decrease accordingly, while the performance of the multi-antenna FC remains constant as long as the number of antennas increases proportionally to the square of the power decrease.

6.4.4 Asymptotic Closed-form Solutions

While convergence to a globally optimal solution is guaranteed for the QCLP problem described above, we show here that direct closed-form solutions can be found for low and high SNR scenarios $P \gg \sigma_n^2$ and $P \ll \sigma_n^2$. When $P \gg \sigma_n^2$, the size of $\mathbf{x}^T\mathbf{B}\mathbf{x}$ in the denominator of the objective function (6.51) will dominate the terms involving M, which are already small for large M. Thus, for $P \gg \sigma_n^2$, another upper bound for (6.51) is given by

$$\frac{\mathbf{x}^T\mathbf{d}\mathbf{d}^T\mathbf{x}}{\mathbf{x}^T\mathbf{B}\mathbf{x} + 2\frac{\sigma_n^2}{M}\mathbf{b}^T\mathbf{x} + \frac{\sigma_n^4}{M}} < \frac{\sigma_\theta^4 \mathbf{x}^T\mathbf{d}\mathbf{d}^T\mathbf{x}}{\mathbf{x}^T\mathbf{B}\mathbf{x}} .$$

We can formulate the problem of maximizing this upper bound as

$$\begin{aligned} \max_{x_i} \quad & \frac{\mathbf{x}^T \mathbf{d}\mathbf{d}^T \mathbf{x}}{\mathbf{x}^T \mathbf{B} \mathbf{x}} \\ s.t. \quad & \mathbf{e}^T \mathbf{x} = P \\ & 0 \leq x_i \ , \ i = 1, \cdots, N \ , \end{aligned}$$

which has a closed-form solution since $\mathbf{B}$ and $\mathbf{d}$ have non-negative elements:

$$\mathbf{x}^* = \frac{P}{\mathbf{e}^T \mathbf{B}^{-1} \mathbf{d}} \mathbf{B}^{-1} \mathbf{d} \ ,$$

and the corresponding a_i are

$$|a_i| = \sqrt{\frac{P}{\sum_{i=1}^{N} \frac{d_i^{\alpha}}{\sigma_{v,i}^4}}} \frac{d_i^{\frac{\alpha}{2}}}{\sigma_{v,i}^2} \ . \tag{6.65}$$

Thus, for high SNR, after normalizing for distance, the sensors with the lowest measurement noise are allocated higher power.

When $P \ll \sigma_n^2$, the terms involving $\mathbf{x}$ in the denominator of (6.51) will decrease faster than $1/M$, and thus the term $\frac{\sigma_n^2}{M}$ will eventually dominate. This leads to the simpler optimization problem

$$\begin{aligned} \max_{x_i} \quad & \mathbf{x}^T \mathbf{d}\mathbf{d}^T \mathbf{x} \\ s.t. \quad & \mathbf{e}^T \mathbf{x} = P \\ & 0 \leq x_i \ , \ i = 1, \cdots, N \ . \end{aligned} \tag{6.66}$$

This is equivalent to maximizing the weighted sum $\mathbf{x}^T \mathbf{d}$ with constraint $\mathbf{e}^T \mathbf{x} = P$, and the optimal solution is to simply allocate all of the power to the sensor that is

closest to the FC:

$$|a_i| = \begin{cases} \sqrt{P} & i = \arg\min_i d_i \\ 0 & \text{otherwise} . \end{cases} \tag{6.67}$$

Later in the simulation results, we will show that the solutions in (6.65) and (6.67) provide good approximations to the optimal solution of problem (6.63) for very large and very small values of the available sum power P, respectively.

6.4.5 Detection Threshold Calculation

Once the transmission gains a_i of the sensor nodes are optimized, we need to find the threshold $\hat{\gamma}$ to achieve the desired PFA. In the following, we will show that asymptotically as $M \to \infty$, the value of $\hat{\gamma}$ can be calculated according to (6.45) without requiring CSI. Under $\mathcal{H}_0$, the eigenvalues of $\mathbf{C}_w$ are given by

$$\lim_{M\to\infty} \lambda_i\{\mathbf{C}_w\} = \begin{cases} M\eta_i + \sigma_n^2 & 1 \le i \le N \\ \sigma_n^2 & N < i \le M , \end{cases} \tag{6.68}$$

where $\eta_i = \frac{|a_i|^2\sigma_{v,i}^2}{d_i^\alpha}$. Substituting (6.68) into (6.45), we have

$$\lim_{M\to\infty} T = \sum_{i=1}^{N} \frac{1}{2}\left(\eta_i + \frac{\sigma_n^2}{M}\right)\chi_i^2(2) + \frac{\sigma_n^2}{2M}\sum_{i=N+1}^{M}\chi_i^2(2) . \tag{6.69}$$

According to the Strong Law of Large Numbers,

$$\lim_{M\to\infty} \frac{\sigma_n^2}{2M}\sum_{i=N+1}^{M}\chi_i^2(2) = \frac{M-N}{M}\sigma_n^2 ,$$

and this equation holds almost surely. Then the right hand side of equation (6.69) can be viewed as the sum of weighted chi-square variables plus a constant, and for a

specific $\hat{\gamma}$, the PFA is calculated as

$$\begin{aligned} P_{FA} =& \Pr\left(\lim_{M\to\infty} T > \hat{\gamma}|\mathcal{H}_0\right) \\ =& \Pr\left(\sum_{i=1}^{N} \frac{1}{2}\left(\eta_i + \frac{\sigma_n^2}{M}\right)\chi_i^2(2) > \hat{\gamma} - \frac{M-N}{M}\sigma_n^2\right) \\ \overset{(k)}{=}& \sum_{i=1}^{N} \frac{\left(\eta_i + \frac{\sigma_n^2}{M}\right)^{N-1}}{\prod_{l\neq i}(\eta_i - \eta_l)} e^{-\frac{1}{\eta_i + \frac{\sigma_n^2}{M}}\left(\hat{\gamma} - \frac{M-N}{M}\sigma_n^2\right)}, \end{aligned} \tag{6.70}$$

where in (k) we used a result from [30], and we assume that the values of η_i are distinct. In the limit the PFA expression is independent of the CSI, and the value of the threshold $\hat{\gamma}$ that achieves the desired PFA can be found numerically using (6.70).

6.5 Simulation Results

In the simulation examples that follow, we assume $\sigma_\theta^2 = 1, \sigma_n^2 = 0.3, \alpha = 2$ and $N = 10$ sensor nodes. The distances d_i were uniformly distributed over $[2, 10]$, and the measurement noise powers $\sigma_{v,i}^2$ were uniformly distributed in the interval $[0.25, 0.5]$. Once generated, d_i and $\sigma_{v,i}^2$ were held fixed for all simulations. Each point in the following plots is the result of averaging over 10000 trials for each of 300 scenarios; each trial involved a new random parameter θ, as well as new noise realizations and each scenario has a new channel. Plots showing probability of detection were computed assuming a false alarm probability of $\epsilon = 0.05$. For the energy detector, both the deflection and modified deflection gave essentially the same performance, so only the results for the deflection are included.

Figs. 6.1 and 6.2 show the NP detection and LMMSE estimation performance for a single-antenna FC and a multi-antenna FC with $M = 50$ as the available power P ranges from 0.1 to 400. As predicted, as P grows, the performance benefit of

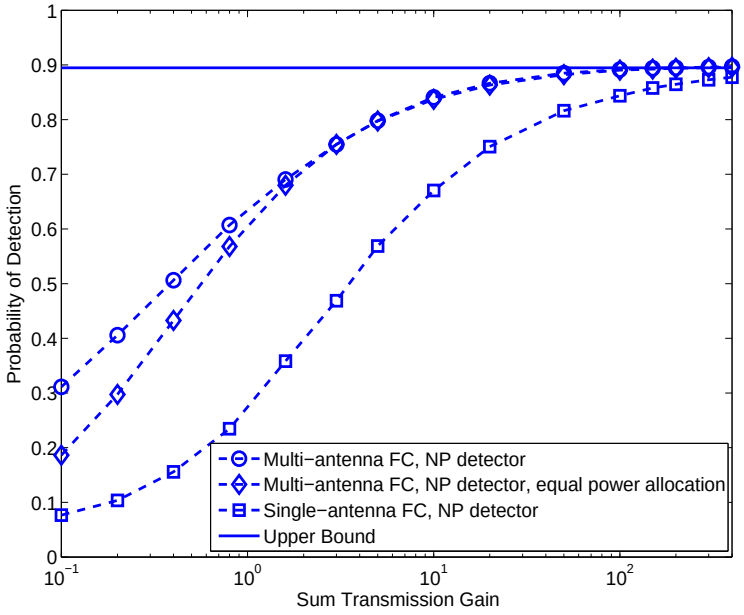

Figure 6.1: Probability of detection for NP detector vs. the value of P, with antenna number $M = 50$. In the plot, for NP detector with Multi- or single-antenna FC, the optimal transmission gains are defined in equations (6.19) and (6.25) respectively, and the upper bound is calculated using equation (6.29).

having multiple antennas at the FC is eventually lost, with both curves in Fig. 6.1 approaching the upper bound in (6.29) and both curves in Fig. 6.2 approaching the lower bound in (6.43). However, in both cases the bound is reached with a much smaller value of P in the multi-antenna case. Note also that for the multi-antenna FC, use of the optimal sensor transmit gains can achieve significantly better performance than equal power allocation when the sum transmit power is low.

Figs. 6.3 and 6.4 respectively present the detection and estimation performance of single- and multi-antenna FCs for increasing M, with the sum power decreasing as $O(1/M)$ according to the formula $P = \sum_{i=1}^{N} \frac{\sigma_n^2 d_i^\alpha}{2\sigma_{v,i}^2 M}$. The energy efficiency of the multi-antenna NP detector is evident, as the MSE and P_D are unchanged as M increases and P decreases; however, the performance of the multi-antenna ED detector degrades with M as the sum power is decreasing at a rate faster than $1/\sqrt{M}$. The lower bound

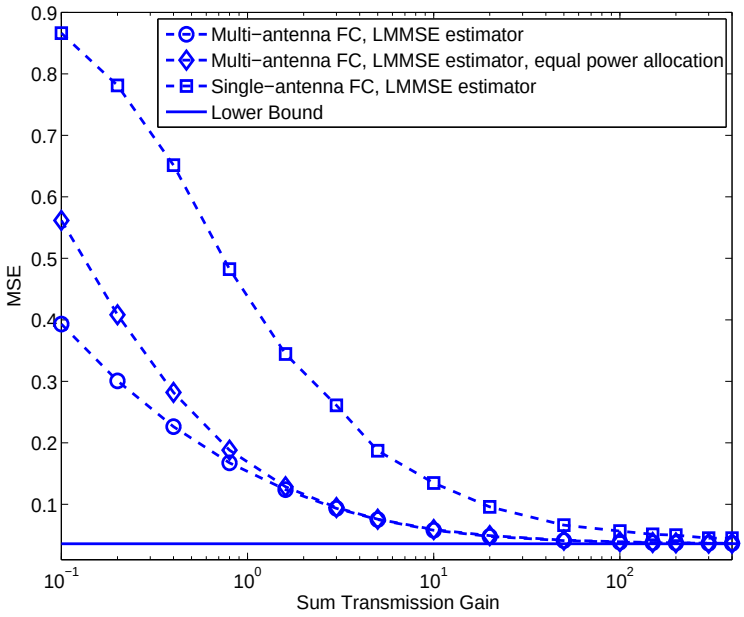

Figure 6.2: Mean-squared error vs. the value of P, with number of antennas $M = 50$. In the plot, for LMMSE estimator with Multi- or single-antenna FC, the optimal transmission gains are defined in equations (6.19) and (6.25) respectively, and the lower bound is given by equation (6.43).

in (6.27) and the upper bound in (6.41) provide tight estimates of the multi-antenna NP probability of detection and LMMSE estimation error, respectively. The value of choosing the optimal sensor gains is evident in comparing the two detection curves for the single-antenna FC, which show a large gap in performance between that achieved with the optimal gains and simply assigning equal gains to all sensors. The latter approach provides a P_D^s that is barely greater than P_{FA}^s, while the optimal sensor gains have much better performance, although P_D^s is decreasing due to the reduction in power. The single-antenna upper bound in (6.28) grows tight as M increases, and is decreasing towards the lower bound ϵ, albeit very slowly.

Fig. 6.5 illustrates the detection performance of the ED approach with P varying from 0.1 to 400. The optimal QCLP approach is plotted along with the low and high SNR approximations in (6.65) and (6.67), the ED implemented with equal power allocation

to all sensors, and the single-antenna FC. The low SNR approximation matches the QCLP approach for $P \leq 1$, while the high SNR solution is optimal for $P \geq 20$; in between these values, the QCLP algorithm provides significantly better performance, although the equal power allocation is close for some values of P. Unlike the NP detector, the single- and multiple-antenna ED solutions do not converge to the same performance for large P; we see in this example that there is a large performance benefit in having a multi-antenna FC, even for large P. In Fig. 6.6, we compare NP and energy detection performance as a function of M assuming that $P = 15/\sqrt{M}$. Consistent with our analytical predictions, the ED with sensor gains chosen via the QCLP to maximize deflection has constant P_D, while the multi-antenna NP detector slowly improves and the single-antenna FC solutions degrade as M increases.

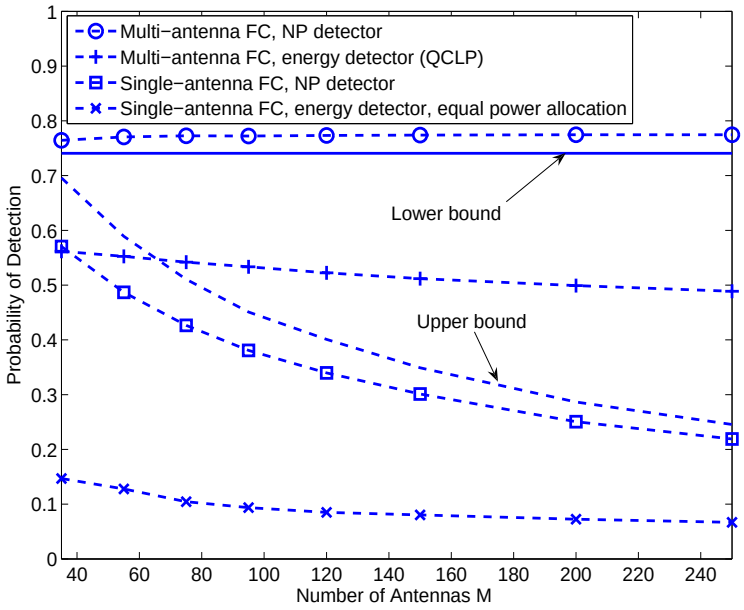

Figure 6.3: Probability of detection vs. number of antennas M. In the above results, the lower and upper bounds are calculated using equations (6.27) and (6.28) respectively.

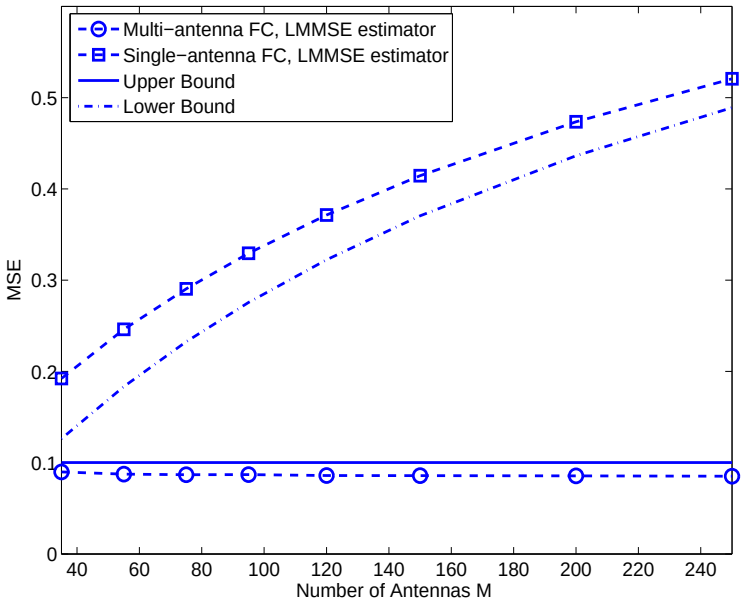

Figure 6.4: Mean-squared error vs. number of antennas M. In the plot, the upper and lower bounds are given by equations (6.41) and (6.42).

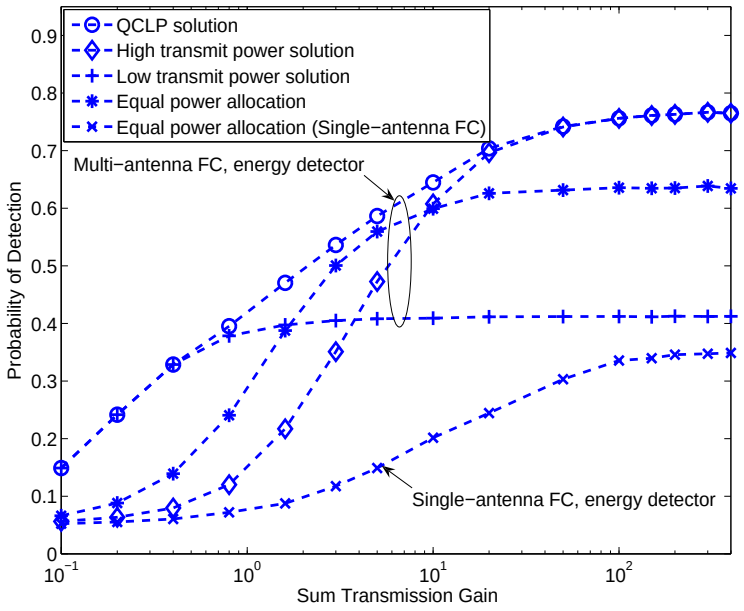

Figure 6.5: Probability of detection for energy detector vs. the value of P, with number of antennas $M = 50$. The high and lower transmit power solutions are given by equations (6.65) and (6.67).

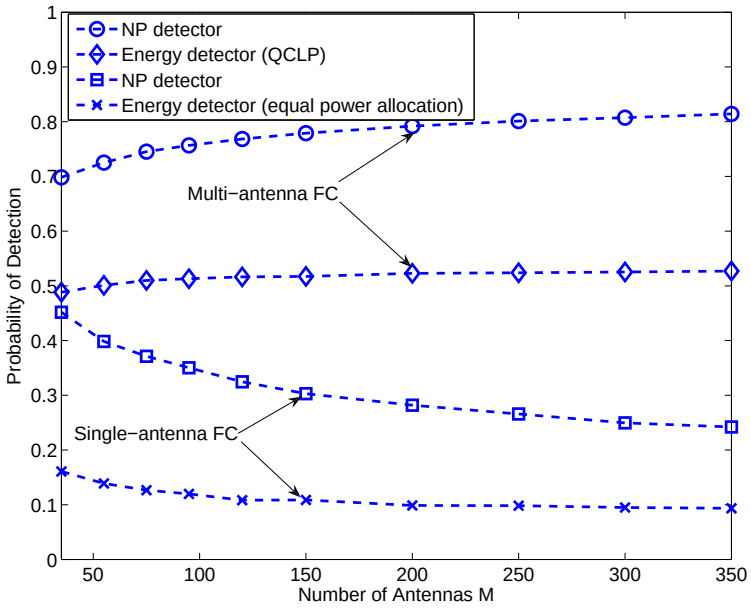

Figure 6.6: Probability of detection vs. number of antennas M.

6.6 Summary

We have studied the detection and estimation performance of a sensor network communicating over a coherent multiple access channel with a fusion center possessing a large number of antennas. We studied Neyman-Pearson and energy detection, derived optimal sensor transmission gains for each case, and showed that the optimal gains are phase-independent as the number of antennas grows large. Similar to properties of massive MIMO wireless communications, one can trade antennas at the fusion center for energy efficiency at the sensors. For the case of Neyman-Pearson detection and LMMSE estimation, which require channel state information, constant levels of performance can be achieved if the transmit power at the sensors is reduced proportional to the gain in the number of antennas. For energy detection, which does not require channel state information, a constant deflection coefficient can be maintained if power is reduced proportional to the inverse square root of the number of antennas. While bounds derived for Neyman-Pearson detection and LMMSE estimation show

performance gains for a multiple-antenna fusion center in low sensor transmit power scenarios, the benefit is shown to disappear when the transmit power is high. However, for the energy detector, having multiple antennas at the fusion center provides a significant advantage even when the sensors have high power.

Chapter 7

Connectivity Optimization for Mobile Network with UAV Relay

7.1 Introduction

There is increasing interest in the use of relatively small, flexible UAVs that fly at lower altitudes to provide relay services for mobile *ad hoc* networks (i.e., networks without a centralized basestation or other infrastructure) composed of ground-based communication nodes [71–78]. A number of different approaches have been proposed in the literature to address the performance of UAV-assisted communication networks. For example, in [71], a throughput maximization protocol for non-real time applications was proposed for a network with UAV relays in which the UAV first loads data from the source node and then flies to the destination node to deliver it. The authors in [72] investigated different metrics for ad hoc network connectivity and propose several approaches for improving the connectivity through deployment of a UAV. In [73], the authors considered a scenario in which multiple UAVs are deployed to relay

data from isolated ground sensors to a base station, and an algorithm was proposed to maintain the connectivity of the links between the sensors and base station.

The work described above assumes that the ground nodes are static and that the UAV is configured with only a single antenna. Given the well-known benefits of employing multiple antennas for communications, it is natural to consider the advantages they offer for UAV-based platforms[79]. The measurement results of [80] showed that using multiple receivers at the UAV can significantly increase the packet delivery rate of the ground-to-air link. A swarm of single antenna UAVs were used as a virtual antenna array to relay data from a fixed ad hoc network on the ground in [74], and the performance of distributed orthogonal space-time block codes (OSTBC) and beamforming were evaluated. A relay system with multi-antenna UAVs and multi-antenna mobile ground terminals was investigated in [75]. The users employ OSTBC to transmit data and the data transmissions are assumed to be interference free. Based on estimates of the user terminals' future position, a headingf optimization approach was proposed that maximizes the uplink sum rate of the network (the sum of the theoretically achievable throughputs of all users) under the constraint that each user's rate is above a given threshold. The restriction of [75] to the interference-free case is a significant drawback, which we address in this chapter. An earlier version of our work [81] discussed the use of an antenna array to improve the throughput of the ground-to-air uplink when the users share the same channel and interfere with one another.

In this chapter, we consider a model similar to [75], with several ground-based users communicating simultaneously with a multi-antenna UAV. The main difference with [75] is that we assume there exists co-channel interference between the different users' data streams. The users are assumed to transmit data with a single antenna and the UAV uses beamforming to separate the co-channel data streams. We assume a corre-

lated Rician fading channel model between each ground node and the UAV, where the channel is represented as the sum of a deterministic LOS component and a correlated Gaussian term to represent Rayleigh fading due to multipath. We then quantify the uplink performance of the relay network by deriving an approximation to the ergodic achievable rate (the achievable throughput of the users averaged over the distribution of the channels), assuming that the UAV uses a maximum SINR beamformer for interference mitigation. The strength of the mutual interference depends on the correlation between the users' channel vectors, which in a channel with a strong LOS component is a function of the signals' AoA. The AoAs depend in turn on the UAV's heading and the relative positions of the UAV and the ground nodes. Consequently, we propose an adaptive algorithm for adjusting the heading of the UAV to minimize the users' mutual interference and improve the uplink communications performance. In particular, the UAV is assumed to fly with a constant velocity, and it adjusts its heading in discrete time steps (assuming a constraint on the maximum turning radius) in order to optimize the approximate achievable rate. At time step n, the UAV uses a Kalman filter driven by feedback from the ground terminals to predict their positions at time $n+1$, and then the UAV computes its heading in order to optimize the approximate sum rate based on these future position estimates.

The main results of this chapter are summarized as follows:

1. We analyze the trajectory optimization problem for a special case involving two static ground nodes. We use a rectangular-path model to characterize the UAV's trajectory, which reduces the problem to one of optimizing only the heading. This problem can be solved using a simple line search, and the results indicate how increasing the size of the UAV array can reduce the system's sensitivity to the heading direction.

2. For the case of a general network of mobile ground-based nodes, we derive an

approximation to the average achievable sum rate to measure the system performance. Based on this approximation, we formulate a heading optimization problem and propose a line-search algorithm to adjust the UAV's heading direction at time step n such that the system performance at time step $n+1$ is optimized. We study the performance of both time-division multiple access (TDMA, where each user accesses the channel at different times) and space-division multiple access (SDMA, where all users access the channel at the same time, but are separated based on the spatial component of their signals, such as AoA), and illustrate via simulation the dramatic improvement offered by SDMA.

3. We derive asymptotic analytical results for the heading optimization problem under the assumption of a Rician channel with a strong LOS component between the ground nodes and UAV. The asymptotic results provide simplified methods for solving the heading optimization problem. A separate approximation method is used for low and high SNR cases, and we show that using the asymptotic expressions for heading optimization results in performance nearly identical to that of the optimal algorithm.

The organization of this chapter is as follows. We present our assumed signal and channel model in Section 7.2, and in Section 7.3 we focus on the UAV heading optimization problem for the special case of two static ground users. In Section 7.4, we first describe the mobility model for the UAV and ground nodes, as well as a standard Kalman filter for predicting the future positions of the ground nodes. Then we formulate the UAV heading optimization problem and propose an adaptive heading adjustment algorithm. We then derive asymptotic expressions for the general heading optimization problem in Section 7.5, assuming a Rician channel with a strong LOS component between the UAV and ground nodes. Simulation results are pro-

vided in Section 7.6 to illustrate the performance of the heading control algorithm, the advantage of SDMA over TDMA, and the validity of the asymptotic results.

7.2 System Model

According to equation (2.3), when the channels from the ground user nodes $\mathbf{h}_{i,n} \in \mathbb{C}^{M\times 1}, i = 1, \ldots, N$ are known to the UAV (e.g., via training data from the ground nodes), the vector $\mathbf{w}_{i,n}$ that maximizes the $SINR_{i,n}$ is given by [82]

$$\mathbf{w}_{i,n} = \mathbf{Q}_{i,n}^{-1}\mathbf{h}_{i,n} \ ,$$

where $\mathbf{Q}_{i,n} = \sum_{j=1,j\neq i}^{N} P_t \mathbf{h}_{j,n}\mathbf{h}_{j,n}^H + \sigma^2\mathbf{I}_M$. The corresponding $SINR_{i,n}$ can be calculated as

$$SINR_{i,n} = P_t \mathbf{h}_{i,n}^H \mathbf{Q}_{i,n}^{-1}\mathbf{h}_{i,n}. \tag{7.1}$$

For the channel between the ground nodes and UAV, we assume a correlated Rician fading channel with consideration of large-scale path loss:

$$\mathbf{h}_{i,n} = \frac{\mathbf{h}_{i,n}^{'}}{d_{i,n}^{\alpha}},$$

where $\mathbf{h}_{i,n}^{'}$ is the normalized channel vector, $d_{i,n}$ is the distance between node i and the UAV during the nth time step, and α is the path loss exponent. Define the three dimensional coordinates of the UAV and node i as $(x_{u,n}, y_{u,n}, h_u)$ and $(x_{i,n}, y_{i,n}, 0)$, so that $d_{i,n}$ is given by

$$d_{i,n} = \sqrt{(x_{u,n} - x_{i,n})^2 + (y_{u,n} - y_{i,n})^2 + h_u^2}.$$

For node i, we write the Rician fading channel vector $\mathbf{h}_{i,n}^{'}$ with two components [83], a LOS component $\bar{\mathbf{h}}_{i,n}$ and a Rayleigh fading component $\tilde{\mathbf{h}}_{i,n}$:

$$\mathbf{h}_{i,n}^{'} = \bar{\mathbf{h}}_{i,n} + \tilde{\mathbf{h}}_{i,n}.$$

The LOS response will depend on the AoA of the signal, which in turn depends on orientation of the array (and hence the heading of the UAV) and the positions of the UAV and user nodes. For example, assume a uniform linear array (ULA) with antennas separated by one-half wavelength, and that at time step n the phase delay between adjacent antenna elements for the signal from the i-th node is $p_{i,n}$, then the LOS component could be modeled as

$$\bar{\mathbf{h}}_{i,n} = \beta(\phi_{i,n})\sqrt{\frac{K}{1+K}}\left[1, e^{jp_{i,n}} \cdots e^{j(M-1)p_{i,n}}\right]^T, \tag{7.2}$$

where K is the Rician K-factor and $\beta(\phi_{i,n})$ is used to account for variations in the antenna gain as a function of the elevation angle $\phi_{i,n}$ to the i-th node. The phase delay $p_{i,n}$ is calculated by [84, chap. 4]

$$p_{i,n} = \pi \cos(\phi_{i,n}) \sin(\theta_{i,n}),$$

where $\theta_{i,n}$ represents the azimuth angle to the i-th ground node. In terms of the UAV and user node positions, these quantities can be calculated as

$$\begin{aligned} \cos(\phi_{i,n}) &= \sqrt{\frac{(x_{u,n}-x_{i,n})^2 + (y_{u,n}-y_{i,n})^2}{(x_{u,n}-x_{i,n})^2 + (y_{u,n}-y_{i,n})^2 + h_u^2}}, \\ \sin(\theta_{i,n}) &= \cos(\delta_n - \epsilon_{i,n}), \end{aligned} \tag{7.3}$$

where δ_n is the heading angle of the UAV, $\delta_n - \epsilon_{i,n}$ denotes the angle between the

UAV heading and the LOS to user i, and

$$\begin{aligned}\epsilon_{i,n} &= \begin{cases} \zeta_{i,n}, & y_{i,n} - y_{u,n} \geq 0 \text{ and } x_{i,n} - x_{u,n} \geq 0, \\ \zeta_{i,n} + \pi, & x_{i,n} - x_{u,n} \leq 0, \\ \zeta_{i,n} + 2\pi, & \text{otherwise.} \end{cases} \\ \zeta_{i,n} &= \arctan\left(\frac{y_{i,n} - y_{u,n}}{x_{i,n} - x_{u,n}}\right) .\end{aligned}$$

Since there is little multipath scattering near the UAV, any Rayleigh fading components will experience high spatial correlation at the receive end of the link. Thus, we model the spatially correlated Rayleigh component as

$$\tilde{\mathbf{h}}_{i,n} = \beta(\phi_{i,n})\sqrt{\frac{1}{1+K}}(\mathbf{R}_r)^{\frac{1}{2}}\mathbf{g}_{i,n},$$

where $\mathbf{g}_{i,n} \in \mathbb{C}^{M\times 1}$ has i.i.d. zero-mean, unit-variance complex Gaussian entries (which we denote by $\mathcal{CN}(0,1)$), and $\mathbf{R}_r$ is the spatial correlation matrix of the channel on the receiver side of the link. In [85], a model for $\mathbf{R}_r$ is proposed under the assumption that the multipah rays are distributed normally in two dimensions around the angle from the source with standard deviation σ_r, assuming a ULA receiver. We can easily extend this model to take into account the third dimension corresponding to the elevation angle, and the resulting $\mathbf{R}_r$ is given by

$$\mathbf{R}_r = \left(1 + \frac{1}{K}\right)\bar{\mathbf{h}}_{i,n}\bar{\mathbf{h}}_{i,n}^H \odot \mathbf{B}(\theta_{i,n}, \sigma_r), \tag{7.4}$$

where $\mathbf{B}(\theta_{i,n}, \sigma_\phi)$ is calculated as

$$\mathbf{B}(\theta_{i,n}, \sigma_\phi)_{k,l} = e^{-\frac{1}{4}(\pi(k-l))^2\sigma_r^2\cos^2(\theta_{i,n})\left(1+\cos(2\phi_{i,n})-\frac{1}{2}\sigma_r^4\sin^2(2\phi_{i,n})(\pi(k-l))^2\cos^2(\theta_{i,n})\right)}. \tag{7.5}$$

The resulting distribution for $\mathbf{h}_{i,n}^{'}$ is thus

$$\mathbf{h}_{i,n}^{'} \sim \mathcal{CN}\left(\bar{\mathbf{h}}_{i,n}, \frac{1}{K+1}\mathbf{R}_r\right). \tag{7.6}$$

The goal of this chapter is to derive an algorithm for adjusting the heading angle δ_n of the UAV in order to optimize the achievable uplink throughput of the network (defined in the next section). For simplicity we consider only UAV heading adjustments, but the same type of approach could be used if UAV speed and altitude were assumed to be adaptive as well. We assume a UAV equipped with a ULA oriented along either the fuselage or wings, the only difference being a 90° change in how we define the heading angle. Extensions of the algorithm and analysis to different array geometries would require one to use a different expression for (7.2) and to derive a different spatial correlation matrix $\mathbf{R}_r$. We will consider both SDMA and TDMA approaches in the following sections. In practice, SDMA would not be used as the only method of providing wireless access to all users on the ground, since the number of antennas is limited and the presence of a (near-)LOS channel would make it difficult to separate users on the ground that are close together. As in the design of terrestrial cellular basestations, SDMA would be a tool to augment the capacity of the network beyond what TDMA and FDMA schemes already provide. The approach described below can be thought of as solving the SDMA problem only for those users that have been scheduled for the same time/frequency slot. Finally, we note that in practice, considerations other than communications performance would likely need to be considered in choosing the heading of the UAV, and these would need to be included as additional constraints to the optimization presented below.

7.3 Results for the Static Two-User Case

To demonstrate the significant impact of the UAV trajectory on the performance of the ground-to-air uplink, we first consider a simple two user scenario. The gross behavior of the UAV would be governed by the distance D between the two users, with three possibilities:

1. $D \gg h_u$ - This is not a particularly useful scenario for a simultaneous uplink from both users since, if the UAV flies near their midpoint, both users would experience low SINR at the UAV due to path loss, and the sum data rate would be quite low. In this case, a better approach would likely involve the UAV serving each ground node separately, circling directly above each user and alternately flying between them.

2. $D \ll h_u$ - This case is also less interesting since the UAV should obviously fly directly above the two users in as tight a pattern as possible to minimize path loss. The effect of the UAV heading would be minimal, since the AoAs to the two users would be nearly identical. If the K-factor of the channel was high (as one would expect when the UAV is essentially directly overhead), then the channels would be highly correlated and a TDMA solution would likely be preferred over SDMA.

3. $D = O(h_u)$ - Since the users transmit with the same power and their channels have the same statistical properties, equalizing the average uplink rates for the two users would require the UAV to fly a symmetric trajectory centered around the midpoint of the two users. If it was desired to minimize the variation in each user's average uplink rate, the bounds of this trajectory would be small relative to the distance to the users. This is the case we consider in this section.

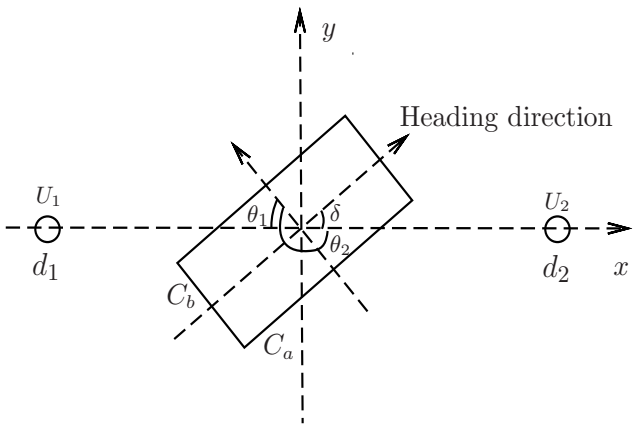

Figure 7.1: Simplified UAV trajectory for the two-user case. C_a and C_b represent the edges of the rectangular trajectory. The angles θ_1 and θ_2 denote the azimuth angles of arrival of the users' signals at the ULA when the UAV is flying over the midpoint of the two users with heading direction δ.

To make the analysis tractable, we focus on a rectangular trajectory as depicted in Fig. 7.1, defined by the side lengths C_a and C_b and the orientation δ. The angle δ is defined to be with respect to the side of the rectangle with greater length. Given the assumptions for scenario (3) above, the side lengths are assumed to satisfy $\max\{C_a, C_b\} \leq C_{\max} \ll d_i$, so the figure is not to scale. Under this assumption, the performance of a rectangular trajectory is expected to be similar to that for other trajectories with similar size and orientation (e.g., an ellipse or figure-8 pattern). We also assume that $\min\{C_a, C_b\} \geq C_{\min}$, which effectively accounts for the turning radius of the UAV. Since the UAV flies near the midpoint between the two users, we assume that the antenna gain factor due to elevation angle is the same for both users, and for simplicity we set it as $\beta(\phi_{i,n}) = 1$ for $i = 1, 2$.

The sum data rate at the UAV averaged along the trajectory is given by

$$\begin{aligned} \bar{R} &= \mathbb{E}\left\{\log_2(1 + SINR_1) + \log_2(1 + SINR_2)\right\} \\ &= \frac{1}{2(C_a + C_b)} \int_{\mathcal{C}} \Big(\log_2(1 + SINR_1(p)) + \log_2(1 + SINR_2(p)) \Big) dp, \quad (7.7) \end{aligned}$$

where $\mathcal{C}$ denotes the rectangular path followed by the UAV, variable p denotes different positions along the trajectory and dp represents the length of the elementary subintervals along the trajectory. The optimization problem we wish to solve is formulated as

$$\begin{aligned} \max_{\delta, C_a, C_b} \quad & \bar{R} \\ s.t. \quad & 0 \leq \delta \leq \frac{\pi}{2} \\ & C_{\min} \leq C_b \leq C_a \leq C_{\max}, \end{aligned} \tag{7.8}$$

where the symmetry of the problem allows us to restrict attention to $0 \leq \delta \leq \pi/2$ and assume $C_b \leq C_a$ without loss of generality. This non-linear optimization problem is difficult to solve directly. In the following, we will derive an approximate problem to (7.8) for high SNR ($\frac{P_t}{d_i^\alpha \sigma^2} \gg 1$) and assuming channels with a large K-factor. To begin with, we observe that, due to the symmetric trajectory centered at the midpoint between the two ground nodes, the expected data rate averaged over the trajectory will be the same for both users:

$$\int_{\mathcal{C}} \log_2(1 + SINR_1(p))dp = \int_{\mathcal{C}} \log_2(1 + SINR_2(p))dp \; .$$

Thus, we can focus on evaluating the SINR for just one of the users. For large K, we can ignore the Rayleigh component of the channel, and assume that $\mathbf{h}_i^{'} \approx \bar{\mathbf{h}}_i$. We replace the explicit dependence of the channel on n with an implicit dependence on a point p along the trajectory defined in Fig. 7.1. At point p, the SINR for user 1 can be expressed as

$$\begin{aligned} SINR_1 &= \frac{P_t}{d_1^\alpha} \bar{\mathbf{h}}_1^H \left(\sigma^2 \mathbf{I}_M + \frac{P_t}{d_2^\alpha} \bar{\mathbf{h}}_2 \bar{\mathbf{h}}_2^H \right)^{-1} \bar{\mathbf{h}}_1 \\ &= \frac{MP_t}{d_1^\alpha \sigma^2} - \frac{P_t^2}{d_1^\alpha d_2^\alpha \sigma^4} \frac{\left| \bar{\mathbf{h}}_1^H \bar{\mathbf{h}}_2 \right|^2}{1 + \frac{MP_t}{d_2^\alpha \sigma^2}}, \end{aligned} \tag{7.9}$$

where, assuming that $\beta(\phi_1) = \beta(\phi_2) = 1$,

$$\left|\bar{\mathbf{h}}_1^H \bar{\mathbf{h}}_2\right| = \left| \frac{\sin\left(\frac{M\pi}{2}\big(\cos(\phi_1)\sin(\theta_1) - \cos(\phi_2)\sin(\theta_2)\big)\right)}{\sin\left(\frac{\pi}{2}\big(\cos(\phi_1)\sin(\theta_1) - \cos(\phi_2)\sin(\theta_2)\big)\right)} \right|,$$

and $\cos(\phi_i)$ and $\sin(\theta_i)$ are defined in (7.3). Note that in addition to $\bar{\mathbf{h}}_1$, the parameters d_i, ϕ_i and θ_i all implicitly depend on p.

Using Jensen's inequality, the following upper bound for $\bar{R}$ can be found:

$$\bar{R} \leq \log_2(1 + \mathbb{E}\{SINR_1\}) + \log_2(1 + \mathbb{E}\{SINR_2)\}. \tag{7.10}$$

We will proceed assuming that an operating point that maximizes the upper bound will also approximately optimize $\bar{R}$. Based on (7.9) and assuming we have a high SNR scenario where $\frac{P_t}{d_i^\alpha \sigma^2} \gg 1$,

$$\begin{aligned} \mathbb{E}\{SINR_1\} &\overset{(d)}{\approx} \frac{P_t}{\sigma^2}\mathbb{E}\left\{\frac{M}{d_1^\alpha} - \frac{|\mathbf{h}_1^H \mathbf{h}_2|^2}{d_1^\alpha M}\right\} \\ &\overset{(e)}{\approx} \frac{P_t}{d_1^\alpha \sigma^2}\left(M - \frac{\mathbb{E}\{|\mathbf{h}_1^H \mathbf{h}_2|^2\}}{M}\right), \end{aligned} \tag{7.11}$$

where (d) is due to the high SNR assumption and (e) follows from the assumption that $C_{\max} \ll d_1$. The dependence of $SINR_1$ on d_2 is thus eliminated, and in what follows we drop the subscript on d_1 and write it simply as d. Substituting equation (7.11) in (7.10), and replacing the objective function in problem (7.8) with the upper bound of (7.10), our optimization problem is approximately given by

$$\begin{aligned} \max_{\delta, C_a, C_b} \quad & \log_2\left(1 + \frac{MP_t}{d^\alpha \sigma^2} - \frac{P_t \mathbb{E}\{|\mathbf{h}_1^H \mathbf{h}_2|^2\}}{M d^\alpha \sigma^2}\right) \\ s.t. \quad & 0 \leq \delta \leq \frac{\pi}{2} \\ & C_{\min} \leq C_b \leq C_a \leq C_{\max}. \end{aligned} \tag{7.12}$$

Since the objective function in (7.12) is monotonically decreasing with $\mathbb{E}\{|\mathbf{h}_1^H\mathbf{h}_2|^2\}$, an equivalent problem is formulated as

$$\begin{aligned} \min_{\delta, C_a, C_b} \quad & \mathbb{E}\{|\mathbf{h}_1^H\mathbf{h}_2|^2\} \\ s.t. \quad & 0 \leq \delta \leq \frac{\pi}{2} \\ & C_{\min} \leq C_b \leq C_a \leq C_{\max}. \end{aligned} \tag{7.13}$$

The interpretation of (7.13) is that the optimal trajectory minimizes the average correlation between the two users' channels. The calculation of $\mathbb{E}\{|\mathbf{h}_1^H\mathbf{h}_2|^2\}$ includes the integral of the function

$$\frac{\sin^2\Big(\frac{M\pi}{2}\big(\cos(\phi_1)\sin(\theta_1) - \cos(\phi_2)\sin(\theta_2)\big)\Big)}{\sin^2\Big(\frac{\pi}{2}\big(\cos(\phi_1)\sin(\theta_1) - \cos(\phi_2)\sin(\theta_2)\big)\Big)}$$

with respect to p, which is difficult to evaluate. To simplify (7.8), we assume that, compared with the distance to the users on the ground, the UAV moves over a small region, and for purposes of analyzing the mathematics, one can assume that the UAV essentially remains fixed at the midpoint between the two users. Only the heading of the UAV changes the uplink rate in this case. Under this assumption, the elevation angles ϕ_1, ϕ_2 are constant and equal $\phi_1 = \phi_2 = \phi^{'}$, and the azimuth angles θ_1, θ_2 are piecewise constant. When UAV flies along C_a, they are equal to θ_1 and θ_2; when the UAV flies along C_b, they are equal to $\theta_1 + \frac{\pi}{2}$, $\theta_2 + \frac{\pi}{2}$. Note that since $\theta_2 = \theta_1 + \pi$ always holds, then $\sin(\theta_2) = -\sin(\theta_1)$ and we have

$$|\mathbf{h}_1^H\mathbf{h}_2|^2 = \frac{\sin^2(M\pi\cos(\phi^{'})\sin(\theta_1))}{\sin^2(\pi\cos(\phi^{'})\sin(\theta_1))}.$$

Note also that $\theta_1 + \delta = \frac{\pi}{2}$, and hence $\sin(\theta_1) = \cos(\delta)$. Thus

$$|\mathbf{h}_1^H\mathbf{h}_2|^2 = \frac{\sin^2(M\pi\cos(\phi^{'})\cos(\delta))}{\sin^2(\pi\cos(\phi^{'})\cos(\delta))}.$$

Along C_a, the UAV flies with heading δ and along C_b, the UAV flies with heading $\delta + \frac{\pi}{2}$, so that $\cos(\delta + \frac{\pi}{2}) = -\sin(\delta)$. Thus, we have

$$\begin{aligned}\mathbb{E}\{|\mathbf{h}_1^H \mathbf{h}_2|^2\} &= \frac{C_a}{C_a + C_b} \frac{\sin^2(M\pi \cos(\phi_i') \cos(\delta))}{\sin^2(\pi \cos(\phi_i') \cos(\delta))} \\ &\quad + \frac{C_b}{C_a + C_b} \frac{\sin^2(M\pi \cos(\phi_i') \sin(\delta))}{\sin^2(\pi \cos(\phi_i') \sin(\delta))}. \end{aligned} \tag{7.14}$$

Substituting (7.14) into the objective function of problem (7.13) yields

$$\begin{aligned} \min_{\delta, C_a, C_b} \quad & \frac{C_a}{C_a + C_b} \frac{\sin^2(M\pi \cos(\phi_i') \cos(\delta))}{\sin^2(\pi \cos(\phi_i') \cos(\delta))} \\ & \quad + \frac{C_b}{C_a + C_b} \frac{\sin^2(M\pi \cos(\phi_i') \sin(\delta))}{\sin^2(\pi \cos(\phi_i') \sin(\delta))} \\ s.t. \quad & 0 \leq \delta \leq \frac{\pi}{2} \\ & C_{\min} \leq C_b \leq C_a \leq C_{\max}. \end{aligned} \tag{7.15}$$

We now show that problem (7.15) is equivalent to an optimization problem over the single variable δ. First define

$$s_1 = \frac{\sin^2(M\pi \cos(\phi_i') \cos(\delta))}{\sin^2(\pi \cos(\phi_i') \cos(\delta))}, \; s_2 = \frac{\sin^2(M\pi \cos(\phi_i') \sin(\delta))}{\sin^2(\pi \cos(\phi_i') \sin(\delta))},$$

$$R_c = \frac{C_{\max}}{C_{\min}}, \; R = \frac{C_a}{C_b} \, ,$$

so that $1 \leq R \leq R_c$. Then the objective function of (7.15) can be rewritten as

$$\frac{R}{1+R} s_1 + \frac{1}{1+R} s_2 = s_1 + \frac{s_2 - s_1}{1+R}.$$

Given a heading direction $\delta \in [0, \frac{\pi}{2}]$, if $s_2 \geq s_1$, then the objective function is minimized when $R = R_c$. Otherwise, if $s_2 < s_1$, $R = 1$ minimizes the objective function. The domain $[0, \frac{\pi}{2}]$ can be divided into two sets $\mathcal{S}_1$ and $\mathcal{S}_2$, such that for $\delta \in \mathcal{S}_1$ we have $s_2 < s_1$, and for $\delta \in \mathcal{S}_2$ we have $s_2 \geq s_1$. Then problem (7.15) can be divided

into two subproblems

$$\begin{aligned} \min_{\delta} \quad & \frac{R_c}{1+R_c} s_1 + \frac{1}{1+R_c} s_2 \\ s.t. \quad & \delta \in \mathcal{S}_2 \ . \end{aligned} \tag{7.16}$$

$$\begin{aligned} \min_{\delta} \quad & \frac{1}{2} s_2 + \frac{1}{2} s_1 \\ s.t. \quad & \delta \in \mathcal{S}_1 \ . \end{aligned} \tag{7.17}$$

Since $s_1(\frac{\pi}{2} - \delta) = s_2(\delta)$, for each $\delta \in \mathcal{S}_2$ and $s_2(\delta) > s_1(\delta)$, we have $\frac{\pi}{2} - \delta \in \mathcal{S}_1$ and vice versa. Thus the following equation holds

$$\frac{R_c}{1+R_c} s_1(\delta) + \frac{1}{1+R_c} s_2(\delta) \leq \frac{1}{2} s_1(\delta) + \frac{1}{2} s_2(\delta),$$

$$\frac{1}{2} s_2\left(\frac{\pi}{2} - \delta\right) + \frac{1}{2} s_1\left(\frac{\pi}{2} - \delta\right) = \frac{1}{2} s_1(\delta) + \frac{1}{2} s_2(\delta).$$

Then the minimum value of (7.16) must be smaller than or equal to the minimum value of (7.17) and problem (7.15) is equivalent to problem (7.16). For each $\delta \in \mathcal{S}_2$, the following equation holds

$$R_c s_1(\delta) + s_2(\delta) \leq R_c s_2(\delta) + s_1(\delta)$$

$$R_c s_1\left(\frac{\pi}{2} - \delta\right) + s_2\left(\frac{\pi}{2} - \delta\right) = R_c s_2(\delta) + s_1(\delta),$$

and problem (7.16) is thus equivalent to

$$\begin{aligned} \min_{\delta} \quad & \frac{R_c}{1+R_c} s_1 + \frac{1}{1+R_c} s_2 \\ s.t. \quad & 0 \leq \delta \leq \frac{\pi}{2} \ . \end{aligned} \tag{7.18}$$

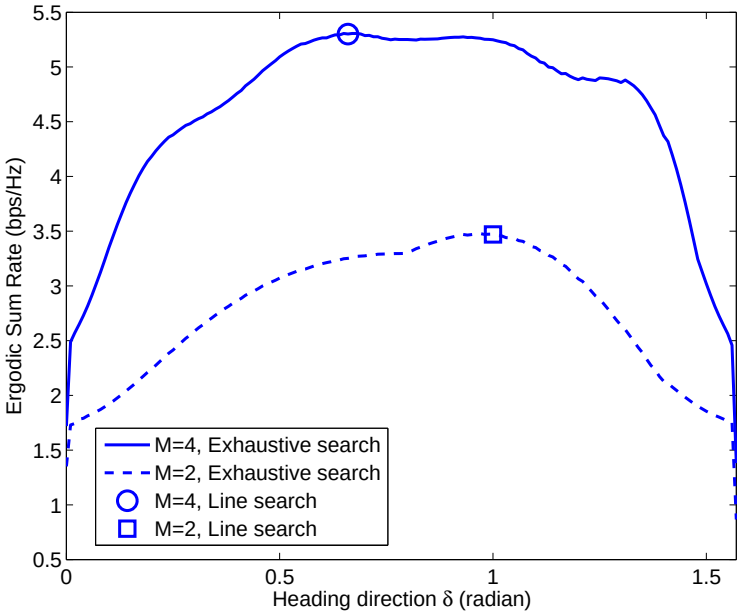

Figure 7.2: Orientation of the rectangular trajectory provided by the line search method in (7.19). For the exhaustive search method, the solid curve and the dashed curve denote the optimal sum rate that can be achieved for different orientations δ. When $M = 4$, the optimal δ are: 0.66 (exhaustive search), 0.69 (line search); when $M = 2$, the optimal δ are: 0.98 (exhaustive search), 1.00 (line search).

Based on (7.18), the solution to (7.8) is approximately given by $C_a = C_{\max}$, $C_b = C_{\min}$ and

$$\delta = \arg \min_{0 \le \delta \le \pi/2} \left\{ \frac{R_c}{1+R_c} \frac{\sin^2(M\pi \cos(\phi') \cos(\delta))}{\sin^2(\pi \cos(\phi') \cos(\delta))} + \frac{1}{1+R_c} \frac{\sin^2(M\pi \cos(\phi') \sin(\delta))}{\sin^2(\pi \cos(\phi') \sin(\delta))} \right\}, \tag{7.19}$$

where $R_c = \frac{C_{\max}}{C_{\min}}$ and ϕ' is the elevation angle to the two users at the center of the rectangle in Fig. 7.1, and satisfies

$$\cos(\phi') = \frac{d_i}{\sqrt{d_i^2 + h_u^2}}.$$

Minimizing (7.19) can be achieved by a simple line search over the interval $[0, \pi/2]$.

To illustrate the validity of the approximate solution, we compare the average system sum rate achieved by maximizing (7.8) using an exhaustive search over $\{C_a, C_b\}$ for each value of δ evaluated in the approximate line search of (7.19). The simulation parameters were $d_1 = d_2 = 1500$m, $h_u = 350$m, $C_{\min} = 200$m, $C_{\max} = 800$m, and $\frac{P_t}{\sigma^2} = 65$dB. The results of the simulation are plotted in Fig. 7.2, which shows the best rate obtained by (7.8) for each value of δ, and the optimal value obtained from minimizing (7.19) for $M = 2$ and $M = 4$. In both cases, the approximate approach of (7.19) finds a trajectory orientation that yields a near-optimal uplink rate. Fig. 7.2 also illustrates the benefit of increasing the number of antennas at the UAV, and that proper choice of the UAV heading can have a very large impact on communications performance.

7.4 Heading Optimization for a Mobile Ground Network

In this section we consider a more general scenario in which several mobile ground nodes are present and the UAV tracks their movement. We will consider both SDMA and TDMA approaches. In the SDMA scheme, all of the ground nodes are transmitting simultaneously and the UAV uses beamforming for source separation. For the TDMA method, each user is allocated an equal time slot for data transmission. It is assumed that at time step $n-1$ all of the users feedback their current position to the UAV, and these data are used to predict the positions at time n. An adaptive algorithm is proposed that calculates the UAV heading at time step $n-1$ so that the network's performance at time step n will be optimized.

7.4.1 Mobility Model and Position Prediction

We adopt a first-order auto-regressive (AR) model for the dynamics of the ground-based nodes [86], and we assume the nodes provide their location to the UAVs at each time step. The UAV in turn uses a Kalman filter to predict the positions of the nodes at the next time step. We define the dynamic state of user i at time step $n-1$ as:

$$\mathbf{s}_{i,n-1} = [x_{i,n-1}, y_{i,n-1}, v^x_{i,n-1}, v^y_{i,n-1}]^T,$$

where $v^x_{i,n-1}$, $v^y_{i,n-1}$ denote the velocities in the x direction and y direction respectively. According to the AR model, the state of node i at time step n is given by

$$\begin{aligned} \mathbf{s}_{i,n} &= \mathbf{T}\mathbf{s}_{i,n-1} + \mathbf{w}^{'}_{i,n} \\ \mathbf{T} &= \begin{bmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \end{aligned}$$

where $\mathbf{w}^{'}_{i,n} \sim \mathcal{N}(0, \sigma^2_w \mathbf{I}_4)$ represents a process noise term. Due to the effects of delay, quantization and possible decoding errors, the UAV's knowledge of the ground nodes' position is imprecise. This effect is described by the measurement model for user i's position:

$$\begin{aligned} \mathbf{z}_{i,n} &= \mathbf{F}\mathbf{s}_{i,n} + \mathbf{u}_{i,n} \\ \mathbf{F} &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \end{aligned}$$

where $\mathbf{u}_{i,n} \sim \mathcal{N}(0, \sigma_u^2 \mathbf{I}_2)$ represents the observation noise. We assume a standard implementation of the Kalman filter, as follows:

Initialization

$$\mathbf{x}_{i,0} = \mathbf{F}\mathbf{s}_{i,0},\ \mathbf{P}_{i,0} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Prediction

$$\begin{aligned} \hat{\mathbf{s}}_{i,n|n-1} &= \mathbf{T}\hat{\mathbf{s}}_{i,n-1|n-1}, \\ \mathbf{P}_{i,n|n-1} &= \mathbf{T}\mathbf{P}_{i,n-1|n-1}\mathbf{T} + \sigma_w^2 \mathbf{I}_4. \end{aligned}$$

Kalman gain

$$\mathbf{K}_{i,n} = \mathbf{P}_{i,n|n-1}\mathbf{F}^T(\mathbf{F}\mathbf{P}_{i,n|n-1}\mathbf{F}^T + \sigma_u^2 \mathbf{I}_2)^{-1}.$$

Measurement update

$$\begin{aligned} \hat{\mathbf{s}}_{i,n|n} &= \hat{\mathbf{s}}_{i,n|n-1} + \mathbf{K}_{i,n}(\mathbf{z}_{i,n} - \mathbf{F}\hat{\mathbf{s}}_{i,n|n-1}), \\ \mathbf{P}_{i,n|n} &= (\mathbf{I}_4 - \mathbf{K}_{i,n}\mathbf{F})\mathbf{P}_{i,n|n-1}. \end{aligned}$$

7.4.2 SDMA Scenario

The average sum rate of the uplink network can be approximated as follows:

$$\begin{aligned} C_n &= \sum_{i=1}^{N} \mathbb{E}\left\{\log_2(1 + SINR_{i,n})\right\} \\ &\simeq \sum_{i=1}^{N} \log_2\Big(1 + \mathbb{E}\{SINR_{i,n}\}\Big). \end{aligned} \tag{7.20}$$

The UAV heading δ_n will impact C_n in two ways. First, it will change the distance between the user nodes and the UAV during time step n, which will impact the received power. Second, and often most importantly, changes in the heading will modify the AoA of the LOS components, which impacts the ability of the beamformer to spatially separate the users. At time step $n-1$, based on the noisy observation $\mathbf{z}_{i,n-1}$, the UAV uses the Kalman filter to predict $(\hat{x}_{i,n}, \hat{y}_{i,n})$ and hence $\mathbb{E}\{SINR_{i,n}\}$. The heading optimization problem can thus be formulated as

$$\begin{aligned} \max_{\delta_n} \quad & \sum_{i=1}^{N} \log_2\left(1+\mathbb{E}\{SINR_{i,n}\}\right) \\ s.t. \quad & |\delta_n - \delta_{n-1}| \leq \Delta\delta \ , \end{aligned} \tag{7.21}$$

where $\Delta\delta$ represents that maximum change in UAV heading possible for the given time step.

The mean value of $SINR_{i,n}$ is calculated by

$$\begin{aligned} \mathbb{E}\{SINR_{i,n}\} &= \mathbb{E}\big\{ P_t \mathbf{h}_{i,n}^H \mathbb{E}\{\mathbf{Q}_{i,n}^{-1}\}\mathbf{h}_{i,n}\big\} \\ &= \frac{P_t}{d_{i,n}^{2\alpha}}\left(\frac{K}{K+1}\bar{\mathbf{h}}_{i,n}^H \mathbb{E}\{\mathbf{Q}_{i,n}^{-1}\}\bar{\mathbf{h}}_{i,n} + \frac{1}{K+1}\mathrm{tr}\Big(\mathbf{R}_r \mathbb{E}\{\mathbf{Q}_{i,n}^{-1}\}\Big)\right). \end{aligned}$$

Instead of working with the complicated term $\mathbb{E}\{\mathbf{Q}_{i,n}^{-1}\}$, we use instead the following approximation based on Jensen's inequality[87, Lemma 4]:

$$\mathbb{E}\{SINR_{i,n}\} \geq \frac{P_t}{d_{i,n}^{2\alpha}}\left(\frac{K}{K+1}\bar{\mathbf{h}}_{i,n}^H \mathbb{E}\{\mathbf{Q}_{i,n}\}^{-1}\bar{\mathbf{h}}_{i,n} + \frac{1}{K+1}\mathrm{tr}\Big(\mathbf{R}_r \mathbb{E}\{\mathbf{Q}_{i,n}\}^{-1}\Big)\right), \tag{7.22}$$

where

$$\mathbb{E}\{\mathbf{Q}_{i,n}\} = \sum_{j=1, j\neq i}^{N} \frac{P_t}{d_{j,n}^{2\alpha}}\left(\frac{K}{K+1}\bar{\mathbf{h}}_{j,n}\bar{\mathbf{h}}_{j,n}^H + \frac{1}{K+1}\mathbf{R}_r\right) + \sigma^2 \mathbf{I}_M.$$

We denote the approximation on the right side of equation (7.22) as $\mathbb{E}_l\{SINR_{i,n}\}$

and substitute it into (7.21), leading to a related optimization problem:

$$\begin{aligned} \max_{\delta_n} \quad & \sum_{i=1}^{N} \log_2(1 + \mathbb{E}_l\{SINR_{i,n}\}) \\ s.t. \quad & |\delta_n - \delta_{n-1}| \leq \Delta\delta. \end{aligned} \tag{7.23}$$

Problem (7.23) requires finding the maximum value of a single-variable function over a fixed interval $\delta_n \in [\delta_{n-1}-\Delta\delta,\ \delta_{n-1}+\Delta\delta]$, and thus can be efficiently solved using a one-dimensional line search. Note that the accuracy of the above sum rate approximations is less important than their ability to accurately predict the impact of changes to the UAV heading. The excellent performance achieved by our simulations based on (7.23) supports its use for this application.

Since problem (7.23) aims at maximizing the sum rate of the system, the algorithm may lead to a large difference in achievable rates between the users. As an alternative, we may wish to guarantee fairness among the users via, for example, the proportional fair method [88]:

$$\begin{aligned} \max_{\delta_n} \quad & \sum_{i=1}^{N} w_{i,n} \log_2\left(1 + \mathbb{E}_l\{SINR_{i,n}\}\right) \\ s.t. \quad & |\delta_n - \delta_{n-1}| \leq \Delta\delta, \end{aligned} \tag{7.24}$$

where $w_{i,n} \propto \bar{R}_{i,n}$ and $\bar{R}_{i,n}$ is user i's average data rate:

$$\bar{R}_{i,n} = \frac{1}{n-1} \sum_{k=1}^{n-1} \mathbb{E}\{\log_2(1 + SINR_{i,k})\}.$$

Based on our experience simulating the behavior of the algorithms described in (7.23) and (7.24), we propose two simple refinements that eliminate undesirable UAV behavior. First, to avoid the UAV frequently flying back and forth between the user nodes in an attempt to promote fairness, the weights $w_{i,n}$ in (7.24) are only updated

every N_w time steps rather than for every n. Second, we expect that the optimal position of the UAV should not stray too far from the center of gravity (CoG) of the ground nodes. This would not be the case if the users were clustered into very widely separated groups, but such a scenario would likely warrant the UAV serving the groups individually anyway (similar to the $D \gg h_u$ case discussed in Section 7.3). To prevent the UAV from straying too far from the CoG, at each time step the UAV checks to see if the calculated heading would put it outside a certain range $d_{\max}$ from the CoG. If so, instead of using the calculated value, it chooses a heading that points towards the CoG (or as close to this heading as possible subject to the turning radius constraint). Appropriate values for N_w and $d_{\max}$ are found empirically.

The proposed adaptive heading algorithm is summarized in the following steps:

1. Use the Kalman filter to predict the user positions $(\hat{x}_{i,n}, \hat{y}_{i,n})$ based on the noisy observations at time step $n-1$, and construct the objective function in (7.23) or (7.24) based on the predicted positions.

2. Use a line search to find the solution of (7.23) or (7.24) for $\delta_n \in [0, 2\pi]$, and denote the solution as $\tilde{\delta_n}$. Calculate the heading interval $\mathcal{O}_n = [\delta_{n-1} - \Delta\delta, \delta_{n-1} + \Delta\delta]$. If $\tilde{\delta}_n \in \mathcal{O}_n$, set $\delta_n = \tilde{\delta}_n$, else set $\delta_n = \arg\min_{\delta} |\delta - \tilde{\delta}_n|$, where $\delta = \delta_{n-1} - \Delta\delta$ or $\delta_{n-1} + \Delta\delta$.

3. Check to see if the calculated heading δ_n will place the UAV at a distance of $d_{\max}$ or greater from the predicted CoG of the users. If so, set $\delta_n = \delta_g$, where δ_g is the heading angle corresponding to the CoG, or set $\delta_n = \arg\min_{\delta} |\delta - \delta_g|$, where $\delta = \delta_{n-1} - \Delta\delta$ or $\delta_{n-1} + \Delta\delta$.

4. UAV flies with heading δ_n during time step n.

Note that the line search in step 2 is over $[0, 2\pi]$ rather than just $[\delta_{n-1} - \Delta\delta,\ \delta_{n-1} +$

$\Delta\delta]$, and the boundary point closest to the unconstrained maximum is chosen rather than the boundary with the maximum predicted rate. Thus, the algorithm may temporarily choose a lower overall rate in pursuit of the global optimum, although this scenario is uncommon.

7.4.3 TDMA Scenario

In the TDMA scenario, each node is assigned one time slot for data transmission. Since there is no interference from other users with TDMA, the beamformer in this case becomes simply the maximum ratio combiner $\mathbf{w}_{i,n} = \mathbf{h}_{i,n}$. Thus, the signal-to-noise ratio (SNR) of user i is given by

$$SNR_{i,n} = \frac{P_t}{\sigma^2}\|\mathbf{h}_{i,n}\|^2,$$

whose mean can be calculated as

$$\mathbb{E}\{SNR_{i,n}\} = \frac{P_t M}{d_{i,n}^{2\alpha}\sigma^2}.$$

For the TDMA scenario, the optimization problem is formulated as

$$\begin{aligned} \max_{\delta_n} \quad & \frac{1}{N}\sum_{i=1}^{N} w_{i,n}\log_2\left(1+\frac{P_t M}{d_{i,n}^{2\alpha}\sigma^2}\right) \\ \text{subject to} \quad & |\delta_n - \delta_{n-1}| \le \Delta\delta\ . \end{aligned} \tag{7.25}$$

where

$$w_{i,n} = \begin{cases} 1, & \text{max sum rate,} \\ \propto \bar{R}_i, & \text{proportional fair.} \end{cases}$$

The objective function in (7.25) can be substituted in step 2 of the adaptive heading algorithm described above to implement the TDMA approach.

7.5 Approximate Heading Algorithms

Under certain conditions, we can eliminate the need for the approximation in (7.22) when defining our adaptive heading control algorithm and thus simplify the algorithm implementation. In this section, we explore the asymptotic form of $SINR_{i,n}$ under both low and high SNR conditions. We show that in the low-SNR case, the optimal heading can be found in closed-form, without the need for a line search. In the high-SNR case, we show that maximizing the sum rate is equivalent to minimizing the sum of the users channel correlations, which can be achieved by checking a finite set of candidate headings. In Section 7.6, we show that the simpler asymptotic algorithms derived here provide performance essentially identical to the line-search algorithm of the previous section. Our discussion here will focus on the max-sum-rate case for SDMA; extensions to the proportional fair and TDMA cases are straightforward. To simplify the analysis, we have assumed $\beta(\phi_{i,n}) = 1$.

7.5.1 Asymptotic Analysis for Low SNR Case

For low SNR $\frac{P_t}{d_{i,n}^{2\alpha}\sigma^2} \ll 1$, the average sum rate in (7.20) is approximated by

$$C_n \approx \sum_{i}^{N} \mathbb{E}\{SINR_{i,n}\}$$

and problem (7.23) can be rewritten as follows

$$\begin{aligned} \max_{\delta_n} \quad & \sum_{i}^{N} \mathbb{E}\{SINR_{i,n}\} \\ s.t. \quad & |\delta_n - \delta_{n-1}| \leq \Delta\delta \ . \end{aligned} \tag{7.26}$$

In this case we can approximate $\mathbf{Q}_{i,n}^{-1}$ with the first order Neumann series [89, Theorem 4.20]:

$$\mathbf{Q}_{i,n}^{-1} \approx \frac{1}{\sigma^2}\left(\mathbf{I}_M - \sum_{j=1,j\neq i}^{N} \frac{P_t}{\sigma^2}\mathbf{h}_{j,n}\mathbf{h}_{j,n}^H\right). \tag{7.27}$$

Substituting (7.27) into (7.1), the $SINR_{i,n}$ for low SNR can be further expressed as

$$SINR_{i,n} = \frac{P_t}{\sigma^2}\mathbf{h}_{i,n}^H\left(\mathbf{I}_M - \sum_{j=1,j\neq i}^{N} \frac{P_t}{\sigma^2}\mathbf{h}_{j,n}\mathbf{h}_{j,n}^H\right)\mathbf{h}_{i,n},$$

and we obtain an approximation of $\mathbb{E}\{SINR_{i,n}\}$ as

$$\begin{aligned}
&\mathbb{E}\{SINR_{i,n}\} \\
= \;& \mathbb{E}\left\{\frac{P_t}{\sigma^2}\mathbf{h}_{i,n}^H\left(\mathbf{I}_M - \sum_{j=1,j\neq i}^{N} \frac{P_t}{\sigma^2}\mathbf{h}_{j,n}\mathbf{h}_{j,n}^H\right)\mathbf{h}_{i,n}\right\} \\
= \;& \frac{P_t}{d_{i,n}^{2\alpha}\sigma^2}\left(\frac{K}{K+1}\bar{\mathbf{h}}_{i,n}^H\left(\mathbf{I}_M - \sum_{j=1,j\neq i}^{N} \frac{P_t}{d_{j,n}^{2\alpha}\sigma^2}\left(\frac{K}{K+1}\bar{\mathbf{h}}_{j,n}\bar{\mathbf{h}}_{j,n}^H + \frac{1}{K+1}\mathbf{R}_r\right)\right)\bar{\mathbf{h}}_{i,n}\right. \\
& \left. + \frac{1}{K+1}\mathrm{tr}\left(\mathbf{R}_r - \sum_{j=1,j\neq i}^{N} \frac{P_t}{d_{j,n}^{2\alpha}\sigma^2}\left(\frac{K}{K+1}\mathbf{R}_r\bar{\mathbf{h}}_{j,n}\bar{\mathbf{h}}_{j,n}^H + \frac{1}{K+1}\mathbf{R}_r^2\right)\right)\right) \\
\overset{(a)}{\approx} \;& \frac{P_t}{d_{i,n}^{2\alpha}\sigma^2}\left(M - \sum_{j=1,j\neq i}^{N} \frac{P_t}{d_{j,n}^{2\alpha}\sigma^2}|\bar{\mathbf{h}}_{i,n}^H\bar{\mathbf{h}}_{j,n}|^2\right),
\end{aligned} \tag{7.28}$$

where in (7.28), (a) is based on the assumption of a large Rician factor K for the ground-to-air channel. When scaled by $\frac{P_t}{d_{i,n}^{2\alpha}\sigma^2} \ll 1$, the term involving $|\bar{\mathbf{h}}_{i,n}^H\bar{\mathbf{h}}_{j,n}|^2$ in the above equation plays a minor role in determining the value of $\mathbb{E}\{SINR_{i,n}\}$. Assuming $\Delta\delta$ and the ratio $\frac{v}{d_{i,n}}$ are small enough, we treat $|\bar{\mathbf{h}}_{i,n}^H\bar{\mathbf{h}}_{j,n}|$ as a constant when δ_n varies in $[\delta_{n-1} - \Delta\delta, \delta_{n-1} + \Delta\delta]$. We then approximate $|\bar{\mathbf{h}}_{i,n}^H\bar{\mathbf{h}}_{j,n}|$ as

$$|\bar{\mathbf{h}}_{i,n}^H\bar{\mathbf{h}}_{j,n}| \approx |\bar{\mathbf{h}}_{i,n}^{'H}\bar{\mathbf{h}}_{j,n}^{'}| = \left|\frac{\sin\left(\frac{M\pi}{2}\left(\cos(\phi_{i,n}^{'})\cos(\delta_{n-1}-\epsilon_{i,n}^{'}) - \cos(\phi_{j,n}^{'})\cos(\delta_{n-1}-\epsilon_{j,n}^{'})\right)\right)}{\sin\left(\frac{\pi}{2}\left(\cos(\phi_{i,n}^{'})\cos(\delta_{n-1}-\epsilon_{i,n}^{'}) - \cos(\phi_{j,n}^{'})\cos(\delta_{n-1}-\epsilon_{j,n}^{'})\right)\right)}\right|, \tag{7.29}$$

where $\phi_{i,n}^{'}$ and $\epsilon_{i,n}^{'}$ are calculated assuming the user nodes are located at $(\hat{x}_{i,n}, \hat{y}_{i,n})$ and the UAV is at $(x_{u,n-1}, y_{u,n-1}, h_u)$ with heading δ_{n-1}. The idea here is to use the UAV's position at time step $n-1$ to calculate the users' AoA at time step n. Moreover, $\frac{1}{d_{i,n}^{2\alpha}}$ can be approximated in the following way

$$\begin{aligned}
\frac{1}{d_{i,n}^{2\alpha}} &= \Big((x_{u,n-1} + v\cos\delta_n - x_{i,n})^2 + (y_{u,n-1} + v\sin\delta_n - y_{i,n})^2 + h_r^2\Big)^{-\alpha} \\
&= \Big((x_{u,n-1} - x_{i,n})^2 + (y_{u,n-1} - y_{i,n})^2 + v^2 + hr^2 + 2(x_{u,n-1} - x_{i,n})v\cos(\delta_n) \\
&\quad +2(y_{u,n-1} - y_{i,n})v\sin(\delta_n)\Big)^{-\alpha} \\
&\approx a_{i,n} - b_{i,n}\cos(\delta_n) - c_{i,n}\sin(\delta_n), \qquad (7.30)
\end{aligned}$$

where $a_{i,n}$, $b_{i,n}$ and $c_{i,n}$ are defined as follows

$$\begin{aligned}
a_{i,n} &= \Big((x_{u,n-1} - x_{i,n})^2 + (y_{u,n-1} - y_{i,n})^2 + v^2 + hr^2\Big)^{-\alpha} \\
b_{i,n} &= 2\alpha v(x_{u,n-1} - x_{i,n})\Big((x_{u,n-1} - x_{i,n})^2 + (y_{u,n-1} - y_{i,n})^2 + v^2 + hr^2\Big)^{-(\alpha+1)} \\
c_{i,n} &= 2\alpha v(y_{u,n-1} - y_{i,n})\Big((x_{u,n-1} - x_{i,n})^2 + (y_{u,n-1} - y_{i,n})^2 + v^2 + hr^2\Big)^{-(\alpha+1)}.
\end{aligned}$$

Substituting (7.29) and (7.30) into (7.28), C_n can be approximated as

$$\begin{aligned}
C_n &\approx \frac{P_t}{\sigma^2}\sum_{i=1}^{N} M(a_{i,n} - b_{i,n}\cos(\delta_n) - c_{i,n}\sin(\delta_n)) - \left(\frac{P_t}{\sigma^2}\right)^2 \sum_{i=1}^{N}\sum_{j=1,j\neq i}^{N} |\bar{\mathbf{h}}_{i,n}^{'H}\bar{\mathbf{h}}_{j,n}^{'}|^2 \Big(a_{i,n}a_{j,n} \\
&\qquad -(a_{i,n}b_{j,n} + b_{i,n}a_{j,n})\cos(\delta_n) - (a_{i,n}c_{j,n} + c_{i,n}a_{j,n})\sin(\delta_n)\Big) \\
&= \frac{MP_t}{\sigma^2}\sum_{i=1}^{N} a_{i,n} - \left(\frac{P_t}{\sigma^2}\right)^2 \sum_{i=1}^{N}\sum_{j=1,j\neq i}^{N} |\bar{\mathbf{h}}_{i,n}^{'H}\bar{\mathbf{h}}_{j,n}^{'}|^2 a_{i,n}a_{j,n} - \Bigg(\frac{MP_t}{\sigma^2}\sum_{i=1}^{N} b_{i,n} \\
&\quad -\left(\frac{P_t}{\sigma^2}\right)^2 \sum_{i=1}^{N}\sum_{j=1,j\neq i}^{N} |\bar{\mathbf{h}}_{i,n}^{'H}\bar{\mathbf{h}}_{j,n}^{'}|^2 (a_{i,n}b_{j,n} + b_{i,n}a_{j,n})\Bigg)\cos(\delta_n) - \Bigg(\frac{MP_t}{\sigma^2}\sum_{i=1}^{N} c_{i,n} \\
&\quad -\left(\frac{P_t}{\sigma^2}\right)^2 \sum_{i=1}^{N}\sum_{j=1,j\neq i}^{N} |\bar{\mathbf{h}}_{i,n}^{'H}\bar{\mathbf{h}}_{j,n}^{'}|^2 (a_{i,n}c_{j,n} + c_{i,n}a_{j,n})\Bigg)\sin(\delta_n). \qquad (7.31)
\end{aligned}$$

Define the first two terms in (7.31) as A_n, and the term multiplying $\cos(\delta_n)$ and $\sin(\delta_n)$ as B_n and D_n, respectively. Then (7.31) can be further expressed as

$$C_n = A_n - \sqrt{B_n^2 + D_n^2}\cos(\delta_n - \psi_n),$$

where

$$\psi_n = \begin{cases} \arctan\left(\dfrac{D_n}{B_n}\right) & \text{if } B_n \geq 0, \\ \arctan\left(\dfrac{D_n}{B_n}\right) + \pi & \text{otherwise.} \end{cases}$$

From this expression, we see that the average sum rate C_n can be written as a sinusoidal function of δ_n, and the maximizing heading δ_n is given by

$$\delta_n^* = \mathrm{mod}_{2\pi}(\psi_n + \pi).$$

So for low-SNR, the approximate closed-form solution to problem (7.26) is given by:

$$\delta_n = \begin{cases} \delta_n^* & \delta_l < \delta_n^* < \delta_u \\ \delta_{n-1} - \Delta\delta & \mathrm{mod}_\pi|\delta_l - \delta_n^*| < \mathrm{mod}_\pi|\delta_u - \delta_n^*| \\ \delta_{n-1} + \Delta\delta & \mathrm{mod}_\pi|\delta_l - \delta_n^*| \geq \mathrm{mod}_\pi|\delta_u - \delta_n^*| \end{cases}$$

where $\delta_l = \delta_{n-1} - \Delta\delta$ and $\delta_u = \delta_{n-1} + \Delta\delta$.

7.5.2 Asymptotic Analysis for High SNR Case

In the high SNR case where $\frac{P_t}{d_{i,n}^{2\alpha}\sigma^2} \gg 1$, the average sum rate maximization problem can be approximated as

$$\begin{aligned} \max_{\delta_n} \quad & \prod_{i=1}^{N} \mathbb{E}\{SINR_{i,n}\} \\ s.t. \quad & |\delta_n - \delta_{n-1}| \leq \Delta\delta \ . \end{aligned}$$

Here, when $\frac{P_t}{d_{i,n}^{2\alpha}\sigma^2} \gg 1$, we approximate $\mathbf{Q}_{i,n}^{-1}$ as follows:

$$\begin{aligned}
\mathbf{Q}_{i,n}^{-1} &= \frac{1}{\sigma^2}\left(\mathbf{I}_M + \frac{P_t}{\sigma^2}\mathbf{H}_{i,n}\mathbf{D}_{i,n}\mathbf{H}_{i,n}^H\right)^{-1} \\
&\overset{(b)}{=} \frac{1}{\sigma^2}\left(\mathbf{I}_M - \frac{P_t}{\sigma^2}\mathbf{H}_{i,n}\mathbf{D}_{i,n}\Big(\mathbf{I}_M + \frac{P_t}{\sigma^2}\mathbf{H}_{i,n}^H\mathbf{H}_{i,n}\mathbf{D}_{i,n}\Big)^{-1}\mathbf{H}_{i,n}^H\right) \\
&\overset{(c)}{\approx} \frac{1}{\sigma^2}\left(\mathbf{I}_M - \mathbf{H}_{i,n}\big(\mathbf{H}_{i,n}^H\mathbf{H}_{i,n}\big)^{-1}\mathbf{H}_{i,n}^H\right), \qquad (7.32)
\end{aligned}$$

where (b) is due to the matrix inversion lemma, (c) is due to the approximation $\big(\mathbf{I}_M + \frac{P_t}{\sigma^2}\mathbf{H}_{i,n}^H\mathbf{H}_{i,n}\mathbf{D}_{i,n}\big)^{-1} \approx \big(\frac{P_t}{\sigma^2}\mathbf{H}_{i,n}^H\mathbf{H}_{i,n}\mathbf{D}_{i,n}\big)^{-1}$, and

$$\begin{aligned}
\mathbf{D}_{i,n} &= \operatorname{diag}\left\{\frac{1}{d_{1,n}^{2\alpha}}\cdots\frac{1}{d_{i-1,n}^{2\alpha}}\frac{1}{d_{i+1,n}^{2\alpha}}\cdots\frac{1}{d_{N,n}^{2\alpha}}\right\} \\
\mathbf{H}_{i,n} &= [\mathbf{h}_{1,n}\ \cdots\ \mathbf{h}_{i-1,n}\ \mathbf{h}_{i+1,n}\ \cdots\ \mathbf{h}_{N,n}]
\end{aligned}$$

are formed by eliminating the terms for user i. Plugging (7.32) into (7.1), we obtain

$$SINR_{i,n} \approx \frac{P_t}{\sigma^2 d_{i,n}^{2\alpha}}\left(\mathbf{h}_{i,n}^H\mathbf{h}_{i,n} - \left\|\mathbf{h}_{i,n}^H\mathbf{H}_{i,n}\big(\mathbf{H}_{i,n}^H\mathbf{H}_{i,n}\big)^{-1}\mathbf{H}_{i,n}^H\right\|^2\right).$$

For large K-factor channels we ignore the contribution of the Rayleigh term, so that

$$\mathbb{E}\{SINR_{i,n}\} \approx \frac{P_t}{\sigma^2 d_{i,n}^{2\alpha}}\left(M - \left\|\bar{\mathbf{h}}_{i,n}^H\bar{\mathbf{H}}_{i,n}\big(\bar{\mathbf{H}}_{i,n}^H\bar{\mathbf{H}}_{i,n}\big)^{-1}\bar{\mathbf{H}}_{i,n}^H\right\|^2\right),$$

where $\bar{\mathbf{H}}_{i,n}$ is defined similarly to $\mathbf{H}_{i,n}$. Thus, the heading optimization problem can be written as

$$\begin{aligned}
\max_{\delta_n}\quad & \prod_{i=1}^{N}\frac{P_t}{d_{i,n}^{\alpha}\sigma^2}\prod_{i=1}^{N}\left(M - \left\|\bar{\mathbf{h}}_{i,n}^H\bar{\mathbf{H}}_{i,n}\big(\bar{\mathbf{H}}_{i,n}^H\bar{\mathbf{H}}_{i,n}\big)^{-1}\bar{\mathbf{H}}_{i,n}^H\right\|^2\right) \qquad (7.33) \\
s.t.\quad & |\delta_n - \delta_{n-1}| \le \Delta\delta.
\end{aligned}$$

At this point we make two further approximations. First, we will ignore the terms

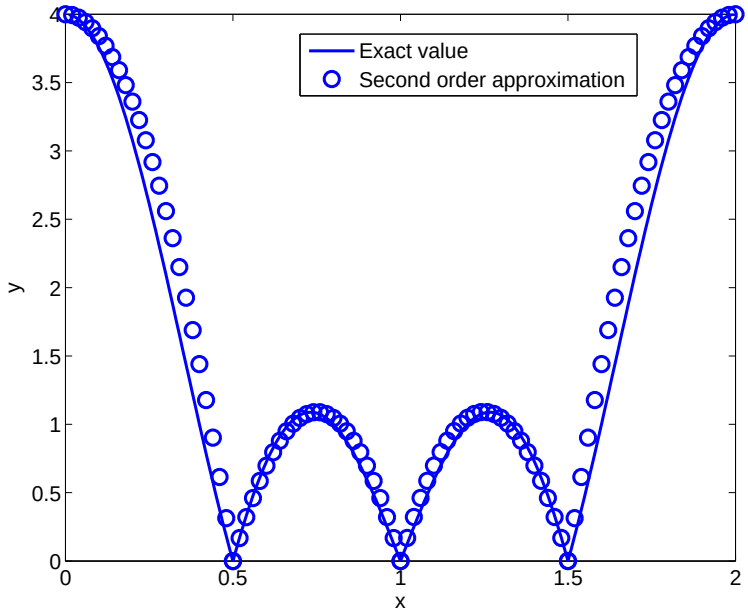

Figure 7.3: Plot of $|\bar{\mathbf{h}}_i^H \bar{\mathbf{h}}_j|$ as a function of the AoA between the two users, along with a set of piecewise quadratic approximations.

in the product involving $1/d_{i,n}$, since $d_{i,n}$ will not change appreciably over one time step compared with the terms involving products of $\bar{\mathbf{h}}_{i,n}$, which are angle-dependent. Second, we will make the assumption that the matrix $\bar{\mathbf{H}}_{i,n}^H \bar{\mathbf{H}}_{i,n}$ is approximately diagonal, which implies that the UAV attempts to orient itself so that the correlation between the mean channel vectors for different users is minimized. If we then apply these two assumptions to (7.33), we find that the heading problem reduces to

$$\begin{aligned} \min_{\delta_n} \quad & \sum_{i=1}^{N} \sum_{j=i+1}^{N} |\bar{\mathbf{h}}_{i,n}^H \bar{\mathbf{h}}_{j,n}| \\ s.t. \quad & |\delta_n - \delta_{n-1}| \leq \Delta\delta \ , \end{aligned} \tag{7.34}$$

which is consistent with the assumption of minimizing inter-user channel correlation.

In Fig. 7.3, we show a plot of $|\bar{\mathbf{h}}_{i,n}^H \bar{\mathbf{h}}_{j,n}|$ for $M = 4$ as a function of the difference in AoA between the two users (variable x in the plot). It is clear that $|\bar{\mathbf{h}}_{i,n}^H \bar{\mathbf{h}}_{j,n}|$ is a piecewise concave function. Since a sum of concave functions is also concave, the

criterion in (7.34) is piecewise concave as well. Since the minimum of a concave function must be located at the boundary of its domain, to find the solution to (7.34) it is enough to evaluate the criterion at the boundary points $\{\delta_{n-1} - \Delta\delta, \delta_{n-1} + \Delta\delta\}$ and the zero points of $|\bar{\mathbf{h}}_{i,n}^H \bar{\mathbf{h}}_{j,n}|$ located within $[\delta_{n-1} - \Delta\delta, \delta_{n-1} + \Delta\delta]$. To find the zero locations, we use the fact that a piecewise quadratic approximation to $|\bar{\mathbf{h}}_{i,n}^H \bar{\mathbf{h}}_{j,n}|$ is very accurate (as depicted in Fig. 7.2). When $\Delta\delta$ is not too large, the phase term $p_{i,n}$ in (7.2) satisfies

$$\begin{aligned} p_{i,n} &\approx \pi \cos(\phi_{i,n}^{'}) \Big(\cos(\epsilon_{i,n}^{'} - \delta_{n-1}) + \sin(\epsilon_{i,n}^{'} - \delta_{n-1})(\delta_n - \delta_{n-1}) \Big) \\ &= e_{i,n} + f_{i,n} x, \end{aligned} \tag{7.35}$$

where $x = \delta_n - \delta_{n-1}$, $e_{i,n} = \pi \cos(\phi_{i,n}^{'}) \cos(\epsilon_{i,n}^{'} - \delta_{n-1})$, $f_{i,n} = \pi \cos(\phi_{i,n}^{'}) \sin(\epsilon_{i,n}^{'} - \delta_{n-1})$, $x \in [-\Delta\delta, \Delta\delta]$ and the calculation of $\phi_{i,n}^{'}$ and $\epsilon_{i,n}^{'}$ follows (7.29). Based on (7.35), we obtain

$$|\bar{\mathbf{h}}_{i,n}^H \bar{\mathbf{h}}_{j,n}| \approx \left| \frac{\sin\Big(\frac{M}{2}\big((f_{i,n} - f_{j,n})x + e_{i,n} - e_{j,n}\big)\Big)}{\sin\Big(\frac{1}{2}\big((f_{i,n} - f_{j,n})x + e_{i,n} - e_{j,n}\big)\Big)} \right|.$$

Then the zero points of $|\bar{\mathbf{h}}_{i,n}^H \bar{\mathbf{h}}_{j,n}|$ in terms of x are approximately given by[1]

$$z_k^{i,j} = \frac{2k\pi/M - e_{i,n} + e_{j,n}}{f_{i,n} - f_{j,n}}, \qquad k = \pm 1, \ldots, \pm 2M - 1.$$

Finally, the asymptotic solution to problem (7.34) can be written as

$$\delta_n = \arg\min_{\delta_n} \sum_{i=1}^{N} \sum_{j=i+1}^{N} |\bar{\mathbf{h}}_{i,n}^H \bar{\mathbf{h}}_{j,n}|, \forall \delta_n \in \{z_k^{i,j} \in [-\Delta\delta, \Delta\delta]\} \cup \{\delta_{n-1} - \Delta\delta, \delta_{n-1} + \Delta\delta\}.$$

[1]Where we assume $\Delta\delta < 1$, $|(f_{i,n} - f_{j,n})\, x + e_{i,n} - e_{j,n}| < 4\pi$ and we only consider the zero points in $[-4\pi, 4\pi]$.

7.6 Simulation Results

A simulation example involving a UAV with a 4-element ULA and four users was carried out to test the performance of the proposed algorithm. The time between UAV heading updates was $\Delta t = 1$s, and the simulation was conducted over $L = 300$ steps. The initial speed of all nodes was 10m/s, and their initial positions in meters were $(0, 25)$, $(240, 20)$, $(610, 30)$, $(1240, 20)$. To describe the user mobility, we assume a state-space model with random process noise on the user's position and velocity, and we assume the UAV uses a standard Kalman filter to predict future user positions. The user's transmit power was set to $\frac{P_t}{\sigma^2} = 45$dB. We assume free space propagation for the large-scale fading, and thus the path loss exponent was chosen as $\alpha = 1$ [90, chap. 3]. Halfway through the simulation, at step 150, all the nodes make a sharp turn and change their velocity according to $v^y_{i,150}/v^x_{i,150} = -1.8856$, where $v^x_{i,150}$ and $v^y_{i,150}$ represent the velocity of the i-th user in the x and y-directions, respectively. The initial position of the UAV was $(x_{u,0}, y_{u,0}) = (50, 100)$m and its altitude was $h_u = 350$m. The speed of the UAV was $v_u = 50$m/s, and the maximum heading angle change was set to be either $\Delta\delta = \frac{\pi}{6}$ or $\frac{\pi}{9}$ depending on the case considered. The angle spread factor in (7.4) was set to $\sigma_r^2 = 0.05$. For the proportional fair case, N_w was set to 4 and for the high SNR case, $d_{\max}$ was set to 300m. For simplicity, we set $\beta(\phi_{i,n}) = 1$.

Figs. 7.4-7.7 show the trajectories of the UAV and mobile nodes for the SDMA and TDMA scenarios assuming either max-sum or proportional fair objective functions and $\Delta\delta = \frac{\pi}{6}$. The decision-making behavior of the UAV is evident from its ability to appropriately track the nodes as they dynamically change position. Due to the relatively high speed of the UAV compared with the ground-based nodes, in some cases the UAV is forced to fly in a tight circular trajectory to maintain an optimal position for the uplink communications signals. In the proportional-fair approach, the

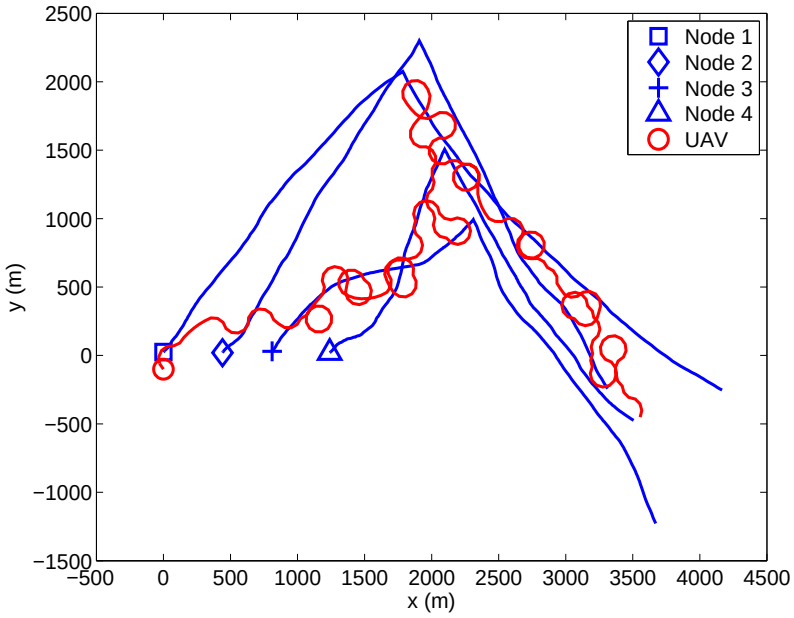

Figure 7.4: Trajectories of the UAV and user nodes for SDMA with $\Delta\delta = \frac{\pi}{6}$, $K = 10$ and $\frac{P_t}{\sigma^2} = 45$dB, maximizing sum rate. The average sum rate is: 1.8185 bps/Hz. The single user data rates are $u_1 = 0.5607, u_2 = 0.6138, u_3 = 0.2406, u_4 = 0.4034$.

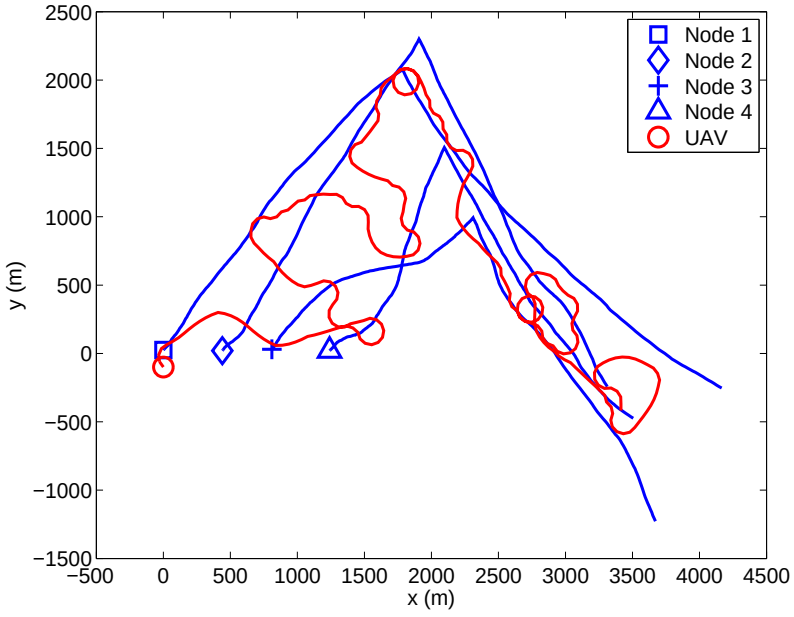

Figure 7.5: Trajectories of the UAV and user nodes for SDMA with $\Delta\delta = \frac{\pi}{6}$, $K = 10$ and $\frac{P_t}{\sigma^2} = 45$dB, proportional fair algorithm. The average sum rate is 1.6968 bps/Hz with $u_1 = 0.4169, u_2 = 0.4084, u_3 = 0.4088, u_4 = 0.4627$.

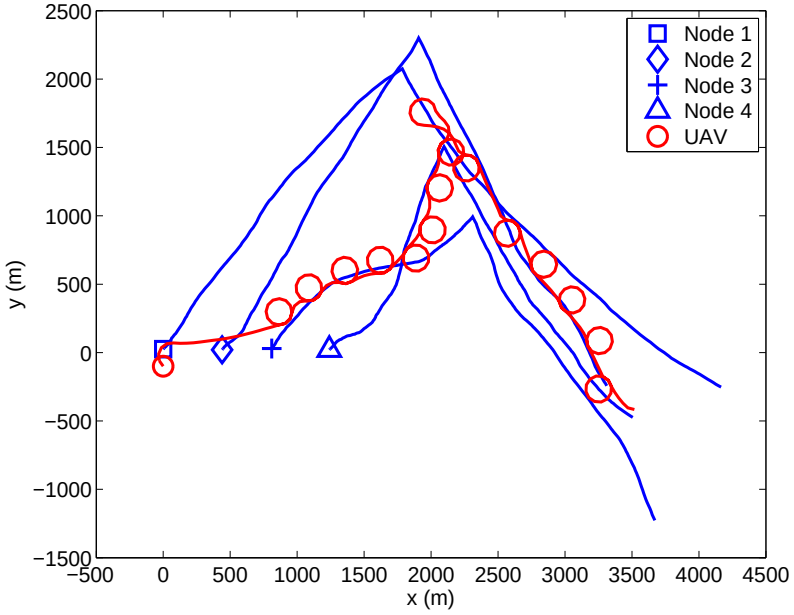

Figure 7.6: Trajectories of the UAV and user nodes for TDMA with $\Delta\delta = \frac{\pi}{6}$, $K = 10$ and $\frac{P_t}{\sigma^2} = 45$dB, maximizing sum rate. The average sum rate is: 0.5294 bps/Hz, with $u_1 = 0.1418, u_2 = 0.1674, u_3 = 0.0895, u_4 = 0.1307$.

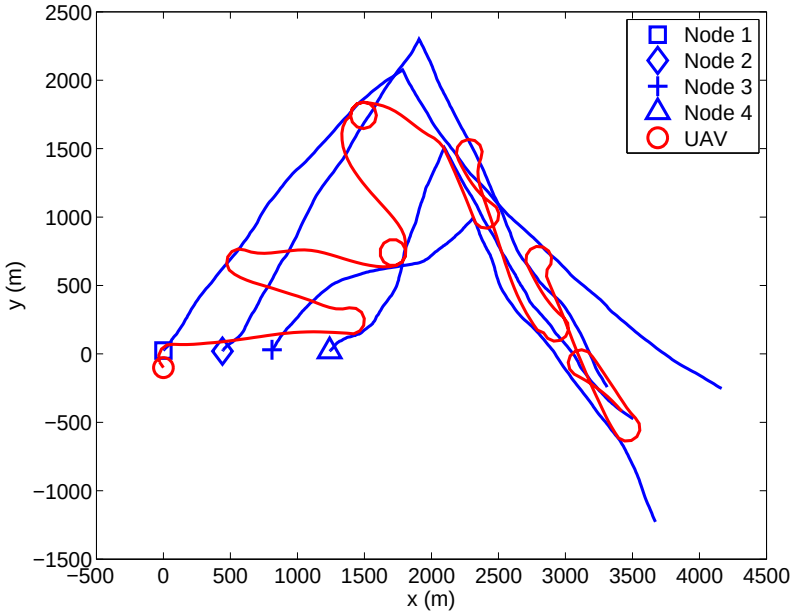

Figure 7.7: Trajectories of the UAV and user nodes for TDMA with $\Delta\delta = \frac{\pi}{6}$, $K = 10$ and $\frac{P_t}{\sigma^2} = 45$dB, proportional fair algorithm. The average sum rate is 0.5139 bps/Hz, with $u_1 = 0.1222, u_2 = 0.1274, u_3 = 0.1193, u_4 = 0.1450$.

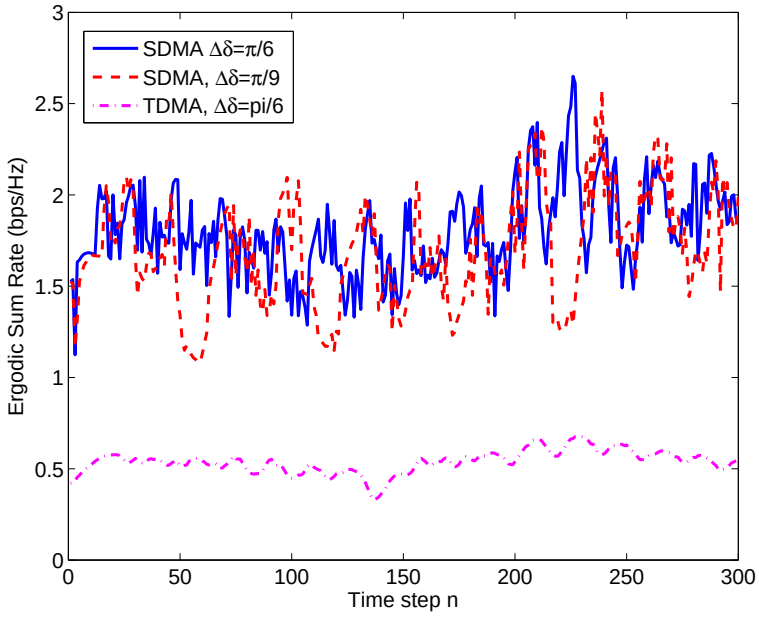

Figure 7.8: Comparison of sum rate performance (bps/Hz) with $K = 10$ and $\frac{P_t}{\sigma^2} =$ 45dB, maximizing sum rate. The average sum rates are: 1.802 for SDMA with $\Delta\delta = \frac{\pi}{6}$, 1.6921 for SDMA with $\Delta\delta = \frac{\pi}{9}$, and 0.5377 for TDMA with $\Delta\delta = \frac{\pi}{6}$.

UAV tends to visit the nodes in turn, while the max-sum rate algorithm leads to the UAV approximately tracking the area where the user node density is highest. Note that in this example the proportional-fair algorithm only suffers a slight degradation in overall sum rate compared with the max-sum rate approach.

Figs. 7.8-7.9 show the ergodic sum rate for the different scenarios. For each time step, the rate is calculated by averaging over 1000 independent channel realizations. Results for both $\Delta\delta = \frac{\pi}{6}$ and $\frac{\pi}{9}$ are plotted. Increasing the maximum turning rate will clearly provide better performance since it gives the UAV more flexibility in choosing its heading. The benefit of using SDMA is also apparent from Figs. 7.8-7.9, where we see that a rate gain of approximately a factor of 3.3 is achieved over the TDMA scheme. We also note that the obtained sum rate is only about 16% less than what would be achieved assuming no interference, indicating the effectiveness of the beamforming algorithm.

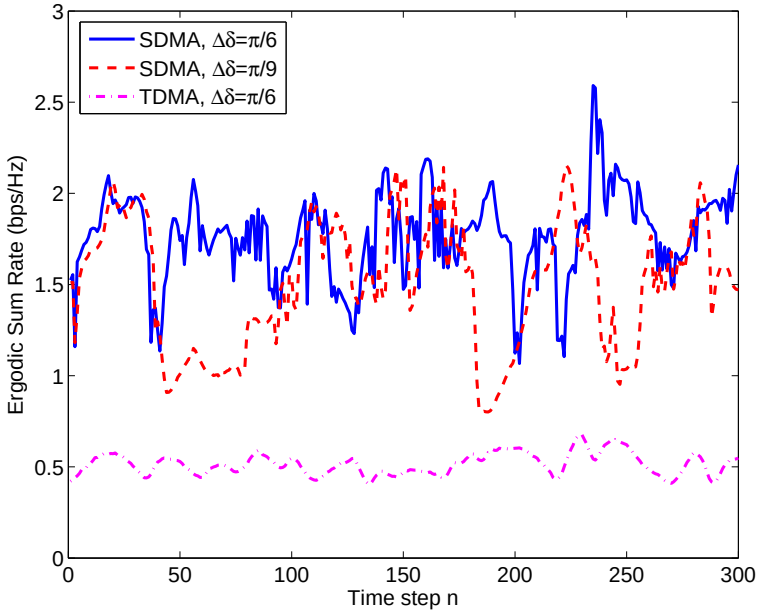

Figure 7.9: Comparison of sum rate performance (bps/Hz) with $K = 10$ and $\frac{P_t}{\sigma^2} =$ 45dB, proportional fair algorithm. The average sum rates are: 1.6968 for SDMA with $\Delta\delta = \frac{\pi}{6}$, 1.6042 for SDMA with $\Delta\delta = \frac{\pi}{9}$, 0.5139 for TDMA with $\Delta\delta = \frac{\pi}{6}$.

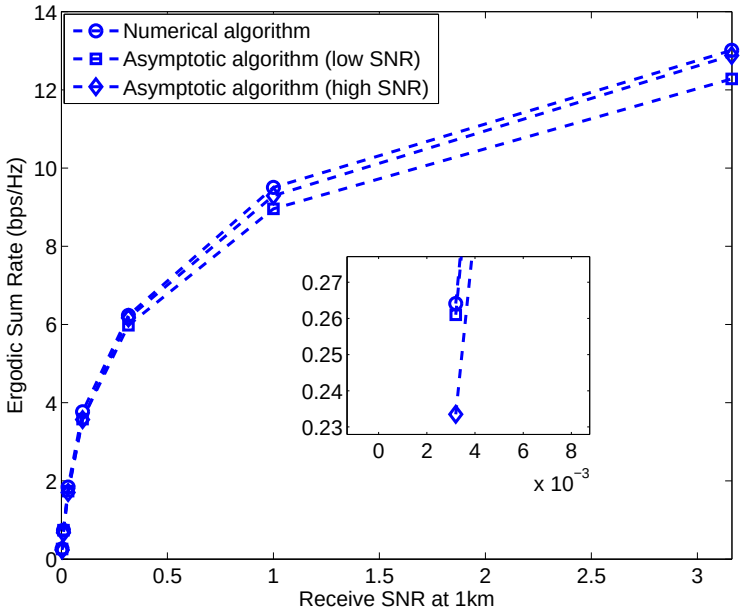

Figure 7.10: Comparison of the average sum rate of the line-search and closed-form approximations with $\Delta\delta = \frac{\pi}{9}$, $K = 1000$, maximizing sum rate. The x-axis denotes the SNR that would be observed at the UAV for a user node at a distance of 1km.

Fig. 7.10 compares the average sum rate of the line-search algorithm in (7.23) with both the low- and high-SNR approximations derived in the previous section. The K-factor for this example was 1000 and $\Delta\delta = \pi/9$. The performance is plotted as a function of the received SNR that would be observed at the UAV from a ground node located at a distance of 1km. Although the approximate algorithms were derived separately under different SNR assumptions, both of them yield performance essentially identical to (7.23) over all SNR values. Each approximate algorithm is slightly better than the other in its respective SNR regime, but the performance difference is small.

7.7 Summary

We have investigated the problem of positioning a multiple-antenna UAV for enhanced uplink communications from multiple ground-based users. We studied the optimal UAV trajectory for a case involving two static users, and derived an approximate method for finding this trajectory that only requires a simple line search. For the case of a network of mobile ground users, an adaptive heading algorithm was proposed that uses predictions of the user terminal positions and beamforming at the UAV to maximize SINR at each time step. Two kinds of optimization problems were considered, one that maximizes an approximation to the average uplink sum rate and one that guarantees fairness among the users using the proportional fair method. Simulation results indicate the effectiveness of the algorithms in automatically generating a suitable UAV heading for the uplink network, and demonstrate the benefit of using SDMA over TDMA in achieving the best throughput performance. We also derived approximate solutions to the UAV heading problem for low- and high-SNR scenarios; the approximations allow for a closed-form solution instead of a line search, but still provide near-optimal performance in their respective domains.

Chapter 8

Conclusion

This dissertation investigated the physical layer optimization for the wireless sensing and network connectivity. For the amplify-and-forward wireless sensor network with a coherent multiple access channel, we studied the problems of how to improve the estimation or detection performance through optimizing the complex transmission gain at sensor nodes or configuring fusion center (FC) with multiple antennas. Additionally, when the network doesn't have a fixed infrastructure in which the sensor nodes are mobile, a multi-antenna unmanned aerial vehicle (UAV) is used as relay to improve the connectivity between the sensor nodes and the FC. The UAV's trajectory is adaptively optimized based on the positions of the sensor nodes.

For the single-antenna FC, we considered the problem of how to allocate the transmit power at sensor nodes when the FC attempts to track a dynamic parameter based on the measurements from the sensor nodes. We formulated problems with either constraints on sensor transmit power or constraints on achieved MSE, and for the power constraint we examined cases involving global sum and individual sensor power constraints. To minimize MSE under global sum power constraint or minimize

the sum transmit power under MSE constraint, we have found the optimal closed-form solutions of sensor transmission gain through maximizing a related Rayleigh quotient. When the sensor node's individual power constraint is taken into account, we have demonstrated that a rank-relaxed semidefinite program (SDP) can be used to find optimal solutions for minimizing MSE under individual power constraints, or minimizing peak sensor power under MSE constraints. Furthermore, we derived asymptotic expressions for the optimal transmission gains for both low and high SNR at the FC, and when the number of sensors is large, we found an approximate expression for the required sum transmit power to achieve a certain value of MSE. For the special case where the sensors transmit with equal power, the exact expression for the MSE outage probability has also been derived. Several simulation results have been presented to verify the performance of the optimal solutions and confirm the analysis. Additionally, we generalized the scalar dynamic model to the vector process, and we investigated the problem of a Kalman filter with a linearly reconfigurable observation matrix. Two kinds of problems were formulated: Minimize-Sum-MSE or Minimize-Maximum-MSE. When the observed signal is a vector, both of the above mentioned optimization problems are difficult to solve directly, and we proposed a method to divide the original problem into two simpler sub-problems that are easier to solve. Simulation results show that when the quadratic constraint has a small value, this approach can provide performance close to the MSE lower bound. For the scalar observation model, we transformed the Minimize-Sum-MSE to an equivalent Rayleigh quotient maximization problem which has an optimal closed-form solution, and to solve the Minimize-Maximum-MSE, the rank-one constraint on the observation parameters is relaxed and the resulting problem becomes an SDP feasibility problem. If the solution to the SDP feasibility problem is not rank-one, a rank-one solution need to be reconstructed. Simulation results show that the solution to the feasibility problem is most likely to be rank-one.

To further improve the detection or estimation performance of WSN, the FC has been configured with multiple antennas. We first formulate a phase-shift and forward scheme in which the sensors nodes apply a phase shift to their noisy measurements before forwarding them to the multi-antenna FC. Two algorithms were proposed for finding the sensor phase shifts that minimize the estimation error at FC, one based on a relaxed SDP and a closed-form heuristic algorithm based on the ACMA approach. When the number of sensors or number of antennas is large, asymptotic estimation error has been analyzed, and we have derived the conditions under which the multi-antenna FC can provide more benefit than single-antenna FC. Due to the feedback channel between the FC and sensor nodes, there may exist errors in the sensor's phase shift. The phase error's impact on the estimation performance was studied and an upper bound for the performance degradation has been derived. To save the battery power and extend the life time of the network, sensor selection problem was formulated and solved, in which only a subset of senor nodes are selected for signal transmission. For the low SNR scenario, a linear programming was used to find the subset of the sensor nodes and for the high SNR scenario, we showed that selecting the sensors with lowest measurement noise for signal transmission was approximately optimal.

When the FC possesses a massive number of antennas, we have shown that the detection or estimation performance only depends on the transmission gain of sensor nodes and the phase can be arbitrary. Under the assumptions that the FC has a perfect knowledge of the channel state information (CSI) from sensor nodes, we studied Neyman-Pearson (NP) detector and linear minimum mean-squared error (LMMSE) estimator, derived optimal sensor transmission gains for each case and showed that for the case of NP detection and LMMSE estimation, constant levels of performance can be achieved if the transmit power at the sensors is reduced proportional to the gain in the number of antennas. If the CSI is unknown to the FC, the energy detector can be

used for signal detection. Optimal solution has been found to maximize the deflection coefficient of the energy detector and it has been proved that for energy detector, a constant deflection coefficient can be maintained if power is reduced proportional to the inverse square root of the number of antennas. The performance bounds derived for NP detection and LMMSE estimation showed that the multi-antenna FC can provide significant benefits over the single-antenna FC in low sensor transmit power scenario, and the benefit disappears when the transmit power is high. However, for the energy detector, compared with single-antenna FC, multi-antenna FC can always provide a better detection performance.

The above mentioned results are based on the assumption that the network has a fixed infrastructure, in which the FC can optimize the network configuration, and then feed the solutions back to the sensor nodes. In military or disaster response scenarios, the network may have a random topology and it's possible that the distance between sensor nodes and FC exceeds their communication limits. For such a scenario, we investigated how to use the multi-antenna UAV as a relay to improve the connectivity between the sensor nodes and the FC. The UAV utilizes beamforming to separate difference users' signal and using the estimation of users' positions as an input, the heading direction of the UAV is adaptively optimized to maximize the uplink sum rate or minimum rate of the user nodes moving on the ground. Several UAV heading optimization algorithms were proposed and simulation results have verified the effectiveness of these algorithms in achieving an optimal throughput.

Bibliography

[1] M. Gastpar, "Uncoded transmission is exactly optimal for a simple Gaussian "sensor" network," *IEEE Trans. Info. Theory*, vol. 54, pp. 5247–5251, Nov. 2008.

[2] M. Gastpar and M. Vetterli, "Power, spatio-temporal bandwidth, and distortion in large sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 23, pp. 745–754, Apr. 2005.

[3] M. Gastpar and M. Vetterli, "Source channel communication in sensor networks," in *Proc. ACM/IEEE IPSN 2003*, pp. 162–177, Apr. 2003.

[4] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, pp. 3590–3600, Nov. 2010.

[5] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, pp. 1436–1449, Apr. 2013.

[6] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?," *IEEE J. Sel. Areas Commun.*, vol. 31, pp. 160–171, Feb. 2013.

[7] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. of Sel. Topics in Signal Processing*, vol. 8, pp. 742–758, Oct. 2014.

[8] F. Jiang, J. Chen, and A. L. Swindlehurst, "Optimal power allocation for parameter tracking in a distributed amplify-and-forward sensor network," *IEEE Trans. Signal Process.*, vol. 62, pp. 2200–2211, May 2014.

[9] J. C. Feng Jiang and A. L. Swindlehurst, "Linearly reconfigurable kalman filtering for a vector process," in *Proc. IEEE ICASSP 2013*, pp. 5725–5729, May 2013.

[10] F. Jiang, J. Chen, and A. L. Swindlehurst, "Estimation in phase-shift and forward wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 61, pp. 3840–3851, Aug. 2013.

[11] F. Jiang, J. Chen, A. L. Swindlehurst, and J. A. Lopez-Salcedo, "Massive mimo for wireless sensing with a coherent multiple access channel," *accepted by IEEE Trans. Signal Process.*, Feb. 2015.

[12] F. Jiang and A. L. Swindlehurst, "Optimization of uav heading for the ground-to-air uplink," *IEEE J. Sel. Areas Commun.*, vol. 30, pp. 993–1005, Jun. 2012.

[13] S. Cui, J.-J. Xiao, A. J. Goldsmith, Z.-Q. Luo, and H. V. Poor, "Estimation diversity and energy efficiency in distributed sensing," *IEEE Trans. Signal Process.*, vol. 55, pp. 4683–4695, Sep. 2007.

[14] J.-J. Xiao, S. Cui, Z.-Q. Luo, and A. J. Goldsmith, "Linear coherent decentralized estimation," *IEEE Trans. Signal Process.*, vol. 56, pp. 757–770, Feb. 2008.

[15] I. Bahceci and A. K. Khandani, "Linear estimation of correlated data in wireless sensor network with optimum power allocation and analog modulation," *IEEE Trans. Commun.*, vol. 56, pp. 1146–1156, Jul. 2008.

[16] J. Matamoros and C. Anton-Haro, "Scaling law of an opportunistic power allocation scheme for amplify-and-forward wireless sensor networks," *IEEE Commun. Letter*, vol. 15, pp. 169–171, Feb. 2011.

[17] J. Fang and H. Li, "Power constrained distributed estimation with correlated sensor data," *IEEE Trans. Signal Process.*, vol. 57, pp. 3292–3297, Aug. 2009.

[18] A. S. Leong and S. Dey, "On scaling laws of diversity schemes in decentralized estimation," *IEEE Trans. Info. Theory*, vol. 57, pp. 4740–4759, Jul. 2011.

[19] H. Senol and C. Tepedelenlioglu, "Performance of distributed estimation over unknown parallel fading channels," *IEEE Trans. Signal Process.*, vol. 56, pp. 6057–6068, Dec. 2009.

[20] C.-H. Wang, A. S. Leong, and S. Dey, "Distortion outage minimization and diversity order analysis for coherent multiaccess," *IEEE Trans. Signal Process.*, vol. 59, pp. 6144–6159, Jan. 2011.

[21] S. Kar and P. K. Varshney, "Linear coherent estimation with spatial collaboration," *IEEE Trans. Info. Theory*, vol. 59, pp. 3532–3553, Jun. 2013.

[22] M. K. Banavar, C. Tepedelenlioglu, and A. Spanias, "Estimation over fading channels with limited feedback using distributed sensing," *IEEE Trans. Signal Process.*, vol. 58, pp. 414–425, Jan. 2010.

[23] A. S. Leong, S. Dey, and J. S. Evans, "Asymptotics and power allocation for state estimation over fading channels," *IEEE Trans. Aero. and Elec. Sys.*, vol. 47, pp. 611–633, Jan. 2011.

[24] A. S. Leong, S. Dey, G. N. Nair, and P. Sharma, "Power allocation for outage minimization in state estimation over fading channels," *IEEE Trans. Signal Process.*, vol. 59, pp. 3382–3397, July 2011.

[25] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. NJ: Prentice Hall, 1993.

[26] A. Beck and M. Teboulle, "On minimizing quadratically constrained ratio of two quadratic functions," *Journal of Convex Analysis*, vol. 17, no. 3 and 4, pp. 789–804, 2010.

[27] W.-C. Liao, T.-H. Chang, W.-K. Ma, and C.-Y. Chi, "QoS-based transmit beamforming in the presence of eavesdroppers: An optimized artificial-noise-aided approach," *IEEE Trans. Signal Process.*, vol. 59, pp. 1202–1216, Mar. 2011.

[28] Y. Zhang and A. M.-C. So, "Optimal spectrum sharing in MIMO cognitive radio networks via semidefinite programming," *IEEE J. Sel. Areas Commun.*, vol. 29, pp. 362–373, Feb. 2011.

[29] J. E. Gentle, *Matrix Algebra: Theory, Computations, and Applications in Statistics*. New York: Springer, 2007.

[30] T. Y. Al-Naffouri and B. Hassibi, "On the distribution of indefinite quadratic forms in Gaussian random variables," in *Proc. IEEE ISIT 2009*, pp. 1744–1748, Jun. 2009.

[31] W. So, "Rank one perturbation and its application to the Laplacian spectrum of a graph," *Linear and Multilinear Algebra*, vol. 46, no. 3, pp. 193–198, 1999.

[32] J. Salmi, A. Richter, and V. Koivunen, "Detection and tracking of MIMO propagation path parameters using state-space approach," *IEEE Trans. Signal Process.*, vol. 57, pp. 1538–1550, Apr. 2009.

[33] P. Zhan, K. Yu, and A. L. Swindlehurst, "Wireless relay communications with unmannned aerial vehicles: performance and optimization," *IEEE Trans. Aero. and Elec. Sys. (to appear)*, 2010.

[34] F. Jiang, J. Chen, and A. L. Swindlehurst, "Parameter tracking via optimal distributed beamforming in an analog sensor network," in *Proc. Asilorma Conference on Signals, Systems and Computers 2012*, pp. 1397–1401, Nov. 2012.

[35] Z.-Q. Luo, T. N. Davidson, G. B. Giannakis, and K. M. Wong, "Transceiver optimization for block-based multiple access through isi channels," *IEEE Trans. Signal Process.*, vol. 52, pp. 1037–1052, Apr. 2004.

[36] V. Havary-Nassab, S. Shahbazpanahi, A. Grami, and Z.-Q. Luo, "Distributed beamforming for relay networks based on second-order statistics of the channel state information," *IEEE Trans. Signal Process.*, vol. 56, pp. 4306–4316, Sep. 2008.

[37] G. Thatte and U. Mitra, "Sensor selection and power allocation for distributed estimation in sensor networks: Beyond the star topology," *IEEE Trans. Signal Process.*, vol. 56, pp. 2649–2661, Jul. 2008.

[38] K. Zarifi, S. Zaidi, S. Affes, and A. Ghrayeb, "A distributed amplify-and-forward beamforming technique in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 59, pp. 3657–3674, Aug. 2011.

[39] A. D. Smith, M. K. Banavar, C. Tepedelenlioglu, and A. Spanias, "Distributed estimation over fading macs with multiple antennas at the fusion center," in *Proc. IEEE Asilomar 2009*, pp. 424–428, Nov. 2009.

[40] M. K. Banavar, A. D. Smith, C. Tepedelenlioglu, and A. Spanias, "On the effectiveness of multiple antennas in distributed detection over fading macs," *IEEE Trans. Wireless Commun.*, vol. 11, pp. 1744–1752, May 2012.

[41] M. K. Banavar, A. D. Smith, C. Tepedelenlioglu, and A. Spanias, "Distributed detection over fading macs with multiple antennas at the fusion center," *arXiv:1001.3173*, 2010.

[42] S. T. Smith, "Optimum phase-only adaptive nulling," *IEEE Trans. Signal Process.*, vol. 47, pp. 1835–1843, Jul. 1999.

[43] P. J. Kajenski, "Phase only antenna pattern notching via a semidefinite programming relaxation," *IEEE Trans. Antennas Propag.*, vol. 60, pp. 2562–2565, May 2012.

[44] M. Xia, W. Wen, and S.-C. Kim, "Opportunistic cophasing transmission in miso systems," *IEEE Trans. Commun.*, vol. 57, pp. 3764–3770, Dec. 2009.

[45] C. Tepedelenlioglu and A. B. Narasimhamurthy, "Universal distributed estimation over multiple access channels with constant modulus signaling," *IEEE Trans. Signal Process.*, vol. 58, pp. 4783–4794, Sep. 2010.

[46] S. Joshi and S. Boyd, "Sensor selection via convex optimization," *IEEE Trans. Signal Process.*, vol. 57, pp. 451–462, Feb. 2009.

[47] V. Gupta, T. H. Chung, B. Hassibi, and R. M. Murray, "On a stochastic sensor selection algorithm with applications in sensor scheduling and sensor coverage," *Automatica*, vol. 42, pp. 251–260, Feb. 2006.

[48] F. Bian, D. Kempe, and R. Govindan, "Utility-based sensor selection," in *Proc. ACM/IEEE IPSN 2006*, pp. 11–18, Apr. 2006.

[49] V. Krishnamurthy, "Dcentralized activation in dense sensor networks via global games," *IEEE Trans. Signal Process.*, vol. 56, pp. 4936–4950, Oct. 2008.

[50] Z.-Q. Luo and W. Yu, "An introduction to convex optimization for communications and signal processing," *IEEE J. Sel. Areas Commun.*, vol. 24, pp. 1426–1438, Aug. 2006.

[51] A.-J. van der Veen and A. Paulraj, "An analytical constant modulus algorithm," *IEEE Trans. Signal Process.*, vol. 44, pp. 1136–1155, May 1996.

[52] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge University Press, 2004.

[53] G. Golub and C. V. Loan, *Matrix Computations*. Baltimore, MD: The Johns Hopkins University Press, 1989.

[54] G. Kramer and M. Gastpar, "Capacity theorems for wireless relay channels," in *Proc. Allerton Conf. Communications, Control and Computing, 2003*, pp. 1074–1083, Oct. 2003.

[55] S. A. Jafar, "The ergodic capacity of phase-fading interference networks," *IEEE Trans. Info. Theory*, vol. 57, pp. 7685–7694, Feb. 2011.

[56] A. Billionnet and A. Faye, "A lower bound for a constainted quadratic 0-1 minimization problem," *Discrete Applied Mathematics*, vol. 74, pp. 135–146, Apr. 1997.

[57] R. Viswanathan and P. K. Varshney, "Distributed detection with multiple sensors: Part i-fundamentals," *Proc. IEEE*, vol. 85, pp. 54–63, Jan. 1997.

[58] R. Niu, B. Chen, and P. K. Varshney, "Fusion of decisions transmitted over Rayleigh fading channels in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 54, pp. 1018–1027, Mar. 2006.

[59] J.-F. Chamberland and V. V. Veeravalli, "Decentralized detection in sensor network," *IEEE Trans. Signal Process.*, vol. 51, pp. 407–416, Feb. 2003.

[60] J.-F. Chamberland and V. V. Veeravalli, "Asymptotic results for decentralized detection in power constrained wireless sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 22, pp. 1007–1015, Aug. 2004.

[61] W. Li and H. Dai, "Distributed detection in wireless sensor networks using a multiple access channel," *IEEE Trans. Signal Process.*, vol. 55, pp. 822–833, Mar. 2007.

[62] Z. Quan, W.-K. Ma, S. Cui, and A. H. Sayed, "Optimal linear fusion for distributed detection via semidefinite programming," *IEEE Trans. Signal Process.*, vol. 58, pp. 2431–2436, Apr. 2010.

[63] D. Ciuonzo, G. Romano, and P. Salvo Rossi, "Channel-aware decision fusion in distributed MIMO wireless sensor networks: Decode-and-fuse vs. decode-then-fuse," *IEEE Trans. Wireless Commun.*, vol. 11, pp. 2976–2984, Aug. 2012.

[64] I. Nevat, G. W. Peters, and I. B. Collings, "Distributed detection in sensor networks over fading channels with multiple antennas at the fusion centre," *IEEE Trans. Signal Process.*, vol. 62, pp. 671–683, Feb. 2014.

[65] D. Ciuonzo, P. Salvo Rossi, and S. Dey, "Massive MIMO meets decision fusion: Decode-and-fuse vs. decode-then-fuse," in *Proc. IEEE SAM 2014*, pp. 265–268, Jun. 2014.

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

[67] B. Picinbono, "On deflection as a performance criterion in detection," *IEEE Trans. Aero. and Elec. Sys.*, vol. 31, pp. 1072–1081, Jul. 1995.

[68] J. Unnikrishnan and V. V. Veeravalli, "Cooperative sensing for primary detection in cognitive radio," *IEEE J. Sel. Topics Signal Process.*, vol. 2, pp. 18–27, Feb. 2008.

[69] Z. Quan, S. Cui, and A. H. Sayed, "Optimal linear cooperation for spectrum sensing in cognitive radio networks," *IEEE J. of Sel. Topics in Signal Processing*, vol. 2, pp. 28–40, Feb. 2008.

[70] L. Martein and S. Schaible, "On solving a linear program with one quadratic constraint," *Rivista di matematica per le scienze economiche e sociali*, vol. 10, no. 1-2, pp. 75–90, 1987.

[71] C. Cheng, P. Hsiao, H. Kung, and D. Vlah, "Maximizing throughtput of UAV-relaying networks with the load-carry-and-deliver paradigm," in *Proc. IEEE WCNC 2007*, pp. 4417–4424, Mar. 2007.

[72] Z. Han, A. L. Swindlehurst, and K. J. R. Liu, "Optimization of MANET connectivity via smart deployment/movement of unmanned air vehicles," *IEEE Trans. Veh. Technol.*, vol. 58, pp. 3533–3546, Sep. 2009.

[73] E. P. de Freitas, T. Heimfarth, I. F. Netto, C. E. Lino, C. E. Pereira, A. M. Ferreira, F. R. Wagner, and T. Larsson, "UAV relay network to support WSN connectivity," in *Proc. IEEE ICUMT 2010*, pp. 309–314, Oct. 2010.

[74] R. Palat, A. Annamalau, and J. Reed, "Cooperative relaying for ad-hoc ground networks using swarm UAVs," in *Proc. IEEE MILCOM 2005*, pp. 1588–1594, Oct. 2005.

[75] P. Zhan, K. Yu, and A. L. Swindlehurst, "Wireless relay communications with unmannned aerial vehicles: performance and optimization," *IEEE Trans. Aero. and Elec. Sys.*, vol. 47, pp. 2068–2085, Jul. 2011.

[76] I. Rubin and R. Zhang, "Placement of UAVs as communication relays aiding mobile ad hoc wireless networks," in *Proc. IEEE MILCOM 2007*, pp. 1–7, Oct. 2007.

[77] J. L. Hillman, S. D. Jones, R. A. Nichols, and I. J. Wang, "Communications network architectures for the army future combat system and objective force," in *Proc. IEEE MILCOM 2002*, pp. 1417–11421, Oct. 2002.

[78] M. F. J. Pinkney, D. Hampel, and S. DiPierro, "Unmanned aerial vehicle (UAV) communications relay," in *Proc. IEEE MILCOM 1996*, pp. 47–51, Oct. 1996.

[79] M. S. Sharawi, D. Aloi, and O. A. Rawashdeh, "Design and implementation of embedded printed antenna arrays in small UAV wing structures," *IEEE Trans. Antennas and Propag.*, vol. 58, pp. 2531–2538, Aug. 2010.

[80] H. T. Kung, C.-K. Lin, T.-H. Lin, S. J. Tarsa, and D. Vlah, "Measuring diversity on a low-altitude UAV in a ground-to-air wireless 802.11 mesh network," in *Proc. IEEE Globecom Workshop on Wireless Networking for Unmanned Aerial Vehicles*, pp. 1799–1804, Dec. 2010.

[81] F. Jiang and A. L. Swindlehurst, "Dynamic UAV relay positioning for the ground-to-air uplink," in *Proc. IEEE Globecom Workshop on Wireless Networking for Unmanned Aerial Vehicles*, pp. 1766–1770, Dec. 2010.

[82] J. H. Winters, "Optimum combining in digital mobile radio with cochannel interference," *IEEE Trans. Veh. Technol.*, vol. VT-33, pp. 144–155, Aug. 1984.

[83] H. Bolcskei, M. Borgmann, and A. J. Paulraj, "Impact of the propagation environment on the performance of space-frequency coded MIMO-OFDM," *IEEE J. Sel. Areas Commun.*, vol. 21, pp. 427–439, Apr. 2003.

[84] K. Varma, "Time-delay-estimate based direction-of-arrival estimation for speech in reverberant environments," Master's thesis, Virginia Polytechnic Institute and State University, 2002.

[85] T. Trump and B. Ottersten, "Estimation of nominal direction of arrival and angular spread using an array of sensors," *Signal Processing*, vol. 50, pp. 57–69, Apr. 1996.

[86] Z. Zaidi, B. Mark, and R. Thomas, "A two-tier representation of node mobility in ad hoc networks," in *Proc. IEEE SECON 2004*, pp. 153–161, Oct. 2004.

[87] X. Zhang, D. P. Palomar, and B. Ottersten, "Statistically robust design for linear MIMO transceivers," *IEEE Trans. Signal Process.*, vol. 56, pp. 3678–3689, Aug. 2008.

[88] J. Holtzman, "Asymptotic analysis of proportional fair algorithm," in *Proc. IEEE PIMRC 2001*, pp. F33–F37, Sep. 2001.

[89] G. W. Stewart, *Matrix Algorithms Volume I: Basic Decompositions*. Philadelphia: SIAM, 1998.

[90] T. S. Rappaport, *Wireless Communications: Principles and Practice*. NJ: Prentice Hall, second ed., 2001.